Determination of cell and drug line sensitive activity with a novel ensemble learning

Dr.P.Selvi Rajendran[1],
Professor, Computer Science and Engineering,
School of Computing,
Hindustan Institute of Technology and Science, Chennai, India
K.R.Kartheeswari[2]
Senior Research Fellow, Computer Science and Engineering,
School of Computing,
Hindustan Institute of Technology and Science, Chennai, India

ABSTRACT:

Despite the fact that cancer has been treated using a variety of surgical and therapeutic techniques such as chemotherapy, target therapy, immunotherapy, and hormone therapy, understanding cancer cell biology and cancer metastatic mechanisms are critical. According to recent studies, up to 30–40% of cancers can be avoided simply by changing one's diet and lifestyle. Diet and nutritional factors are vital in the prevention of a variety of diseases, and they have a substantial impact on patients' disease outcomes both during and after therapy. In this, mechanical response of food particles play a major role to modulate the muscle activity and the dielectric properties of lossy materials are affected by frequency, temperature, and material composition. Hence a unique, network machine learning approach is used for finding mechanical and electrical features that are proposed, that intern used for cancer-fighting. The groundwork models that planned to employ stacked ensemble as learning methods. Elastic Net, pairwise support vector regression, Kernelized Bayesian Multitask Learning and Neural networks are used in this prediction. In terms of sensitivity and activity, the two signatures indicated cell lines and drugs. A sensitivity signature shows the changes in gene expression that cause cell death in a particular cell line. During the action, mechanical, electrical properties and expressions are taken to suppress the cancer in the sensitive line. To improve the fitness analysis in the drug response, cost function is calculated in this research.

Keywords: Anti-cancer, Prediction Model, Stacked Ensemble Neural networks, Machine learning

I.INTRODUCTION

The world's second most-scared disease is cancer as per the report obtained by statistical analysis of the United States National Cancer Institute. In that report, it was stated that in 2020, 1.8 million new cancer cases are projected, with 600,000 fatalities. [1]. In addition, lung cancer is expected to be the second most lethal malignancy. The analysis of the United States specified that estimated new cases were 2,28,820 in the year 2020. In the same year, predictable death cases were 1,35720. So, 25% of the population in the United States were expected to be affected by lung cancer in 2020. A sustained five years virtual survey report stated that in a lifetime, 1 out of 15 in men, 1 out of 17 women, and smokers had a high-risk factor for lung cancer. This five years report indicated that 20% of lung cancer patients were badly affected [2The utilization of antiePD-1 therapy has identified experimental aspects of the treatment response in those with NSCLC. In clinical practice, immunotherapy, smoking history, performance status, sex, the presence of meta states, mutations, and pathology [6e14] are all frequently assessed. These are recorded in electronic health records (EMRs). No clinical factor can accurately predict the response of antiePD-1 therapy (programmed cell death protein-1). To improve the response of antiePD-1 models, integrating factors are needed. Machine learning plays a vital role in disease progression and treatment response [3,4]. In the early stages, an insignificant amount of molecule inhibitors produced hopeful results in drug therapy. So, the lifetime of the patient is also increased. But unexpected secondary mutation(s) restricted the effectiveness of the drug due to drug resistivity [5]. Many computational studies have been directed in molecular dynamics to crack the drug resistivity simulations [6,7]. Mechanical properties play a vital role and are strictly interrelated to human health conditions.

The physiological function of the cells depends upon the mechanical property. The changes in mechanical property illustrate the damage in the physiological functions of the cell. Therefore, it yields to an increase in the disease state [8,9].

II. LITERATURE SURVEY

The Drug Target Integration (DTI) is analysed by Ding et. al [13], through a similarity-based machine learning algorithm. In this work, chemical and genomic spaces are used in the similarity-based analysis. In this approach, similarity-based analysis easily merged with Kernel-based learning methods. In [14], Lin Zhang et. al examined that with the help of cell and drug line similarities, we were able to determine the drug's effective response. To make this prediction we used collaborative filtering models, and from the average of the projected values predicted sensitivity was calculated. The Hyper Interpolation Weighted Collaborative Filter (HIWCF) is based on set of known drug responses. The performance of the system highly dependent on the sparsity of the response matrix. This resulted in a reduction of performance in the drug. The result of PCC was higher than that of other integrated cell line. Chuanying Liu[15] elaborated a cross validation with 10 fold .For the GDSC dataset, this model achieved a better drug response. From the analysis, they found that the model was more biologically interpretable, and feature selection possibility was lagging in the analysis. In Yitan Zhu et. al [16] approach, power transfer learning has been used to predict both new tumour cells and new drugs. This analysis stated that this new powerful learning improved the prediction performance in the drug response prediction applications. From the analysis, it was observed that a single drug response was not available, and a combination of drug responses reduced the performance. Chen *et al.* [17] elaborated DTI prediction machine learning methods. In this technique, chemogenomic databases were used. Negative samples were also handled. A chemogenomic method is categorized into supervised and semi-supervised learning methods. This supervised learning is further subdivided into similarity and feature-based methods. In this research [18,32,33], the author Guosheng Lianga et al. has shown that Artificial Intelligence plays a significant role in the discovery of new materials and accelerating anticancer drug development. These services help the doctors to make the right treatment decisions and reduce unnecessary surgery. Machine learning analysis was used to predict the sensitivity of the drug. Sachdev *et al.* [19,33]. The author examined that DTI prediction using feature-based chemogenomic approaches was described. Based on the results of this survey,these are classified as SVMs and ensemble-based methods. Here, the similarity approach is excluded by the chemogenomic approach. In this work, Raghvi Bhardwaj et. al [20,32,34] developed a double ensemble machine learning algorithm to predict the pathological response after neoadjuvant chemotherapy. This double ensemble was used to predict multi-criteria decision - making. 99.08% accuracy is obtained using the k-fold cross validation technique. In [21] Aman Sharma et al. proposed an ensemble and multitask learning framework for predicting drug sensitivity, which implied that the mean square error(MSE) on a CGP data set was 3.28. In CCLE data set, the MSE was 0.49 and 0.54 for NCI-DREAM. Mehmet Tan, IN [22], proposed an ensemble method for drug response prediction. Drug-induced gene expression has been added to cytotoxicity databases in large quantities. Vitro experiments, in addition to data set testing, are used to develop the prediction. Fanfang Xia et al.[23,29,20,31,35] employed five publicly available cell-line-based data sets in this study. This crossed study is based on the machine learning models. The Best cross-study was achieved through deep neural networks. In this research initially stated that CTRP yields a better prediction on the test data. GCSI was the most predictable data set among all other cell line data set. In this cross study, GCSI combined with CTRP provide a more accurate prediction. The research of M. B. Senousyet et al. [24], provided a micro-array technology for cancer prediction. This micro-array can overcome the missing values or imbalanced bio-medical data problem. Ensemble Learning based on Ranking Attribute Value (ELBRAV) has been used for the research. Aman Sharma et al. [25] established an anti-cancer drug response prediction model using a similarity-based regularization matrix. In this similarity analysis, drug and tissue were the main objective and GDSC and CCLE data sets were used. The average MSE was 3.24 and 0.504.

III. CURRENT CHALLENGES

Lot of more bio-agents and chemicals are waiting for pharmacological evaluation and mechanism investigation following a significant promotion of chemistry and genetic engineering approaches. The medicine's worth and efficiency promotion may be important for anti-tumour drug development as large quantities of chemicals, bio-agents, and herbal medications are created. Past conventions must be reshuffled or perhaps broken down in the future to upgrade the drug development system. Key biological or pathological reasons that lead to clinical therapeutic failures, such as neoplasm metastasis (unpredictable nature), the enigma of cancer stem cells, and drug resistance (after long-term drug utility and tumour evolution), have remained

unresolved till now [36-40]. In order to increase evaluation outcomes in this field, several platforms and techniques must be updated.

Tubulin:

Medicines that interfere with tubulin are grouped together in this categorization because tubulin is involved in cell shape maintenance, intracellular transport, and mitosis. The vinca alkaloids bind to certain tubulin sites and prevent tubulin dimers from polymerizing, blocking microtubule formation. The taxanes bind to microtubules in diverse ways and stabilise them, limiting normal microtubule network rearrangement. Oral taxanes will be more convenient if they prove to be as effective as the parent drugs [41,71]. Tubulin stabilisers known as epothilones are a novel type of stabiliser. Although preclinical research has shown that this type of medications have potential, the results of phase II ,phase III clinical trials are currently unavailable [42,43,71].

IV. STAGES OF CANCER

The size of a tumour and how far its cells have spread are determined by the cancer stage. A doctor can utilize a variety of approaches to assess the stage of cancer.They can employ the 0–5 scalability system.

Stage 0: There are abnormal cells present, but they have not spread to neighboring cells. Cancer is found in stages 1, 2, and 3. Larger tumours and more extensive dissemination into adjacent tissues are associated with higher stages. The cancer has spread to other parts of the body at this stage.

V. NEED OF FOCUSING METABOLIC PATHWAYS IN CANCER RESEARCH

To present, cancer research has focused on two key metabolic pathways: glucose metabolism via glycolysis and glutamine metabolism via the Krebs (TCA) cycle. Because glucose is readily taken up by many cancers, it has been a major focus, as indicated by the use of glucose as a tracer in positron-emission tomography (PET) studies. Despite this, researchers are still unsure about the relevance of each pathway in cancer pathogenesis.[44]

VI. MECHANISMS USED IN CANCER THERAPIES

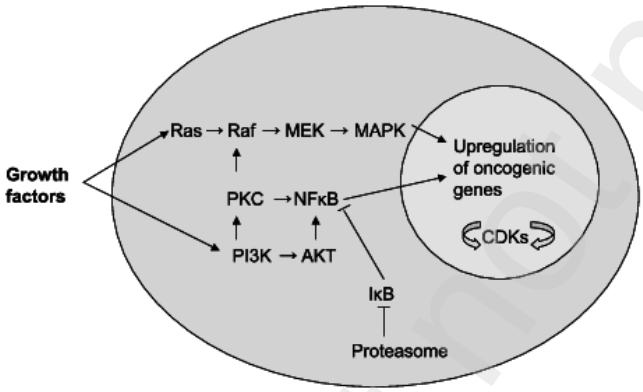


Fig 1. Pathways of metabolism

Figure 1 depicts a diagram of the mechanisms employed in cancer therapy. The most well-known drug in this class is imatinib, which inhibits the tyrosine kinases bcr/abl and c-kit(40,41,71). In chronic myeloid leukaemia and gastrointestinal stromal tumours, it is one of the most effective medications. Other medications target the ras, phosphatidylinositol, proteasome and cyclin-dependent kinases pathways, as well as the proteasome and cyclin-dependent kinases. These medications are currently in the early stages of clinical studies with a few exceptions. Farnesyl transferase activates Ras. Raf and MEK are active when the ras protein is activated. Farnesyl-transferase inhibators, such as lonafarnib and R115,777, operate as fake metabolites of this enzyme [42,43,71]. Raf (BAY 43-9006) and MEK(CI-1040) inhibators are also available[45,46,71].

VII. EFFECTS OF ANTI-CANCER DRUGS ON THE METABOLISM

At various levels, anticancer drugs can affect cancer cells, endothelium, extracellular matrix, immune system, or host cells. The tumour cell's DNA, RNA, or protein levels can all be targeted. Traditional chemotherapy medications mostly target tumour DNA, whereas monoclonal antibodies and nanoparticles primarily target proteins. Specific antibodies and tiny compounds may potentially influence the endothelium and extracellular matrix. [47,71]

A. Anticancer Drugs: A Classical Classification:

I. Chemotherapy

It is an effective treatment that comprises the use of chemicals to remove cancer cells. It is used to avoid cell multiplication and division.[47]

The following are the chemotherapy therapy drugs

1. Alkylators 2Antibiotics 3. Mitosis inhibitors 4. Antimetabolites 5. Topisomerases inhibitors

II. Hormone replacement therapy

It is a type of treatment that incorporates female hormones. Hormone therapy is most commonly used to address common menopausal symptoms such as hot flashes and vaginitis.

Hormone replacement therapies are

1. Anti-aromatase agents 2. Steroids 3. Anti-estrogens 4. LH-RH analogues 5. Anti-androgens

III. Immunotherapy

The treatment of disease by stimulating or inhibiting unaffected cells is known as biological therapy. Immunotherapy also called as biological therapy. It is a process of treating cancer by stimulating else inhibiting the immune system. Activation immunotherapies enhance an immune response, whereas suppression immunotherapies reduce or inhibit the immune response. The vaccines are as follows: 1. Interferon 2. Interleukin 2 (IL-2) 3.Vaccines.

A drug that targets the DNA of tumours:

Breaking the twofold helix, interfering with DNA-related proteins, or affecting gene expression are all possible effects of the drugs. The majority of traditional anticancer medicines rely on one of these processes, and new drugs are released all time[71].

DNA helix:

The first compounds revealed to be beneficial in the treatment of cancer were alkylating agents. The most common site of alkylation is at N-7 position of guanine, but this varies depending on the drug family. Alkylators include nitrogen mustards, nitrosoureas, triazenes, platinum compounds and antibiotics[71].

Drugs that Activate the Endothelium and Extracellular Matrix:

Compounds that target the endothelium inhibit endothelial growth factors or their receptors. The majority of extracellular matrix drugs, on the other hand, block metalloproteinases (MMPs). Antiangiogenic characteristics are present in all of them[71].

Endothelium:

The principal endothelial growth factors, vascular endothelial growth factor (VFGF) and basic fibroblast factor(bFGF) are inhibited by thalidomide(48,49). Another VEGF inhibitor is carboxyamido-trizole [50,51,71]. In tumour cells, interferon also reduces the formation of VEGF, although this effect appears to be medicated by interferon gamma [52,53,54,71]. As a result, one of COX-2 inhibitors possible modes of action is to promote endothelium growth [55,56,71].

Bevacizumab, a VEGF receptor-binding monoclonal antibody, binds to all of them [57,58,71]. SU-5416 is a small molecule that interacts to the tyrosine kinases VEGFR-1 and VEGFR-2 [48,59,71]. C-kit and platelet-derived growth factor receptor are two more proteins with which it interacts. Clinical studies of SU-5416 in haematological malignancies and colorectal cancer have commenced. SU-6668 is a small molecule that binds to the VEGFR, bFGFR and PDGFR(platelet-derived growth factor receptor) proteins.Finally, combretastatin suppresses the mitotic spindle of the endothelium, resulting in apoptosis [60,61,71].

Extracellular matrix:

MMP activation increases invasion in tumours and in a crucial stage in angiogenesis. MMPs can also cause the release of VEGE, bFGF and insulin growth factor. Clinical trails are now being conducted on a number of MMP inhibitors [63,71]. Marimastat [64-66,71], prinomastat and BAY 12-9566 are among the most common enzyme activity inhibitors[63,71].MMP production is suppressed, activation is blocked and MMP breakdown is promoted by tetracycline derivatives like neovastat [67,68,71].In addition to MMPs, other extracellular matrix elements such integrin, endothelin and thrombospondin[69,70,71] could be targeted as anti-cancer therapy.

HOST CELL INHIBITORS AND OTHER DRUGS:

Some drugs are designed to attack organs that may harbour cancer cells. Bisphosphonates [40,41] osteoprotogerin [37] and PTHRP antibodies are now the only substances that inhibit bone cell activity and the micro environment. More drugs to target other organs at risk of metastasis could be developed in the future. Finally, cytokines such as interferon and interleukin2 are well-known for increasing the immune system's anticancer activity[71].

The goal of this paper is to present state-of-art machine learning methods for anti-cancer drug response modelling and prediction, with the goal of better utilizing high-dimensional multi-omics profiles, as well as knowledge of cancer pathways targeted by anti-cancer compounds, when predicting phenotypic responses.

2.3. Importance of the inclusion of Mechanical and Electrical Property Analysis in the Drug Response:

The relevance of drug-polymer interactions is revealed by the mechanical properties of drug-incorporated fibres, and a mechanical model was used to examine drug partition in mix fibers. Measurements of cancer cell's mechanical characteristics give new information that Cancer cells are softer than non-transformed cells.

The effective stiffness and time-dependent deformation characteristics of the cell, the extend of stretching or contraction it undergoes during mechanical testing and the force ranges typically achieved during such deformation all influence the type of biophysical assay used to investigate a specific cell type. The relevant force ranges and length(displacement) scales in selected cell and molecular processes of relevance in biological systems are depicted in Figures 2a and 2b. The force and displacement ranges achieved by various biomechanical assays are shown in these diagrams.

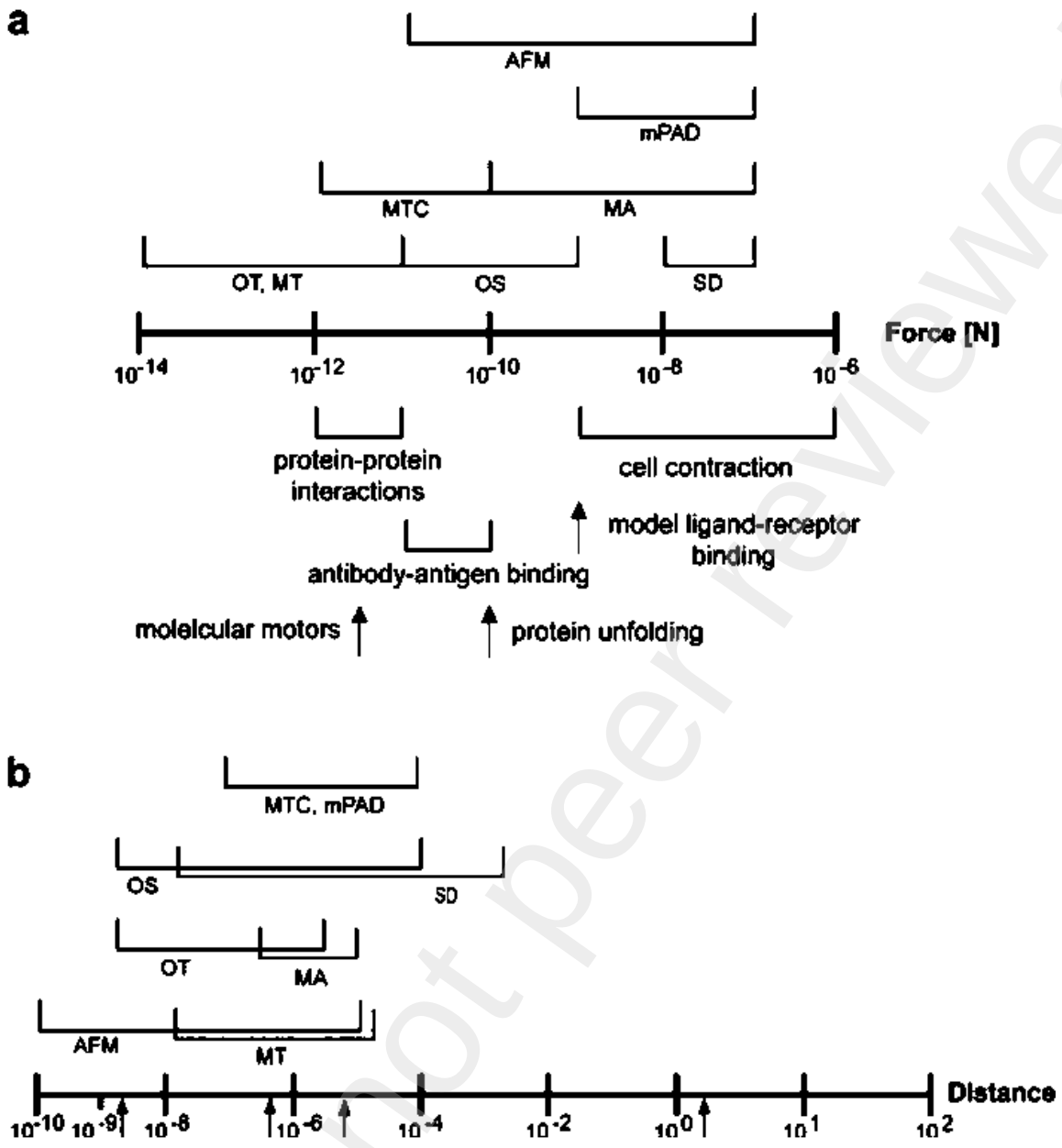


Fig.2. An overview of the forces and displacements investigated by various biomechanical arrays

MTC-Medullary thyroid carcinoma

AFM-Atomic force microscopy

OS-Overall survival

OT-Occupational Therapy

Mpad-Microfabricated Post-Array-Dectectors

Mechanical Factors facilitate the spread of malignancy:

Cell behavior is powerfully inclined by the mechanical influences of the tumour micro environment. Carcinogenesis & cancer are influenced by intracellular signaling processes, which can be influenced by mechanics, progression, as well as the tumour's reaction to treatment. The behavior of normal mammary glands is an example of the effect of the mechanical microenvironment on tumorigenesis.

Potential and kinetic energy are the two forms of energy. The position of an object determines its potential energy, whereas its mobility determines its kinetic energy.

Kinetic and Dynamic:

Drug kinetics (pharmacokinetics) describes how the body processes a drug and accounts for absorption, distribution, metabolism, and elimination.

METHODS AND MATERIALS

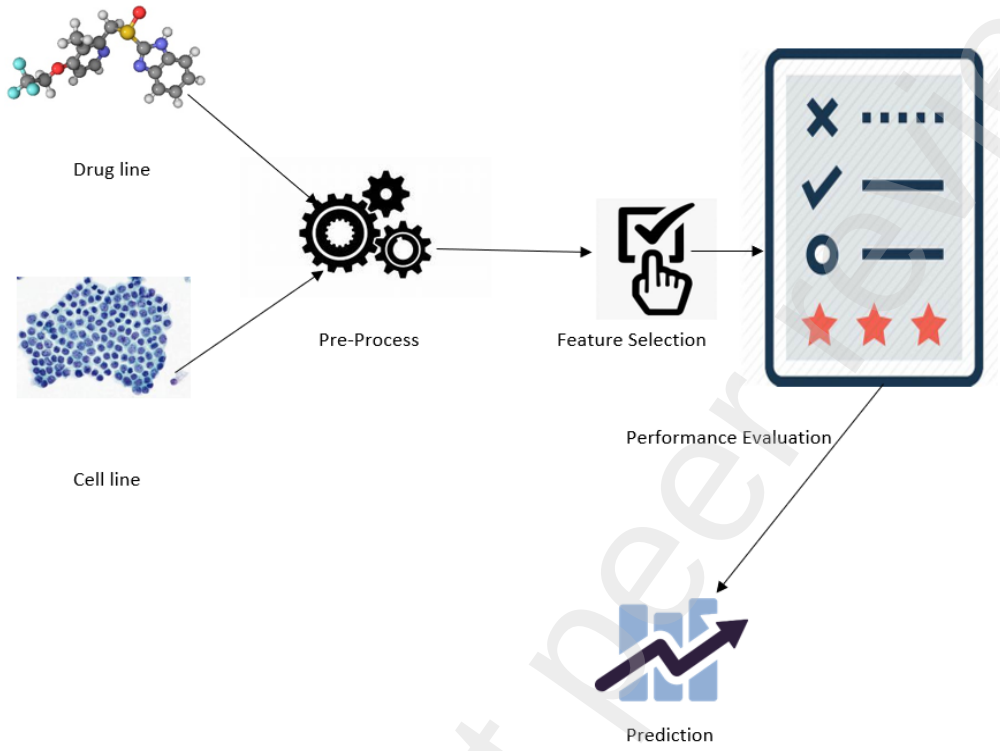
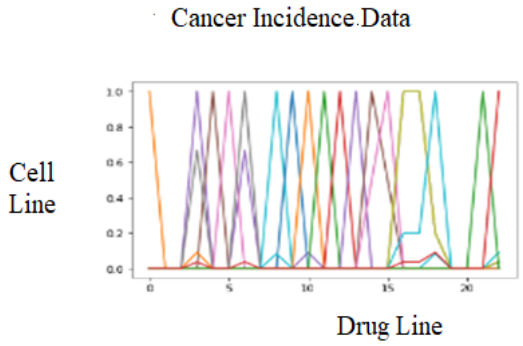


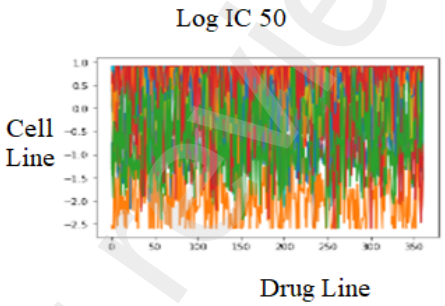
Fig 3. Performance evaluation procedure

A. Dataset:

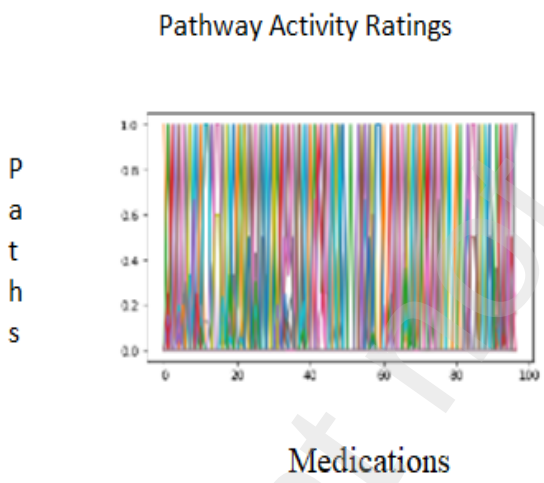
In this study, the cancer statistics are derived from both the properties of the drugs and cell lines, as well as an ensemble voting method applied to the end results of the study. This section contains statistics on anti-cancer drug response prediction in order to reduce the cancer burden. Cancer incidence data were gathered from the CCLE and GDSC datasets for this study. In the cell line, this data predicts the patient's drug response. To protect the information, they are distinguished and encrypted using a set of pre-defined codes.



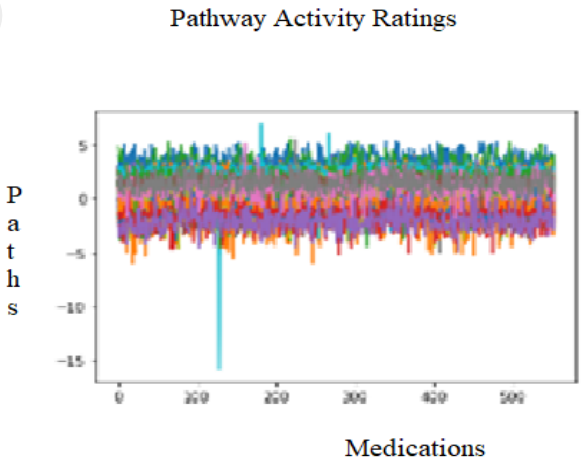
a) CCLE dataset



b) A.Log IC50(CCLE)



c) GDSC Dataset



(d) Log IC50 (GDSC)

FIG 4. a, b, c, d. CCLE, GDSC drug and cell line feature-based dataset

In fig 4., Pathway activity ratings and pharmacological reactions are similar. The medications are arranged in rows, while the paths are arranged in columns. The positive and negative values represent their positive and negative similarity.

B. Preprocessing the Data:

The original compound data is too large for a limited device to handle, particularly in terms of the number of unlabeled compounds, preventing us from building a well-trained machine. With the primary goal of improving the prediction accuracy of supervised learning, we reduced the unlabeled data from thousands to tens of thousands. We discovered that unlabeled data had no effect on the accuracy of the final operation by running tests on reducing data.

Inhibitory Concentration (IC50) Data after Preprocess

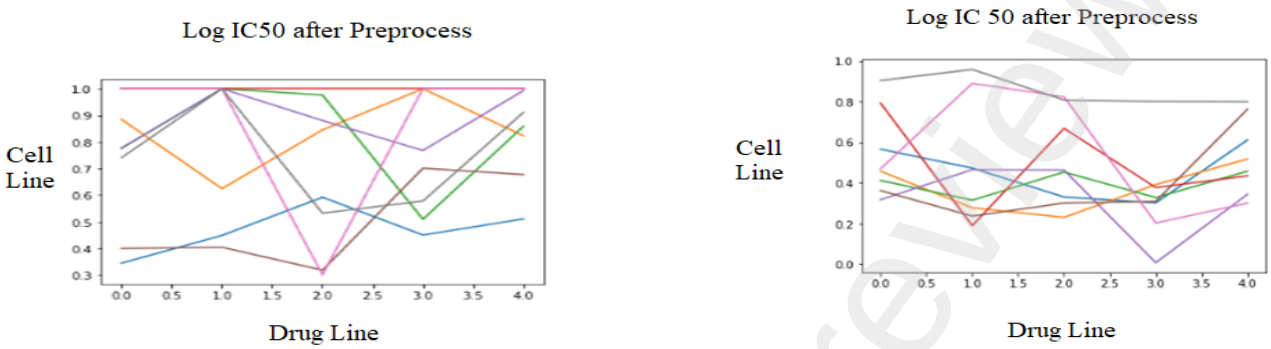


Fig 5. Log IC50 dataset (CCLE and GDSC) after Preprocess

B.Similarity:

CSS-Cosine similarity signature

DL=Drug Line

GenExp → Genetic Expression

TarPro → Target Protein

LIC50 → Log IC50

$$CSS\begin{pmatrix} DL, CL \\ \end{pmatrix} = \begin{cases} \text{GenExp CL=Cell Line} & (CL)\sum CCLE,GDSC \\ \text{TarPro DL=Drug Line} & (DL)\sum CCLE,GDSC \\ \text{LIC50 (Response)} & \\ \sum CCLE, GDSC & \end{cases}$$

1.Cosine Similarity:

In a dataset, the cosine similarity can be used to locate groups of related things.Similarity is measured in machine learning by a distance with dimension representing attributes of objects. When the distance between the feature is small, they are highly comparable, and vice versa.

If two vector features x and y in The dataset,then the similarity index is

Similarity(x,y)=cos(Q)=x.y/|x||y|

Advantages

- Greater the similarity, the smaller the angle.

Fig 6 describes the two vector features x and y of CCLE and GDSC dataset.

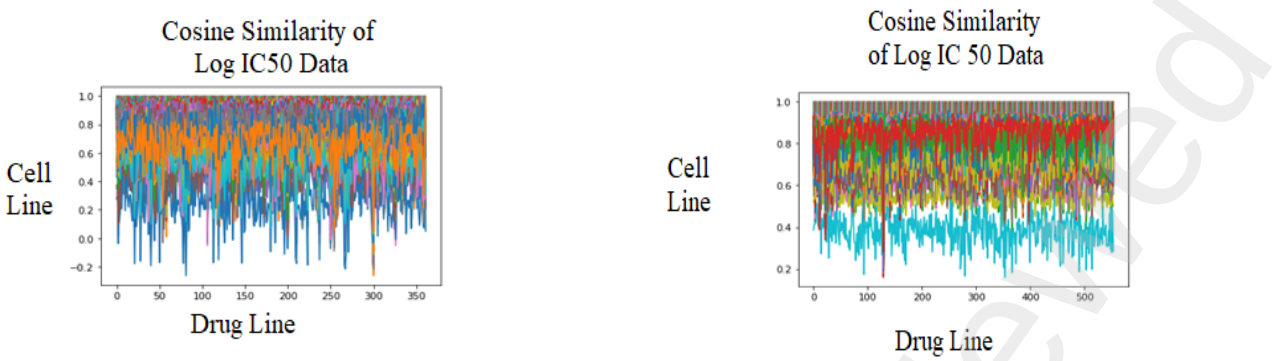


Fig.6.GDSC cosine similarity and CCLE cosine similarity(LogIC50)

It is used to compute the cosine distance between 1-D arrays.

```
from scipy import spatial
List1 =np.array([[1. ],
       [1. ],
       [1.2],
       [1.4],
       [3.4],
       [2.2],
       [ 1.1],
       [ 1.2]])

List2=np.array([[ 5],
       [ 7],
       [ 9],
       [11],
       [13],
       [0],
       [0],
       [0]])
result = 1 - spatial.distance.cosine(List2, List1)
print(result)
```

```
0.7899648626818542
```

From the above analysis it is observed the similarity index is more than 75%. Correlation of pathway activity scores and drug responses. Rows of medications are displayed, whereas columns of pathways are displayed. The positive and negative values represent the positive and negative correlation between the two variables.

The variables are travelling in the same direction if the correlation is positive. In other words, as one variable rises, the other falls, and vice versa. The variables are moving in opposite directions when there is a negative correlation.

2. Procedure for cosine analysis:

Step 1: import necessary library first

Step2: read.csv

Step3: import library cosine similarity

Step4: def cossim (x, y):
numpy. inner (x, y)/ (numpy. linalg. norm(x)*numpy. linalg. norm(y))
Step 5: Print
Step 6: end

C.PCA analysis:

PCA is a method of taking data from high-dimensional spaces and projecting it into lower-dimensional subspaces. It tries to maintain the data's key bits with the most variation and delete the non-essential parts with the least variation. Data dimensions are simply features.

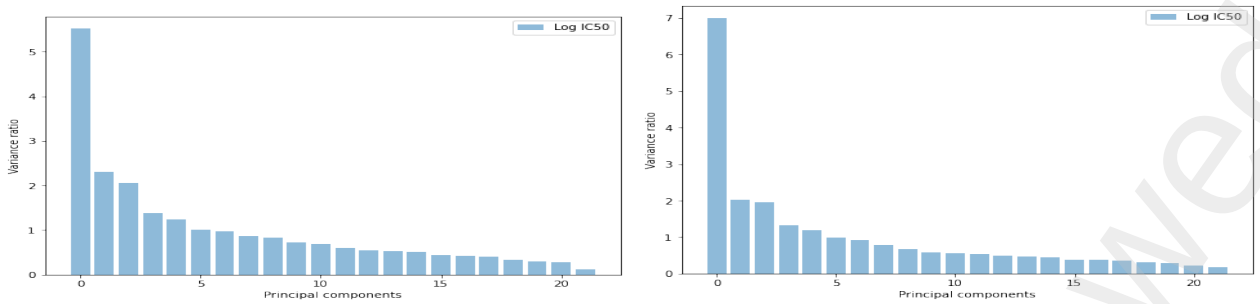


Fig7.PCA of Log IC50 for CCLE and GDSC Dataset

The Dimensionality reduction in Log IC50 is illustrated in Fig. The x axis considers the principal components, whereas the y axis considers the variance ratio.

1.Procedure for PCA Analysis:

Step 1: Read csv file

Step 2: Separate the csv file into x,y where x will be training set.y is used to verify that our study is correct or not.

Step 3:Row represent data items and column represent features.Number of column is number of dimension.

Step 4: Given the columns of X, are features with higher variance more important than features with lower variance.

Step 5: Covariance $z=z^{T}. z$.

Step 6: Eigen values and eigen vectors of z.

Step 7: Decompose into DPD-1, where D is the eigen vector. The diagonal matrix P has eigen values on both sides.

Step 8: Calculate new features $z^{*}=z_{p}^{*}$.

Step 9: Drop unimportant features from the new dataset.

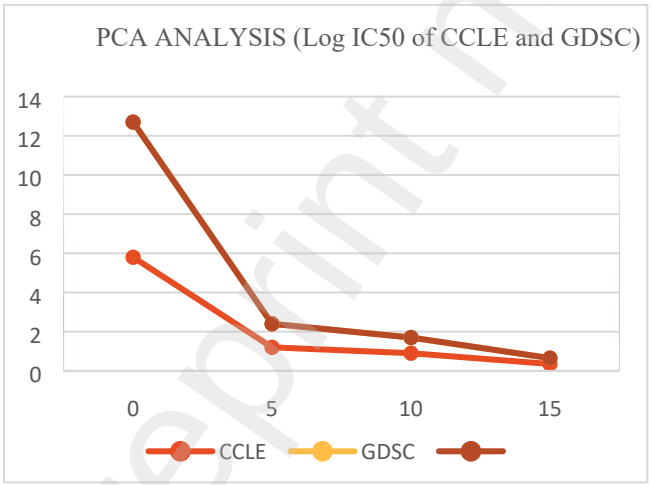


Fig 8.Comparision of variance for Log IC50 in CCLE and GDSC dataset.

D. Cost function (Gradient Decent method) (CCLE Data with Mechanical and Electrical Property)

The cost function is a method for assessing the performance of our algorithm/model. This calculation compares the predicted outputs and actual outputs of the model to see how far off it was. If our prediction differs significantly from the actual value, it returns a higher value.

Minimising cost function:
Cost function can have a negative value. It should be more negative because cost is being measured, and the goal is to minimise it.

Pseudo Code for Loss Function:
Inputs:

Learning_rate = 0.001
Iterations = 500
Output : Y_{pred}

N -Number of features
W_i-Drug similarity matrix
X_i . Cell line similarity matrix
$Y=\sum W_i.X_i = (W_0.X_0) + (W_0.X_1) + \ldots\ldots (W_n. X_n)$

where
learning rate co-efficient =0.001 #static,Dynamic
iteration=1000
W_i= [W_0 W_1 ……W_n]
X_j = [X_0 X_1 ……X_n]
Compare with threshold
k=0.3 # hyperparameter tuning
if Y>k, then Y=1#Drug is responding
Otherwise, Y=0 #Drug is not responding
Where Y is output

$Y_{predicted}$ = [Y_0 Y_1 ……Y_z]

probability = [a. fit (X, Y) predict _ proba (X) for a in (alf1, alf2, alf3,alf4, bclf)] where

alf1-L Elastic Net
alf2- pairwise support vector regression
alf3- Kernelized Bayesian Multitask Learning
alf4-Neural Networks
bclf -Voting Classifier
Y_pred = m * x_i+b #from statistical analysis y=mx+b

Mean Square Error (MSE) = $1/n(\sum(y_i-(mx_i+b))^2$
Where,

m=m-learning rate*md

md = -(2/n)*sum(x*(y-y_pred)) & b= b = -(2/n)*sum(y-y_pred)

If the cost is larger, it indicates that the model did not do well on the training set. We must identify ideal parameters such as m (coefficient) and b (cost) to find the lowest cost (intercept). Now, we may use gradient descent, which is an iterative process, to identify the best parameters and apply the "update rule."

Inputs:
X_i= $SimC_{nxn}$: Cell line similarity matrix
W_j=$SimD_{mxm}$: Drug similarity matrix
k: Model hyper-parameter

Output :Cost function

E. Loss formula for computation:
Gradient descent method:

This method is used to optimize a cost function that is parameterized by model parameter W.Cost function's incline or slope is determined by the gradient(derivative). As a result, we proceed in the opposite direction of the gradient to minimize the cost function.

Step 1: Set the weights W at random.
Step 2: Determine the gradients G of the cost function with respect to the parameters. Partial differentiation is used to accomplish this. G=J(W)/W.G.
Step 3: Adjust the weights by a factor that is proportional to G. Repeat the procedure until the criteria is reached.

The learning rate determines how many steps we must take to reach a minimum in step 3.

Loss function:

An event loss function, also known as a cost function, is a function that associates a real number with an event or a group of variables and is used in mathematical optimization and decision theory to understand the event's "cost." It's a metric for evaluating the algorithm's performance. The discrepancy between the present output and the expected output is known as loss in algorithmic terms.

RESULTS AND DISCUSSION:

```
    iterations = 1000
    #Number of data points n
    n = len(x)
    #Initialize learning rate
    learning_rate = 0.001

    for i in range(iterations):
        y_pred = m_curr * x + b_curr
        cost = (1/n) * sum([val**2 for val in (y-y_pred)])
        md = -(2/n)*sum(x*(y-y_pred))
        bd = -(2/n)*sum(y-y_pred)
        m_curr = m_curr - learning_rate * md
        b_curr = b_curr - learning_rate * bd
        print("m {}, b {}, cost {}, iteration {}".format(m_curr, b_curr, cost, i))
        return cost
        cost = training (x, y, 0.001, 1000)
x = np.array([[[-1. ],
       [-1. ],
       [-1.2],
       [-1.4],
       [-3.4],
       [-2.2],
       [ 1.1],
       [ 1.2]]])
y = np.array([[[[[ 5],
       [ 7],
       [ 9],
```

Table 1: cost estimation of y-predction

S.no	M(weight)	B(bias)	Cost
1	-2.499	2.49	0.05
2	-2.4	2.4	0.06
3	-1.67	1.67	0.0271
4	-1.39	3.39	0.048
5	-1.388	3.388	0.0591

To minimise our loss (or cost) function, we must first determine its slope. Our cost function, we discovered, is:

$$C(\mathbf{y}, \mathbf{w}, \mathbf{X}, b) = \frac{1}{N}\sum_{i=1}^{N}(y_i - max(0, \mathbf{w} \cdot \mathbf{X}_i + b))^2 \quad \text{---(1)}$$

w1,w2,w….wn are weights
X_i -inputs
Y_i-output
b-bias
Ensembles Performance on CCLE DATASET for the algorithms are

Table 2: MSE of Ensemble Algorithm with and without adding mechanical and electrical properties in CCLE dataset

Algorithm	PREVIOUS MSE	CURRENT MSE
KBMTL	4.073±0.522	4.07±0.5
PSVR	4.806±1.040	4.706±1.03
Elastic-Net	4.233±0.736	4.22+0.636
NN	4.087±0.695	4.03±0.595

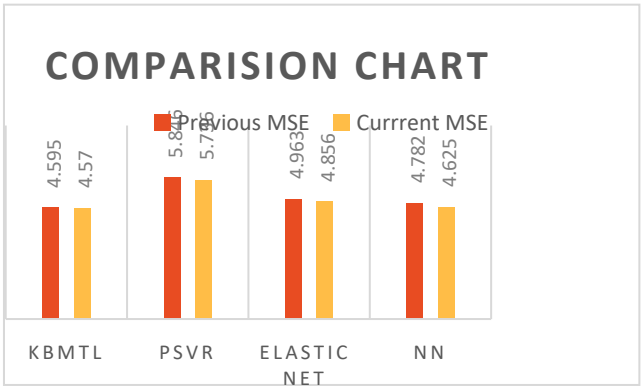


Fig.9 Comparison of MSE value with and without adding mechanical electrical properties in the CCLE dataset using Stacked Ensemble

Comparison of MSE using stacked ensemble algorithm before and after adding mechanical and electrical properties in the dataset is given in the chart. From the result it is understood the MSE value is reduced using the ensemble technical along with mechanical and electrical properties and prediction is improved.

CONCLUSION

Health-related issues are constantly promoted and must be addressed as quickly and easily as possible. To improve medical diagnosis while taking into account the cost of providing these services, while keeping in mind that each individual's health is of the utmost importance. Cancer has been a significant life-threatening disease, and early prediction of cancer cell elongation is critical. Early detection of cancer can help patients live longer lives. The ensemble technique using with Elastic Net,pairwise support vector regression, Kernelized Bayesian Multitask Learning and Neural networks are used in this study to predict drug response. To improve the fitness analysis in the drug response cost function is evaluated. As a result, cell elongation in cancer cell lines is avoided, and cell stability is preserved. This prediction is used in drug analysis to improve the inhibitory concentration, which is then used to recover from cancer sooner, making use of multiple stacked.

FUTURE SCOPE

To expand on this work, planned to include genomic markers of drug sensitivity dataset, that can have an impact on the outcomes in the cell line.

REFERENCES

1.National Cancer Institute (NCI): Cancer Stat Facts: Cancer of Any Site, (2019). https://seer.cancer.gov/statfacts/html/all.html.
2.American Cancer Society (ACS), Key Statistics for Lung Cancer, (2020). https ://www.cancer.org/cancer/lung-cancer/about/key-statistics.html.
3.Wiesweg M.,(2019). Machine learning-based predictors for immune checkpoint inhibitor therapy of non-small-cell lung cancer, Ann Oncol:30(4):655e7.
4.Heo, J. (2019). Machine learning based model for prediction of outcomes in acute stroke, Stroke :50(5):1263e5.
5.Gunther,M., Juchum, M., Kelter,G., Fiebig, H., & Laufer, S.(2016). Lung cancer: Egfr inhibitors with low nanomolar activity against a therapy resistant l858r/t790m/c797s mutant Angewandte Chemie, International Edition, vol. 55, no. 36, pp. 10890–10894.

6.Qureshi,R., Nawaz,M., Ghosh, A., & Yan, H.(2019). Parametric models for understanding atomic trajectories in different domains of lung cancer causing protein, IEEE Access , vol. 7, pp. 67551–67563.
7.Ikemura,S.,Yasuda.H., Matsumoto,S., Kamada,M., Hamamoto, J., Masuzawa,K., Kobayashi,K., Manabe,T.,D., Arai, I., & Nakachi.(2019). Molecular dynamics simulation-guided drug sensitivity prediction for lung cancer with rare egfr mutations. Proceedings of the National Academy of Sciences , vol. 116, no. 20, pp. 10025–10030.
8.Lee, G.,Y., H., & Lim, C., T.(2007). Biomechanics approaches to studying human diseases. Trends Biotechnol , 25: 111–118.
9.Lim, C.,T., Zhou, E., H., Li, A., Vedula, S., R., K., & Fu HX.(2006). Experimental techniques for single cell and single molecule biomechanics, Mater Sci Eng C, 26: 1278–1288.
10. Garnett, M., J.(2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature, 483, 570–575.
11. Barretina, J., (2012). The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity Nature, 483, 603–607.
12. James Surowiecki. (2014). The Wisdom of Crowds.
13. Ding, H., Takigawa, I.& Mamitsuka, H.(2013). Similarity-based machine learning methods for predicting drug-target interactions: a brief review. Brief Bioinform, 5(5):734-47.
14. Lin Zhang, Xing Chen, Na-Na Guan, Hui Liu & Jian-Qiang Li.(2018). A Hybrid Interpolation Weighted Collaborative Filtering Method for Anti-cancer Drug Response Prediction, vol no.9, article 1017.
15. Chuanying Liu, Dong Wei, Ju Xiang, Fuquan Ren, Li Huang, Jidong Lang, Geng Tian, Yushuang Li & Jialiang Yang.(2020).An Improved Anticancer Drug-Response Prediction Based on an Ensemble Method Integrating Matrix Completion and Ridge Regression, vol no.21.
16. Yitan Zhu, Thomas Brettin, Yvonne.A., Evrard, Alexander Partin, Fangfang Xia, Maulik Shukla, Hyun seung Yoo, James, H., Doroshow, Rick, L., & Stevens. (2020), Ensemble transfer learning for the prediction of anticancer drug response, vol.10,18040.
17. Chen, R., Liu, X., & Jin, S.(2018). Machine learning for drug-target interaction predicition. Molecules,23(9):2208.
18. Guosheng Lianga, Wenguo Fanb, Hui Luoa & Xiao Zhua.(2020). The emerging roles of artificial intelligence in cancer drug development and precision therapy, vol.128,110255.
19. Sachdev, K., & Gupta, M., K. (2019). A comprehensive review of feature based methods for drug target interation predicition. J Biomed Inform,93:103159.
20. Raghvi Bhardwaj & Nishtha Hooda. (2009). Prediction of Pathological Complete Response after Neoadjuvant Chemotherapy for breast cancer using ensemble machine learning, 2352-9148.
21. Aman Sharma & Rinkle Rani. (2019). Drug sensitivity prediction framework using ensemble and multitask learning.
22. Mehmet Tana, Ozan Fırat Özgüla, Batuhan Bardaka, Işıksu Ekşioğlua & Suna Sabuncuoğlu. (2018). Drug response prediction by ensemble learning and drug-induced gene expression signatures. Grand No.115E274.
23. Fangfang Xia et.al: “A cross-study analysis of drug response prediction in cancer cell lines.” (2021):1-14.
24. Senousy, M.B., El-Deeb, H.M., Badran, K., & Ibrahim Ali Al-Khlil. (2012). Ensample Learning Based on Ranking Attribute Value (ELBRAV) for imbalanced Biomedical Data Classification. ISSN-1110-2586: vol. 36 No. 1.

25. Aman Sharma, Rinkle Rani. (2018). Kernelized similarity based regularized matrix factorization framework for prediction anti-cancer drug responses. KSRMF: vol.1779-1790.

26. Emdadi Akram, Eslahchi Changiz. Dsplmf. (2020). A method for cancer drug sensitivity prediction using a novel regularization approach in logistic matrix factorization. *Frontiers in Genetics*, 11:75. pmid:32174963.
27. Suphavilai Chayaporn, Bertrand Denis & Nagarajan Niranjan (2018). Predicting cancer drug response using a recommender system *Bioinformatics.,* 34(22):3907–3914, pmid:29868820.
28. Zhang Naiqian, Wang Haiyun, Fang Yun, Wang Jun, Zheng Xiaoqi, & Shirley Liu X.(2015). Predicting anticancer drug responses using a dual-layer integrated cell line-drug network model, *PLoS Comput Biol* :11(9): e1004498: pmid:26418249.
29. Lakshman Kumar pappala, Dr. Selvi Rajendran.P.(2021). A Novel Music Genre Classification using Convolution Neural Networks: IEEE conference on communication and electronics system,8-10.
30. Dr. Selvi Rajendran,P., & Dr. Anithaashri.T.P. CNN based Framework for identifying the Indian Currency Denomination for Physically Challenged People. : IOP Conference Series: Materials Science and Engineering for the publication.
31. Dr. Selvi Rajendran,P., & Dr. Angilna Geetha.(2021). Optimization of hospital Bed Occupancy in Hospital Using Double Deep Q Network. International Conference on Intelligent Communication Technologies and Virtual Mobile Network : (ICICV-2021):4-6.
32. Chen, Joy Iong Zong, Hengjinda, P. (2021). Early Prediction of Coronary Artery Disease (CAD) by Machine Learning Method-A Comparative Study.Journal of Artificial Intelligence 3, no. 01: 17-33.
33. Balasubramaniam & Vivekanadam. (2021). Artificial Intelligence Algorithm with SVM Classification using Dermascopic Images for Melanoma Diagnosis. Journal of Artificial Intelligence and Capsule Networks 3, no. 1: 34-42.

34. Manoharan & Samuel. (2019). Study on Hermitian graph wavelets in feature detection, Journal of Soft Computing Paradigm (JSCP) 1, no. 01 (2019): 24-32.
35. Smys, S., Joy Iong Zong Chen & Subarna Shakya, (2020). Survey on Neural Network Architectures with Deep Learning, Journal of Soft Computing Paradigm (JSCP) 2, no. 03 : 186-194.

[36]. Lu, D.,Y. (2017). "Anticancer drug development, breakthroughs are waiting". Advances in Pharmacology and Clinical Trials 2.1: 119.
[37]. Gupta GP and Massague J. "Cancer metastasis: building a framework". Cell 127 (2006): 679-695.
[38]. Valastyan,S. & Weinberg, R.,A.(2011).Tumour metastasis: molecular insights and evolving paradigms, Cell 147.2: 275-292.
[39]. Nieto, M., A.(2016), EMT. Cell 166.1 : 21-45.
[40]. Lambert, A.,W.(2017), Emerging biological principles of metastasis, Cell 168: 670-691.
[41]. Rose,W.,C., Long, B.,H., Fairchild, C.,R., Lee,F.,Y., Kadow, J.,F. (2001). Preclinical pharmacology of BMS-275183, an orally activetaxane. Clin Cancer Res 2001; 7(7): 2016–2021.
[42]. Lee, F.,Y., Borzilleri, R., Fairchild, C.,R., Kim, S., H., Long, B., H.,& Reventos-Suarez C. (2001). BMS-247550: a novel epothilone analog with a mode of action similar to paclitaxel but possessing superior antitumour efficacy. Clin Cancer Res,7(5): 1429–1437.
[43]. Stachel, S.,J., Biswas, K., Danishefsky, S., J.(2001). The epothilones, eleutherobins, and related types of molecules. Curr Pharm Des, 7(13): 1277–1290.
[44] Gunjan Jeswani, Swarnali, D., & Paul,(2017). Nano- and Microscale Drug Delivery Systems.
[45]. Hotte, S.J., Hirte. & H.W.(2002) BAY 43-9006: early clinical data in patients with advanced solid malignancies. Curr Pharm Des, 8(25): 2249–2253.
[46]. Dancey, J.E.(2002). Agents targeting ras signaling pathway.Curr Pharm Des, 8(25):2249-2253.
[47] Chemotherapy: Success rates for different cancers (medicalnewstoday.com)
[48] Neben, K., Moehler, T., Egerer, G., Kraemer, A., Hillengass, J., & Benner, A.(2001). High plasma basic fibroblast growth factor concentration is associated with response to thalidomide in progressive multiple myeloma. Clin Cancer Res 2001; 7(9):2675–2681.
[49] Raje, N., Anderson, K.C.(2002). Thalidomide and immunomodulatory drugs as cancer therapy. Curr Opin Oncol, 14(6): 635–640.
[50] Figg, W.D., Kruger, E.A., Price, D.K., Kim, S., & Dahut, W., D. (2002). Inhibition of angiogenesis: treatment options for patients with metastatic prostate cancer. Invest New Drugs, 20(2):183–194
[51] Waselenko, J.K., Shinn, C.A., Willis, C.R., Flinn, I.W., Grever, M.R., & Byrd, J., C.(2001). Carboxyamido-triazole (CAI)—a novel ''static''signal transduction inhibitor induces apoptosis in human B-cell chronic lymphocytic leukemia cells. Leukemia Lymphoma, 42(5): 1049–1053.
[52] Jonasch, E., Haluska, F., G. (2001) Interferon in oncological practice: review of interferon biology, clinical applications, and toxicities. Oncologist, 6(1): 34–55.
[53] Dreau, D., Foster, M., Hogg, M., Swiggett, J., Holder., & W., D. (2000). White RL. Angiogenic and immune parameters during recombinant interferon-a2b adjuvant treatment in patients with melanoma. Oncol Res,12(5): 241–251.
[54] Kawano, Y., Matsui, N., Kamihigashi, S., Narahara, H., & Miyakawa, I. (2000). Effects of interferon-gamma on secretion of vascular endothelial growth factor by endometrial stromal cells. Am J Reprod Immunol, 43(1): 47–52.
[55] Leahy, K.,M., Ornberg, R.L, Wang, Y., Zweifel, B.S., Koki, A.T., & Masferrer, J.,L.(2002). Cyclooxygenase-2 inhibition by celecoxib reduces proliferation and induces apoptosis in angiogenic endothelial cells in vivo. Cancer Res , 62(3): 625–631.
[56] Wang, Z., Fuentes, C.F., Shapshay, S.M.(2002). Antiangiogenic and chemopreventive activities of celecoxib in oral carcinoma cell.Laryngoscope, 112(5): 839–843.
[57] Bevacizumab. Anti-VEGF monoclonal antibody, avastin, rhumab-VEGF. Drugs R D 2002; 3(1): 28–30.
[58] Chen, H.X., Gore-Langton, R.E., & Cheson, B.D. Clinical trials referral resource: Current clinical trials of the anti-VEGF monoclonal antibody bevacizumab. Oncology (Huntingt),15(8): 1017, 1020, 1023–1026.
[59] Stopeck, A., Sheldon, M., Vahedian, M., Cropp, G., Gosalia,R., & Hannah, A.(2002). Results of a Phase I dose-escalating study of the antiangiogenic agent, SU5416, in patients with advanced malignancies. Clin Cancer Res, 8(9): 2798–2805.
[60] Fabbro, D., & Manley, P.(2001). Su-6668. SUGEN. Curr Opin Investig Drugs, 2: 1142–1148.
[61] Hoekman, K.(2001). SU6668, a multitargeted angiogenesis inhibitor.Cancer J , 7(Suppl 3): S134–S138.
[62] Dowlati, A., Robertson, K., Cooney, M., Petros, W.P., Stratford,M.,& Jesberger. J.(2002). A phase I pharmacokinetic and translational study of the novel vascular targeting agent combretastatin a-4phosphate on a single-dose intravenous schedule in patients with advanced cancer. Cancer Res, 62(12): 3408–3416.
[63] Ohno, T., Kawano, K., Sasaki, A., Aramaki, M., Tahara, K., & Etoh, T.(2002). Antitumour and antivascular effects of AC-7700, a combretastatin A-4 derivative, against rat liver cancer. Int JClin Oncol 2, 7(3): 171–176.

[64] Bramhall, S.R., Hallissey, M.T., Whiting, J., Scholefield, J., Tierney, G., Stuart, R.C.(2002). Marimastat as maintenance therapy for patients with advanced gastric cancer: a randomised trial. Br J Cancer , 86(12): 1864–1870.
[65] Bramhall, S.R., Schulz,J., Nemunaitis, J., Brown, P.D., Baillet, M.,& Buckels J.A.(2002). A double-blind placebo-controlled, randomised study comparing gemcitabine and marimastat with gemcitabine and placebo as first line therapy in patients with advanced pancreatic cancer. Br J Cancer; 87(2):161–167.
[66] Groves.M.D., Puduvalli,V.K., Hess,K.R., Jaeckle, K.A., Peterson, P., & Yung, W.K., (2002). Phase II trial of temozolomide plus the matrix metalloproteinase inhibitor, marimastat, in recurrent and progressive glioblastoma multiforme. J Clin Oncol; 20(5):1383–1388.
[67] Batist, G., Patenaude, F., Champagne, P., Croteau, D., Levinton, C., & Hariton, C. (2002). Neovastat (AE-941) in refractory renal cell.carcinoma patients: report of a phase II trial with two dose levels. Ann Oncol, 13(8): 1259–1263.
[68] Falardeau, P., Champagne, P., Poyet, P., Hariton, C., Dupont, E.(2001). Neovastat, a naturally occurring multifunctionalantiangiogenic drug, in phase III clinical trials. Semin Oncol, 28(6): 620–625.
[69] Gutheil,J.C., Campbell,T.N., Pierce, P.R., Watkins,J.D., Huse,W.D., & Bodkin D.J.(2000). Targeted antiangiogenic therapy for cancer using Vitaxin: a humanized monoclonal antibody to the integrin avb3. Clin Cancer Res; 6(8), 3056–3061.
[70] Patel, S.R., Jenkins, J., Papadopolous, N., Burgess, M.A., Plage, C., & Gutterman, J. Pilot study of vitaxin—an angiogenesis.
[71] Enrique Espinosa, Pilar Zamora, Jaime Feliu & Manuel González Barón(2003). Classification of anticancer drugs—a new system based on therapeutic targets, Cancer Treatment Reviews.