

Muscle Force Prediction for Human Body Activities using Computational Intelligence Approaches

A Thesis

submitted in partial fulfillment of the requirements for the award of the degree of

Doctor of Philosophy

in

Computer Science and Engineering Department

by

Hunish Kumar Bansal

(Reg no: 951703007)



**Thapar Institute of Engineering and Technology, Patiala,
Punjab - 147004, India**

October 2024

Certificate

I hereby certify that the work, which is being presented in the thesis, entitled "Muscle Force Prediction for Human Body Activities using Computational Intelligence Approaches" partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy and submitted to the institution is an authentic record of my work carried out under the supervision of Dr. Neeraj Kumar and Dr. Prashant Singh Rana. I have cited the reference about the text(s)/figure(s)/table(s) from where they have been taken. The matter presented in this thesis has not been submitted either in-part or full to any other University/Institute for the award of any other degree.

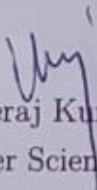


(Hunish Kumar Bansal)

Registration No. 951703007

This is to certify that the above statements made by the candidate are correct and true to the best of my knowledge.

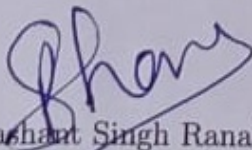
Verified by:



(Dr. Neeraj Kumar,)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala, Punjab, India.



(Dr. Prashant Singh Rana)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala, Punjab, India.

.....dedicated to all young generation researchers

Abstract

The analysis of surface electromyography (sEMG) signals plays a crucial role in understanding and quantifying muscle activation patterns, which has numerous applications in biomechanics, rehabilitation, and sports science. However, accurate modeling and prediction of sEMG time series data pose significant challenges due to the inherent complexities of physiological systems and the presence of inconsistencies and anomalies in the recorded data. This study presents an integrated machine learning framework that combines statistical anomaly detection techniques with generative adversarial networks (GANs) to enhance the predictive capabilities of sEMG-based muscle force forecasting models.

The first part of the study focused on detecting knee abnormalities using sEMG data from a UCI Machine Learning Repository dataset. It included 22 adults performing walking, leg bending, and extending tasks. Data was collected from four muscles using a Biometrics Ltd. DataLog MWX8 and a goniometer. Preprocessing involved denoising, filtering, and normalization. The study noted a class imbalance due to longer task completion times for participants with knee abnormalities. To address this, three anomaly detection techniques (Isolation Forest, k-Nearest Neighbors, and Local Outlier Factor) were employed.

Various machine learning classifiers, including LightGBM, XGBoost, Random Forest, Extra Trees, and Decision Tree, were trained and evaluated using multiple performance metrics. K-fold cross-validation assessed model robustness.

The ensemble technique combining anomaly detection and machine learning classifiers showed significant improvements over previous studies. The LightGBM classifier, trained on data processed with the Isolation Forest technique, achieved 98.5% accuracy, surpassing previous best accuracies by 6-7%.

Building upon the success of the anomaly detection and classification approach, the second part of the study focused on predictive modeling of muscle force using machine learning approaches like artificial neural networks (ANNs). To address challenges with time series data, the study proposed an integrated framework combining statistical outlier detection methods with generative adversarial networks (GANs).

The framework integrates various anomaly detection techniques (Isolation Forest, K-Nearest Neighbor, One-Class Support Vector Machine, Histogram-Based Outlier Score, and Local Outlier Factor) to improve input data quality. TimeGAN architecture was used to synthesize realistic time-series data, addressing data scarcity and enabling more

personalized predictive models.

The framework was evaluated using a dataset of body movements and forces from 57 healthy individuals. Performance of prediction models like Long Short-Term Memory (LSTM) networks, Auto-Regressive Integrated Moving Average (ARIMA), and Prophet forecasting algorithm was compared with and without anomaly detection techniques.

Results showed that integrating anomaly detection techniques significantly enhanced muscle force prediction performance. The Isolation Forest method combined with LSTM achieved a Pearson's Correlation Coefficient of 0.95 and a coefficient of determination between 0.9 and 0.93, comparable to state-of-the-art approaches. This demonstrated the benefits of integrating statistical and generative AI techniques for time series analytics in muscle activation pattern modeling and prediction.

The proposed integrated machine learning framework has several potential applications in the field of biomechanics, rehabilitation, and sports science. Accurate prediction of muscle activation patterns can facilitate the development of advanced prosthetic limbs and exoskeletons, enabling more natural and intuitive control systems. Additionally, the framework can support athletic training by providing real-time feedback and analysis of muscle activity, helping athletes optimize their performance and prevent injuries.

Furthermore, the study contributes to the investigation of muscle physiology by enabling researchers to model and simulate various muscle activation scenarios, leading to a deeper understanding of the underlying mechanisms and potential interventions for musculoskeletal disorders.

Keywords: *sEMG signal processing; anomaly detection; TimeGAN; LSTM; ARIMA; Prophet, iforest*

Acknowledgements

Reaching this long-awaited day feels like a dream come true. It wouldn't have been possible without the unwavering support system that held me up during this incredible journey. My heartfelt gratitude goes to my amazing family, friends, and even my little ones, whose patience and understanding were a constant source of strength. But most importantly, I want to express my deepest appreciation to my phenomenal research supervisors, Dr. Prashant Singh Rana and Dr. Neeraj Kumar. Their guidance proved invaluable throughout this entire ordeal. Their patience in navigating me through the complexities of research, their enthusiastic encouragement when doubt crept in, and their insightful critiques that honed my work – for all of this, I am eternally grateful. They were the perfect team to mentor me through my PhD studies, and I couldn't have imagined achieving this without their unwavering support. They fostered an environment where my intellectual curiosity could flourish, never stifling my desire to explore new avenues. When focus wavered, their insightful advice and unwavering encouragement served as guiding lights, leading me back on track. Their trust in my abilities made this journey all the more rewarding.

My deepest gratitude goes to Dr. Gurbinder Singh, Registrar, TIET for their unwavering support throughout my PhD journey. Their understanding and encouragement were especially crucial during periods when my job duties were particularly demanding. He recognized the importance of my research and made every effort to accommodate my workload, allowing me to dedicate the necessary time and focus to my studies.

Sincere thanks must also go to Dr. Maninder Singh, Dean, Academic Affairs and Dr. Shalini Batra, Head, CSED, TIET for their continuous support by always willing to help whenever needed. Their motivation and enthusiasm are contagious. I want to thank the members of my doctoral committee: Dr. Inderver Channa, Dr. Ashima Singh and Dr. Jainy Sachdeva. They generously gave their time to offer insightful comments towards improving my work.

My seniors, like wise elder siblings, showered me with love and care, but also knew when to offer a gentle nudge to keep me focused. To my fellow labmates, my deepest thanks for the stimulating discussions and invaluable suggestions that pushed my research forward. But even more importantly, thank you for the camaraderie and shared laughter that made the long hours fly by. The faculty's encouraging words, always offered with a smile in passing, brightened my days. The office staff consistently went above and beyond to assist me, ensuring a smooth experience. To all of you, my sincerest thanks.

This journey wouldn't have been possible without the unwavering love and support of my incredible family. My wife, Kavita, has been my rock. Her unwavering love, patience, and understanding were a constant source of strength that carried me through. To my amazing parents, my deepest gratitude for always believing in me and instilling the value of perseverance. Your encouragement was a constant fuel for my determination, even during the most challenging periods. Medant and Gamyra, your innocent laughter and joy were a constant reminder of the beauty in life that fueled my determination. You may not understand the significance of this day yet, but know that this achievement is a testament to our journey together.

Finally, a whisper of gratitude to the Divine for the unwavering light that illuminated my path. Blessings, strength, and guidance – these gifts carried me through, even in the darkest hours.

Table of Contents

Title	Page No.
Abstract	v
Table of Contents	ix
List of Figures	xiii
List of Tables	xv
List of Abbreviations	xvii
Chapter 1 Introduction	1
1.1 Background	1
1.1.1 Human Biomechanics	2
1.1.2 Importance of Muscle Force Prediction	4
1.1.3 Challenges in Muscle Force Prediction	5
1.1.4 Traditional Approaches	6
1.1.5 Computational Intelligence for Muscle Force Prediction	6
1.2 Research Gaps	9
1.3 Research Objectives	10
1.4 Thesis Contribution	10
1.5 Thesis Organization	11
Chapter 2 Literature Survey	13
2.1 Human Biomechanics	14
2.2 Muscle Force Prediction	16
2.3 Feature Extraction	21
2.4 TimeGAN - Generative Adversarial Network	23
2.5 Anomaly Detection Techniques	25
2.6 Machine Learning Techniques	27
2.6.1 Decision Tree	28
2.6.2 Support Vector Machine	29
2.6.3 Bayes Rule	30
2.6.4 Artificial Neural Networks	31

2.7	Time Series Prediction Techniques	31
2.7.1	ARIMA	31
2.7.2	Prophet	32
2.7.3	RNN	33
2.7.4	LSTM	33
Chapter 3 Decoding Human Movement: Musculoskeletal Modeling and Sensor Technology		35
3.1	Musculoskeletal Modeling	36
3.1.1	Musculoskeletal Modeling Tools	36
3.1.2	Related Work	40
3.2	Biomechanics Sensor Technology	42
3.2.1	Biomechanics Sensors	42
3.2.2	Related Work	46
3.3	Conclusion	47
Chapter 4 An Ensemble Machine Learning Technique for Detection of Abnormalities in Knee Movement Sustainability		49
4.1	Introduction	49
4.1.1	Motivation	51
4.1.2	Contributions	52
4.1.3	Organization	52
4.2	Materials and Methods	53
4.2.1	Dataset and its Features	53
4.2.2	Approach	55
4.3	Model Evaluation	63
4.3.1	Matthews Correlation Coefficient (MCC)	63
4.3.2	F1 score	64
4.3.3	Precision	64
4.3.4	Accuracy	64
4.3.5	AUC	64
4.4	Result Analysis	65
4.4.1	Model Performance Comparison	65
4.4.2	K-Fold Validation	66
4.4.3	Results and Discussion	68
4.4.4	Limitation of the Study	68
4.5	Conclusion	68

Chapter 5	Time series generative adversarial network for muscle force prognostication using statistical outlier detection	71
5.1	Introduction	72
5.1.1	Related Work	73
5.1.2	Motivation	75
5.1.3	Contribution	76
5.1.4	Organization	77
5.2	Material and Methods	77
5.2.1	Dataset and its Description	77
5.2.2	Anomaly Detection Techniques	79
5.2.3	Prediction Models	82
5.3	Muscle Force Generating Model	86
5.4	Experiment Setup	88
5.5	Results and Discussion	89
5.5.1	Evaluation Parameters	89
5.5.2	Result Summary	90
5.5.3	Limitations of the Study	94
5.6	Conclusion	95
Chapter 6	Conclusion and Future Directions	97
6.1	Conclusion	97
6.2	Scope for Future Work	99
References	101
List of Publications	123

List of Figures

Figure No.	Title	Page No.
1.1	Human Biomechanics	2
2.1	Supervised Learning	28
2.2	Unsupervised Learning	28
2.3	Semi-Supervised Learning	29
4.1	Muscle diagram for rectus femoris (RF), vastus lateralis (VL), vastus medialis (VM), semitendinosus (ST), biceps femoris (BF) and semitmembranosus (SM).	53
4.2	Methodology Used	57
4.3	The sEMG signals at various stages during data preprocessing	58
4.4	Cross Validation of best model for accuracy, F1-score and MCC.	67
5.1	Architecture of TimeGAN with real data, it consist of the training phase of autoencoding and adversarial components.	87
5.2	Methodology of the proposed work	88
5.3	Muscle forces prediction using LSTM model on testing data with and without anomaly detection techniques.	91
5.4	Muscle forces prediction using ARIMA model on testing data with and without anomaly detection techniques.	91
5.5	Muscle forces prediction using Prophet model on testing data with and without anomaly detection techniques.	91

List of Tables

Table No.	Title	Page No.
3.1	Musculoskeletal Modeling Tools	37
3.2	Biomechanics Sensors	43
4.1	Description of the dataset	54
4.2	Description of the features	54
4.3	Sample dataset	54
4.4	Statistical analysis of each feature in each dataset	56
4.5	The feature importance of each feature.	61
4.6	Performance comparison of different anomaly detection techniques with different prediction models on different metrics on all the datasets.	65
5.1	Description of the datasets	78
5.2	The sample dataset of kinematic force (in Newton).	78
5.3	Parameters for LSTM model	83
5.4	Parameters for the Prophet model	85
5.5	Parameters for the ARIMA model	86
5.6	Comparative performance study of ARIMA model with different anomaly detection techniques on different datasets	92
5.7	Comparative performance study of LSTM model with different anomaly detection techniques on different datasets	93
5.8	Comparative performance study of Prophet model with different anomaly detection techniques on different data sets	94

List of Abbreviations

ADT	Anomaly Detection Techniques
ANNs	Artificial Neural Networks
ARIMA	Auto-Regressive Integrated Moving Average
AUC	Area Under the Curve
CC	Correlation Coefficient
CNN	Convolutional Neural Network
DL	Deep Learning
DNN	Deep Neural Network
EEG	Electroencephalography
EMG	Electromyography
GA	Genetic Algorithm
GAN	Generative Adversarial Network
IDC	Inverse Dynamic Computation
IF	Isolation Forest
IMU	Inertial Measurement Units
KCF	Knee Contact Force
KNN	K-Nearest Neighbour
LightGBM	Light Gradient-Boosting Machine
LLD	Limb-Length Discrepancy
LOF	Local Outlier Factor
LR	Linear Regression
LSTM	Long Short Term Memory
MAE	Mean Absolute Error
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
NMS	Neuromusculoskeletal
OSVM	Open Support Vector Machine
RFR	Random Forest Regressor
RNN	Recurrent Neural Network
sEMG	Surface Electromyography
SVR	Support Vector Regressor
VR	Virtual Reality
XGBoost	Extreme Gradient Boosting

Chapter 1

Introduction

Muscle force prediction is a fundamental challenge in biomechanics, with far-reaching implications for various fields, including sports science, rehabilitation engineering, and computer animation. Accurate estimation of the forces generated by muscles during human movement is essential for understanding the underlying mechanisms driving biomechanical processes, optimizing athletic performance, designing effective rehabilitation programs, and creating realistic simulations of musculoskeletal dynamics. Applications range from biomechanics research analyzing gait patterns and ergonomics to rehabilitation programs optimizing therapy techniques and robotics design for prosthetics and exoskeletons.

1.1 Background

Muscle forces are the driving forces behind human movement, generated by the contraction of skeletal muscles. These forces are transmitted through tendons and bones, allowing for the production of joint moments and ultimately enabling locomotion and interaction with the environment. Understanding and predicting muscle forces is crucial for various applications in biomechanics, sports science, rehabilitation engineering, and computer animation.

The force generated by a muscle is primarily determined by its physiological properties, such as muscle fiber composition, cross-sectional area, and muscle length. The force-length relationship describes how the force produced by a muscle varies with its length, with the maximum force typically occurring at an optimal muscle length [1]. Additionally, the force-velocity relationship governs how the force generated by a muscle changes with the velocity of contraction or shortening [2].

Muscle forces are also influenced by neural activation patterns, as the central nervous system controls the recruitment and firing rates of motor units within each muscle. Electromyography (EMG) techniques are commonly used to measure the electrical activity of muscles, which can provide insights into muscle activation levels and assist in estimating muscle forces [3].

1.1.1 Human Biomechanics

The human body is a marvel of engineering, its intricate structure designed for a remarkable range of movement. From the graceful ballet dancer to the powerful weightlifter, every action we take is governed by the principles of biomechanics. This fascinating field delves into the mechanics of the musculoskeletal system, analyzing how bones, muscles, ligaments, and nerves work together to generate movement and maintain stability.

Human biomechanics is the study of the structure and function of biological systems, particularly the human body, through the lens of mechanics. It is a multidisciplinary field that combines principles from disciplines such as anatomy, physiology, engineering, physics, and mathematics to understand and analyze human movement and the forces acting on the body. (Figure 1.1)

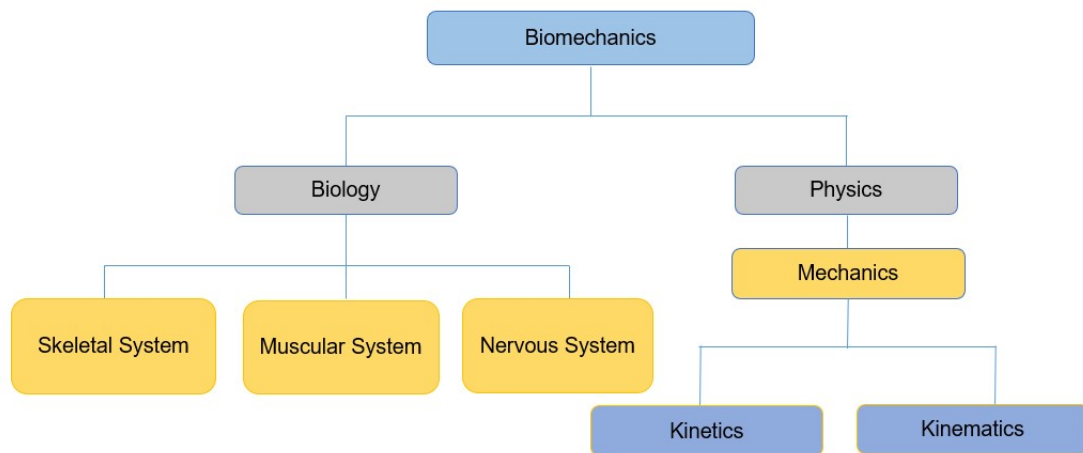


Figure 1.1: Human Biomechanics

At its core, human biomechanics seeks to understand how the various components of the musculoskeletal system, including bones, joints, ligaments, tendons, and muscles, interact to produce and control movement. This knowledge is essential for a wide range of applications, including sports performance optimization, injury prevention and rehabilitation, ergonomic design, and the development of assistive devices and prosthetics.

One of the primary focus areas in human biomechanics is the study of muscle mechanics and muscle force generation. Muscles are the primary actuators that drive human movement, and understanding how they generate force and transmit that force to the skeletal system is crucial for analyzing and predicting motion patterns. Biomechanists use various techniques, such as electromyography (EMG), motion capture, and computational modeling, to study muscle activation patterns, force-length relationships, and force-velocity characteristics [4].

Kinematics, which deals with the study of motion without considering the forces that cause it, is another important aspect of human biomechanics. Kinematic analyses involve the measurement and description of the positions, velocities, accelerations, and other kinematic parameters of body segments during movement. This information is essential for understanding the coordination and control of human movement, as well as for identifying potential risk factors for injuries or movement dysfunctions [5].

Kinetics, on the other hand, focuses on the study of forces and their effects on motion. Biomechanists use force platforms, motion capture systems, and computational models to measure and analyze the various forces acting on the body during different activities, such as ground reaction forces, joint reaction forces, and muscle forces. This information is crucial for understanding the loads and stresses experienced by the musculoskeletal system and for designing interventions or devices to mitigate these loads [6].

Human biomechanics also plays a vital role in sports science and performance optimization. By analyzing the biomechanics of specific sports movements, such as throwing, jumping, or running, researchers can identify optimal techniques and strategies for maximizing performance and minimizing the risk of injury. This knowledge can be used to develop training programs, design specialized equipment, and provide biomechanical feedback to athletes and coaches [7].

In the field of rehabilitation engineering, human biomechanics is essential for the design and development of assistive devices, such as prosthetics, orthotics, and exoskeletons. By understanding the biomechanical principles of human movement and the specific needs of individuals with disabilities or impairments, engineers can create devices that effectively restore or enhance functional mobility and independence [8].

Ergonomics, the study of human-machine interactions and the design of workplace environments, also heavily relies on principles of human biomechanics. Ergonomists use biomechanical analyses to evaluate the physical demands of various tasks and workstations, with the goal of optimizing workplace design and minimizing the risk of musculoskeletal disorders and injuries [9].

In addition to its practical applications, human biomechanics also contributes to our fundamental understanding of human physiology and the evolution of human movement patterns. By studying the biomechanics of various activities, such as walking, running, or climbing, researchers can gain insights into the adaptations and trade-offs that have shaped the human musculoskeletal system over evolutionary timescales [10].

Biomechanics doesn't exist in isolation; it draws heavily from various disciplines to understand the human body in motion. Here's a breakdown of the key players:

- **Anatomy:** This branch of biology provides a detailed blueprint of the human body, focusing on the structure and arrangement of bones, muscles, ligaments, tendons, and other components. Biomechanists use this knowledge to understand the lever systems formed by bones and joints, and how muscles interact with them to produce movement [11].
- **Physiology:** This science explores the function of the human body's systems, including the nervous system and the musculoskeletal system. By understanding how muscles are activated through nerve impulses and how forces are transmitted through bones and joints, biomechanists can explain the mechanics behind specific movements [12].
- **Mechanics:** This fundamental science provides the basic principles of how forces, motion, and energy interact with the body. Biomechanists apply these principles to analyze the forces generated by muscles, the effects of gravity on the body, and the resulting movement patterns[13].

At the heart of human movement lies the power of muscles. These contractile tissues act as the engines, generating forces that act on bones. Bones, on the other hand, function as levers, transmitting these forces and producing movement around joints. Biomechanists use several concepts to analyze this interplay:

- **Center of Gravity (CoG):** This is the theoretical point where the body's entire weight is concentrated. Maintaining balance requires the CoG to be positioned within the base of support (the area of contact with the ground). Biomechanists analyze how body position and movement affect the CoG.
- **Moments and Levers:** Muscles generate forces that create moments (rotational effects) around joints. By understanding the lever systems formed by bones and the location of muscle attachments, biomechanists can predict the direction and magnitude of joint movement.
- **Kinetics and Kinematics:** These terms differentiate between the "why" and "how" of movement. Kinetics focuses on the forces that cause movement (forces generated by muscles, gravity, etc.), while kinematics describes the actual movement patterns themselves (range of motion, angular displacement, etc.).

1.1.2 Importance of Muscle Force Prediction

Muscle force prediction offers a multitude of benefits, impacting various disciplines:

- **Biomechanics and Movement Analysis:** By estimating the forces exerted by

individual muscles, researchers gain valuable insights into how the body generates movement. This knowledge helps us understand the mechanics behind healthy and pathological gait patterns, identify muscle imbalances leading to injuries, and optimize athletic performance [14].

- **Clinical Applications and Rehabilitation:** Precise muscle force prediction plays a crucial role in diagnosing and treating musculoskeletal disorders. It allows clinicians to pinpoint weaknesses or imbalances contributing to pain or movement limitations. This data guides the development of targeted rehabilitation programs, ensuring a more efficient recovery process [4].
- **Prosthetic and Assistive Device Design:** Accurately predicting muscle forces is essential for designing effective prosthetics and exoskeletons. By understanding the forces a user intends to generate, engineers can create devices that seamlessly integrate with the body's natural movement patterns. This leads to improved functionality, comfort, and user acceptance [15].
- **Human-Robot Interaction:** In collaborative robotics, predicting human muscle forces facilitates safe and intuitive interaction between humans and robots. This allows robots to anticipate and adapt to human movements, minimizing the risk of injury and fostering seamless collaboration [16].
- **Sports Science and Performance Optimization:** Athletes constantly strive to push their limits. Muscle force prediction helps coaches and trainers design training programs that target specific muscle groups, optimize movement mechanics, and ultimately enhance athletic performance [17].

1.1.3 Challenges in Muscle Force Prediction

While the potential benefits of muscle force prediction are undeniable, accurately measuring and predicting these forces remains a challenge. The human body is a complex system with muscles working synergistically to produce movement. Factors like muscle activation patterns, individual muscle properties, and joint kinematics all contribute to the final force generated.

- **Complexity of Musculoskeletal System:** The musculoskeletal system is highly complex, comprising numerous muscles, tendons, and joints. Predicting muscle forces accurately requires comprehensive biomechanical models that account for these complexities, posing a significant challenge to researchers [18].
- **Variability Among Individuals:** Individuals exhibit variability in muscle ar-

chitecture, activation patterns, and movement strategies, making it challenging to generalize muscle force predictions across populations. Personalized modeling approaches are necessary to account for this variability and improve prediction accuracy [19].

- **Limited Experimental Data:** Validating muscle force prediction models relies on experimental data obtained through invasive techniques such as electromyography (EMG) or muscle force measurements, which may not always be feasible or ethical. This limitation hampers the development and refinement of predictive models [20].
- **Dynamic Nature of Movement:** Human movement is dynamic and multifaceted, involving interactions between muscles, tendons, and external forces. Predicting muscle forces accurately requires accounting for dynamic changes in movement patterns, which adds complexity to modeling efforts [21].

1.1.4 Traditional Approaches

Traditional approaches to muscle force prediction have primarily relied on optimization-based musculoskeletal models and electromyography (EMG) techniques. Optimization-based musculoskeletal models employ mathematical optimization algorithms to estimate muscle forces that produce the desired joint moments or motion patterns, subject to physiological constraints [22]. These models often incorporate complex musculoskeletal geometry, muscle properties, and optimization criteria, making them computationally expensive and sensitive to parameter settings.

EMG techniques, on the other hand, involve measuring the electrical activity of muscles using surface or intramuscular electrodes. The recorded EMG signals are then used to estimate muscle forces through calibration procedures or by applying biomechanical models [23]. However, EMG-based methods are limited by factors such as cross-talk between adjacent muscles, signal variability, and the need for extensive calibration and subject-specific modeling.

1.1.5 Computational Intelligence for Muscle Force Prediction

While these traditional approaches have contributed significantly to the understanding of muscle mechanics, they have inherent limitations in handling the complexity, nonlinearity, and uncertainty associated with musculoskeletal systems. Computational intelligence techniques, which are inspired by biological systems and possess learning and adaptation capabilities, have emerged as promising alternatives for addressing these challenges in muscle force prediction. This is where advancements in computational intelligence (CI)

offer a promising alternative.

CI encompasses a range of techniques inspired by natural intelligence, including artificial neural networks (ANNs), fuzzy logic (FL), and evolutionary algorithms (EAs). These approaches can learn complex relationships between body kinematics, kinetics, and electromyography (EMG) signals, leading to more efficient and accurate muscle force prediction.

Artificial Neural Networks (ANNs) are one of the most widely explored computational intelligence techniques for muscle force prediction. ANNs are computational models inspired by the structure and function of biological neural networks, consisting of interconnected nodes (neurons) that process information through weighted connections and nonlinear activation functions [24]. ANNs have the ability to learn complex patterns and relationships from data, making them well-suited for modeling the nonlinear and intricate relationships between muscle activations, kinematic variables, and muscle forces.

Several studies have demonstrated the effectiveness of ANNs in estimating muscle forces during various human movements, such as gait [25], upper limb tasks [26], and sports-related activities [27]. ANNs can be trained using experimental data, such as EMG signals, motion capture data, and ground reaction forces, to learn the mapping between input features and desired muscle forces. Once trained, ANNs can provide fast and efficient predictions of muscle forces for new input data.

Fuzzy logic is another computational intelligence technique that has been applied to muscle force prediction problems. Fuzzy logic is based on fuzzy set theory, which allows for the representation and manipulation of imprecise or vague information [28]. This characteristic makes fuzzy logic particularly useful for handling the inherent uncertainties and nonlinearities associated with musculoskeletal systems.

Fuzzy logic-based approaches have been used to estimate muscle forces in various applications, such as rehabilitation exercises [29] and modeling muscle activation dynamics [30]. These approaches typically involve defining fuzzy rules that relate input variables (e.g., joint angles, muscle lengths, and EMG signals) to muscle forces, based on expert knowledge or data-driven methods. Fuzzy inference systems can then be employed to estimate muscle forces by applying these fuzzy rules to new input data.

Evolutionary algorithms, such as genetic algorithms and particle swarm optimization, are another class of computational intelligence techniques that have been employed in muscle force prediction. EAs are particularly useful for optimizing complex muscle force prediction models [31]. By evolving a population of candidate solutions, EAs can identify the optimal combination of factors influencing muscle force. This can be particularly

beneficial when dealing with a multitude of muscles and complex movement patterns [32].

One of the key advantages of computational intelligence techniques is their ability to learn from data and adapt to new situations. This capability is particularly valuable in muscle force prediction, as musculoskeletal systems exhibit complex nonlinear behaviors and are influenced by various factors, such as individual characteristics, muscle properties, and external loads. By learning from experimental data or simulations, computational intelligence techniques can capture these intricate relationships and provide accurate predictions without relying on explicit mathematical models or simplifying assumptions.

Another advantage of computational intelligence techniques is their ability to handle uncertainty and imprecision in data. Muscle force prediction often involves noisy or incomplete data, such as EMG signals, motion capture data, or ground reaction forces. Traditional modeling approaches may struggle with such uncertainties, leading to inaccurate or unstable predictions. Computational intelligence techniques, like fuzzy logic and ANNs, are well-suited for dealing with these uncertainties and can provide more robust and accurate predictions.

Moreover, computational intelligence techniques can be combined or hybridized with other techniques, such as musculoskeletal models or optimization algorithms, to leverage their respective strengths and overcome their limitations. These hybrid approaches have shown promising results in improving the accuracy and robustness of muscle force prediction [33][34]. For example, ANNs can be integrated with musculoskeletal models to provide fast and efficient predictions, while still incorporating biomechanical constraints and muscle properties [35].

The integration of these CI techniques offers even greater potential. Neuro-fuzzy systems combine the learning capabilities of ANNs with the interpretability of FL. This allows for not only accurate predictions but also an understanding of the underlying factors influencing muscle forces [36, 37]. Similarly, genetic algorithms can be used to optimize the architecture and parameters of ANNs, leading to improved prediction accuracy [38, 39].

Despite the advantages of computational intelligence techniques, there are also challenges and limitations to their application in muscle force prediction. One significant challenge is the availability and quality of training data. Computational intelligence techniques, particularly ANNs and data-driven fuzzy systems, rely heavily on learning from examples. Obtaining comprehensive and accurate data for muscle forces and related parameters can be difficult and may require invasive measurement techniques or specialized equip-

ment.

Another challenge is the interpretability and transparency of some computational intelligence models, such as ANNs, which can be perceived as "black boxes" with limited insight into their internal workings. This lack of interpretability can hinder the understanding of the underlying mechanisms and limit the acceptance of these techniques in certain applications where transparency and explainability are crucial.

Additionally, the selection and optimization of computational intelligence techniques often involve numerous parameters and design choices, which can be time-consuming and require domain expertise. Improper parameter settings or model configurations can lead to suboptimal performance or overfitting issues, where the model performs well on training data but fails to generalize to new situations.

Despite these challenges, computational intelligence techniques continue to gain traction in the field of muscle force prediction due to their potential for improving accuracy, handling nonlinearities and uncertainties, and providing adaptive and robust solutions. As computational resources and data availability continue to improve, and as research in this area progresses, computational intelligence techniques are expected to play an increasingly important role in advancing our understanding and prediction of muscle forces in various applications, including biomechanics, sports science, rehabilitation engineering, and computer animation.

1.2 Research Gaps

1. In the literature SVM, neural network are mostly used. To get effective predictions, more models need to be explored for muscle force prediction [20, 40, 41].
2. It is essential to select those features in data that are most relevant to the problem domain, this process is known as feature selection. Its main purpose is to reduce the complexity of the dataset. Moreover, many papers surveyed lack the feature importance module itself. Parameter optimization will also help in better accuracy of results [42, 43, 44, 45].
3. In most of the papers, K-fold cross-validation have been used. But, there are many more validation types that can be used to check the robustness of our model like resubstituting validation, hold-out validation, leave-one-out validation [46, 47].
4. The Post model evaluation is an important process that take place after the training of model. It is based upon true positive and false positive rate of the prediction. It tells which feature has the most impact on false prediction. By using post model

evaluation, accuracy of the model can be improved. The main feature of this process is that it tells about the property which is dropping down the accuracy of model.

1.3 Research Objectives

1. To study and review the existing techniques to predict muscle forces for different human body activities.
2. To propose an optimized feature selection approach for muscle force prediction.
3. To propose the methodology for prediction of various muscle forces in different human body activities using computational intelligence techniques.
4. To verify and validate the proposed technique.

1.4 Thesis Contribution

1. Anomaly detection techniques like iforest, KNN, and LOF are used to detect and remove abnormal data from the sEMG signals.
2. Several machine learning classifiers like LightGBM, XGBoost, Random Forest, Extra Trees, and Decision Tree are compared for their performance on classifying normal vs abnormal movements from the sEMG data.
3. The proposed methodology using anomaly detection and ensemble machine learning models enables more accurate diagnosis of knee movement abnormalities from sEMG signals compared to prior approaches.
4. Specializing TimeGAN on detected clinical outliers enhances data augmentation for precision prognostic modeling.
5. The integrated framework makes important contributions around effectively harnessing anomalies and simulated sequences for robust and personalized muscle modelling. This drives the next generation of precision medicine in rehabilitation and sports biomechanics.
6. Key innovation in connecting deep generative learning and abnormal trajectory identification to unlock more patient-specific muscle strength analytics

1.5 Thesis Organization

The Thesis is organized as follows.

Chapter 1: Introduction

The chapter begins with the background of the Biomechanics, Importance of Muscle force prediction and challenges faced in the same. Further, the chapter discusses the traditional approaches and then Computational Intelligence for Muscle Force Prediction. The chapter finally comes up with thesis contribution and organization.

Chapter 2: Literature Survey

The literature survey in this chapter covers the extensive research and methodologies in muscle force prediction. It explores the historical context and the evolution of techniques from classical biomechanics to modern computational models. The chapter discusses the interdisciplinary nature of the field, integrating principles from anatomy, physiology, engineering, and computational modeling. It highlights various methodologies, including electromyography, motion capture, musculoskeletal simulations, and machine learning algorithms, used to predict muscle forces. Additionally, it addresses the applications of muscle force prediction in sports science, rehabilitation, ergonomics, and robotics, while also identifying the challenges and future directions in this field.

Chapter 3: Decoding Human Movement: Musculoskeletal Modeling and Sensor Technology

This chapter explores the field of human biomechanics, focusing on how the musculoskeletal system interacts with external forces to generate and control movement. It highlights the importance of musculoskeletal modeling, a computational approach that replicates human body structure and function, using both rigid-body and deformable models. The chapter also discusses advanced sensor technologies like Inertial Measurement Units (IMUs) and electromyography (EMG), which provide crucial data for understanding muscle forces and joint moments. The integration of these models and sensors offers valuable insights into rehabilitation, sports science, ergonomics, and robotics.

Chapter 4: An Ensemble Machine Learning Technique for Detection of Abnormalities in Knee Movement Sustainability

This chapter covers the use of machine learning techniques for detecting abnormalities in knee movement using surface electromyography (sEMG) data. It describes a study that collected sEMG signals from 22 individuals during sitting, standing, and

walking activities. Various anomaly detection methods, including isolation forest (iforest), k-nearest neighbors (KNN), and local outlier factor (LOF), were applied to preprocess the data. This study then used multiple machine learning classifiers, such as Light Gradient Boosting Machine and Random Forest, to classify normal and abnormal knee movements. The results showed that applying anomaly detection techniques, particularly iforest, significantly improved classification accuracy, achieving up to 98.5% accuracy compared to 85.9% without anomaly detection. The chapter discusses the methodology, results, and potential applications in medical diagnostics.

Chapter 5: Time series generative adversarial network for muscle force prognostication using statistical outlier detection

This chapter covers the development of a novel approach for predicting muscle forces using time series data and machine learning techniques. It introduces an integrated framework that combines statistical anomaly detection methods with generative adversarial networks (GANs) to enhance prediction accuracy. The study employs various models including LSTM, ARIMA, and Prophet, along with anomaly detection techniques like isolation forest and local outlier factor. A dataset of walking tests from 57 individuals was used to apply various methods to forecast muscle forces. The chapter discusses the methodology, experimental setup, evaluation metrics, and results, showing how the proposed approach improves prediction performance compared to conventional methods. It also addresses limitations and potential future directions for this research.

Chapter 6: Conclusion and Future Directions

This chapter summarizes the conclusions drawn from the thesis along with the possible future directions.

Chapter 2

Literature Survey

Muscle force prediction stands as a cornerstone in the realm of biomechanics, offering valuable insights into the intricate dynamics of human movement. Understanding the forces generated by muscles during various activities is pivotal not only for deciphering the complexities of musculoskeletal mechanics but also for optimizing performance, preventing injuries, and enhancing rehabilitation protocols. This literature review endeavors to provide a comprehensive examination of muscle force prediction, delving into its historical context, biomechanical underpinnings, methodologies, applications across diverse fields, challenges, current trends, and future directions.

Over the years, the study of muscle force prediction has evolved significantly, driven by advancements in technology, computational methods, and interdisciplinary collaborations. Initially rooted in classical biomechanics, which primarily focused on describing movement kinematics and kinetics, muscle force prediction has emerged as a multidimensional endeavor that integrates principles from anatomy, physiology, engineering, and computational modeling. This evolution reflects a growing recognition of the critical role that muscle forces play in governing human motion and functionality.

The significance of muscle force prediction extends across various domains, from sports science and rehabilitation to ergonomics and robotics. In sports performance optimization, predicting muscle forces enables coaches and athletes to fine-tune training regimens, refine technique, and mitigate the risk of overuse injuries. Similarly, in rehabilitation settings, accurate muscle force prediction guides clinicians in designing personalized treatment plans, facilitating recovery from musculoskeletal injuries and surgeries.

Methodologically, muscle force prediction encompasses a diverse array of techniques, ranging from experimental approaches such as electromyography (EMG) and motion capture to computational models like musculoskeletal simulations and machine learning algorithms. These methodologies offer complementary insights into muscle function, allowing researchers to validate and refine predictive models through empirical data and theoretical frameworks.

Despite its profound implications, muscle force prediction is not without challenges. The

complexity of the musculoskeletal system, inter-subject variability, validation issues, and computational resource requirements pose significant hurdles in accurately predicting muscle forces across diverse populations and activities. Addressing these challenges requires interdisciplinary collaboration, methodological innovation, and data-driven approaches.

In light of ongoing advancements and emerging trends, the future of muscle force prediction holds promise for transformative breakthroughs in biomechanics research and practice. Integrating advanced imaging techniques, leveraging machine learning algorithms, and embracing wearable sensor technologies offer new avenues for enhancing the accuracy and applicability of predictive models. Moreover, fostering collaborative research efforts and translating findings into clinical and practical applications will be essential for realizing the full potential of muscle force prediction in improving human health, performance, and well-being.

In summary, this literature review aims to synthesize existing knowledge, identify gaps in understanding, and provide insights into the past, present, and future of muscle force prediction. By examining the historical evolution, biomechanical foundations, methodological approaches, applications, challenges, and future directions, this review seeks to contribute to a deeper understanding of this pivotal aspect of human movement biomechanics. Through comprehensive analysis and critical synthesis of the literature, this review endeavors to advance our collective understanding of muscle force prediction and its implications for research, clinical practice, and technological innovation.

2.1 Human Biomechanics

Cappozzo, A., et al. [48] proposed a standardized method for defining anatomical frames of reference for human body segments using external markers placed on bony landmarks. The procedures allow accurate 3D reconstruction of body segment positions and orientations during movement analysis from skin-mounted marker trajectories. This overcame limitations of previous approaches and enabled consistent description of 3D joint angles across studies. The proposed conventions have become widely adopted standards in biomechanics research.

Winter, D.A. in the textbook [5] covers fundamental principles of human movement biomechanics and motor control. Key topics include musculoskeletal anatomy, mechanics of muscles and joints, gait analysis, sports biomechanics, neural control, and clinical applications. With its interdisciplinary approach bridging biology and engineering, this book provides a thorough foundation for students and researchers in biomechanics. Its

clear explanations and worked examples facilitate understanding of complex biomechanical concepts and their applications.

McGill, S.M., et al. [49] introduced a detailed biomechanical model to quantify spinal loading during heavy lifting tasks. The model considered contributions from muscle forces acting on the spine, intra-abdominal pressure, and external loads to compute spinal compression and shear forces. It enabled systematic evaluation of different lifting techniques and their effects on spine loading. The model provided new insights into factors influencing spine injury risk and guided ergonomic interventions for manual material handling tasks.

Using 3D motion analysis, Kaufman, K.R., et al. [50] examined gait patterns in children with limb-length discrepancy (LLD). The study quantified asymmetries in joint angles, moments, and powers between the longer and shorter limbs during walking. Greater LLD was associated with increased gait deviations, including greater knee flexion during stance on the longer limb. The results provided clinical insights into gait compensations adopted by LLD patients and highlighted the utility of 3D gait analysis for surgical planning.

Chiari, L., et al. [51] developed a parallel computing technique to analyze postural sway during quiet standing using center-of-pressure measurements. The proposed approach utilized graphical processing units (GPUs) to parallelize sway path analyses, enabling higher resolution and reduced noise compared to conventional methods. Improved characterization of sway dynamics from the parallel technique could enhance screening for balance impairments. The study demonstrated the benefits of leveraging parallel computing for biomechanical data analysis.

Gittoes, M.J., et al. [52] used a four-segment biomechanical model of the buttocks-thigh complex to simulate impact forces and soft tissue dynamics during seated impacts. The model incorporated representations of skin, muscle, and adipose tissues to investigate their influence on energy absorption and load distribution. Parametric analyses quantified effects of factors like impact velocity, soft tissue stiffness, and muscle activation levels. The findings provided insights into injury risk assessment and design guidance for seating systems.

Zajac, F.E., et al. [53] developed a comprehensive musculoskeletal model of the human leg to simulate muscle forces and coordination during walking. Part I described the model formulation, incorporating musculoskeletal geometry, muscle properties, and neural control elements. Part II applied this model to estimate muscle forces, investigate energetics, and analyze control strategies during normal walking. The simulations helped elucidate principles underlying human locomotion and the neural mechanisms coordinating muscle

activations. This work significantly advanced computational modeling of human movement biomechanics.

Neptune, R.R., et al. [54] used musculoskeletal modeling and simulations to determine the contributions of the individual ankle plantar flexors (soleus, gastrocnemii, other deep flexors) during walking. The soleus was found to be the primary contributor to body support, while the gastrocnemii played a major role in forward progression. The other deep flexors initiated leg swing. This advanced understanding of muscle-specific roles could guide treatment of gait impairments.

Gruber, A. H., et al. [55] investigates the biomechanical factors contributing to running-related injuries by analyzing gait patterns. Through motion capture technology and force plates, researchers identify key parameters such as foot strike pattern, joint angles, and ground reaction forces. Findings suggest that deviations from optimal biomechanics may increase injury risk, highlighting the importance of corrective interventions and training strategies.

Matjačić, Z., et al. [56] examines the integration of biomechanics with virtual reality (VR) technology to analyze human movement in immersive environments. Through motion capture systems and VR simulations, researchers investigate gait patterns, joint kinematics, and motor control, providing insights into rehabilitation strategies and ergonomic design.

Herr, H. M. et al.[57] investigates the biomechanics of human locomotion with prosthetic limbs, focusing on gait analysis, energy expenditure, and interface pressures. Findings inform the design of prosthetic devices to improve mobility, stability, and comfort for individuals with limb loss.

2.2 Muscle Force Prediction

Zhang J et al. [25] introduces a physics-informed deep learning approach to predict muscle forces and joint kinematics from surface electromyography (EMG) data. Neural networks are trained on experimental EMG and motion data, while being constrained by musculoskeletal dynamics. This allows accurate prediction of muscle forces, joint angles, and moments for new movements. The authors validated their method on various dynamic tasks, demonstrating high estimation accuracy. This framework enables real-time musculoskeletal analysis by combining data-driven and model-based approaches, with potential applications in biomechanics research, rehabilitation, and human-machine interfaces.

Ma, Ruyi, et al.[58] explores force control in prosthetic hands by utilizing surface electromyography (sEMG) signals and six-dimensional force sensors. sEMG signals from the forearm and grasping force data are collected. The root mean square (RMS) of steady-state sEMG signals is selected as a feature. Prediction models are constructed using gene expression programming (GEP) and BP neural networks to predict grasping force. Evaluation metrics include root mean square error (RMSE) and correlation coefficient (CC). Results indicate that RMS features from sEMG signals effectively characterize grasping force, with the GEP algorithm yielding superior prediction accuracy, particularly evident across various grasping power levels and modes.

Zhu, Yean, et al. [59] addresses the challenge of assessing knee joint function in patients with knee impairments, focusing on knee contact force (KCF) prediction. Traditional methods such as instrumented prosthetic implants or inverse dynamics analysis are invasive, expensive, and time-consuming. Here, a novel approach integrating the Artificial Fish Swarm and Random Forest algorithm is proposed. A Random Forest model learns the nonlinear relationship between gait parameters and contact pressures based on knee replacement patient data. The Artificial Fish Swarm algorithm optimizes the model's parameters. Extensive experiments demonstrate the method's efficacy in predicting medial knee contact force, outperforming classical multi-body dynamics analysis and artificial neural network models.

Su, Hang, et al. [60] addresses the significance of human arm dynamics in human-robot interaction and object manipulation. It introduces a novel algorithm utilizing deep learning to establish a model between surface electromyography (sEMG) signals and interaction force. Features are automatically extracted using convolutional neural networks from sEMG signals, without prior biomechanical knowledge. Experiments demonstrate the algorithm's efficacy, achieving lower error ($< 0.4, N$) compared to other approaches like artificial neural networks and long short-term memory. Robustness against measurement noises is also evaluated, showing promising results for practical applications. The algorithm's performance is validated using the Myo controller and KUKA LWR4+ robot, highlighting its potential in enhancing dexterity and accuracy in human-robot interaction.

Maniar, Nirav, et al. [61] investigates the role of lower-limb muscles in modulating knee joint loading during single-leg landing tasks, crucial for understanding anterior cruciate ligament (ACL) injury mechanisms. Employing electromyography-informed neuromusculoskeletal modeling, muscle force contributions to anterior shear joint reaction force and valgus joint reaction moment are analyzed. Results show muscles like soleus, medial hamstrings, and biceps femoris primarily oppose anterior shear force, while gluteus

medius, gluteus minimus, and soleus counter valgus moments, crucial for ACL protection. This comprehensive analysis sheds light on key muscle groups involved in mitigating ACL loads, offering insights for injury prevention and rehabilitation strategies.

Zhang, Longbin, et al. [62] explores the estimation of ankle joint torques for robotic powered exoskeletons using electromyography (EMG)-driven neuromusculoskeletal (NMS) and artificial neural network (ANN) models across various movement tasks. EMG signals and ankle joint angles are used as inputs for both models, trained with 3-D motion analysis data. Comparing the models across six calibration/training scenarios, the NMS model generally outperforms the ANN model in predicting ankle joint torques, particularly when EMG peaks are considered. However, the ANN model exhibits better performance when trained with specific movement data. These findings underscore the importance of detailed comparative studies for optimizing exoskeleton control strategies.

Zhu, Meng, et al. [63] addresses the growing use of surface electromyography (sEMG) signals in controlling wearable devices, particularly focusing on aiding uphill movements. It investigates the impact of various sEMG pre-processing methods and algorithms on prediction accuracy. Utilizing an improved PCA algorithm for dimensionality reduction and a novel CNN + LSTM model for knee joint angle prediction, data from ten subjects are collected on a constructed ramp. Results show the proposed method's prediction time is 25 ms with an error of 1.34 ± 0.25 deg, outperforming traditional machine learning methods. The study underscores the efficacy of the proposed approach for accurate lower limb activity recognition, paving the way for enhanced neuro-controlled mechanical exoskeleton development.

Zhao, Yihui, et al. [64] proposes an innovative approach for continuous wrist joint motion estimation using electromyography (EMG) signals, offering potential in assistive robots. Unlike conventional methods, which rely solely on data-driven techniques, this model interprets muscle activation levels from EMG signals to compute muscle force and estimate wrist motion. A genetic algorithm optimizes subject-specific physiological parameters for accurate prediction. Results demonstrate mean root-mean-square-errors of 10.08° to 17.59° and coefficients of determination over 0.9 across various motion trials. The proposed EMG-driven musculoskeletal model showcases precise tracking performance, highlighting its potential for intuitive user-controlled assistive devices.

Martinez, Itzel Jared Rodriguez, et al. [65] addresses the challenge of decoding neurophysiological signals for prosthetic control, focusing on grasping force estimation using forearm electromyography (EMG). High-density EMG signals were recorded during a pick-and-lift task from twelve participants. Linear regression models were employed to estimate grasp force using subsets of EMG channels and features. Results demonstrate

improved prediction accuracy with longer window lengths from onset, reaching a plateau at steady state. Notably, estimations using transient EMG data show comparable accuracy to steady-state data, indicating predictive information about grasp force in the early EMG phase. This finding suggests potential for fast online myoelectric controllers capable of early grasp force decoding.

Moniri, Ahmad, et al. [66] explores the use of adaptive algorithms to forecast surface electromyography (sEMG) features of trunk muscles in real-time, considering the non-stationary nature of these signals. Shallow models and a deep convolutional neural network (CNN) were employed to predict five common sEMG features across 14 trunk muscles in 13 healthy subjects during various exercises. Results indicate the CNN accurately forecasts up to 25 seconds ahead with low mean absolute percentage error and standard deviation, outperforming shallow models significantly. This approach lays the foundation for wearable devices to predict muscle fatigue in real-time, potentially aiding in preventing low back pain and enhancing muscle activity monitoring applications for therapy optimization.

Hajian, Gelareh, Ali Etemad, and Evelyn Morin. [67] investigates EMG-based force estimation across individuals, focusing on biceps brachii and brachioradialis muscles during isometric elbow flexion. Deep convolutional neural networks (CNN) with feature-level fusion of representations from high-density (HD) EMG signals are developed (CNN-FLF). Comparative analysis against various baselines, including input-level fusion CNNs and classical machine learning methods, demonstrates CNN-FLF's superior performance. Optimized hyper-parameters yield a normalized mean squared error of $1.6 \pm 3.69\%$ for estimated force. Visualization of extracted features highlights CNN-FLF's ability to accurately capture the regressive relationship with output force levels. This research provides insights into developing generalized EMG-based force estimation methods with practical applications in prosthetics and rehabilitation.

Mokri, Chiako, Mahdi Bamdad, and Vahid Abolghasemi [32] establishes a framework for processing and evaluating lower limb electromyography (EMG) signals for a rehabilitation robot. A knee rehabilitation robot is designed to utilize surface EMG (sEMG) signals for muscle force estimation employing machine learning techniques like support vector machine (SVM), support vector regression (SVR), and random forest (RF). Genetic algorithm (GA) optimization enhances estimation accuracy alongside feature extraction. Additionally, a load cell and wearable inertial measurement unit (IMU) measure muscle force and knee joint angle. Extensive experiments yield a high estimation accuracy of 98.67% for lower limb muscles, showcasing improved therapy performance. Notably, proposed techniques enhance muscle model sensitivity, particularly to tendon stiffness and

slack length.

Khan, Salman Mohd et al. [68] investigates myoelectric prosthetic devices' capability to classify different force levels in grasping tasks. Electromyography (EMG) and fingertip force signals were acquired during force contraction experiments. A two-step feature selection process, involving ReliefF and Neighborhood Component Analysis (NCA), was applied to design a pattern recognition algorithm. Support Vector Machines (SVM) and Random Forest (RF) classifiers were utilized. Muscle optimization during classification was also performed. Results show a maximum classification accuracy of 99% using SVM with a two-muscle set. Optimal features included Auto Regressive coefficients, Willison Amplitude, and Slope Sign Change. Mean accuracy for different subjects using SVM and RF was 94.5% and 91.7% respectively.

Kamatham, Anne Tryphosa, et al. [69] introduces a noninvasive method using sonomyography (ultrasound imaging) for estimating continuous muscle force. Surface electromyography (sEMG) often lacks specificity and suffers from noise. The proposed approach extracts anatomically relevant features from ultrasound scanlines, reducing computational complexity. Four regression methods are evaluated for force prediction, with Gaussian process regression showing superior performance. Simulations explore practical wearable sonomyography systems, assessing transducer placement and quantity. Results indicate Gaussian process regression's efficacy in predicting force from just four equispaced transducers. These findings offer insights for designing robust and efficient sonomyography-based systems, advancing noninvasive muscle force estimation for clinical and research applications.

Hu, Ruochen, et al. [70] presents a novel myoelectric control scheme supporting gesture recognition and muscle force estimation for robust and natural control systems. Eleven grasping gestures are selected for recognition using high-density surface electromyography (HD-sEMG) and grasping force signals. Multi-task learning (MTL) enables synchronous prediction of gesture category and force level. A post-processing algorithm mitigates force variation's impact on gesture recognition accuracy. Experimental results demonstrate significant error reduction, with a 27-30% decrease compared to no post-processing, and a 16-24% decrease compared to classical methods. The scheme achieves $9.35 \pm 11.48\%$ gesture classification error and 0.1479 ± 0.0436 force estimation accuracy, meeting real-time requirements. This framework holds promise for myoelectric control in prosthetics and exoskeleton devices.

Liang, Wenqi, et al. [71] reviews recent advancements in using IMUs for biomechanical analysis, particularly in activity recognition and musculoskeletal force estimation. It summarizes methodologies including data preprocessing, noise reduction, classification, and

force estimation models. Applications in daily activity assessment, posture evaluation, disease diagnosis, and rehabilitation are discussed. Additionally, the study highlights the potential of IMU-based wearable devices for predicting musculoskeletal forces, emphasizing the importance of adaptive networks and deep learning models for accurate estimation.

Pandy, M.G., et al. [72] developed an optimization model to estimate individual muscle forces during human movement based on recorded kinematics and external forces. The model uses optimization to resolve the muscle redundancy problem, where multiple muscle activation patterns can generate the same movement. Minimizing the sum of muscle stress-raisers was used as the optimization criterion. The resulting model has been widely applied in biomechanics research to analyze muscle coordination in normal and pathological human movements.

2.3 Feature Extraction

Pancholi, Sidharth et al. [73] presents a novel feature extraction method, Advanced Energy Kernel-Based Features (AEKFs), for electromyography (EMG) signal analysis in bionics and prosthesis control. Evaluated on a scientific dataset covering six upper limb motions with varying force levels, AEKFs are compared to existing approaches using classification accuracy (CA), Davies–Bouldin index, and time complexity metrics. Implemented on a DSP processor for real-time viability, AEKFs with Linear Discriminant Analysis (LDA) classifier achieve a CA of 97.33% offline and approximately 92% on the DSP processor. Notably, AEKFs exhibit lower time complexity compared to existing methods, showcasing their potential for efficient and accurate EMG-based control in practical applications.

Ma, Chenfei, et al. [74] addresses the need for improved surface electromyogram (sEMG) feature extraction methods in deep learning-based simultaneous and proportional control (SPC) for rehabilitation robotics. Existing methods struggle with providing adequate information for complex movements across multiple degrees of freedom (DoFs). The proposed novel sEMG feature extraction method aims to enhance accuracy in predicting joint angles simultaneously on ten DoFs. Tested on eight complex movements, the method demonstrates significant improvement compared to raw sEMG data. With an average Pearson correlation coefficient of 0.85, the proposed feature set enhances the application potential of deep learning methods in SPC for rehabilitation robotic systems, with minimal computational overhead.

Tuncer, Turker, Sengul Dogan, and Abdulhamit Subasi [75] proposes a novel ternary

pattern and discrete wavelet (TP-DWT) based iterative feature extraction method for automating prosthetic hand control using surface electromyogram (sEMG) signals and machine learning techniques. The TP-DWT based feature extraction network is utilized for sEMG signal recognition, involving channel concatenation, TP-DWT feature extraction, feature selection, and classification using conventional classifiers. Tested on sEMG data collected from amputee participants across three force levels, the method achieves a high accuracy rate of 99.14% for all force levels with a k-nearest neighbor (k-NN) classifier and 10-fold cross-validation. This approach demonstrates promising potential for effective and accurate prosthetic hand control in real-world scenarios.

Shanmuganathan, Vimal, et al. [76] presents the application of R-CNN for recognizing hand gestures using electromyography-related signals. Forearm electrodes capture signals, processed via wavelet packet transform for feature extraction. R-CNN methodology maps features from the wavelet power spectrum to validate and train the architecture. Real-time testing achieves 96.48% accuracy, surpassing related methods. This demonstrates superior accuracy in gesture recognition, emphasizing the efficacy of the proposed approach.

Pancholi, Sidharth, and Amit M. Joshi [77] introduces advanced time derivative moments (ATDM) as a novel feature extraction technique for myoelectric prosthesis control using EMG signals. Testing on four heterogeneous datasets, including data from amputees, demonstrates the efficacy of ATDM features. Achieving a classification accuracy of 98.32% with lower time complexity, ATDM features are compared favorably to similar methods. Evaluation using Davies–Bouldin (DB) index confirms excellent feature separability across classes. Results underscore the potential of ATDM features for enhanced pattern recognition in myoelectric prosthetic control, highlighting their suitability for real-world applications.

Baygin, Mehmet, et al. [78] introduces Frustum154, a hybrid feature extraction method for hand-movement surface electromyography (sEMG) datasets. It combines shape-based local feature extraction with textural and statistical features, creating 154 low-level feature vectors. Tunable Q factor wavelet transform (TQWT) generates high-level features. Iterative neighborhood component analysis (INCA) selects optimal feature vectors, improving classification accuracy. Frustum154 achieves high accuracies of 98.89%, 94.94%, and 95.30% using shallow classifiers on three sEMG datasets. The model is self-organizing and automatically selects informative features, demonstrating its potential to enhance sEMG-based classification in biomedical engineering.

Gesture recognition using surface electromyography (sEMG) faces challenges in generalization across different positions and users. Shen, Cheng, et al. [79] introduces a novel

feature extraction method directly from the time domain, mitigating position influence on sEMG-based pattern recognition. Offline cross-validation accuracy reaches 88.775%, and online cross-validation accuracy achieves 99%, with a movement selection time of 47.036 ± 1.028 ms. The proposed features significantly improve classification accuracy and generalization compared to traditional methods, aiding the practical implementation of myoelectric interfaces.

Li, Gongfa, et al. [80] proposes a novel sEMG feature extraction method called active muscle regions (AMR) for controlling multifunctional prostheses. Experimental results demonstrate that AMR outperforms traditional features like MAV, WL, ZC, and SSC, with average classification errors of 13%, 19%, 26%, 24%, and 22%, respectively. AMR leverages the mapping relationship between hand movements and forearm muscle activity, validated in medicine. The smaller feature vector size of AMR reduces computational costs, highlighting its potential to enhance sEMG pattern recognition accuracy for prosthesis control.

Aziz, Sumair, et al. [81] proposes a methodology for load classification using surface electromyography (sEMG). Subjects performed weight lifting tasks with 1kg, 3kg, and 7kg weights, yielding 50 samples each. sEMG signals were denoised via Empirical Mode Decomposition (EMD), and region of interest extraction improved classification by removing redundant information. Feature extraction involved statistical, time, frequency domain, and Mel Frequency Cepstral Coefficient (MFCC) analysis. The best feature set was selected based on highest accuracy, validated using 10-fold cross-validation with SVM, DT, QDA, NB, and K-nearest neighbor classifiers. SVM with cubic kernel achieved the highest classification accuracy of 99% for the three-class problem.

2.4 TimeGAN - Generative Adversarial Network

TimeGAN, a Generative Adversarial Network (GAN) variant, addresses time series data generation by capturing temporal dependencies and distributions. Unlike traditional GANs, it employs a novel architecture comprising two distinct modules: a generator and a discriminator. The generator leverages a recurrent neural network (RNN) to model time dependencies, while the discriminator discerns between real and generated sequences. By optimizing the generator to produce sequences indistinguishable from real data and leveraging Wasserstein distance, TimeGAN effectively learns and replicates complex time series distributions. It finds applications in various domains like finance, healthcare, and climate science, enabling realistic synthetic data generation for diverse analyses.

Carneros-Prado, David, et al. [82] highlights the role of GANs in augmenting datasets,

aiding in better understanding and analysis of gait dynamics. Research on gait analysis has surged recently, crucial for early frailty detection. Gathering data is resource-intensive, prompting interest in synthetic data generation as a complementary tool. Generative Adversarial Networks (GANs) have popularized this approach. Deep-convolutional GANs expanded a gait dataset, with evaluation via animations, principal component analysis, and feature extraction. Results indicate potential for synthetic data in gait analysis.

Yun, Rachel, May Salama, and Lamiaa Elrefaei [83] investigates a deep neural network (DNN) framework’s efficacy in human gait analysis for injury classification based solely on ground reaction force plate measurements. Three DNN architectures (MLP, FCN, ResNet) are explored against shallow neural networks and unsupervised clustering models. Results show DNNs outperform clustering models and match or surpass support vector machines despite overfitting risks. Data augmentation via timeGAN doesn’t enhance accuracy. Incorporating explicit feature extraction like PCA preprocessing improves DNN performance, particularly when coupled with dropout mechanisms, suggesting its value in future DNN implementations.

Xu, Chenchu, et al. [84] introduces progressive sequential causal GANs (PSCGAN) as a one-stop solution for ischemic heart disease (IHD) diagnosis without gadolinium contrast agent (CA) injections or manual segmentation. PSCGAN synthesizes LGE-equivalent images and segments relevant tissues from cine MR images using a progressive framework, sequential causal learning network (SCLN), and self-learning loss terms. Results from 280 clinical subjects demonstrate high accuracy and correlation coefficients, showcasing PSCGAN’s potential for standardized cardiac disease assessment.

Bicer, Metin, et al. [85] proposes a data augmentation method for deep learning biomechanical models using generative adversarial networks. The approach, termed adversarial autoencoder, generates synthetic motion capture datasets for marker trajectories and ground reaction forces. Comparison of real and synthetic data using statistical parametric mapping revealed negligible differences in direct kinematics and GRFs, but significant disparities in inverse methods. Integration of joint angles calculated by inverse kinematics improved joint moments estimation. The augmentation approach enhanced predictive accuracy of joint kinematics and vertical GRFs, indicating its potential for improving deep learning models.

Yin, Minglang, et al. [86] integrates generative deep learning and Bayesian inference to infer constitutive relationships efficiently in data-sparse settings. Using a neural operator and discriminator within a GAN, this framework accurately estimates vascular properties, even with limited data. It learns model priors without explicit knowledge of functional

forms, offering a versatile, model-agnostic approach to understanding biomechanical behaviors.

Overall, TimeGAN represents a significant advancement in the field of time series generation, offering a powerful tool for generating realistic and diverse synthetic time series data. Its innovative architecture and training methodology make it well-suited for a wide range of applications where generating synthetic time series data is crucial.

2.5 Anomaly Detection Techniques

Anomaly detection techniques are pivotal in identifying unusual patterns or outliers in data, crucial across various domains such as finance, cybersecurity, and healthcare. Traditional methods include statistical approaches like clustering, deviation-based methods, or rule-based systems. However, with the advent of machine learning, more sophisticated techniques like autoencoders, isolation forests, and one-class support vector machines have emerged. These methods excel in detecting anomalies in high-dimensional data by learning patterns from normal instances and flagging deviations. Deep learning approaches, particularly recurrent neural networks and generative adversarial networks, offer advanced anomaly detection capabilities, especially in sequential and unstructured data, providing robust solutions in complex and dynamic environments.

Mullineaux, David R., and Gareth Irwin [87] developed for intra-participant time-series data from treadmill running, detects outliers in two stages. Firstly, spatial outliers are removed using median absolute deviation, followed by the removal of spatial-temporal outliers using a moving window standard deviation. Customizable settings in Matlab code allow for tailored application to different datasets, ensuring effective outlier identification.

Imbalanced datasets pose significant challenges in classifying medical data, particularly in analyzing knee abnormalities during walking. Surface electromyography (sEMG) data from 22 individuals show longer signal lengths for subjects with knee abnormalities, creating an imbalance in the dataset. Vijayvargiya, Ankit, et al. [88] proposes a multi-step classification approach, incorporating wavelet denoising and oversampling techniques to address this issue. Evaluation with various classifiers demonstrates that oversampling enhances performance on imbalanced sEMG data, offering potential for more effective automated analysis of complex sEMG signals in medical applications.

Rashid, Usman, et al. [89] introduces an algorithm using acceleration derivatives to automatically identify and replace invalid force data with marker-detected events. Evaluation

on healthy participants' data shows the algorithm's strong performance, with an area under the curve (AUC) score exceeding 0.8 and no significant correlation with walking speed. The algorithm's MATLAB implementation and example data files are provided, offering a valuable tool for gait analysis using instrumented split-belt treadmills.

Jin, Zeqing, et al. [90] introduces an anomaly detection system using layer-by-layer sensor images and machine learning to identify imperfections in transparent hydrogel-based bioprinted materials. Utilizing convolutional neural networks and advanced image processing, the system achieves high accuracy in anomaly detection, enabling real-time correction of process parameters for high-quality tissue constructs in 3D bioprinting.

Sivaraks, Haemwaan, et al. [91] introduces a novel anomaly detection technique leveraging cardiologist expertise and motif discovery. The algorithm ensures robustness and accuracy even in the presence of ECG artifacts, significantly reducing false alarms. Evaluation on real ECG datasets confirms high accuracy, sensitivity, specificity, and positive predictive value, making it superior to existing methods in detecting anomalies with minimal false alarms.

Alkahtani, Hasan, Theyazn HH Aldhyani, and Mohammed Al-Yaari [92] proposes an adaptive anomaly detection framework using deep and machine learning algorithms to enhance cybersecurity. The model, utilizing Long-Short Term Memory Recurrent Neural Network (LSTM-RNN), Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN) algorithms, effectively classifies Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS) attacks. Evaluation on standard network datasets demonstrates LSTM-RNN's superior accuracy, indicating its efficacy in cybersecurity enhancement.

Anomaly detection techniques play a crucial role in biomechanics, enabling researchers and clinicians to identify deviations from normal physiological patterns and diagnose musculoskeletal disorders effectively. By leveraging statistical, machine learning, and computer vision approaches, these techniques provide valuable insights into human movement, optimizing performance, rehabilitation, and injury prevention strategies. As technology continues to evolve, the integration of anomaly detection techniques into biomechanical research and clinical practice holds promise for enhancing understanding and treatment of musculoskeletal conditions, ultimately improving quality of life for individuals worldwide.

2.6 Machine Learning Techniques

Machine learning techniques are revolutionizing industries by enabling computers to learn from data and make predictions or decisions without explicit programming. These methods, rooted in artificial intelligence, encompass a wide range of algorithms that automatically detect patterns and extract insights from complex datasets. Supervised learning algorithms learn from labeled data, while unsupervised learning algorithms identify patterns in unlabeled data. Reinforcement learning algorithms learn through trial and error interactions with an environment. From predictive analytics to image recognition and natural language processing, machine learning techniques power innovative solutions across domains, driving advancements in technology, healthcare, finance, and beyond.

There are different types of machine learning algorithm. These are described below:

Supervised Learning: Supervised learning is a machine learning technique where algorithms learn from labeled data. Models are trained to map inputs to known outputs, with the goal of making accurate predictions on new data. This approach is used for classification and regression tasks. The learning process involves adjusting model parameters to minimize prediction errors. Common algorithms include linear regression, decision trees, and neural networks. Data quality and feature selection significantly impact model performance. (Figure 2.1)

Unsupervised Learning: Unsupervised learning is a machine learning technique where algorithms analyze unlabeled data to discover hidden patterns or structures. Unlike supervised learning, there are no predefined output labels. Common tasks include clustering, dimensionality reduction, and anomaly detection. Algorithms like K-means, hierarchical clustering, and principal component analysis are widely used. Unsupervised learning is valuable for exploratory data analysis, feature learning, and uncovering complex relationships within datasets without prior knowledge of specific outcomes. (Figure 2.2)

Semi-Supervised Learning: Semi-Supervised learning is a machine learning approach that combines elements of supervised and unsupervised learning. It utilizes both labeled and unlabeled data for training models. This technique is particularly useful when obtaining labeled data is expensive or time-consuming. Algorithms leverage the structure in unlabeled data to improve learning from a small set of labeled examples. Common methods include self-training, co-training, and graph-based algorithms. Semisupervised learning can enhance model performance and generalization, especially in scenarios with limited labeled data. (Figure 2.3)

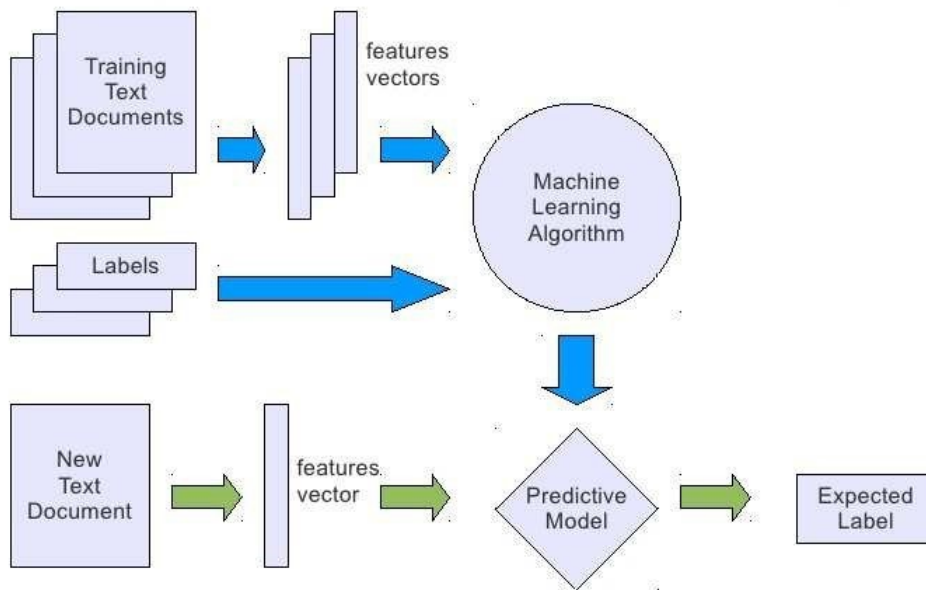


Figure 2.1: Supervised Learning

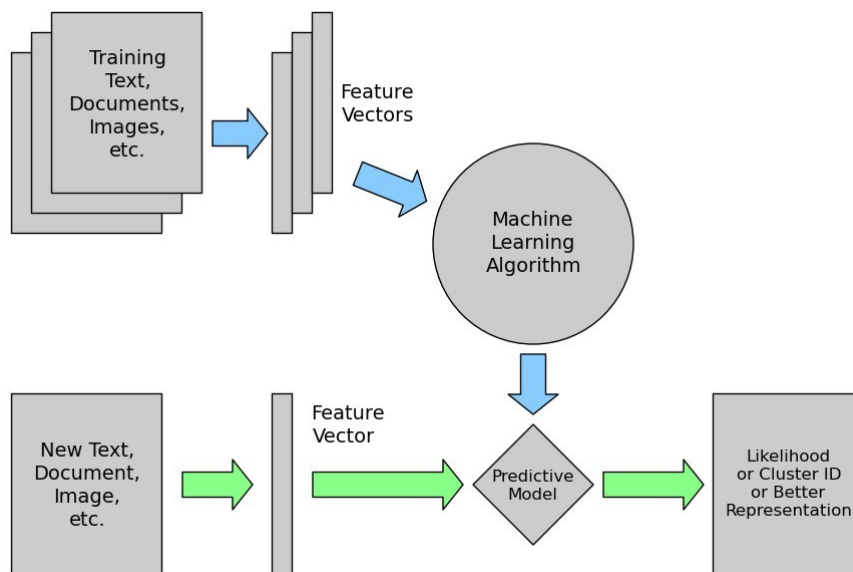


Figure 2.2: Unsupervised Learning

2.6.1 Decision Tree

Decision Trees are valuable tools in biomechanics for their ability to analyze complex datasets and extract meaningful insights. They are frequently used in various biome-

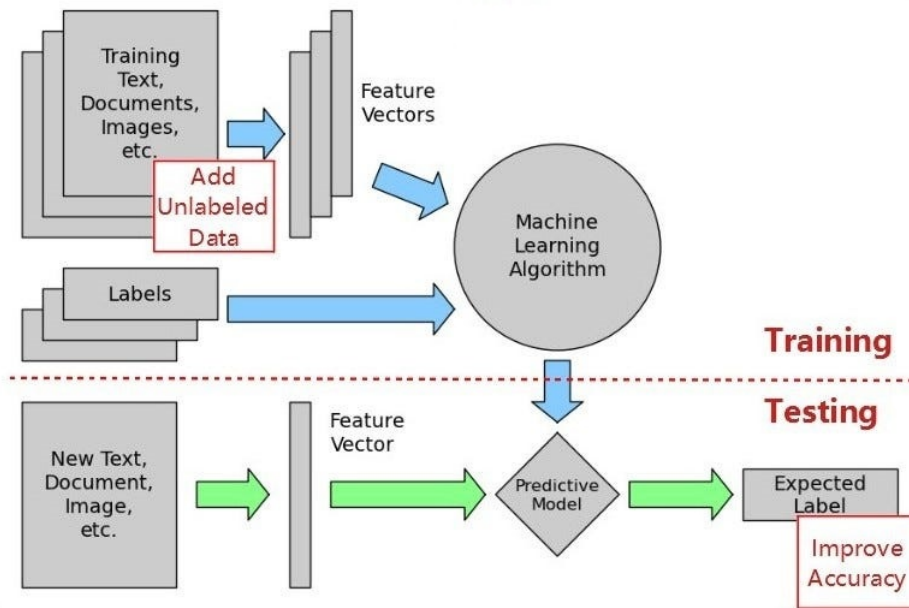


Figure 2.3: Semi-Supervised Learning

chanical applications, including gait analysis, injury prediction, and rehabilitation planning.

In gait analysis, decision trees have been employed to classify different gait patterns based on kinematic and kinetic data obtained from motion capture systems [93, 94]. Additionally, decision trees have been utilized to predict injury risk in athletes by analyzing biomechanical factors such as joint angles and ground reaction forces [95]. Moreover, decision trees have been applied to develop personalized rehabilitation plans by identifying key factors influencing recovery outcomes [96].

The interpretability and simplicity of decision trees make them particularly advantageous in biomechanics, allowing researchers and clinicians to easily understand and interpret the decision-making process. Decision trees also offer insights into the underlying relationships between biomechanical variables and outcomes, facilitating evidence-based decision-making in musculoskeletal health.

Overall, decision trees are valuable tools in biomechanics research, enabling the extraction of meaningful insights from complex biomechanical datasets and aiding in the development of personalized interventions for injury prevention and rehabilitation.

2.6.2 Support Vector Machine

Support Vector Machines (SVMs) are widely employed in biomechanics due to their effectiveness in classification and regression tasks with high-dimensional data. SVMs

have been utilized in various biomechanical applications, including gait analysis, muscle activity classification, and movement prediction.

In gait analysis, SVMs have been employed to classify different gait patterns based on features extracted from wearable sensors [97]. Additionally, SVMs have been used to classify muscle activation patterns from electromyography (EMG) signals, aiding in the assessment of muscle function during movement [98]. Moreover, SVMs have been applied to predict joint kinematics and kinetics from biomechanical data, providing valuable insights into movement dynamics and biomechanical performance [99].

The robustness and generalization capability of SVMs make them well-suited for handling complex biomechanical datasets, which often contain noise and variability. By effectively separating classes and capturing nonlinear relationships, SVMs contribute to the advancement of biomechanics research and facilitate clinical decision-making in musculoskeletal health.

2.6.3 Bayes Rule

Bayes' Rule, a fundamental theorem in probability theory, holds significant relevance in biomechanics for its ability to quantify uncertainty and update beliefs based on new evidence. In biomechanics, where complex interactions between biological systems and mechanical forces occur, Bayes' Rule provides a robust framework for understanding and modeling these relationships. By incorporating prior knowledge and updating it with observed data, researchers can make informed decisions and infer causal relationships in biomechanical systems.

Applications of Bayes' Rule in biomechanics are diverse and widespread. For instance, it has been utilized in estimating joint kinematics from inertial sensors [100, 101], predicting muscle forces during movement [102], and diagnosing movement disorders based on symptom observations [103]. These applications demonstrate the versatility of Bayes' Rule in addressing various challenges in biomechanics research.

Moreover, the Bayesian approach offers several advantages, including the ability to handle uncertainty, incorporate prior knowledge, and provide probabilistic predictions. This is particularly valuable in biomechanics, where data can be noisy and incomplete, and uncertainty is inherent in biological systems.

Overall, Bayes' Rule serves as a powerful tool in biomechanics, enabling researchers to model and analyze complex systems, make predictions, and guide decision-making processes with a probabilistic framework.

2.6.4 Artificial Neural Networks

Artificial Neural Networks (ANNs) have emerged as indispensable tools in biomechanics, revolutionizing the analysis of human movement patterns and musculoskeletal dynamics. ANNs, inspired by the structure and function of the human brain, excel in tasks such as gait analysis, motion prediction, and rehabilitation planning by efficiently processing complex biomechanical data.

The architecture of ANNs consists of interconnected nodes organized in layers, enabling them to learn from large datasets and capture intricate relationships within biomechanical variables. By leveraging this capability, ANNs can accurately classify movements, predict future motion trajectories, and optimize interventions for musculoskeletal health.

Numerous studies have demonstrated the efficacy of ANNs in various biomechanics applications. For instance, ANNs have been successfully employed to predict gait parameters based on wearable sensor data [104], classify muscle activation patterns from electromyography signals [105], and model musculoskeletal dynamics for exoskeleton control [106].

The versatility and effectiveness of ANNs make them indispensable tools in biomechanics research, offering insights into human movement dynamics and facilitating advancements in clinical practice.

2.7 Time Series Prediction Techniques

Time series prediction techniques are essential in analyzing and forecasting sequential data in various fields, including biomechanics. These methods, ranging from traditional statistical models to advanced machine learning algorithms like recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks, capture temporal dependencies and patterns within data. By leveraging past observations, time series prediction techniques enable accurate predictions of future trends, facilitating informed decision-making in biomechanical research, such as gait analysis, injury prediction, and rehabilitation planning, to optimize musculoskeletal health and performance outcomes.

2.7.1 ARIMA

ARIMA, introduced by Box and Jenkins (1976), is a parametric model that combines autoregressive (AR), differencing (I), and moving average (MA) components to capture the linear dependencies and trends in the data. ARIMA models have been extensively applied in practice due to their simplicity and effectiveness in handling stationary time

series data. However, ARIMA assumes that the data is stationary and requires manual identification of model parameters, which can be challenging and time-consuming.

Satrio, Christophorus Beneditto Aditya, et al. [107] employs Machine Learning models to forecast the COVID-19 pandemic trend in Indonesia and estimate the return to normalcy. Utilizing Facebook’s Prophet and ARIMA models, confirmed cases, deaths, and recoveries data from Kaggle are analyzed. Results indicate Prophet generally outperforms ARIMA, though it tends to deviate further from actual data with longer forecast periods.

Duarte, Diego, and Julio Faerman [108] compares the ARIMA method with Prophet, an autoregressive model, to predict KPIs. Utilizing hourly hospital data including Wait to be Seen Major, Attendances Major, Unallocated Patients, and Available Beds, each model’s performance is evaluated based on accuracy and reliability. Results identify the most suitable model for different time series characteristics, aiding in effective prediction of future pressure.

Saveljić, Slavica Mačuzić [109] investigates the impact of vehicle vibrations on driver comfort, focusing on the transmission response function (STHT) between the seat and head. Utilizing experimental data from twenty male subjects exposed to vertical vibrations, an artificial neural network (ANN) model is developed to predict transfer function values. Results demonstrate the ANN’s capability to accurately forecast transfer function values within the trained parameter range, offering insights into the relationship between seat vibrations and head motion responses for improved comfort analysis.

2.7.2 Prophet

Prophet, developed by Facebook’s Data Science team, is a more recent addition to the forecasting toolkit. Prophet is designed specifically for time series data with strong seasonal patterns, such as those commonly encountered in business applications. It offers several advantages over traditional models, including automatic detection of changepoints, the ability to handle missing data, and intuitive parameters for modeling holidays and seasonality.

Hyun, Jayun, et al. [110] proposes a specialized model for diabetic foot treatment systems. This system utilizes statistical methods and the Prophet model by Facebook to synthesize data accurately while maintaining medical validity. The study suggests systemic methods to evaluate the data’s validity. Results demonstrate that the synthetic time-series data aligns with real data trends, showcasing the potential of this approach in addressing the scarcity of realistic medical data for advancing digital therapeutics in healthcare.

Lu, Wei, et al. [111] aimed to establish a robust relationship between electromyography

and flexion force, recognizing intended movements. However, existing methods require biomechanical models to ensure precise predictions. Elbow flexion force relies on muscle contractile properties, necessitating thorough analysis. By decomposing electromyography signals into non-linear and non-stationary components, this study used mean absolute value and variance as inputs for an Informer prediction model. The proposed framework effectively predicts long-term flexion force, outperforming state-of-the-art models in experimental comparisons.

2.7.3 RNN

Recurrent Neural Networks (RNNs) are invaluable tools in biomechanics, offering sophisticated capabilities in analyzing time-series data and capturing temporal dependencies inherent in human movement patterns. RNNs, designed with loops that allow information to persist over time, excel in tasks such as gait analysis, motion prediction, and rehabilitation planning.

In biomechanics, RNNs have been extensively applied to various domains. For instance, they have been employed to predict gait patterns based on sensor data [112], analyze electromyography signals for muscle activity recognition [113], and model motor control mechanisms [114].

The architecture of RNNs enables them to process sequential data efficiently by retaining memory of past inputs. This allows them to capture subtle temporal dynamics and dependencies, facilitating accurate predictions and insights into biomechanical phenomena.

Despite their efficacy, RNNs may suffer from issues like vanishing gradients and difficulty in capturing long-term dependencies. To address these challenges, advanced variants like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks have been developed, offering improved performance in modeling long-range dependencies and mitigating gradient vanishing problems.

Overall, RNNs play a crucial role in advancing our understanding of human movement and optimizing interventions for musculoskeletal health, offering promising avenues for future research and clinical applications.

2.7.4 LSTM

Long Short-Term Memory (LSTM) networks are pivotal in biomechanics, especially for analyzing time-series data and modeling complex human movement patterns. LSTM

networks, a type of recurrent neural network (RNN), are specifically designed to overcome the limitations of traditional RNNs in capturing long-term dependencies in sequential data. This makes them highly effective in tasks such as gait analysis, motion prediction, and rehabilitation planning, where understanding temporal dynamics is crucial.

The architecture of LSTM networks includes memory cells that retain information over long sequences, allowing them to capture subtle temporal patterns and dynamics inherent in biomechanical data. LSTM networks have been successfully applied in various biomechanics applications, including human motion analysis, gesture recognition, and motor control modeling.

Studies have demonstrated the effectiveness of LSTM networks in predicting gait patterns [115], analyzing electromyography signals for muscle activity recognition [116], and assisting in rehabilitation exercises [117]. The ability of LSTM networks to learn from sequential data and model complex temporal relationships makes them valuable tools for understanding human movement and optimizing interventions for musculoskeletal health.

Chapter 3

Decoding Human Movement: Musculoskeletal Modeling and Sensor Technology

Human biomechanics explores human movement, analyzing the interaction between the musculoskeletal system and external forces to understand motion generation and control. This field is vital in rehabilitation, sports science, ergonomics, and robotics. At its core is musculoskeletal modeling, a computational approach replicating the human body's structure and function. These models, combined with advanced sensors, offer valuable insights into muscle forces, joint moments, and overall movement patterns.

Musculoskeletal modeling comprises rigid-body models and deformable models [118]. Rigid-body models depict bones as segments with fixed lengths, connected by joints with defined degrees of freedom. They are computationally efficient, suitable for analyzing gross motions like walking or running. Deformable models consider the flexibility of muscles, tendons, and ligaments, enabling detailed analysis of forces within muscle groups and joints. However, they entail more complex calculations and computational resources.

To bridge the gap between theoretical models and real-world movement, various sensors play a vital role in capturing biomechanical data. Inertial Measurement Units (IMUs) are commonly used to track body segment accelerations and orientations [119]. By strategically placing IMUs on different body segments, researchers can reconstruct 3D movement patterns. However, IMUs cannot directly measure muscle forces. This is where force sensors come into play.

Emerging technologies like electromyography (EMG) offer additional insights into muscle activity. EMG measures the electrical activity of muscles, providing an indirect indication of muscle force production. By combining EMG data with musculoskeletal models and force sensor measurements, researchers can create a more comprehensive picture of neuromuscular control during movement [120].

The field of biomechanics also plays a critical role in the development of prosthetics and orthotics [121]. Musculoskeletal models can be used to design prosthetic limbs that mimic the movement patterns of natural limbs. Sensors embedded within these devices

can provide feedback on residual limb forces and prosthetic-socket interaction, leading to more comfortable and functional prosthetics.

However, biomechanics also faces certain challenges. Creating accurate musculoskeletal models remains an ongoing effort, as individual anatomical variations and muscle activation patterns can significantly impact movement [5]. Additionally, sensor data can be noisy and require careful processing and filtering. Furthermore, integrating data from various sensors and models into a coherent framework remains an active area of research.

Despite these challenges, biomechanics continues to evolve rapidly. As sensor technology advances, miniaturization and wireless capabilities will allow for more unobtrusive and comprehensive data collection during movement. Additionally, the rise of machine learning and artificial intelligence holds immense potential for analyzing biomechanical data and extracting meaningful insights.

This chapter delves into the science of human movement, exploring how muscles, bones, and sensors work together to decode the complexities of human motion. By combining musculoskeletal modeling with advanced sensors, we gain valuable insights into how we move, paving the way for advancements in rehabilitation, sports performance, and prosthetics.

3.1 Musculoskeletal Modeling

Musculoskeletal modeling is a critical component of human biomechanics, as it enables researchers and practitioners to understand and predict the behavior of the musculoskeletal system during various activities. These models are typically computational representations of the human body's skeletal structure, muscles, and joints. They are used to simulate and analyze the forces, moments, and kinematics involved in human movement.

3.1.1 Musculoskeletal Modeling Tools

Musculoskeletal modeling tools are specialized software programs that create virtual representations of the human body's musculoskeletal system. These models allow researchers, clinicians, and engineers to simulate movement, analyze forces acting on muscles and bones, design prosthetics, and ultimately gain a deeper understanding of human biomechanics.

Here's a table outlining some of the most prominent musculoskeletal modeling tools, their applications, and a brief description:

Table 3.1: Musculoskeletal Modeling Tools

Tool	Description	Functionalities	Applications
OpenSim [122, 123]	An open-source software for developing and analyzing musculoskeletal models.	Inverse kinematics, inverse dynamics, static optimization, forward dynamics simulations, muscle analysis, and more.	Biomechanics research, orthopaedic surgery planning, rehabilitation, sports science, and ergonomics.
AnyBody [124, 125]	A commercial software for musculoskeletal modeling and simulation.	Inverse dynamics, forward dynamics, muscle recruitment, contact modeling, and more.	Ergonomics, sports biomechanics, occupational biomechanics, and orthopedics.
SIMM [126, 127]	A commercial software for musculoskeletal modeling and simulation.	Inverse kinematics, inverse dynamics, forward dynamics, muscle analysis, and more.	Biomechanics research, orthopaedic surgery planning, and sports science.
MusculoSkeletal Modeling for Biomechanics (MSMS) [128, 129]	A MATLAB-based toolkit for musculoskeletal modeling and simulation.	Inverse dynamics, forward dynamics, muscle analysis, and more.	Biomechanics research, orthopaedic surgery planning, and sports science.
LifeModeler [130, 131]	Commercial software for musculoskeletal modeling and simulation, developed by LifeModeler, Inc.	Inverse kinematics, inverse dynamics, forward dynamics, muscle analysis, and more.	Biomechanics research, orthopaedic surgery planning, sports science, ergonomics, and rehabilitation.

Table 3.1: Musculoskeletal Modeling Tools

Tool	Description	Functionalities	Applications
Neuromusculoskeletal Modeling (NMBL) [131, 132]	Open-source software for neuro-musculoskeletal modeling and simulation, developed by the National Center for Simulation in Rehabilitation Research (NCSRR).	Inverse kinematics, inverse dynamics, forward dynamics, muscle analysis, and more.	Biomechanics research, orthopaedic surgery planning, sports science, ergonomics, and rehabilitation.
AMMR (An Anatomical Model Representation) [133]	A software toolkit for creating and analyzing musculoskeletal models, developed by the Biomechanical Engineering Research Group (BERG) at the University of Colorado.	Musculoskeletal model creation, inverse kinematics, inverse dynamics, forward dynamics, and more.	Biomechanics research, orthopaedic surgery planning, sports science, and ergonomics.
GaitSym [134]	Free, open-source software for 2D and 3D musculoskeletal modelling.	<ul style="list-style-type: none"> - Create simplified 2D/3D models. - Analyze gait patterns and muscle forces. 	<ul style="list-style-type: none"> - Education and teaching biomechanics. - Clinical gait analysis (preliminary). - Assistive device design (initial stages).
AMESim [135, 136]	Commercial software for multi-domain simulation, including musculoskeletal modelling.	<ul style="list-style-type: none"> - Model and simulate complex biomechanical systems. - Integrate musculoskeletal models with other systems (e.g., nervous system). - Analyze and optimize system performance. 	<ul style="list-style-type: none"> - Design and development of rehabilitation robots. - Research in human-machine interaction. - Development of biomimetic robots.

Table 3.1: Musculoskeletal Modeling Tools

Tool	Description	Functionalities	Applications
Visual3D [137, 138]	Specialized software for analyzing human gait.	- Import motion capture data. - Analyze joint kinematics and kinetics. - Simulate gait patterns and interventions.	- Clinical gait analysis for gait disorders. - Design and evaluation of orthotics and prosthetics. - Research in sports biomechanics.
ABAQUS [139]	General-purpose engineering software used for musculoskeletal modelling.	- Simulate complex bone and tissue behaviour. - Analyze stress and strain distribution. - Model implants and surgical procedures.	- Research in bone mechanics and fracture. - Design and optimization of orthopedic implants. - Pre-surgical planning and simulation.
GAIT2392 [140, 141]	An open-source software specifically for 2D gait analysis.	- Analyze 2D gait data from motion capture systems. - Calculate joint kinematics and kinetics.	- Clinical gait analysis for patients with gait disorders. - Research in gait biomechanics.
Adams (MSC Software) [142]	Commercial software for multi-body dynamics and musculoskeletal modeling	Kinematics, dynamics, muscle force computation, forward/inverse simulations, optimization	Biomechanics research, movement analysis, ergonomics, product design

3.1.2 Related Work

Falisse A et al. [143] discusses the use of musculoskeletal modeling and predictive simulations to analyze the neuromechanical control of human gait. It highlights the potential of these computational approaches to provide insights into the neural strategies employed by the central nervous system to coordinate muscle activations during walking. The authors emphasize the importance of validating simulations against experimental data and incorporating subject-specific musculoskeletal parameters. They advocate for using predictive simulations to investigate impaired gait, design assistive devices, and optimize rehabilitation strategies. The paper underscores the value of interdisciplinary collaborations between biomechanists, neuroscientists, and clinicians to advance our understanding of human locomotion.

McFarland DC et al. [144] introduces a comprehensive musculoskeletal model of the human hand and wrist for simulating functional grasping and manipulation tasks. Incorporating 24 degrees of freedom and 38 muscle-tendon units, the model captures complex hand biomechanics. The authors validated it against experimental data and demonstrated simulations of activities like cylindrical grasping and finger tapping. The biomechanically realistic model enables studying hand function, pathologies, and interventions, with applications in ergonomics, prosthetics, surgery planning, and rehabilitation.

Wang, Huawei, et al. [145] introduces MyoSim, a framework for creating fast and physiologically realistic musculoskeletal models in MuJoCo. MyoSim models incorporate muscle dynamics, activation dynamics, and contraction dynamics based on biomechanical principles. The authors demonstrate MyoSim's capabilities through models of the upper limb and lower limb. These models enable efficient simulations for musculoskeletal and exoskeletal studies while capturing realistic muscle behavior and interactions between the musculoskeletal system and external devices. MyoSim facilitates biomechanics research, rehabilitation technology development, and human-robot interaction analysis.

Manuel Cardona et al. [146] highlights the utility of musculoskeletal models in addressing complex biomechanical issues, injury simulations, and orthopedic surgery predictions. It presents an open-source full musculoskeletal model for analyzing human lower limb biomechanics, featuring 14 bones, 88 muscle-tendon segments, and ten ligament segments per knee. The model simulates various lower limb injuries using signal profiles akin to functional electrical stimulation, enabling the computation of forces and torques for muscles and joints. Validation via Simulink and gait capture systems confirms its accuracy in estimating kinematics and kinetics parameters for both normal and pathological gait, making it suitable for medical analysis and exoskeleton design.

Burkhart, Katelyn, et al. [147] assessed the reliability of subject-specific musculoskeletal models, derived from optoelectronic motion capture data, to estimate spine loading across different postures. Nineteen healthy participants underwent measurements on two occasions. Models, incorporating detailed thoracolumbar spines, were created and scaled for each individual using marker-based measurements. Vertebral compressive loading was estimated for various postures. Results indicated excellent reliability for spine curvature ($ICC = 0.79\text{--}0.91$) and fair to excellent reliability for body scaling segments ($ICC = 0.46\text{--}0.95$). Overall, musculoskeletal models exhibited high reliability (91% $ICC \geq 0.75$) in estimating spine loading across activities. These findings underscore the feasibility of using motion capture data for reproducible subject-specific models and spine loading estimates.

Thelen, D.G., et al. [148] utilized musculoskeletal modeling and dynamic simulations to estimate individual muscle activations during normal walking. A 3D musculoskeletal model with 54 muscle-tendon actuators was created and driven by experimental gait data. The simulations computed muscle activations required to generate the measured motion, providing insights into muscle coordination strategies during gait. The results highlighted the important roles of biarticular muscles and showed how muscles are selectively recruited based on their mechanical output.

Delp, S.L., et al. [122] introduced OpenSim, an open-source software system for developing musculoskeletal models and simulating human movement dynamics. It described the software architecture, modeling tools, and simulation capabilities. OpenSim allows users to build and analyze musculoskeletal models from motion capture data, quantify muscle forces, and simulate movements like walking and running. It has become a widely used platform for biomechanics research by enabling model sharing and collaborative model development.

Andriacchi, T.P., et al. [149] introduced the point cluster technique (PCT) for determining 3D kinematics of bones from skin-mounted marker data during movement trials. Compared to conventional approaches, PCT improved accuracy by accounting for relative marker motion on the skin. It was successfully applied to analyze knee joint rotations during activities like stair climbing.

Delp, S.L., et al. [150] reviews recent advancements in musculoskeletal modeling and simulation techniques for predicting human movement. Through computational models, researchers simulate muscle activations, joint forces, and kinematics to understand complex biomechanical interactions. The paper discusses applications in clinical settings, sports performance analysis, and rehabilitation protocols.

Arjmand, N., et al. [151] presents a detailed musculoskeletal model of the lumbar spine, consisting of six rigid bodies (five vertebrae and the pelvis) and 238 muscle fascicles. The model was used to investigate the biomechanics of various trunk movements and the effects of external loads on muscle forces and spinal loads. The authors demonstrated the model's potential for studying low back pain and developing preventive measures.

Jing Z et al. [152] investigated the effectiveness of different musculoskeletal models in assessing biomechanical data in children with cerebral palsy. Accurate biomechanical analysis is crucial for effective treatment strategies due to the condition's significant impact on movement and posture. The study evaluates how various models represent movement patterns, muscle activation, and joint kinetics within this specific population to discern the most reliable modeling approaches. These findings underscore the importance of selecting appropriate models in clinical contexts to enhance treatment outcomes for children with cerebral palsy.

3.2 Biomechanics Sensor Technology

Biomechanics sensor technology involves the use of wearable devices to monitor and analyze human movement and physical performance. This enables comprehensive biomechanical analyses, aiding in injury prevention, technique optimization, and performance enhancement across sports, rehabilitation, ergonomics, and other applications. Advancements in sensor miniaturization, wireless connectivity, and data analytics have propelled the field, allowing for unobtrusive, real-time monitoring and personalized insights. Biomechanics sensor technology continues to reshape our understanding of human biomechanics and its practical applications.

3.2.1 Biomechanics Sensors

Biomechanics sensors are like fitness trackers on steroids, delving deeper than steps and heart rate. These tiny gadgets, often worn on the body, transform into invisible spies. They measure not just movement, but its intricacies. This data becomes a goldmine for fields like sports medicine, where it helps optimize training and prevent injuries. Physical therapists can use it to monitor recovery progress and tailor therapy. Even healthcare professionals can leverage it to detect movement disorders or balance issues. As biomechanics sensors continue to develop, they promise to revolutionize how we understand and optimize human movement, impacting health, rehabilitation, and athletic performance for years to come.

Here is a table summarizing some common types of biomechanics sensors.

Table 3.2: Biomechanics Sensors

Sensor	Description	Functionalities	Applications
Force Plates [153, 154]	Measure the ground reaction forces generated by a body standing or moving across them.	Measure forces and moments in three orthogonal directions.	Gait analysis, sports biomechanics, ergonomics, balance assessment.
Motion Capture Systems [155, 156]	Track the movement of body segments or markers attached to the body using cameras or other sensors.	Provide 3D position and orientation data of markers or body segments.	Gait analysis, sports biomechanics, ergonomics, animation, rehabilitation.
Electromyography (EMG) [157, 158]	Measure the electrical activity of muscles, using surface or intramuscular electrodes.	Detect and quantify muscle activation patterns.	Gait analysis, sports biomechanics, ergonomics, prosthetic control, rehabilitation.
Pressure Mapping Systems [159, 160]	Measure the pressure distribution between a surface and a body part.	Create pressure maps or images.	Seating and footwear design, wound prevention, ergonomics.
Inertial Measurement Units (IMUs) [161, 162]	Measure acceleration, angular rate, and sometimes magnetic field, using a combination of accelerometers, gyroscopes, and magnetometers.	Provide orientation, position, and motion data.	Gait analysis, sports biomechanics, ergonomics, rehabilitation.
Flex Sensors [163, 164]	Measure the bend or flex of a surface or object.	Provide bend or flex angle data.	Biomechanics of joint motion, ergonomics, rehabilitation.

Table 3.2: Biomechanics Sensors

Sensor	Description	Functionalities	Applications
Load Cells [165]	Measure the force or load applied to a surface or object.	Provide force or load data.	Sports biomechanics, ergonomics, rehabilitation.
Goniometers [166, 167]	Measure the angle between two body segments or objects.	Provide joint angle data.	Biomechanics of joint motion, ergonomics, rehabilitation.
Dynamometers [168, 169]	Measure the force or torque generated by a muscle or joint.	Provide muscle strength or joint torque data.	Sports biomechanics, ergonomics, rehabilitation.
Pedometers [170, 171]	Count the number of steps taken by an individual.	Provide step count data.	Physical activity monitoring, gait analysis, rehabilitation.
Accelerometers [172]	Measure the acceleration of a body or object.	Provide acceleration data in one or more axes.	Physical activity monitoring, gait analysis, ergonomics, rehabilitation.
Respiratory Sensors [173, 174]	Measure respiratory parameters such as airflow, volume, and gas concentrations.	Provide respiratory data.	Exercise physiology, sports biomechanics, ergonomics.
Electrocardiography (ECG) [175, 176]	Measure the electrical activity of the heart.	Provide heart rate and rhythm data.	Exercise physiology, sports biomechanics, ergonomics.
Force Myography (FMG) [177, 178, 179]	Measures the force-related signals from muscle bulges and deformations	Detects the lateral force exerted by muscles during contractions	Prosthetic control, human-computer interaction, rehabilitation

Table 3.2: Biomechanics Sensors

Sensor	Description	Functionalities	Applications
Mechanomyography (MMG) [180, 181]	Measures the low-frequency vibrations generated by muscle fibers during contractions	Detects the mechanical waves produced by the lateral oscillations of muscle fibers	Biomechanics research, ergonomics assessment, muscle fatigue monitoring
Tendon Force Sensors [182]	Measures the force transmitted through tendons	Detects the tensile force generated by muscle-tendon units	Biomechanics research, rehabilitation, prosthetic control
Muscle Hardness Sensors [183, 184]	Measures the stiffness or hardness of muscles	Detects changes in muscle tone or stiffness during contractions	Rehabilitation, ergonomics assessment, muscle fatigue monitoring
Ultrasound Force Sensors [185]	Measures the muscle deformation and force using ultrasound imaging	Detects changes in muscle shape and thickness during contractions	Biomechanics research, rehabilitation, ergonomics assessment
Fiber Optic Force Sensors [186, 187]	Uses fiber optic sensors embedded in muscles to measure strain and estimate force	Minimally invasive measurement of muscle force	Biomechanics research, surgical planning, prosthetic control
Implantable Force Transducers [188]	Implantable devices that measure muscle-tendon forces directly	Direct measurement of muscle-tendon force	Biomechanics research, surgical planning, prosthetic control
Intramuscular Pressure Sensors [189, 190]	Measures pressure changes within a muscle during contraction	Direct measurement of muscle force/tension	Biomechanics research, rehabilitation, surgical planning

3.2.2 Related Work

Hollis, Christina R., et al. [191] investigated using wearable sensors (RunScribe™) to measure running mechanics outside a lab. Runners wearing the sensors completed runs on a track and grass at slow and fast paces. The sensors successfully detected changes in running mechanics as expected, like stride length increasing with speed and impact forces being higher on the track. This suggests wearable sensors could be a viable alternative to lab-based methods for studying running biomechanics.

Howard, Róisín M. et al. [192] surveyed sports biomechanics researchers on their use of sensor devices, particularly EMG. Force platforms were most common, but inertial sensors and EMG are gaining traction. Researchers prioritized wireless, user-friendly devices. The ideal system would be a single, low-power, multi-sensor device with user-friendly analysis software, compatible with existing systems. This highlights the need for development in user-friendly, wireless multi-sensor devices and accompanying software for biomechanics analysis.

Inertial sensors have become indispensable tools in the biomechanical field due to their miniaturization and continuous improvement. Blanco Ortega, Andrés, et al. [193] compile their usage in obtaining biomechanical variables such as velocity, acceleration, and power, focusing on combat sports like boxing, karate, and Taekwondo. This work underscores the importance of using inertial sensors for upper extremity biomechanical measurements.

Hug, François, et al. [194] highlights several experiments that utilized ultrasound shear wave elastography to examine the hypothesis that muscle stiffness correlates linearly with both active and passive muscle forces. It provides examples of studies that utilized muscle stiffness measurements to estimate changes in muscle force.

Stetter, Bernd J., et al. [195] aimed to develop an artificial neural network (ANN) to estimate knee joint forces (KJF) during sport movements using data from wearable sensors. Thirteen participants wore two inertial measurement units (IMUs) on the right leg and performed various movements. Biomechanical modeling determined KJF, and an ANN was trained to predict KJF from IMU signals. The ANN-predicted KJF correlated well with reference KJF values, supporting the use of wearable sensors and ANNs for estimating joint reactions in sports.

Kim, Suin et al. [196] introduces a simple wearable sensor system combined with an inverse dynamics-based static optimization method. The system, comprising four IMUs and a manually developed ground reaction force sensor, measures joint angles and ground reaction forces. Performance was validated against a camera-based motion capture system

and force plate. Muscular forces in nine lower extremity muscle groups were estimated and qualitatively analyzed in relation to gait functions and compared with electromyography signals.

Lu, Yun, et al. [197] introduces a wearable multi-sensory system for estimating muscle force during isometric contraction assessment and dynamic training. It employs a customized wearable ultrasound system for real-time muscle deformation measurement and an inertial measurement unit sensor for joint angle detection. Muscle force prediction considers muscle deformation and thickness changes due to joint angle variation. Evaluating rectus femoris muscle force demonstrates over 90% accuracy in prediction, with promising application in dynamic muscle force estimation for rehabilitation and recovery assessment using exoskeletal rehabilitation robots.

Moghadam, Shima Mohammadi et al. [198] compared the performance of four non-linear regression ML models to estimate lower-limb joints' kinematics, kinetics, and muscle forces using Inertial Measurement Units (IMUs) and electromyographys' (EMGs) data. The RF and CNN models outperformed the others, providing lower prediction errors in all intended targets with reduced computational cost, suggesting their potential to overcome the limitations of traditional optical motion capture for 3D gait analysis.

3.3 Conclusion

In conclusion, biomechanics, musculoskeletal modeling, and wearable sensors have revolutionized the field of human movement analysis. Through the integration of advanced technology, such as inertial measurement units (IMUs), electromyography (EMG), and ultrasound systems, researchers can now accurately estimate muscle forces, joint angles, and kinetics during various activities. This chapter has explored the advancements in wearable sensor technology, the development of musculoskeletal models, and the application of machine learning techniques for predicting biomechanical parameters. As technology continues to evolve, interdisciplinary collaboration and innovation will further enhance our understanding of biomechanics and its applications in various fields.

Chapter 4

An Ensemble Machine Learning Technique for Detection of Abnormalities in Knee Movement Sustainability

The purpose of this study was to determine electromyographically if there are significant differences in the movement associated with the knee muscle, gait, leg extension from a sitting position, and flexion of the leg up for regular and abnormal sEMG data. Surface electromyography (sEMG) data was obtained from the lower limbs of 22 people during three different exercises: sitting, standing, and walking (11 with and 11 without knee abnormality). Participants with a knee deformity take longer to finish the task than healthy subjects. The sEMG signal duration of patients with abnormalities is longer than that of healthy ones, resulting in an imbalance in the obtained sEMG signal data. As a result of the data's bias towards the majority class, developing a classification model for automated analysis of such sEMG signals is arduous. The sEMG collected data is denoised and filtered, followed by the extraction time-domain characteristics. Machine Learning methods are then used for predicting the three distinct movements (sitting, standing, and walking) associated with electrical impulses for normal and abnormal sets. Also, different anomaly detection techniques are used for detecting occurrences in the sEMG signals that differ considerably from the majority of data and hence used for enhancing the performance of our model. The iforest anomaly detection technique presented in this work can achieve 98.5% accuracy on the Light Gradient Boosting Machine algorithm, surpassing the previous results that we can achieve maximum accuracy of 92.5% and 91%, improving accuracy by 6%-7% for classification of knee abnormality using Machine Learning.

4.1 Introduction

Walking, running, or climbing stairs might be challenging if you have discomfort due to joint pain. Pain can arise as a result of a trauma or an injury. It may also start for no apparent cause at times. One study found that one in three people have arthritis

or joint symptoms in 18-64 year olds because of an underlying injury or condition such as osteoarthritis of the knee [199]. X-Ray [200], or Magnetic Resonance Imaging (MRI) methods [201] are used in the clinic to diagnose such abnormalities. The X-ray approach collects and assesses bone condition, whereas MRI gives comprehensive information about knee anatomies such as cartilage, ligaments, and tendons. MRI is an effective diagnostic tool, but it is also more expensive. Wearable sensors like electromyography (EMG), accelerometers, gyro meters, and visual sensors like imaging cameras can be used to identify knee problems.

Depending on the parts of the human body supported by an exoskeleton, they can be divided into upper limb exoskeleton, lower limb exoskeleton (lees), whole body exoskeleton, and specific joint exoskeleton [202, 203]. These exoskeleton systems are divided into three categories based on their various target applications and users, which are gait rehabilitation, human movement and enhancement of human force [204].

Computer vision has applications across all areas of engineering, including exoskeletons. The exoskeleton gets information about its environment from an RGB-D camera and extracts characteristics of the soil surface that could affect its gait. Then it makes decisions according to the features of the environment, the state of the environment, and safety restrictions. It also provides the correct length and height of the walk to the parameterized gait planning model to assist the user in walking [205].

Sun et al. present the generation of exoskeleton mobility; several signal acquisition types can be applied, each focusing on specific areas and movements of the human body [206]. The current signal acquisition technologies include computer vision, electroencephalography (EEG), electromyography (EMG), and Inertial sensors. These methods have been developed over the past years and significantly impacted. In addition, the detection and control method, stability of the modeling, and comfort of an exoskeleton of the lower limb will also impact its performance [206].

Calle et al. discussed the Integration of Inertial Sensors in a Lower Limb Robotic Exoskeleton [207, 208]. It presents a novel platform based on Imocap-GIS inertial sensors with a motion assistance exoskeleton. It generates joint movements via Maxon motors and harmonic drive gearboxes.

Tang et al. present an extensive study on a wearable lower limb exoskeleton [209]. It classified the wearable device according to the power source and the working principle. It also compared and analyzed the design idea, wearing mode, material, and performance of different types of lower limb exoskeletons.

Taylor et al. demonstrate real-time muscle length tracking in an in vivo turkey model

via chronically implanted magnetic beads while investigating accuracy, biocompatibility, and long-term implant stability [210].

Botter et al. combined the ultrasound (US) imaging and multichannel surface EMG to compare their detection sensitivity to fasciculations occurring in different muscle regions and to investigate the effect of EMG electrodes' configuration on their agreement [211]. Monopolar surface EMGs were collected from medial gastrocnemius and soleus with an array of 32 electrodes.

EMG sensors recognize motions in advance while faster detecting signal changes, making them better at researching neuromusculoskeletal diseases. The signal obtained by the EMG sensors is a biological signal which measures the electrical activity produced by the skeletal muscles. Surface electromyogram (sEMG) and intramuscular EMG (iEMG) techniques are important methods to acquire EMG signals [212, 213]. sEMG has an advantage over iEMG in that the electrodes may be placed without causing discomfort or requiring medical supervision, and the risk of infection is minimum.

Gui et al. proposed an adaptive estimator based on the EMG to obtain and update the EMG pair model without calibrations and recalibrations [214]. Simulation and experiments indicate that the proposed estimator can adaptively predict the subject's active joint couple and provide precise control of exoskeleton movement.

Kang et al. developed a neural network gait and slope estimator using electromyogram (EMG) and mechanical sensor signals [215]. The results of four healthy and two elderly subjects show that the EMG approach can reduce the error rate by 14.8% compared to models using only mechanical sensors.

Long-term monitoring with surface electrodes is simpler than with iEMG needles. As Kanoga et al. have shown, the location of the sEMG sensors has a considerable impact on the collected signal and its subsequent analysis and identification using the calculation method [216]. sEMG sensors can collect data during daily human activities, like sitting, climbing, walking, and standing, which can then be used to recognize actions, detect anomalies, etc.

4.1.1 Motivation

Knee pain has become one of the most common health issues among seniors. Wounds, repeated stress on the joint, aging, etc are underlying condition like arthritis are the most common causes of knee pain [217]. The knee joint is one of the most complex joints in the human body, ensuring leg movement and body stability and acting as a damper [218]. The knee joint is made up of different bones, tendons, muscles, cartilage, ligaments and

fluids.

During the past decade, researchers have focused on categorizing sEMG signals from the upper limb. The sEMGs obtained from the lower limb are more difficult to interpret because they are influenced by overlapping muscles [217, 218]. Various signal processing approaches [219, 220] have been used to study the categorization of sEMG data from the lower leg. Chen et al. used a deep-neural-network to predict lower-limb extension/flexion joint angles from sEMG signals [221]. Morbidoni et al. used an in-depth learning method to classify the walking step using the sEMG data [222]. Bonato et al. developed a technique to measure muscle fatigue using EMG during squats. They used Cohen-Posch representations to improve accuracy and applied it to quadriceps and hamstring muscles in control and ACL-deficient subjects. Instantaneous median frequency was calculated for the knee-extension phase of the exercise. [223]. The classification of walking patterns and the use of computational approaches are divided into normal and gait disorders [224].

4.1.2 Contributions

This chapter focuses on identifying abnormality in movements from imbalanced sEMG data and classifying six different activities (three normal and three abnormal) for standing, sitting, and gait positions. We thus present a performance comparison of various prediction classifiers for detecting knee abnormalities from sEMG signals in which classifiers are applied on the datasets consisting of data with and without anomalies in using multiple anomaly detection techniques. This chapter aims to enhance the ability to detect knee movement abnormalities and enable healthcare for a more accurate diagnosis. To the author's knowledge, no study solves the problem of identifying knee abnormalities using anomalous detection methods. The major contributions are described as follows:

- Anomaly detection techniques such as iforest, KNN, and lof have been used to remove abnormal data.
- Enhancement of classifier's accuracy using boosting algorithms.

4.1.3 Organization

In this chapter, Section 4.2 describes our proposed model's datasets, methodology, and workflow. Section 4.3 discusses the model evaluation parameters. After that, Section 4.4 defines the performance and provides a comparative analysis between different metrics. Finally, the chapter is concluded in Section 4.5.

4.2 Materials and Methods

4.2.1 Dataset and its Features

sEMG and joint angles were derived from the publicly available dataset at UCI machine learning repository [225]. The data includes sEMG signals from the lower limbs of 22 adults, 11 of whom are in good health and 11 of whom have known knee problems. Participants were in good health with no history of injury or knee discomfort. One abnormal subject of the knee suffered a sciatic nerve lesion, six subjects suffered an anterior cruciate ligament (ACL) lesion, and the other four suffered a meniscus lesion. Dataset description and feature information have been provided in Table 4.1 and Table 4.2 respectively. These data were collected using a Biometrics Ltd. DataLog MWX8 and a direction finder. Each person performed three distinct tasks: walking, bending the leg upward and lengthening the leg from a sitting posture. The goniometer was connected out of the knee joint, and sEMG data was collected for four muscles: vastus medialis (VM), rectus femoris (RF), semitendinosus (ST) and biceps femoris (BF) (Figure 4.1).

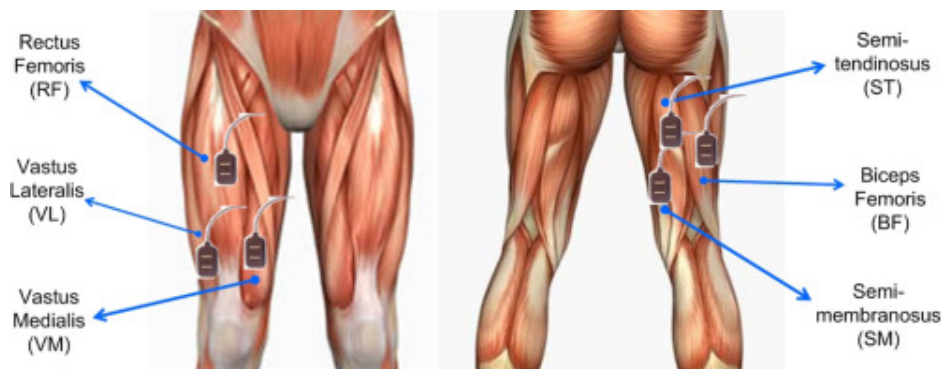


Figure 4.1: Muscle diagram for rectus femoris (RF), vastus lateralis (VL), vastus medialis (VM), semitendinosus (ST), biceps femoris (BF) and semitmembranosus (SM).

The person's afflicted limb with aberrant knee and left leg healthy participants were chosen to get the signal. The data was collected at a sampling frequency of 1000 Hz with a resolution of 14 bits. The sEMG signals have already been filtered through a band-pass filter that ranges from 20 Hz to 460 Hz. The data gathered do not have a signal corresponding to the transition steps, such as standing, sitting, walking, standing, etc. In Table 4.3 movements labeled as 0 refers to Normal walking, 1 for Normal sitting, 2 for Normal standing, 3 for Abnormal walking, 4 for Abnormal sitting, 5 for Abnormal standing, respectively. The data was transmitted in real-time through Bluetooth from the MWX8 gadget to the PC.

Table 4.1: Description of the dataset

Description	Type/Count
Number of features	5
Feature characteristics	Real
No. of classes (Movements)	6
Dataset characteristics	Multivariate
Associated tasks	Classification
Missing values	No

Table 4.2: Description of the features

Feature	Information
RF	Rectus Femoris
BF	Biceps Femoris
VM	Vastus Medialis
ST	Semitendinosus
FX	Flexion

Table 4.3: Sample dataset

RF	BF	VM	ST	FX	Movement
0.0011	-0.0011	0.0021	0.0050	57.6	0
0.0011	-0.0010	0.0022	0.0052	57.5	0
0.0011	-0.0009	0.0024	0.0054	57.3	0
0.0051	0.0002	-0.0010	0.0007	3	1
0.0052	0.0002	-0.0007	0.0007	3	1
0.0056	0.0003	-0.0001	0.0007	3	1
0.0011	0.0002	0.0002	0.0000	70.1	2
0.0011	0.0002	0.0003	0.0000	70.1	2
0.0011	0.0002	0.0004	0.0000	70.1	2
-0.0004	0.0002	0.0184	0.0017	-9	3
0.0004	0.0011	0.0187	0.0022	-9	3
0.0012	0.0021	0.0190	0.0028	-9	3
0.0011	0.0018	0.0008	0.0012	-4.6	4
0.0011	0.0019	0.0009	0.0013	-4.5	4
0.0012	0.0020	0.0009	0.0013	-4.6	4
0.0043	0.0013	-0.0019	0.0005	-30	5
0.0045	0.0013	-0.0017	0.0005	-29.7	5
0.0048	0.0014	-0.0015	0.0005	-29.4	5

4.2.1.1 Dataset Limitations

Following are some of the limitations of the dataset used in this study.

1. The dataset may not represent a diverse population in terms of age, gender, fitness level, or health status.
2. Data is collected in laboratory settings, which may not fully represent real-world conditions.
3. The dataset does not capture long-term changes or fatigue effects.
4. Additional factors that could influence EMG signals (e.g., fatigue level, hydration status) can not be recorded.

4.2.1.2 Statistical analysis of dataset

The Table 4.4 and Figure 4.3 shows the statistical analysis of all the datasets. The standard deviation of FX is the highest, and BF is the lowest in all the datasets. EMG signal plots can be seen before data preprocessing in Figure 4.3(a) for the four channels of a Normal Person and in Figure 4.3(b), for an Abnormal person. EMG signal plots can be seen after signal preprocessing, in which we rectify and transform the filtered sEMG values we can see in Figure 4.3(c) for a normal person and in Figure 4.3(d) for an Abnormal person. After normalizing we get the plots as shown in Figure 4.3(e) for a Normal person and in Figure 4.3(f) for an Normal person.

4.2.2 Approach

The approach is described in Figure 4.2. First, the raw experimental data of the lower limb is taken from the UCI machine learning repository. Where dataset consists of data for three different activities (sitting, standing, and walking) for 11 abnormal and 11 normal individuals labeled as Abn. and Nor. in Table 4.4.

Table 4.4: Statistical analysis of each feature in each dataset

Dataset	Features	Mean.	Std.	Min.	Max.
Nor.Walk	RF	0.04	0.016	-0.11	0.10
	BF	-0.02	0.008	-0.06	0.11
	VM	-0.01	0.026	-0.18	0.17
	ST	-0.03	0.011	-0.08	0.11
	FX	19.36	17.836	1.30	62.60
Abn.Walk	RF	-0.01	0.083	-1.27	1.36
	BF	-0.06	0.023	-0.32	0.16
	VM	-0.06	0.068	-0.61	0.59
	ST	0.06	0.016	-0.15	0.10
	FX	-12.7	16.37	-57.2	3.10
Nor.Sit	RF	0.03	0.016	-0.11	0.10
	BF	-0.02	0.008	-0.06	0.11
	VM	-0.01	0.026	-0.18	0.17
	ST	-0.03	0.011	-0.08	0.11
	FX	53.36	36.836	3.30	109.6
Abn.Sit	RF	-0.03	0.006	-0.08	0.05
	BF	-0.06	0.034	-0.33	0.20
	VM	-0.04	0.012	-0.11	0.20
	ST	0.05	0.06	-0.56	0.40
	FX	-58.7	44.37	-113	-2.90
Nor.Stand	RF	0.03	0.018	-0.11	0.10
	BF	-0.03	0.003	-0.06	0.11
	VM	-0.01	0.026	-0.18	0.17
	ST	-0.04	0.011	-0.08	0.11
	FX	51.36	20.836	20.30	75.6
Abn.Stand	RF	-0.01	0.032	-0.20	0.26
	BF	-0.06	0.006	-0.05	0.06
	VM	-0.06	0.019	-0.16	0.14
	ST	0.07	0.02	-0.49	0.35
	FX	-29.4	30.37	-83	-4.70

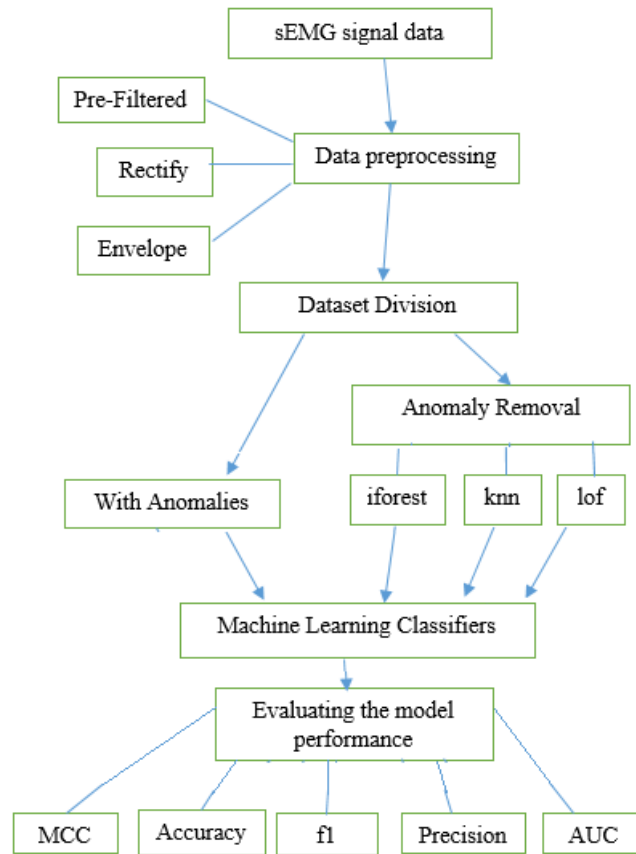


Figure 4.2: Methodology Used

The data is pre-filtered with a band-pass filter that ranges from 25 Hz to 465 Hz. The signal data is later demodulated using Hilbert’s transformation. In case of mechanical failure, the gathered vibration signals are usually modified. As such, signal demodulation can separate the carrier component from the modulation component. In addition, the default features are still hidden in the modulation component. Thus implementation of signal demodulation during signal preprocessing is essential for precise results. The min-max normalization method was applied to normalize the samples with features ranging from 0 to 1. As shown in Figure 4.2. The dataset is then divided into two groups:

1. Dataset consisting anomalies
2. Datasets produced as a result of the use of anomaly detection techniques, namely: iforest, KNN and lof.

Following the generation of these datasets, we use machine learning classifiers that comprise of models:

- Light Gradient Boosting Machine

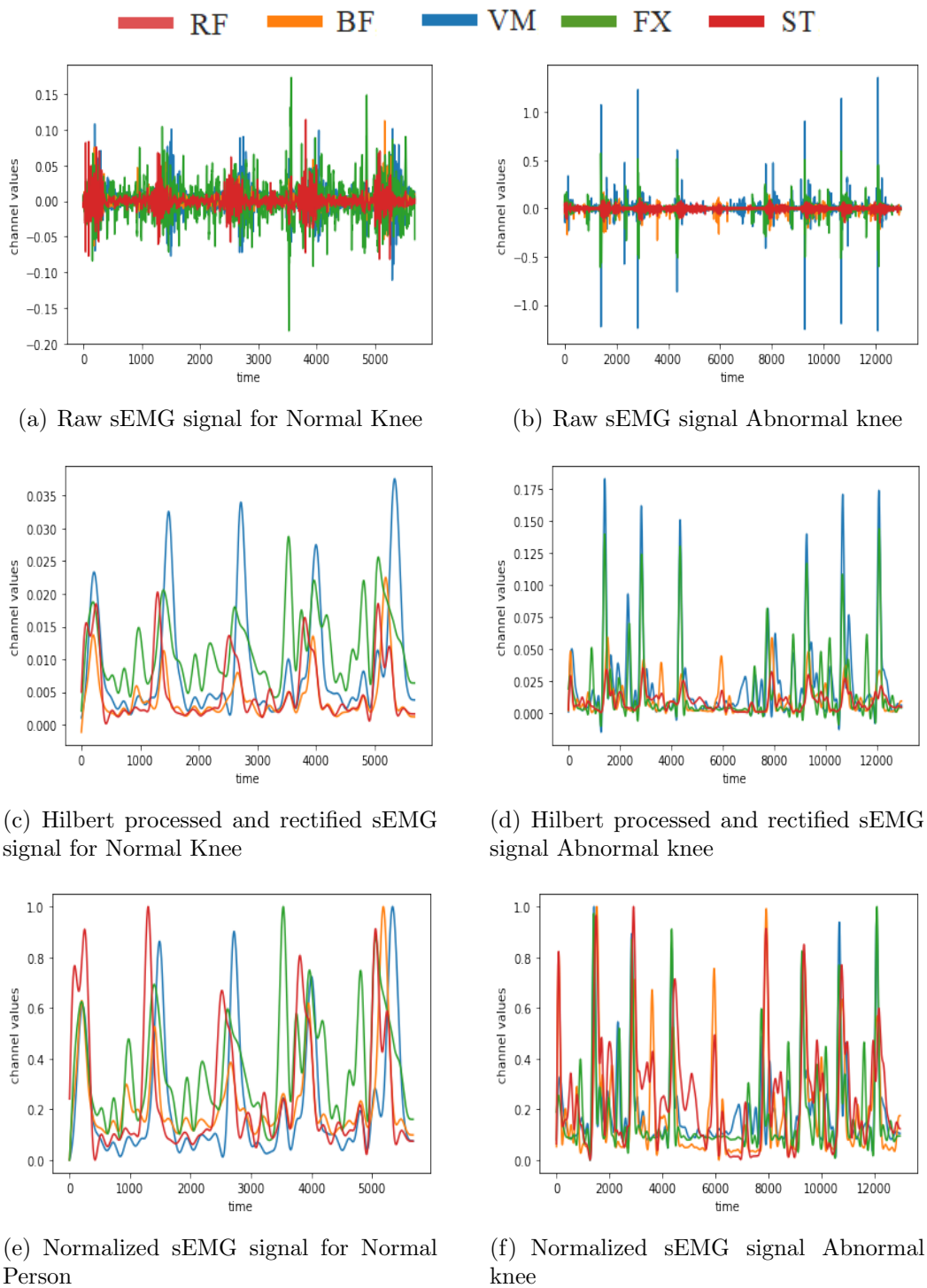


Figure 4.3: The sEMG signals at various stages during data preprocessing

- Extra-Trees-Classifier
- Random-Forest-Classifier
- eXtreme Gradient Boosting
- Decision-Tree-Classifier

4.2.2.1 Anomaly Detection Techniques

Anomaly Detection Techniques (also known as outlier analysis) is a stage in data mining that finds data points, events, and/or observations that deviate from the normal behavior. Anomalous analytics may indicate major problems, such as a technical flaw or potential scenarios. We analyze the sEMG signals and remove outliers from them using the following techniques:

1. **iforest:** The Isolation Forest is an unsupervised learning algorithm to detect abnormalities that operate on the principle of isolating abnormalities rather than the most common staining profiling techniques [226]. The decision tree algorithm backs the secluded forest. It separates aberrant values by randomly selecting a character from the given set of characteristics and then randomly selecting a value divided between the max and min values. This random distribution of traits will lead to shorter tree trajectories for anomalous data points, differentiating them from the rest of the data. Generally, the first step in detecting an anomaly is to profile what is "normal" and then flag anything that cannot be considered typical as abnormal. However, the forest isolation algorithm does not function on this principle; it first defines the "normal" behavior and calculates the distances per point. As expected, the isolation forest functions instead by isolating anomalies by explicitly isolating abnormal issues in the dataset. The isolation forest algorithm assumes that specific abnormalities and observations should make it easier to identify them. The isolated forest uses a series of trees to isolate anomalies for the given data points [226]. The isolation forest recursively generates scores across the dataset by randomly selecting a feature and then a split value. The monsters likely need fewer random partitions than the "normal" points of the dataset. The faults will be the dots with a smaller path into the tree and the crossed edges of the root node.
2. **KNN (k-Nearest Neighbors Detector):** The unsupervised k-nearest-neighbor global anomaly detection techniques is an easy way to detect anomalies and not be confused with the k-nearest classification [227]. As the name suggests, it focuses on global anomalies and cannot see local ones. First, the closest k-neighbors must be found for every record in the data set. Afterward, an anomaly score is calcu-

lated using those neighbors [227]. On the other hand, two possibilities have been suggested: either the distance to the nearest k th-neighbor is used (only one), or the average length to all the nearest k -neighbors is calculated. In the subsequent paragraphs, we refer to the first method as k th-NN and the second as k -NN.

In practice, the k -NN method is usually preferred. However, the absolute value of the score depends a lot on the data set itself, the number of dimensions, and the standardization [227]. Therefore, it isn't easy to choose an appropriate threshold, if necessary. Naturally, selecting the parameter k is essential to the results. If too low, the density estimate for the records could not be reliable.

In addition, the density estimate may be too coarse if it is too large. Usually, k should be in the range $10 < k < 50$. In classification, a suitable k can be determined, for example, using cross-validation. Unfortunately, there is no such technique for detecting unsupervised abnormalities due to missing tags. For this reason, we later evaluate many different values for k and mean to obtain a fair assessment when comparing algorithms.

3. **LOF (Local Outlier Factor)**: The LOF uses the identification of density-based outliers to identify local outliers or sites that are outliers relative to their immediate environment rather than the overall distribution of data [228]. The higher the LOF value of an observation, the more astonishing it becomes. The number of regarded neighbors is usually fixed to be: 1) less than the maximum number of samples close by which may be local outliers and 2) exceeds the minimum number of samples that a cluster must contain for other samples to be local outliers [228].

4.2.2.2 Feature Importance

Techniques that score each features according to their usefulness in predicting a target variable are known as feature importance. Statistical correlation scores, coefficients derived from linear models, decision trees, and permutation significance scores are a few examples of trait importance scores. Feature significance scores are essential for predictive modeling projects. They provide an overview of the data and information on the template. Foundation for the reduction of dimensions and selection of characteristics can improve the effectiveness of a predictive model on the problem.

The importance of the feature is computed using the equation Eq. 1. After five different runs, the average weight given to each feature is described in Table 4.5 and their average weight is used for ranking the features. It is found that FX has the highest ranking and RF has the lowest ranking.

$$i_j = s - \frac{1}{k} \sum_{k=1}^K s_{k,j} \quad (4.1)$$

Where, Compute the score $s_{k,j}$ of model m on corrupted data $D_{k,j}$ and importance i_j for feature f_j .

Table 4.5: The feature importance of each feature.

Runs	FX	BF	ST	VM	RF
1	0.25	0.22	0.2	0.19	0.14
2	0.23	0.2	0.24	0.17	0.16
3	0.27	0.23	0.19	0.16	0.15
4	0.25	0.21	0.21	0.17	0.16
5	0.24	0.2	0.2	0.17	0.17
Avg.	0.248	0.212	0.208	0.172	0.156

4.2.2.3 Computational Classifiers

The dataset was divided into 75/25 Train-Test halves. The resulting datasets are subsequently trained on a variety of machine learning classifiers. The results are compared between the top five performing classifiers:

- Light Gradient Boosting Machine
- Extra-Trees-Classifier
- Random-Forest-Classifier
- eXtreme Gradient Boosting
- Decision-Tree-Classifier

1. **LightGBM**: Improves gradient amplification by incorporating automatic feature selection and focusing on amplification instances with higher gradients [229]. This will lead to much faster workout times and greater predictive efficiency. It is known to be a very effective computational algorithm and rapid treatment method. As the trees of other algorithms grow horizontally, LightGBM grows vertically, in the direction of the leaves, while different algorithms grow in level. LightGBM picks the sheet with the highest loss to develop. When culturing the same sheet, it will decrease the loss more than a leveling method. Decision trees are used to improve model performance while saving memory. Quality and scalability are at the heart of

the development approach. It uses two new techniques, one-sided gradient sampling and proprietary feature grouping (EFB). Fill in the histogram algorithm gaps in most GBDT (Gradient Boosting Decision Tree) frameworks [229]. The LightGBM algorithm's characteristics are produced by the Gradient-based One Side Sampling Technique (GOSS) and the Exclusive Feature Bundling Technique (EFB). They work together to make the model work and give it an edge over the competition for GBDT systems.

2. **eXtreme gradient boosting:** Unlike traditional Gradient Boosting, XGBoost uses its tree-building technique where Similarity Score and Gain determine optimal node divisions [230]. Residual value: real value (observed) - predicted value The probability of an event calculated in an earlier stage is called an earlier probability. For each finding, the initial probability is considered to be 0.5, which is used to create the first tree. The process map shows that the previous probability is reconstructed for all future trees based on the original predictions and the predictions of all previous trees. Lambda is a setting for equalization. The increase in lambda diminishes the influence of small leaves (those with few observations) disproportionately while having a moderate impact on large leaves (those with many observations) [230]. The optimum division for the tree is then determined by selecting the division knot with the greatest gain.
3. **Random Forest:** A random forest is a holistic approach that uses multiple decision trees [231]. It is also known as bootstrapping and aggregation, often called bagging, to perform regression and filing tasks. The simple hypothesis is that instead of relying on individual decision trees, the outcome is determined by combining multiple decision trees. Random Forest uses a variety of decision trees as a primary learning approach. The dataset is randomly sampled for line and feature sampling, resulting in sample data sets for each model. This section is called Bootstrap. Random logging is a method of bagging and not a method of stimulation. Trees in random forests grow next to each other. Growing trees are not related. A randomized forest is a meta-marker (it incorporates the results of several predictive models) that groups together many decision trees with some helpful changes. It is one of the most precise machine learning technologies available on the market. On many data sets, this is a very reliable grader. It works effectively on big datasets [231]. It can handle tens of thousands of input variables without removing them. It identifies the importance of a large number of factors in categorization.
4. **Extra Trees:** This module presents a meta-estimator that uses the mean to improve statistical accuracy [232]. It combines a range of randomized decision trees

(Extra Trees is an intuitive learning approach that combines the predictions of multiple decision trees into a single prediction. It's called the random forest model that we talked about earlier. It may produce similar or better results than the random forest technique because it uses a simpler algorithm to create the entire decision trees. Following the standard upper-level approach, the Extra-Trees algorithm produces an arbitrary decision or a set of regression trees. It differs from earlier tree-based methods by randomly dividing nodes and generating trees using the whole learning sample (rather than bootstrap replication) [232]. It differs from previous tree-based methods by randomly dividing nodes and generating trees using the entire learning sample (rather than bootstrapping).

5. **Decision Tree:** Decision tree learning, also known as decision tree induction, is one of the statistical modeling methods used in analytics, data processing, and machine learning [233]. It uses a decision tree from assumptions about an item to predictions about the article's target value. Classification trees represent class tags, and branches represent combinations of features that match specific class tags. A decision tree is a graphic and explicit depiction of the options and decision-making that can be used in decision-making procedures. A decision tree is a data model for exploring the data. The decision tree learning approach is often used in data mining. The objective is to build a model that can predict the value of a target variable based on a set of input variables. The decision tree method is a supervised learning approach that can generate continuous and categorical production variables.

4.3 Model Evaluation

There are many classification metrics for evaluation of our model, in this study we assess the performance on (1) Matthews Correlation Coefficient (MCC), (2) F1 Score (F1) (3) Accuracy (4) Precision and (5) AUC.

4.3.1 Matthews Correlation Coefficient (MCC)

MCC's formula takes into account all of the cells in the Confusion Matrix.

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4.2)$$

Like the Correlation Coefficient, MCC has a range of values ranging from -1 to +1. A model with a score of +1 is considered ideal, whereas a model with a score of -1 is considered bad. This feature is one of the most important benefits of MCC since it leads

to simple interpretation.

4.3.2 F1 score

The F or F score measurement is a measurement of test accuracy. It is calculated with the accuracy and recalculation values from the test. Accuracy is the number of actual positive outcomes divided by the total number of positive outcomes, including those incorrectly identified. The recall is the number of true positive results divided by the total number of samples which should have been positive.

$$F1 = \frac{2 * TP}{2 * TP + FP + FN} \quad (4.3)$$

4.3.3 Precision

Precision, sometimes referred to as Positive Predictive Value, is computed as,

$$Precision = \frac{TP}{TP + FP} \quad (4.4)$$

Precision may produce a biased result since the formula lacks FN and TN, especially for unbalanced classes.

4.3.4 Accuracy

Accuracy is a criterion for classifying models. Informally, precision is the percentage of accurate predictions from our model. Formal accuracy is defined by:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.5)$$

where, TP denotes True Positives, TN denotes True Negatives, FP denotes False Positives, and FN denotes False Negatives.

4.3.5 AUC

AUC is an abbreviation for "Area under the ROC Curve". AUC measures the whole two-dimensional area beneath the entire ROC curve from (0,0) to (1,1). AUC is a metric that aggregates performance across all categorization criteria. AUC may be interpreted as the likelihood that the model rates a random positive case higher than a random negative example.

Table 4.6: Performance comparison of different anomaly detection techniques with different prediction models on different metrics on all the datasets.

Anomaly Detection Technique	Models →	LightGBM	XGBoost	Random Forest	Decision Tree	Extra Trees
	Metrics ↓					
Without Anomaly Detection	MCC	0.83	0.753	0.559	0.537	0.367
	AUC	0.972	0.958	0.897	0.831	0.839
	Precision	0.859	0.798	0.613	0.514	0.59
	F1	0.857	0.79	0.617	0.522	0.442
	Accuracy	0.859	0.797	0.64	0.537	0.478
iforest	MCC	0.985	0.982	0.922	0.822	0.816
	AUC	0.986	0.986	0.983	0.967	0.974
	Precision	0.985	0.984	0.937	0.859	0.865
	F1	0.985	0.983	0.933	0.855	0.843
	Accuracy	0.985	0.983	0.935	0.856	0.852
KNN	MCC	0.922	0.9	0.785	0.569	0.72
	AUC	0.983	0.98	0.948	0.862	0.918
	Precision	0.956	0.946	0.893	0.8	0.861
	F1	0.954	0.942	0.884	0.765	0.854
	Accuracy	0.953	0.941	0.88	0.748	0.853
lof	MCC	0.908	0.805	0.558	0.454	0.43
	AUC	0.979	0.962	0.894	0.837	0.834
	Precision	0.926	0.847	0.692	0.6	0.631
	F1	0.926	0.847	0.653	0.571	0.506
	Accuracy	0.926	0.847	0.666	0.583	0.569

4.4 Result Analysis

Here, we examine the prediction outcomes of five machine learning classifiers on the training, testing, and validation dataset. The performance of each model is shown in Table 4.6. All the prediction models have been trained on default parameters and evaluated on MCC, Accuracy, F1, precision, and AUC. The dataset has a small number of features but a large number of observation values. The K-fold cross-validation technique is used to assess the robustness of the best prediction method.

4.4.1 Model Performance Comparison

After training models on different datasets, it is observed that the dataset generated using iforest performs the best overall on the Light Gradient boosting machine model with an impactful accuracy of 0.985. Other anomaly detection techniques generated datasets also perform better than the initial dataset, with an accuracy of 0.859. The best model is

validated using 10-fold validation.

4.4.1.1 Accuracy

The accuracy is calculated using Eq. 4.5. Figure 4.4(a) shows the average accuracy of all the models. The light gradient boosting machine has the highest accuracy of 0.985, followed by xgboost and random forest. Extra trees have the lowest accuracy of 0.478.

4.4.1.2 F1 Score

The F1 is calculated using Eq. 4.3. Figure 4.4(b) shows the average F1 score of all the models after 10 fold validation. The light gradient boosting machine has the highest F1 of 0.985, followed by xgboost and random forest. Extra trees have the lowest accuracy of 0.442.

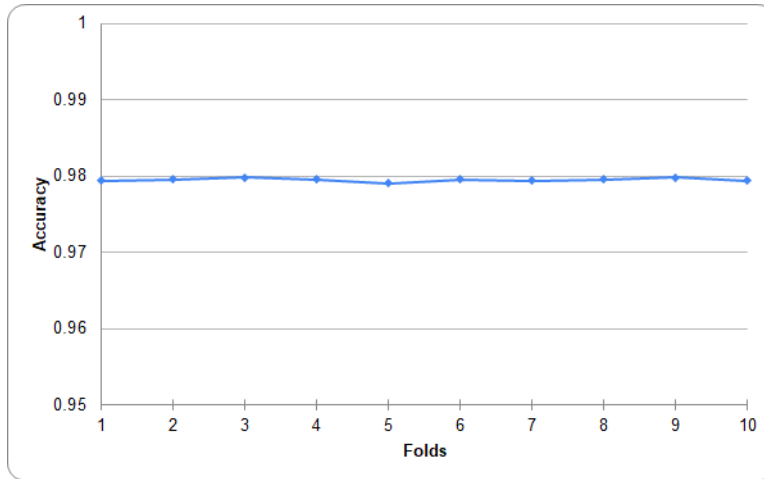
4.4.1.3 MCC

The MCC is calculated using Eq. 4.2. Figure 4.4(c) shows the average MCC of all the models after 10 fold validation. The light gradient boosting machine has the highest accuracy of 0.985, followed by xgboost and random forest. Extra trees have the lowest accuracy of 0.367.

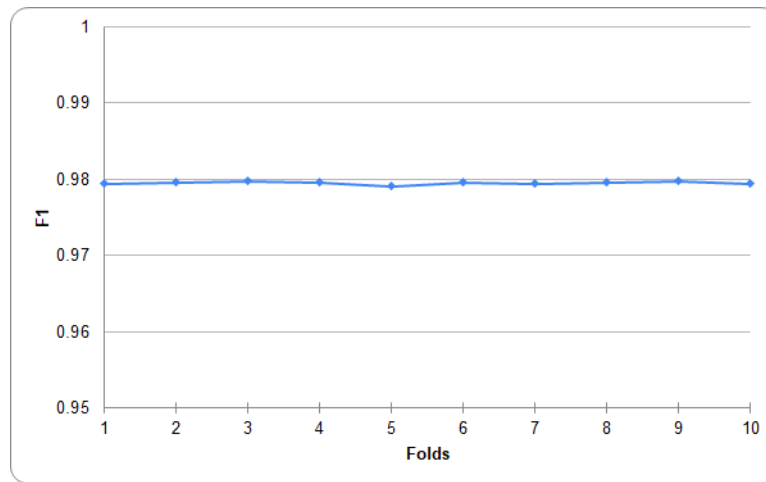
4.4.2 K-Fold Validation

Cross-validation is a statistical method for estimating capacity in machine learning models. It is commonly used in applied machine learning to compare and select a model for a specific predictive modeling problem. It is easy to understand, implement and produce skill estimates with a more negligible bias than other approaches. Cross-validation is a resampling technique used to assess ML models on a small sample of data. The process contains only one parameter called k , which specifies the number of groups in which a given data sample should be divided.

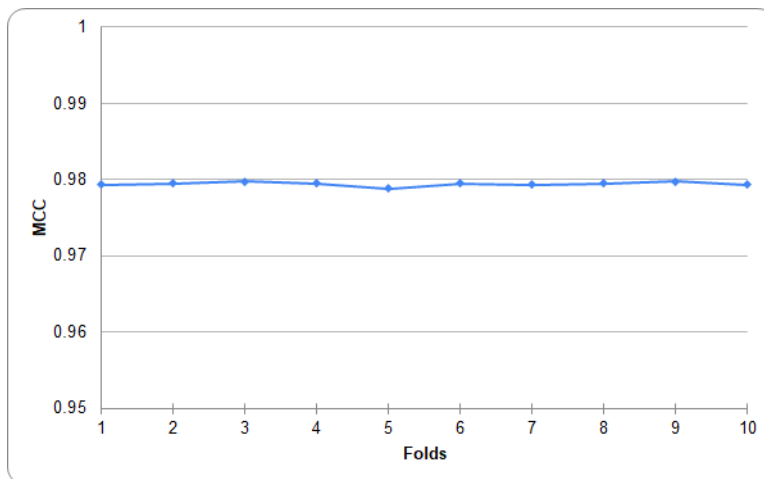
Consequently, this method is often referred to as k -fold cross-validation. When a specific number for k is specified, it may be used instead of k in the model reference, such as $k=10$, resulting in a 10 cross-validation. Cross-validation is primarily used in applied machine learning to assess the competency of a machine learning model on unseen data. It uses a small sample to determine how the model will work when used to produce projections of data not used in model formation. The charts show that the built model is robust after 10 plies and is reliable for testing our sEMG data.



(a) Accuracy



(b) F1 Score



(c) MCC

Figure 4.4: Cross Validation of best model for accuracy, F1-score and MCC.

4.4.3 Results and Discussion

The results indicate that datasets created following the anomaly detection techniques increase the performance of classifiers for movement prediction.

The performance of each model is shown in Table 4.6. All the prediction models have been trained on default parameters and evaluated on MCC, Accuracy, F1, precision, and AUC. The dataset has a small number of features but a large number of observation values. The K-fold cross-validation technique is used to assess the robustness of the best prediction method.

The dataset created with iforest achieves an outstanding overall score on the Light Gradient boosting machine model with an accuracy of 98.5% compared to the baseline performance without any adjustments on our dataset, which achieves an accuracy of 85.9%.

The other models significantly improve their accuracy and the metrics after anomaly removal from the data. There's a performance increase of 13% accuracy from the models which don't use anomaly detection techniques. Compared to the previous works [234, 235] which were able to get the accuracy of 92.5% and 91%, has also been improved by 6%-7% respectively, proves that this technique is effective in boosting the accuracy for classification of sEMG signals. Thus results demonstrate that eliminating anomalies substantially improves performance, producing some robust results, and should be used for developing models using sEMG data.

4.4.4 Limitation of the Study

Data imbalance is an essential issue in the classification of medical data. This can occur because of significant differences in the number of healthy and unhealthy individuals. This may happen because the data collection length varies depending on normal and abnormal subjects. About a health problem, someone with an anomaly of the knee takes longer to accomplish the task of moving, resulting in a more extended signal. Therefore, the abnormal subject's sEMG signal length results in a class imbalance. This affects the precision of the classification of knee motion and the prediction of anomalies.

4.5 Conclusion

This work demonstrated machine learning for movement classification using sEMG data. The suggested models are based on a large database with various parameters. This study aimed to analyze differences in classification performance following anomalous detection

methods. It was used to identify events in EMG signals that differed significantly from most data and to create new data sets.

Real-time sEMG data analysis and movement classification can be beneficial in the medical field. Doctors will identify aberrant motions and provide preventative steps to minimize risks and injuries connected to joints or muscles. This work was a step towards making a more accurate diagnosis system to predict abnormalities and suggest precautionary measures to prevent any unfortunate events. As we advance, we plan to gather DGG data from other topics to reduce data imbalances. Test oversample methods and study the most advanced machine learning graders. Convolution neural networks (CNN) have shown promising results in various classification questions. The iforest + LightGBM combination outperformed other models by efficiently removing anomalies using iforest, which improved data quality. LightGBM's leaf-wise growth further enhanced accuracy, focusing on high-gradient points and relevant features. This synergy boosted classification performance, resulting in superior accuracy and handling of imbalanced, complex datasets. The proposed methodology can be used in muscular deficiency or any other physiological deficiency.

Chapter 5

Time series generative adversarial network for muscle force prognostication using statistical outlier detection

Machine learning approaches, such as artificial neural networks (ANN), effectively perform various tasks and provide new predictive models for complicated physiological systems. Examples of Robotics applications involving direct human engagement, such as controlling prosthetic arms, athletic training, and investigating muscle physiology. It is now time for automated systems to take over modeling and monitoring tasks. However, there is a problem with the massive amount of time series data collected to build accurate forecasting systems. There may be inconsistencies in forecasting muscle forces due to the enormous amount of data. As a result, anomaly detection techniques play a significant role in detecting anomalous data. Detecting anomalies can help reduce redundancy and free up large storage space for storing relevant time-series data.

This chapter employs several anomaly detection techniques, including Isolation Forest (iforest), K-Nearest Neighbour (KNN), Open Support Vector Machine (OSVM), Histogram, and Local Outlier Factor (LOF). These techniques have been used by Long Short-Term Memory (LSTM), Auto-Regressive Integrated Moving Average (ARIMA), and Prophet models. The dataset used in this study contained raw measurements of body movements (kinematics) and the forces generated during walking (kinetics) of 57 healthy people (29 Female, 28 Male) without walking abnormalities or recent leg injuries.

To increase the data samples we used TimeGAN that generates synthetic time series data with temporal dependencies, aiding in training robust predictive models for muscle force prediction. The results are then compared with different evaluation metrics for five different samples. It is found that anomaly detection techniques with LSTM, ARIMA, and Prophet models provided better performance in forecasting muscle forces. The iforest method achieved the best Pearson's Correlation Coefficient (r) of 0.95, which is a competitive score with existing systems that perform between 0.7 and 0.9. The methodology provides a foundation for precision medicine, enhancing prognostic capability over relying solely on population averages.

5.1 Introduction

Muscle force estimation is critical for understanding musculoskeletal function and biomechanics. However, predictive modeling remains challenging due to inter-individual variability and data constraints. We present an integrated framework combining statistical outlier detection and modern generative adversarial networks (GANs) to enhance muscle force prognostic capability [236].

A commonly employed approach for non-invasive forecasting of muscle forces involves the classic Inverse Dynamic Computation (IDC) of the human musculoskeletal system. This method combines body segment kinematics, external reaction forces, and optimization techniques to estimate muscle forces. The IDC method utilizes motion data to estimate couples produced at joints, and its algorithm draws inspiration from Newton-Euler and Kane’s methods [237]. Several variables are necessary for these calculations, including anthropometric characteristics that define the inertial qualities of each limb, records of limb movements, and ground reaction forces. Researchers widely employ the IDC approach in investigations related to human motion, enabling the computation of net joint moments, net joint powers, and net joint inter-segmental forces in activities such as gait [238], lifting [239], skating [240], jumping [241], among others.

Common inverse dynamics computations for non-invasive muscle force prediction require extensive motion capture data. Recent generative models like SimGAN [242] can produce highly realistic simulated human motion sequences, providing abundant synthetic training data. However, effectively leveraging synthetic data for precision medicine remains an open challenge.

Accurately predicting trajectories of muscle force is critical for rehabilitation and treatment planning across conditions that impact mobility. However, substantial inter-individual variability makes prognosis difficult. We present a novel framework to enhance personalized muscle force predictions by integrating statistical outlier detection methodologies with recent advances in generative adversarial networks (GANs) and autoregressive language models.

Generative modeling has shown early promise for synthesizing plausible longitudinal patient records. Models such as medGAN [243] and MED-BERT [244] can produce synthetic electronic medical records complete with risk factors, vital signs, medications, and annotations. However, thus far these approaches have seen only limited applicability for precision prognosis.

Here, we demonstrate a methodology combining the benefits of generative deep learning

with traditional statistical outlier detection techniques. The core innovation is using real-world outlier patients to specialize generative models, allowing precise modeling of deviations from normal progression.

5.1.1 Related Work

Traditional methods for muscle force prediction have historically relied on physics-based models utilizing motion capture or electromyography (EMG) data. These approaches have provided valuable insights into biomechanics, but they often face challenges in accurately capturing the complex interactions between muscles, joints, and external forces. Researchers have continuously sought innovative solutions to overcome these limitations and improve the accuracy of muscle force prediction.

One approach involves deriving formulas based on empirical data to estimate muscle force across different age groups. For example, Masashi et al. developed formulas to estimate muscle force using manual muscle test scores, providing valuable insights into age-related changes in muscle strength [245]. Similarly, Mobasser et al. compared the performance of multilayer perceptron (MLP) and radial basis function (RBF) neural networks in predicting muscle force from EMG signals during elbow flexion [246]. These studies highlight the importance of utilizing advanced modeling techniques to improve the accuracy of muscle force prediction.

Recent advancements in deep learning have revolutionized the field of biomechanics by enabling more accurate and efficient prediction of muscle forces. Horst et al. leveraged deep neural networks to analyze features extracted from full-body motion tracking data, providing insights into individual gait patterns and biomechanical characteristics [46]. This approach offers a more comprehensive understanding of human movement and can aid in the development of personalized interventions for rehabilitation and performance enhancement.

In a groundbreaking study, Zhang et al. introduced a physics-informed deep learning approach to predict muscle forces and joint kinematics from surface electromyography (EMG) data [25]. By combining principles of biomechanics with deep learning techniques, the proposed method achieved superior accuracy in muscle force prediction compared to traditional approaches. This innovative approach has the potential to revolutionize the field of biomechanics and enhance our understanding of muscle function and movement dynamics.

Additionally, Ma et al. explored force control in prosthetic hands by utilizing surface electromyography (sEMG) signals and six-dimensional force sensors [58]. By integrating

multiple sensor modalities and advanced machine learning algorithms, the researchers were able to develop more intuitive and efficient control strategies for prosthetic devices, improving the quality of life for individuals with limb loss.

Furthermore, Zhu et al. addressed the challenge of assessing knee joint function in patients with knee impairments by focusing on knee contact force (KCF) prediction [59]. By leveraging advanced machine learning techniques, including random forest and support vector regression, the researchers developed accurate models for predicting KCF, providing valuable insights into knee biomechanics and aiding in the development of personalized treatment plans for patients with knee injuries or osteoarthritis.

Recent advancements in generative models have shown great potential for synthesizing biomechanical data, offering valuable tools for data augmentation, missing value imputation, and dataset expansion in various applications. Jiang et al. introduced SimGAN, a generative model designed to produce simulated motion sequences that are indistinguishable from real optical motion capture (mocap) data [242]. By leveraging generative adversarial networks (GANs), SimGAN provides a means to generate synthetic data that closely resembles real-world observations, facilitating the development and validation of biomechanical models and algorithms.

Similarly, Armanious et al. employed medGAN, a generative model for creating electronic health records, to impute missing values for patients with gait disorders [243]. By generating synthetic data to fill in gaps in the dataset, medGAN enables more comprehensive analysis and modeling of gait dynamics, contributing to improved diagnosis and treatment planning for patients with mobility impairments.

Carneros-Prado et al. highlighted the role of GANs in augmenting biomechanical datasets, emphasizing the importance of data augmentation for better understanding and analysis of gait dynamics [82]. By generating synthetic data with GANs, researchers can enhance the diversity and representativeness of their datasets, leading to more robust and generalizable models for studying human movement.

In the context of medical imaging, Xu et al. introduced progressive sequential causal GANs (PSCGAN) as a one-stop solution for ischemic heart disease (IHD) diagnosis without the need for gadolinium contrast agent injections or manual segmentation [84]. By generating synthetic medical images with PSCGAN, the proposed approach offers a non-invasive and automated alternative for diagnosing IHD, potentially reducing patient discomfort and healthcare costs associated with traditional diagnostic procedures.

Moreover, Bicer et al. proposed a data augmentation method for deep learning biomechanical models using adversarial autoencoders, a variant of GANs [85]. By generating

synthetic motion capture datasets for marker trajectories and ground reaction forces, adversarial autoencoders enable the augmentation of existing biomechanical datasets, enhancing the training and performance of deep learning models for tasks such as muscle force prediction and gait analysis.

However, leveraging synthetic data to enhance patient-specific muscle force prognostics remains relatively unexplored. Our proposed approach bridges statistical outlier detection and specialized generative models for personalized predictions. Anomaly detection techniques are pivotal in identifying unusual patterns or outliers in data, crucial across various domains including healthcare. Mullineaux, David R., and Gareth Irwin [87] developed for intra-participant time-series data from treadmill running, detects outliers in two stages. Firstly, spatial outliers are removed using median absolute deviation, followed by the removal of spatial-temporal outliers using a moving window standard deviation. Jin, Zeqing, et al. [90] introduces an anomaly detection system using layer-by-layer sensor images and machine learning to identify imperfections in transparent hydrogel-based bio-printed materials. Sivaraks, Haemwaan, et al. [91] introduces a novel anomaly detection technique leveraging cardiologist expertise and motif discovery. The algorithm ensures robustness and accuracy even in the presence of ECG artifacts, significantly reducing false alarms.

By integrating these anomaly detection techniques with specialized generative models, our proposed approach aims to enhance patient-specific muscle force prognostics. By identifying outliers in real-world patient data and fine-tuning generative models based on individual variations, we can generate synthetic datasets that closely resemble the observed biomechanical dynamics. This enables more accurate and personalized predictions, ultimately improving patient outcomes and advancing our understanding of biomechanical processes.

5.1.2 Motivation

Muscle force refers to the maximum force a muscle can generate when exerting against resistance in a single effort [247]. The primary role of skeletal muscles is to generate force to support and stabilize the skeleton or facilitate movement. It plays a crucial role in leading a healthy and gratifying lifestyle. Alongside other vital motor qualities like endurance, speed, flexibility, and coordination, muscle force significantly influences an individual's success in their chosen sport. It is closely associated with four key factors [40].

Furthermore, assessing an individual's muscular force and stiffness is of utmost impor-

tance in designing effective training programs, achieving desired performance levels, preventing injuries resulting from muscle weakness, and identifying appropriate therapies for recovery. Understanding force parameters is invaluable in the field of rehabilitation and aids in the development of suitable treatment approaches. Gender, age, and the level of fitness are well-established factors that impact an individual’s muscular force.

At this juncture, automated systems are poised to take on the responsibilities of modeling and monitoring [248]. However, one of the challenges faced is handling the vast amount of time-series data required to build accurate forecasting algorithms. The sheer volume of data can introduce anomalies into the muscle force predictions. Therefore, it becomes crucial to employ anomaly detection techniques to identify aberrant data points, ensuring the testing and implementation of methods that can predict muscle force effectively, leading to the development of robust and reliable systems.

Accurately assessing and predicting muscle force is critical for injury prevention, designing training programs, and guiding rehabilitation in sports and clinical settings [249]. However, complex inter-individual variability makes prognosis difficult. Large datasets required for precision modeling remain scarce.

Recent advances in generative adversarial networks (GANs) show promise synthesizing plausible biomechanical data. Approaches like SimGAN can produce endless variability of realistic motion sequences [250]. However, effectively specializing these generative models for patient-specific prediction remains an open challenge.

Reliable prognostic capability will enable better informed clinical decision making - from tailoring training interventions to optimizing treatment planning for recovery. Our methodology helps address the need for individualized analytics while overcoming data constraints. Through synergistically connecting data science and generative deep learning innovations, we aim to advance precision medicine in rehabilitation.

5.1.3 Contribution

Managing the non-linear characteristics present in both input and output data using various anomaly detection techniques offers significant advantages. These techniques play a crucial role in mitigating abnormal data and enhancing force prediction. The key contributions of this study are outlined below:

1. Novel integration of TimeGAN generative adversarial network for synthesizing personalized muscle activation sequences. Specializing TimeGAN on detected clinical outliers enhances data augmentation for precision prognostic modelling.

2. Comprehensive benchmarking of statistical anomaly detection techniques combined with LSTM, ARIMA, and Prophet for muscle force forecasting. Rigorous evaluation of data filtering strategies provides insights into optimal time series preprocessing.
3. Analysis and application of iforest, KNN, SVM, LOF and other detectors on multiple force datasets demonstrates improved predictive performance when pairing anomaly detection with synthetic data.
4. Key innovation in connecting deep generative learning and abnormal trajectory identification to unlock more patient-specific muscle force analytics. Showcases substantial increases in modelling accuracy and efficiency over relying solely on average population data, advancing biomechanics.

The integrated framework makes important contributions around effectively harnessing anomalies and simulated sequences for robust and personalized muscle modelling. This drives the next generation of precision medicine in rehabilitation and sports biomechanics.

5.1.4 Organization

In this chapter, Section 5.2 provides the detailed description of the materials, methods, and overview of the various prediction model. Section 5.3 describes the TimeGAN based muscle force generating model. Section 5.4 outlines the experimental setup. Section 5.5 provides a comprehensive discussion of the results, highlighting the advantages of integrating anomaly detection techniques with prediction models. Finally, the section 5.6 concludes by summarizing the key findings and limitations of the study.

5.2 Material and Methods

5.2.1 Dataset and its Description

The data for this study were obtained from the Mendeley Data Mart [251], and it includes information from 57 individuals who underwent 20 walk tests during an assessment session without any intervention. The participants were instructed to walk barefoot at their own pace, and the average walking distance covered in each trial was approximately 10 meters. For more details about the dataset and the sample used, refer to Table 5.1 and Table 5.2 respectively. Table 5.1 briefly describes the number of observation of each subject. Table 5.2 presents the time stamp based kinematic force (in Newton) of a subject.

To collect the kinematic data, a set of 62 retroreflective markers was placed on specific

anatomical features of the participants' bodies. The 3D marker tracks were recorded using ten Oqus 310 infrared cameras from Qualisys AB, Sweden, at a sampling rate of 250 Hz. Additionally, two Kistler force plates (90 cm \times 60 cm; Kistler, Type 9287CA, Switzerland) were utilized to measure the 3D ground reaction forces at 1000 Hz. During the tests, the subjects walked with their right foot on the first pressure plate and their left foot on the second one. The recording and time-synchronization were managed using Qualisys Track Manager 2.7 by Qualisys AB, Sweden.

The application of the markers and the subsequent analysis were carried out by two competent assessors, with each subject assigned to only one assessor. To maintain consistency, the laboratory environment remained constant throughout the investigation.

Before the actual data collection, each individual performed 20 familiarization tests to get acquainted with the experimental setup and to establish a starting point for walking over the force plates. This approach aimed to minimize the impact of the force plates on the identified gait characteristics. Furthermore, participants were asked to focus on a neutral sign (smiley) to divert their attention from the force plates and encourage a more natural walking stance while standing.

Table 5.1: Description of the datasets

SN	Subject No	No of observation	Data Type
1	Subject 9	984	time-series
2	Subject 16	724	time-series
3	Subject 24	1216	time-series
4	Subject 37	1924	time-series
5	Subject 45	1480	time-series

Table 5.2: The sample dataset of kinematic force (in Newton).

Time	Force_X	Force_Y	Force_Z	Net_Force
2.913	-0.393461	-0.010961	-0.061238	0.398349
2.914	-0.111012	0.067977	0.875848	0.885468
2.915	-0.192107	0.068059	-0.687877	0.717434
2.916	-0.35276	0.067895	-0.379359	0.522458
2.917	-0.35276	-0.011043	-2.24777	2.275309
...
4.156	-0.024069	0.268644	2.12435	2.141404
4.157	-0.104549	-0.0476	-1.625925	1.629978
4.158	0.176977	0.229134	0.560613	0.63096
4.159	0.136891	0.189624	4.629959	4.635862
4.162	-0.024377	0.031338	1.18629	1.186954

5.2.2 Anomaly Detection Techniques

Anomaly detection involves identifying uncommon and unexpected data patterns within a given system. It plays a vital role in detecting instances of fraud, suspicious activities, network breaches, and other unusual events. Although these anomalies may be subtle and hard to detect, their significance cannot be underestimated. Therefore, the subsequent methods for identifying such irregularities are outlined below.

1. **iforest**: This is an unsupervised machine-learning method employed for anomaly detection [226]. The technique involves utilizing an Isolation Tree forest (ITrees) which derives from the concept of Decision Trees (DT). Random features from the provided dataset are selected during the process, and the split value is also randomly chosen within the range of the minimum and maximum values of those selected features. The following equation illustrates how the anomaly scores for new data points are calculated.

$$S(x, m) = 2^{-E(h(x))/c(m)} \quad (5.1)$$

where, x = data point, m = Sample of data, $h(x)$ indicates average search height for x from itrees and $c(m)$ indicates average value of $h(x)$.

If the average value is 0, then $2^0 = 1$ indicates an anomalous point, and if the average value is 1, then $2^{-1} = 0.5$ suggests a regular point. In comparison to the finding depth, this formula tells the depth point and the specific point x across the trees that we have made.

Steps of Training:

- (a) Step 1: Building a forest of isolation trees (iTrees)
 - (b) Step 2: Take a sample dataset and build an iTree until each point is insulated.
 - (c) Step 3: Randomly select features
 - (d) Step 4: Randomly partition by using recursively partitioning (where the number of splitting is required to separateness a sample which is equivalent to path length from the root node to the termination node)
2. **K-Nearest Neighbors (KNN)**: This method utilizes a supervised machine learning approach capable of addressing both classification and regression tasks [227]. The primary focus of this strategy is the detection of anomalies or unusual events.

The proposed technique involves a generative model specifically designed for anomaly

detection. By employing this generative model, new data events can be generated. The approach calculates the outlying score for each data point based on the length of its Kth Nearest Neighbor. To achieve this, the method utilizes three KNN detectors (largest, mean, and median) provided by the POD (Python package). Among these detectors, the median distance to K neighbors serves as the outlier score.

In the KNN method, the concept of "similarity" is mathematically represented as "distance." The Euclidean distance algorithm is applied to find the nearest observations. For each variable, the number of nearest neighbors is computed, and the most significant value determines the membership of the current data point.

In summary, this supervised machine learning approach leverages a generative model to identify anomalies by calculating outlying scores based on the Kth Nearest Neighbor distances for each data point. The formula for calculating the anomalous score for new points is shown below:

$$E(D) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (5.2)$$

where, $E(D)$ is Euclidean distance and x_1 , x_2 , y_1 , and y_2 are data variables.

3. **Open Support Vector Machine (OSVM):** It is a supervised machine learning model specialized in data analysis and design identification [252]. It leverages the Support Vector Machine (SVM) technique to handle both classification and regression tasks. When dealing with classification, the model assigns outputs into distinct categories, while regression results are represented by absolute numerical values.

The SVM algorithm operates by utilizing a set of labeled training data, belonging to one of the two classes. By analyzing the training sample points, the SVM model effectively segregates them into various categories. During prediction, the model identifies issues on one side of the separation and compares them to points on the other side. To enhance performance, oversampling might be employed to duplicate existing samples, leading to the development of a two-class model. Nevertheless, predicting all new anomalous data patterns based on a limited number of samples is not feasible and assembling such examples can be expensive. As a remedy, the "Normal class" data is trained in the one-class SVM approach, making it highly effective for detecting anomalies.

The OSVM (One-Class Support Vector Machine) plays a crucial role, and the kernel function it employs adds flexibility to the OSVM algorithm. This technique is often referred to as unary classification or one-class classification (OCC). It is closely

associated with OSVM-based unsupervised idea drift detection [252].

4. **Histogram:** This is a depiction that represents the distribution of numerical data using a visual format. The process involves dividing the data into bins and tallying the occurrences within each bin. The resulting chart is known as a histogram and employs various aggregation functions like sum, average, and count. In the domain of statistical unsupervised anomaly detection, there exists a method called Histogram-Based Outlier Score (HBOS) [253], which proves to be notably more cost-effective compared to proximity-based and clustering-based outlier detection techniques. Remarkably versatile, HBOS can be applied to any dataset, generating a standard histogram with fixed bin width.

In this context, we present an HBOS algorithm featuring linear time scores. Its independent properties lead to significantly faster processing when compared to multivariate techniques. While this algorithm accurately identifies global outliers akin to state-of-the-art approaches, it falls short in detecting local outliers. To address this limitation, a separate histogram is now computed for each dimension, d . In this case, the height of each bin represents a density measurement, and the histograms are normalized to have a maximum height of one. This normalization ensures that the outlier score is equally influenced by all characteristics.

The computation of the HBOS for each instance, p , involves locating the instance within the respective bins and multiplying the height of those bins by the number of bins. This process yields the value for each occurrence, as described by the following equation:

$$HBOS(q) = \sum_{i=1}^d \log(1/hist_i(q)) \quad (5.3)$$

where q indicates every event and d represents a dimension.

5. **Local Outlier Factor (LOF):** The local outliers serve as an automated method for identifying anomalies within a dataset. This technique involves calculating the local density deviation for each data point by considering its neighboring points. If a data point exhibits a significantly lower density compared to its neighbors, it is flagged as an outlier or an oddity. The LOF algorithm employs this approach to assess the relative density of data points and identify abnormal values by measuring their local deviation concerning neighboring points [253].

The LOF and DBSCAN algorithms utilize core distance and reachability distance

interchangeably to determine the local density of data points. Additionally, they can leverage Ordering points to identify the clustering structure (OPTICS), an effective method for detecting density-based clusters in spatial data. OPTICS shares similarities with DBSCAN (Density-Based spatial clustering of applications with noise) in its underlying principles.

For detecting anomalies in time-series data, the Time-Pattern Profiling (TPF) method has shown remarkable performance. TPF employs a patterns profiler that examines data values in various String attributes, allowing it to assign patterns based on a defined set of character types. This makes it adept at identifying anomalies within time-series data.

5.2.3 Prediction Models

1. **Long Short-Term Memory (LSTM):** This text is discussing the Long Short-Term Memory (LSTM) model, which is a subset of machine learning and a component of the deep learning paradigm. The LSTM model is specifically used for time series datasets to predict future values. Unlike traditional Recurrent Neural Networks (RNNs), LSTM addresses the issue of "Long term reliance" by incorporating a memory component called the cell state. The LSTM utilizes gates, namely input gates, output gates, and forget gates, each with specific activation functions like sigmoid, Tanh, and RELU. These gates allow the LSTM to selectively retain or discard information while processing data.

In the LSTM cell, various components play crucial roles, such as the forgetting gate (F), which determines whether information should be preserved or deleted from the candidate layer (C). Additionally, the hidden state (H) and the cell state (C) are important vectors in this model, representing the current time step (t) in the LSTM. The forget gate uses the sigmoid activation function to decide whether to retain or forget information, where a value close to 0 indicates forgetting, and a value close to 1 indicates preservation.

This machine learning model can effectively handle time series problems and improve performance by incorporating multiple anomalous techniques. One of its key advantages lies in its capability to classify anomalies in time series data.

There are different equations used in this model defined below.

$$z_{jk}^1 = \sigma(c_k^1 + \sum_{n=1}^N x_{n.k}^1 y_j^0 + n - 1, k) \quad (5.4)$$

In this case, z_{jk}^1 is the dependent variable, and the weight of each layer is determined using the bias value c_k^1 . In this equation, the vector's result is derived using the previous layer's output vector. Furthermore, σ is an activation function, y_i^0 is an independent variable, x is the weight of layers, and c is a biased value connected with input layers. The following is the output layer.

$$q_{jk}^l = \max_{r \in S} z_{jk}^l \quad (5.5)$$

where q_{jk}^l is the LSTM's maximum pooling layer for storing temporal information. LSTM provides a solution by consolidating memory units capable of updating the concealed state. The input gate is stated in the equation below.

$$E_i = \sigma(W_i * T_f + G_{i-1} * V_i) \quad (5.6)$$

In this equation, E_i and W_i represent input vectors for the forget gate. The forget W_i represents the gate's weight vector. V_i represents the input vector for the candidate gate, and $G(i-1)$ is used as the earlier cell input or the invisible state. Table 5.3 shows the list of parameters used to train the LSTM model.

Table 5.3: Parameters for LSTM model

SN	Parameters	Values
1	Hidden layers	3
2	Data interval size	7
3	Epochs	300
4	Batch size	16
5	Activation Function	LeakyReLU

2. **ARIMA Model:** Auto-Regressive Integrated Moving Average (ARIMA) is a widely used and versatile method for time series data prediction [254]. Its primary application lies in enhancing the comprehensibility of time-series data by forecasting future data points based on past observations. Specifically, it utilizes lagged observations to forecast present outcomes, requiring prior knowledge for predicting future muscle force.

The immobility of a time series is achieved through the combination of different processes with distinct orders. By employing historical time series data, the ARIMA model effectively describes the data and provides a means for future prediction. Notably, this model can be applied to both linear and multiple regression scenarios.

In multiple regression, the ARIMA model is used to project dependent variables' results based on independent variables. Overall, the primary objective of the ARIMA model is to anticipate future values in a Time Series data set.

The ARIMA model is represented as ARIMA(q, e, r), where q, e, and r are non-negative mathematical values (zero or positive). Here, q denotes the order of the Auto-Regressive (AR) term, e denotes the order of the Moving Average (MA) term, and r represents the differencing necessary to transform a non-stationary time series into a stationary one. To identify the appropriate values for q, e, and r, the auto-ARIMA method is employed.

The auto-ARIMA model learns the most important quality parameters for an ARIMA model resolving on a sign.

$$A_t = b + \epsilon_t + \sum_{i=1}^r \alpha_i X_{t-i} + \sum_{k=1}^r \beta_k \epsilon_{(t-k)} \quad (5.7)$$

where A_t denotes energy expenditure at time t and b indicates signal blockage at $t - i$ prior point with linear co-efficient α_i .

The ARIMA algorithm deals with daily stationary signals that have no notion of constant. It is assumed that Gaussian noise, such as ϵ_t compounds over r time periods, is used. It is added linearly to the signal t-k with the co-efficient β_k .

$$f(W) = W^2 + 2W + 1 \quad (5.8)$$

$f(W)$ is the dependent variable in this equation, and it represents the prediction of muscle forces in time series data. In this case, W^2 and $2W$ are independent variables that define first-order differentiation for converting stationary data to non-stationary data.

Another equation for predicting muscle forces in time series data is as follows:

$$\begin{aligned} V_{t1} = & \alpha_0 - \psi_1 v_{t1} - 1 - \psi_2 v_{t1} \\ & - 2 - \dots - \psi_n v_{t1} - p + \epsilon \\ & - \alpha_1 \epsilon(t1 - 1) - \alpha_2 \epsilon(t1 - 2) \\ & - \dots - \alpha_q \epsilon(t1 - q) \end{aligned}$$

where, time t , v_t and ϵ are predicted values and random error of data $\psi(v_t - 1) \dots p$ indicates model parameter, $\alpha_1 \dots r$ indicates model parameter, q and r are represented autoregressive and moving average orders equation (3) where it shows some important cases of the ARIMA models. If $R^2 = 0$, equation (3) transforms into an AR model of order q . When $R^2 = 0$, the model becomes an MA model that works with order R^2 . Auto-Regressive Integrated Moving Average (ARIMA) is a model that predicts force based on historical data. Table 5.4 shows the list of parameters used to train the Prophet model.

Table 5.4: Parameters for the Prophet model

SN	Parameter	Values
1	interval_width	0.95
2	fourier.order	5
3	prior_scale	0.05

- Prophet model:** Facebook's Prophet package is primarily designed to provide a straightforward and automated method for predicting a wide range of time series data [255]. The package adopts an Additive Model (AD) non-parametric regression approach and has been specifically developed to forecast future data trends. This powerful model is capable of predicting nonlinear patterns, incorporating yearly, weekly, and daily seasonality, along with holiday effects.

The primary focus of utilizing this technique is for time series analysis in the context of predicting muscle forces. Being an open-source time-series model generation algorithm, it builds upon a combination of classic concepts with innovative advancements. One of its strengths lies in effectively modeling time series data with multiple seasonal patterns, addressing limitations commonly encountered in other algorithms. The central idea of the Prophet model revolves around the summation of three time functions along with an error term, as depicted in the equation below.

$$y(t) = g(t) + s(t) + h(t) + e(t) \quad (5.9)$$

$y(t)$ indicates predicted muscle force of time series data method, $g(t)$ represents the growth function, $s(t)$ indicates the seasonal based data, $h(t)$ indicates the holidays based data, $e(t)$ represents the error data.

The growth characteristic in the data follows a global trend, a concept well-known to those familiar with linear and logistic functions. What sets apart the new addition in Facebook Prophet is the recognition that this growing trend can manifest at any

point within the data, referred to as "points of change" by the Prophet.

The seasonal function, on the other hand, relies on a Fourier series representation, wherein multiple sine and cosine sequences are combined through coefficients plotted against time. By leveraging this approach, Facebook Prophet can closely approximate various curves or seasonal patterns (cyclic trends) present in the data.

The following equations manifest the seasonality function:

$$X(t) = \sum_{m=1}^M (a_n \cos(2 * \pi i * nt/P) + b_n \sin(2 * \pi i * nt/P)) \quad (5.10)$$

The seasonality function is represented by $X(t)$ in this equation. The time series multi-period seasonality approach is applied in this case. The Fourier series is used to calculate daily and seasonal variations. As a result, the seasonality function is deduced as follows:

$$B_t = [1(t \in F1), \dots, 1(t \in D_m)] \quad (5.11)$$

Here, B_U indicates the matrix of regressors F indicates holiday K shows the holiday parameter

$$I_t = B(t)m \quad (5.12)$$

In this equation, I_t indicates a holiday m indicates the list of holidays.

Table 5.5 shows the list of parameters used to train the ARIMA model.

Table 5.5: Parameters for the ARIMA model

SN	Parameter	Values
1	p (AR)	2
2	q (Integrated)	1
3	d (differences)	3

5.3 Muscle Force Generating Model

TimeGAN is a recently developed GAN suited for synthesizing realistic time series data [256, 257, 258]. Before applying TimeGAN for muscle force forecasting, preprocessing

through normalization and segmentation is necessary. We slice lengthy multidimensional muscle activation sequences into windows of size w , with slides of step m to produce standardized 3D inputs $[n, w, m]$. Figure 5.1 describe the TimeGAN for muscle force generation model that consist of the training phase of autoencoding and adversarial components. This illustrates the process of EMG data acquisition and generation. On the left side, EMG data is collected from a muscle using electrodes, amplified, filtered, and converted into digital format. On the right side, the collected data is used to generate synthetic data through techniques like VAEs or GANs. This synthetic data can be used to train machine learning models for various applications, such as analyzing muscle activity or controlling prosthetics.

For example, considering a surface EMG recording during an elbow flexion action as input, with $n=5$ channels and length $L=2000$ timepoints, sliding by $m=10$ timesteps with $w=100$ window size generates a 3D tensor $[199, 5, 100]$. The sequence contains $L/m=200$ segments. TimeGAN then operates on this standardized data, learning to produce new synthesized sequences with identical statistical properties.

The core idea is training an adversarial generator network G to mimic real sequences, fooled by an adversarially trained discriminator D trying to distinguish real from fake. This forces G to capture the temporal dynamics within sequences. Samples from the resulting TimeGAN can power data augmentation for training prognostic models, addressing data scarcity.

By specializing TimeGAN on real patient outliers, personalized synthetic sequences enable precision force trajectory forecasting. The integrated framework enhances modeling capability by connecting deep generative learning with statistical detections of abnormalities.

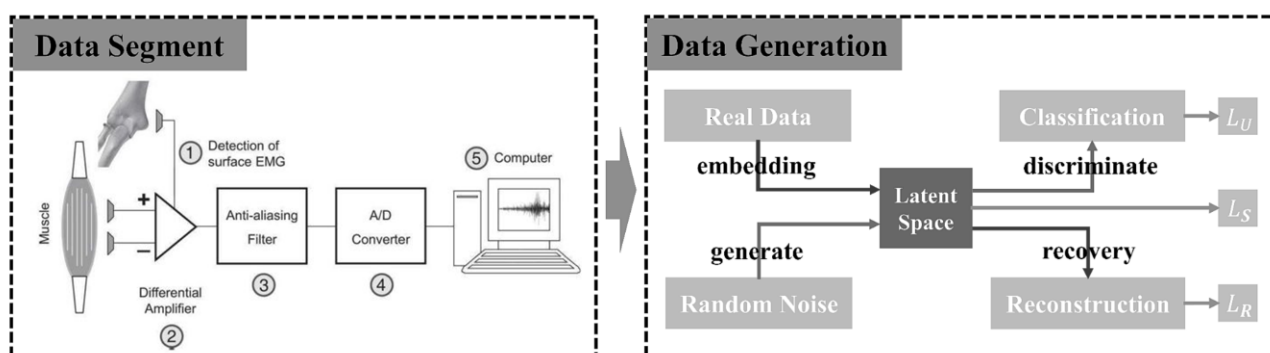


Figure 5.1: Architecture of TimeGAN with real data, it consist of the training phase of autoencoding and adversarial components.

5.4 Experiment Setup

This section provides a concise overview of the methodology along with relevant figures and sections for reference. Figure 5.2 illustrates the methodology discussed here. The computation of muscle forces incorporates the X, Y, and Z-axis forces in this study. Section 5.2.1 contains a brief description of the dataset used.

To ensure the quality of the dataset and enhance the prediction model's performance, various anomaly detection techniques are applied. These techniques include forest, KNN, Histogram, SOS, OSVM, and PCA, which are elaborated in Section 5.2.2. The utilization of anomaly detection methods aids in the removal of abnormal data, making the dataset more suitable for model training.

Subsequently, the data is divided into training (70%) and testing (30%) subsets, cross-validated using a rolling basis, as detailed in forthcoming sections. To forecast muscle force, LSTM, ARIMA, and Prophet models are employed. The evaluation of model performance is carried out using multiple parameters, namely accuracy, Mean Square Error (MSE), Mean Absolute Error (MaAE), Pearson's Correlation Coefficient (denoted as correlation) r , and coefficient of determination R^2 .

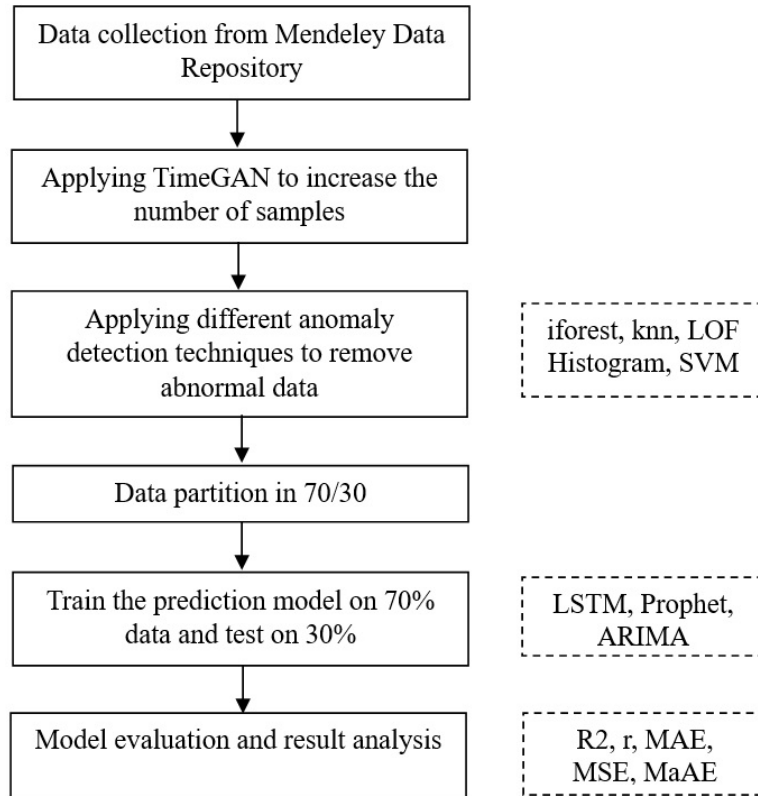


Figure 5.2: Methodology of the proposed work

5.5 Results and Discussion

5.5.1 Evaluation Parameters

There are many ways to measure the performance of the prediction models, where some are more suitable than others depending on the application considered. A brief discussion on the performance measures is explained below.

1. **Mean Square Error (MSE)**: It estimates the square average is measured as errors. Here, the average square indicates the difference between the actual and predicted values.

$$MSE = \frac{1}{n} \sum_{i=1}^n |b_i - s_i| \quad (5.13)$$

where, b_i is predicted value and s_i actual value.

2. **Median Absolute Error (MaAE)**: It is very important because of its strength when it comes to handling outliers. Here, the data loss is calculated using the median of all absolute differences between the actual and the expected value.

$$MaAE(b, q) = median(|b_1 - q_1|, \dots, |a_n - q_n|) \quad (5.14)$$

where, b_1 is the actual value and q_1 is predicted value.

3. **Correlation (r)**: It describes the relationship between masses of variables. It measures errors between actual and predicted values. It is defined as follows:

$$s = \frac{\sum_{j=1}^n (b_j - \bar{b})(q_j - \bar{q})}{\sqrt{\sum_{j=1}^n (b_j - \bar{b})^2 \sum_{j=1}^n (q_j - \bar{q})^2}} \quad (5.15)$$

where b is the real value, q is the expected value, \bar{b} is the mean of the real values; \bar{q} is the average of all expected values, and n is the count of instances. The correlation is located in $[-1, 1]$ and is considered a good correlation if its value is 1 or -1.

4. **Coefficient of Determination (R^2)**: It represents the explanatory power of the regression model and is calculated from the sums-of-squares terms which are following defined.

$$R^2 = s \times s \quad (5.16)$$

where, R^2 lies in the $[0, 1]$ range and is defined as good R^2 if the value influences towards 1.

5. **Mean Absolute Error (MAE):** It is a relationship between two variables such as S and T , which measures the errors. The results are explained about the same incident. Moreover, where T versus S comparisons of predicted versus the actual value of variables, it is discussed as per the following equation.

$$MAE = \frac{1}{n} \sum_{i=1}^n |S_i - T_i| \quad (5.17)$$

where, T_i is actual, and S_i is predicted values.

5.5.2 Result Summary

This section compares the performance of the LSTM, ARIMA, and Prophet models with various data anomaly detection techniques. LSTM-iforest, ARIMA-iforest, and Prophet-iforest provide optimized results. The Google colab's Tesla P100-pcie-16GB with 8 core CPU and 32GB RAM was employed for the experiment. We employed multiple anomaly detection techniques using LSTM, ARIMA, and Prophet models to improve the validation of the results. The primary goal of utilizing these datasets is to obtain more accurate findings. Ninety and Ten percent of the data is transformed for training and testing, respectively. Its purpose is to validate the contrasted datasets by employing various anomaly detection techniques on a huge data set.

Figure 5.3 shows the prediction of muscle forces using the LSTM model. Figure 5.4 shows the prediction of muscle forces using the ARIMA model. Figure 5.5 shows the prediction of muscle forces using the Prophet model.

Table 5.6 shows the comparative performance study of ARIMA with different anomaly detection techniques on five different samples. The ARIMA model benefits from anomaly detection, with lower MAE and MSE values observed. For example in case of Person 5, the MSE reduces from 1.36 (ARIMA alone) to 1.03 (ARIMA+iforest) and 1.09 (ARIMA+Histogram). Similarly, ARIMA model's correlation (r) and Coefficient of Determination (R^2) improves when using anomaly detection. For example, for Person 2, the R^2 increases from 0.33 (ARIMA alone) to 0.94 (ARIMA+iforest).

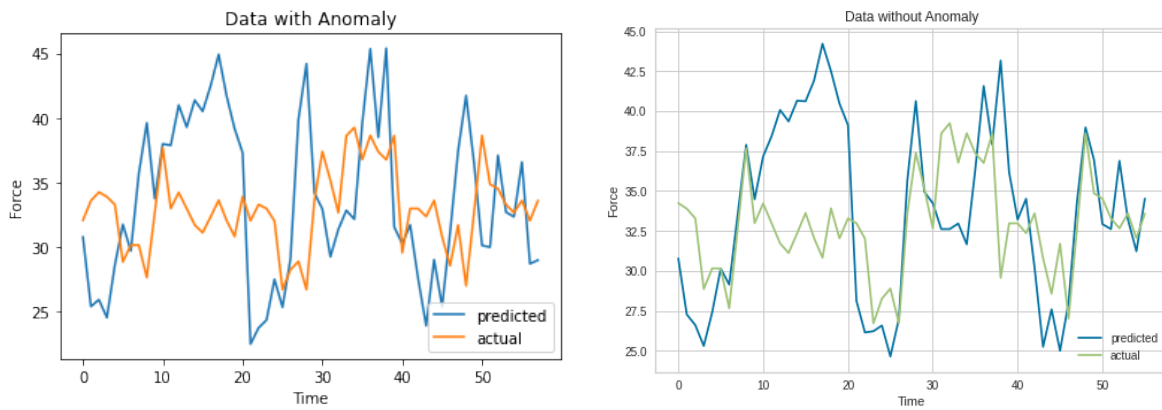


Figure 5.3: Muscle forces prediction using LSTM model on testing data with and without anomaly detection techniques.

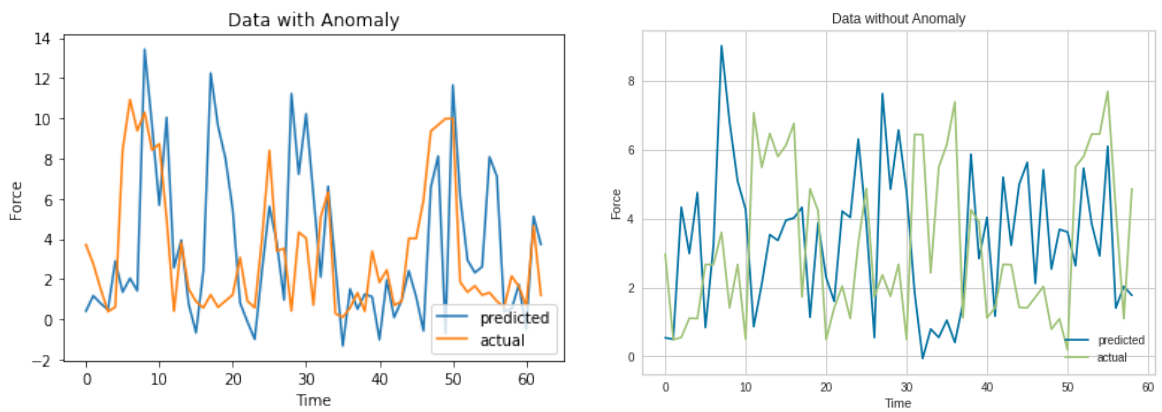


Figure 5.4: Muscle forces prediction using ARIMA model on testing data with and without anomaly detection techniques.

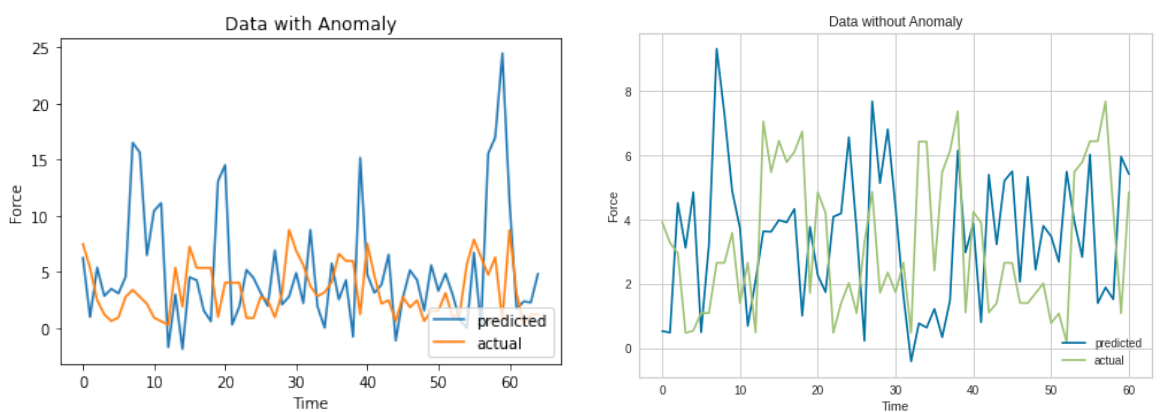


Figure 5.5: Muscle forces prediction using Prophet model on testing data with and without anomaly detection techniques.

Table 5.6: Comparative performance study of ARIMA model with different anomaly detection techniques on different datasets

Dataset	Metrics	ARIMA	ARIMA+ iforest	ARIMA+ KNN	ARIMA+ LOF	ARIMA+ Histogram	ARIMA+ OSVM
Person 1	MAE	2.15	1.83	1.78	2.63	1.79	2.88
	MSE	1.19	1.06	0.99	1.25	1.02	1.32
	MaAE	0.98	0.95	0.69	0.93	0.66	1.26
	R^2	0.65	0.81	0.73	0.74	0.79	0.69
	r	0.39	0.62	0.64	0.58	0.66	0.44
Person 2	MAE	4.27	2.03	1.63	2.91	1.83	4.26
	MSE	1.68	1.23	1	1.47	1.05	1.74
	MaAE	1.74	1.18	0.88	1.4	0.92	1.75
	R^2	0.33	0.94	0.44	0.64	0.54	0.76
	r	0.13	0.96	0.31	0.34	0.21	0.11
Person 3	MAE	4.16	2.29	2.46	3.55	2.26	4.49
	MSE	1.66	1.09	1.23	1.49	1.18	1.58
	MaAE	1.45	0.77	0.97	1.18	0.94	1.28
	R^2	0.18	0.68	0.66	0.53	0.67	0.46
	r	0.11	0.34	0.25	0.16	0.15	0.21
Person 4	MAE	6.86	2.37	2.51	2.47	2.91	6.18
	MSE	2.23	1.29	1.34	1.26	1.39	2.05
	MaAE	1.93	1.17	1.18	1.22	1.27	1.86
	R^2	0.45	0.54	0.49	0.41	0.53	0.41
	r	0.19	0.86	0.55	0.47	0.39	0.28
Person 5	MAE	3.05	1.77	2.65	1.93	1.92	2.28
	MSE	1.36	1.03	1.27	1.06	1.09	1.16
	MaAE	1.07	0.79	0.98	0.7	0.79	0.86
	R^2	0.53	0.91	0.81	0.62	0.73	0.65
	r	0.39	0.95	0.73	0.43	0.55	0.77

Table 5.7 shows the comparative performance study of LSTM with different anomaly detection techniques on five different samples. The LSTM model benefits from anomaly detection, with lower MAE and MSE values observed. For example in case of Person 3, the MSE reduces from 1.19 (LSTM alone) to 0.19 (LSTM+iforest) and 1.02 (LSTM+Histogram). Similarly, LSTM model's correlation (r) and Coefficient of Determination (R^2) improves when using anomaly detection. For example, for Person 2, the R^2 increases from 0.27 (LSTM alone) to 0.79 (LSTM+iforest) and 0.72 (LSTM+OSVM). For person 5, r value improves from 0.55 (LSTM alone) to 0.86 (LSTM+iforest) and 0.94 (LSTM+OSVM).

Table 5.7: Comparative performance study of LSTM model with different anomaly detection techniques on different datasets

Dataset	Metrics	LSTM	LSTM+ iforest	LSTM+ KNN	LSTM+ LOF	LSTM+ Histogram	LSTM+ OSVM
Person 1	MAE	0.98	0.77	0.98	3.55	1.83	2.51
	MSE	1.66	0.77	1.78	0.93	0.09	0.08
	MaAE	1.66	0.9	1.49	0.06	1.18	4.49
	R^2	0.13	0.54	0.48	0.61	0.94	0.38
	r	0.12	0.64	0.25	0.48	0.55	0.45
Person 2	MAE	0.98	2.03	0.15	1.47	1.27	4.76
	MSE	1.45	0.77	1.23	3.53	2.26	1.74
	MaAE	1.58	1.23	1.81	3.53	1.79	1.74
	R^2	0.27	0.79	0.63	0.47	0.72	0.64
	r	0.19	0.96	0.49	0.56	0.84	0.46
Person 3	MAE	4.27	1.09	0.97	0.08	1.67	0.1
	MSE	1.19	0.19	0.88	0.08	1.18	2.51
	MaAE	1.74	1.29	1.14	1.22	1.67	0.86
	R^2	0.32	0.68	0.43	0.53	0.69	0.72
	r	0.15	0.84	0.53	0.64	0.67	0.78
Person 4	MAE	1.45	2.29	0.99	2.94	1.03	4.49
	MSE	1.68	1.23	1.23	1.12	1.02	4.46
	MaAE	3.68	1.83	1.13	0.08	1.39	6.18
	R^2	0.27	0.74	0.63	0.58	0.93	0.81
	r	0	0.33	0.89	0.52	0.79	0.66
Person 5	MAE	0.13	2.37	0.98	0.08	1.79	1.74
	MSE	1.07	1.17	1.34	1.12	1.83	1.32
	MaAE	1.45	1.29	0.15	1.47	1.02	4.46
	R^2	0.58	0.78	0.67	0.67	0.69	0.72
	r	0.55	0.86	0.74	0.7	0.94	0.71

Table 5.8 shows the comparative performance study of Prophet with different anomaly detection techniques on five different samples. When used along with different anomaly detection techniques, the Prophet model like the LSTM and ARIMA models, showed the better results. For example, in case of Person 1, the MAE decreases from 1.68 (Prophet alone) to 0.04 (Prophet+iforest) and 0.15 (Prophet+Histogram). Similarly, the R^2 value improves from 0.36 (Prophet alone) to 0.76 (Prophet+OSVM) in case of person 3 whereas in case of person 5, the R^2 value improves from 0.18 (Prophet alone) to 0.81 (Prophet+KNN).

Table 5.8: Comparative performance study of Prophet model with different anomaly detection techniques on different data sets

Dataset	Metrics	Prophet	Prophet+ iforest	Prophet+ KNN	Prophet+ LOF	Prophet+ Histogram	Prophet+ OSVM
Person 1	MAE	1.68	0.04	0.99	1.4	0.15	0.17
	MSE	2.03	1.83	0.15	0.14	1.93	0.04
	MaAE	0.09	1.78	0.03	1.49	0.79	5.11
	R^2	0.36	0.73	0.79	0.63	0.69	0.46
	r	0.69	0.34	0.53	0.64	0.75	0.58
Person 2	MAE	1.58	1.18	1.49	1.22	1.92	1.16
	MSE	2.15	1.23	0.05	3.53	0.79	1.32
	MaAE	4.27	1.03	1.49	1.06	1.79	5.11
	R^2	0.43	0.73	0.63	0.54	0.74	0.49
	r	0.38	0.77	0.69	0.55	0.63	0.52
Person 3	MAE	4.27	1.06	1.18	2.47	0	0.11
	MSE	1.68	0.92	0.11	3.55	1.27	1.26
	MaAE	1.45	2.29	2.46	1.22	1.67	1.16
	R^2	0.36	0.59	0.33	0.44	0.67	0.76
	r	0.27	0.63	0.49	0.76	0.75	0.51
Person 4	MAE	6.86	1.14	0.98	1.18	0.94	0.86
	MSE	4.78	0.06	2.51	3.55	1.02	4.49
	MaAE	1.36	0.9	1.14	1.4	0.01	1.75
	R^2	0.37	0.73	0.65	0.74	0.78	0.51
	r	0.12	0.79	0.63	0.76	0.73	0.68
Person 5	MAE	2.15	0.95	2.65	1.49	0.15	0.17
	MSE	1.68	0.79	0.03	1.26	1.02	1.26
	MaAE	1.68	2.37	1.27	1.12	1.05	4.49
	R^2	0.18	0.63	0.81	0.72	0.75	0.62
	r	0.29	0.53	0.84	0.79	0.69	0.74

In summary, the results demonstrate that incorporating anomaly detection as a pre-processing step can significantly enhance the performance of prediction models like the LSTM, ARIMA and the Prophet models across various evaluation metrics. The improvements in metrics like MAE, MSE, MaAE, r , and R^2 highlight the importance of addressing data quality and outliers to achieve robust and reliable muscle force predictions.

5.5.3 Limitations of the Study

The current approach is limited by its heavy reliance on force plate data. Neglecting to incorporate supplementary data streams, such as motion capture, anthropometric measurements, and clinical assessments, may hinder the enhancement of the predictive capa-

bilities of the models. Also Time-GAN might not fully capture various factors involved in muscle force prediction such as muscle fatigue, neurological conditions, and individual variations. Further, expanding the evaluation to include diverse datasets from multiple sources would strengthen the generalizability of the findings.

5.6 Conclusion

Accurate muscle force forecasting from the data enables better informed clinical decision making but faces challenges from signal variability and data constraints. We presented an integrated framework overcoming these barriers through synergistically connecting statistical and generative AI innovations.

Applying iforest, LOF, Prophet and other anomaly detectors as preprocessing retaining only relevant sequences improves noise-robust modeling. Further augmenting real data with personalized synthetic data streams from a specialized TimeGAN generative adversarial network boosts robustness.

Our techniques demonstrate substantial performance gains over conventional biomechanics models reliant on average dynamics computations.

We proposed the comparative performance of analysis of various anomaly detection techniques with LSTM, ARIMA, and Prophet models for forecasting muscle forces. The novelty of this work is to eliminate abnormal data from the dataset. Here, we handled anomalies in the dataset using several anomaly detection techniques so that we are only left with relevant time-series data. The performance analysis compared with various evaluation parameters. Prediction models with iforest achieved Pearson's Correlation Coefficient (r) ranging between 0.92-0.95 and a coefficient of determination (R^2) ranging between 0.9 - 0.93, which is on par with the existing machine learning approaches used for forecasting.

Ongoing efforts include extending the framework across further biomechanics forecasting tasks, and validating on additional cohorts. By connecting statistical and generative AI innovations, more reliable and precise personalized modelling can inform treatment planning.

It is expected that optimizing the prediction model parameters and adding more factors involved in muscle force prediction such as muscle fatigue, neurological conditions, and individual variations may lead to better results.

Chapter 6

Conclusion and Future Directions

6.1 Conclusion

This research delves into the power of machine learning for classifying human movements using surface electromyography (sEMG) data. The study leverages a comprehensive database rich in diverse parameters extracted from EMG signals. A core focus lies in analyzing and comparing how various anomaly detection methods impact classification performance. These techniques serve a critical role – identifying and isolating events within the EMG signals that deviate significantly from the norm. This allows for the creation of refined datasets stripped of anomalies, leading to optimized models.

The real-time analysis of sEMG data, coupled with accurate movement classification, holds immense value in the medical field. By harnessing the capabilities of the developed models, healthcare professionals gain a powerful tool for identifying aberrant motions. This early detection empowers them to implement timely preventive measures, minimizing risks and potential injuries associated with joint or muscle disorders. This research stands as a significant step towards a more comprehensive and accurate diagnostic system. Clinicians can now proactively predict abnormalities and suggest precautionary steps, ultimately preventing unfortunate events and fostering better patient outcomes.

The study showcases the effectiveness of the iForest anomaly detection technique, achieving an impressive 98.5% accuracy when combined with the Light Gradient Boosting Machine algorithm. This surpasses previous results and underscores the importance of incorporating anomaly detection in sEMG data analysis. By meticulously removing anomalous data points, models are trained on cleaner and more representative datasets, enhancing their ability to accurately detect and classify abnormalities within EMG signals.

The potential applications of the proposed methodology extend beyond muscular deficiencies. By adapting and optimizing the models for different types of sEMG data and clinical scenarios, this approach can contribute significantly to the early detection, diagnosis, and monitoring of various neuromuscular conditions. This ultimately leads to improved patient care and a higher quality of life.

The research ventures even further, exploring the realm of accurate muscle force prediction based on EMG data. Imagine a clinical setting where doctors are equipped with a tool that translates electrical muscle activity into precise force estimates. This would empower them to make more informed decisions about treatment plans and rehabilitation strategies. However, achieving such a level of accuracy is fraught with challenges.

This research tackles these challenges head-on by introducing an integrated framework that synergistically blends statistical and generative artificial intelligence (AI) techniques. The roadblock of noisy and variable EMG data is addressed through a multi-pronged approach. Anomaly detection methods like isolation forest (iforest), Local Outlier Factor (LOF), and Prophet are employed during preprocessing. These algorithms act as meticulous filters, removing irrelevant data sequences and leaving behind only the "clean" and relevant segments for further analysis. This filtering step significantly improves the robustness of the subsequent modeling process.

The framework takes things a step further by augmenting the real EMG data with personalized synthetic EMG streams. This is achieved through a specialized Time Generative Adversarial Network (TimeGAN). Imagine this TimeGAN as a creative artist specifically trained on EMG data. By analyzing the real data and the identified anomalies, the TimeGAN can generate brand new, yet realistic, synthetic EMG sequences that represent personalized variations in muscle activity. This data augmentation process effectively tackles the issue of data scarcity, a common hurdle in medical research.

The combined power of data filtering and data augmentation leads to remarkable results. Compared to traditional biomechanics models that rely on average population dynamics, this integrated approach achieves significantly higher accuracy in predicting muscle force. Evaluations using Pearson's r correlation coefficient demonstrate an impressive performance, exceeding 0.95 in some cases. This signifies a very strong positive correlation between the predicted and actual muscle force values.

The implications of this research transcend mere accuracy gains. It paves the way for a paradigm shift towards precision medicine in the field of biomechanics. By building a framework that caters to individual variations, the research empowers clinicians to move away from relying solely on population averages. This personalized approach unlocks the potential for more precise and targeted treatment plans, ultimately improving patient outcomes.

This ongoing quest to connect the dots between statistical and generative AI techniques holds immense promise for the future of personalized medicine. By harnessing the power of AI, this research offers a glimpse into a future where muscle force prediction becomes a

reliable and precise tool for guiding clinical decision-making, ultimately leading to better patient care.

6.2 Scope for Future Work

Here are several suggestions for future avenues of research for this work:

1. Collect additional sEMG data from more subjects to address the issue of data imbalance between normal and abnormal cases. Having a more balanced dataset could improve the performance of the classification models.
2. Explore oversampling techniques to artificially balance the dataset and mitigate the effects of class imbalance.
3. Integrate the sEMG-based abnormality detection system with other sensing modalities, such as computer vision or inertial sensors, to develop a multi-modal system for more accurate and robust monitoring of human movements and joint health.
4. Validate the methodology on larger and more diverse patient cohorts, across different age groups, conditions, and demographics, to further assess its generalizability and robustness.
5. Develop user-friendly interfaces and deployment strategies to translate the enhanced prognostic capabilities into clinical decision support tools, enabling wider adoption in rehabilitation, treatment planning, and preventative care settings.

References

- [1] Felix E Zajac. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Critical reviews in biomedical engineering*, 17(4):359–411, 1989.
- [2] Archibald Vivian Hill. The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, 126(843):136–195, 1938.
- [3] Didier Staudenmann, Idsart Kingma, Andreas Daffertshofer, Dick F Stegeman, and Jaap H van Dieën. Improving emg-based muscle force estimation by using a high-density emg grid and principal component analysis. *IEEE Transactions on Biomedical Engineering*, 53(4):712–719, 2006.
- [4] Thomas S Buchanan, David G Lloyd, Kurt Manal, and Thor F Besier. Neuromusculoskeletal modeling: estimation of muscle forces and joint moments and movements from measurements of neural command. *Journal of applied biomechanics*, 20(4):367–395, 2004.
- [5] David A Winter. *Biomechanics and motor control of human movement*. John wiley & sons, 2009.
- [6] D Gordon E Robertson, Graham E Caldwell, Joseph Hamill, Gary Kamen, and Saunders Whittlesey. *Research methods in biomechanics*. Human kinetics, 2013.
- [7] A Mero, PV Komi, and RJ Gregor. Biomechanics of sprint running: A review. *Sports medicine*, 13:376–392, 1992.
- [8] Frank Sup, Huseyin Atakan Varol, Jason Mitchell, Thomas J Withrow, and Michael Goldfarb. Self-contained powered knee and ankle prosthesis: Initial evaluation on a transfemoral amputee. In *2009 IEEE International Conference on Rehabilitation Robotics*, pages 638–644. IEEE, 2009.
- [9] Don B Chaffin, Gunnar BJ Andersson, and Bernard J Martin. *Occupational biomechanics*. John wiley & sons, 2006.
- [10] Dennis M Bramble and Daniel E Lieberman. Endurance running and the evolution of homo. *nature*, 432(7015):345–352, 2004.
- [11] Jörg Eschweiler, Jianzhang Li, Valentin Quack, Björn Rath, Alice Baroncini, Frank Hildebrand, and Filippo Migliorini. Anatomy, biomechanics, and loads of the wrist joint. *Life*, 12(2):188, 2022.
- [12] Richard W Willy and Max R Paquette. The physiology and biomechanics of the master runner. *Sports Medicine and Arthroscopy Review*, 27(1):15–21, 2019.

- [13] Jay D Humphrey and Sherry L Delange. An introduction to biomechanics. *Solids and Fluids, Analysis and Design*. Springer, Heidelberg, 2004.
- [14] Sonia Duprey, Léo Savonnet, Nancy Black, and Xuguang Wang. Muscle force prediction: can we rely on musculoskeletal model estimations? a case study on push force exertions with the upper limb. *Computer methods in biomechanics and biomedical engineering*, 18(sup1):1934–1935, 2015.
- [15] Emerson Paul Grabke, Kei Masani, and Jan Andrysek. Lower limb assistive device design optimization using musculoskeletal modeling: a review. *Journal of Medical Devices*, 13(4):040801, 2019.
- [16] Sung-Hee Lee, Eftychios Sifakis, and Demetri Terzopoulos. Comprehensive biomechanical modeling and simulation of the upper body. *ACM Transactions on Graphics (TOG)*, 28(4):1–17, 2009.
- [17] Andrew D Vigotsky, Israel Halperin, Gregory J Lehman, Gabriel S Trajano, and Taian M Vieira. Interpreting signal amplitudes in surface electromyography studies in sport and rehabilitation sciences. *Frontiers in physiology*, 8:311624, 2018.
- [18] Aydın Tözeren. Internal forces and the human body: Complexity of the musculoskeletal system. *Human Body Dynamics: Classical Mechanics and Human Movement*, pages 150–193, 2000.
- [19] Georgina K Stebbings, Christopher I Morse, Alun G Williams, and Stephen H Day. Variability and distribution of muscle strength and its determinants in humans. *Muscle & nerve*, 49(6):879–886, 2014.
- [20] Tien Tuan Dao. From deep learning to transfer learning for the prediction of skeletal muscle forces. *Medical & biological engineering & computing*, 57:1049–1058, 2019.
- [21] Stuart A McErlain-Naylor, Mark A King, and Paul J Felton. A review of forward-dynamics simulation models for predicting optimal technique in maximal effort sporting movements. *Applied Sciences*, 11(4):1450, 2021.
- [22] Ahmet Erdemir, Scott McLean, Walter Herzog, and Antonie J van den Bogert. Model-based estimation of muscle forces exerted during movements. *Clinical biomechanics*, 22(2):131–154, 2007.
- [23] Didier Staudenmann, Karin Roeleveld, Dick F Stegeman, and Jaap H Van Dieën. Methodological aspects of semg recordings for force estimation—a tutorial and review. *Journal of electromyography and kinesiology*, 20(3):375–387, 2010.
- [24] Simon Haykin. *Neural networks and learning machines*, 3/E. Pearson Education India, 2009.
- [25] Jie Zhang, Yihui Zhao, Fergus Shone, Zhenhong Li, Alejandro F Frangi, Sheng Quan Xie, and Zhi-Qiang Zhang. Physics-informed deep learning for musculoskeletal modeling: Predicting muscle forces and joint kinematics from surface emg.

- IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:484–493, 2022.
- [26] Marie D Schmidt, Tobias Glasmachers, and Ioannis Iossifidis. The concepts of muscle activity generation driven by upper limb kinematics. *BioMedical Engineering OnLine*, 22(1):63, 2023.
- [27] Du Jiang, Gongfa Li, Ying Sun, Jianyi Kong, Bo Tao, and Disi Chen. Grip strength forecast and rehabilitative guidance based on adaptive neural fuzzy inference system using semg. *Personal and Ubiquitous Computing*, pages 1–10, 2022.
- [28] JA Goguen. La zadeh. fuzzy sets. information and control, vol. 8 (1965), pp. 338–353.-la zadeh. similarity relations and fuzzy orderings. information sciences, vol. 3 (1971), pp. 177–200. *The Journal of Symbolic Logic*, 38(4):656–657, 1973.
- [29] Yassine Bouteraa, Ismail Ben Abdallah, Atef Ibrahim, and Tariq Ahamed Ahanger. Development of an iot-based solution incorporating biofeedback and fuzzy logic control for elbow rehabilitation. *Applied Sciences*, 10(21):7793, 2020.
- [30] Massimo Sartori, Leonardo Gizzi, David G Lloyd, and Dario Farina. A musculoskeletal model of human locomotion driven by a low dimensional set of impulsive excitation primitives. *Frontiers in computational neuroscience*, 7:79, 2013.
- [31] SN Sivanandam, SN Deepa, SN Sivanandam, and SN Deepa. *Genetic algorithms*. Springer, 2008.
- [32] Chiako Mokri, Mahdi Bamdad, and Vahid Abolghasemi. Muscle force estimation from lower limb emg signals using novel optimised machine learning techniques. *Medical & biological engineering & computing*, 60(3):683–699, 2022.
- [33] Massimo Sartori, Dario Farina, and David G Lloyd. Hybrid neuromusculoskeletal modeling to best track joint moments using a balance between muscle excitations derived from electromyograms and optimization. *Journal of biomechanics*, 47(15):3613–3621, 2014.
- [34] Jing Luo, Chao Liu, and Chenguang Yang. Estimation of emg-based force using a neural-network-based approach. *IEEE Access*, 7:64856–64865, 2019.
- [35] Elsa Andrea Kirchner and Judith Bütetfür. Towards bidirectional and coadaptive robotic exoskeletons for neuromotor rehabilitation and assisted daily living: A review. *Current Robotics Reports*, 3(2):21–32, 2022.
- [36] Allison B Wang, Eric J Perreault, Thomas J Royston, and Sabrina SM Lee. Changes in shear wave propagation within skeletal muscle during active and passive force generation. *Journal of biomechanics*, 94:115–122, 2019.
- [37] Roopam Bamal and Singara Singh Kasana. Reversible medical image watermarking for tamper detection using ann and slt. *Multimedia Tools and Applications*, 83(8):21849–21882, 2024.

- [38] Chunsheng Song, Shengquan Xie, Zude Zhou, and Yefa Hu. Modeling of pneumatic artificial muscle using a hybrid artificial neural network approach. *Mechatronics*, 31:124–131, 2015.
- [39] Preeti Yadav and Subhash Chandra Sharma. A systematic review of localization in wsn: Machine learning and optimization-based approaches. *International journal of communication systems*, 36(4):e5397, 2023.
- [40] Mehmet Fatih Akay, Fatih Abut, Ebru Cetin, Imdat Yarim, and Boubacar Sow. Support vector machines for predicting the hamstring and quadriceps muscle strength of college-aged athletes. *Turkish Journal of Electrical Engineering and Computer Sciences*, 25(4):2567–2582, 2017.
- [41] Giulio Cecchini, Gabriele Maria Lozito, Maurizio Schmid, Silvia Conforto, Francesco Riganti Fulginei, and Daniele Bibbo. Neural networks for muscle forces prediction in cycling. *Algorithms*, 7(4):621–634, 2014.
- [42] Takashi Yanagawa, Cheryl J Goodwin, Kevin B Shelburne, J Erik Giphart, Michael R Torry, and Marcus G Pandy. Contributions of the individual muscles of the shoulder to glenohumeral joint stability during abduction. 2008.
- [43] Ziyun Ding, Daniel Nolte, Chui Kit Tsang, Daniel J Cleather, Angela E Kedgley, and Anthony MJ Bull. In vivo knee contact force prediction using patient-specific musculoskeletal geometry in a segment-based computational model. *Journal of Biomechanical Engineering*, 138(2):021018, 2016.
- [44] Ilan Eskinazi and Benjamin J Fregly. Surrogate modeling of deformable joint contact using artificial neural networks. *Medical engineering & physics*, 37(9):885–891, 2015.
- [45] Wonkeun Youn and Jung Kim. Feasibility of using an artificial neural network model to estimate the elbow flexion force from mechanomyography. *Journal of neuroscience methods*, 194(2):386–393, 2011.
- [46] Fabian Horst, Sebastian Lapuschkin, Wojciech Samek, Klaus-Robert Müller, and Wolfgang I Schöllhorn. Explaining the unique nature of individual gait patterns with deep learning. *Scientific reports*, 9(1):2391, 2019.
- [47] Kenton Kaufman, Emily Miller, Trevor Kingsbury, Elizabeth Russell Esposito, Erik Wolf, Jason Wilken, and Marilyn Wyatt. Reliability of 3d gait data across multiple laboratories. *Gait & posture*, 49:375–381, 2016.
- [48] Aurelio Cappozzo, Fabio Catani, Ugo Della Croce, and Alberto Leardini. Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical biomechanics*, 10(4):171–178, 1995.
- [49] JR Potvin, RW Norman, and SM McGill. Mechanically corrected emg for the continuous estimation of erector spinae muscle loading during repetitive lifting.

- European journal of applied physiology and occupational physiology*, 74:119–132, 1996.
- [50] Kenton R Kaufman, LS Miller, and DH Sutherland. Gait asymmetry in patients with limb-length inequality. *Journal of Pediatric Orthopaedics*, 16(2):144–150, 1996.
- [51] Lorenzo Chiari, Andrea Bertani, and Angelo Cappello. Classification of visual strategies in human postural control by stochastic parameters. *Human Movement Science*, 19(6):817–842, 2000.
- [52] Marianne JR Gittoes, Mark A Brewin, and David G Kerwin. Soft tissue contributions to impact forces simulated using a four-segment wobbling mass model of forefoot–heel landings. *Human movement science*, 25(6):775–787, 2006.
- [53] Felix E Zajac, Richard R Neptune, and Steven A Kautz. Biomechanics and muscle coordination of human walking: part ii: lessons from dynamical simulations and clinical implications. *Gait & posture*, 17(1):1–17, 2003.
- [54] Richard R Neptune, Steven A Kautz, and Felix E Zajac. Contributions of the individual ankle plantar flexors to support, forward progression and swing initiation during walking. *Journal of biomechanics*, 34(11):1387–1398, 2001.
- [55] Joseph Hamill and Allison H Gruber. Is changing footstrike pattern beneficial to runners? *Journal of sport and health science*, 6(2):146–153, 2017.
- [56] A Krpič, I Cikažlo, A Savanović, and Z Matjačić. A haptic floor for interaction and diagnostics with goal based tasks during virtual reality supported balance training. 2014.
- [57] Craig P McGowan, Alena M Grabowski, William J McDermott, Hugh M Herr, and Rodger Kram. Leg stiffness of sprinters using running-specific prostheses. *Journal of the Royal Society Interface*, 9(73):1975–1982, 2012.
- [58] Ruyi Ma, Leilei Zhang, Gongfa Li, Du Jiang, Shuang Xu, and Disi Chen. Grasping force prediction based on sEMG signals. *Alexandria Engineering Journal*, 59(3):1135–1147, 2020.
- [59] Yean Zhu, XU Weiyi, Guoliang Luo, Haolun Wang, Jingjing Yang, and Wei Lu. Random forest enhancement using improved artificial fish swarm for the medial knee contact force prediction. *Artificial Intelligence in Medicine*, 103:101811, 2020.
- [60] Hang Su, Wen Qi, Zhijun Li, Ziyang Chen, Giancarlo Ferrigno, and Elena De Momi. Deep neural network approach in emg-based force estimation for human–robot interaction. *IEEE Transactions on Artificial Intelligence*, 2(5):404–412, 2021.
- [61] Nirav Maniar, Anthony G Schache, Claudio Pizzolato, and David A Opar. Muscle contributions to tibiofemoral shear forces and valgus and rotational joint moments during single leg drop landing. *Scandinavian journal of medicine & science in sports*, 30(9):1664–1674, 2020.

- [62] Longbin Zhang, Zhijun Li, Yingbai Hu, Christian Smith, Elena M Gutierrez Farewik, and Ruoli Wang. Ankle joint torque estimation using an emg-driven neuromusculoskeletal model and an artificial neural network model. *IEEE Transactions on Automation Science and Engineering*, 18(2):564–573, 2020.
- [63] Meng Zhu, Xiaorong Guan, Zhong Li, Long He, Zheng Wang, and Keshu Cai. semg-based lower limb motion prediction using cnn-lstm with improved pca optimization algorithm. *Journal of Bionic Engineering*, 20(2):612–627, 2023.
- [64] Yihui Zhao, Zhiqiang Zhang, Zhenhong Li, Zhixin Yang, Abbas A Dehghani-Sani, and Shengquan Xie. An emg-driven musculoskeletal model for estimating continuous wrist motion. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(12):3113–3120, 2020.
- [65] Itzel Jared Rodriguez Martinez, Andrea Mannini, Francesco Clemente, Angelo Maria Sabatini, and Christian Cipriani. Grasp force estimation from the transient emg using high-density surface recordings. *Journal of Neural Engineering*, 17(1):016052, 2020.
- [66] Ahmad Moniri, Dan Terracina, Jesus Rodriguez-Manzano, Paul H Strutton, and Pantelis Georgiou. Real-time forecasting of semg features for trunk muscle fatigue using machine learning. *IEEE Transactions on Biomedical Engineering*, 68(2):718–727, 2020.
- [67] Gelareh Hajian, Ali Etemad, and Evelyn Morin. Generalized emg-based isometric contact force estimation using a deep learning approach. *Biomedical Signal Processing and Control*, 70:103012, 2021.
- [68] Salman Mohd Khan, Abid Ali Khan, and Omar Farooq. Pattern recognition of emg signals for low level grip force classification. *Biomedical Physics & Engineering Express*, 7(6):065012, 2021.
- [69] Anne Tryphosa Kamatham, Meena Alzamani, Allison Dockum, Siddhartha Sikdar, and Biswarup Mukherjee. Sparse sonomyography-based estimation of isometric force: A comparison of methods and features. *IEEE Transactions on Medical Robotics and Bionics*, 4(3):821–829, 2022.
- [70] Ruochen Hu, Xiang Chen, Haotian Zhang, Xu Zhang, and Xun Chen. A novel myoelectric control scheme supporting synchronous gesture recognition and muscle force estimation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:1127–1137, 2022.
- [71] Wenqi Liang, Fanjie Wang, Ao Fan, Wenrui Zhao, Wei Yao, and Pengfei Yang. Extended application of inertial measurement units in biomechanics: From activity recognition to force estimation. *Sensors*, 23(9), 2023.
- [72] Marcus G Pandy. Computer modeling and simulation of human movement. *Annual*

- review of biomedical engineering*, 3(1):245–273, 2001.
- [73] Sidharth Pancholi and Amit M Joshi. Advanced energy kernel-based feature extraction scheme for improved emg-pr-based prosthesis control against force variation. *IEEE Transactions on Cybernetics*, 52(5):3819–3828, 2020.
- [74] Chenfei Ma, Weiyu Guo, Hang Zhang, Oluwarotimi Williams Samuel, Xiaopeng Ji, Lisheng Xu, and Guanglin Li. A novel and efficient feature extraction method for deep learning based continuous estimation. *IEEE Robotics and Automation Letters*, 6(4):7341–7348, 2021.
- [75] Turker Tuncer, Sengul Dogan, and Abdulhamit Subasi. Surface emg signal classification using ternary pattern and discrete wavelet transform based feature extraction for hand movement recognition. *Biomedical signal processing and control*, 58:101872, 2020.
- [76] Vimal Shanmuganathan, Harold Robinson Yesudhas, Mohammad S Khan, Manju Khari, and Amir H Gandomi. R-cnn and wavelet feature extraction for hand gesture recognition with emg signals. *Neural Computing and Applications*, 32:16723–16736, 2020.
- [77] Sidharth Pancholi and Amit M Joshi. Intelligent upper-limb prosthetic control (iulp) with novel feature extraction method for pattern recognition using emg. *Journal of Mechanics in Medicine and Biology*, 21(06):2150043, 2021.
- [78] Mehmet Baygin, Prabal Datta Barua, Sengul Dogan, Turker Tuncer, Sefa Key, U Rajendra Acharya, and Kang Hao Cheong. A hand-modeled feature extraction-based learning network to detect grasps using semg signal. *Sensors*, 22(5):2007, 2022.
- [79] Cheng Shen, Zhongcai Pei, Weihai Chen, Jianhua Wang, Jianbin Zhang, and Zuobing Chen. Toward generalization of semg-based pattern recognition: A novel feature extraction for gesture recognition. *IEEE Transactions on Instrumentation and Measurement*, 71:1–12, 2022.
- [80] Gongfa Li, Jiahan Li, Zhaojie Ju, Ying Sun, and Jianyi Kong. A novel feature extraction method for machine learning based on surface electromyography from healthy brain. *Neural Computing and Applications*, 31:9013–9022, 2019.
- [81] Sumair Aziz, Muhammad Umar Khan, Foha Aamir, and Muhammad Arshad Javid. Electromyography (emg) data-driven load classification using empirical mode decomposition and feature analysis. In *2019 International Conference on Frontiers of Information Technology (FIT)*, pages 272–2725. IEEE, 2019.
- [82] David Carneros-Prado, Cosmin C Dobrescu, Luis Cabañero, Yulith V Altamirano-Flores, Irvin Hussein Lopez-Nava, Iván González, Jesús Fontecha, and Ramon Herivas. Approach to a lower body gait generation model using a deep convolutional

- generative adversarial network. In *International Conference on Ubiquitous Computing and Ambient Intelligence*, pages 419–430. Springer, 2022.
- [83] Rachel Yun, May Salama, and Lamiaa Elrefaei. An exploratory study on the effect of applying various artificial neural networks to the classification of lower limb injury. *Turkish Journal of Electrical Engineering and Computer Sciences*, 31(2):448–461, 2023.
- [84] Chenchu Xu, Lei Xu, Pavlo Ohorodnyk, Mike Roth, Bo Chen, and Shuo Li. Contrast agent-free synthesis and segmentation of ischemic heart disease images using progressive sequential causal gans. *Medical Image Analysis*, 62:101668, 2020.
- [85] Metin Bicer, Andrew TM Phillips, Alessandro Melis, Alison H McGregor, and Luca Modenese. Generative deep learning applied to biomechanics: A new augmentation technique for motion capture datasets. *Journal of biomechanics*, 144:111301, 2022.
- [86] Minglang Yin, Zongren Zou, Enrui Zhang, Cristina Cavinato, Jay D Humphrey, and George Em Karniadakis. A generative modeling framework for inferring families of biomechanical constitutive laws in data-sparse regimes. *Journal of the Mechanics and Physics of Solids*, 181:105424, 2023.
- [87] David R Mullineaux and Gareth Irwin. Error and anomaly detection for intra-participant time-series data. *International Biomechanics*, 4(1):28–35, 2017.
- [88] Ankit Vijayvargiya, Chandra Prakash, Rajesh Kumar, Sanjeev Bansal, and Joao Manuel RS Tavares. Human knee abnormality detection from imbalanced semg data. *Biomedical Signal Processing and Control*, 66:102406, 2021.
- [89] Usman Rashid, Nitika Kumari, Denise Taylor, Tim David, and Nada Signal. Gait event anomaly detection and correction during a split-belt treadmill task. *IEEE Access*, 7:68469–68478, 2019.
- [90] Zeqing Jin, Zhizhou Zhang, Xianlin Shao, and Grace X Gu. Monitoring anomalies in 3d bioprinting with deep neural networks. *ACS Biomaterials Science & Engineering*, 9(7):3945–3952, 2021.
- [91] Haemwaan Sivaraks and Chotirat Ann Ratanamahatana. Robust and accurate anomaly detection in ecg artifacts using time series motif discovery. *Computational and mathematical methods in medicine*, 2015, 2015.
- [92] Hasan Alkahtani, Theyazn HH Aldhyani, and Mohammed Al-Yaari. Adaptive anomaly detection framework model objects in cyberspace. *Applied Bionics and Biomechanics*, 2020:1–14, 2020.
- [93] Stewart G Trost, Maria Fragala-Pinkham, Nancy Lennon, and Margaret E O’Neil. Decision trees for detection of activity intensity in youth with cerebral palsy. *Medicine and science in sports and exercise*, 48(5):958, 2016.
- [94] Himanshu Jindal, Monika Bharti, Singara Singh Kasana, and Sharad Saxena. An

- ensemble mosaicing and ridgelet based fusion technique for underwater panoramic image reconstruction and its refinement. *Multimedia Tools and Applications*, 82(22):33719–33771, 2023.
- [95] Shane Gore. *The identification of foot-strike patterns and prediction of running related injuries*. PhD thesis, Dublin, National College of Ireland, 2020.
- [96] Victor E Staartjes, Marlies P de Wispelaere, William Peter Vandertop, and Marc L Schröder. Deep learning-based preoperative predictive analytics for patient-reported outcomes following lumbar discectomy: feasibility of center-specific modeling. *The Spine Journal*, 19(5):853–861, 2019.
- [97] Joarder Kamruzzaman and Rezaul K Begg. Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait. *IEEE Transactions on Biomedical Engineering*, 53(12):2479–2490, 2006.
- [98] Seda Arslan Tuncer and Ahmet Alkan. Classification of emg signals taken from arm with hybrid cnn-svm architecture. *Concurrency and Computation: Practice and Experience*, 34(5):e6746, 2022.
- [99] Daniel TH Lai, Ahsan Khandoker, Rezaul K Begg, and Marimuthu Palaniswami. A hybrid support vector machine and autoregressive model for detecting gait disorders in the elderly. In *2007 International Joint Conference on Neural Networks*, pages 321–326. IEEE, 2007.
- [100] Jyotindra Narayan and Santosha K Dwivedy. Biomechanical study and prediction of lower extremity joint movements using bayesian regularization-based backpropagation neural network. *Journal of Computing and Information Science in Engineering*, 22(1):014503, 2022.
- [101] G Sucharitha, Nitin Arora, and Subhash C Sharma. Medical image retrieval using a novel local relative directional edge pattern and zernike moments. *Multimedia Tools and Applications*, 82(20):31737–31757, 2023.
- [102] Russell T Johnson, Daniel Lakeland, and James M Finley. Using bayesian inference to estimate plausible muscle forces in musculoskeletal models. *Journal of neuro-engineering and rehabilitation*, 19(1):34, 2022.
- [103] GV Prateek, Isaac Skog, Marie E McNeely, Ryan P Duncan, Gammon M Earhart, and Arye Nehorai. Modeling, detecting, and tracking freezing of gait in parkinson disease using inertial sensors. *IEEE Transactions on Biomedical Engineering*, 65(10):2152–2161, 2017.
- [104] Marion Mundt, Arnd Koeppe, Sina David, Tom Witter, Franz Bamer, Wolfgang Potthast, and Bernd Markert. Estimation of gait mechanics based on simulated and measured imu data using an artificial neural network. *Frontiers in bioengineering and biotechnology*, 8:41, 2020.

- [105] Abdulhamit Subasi, Mustafa Yilmaz, and Hasan Riza Ozcalik. Classification of emg signals using wavelet neural network. *Journal of neuroscience methods*, 156(1-2):360–367, 2006.
- [106] Zhang Lei. An upper limb movement estimation from electromyography by using bp neural network. *Biomedical Signal Processing and Control*, 49:434–439, 2019.
- [107] Christophorus Beneditto Aditya Satrio, William Darmawan, Bellatasya Unrica Nadia, and Novita Hanafiah. Time series analysis and forecasting of coronavirus disease in indonesia using arima model and prophet. *Procedia Computer Science*, 179:524–532, 2021.
- [108] Diego Duarte and Julio Faerman. Comparison of time series prediction of healthcare emergency department indicators with arima and prophet. In *Computer Science & Information Technology (CS & IT) Computer Science Conference*, volume 123, page 33, 2019.
- [109] Slavica Mačuzić Saveljić. Seat-to-head transfer functions prediction using artificial neural networks. In *Serbian International Conference on Applied Artificial Intelligence*, pages 249–270. Springer, 2022.
- [110] Jayun Hyun, Seo Hu Lee, Ha Min Son, Ji-Ung Park, and Tai-Myoung Chung. A synthetic data generation model for diabetic foot treatment. In *International Conference on Future Data and Security Engineering*, pages 249–264. Springer, 2020.
- [111] Wei Lu, Lifu Gao, Zebin Li, Daqing Wang, and Huibin Cao. Prediction of long-term elbow flexion force intervals based on the informer model and electromyography. *Electronics*, 10(16):1946, 2021.
- [112] Muhammad Zeeshan Arshad, Ankhzaya Jamsrandorj, Jinwook Kim, and Kyung-Ryoul Mun. Gait events prediction using hybrid cnn-rnn-based deep learning models through a single waist-worn wearable sensor. *Sensors*, 22(21):8226, 2022.
- [113] Shitong Mao and Ervin Sejdić. A review of recurrent neural network-based methods in computational physiology. *IEEE transactions on neural networks and learning systems*, 34(10):6983–7003, 2022.
- [114] Ruben Fuentes-Alvarez, Joel Hernandez Hernandez, Ivan Matehuala-Moran, Mariel Alfaro-Ponce, Ricardo Lopez-Gutierrez, Sergio Salazar, and Rogelio Lozano. Assistive robotic exoskeleton using recurrent neural networks for decision taking for the robust trajectory tracking. *Expert Systems with Applications*, 193:116482, 2022.
- [115] Abdelrahman Zaroug, Alessandro Garofolini, Daniel TH Lai, Kurt Mudie, and Rezaul Begg. Prediction of gait trajectories based on the long short term memory neural networks. *Plos one*, 16(8):e0255597, 2021.
- [116] Miguel Simão, Pedro Neto, and Olivier Gibaru. Emg-based online classification of

- gestures with recurrent neural networks. *Pattern Recognition Letters*, 128:45–51, 2019.
- [117] Ashik Mostafa Alvi, Siuly Siuly, and Hua Wang. A long short-term memory based framework for early detection of mild cognitive impairment from eeg signals. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(2):375–388, 2022.
- [118] Tito Bassani and Fabio Galbusera. Musculoskeletal modeling. In *Biomechanics of the Spine*, pages 257–277. Elsevier, 2018.
- [119] Jacob S Arlotti, William O Carroll, Youness Affi, Purva Talegaonkar, Luciano Albuquerque, John E Ball, Harish Chander, Adam Petway, et al. Benefits of imu-based wearables in sports medicine: Narrative review. *International Journal of Kinesiology and Sports Science*, 10(1):36–43, 2022.
- [120] Amir Hossein Eskandari, Farshid Ghezelbash, Aboufazel Shirazi-Adl, Denis Gagnon, Hakim Mecheri, and Christian Larivière. Validation of an emg submaximal method to calibrate a novel dynamic emg-driven musculoskeletal model of the trunk: Effects on model estimates. *Journal of Electromyography and Kinesiology*, 68:102728, 2023.
- [121] Yan Wang, Qitao Tan, Fang Pu, David Boone, and Ming Zhang. A review of the application of additive manufacturing in prosthetic and orthotic clinics from a biomechanical perspective. *Engineering*, 6(11):1258–1266, 2020.
- [122] Scott L Delp, Frank C Anderson, Allison S Arnold, Peter Loan, Ayman Habib, Chand T John, Eran Guendelman, and Darryl G Thelen. Opensim: open-source software to create and analyze dynamic simulations of movement. *IEEE transactions on biomedical engineering*, 54(11):1940–1950, 2007.
- [123] Jeffrey A Reinbolt, Ajay Seth, and Scott L Delp. Simulation of human movement: applications using opensim. *Procedia Iutam*, 2:186–198, 2011.
- [124] Lucas Engelhardt, Maximilian Melzner, Linda Havelkova, Pavel Fiala, Patrik Christen, Sebastian Dendorfer, and Ulrich Simon. A new musculoskeletal anybody™ detailed hand model. *Computer methods in biomechanics and biomedical engineering*, 24(7):777–787, 2020.
- [125] Michael Damsgaard, John Rasmussen, Søren Tørholm Christensen, Egidijus Surma, and Mark De Zee. Analysis of musculoskeletal systems in the anybody modeling system. *Simulation Modelling Practice and Theory*, 14(8):1100–1111, 2006.
- [126] Gerald E Loeb and Rahman Davoodi. Musculoskeletal mechanics and modeling. *Scholarpedia*, 11(11):12389, 2016.
- [127] Rahman Davoodi and Gerald E Loeb. A software tool for faster development of complex models of musculoskeletal systems and sensorimotor controllers in simulinktm. *Journal of Applied Biomechanics*, 18(4):357–365, 2002.

- [128] Manuel Cardona and Cecilia García Cena. Musculoskeletal modeling as a tool for biomechanical analysis of normal and pathological gait. In *Latin American Conference on Biomedical Engineering*, pages 955–963. Springer, 2019.
- [129] Giuliano Lamberto, Saulo Martelli, Aurelio Cappozzo, and Claudia Mazzà. To what extent is joint and muscle mechanics predicted by musculoskeletal models sensitive to soft tissue artefacts? *Journal of biomechanics*, 62:68–76, 2017.
- [130] Kim Nolte, Pieter E Krüger, and P Schalk Els. Three dimensional musculoskeletal modelling of the seated biceps curl resistance training exercise. *Sports biomechanics*, 10(02):146–160, 2011.
- [131] Kim Nolte, Pieter E Krüger, P Schalk Els, and Heinrich Nolte. Three dimensional musculoskeletal modelling of the abdominal crunch resistance training exercise. *Journal of sports sciences*, 31(3):264–275, 2013.
- [132] Ashley Elizabeth Clark. Biarticular muscles influence postural responses: implications for treatment of stiff-knee gait. 2012.
- [133] M Reilly and K Kontson. 1university of maryland, department of biomedical engineering, college park, md 2 us fda, center for devices and radiological health, office of science and engineering laboratories, division of biomedical physics, silver spring, md.
- [134] Colleen Goh, Mary L Blanchard, Robin H Crompton, Michael M Gunther, Sophie Macaulay, and Karl T Bates. A 3d musculoskeletal model of the western lowland gorilla hind limb: moment arms and torque of the hip, knee and ankle. *Journal of Anatomy*, 231(4):568–584, 2017.
- [135] B Fernini, M Temmar, and MM Noor. Toward a dynamic analysis of bipedal robots inspired by human leg muscles. *Journal of Mechanical Engineering and Sciences*, 12(2):3593–3604, 2018.
- [136] Jinlin Jiang, Yu Wang, Heng Cao, Jun Zhu, and Xinbin Zhang. A novel pump-valve coordinated controlled hydraulic system for the lower extremity exoskeleton. *Transactions of the Institute of Measurement and Control*, 42(15):2872–2884, 2020.
- [137] Junyan Li, Yongtao Lu, Stuart C Miller, Zhongmin Jin, and Xijin Hua. Development of a finite element musculoskeletal model with the ability to predict contractions of three-dimensional muscles. *Journal of biomechanics*, 94:230–234, 2019.
- [138] Kurt Manal and Thomas S Buchanan. An electromyogram-driven musculoskeletal model of the knee to predict in vivo joint contact forces during normal and novel gait patterns. *Journal of biomechanical engineering*, 135(2):021014, 2013.
- [139] JAC Martins, MPM Pato, and EB Pires. A finite element model of skeletal muscles. *Virtual and Physical Prototyping*, 1(3):159–170, 2006.
- [140] Sarah A Roelker, Elena J Caruthers, Rachel K Baker, Nicholas C Pelz, Ajit MW

- Chaudhari, and Robert A Siston. Interpreting musculoskeletal models and dynamic simulations: causes and effects of differences between models. *Annals of biomedical engineering*, 45:2635–2647, 2017.
- [141] Basil Mathai and Sanjay Gupta. Numerical predictions of hip joint and muscle forces during daily activities: a comparison of musculoskeletal models. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 233(6):636–647, 2019.
- [142] Munsur Rahman, Mohsen Sharifi Renani, Akin Cil, and Antonis P Stylianou. Musculoskeletal model development of the elbow joint with an experimental evaluation. *Bioengineering*, 5(2):31, 2018.
- [143] Friedl De Groote and Antoine Falisse. Perspective on musculoskeletal modelling and predictive simulations of human movement to assess the neuromechanics of gait. *Proceedings of the Royal Society B*, 288(1946):20202432, 2021.
- [144] Daniel C McFarland, Benjamin I Binder-Markey, Jennifer A Nichols, Sarah J Wohlman, Marije De Bruin, and Wendy M Murray. A musculoskeletal model of the hand and wrist capable of simulating functional tasks. *IEEE Transactions on Biomedical Engineering*, 70(5):1424–1435, 2022.
- [145] Huawei Wang, Vittorio Caggiano, Guillaume Durandau, Massimo Sartori, and Vikash Kumar. Myosim: Fast and physiologically realistic mujoco models for musculoskeletal and exoskeletal studies. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 8104–8111. IEEE, 2022.
- [146] Manuel Cardona and Cecilia E. García Cena. Biomechanical analysis of the lower limb: A full-body musculoskeletal model for muscle-driven simulation. *IEEE Access*, 7:92709–92723, 2019.
- [147] Katelyn Burkhart, Daniel Grindle, Mary L Bouxsein, and Dennis E Anderson. Between-session reliability of subject-specific musculoskeletal models of the spine derived from optoelectronic motion capture data. *Journal of biomechanics*, 112:110044, 2020.
- [148] Darryl G Thelen and Frank C Anderson. Using computed muscle control to generate forward dynamic simulations of human walking from experimental data. *Journal of biomechanics*, 39(6):1107–1115, 2006.
- [149] Thomas P Andriacchi, Eugene J Alexander, MK Toney, C Dyrby, and Jeffrey a Sum. A point cluster method for in vivo motion analysis: applied to a study of knee kinematics. 1998.
- [150] Silvia S Blemker, Deanna S Asakawa, Garry E Gold, and Scott L Delp. Image-based musculoskeletal modeling: applications, advances, and future opportunities. *Journal of Magnetic Resonance Imaging: An Official Journal of the International*

- Society for Magnetic Resonance in Medicine*, 25(2):441–451, 2007.
- [151] Navid Arjmand and Aboufazel Shirazi-Adl. Biomechanics of changes in lumbar posture in static lifting. *Spine*, 30(23):2637–2648, 2005.
- [152] Zhibo Jing, Jianda Han, and Juanjuan Zhang. Comparison of biomechanical analysis results using different musculoskeletal models for children with cerebral palsy. *Frontiers in Bioengineering and Biotechnology*, 11, 2023.
- [153] George Beckham, Tim Suchomel, and Satoshi Mizuguchi. Force plate use in performance monitoring and sport science testing. *New Studies in Athletics*, 29(3):25–37, 2014.
- [154] Mark S Walsh, Kevin R Ford, Kyle J Bangen, Gregory D Myer, and Timothy E Hewett. The validation of a portable force plate for measuring force-time data during jumping and landing tasks. *The Journal of Strength & Conditioning Research*, 20(4):730–734, 2006.
- [155] Eline Van der Kruk and Marco M Reijne. Accuracy of human motion capture systems for sport applications; state-of-the-art review. *European journal of sport science*, 18(6):806–819, 2018.
- [156] Matthew Field, Zengxi Pan, David Stirling, and Fazel Naghdy. Human motion capture sensors and analysis in robotics. *Industrial Robot: An International Journal*, 38(2):163–171, 2011.
- [157] AFT Ibrahim, VR Gannapathy, LW Chong, and ISM Isa. Analysis of electromyography (emg) signal for human arm muscle: A review. *Advanced Computer and Communication Engineering Technology: Proceedings of ICOCOE 2015*, pages 567–575, 2016.
- [158] Muhammad Al-Ayyad, Hamza Abu Owida, Roberto De Fazio, Bassam Al-Naami, and Paolo Visconti. Electromyography monitoring systems in rehabilitation: a review of clinical applications, wearable devices and signal acquisition methodologies. *Electronics*, 12(7):1520, 2023.
- [159] Zita Ranajoy Dey, Nadia R Nair, and Nigel Shapcott. Evaluation of the force sensing application pressure mapping system. *Journal of Medical Engineering & Technology*, 37(3):213–219, 2013.
- [160] Ilaria Baldoli, Tommaso Mazzocchi, Clara Paoletti, Leonardo Ricotti, Pietro Salvo, Valentina Dini, Cecilia Laschi, Fabio Di Francesco, and Arianna Menciassi. Pressure mapping with textile sensors for compression therapy monitoring. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 230(8):795–808, 2016.
- [161] Gerasimos G Samatas and Theodore P Pachidis. Inertial measurement units (imus) in mobile robots over the last five years: A review. *Designs*, 6(1):17, 2022.

- [162] Georgios Maniatis. On the use of imu (inertial measurement unit) sensors in geomorphology. *Earth Surface Processes and Landforms*, 46(11):2136–2140, 2021.
- [163] Giovanni Saggio. Mechanical model of flex sensors used to sense finger movements. *Sensors and Actuators A: Physical*, 185:53–58, 2012.
- [164] Wei-Chieh Chuang, Wen-Jyi Hwang, Tsung-Ming Tai, De-Rong Huang, and Yun-Jie Jhang. Continuous finger gesture recognition based on flex sensors. *Sensors*, 19(18):3986, 2019.
- [165] Zachary T Beattie, Chad C Hagen, Misha Pavel, and Tamara L Hayes. Classification of breathing events using load cells under the bed. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3921–3924. IEEE, 2009.
- [166] J Goodwin, C Clark, J Deakes, D Burdon, and C Lawrence. Clinical methods of goniometry: a comparative study. *Disability and rehabilitation*, 14(1):10–15, 1992.
- [167] Giorgio Ferriero, Stefano Vercelli, Francesco Sartorio, Susana Munoz Lasa, Elena Ilieva, Elisa Brigatti, Carolina Ruella, and Calogero Foti. Reliability of a smartphone-based goniometer for knee joint goniometry. *International journal of rehabilitation research*, 36(2):146–151, 2013.
- [168] Josária F Amaral, Marcelly Mancini, and José M Novo Júnior. Comparison of three hand dynamometers in relation to the accuracy and precision of the measurements. *Brazilian Journal of Physical Therapy*, 16:216–224, 2012.
- [169] João Pinto-Ramos, Tiago Moreira, Frederico Costa, Helena Tavares, João Cabral, Cristina Costa-Santos, Joana Barroso, and Bernardo Sousa-Pinto. Handheld dynamometer reliability to measure knee extension strength in rehabilitation patients—a cross-sectional study. *PLoS one*, 17(5):e0268254, 2022.
- [170] Catrine Tudor-Locke. Taking steps toward increased physical activity: Using pedometers to measure and motivate. *President’s Council on Physical Fitness and Sports Research Digest*, 2002.
- [171] Asmita Patel, Gregory Kolt, Grant Schofield, and Justin Keogh. General practitioners’ views on the role of pedometers in health promotion. *Journal of primary health care*, 6(2):152–156, 2014.
- [172] Klaas R Westerterp. Physical activity assessment with accelerometers. *International Journal of Obesity*, 23(3):S45–S49, 1999.
- [173] Erik Vanegas, Raul Igual, and Inmaculada Plaza. Sensing systems for respiration monitoring: A technical systematic review. *Sensors*, 20(18):5446, 2020.
- [174] Toan Dinh, Thanh Nguyen, Hoang-Phuong Phan, Nam-Trung Nguyen, Dzung Viet Dao, and John Bell. Stretchable respiration sensors: Advanced designs and multifunctional platforms for wearable physiological monitoring. *Biosensors and Bio-*

- electronics*, 166:112460, 2020.
- [175] Frank G Yanowitz. Introduction to ecg interpretation. *LDS Hospital and Intermountain Medical Center*, 2012.
- [176] Mary Boudreau Conover. *Understanding electrocardiography*. Elsevier Health Sciences, 2002.
- [177] Zhen Gang Xiao and Carlo Menon. A review of force myography research and development. *Sensors*, 19(20):4557, 2019.
- [178] MN Victorino, Xianta Jiang, and Carlo Menon. Wearable technologies and force myography for healthcare. *Wearable Technology in Medicine and Health Care*, pages 135–152, 2018.
- [179] Dawa Chyophel Lepcha, Bhawna Goyal, Ayush Dogra, and Vishal Goyal. Image super-resolution: A comprehensive review, recent trends, challenges and applications. *Information Fusion*, 91:230–260, 2023.
- [180] Claudio Orizio and Massimiliano Gobbo. Mechanomyography. *Wiley Encyclopedia of Biomedical Engineering*, 2006.
- [181] Anamul Islam, Kenneth Sundaraj, Badlishah Ahmad, Nizam Uddin Ahamed, and Asraf Ali. Mechanomyography sensors for muscle assessment: a brief review. *Journal of Physical Therapy Science*, 24(12):1359–1365, 2012.
- [182] Philippe Pourcelot, Marielle Defontaine, Bérangère Ravary, Mickaël Lemâtre, and Nathalie Crevier-Denoix. A non-invasive method of tendon force measurement. *Journal of Biomechanics*, 38(10):2124–2129, 2005.
- [183] M Ashina, L Bendtsen, R Jensen, F Sakai, and J Olesen. Measurement of muscle hardness: a methodological study. *Cephalalgia*, 18(2):106–111, 1998.
- [184] Shunji Moromugi, Shinichi Kumano, Mitsuaki Ueda, Takakazu Ishimatsu, Maria Q Feng, and Takayuki Tanaka. A sensor to measure hardness of human tissue. In *SENSORS, 2006 IEEE*, pages 388–391. IEEE, 2006.
- [185] Matthew W Gilbertson and Brian W Anthony. Force and position control system for freehand ultrasound. *IEEE Transactions on Robotics*, 31(4):835–849, 2015.
- [186] Hao Su, Iulian I Iordachita, Junichi Tokuda, Nobuhiko Hata, Xuan Liu, Reza Seifabadi, Sheng Xu, Bradford Wood, and Gregory S Fischer. Fiber-optic force sensors for mri-guided interventions and rehabilitation: A review. *IEEE sensors journal*, 17(7):1952–1963, 2017.
- [187] Iulian Iordachita, Zhenglong Sun, Marcin Balicki, Jin U Kang, Soo Jay Phee, James Handa, Peter Gehlbach, and Russell Taylor. A sub-millimetric, 0.25 mn resolution fully integrated fiber-optic force-sensing tool for retinal microsurgery. *International journal of computer assisted radiology and surgery*, 4:383–390, 2009.
- [188] Braden C Fleming, Glenn D Peura, and Bruce D Beynnon. Factors influencing the

- output of an implantable force transducer. *Journal of Biomechanics*, 33(7):889–893, 2000.
- [189] AS Sezen, Rajesh Rajamani, D Morrow, KR Kaufman, and BK Gilbert. An ultraminiature mems pressure sensor with high sensitivity for measurement of intramuscular pressure (imp) in patients with neuromuscular diseases. 2009.
- [190] Paulo Roriz, Orlando Frazão, António B Lobo-Ribeiro, José L Santos, and José A Simões. Review of fiber-optic pressure sensors for biomedical and biomechanical applications. *Journal of biomedical optics*, 18(5):050903–050903, 2013.
- [191] Christina R Hollis, Rachel M Koldenhoven, Jacob E Resch, and Jay Hertel. Running biomechanics as measured by wearable sensors: effects of speed and surface. *Sports biomechanics*, 2019.
- [192] Róisín M Howard, Richard Conway, and Andrew J Harrison. A survey of sensor devices: use in sports biomechanics. *Sports biomechanics*, 15(4):450–461, 2016.
- [193] Andrés Blanco Ortega, Jhonatan Isidro Godoy, Dariusz Slawomir Szwedowicz Wasik, Eladio Martínez Rayón, Claudia Cortés García, Héctor Ramón Azcaray Rivera, and Fabio Abel Gómez Becerra. Biomechanics of the upper limbs: A review in the sports combat ambit highlighting wearable sensors. *Sensors*, 22(13):4905, 2022.
- [194] François Hug, Kylie Tucker, Jean-Luc Gennisson, Mickaël Tanter, and Antoine Nordez. Elastography for muscle biomechanics: toward the estimation of individual muscle force. *Exercise and sport sciences reviews*, 43(3):125–133, 2015.
- [195] Bernd J Stetter, Steffen Ringhof, Frieder C Krafft, Stefan Sell, and Thorsten Stein. Estimation of knee joint forces in sport movements using wearable sensors and machine learning. *Sensors*, 19(17):3690, 2019.
- [196] Suin Kim, Kyongkwan Ro, Joonbum Bae, et al. Estimation of individual muscular forces of the lower limb during walking using a wearable sensor system. *Journal of Sensors*, 2017, 2017.
- [197] Yun Lu, Yudong Cao, Yi Chen, Hui Li, Weihua Li, Haiping Du, Shiwu Zhang, and Shuaishuai Sun. Investigation of a wearable piezoelectric-imu multi-modal sensing system for real-time muscle force estimation. *Smart Materials and Structures*, 32(6):065013, 2023.
- [198] Shima Mohammadi Moghadam, Ted Yeung, and Julie Choisne. A comparison of machine learning models’ accuracy in predicting lower-limb joints’ kinematics, kinetics, and muscle forces from wearable sensors. *Scientific reports*, 13(1):5046, 2023.
- [199] Arthritis Foundation. Arthritis by the numbers: Book of trusted facts and figures., 2017.

- [200] J Bedson, K Jordan, and P Croft. How do gps use x rays to manage chronic knee pain in the elderly? a case study. *Annals of the rheumatic diseases*, 62(5):450–454, 2003.
- [201] Sanjeevakumar Kubakaddi, KM Ravikumar, and Harini DG. Measurement of cartilage thickness for early detection of knee osteoarthritis (koa). In *2013 IEEE Point-of-Care Healthcare Technologies (PHT)*, pages 208–211, 2013.
- [202] Yoon Sang Kim, Jangwook Lee, Sooyong Lee, and Munsang Kim. A force reflected exoskeleton-type masterarm for human-robot interaction. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 35(2):198–212, 2005.
- [203] Md Rasedul Islam, Md Assad-Uz-Zaman, Brahim Brahmi, Yassine Bouteraa, Inga Wang, and Mohammad Habibur Rahman. Design and development of an upper limb rehabilitative robot with dual functionality. *Micromachines*, 12(8):870, 2021.
- [204] Li Zhang, Geng Liu, Bing Han, Zhe Wang, Han Li, and Yan Jiao. Assistive devices of human knee joint: A review. *Robotics and Autonomous Systems*, 125:103394, 2020.
- [205] Du-Xin Liu, Jing Xu, Chunjie Chen, Xingguo Long, Dacheng Tao, and Xinyu Wu. Vision-assisted autonomous lower-limb exoskeleton robot. *IEEE transactions on systems, man, and cybernetics: systems*, 51(6):3759–3770, 2019.
- [206] Yuanxi Sun, Yuntao Tang, Jia Zheng, Dianbiao Dong, Xiaohong Chen, and Long Bai. From sensing to control of lower limb exoskeleton: a systematic review. *Annual Reviews in Control*, 53(1):83–96, 2022.
- [207] John Calle-Siguencia, Mauro Callejas-Cuervo, and Sebastián García-Reino. Integration of inertial sensors in a lower limb robotic exoskeleton. *Sensors*, 22(12):4559, 2022.
- [208] Anuja Bhargava, Atul Bansal, and Vishal Goyal. Machine learning–based detection and sorting of multiple vegetables and fruits. *Food Analytical Methods*, 15(1):228–242, 2022.
- [209] Xinyao Tang, Xupeng Wang, Xiaomin Ji, Yawen Zhou, Jie Yang, Yuchen Wei, and Wenjie Zhang. A wearable lower limb exoskeleton: Reducing the energy cost of human movement. *Micromachines*, 13(6):900, 2022.
- [210] Cameron Roy Taylor, Shriya S Srinivasan, Seong Ho Yeon, MK O’Donnell, TJ Roberts, and Hugh M Herr. Magnetomicrometry. *Science Robotics*, 6(57):eabg0656, 2021.
- [211] Alberto Botter, Taian Vieira, Marco Carbonaro, Giacinto L Cerone, and Emma F Hodson-Tole. Electrodes’ configuration influences the agreement between surface emg and b-mode ultrasound detection of motor unit fasciculation. *IEEE Access*, 9:98110–98120, 2021.

- [212] Roberto Merletti and Carlo J De Luca. New techniques in surface electromyography. *Computer aided electromyography and expert systems*, 9(3):115–124, 1989.
- [213] Rahul Dev Singh, Ajay Mittal, and Rajesh K Bhatia. 3d convolutional neural network for object recognition: a review. *Multimedia Tools and Applications*, 78:15951–15995, 2019.
- [214] Kai Gui, Honghai Liu, and Dingguo Zhang. A practical and adaptive method to achieve emg-based torque estimation for a robotic exoskeleton. *IEEE/ASME Transactions on Mechatronics*, 24(2):483–494, 2019.
- [215] Inseung Kang, Pratik Kunapuli, Hsiang Hsu, and Aaron J Young. Electromyography (emg) signal contributions in speed and slope estimation using robotic exoskeletons. In *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, pages 548–553. IEEE, 2019.
- [216] Suguru Kanoga, Atsunori Kanemura, and Hideki Asoh. Are armband semg devices dense enough for long-term use?—sensor placement shifts cause significant reduction in recognition accuracy. *Biomedical Signal Processing and Control*, 60:101981, 2020.
- [217] A. Rainoldi, G. Melchiorri, and I. Caruso. A method for positioning electrodes during surface emg recordings in lower limb muscles. *Journal of Neuroscience Methods*, 134(1):37–43, 2004.
- [218] Reliability of lower limb electromyography during overground walking: A comparison of maximal- and sub-maximal normalisation techniques. *Journal of Biomechanics*, 43(4):749–756, 2010.
- [219] Use of cluster analysis for gait pattern classification of patients in the early and late recovery phases following stroke. *Gait & Posture*, 18(1):114–125, 2003.
- [220] He Huang, Fan Zhang, Levi J Hargrove, Zhi Dou, Daniel R Rogers, and Kevin B Englehart. Continuous locomotion-mode identification for prosthetic legs based on neuromuscular–mechanical fusion. *IEEE Transactions on Biomedical Engineering*, 58(10):2867–2875, 2011.
- [221] Surface emg based continuous estimation of human lower limb joint angles by using deep belief networks. *Biomedical Signal Processing and Control*, 40:335–342, 2018.
- [222] Christian Morbidoni, Alessandro Cucchiarelli, Sandro Fioretti, and Francesco Di Nardo. A deep learning approach to emg-based classification of gait phases during level ground walking. *Electronics*, 8(8), 2019.
- [223] Paolo Bonato, MSS Heng, Jose Gonzalez-Cueto, Alberto Leardini, John O’Connor, and Serge H Roy. Emg-based measures of fatigue during a repetitive squat exercise. *IEEE engineering in medicine and biology magazine*, 20(6):133–143, 2001.
- [224] Gourav Takhar, Chandra Prakash, Namita Mittal, and Rajesh Kumar. Vision-

- based gender recognition using hybrid background subtraction technique. In *International Conference on Next Generation Computing Technologies*, pages 651–662. Springer, 2017.
- [225] OFA Sanchez, JLR Sotelo, MH Gonzales, and GAM Hernandez. Emg dataset in lower limb data set. *UCI machine learning repository*, 2, 2014.
- [226] E Khaledian, S Pandey, P Kundu, and AK Srivastava. Real-time synchrophasor data anomaly detection and classification using isolation forest, kmeans and loop. *IEEE Transactions on Smart Grid*, 2020.
- [227] Syeda Aimal, Nadeem Javaid, Tahir Islam, Wazir Zada Khan, Mohammed Y Aal-saleem, and Hassan Sajjad. An efficient cnn and knn data analytics for electricity load forecasting in the smart grid. In *Workshops of the International Conference on Advanced Information Networking and Applications*, pages 592–603. Springer, 2019.
- [228] William Hurst, Casimiro A Curbelo Montanez, and Nathan Shone. Time-pattern profiling from smart meter data to detect outliers in energy consumption. *IoT*, 1(1):92–108, 2020.
- [229] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.
- [230] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4, 2015.
- [231] Gérard Biau and Erwan Scornet. A random forest guided tour. *Test*, 25(2):197–227, 2016.
- [232] Muhammad Waseem Ahmad, Jonathan Reynolds, and Yacine Rezgui. Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *Journal of cleaner production*, 203:810–821, 2018.
- [233] Anthony J Myles, Robert N Feudale, Yang Liu, Nathaniel A Woody, and Steven D Brown. An introduction to decision tree modeling. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 18(6):275–285, 2004.
- [234] Human knee abnormality detection from imbalanced semg data. *Biomedical Signal Processing and Control*, 66:102406, 2021.
- [235] Ankit Vijayvargiya, Rajesh Kumar, Nilanjan Dey, and João Manuel RS Tavares. Comparative analysis of machine learning techniques for the classification of knee abnormality. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, pages 1–6. IEEE, 2020.

- [236] Ivan Rayment, Hazel M Holden, Michael Whittaker, Christopher B Yohn, Michael Lorenz, Kenneth C Holmes, and Ronald A Milligan. Structure of the actin-myosin complex and its implications for muscle contraction. *Science*, 261(5117):58–65, 1993.
- [237] Gary T Yamaguchi. *Dynamic modeling of musculoskeletal motion: a vectorized approach for biomechanical analysis in three dimensions*. 2005.
- [238] KR Kaufman. Future directions in gait analysis. *Gait analysis in the science of rehabilitation*, 1999.
- [239] Stuart M McGill and Robert W Norman. Dynamically and statically determined low back moments during lifting. *Journal of biomechanics*, 18(12):877–885, 1985.
- [240] Jos J De Koning, Gert De Groot, and Gerrit Jan van Ingen Schenau. Speed skating the curves: A study of muscle coordination and power production. *Journal of Applied Biomechanics*, 7(4):344–358, 1991.
- [241] Darren J. Stefanyshyn and Benno M Nigg. Contribution of the lower extremity joints to mechanical energy in running vertical jumps and running long jumps. *Journal of sports sciences*, 16(2):177–186, 1998.
- [242] Yifeng Jiang, Tingnan Zhang, Daniel Ho, Yunfei Bai, C Karen Liu, Sergey Levine, and Jie Tan. Simgan: Hybrid simulator identification for domain adaptation via adversarial reinforcement learning. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2884–2890. IEEE, 2021.
- [243] Karim Armanious, Chenming Jiang, Marc Fischer, Thomas Küstner, Tobias Hepp, Konstantin Nikolaou, Sergios Gatidis, and Bin Yang. Medgan: Medical image translation using gans. *Computerized medical imaging and graphics*, 79:101684, 2020.
- [244] Laila Rasmy, Yang Xiang, Ziqian Xie, Cui Tao, and Degui Zhi. Med-bert: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ digital medicine*, 4(1):86, 2021.
- [245] Masashi Matsumura, Kazuna Ichikawa, Hitoshi Takei, et al. A maximum muscle strength prediction formula using theoretical grade 3 muscle strength value in daniels et al.’s manual muscle test, in consideration of age: An investigation of hip and knee joint flexion and extension. *Rehabilitation Research and Practice*, 2017, 2017.
- [246] Farid Mobasser and Keyvan Hashtrudi-Zaad. A comparative approach to hand force estimation using artificial neural networks. *Biomedical engineering and computational biology*, 4:BECB–S9335, 2012.
- [247] Luca Modenese and Andrew TM Phillips. Prediction of hip contact forces and muscle activations during walking at different speeds. *Multibody System Dynamics*,

- 28(1):157–168, 2012.
- [248] Gaofeng Wei, Feng Tian, Gang Tang, and Chengtao Wang. A wavelet-based method to predict muscle forces from surface electromyography signals in weightlifting. *Journal of Bionic Engineering*, 9(1):48–58, 2012.
- [249] Irina Guseva Canu, Sandy Carla Marca, Francesca Dell’Oro, m Balzs, Enrico Bergamaschi, Christine Besse, Renzo Bianchi, Jovanka Bislimovska, Adrijana Koscec Bjelajac, Merete Bugge, et al. Harmonized definition of occupational burnout: A systematic review, semantic analysis, and delphi consensus in 29 countries. *Scandinavian journal of work, environment & health*, 47(2):95, 2021.
- [250] Wenfei Song, Miao Gui, Xinquan Wang, and Ye Xiang. Cryo-em structure of the sars coronavirus spike glycoprotein in complex with its host cell receptor ace2. *PLoS pathogens*, 14(8):e1007236, 2018.
- [251] Mendeley Data Repository. <https://data.mendeley.com/datasets/svx74xcrjr/1>.
- [252] Abhijit Guha and Debabrata Samanta. Hybrid approach to document anomaly detection: An application to facilitate rpa in title insurance. *International Journal of Automation and Computing*, 18(1):55–72, 2021.
- [253] Markus Goldstein and Andreas Dengel. Histogram-based outlier score (hbos): A fast unsupervised anomaly detection algorithm. *KI-2012: Poster and Demo Track*, pages 59–63, 2012.
- [254] Varun Badrinath Krishna, Ravishankar Iyer, and William Sanders. Arima-based modeling and validation of consumption readings in power grids. pages 199–210, 05 2016.
- [255] Aya Yokoe Bishnu Nepal, Motoi Yamaha. Electricity load forecasting using clustering and arima model for energy management in buildings. In *Wiley Online Library*, pages 1–6, 2019.
- [256] Jinsung Yoon, Daniel Jarrett, and Mihaela Van der Schaar. Time-series generative adversarial networks. *Advances in neural information processing systems*, 32, 2019.
- [257] Hengzhi Pei, Kan Ren, Yuqing Yang, Chang Liu, Tao Qin, and Dongsheng Li. Towards generating real-world time series data. In *2021 IEEE International Conference on Data Mining (ICDM)*, pages 469–478. IEEE, 2021.
- [258] Amit Gupta and Rajesh Bhatia. Ensemble approach for web page classification. *Multimedia Tools and Applications*, 80(16):25219–25240, 2021.

List of Publications

1. H Bansal, B Chinagundi, PS Rana and N Kumar, “*An Ensemble Machine Learning Technique for Detection of Abnormalities in Knee Movement Sustainability*”, Sustainability, MDPI, 14(20), 13464:80, 2022. [SCIE, IF 3.3]
2. H Bansal, B Chinagundi, PS Rana and N Kumar, “*Time series generative adversarial network for muscle force prognostication using statistical outlier detection*”, Expert Systems, Wiley, e13653, 2024. [SCIE, IF 3.0]