International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN :2147-6799 www.ijisae.org Original Research Paper

Feature-Based Machine Intelligent Mapping of Cancer Beating Molecules

[1]P.Selvi Rajendran, [2]K.R.Kartheeswari

Submitted: 10/11/2022 **Accepted**: 11/02/2023

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,the mechanical response of food particles plays a major role to modulate 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 learning methods. ElasticNet, pairwise support vector regression, Kernelized Bayesian Multitask Learning,and Neural networks use 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 cancer in the sensitive line. To improve the fitness analysis in the drug response,the cost function is calculated in this research.

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

1. Introduction

The world's second most-frightened sickness is cancer growth according to the report acquired by factual investigation of the US National Cancer Institute. In that report, it was expressed that in 2020, 1.8 million new malignant growth cases are anticipated, with 600,000 fatalities. [1]. What's more, cellular breakdown in the lungs is supposed to be the second generally deadly censured. The examination of the US determined that assessed new cases were 2,28,820 in the year 2020. Around the same time, unsurprising demise cases were 1,35720. Thus, 25% of the populace in the US were supposed to be impacted by cellular breakdown in the lungs in 2020. A supported five years virtual review report expressed that in a lifetime, 1 out of 15 in men, 1 out of 17 ladies, and smokers had a high-risk factor for cellular breakdown in the lungs.

This five years report showed that 20% of cellular breakdown in the lungs patients were seriously impacted [2].The use of antiePD-1 treatment has recognized exploratory parts of the treatment reaction in those with NSCLC. In clinical practice, immunotherapy, smoking history, execution status, sex, the presence of meta states, transformations, and pathology [6e14] are regularly evaluated. These are kept in electronic wellbeing records (EMRs). No clinical element can precisely foresee the reaction of antiePD-1 treatment (customized cell passing protein-1).To work on the reaction of antiePD-1 models, it are expected to incorporate variables. AI assumes an imperative part in sickness movement and therapy reaction [3,4]. In the beginning phases, an immaterial measure of particle inhibitors created confident outcomes in drug treatment. Thus, the lifetime of the patient is likewise expanded. However, startling optional mutation(s) limited the adequacy of the medication because of medication resistivity [5]. Numerous computational examinations have been guided in sub-atomic elements to break the medication resistivity reproductions [6,7]. Mechanical properties assume an essential part and are completely interrelated to human medical issue. The physiological capability of the cells relies on the mechanical property. The progressions in mechanical property outline the harm in the physiological elements of the cell. In this way, it yields an expansion in the sickness state [8,9].

[1]*Professor, Computer Science and Engineering,School of Computing,Hindustan Institute of Technology and Science, Chennai, India*
[2]*Senior Research Fellow, Computer Science and Engineering,School of Computing, Hindustan Institute of Technology and Science, Chennai, India*

2. Related Works

The Medication Target Joining is investigated [13], through a similitude based AI calculation. In this work, synthetic and genomic spaces are utilized in the similitude based examination. In this methodology, similitude based examination handily converged with Portion based learning techniques. In [14], inspected that with the assistance of cell and medication line likenesses, we had the option to decide the medication's impact reaction. To make this forecast we utilized cooperative separating models, and from the normal of the projected qualities anticipated awareness was calculated.

The Hyper Interjection Weighted Cooperative Channel depends on the set ofknown drug responses.The execution of the framework is exceptionally reliant upon the sparsity of the reaction lattice. This brought about a decrease of execution in the drug.The consequence of PCC was higher than that of the other coordinated cell line. [15] explained a cross-approval with 10 overlap. From the examination, they observed that the model was all the more naturally interpretable, and highlight choice chance was slacking in the examination.

In [16] approach, power move learning has been utilized to anticipate both new growth cells and new medications. This examination expressed that this new strong learning further developed the expectation execution in the medication reaction forecast applications. From the examination, it was seen that a solitary medication reaction was not accessible, and a mix of medication reactions decreased the presentation. [17] expounded DTI forecast AI techniques. In this strategy, chemogenomic data sets were utilized. Negative examples were likewise taken care of. A chemogenomic strategy is ordered into managed and semi-regulated learning techniques. This administered learning is additionally partitioned into comparability and element based techniques. In this exploration [18,32], the author has shown that Computerized reasoning assumes a huge part in the disclosure of new materials and speeding up anticancer medication development.These administrations assist the specialists with settling on the right therapy choices and diminish superfluous medical procedure. AI examination was utilized to foresee the awareness of the medication. [19].The creator inspected that DTI expectation utilizing highlight put together chemogenomic approaches was described.Based with respect to the aftereffects of this survey,these are delegated SVMs and group based methods.

Here, the similitude approach is avoided by the chemogenomic approach. In this work,[20] fostered a twofold gathering AI calculation to foresee the neurotic reaction after neoadjuvant chemotherapy. This twofold outfit was utilized to foresee multi-standards direction. 99.08% precision is gotten utilizing the k-overlay cross-approval method. In [21]proposed a group and perform various tasks learning system for anticipating drug sensitivity,which suggested [22], proposed a group technique for drug reaction prediction.Drug-prompted quality articulation has been added to cytotoxicity data sets in huge quantities.Vitro tests, notwithstanding informational collection testing, are utilized to foster the expectation. [23,29,20,31] utilized five openly accessible cell-line-based informational indexes in this review. This cross review depends on themachine learning models.

The Best get study was accomplished through profound brain organizations. This examination at first expressed that CTRP yields a superior forecast on the test information. GCSI was the most unsurprising informational collection among any remaining cell line informational indexes. In this cross-study, GCSI joined with CTRP gives a more exact forecast. The exploration of [24], gave a miniature cluster innovation to malignant growth forecast. This miniature cluster can conquer the missing qualities or imbalanced bio-clinical information issue. Gathering Learning in view of Positioning Trait Worth (ELBRAV) has been utilized for the examination. [25] laid out an enemy of malignant growth drug reaction expectation model utilizing a likeness based regularization network. In this likeness examination, medication and tissue were the principal goals and GDSC and CCLE informational indexes were utilized. The typical MSE was 3.24 and 0.504.

3. Current Challenges

A Lot of bio-agents and chemicals are waitingfor pharmacological evaluation and mechanism investigation following a significant promotion of chemistry and genetic engineering approaches. Themedicine’sworth 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.

Tubulin:

Medicines that interfere with tubulin are grouped 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. Tubulin stabilisers known as epothilones are a novel type of stabiliser. Although preclinical researchhas shown that this type ofmedicationhas potential, the results of phase II,phase III clinical trials are currently unavailable.

4. 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. Cancer has spread to other parts of the body at this stage.

5. 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.

6. Mechanisms Used in Cancer Therapies

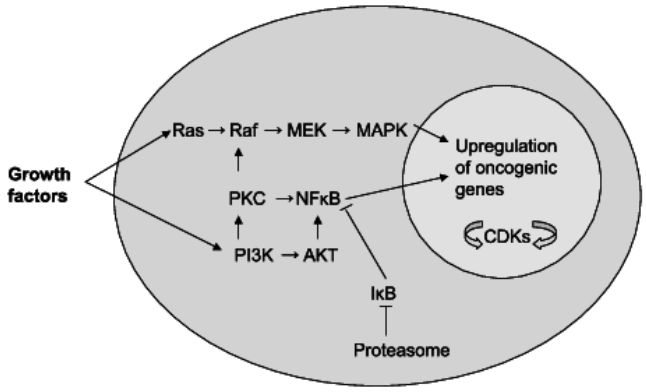


Fig 1. Pathways of metabolism

Figure 1 depicts a diagram of the mechanismsemployed in cancer therapy. The most well-known drug in this class is imatinib, which inhibits the tyrosine kinases bcr/abl and c-kit. 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 inhibitors, such as lonafarnib and R115,777, operate as fake metabolites of this enzyme.Raf (BAY 43-9006) and MEK(CI-1040) inhibitors are also available.

7. 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.

Anticancer Drugs: A Classical Classification:

Chemotherapy

It is aneffective treatment that comprises the use of chemicals to remove cancer cells. It is used to avoid cell multiplication and division.

The following are the chemotherapy therapy drugs

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

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 analogs 5. Anti-androgens

Immunotherapy

The treatment of disease by stimulating or inhibiting unaffected cells is known as biological therapy.Immunotherapy,is also called biological therapy.It is a process of treating cancer by stimulating else

inhibiting the immune system. Activation immunotherapies enhance 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 tumors:

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.

DNA helix:

The main mixtures uncovered to be useful in the therapy of malignant growth were alkylating specialists. The most well-known site of alkylation is inanN-7 place of guanine, yet this differs relying upon the medication family. Alkylators incorporate nitrogen mustards, nitrosoureas, triazenes, platinum mixtures, and anti-infection agents.

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.

Endothelium:

In tumour cells, interferon also reduces the formation of VEGF, although this effect appears to be medicated by interferon gamma. As a result, one of the COX-2 inhibitors possible modes of action is to promote endothelium growth.

Bevacizumab, a VEGF receptor-binding monoclonal antibody, binds to all of them . SU-5416 is a small molecule that interacts with the tyrosine kinases VEGFR-1 and VEGFR-2. C-kit and platelet-derived growth factor receptor are two more proteins with which it interacts. Clinical studies of SU-5416 in hematological 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.

Extracellular matrix:

MMP activation increases invasion in tumours and is a crucial stage in angiogenesis.MMPs can also cause the release of VEGE,bFGF, and insulin growth factors. Clinical trials are now being conducted on several MMP inhibitors .Marimastat ,prinomastat and BAY 12-9566 are among the most common enzyme activity inhibitors .MMP production is suppressed, activation is blocked and MMP breakdown is promoted by tetracycline derivatives like neovastat .In addition to MMPs, other extracellular matrix elements such integrin,endothelin and thrombospond in could be targeted as anti-cancer therapy.

8. Host Cell Inhibitors and Other Drugs

A few medications are intended to go after organs that might hold onto malignant growth cells.Bisphosphonates osteoprotogerin and PTHRP antibodies are presently the main substances that restrain bone cell movement and the microenvironment. More medications to target different organs in danger of metastasis could be created from here on out. At long last, cytokines like interferon and interleukin2 are notable for expanding the resistant framework's anticancer action.

The objective of this paper is to current situation with workmanship AI techniques for hostile to malignant growth drug reaction displaying and expectation, with the objective of better using high-layered, multi-omics profiles, as well as information on disease pathways designated by against malignant growth compounds while anticipating phenotypic reactions.

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 mixedfibres. Measurements of the cancer cell’s mechanical characteristics give new information that Cancer cells are softer than non-transformed cells.

The compelling firmness and time-subordinate deformity qualities of the cell,the degree of extending or compression it goes through during mechanical testing, and the power runs regularly accomplished during such disfigurement all impact the kind of biophysical measure used to research a particular cell type.The significant power reaches and length(displacement) scales in the chose cell and sub-atomic cycles of pertinence in natural frameworks are portrayed in Figures 2a and 2b.The power and dislodging ranges accomplished by different biomechanical examines are displayed in these graphs.

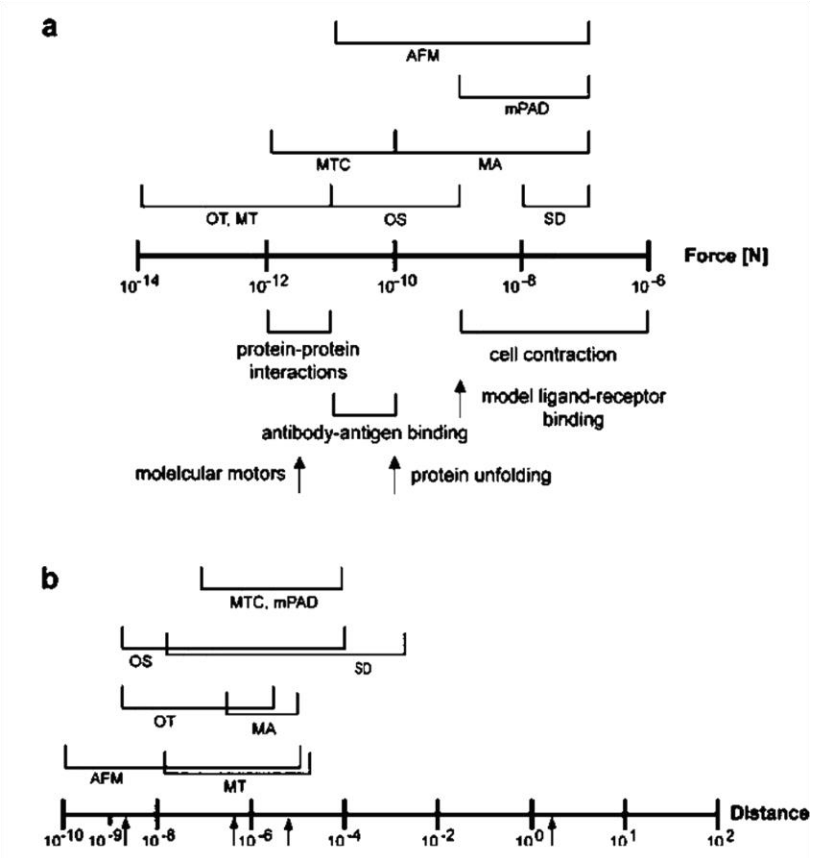


Fig.2. An overview of the forces and displacements investigated by various biomechanical arrays

MTC-Medullary thyroid carcinoma
AFM-Atomic force microscopy
OS-Overall survival
OT-Occupational Therapy
Mpad-Microfabricated Post-Array-Detectors
Mechanical Factors facilitate the spread of malignancy:
Cell behavior is powerfullyinclined by the mechanical influences of the tumourmicroenvironment. 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.

9. Methods And Materials

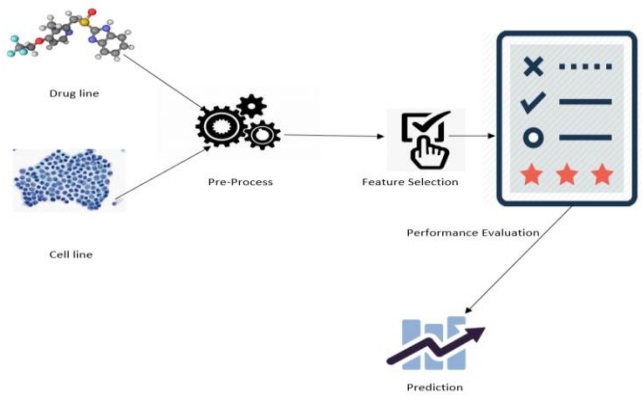
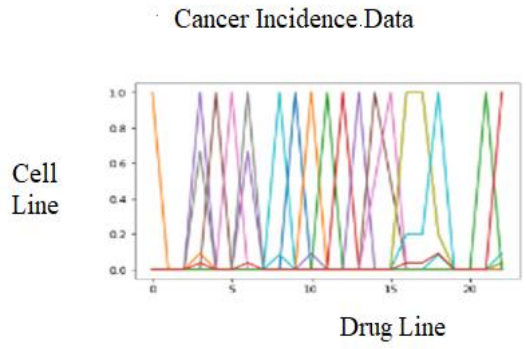


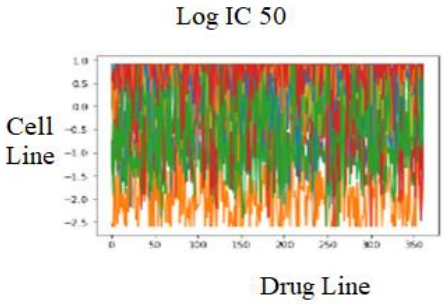
Fig 3. Performance evaluation procedure

A. Dataset:

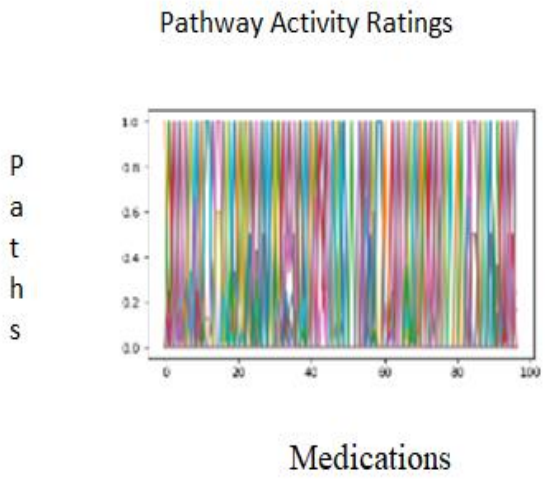
In this study, the cancer statistics are derived from both the properties of the drugs and cell lines, as well as an ensemble voting method applied to the results of the study. This section contains statistics on anti-cancer drug response prediction to reduce the cancer burden. Cancer incidence data 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.



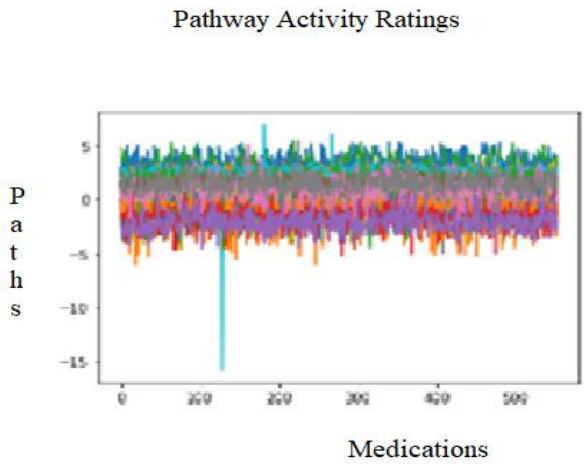
a) CCLE dataset



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's 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 first compound information is excessively enormous for a restricted gadget to deal with, especially regarding the quantity of unlabeled mixtures, keeping us from building a thoroughly prepared machine. With the essential objective of further developing the forecast exactness of regulated learning, we diminished the unlabeled information from thousands to many thousands. We found that unlabeled information didn't influence the precision of the last activity by running tests on decreasing information.

Inhibitory Concentration (IC50) Data after Preprocess

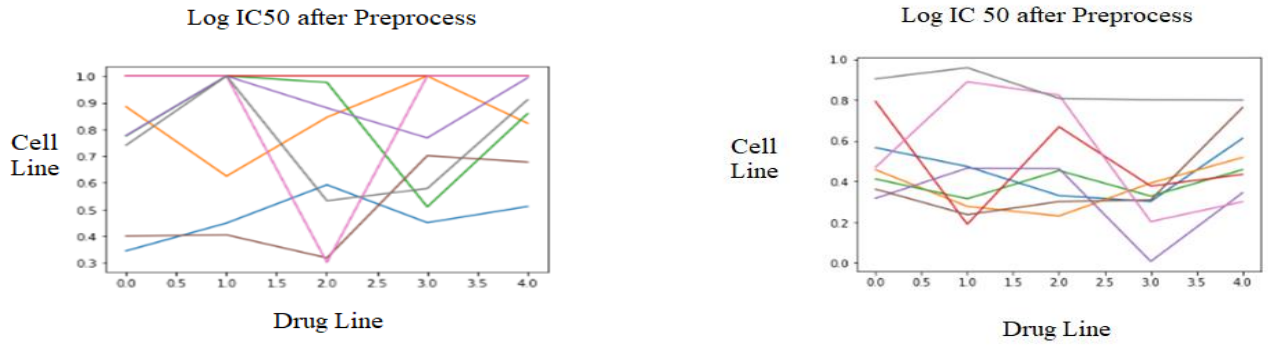


Fig 5. Log IC50 dataset (CCLE and GDSC) after Preprocess

B. Similarity:

CSS-Cosine similarity signature

DL=Drug Line

GenExpGenetic Expression

TarProTarget Protein

LIC50 → Log IC50

- GenExpCL=Cell Line (CL)∑CCLE,GDSC
- TarProDL=Drug Line (DL) ∑CCLE,GDSC
- LIC50(Response)

(CSSDL, CL) =∑CCLE, GDSC

1. Cosine Similarity:

In a dataset, the cosine similarity can be used to locate groups of related things.The similarity is measured in machine learning from a distance with dimensions 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 the CCLE and GDSC dataset.

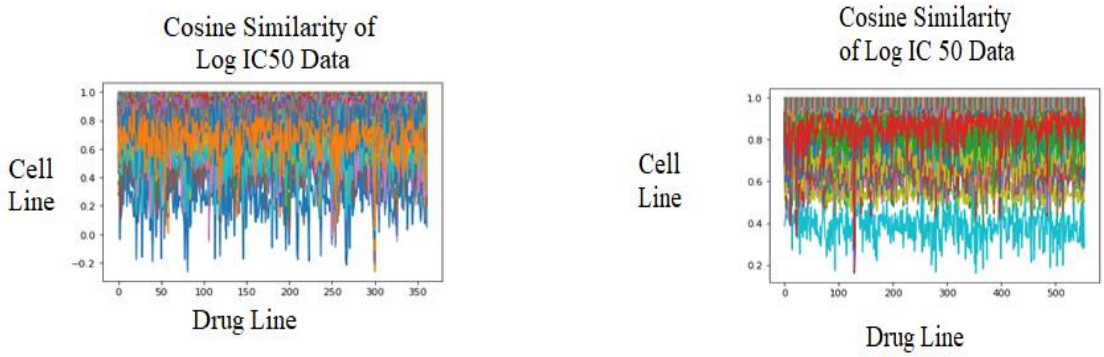


Fig.6.GDSC cosine similarity and CCLE cosine similarity(LogIC50)

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%. Relationship of pathway action scores and medication reactions. Lines of drugs are shown, though segments of pathways are shown. The positive and negative qualities address the positive and negative connection between's the two factors.

The factors are going in a similar bearing on the off chance that the relationship is positive. At the end of the day, as one variable ascent, different falls, as well as the other way around. The factors are moving in inverse bearings when there is a negative relationship.

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)

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.

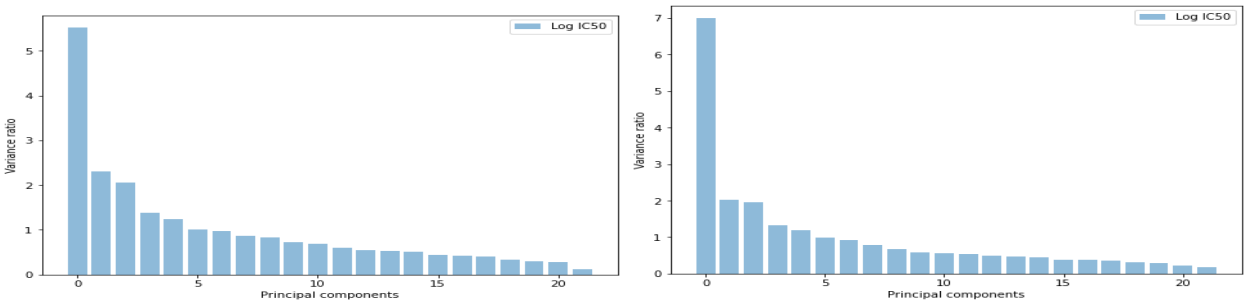


Fig7.PCA of Log IC50 for CCLE and GDSC Dataset

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 the training set.Yis used to verify whether our study is correct or not.

Step 3:Row represents data items and column represents features.A number of columns is the number of dimensions.

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

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

Step 6:Eigenvalues and eigenvectors of Z.

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

Step 8:Calculate new features $Z^{*}= Z_{p}^{*}$.

Step 9: Drop unimportant features from the new dataset.

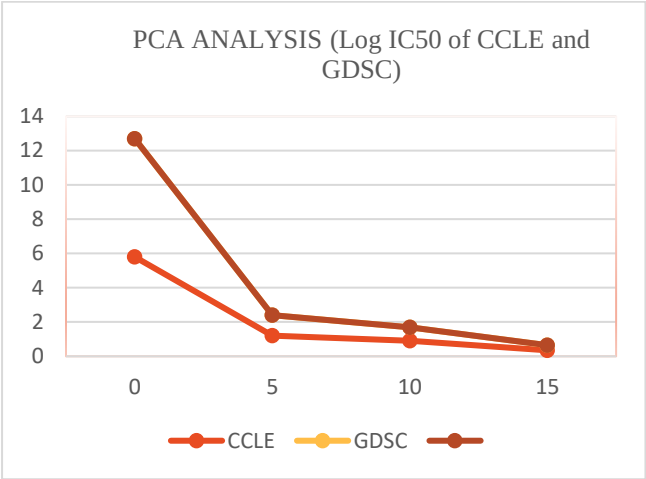


Fig 8. Comparison of variance for Log IC50 in CCLE and GDSC dataset.

D.Cost function(Gradient Decent method) (CCLE Datawith 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.

Minimizing cost function:

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

Pseudo Code for Loss Function:

Inputs:

Learning_rate = 0.001

Iterations = 500

Output:b_{pred}

$c=\sum a_i.b_i=(a_0.b_0) +(a_0.b_1) +\ldots\ldots(a_n. b_n)$

where

learning rate co-efficient =0.001 #SStatic,Dynamic

iteration=1000

a_i= [a_0 a_1 ……a_n]

b_j= [b_0 b_1 ……b_n]

Comparewiththreshold

d=0.3 # hyperparametertuning

if c>d, then =c1#Drugisresponding

Otherwise, c=0 #Drug is not responding

Wherecisoutput

$c_{predicted}$ = [c_0 c_1 ……c_z]

probability=[a.fit(b,c)predict_proba(b)fora in (elf1, elf2, elf3,elf4,gclf)] where

elg1-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

If the cost is larger, 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."

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 concerning 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 are 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.

10. 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-prediction

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….we are weights

X_i -inputs

Y_i-output

b-bias

EnsemblesPerformance on CCLE DATASET for the algorithms are given in tabulation 2.

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

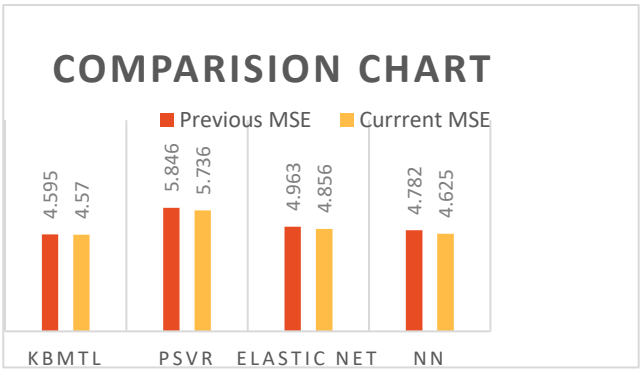


Fig.9 Comparison of MSE value with and without adding mechanical electrical properties in the CCLE dataset using Stacked Ensemble

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 techniques along with mechanical and electrical properties, and prediction is improved.

11. 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 Elastic Net,pairwise support vector regression, Kernelized Bayesian Multitask Learning, and Neural networksis 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 includegenomic markers of drug sensitivity dataset, that can have an impact on the outcomes in the cell line.

Acknowledgments

This research is funded by the Indian Council of Medical Research (ICMR). (Sanction no:
ISRM/12(125)/2020 ID NO.2020-5128 dated 10/01/21)

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 AngewandteChemie, 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 modeling 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-QiangLi.(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 &JialiangYang.(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 seungYoo, 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 prediction. Molecules,23(9):2208.

[18] GuoshengLianga, WenguoFanb, 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 interactionprediction.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 FıratÖzgüla, BatuhanBardaka, IşıksuEkşioğlua&SunaSabuncuoğlu.(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).Ensemble 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] EmdadiAkram, EslahchiChangiz. 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] SuphavilaiChayaporn, 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, *PLoSComput Biol*:11(9): e1004498: pmid:26418249.

[29] Lakshman Kumar pappala, Dr. SelviRajendran.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. SelviRajendran,P.,&Dr. Angilna Geetha.(2021). Optimization of hospital Bed Occupancy in Hospital Using a Double Deep Q Network.International Conference on Intelligent Communication Technologies and Virtual Mobile Network : (ICICV-2021):4-6.

[32] Chen, Joy IongZong, Hengjinda, P.(2021). Early Prediction of Coronary Artery Disease (CAD) by Machine Learning Method-A Comparative Study.Journal of Artificial Intelligence 3, no. 01: 17-3