

Drug Synergy for Cancer using Computational Intelligence Techniques

**A Thesis submitted in fulfillment of the requirement for the award of
the degree of**

Doctor of Philosophy

in

Computer Science and Engineering

Submitted By

Harpreet Singh

(Registration No: 951503005)

Under the guidance of

Dr. Prashant Singh Rana
Assistant Professor

Dr. Urvinder Singh
Associate Professor



Computer Science and Engineering Department
Thapar Institute of Engineering and Technology,
Patiala, Punjab – 147004, India.

December, 2020

Certificate

I, Harpreet Singh, Regn. No. 951503005, hereby declare that the thesis entitled **"Drug synergy for cancer using computational intelligence techniques"** submitted to the Computer Science and Engineering Department at Thapar Institute of Engineering and Technology, Patiala, Punjab, India is an authenticated record of my own work for the award of the degree of "Doctor of Philosophy" under the supervision of Dr. Prashant Singh Rana and Dr. Urvinder Singh. This report has not been submitted to any other institution for award of any other degree.

Place: Patiala

Date: 28-12-2020


Harpreet Singh

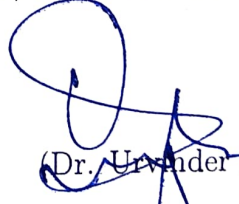
Regn. No. 951503005

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Verified by:


(Dr. Prashant Singh Rana)

Computer Science and Engineering Department,
TIET, Patiala-147001, India.


(Dr. Urvinder Singh)
Electronics and Communication Engineering Department,
TIET, Patiala-147001, India.

Acknowledgements

I express my heartfelt gratitude to all the ingredients of composition, who, while writing this thesis, have played their crucial parts, and become a part of this effort. First and foremost, I wish to acknowledge the benevolence of omnipotent Almighty whose blessings provided me strength, courage and patience to accomplish the journey and overcome all the bumpy rides.

With a great sense of appreciation, I express my regard and sincere feelings of indebtedness to my supervisors Dr. Prashant Singh Rana, Assistant Professor, Computer Science Engineering Department and Dr. Urvinder Singh, Associate Professor, Electronics and Communication Engineering Department, Thapar Institute of Engineering and Technology for their invaluable guidance, motivation, encouragement, moral support and precious co-operation. Their generous and positive attitude for readily helping me with queries, clearing doubts and finding solutions has facilitated the process to a great extent. Performing research and tasks under their guidance has always provided a steady direction to this work. It has been a great pleasure and experience to work under their sanctuary.

I would always be grateful to Head of Department, Prof. Maninder Singh and Prof. Inderveer Channa (Associate Head and DOSA), whose support has always been prominent throughout the process. I am much obligated to the Director, Dean (RSP) and the Management of Thapar Institute of Engineering and Technology, who provided me all the necessary resources and encouragement for success since the beginning, and towards the achievement of end results. I feel appreciative towards the esteemed members of the doctoral committee Dr. Ajay Kumar Loura, Dr. Sushma Jain and Dr. Mandeep Singh who have been providing continuous feedback for my work, and backed its procession. I owe my thankfulness to the faculty and support staff of Computer Science and Engineering department for their consistent support whenever it was asked for. I am grateful to thank Dr. Prashant Singh Rana and Dr. Urvinder Singh, who offered great support in the beginning,

and beyond.

I pay deep gratitude to my parents S. Amarjit Singh and Mrs. Paramjit Kaur for the immense love and affection, and for encouraging despite all the ups and downs in the journey. I am greatly appreciative of the unconditional love and support of my sister Jaspreet Kaur (Mini), and her being so encouraging throughout the journey. I am utterly thankful to Mrs. Deepika Rana and Dr. Sarabjit Kaur for whose support was an added fillip. I would also like to thank Dr. Seema Bawa, Dr. Neeraj Kumar and Mr. Mandeep Singh Tinna whose encouragement was an added strength for me, always.

I fall short of words expressing great appreciation towards my friends and colleagues Harman Brar, Sumit Dey, Deepika Sirohi, Jasleen Bhullar, Gagangeet, Prabhleen, Navneet, Rajanpreet and Ravneet who were my citations for knowledge, suggestions and motivation during the journey, and made the process aesthetic and memorable. My appreciative heart must have missed the names of lots of well-wishers, critics, friends and beloved ones; my sincerest regards to everyone, who kindly supported me enthusiastically during this journey of knowledge.

Harpreet Singh

Special Thanks to NVIDIA Graphics Pvt. Ltd.

I would like to acknowledge the support from NVIDIA Corporation to sponsor the Quadro P6000 GPU (worth of Rs 5,00,000), under the scheme of "NVIDIA GPU Grant for Academic" for this research.

Dedicated to my world-my mother

"God could not be everywhere, and therefore he made mothers."

-Rudyard Kipling

Having started researching on a disease, which is a nightmare to humans, I felt a yearning to contribute something significant towards the on-going work on cancer. Never had I thought, I was going to witness the dreaded disease so closely. Everyone wants law of attraction to work, in positive ways, without a say. But life, owing to its unpredictable nature, never happens the way we want it to. The world fell apart for me, when my mother was diagnosed with breast cancer. With all sorts of bad thoughts striking your mind, the real wars are the ones that you fight in your head. Recollecting yourself becomes a necessity, and tough times only bring out the stronger you! I express heartiest thanks to the angelic souls, Dr. Ravimohan and Dr. Smita Patnaik, PGI, Chandigarh, also my research collaborators, who backed me during those rugged times. With their benign facilitation, treatment could be started instantly, which endowed the necessary courage and positivity in me. This thesis also involves my personal experiences related to the disease and its treatment, which is even more intense than the pain associated with the disease. My mother underwent 8 chemotherapy sessions, a breast removal surgery, and 15 radiation therapy sessions, which, besides being painful, resulted in a number of side-effects. With the blessings of the Almighty, and support of the doctors and hospital staff, my sister and I were able to assist my mother tackle the psychological pressure of the disease, and strengthen her will power. I dedicate this research to my mother, as this work shall honour me with the title of a Doctor, and this idea of PhD started with my mother's dream for me. My success belongs to my mother, and without a mention of her, the work would have been incomplete.

Abstract

Cancer still remains a major health concern in the world. It is still a challenge for researchers to develop an effective cancer therapy in most types of cancers. Although cancer presents with a very high mortality rate, there are certain treatments available and the number is growing with new breakthroughs every day. A variety of cancer treatments exist, and the type of treatment a patient receives mainly depends upon the type of cancer they present with and its current stage. Various treatments of cancer exist such as radiation therapy, chemotherapy, drug synergy, hormonal therapy and surgery. Some patients need only one kind of treatment, but most of the patients will receive a combination of these. There are many side effects of aforementioned treatments including hair loss, vomiting, nausea, blood disorders, itchiness, constipation, diarrhea, vaginal dryness, enlarged and tender breast. Researchers are trying very hard to reduce the side effects caused by cancer treatment. Now days, more focus is being given by researchers towards targeted therapy such as drug synergism. It mainly targets the cancerous cells. Drug synergism is a branch which finds the optimal combination of two or more drugs on the basis of DDI (Drug Drug Interaction) to kill the cancerous cells in a more efficient manner as compared to their individual effects. However, there are huge number of drugs which leads to the combinatorial explosion problem. To handle the numerous amounts of drugs, machine learning (ML) plays a major role in this area. ML is categorized in supervised and unsupervised learning. In supervised learning, different problems can be categorized under classification and regression whereas in unsupervised learning, clustering problems are considered. Problem of drug synergy falls under the category of regression problems in which synergy score (considered as target) of different drugs is predicted from the given set of input parameters. ML techniques can be used to improve the results of such type of problems. There are some studies exist to predict the synergy score. But these studies have some limitations such as use of

single model, a smaller number of input parameters, and lack of drug data to train the models. Because of such issues, the trained model may or may not produce a reliable and efficient prediction. Single model can be replaced with the ensemble models to predict synergy score. Ensemble learning is a process of combining more than one model or combination of one model with some optimization algorithm to solve a given computational intelligence problem. Generally, it is used to enhance the predictability as well as to improve the robustness of a model. The ensemble approach among different models is used because it is capable of boosting the weak learners. This approach has advantage that the ensemble model can adapt any diversity in the data more correctly as compared to single model. Neuro-fuzzy based ensembling technique is designed using biased-weighted aggregation (addition of more weights to model with higher prediction score). Four with the highest accuracy among nine ML models have been selected for developing ensemble-based machine learning model. On the other fold, ensembling of model with optimization algorithm gives the best result by finding the best combinations of parameters to be tuned for the given model. An effective ensemble based machine learning techniques are defined to predict drug synergy, such as SVM based differential evolution and evolutionary based neural network structure using Adaptive Lévy method to overcome the parameter tuning issue. It regulates the precision in prediction. It suggests that the ensemble approach is more efficient than the single model. Therefore, ensemble models have been proposed to predict the drug synergy score. To check the consistency of proposed ensemble model prediction, repeated k-fold cross validation has been performed. The existing machine learning techniques and this evolutionary based technique are tested on drug synergy score data. An extensive analysis shows the better performance of proposed method over existing techniques, in terms of following evaluation parameters such as accuracy, coefficient of correlation and root mean squared error.

Contents

Certificate	i
Acknowledgment	iii
Abstract	ix
Contents	xi
List of Figures	xv
List of Tables	xvii
List of Abbreviations	xix
1 Introduction	1
1.1 Background	6
1.2 Cancer Treatments	7
1.2.1 Surgery	8
1.2.2 Chemotherapy	9
1.2.3 Radiotherapy	9
1.2.4 Bone Marrow Transplant	9
1.2.5 Immunotherapy	9
1.2.6 Hormone Therapy	10
1.2.7 Synergy	10
1.3 Drug Synergy	10

1.4	Machine Learning	12
1.4.1	Supervised Learning	13
1.4.2	Unsupervised learning	14
1.4.3	Reinforcement learning	15
1.5	Machine Learning Approaches	15
1.6	Ensemble Machine Learning	19
1.6.1	Need of Ensemble Machine Learning	19
1.6.2	Techniques of Ensemble Machine Learning	20
1.7	Metaheuristic Approaches	22
1.8	Metaheuristic with Machine Learning	22
1.9	Performance Evaluation Parameters	23
1.9.1	Classification Performance Parameters	23
1.9.2	Regression Performance Parameters	26
1.10	Motivation	27
1.11	Thesis Organization	28
2	Literature Review	31
2.1	Drug Synergy Methods	34
2.1.1	Kinetic Methods	35
2.1.2	System Biology Methods	35
2.1.3	Computational Methods	36
2.1.4	Mixed Integer Linear Programming Methods	37
2.1.5	Machine Learning-based Methods	38
2.1.6	Drug Synergy Metrics	40
2.1.7	Surgery Based Methods	46
2.1.8	Drug synergy effects in Antimicrobial Treatments	46
2.1.9	Drug synergy based on Anti-tumor Agents	48
2.1.10	Drug Synergy effects in Cancer Treatments	48
2.2	Machine Learning in the field of Drug Synergy	51

2.3	Limitations	58
2.4	Comparison	58
2.5	Gaps in Literature	60
2.6	Research Contributions	60
2.7	Research Objectives	62
3	Ensemble based DE scheme	63
3.1	Introduction	64
3.1.1	SVM for regression	66
3.1.2	Conventional differential evolution	67
3.2	Proposed technique	70
3.2.1	Ensemble mutation operator based differential evolution . . .	70
3.2.2	Drug synergy prediction using ensemble based differential evolution	73
3.3	Materials and Methods	74
3.3.1	Experimental Setup	74
3.3.2	Dataset	74
3.3.3	Results and discussions	76
4	Ensemble based machine learning Scheme	83
4.1	Introduction	84
4.1.1	Contributions	87
4.2	Proposed technique	87
4.2.1	ANFIS	87
4.2.2	Random forest	88
4.2.3	Designed Ensembling based model	89
4.3	Performance analysis	91
4.3.1	Accuracy analysis	92
4.3.2	Root mean square error analysis	94
4.3.3	Coefficient of correlation analysis	95

4.4	Receiver operating characteristic curve	98
5	Adaptive Lévy based Neural Network Scheme	99
5.1	Introduction	100
5.1.1	Contributions	102
5.2	Artificial Neural Networks(ANN)	103
5.2.1	The Process of Learning in ANN	103
5.3	Evolutionary Algorithms	105
5.4	The optimization of ANNs based on GAs	106
5.5	Improved Evolutionary Neural Network	107
5.5.1	Combined Mutation	107
5.5.2	Adaptive Lévy	108
5.5.3	Structure and Parameter's Tuning	109
5.6	Dataset	111
5.7	Results and discussions	112
5.7.1	Accuracy analysis	113
5.7.2	Root mean square error analysis(RMSE)	115
5.7.3	Correlation analysis	116
5.7.4	Comparative Result Analysis	118
6	Conclusion and Future Scope	119
6.1	Conclusion	119
6.2	Future Scope	121
	Bibliography	123
	List of Publications	141

List of Figures

1.1	Cancer occurrences in different geographic regions.	2
1.2	Cancer deaths in different geographic regions.	2
1.3	Distribution of types of cancer in 2018.	3
1.4	Cases of cancer in USA from 2012-2016 based on age.	4
1.5	Challenges and issues in existing treatments of cancer	7
1.6	Prediction methods for drug synergy	12
1.7	Basic structure of machine learning process	13
1.8	Categorization of Machine Learning Techniques	13
1.9	Bagging ensemble technique	21
1.10	Stacking ensemble technique	21
2.1	Synergistic/antagonistic effects of different drug combinations	31
2.2	Targeted diseases using drug combinations	32
2.3	The checkerboard method depicting synergy of two-drug combination	44
2.4	Proposed computational solutions for drug synergy	61
3.1	Ensemble based mutation	71
3.2	Dose response surface	75
3.3	Dose response curves of different effects	75
4.1	Flowchart of the proposed technique	85
4.2	Flowchart of the proposed technique	90
4.3	ROC curve	97

5.1	Evolutionary Neural Network Model	110
5.2	Adaptive Lévy Technique For Mutation	110
5.3	Accuracy Comparison	114

List of Tables

2.1	Various approaches in anti cancer synergy prediction	35
2.2	Statistical and computational models used for drug synergy	41
2.3	Various methods for measuring effectiveness of synergism	42
2.4	Summary of cancer signature and surgery methods	47
2.5	Summary of anti-tumor agents and methodologies	49
2.6	Limitations of statistical/computational techniques	53
2.7	Limitations of anti-tumor agents and its methodologies	57
2.8	Limitations based on cancer surgery/ signature based methodologies .	58
2.9	Comparison of different drug synergy methods	59
3.1	Features of Drug Synergy Dataset	76
3.2	Comparative analysis of accuracy on different data partitions	78
3.3	Comparative analysis of RMSE (root mean square error) on different data partitions	79
3.4	Comparative analysis of coefficient of correlation on different data partitions	81
4.1	Comparative analysis of Accuracy on different data partitions	94
4.2	Comparative analysis of RMSE (root mean square error) on different data partitions	95
4.3	Comparative analysis of correlation on different data partitions	97
5.1	Comparative analysis of Accuracy on different data partitions	113

5.2	Comparative analysis of RMSE(root mean square error) on different data partitions	116
5.3	Comparative analysis of correlation on different data partitions	116
5.4	Comparitive Analysis	118

List of Abbreviations

ACO	Ant colony optimization
ADMM	Alternating Direction Methods of Multipliers
ADR	Adverse Drug Reaction
AIS	Artificial Immune Systems
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
ASDCD	Antifungal Synergistic Drug Combination Database
AUC	Area Under Curve
AZT	Azidothymidine
BN	Bayesian Network
CLSI	Clinical Laboratory Standards Institute
CI	Combination Index
COP	Complex Optimization Problems
DDIs	Drug-Drug Interactions
DCDB	Digital Cadastral DataBase
DE	Differential Evolution
DENFIS	Dynamic Evolving Neural-Fuzzy Inference System
DLBCL	Diffuse Large B Cell Lymphoma
DNA	Deoxyribo Nucleic Acid
DREAM	Dialogue on Reverse Engineering Assessment And Approaches
DT	Decision Tree

ER	Error Rate
FP	False Positive
GA	Genetic Algorithm
GFS.GCCL	Genetic Cooperative-Competitive Learning Method
GRASP	Greedy Randomized Adaptive Search Procedure
HNAI	Hetrogenous Network Assisted Inference
HTS	High-Throughput Screening
ILS	Iterated Local Search
KNN	K-Nearest Neighbors
M5P	M5 Model Trees
MA	Metaheuristic Algorithm
MER	Minimum Error Rate
MIC	Minimum Inhibitory Concentration
ML	Machine Learning
MWL	Minimum Cost-Weighted Error Rate
NN	Neural Network
PCA	Principle Component Analysis
PCI	Probability Concordance Index
PD	Pharmacodynamics
PK	Pharmacokinetics
PSO	Particle Swarm Optimization
r	Coefficient of Correlation
RF	Random Forest
RMSE	Root Mean Square Error
SA	Simulated Annealing
SEER	Surveillance, Epidemiology, and End Results
Sens	Sensitivity

SI	Synergy Index
SMILES	Simplified Molecular Input Line Entry System
Spec	Specificity
SS	Scatter Search
SVM	Support Vector Machine
TP	True Positive
TS	Tabu Search
VNS	Variable Neighborhood Search

Chapter 1

Introduction

Human body is composed of cells at an elementary level. New cells emerge out of existing ones when body demands, and they die upon ageing or when they get damaged. They are then replaced by new cells. If any genetic changes interfere with the standard process described above, cells start growing in an uncontrolled manner, which could result in the formation of a mass known as tumour. Hence, it could mark the onset of cancer, depending upon the type: cancerous or benign. Cancerous tumours may spread across various body parts via blood or lymphatic systems, and are malignant. When these genetically abnormal cells adapt to a tissue micro environment that is distant from the primary tumour, it is called metastasis and is extremely fatal.

Cancer is one of the most dreadful diseases and second leading cause of death globally, and according to WHO in 2018 alone 9.6 million deaths i.e. every 1 in 6 deaths, were caused due to cancer. Global patterns in Figure 1.1 show that for men and women combined, more than half of the cancer occurrences and cancer related deaths are from Asia and Africa [1], mainly because they contain nearly 60% of the world's population.

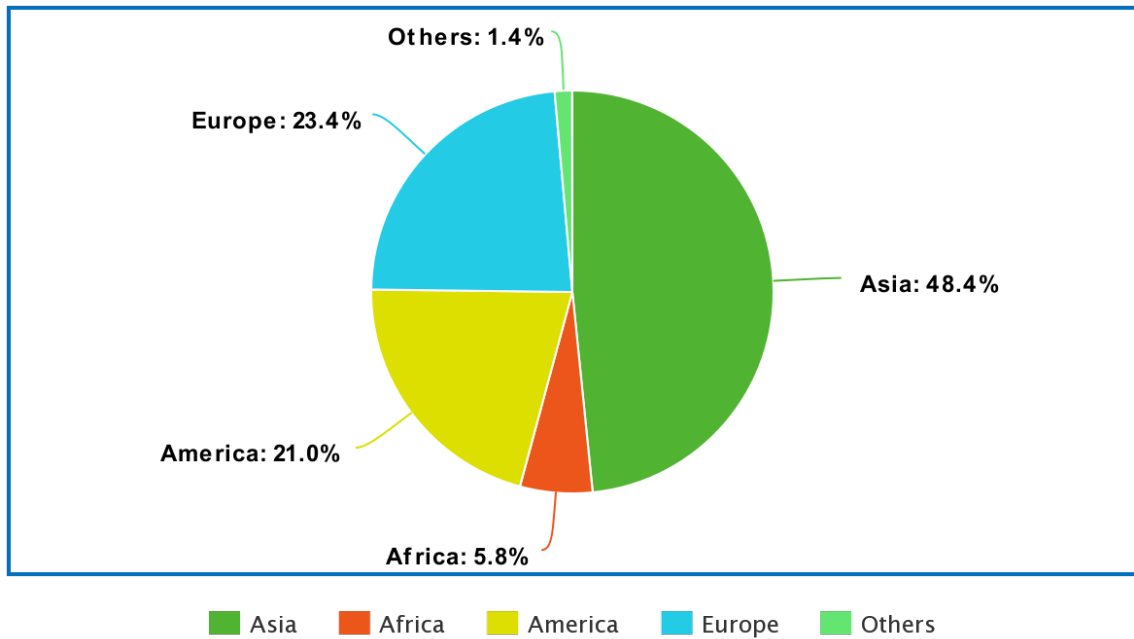


Figure 1.1: Cancer occurrences in different geographic regions.

Out of these, nearly 70% of deaths take place in countries with low or middle incomes, primarily due to inaccessible diagnoses and treatment. Asia and Africa alone account for 64.6% of cancer deaths as can be seen in Figure 1.2.

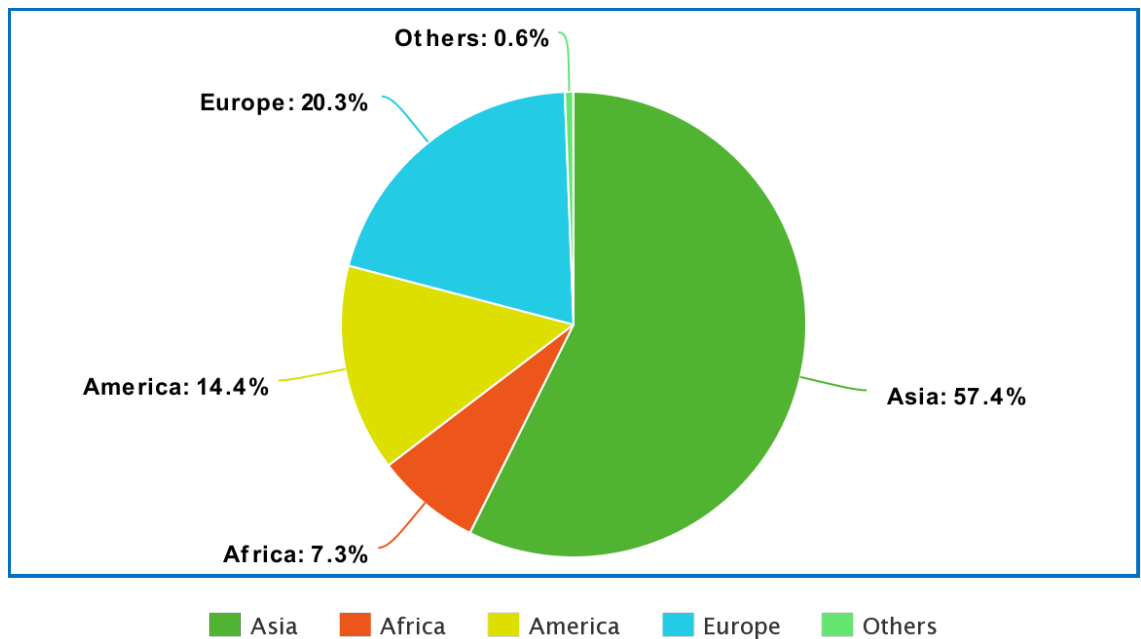


Figure 1.2: Cancer deaths in different geographic regions.

Cancer is generally classified by their point of origin in the body or the type of

cell they originate in. According to the 2018 WHO report, the largest number of cases were of lung and breast cancer, and subsequently, lung cancer was also the most common cause of cancer deaths.

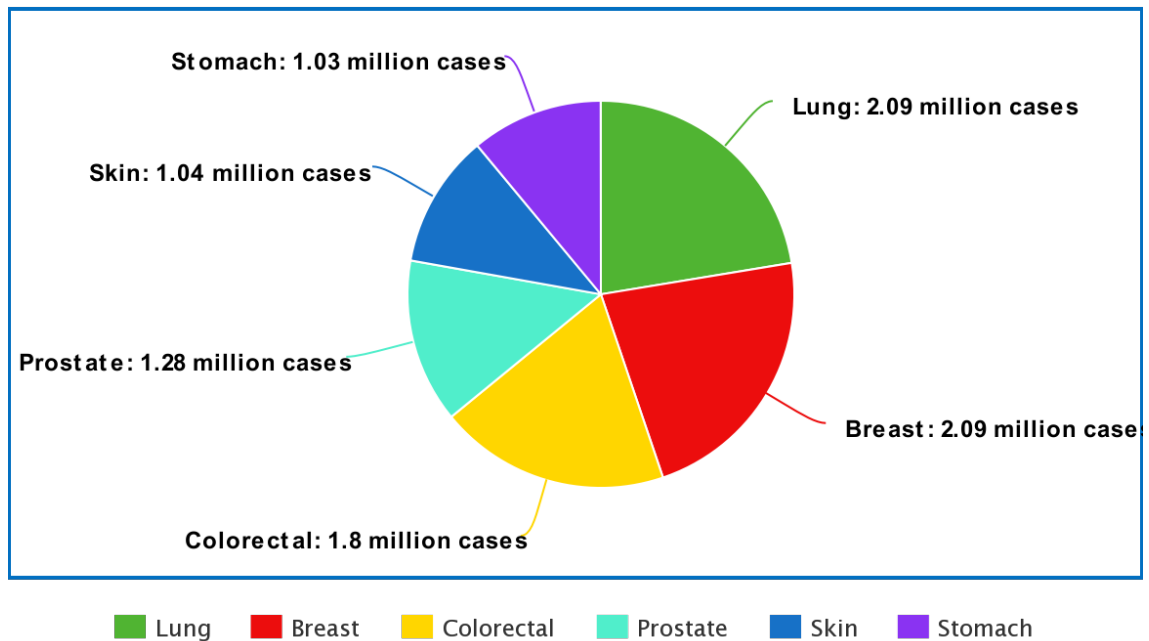


Figure 1.3: Distribution of types of cancer in 2018.

We can see from Figure 1.3 that lung, breast and colorectal cancer together make up one third of the cancer occurrences. However, cancer does not affect every person equally and the chances of a cancerous tumour and the mortality rate differ depending upon the age, gender and other factors. From Figure 1.4, it can be seen that in USA, between 2012-2016 cancer mainly affected people in the 60-80 age group.

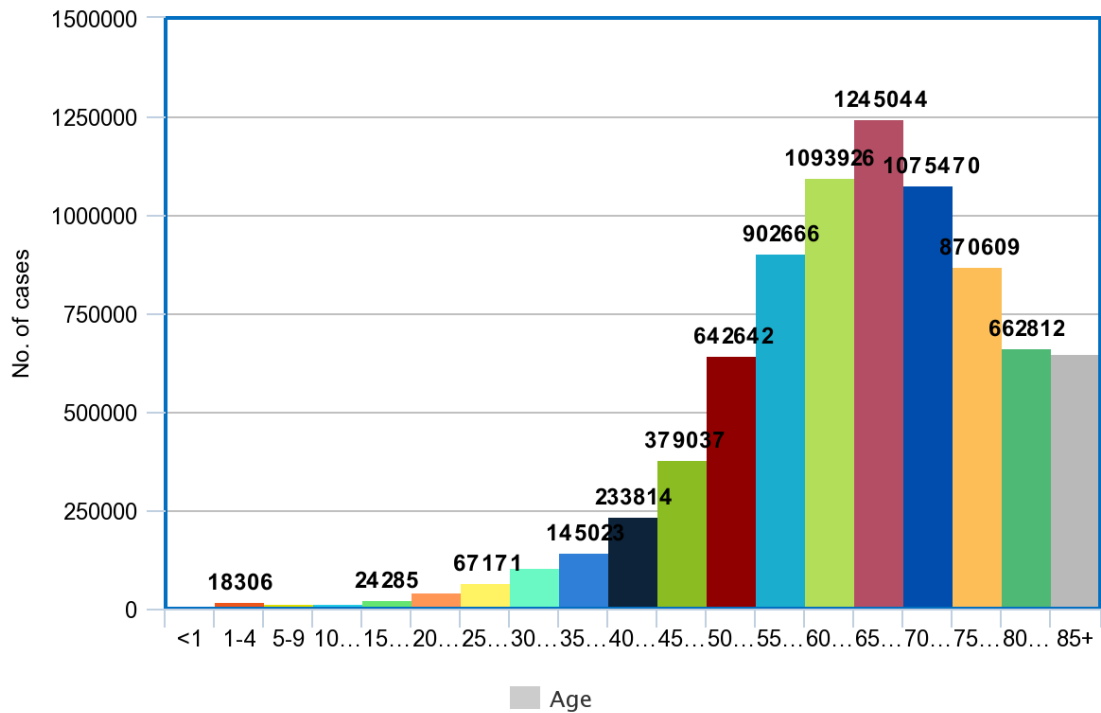


Figure 1.4: Cases of cancer in USA from 2012-2016 based on age.

Although cancer causes a very high mortality rate, there are certain treatments available and the number is growing with new breakthroughs every day. A variety of cancer treatments exist, and the type of treatment a patient receives mainly depends upon the type of cancer he is suffering from and its current stage. Various treatments of cancer exist such as radiation therapy, chemotherapy, drug synergy, hormonal therapy and surgery. Some patients need only one kind of treatment, but most of the patients will receive a combination of these. There are many side effects of aforementioned treatments like hair loss, vomiting, nausea, blood disorders, itchiness, constipation, diarrhea, vaginal dryness, enlarged and tender breast. Researchers are trying very hard to reduce the side effects caused by cancer treatment. Now days, more focus is being given by researchers towards targeted therapy such as drug synergism which mainly targets the cancerous cells. Drug synergism is a branch which finds the optimal combination of two or more drugs on the basis of DDI (Drug Drug Interaction) to kill the cancerous cells in a more efficient manner as compared to their individual effects. However, there are large number of drugs

which lead to the combinatorial explosion problem. To handle the huge amount of drugs, machine learning (ML) plays a major role in this area. Apart from prediction of drug synergy in medical field, important examples of ML include prediction of heart disease, diabetes, varieties of cancer, like breast, lung, colon and many more which were primarily passed down generation to generation through genetic, communicable and other effects. The impact of ML in medical field has been increased in our society in recent years.

Different problems are mainly categorized under supervised and unsupervised machine learning based upon the respective dataset. Dataset contains the collection of tuples under different parameters of respective problem. Dataset is divided into two parts; one part is used to train the classifier and the second part is used for the testing of trained classifier. Supervised and unsupervised machine learning can also be defined in terms of patient dependent and patient independent. If the system uses the training data in predicting the disease of a particular person, it is called patient dependent or supervised and if no training data is provided of specific person but system has to train itself for the process of predicting disease, it is called patient independent or unsupervised. Patient dependence determines the range till which system requires a particular information of one's individual features to process the parameters of different disease. In patient independence, system of medical diagnosis learns the various parameters of different types of data in case of cancer such as SEER (Surveillance, Epidemiology, and End Results), clinical data, microarray, gene expression, molecular SEER, genomic data and based on that learning or training of individual's features, it gives the relevant output.

In supervised learning, different diseases can be categorized under classification and regression problems. Classification problem is basically predicting a discrete output like either 'disease' or 'no disease' and regression problem is predicting a continuous output like 'degree of disease'. Various terms used under classification and regression for medical system are prognosis, susceptibility, recurrence and survival. If we classify the patients in high risk and low risk, it comes under classification

technique [2]. Based upon classification i.e. binary or multi classification of disease at different levels or stages, proper treatment is recommended for the given disease. Regression can be used to predict the health care costs in future from previous history such as diagnoses, expenditure etc which can help to choose a better health insurance plan [3]. There are different evaluation parameters of regression and classification. For regression, various parameters are root mean square error, accuracy, R^2 and correlation. On the other fold, evaluation parameters of classification are specificity, sensitivity, precision, recall, receiver operating characteristics, area under curve. All the classification parameters are calculated with the help of confusion matrix or error matrix.

Another application of regression is drug synergy as discussed previously. It is commonly used as treatment method for most dreaded diseases like cancer. The aim of drug synergy is to achieve reduction in dose and toxicity, to delay/reduce the induction of drug resistance and reduce synergistic therapeutic effect [2]. The major outcomes of synergism are minimized drug resistance and reduced toxicity. A recent study in cancer has revealed that mere analysis of genetic mutation is insufficient to study the growth of cancer. Instead, cancer should be considered as a complex tissue with evolving cells. These cells communicate with the surrounding micro environment, study of which would be beneficial [2].

1.1 Background

Cancer still remains a major health concern in the world. It is still a challenge for researchers to develop an effective cancer therapy in most types of cancers. Cancer can affect any part of the human body such as breasts, lungs, skin, pancreas and other organs and glands. The human body is made up of cells. For normal growth, these cells grow and divide into new ones as per body requirement. These new cells replace the aged and blemished cells which die with time. But this process breaks

down when cancer starts to develop in the body. As a result abnormal, damaged or old cells survive instead of dying, and more new cells are created abruptly. This extreme growth of cells become uncontrollable and takes the form of cancer.

The cancerous cells are different in behaviour from normal cells in many ways. Cancer cells are less specialized as compared to normal cells which means cancer cells do not mature into cell having specific functions whereas normal cells do. This is a major reason due to which cancer cells continue to divide and grow uncontrollably. In addition, cancer cells ignore the signals that gives the indication to the cells to stop growing further, thereby making the situation worse.

1.2 Cancer Treatments

A number of cancer treatments are available now a days, depending upon the type of cancer and its current stage that it has reached. In some cases, patients can be treated with only one treatment but in some cases, patients have to be given combination of different treatments as discussed previously, like radiation and/or chemotherapy therapy followed by surgery. The major treatments categorizations are shown in Figure 1.5.

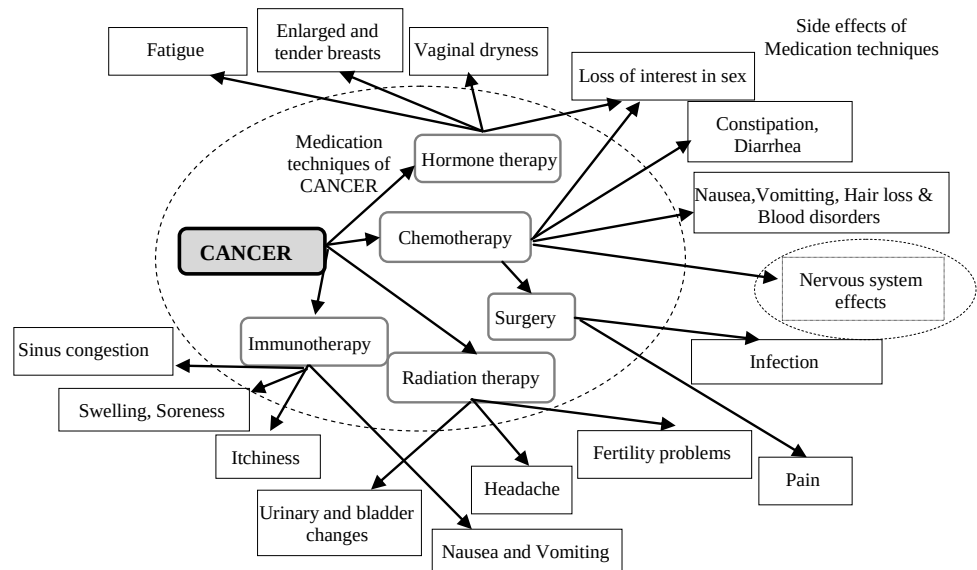


Figure 1.5: Challenges and issues in existing treatments of cancer

There are three ways in which we can broadly classify cancer treatments:

i. **Primary treatment**

The goal of this treatment is to completely remove the cancer from a patient's body. Any treatment can be administered as a primary treatment depending upon the type and stage of cancer. Surgery comes under the category of primary treatment.

ii. **Adjuvant treatment**

This treatment is used to minimize the chances of recurrence of cancer after the primary treatment. Commonly, radiotherapy and chemotherapy are used as adjuvant treatment.

iii. **Palliative treatment**

This treatment is used to relieve the patient of the side effects that remain after the patient has been subjected to the above two treatments. Simple medications are used to relieve symptoms such as nausea, shortness of breath etc.

The different types of treatments that are utilized today are categorized as given below:

1.2.1 Surgery

Surgery is mainly used when the tumour can safely be excised from the affected organ. Surgery can be effective but it has side effects too, like some blood may leak from small blood vessels, infection etc. In case of advanced stage of cancer, it may spread to other body parts. So, surgery might not help as the best solution to remove the cancerous cells from multiple body parts [4]. However, this treatment is only possible when the tumour is small and will not affect the body after being removed. Generally, surgery is done after shrinking the tumour using chemotherapy or radiation.

1.2.2 Chemotherapy

This treatment deals with anti-cancer drugs as part of a chemotherapy treatment. It uses drugs to kill the cancerous cells. It is an aggressive form of treatment that can even kill metastasized cancer cells as the drugs flow through the patient's entire body. It aims to reduce the symptoms and prolong life. But it also destroys the normal cells. Though, chemo treatment has improved in last years, but still patient has to face severe side effects during chemotherapy such as extreme weakness, hair loss, nausea etc.

1.2.3 Radiotherapy

This treatment uses high energy beams like X-Rays based on ionizing radiation, in which a high dose of radiation is given for killing cancer cells and to cause tumours to shrink. The radiation can be performed either externally or it can be placed inside the body. It damages both healthy and cancerous cells. It destroys the genetic material of the cells which controls the normal growth of cells which leads to many side effects such as urinary and bladder changes, headache, fertility problem etc.

1.2.4 Bone Marrow Transplant

Blood cells are produced inside the bone marrow; hence, this procedure is mainly used for leukaemia. A donor is needed for bone marrow transplant and the chances of an immediate family member being a marrow match are higher than an unrelated donor. This treatment also allows the doctor to administer higher doses of chemotherapy.

1.2.5 Immunotherapy

The treatment aims to inhibit the cancer cells from spreading to other body parts by strengthening the defense mechanism of body. It does not directly attack the cancer cells but it helps the immune system to destroy the cancer cells. This treatment is effective only in pre stage or along with other treatments in advanced stage.

1.2.6 Hormone Therapy

This treatment focuses on stopping or slowing down the growth of cancer cells. It is also called hormone withdrawal therapy or hormone manipulation or endocrine therapy. This therapy falls into two major categories, those that interfere with hormones behaviour in the body and those that block the ability of hormonal production.

1.2.7 Synergy

Synergy basically means interaction of two or more forces or agents or drugs so that a combination of these effects surpasses the individual ones. It is generally quantified on the basis of potency and efficacy. A higher synergy of potency means a lower drug dosage and hence lesser side effects. A higher synergy of efficacy means a better chance of the drug being effective and a lesser chance for the disease to relapse. Synergy is mainly gaining importance in medical fields for treating those patients having complex diseases and who respond poorly to single agent therapies [5].

Therapies that specifically target cancer cells such as pathways or signalling proteins have been remarkably successful in cancer treatment. But unfortunately, sometimes tumours frequently develop resistance leading to even more aggressive forms of disease [3]. Cancer cells grow very fast, and these drugs often work against them very effectively. In chemo therapy, most of the drugs kill the malignant cells which grow and divide quickly in the body. But on the other fold, it also kill the healthy cells too which leads to some side effects of these drugs as discussed in the Section 1.2.2.

1.3 Drug Synergy

The term synergy basically means "working together". This is a concept that can be applied in many situations of life and various settings, like technical systems, mechanics, medical field, and many more. An interaction of biological active agents is an important feature of biomedicine and pharmacology. In this perspective, interaction represents the biological activity of combining several drugs at the same time.

This field basically deals with the study of effect of combination of drugs collectively for achieving a significantly superior result as compared to a single drug.

Synergy describes a system which represents combination or interaction of two or more agents together which leads to better result [6]. There can be a number of possible ways of such interaction of these drugs or agents i.e. their combination can produce a better result or may not affect each other, or result could be less as compared to summation of individual effects. These compounded drug effects are generally classified into three main types: additive which means the combination effect is equivalent to that of the summation of the individual components; synergism refers to the effect that exceeds the combined effect of two individual components, and antagonism, in which final combined effect is lower as compared to individual component.

It has been seen that combination therapies have higher efficiency than single drug treatments when complex diseases such as cancer, asthma and AIDS are considered [7]. Since the synthesis of new molecular entities is becoming stagnant, combining drugs together can prove to be an effective strategy for untreatable diseases [8]. However, the properties of FDA(Food and Drug Administration)-approved drugs are still not studied enough to deduce the pharmacokinetic interactions between them. Hence, other methods need to be utilized to predict the synergism between the drugs [9]. The major advantage of using synergistic combinations is that the effectiveness of the treatment remains the same while reducing the amount of dosage. The standard procedure is through drug screening experiments, but these become highly complex and costly to perform in case the number of drugs increase. Finding the combined effect of two individual drugs is a challenge. During this drug discovery process, finding the effect of drug combination in vitro relies heavily upon the experiments. This, however, makes it difficult to test all possible combinations due to larger combinatorial search space.

In recent years, proposals of multiple computational approaches have been made

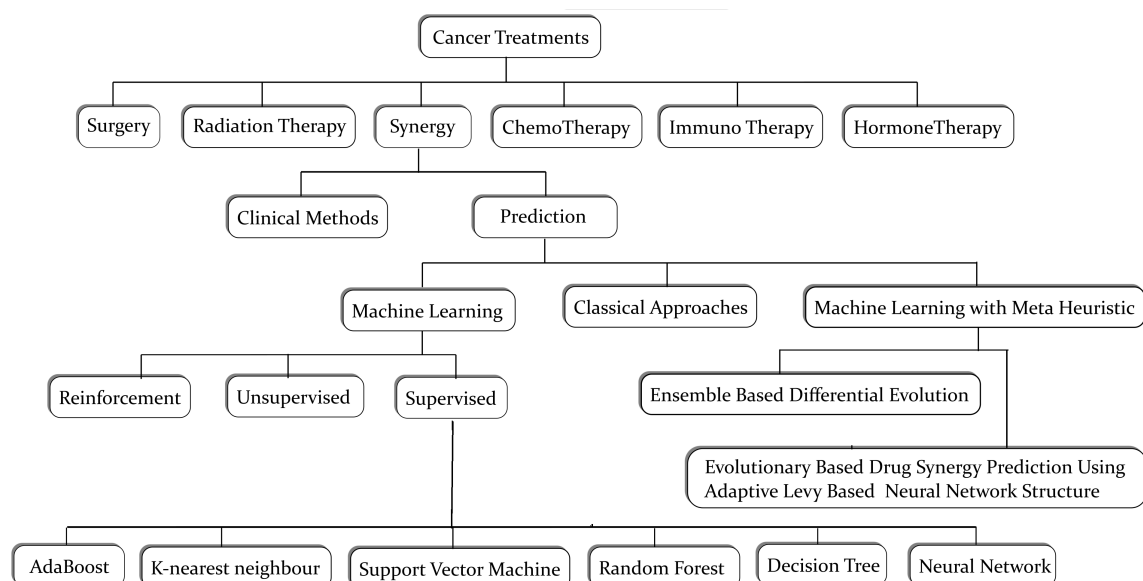


Figure 1.6: Prediction methods for drug synergy

to predict some effective drug combinations as shown in Figure 1.6. Computational prediction methods might assist in recognition of drug combination effects at molecular levels. For instance, machine learning-based drug-drug interactions (DDIs) prediction is done by integrating drug phenotypic, and combining molecular and pharmacological data, chemical, therapeutic, and genomic properties. These are promising methods for predicting different unknown drug-drug interactions. Thus, it is crucial to develop data-driven techniques to computationally identify the optimal drug combination for a patient. Deep learning and machine learning algorithms are continually being developed for the prediction of interactions.

1.4 Machine Learning

Machine Learning is a branch of artificial intelligence (AI) in which computer system possesses the ability to perform task on the basis of its learning based on training datasets/patterns and improve further from its past experience without any programming explicitly. In other words, training data is fed into the classifier i.e. machine learning algorithms and statistical models which train the classifier. New data is given as input to the trained classifier which is ready to do the prediction as shown in Figure 1.7. It provides ability to the systems to learn automatically from

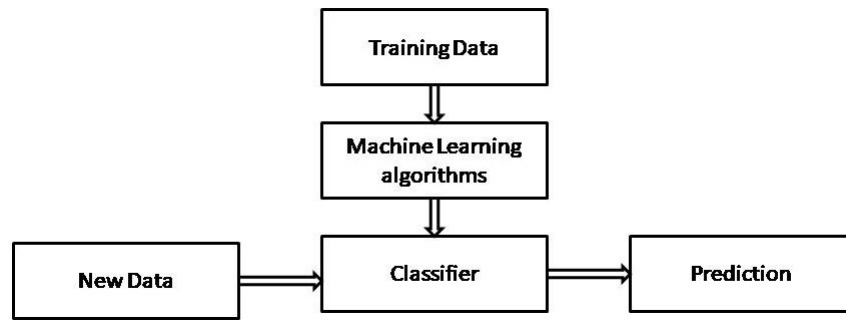


Figure 1.7: Basic structure of machine learning process

huge amount of training data.

Three major recognized classification of machine learning are supervised, unsupervised, and reinforcement learning as shown in Figure 1.8.

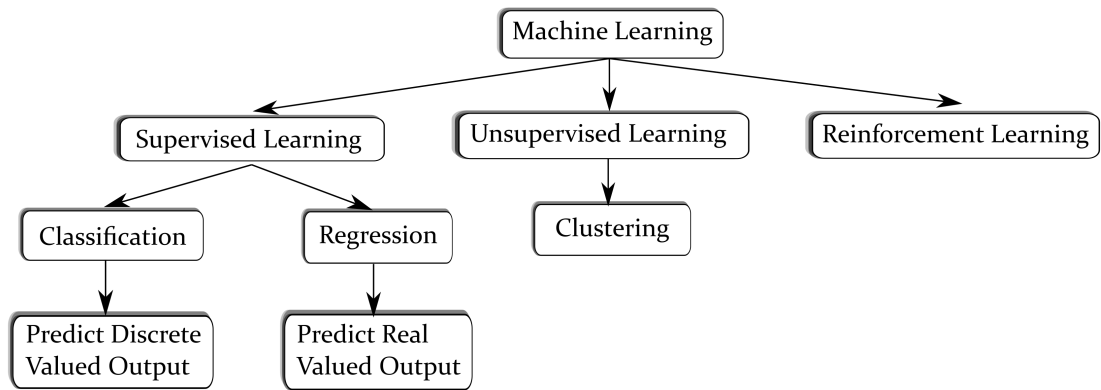


Figure 1.8: Categorization of Machine Learning Techniques

1.4.1 Supervised Learning

This is very basic type of learning in which the machine is trained using large amount of well labelled data. Here labelled represents that few data items are tagged with correct answers too. When training completes, the supervised learning algorithm analyses the new data given to it. Based on facts, it learns and predicts a correct outcome. Suppose there is a basket containing some fruits. Initially, training data of different types of fruits has been provided to the system for training purpose. If the shape of an object is round and color is red, then supervised learning indicates it as apple and if color of an object is yellow and it has curve like cylindrical shape, then signifies the object as banana. After the completion of training, when some input

is given, it will recognize its features through pattern matching. Since machine has learned things from the previous data, it will use that experience for further prediction. Classification of the object will be performed on the basis of color and shape of the given object. It also enhances its knowledge by addition of some new identified features which will be used for classification of any new object.

Supervised learning is further classified as follows:

- i. **Classification:** In this category of supervised learning, the output belongs to a particular class, for example output can be either a "dog" or a "cat" or "disease" or "no disease".
- ii. **Regression:** In this category, the real value is depicted as output, such as "price of house" or "weight".

1.4.2 Unsupervised learning

In this technique, machine is trained from the given dataset which is neither labelled nor classified i.e. it is performed in the absence of a teacher. The system/model has to collate and group the unsorted data on the basis of pattern matching, similarities and differences among each other irrespective of any training dataset. For example, consider a dataset of images (containing cats and dogs) which has not been fed into the system/model i.e. no feature of cat and dog has been fed for the training. The system gets trained itself by analyzing the patterns, differences and similarities among each other. Here machine didn't learn anything before, meaning that no training data and examples were supplied beforehand. So in this learning technique, the algorithm learns without any guidance. The following concept comes under unsupervised learning.

Clustering: It comes under unsupervised learning in which different data points are grouped into clusters based on various characteristics of data like patterns, similarities, and dissimilarities without any prior training of data.

1.4.3 Reinforcement learning

It is a branch of ML in which maximizing the reward is the main goal for any given situation. Various machines and applications have used it to find the best path behaviour. Reinforcement learning focuses on finding an optimized solution subject to maximizing some reward value. The problem related to Reinforcement Learning is as follows: Suppose there is an agent and a target having various hurdles between them along different paths. The job of agent is to find best and optimal path in order to reach the target with minimum possible hurdles.

The agent learns by trying each and every possible paths, then choosing the path which allows it to reach the target with minimum number of hurdles. Reward will be awarded for every correct step and penalty will be imparted for every wrong step. Thus, total reward shall be calculated when agent reaches its final target.

1.5 Machine Learning Approaches

There are various machine learning approaches used by a number researchers to solve a variety of problems. These problems require datasets to train the model (approach). Due to the unique nature of datasets that needed to be computed upon, there were many machine learning approaches created which performed better on different datasets individually. Additionally, this observation lead William Macready and David Wolpert to define a theorem called "No Free Lunch Theorem" in 1997 [10], stating that no single machine learning approach can be used on different datasets whilst achieving the highest efficiency. The brief details of general machine models are given below:

- i. **Fuzzy Rules Using Genetic Cooperative-Competitive Learning:** Fuzzy rules containing dissimilar number of variables can be processed with the help of Genetic Programming (GP) for obtaining accurate and compact rule bases for high dimensional problems [11]. It makes the rules cooperate and compete amongst themselves and at the same time, resulting in fuzzy set of rules that

leads to better performance [12]. Genetic programming is a type of evolutionary technique which shows the population's individuals using variable-length tree structures, thereby making it convenient for memorizing the rules [13]. In GCCL (Genetic Cooperative-Competitive Learning) technique, first of all, the entire population is codified into rules. So, it becomes crucial to propose a kind of framework which avoids the convergence of population to the same search space and keep it diverse in nature [14].

- ii. **Adaptive-Network-Based Fuzzy Inference System (ANFIS):** It is a neuro fuzzy based technique in which ensembling of fuzzy inference system and the neural network has been done. In modelling nonlinear methods, ANFIS has achieved substantial results [15]. The parameters of membership functions drawn from dataset represent the system's behaviour. They train the system from the data set and then system parameters are adjusted according to a given error criterion. ANFIS has been successfully used in identification and modelling of numerous systems. The major factor to obtain effective results with ANFIS is the selection of optimization method utilized in training [16].
- iii. **Dynamic Evolving Neural-Fuzzy Inference System (DENFIS):** It is based on fuzzy rules and inference used to propose the systematic approach from given dataset. It uses a number of highly activated fuzzy rules to dynamically compose an inference system to calculate the output values [17]. In addition to updating fuzzy rules during training period, the fuzzy rules are randomly selected from the existing fuzzy rule set. Depending on the basis of current input vector's position, these rules participate in the inference for each new input vector [18].
- iv. **AdaBoost:** AdaBoost is an ensemble technique in which several weak classifiers attempt to create a strong classifier [19]. This is done by building a model using training data. For correcting error(s) from first model, second model is created that attempts to correct the error of the first one and this process of adding

models for correction in previous one goes on until the training set is predicted perfectly.

$$\text{Weak Classifiers} \left\{ \begin{array}{l} h_1(x) \in \{-1, +1\} \\ h_2(x) \in \{-1, +1\} \\ h_3(x) \in \{-1, +1\} \\ \dots\dots\dots \\ h_T(x) \in \{-1, +1\} \end{array} \right. \quad (1.1)$$

$$\text{Strong Classifier } H_T(x) = \sum_{t=1}^T \alpha_t h_t(x) \quad (1.2)$$

where h_T is hypothesis and α_t is the weight of the classifier.

- v. **k-nearest neighbour (k-NN):** It comes under supervised machine learning algorithm which is generally used due to its less complexity in solving various regression and classification problems. Here, "k" in k-NN algorithm represents the nearest neighbors in training examples which forms input to the algorithm. This algorithm works on the assumption of feature similarity i.e. there exists close proximity in similar things. This algorithm can be used in classification problems. Majority of neighbouring votes are used for classification of an object, which is assigned to the most common class which exists amongst the k nearest neighbours. In k-NN regression problem, property value, the average of values of k nearest neighbours, emerges as the output. It is widely utilized in real-life scenarios since it is a non-parametric learning algorithm which means it does not make any assumptions about the distribution of data, and works quite well in comparison to more complex and more recent techniques.
- vi. **Decision Tree:** A decision tree [20] [21] is a tool with decision support which uses a tree-like model for representing decisions and the possible consequences of each outcome. It also includes some parameters for better decision making

like resource costs, chance event outcomes, utility and so on. The main benefit of tree-structure is that it is quite efficient in representing real life situations and thus, has influenced a wide area of machine learning approaches, covering both regression and classification problems. In decision making and analyzing, a decision tree can be very useful to represent decisions visually, explicitly and with better understanding. The decision tree is incrementally developed by breaking down a dataset into smaller subsets and the final result is a tree having decision nodes which have two or more branches and leaf nodes which exhibit a classification or decision. To calculate the homogeneity of a sample, it uses entropy. The entropy is zero if the sample is completely homogeneous and it has an entropy of one if the sample is equally divided.

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i \quad (1.3)$$

where $E(S)$ is entropy and p_i is probability of one attribute.

vii. **Random Forest:** Random Forest [22] [23] comes under supervised learning algorithm. In this, a forest of random information is created. The forest which builds as an ensemble of decision trees get trained most of the time using the bagging method. The main idea behind bagging is that learning achieved by combination of models will definitely increase the final outcome [24]. In simple words, the Random Forest is built up of multiple decision trees and it uses them by merging together to achieve more accurate and stable prediction.

It contains input vector $Y = y_1, y_2, \dots, y_q$, where Y is an input vector with q -dimensions which is used in building a forest. Inside the forest, a set of M trees such as $T_1(y), T_2(y), \dots, T_m(y)$, the output of each tree estimates the actual value $\hat{Y}_1 = \hat{T}_1(Y), \hat{T}_2(Y), \dots, \hat{T}_n(Y)$, where $n = 1, \dots, M$. The final result of it is the mean of all values which has been predicted using various trees.

$$Estimate_{RF}(Y) = \frac{1}{m} \sum_{m=1}^m y_m(Y) \quad (1.4)$$

- viii. **Support Vector Machine:** SVM is a discriminative classifier which generates the hyperplane that divides the whole data into classes. It is an algorithm that receives data as input and a line is produced as an output for separating these classes. The points close to the hyperplane from both the classes are known as support vectors. After this, margin is computed which is the distance between this hyperplane and support vectors. The main objective of the SVM is to maximize this margin and it creates a decision boundary in such a way which separates the two classes.
- ix. **Neural Network:** In machine learning, the neural networks [25] are the subsets of the algorithm constructed around the model of artificial neurons spreading across three or more layers. This machine learning technique totally relies on human neural network. It has basically two types of networks such as feed forward and back propagation. Feed forward propagation is preferred when there are multiple number of outputs and back propagation is used in case of single output.

1.6 Ensemble Machine Learning

Ensemble learning revamps the prediction performance of a learning model by combining different algorithms together. Researchers around the globe are working efficiently by applying ensemble machine learning algorithms for modeling prediction and analytics problems. Dietterich [26] reviewed various ensemble methods like adaboost, random forest, etc. It is revealed in the experiments that an ensemble works more efficiently when compared to a single learner in terms of their accuracy in classification and prediction.

1.6.1 Need of Ensemble Machine Learning

As stated by Wolpert [27], there does not exist a single algorithm available which can be applied to all the possible cases of problems. Furthermore, many researchers

strove to improve the performance of diverse models of machine learning using ensemble-classifier.

Now in a practical way, it is much better to have multiple-opinions in any decision making process than a single opinion. Ensemble learning proved to be an effective tool in the machine learning domain in which it utilizes classifiers to lay a focus on optimizing performance of base classifiers distinctly. But ensemble technique does not guarantee a sure success, it just reduces the variance and provides a better performance as compared to a single classifier solution [28]. By selecting a particular aggregation technique, for instance, voting, boosting, bagging, risk factor of inadequate results arising out of single classifier system can be reduced with the help of an ensemble classifier.

1.6.2 Techniques of Ensemble Machine Learning

The three most popular techniques for building an ensemble are bagging, boosting and stacking. These are known as "meta algorithms" and are used for combining different ML models into one efficient predictive model. They are explained below:

- i. **Bagging:** It stands for bootstrap aggregation. In this technique, an ensemble model is constructed by utilizing the random sub-samples of the training data and by combining multiple homogeneous models i.e. the models typically of the same type. The sub-samples are randomly generated and sampling with replacement from the original dataset known as bootstrap samples as shown in the Figure 1.9. The advantage of decreasing the size of training data helps in reducing the variance of the prediction. It is generally used when the base classifiers are unstable like neural network, decision tree, etc. and significant changes can emerge in the learning classifier as a resultant of minute changes in training dataset(s). Example of bagging based ensemble is random forest.
- ii. **Boosting:** In this technique, an ensemble model is constructed using continuous subsets of the training data and combining multiple homogeneous models i.e.

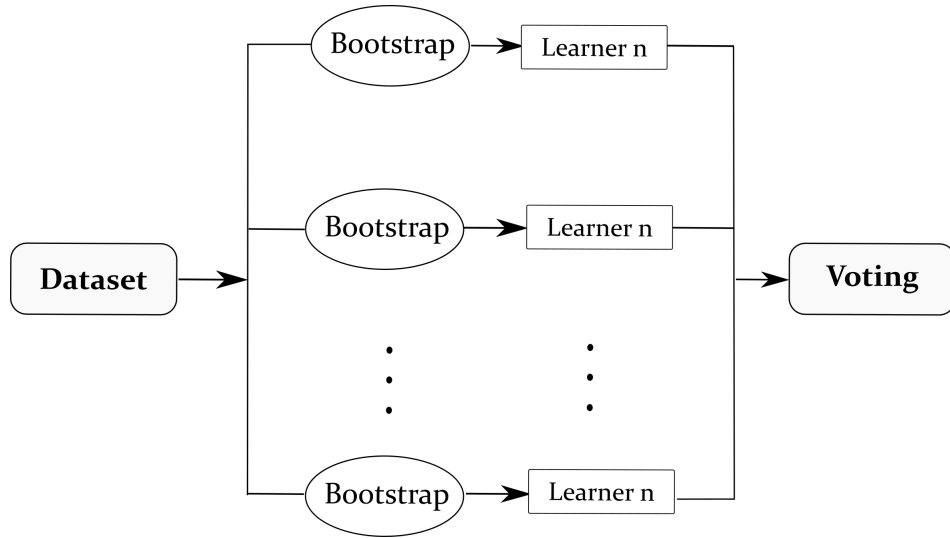


Figure 1.9: Bagging ensemble technique

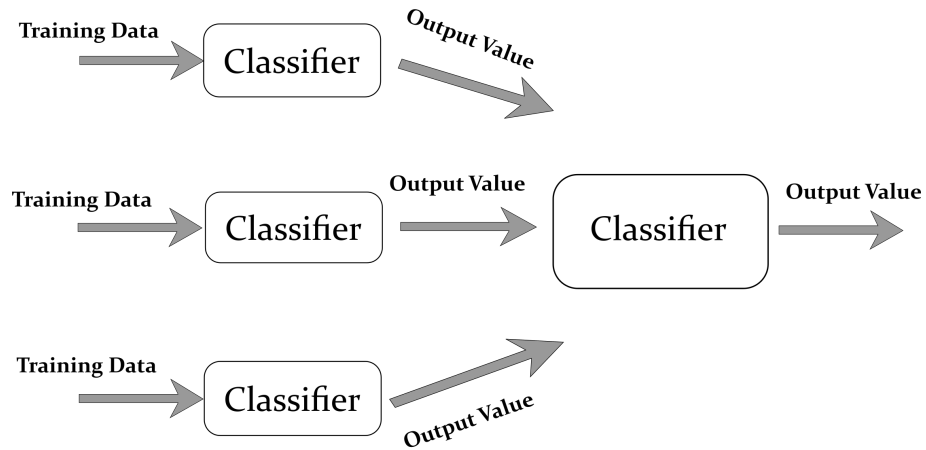


Figure 1.10: Stacking ensemble technique

the models typically of the same type. But, unlike bagging, the sub-samples generated from the training data are not random. The generation of every new sub-sample depends entirely upon the performance of the previous models i.e. it contains the records that were misclassified by the previous models. The sub-samples of the training data are utilized for the production of a series of models that yield average performance, which is boosted by clubbing them using a specific cost function like majority voting. The advantage of boosting is to improve the predictive force. Example of boosting based ensemble is Adaboost.

- iii. **Stacking:** In this technique, multiple models typically of different types are used for combination. The combining mechanism is different in the sense that

the output of the classifiers of Level (N-1) will be used as training data for another classifier of Level N to approximate the very same target function as shown in the Figure 1.10.

1.7 Metaheuristic Approaches

Metaheuristic approaches are defined as strategies that are used for providing a guide for the search process. These approaches are generally based on an iterative process to combine different concepts in an intelligent way for exploiting and exploring the search space. Major objective of metaheuristic approaches is to provide an effective exploration search space so as to provide solutions that are close to optimal values. This is attained by devising new learning strategies and structuring information such that they yield efficient and near-optimal solutions. Generally, metaheuristic based algorithms are non-deterministic and approximate. The main techniques used for such algorithms span across plain local search methods and complex learning procedures. They incorporate mechanisms to avoid approaches that are focused on limited search space [29].

The fundamentals of metaheuristics do not depend on problem, and they allow abstract-level descriptions. The key principle behind metaheuristics is to tackle complex optimization problems (COPs) [26] efficiently and to form heuristics which are controllable using upper-level strategies. These heuristics are based on domain-specific knowledge. However today's more advanced metaheuristics are based on application of memorized search experiences to guide the search and most of the existing metaheuristic approaches describe empirical results based on computational algorithms.

1.8 Metaheuristic with Machine Learning

Metaheuristic can also be considered as a search methodology which has been used for solving combinatorial optimization problems for which the algorithm is not able to generate solutions within a reasonable amount of time. Such problems are consid-

ered as computationally hard and a solution for these problems cannot be generated within a real time. Therefore combining metaheuristics with machine learning techniques has emerged as a widely used method aimed at improving effectiveness and efficiency of conventional optimization algorithms and is known as hybridization of metaheuristics with machine learning technique.

As the scope of machine learning is expanding and it is finding solutions to problems in large number of fields, researchers have initiated the process of integrating metaheuristics and machine learning techniques further to improve the system. This hybridization can be considered to work in both ways, using metaheuristics for improving machine learning performance, and employing machine learning for enhancing metaheuristic techniques. In metaheuristics, there exists a heterogeneous family of algorithms that has been specially designed to solve COPs. These algorithms are not guaranteed to generate optimal solutions. However, they can provide near-optimal solutions in an expected computing time. Some of the more popular metaheuristics techniques are artificial immune systems (AIS), ant colony optimization (ACO), genetic algorithms (GA), iterated local search (ILS), particle swarm optimization (PSO), greedy randomized adaptive search procedure (GRASP), scatter search (SS), tabu search (TS) simulated annealing (SA), and variable neighborhood search (VNS) etc.

1.9 Performance Evaluation Parameters

A number of parameters are there for performance evaluation of ML problems depending upon whether the problem is classification or regression.

1.9.1 Classification Performance Parameters

- i. **Accuracy:** Accuracy of any model is the quality or state of being correct and it is calculated as:

$$Accuracy = \frac{T_P + T_N}{TotalData} * 100 \quad (1.5)$$

- ii. **Sensitivity (Sens):** It is also known as true positive rate or recall. It shows the proportion of actual positive values correctly identified by the classifiers and is computed as:

$$Sensitivity = \frac{T_P}{T_P + F_N} \quad (1.6)$$

- iii. **Specificity:** It is the measure of accurately identified negative results by the classifier. It is calculated as:

$$Specificity = \frac{T_N}{T_N + F_P} \quad (1.7)$$

- iv. **Spec.Sens95:** Spec.Sens95 is the calculation of specificity when sensitivity is fixed at 95%.

- v. **Precision:** Precision of a model is when model predicts yes and how often is it correct.

$$Precision = \frac{T_P}{T_P + F_P} \quad (1.8)$$

- vi. **True Positive Rate:** True positive rate (TPR) is also known as recall, or probability of detection in some fields. It calculates the proportion of correctly predicted actual positives. It is computed as:

$$TPR = \frac{T_P}{T_P + F_N} \quad (1.9)$$

- vii. **False Positive Rate:** False positive rate is measured as the ratio between number of negative cases inaccurately predicted as positive F_p (false positives) and the total number of F_p and T_n .

$$FPR = \frac{F_P}{F_P + T_N} \quad (1.10)$$

- viii. **F Score:** F Score is the calculation of test's accuracy which is the weighted mean of the recall (R) and precision (P). The best value and the worst value of

the F score are 1 is 0 respectively.

$$F = 2 * \frac{R * P}{R + P} \quad (1.11)$$

- ix. **AUC:** This parameter measures quality of the classifier. The value of AUC ranges between 0 and 1. Model scoring high AUC is an efficient model compared to other models. If the AUC value of the model is near to 1 then the quality of that model is considered better.
- x. **Youden:** It measures the likelihood of positive results when tested on subjects who possess a condition, against those who do not. It clubs sensitivity and specificity to form a single measure, which has value between 0 and 1. $Youden = 1$ indicates perfect test.

$$Youden = Sensitivity + Specificity - 1 \quad (1.12)$$

- xi. **H Measure:** Classification performances [30] are reported by the hmeasure package which is the most popular measure. It also includes the H and AUC measure, for one or more binary classifiers.
- xii. **Gini Coefficient:** Gini coefficient measures the inequality in the distribution. The value ranges between 0 and 1.
- xiii. **Kolmogorov-Smirnoff:** It is a nonparametric test to measure of the degree of separation between two distributions or to check if two datasets differ significantly. Best thing about this test is that it does not make any assumption about the data distribution.
- xiv. **Error Rate:** The most widely used scalar summary in machine learning and statistics is Error Rate (ER). It is total misclassification count divided by the total count.

$$ER = \frac{F_N + F_P}{F_N + F_P + T_N + T_P} \quad (1.13)$$

xv. **Minimum Error Rate:** Minimum Error Rate (MER) is an alternate to Error rate. It is a thresholding methodology focused on minimizing the error probability instead of minimization of the miss probability for a Constant False Alarm Rate.

xvi. **Minimum cost-Weighted Error Rate:** Minimum Error Rate (MER) can be generalized into Minimum cost-Weighted Error Rate (MWL) by choosing the threshold that minimizes the total cost.

T_N :True negative, F_P :False positive, T_P :True positive, F_N :False negative

1.9.2 Regression Performance Parameters

i. **Root Mean Square Error (RMSE):** RMSE is utilized as a standard statistical parameter for measuring error rate of a regression model. The RMSE basically represents the square root of the differences between observed values and predicted values with below formula being used for calculation :

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (\alpha_t - a_t)^2}{n}} \quad (1.14)$$

where, α is predicted target, a specifies actual target and n signifies the total number of instances.

ii. **Coefficient of Correlation (r):** Correlation describes the measure to show how strong a relation is there between two variables or between actual and predicted values. Mathematically it is defined as follows:

$$r = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2 (b_i - \bar{b})^2}} \quad (1.15)$$

where n specifies the number of instances, a signifies the actual value, $\bar{a}$ stands for mean of the all actual values, b is the predicted value, $\bar{b}$ is the mean of the all predicted values. Correlation lies in $[-1,1]$ and its value towards 1 or -1 considered to be highly correlated.

- iii. **Accuracy:** Accuracy represents the closeness of a measured value to a standard value. It represents the deviation of predicted values from the actual values in percentage with acceptable error. Accuracy can be computed as follows:

$$Accuracy = \frac{\sum_{i=1}^N d_i}{N} \times 100 \quad (1.16)$$

Here, N represents total number of records in the predicted data. A binary difference (d_i) can be computed as follows:

$$d_i = \begin{cases} 1 & \text{if } |a_i - p_i| \leq A_e \\ 0 & \text{Otherwise} \end{cases} \quad (1.17)$$

Here, a_i is the actual target, p_i is predicted target and A_e is the acceptable error. Range of accuracy is 0 to 100. The technique with accuracy approaches to 100 is more efficient compared with other techniques.

1.10 Motivation

Most of the classical chemo drugs destroy rapidly-growing cells. As cancer cells have a fast-growing pace, these drugs are effective against them. As a side-effect, these can damage other rapidly-sharing body cells too. In the entire discovery process, estimation of drug amalgamation consequence largely depends upon examination. However, testing complete in-vitro cancer cell set is not feasible due to significant search space combinations. Computation-based estimation can be helpful for the recognition of probable drug mixture effects at molecular levels. e.g. hybridization of pharmacological and molecular data. Regression-based estimates of DDIs are

obtained with combinations of drug phenotypic, therapeutic, genomic and chemical properties and large-scale pharmacogenomic cancer cell lines screens. Such methods are effective for estimating DDIs that are not known. Hence, computational approaches for guiding readily testable candidates and reliably predicting effects of drug groupings are expected.

Predicting efficient drug synergy is very critical in the medical field. The utilization of machine learning techniques prove to be of high importance while decisions undertaken by medical practitioners. Now, due to an unfortunate rise in the rate of cancer cases, a significant amount of data is also produced for further clinical and medical research, which is very important for application of machine learning and data science in the medical domain.

Drug synergy score is the main factor to examine different drug-drug interaction for various drug combinations. However, drug synergy score prediction is an ill-posed problem. It's value is crucial for inhibiting specific cancer agents in various medical domains. In order to minimize the drug synergy prediction errors, the use of efficient regression-based machine learning technique will be a great solution. Therefore, the main motivation behind this research work is to design novel computational machine learning techniques for drug synergy prediction for cancer patients.

1.11 Thesis Organization

This thesis is devoted to design computational machine learning techniques to predict drug synergy for cancer. The complete chapter-wise summary of thesis work is given as follows:

Chapter 1: Introduction

In this Chapter, various types of cancer treatments are discussed. the motivation behind this thesis is presented by introducing the concept of drug synergy score prediction using machine learning techniques. Different machine learning techniques that can be utilized for drug synergy prediction are discussed. The performance

measures related to regression based machine learning techniques are also described.

Chapter 2: Literature review

In Chapter 2, a comprehensive review of the existing machine learning techniques is discussed. The drug synergy score prediction is generally a regression problem. Therefore, in this chapter, various regression based machine learning models have been compared with each other. Thereafter, the research gaps are also formulated. Finally, the objectives of this thesis are discussed.

Chapter 3: Ensemble based Differential Evolution

In this Chapter, an effective machine learning technique is defined to predict drug synergy, by utilizing ensemble based differential evolution in order to tune attributes of the SVM. It regulates the precision in prediction. The ensemble based differential evolution employs two trial vector generation techniques. Proposed and existing machine learning processes have been applied to drug synergy data. A rigorous analysis points out that the proposed technique outperforms others in terms of accuracy, root mean square error, and coefficient of correlation.

Chapter 4: Neuro-fuzzy based Ensembling Technique

In Chapter 4, a neuro-fuzzy based ensembling technique is designed. Nine popular machine learning methods have been used by utilizing drug synergy data to compute the drug synergy score. Out of these, four with the highest accuracy have been selected for developing ensemble-based machine learning model. These models are Random forest, Fuzzy Rules Using Genetic Cooperative Competitive Learning method (GFS.GCCL), Adaptive-Network-Based Fuzzy Inference System (ANFIS) and Dynamic Evolving Neural-Fuzzy Inference System method (DENFIS). Biased-weighted aggregation(addition of more weights to model with higher prediction score) is evaluated to achieve ensembling. Drug synergy score data has been used for evaluating existing and proposed machine learning models. Analysis performed out of a comparison reveals a better performance of the proposed method over the others when it comes to accuracy, RMSE and correlation coefficient.

Chapter 5: Adaptive Lévy based Neural Network Scheme

In Chapter 5, tuning of machine learning techniques has been explained, which outlines improvement of prediction score of the drug synergy data. To overcome this issue, an efficient and robust Evolutionary based Neural Network Structure using Adaptive Levy technique has been proposed to overcome the parameter tuning issue. The existing machine learning techniques and proposed evolutionary based technique are tested on drug synergy score data. An extensive analysis shows the better performance of proposed method over existing techniques, in terms of accuracy, coefficient of correlation and root mean squared error.

Chapter 6: Conclusions and future directions

This chapter gives the concluding remarks on the research by highlighting the significant contributions of the work done. The chapter concludes the thesis by explaining the outcome of each experiment. Future directions of the implemented research work have also been presented towards the end of this chapter.

Chapter 2

Literature Review

Over the past decades, enormous research has been done in the field of medical science. The interest in medication combinations is increasing rapidly. It is because of two reasons. First, medication combinations are more likely to boost the therapeutic impact and second, to decrease the frequency of undesirable side effects when solitary drugs don't deliver satisfactory outcomes. Drug combinations can display synergistic, additive or antagonistic effects as shown in Figure 2.1.

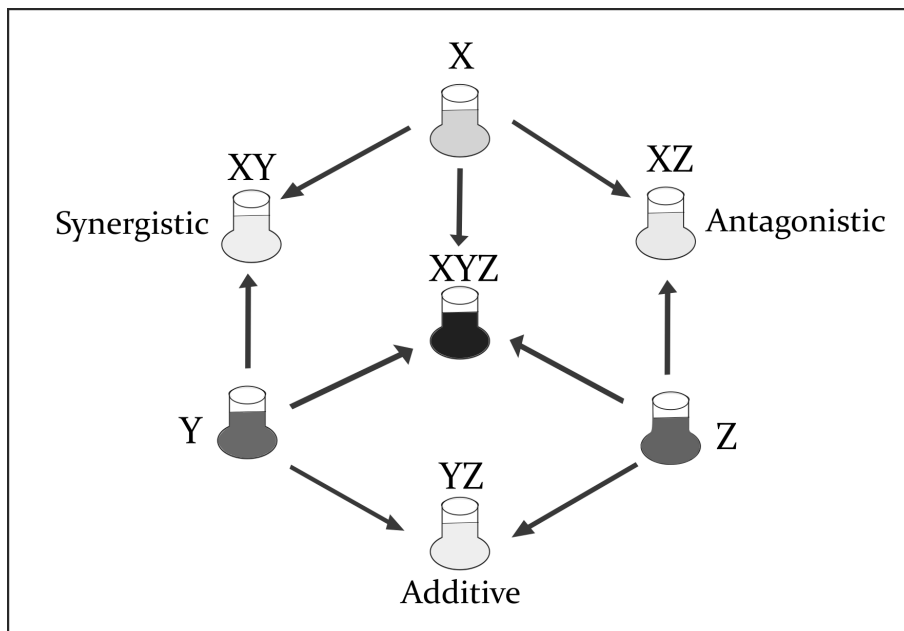


Figure 2.1: Synergistic/antagonistic effects of different drug combinations [31]

Drug combinations that are synergistic have a rational design, but there remains

a challenge which consists of heavy computational efforts in experimentation. Substantial effort in analyzing the benefits of this joint action of medication has been matched with the growth of applicable statistical procedures. It further generates the resources for statistical evaluation of the information obtained in this research.

Some of the most dreadful diseases where drug synergy has been useful are symbolized in Figure 2.2.

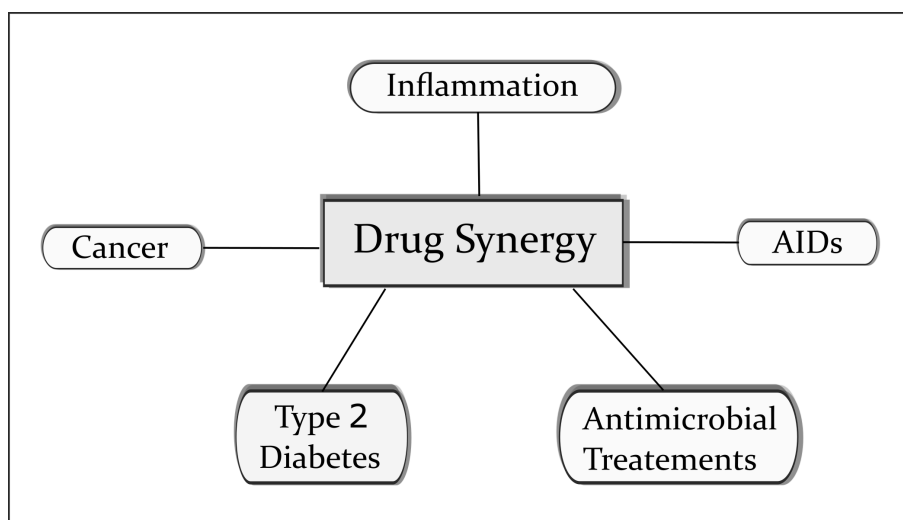


Figure 2.2: Targeted diseases using drug combinations

Researchers collect data sets from various available data repositories concerning a particular disease prevalent in a specific region of interest. Then, Machine Learning (ML) techniques are applied, which learn from the existing available data set. The applicability of these techniques increases the predictability of a particular disease. The following related work explores the various ML techniques applied to predict the best combination of drugs for multiple diseases.

Synergy in drugs, the joint increase in medication efficacy, is a main aim of combinatorial medication development process. Synergy in drug combinations plays crucial role to demonstrate the positive results. Drug's antagonistic behaviour, by comparison, cannot be desirable but will be useful in picking against drug contrary mutations. Experimentally, combinational extensive throughput screening has been invented to hunt for drug-drug synergy pairs in many systems. Determination of

drug synergism requires a lot of computational resources to find out the optimal combination. Li et al. [32] utilized an abstract network distribution-based strategy to predict drug synergism. According to the different combinations that drugs make with each other, they formulated a research to rate and choose potential antipsychotic medication pairs. A compound genome-related strategy was formed by Jansen et al. [33] to discover anti-fungal synergistic effects. These effects were based on the approach that medicines with same chemogenomic profiles are bound to show synergy. They demonstrated that a support vector machine(SVM) classifier recovers and asserts some synergistic combinations. Considering the connection among medication synergism and interactivity in genetics, Cokol et al. [34] showed that such receptor pairs making adverse genetic interactions might lead to the possibility of antipsychotic medication target.

Computational intelligence techniques offer the option to go through the big synergistic space effectively. The performance of the prediction model can be increased by available synergism data.

Cancer is a life-threatening dominant disorder that comes into existence by a complicated interplay of genetic aspects. Cancer drug formulators and developers are rewarded and doomed with plenty of drug targets from the tumor cells and the surrounding components. This reward of targets has motivated the fast gain for the quantity of experimentally accessible little molecules, zoological, and cell-based therapies like solid and blood-based altercations. The hardest questions in cancer therapeutics, especially for a small group of atoms, is what would be the best drug combinations exactly? The impact of combination drugs (mix) on many targets at once is good at restraining the multiplex system of diseases in contrast to single drug. Several studies show that combination drug therapy prevents cell multiplication to a great extent in comparison with a single drug in cancer.

Cancer treatment using combination of drugs is an important method to increase the rates of medication reaction. Results of combination treatments are generally interpreted as proof that human tumors react more firmly to a mix than the usual

monotherapy. Mixtures are usually developed according to pre-clinical observations of medication additivity or synergy. Independent drug actions, without invoking synergistic or additive consequences, provide a sufficient explanation for the high quality of several clinically accepted combination cancer treatments. These actions also affect how trial information is to be interpreted. Even if the medication mixture exhibits combination or synergism in clinical versions, patient dependent variations frequently make an independent activity. It is the most effective method in clinical trials [35].

Computational intelligence procedures based on the screened information of DDI have been able to provide the cost-effective strategy. DeepSynergy [36] is a deep learning based technique for the prediction of drug synergy. It outperformed many other available techniques. It is discussed in detail in Section 2.1.5.

The Chou-Talalay process relies upon method derived from the principle of mass-action law, which yields relationship between individual items and numerous objects in a highly correlated order. The method comprehends the biochemistry and biophysics namely, Henderson-Hasselbalch, Michaelis-Menten, Hill and Scatchard methods. Hence, it leads to a new theorem of the Chou-Talalay process, called as combination index (CI). It qualitatively defines in which $CI < 1$ indicated as synergistic effect, $CI > 1$ as antagonistic effect, and $CI = 1$ as additive effect in medicine permutations. This theory also provides an algorithm for various existing computational methods that predict synergistic or antagonistic effects in any given impact and level of doses [37] [38].

2.1 Drug Synergy Methods

There are several methodologies available for anti-cancer synergistic models. They are described in Table 2.1:

Table 2.1: Various approaches in anti cancer synergy prediction

Year	Method Name	Author	Summary	Ref
2017	Machine learning approaches	Richard et al.	Deep Learning has impacted many research areas by giving models having very high performance. They have presented Deep learning-based drug synergy prediction called DeepSynergy.	[36]
2016	Mixed-integer linear programming methods	Poos et al.	Integrated Integer Mixed Linear Programming simulations to a relative ML-based way to determine regulatory connections describe the discrepancy of telomerase levels of transcript in the best possible way in mutations(yeast) with squeezed regulators demonstrating divergent telomere length, in contrast to the ones having normal telomere length.	[39]
2015	Computational methods	Goswami et al.	Computational methods for computation relying upon gene expression profiles following therapy with one drug and response curve of dose provided. A signature with a drug-drug activity scoring algorithm is created to account for antagonism or the synergism between drugs. Various strategies that contained a complete set of genes were considered for selecting core genes. Gene collection was affected by medication whereas, drug-resistant gene set and known gene set for drug targets were affected by the drug.	[2]
2015	Systems biology methods	Ryall et al.	It engages cutting-edge system approaches, such as large-scale modeling of cell marking networks, network theme manipulation, statistical models that have been based on association, identifying similarities in the signature of genes, functional genomics, and high-output blend displays.	[40]
2010	Kinetic Methods	Kosalaram et al.	In order to predict tumor growth rate in mice, novel PK/PD versions have been proposed. Synergistic changes are mimicked in the rate of tumor growth using anticancer combination treatments for tumor growth inhibition. Both models predict under assumptions and constraints.	[41]

2.1.1 Kinetic Methods

Pharmacokinetic/pharmacodynamic (PK/PD) methods are proven to be helpful in predicting tumor growth levels in mouse xenografts. PK indicates the circulation of the drug throughout the body, whereas PD indicates the effect of drugs on the body in biological terms. Published PK/PD versions are implemented to the anti-cancer combination treatments to mimic synergistic changes in the rate of tumor growth inhibition. The parameters in the PK/PD version were further utilized to estimate clinical doses of the combined drugs [42].

2.1.2 System Biology Methods

Existing cancer therapies mostly depend upon the outdated approach of tackling one disease by only one drug. Permutations of drugs may conquer resistance by restricting the cell mutations and the creation of pathways to escape. But due to

large pool of potential drug combinations, some approaches are required to constrain the research region which can further lead to experimental prioritization of different drugs. Here, researchers focused upon the use of computation based models, bioinformatics and high-throughput for the experimental processes to find a new combination of drugs that can be used. In addition to that, they also have a list of resources that are available to the public and information to aid in the discovery of new drug-drug effects of the combination. These approaches integrate better and robust techniques to discover drug-drug interactions [40].

2.1.3 Computational Methods

Gene expression information before and after therapy with a single medication as well as the dosage responses for the IC₂₀ statistics are used to predict the interaction between drugs and how they impact Diffuse Large B Cell Lymphoma (DLBCL) cancer cells. A novel algorithm was created to consider the synergism and antagonism between different drug pairs. To identify the responsible genes, different schemes are available which consist of whole gene set, medication affected receptor sets, difference of medication affected and unaffected receptor groups, and target receptor groups. The result shows that the consistency among predicted and observed drug activity achieves a probability concordance indicator of 0.605. The synergy score of five drug activities based upon DDIs is available in GEO database [2].

In Lehar et al. [43], a combination of drugs is considered as a promising technique to minimize the undesirable off-target impacts that limit the utility of numerous potential medications. Large number of multi-dose experiments of bacterial metabolism shows that antipsychotic drug interactions are usually more relevant to particular regions of cells than the individual drug action. Synergy Index (SI) is defined as the effect of combined drugs with respect to the effect of individual drugs. The SI of unfiltered interactions were similar to the individual agents. They have a positive skew due to the effective combination. Authors took into consideration 13 sets of drug data comprising dose matrix that had 94,110 values, related to around

six different diseases. Synergism (S) was calculated by

$$S = \ln f(X) * \ln f(Y) \quad (2.1)$$

where $f(X)$ and $f(Y)$ are drug dilution factors. Medicines have serially diluted, and most likely diluted pairs of DDIs were examined in vitro [43].

In Cheng et al. [44], authors forecast interactions for new drug formulation and applications in clinical trials. They introduced an Heterogeneous Network Assisted Interactions (HNAI) frame to forecast interactions between drug pairs by using the drug observables, curative, the defined structure, and genomic similarities. The SVM variation had an area under the curve of 0.67 based on five-fold cross-validation. They gathered DDI information from the DrugBank database. According to the observations, at least 16,453 drug interaction pairs exist for both experimental and approved medication drugs. Hence, these interactions were divided into three specific types: pharmacokinetic, pharmacodynamic, and pharmaceutical.

The area under the curve of the HNAI variants is also 0.67, as calculated with five-fold cross-validation. Several HNAI utilizing antipsychotic drugs predicted some DDIs which include gain in weight and inhibition using drugs cytochrome P450 [44].

2.1.4 Mixed Integer Linear Programming Methods

Identifying mechanisms to maintain telomere length is significant in cancer biology. Its dysregulation comes as a hallmark that signifies the immortality of cancer cells. It is difficult to balance and control the regulation for the transcription of genes of telomerase. Poos et al. [39] integrated Linear Programming for mixed integers types and comparative ML-based system to determine regulatory relations. These relations regulate and significantly explain the variations present in telomerase transcript levels and mutations in yeast. Deleted regulators depict irregular telomere intervals against the ones possessing regular telomere length. They found unique regulators of telomerase expression and also formulated an easy to use learn-

ing system with the help of R programming. Its applications are such types of problems that have similar combining gene intrusion transmitting information and sample profiles for the expressions, e.g., different ailments or remedies.

2.1.5 Machine Learning-based Methods

Medication therapies involving drug combination [45] have been widespread in cancer treatments. Identifying unique synergy mixtures is very difficult because of large combination space. But with the advent of new techniques, new processes have been developed. Recently, Deep Learning has been used in many research areas efficiently. Deep Learning techniques have been continuously applied for drug admissions prediction. Researchers present a technique termed DeepSynergy. This new technique uses chemical and genomic data as the input for the model; it tackles data heterogeneity by using a normalization layer and conical layers to find drug synergies [36].

Gustavo Stolovitzky and Andrea Califano [46] in 2006 started the DREAM project. The primary objective of the DREAM was to supply a consistent and dependable infrastructure to take into consideration the latest computational approaches in biology. In the Critical Assessment of Structure Competition (CASP) contest [2], an arrangement of amino acids is provided. The challenge is that how can we forecast the 3D structure of the protein. In DREAM7-4 subchallenge, the challenge is to predict the combined action of different chemicals in the DLBCL along with OCILY3 cellular line. All participants were given the dataset of gene expression after dividing treatments with 14 distinct medications at different periods. Dose-response curves for 9 of the total 14 for the medications at periods of 24 hours has been taken. Prediction of group-wise drug interactions arranged in the antagonistic to the synergistic was the primary objective. Similarly, researchers purposed a strategy for identifying the medication combinations to acquire over the undesirable off-target effects in addition to compensatory mechanisms that limit the usefulness of several prospective drugs. The results shows that synergism in drug

interactions are more particular to combined medication than single representative actions. They primarily focus on six mixtures, whose discerning synergy depends upon multi-target drug action. Further, they revealed that the notion of selectivity achieved via differential expression of distinct medication targets in cell forms are not toxic. Cell forms are related to curative measure whose effect have a confirmation for its healing significance in vivo experiments i.e. asthma in a rat [47] [48].

Because of the latest development of significant frequency output processes, In recent years, the quantity of readily accessible synergistic data pivots has improved significantly. For eg. general public databases such as Antifungal Synergistic Drug Combination Database (ASDCD) comprising antifungal mixtures and Digital Cadastral DataBase (DCDB) are available. Drug combinations with better synergy score play a very significant role in giving training advice for producing virtual predictive procedures. Recently, there has been a massive High-Throughput Screening (HTS) synergy research with over 20000 synergistic recordings, which offers the chance to rate machine-based methods for predicting novel medication combinations [49].

In Yin et al. [50], medication for cancer required frequently screening of a combinational dose on various testing platforms to acquire the best anticancer effects. To create successful computational models, they analyzed the newest strategy in hunting optimal dose blend from three viewpoints: (1) Experimental-based strategy; (2) Computational-guided experimental approach; and (3) Computational-based policy. Besides the debut of every plan, crucial discussion of merits and demerits are also included, along with reliable focusing on the present applications and future advancements. For largely experimental established techniques, the synergistic impact of Irinotecan and 5-Fluorouracil was analyzed in the rat model of colon cancer, in the dose of Maximum Tolerated Doses (MTDs), 12.5% MTDS, 50% MTDS, 75% MTDS, respectively. For computational guided experimental strategy, lots of medications with various dosages (e.g., 0, IC25, IC50, IC75) were examined to get a particular disorder; this was utilized to form a sizable searching distance. The aforementioned

three strategies outlined the most recent developments in the region of synergistic dose mixtures for anticancer drugs. Author's gathered work will pave the way to get a detailed predictive model of optimum dose combination. It needs to know about the present hunting methods continue to be confined to neighborhood optimization, whereas more experimental results are required to confirm the performance of computational models.

Zhang et al. [51], suggested a "network goal"-established technique in place of the conventional "single target"-based method for computer simulations. They created the NIMS algorithm (Network goal developed analysis of Multicomponent Synergy) to re-evaluate synergism at high frequency. Out of the five agent pairs, one of them reported synergism and was taken as benchmark for NIMS outputs. They took a note of Gene Ontology (GO) co-annotations for progressing process understanding that caused synergism among pairs called by NIMS. Results demonstrated that brokers with synergy might not target the specific same operational procedures. Their work indicates that system target-based approaches can be used effectively to target synergism among pairs and to make the combinational medication development process easy. It can be the process of identifying multicomponent synergy at high frequency.

Related work is tabulated based on Statistical/Computational models discussed in the Table 2.2:

2.1.6 Drug Synergy Metrics

A corresponding pharmacological and therapeutic excellence is desirable, to support the sturdy and effective herbal medication, which is part of the conventional cure. Berenbaum's isobole assesses and confirms the synergistic effectiveness. Various clinical trials are conducted and compared to conventional synthetic drugs. Lots of mono and multi-extract combinational cases are taken as case studies. These depict synergistic efficacy depending upon multi-target activity mechanics [52] [53].

Many strategies are developed to find experimental synergism with time-kill

Table 2.2: Statistical and computational models used for drug synergy

Year	Author	Focus Area	Approach	Input Data	Output	Ref
2018	Yin et al.	Searching Synergistic Dose Combinations for Anti-Cancer Drugs	Review and Analysis	Irinotecan and 5-Fluorouracil and other drugs with several doses(0, IC25, etc.)	The initial phase for a comprehensive predictive model of optimal dose combination	[50]
2017	Preuer et al.	Anti-Cancer Drug Synergy Using Deep Learning	Deep Learning	DeepSynergy utilizes chemical and genomic data as input data, a normalization approach to take into consideration the entered information heterogeneity, and conical layers to simulate medication synergies.	DeepSynergy significantly outperformed SVM, GB, RF, and Elastic Nets with an improvement of 7.2	[36]
2015	Goswami et al.	A new method was developed i.e. new drug interaction scoring algorithm for any synergistic and antagonistic effects produced as a result of drug combinations.	Principal Component Analysis, Differential Expression, and ANOVA method.	Gene expression profiles	Probability Concordance Index	[2]
2013	Cheng et al.	Predicting Drug-drug interactions (DDIs) for drug development and clinical application and co-administered medications	HNAI model	DDI data from the DrugBank database classified into three groups: pharmaceutical, pharmacokinetic, and pharmacodynamic	HNAI framework: k-nearest neighbor, logistic regression, support vector machine, decision tree and naive Bayes respectively	[44]
2010	Zhang et al.	Screening synergistic drug combinations in-network target with it's application to traditional Chinese medicine.	Network target-based Identification of Multicomponent Synergy	Forty-nine traditional Chinese medicine herbs and 12 herb-derived compounds. Two chemicals Rapamycin and 5-fluorouracil were also included	TGene Ontology (GO) Co-annotations	[51]
2009	Lehar et al.	Synergistic drug combinations are more in accordance with the cellular context, as against lone agent activities.	A dose matrix of inhibitory values is used	94,110 dose matrices divided into 13 sets of combination data	Synergy Score (S) prediction having values in range of 0 to 1 for different drug combinations.	[43]

method and checkerboard tests [54]. The new procedure known as Epsilometer test (Etest) is used for testing the multiple antimicrobials or organisms [55] [56].

In checkerboard and time-kill testing, DDIs possess similar dose-response curves and a timeline of their effects [57] [58]. In the recent past, the Etest was taken as a method for testing antifungal susceptibility, and generated considerable results with both macro and microdilution testing procedures [59] [60] [61] [62].

Additionally, Etest was successful in examining combinations of different drugs against aerobic Gram-negative bacilli and mycobacterium [61]. If the Etest method is taken as a standard for testing out antifungal mixtures, it is going to represent a

Table 2.3: Various methods for measuring effectiveness of synergism

Sn	Method Name
1	Bliss Independence
2	Berenbaum's Isobole Method
3	Checkerboard Test
4	Time-kill Test
5	Etest

very efficient testing methodology for the microbiology lab in the clinical trials [62] [63].

2.1.6.1 Bliss Independence

The Bliss independence version has been utilized to examine the different combinations of drugs during the extraction of the productive combinations [64]. This approach performs the comparison between actual score ($X(A)$) and predicted score ($X(P)$), considering the independent effect of drug combinations. Generally, the mixing phenomenon can be said to have synergy if $X(A)$ is higher than $X(P)$ [65]. But this procedure is not supported by any statistical measures since it doesn't consider the variability of the response steps and may often lead to false-positive claims [66].

2.1.6.2 Berenbaum's Isobole Method

The isobole is utilized in the measurement of protagonist drug interactions. The isobole resembles the linear plot in cartesian coordinates with x and y. Here x and y-intercepts symbolize drug effects. The nonlinearity of isobole is proved by Loewe, the creator of isobole, but disapproved by almost all other users. Isobole is helpful in discovering antagonism and synergism for drug mixtures despite the unknown nature, and its theoretical foundation. Various applications of isoboles in experimental testing has shown that dose ratio can also be significant along with synergism or antagonism. Specific results of regression methods for every drug show a smooth curve. Individual doses provide an impact of interest from the set of curves that have been derived from regression methods, which can usually be half of E_{max}

(Drug's max effect) [67].

Considering that the isobole is developed for a particular dose blend, different DDIs lays focus on selecting combinations of doses that supply a specific effect. To ascertain the dose blend, which provides the 50% of Emax effect, the doses set is administered in such a way that preserves constant dose ratio. The dose-response information resultant of isobole method is subject to curve-fitting(regression, for instance), as well as the smoothened dose-response curve. In this method, the dose blend yields a 50% of Emax effect.

Lack of interaction by the detected blend of doses has been represented by point on a isobole. Observed point being off and beneath the isobole signifies synergism, while it being over the isobole symbolizes subadditivity. It shows that synergy isn't only a characteristic of both drugs; additionally, it is determined by the dose ratio [68].

2.1.6.3 Checkboard Method

It is one of the methods that has been widely used for clinical trials and is able to give significantly accurate and reproducible results. The Time-kill method and checkerboard method are one of the few sought after techniques used for estimating synergy. However, these methods consume a lot of time and effort [69].

Microdilution checkerboard is a traditional method for a quantitative record of antibiotic synergy. As against this method, synergism provides a four-fold decrease in Minimum Inhibitory Concentration (MIC) of the two combined drugs. It leads to a Fractional inhibitory concentration of 0.5. The variation being affected by the 1-dilution variance in concluding MIC of every individual antibiotic keeps adding and might get proportionately high for the mixture wells [71]. The micro-dilute checkerboard plates can be initiated using two distinct procedures, along with five *Pseudomonas aeruginosa* strains, in addition to mixtures of four antibiotics. The synergy of two-drug combination at different concentrations of two drugs has been shown in Figure 2.3.

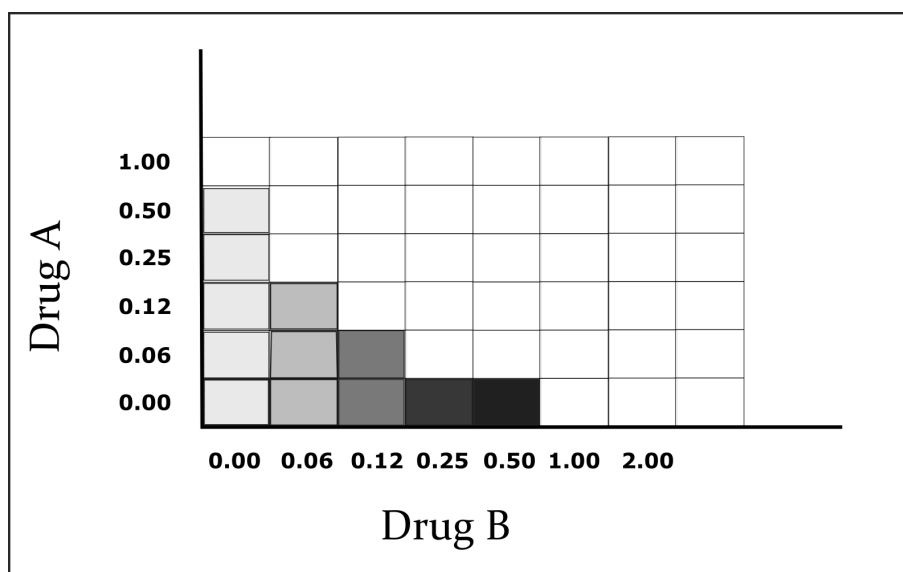


Figure 2.3: The checkerboard method depicting synergy of two-drug combination [70]

A novel procedure to dilute the MIC at a micro level is used in which disks comprising proper levels of antibiotics along with a proprietary colorimetric development indicator are dispersed into microdilution wells [72]. This process had been utilized in the preparation of panels of micro-diluted antibiotic synergy at Alamar Biosciences [73].

2.1.6.4 Time-Kill Method

Due to lack of standardized Clinical Laboratory Standards Institute (CLSI) procedures for anti-inflammatory time-kill tests, anti-fungal time-kill research can be accomplished using methods which were described and evaluated by Klepser et al. [74]. Every isolate being considered was tested alongside fluconazole and plymyxin B. In that interaction, time-kill data utilized MIC measurements which had been recorded using broth microdilution. A tube of water(sterile using expansion result of 24-hour Sabouraud dextrose agar plate which had been adapted as per turbidity of 0.5 McFarland standard) was used to form the starting inoculum. Aliquot addition was done to every tube with RPMI 1640 medium with MOPS, following initial mixing, expecting a resultant inoculum of approximately 105 CFU/ml [75].

A sample not containing medication was selected as a control for each isolate [76].

A control lesser than zero implied growth without drugs and organisms. It was also considered during all days of testing. The incubation temperature for tubes was 35°C . They were then mixed and samples for measuring colony counts were removed. Samples were collected at fixed time intervals. A serial dilution in saline was performed on requirement, which were further plated to Sabourand agar plates with the help of spiral platters. While being placed at 35°C , the plates were counted with the help of QCount counter, which was capable of discovering counts like 20CFU/ml, or below [77]. Mean count of colony(CFU/ml) from the other available specimens was used to determine synergism. Antifungal carryover was determined thereafter as per the Klepser et al. method. These samples showed no signs of antifungal carryover for fluconazole at MIC [78].

Minimal bactericidal concentration derives Time-kill assays (TKA). It provides concentration of antimicrobes which is supposed to kill 99.9% of inoculum after a complete day being into vulnerable period. In addition, this is the basic principle employed, and its usefulness is restricted because it depends on some random (99.9%) definition of total killing. There has been a doubt on the usefulness of MCBTs due to their highly specific way of testing and also the constraints that come up with controlling factors [71]. TKAs works in this method of MCBT. But instead of considering cidal action at a specific moment, the action is evaluated in order of a two day incubation period. As compared to checker board method, which considers killing at a predetermined moment and determines the best dose, TKAs fix killing speed that could be a significant metric to judge the outcome of patient [73].

2.1.6.5 Etest Method

Traditionally, one drug testing together with the Etest depends upon antimicrobial constant concentration slope. It is diffused in the affected strip towards solid agar. Strips used for Etest are placed on agar's medium, which had been inoculated using a yard belonging to the organism being tested. Etest is left overnight. In this stage, zero growth zone rolling in which one of the strips is visible as the MIC. This process

has two versions to get synergism evaluated [71].

For the first procedure, two strips were used for the Etest were placed perpendicular to one another. Strips comprising antimicrobials intersect in the MIC for each antimicrobial when analyzed independently [76]. In checkerboard procedure, the translation of Etest synergistic effects are dependent on the FIC (Fractional Inhibitory Concentration) calculation. Another method going about Etest synergism puts the first strip of the Etest on agar, consisting of drug number one, which was mixed with a sample of test organisms. It is then left about an hour before elimination. Another sample is subsequently put in the same position. It signifies the synergism [77].

Controlled Etests for specific medication are set in the same position so that they don't intermingle with the synergism test. Applying this technique, synergism is explained as a factor of dilution reduction in MIC [79].

2.1.7 Surgery Based Methods

Signature/Surgery based methods are tabulated below that list the various methodologies that have been studied for drug synergy in Table 2.4.

2.1.8 Drug synergy effects in Antimicrobial Treatments

Drug pairs may interact synergistically or antagonistically, producing inhibitory effects greater or smaller than anticipated from the medication's human potencies. Clinical approaches often prefer synergistic interactions since they maximize the pace at which the disease is eliminated from a person. But it's uncertain how such interactions affect the growth of multi-drug immunity. The speedy proliferation of antibiotic-resistant pathogens has resisted the use of medication combinations to keep clinical efficacy and fight the growth of resistance. The mathematical model of in vivo disease dynamics is utilized to ascertain the best treatment approach for preventing the growth of multi-drug immunity [86].

Synergy has two contradictory effects: it clears the disease faster and thus reduces

Table 2.4: Summary of cancer signature and surgery methods

Year	Author	Focus Area	Input Data	Output
2018	Pemovska et al. [80]	Advances in combinatorial drug screening and synergy scoring.	Combinations of drugs that were successful, as well as models in data analysis were used.	Author discussed the latest developments remains the same by including advancement in technology in cellular model systems, screening strategies, and new drug classes.
2018	Stathias et al. [81]	The glioblastoma synergistic combinations are identified by integrating drug and disease signatures.	SynergySeq platform to identify drug combinations for the treatment of glioblastoma. They used glioblastoma samples consisting of differentially expressed genes and devise a consensus gene expression signature.	They demonstrated that by using disease-specific gene expressions, they could identify preclinical combinations for glioblastoma, which are a potential for human tests.
2017	Kakushadze et al. [82]	Extraction of cancer signatures by application statistical risk models.	Author used 14 cancers samples producing 1389 whole genome.	Three new cancer signatures, including liver cancer showing dominance along with a renal cell carcinoma signature were discovered.
2016	Wong et al. [83]	Cancer therapy using nanomedicine of synergistic drug combinations	Nanomedicines of synergistic drug combinations provide superior therapeutic benefits. Here they discuss the rationale, principles, and advantages to apply nanocarriers for improving anti-cancer drug combination therapies.	They provide a better understanding of this novel cancer treatment paradigm along with primary considerations to aid the rational design of nanomedicine relating to synergistic combinations of drugs, to assist cancer therapy.
2016	Azuaje et al. [84]	Analysis of multiple types of genome-wide molecular data	Drug sensitivity cell lines prediction from samples of different patients.	Highlight major trends that will help people to select recent oncology application approaches.
2011	Wistuba et al. [85]	Molecular characteristics of a tumor based on treatment.	A comprehensive review of tumor relevant as per clinical proceedings, biomarker discovery along with biomarker imaging development was taken as input.	Proposed that next-generation clinically relevant trails which require fresh tissue biopsies for identification of appropriate biomarkers, and new statistical designs for testing personalized therapies.

the time through which resistant mutants may appear. But it raises the selective benefit of the mutants over wild-type cells. When competition for funds is feeble, the former effect is notable, and increased synergy more efficiently prevents multi-drug resistance [87]. But under conditions of persuasive resource contest, a trade-off emerges by which higher synergy increases the speed of disease clearance, but also raises the chance of multi-drug resistance. This trade-off breaks down in a vital degree of medication interaction, over which higher synergy doesn't have any impact on disease clearance, but raises the danger of multi-drug resistance. The perfect strategy for restraining multi-drug immunity isn't necessarily to optimize synergy. But, it can also result in medication antagonism in some instances. It may significantly suppress the development of multi-drug immunity [88].

For anti-inflammatory medications using paired and triple mixtures of inhibitory

agents at the practice, the evaluations are frequently done in vitro. It reveals that favourable interactions were preventing the addition of specific microorganisms. It includes lots of versions for clinical models to quantify how each mix impacts. Among the best understood and elementary types of these evaluations is the 'checkerboard' experimentation, where a 2-d selection of sequential dosage of test chemicals can be used for measurement of fractional inhibitory concentration index (FICI). It shows DDI of agents that could exert inhibitory effects clock higher compared to the independent impacts (synergy; $FICI < 1.0$), or lesser compared to the sum of individual results (antagonism; $FICI > 1.0$).

Combinations of antibiotics are widely utilized in medicine to expand the antimicrobial spectrum and create synergistic outcomes [89]. Alternately, a mix of nonantibiotic drugs with antibiotics provides a chance for sampling former untapped expanse of bioactive chemical distance. Selected medication was screened for identifying substances that strengthen the activity of the antibiotic minocycline. Unexpected synergistic drug combinations shown in vitro and in vivo activity against bacterial pathogens, such as multi-drug-resistant isolates [90].

2.1.9 Drug synergy based on Anti-tumor Agents

In Table 2.5 illustrates the effect of anti-tumour agents and methodologies on drug synergy.

2.1.10 Drug Synergy effects in Cancer Treatments

New agents have traditionally been introduced into cancer therapy by adding new drugs to accepted and established treatment regimens. The process is empiric. Though it may seem useful, but it is not frequently advantageous when it comes to potential positive molecular interactions that take place amongst drugs that are used. In fact, sequencing and drug scheduling are investigated by the majority of studies in the area of determination of optimal molecule-level interactions amongst agents [97]. A significant number of molecularly targeted drugs are being introduced

Table 2.5: Summary of anti-tumor agents and methodologies

Year	Author	Focus Area	Approach	Input Data	Output
2016	Martignetti et al. [91]	Using liquid biopsy as an important transformative precision oncology tool for endometrial cancer.	Clinical Review and Analysis	Molecule-level analysis of uterine lavage fluid from asymptomatic females, barring histopathologic evidence supporting pre-malignant lesions or cancer.	Returning patients possessing post-menopausal bleeding along with a microscopic focus on endometrial cancer. Isolation of the DNA and sequencing arising out of laser-captures microdissected tumor tissues yielded two similar PTEN mutations that had been identified previously.
2014	Jeon et al. [92]	Computational framework development that inculcates high throughput data in majority, to identify therapeutic cancer targets at genome levels.	DNA AND mRNA cell lines from Cancer Cell Line Encyclopedia, gene essentiality data, Mutation data from Cosmic Database	drug dataset, especially for ovarian, pancreatic , and breast cancer.	Drug targets using ROC(Receiver operating characteristic) and AUC(Area Under Curve).
2014	Geeleher et al. [93]	Utilized before-treatment baseline tumor gene expression data for chemotherapeutic response prediction in patients. .	Machine Learning using Fit Ridge Regression Model in R	Drug IC50 and baseline expression in cell lines and in vivo tumor gene expressions.	Performing ROC curve analysis by utilizing ROCR package of R. Using the base functions in RLinear regression, t-tests, and Spearmanâs correlation tests were performed.
2014	Guinney et al. [94]	Predictor of sensitivity to anti-EGF receptor monoclonal antibodies in colorectal cancer	PCA(Principal Component Analysis) followed by Machine Learning Regression and Classification Models	334 fresh-frozen colorectal cancer tissues provided by the Koo Foundation Sun-Yat-Sen Cancer Center.	They used a predictive model. The output of the model was a RAS Index Score having continuous value between 0 and 1.
2013	Menden et al. [95]	Machine learning model for predicting response by lines of the cancer cell for drug treatments, which have been quantified via IC50 values.	Machine Learning using Neural Networks and Random forest	Two major input streams : (1) mutation state and cancer cell line features of 77 oncogenes. (2) PaDEL software drug feature from the SMILES	The continuous IC50 value.
2013	Daemen et al. [96]	Best combination of various data types to produce optimal predictors for human breast cancer.	Machine Learning using Random Forest and SVM algorithms	84 breast cancer cell lines composed of 27 basal, 35 luminal, 7 normal-like,10 claudin-low, 3 of unknown subtype, 2 matched normal, 14 luminal, and 7 basal cell lines.	Treatment response predictions using the area under the curve.

in oncology; because of which the importance of understanding of DDIs' effects is increasing [98].

Past studies have shown that in human breast and ovarian cancer cells, cytotoxic drug cisplatin shows the interaction between rhuMAb HER2 [99]. To define the

significance of the other types of cytotoxic drugs along with the interaction between rhuMAb HER2, authors experimented with multiple drug CI isobologram analysis [100].

Synergy, when applied to DDIs, can be defined using interaction amongst multiple drugs, which achieve therapeutic impact much higher as compared to the standard inclusion of the benefits of the element drugs [101]. These interactions among medications improve clinical outcome in cancer treatment. Synergy is targeted for particular tumor cells. For seeing the effects of interactions of rhuMAb HER2 and cytotoxic chemotherapeutic drugs, they were generally used for cancer treatment of breast [102]. CI values were calculated using this technique for a large number of dose-effect levels. These are dependent on parameters which were based on median effect plots resultant of rhuMAb HER2 and cytotoxic chemotherapeutic drugs, as well as combining both at predetermined molar ratios. $CI = 1$ denotes antagonism [103].

It took 20 decades for chemotherapy to develop as a reliable option in cancer treatment [104]. This development is a result of increasing awareness regarding cancer biology, pharmacology and enhancement of clinical trial designs, apart from considerations for intricacies of the natural history related to diseases under concern [105]. With the sharpening of chemotherapeutic tools, the utilization in parallel to modalities has been increasing [106].

Graphene reveals immense potential in photothermal treatment alongside drug delivery. Tumor ablation in vivo and in vitro employs practical graphene oxide from photothermal treatment and chemotherapy combination. NGO_PEG_DOX nanoparticle capabilities for uniting region-specific chemotherapy utilising NIR(near-infrared) enhances the therapeutic effectiveness of cancer therapy significantly. When compared to chemotherapy or photothermal treatment, combo treatment reveals a synergistic effect. It leads to greater therapeutic effectiveness [107].

2.2 Machine Learning in the field of Drug Synergy

The applicability of Machine Learning techniques can increase the predictability of a particular disease [108] [109]. The following related work explores the various Machine Learning techniques applied to predict the best combination of drugs for multiple diseases [110]. Cancer is a life-threatening and highly dominant disorder that has genetic factor complexity associated with it [111]. The effect of combination drugs on different targets parallelly is more useful to control complex diseases than single drug [112] [113]. Many types of research proposed that Drug-combination therapy depicted inhibition of cell growth and population to a great extent than individual agents [114] [115].

DDIs [116] is a crucial factor in the clinical study and drug development, specifically for medications taken together. Since it is compulsory to characterize every DDIs for clinical experiments. Drug interactions majorly follow the medications that were accepted for experimentation. This is the reason for Adverse Drug Reaction(ADR) [117] [118] and a linear increase in costs [119] [120] [121] [122]. The authors suggest HNAI frame to forecast DDIs. [123]. Firstly, author constructed a Drug Interaction topology [124] that included 6946 DDI couples linking 721 approved medication based on data provided by DrugBank. Secondly, the resemblances in drug interactions are measured on four characteristics: phenotypic similarity according to a prescription, therapeutic concordance having a basis in medication of Anatomical Therapeutic Chemical Classification (ATCC) system, compound structural degradation from SMILES data, and genomic similarity according to a sizable medication target interaction network built with the DrugBank and Therapeutic Target Database. In the end, five predictive models have been implemented in the HNAI frame: NB, DT, KNN, logistic regression, and SVM, respectively [44].

Cancer is a grave and dominant disorder that is triggered as a result of the

intricate interplay of genetic aspects [125]. Drugs prepared for tackling individual molecule-level goals are incapable of combating mutagenic disease disorder [126]. Combination drugs that affect many goals at the same time are better at controlling complicated disease systems compared to single medication [127]. Throughout the drug discovery process, forecasting the medication mix effect in vitro significantly depends on the test. Larger combinational search area makes it infeasible to test all possible combinations in vitro in cancer. Prediction based on computation might help in the identification of possible combinations of drugs that affect molecular levels, like incorporating procedure by pancreatic and molecular data. A system based on learning forecast DDIs by incorporating medication phenotypic, curative, compound, and genomic properties, along with large-scale pharmacogenetic displays of cell lines of cancer. These approaches are effective in the prediction of unidentified DDIs [128].

Inferring prediction of specific cancer response in therapies is a major goal of modern oncology [129], which in turn, should result into personalised therapy.

Non homogenous cell lines for cancer have introduced multiple connections between alterations in the genome and responses to different drugs. Many different machine based approaches are suggested to anticipate response based on features related to the genome, while others have taken chemical effect of the drug into consideration to showcase their effect. With the aim of combining these complementary approaches, ML models are developed to forecast the sensitivity of cancer cell lines into medication therapy. Sensitivity is measured by IC50 values which is dependent on either the cell lines genomic features or drugs' chemical properties. Models anticipated the IC50 values in an eight-fold cross validation, in addition to a unique and unbiased blind evaluation with co-efficient of determination i.e. R^2 of 0.64 and 0.72. Additionally, models could make predictions with similar accuracy (R^2 of 0.61). IC50s of the cell lines in tissues is not utilized during practice phase. In-silico models could be made use of, in optimizing experimental design of drug-cell screenings [130] [131], with the estimation of greater proportion of lost values of IC50,

instead of experimental measurements. This further leads to providing virtual drug screening designs. A great number of medication can be probed in-silico, for testing the potential effectiveness in the form of anti-tumour agents as per the structure. Hence, it gave a computational framework for identification of new drug repositioning opportunities, in parallel with being handy for personalized medication by establishing links between genomic characteristics of patients and drug sensitivity.

Table 2.6: Limitations of statistical/computational techniques

Year	Author	Focus Area	Limitation
2011	Zhang et al. [51]	Screening synergistic drug combinations in-network target with it’s an application to traditional Chinese medicine.	Quantitative analysis of synergy is not taken into consideration. The same goes for tissue-level analysis and pattern comparison between synergism and antagonism by integrating multilayer-omic data and spatio-temporal information. This would have resulted in better performance.
2014	Cheng et al. [44]	Predicting drug-drug interactions (DDIs) for both clinical application and drug development, specifically for medications that are co-administered.	This falls short of standardised non-DDI datasets, which would ideally be a negative DDI for building the model. The interaction between non-DDI datasets could not, however, be concluded. Additionally, the accuracy of four similarities in drugs is also limited.
2015	Goswami et al. [2]	A novel drug interaction scoring algorithm was developed to support synergistic or antagonistic influences on combinations of drugs.	Their best drug interaction performance is a PC index of 0.605, which is significantly lesser than optimally predicted value 0.90.
2017	Preuer et al. [36]	Deep Learning in Anti-Cancer Drug Synergy	The dataset has inadequate distinct drug and cell lines. This creates hurdles in generalizing novel drugs and cell lines. The increase in size of datasets and advancing the algorithms could further improve Deep-Synergy model.
2018	Yin et al. [50]	Searching Synergistic Dose Combinations for Anticancer Drugs	The searching methods used are limited to local optimization, while more experimental results are needed to validate the computational models. Cancer heterogeneity should be considered in order to achieve better results.

The objective of the cancer precision medication is to provide an individual presence to therapy with the choice of therapeutics that might succeed in delivering a molecular profile of patients’ diseases. Specifically, a new pathway targeted therapeutics factor receptors may attain robust response out of cancers that harbour specific activating mutations from targeted signalling protein, or display deregulated action from the pathway that is targeted. But the prediction of clinical responses of targeted treatments in the mutational status of pathway genes is tough. There has been limited success in the prediction of responses of patients towards conventional cytotoxic treatments from molecular dimensions like gene expression levels.

Preuer et al. [36], found medication synergy for anticancer effects using Deep

Learning. DeepSynergy was implemented using publicly available synergy dataset which includes 23,062 samples. Each sample contains a cell line and two compounds. Dataset has 583 different combinations from individual 38 anticancer drugs. These combinations are examined on 39 human cancer cell lines deduced from seven different types of tissues. DeepSynergy significantly outperformed the other procedures by more than 7.2% within the region of drug research and cellular medications. So using deep learning-based technique for classification of those unique drug interactions had an Area Under the Curve(AUC) of 0.90 and mean Pearson correlation coefficient is .73. Similarly, AUC of Random forests, SVM and elastic nets is .87, .81 and .78 on the same dataset [36]. But still lot of changes are required to improve the performance of the models. Many important parameters are still not taken in to consideration in the field of drug synergy such as parameter tuning, meta heuristics, ensemble learning and overfitting. Parameter Tuning refers to choosing a set of optimal hyperparameters for a learning algorithm. These parameters control the complete learning process of our model. The performance of our algorithm is highly reliant on the type of parameters which are used to train it. Manual tuning of parameters requires a considerable time because a lot of time is wasted in evaluating unpromising areas of the search space. Whereas, hyperparameter tuning is done by automated methods whose aim is to search optimal hyperparameters in comparatively less time using an informed search without manual effort. Huang et. al. [132] developed a model for the diagnosis of diabetic patients related to GC-MS metabolomics fingerprints by using the principle of random forest. The proposed model for distinguishing diabetes patients from normal patients' has been done by using a decision support system which provides a classification among patients with the help of optimal random forests. In order to find the optimal number of trees to be used in a random forest, they performed various tests on random forest by varying the parameter of number of trees and found out which model produced optimal results by evaluating their performance manually. So, this manual technique took a lot of time to reach the optimal model. Furthermore, it was also mentioned in the

research that other metrics like classification time could also be considered in future to check if the accuracy of their model can be increased. In another study of Panda et. al., an algorithm was developed for the prediction of protein structural class [2]. They used deep recurrent neural networks to formulate the prediction algorithm. Various optimisers like adam optimisers were also used in this work and they were trained for the default parameters of the model. It was also mentioned that performance of this algorithm can be increased by considering the Atchley's factors with its optimal value into the representation framework in the future and may compare its performance by the proposed methods [133].

A metaheuristic is a set of various algorithmic techniques which can be used to define heuristic methods that can be applicable to a wide range of problems. Nowadays, metaheuristics is widely used to solve healthcare problems which need to be solved quickly by using optimal methods as compared to various search algorithms. They can also provide better results as compared to various algorithms which are currently used like rule-based algorithms and deterministic algorithms. In the study of Chun et. al., they researched on how metaheuristics can be a useful tool for tackling different problems that exist at different levels of a healthcare system. Various heuristic algorithms were applied and it found out how metaheuristic can be used to improve performance of healthcare services [134]. In an another study of Mesejo et. al., they conducted a survey on how image segmentation could be carried out using metaheuristic based models and thereafter, they concluded that metaheuristics can be very useful in configuring computer vision algorithms automatically, instead of relying on expensive and less efficient techniques [135]. In another research by Wang et. al. [136] on how to extract useful features for biomedicine industry using metaheuristics, they found that the metaheuristics plays a very important role in optimization. It also concluded that it is a very powerful tool which can be used in feature selection of low dimensional problems. In order to produce accurate results for high dimensional biomedical problems, it is necessary to choose metaheuristics carefully for feature selection.

Ensemble learning is a type of machine learning technique in which several machine learning models combined in such a manner that it can produce an optimal prediction model. Individual models suffer from the problem of local maxima and minima. Khanna et al. [137] developed an ensemble model in which multiple machine learning models were used and placed in layer structure where results of one layer were used as an input for another layer for prediction of IgA and IgG antibodies. This multilevel ensemble method used true predictions and false predictions from previous layers of the model to produce the proposed model and increase the overall accuracy of the model. It was also proved that this proposed model has increased the prediction accuracy as compared to various other existing techniques for the prediction of IgG and IgA antibodies and overcome the problem to stuck in local minima/maxima. In Onan et al., they developed a model to diagnose a breast cancer using various ML classification algorithms. After performing various tests on the model, they concluded that ensemble learning methods is a feasible method for improving the performance of classifiers used for medical diagnosis.

Overfitting refers to a model which models the training data too well as compared to testing data. Overfitting occurs when a particular model learns the noise in the dataset while in training to the extent that it impacts the performance of the model on validation data negatively. In the study by Aleksandar et al. [138], performance of a model differed among various multipliers of regularisation and also among different approaches for data partitioning. The randomly partitioned approach increased predictions for performance of model and in case of geographically structured approach, more chances of overfitting were seen. But, the regularization levels produced were twice fold higher as compared to the default technique which produce considerably lower overfitting for a model. In another study by Jenni et al. [139], they found that ANNs are more prone to overfitting since the models are highly parameterized. The performance of models can be improved by using a regularization technique. In the research it was also concluded that the performance of the model can also be increased when the individual contribution of each feature

which is included to the model is reduced using regularisation technique. This also helps to reduce chances of occurrence of over-fitting.

Precision oncology [140] is aimed at anticipating the best drug-based method for specific patients based on their tumour’s molecular structure. To measure the experimental feasibility of drug reception, several people have noted the responses for cytotoxic and particular methods for substantial cancer cell lines and given data for the anticipation models that use generic methods such as elastic net regression [141] [110]. Using already available data sets of drug responses for showing the multiple tasks learning across substances significantly affects how accurate the drug prediction models are. It utilizes trace norm regularization alongside a coherent and scalable ADMM (alternating direction method of multipliers) optimization algorithm. This leads to a prediction implying that the approach was bound to improve efforts that would exploit the increase in drug response compendia, for the advancement of personalized therapy.

Table 2.7: Limitations of anti-tumor agents and its methodologies

Year	Author	Focus Area	Limitation
2013	Daemen et al. [96]	The best combination of various data types to produce optimal predictors for human breast cancer.	Specifically developed response signatures based on profiling at multiple omic levels could lead to improved predictions.
2013	Menden et al. [95]	The machine learning model for prediction of response of cancer cell lines for treatment using drugs, which is quantified via IC50 values.	Cell lines used are not similar to real tumours. Expanded genomic and epigenomic datasets, together with better predictive models, maybe best to help the development of new therapeutic strategies tailored to individual patients.
2014	Jeon et al. [92]	Computational framework development that inculcates major high-throughput categories to identify therapeutic targets of cancer-based on genomes.	Organised mapping of tumour-specific transcriptional networks and identifying the negative genetic interactions are used for drug target identification which having low-throughput screening leads to relatively low coverage of the human genome along with their adequacy for genome-wide identification of drug target falls.
2014	Geeleher et al. [93]	Utilising the pre-treatment baseline expression data of the tumour gene for chemotherapeutic response prediction in patients.	Predictions were not as accurate as compared to optimal because of not using rigorous transcriptome quantification and further testing in prospective clinical trials.
2014	Guinney et al. [94]	Predictor of sensitivity to anti-EGF receptor monoclonal antibodies in colorectal cancer.	Mostly patients in their datasets received conventional chemotherapy plus cetuximab which has influenced their model performance and made it specific to an extent.
2016	Martignetti et al. [91]	Using liquid biopsy as a vital transformative precision oncology tool for endometrial cancer.	Time and costs emerge as primary limitations linked to uterine sonography, apart from the discomfort, risk and anaesthesia being a necessity for a room hysteroscopy.

Table 2.8: Limitations based on cancer surgery/ signature based methodologies

Year	Author	Focus Area	Limitation
2011	Wistuba et al. [85]	Molecular characteristics of a tumor based on treatment	The major bottlenecks to the work is advancement of the field and which represent challenges and opportunities for interventional oncologists participating in image-guided biopsies.
2015	Azuaje et al. [84]	Analysis of multiple types of genome-wide molecular data	Cell line selection that inadequately reflect tumor biology, and the deficient number of cell lines for modelling response in specific cancers.
2016	Wong et al. [83]	Cancer therapy using nanomedicine of synergistic drug combinations	They didn't take note of rational consideration of drug combinations, required at both micro- and macroscopic levels, which could have to enhance efficacy, improve host antitumor immunity and reduce normal tissue toxicity.
2017	Kakushadze et al. [82]	Extraction of cancer signatures by application statistical risk models	Depiction of published genome samples limits the work. Time consuming issue of meticulously ascertaining published samples.
2018	Pemovska et al. [80]	Advances in combinatorial drug screening and synergy scoring.	Hematopoietic malignancies access to tumor cells and transmitting solid tumors accompanied by cellular complexity, and three-dimensional architecture involves a lot of effort under laboratory culture conditions. Advancing screening strategies (e.g., CRISPR genome editing), developing novel drug classes (e.g., immunotherapeutics, PROTACs, etc.) and optimizing cellular animal models (e.g., organoids, PDXs, liquid biopsies) is required to improve the efficiency of the model.
2018	Stathias et al. [81]	Signatures of drug and diseases identify synergistic combinations in glioblastoma.	SynergySeq could not associate single-cell sequencing data for identification of mixtures of drug synergies which targets the cells accountable for tumor formation for provided types of GBM.

2.3 Limitations

There are many aspects of drug synergy for targeting the limitations. Various findings were drawn during a literature survey and that are categorized according to the different elements such as drug synergy based on statistical/computational techniques, Anti-Tumor Agents and Methodologies, and Cancer Surgery/ Signature-based methodologies.

Various limitations have been tabulated in Table 2.6, 2.7 and 2.8.

2.4 Comparison

The generic comparison has been drawn based on different methods, approaches and analyses such as computational approach, experimental approach, clinical review and analysis, surgery based methods, deep learning, ensemble-based machine learning approach and machine learning models. Comparison table has been tabulated in Table 2.9.

Table 2.9: Comparison of different drug synergy methods

Approach	Computational Approach	Experimental Approach	Clinical Review and Analysis	Surgery Based Methods	Deep Learning	Ensemble based Machine Learning Approach	Machine Learning models
Goswami et al. [2]	✓	✗	✗	✗	✗	✗	✗
Lehar et al. [43]	✓	✗	✗	✗	✗	✗	✗
Cheng et al. [44]	✗	✗	✗	✗	✗	✗	✓
Preuer et al. [36]	✗	✗	✗	✗	✓	✗	✗
Yin et al. [50]	✓	✓	✗	✗	✗	✗	✗
Zhang et al. [51]	✓	✗	✗	✗	✗	✗	✗
Kakushadze et al. [82]	✓	✗	✗	✓	✗	✗	✗
Wistuba et al. [85]	✗	✗	✓	✓	✗	✗	✗
Azuaje et al. [84]	✗	✗	✓	✗	✗	✗	✗
Pemovska et al. [80]	✗	✗	✓	✗	✗	✗	✗
Stathias et al. [81]	✓	✗	✓	✗	✗	✗	✗
Wong et al. [83]	✗	✗	✓	✗	✗	✗	✗
Menden et al. [95]	✗	✗	✗	✗	✗	✗	✓
Jeon et al. [92]	✓	✗	✗	✗	✗	✗	✗
Geeleher et al. [93]	✗	✗	✗	✗	✗	✗	✓
Daenen et al. [96]	✗	✗	✗	✗	✗	✗	✓
Guinney et al. [94]	✓	✗	✗	✗	✗	✗	✓
Martignetti et al. [91]	✗	✗	✓	✗	✗	✗	✗

2.5 Gaps in Literature

- i. **Parameter tuning:** The majority of existing machine learning techniques suffer from the parameter tuning issue. An efficient tuning of the machine learning techniques can improve the accuracy of existing methods.
- ii. **Ensembelling:** The majority of existing machine learning techniques cannot individually yield noticeable accuracy for drug synergy score prediction, which generally suffers from low accuracy rate as discussed above. Therefore, ensemble based machine learning techniques need to be designed to predict the drug synergy score efficiently, which leads to an optimal solution.
- iii. **Metaheuristic:** Machine Learning models used for the prediction stuck into the local minima or local maxima. To overcome this problem, metaheuristic techniques are used to escape from local minima or local maxima. But, these techniques may perform slightly better in a particular issue and may not be suitable for other sets of problems. So, metaheuristic algorithms need to be designed to address the broad domain of the problem.
- iv. **Over-fitting issue:** Existing machine learning techniques suffer from the over-fitting issue. Therefore, k-fold validation is required to overcome this issue.

2.6 Research Contributions

The main contribution of the research is made in several phases, and they are as follows: Figure 2.4 shows the contribution diagrammatically for better understanding.

- i. The AstraZeneca-Sanger Drug Combination [142] dataset is considered for experimental purposes. Data cleansing is also done through several pre-processing techniques such as removal of duplicate entries, handling of missing values, feature extraction on the obtained dataset.
- ii. An analysis of current regression-based machine learning techniques has been carried out on the drug synergy dataset.

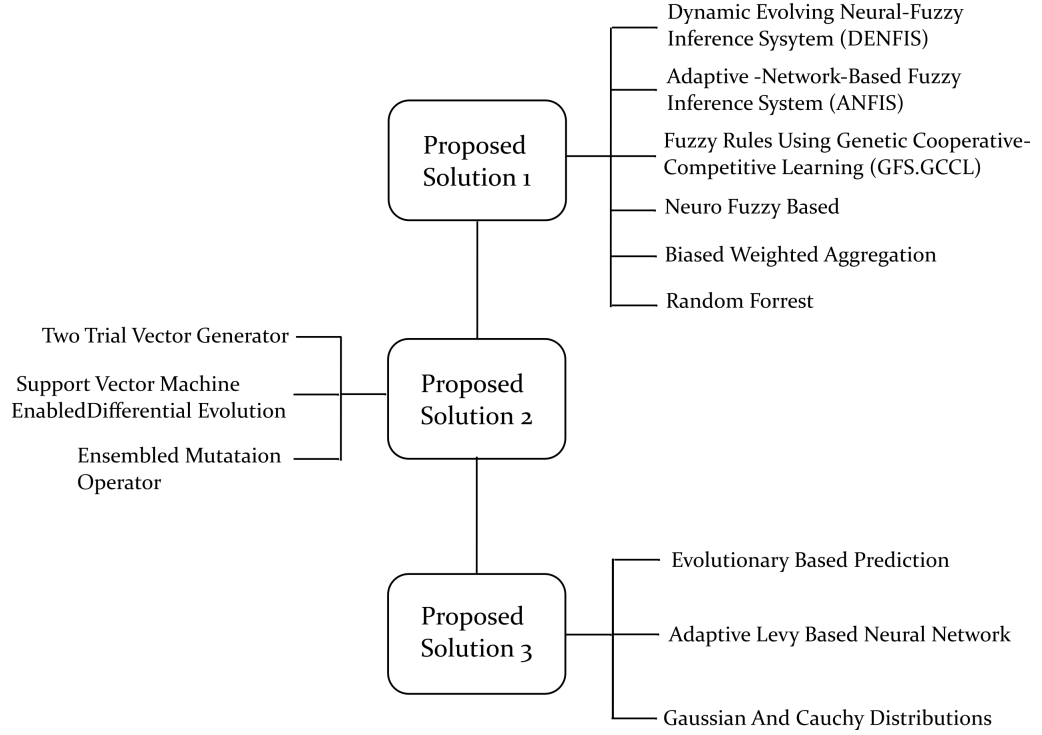


Figure 2.4: Proposed computational solutions for drug synergy

- iii. In the third phase, a productive machine learning technique to predict drug synergy is designed with the help of ensemble-based differential evolution to optimize the attributes of the support vector machine (SVM).
- iv. In this phase, Drug synergy data has been utilized for implementing nine popular machine learning techniques. Depending upon the accuracy of each of them, the ensemble-based machine learning model is developed by selecting four with the highest efficiency. These models are Random forest, Fuzzy Rules Using Genetic Cooperative-Competitive Learning method (GFS.GCCL), Adaptive-Network-Based Fuzzy Inference System (ANFIS), and Dynamic Evolving Neural-Fuzzy Inference System method (DENFIS). Biased weighted aggregation of predicted data is evaluated to achieve ensembling, which means the addition of more weights that have a high prediction score.
- v. In this phase, an efficient evolutionary-based drug synergy prediction using an Adaptive Lévy based neural network structure has been proposed and implemented to improve the prediction of drug synergy prediction. Drug synergy

score data is used to test the existing machine learning techniques and Adaptive Lévy based technique.

- vi. Finally, the obtained drug synergy dataset is used to evaluate the proposed as well as existing machine learning techniques. The comparative analysis reveals that the proposed techniques outperform others in terms of various performance metrics.

2.7 Research Objectives

The following objectives are set for the research proposal:

- i. To obtain valid and authenticated large datasets which contain the molecular profiles and drug synergy response.
- ii. To explore existing statistical and machine learning techniques for obtained datasets.
- iii. To propose computational intelligence based framework for predicting the synergy of drugs.
- iv. To test and validate the effectiveness of proposed framework using validations techniques.

Chapter 3

Ensemble based differential evolution (DE) scheme

Prediction of drug synergy score is an ill-posed problem. It plays an efficient role in the medical field for inhibiting specific cancer agents. An efficient regression-based machine learning technique has an ability to minimize the drug synergy prediction errors. In this chapter ¹, an efficient machine learning technique for drug synergy prediction technique is designed by using ensemble based differential evolution for optimizing the support vector machine (SVM). Because the tuning of the attributes of SVM kernel regulates the prediction precision. The ensemble based differential evolution employs two trial vector generation techniques and two control attributes settings. The initial generation technique has the best solution and the other is without best solution. The proposed and existing competitive machine learning techniques have been applied on drug synergy data. The extensive analysis demonstrate that the proposed technique outperforms others in terms of accuracy, root mean square error, and coefficient of correlation.

¹Harpreet Singh, Prashant Singh Rana and Urvinder Singh, "*Prediction of drug synergy score using ensemble based differential evolution*", IET Systems Biology, IET, 13(1):24-29, 2018.

3.1 Introduction

Rational design of multi-targeted drug combinations is an efficient way to handle the drug resistance issue for several complex disorders [143]. A combination of drug is generally decomposed into synergistic, antagonistic and additive based upon the deviation of the evaluated combination response from the probable effect computed using a reference technique of non-interaction [144]. The majority of standard reference techniques were designed for low-throughput drug combination analysis that make the technique expectations frequently unsuited with the complex drug interaction patterns across several dose pairs [145]. These dose pairs are computed in large-scale dose response matrix analysis [146].

To compute the interaction between different drugs, the experiential combination response is generally compared to the probable effect under the hypothesis of non-interaction estimated by a reference technique [147]. In case, the combination response is more than as expected, then the combinations is recognized as synergistic [148]. The antagonism is considered when the combination evaluated is less as compared to expected effect and if there is no change in combination response with respect to expected, then it is additive [149].

There are several recent studies on drug sensitivity and anti-cancer drug response prediction. In [150], random forest regression technique was used for obtaining a single drug or a group of drugs for a certain cancer. Zhang et al. proposes a technique with cell line similarity network (CSN) and drug similarity network (DSN) data to predict anti-cancer drug responses of a given cell line by using dual-layer integrated cell line-drug network [151]. To predict drug activity, Turki et al. proposes a novel technique that uses link prediction algorithms and compare their results with existing drug sensitivity prediction techniques [152]. Based on their experimental results, link prediction algorithms perform better than the methods used in comparison. Multi-task learning schemes have come into prominence as a consequence of their capability of exploiting inter-drug relationships during training. While multi-

task learning for the drug activity prediction was largely limited to regularized linear techniques before, recently more sophisticated techniques have emerged. An technique called Kernelized Bayesian Multi-task Learning (KBMTL) [153] relying on kernel-based dimension reduction and this multi-task learning demonstrated notable performance for drug response prediction. The study by Tan has shown that the trace-norm regularized multi-task learning followed by a kernel transform of the input gene expression data is superior to several other methods proposed for this task [154]. A similar work by Yuan et al. [155], also demonstrates the improved predictive power of multi-task learning with trace-norm regularization. In another study [156], the authors premeditates that the similar drugs should possess similar responses and define drug-similarity by Pearson correlation of drug responses. This is followed by application of a novel method called similarity-regularized matrix factorization (SRMF) which maps drugs and cell lines to a lower dimensional space and reconstruct the drug response matrix in this latent space. The usage of ensemble methods were also found promising in drug response prediction task, as in many other problems. These low-variance techniques surpassed all other schemes in a recent DREAM challenge on predicting human responses to toxic compounds [157]. Winners of both sub-tasks utilized ensemble of random forest techniques where each forest was devoted to a single cluster of cell lines. In addition to ensemble learning, biologically-plausible preprocessing techniques such as filtrating irrelevant genetic features, clustering cell lines and chemical using external datasets were found to be efficient strategies. None of these methods consider to use both multi-task and single-task methods together in an ensemble technique.

Our contributions in this chapter can be summarized as follows. First, we propose to use an ensemble based differential evolution to predict drug synergy by using SVM. Second, we apply proposed technique to test drug synergy data set. Afterwards, 10-fold cross validation is also considered to prevent over-fitting issue. The proposed technique is compared with five competitive techniques on drug synergy data by considering different performance metrics.

3.1.1 SVM for regression

Support vector machine (SVM) is extensively utilized to explore small samples under statistical learning rules [158]. The sample data set (D_s) is defined as: $D_s = (m_j, n_j) | j = 1, 2, \dots, N$, Here, $m_j \in \mathbb{R}^n$ denotes the input parameters. $n_j \in \mathbb{R}^n$ represents the output parameters.

The SVM search for one misalignment mapping from the input space to output space Ψ . By using this mapping, data Ψ is mapped to a feature space (τ) , and linear regression is carried out in τ as:

$$f(m) = [\psi \times \Psi(m)] \times \mathcal{Y} \quad (3.1)$$

$$\gamma : \mathbb{R}^m \rightarrow \tau$$

Here, $\mathcal{Y}$ is a threshold value in Eq. (3.1). The SVM determines the regression equation by using the objective function minimization as:

$$\begin{aligned} \min & \left\{ \frac{1}{2} \|\psi\|^2 + \omega \sum_{j=1}^N (\beta_j^* + \beta_j) \right\} \\ \text{s.t. } & n_j - \psi \times \gamma(m) - \mathcal{Y} \leq \chi + \\ & \psi \times \gamma(m) + \mathcal{Y} - n_j \leq \chi + \beta_j \\ & \beta_j, \beta_j^* \geq 0 \end{aligned} \quad (3.2)$$

Here, ω is used to balancing the complex items of the SVM and training error. The relaxation factors are denoted by β_j^* and β_j . χ represents insensitive loss function. β_j^* is evaluated by:

$$\beta_j^* = \begin{cases} 0, & |g(m) - n_j| \leq \chi \\ |g(m) - n_j| - \chi, & |g(m) - n_j| > \chi \end{cases} \quad (3.3)$$

By normalizing Eq. (3.2), lag range factors (i.e., a_j, a_j^*) can be evaluated. There-

fore, the regression coefficient (ψ) is computed as:

$$\psi = \sum_{j=1}^N (a_j - a_j^*) m_j \quad (3.4)$$

The SVM regression equation is obtained as:

$$g(m) = \sum_{j=1}^N (a_j - a_j^*) k_f(m_j, m) + \Upsilon \quad (3.5)$$

Here, $k_f(m_j, m)$ represents SVM kernel function. Kernel function parameter (k_f) and penalty factor (ω) evaluate the performance of SVM. k_f estimates the complexity of the solution. ω estimates the penalty to large fitting deviation. However, large value of ω may cause over-fitting. Also, too small value of ω causes under-fitting issue. Therefore, an efficient tuning of these attributes can improve the performance of SVM.

3.1.2 Conventional differential evolution

Differential evolution is a well-known metaheuristic technique. It is a simple, robust and effective global optimization technique. From the literature, differential evolution has shown better performance as compared to existing metaheuristic technique in terms of robustness and convergence speed over many real-world applications [159], [160], [161], [162].

Differential evolution utilizes mutation, recombination and selection operators to optimize given problem. Initially, randomly population is developed. Afterwards, mutation, recombination and selection operators are utilized to develop new solutions. The selection operator is then utilized to evaluate the best trial vector for next generation [163].

DE starts with a population of n_p D-dimensional candidate solutions which may be represented as:

$$\Psi_{j,\alpha}(j = 1, 2, \dots, n_p) = m_{j,\alpha}^1, m_{j,\alpha}^2, \dots, m_{j,\alpha}^D \quad (3.6)$$

Here, j represents the j^{th} solution of the population. α shows current generation. Γ represents the dimension of population.

Initial population try to cover the entire search space and constrained by minimum bound as:

$$\Psi_{lb} = m_{lb}^1, m_{lb}^2, \dots, m_{lb}^\Gamma \quad (3.7)$$

Also, initial population is constrained using maximum bound as:

$$\Psi_{ub} = m_{ub}^1, m_{ub}^2, \dots, m_{ub}^\Gamma \quad (3.8)$$

Therefore, the initial population can be represented as:

$$m_{j,0} = m_{lb} + rand(0, 1) * (m_{ub} - m_{lb}) \quad (3.9)$$

Here, $rand(0, 1) \in [0, 1]$ represents a uniformly distributed random variable [163, 164].

i. **Mutation**

After initialization, mutation operation to used for computing a trial vector as:

$$V_{j,\alpha} = v_{j,\alpha}^1, v_{j,\alpha}^2, \dots, v_{j,\alpha}^\Gamma \quad (3.10)$$

The $V_{j,\alpha}$ can be can be computed using various mutation techniques such as [162, 165]:

M_b^1 :

$$V_{j,\alpha} = \Psi_{best,\alpha} + F.(\Psi_{r_1^j,\alpha} - \Psi_{r_2^j,\alpha}) \quad (3.11)$$

M_b^2 :

$$V_{j,\alpha} = \Psi_{best,\alpha} + F.(\Psi_{r_1^j,\alpha} - \Psi_{r_2^j,\alpha}) + F.(\Psi_{r_3^j,\alpha} - \Psi_{r_4^j,\alpha}) \quad (3.12)$$

M_b^3 :

$$V_{j,\alpha} = \Psi_{j,\alpha} + F.(\Psi_{\text{best},\alpha} - \Psi_{j,\alpha}) + F.(\Psi_{r_1^j,\alpha} - \Psi_{r_2^j,\alpha}) \quad (3.13)$$

M_r^1 :

$$V_{j,\alpha} = \Psi_{r_1^j,\alpha} + F.(\Psi_{r_2^j,\alpha} - \Psi_{r_3^j,\alpha}) \quad (3.14)$$

M_r^2 :

$$V_{j,\alpha} = \Psi_{r_1^j,\alpha} + F.(\Psi_{r_2^j,\alpha} - \Psi_{r_3^j,\alpha}) + F.(\Psi_{r_4^j,\alpha} - \Psi_{r_5^j,\alpha}) \quad (3.15)$$

M_r^3 [166]:

$$v_{j,\alpha} = \Psi_{j,\alpha} + K.(\Psi_{r_1^j,\alpha} - \Psi_{r_j,\alpha}) + F.(\Psi_{r_2^j,\alpha} - \Psi_{r_3^j,\alpha}) \quad (3.16)$$

Here, $r_1^j, r_2^j, r_3^j, r_4^j, r_5^j \in [0, 1]$ represent mutually exclusive indexes [165]. F shows mutation scale factor that controls the amplification of DE [163].

ii. Recombination

After applying the mutation process, the recombination operator is utilized to maximizes the diversity of target vectors. The recombination operation is implemented as:

$$\begin{aligned} V_{j,\alpha} &= v_{j,\alpha}^1, v_{j,\alpha}^2, \dots, v_{j,\alpha}^\Gamma \\ \Psi_{j,\alpha} &= \Psi_{j,\alpha}^1, \Psi_{j,\alpha}^2, \dots, \Psi_{j,\alpha}^\Gamma \end{aligned} \quad (3.17)$$

Eq. 3.17 is utilized to develop a trial vector as:

$$v_{j,\alpha} = v_{j,\alpha}^1, v_{j,\alpha}^2, \dots, v_{j,\alpha}^\Gamma \quad (3.18)$$

Therefore, a new trial vector ($v_{j,\alpha}^j$) is evaluated as:

$$v_{j,\alpha}^j = \begin{cases} v_{j,\alpha}^j & \text{if rand}_j[0, 1] \leq \text{CR or } (j = j_{\text{rand}}) \\ m_{j,\alpha}^j & \text{others} \end{cases} \quad (3.19)$$

Here, $CR \in [0, 1]$ shows the recombination constant. $jrand \in [1, \Gamma]$ is a randomly chosen index that ensures $v_{j,\alpha}$ will differ from $\Psi_{j,\alpha}$ by at least one parameter.

iii. Selection

In case $v_{j,\alpha}^j$ exceeds the upper or lower bounds, then it is randomly re-initiated within the search range. The fitness values of all trail vectors ($v_{j,\alpha}^j$) are then computed as:

$$m_{j,\alpha+1} = \begin{cases} v_{j,\alpha} & \text{if } f(v_{j,\alpha}) \leq f(m_{j,\alpha}) \\ m_{j,\alpha} & \text{otherwise} \end{cases} \quad (3.20)$$

In case, if $v_{j,\alpha}$ provides a better fitness as compared to $m_{j,\alpha}$, then $v_{j,\alpha}$ will be replaced with $m_{j,\alpha}$ and proceeds for next generation; continue otherwise.

3.2 Proposed technique

3.2.1 Ensemble mutation operator based differential evolution

The performance of differential evolution depends upon a mutation and recombination operators. In this work, mutation operators are then divided into two types (i.e, M_r^k and M_b^k). The M_b^k is with best solution in Eqs. (3.11 - 3.13), and the M_r^k is without best solution including Eqs. (3.14 - 3.16). To balance the diversity of the population and the convergence speed, the proposed technique considers two mutation operators. The first one is obtained from M_r^k , and second one is obtained from M_b^k . Two mutation operators are randomly selected from M_r^k and M_b^k , respectively, as shown in Figure 3.1.

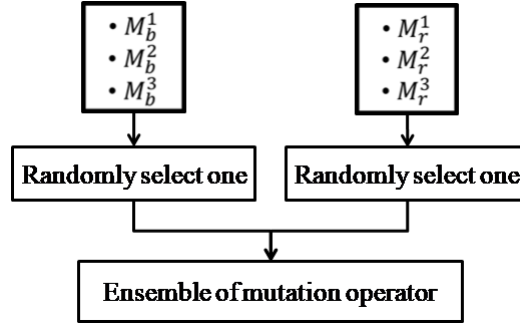


Figure 3.1: Ensemble based mutation

Algorithm 1: Ensemble based differential evolution based drug synergy prediction

Step 1: First of all, parameters initialization of proposed ensemble based differential evolution is done (i.e, population size (n_p), function evaluations (ρ), two mutation operators obtained from Eq. (3.11 - 3.13) and Eq. (3.14-3.16), respectively. Also, maximum number of ρ are represented by M_ρ .

Step 2: Initialize a counter $\alpha = 0$. Randomly initialize a population with size n_p (i.e, $\nu = \Psi_{1,\alpha}, \dots, \Psi_{n_p,\alpha}$) with $\Psi_{j,\alpha} = m_{j,\alpha}^1, \dots, m_{j,\alpha}^\Gamma, j = 1, \dots, n_p$ uniformly distributed within $[\Psi_{lb}, \Psi_{ub}]$, where $\Psi_{lb} = m_{lb}^1, m_{lb}^2, \dots, m_{lb}^\Gamma$ and $\Psi_{ub} = m_{ub}^1, m_{ub}^2, \dots, m_{ub}^\Gamma$.

Step 3: Evaluate the fitness of each population and determine the solution with best fitness (b_t)

Step 4: $\rho = \rho + n_p$;

Step 5: While $\rho \leq M_\rho$

for $j = 1 : n_p$

 Compute $V_{j1,\alpha}$ by the first technique

 Compute $V_{j2,\alpha}$ by the second technique

 Compute trial vector $v_{j1,\alpha}$ by Eq.(3.19)

 Compute trial vector $v_{j2,\alpha}$ by Eq.(3.19)

$$\rho = \rho + 2$$

End for

Step 6: Randomly reinitialize the trial vector $v'_{j,\alpha}$ (including $v_{j-1,\alpha}$ and $v_{j-2,\alpha}$) within the search space if any variable is outside its range.

Step 7: Selection

for $j=1$ to n_p

Evaluate the trial vector $v_{j,\alpha}$

If $f(v'_{j,\alpha}) \leq f(\Psi_{j,\alpha})$

$$\Psi_{j,\alpha+1} = v'_{j,\alpha}, f(\Psi_{j,\alpha+1}) = f(v'_{j,\alpha})$$

$$P_{j,pbest} = v'_{j,\alpha}, f(P_{j,pbest}) = f(v'_{j,\alpha})$$

If $f(v'_{j,\alpha}) < f(\Psi_{best,\alpha})$

$$\Psi_{best,\alpha} = v'_{j,\alpha}, f(\Psi_{best,\alpha}) = f(v'_{j,\alpha})$$

$$p_g = v'_{j,\alpha}, f(p_g) = f(v'_{j,\alpha})$$

End if

End if

End for

Step 8: Increment the generation count $\alpha = \alpha + 1$;

Step 9: End while

Ensemble based differential evolution uses adaptive $F \in [0.1, 1.0]$ and $CR \in [0, 1]$. The mutation technique and the attributes of CR are self-adapted based upon their previous performance. The proposed technique decomposes mutation into two types. First one provides best population and the other does not have. During evolution, proposed technique selects two techniques from these types in random fashion. Therefore, the ensemble based differential evolution outperforms existing variants of differential evolution.

3.2.2 Drug synergy prediction using ensemble based differential evolution

This section describes various steps which are required to be implement to evaluate the drug synergy score in an efficient manner. These steps are briefly discussed in the succeeding section.

- i. **Parameter selection:** To achieve the proposed technique, RBF kernel function of SVM is considered for regression. Because it demands only two attributes. These attributes will be optimized using the proposed ensemble based differential evolution.
- ii. **Scaling:** It is utilized is to avoid features in maximum numeric ranges dominating those with minimum numeric ranges. It also reduces the computational complexity for machine learning techniques. In this work, we have scaled the numeric attributes between either $[-1, 1]$ or $[0, 1]$ as:

$$v' = \frac{v - l_b^a}{u_b^a - l_b^a} \quad (3.21)$$

Here, v represents original value. v' is a scaled value. u_b and l_b represent upper and lower bound of feature values, respectively.

- iii. **Training and testing data:** The synergy data obtained from [142] is decomposed into training and testing data. The training data is used to train the proposed SVM based machine learning technique. Afterwards, testing data is used to monitor the effectiveness of the proposed technique.
- iv. **Ensemble based differential evolution:** The ensemble based differential evolution is used to tune the required attributes of SVM. The root mean squared error is used as a fitness function. The overall objective is to find optimal parameters for SVM based drug synergy prediction technique.

- v. **Stopping criteria:** When the stopping criteria achieved, the ensemble based differential evolution terminates itself and return tuned attributes; otherwise, continue for next iterations of ensemble based differential evolution.
- vi. **Technique building:** The results obtained from ensemble based differential evolution are used as initial attributes of SVM technique. It is then utilized to build the trained SVM based drug synergy prediction technique.
- vii. **Performance analysis:** In this step, trained technique obtained using ensemble differential evolution based SVM is applied on the training data to predict the drug synergy values. Afterwards, these values are used to evaluate the performance of proposed machine learning technique.

3.3 Materials and Methods

This section describes the experimental setup, used dataset and experimental analysis of existing and proposed machine learning techniques for drug synergy prediction.

3.3.1 Experimental Setup

The proposed and other methods have been implemented on Intel core *i7* processor with 16 GB RAM. The simulation environment is designed using MATLAB 2017a software with the help of matplotlib library. In succeeding section, the description of used datasets (i.e., AstraZeneca-Sanger Drug Combination [142]) along with its features.

3.3.2 Dataset

The drug effects on cells are measured at different concentrations can also be performed with drug combinations. In the presence of two drugs, the concentration space has been increased by one dimension and the induced effect has been described by a dose-response surface rather than a curve. A dose-response surface will generally look like this:

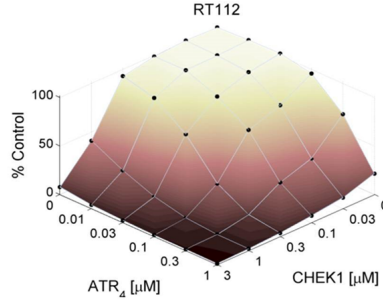


Figure 3.2: Dose response surface

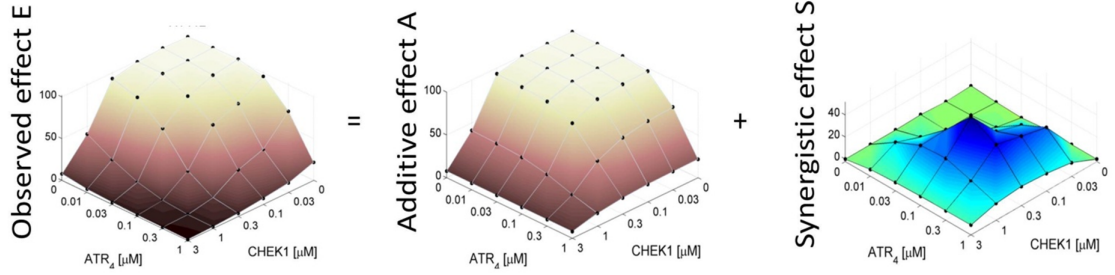


Figure 3.3: Dose response curves of different effects

The benefit of a drug combination has been partly assessed in terms of the amount of “extra-effect” (i.e., obtained when combining the drugs). Thus effects can be decomposed as following:

$$E(a, b) = A(a, b) + S(a, b) \quad (3.22)$$

Once the synergy distribution is obtained, it is summarized via a unique score number which we called total-synergy. The different features of dataset have been shown in Figure 3.1.

The synergy score data, which has been used in this research work for validation purpose contain two terms i.e. (a) Maximum concentration (uM) of drug A and (b) Maximum concentration (uM) of drug B which signifies that synergy score data depends upon the different concentration of different drugs. Therefore, the proposed machine learning technique has been implemented on the train data set containing the different concentration of different drugs to predict the synergy score [142]. The different concentration of different drugs has been evaluated using coefficient of drug

Table 3.1: Features of Drug Synergy Dataset

Features	Description
CELL_LINE	Normalised cell line name.
COMPOUND_A	Name of drug/compound A described by its target.
COMPOUND_B	Name of drug/compound B described by its target.
MAX_CONC_A	Maximum concentration (uM) of drug A.
MAX_CONC_B	Maximum concentration (uM) of drug B.
IC50_A	Concentration where half of the maximum kill is obtained with drug A.
H_A	Slope of the dose-response curve for drug A.
Einf_A(Potency)	Maximum cells killed (percentage) with drug A.
IC50_B	Concentration where half of the maximum kill is obtained with drug B.
H_B	Slope of the dose-response curve for drug B.
Einf_B	Maximum cells killed (percentage) with drug B.
SYNERGY_SCORE	Estimated total synergy between drug A and drug B in combination.
QA	Quality assurance (assessment) flag of combination assays
COMBINATION_ID	Name for drug A and drug B combination

interaction (CDI). It is evaluated as follows:

$$CDI = \frac{AB}{A \times B} \quad (3.23)$$

Here, AB is the ratio of the 2-drug combination group to the control group and A or B is the ratio of the single drug group to the control group. The $CDI < 1$ indicates synergism, especially $CDI < 0.7$ indicates a significantly synergistic effect, $CDI = 1$ indicates additivity and $CDI > 1$ indicates antagonism.

Subsequent sections describe experimental results of proposed and other techniques when they are applied on drug synergy score data [142].

3.3.3 Results and discussions

In this research work, 10-fold cross-validation has been used to test the proposed ensemble based differential evolution during the training phase to evaluate overfitting issue. To achieve 10-fold cross-validation, initially training data has been divided

into 10 equal subsets (fold). Keep the 1 fold as validation set and keep other 9 folds in the cross validation training set. Train proposed technique using the cross-validation training set and evaluate the accuracy of the proposed technique by validating the predicted values against the validation set. Similarly, accuracy of all 10 folds have been evaluated. To overcome the issue of overfitting, mean of calculated accuracies have been evaluated. Each fold has been utilized for validation just once. Therefore, 10-fold cross-validation guarantees that the proposed trained technique does not suffers from the overfitting issue.

- i. **Accuracy analysis:** Accuracy is a performance metric of a machine learning technique that reflects the proportionate number of times that the developed technique is correct when implemented to data. In regression based machine learning techniques, accuracy depends upon a factor so-called acceptance error (A_e). It allows us to accept the predicted data as accurate with some amount of error. A_e converts the difference between actual and predicted data to the binary form. In which some 1s determine the accuracy of data. In the same way, number of 0s determine the error rate. The accuracy can be computed as follows:

$$Accuracy = \frac{\sum_{i=1}^N d_i}{N} \times 100 \quad (3.24)$$

Here, N represents total number records in the predicted data. Also, a binary difference (d_i) can be computed as follows:

$$d_i = \begin{cases} 1 & \text{if } |a_i - p_i| \leq A_e \\ 0 & \text{Otherwise} \end{cases} \quad (3.25)$$

Here, a_i and p_i represented actual and predicted drug energy scores, respectively.

Thus, accuracy ranges from 0 to 100. The technique with accuracy techniques to 100 is more efficient compared with other techniques. In our experiment we have taken $A_e = 2$.

Table 3.2 depicts the accuracy analysis of the proposed ensemble based machine learning technique. It has been observed that the proposed technique outperforms others because it has better accuracy rate compared to other machine learning techniques. In Table 3.2, data is trained and tested on same drug synergy score data by dividing it into training and testing data. It has been observed that the proposed technique outperforms other machine learning techniques because the proposed technique has shown better accuracy compared to different techniques.

Table 3.2: Comparative analysis of accuracy on different data partitions

	Accuracy on different data partitions			
Model	50-50%	60-40%	70-30%	80-20%
Linear Model	89.6±4.9	90.1±4.9	92.4±3.6	92.1±4.1
J48	88.3±4.7	86.3±4.4	87.9± 4.5	89.2±3.4
SVM	90.5±3.7	91.3±3.1	90.3±3.9	91.3±3.4
Neural Network	87.2±5.4	86.2±4.8	87.2 ± 4.9	90.1±4.7
Random Forest	91.2±4.1	90.9±4.1	90.1±3.4	90.1±3.1
Proposed	94.1± 2.8	93.6±2.7	94.4±2.1	95.7±2.1

- ii. **Root mean square error analysis:** Root mean square error (RMSE) is a measure of the differences between actual and predicted values. RMSE denotes standard deviation of the error between predicted and actual values so-called residuals or prediction errors.

RMSE of predicted values $\hat{y}_t$ for attributes t of a regression's dependent attribute y_t is evaluated for n different predictions as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} \quad (3.26)$$

Table 3.3 shows the comparison of RMSE analysis between the proposed and

other machine learning techniques. It depicts that the proposed technique outperforms others because it has lesser RMSE compared to others. In Table 3.3, data is trained and tested on same drug synergy score data by dividing it into training and testing data. It has been observed that if we increase the size of training data, i.e., from 40 % to 80 % even then proposed technique outperforms others. Therefore, proposed technique has more realistic results compared to other machine learning techniques.

Table 3.3: Comparative analysis of RMSE (root mean square error) on different data partitions

	RMSE on different data partitions			
Model	50-50%	60-40%	70-30%	80-20%
Linear Model	13.89±2.6	14.21±2.6	13.56±3.1	13.82±2.7
J48	13.41±3.4	14.97±3.1	14.53±3.1	13.56±2.9
SVM	14.04±2.8	14.01±3.6	13.88±3.3	14.08±2.9
Neural network	14.99±3.1	15.21±3.1	14.98±3.4	14.34±2.5
Random forest	13.79±2.5	14.20±2.6	13.47±2.4	13.89±2.6
Proposed	12.42±1.6	12.14±1.5	11.64±1.3	11.79±1.4

- iii. **Coefficient of correlation analysis:** The coefficient of correlation (r) so-called Pearson correlation coefficient measures degree of linear correlation between two attributes A and B . $r \in [-1, 1]$, where 1 is represented positive linear correlation, 0 indicates no linear correlation and -1 represents negative linear correlation. Mathematically, r can be computed as follows:

$$r = \frac{c_v(A, B)}{\sigma_A \sigma_B} \quad (3.27)$$

Here, c_v represents covariance between A and B . σ_A shows standard deviation of A and σ_B shows standard deviation of B .

c_v can be computed as follows:

$$c_v(A, B) = \epsilon[(A - \mu_A)(B - \mu_B)], \quad (3.28)$$

Here, μ_A and μ_B represent mean of A and y , respectively. ϵ shows expectation.

Now, r can be rewritten as follows:

$$r = \frac{\epsilon[(A - \mu_A)(B - \mu_B)]}{\sigma_A \sigma_B} \quad (3.29)$$

Therefore formula for r can be represented in terms of uncentered moments.

Since,

$$\mu_A = \epsilon[A] \quad (3.30)$$

$$\mu_B = \epsilon \quad (3.31)$$

$$\sigma_A^2 = \epsilon[(A - \epsilon[A])^2] = \epsilon[A^2] - [\epsilon[A]]^2 \quad (3.32)$$

$$\sigma_B^2 = \epsilon[(B - \epsilon[B])^2] = \epsilon[B^2] - [\epsilon[B]]^2 \quad (3.33)$$

$$\epsilon[(A - \mu_A)(B - \mu_B)] = \epsilon[(A - \epsilon[A])(B - \epsilon[B])] = \epsilon[AB] - \epsilon[A] \epsilon[B], \quad (3.34)$$

Therefore, r can also be rewritten as follows:

$$r = \frac{\epsilon[AB] - \epsilon[A] \epsilon[B]}{\sqrt{\epsilon[A^2] - [\epsilon[A]]^2} \times \sqrt{\epsilon[B^2] - [\epsilon[B]]^2}}. \quad (3.35)$$

Table 3.4: Comparative analysis of coefficient of correlation on different data partitions

	Correlation on different data partitions			
Model	50-50%	60-40%	70-30%	80-20%
Linear Model	0.89±0.06	0.89±0.08	0.91±0.04	0.87±0.04
J48	0.86±0.07	0.84±0.06	0.87±0.06	0.91±0.06
SVM	0.91±0.08	0.90±0.05	0.89±0.07	0.91±0.06
Neural network	0.85±0.20	0.87±0.09	0.89±0.08	0.90±0.07
Random forest	0.90±0.06	0.91±0.07	0.92±0.08	0.90±0.07
Proposed	0.94±0.04	0.93±0.02	0.96±0.02	0.95±0.02

Table 3.4 demonstrates the coefficient of correlation analysis between the proposed ensemble based machine learning technique with other techniques. It has been observed that the proposed technique has significant correlation coefficient than other machine learning techniques. It has also been observed that the proposed technique has better as well as more consistent results compared to other techniques, as proposed techniques show minor variation than others.

Chapter 4

Ensemble based machine learning Scheme

Drug Synergy prediction plays a significant role in the medical field for inhibiting specific cancer agents. It can be developed as a pre-processing technique for therapeutic successes ¹. Examination of different drug-drug interaction can be done by drug synergy score. It needs efficient regression based machine learning approaches to minimize the prediction errors. Numerous machine learning techniques such as neural networks, support vector machines, random forests, LASSO, ElasticNets, etc., have been used in the past to realize requirement as mentioned above. However, these techniques individually do not provide significant accuracy in drug synergy score. In this chapter, we have discussed to design a neuro-fuzzy based ensemble approach. To achieve this, nine well-known machine learning techniques have been implemented by considering the drug synergy data. Based on the accuracy of each model, four techniques with high accuracy are selected to develop ensemble based machine learning model. These models are Random forest, Fuzzy Rules Using Genetic Cooperative-Competitive Learning method (GFS.GCCL), Adaptive-Network-Based Fuzzy Inference System (ANFIS) and Dynamic Evolving Neural-

¹Harpreet Singh, Prashant Singh Rana and Urvinder Singh, "*Prediction of drug synergy in cancer using ensemble-based machine learning techniques*", Modern Physics Letters B, World Scientific, 32(11):1-12, 2018.

Fuzzy Inference System method (DENFIS). Ensembling is achieved by evaluating the biased weighted aggregation (i.e., adding more weights to the model with a higher prediction score) of predicted data by selected models. The proposed and existing machine learning techniques have been evaluated on drug synergy score data. The comparative analysis reveals that the proposed method outperforms others in terms of accuracy, root mean square error and coefficient of correlation.

4.1 Introduction

Drug synergy or combination of different drugs is extensively utilized in treating the most dreadful diseases, such as cancer [167]. The main objective is to achieve synergistic therapeutic effect, dose and toxicity reduction and to decrease the induction of drug resistance [163]. Recent advancements in the study of cancer have shown that the disease cannot be explored only with the help of genetic mutations within the cancer cells [164]. Instead, tumors should be considered as compound tissues in which cancer cells grow and interconnect with neighboring micro-environment to endorse their survival and distribution [165]. Though, remedies that target specific signaling proteins have been effective for cancer treatment [166].

Majority of classical chemo drugs kill cells in the body that produce and divide rapidly [168]. Cancer cells divide at a faster speed. Therefore these drugs work efficiently against them [169]. However, it can also harm other cells in the body that share rapidly. Therefore, it leads specific side effects to human beings [170]. Throughout, the discovery process, estimating the drug amalgamation consequence in vitro profoundly depends on the examination. But, the testing of entire in-vitro set in cancer is not possible because of large combinatorial search space [171]. The computational estimation may help in recognizing probable drug mixture effect at a molecular level, such as hybridization of molecular and pharmacological data [2]. A regression-based estimate of drug-drug interactions (DDIs) by combining drug phenotypic, therapeutic, chemical and genomic properties and large scale pharmacogenomic screens of cancer cell lines [172]. These techniques are efficient to estimate

unknown DDIs [173]. Therefore, any computational approach to either guide readily testable candidates or reliably predict the effect of drug groupings would be expectable [174].

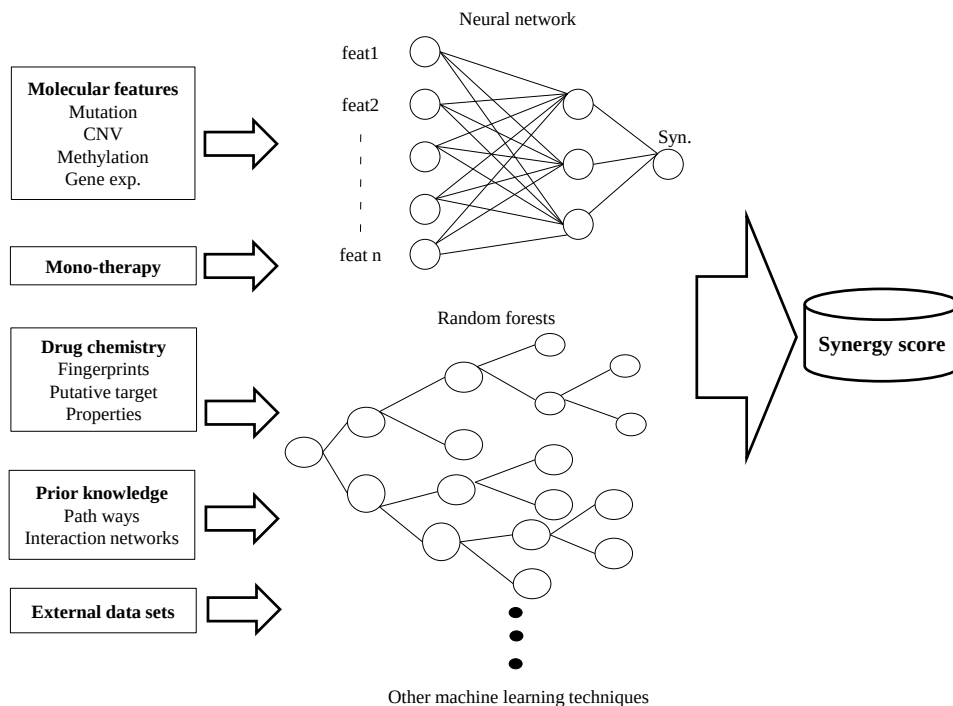


Figure 4.1: Flowchart of the proposed technique

Figure 4.1 shows the diagrammatically flow of the machine learning based drug synergy score detection. Initially, medical data will be loaded into the system's workspace and partition of data will be done between training and testing data [175]. Then, machine learning based techniques come in action to train a drug synergy model by considering training data [176]. In the end, the trained model is utilized to predict the synergy score from testing data [3].

Taghizadeh et al. utilized adaptive neuro-fuzzy inference system (ANFIS) to predict the surgery data. In this approach, attributes were reduced by considering a univariant analysis to make the machine learning model more efficient [177]. Chahkandi et al. designed a gradient-based iterative learning technique for parameter tuning of neuro-fuzzy based machine learning model. The premise and following attributes of neuro-fuzzy were taken randomly and then optimized using an iterative

technique. This technique utilizes the first order partial derivative of the predicted class concerning the structure attributes [178]. It provides better accuracy compared to standard ANFIS and neuro-fuzzy model [3].

Keeley et al. studied that to develop an efficient fuzzy classifier, substantial effort is required to optimize fuzzy sets [179]. But, these models do not focus on the method in which membership functions are integrated within fuzzy rules [180]. Therefore, standard fuzzy techniques have no control over how powerfully or inadequately the inference is implemented within these regulations [181]. A genetic-based fuzzy inference system was applied to improve the performance of fuzzy classifier system [182]. Each fuzzy system is developed by considering well-established decision tree techniques such as C4.5 and CHAID [183]. The proposed method provides efficient results compared to earlier techniques [183].

Therefore, the main issue associated with the existing machine learning techniques for drug synergy prediction is suffering from low accuracy rate [184]. Shuang and Antony proved that neural network based ensembles are effective approaches to enhance the accuracy of neural network based machine learning techniques. Integrating a set of neural network techniques whose error distributions are diverse can achieve better prediction outcomes than standard neural network classifiers [185].

Frederic and Max demonstrated that the ensemble based machine learning techniques improve the performance by integrating multiple models and their respective outcomes [186]. Yigit and Ufuk utilize six well-known machine learning techniques, such as decision tree, random forest, Bayesian network [187], Naïve Bayes, support vector machine and K^* , to design an ensemble based machine learning model. The voting technique is utilized to ensemble the output of other models [188]. Akin studied that the ensemble based machine learning technique integrate different machine learning models that make errors in different fragments of analyzed data. [189]. Therefore, we have proposed a novel ensembling based drug synergy prediction technique that overcomes the above-discussed issues.

4.1.1 Contributions

Following are our main contributions in this chapter:

- i. Initially, we have implemented nine well-known machine learning techniques on the drug synergy dataset. Among them, four techniques are selected with high accuracy rate.
- ii. After that, these techniques such as Random forest, Fuzzy Rules Using Genetic Cooperative-Competitive Learning method (GFS.GCCL), Adaptive-Network-Based Fuzzy Inference System (ANFIS) and Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) are used to propose an ensembling based drug synergy prediction technique.
- iii. To validate the proposed technique, the comparisons have been drawn between proposed and existing techniques by considering various sizes of training data, i.e., 40%, 50%, 60%, etc.

4.2 Proposed technique

In this section, proposed technique is described using mathematical formulation. Initially, ANFIS and Random forest-based machine learning models have been explained. Then, a proposed ensemble based model is described.

4.2.1 ANFIS

The fuzzy and neural networks come up with certain pros and cons. Therefore, researchers have merged them to utilize the features of both techniques. The integrated model is called Adaptive neuro-fuzzy system (ANFIS) [190]. It contains multiply layered feed-forward neural network. First-order fuzzy logic is utilized as a rule-based with two inputs V_α and V_β with a function response $f_1, f_2, \dots, f_n$. The

rule-based can be defined as follows:

Rule 1: If V_α is A_1 and V_β is B_1 , then $f_1 = \alpha_1 V_\alpha + b_1 V_\beta + c_1$

Rule 2: If V_α is A_2 and V_β is B_2 , then $f_2 = \alpha_2 V_\alpha + b_2 V_\beta + c_2$

$\vdots$

$\vdots$

Rule n : If V_α is A_i and V_β is B_n , then $f_n = \alpha_n V_\alpha + b_n V_\beta + c_n$

Here, $n = 1, 2, \dots$ and $i = 1, 2, \dots$, respectively. A_i and B_i represent fuzzy membership functions (*MFs*) for inputs V_α, V_β . a_n, b_n, c_n shows attributes of fuzzy rule base. ANFIS has five layers such as fuzzy layer, product layer, normalized layer, defuzzification layer, total output layer.

The ANFIS inputs are V_α and V_β that interference for each node from A_i and B_i . A_i and B_i are the linguistic functions utilized to define *MFs*. ANFIS utilize fuzzy triangular membership function. It is evaluated as follows:

$$f(x; z, c, m) = \begin{cases} \frac{x-z}{c-z} z \leq x < c \\ \frac{x-m}{c-m} c \leq x < m \end{cases} \quad (4.1)$$

Here, z and m represent bound to the width of *MFs*. c shows centre of *MFs* [190].

4.2.2 Random forest

The random forest contains several trees and selects random subsets of the number of different predictors tested at each node [191]. To grow the trees, a deterministic technique is used to select each tree from a random set of attributes. The node is utilized to decompose the nodes for decreasing the total number through description available for analysis. The sampled random vector and standard random forest are mixed as an estimator for each tree by using similar distribution for all trees.

It contains input vector $X = x_1, x_2, \dots, x_p$, where a p -dimensional input vector that works in building a forest. Inside the forest a set of K trees $T_1(x), T_2(x), \dots, T_k(x)$,

the output of each tree estimates the actual value $\hat{Y}_1 = T_1(X), \dots, \hat{Y}_m = T_m(X)$, where $m = 1, \dots, K$. The final result of it is the mean of all the values predicted by different trees.

$$Estimate_{RF}(X) = \frac{1}{k} \sum_{k=1}^k \hat{y}_k(X) \quad (4.2)$$

The training dataset is independently taken from the input and output $D = D_1, D_2, \dots, D_n = (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, where $x_i, i = 1, \dots, n$, is the training dataset for input vector and $y_i, i = 1, \dots, n$, is training dataset for output vector. Each tree is grown using the training approach. The estimated error and accuracy is evaluated for the random forest using the minimization of mean square error (MSE). Determining the optimum trees in forest focuses on MSE . Testing data is communicated using each split node, by sending it either to right or to left child until ending up at a leaf node.

$$MSE \approx MSE^{OOB} = \frac{1}{n} \sum_{i=1}^n (\hat{Y}(X_i) - Y_i)^2 \quad (4.3)$$

Here, $\hat{Y}(X_i)$ represents estimated output of trees in forest corresponding to a given input sample, Y_i shows observed output and n shows a total number of samples.

However, a random forest has difficulties in selecting the number of trees. In this work, we have taken a number of trees =100 [191].

4.2.3 Designed Ensembling based model

This step describes the detail of the proposed ensembling based machine learning model for drug synergy prediction. Figure 4.2 demonstrates the different stages of the proposed ensembling based machine learning technique. Every stage shows a building block of the proposed model. The following section provides brief description of every stage.

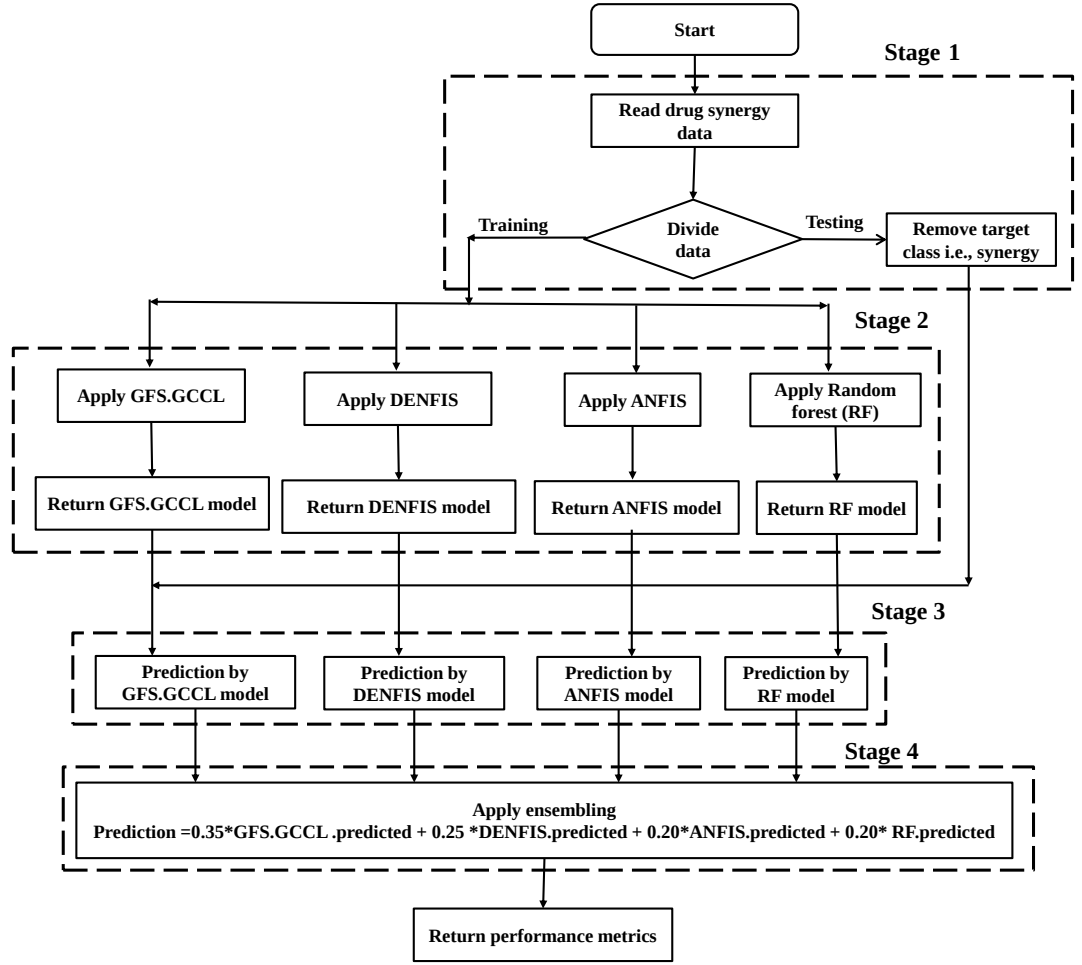


Figure 4.2: Flowchart of the proposed technique

Stage 1: Initially, actually drug synergy score dataset is loaded into the R-tool’s memory. This data contains multiple attributes along with the target class, i.e., drug synergy score. Then, this dataset is decomposed into training and testing data. Then, the target class ‘drug synergy score’ will remove it for validation purpose.

Stage 2: In this stage, well-known machine learning techniques such as Random forest, Fuzzy Rules Using Genetic Cooperative-Competitive Learning method (GFS.GCCL), Adaptive-Network-Based Fuzzy Inference System (ANFIS) and Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) are implemented on the drug synergy score training dataset to form their respective models.

Stage 3: In this stage-trained machine learning models are implemented on the testing data to predict drug synergy score.

Stage 4: This stage is an ensembling stage. Drug score predicted by Random forest, Fuzzy Rules Using Genetic Cooperative-Competitive Learning method (GFS.GCCL), Adaptive-Network-Based Fuzzy Inference System (ANFIS) and Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) models are integrated into this stage. The integration is achieved by evaluating the biased weighted aggregation (i.e., adding more weights to the model with a higher prediction score) of predicted data by selected models. The biased weighted aggregation is computed as follows:

$$\begin{aligned} Prediction = & .35 * GFS.GCCL.pred + .25 * DENFIS.pred \\ & + .20 * ANFIS.pred + .20 * RF.pred \quad (4.4) \end{aligned}$$

Here, pred represents predicted values by given machine learning models.

4.3 Performance analysis

This section provides the comparative analysis of proposed and existing machine learning techniques. The proposed and other methods have been implemented on Intel core *i7* processor with 16 GB RAM. The simulation environment is designed in well-known R-tool. Subsequent sections describe experimental results of proposed and other techniques when they are applied on drug synergy score data [142].

The synergy score data, which has been used in this research work for validation purpose contain two terms i.e. (a) Maximum concentration (uM) of drug A and (b) Maximum concentration (uM) of drug B. Therefore, the significance of synergy score data depends upon the different concentration of different drugs. Therefore, the proposed machine learning model has been implemented on the train data set containing the different concentration of different drugs to predict the synergy score [142]. The

different concentration of different drugs has been evaluated using coefficient of drug interaction (CDI). It is evaluated as follows:

$$CDI = \frac{AB}{A \times B} \quad (4.5)$$

Here, AB is the ratio of the 2-drug combination group to the control group and A or B is the ratio of the single drug group to the control group. The $CDI < 1$ indicates synergism, especially $CDI < 0.7$ indicates a significantly synergistic effect, $CDI = 1$ indicates additivity and $CDI > 1$ indicates antagonism.

In this research work, 10-fold cross-validation has been used to test the proposed ensemble based machine learning model during the training phase to evaluate overfitting issue. To achieve 10-fold cross-validation, initially training data has been divided into 10 equal subsets (fold). Keep the 1 fold as validation set and keep other 9 folds in the cross validation training set. Train proposed model using the cross-validation training set and evaluate the accuracy of the proposed model by validating the predicted values against the validation set. Similarly, accuracy of all 10 folds have been evaluated. To overcome the issue of overfitting, mean of calculated accuracies have been evaluated. Each fold has been utilized for validation just once. Therefore, 10-fold cross-validation guarantees that the proposed trained model does not suffers from the overfitting issue.

4.3.1 Accuracy analysis

Accuracy is a performance metric of a machine learning model that reflects the proportionate number of times that the developed model is correct when implemented to data. In regression based machine learning models, accuracy depends upon a factor so-called acceptance error (Ae). It allows us to accept the predicted data as accurate with some amount of error. Ae converts the difference between actual and predicted data to the binary form. In which some 1s determine the accuracy of data. In the same way, number of 0s determine the error rate. The accuracy can be

computed as follows:

$$Accuracy = \frac{\sum_{i=1}^N d_i}{N} \times 100 \quad (4.6)$$

Here, N represents total number records in the predicted data. Also, a binary difference (d_i) can be computed as follows:

$$d_i = \begin{cases} 1 & \text{if } |a_i - p_i| \leq A_e \\ 0 & \text{Otherwise} \end{cases} \quad (4.7)$$

Here, a_i and p_i represented actual and predicted drug energy scores, respectively. Thus, accuracy ranges from 0 to 100. The technique with accuracy approaches to 100 is more efficient compared with other techniques. In our experiment we have taken $A_e = 2$.

Table 4.1 depicts the accuracy analysis of the proposed ensemble based machine learning technique. It has been observed that the proposed technique outperforms others because it has better accuracy rate compared to other machine learning techniques. In Table 4.1, data is trained and tested on same drug synergy score data by dividing it into training and testing data. It has been observed that the proposed technique outperforms other machine learning models because the proposed technique has shown better accuracy compared to different techniques.

Table 4.1: Comparative analysis of Accuracy on different data partitions

	Accuracy on different data partitions				
Model	40-60%	50-50%	60-40%	70-30%	80-20%
Linear Model	86.41 $\pm$ 5.9	88.78 $\pm$ 5.1	89.23 $\pm$ 4.7	91.54 $\pm$ 3.8	91.92 $\pm$ 3.6
J48	88.90 $\pm$ 5.2	87.24 $\pm$ 4.8	85.82 $\pm$ 4.7	87.74 $\pm$ 4.3	88.48 $\pm$ 3.7
SVM	90.94 $\pm$ 3.8	91.40 $\pm$ 3.6	91.01 $\pm$ 2.9	89.18 $\pm$ 3.7	90.43 $\pm$ 2.7
Neural Network	86.49 $\pm$ 4.3	86.09 $\pm$ 6.3	85.04 $\pm$ 5.6	86.68 $\pm$ 5.8	89.44 $\pm$ 4.5
Random Forest	91.48 $\pm$ 3.1	90.13 $\pm$ 3.7	91.16 $\pm$ 3.6	89.18 $\pm$ 3.1	89.30 $\pm$ 2.8
CART	87.33 $\pm$ 3.9	90.94 $\pm$ 3.7	89.64 $\pm$ 3.3	90.36 $\pm$ 3.6	89.50 $\pm$ 3.2
ANFIS	91.04 $\pm$ 3.2	88.38 $\pm$ 3.6	90.35 $\pm$ 3.1	90.29 $\pm$ 3.2	90.77 $\pm$ 2.7
DENFIS	92.90 $\pm$ 2.9	93.95 $\pm$ 2.8	93.73 $\pm$ 2.7	93.62 $\pm$ 2.8	94.51 $\pm$ 2.3
GFS.GCCL	92.36 $\pm$ 2.6	93.57 $\pm$ 2.7	94.64 $\pm$ 2.7	94.78 $\pm$ 2.3	95.81 $\pm$ 2.1
Proposed	93.11 $\pm$ 2.3	94.05 $\pm$ 2.6	95.45 $\pm$ 2.5	96.08 $\pm$ 2.1	96.79 $\pm$ 2.1

4.3.2 Root mean square error analysis

Root mean square error (RMSE) is a measure of the differences between actual and predicted values. RMSE denotes standard deviation of the error between predicted and actual values so-called residuals or prediction errors.

RMSE of predicted values $\hat{y}_t$ for attributes t of a regression's dependent attribute y_t is evaluated for n different predictions as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} \quad (4.8)$$

Table 4.2 shows the comparison of RMSE analysis between the proposed and other machine learning techniques. It depicts that the proposed technique outperforms others because it has lesser RMSE compared to others. In Table 4.2, data is trained and tested on same drug synergy score data by dividing it into training and testing data. It has been observed that if we increase the size of training data, i.e., from 40 % to 80 % even then proposed technique outperforms others. Therefore,

proposed technique has more realistic results compared to other machine learning techniques.

Table 4.2: Comparative analysis of RMSE (root mean square error) on different data partitions

	RMSE on different data partitions				
Model	40-60%	50-50%	60-40%	70-30%	80-20%
Linear Model	15.12 $\pm$ 2.9	14.33 $\pm$ 2.7	14.18 $\pm$ 2.8	13.41 $\pm$ 2.9	13.28 $\pm$ 3.1
J48	14.29 $\pm$ 3.2	14.84 $\pm$ 3.2	15.32 $\pm$ 2.9	14.67 $\pm$ 2.8	14.43 $\pm$ 2.6
SVM	13.61 $\pm$ 2.8	13.46 $\pm$ 2.9	13.58 $\pm$ 3.4	14.19 $\pm$ 3.1	13.78 $\pm$ 2.8
Neural network	15.09 $\pm$ 2.4	15.22 $\pm$ 2.6	15.58 $\pm$ 2.9	15.03 $\pm$ 2.4	14.11 $\pm$ 2.7
Random forest	13.43 $\pm$ 2.3	13.88 $\pm$ 2.6	13.53 $\pm$ 2.8	14.20 $\pm$ 1.9	14.15 $\pm$ 2.2
CART	14.81 $\pm$ 2.6	13.61 $\pm$ 2.4	14.04 $\pm$ 2.6	13.80 $\pm$ 2.4	14.09 $\pm$ 2.3
ANFIS	13.57 $\pm$ 2.1	15.46 $\pm$ 1.9	13.81 $\pm$ 1.7	13.82 $\pm$ 1.8	13.66 $\pm$ 1.4
DENFIS	12.96 $\pm$ 1.6	12.60 $\pm$ 1.8	12.68 $\pm$ 1.4	12.71 $\pm$ 1.7	12.42 $\pm$ 1.3
GFS.GCCL	13.13 $\pm$ 1.4	12.73 $\pm$ 1.7	12.37 $\pm$ 1.5	12.33 $\pm$ 1.5	11.98 $\pm$ 1.2
Proposed	12.89 $\pm$ 1.3	12.57 $\pm$ 1.5	12.11 $\pm$ 1.3	11.91 $\pm$ 1.2	11.66 $\pm$ 1.3

4.3.3 Coefficient of correlation analysis

The coefficient of correlation (r) so-called Pearson correlation coefficient measures degree of linear correlation between two attributes A and B . $r \in [-1, 1]$, where 1 is represented positive linear correlation, 0 indicates no linear correlation and -1 represents negative linear correlation. Mathematically, r can be computed as follows:

$$r = \frac{c_v(A, B)}{\sigma_A \sigma_B} \quad (4.9)$$

Here, c_v represents covariance between A and B . σ_A shows standard deviation of A and σ_B shows standard deviation of B .

c_v can be computed as follows:

$$c_v(A, B) = \epsilon[(A - \mu_A)(B - \mu_B)], \quad (4.10)$$

Here, μ_A and μ_B represent mean of A and y , respectively. ϵ shows expectation.

Now, r can be rewritten as follows:

$$r = \frac{\epsilon[(A - \mu_A)(B - \mu_B)]}{\sigma_A \sigma_B} \quad (4.11)$$

Therefore formula for r can be represented in terms of uncentered moments.

Since,

$$\mu_A = \epsilon[A] \quad (4.12)$$

$$\mu_B = \epsilon \quad (4.13)$$

$$\sigma_A^2 = \epsilon[(A - \epsilon[A])^2] = \epsilon[A^2] - [\epsilon[A]]^2 \quad (4.14)$$

$$\sigma_B^2 = \epsilon[(B - \epsilon[B])^2] = \epsilon[B^2] - [\epsilon[B]]^2 \quad (4.15)$$

$$\epsilon[(A - \mu_A)(B - \mu_B)] = \epsilon[(A - \epsilon[A])(B - \epsilon[B])] = \epsilon[AB] - \epsilon[A] \epsilon[B], \quad (4.16)$$

Therefore, r can also be rewritten as follows:

$$r = \frac{\epsilon[AB] - \epsilon[A] \epsilon[B]}{\sqrt{\epsilon[A^2] - [\epsilon[A]]^2} \times \sqrt{\epsilon[B^2] - [\epsilon[B]]^2}}. \quad (4.17)$$

Table 4.3 demonstrates the coefficient of correlation analysis between the proposed ensemble based machine learning technique with other techniques. It has been observed that the proposed technique has significant correlation coefficient

than other machine learning techniques. It has also been observed that the proposed technique has better as well as more consistent results compared to other techniques, as proposed techniques show minor variation than others.

Table 4.3: Comparative analysis of correlation on different data partitions

	Correlation on different data partitions				
Model	40-60%	50-50%	60-40%	70-30%	80-20%
Linear Model	0.87 $\pm$ 0.09	0.90 $\pm$ 0.07	0.91 $\pm$ 0.07	0.92 $\pm$ 0.05	0.93 $\pm$ 0.04
J48	0.90 $\pm$ 0.07	0.88 $\pm$ 0.09	0.87 $\pm$ 0.08	0.89 $\pm$ 0.07	0.89 $\pm$ 0.08
SVM	0.92 $\pm$ 0.05	0.92 $\pm$ 0.06	0.92 $\pm$ 0.05	0.90 $\pm$ 0.07	0.91 $\pm$ 0.05
Neural network	0.87 $\pm$ 0.11	0.87 $\pm$ 0.10	0.86 $\pm$ 0.13	0.88 $\pm$ 0.09	0.90 $\pm$ 0.08
Random forest	0.92 $\pm$ 0.06	0.91 $\pm$ 0.07	0.92 $\pm$ 0.05	0.90 $\pm$ 0.07	0.90 $\pm$ 0.05
CART	0.88 $\pm$ 0.08	0.92 $\pm$ 0.06	0.91 $\pm$ 0.05	0.91 $\pm$ 0.06	0.90 $\pm$ 0.07
ANFIS	0.92 $\pm$ 0.05	0.86 $\pm$ 0.11	0.91 $\pm$ 0.07	0.91 $\pm$ 0.06	0.92 $\pm$ 0.05
DENFIS	0.94 $\pm$ 0.04	0.94 $\pm$ 0.04	0.94 $\pm$ 0.04	0.93 $\pm$ 0.05	0.94 $\pm$ 0.04
GFS.GCCL	0.93 $\pm$ 0.04	0.95 $\pm$ 0.04	0.96 $\pm$ 0.03	0.96 $\pm$ 0.03	0.97 $\pm$ 0.02
Proposed	0.94 $\pm$ 0.04	0.95 $\pm$ 0.03	0.96 $\pm$ 0.03	0.97 $\pm$ 0.02	0.98 $\pm$ 0.01

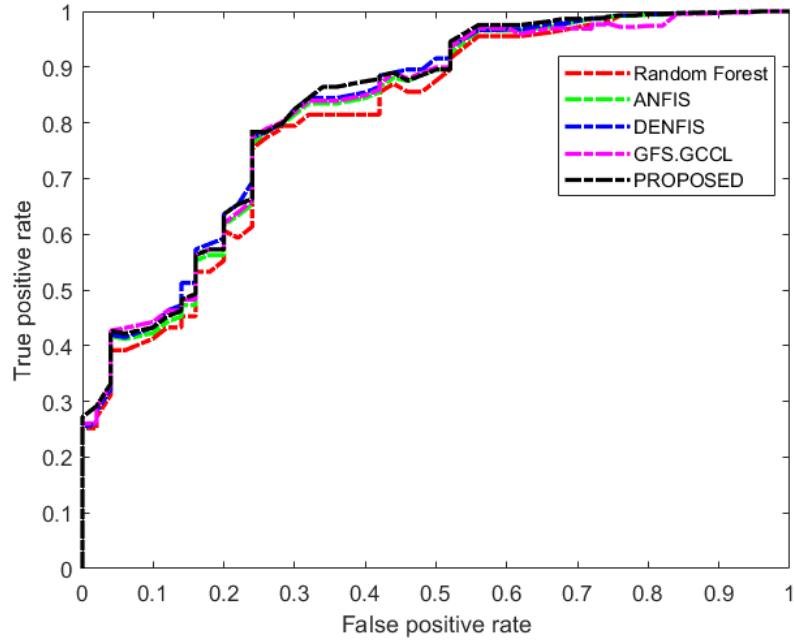


Figure 4.3: ROC curve

4.4 Receiver operating characteristic curve

A receiver operating characteristic curve (ROC curve) is a graphical plot that shows the diagnostic ability of a binary classifier model as its discrimination threshold is varied. Since, our data is regression based, therefore, we have used threshold value ($t_1 = 0.5$) to convert synergy score into classification data. The ROC curve is developed by plotting the true positive rate (sensitivity) against the false positive rate ($1 - \text{specificity}$) at various threshold settings. Figure 4.3 demonstrates that the proposed model has significantly more sensitivity than existing models.

Chapter 5

Adaptive Lévy based Neural Network Scheme

Medical domain is getting reshaped with continuous evolution of machine learning techniques that empowers doctors and medical practitioners to take data driven decisions. As the cancer cases are looming large worldwide, the applications of data science and machine learning in these fields possesses a greater scope especially when availability of major data breakthrough is on the horizon. Predictive analytics holds ability to generate more efficient and accurate drug synergy score, that describes the synergies between interactions of two drugs. Various machine learning techniques have been designed so far for the prediction of drug synergy score. However, parameter tuning is the primary issue for these techniques. A robust tuning of machine learning techniques can improve the prediction score of drug synergy data. To overcome this issue, an efficient Evolutionary based Neural Network Structure using Adaptive Lévy technique has been proposed and discussed in this chapter. The existing machine learning techniques and this evolutionary-based technique are tested on the drug dataset to predict synergy score. As supported by comparative analysis, the findings of the proposed research intends to extend outperforms all the known techniques in terms of parameters like accuracy, coefficient of correlation and RMSE.

5.1 Introduction

Drug synergy plays a critical role in cancer cure. As per earlier belief, genetic mutations were a pivot in curing the deadly disease, but the latest scientific discoveries unveil that mutations cannot solely help in examination of disease [164]. Rather, tumors should be well thought out that describes the neighbouring cells where cancerous cells breed within the micro-environment [165]. The primary goal would be to accomplish synergistic therapeutic effect, toxicity and dose reduction, and diminutive the development of drug resistance in host. Therefore, the efficacious treatment for cancer are the cures which target particular signalling proteins [166]. A cancerous tumour results in a higher rate of division of cells thereby ensuing cancer to spread in host at a very fast rate. When a patient undergoes chemotherapy, cytotoxic and other drugs target actively spreading cells [169]. This way cancer causing cells are eliminated from the host, but the process also damages non-cancerous cells that are critical for other important functions. These are some of the adverse side effects of the therapy [170].

During discovery phase, actual examination conducted in lab experiments greatly defines the effects of drug amalgamation in vitro. But it is not likeable to test the whole in-vitro set in cancer due to wide combinatorial search space [171]. Hybridization of molecular and pharmacological data [2] are common examples of computational estimation techniques that assist in spotting feasible drug mixture effect at a molecular level. DDIs have been calculated using regression based approach by integrating genomic, chemical, drug phenotypic and therapeutic properties. Pharmacogenomic screens of cancer cell lines are also calculated using regression-based approach at large scale. Hence, computational approaches are used to guide the search in the right direction for the suitable candidates i.e. drugs to be examined in vitro [174].

One of the primary problems with the conventional computational techniques to predict drug synergy for cancer cure is that the accuracy rate is significantly

low [184]. Shuang and Antony has proposed to overcome this drawback of low accuracy in a practical manner with neural network-based ensembles. Further, when neural networks with diverse set of error distributions are integrated the results are strengthened as compared to usual NN classifiers [192, 193].

In 1975, Holland invented the Genetic algorithm (GA), which is broadly used in optimization problems [194]. Complex optimization and severe analytical problems are the primary targets for the GAs, which consists of many parameters to be tuned. So, GA can provide a globally optimal solution over an ample search space. It has used among different fields such as classification, modelling [195], fuzzy control [196] [197], greenhouse climate control [198], path planning [194] etc.

The neural system was turned out to be an all-inclusive estimator [192]. Different neurons in network structure are connected in different patterns according to the different problems such as forecast, system demonstrating and control etc [199]. The learning of these structures can be optimized using some meta-heuristics approaches, for example, GA, PSO (Particle Swarm Optimization) and back proliferation calculation [200]. Furthermore, calculations used for the neural network learning are used to understand the learning procedure. Neural system with default parameters may not give ideal execution inside a characterized preparing slot. Sometimes, the used neural network can produce better results than it ought to have. At the point when this occurs, the preparation time frame may be significant (as the performance isn't met) and the cost of usage will increase.

Computational techniques and machine learning models offer the likelihood to investigate the vast synergistic space effectively. Precise prescient models can be produced by utilizing accessible cooperative energy information. Reliable forecasts give direction to in-vitro and in-vivo drug generation. Besides, in an in-vivo environment, techniques created to use genomic data for their expectations provide the chance to apply them additionally. In this way, prescient models constitute a significant advance, which leads to exactness drug [201]. The collaboration forecast has just been handled with a wide assortment of methodologies. Techniques

run from frameworks science strategies, motor models, blended whole number direct programming strategies dependent on the infected quality set [202], quality articulation profiles with individual medications and portion reaction bends [203], to computational techniques including Naive Bayes and Decision Tree strategies [203]. Nonetheless, these techniques are confined to some fixed cell lines or pathways. While the space of possible drug combinations is too large to be tested exhaustively, there have been significant efforts to measure the efficacy of drug combinations via high-throughput screening. However, it is unfeasible to comprehensively examine the immense space of possible combinations, which motivates the need for a data-driven approach to discovering effective combinations of drugs. There have been a variety of prior attempts to use machine learning methods to predict the most synergistic combinations of anti-cancer drugs. A recent study made synergy predictions by applying deep learning to a large dataset of drug combinations [204]. But since the accuracy of deep learning models largely depends on its architecture so the accuracy of this model can fall into local maxima without reaching the global maxima.

Therefore, a novel evolutionary-based drug synergy prediction using adaptive lévy based neural network structure has been proposed that overcomes the aforementioned issues.

5.1.1 Contributions

Following are the main contributions in this chapter:

- i. Initially, we have implemented eight existing machine learning techniques on the drug synergy dataset.
- ii. Thereafter, an evolutionary-based neural network using the Adaptive Lévy technique is proposed for drug synergy prediction. It tunes the parameters of a neural network by considering the genetic algorithm with an Adaptive Lévy

technique for mutation.

- iii. At last, to validate the proposed technique, the comparisons are drawn between this proposed and existing techniques by considering various sizes of training data, i.e., 20%, 40%, 60%, 80%, and 90%.

5.2 Artificial Neural Networks(ANN)

ANN has potential to model the complex and non linear problems because of their flexible network design and algorithms. Currently, undergoing research on ANN principally focuses on both of these aspects. ANN is made up of nerves, which can be primarily set of processing components, and they're connected to one another. It may also be considered as a multi-graph and every neuron act as transport method [205]. A neuron is generally a single-output and multi-input nonlinear component. The structure of an ANN depends upon all connections of the system as well as the transport functions of the nerves.

5.2.1 The Process of Learning in ANN

The learning procedure in an ANN is generally implemented by illustrations known as 'training' since the learning process is accomplished by progressively altering the assigned weights of neurons. The gradient descent is responsible for the optimization calculations called a backpropagation algorithm, so it may help to change weights consequently progressively. It further leads to reduce the error gradients. A generalized delta rule is used to update the weights in different iterations during backpropagation [206]. Backpropagation neural networks can act as an instrument to resolve function approximation issues rather than pattern recognition issues. Consider a neural network with training group P . The error function in the neural network is given that as:

$$F = \sum_{p \in P} {}^p E \quad (5.1)$$

where p being pattern in training set P

$${}^pE = \frac{1}{2} \sum_{j \in M} (\hat{y}_m - y_m)^2 \quad (5.2)$$

with $\hat{y}_m$ being the predicted output of m -th output of y_m

The weight function of neural network is given as:

$${}^p \Delta w_{hm} \sim -\nabla_w \cdot {}^pE \quad (5.3)$$

where w_{hm} being the weight from neuron h to neuron m and ∇_w being the gradient with respect to w .

$$\Delta w_{hm} = -\eta \frac{\partial E(w_{hm})}{\partial w_{hm}} \quad (5.4)$$

where η being the learning rate.

The equation for the relative change in weights of output neuron with respect to error function is given as:

$$net_m = \sum_{i=0}^H w_{im} \tilde{o}_i \quad (5.5)$$

where o_i being the predicted output of the i^{th} neuron and H being the number of hidden layers neurons it is connected to.

$$o_m = y_m = f_m(net_m) = \frac{\partial E(w_{hm})}{\partial w_{hm}} = \frac{\partial E}{\partial net_m} \cdot \frac{\partial net_m}{\partial w_{hm}} \quad (5.6)$$

The equation for the relative change in weights for hidden neurons is given as:

$$net_h = \sum_{i=0}^H w_{ih} \tilde{o}_i \quad (5.7)$$

$$\frac{\partial E}{\partial w_{kh}} = \frac{\partial E}{\partial net_h} \cdot \frac{\partial net_h}{\partial w_{kh}} \quad (5.8)$$

5.3 Evolutionary Algorithms

Evolutionary algorithm (EA) is a subdivision of Evolutionary computation and it is a part of various stochastic search algorithms. EA is a metaheuristic optimization algorithm which takes population into consideration in its search algorithms. Metaheuristics are those high-level procedures which are used to find or select lower-level procedures or heuristics which would conduct partial search. EA is preferred over traditional methods because it is simple and flexible since it is inspired by the concept of natural evolution. These algorithms are even robust and they have been developed in such a manner that they adapt solutions even in changing environments but this does not happen in traditional optimization methods where it changes with changes occurring in the problem environment. The major advantage of EA over traditional methods is its ability to consider prior information about problems and it helps the Evolutionary Algorithms to perform better in comparison to a traditional method which uses less information. Evolutionary Algorithms can also be further classified into various subclasses like Evolutionary Programming, Genetic Algorithm, Evolutionary Strategies and Genetic Programming [207]. The most important type of EA is Genetic Algorithms. It depends upon the phenomenon of natural selection along with genetics. This phenomenon is the same as a biological process where individuals which are strong and fit are given higher preference of winning in a competitive environment. These algorithms truly perform the calculations based upon the population size by mimicking the organic growth of biological programs. GA assumes that a set of parameters can be used to represent an individual which can be a possible solution of a problem. These sets of parameters are usually known as genes of a chromosome and are represented by binary numbers in a form of strings. Evaluation parameter to calculate the goodness values of particular chromosome is fitness function. A higher positive fitness value means that the chromosome is more fit. It leads to produce offspring of better quality indicating that a better solution would be produced for a problem. In a genetic algorithm, the first step is

to define a population pool of chromosomes whose size can be varied depending on the problem statement [208]. An Evolving process refers to different iterations of genetic operations in which next generation species are generated from the chromosomes of previous generations. Now, to produce offspring in the next generation, genes of previous generation chromosomes are mixed and recombined. So, if there is need to produce fit offsprings, parents are selected from the mating pool via a selection technique so that genes of only fit chromosomes are mixed or recombined together to produce next generation offsprings. Therefore, better chromosome has a higher chance of being selected in the mutation process. Subsequently, it would also produce many offsprings and possessing a higher chance of surviving in the coming generations and imitating the nature's law i.e. survival of the fittest. Different individuals among the population communicate with each other and compete to get their own survival.

Algorithm 1: Genetic Algorithm

The procedure of Genetic Algorithm is:

- 1** : Initialize random variable $q = 0$.
- 2** : Initialize absolute population $R(q)$.
- 3** : Calculate fitness of every chromosome from $R(q)$.
- 4** : Choose parents among chromosomes in $R(q+1)$ on the basis of fitness values of individuals.
- 5** : Perform crossover and mutation and subsequently create $R(q+1)$.
- 6** : Increment $q = q + 1$.
- 7** : Loop step 3 to step 6 before termination condition is reached.

5.4 The optimization of ANNs based on GAs

Neural-network, even though being powerful, has limits too. It can readily fall into the trap of local minimum and sometimes tough to adjust the architecture. Evolutionary Algorithm might well not be of use when finding in simple issues due to its computation cost, and its own performance might not be equal with all the distinctive calculations for a specific issue. However, GA is powerful, non-linear and parallel, so it could be utilized to find the optimal solution. EAs may also used for

training of connection weights that could eliminate the issue of local maxima and also make it reach global maxima [209]. To clarify neural network structure, the principal points include: how many layers, the number of neurons of each coating, the linking among nerves, etc. The blend of principle points used to design the neural network can be optimized with the help of genetic algorithm. The ensembling of GA with the neural network can give the best possible results.

Algorithm 2: Evolutionary Neural Network

- STEP 1** : Randomly generate initial weights.
- STEP 2** : Train the structure from the respective set among initial weights.
- STEP 3** : Calculate the fitness of every individual as obtained in the resultant after training.
- STEP 4** : Perform the selection procedure among several individuals whose fitness is above certain criterion.
- STEP 5** : Implement crossover among individuals. and mutation to present population to yield next generation.
- STEP 6** : Implement mutation among present population to yield next generation.
- STEP 7** : Loop Step 2 to Step 5, until accuracy criteria is fulfilled.

5.5 Improved Evolutionary Neural Network

The training of the neural network is improved by ensembling the evolutionary algorithm i.e. genetic algorithm with the neural network. Weights to be assigned in the neural network is optimized with the help of population generated through different phases of genetic algorithm as shown in Figure 5.1. Diversity in the population generated with the help of Mutation operator phase as discussed below.

5.5.1 Combined Mutation

The combined mutation technique is utilised to enhance the performance of ANN. A combined mutation procedure is the usage of a combination of a few mutation operators. This procedure of combined mutation was introduced by Chellapilla (1998). He introduced two mutation operators i.e. adaptive mean mutated (AMM) operator and mean mutation (MM) operator. The process involves the use of a combination of Gaussian and Cauchy distributions, wreathe them to one distribution and on the

basis of this generate a brand-new solution. Both joint mutation operators devised via this technique are called adaptive mean mutated (AMM) operator and also mean mutation (MM) operator introduced by Koenig in 2002. The overall equations are given by:

For MM:

$$x_i^{t+1} = x_i^t + 0.5\sigma'(H(\lambda) + D(\delta)) \quad (5.9)$$

Here, $H(\lambda)$ denotes the random number following the gaussian distribution with variance of $(\sigma^2 = 1)$. and $D(d)$ denotes the random number following the Cauchy distribution. Mean for both the distribution is 0.

For AMM:

$$x_i^{t+1} = x_i^t + \sigma'_1 H(\lambda) + \sigma'_2 D(\delta) \quad (5.10)$$

σ'_1 corresponds to member of population and σ'_2 is similar to performing the mutations as:

$$\sigma'_i = \sigma_i \exp(p' \cdot H(\lambda) + p H_t(\lambda)) \quad (5.11)$$

where i corresponds to size of the population, Gaussian distribution generated the $H(\lambda)$ and $H_t(\lambda)$ random numbers with $H_t(\lambda)$ generated only once. The value of p' and p are given by $\sqrt{2 \ln n}^{-1}$ respectively. In MM, Eq. (5.9) shows the creation of offspring by adding two distribution and assigning the weight to it where as in AMM, weights are assigned to each offspring before addition which helps to adapt the shape and the parameters of the distribution as shown in Eq. (5.10).

5.5.2 Adaptive Lévy

The technique of Adaptive Lévy for mutation is given by Lee and Yao [210] as a way to enhance the speed of search. It is accomplished by deducing the fittest mutated offspring of four distinct distributions of the candidate pool as shown in Figure

5.2 [211]. Lévy distribution probability density function is represented as:

$$f_{Levy(\mu,\nu)}(\mu) = \frac{1}{\pi} \int_0^{\inf} \exp^{-\nu q^\mu} \cos(q\mu) dq \quad (5.12)$$

where $1 < \mu < 2$, $\nu > 0$ and is set to 1. Four different equations are designed based on μ values. The distribution corresponds to Cauchy when $\mu = 1$ and it corresponds to Gaussian distribution when $\mu = 2$. The other two Lévy random 1 and Lévy random 2 distributions have values of 1.3 and 1.7. The different equations designed based on μ are:

$$x_{i,1}^{t+1} = x_i^t + P(\mu)(T^* - x_i^t) \quad (5.13)$$

$$x_{i,2}^{t+1} = x_i^t + D(\mu)(T^* - x_i^t) \quad (5.14)$$

$$x_{i,3}^{t+1} = x_i^t + R_1(\mu)(T^* - x_i^t) \quad (5.15)$$

$$x_{i,4}^{t+1} = x_i^t + R_2(\mu)(T^* - x_i^t) \quad (5.16)$$

where $P(\mu)$, $D(\mu)$, $R_1(\mu)$ and $R_2(\mu)$ reduces to Lévy distributed random numbers relative to distribution of Gaussian, Cauchy, Lévy random 1 and Lévy random 2. Selection will be made based on unique results provided with the help of the four aforementioned equations.

5.5.3 Structure and Parameter's Tuning

Tuning of the neural network is performed by the GA for learning the Input and Output pattern of a given problem. The input/output relation is as follow:

$$q^d(x) = s(f^d(x)) \quad (5.17)$$

$x=1,2,\dots,n_d$ where

$$q^d(x) = [q_1^d(x), q_2^d(x), q_3^d(x), \dots, q_{n_{out}}^d(x)] \quad (5.18)$$

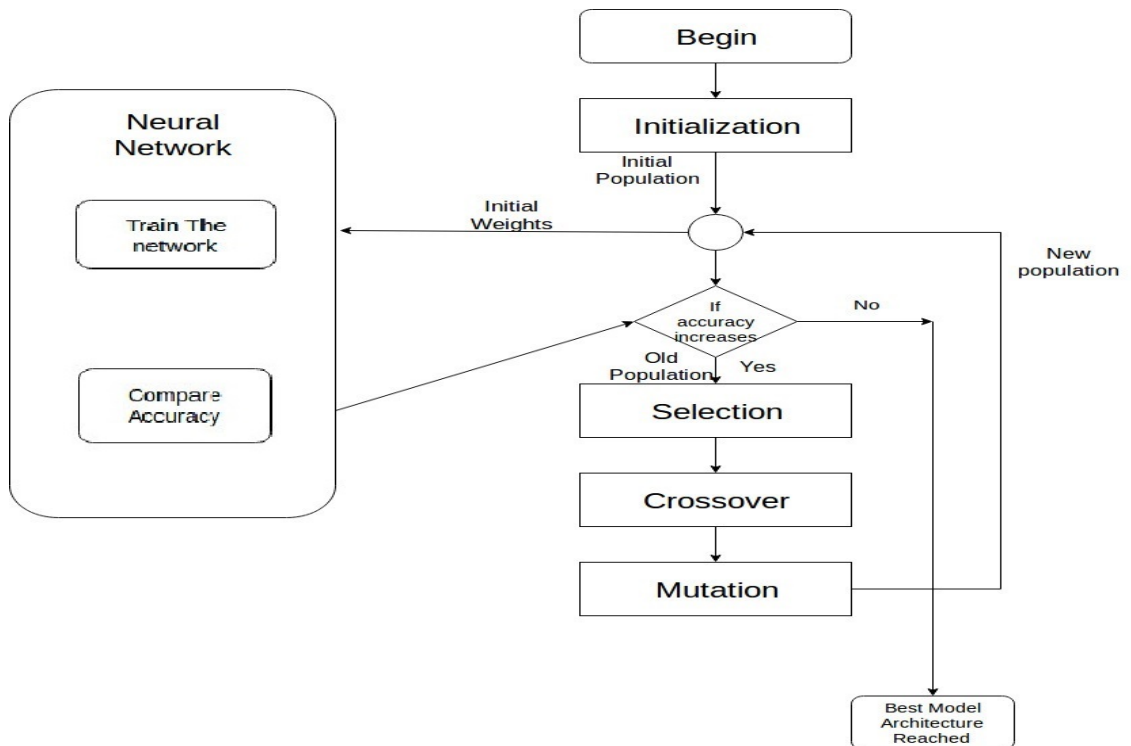


Figure 5.1: Evolutionary Neural Network Model

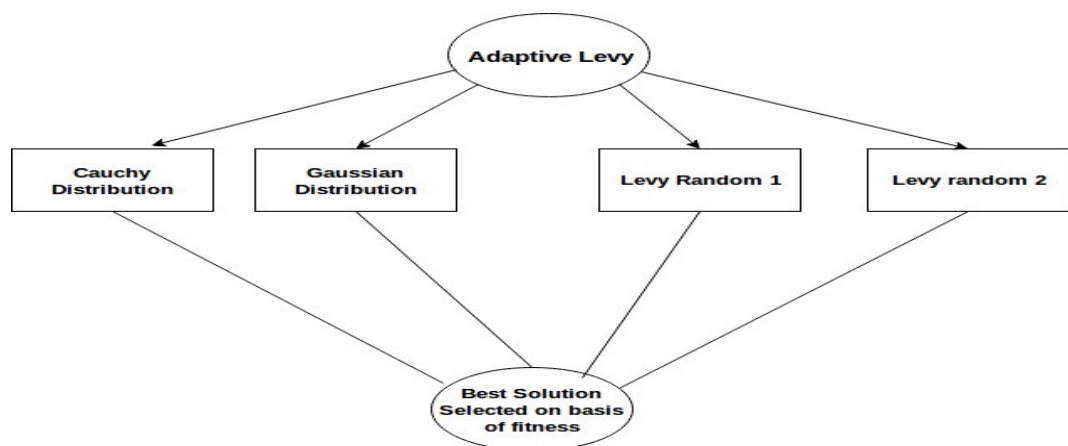


Figure 5.2: Adaptive Lévy Technique For Mutation

and

$$f^d(x) = [f_1^d(x), f_2^d(x), f_3^d(x), \dots, f_{n_{in}}^d(x)] \quad (5.19)$$

The terms used in Eq. (5.19) are the desired outputs and the given inputs. These inputs and outputs belong to unknown non-linear function $s(f^d(x))$. Data pairs of input and output are represented by n_d . The trade-off between fitness function and the error value is as given below:

$$fitness = \frac{1}{1 + error} \quad (5.20)$$

$$error = \sum_{k=1}^{n_{out}} \frac{n_k \sum_{x=1}^{n_d} |q_k^d(x) - q_k(x)|}{n_d} \quad (5.21)$$

In Eq. (5.20), the goal is to reach the global maxima by initializing the chromosome to $[s_{jk}^2, w_{jk}, s_{ij}^1, v_{ij}, s_j^1, b_j^1, s_k^2, b_k^2]$ for all values of i, j and k. Referring to Eq. (5.21), the error value may affect the maximum fitness value for some outputs, which are not of great interest. To reduce the amount of error, $n_k, k = 1, 2, \dots, n_d$, is chosen among dominated outputs. The trade-off between Eq. (5.20) and (5.21) shows that smaller error value leads to larger fitness value.

5.6 Dataset

The effect of drugs and drug combination on cells are estimated at various concentrations. The multiple attributes of a dataset have been shown in Table 3.1. The advantage of a drug combination is the measure of "extra-effect" (i.e., acquired during integrating of drugs) which results in it into the decomposition of effects as following:

$$E(u, z) = A(u, z) + S(u, z) \quad (5.22)$$

$E(u, z)$ denotes observed impact. $A(u, z)$ is frequently named the "additive" impact [212]. $S(u, z)$ is the synergistic effect and antagonistic in case it is negative. Loewe method is used to infer the added impact. This strategy depends upon isobole con-

dition which was numerically fathomed for different values of all drug combinations. It further leads to calculate $A(u, z)$ which helps to determine Eq. (5.22) [213].

After the achievement of synergy dispersion, summarization can be done with the help of drug synergy score. The synergy score information which has been utilized for the validation purpose consist of two attributes i.e. (a) Drug A's maximum concentration in uM and (b) Drug B's maximum concentration in uM. Hence, the criticalness of synergy score information relies on the diverse concentration of various drugs. In this manner, the proposed system has been executed on the trained data set consisting diverse convergence of various drugs for the prediction of synergy score [142]. The diverse concentrations of various drugs has been assessed utilizing coefficient of drug interaction(CDI) [193]. It is assessed as pursues:

$$CDI = \frac{uz}{u \times z} \quad (5.23)$$

Here, combination of two drug's proportion is denoted by uz and u or z represents proportion of the single drug. The $CDI < 1$ shows synergism, particularly $CDI < 0.7$ demonstrates an altogether synergistic impact, $CDI = 1$ shows additivity and $CDI > 1$ shows antagonism [214].

5.7 Results and discussions

In this proposed work, 10-fold cross-validation has been utilized in the preparation stage to assess overfitting issue. To accomplish 10-fold cross-validation, at first training data has been separated into 10 same size subfolds. After segregation, cross validation has been performed by keeping the 9 folds in the training section and 1 fold in the validation section. This process is repeated at different runs by shuffling the different folds among training as well as validation section. Accuracy of the proposed method has been accessed by validating the predicted values against the actual values. Correspondingly, accuracy of all 10 folds has been assessed. Every fold has participated in validation once. Examination of mean of calculated accuracies

Table 5.1: Comparative analysis of Accuracy on different data partitions

Model	Accuracy on different data partitions				
	40-60%	50-50%	60-40%	70-30%	80-20%
LM	86.41 $\pm$ 5.9	88.78 $\pm$ 5.1	89.23 $\pm$ 4.7	91.54 $\pm$ 3.8	91.92 $\pm$ 3.6
J48	88.90 $\pm$ 5.2	87.24 $\pm$ 4.8	85.82 $\pm$ 4.7	87.74 $\pm$ 4.3	88.48 $\pm$ 3.7
SVM	90.94 $\pm$ 3.8	91.40 $\pm$ 3.6	91.01 $\pm$ 2.9	89.18 $\pm$ 3.7	90.43 $\pm$ 2.7
NN	86.49 $\pm$ 4.3	86.09 $\pm$ 6.3	85.04 $\pm$ 5.6	86.68 $\pm$ 5.8	89.44 $\pm$ 4.5
RF	91.48 $\pm$ 3.1	90.13 $\pm$ 3.7	91.16 $\pm$ 3.6	89.18 $\pm$ 3.1	89.30 $\pm$ 2.8
CART	87.33 $\pm$ 3.9	90.94 $\pm$ 3.7	89.64 $\pm$ 3.3	90.36 $\pm$ 3.6	89.50 $\pm$ 3.2
ANFIS	91.04 $\pm$ 3.2	88.38 $\pm$ 3.6	90.35 $\pm$ 3.1	90.29 $\pm$ 3.2	90.77 $\pm$ 2.7
DENFIS	92.90 $\pm$ 2.9	93.95 $\pm$ 2.8	93.73 $\pm$ 2.7	93.62 $\pm$ 2.8	94.51 $\pm$ 2.3
GFS.GCCL	92.36 $\pm$ 2.6	93.57 $\pm$ 2.7	94.64 $\pm$ 2.7	94.78 $\pm$ 2.3	95.81 $\pm$ 2.1
Proposed	93.11 $\pm$ 2.3	93.56 $\pm$ 2.4	94.29 $\pm$ 2.3	95.28 $\pm$ 1.82	95.91 $\pm$ 1.10

have been done to get over the problem of overfitting.

5.7.1 Accuracy analysis

Accuracy is one of the crucial evaluation parameter of machine learning that ensures the consistency of the proposed system. In regression based problems, precision relies on a parameter called acceptance error (A_e) which acknowledge the predicted information by allowing the minute window of measurable error. A_e changes over the deviation from actual data to the predicted data and converted in to binary form. Exactness is defined by:

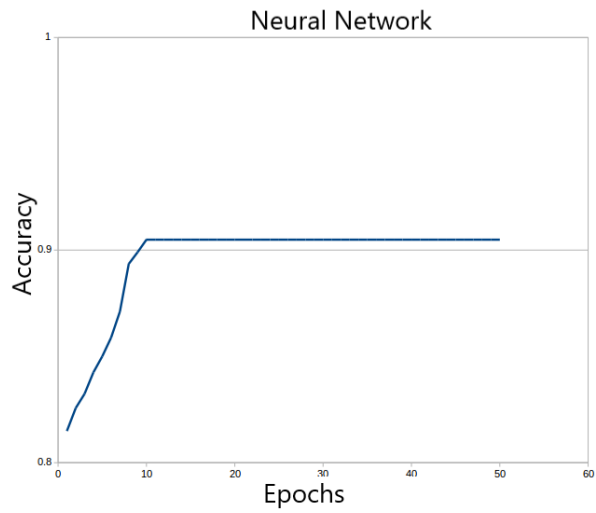
$$Accuracy = \frac{\sum_{i=1}^N d_i}{N} \times 100 \quad (5.24)$$

where, N indicates complete set of values in the predicted data and (d_i) represents the binary difference which can be calculated as:

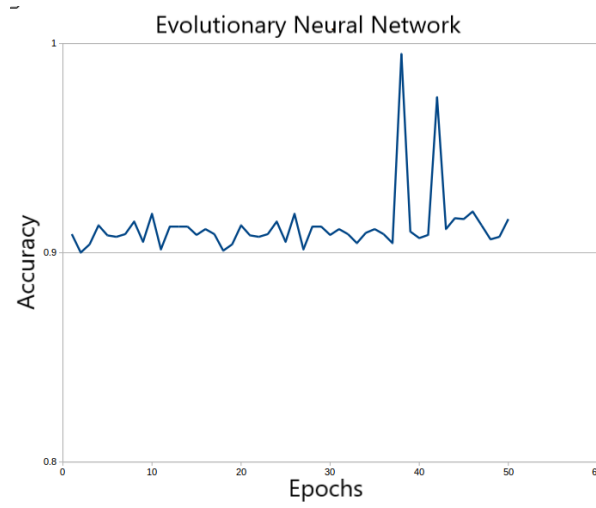
$$d_i = \begin{cases} 1 & \text{if } |a_i - p_i| \leq A_e \\ 0 & \text{Otherwise} \end{cases} \quad (5.25)$$

Here, a_i and p_i represented predicted and actual drug synergy score.

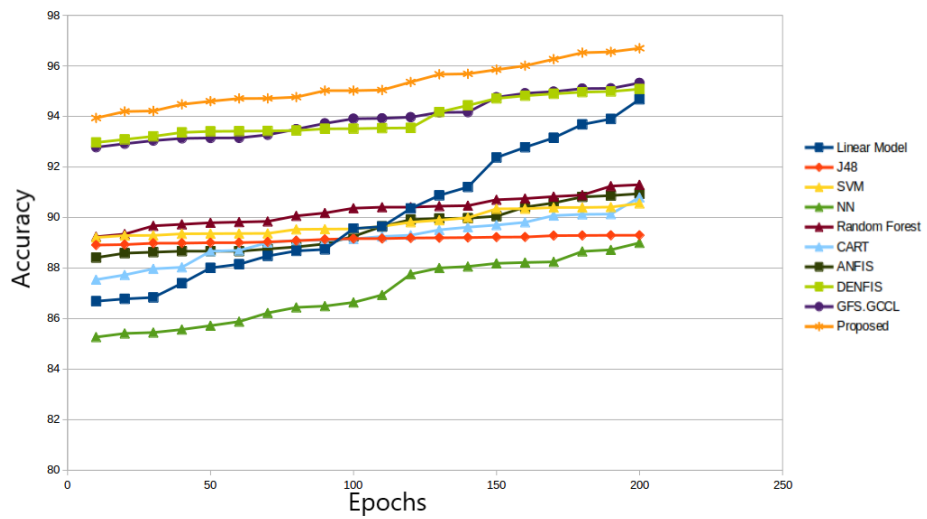
Accordingly, range of accuracy is 0 to 100. The proposed techniques whose accuracy is 100 or approaching to 100 is progressively effective when contrasted



(a) Accuracy Curve For Neural Network



(b) Accuracy Curve For Evolutionary Neural Network



(c) Comparison of all techniques

Figure 5.3: Accuracy Comparison

with different strategies. Here, the value of A_e has been taken = 1.5.

Table 5.1 describes comparative analysis of accuracy of the evolutionary based drug synergy prediction using adaptive lévy based neural network structure. It has been seen that the proposed strategy outperforms others since it has better precision rate contrasted with other machine learning systems. Training and Testing has been done on same synergy data on randomly chosen different segments of data as depicted in Table 5.1. It has been seen that the evolutionary system outflanks other computational methods in light of the fact that the proposed strategy has demonstrated better accuracy contrasted with various procedures. As appeared in Figure 5(a), the accuracy of neural system ends up consistent when maximum accuracy is reached after certain epochs. Since the evolutionary neural system randomly initialize the weights which are optimized iteratively demonstrated as peaks in Figure 5(b) which further leads to the proposed model robust among all other models as shown in Figure 5(c).

5.7.2 Root mean square error analysis(RMSE)

It is defined as the divergence of the predicted values from actual values. In mathematical terms, its a standard deviation of prediction errors or residuals between predicted and actual values.

RMSE of predicted values $\hat{y}_t$ for characteristics t of a regression's dependent attribute y_t is assessed for j diverse forecasts as pursues:

$$RMSE = \sqrt{\frac{\sum_{t=1}^j (\hat{y}_t - y_t)^2}{n}} \quad (5.26)$$

where $\hat{y}_t$ describes the predicted values, Comparative analysis of RMSE has been demonstrated in Table 5.2 among proposed and other existing methods. It delineates that the proposed system performs better than others since it has lesser RMSE contrasted among all. Same data of drug synergy is used for training in Table 5.2. It has been seen that on the off chance that the training data has been increased

Table 5.2: Comparative analysis of RMSE(root mean square error) on different data partitions

	RMSE on different data partitions				
Model	40-60%	50-50%	60-40%	70-30%	80-20%
LM	15.12 $\pm$ 2.9	14.33 $\pm$ 2.7	14.18 $\pm$ 2.8	13.41 $\pm$ 2.9	13.28 $\pm$ 3.1
J48	14.29 $\pm$ 3.2	14.84 $\pm$ 3.2	15.32 $\pm$ 2.9	14.67 $\pm$ 2.8	14.43 $\pm$ 2.6
SVM	13.61 $\pm$ 2.8	13.46 $\pm$ 2.9	13.58 $\pm$ 3.4	14.19 $\pm$ 3.1	13.78 $\pm$ 2.8
NN	15.09 $\pm$ 2.4	15.22 $\pm$ 2.6	15.58 $\pm$ 2.9	15.03 $\pm$ 2.4	14.11 $\pm$ 2.7
RF	13.43 $\pm$ 2.3	13.88 $\pm$ 2.6	13.53 $\pm$ 2.8	14.20 $\pm$ 1.9	14.15 $\pm$ 2.2
CART	14.81 $\pm$ 2.6	13.61 $\pm$ 2.4	14.04 $\pm$ 2.6	13.80 $\pm$ 2.4	14.09 $\pm$ 2.3
ANFIS	13.57 $\pm$ 2.1	15.46 $\pm$ 1.9	13.81 $\pm$ 1.7	13.82 $\pm$ 1.8	13.66 $\pm$ 1.4
DENFIS	12.96 $\pm$ 1.6	12.60 $\pm$ 1.8	12.68 $\pm$ 1.4	12.71 $\pm$ 1.7	12.42 $\pm$ 1.3
GFS.GCCL	13.13 $\pm$ 1.4	12.73 $\pm$ 1.7	12.37 $\pm$ 1.5	12.33 $\pm$ 1.5	11.98 $\pm$ 1.2
Proposed	12.67 $\pm$ 1.3	12.27 $\pm$ 1.7	11.95 $\pm$ 1.1	11.85 $\pm$ 0.9	11.78 $\pm$ 1.2

Table 5.3: Comparative analysis of correlation on different data partitions

	Correlation on different data partitions				
Model	40-60%	50-50%	60-40%	70-30%	80-20%
LM	0.87 $\pm$ 0.09	0.90 $\pm$ 0.07	0.91 $\pm$ 0.07	0.92 $\pm$ 0.05	0.93 $\pm$ 0.04
J48	0.90 $\pm$ 0.07	0.88 $\pm$ 0.09	0.87 $\pm$ 0.08	0.89 $\pm$ 0.07	0.89 $\pm$ 0.08
SVM	0.92 $\pm$ 0.05	0.92 $\pm$ 0.06	0.92 $\pm$ 0.05	0.90 $\pm$ 0.07	0.91 $\pm$ 0.05
NN	0.87 $\pm$ 0.11	0.87 $\pm$ 0.10	0.86 $\pm$ 0.13	0.88 $\pm$ 0.09	0.90 $\pm$ 0.08
RF	0.92 $\pm$ 0.06	0.91 $\pm$ 0.07	0.92 $\pm$ 0.05	0.90 $\pm$ 0.07	0.90 $\pm$ 0.05
CART	0.88 $\pm$ 0.08	0.92 $\pm$ 0.06	0.91 $\pm$ 0.05	0.91 $\pm$ 0.06	0.90 $\pm$ 0.07
ANFIS	0.92 $\pm$ 0.05	0.86 $\pm$ 0.11	0.91 $\pm$ 0.07	0.91 $\pm$ 0.06	0.92 $\pm$ 0.05
DENFIS	0.94 $\pm$ 0.04	0.94 $\pm$ 0.04	0.94 $\pm$ 0.04	0.93 $\pm$ 0.05	0.94 $\pm$ 0.04
GFS.GCCL	0.93 $\pm$ 0.04	0.95 $\pm$ 0.04	0.96 $\pm$ 0.03	0.96 $\pm$ 0.03	0.97 $\pm$ 0.02
Proposed	0.93 $\pm$ 0.05	0.94 $\pm$ 0.03	0.95 $\pm$ 0.04	0.96 $\pm$ 0.01	0.97 $\pm$ 0.02

consistently, and still, at the end, proposed system outperforms others. Along these lines, proposed strategy has increasingly reasonable outcomes contrasted with other machine learning systems.

5.7.3 Correlation analysis

Correlations(*corr*) estimates the linear relationship within 2 characteristics X & Y . $corr \in [-1, 1]$, in which 1 represents positive relationship, -1 shows negative relationship 0 means no correlation. Numerically, *corr* can be processed as pursues:

$$corr = \frac{c_v(X, Y)}{\sigma_X \sigma_Y} \quad (5.27)$$

Here, c_v denotes covariance between X and Y . σ_X and σ_Y shows standard deviation of X and Y .

c_v can be computed as:

$$c_v(X, Y) = \epsilon[(X - \mu_X)(Y - \mu_Y)], \quad (5.28)$$

Here, ϵ represents expectation. μ_X and μ_Y represent mean of X and Y . Now, $corr$ will be:

$$corr = \frac{\epsilon[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (5.29)$$

Therefore, $corr$ is shown with reference to uncentered moments. Since,

$$\mu_X = \epsilon[X] \quad (5.30)$$

$$\mu_Y = \epsilon[Y] \quad (5.31)$$

$$\sigma_X^2 = \epsilon[(X - \epsilon[X])^2] = \epsilon[X^2] - [\epsilon[X]]^2 \quad (5.32)$$

$$\sigma_Y^2 = \epsilon[(Y - \epsilon[Y])^2] = \epsilon[Y^2] - [\epsilon[Y]]^2 \quad (5.33)$$

$$\epsilon[(X - \mu_X)(Y - \mu_Y)] = \epsilon[(X - \epsilon[X])(Y - \epsilon[Y])] = \epsilon[XY] - \epsilon[X] \epsilon[Y], \quad (5.34)$$

Therefore, $corr$ can be rephrased as:

$$corr = \frac{\epsilon[XY] - \epsilon[X] \epsilon[Y]}{\sqrt{\epsilon[X^2] - [\epsilon[X]]^2} \times \sqrt{\epsilon[Y^2] - [\epsilon[Y]]^2}}. \quad (5.35)$$

Table 5.3 demonstrates the coefficient of correlation analysis among proposed technique with other existing techniques. It has been noticed that the evolution-

ary approach has substantial correlation coefficient and more robust outcomes as compared to other computational approaches.

Table 5.4: Comparative Analysis

Cell_Line	Combination_ID	Synergy_Score	Predicted Synergy_Score
BT-20	ADAM17.AKT	29.54039	28.33218805
CAMA-1	AKT.AKT_1	13.01508	12.48276322
BT-549	AKT_1.PIK3C	22.94989	24.07443461
HCC1428	ALK_IGFR.MTOR_1	-5.774149	-5.5379863
DU-4475	ATR_4.MAP2K_1	-13.59996	-14.1561836
A549	BCL2.Carbotaxol	21.7906	22.68183554
CAL-51	BCL2_BCL2L1.FGFR	4.53158	4.346238378
BT-474	BRAF_M2.MAP2K_1	57.90503	60.273345727
647-V	CHEK1.IAP	29.25684	28.060235244
BT-474	Chloroquine.PIK3C	-35.51228	-36.964732252

5.7.4 Comparative Result Analysis

The above result shows that the proposed technique is 95.90% accurate in predicting the synergy score which is higher than any other regression techniques used above for comparison. Table 5.4 has shown the comparative results among actual and predicted synergy score.

Chapter 6

Conclusion and Future Scope

6.1 Conclusion

Drug synergy is commonly utilized in treating the most terrible diseases, such as cancer. The primary significance of drug synergy score prediction is to realize synergistic therapeutic effect, dose and toxicity reduction and to decrease the induction of drug resistance. Current advancements in cancer have discovered that the disease cannot be understood simply through the examination of genetic mutations within cancer cells. Therefore, the examination of the drug synergy score needs well-organized regression models to decrease the prediction errors. Various machine learning techniques have been designed so far for the prediction of drug synergy score. The majority of existing machine learning techniques individually lack the ability to yield noticeable accuracy for drug synergy score prediction, which generally suffer from low accuracy rate. Many ML models used for the prediction of synergy score among the combination of different drugs stuck into the local minima or local maxima and the majority of existing machine learning techniques suffer from the parameter tuning issue, as discussed in chapter 2. In this work, the main focus is to improve the drug synergy score using computational intelligence techniques by eliminating the aforementioned issues. For this reason, three approaches have been proposed.

In the first approach, a neuro-fuzzy based ensembling approach has been designed. Ensembling has been achieved by evaluating the biased weighted aggregation of predicted data by selected models i.e., Random forest, GFS.GCCL, ANFIS and DENFIS based machine learning models. The proposed technique has been evaluated on drug synergy score data and also compared with nine well-known existing machine learning models. It has been assessed that improvement of the proposed technique over other machine learning models in terms of accuracy, root mean square error and correlation coefficient are 4.79%, 2.75% and 3.13% respectively.

In the second approach, an ensemble-based differential evolution has been proposed to tune the attributes of SVM to predict drug synergy score in a more efficient manner. The proposed technique has been evaluated on drug synergy score data and also compared with five well-known existing machine learning techniques. From experimental results, it has been observed that improvement of the proposed method over other machine learning models regarding the accuracy, root mean square error, and coefficient of correlation are 3.93%, 1.70% and 1.70%, respectively.

In the third approach, to overcome the issue of parameter tuning, an efficient evolutionary-based neural network structure using an adaptive lévy technique has been implemented. The existing machine learning techniques and this evolutionary-based technique have been tested on drug synergy score data. The comparative analysis reveals that the evolutionary approach outperforms the existing methods in connection with various performance measures.

The consistency of the prediction of proposed approaches has been validated by performing repeated k-fold cross-validation. Therefore, the proposed techniques are more efficient for designing a real-time drug synergy predictor. The concept of ensembling and metaheuristic techniques have brought in the field of drug synergy by proposing novel approaches using computational intelligence techniques.

The proposed work can help to reduce the sizeable combinational drug screen search space and target to find the best combination of drugs with a valid synergistic score for different cancers. It is an attempt to model a drug synergism mechanism

using computational intelligence techniques. Synergistic and effective combinations of drugs could fill in as the ideal cancer treatment choice.

6.2 Future Scope

The research follows a systematic and iterative approach to analyse and solve the problems. Suggestions for future work are mentioned below. In this thesis, the focus is on the prediction of drug synergy score by using ensemble machine learning and metaheuristic techniques. Proposed work can be used as well as extended to different fields of bioinformatics including drug designing, prediction of chronic diseases, protein structure prediction, prediction of variant classification, protein change in mutation in compliance with the requirements.

In the future, it is required to disengage features of the potential drug combination to comprehend disease-drug interaction in an all-encompassing way for expanding the viability of cancer treatment. Besides, ensemble machine learning techniques can be exploited to upgrade the expectation execution. Our methodologies need validation in heterogeneous settings including the assortment of malignancy patients with unmistakable genomic profiles.

The potency of the drug combinations can be increased by integrating more features of the drug-disease interaction at a different concentration level of drugs or by adding more properties of drugs. The results may be further authenticated by inculcating the dataset feature i.e., Quality Assurance(QA) of the drug along with the synergy score in prediction which can result in more potential drug combinations.

Bibliography

- [1] Rebecca L Siegel, Kimberly D Miller, and Ahmedin Jemal. Cancer statistics, 2016. *CA: a cancer journal for clinicians*, 66(1):7–30, 2016.
- [2] C Pankaj Goswami, Liang Cheng, PS Alexander, A Singal, and L Li. A new drug combinatory effect prediction algorithm on the cancer cell based on gene expression and dose–response curve. *CPT: pharmacometrics & systems pharmacology*, 4(2):80–90, 2015.
- [3] Bao-Hua Zheng. Material procedure quality forecast based on genetic bp neural network. *Modern Physics Letters B*, 31(19-21):1740080, 2017.
- [4] Hyun Woo Jeon, Deog Gon Cho, Jae Kil Park, Kwan Yong Hyun, Si Young Choi, Jong Hui Suh, and Young-Du Kim. Prognostic factors affecting survival of patients with cancer-related pericardial effusion managed by surgery. *World journal of surgical oncology*, 12(1):249, 2014.
- [5] Hongyang Li, Tingyang Li, Daniel Quang, and Yuanfang Guan. Network propagation predicts drug synergy in cancers. *Cancer research*, 78(18):5446–5457, 2018.
- [6] Ali Cuvitoglu, Joseph Xu Zhou, Sui Huang, and Zerrin Isik. Predicting drug synergy for precision medicine using network biology and machine learning. *Journal of bioinformatics and computational biology*, 17(2):1950012–1950012, 2019.
- [7] Julie Fouquier and Mickael Guedj. Analysis of drug combinations: current methodological landscape. *Pharmacology research & perspectives*, 3(3):e00149, 2015.
- [8] Tara M Herrick and Ryan P Million. Tapping the potential of fixed-dose combinations, 2007.
- [9] Ting-Chao Chou. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacological reviews*, 58(3):621–681, 2006.
- [10] Darrell Whitley and Soraya B Rana. Representation, search and genetic algorithms. In *AAAI/IAAI*, pages 497–502, 1997.
- [11] John R Koza and John R Koza. *Genetic programming: on the programming of computers by means of natural selection*, volume 1. MIT press, 1992.

- [12] Alberto Fernández, Francisco José Berlanga, María José del Jesús, and Francisco Herrera. Genetic cooperative-competitive fuzzy rule based learning method using genetic programming for highly imbalanced data-sets. In *IFSA/EUSFLAT Conf.*, pages 42–47. Citeseer, 2009.
- [13] Francisco Jos Berlanga, Maria Jos del Jesus, and Francisco Herrera. A novel genetic cooperative-competitive fuzzy rule based learning method using genetic programming for high dimensional problems. In *2008 3rd International Workshop on Genetic and Evolving Systems*, pages 101–106. IEEE, 2008.
- [14] Chia-Feng Juang, Jiam-Yow Lin, and Chin-Teng Lin. Genetic reinforcement learning through symbiotic evolution for fuzzy controller design. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 30(2):290–302, 2000.
- [15] Ulaş Çaydaş, Ahmet Hasçalık, and Sami Ekici. An adaptive neuro-fuzzy inference system (anfis) model for wire-edm. *Expert Systems with Applications*, 36(3):6135–6139, 2009.
- [16] J-SR Jang. Anfis: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3):665–685, 1993.
- [17] N Kasabov and Q Song. Dynamic evolving neuro-fuzzy inference system (denfis): On-line learning and application for time-series prediction. In *Proceedings of the 6th International Conference on Soft Computing, October 1-4, 2000, Fuzzy Logic Systems Institute, Izuka, Japan*, pages 696–702, 2000.
- [18] Kasabov Nikola and Qun Song. Dynamic evolving neural-fuzzy inference system and its application for time series prediction. 2002.
- [19] Adnan Idris and Asifullah Khan. Churn prediction system for telecom using filter-wrapper and ensemble classification. *The Computer Journal*, 60(3):410–430, 2016.
- [20] Ahmad Taherkhani. Using decision tree classifiers in source code analysis to recognize algorithms: An experiment with sorting algorithms. *The Computer Journal*, 54(11):1845–1860, 2011.
- [21] David M Magerman. Statistical decision-tree models for parsing. In *Proceedings of the 33rd annual meeting on Association for Computational Linguistics*, pages 276–283. Association for Computational Linguistics, 1995.
- [22] Adnan Idris and Asifullah Khan. Churn prediction system for telecom using filter-wrapper and ensemble classification. *The Computer Journal*, 60(3):410–430, 2016.
- [23] Mahesh Pal. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1):217–222, 2005.
- [24] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

- [25] El-Sayed M El-Alfy. Detection of phishing websites based on probabilistic neural networks and k-medoids clustering. *The Computer Journal*, 60(12):1745–1759, 2017.
- [26] Guohua Wu, Xin Shen, Haifeng Li, Huangke Chen, Anping Lin, and Pon-nuthurai N Suganthan. Ensemble of differential evolution variants. *Information Sciences*, 423:172–186, 2018.
- [27] MJ Bayarri, JO Berger, ES Calder, Abani K Patra, E Bruce Pitman, Elaine T Spiller, and Robert L Wolpert. Probabilistic quantification of hazards: a methodology using small ensembles of physics-based simulations and statistical surrogates. *International Journal for Uncertainty Quantification*, 5(4), 2015.
- [28] Timothy I Cannings and Richard J Samworth. Random-projection ensemble classification. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4):959–1035, 2017.
- [29] Jui-Sheng Chou and Julian Pratama Putra Thedja. Metaheuristic optimization within machine learning-based classification system for early warnings related to geotechnical problems. *Automation in Construction*, 68:65–80, 2016.
- [30] Timo Ojala, Matti Pietikainen, and David Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on*, volume 1, pages 582–585. IEEE, 1994.
- [31] M. Dayarathna, Yonggang Wen, and Rui Fan. Data center energy consumption modeling: A survey. *IEEE Communications Surveys Tutorials*, 18(1):732–794, 2016.
- [32] Ning Yin, Wenzhe Ma, Jianfeng Pei, Qi Ouyang, Chao Tang, and Luhua Lai. Synergistic and antagonistic drug combinations depend on network topology. *PloS one*, 9(4):e93960, 2014.
- [33] Gregor Jansen, Anna Y Lee, Elias Epp, Amélie Fredette, Jamie Surprenant, Doreen Marcus, Michelle Scott, Elaine Tan, Tamiko Nishimura, Malcolm Whiteway, et al. Chemogenomic profiling predicts antifungal synergies. *Molecular systems biology*, 5(1), 2009.
- [34] Murat Cokol, Hon Nian Chua, Murat Tasan, Beste Mutlu, Zohar B Weinstein, Yo Suzuki, Mehmet E Nergiz, Michael Costanzo, Anastasia Baryshnikova, Guri Giaever, et al. Systematic exploration of synergistic drug pairs. *Molecular systems biology*, 7(1), 2011.
- [35] Adam C Palmer and Peter K Sorger. Combination cancer therapy can confer benefit via patient-to-patient variability without drug additivity or synergy. *Cell*, 171(7):1678–1691, 2017.
- [36] Kristina Preuer, Richard PI Lewis, Sepp Hochreiter, Andreas Bender, Krishna C Bulusu, and Günter Klambauer. DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics*, 34(9):1538–1546, 2017.

- [37] Ting-Chao Chou. Drug combination studies and their synergy quantification using the chou-talalay method. *Cancer research*, 70(2):440–446, 2010.
- [38] John C Ashton. Drug combination studies and their synergy quantification using the chou-talalay method. *Cancer research*, 75(11):2400–2400, 2015.
- [39] Alexandra M Poos, André Maicher, Anna K Dieckmann, Marcus Oswald, Roland Eils, Martin Kupiec, Brian Luke, and Rainer König. Mixed integer linear programming based machine learning approach identifies regulators of telomerase in yeast. *Nucleic acids research*, 44(10):e93–e93, 2016.
- [40] Karen A Ryall and Aik Choon Tan. Systems biology approaches for advancing the discovery of effective drug combinations. *Journal of cheminformatics*, 7(1):7, 2015.
- [41] Kosalaram Goteti, C Edwin Garner, Lucas Utley, Jing Dai, Susan Ashwell, Demetri T Moustakas, Mithat Gönen, Gary K Schwartz, Steven E Kern, Sonya Zabudoff, et al. Preclinical pharmacokinetic/pharmacodynamic models to predict synergistic effects of co-administered anti-cancer agents. *Cancer chemotherapy and pharmacology*, 66(2):245–254, 2010.
- [42] Johan W Mouton, Michael N Dudley, Otto Cars, Hartmut Derendorf, and George L Drusano. Standardization of pharmacokinetic/pharmacodynamic (pk/pd) terminology for anti-infective drugs: an update. *Journal of antimicrobial Chemotherapy*, 55(5):601–607, 2005.
- [43] Joseph Lehar, Andrew S Krueger, William Avery, Adrian M Heilbut, Lisa M Johansen, E Roydon Price, Richard J Rickles, Glenn F Short Iii, Jane E Staunton, Xiaowei Jin, et al. Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nature biotechnology*, 27(7):659, 2009.
- [44] Feixiong Cheng and Zhongming Zhao. Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *Journal of the American Medical Informatics Association*, 21(e2):e278–e286, 2014.
- [45] Ruolan Chen, Xiangrong Liu, Shuting Jin, Jiawei Lin, and Juan Liu. Machine learning for drug-target interaction prediction. *Molecules*, 23(9):2208, 2018.
- [46] Adam A Margolin, Ilya Nemenman, Katia Basso, Chris Wiggins, Gustavo Stolovitzky, Riccardo Dalla Favera, and Andrea Califano. Aracne: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. In *BMC bioinformatics*, volume 7, page S7. Springer, 2006.
- [47] SM Puddicombe, R Polosa, A Richter, MT Krishna, PH Howarth, ST Holgate, and DE Davies. Involvement of the epidermal growth factor receptor in epithelial repair in asthma. *The FASEB Journal*, 14(10):1362–1374, 2000.
- [48] Scott L Pomeroy, Pablo Tamayo, Michelle Gaasenbeek, Lisa M Sturla, Michael Angelo, Margaret E McLaughlin, John YH Kim, Liliana C Goumnerova, Peter M Black, Ching Lau, et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*, 415(6870):436, 2002.

- [49] Lixin Zhang, Kezhi Yan, Yu Zhang, Ren Huang, Jiang Bian, Chuansen Zheng, Haixiang Sun, Zhihui Chen, Nuo Sun, Rong An, et al. High-throughput synergy screening identifies microbial metabolites as combination agents for the treatment of fungal infections. *Proceedings of the National Academy of Sciences*, 104(11):4606–4611, 2007.
- [50] Zuoqing Yin, Zeliang Deng, Wenyan Zhao, and Zhiwei Cao. Searching synergistic dose combinations for anticancer drugs. *Frontiers in pharmacology*, 9, 2018.
- [51] Shao Li, Bo Zhang, and Ningbo Zhang. Network target for screening synergistic drug combinations with application to traditional chinese medicine. *BMC systems biology*, 5(1):S10, 2011.
- [52] Hildebert Wagner and G Ulrich-Merzenich. Synergy research: approaching a new generation of phytopharmaceuticals. *Phytomedicine*, 16(2-3):97–110, 2009.
- [53] MA Pfaller and AL Barry. Evaluation of a novel colorimetric broth microdilution method for antifungal susceptibility testing of yeast isolates. *Journal of clinical microbiology*, 32(8):1992–1996, 1994.
- [54] Charles R Bonapace, Roger L White, Lawrence V Friedrich, and John A Bosso. Evaluation of antibiotic synergy against acinetobacter baumannii: a comparison with etest, time-kill, and checkerboard methods. *Diagnostic microbiology and infectious disease*, 38(1):43–50, 2000.
- [55] Wei Zhao, Kris Sachsenmeier, Lanju Zhang, Erin Sult, Robert E Hollingsworth, and Harry Yang. A new bliss independence model to analyze drug combination data. *Journal of biomolecular screening*, 19(5):817–821, 2014.
- [56] Roger L White, David S Burgess, Madhavi Manduru, and John A Bosso. Comparison of three different in vitro methods of detecting synergy: time-kill, checkerboard, and e test. *Antimicrobial agents and chemotherapy*, 40(8):1914–1918, 1996.
- [57] GM Eliopoulos, RC Moellering, and V Lorian. Antibiotics in laboratory medicine. by Lorian V., *The Williams & Wilkins Co., Baltimore*, pages 432–492, 1996.
- [58] DW Warnock, EM Johnson, and TR Rogers. Multi-centre evaluation of the etest method for antifungal drug susceptibility testing of candida spp. and cryptococcus neoformans. bsac working party on antifungal chemotherapy. *The Journal of antimicrobial chemotherapy*, 42(3):321–331, 1998.
- [59] MA Pfaller, SA Messer, A Bolmström, FC Odds, and JH Rex. Multisite reproducibility of the etest mic method for antifungal susceptibility testing of yeast isolates. *Journal of Clinical Microbiology*, 34(7):1691–1693, 1996.

- [60] MA Pfaller, SA Messer, and A Bolmström. Evaluation of etest for determining in vitro susceptibility of yeast isolates to amphotericin b. *Diagnostic microbiology and infectious disease*, 32(3):223–227, 1998.
- [61] A Bolmström and U Nordström. Etest for drug combination studies with mycobacteria. In *International Congress of Chemotherapy (ICC)*, 1995.
- [62] MA Pfaller, SA Messer, Å Karlsson, and A Bolmström. Evaluation of the etest method for determining fluconazole susceptibilities of 402 clinical yeast isolates by using three different agar media. *Journal of clinical microbiology*, 36(9):2586–2589, 1998.
- [63] Razan Hamoud, Stefan Zimmermann, Jürgen Reichling, and Michael Wink. Synergistic interactions in two-drug and three-drug combinations (thymol, edta and vancomycin) against multi drug resistant bacteria including e. coli. *Phytomedicine*, 21(4):443–447, 2014.
- [64] Hildebert Wagner and G Ulrich-Merzenich. Synergy research: approaching a new generation of phytopharmaceuticals. *Phytomedicine*, 16(2-3):97–110, 2009.
- [65] Matteo Goldoni and Carolina Johansson. A mathematical approach to study combined effects of toxicants in vitro: evaluation of the bliss independence criterion and the loewe additivity model. *Toxicology in vitro*, 21(5):759–769, 2007.
- [66] Wei Zhao, Kris Sachsenmeier, Lanju Zhang, Erin Sult, Robert E Hollingsworth, and Harry Yang. A new bliss independence model to analyze drug combination data. *Journal of biomolecular screening*, 19(5):817–821, 2014.
- [67] Ronald J Tallarida. Drug combinations: tests and analysis with isoboles. *Current protocols in pharmacology*, 72(1):9–19, 2016.
- [68] Ronald J Tallarida. Revisiting the isobole and related quantitative methods for assessing drug synergism. *Journal of Pharmacology and Experimental Therapeutics*, 342(1):2–8, 2012.
- [69] Gani Orhan, Aysen Bayram, Yasemin Zer, and Iclal Balci. Synergy tests by e test and checkerboard methods of antimicrobial combinations against brucella melitensis. *Journal of clinical microbiology*, 43(1):140–143, 2005.
- [70] Gani Orhan, Aysen Bayram, Yasemin Zer, and Iclal Balci. Synergy tests by e test and checkerboard methods of antimicrobial combinations against brucella melitensis. *Journal of clinical microbiology*, 43(1):140–143, 2005.
- [71] Emilia Cantón, Javier Pemán, Miguel Gobernado, Angel Viudes, and Ana Espinel-Ingroff. Synergistic activities of fluconazole and voriconazole with terbinafine against four candida species determined by checkerboard, time-kill, and etest methods. *Antimicrobial agents and chemotherapy*, 49(4):1593–1596, 2005.

- [72] Michael H Hsieh, M Yu Chen, L Yu Victor, and Joseph W Chow. Synergy assessed by checkerboard a critical analysis. *Diagnostic microbiology and infectious disease*, 16(4):343–349, 1993.
- [73] Charles E Smith, Barbara E Foleno, John F Barrett, and Mary Beth Frosco. Assessment of the synergistic interactions of levofloxacin and ampicillin against enterococcus faecium by the checkerboard agar dilution and time-kill methods. *Diagnostic microbiology and infectious disease*, 27(3):85–92, 1997.
- [74] Michael E Klepser, Erika J Ernst, Russell E Lewis, Michael E Ernst, and Michael A Pfaller. Influence of test conditions on antifungal time-kill curve results: proposal for standardized methods. *Antimicrobial agents and chemotherapy*, 42(5):1207–1212, 1998.
- [75] Ellen E Roling, Michael E Klepser, Ashley Wasson, Russell E Lewis, Erika J Ernst, and Michael A Pfaller. Antifungal activities of fluconazole, caspofungin (mk0991), and anidulafungin (ly 303366) alone and in combination against candida spp. and cryptococcus neoformans via time-kill methods. *Diagnostic microbiology and infectious disease*, 43(1):13–17, 2002.
- [76] Marisa B Marques, Eneida S Brookings, Stephen A Moser, Phillip B Sonke, and Ken B Waites. Comparative in vitro antimicrobial susceptibilities of nosocomial isolates of acinetobacter baumannii and synergistic activities of nine antimicrobial combinations. *Antimicrobial agents and chemotherapy*, 41(5):881–885, 1997.
- [77] Thean Yen Tan, Tze Peng Lim, Winnie Hui Ling Lee, Suranthran Sasikala, Li Yang Hsu, and Andrea Lay-Hoon Kwa. In vitro antibiotic synergy in extensively drug-resistant acinetobacter baumannii: the effect of testing by time-kill, checkerboard, and etest methods. *Antimicrobial agents and chemotherapy*, 55(1):436–438, 2011.
- [78] George Pankey, Deborah Ashcraft, Heather Kahn, and Abdulrahim Ismail. Time-kill assay and etest evaluation for synergy with polymyxin b and fluconazole against candida glabrata. *Antimicrobial agents and chemotherapy*, 58(10):5795–5800, 2014.
- [79] Christopher D Doern. When does 2 plus 2 equal 5? a review of antimicrobial synergy testing. *Journal of clinical microbiology*, 52(12):4124–4128, 2014.
- [80] Tea Pemovska, Johannes W Bigenzahn, and Giulio Superti-Furga. Recent advances in combinatorial drug screening and synergy scoring. *Current opinion in pharmacology*, 42:102–110, 2018.
- [81] Vasileios Stathias, Anna M Jermakowicz, Marie E Maloof, Michele Forlin, Winston Walters, Robert K Suter, Michael A Durante, Sion L Williams, J William Harbour, Claude-Henry Volmar, et al. Drug and disease signature integration identifies synergistic combinations in glioblastoma. *Nature communications*, 9(1):5315, 2018.

- [82] Zura Kakushadze and Willie Yu. Mutation clusters from cancer exome. *Genes*, 8(8):201, 2017.
- [83] Rui Xue Zhang, Ho Lun Wong, Hui Yi Xue, June Young Eoh, and Xiao Yu Wu. Nanomedicine of synergistic drug combinations for cancer therapy—strategies and perspectives. *Journal of Controlled Release*, 240:489–503, 2016.
- [84] Francisco Azuaje. Computational models for predicting drug responses in cancer research. *Briefings in bioinformatics*, 18(5):820–829, 2016.
- [85] Ignacio I Wistuba, Juri G Gelovani, Jörg J Jacoby, Suzanne E Davis, and Roy S Herbst. Methodological and practical challenges for personalized cancer therapies. *Nature reviews Clinical oncology*, 8(3):135, 2011.
- [86] Kavitha Palaniappan and Richard A Holley. Use of natural antimicrobials to increase antibiotic susceptibility of drug resistant bacteria. *International journal of food microbiology*, 140(2-3):164–168, 2010.
- [87] Terry Roemer and Charles Boone. Systems-level antimicrobial drug and drug synergy discovery. *Nature chemical biology*, 9(4):222, 2013.
- [88] Linda Ejim, Maya A Farha, Shannon B Falconer, Jan Wildenhain, Brian K Coombes, Mike Tyers, Eric D Brown, and Gerard D Wright. Combinations of antibiotics and nonantibiotic drugs enhance antimicrobial efficacy. *Nature chemical biology*, 7(6):348, 2011.
- [89] Charles R Bonapace, John A Bosso, Lawrence V Friedrich, and Roger L White. Comparison of methods of interpretation of checkerboard synergy testing. *Diagnostic microbiology and infectious disease*, 44(4):363–366, 2002.
- [90] Juan J Martinez-Irujo, Maria L Villahermosa, Elena Alberdi, and Esteban Santiago. A checkerboard method to evaluate interactions between drugs. *Biochemical pharmacology*, 51(5):635–644, 1996.
- [91] Navya Nair, Olga Camacho-Vanegas, Dmitry Rykunov, Matthew Dashkoff, Sandra Catalina Camacho, Cassie A Schumacher, Jonathan C Irish, Timothy T Harkins, Elijah Freeman, Isaac Garcia, et al. Genomic analysis of uterine lavage fluid detects early endometrial cancers and reveals a prevalent landscape of driver mutations in women without histopathologic evidence of cancer: a prospective cross-sectional study. *PLoS medicine*, 13(12):e1002206, 2016.
- [92] Jouhyun Jeon, Satra Nim, Joan Teyra, Alessandro Datti, Jeffrey L Wrana, Sachdev S Sidhu, Jason Moffat, and Philip M Kim. A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening. *Genome medicine*, 6(7):57, 2014.
- [93] Paul Geeleher, Nancy J Cox, and R Stephanie Huang. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *Genome biology*, 15(3):R47, 2014.

- [94] Justin Guinney, Charles Ferté, Jonathan Dry, Robert McEwen, Gilles Manceau, KJ Kao, Kai-Ming Chang, Claus Bendtsen, Kevin Hudson, Erich Huang, et al. Modeling ras phenotype in colorectal cancer uncovers novel molecular traits of ras dependency and improves prediction of response to targeted agents in patients. *Clinical Cancer Research*, 20(1):265–272, 2014.
- [95] Michael P Menden, Francesco Iorio, Mathew Garnett, Ultan McDermott, Cyril H Benes, Pedro J Ballester, and Julio Saez-Rodriguez. Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. *PLoS one*, 8(4):e61318, 2013.
- [96] Obi L Griffith, Malachi Griffith, et al. Modeling precision treatment of breast cancer. 2013.
- [97] Mark D Pegram, Gottfried E Konecny, Carmina O’Callaghan, Malgorzata Beryt, Richard Pietras, and Dennis J Slamon. Rational combinations of trastuzumab with chemotherapeutic drugs used in the treatment of breast cancer. *Journal of the National Cancer Institute*, 96(10):739–749, 2004.
- [98] Bissan Al-Lazikani, Udai Banerji, and Paul Workman. Combinatorial drug therapy for cancer in the post-genomic era. *Nature biotechnology*, 30(7):679, 2012.
- [99] Mark Pegram, Sheree Hsu, Gail Lewis, Richard Pietras, Malgorzata Beryt, Mark Sliwkowski, Daniel Coombs, Deborah Baly, Fairouz Kabbinavar, and Dennis Slamon. Inhibitory effects of combinations of her-2/neu antibody and chemotherapeutic agents used for treatment of human breast cancers. *Oncogene*, 18(13):2241, 1999.
- [100] Sharmila Shankar, Xufen Chen, and Rakesh K Srivastava. Effects of sequential treatments with chemotherapeutic drugs followed by trail on prostate cancer in vitro and in vivo. *The Prostate*, 62(2):165–186, 2005.
- [101] Mary Ann Jordan and Leslie Wilson. Microtubules as a target for anticancer drugs. *Nature Reviews Cancer*, 4(4):253, 2004.
- [102] Seethalakshmi Sreenivasan and Subramanian Krishnakumar. Synergistic effect of curcumin in combination with anticancer agents in human retinoblastoma cancer cell lines. *Current eye research*, 40(11):1153–1165, 2015.
- [103] Sara Casado-Zapico, Jezabel Rodriguez-Blanco, Guillermo García-Santos, Vanesa Martín, Ana M Sánchez-Sánchez, Isaac Antolín, and Carmen Rodriguez. Synergistic antitumor effect of melatonin with several chemotherapeutic drugs on human ewing sarcoma cancer cells: potentiation of the extrinsic apoptotic pathway. *Journal of pineal research*, 48(1):72–80, 2010.
- [104] Vincent T Devita Jr, Robert C Young, and George P Canellos. Combination versus single agent chemotherapy: a review of the basis for selection of drug treatment of cancer. *Cancer*, 35(1):98–110, 1975.

- [105] Mark D Pegram, Angela Lopez, Gottfried Konecny, and Dennis J Slamon. Trastuzumab and chemotherapeutics: drug interactions and synergies. In *Seminars in oncology*, volume 27, pages 21–5, 2000.
- [106] B Drewinko, TL Loo, B Brown, JA Gottlieb, and EJ Freireich. Combination chemotherapy in vitro with adriamycin. observations of additive, antagonistic, and synergistic effects when used in two-drug combinations on cultured human lymphoma cells. *Cancer biochemistry biophysics*, 1(4):187–195, 1976.
- [107] Wen Zhang, Zhouyi Guo, Deqiu Huang, Zhiming Liu, Xi Guo, and Huiqing Zhong. Synergistic effect of chemo-photothermal therapy using pegylated graphene oxide. *Biomaterials*, 32(33):8555–8561, 2011.
- [108] Han Yuan, Ivan Paskov, Hristo Paskov, Alvaro J González, and Christina S Leslie. Multitask learning improves prediction of cancer drug sensitivity. *Scientific reports*, 6:31619, 2016.
- [109] Jan Wildenhain, Michaela Spitzer, Sonam Dolma, Nick Jarvik, Rachel White, Marcia Roy, Emma Griffiths, David S Bellows, Gerard D Wright, and Mike Tyers. Prediction of synergism from chemical-genetic interactions by machine learning. *Cell systems*, 1(6):383–395, 2015.
- [110] James C Costello, Laura M Heiser, Elisabeth Georgii, Mehmet Gönen, Michael P Menden, Nicholas J Wang, Mukesh Bansal, Petteri Hintsanen, Suleiman A Khan, John-Patrick Mpindi, et al. A community effort to assess and improve drug sensitivity prediction algorithms. *Nature biotechnology*, 32(12):1202, 2014.
- [111] Russ B Altman. Predicting cancer drug response: advancing the dream. *Cancer discovery*, 5(3):237–238, 2015.
- [112] Che-Ming Jack Hu and Liangfang Zhang. Nanoparticle-based combination therapy toward overcoming drug resistance in cancer. *Biochemical pharmacology*, 83(8):1104–1111, 2012.
- [113] Francesca Greco and María J Vicent. Combination therapy: opportunities and challenges for polymer–drug conjugates as anticancer nanomedicines. *Advanced drug delivery reviews*, 61(13):1203–1213, 2009.
- [114] Xiangyi Li, Guangrong Qin, Qingmin Yang, Lanming Chen, and Lu Xie. Biomolecular network-based synergistic drug combination discovery. *BioMed research international*, 2016, 2016.
- [115] Zohar B Weinstein, Andreas Bender, and Murat Cokol. Prediction of synergistic drug combinations. *Current Opinion in Systems Biology*, 4:24–28, 2017.
- [116] Dhanya Sridhar, Shobeir Fakhraei, and Lise Getoor. A probabilistic approach for collective similarity-based drug–drug interaction prediction. *Bioinformatics*, 32(20):3175–3182, 2016.
- [117] Fred E Karch and Louis Lasagna. Adverse drug reactions: a critical review. *Jama*, 234(12):1236–1241, 1975.

- [118] Fred E Karch and Louis Lasagna. Toward the operational identification of adverse drug reactions. *Clinical Pharmacology & Therapeutics*, 21(3):247–254, 1977.
- [119] David N Juurlink, Muhammad Mamdani, Alexander Kopp, Andreas Laupacis, and Donald A Redelmeier. Drug-drug interactions among elderly patients hospitalized for drug toxicity. *Jama*, 289(13):1652–1658, 2003.
- [120] Mark N Prichard and Charles Shipman Jr. A three-dimensional model to analyze drug-drug interactions. *Antiviral research*, 14(4-5):181–205, 1990.
- [121] K Ito, T Iwatsubo, S Kanamitsu, K Ueda, H Suzuki, and Y Sugiyama. Prediction of pharmacokinetic alterations caused by drug-drug interactions: metabolic interaction in the liver. *Pharmacological reviews*, 50(3):387–412, 1998.
- [122] Elena Lopez-Gonzalez, Maria T Herdeiro, and Adolfo Figueiras. Determinants of under-reporting of adverse drug reactions. *Drug safety*, 32(1):19–31, 2009.
- [123] Cecile A Feldman, Martin Giniger, Michael Sanders, Robert Saporito, Harry K Zohn, and Steven P Perlman. Special olympics, special smiles: Assessing the: Feasibility of epidemiologic data collection. *The Journal of the American Dental Association*, 128(12):1687–1696, 1997.
- [124] David S Wishart, Craig Knox, An Chi Guo, Dean Cheng, Savita Shrivastava, Dan Tzur, Bijaya Gautam, and Murtaza Hassanali. Drugbank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic acids research*, 36(suppl_1):D901–D906, 2007.
- [125] Ting-Chao Chou. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacological reviews*, 58(3):621–681, 2006.
- [126] Jonathan B Fitzgerald, Birgit Schoeberl, Ulrik B Nielsen, and Peter K Sorger. Systems biology and combination therapy in the quest for clinical efficacy. *Nature chemical biology*, 2(9):458, 2006.
- [127] Ivana Bozic, Johannes G Reiter, Benjamin Allen, Tibor Antal, Krishnendu Chatterjee, Preya Shah, Yo Sup Moon, Amin Yaqubie, Nicole Kelly, Dung T Le, et al. Evolutionary dynamics of cancer in response to targeted combination therapy. *elife*, 2:e00747, 2013.
- [128] Cláudio A Naranjo, Usoa Busto, Edward M Sellers, P Sandor, I Ruiz, EA Roberts, E Janecek, C Domecq, and DJ Greenblatt. A method for estimating the probability of adverse drug reactions. *Clinical Pharmacology & Therapeutics*, 30(2):239–245, 1981.
- [129] Jouhyun Jeon, Satra Nim, Joan Teyra, Alessandro Datti, Jeffrey L Wrana, Sachdev S Sidhu, Jason Moffat, and Philip M Kim. A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening. *Genome medicine*, 6(7):57, 2014.

- [130] Peter O Krutzik, Janelle M Crane, Matthew R Clutter, and Garry P Nolan. High-content single-cell drug screening with phosphospecific flow cytometry. *Nature chemical biology*, 4(2):132, 2008.
- [131] Christos Hatzis, Philippe L Bedard, Nicolai J Birkbak, Andrew H Beck, Hugo JWL Aerts, David F Stern, Leming Shi, Robert Clarke, John Quackenbush, and Benjamin Haibe-Kains. Enhancing reproducibility in cancer drug screening: how do we move forward? *Cancer research*, 74(15):4016–4023, 2014.
- [132] Jian-Hua Huang, Hua-Lin Xie, Jun Yan, Dong-Sheng Cao, Hong-Mei Lu, Qing-Song Xu, and Yi-Zeng Liang. Correction: Interpretation of type 2 diabetes mellitus relevant gc-ms metabolomics fingerprints by using random forests. *Analytical methods*, 8(8):1950–1951, 2016.
- [133] Bishnupriya Panda and Babita Majhi. A novel improved prediction of protein structural class using deep recurrent neural network. *Evolutionary Intelligence*, pages 1–8, 2018.
- [134] Chun-Wei Tsai, Ming-Chao Chiang, Adlen Ksentini, and Min Chen. Meta-heuristic algorithms for healthcare: Open issues and challenges. *Computers & Electrical Engineering*, 53:421–434, 2016.
- [135] Pablo Mesejo, Oscar Ibáñez, Oscar Cordon, and Stefano Cagnoni. A survey on image segmentation using metaheuristic-based deformable models: state of the art and critical analysis. *Applied Soft Computing*, 44:1–29, 2016.
- [136] Ling Wang, Haoqi Ni, Ruixin Yang, Vijay Pappu, Michael B Fenn, and Panos M Pardalos. Feature selection based on meta-heuristics for biomedicine. *Optimization Methods and Software*, 29(4):703–719, 2014.
- [137] Divya Khanna and Prashant Singh Rana. Multilevel ensemble model for prediction of iga and igg antibodies. *Immunology letters*, 184:51–60, 2017.
- [138] Aleksandar Radosavljevic and Robert P Anderson. Making better maxent models of species distributions: complexity, overfitting and evaluation. *Journal of biogeography*, 41(4):629–643, 2014.
- [139] Jenni AM Sidey-Gibbons and Chris J Sidey-Gibbons. Machine learning in medicine: a practical introduction. *BMC medical research methodology*, 19(1):64, 2019.
- [140] Mehreen Ali and Tero Aittokallio. Machine learning and feature selection for drug response prediction in precision oncology applications. *Biophysical reviews*, 11(1):31–39, 2019.
- [141] Marie Suggitt and Michael C Bibby. 50 years of preclinical anticancer drug screening: empirical to target-driven approaches. *Clinical Cancer Research*, 11(3):971–981, 2005.

- [142] Astrazeneca-sanger drug combination prediction dream challenge. *AstraZeneca-Sanger Drug Combination Prediction DREAM Challenge, Synapse*, (25):235671–235672, 2015.
- [143] Seyed Ali MadaniTonekaboni, Laleh SoltanGhoraie, Venkata Satya Kumar Manem, and Benjamin Haibe-Kains. Predictive approaches for drug combination discovery in cancer. *Briefings in bioinformatics*, page bbw104, 2016.
- [144] Liye He, Evgeny Kuleskiy, Jani Saarela, Laura Turunen, Krister Wennerberg, Tero Aittokallio, and Jing Tang. Methods for high-throughput drug combination screening and synergy scoring. In *Cancer Systems Biology*, pages 351–398. Springer, 2018.
- [145] Aleksandr Ianevski, Liye He, Tero Aittokallio, and Jing Tang. Synergyfinder: a web application for analyzing drug combination dose–response matrix data. *Bioinformatics*, 33(15):2413–2415, 2017.
- [146] Xiangyi Li, Guangrong Qin, Qingmin Yang, Lanming Chen, and Lu Xie. Biomolecular network-based synergistic drug combination discovery. *BioMed research international*, 2016, 2016.
- [147] Jing Tang. Informatics approaches for predicting, understanding, and testing cancer drug combinations. In *Kinase Signaling Networks*, pages 485–506. Springer, 2017.
- [148] Terry Roemer and Charles Boone. Systems-level antimicrobial drug and drug synergy discovery. *Nature chemical biology*, 9(4):222, 2013.
- [149] Bhagwan Yadav, Krister Wennerberg, Tero Aittokallio, and Jing Tang. Searching for drug synergy in complex dose–response landscapes using an interaction potency model. *Computational and structural biotechnology journal*, 13:504–513, 2015.
- [150] Qian Wan and Ranadip Pal. An ensemble based top performing approach for nci-dream drug sensitivity prediction challenge. *PloS one*, 9(6):e101183, 2014.
- [151] Naiqian Zhang, Haiyun Wang, Yun Fang, Jun Wang, Xiaoqi Zheng, and X Shirley Liu. Predicting anticancer drug responses using a dual-layer integrated cell line-drug network model. *PLoS computational biology*, 11(9):e1004498, 2015.
- [152] Turki Turki and Zhi Wei. A link prediction approach to cancer drug sensitivity prediction. *BMC systems biology*, 11(5):94, 2017.
- [153] Mehmet Gönen and Adam A Margolin. Drug susceptibility prediction against a panel of drugs using kernelized bayesian multitask learning. *Bioinformatics*, 30(17):i556–i563, 2014.
- [154] Mehmet Tan. Prediction of anti-cancer drug response by kernelized multi-task learning. *Artificial intelligence in medicine*, 73:70–77, 2016.

- [155] Han Yuan, Ivan Paskov, Hristo Paskov, Alvaro J González, and Christina S Leslie. Multitask learning improves prediction of cancer drug sensitivity. *Scientific reports*, 6:31619, 2016.
- [156] Lin Wang, Xiaozhong Li, Louxin Zhang, and Qiang Gao. Improved anticancer drug response prediction in cell lines using matrix factorization with similarity regularization. *BMC cancer*, 17(1):513, 2017.
- [157] Federica Eduati, Lara M Mangravite, Tao Wang, Hao Tang, J Christopher Bare, Ruili Huang, Thea Norman, Mike Kellen, Michael P Menden, Jichen Yang, et al. Prediction of human population responses to toxic compounds by a collaborative competition. *Nature biotechnology*, 33(9):933, 2015.
- [158] Claudia Canongia. Synergy between competitive intelligence (ci), knowledge management (km) and technological foresight (tf) as a strategic model of prospecting – the use of biotechnology in the development of drugs against breast cancer. *Biotechnology Advances*, 25(1):57 – 74, 2007.
- [159] Yuri A. Kolesnikov, Galina Oksman, and Gavril W. Pasternak. Topical methadone and meperidine analgesic synergy in the mouse. *European Journal of Pharmacology*, 638(1–3):61 – 64, 2010.
- [160] Farzana Kabir Ahmad, Safaai Deris, and Mohd. Syazwan Abdullah. Synergy network based inference for breast cancer metastasis. *Procedia Computer Science*, 3:1094 – 1100, 2011.
- [161] Yong Yang, Zaiqi Zhang, Shuping Li, Xiaoli Ye, Xuegang Li, and Kai He. Synergy effects of herb extracts: Pharmacokinetics and pharmacodynamic basis. *Fitoterapia*, 92:133 – 147, 2014.
- [162] G. Ulrich-Merzenich, D. Panek, H. Zeitler, H. Wagner, and H. Vetter. New perspectives for synergy research with the –omic–technologies. *Phytomedicine*, 16(6–7):495 – 508, 2009.
- [163] Muneeb Ahmed, Marwan Moussa, and S. Nahum Goldberg. Synergy in cancer treatment between liposomal chemotherapeutics and thermal ablation. *Chemistry and Physics of Lipids*, 165(4):424 – 437, 2012.
- [164] Ubaldo Martinez-Outschoorn, Federica Sotgia, and Michael P. Lisanti. Tumor microenvironment and metabolic synergy in breast cancers: Critical importance of mitochondrial fuels and function. *Seminars in Oncology*, 41(2):195 – 216, 2014.
- [165] George Lazaridis, George Pentheroudakis, and Nicholas Pavlidis. Integrating trastuzumab in the neoadjuvant treatment of primary breast cancer: Accumulating evidence of efficacy, synergy and safety. *Critical Reviews in Oncology/Hematology*, 66(1):31 – 41, 2008.
- [166] Jianing Fu, Ning Zhang, Joseph H. Chou, Hua-Jin Dong, Shu-Fu Lin, Gudrun S. Ulrich-Merzenich, and Ting-Chao Chou. Drug combination in vivo using combination index method: Taxotere and {T607} against colon carcinoma hct-116 xenograft tumor in nude mice. *Synergy*, 3(3):15 – 30, 2016.

- [167] Georg E Winter, Uwe Rix, Scott M Carlson, Karoline V Gleixner, Florian Grebien, Manuela Gridling, André C Müller, Florian P Breitwieser, Martin Bilban, Jacques Colinge, et al. Systems-pharmacology dissection of a drug synergy in imatinib-resistant cml. *Nature chemical biology*, 8(11):905, 2012.
- [168] Richard A Lake and Bruce WS Robinson. Immunotherapy and chemotherapy—A practical partnership. *Nature Reviews Cancer*, 5(5):397, 2005.
- [169] Terry Roemer and Charles Boone. Systems-level antimicrobial drug and drug synergy discovery. *Nature chemical biology*, 9(4):222, 2013.
- [170] Xiaolong Jia, Yiyang Li, Alok Sharma, Yulong Li, Guohai Xie, Guoyao Wang, Junhui Jiang, Yue Cheng, and Xianting Ding. Application of sequential factorial design and orthogonal array composite design (oacd) to study combination of 5 prostate cancer drugs. *Computational Biology and Chemistry*, 67:234 – 243, 2017.
- [171] Mikael Larsson, Wei-Ting Huang, Dean-Mo Liu, and Dusan Losic. Local co-administration of gene-silencing {RNA} and drugs in cancer therapy: State-of-the art and therapeutic potential. *Cancer Treatment Reviews*, 55:128 – 135, 2017.
- [172] Konstantina Kourou, Themis P Exarchos, Konstantinos P Exarchos, Michalis V Karamouzis, and Dimitrios I Fotiadis. Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal*, 13:8–17, 2015.
- [173] Eleftheria Kalogera, Debarshi Roy, Ashwani Khurana, Susmita Mondal, Amy L. Weaver, Xiaoping He, Sean C. Dowdy, and Viji Shridhar. Quinacrine in endometrial cancer: Repurposing an old antimalarial drug. *Gynecologic Oncology*, 146(1):187 – 195, 2017.
- [174] J. Brian Cassel, Egidio Del Fabbro, Tobias Arkenau, Irene J. Higginson, Samia Hurst, Lynn A. Jansen, Andrew Poklepovic, Annette Rid, Jordi Rod  n, Florian Strasser, and Franklin G. Miller. Phase i cancer trials and palliative care: Antagonism, irrelevance, or synergy? *Journal of Pain and Symptom Management*, 52(3):437 – 445, 2016.
- [175] C Pankaj Goswami, L Cheng, PS Alexander, A Singal, and Lang Li. A new drug combinatory effect prediction algorithm on the cancer cell based on gene expression and dose–response curve. *CPT: pharmacometrics & systems pharmacology*, 4(2):80–90, 2015.
- [176] Zili Yang. An automatic and accurate method of full heart segmentation from ct image based on linear gradient model. *Modern Physics Letters B*, 31(19-21):1740045, 2017.
- [177] Mahyar Taghizadeh Nouei, Ali Vahidian Kamyad, MahmoodReza Sarzaeem, and Somayeh Ghazalbash. Fuzzy risk assessment of mortality after coronary surgery using combination of adaptive neuro-fuzzy inference system and k-means clustering. *Expert Systems*, 33(3):230–238, 2016.

- [178] Hadi Chahkandi Nejad, Mohsen Farshad, Fereidoun Nowshiravan Rahatabad, and Omid Khayat. Gradient-based back-propagation dynamical iterative learning scheme for the neuro-fuzzy inference system. *Expert Systems*, 33(1):70–76, 2016.
- [179] Keeley A. Crockett, Zuhair Bandar, Jay Fowdar, and James O’Shea. Genetic tuning of fuzzy inference within fuzzy classifier systems. *Expert Systems*, 23(2):63–82, 2006.
- [180] MURAT BEKEN. Artificial neural network prediction for thermal decomposition of potassium nitrate (kno3) and benzoic acid (c6h5cooh). *Modern Physics Letters B*, 24(17):1855–1868, 2010.
- [181] Xin Zhou, Sun Jun, Bing Zhang, and Wu Jun. Classification of different kinds of pesticide residues on lettuce based on fluorescence spectra and wtâĀšbcâĀšsvm algorithm. *Modern Physics Letters B*, 31(19-21):1740082, 2017.
- [182] Yanhua Wang, Zhiyuan Zhang, and Weigang Huo. Research on aviation unsafe incidents classification with improved tf-idf algorithm. *Modern Physics Letters B*, 30(12):1650184, 2016.
- [183] J. L. TANG, C. Z. CAI, and S. ZHAO. Electrical power prediction of proton exchange membrane fuel cell by using support vector regression. *Modern Physics Letters B*, 26(19):1250121, 2012.
- [184] Hui Li, Jie Sun, Ji-Cai Li, and Xiu-Ying Yan. Forecasting business failure using two-stage ensemble of multivariate discriminant analysis and logistic regression. *Expert Systems*, 30(5):385–397, 2013.
- [185] Shuang Yang and Antony Browne. Neural network ensembles: combining multiple models for enhanced performance using a multistage approach. *Expert Systems*, 21(5).
- [186] Frederic Stahl and Max Bramer. Random prism: a noise-tolerant alternative to random forests. *Expert Systems*, 31(5):411–420, 2014.
- [187] Jiang Xie, Yan Liu, Xu Zeng, Wu Zhang, and Zhen Mei. A bayesian network model for predicting type 2 diabetes risk based on electronic health records. *Modern Physics Letters B*, 31(19-21):1740055, 2017.
- [188] Yigit Kultur and Mehmet Ufuk ÅĖaglayan. Hybrid approaches for detecting credit card fraud. *Expert Systems*, 34(2):n/a–n/a, 2017.
- [189] AkÅŹn OzÅĖift. Forward stage-wise ensemble regression algorithm to improve base regressors prediction ability: an empirical study. *Expert Systems*, 31(1):1–8, 2014.
- [190] J-SR Jang. Anfis: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3):665–685, 1993.
- [191] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

- [192] Simon X Yang and Chaomin Luo. A neural network approach to complete coverage path planning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(1):718–724, 2004.
- [193] Harpreet Singh, Prashant Singh Rana, and Urvinder Singh. Prediction of drug synergy score using ensemble based differential evolution. *IET Systems Biology*, 13(1):24–29, 2018.
- [194] H. Juidette and H. Youlal. Fuzzy dynamic path planning using genetic algorithms. *Electronics Letters*, 36(4):374–376, Feb 2000.
- [195] M. Setnes and H. Roubos. Ga-fuzzy modeling and classification: complexity and performance. *IEEE Transactions on Fuzzy Systems*, 8(5):509–522, Oct 2000.
- [196] I. Icke and J. C. Bongard. Improving genetic programming based symbolic regression using deterministic machine learning. pages 1763–1770, June 2013.
- [197] K. Belarbi and F. Titel. Genetic algorithm for the design of a class of fuzzy controllers: an alternative approach. *IEEE Transactions on Fuzzy Systems*, 8(4):398–405, Aug 2000.
- [198] R. Caponetto, L. Fortuna, G. Nunnari, L. Occhipinti, and M. G. Xibilia. Soft computing for greenhouse climate control. *IEEE Transactions on Fuzzy Systems*, 8(6):753–760, Dec 2000.
- [199] M. Li, K. Mehrotra, C. Mohan, and S. Ranka. Sunspot numbers forecasting using neural networks. In *Proceedings. 5th IEEE International Symposium on Intelligent Control 1990*, pages 524–529 vol.1, Sep. 1990.
- [200] Haoran Chen, Yanfeng Sun, Junbin Gao, Yongli Hu, and Baocai Yin. Solving partial least squares regression via manifold optimization approaches. *IEEE transactions on neural networks and learning systems*, (99):1–13, 2018.
- [201] Krishna C. Bulusu, Rajarshi Guha, Daniel J. Mason, Richard P.I. Lewis, Eugene Muratov, Yasaman Kalantar Motamedi, Murat Cokol, and Andreas Bender. Modelling of compound combination effects and applications to efficacy and toxicity: state-of-the-art, challenges and perspectives. *Drug Discovery Today*, 21(2):225 – 238, 2016.
- [202] Kaifang Pang, Ying-Wooi Wan, William T. Choi, Lawrence A. Donehower, Jingchun Sun, Dhruv Pant, and Zhandong Liu. Combinatorial therapy discovery using mixed integer linear programming. *Bioinformatics*, 30(10):1456–1463, 2014.
- [203] C Pankaj Goswami, L Cheng, PS Alexander, A Singal, and L Li. A new drug combinatory effect prediction algorithm on the cancer cell based on gene expression and dose–response curve. *CPT: Pharmacometrics & Systems Pharmacology*, 4(2):80–90, 2015.

- [204] Kristina Preuer, Richard P I Lewis, Sepp Hochreiter, Andreas Bender, Krishna C Bulusu, and GÄijnter Klambauer. Deepsynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics*, 34(9):1538–1546, 2018.
- [205] Shangce Gao, MengChu Zhou, Yirui Wang, JiuJun Cheng, Hanaki Yachi, and Jiahai Wang. Dendritic neuron model with effective learning algorithms for classification, approximation, and prediction. *IEEE transactions on neural networks and learning systems*, (99):1–14, 2018.
- [206] Hinton Geoffrey E. Williams Ronald J. Rumelhart, David E. Learning representations by back-propagating errors. *Nature*, pages 323–533, 1986.
- [207] Pradnya A Vikhar. Evolutionary algorithms: A critical review and its future prospects. In *2016 International conference on global trends in signal processing, information computing and communication (ICGTSPICC)*, pages 261–265. IEEE, 2016.
- [208] S Kwong, KF Man, and KS Tang. Genetic algorithms: Concepts and applications. *IEEE Transactions on Industrial Electronics*, 43(5):519–534, 1996.
- [209] Shifei Ding, Hui Li, Chunyang Su, Junzhao Yu, and Fengxiang Jin. Evolutionary artificial neural networks: a review. *Artificial Intelligence Review*, 39(3):251–260, Mar 2013.
- [210] Chang-Yong Lee and Xin Yao. Evolutionary algorithms with adaptive levy mutations. 1:568–575, 2001.
- [211] UrvinderSingh RohitSalgotra. Application of mutation operators to flower pollination algorithm. *Expert Systems with Applications*, 79:112–129, 2017.
- [212] Jia Jia, Feng Zhu, Xiaohua Ma, Zhiwei W Cao, Yixue X Li, and Yu Zong Chen. Mechanisms of drug combinations: interaction and network perspectives. *Nature reviews Drug discovery*, 8(2):111, 2009.
- [213] Nori Geary. Understanding synergy. *American Journal of Physiology-Endocrinology and Metabolism*, 304(3):E237–E253, 2012.
- [214] Harpreet Singh, Prashant Singh Rana, and Urvinder Singh. Prediction of drug synergy in cancer using ensemble-based machine learning techniques. *Modern Physics Letters B*, 32(11):1850132, 2018.

List of Publications

1. Harpreet Singh, Prashant Singh Rana and Urvinder Singh, "*Prediction of drug synergy in cancer using ensemble-based machine learning techniques*", Modern Physics Letters B, World Scientific, 32(11):1-12, 2018. [Impact Factor 0.731]
2. Harpreet Singh, Prashant Singh Rana and Urvinder Singh, "*Prediction of drug synergy score using ensemble based differential evolution*", IET Systems Biology, IET, 13(1):24-29, 2018. [Impact Factor 1.392]
3. Harpreet Singh, Prashant Singh Rana and Urvinder Singh, "*Drug Synergy prediction using Adaptive Lévy based Neural Network Structure*", IEEE Transactions on Evolutionary Computation. [Impact Factor 8.508, Communicated]