Personalized Breast Cancer Prognosis through Data Mining Innovations

Dr.B Rama Ganesh[1]., Praveen B M[2]., Krishna Prasad K [3]., Viswanath G[4]
[1]Post-Doctoral Fellow, Srinivas University, Mangalore, Karnataka, India.
[2]Professor, Institute of Engineering and Technology, Srinivas University, Mukka-574146, Karnataka, India.
ORCID-ID: 0000-0003-2895-5952,
[3] Professor, Institute of Engineering and Technology, Srinivas University, Mukka-574146, Karnataka, India.
ORCID-ID: 0000-0001-5282-9038,
[4]Associate Professor, Department of CSE-AIML,
Sri Venkatesa Perumal College of Engineering and Technology, Puttur, India.
ORCID-ID: 0009-0001-7822-4739,

ABSTRACT: Progress in medical research on cancer diagnosis and prognosis, especially in breast cancer, has imposed considerable demands on oncologists due to the disease's complex and varied characteristics. Research aimed at estimating breast cancer survival has been proposed to tackle this difficulty. This study seeks to integrate histological and genetic data to improve prognostic precision, reduce superfluous therapy measures, and provide personalized patient care. A comprehensive tool for individualized breast cancer survival forecasts is built by employing various machine learning techniques, including "SVM, Random Forest, and neural networks". This technology enables physicians to make educated treatment decisions while efficiently optimizing healthcare resource allocation. Furthermore, utilizing ensemble approaches, especially the "Voting Classifier", improves forecast accuracy. Extending the project to incorporate a user-friendly frontend with the Flask framework enhances user testing and authentication, hence assuring seamless interaction and practical implementation in clinical environments. Ultimately, individualized breast cancer survival forecasts enhance patient outcomes and optimize healthcare delivery, tackling the escalating worldwide issue of breast cancer.

"Index Terms—Breast cancer survival estimation, gene expression, copy number variation, histopathological whole slide images, utility kernel, support vector machine, machine learning, deep neural networks".

1. INTRODUCTION:

Breast cancer, defined by the unregulated proliferation of breast cells, continues to be a major worldwide health issue, especially for women. The complex makeup of breast tissue and the course of cancer highlight the essential requirement for precise diagnostic and prognostic techniques. Progress in medical science and technology has allowed oncologists to combine qualitative histology data with quantitative genetic information to forecast clinical outcomes and customize treatment options successfully [1].

The "Global Cancer Observatory (GCO)", a undertaking of the "World Health Organization (WHO)", has launched regarding figures about the worldwide incidence of cancer. In 2020, round "19,292,789" new cancer cases have been documented international, with breast most cancers representing 11.7% of the total cases. Moreover, the occurrence of breast cancer in girls is anticipated to rise to 24.Five% by means of 2040, underscoring the important necessity for more suitable prognostic and diagnostic instruments [2].

Conventional survival prediction algorithms frequently encounter difficulties in precisely forecasting clinical outcomes due to the heterogeneity of breast cancer and the diverse treatment responses exhibited by patients [3]. Recent breakthroughs in medical imaging and next-generation sequencing technologies, like METABRIC and "The Cancer Genome Atlas (TCGA)", have substantially enhanced the field by offering comprehensive genome-scale transcriptome data for breast cancer research.

Survival prediction, which refers to the period of time a affected person survives following a cancer diagnosis, is essential for medical decision-making and remedy planning. Standardizing reporting and assessing survivability frequently relies upon on a five-12 months criteria, given that not less than 5 years is needed to categorise a affected person record as both survived or now not survived [5]. Due to the intricacies of breast cancer and the disparities in survival rates, novel methodologies are crucial for accurately categorizing patients as short-term or long-term survivors according to various survival benchmarks [6].

"Machine Learning (ML)" has been an effective instrument in the creation of survival prediction models for breast cancer. ML algorithms may autonomously learn from data, discern complex correlations among variables, and exhibit proficiency in handling substantial quantities of medical data [7]. Using ML in medical research has significant opportunities to increase the accuracy of the patient and better patient treatment results.

This work is trying to meet the immediate need for accurate survival models in breast cancer research. We aim to create strong and sewn survival tool tools by combining different datasets and using modern machine learning algorithms. 10 These technologies can help the oncologist make decisions about educated treatments, increase the distribution of resources for health services and eventually improve patient results. By using an intensive literature analysis and new approaches, we aim to increase breast cancer research and clinical practice.

2. LITERATURE SURVEY

Breast cancer is still a major health problem globally, which requires effective immunity and clinical units to increase the patient's results and inform treatment options. This literature observation examines several studies and research publications, which has increased the effect of large data and machine learning when the diagnosis of breast cancer, the difficulties with compliance with patient treatment, genomic and transcripted structure, and predict the survival results.

"Clarke (1994) [1]" examines the necessity of a variable marker in breast cancer and emphasizes the importance of establishing dependable indications to inform medical decisions and enhance patient outcomes. This study underscores the difficulties in forecasting breast cancer advancement and evaluating disease prognosis.

"Martin et al. (2005) [2] "Investigate the patient's following problems with breast cancer, highlight the importance of the patient's participation and follow the treatment protocol for favorable results.

"Curtis et al. (2012) [3]" examine the genomic and transcriptome structure of breast cancer, uncovering novel subgroups based on molecular traits. Their discovery offers significant insights into the heterogeneity of breast cancer and underscores the possibility for tailored treatment options guided by molecular profiling.

"Tomkzak et al. (2015) [4] presents a comprehensive evaluation of" Cancer Genome Atlas (TCGA) ", which emphasizes its significance as an essential resource for cancer research. This research emphasizes the importance of wide genomic data in indicating molecular foundations and possible treatment goals.

"Delin et al. (2005) [5]" consider three data mining functions to predict the existence of breast cancer, reflecting the effect of machine learning in programming. This study emphasizes the effect of data operation function on predicting clinical consequences and leading decisions on breast cancer treatment.

"Polyek (2011) [6]" examines the diversity of breast cancer, which highlights many genetic sub -factories and clinical properties shown by individuals. This study emphasizes the importance of individual medical strategies adapted to unique patient profiles to increase the results of the treatment.

"Obermeyer and Emanuel (2016) [7]" Check the effect of large data and machine learning on clinical medicine, and emphasizes the ability to change the health care distribution and increase patient results. This work emphasizes the need to use modern analytical approaches to gain valuable insights from a broad dataset for individual patient care.

"Van Veer et al. (2002) [8]" examine gene expression profiling as a pathological indicator of clinical consequences in breast cancer. This study reflects the effect of molecular biomarkers in predicting the disease forecast and indicates treatment strategies, facilitates individual medicine methods in breast cancer care.

Literature analysis emphasizes the versatile character of breast cancer prediction and the need to use different data sources and sophisticated analysis methods to improve the accuracy and patient results that are predicted. By using genetic, transcriptptomic and clinical data, researchers can create individual -related models that improve the decision on treatment and optimize the health care distribution for breast cancer patients.

3. METHODLOGY

a) Proposed work:

The proposed breast cancer Prophial cancer system combines the methods of learning a sophisticated machine with different data sources to improve pregnancy accuracy. This appoints a tool core for "Support Vector Machines (SVMS)", which integrates gene expression data with histopathology images to increase performance in relation to the current function. The analysis is enriched by experimenting with "naive bees, decision trees and random forests". The DL technology, which includes varial autoncoders and transmission learning, improves the model's functionality, using the "Convisional Neural Network (CNNS)".

A "voting classifies, integration of random forest, supportive vector classifies and decision-making three model", is used to use multimodal data combinations to predict survival at different intervals. A "Bottle Framework" connected to SQLITE is designed to provide user deposits and SININ features, which enable user tests through input submissions and evaluation of the proposed model performance. These pre -transdictment prediction systems for breast cancer want to improve the accuracy and functionality of the system.

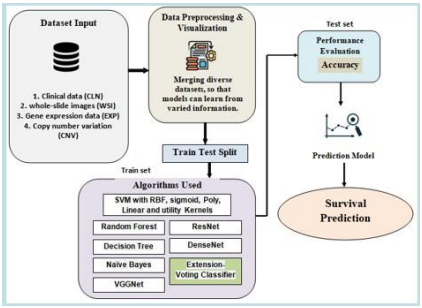
b) System Architecture:

The system architecture starts with the entry of breast cancer datasets, encompassing gene expression data and histology. The data is subjected to preparation stages including cleaning,

normalization, and feature extraction, then followed by visualization to derive insights into the dataset's features. The processed data is subsequently divided into training and testing sets for model building and assessment.

Diverse machine learning methodologies, such as “Random Forest, Decision Tree, Naïve Bayes, VGGNet, and DenseNet”, are utilized for survival prediction. These models are trained with the training data and evaluated on the test set to determine their performance.

The predictive model incorporates trained algorithms to forecast survival outcomes for breast cancer patients. This approach employs characteristics derived from gene expression data and histopathology pictures to generate precise predictions. The system design incorporates a Flask framework connected with SQLite for user registration, authentication, and data submission, facilitating user testing and assessment of the prediction model's efficacy.



“Fig 1 Proposed Architecture”

c) Data processing

1. Data Sampler:

- Subsampling the extensive dataset to provide a balanced training dataset.

- Producing balanced class data to mitigate class imbalance difficulties, guaranteeing equal representation of each class in the training set.

2. Define Utility Kernel:

- A utility kernel for “Support Vector Machines (SVM)” is defined as a mechanism that augments the SVM's capacity to manage intricate and non-linear connections within the data.

3. Applying PCA:

- “Principal Component Analysis (PCA)” [16] is used on the dataset to diminish dimensionality while retaining 95% of the explained variance ratio, facilitating feature selection and enhancing model effectiveness.

4. Exploring the Dataset:

- Analysis of different data components:

- CLN: Clinical data

- WSI: Whole slide images

- EXP: Gene expression data

- CNV: Copy number variation data

- Comprehending the attributes and distributions of each data element.

5. Combining the Data:

- Diverse combinations of data elements are generated for analytical purposes and model training:

- “CLN_EXP, CLN_CNV, CLN_WSI, EXP_CNV, EXP_WSI, CNV_WSI, CLN_EXP_CNV, CLN_EXP_WSI, EXP_CNV_WSI, CNV_CLN_WSI, multimodal (WSI, CLN, CNV, EXP)”.

6. Visualization:

- Employing Seaborn and Matplotlib packages for data visualization to elucidate the dataset's distributions, correlations, and trends.

- Visual representations facilitate comprehension of data linkages and the identification of useful predictive characteristics for model training.

d) TRAINING AND TESTING

The dataset is divided into a training kit and a test set to train the SVM-based core model to predict the existence of breast cancer. The training kit contains a balanced classification of breast samples, which includes both short -term and long -lasting remaining, so that the model can learn from a wide range of examples. The SVM model is trained on this dataset, which improves the ability to identify complex data correlations.

After completion of model training, it is evaluated using a test set to evaluate its performance and generalization functions. The test set includes new samples that were excluded from the training phase, which enables a fair assessment of the estimated performance of the model. Performing indicators, including accuracy, accurate, recall and F1 score, are calculated to assess the effectiveness of the model to predict the existence of breast cancer.

e) ALGORITHMS:

CLN - SVM – rbf

CLN-SVM-RBF trainer The coach on a "Support Vector Machine (SVM)" Model "model" Radial Basis (RBF) core ", clinical (CLN) Data. This model is used to predict the results of the existence of breast cancer based on clinical properties in the project, including patient demographics, tumor reluctance.

CLN - SVM – Poly

CLN-SVM-Poly represents the "Support Vector Machine (SVM)" model with a polymorphic core, which is trained on clinical data. This model is used in the project to predict the existence of breast cancer based on clinical factors such as patient demographics, tumor functions and treatment history. Polynomial Curnell lets SVM identify non-led correlations in CLN data, so individual survival improves its ability to identify patterns associated with results. Research should use "CLN-SVM poly" to create an effective prediction model that helps the oncologist decide on informed treatment and increase patient treatment in breast handling.

CLN - SVM – Utility

CLN-SVM-Youtelity reflects a "Support Vector Machine (SVM)" model, which uses a tool trained on clinical (CLN) data. The model is used to predict the existence of breast cancer based on the clinical variable of the project, including patient demographics, tumor properties and treatment history. The tool nucleus increases the SVMS opportunity to handle complex and non-unique interaction within CLN data, so that the prognosis increases accuracy. Research will use "CLN-SVM-Youtelity" to create a strong prediction model that helps oncologist decisions on educated treatment and improves patient treatment in breast cancer management.

CLN - SVM – Linear

"CLN -SVM Linier" refers to a "Support Vector Machine (SVM) model, which uses a linear core, which is trained on the clinical data. This model is used to predict the existence of breast cancer based on clinical factors such as patient demographics, tumor properties and treatment history in the project. Linear core -Svm -er enables SVM -to separate

linear correlations within CLN data, and help with the classification of live results. The project aims to use the "CLN-SVM linear" to create a strong prediction model that helps the oncologist make decisions and improves the patient's care management by distributing accurate survival projections from "clinical data".

CLN - SVM - Sigmoid

"CLN-SVM-Sigmoid" refers to a "Support Vector Machine (SVM)" model, which uses a sigmoid core, which is trained on clinical data. This model is used to predict the results of the existence of breast cancer based on the clinical variable of the project, including patients, patients, demographics, tumor properties and treatment history. Sigmoid-Core enables SVMs to identify non-led CLN data correlations, which facilitates the recognition of complex patterns associated with different survival results. Research wants to use "CLN-SVM-Sigmoid" to create a strong prediction model that helps the oncologist make decisions and improves the patient's care by distributing accurate survival estimates from "clinical data".

“CLN-WSI-EXP-CNV – NB”

"CLN-WSI-EXP-CNV-NB" reflects a "Naiive Bayes (NB)" trained on the merger of the “entire slideshow (WSI), gene expression (EXP) and Copy Number variation (CNV) Data”. This model is used in the project to predict the existence of the Breast Creation. Synthetizes information from many data sources, such as clinical functions, imaging data and molecular profiles, to gain complete insight into the disease.

“CLN-WSI-EXP-CNV- voting”

"CLN-WSI-EXP-CNV-Voting" represents a clothing model, which uses a voting-eligible classifies that is trained on the synthesis of "Clinical (CLN), the entire slideshow (WSI), Jean Expression (EXP) and Copy Number Variety (CNV) data". This model integrates various independent classification assumptions including "Random Forest, Support Wctor Machine and Decision Tree" to provide final predictions. "CLN-WSI-Exp-CNV-Voting" tries to increase the accuracy and flexibility of prognosis by consolidating the results of several models that are trained on different data. This outfit method improves the prediction of the existence of breast cancer using accurate information from many data sources, and helps oncologists in educated decisions on treatment and increases the results for patient treatment.

“CLN-WSI-EXP-CNV – VGGNets”

“CLN-WSI-EXP-CNV-VGGNets” denotes a deep getting to know model using the VGGNet architecture, skilled on an amalgamation of “medical (CLN), entire slide images (WSI), gene expression (EXP), and replica range variation (CNV) data”. This model is applied in the mission for predicting breast cancer survival through extracting traits from several statistics modalities. VGGNet, identified for its profound design and sturdy characteristic extraction talents, improves the model's capability to determine complex patterns and correlations amongst multi-modal facts. The challenge seeks to make use of “CLN-WSI-EXP-CNV-VGGNets” to create a rather specific prediction model that helps tailor-made remedy planning and enhances patient consequences in breast most cancers care.

“CLN-WSI-EXP-CNV – ResNets”

“CLN-WSI-EXP-CNV-ResNets” denotes a deep learning model using ResNet structure, educated on an amalgamation of “clinical (CLN), whole slide pictures (WSI), gene expression (EXP), and copy wide variety version (CNV) data”. This model is

hired inside the effort to predict breast cancer survival by making use of characteristics amassed from numerous records modalities. ResNet, outstanding for its profound layout and residual connections, augments the model's capability to discern complex patterns and correlations inside multi-modal facts. The assignment is to create a highly correct prediction model making use of "CLN-WSI-EXP-CNV-ResNets" to decorate individualized remedy making plans and improve affected person effects in breast cancer care.

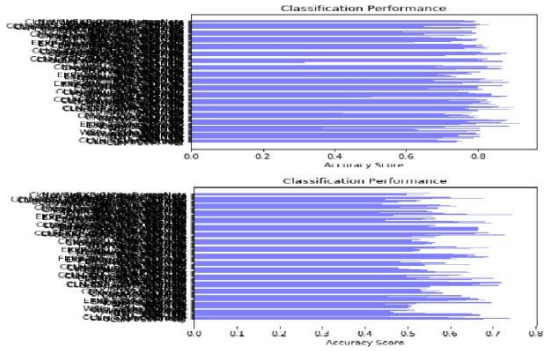
"CLN-WSI-EXP-CNV – DenseNets"

"CLN-WSI-Exp-CNV-Densnet" refers to an intensive learning model, used by the densen Architecture, which includes "CLN), the entire slideshow (WSI), gene expression (EXP) and copy number variations". This model is used to predict the existence of breast cancer by removing information from different data sources in the project. Inches, their dense compounds promote efficient information transfer with patterns and reuse the function of layers, then the model's ability to understand complex conditions in multimodal data improves. Research wants to use the "CLN-WSI-Exp-CNV densan's" to create a strong prediction model that facilitates analog treatment options and increases the patient's consequences in breast cancer care.

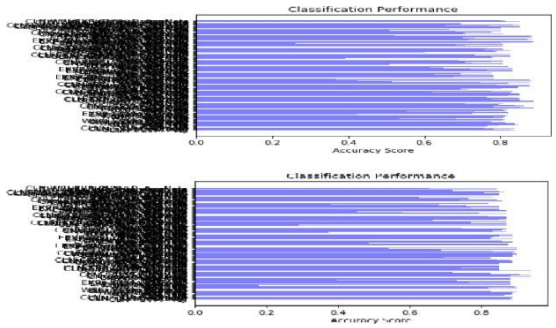
4. EXPERIMENTAL RESULTS

Accuracy: The accuracy of a take a look at refers to its capability to as it should be distinguish among sick and healthful instances. To compare the accuracy of a test, one have to compute the ratio of true positives and proper negatives across all assessed times. This can be expressed mathematically as:

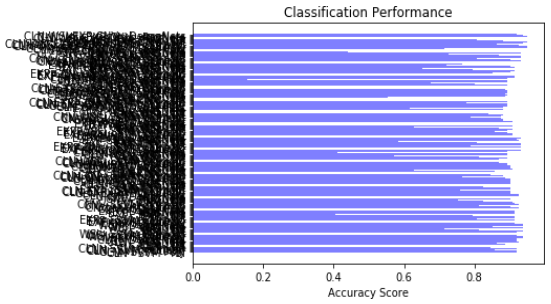
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$



"Fig 3 COMPARISON GRAPH 5,6Year Survival"



"Fig 4 COMPARISON GRAPH 7,8 Year Survival"



"Fig 5 COMPARISON GRAPH 9 Year Survival"

	ML Model	Accuracy	Precision	Recall	F1-Score
0	CLN - SVM - rbf	0.739	0.073	0.429	0.125
1	CLN - SVM - Poly	0.739	0.171	0.467	0.250
2	CLN - SVM - Utility	0.758	0.463	0.528	0.494
3	CLN - SVM - Linear	0.745	0.000	0.000	0.000
4	CLN - SVM - Sigmoid	0.689	0.049	0.154	0.074
...	...	...	...	...	...
175	CLN-WSI-EXP-CNV - NB	0.752	0.727	0.436	0.545
176	CLN-WSI-EXP-CNV- voting	0.863	0.455	0.789	0.577
177	CLN-WSI-EXP-CNV - VGGNets	0.789	0.182	0.462	0.261
178	CLN-WSI-EXP-CNV - ResNets	0.795	0.000	0.000	0.000
179	CLN-WSI-EXP-CNV - DenseNets	0.764	0.091	0.273	0.136

“Fig 6 PERFORMANCE EVALUATION- Survival Prediction Models at 5-Year Survival”

	ML Model	Accuracy	Precision	Recall	F1-Score
0	CLN - SVM - rbf	0.677	0.446	0.750	0.559
1	CLN - SVM - Poly	0.739	0.608	0.776	0.682
2	CLN - SVM - Utility	0.596	0.541	0.563	0.552
3	CLN - SVM - Linear	0.634	0.257	0.826	0.392
4	CLN - SVM - Sigmoid	0.460	0.324	0.393	0.356
...	...	...	...	...	...
175	CLN-WSI-EXP-CNV - NB	0.584	0.534	0.542	0.538
176	CLN-WSI-EXP-CNV- voting	0.615	0.479	0.593	0.530
177	CLN-WSI-EXP-CNV - VGGNets	0.540	0.096	0.467	0.159
178	CLN-WSI-EXP-CNV - ResNets	0.497	0.548	0.455	0.497
179	CLN-WSI-EXP-CNV - DenseNets	0.553	0.014	1.000	0.027

“fig 7 PERFORMANCE EVALUATION- Survival Prediction Models at 6-Year Survival”

	ML Model	Accuracy	Precision	Recall	F1-Score
0	CLN - SVM - rbf	0.839	1.000	0.839	0.912
1	CLN - SVM - Poly	0.826	0.985	0.836	0.905
2	CLN - SVM - Utility	0.783	0.889	0.857	0.873
3	CLN - SVM - Linear	0.839	1.000	0.839	0.912
4	CLN - SVM - Sigmoid	0.758	0.889	0.833	0.860
...	...	...	...	...	...
175	CLN-WSI-EXP-CNV - NB	0.696	0.723	0.900	0.802
176	CLN-WSI-EXP-CNV- voting	0.826	0.964	0.852	0.904
177	CLN-WSI-EXP-CNV - VGGNets	0.851	1.000	0.851	0.919
178	CLN-WSI-EXP-CNV - ResNets	0.814	0.956	0.845	0.897
179	CLN-WSI-EXP-CNV - DenseNets	0.801	0.942	0.843	0.890

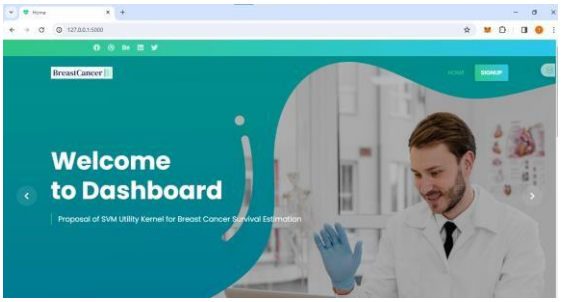
“Fig 8 PERFORMANCE EVALUATION- Survival Prediction Models at 7-Year Survival”

	ML Model	Accuracy	Precision	Recall	F1-Score
0	CLN - SVM - rbf	0.888	1.000	0.888	0.941
1	CLN - SVM - Poly	0.888	1.000	0.888	0.941
2	CLN - SVM - Utility	0.882	0.944	0.925	0.934
3	CLN - SVM - Linear	0.888	1.000	0.888	0.941
4	CLN - SVM - Sigmoid	0.851	0.951	0.889	0.919
...	...	...	...	...	...
175	CLN-WSI-EXP-CNV - NB	0.720	0.788	0.871	0.828
176	CLN-WSI-EXP-CNV- voting	0.851	1.000	0.851	0.919
177	CLN-WSI-EXP-CNV - VGGNets	0.845	0.993	0.850	0.916
178	CLN-WSI-EXP-CNV - ResNets	0.652	0.730	0.840	0.781
179	CLN-WSI-EXP-CNV - DenseNets	0.839	0.985	0.849	0.912

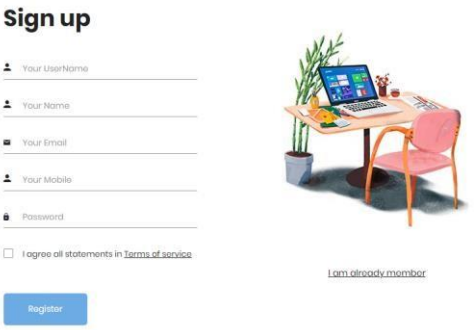
“Fig 9 PERFORMANCE EVALUATION- Survival Prediction Models at 8-Year Survival”

	ML Model	Accuracy	Precision	Recall	F1-Score
0	CLN - SVM - rbf	0.919	1.000	0.919	0.958
1	CLN - SVM - Poly	0.919	1.000	0.919	0.958
2	CLN - SVM - Utility	0.857	0.899	0.943	0.920
3	CLN - SVM - Linear	0.919	1.000	0.919	0.958
4	CLN - SVM - Sigmoid	0.845	0.912	0.918	0.915
...	...	...	...	...	...
175	CLN-WSI-EXP-CNV - NB	0.807	0.837	0.955	0.892
176	CLN-WSI-EXP-CNV- voting	0.950	1.000	0.950	0.975
177	CLN-WSI-EXP-CNV - VGGNets	0.950	1.000	0.950	0.975
178	CLN-WSI-EXP-CNV - ResNets	0.938	0.987	0.950	0.968
179	CLN-WSI-EXP-CNV - DenseNets	0.919	0.967	0.949	0.958

“Fig 10 PERFORMANCE EVALUATION- Survival Prediction Models at 9-Year Survival”

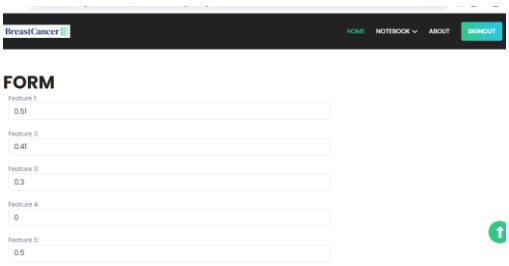


“Fig 11 Home Page”

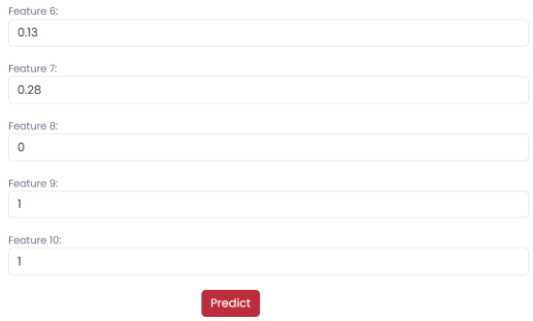


“Fig 12 Sign Up”

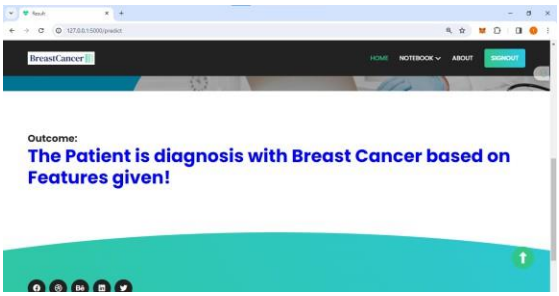
“Fig 13 Sign In”



“Fig 14 Upload Input Data”



"Fig 15 Upload Input Data"



"Fig 16 Predicted Result"

5. CONCLUSION

This study underscores the need of comprehensive and individualized strategies in forecasting breast cancer survival. The research underscores the imperative of algorithmic customisation for precise predictions through a comprehensive evaluation of several machine learning algorithms, including tailored SVMs with utility kernels, conventional models, and deep learning architectures such as VGG-Nets and ResNets.

The integration of many data sources, including clinical records, whole-slide pictures, and genetic databases, facilitates a comprehensive knowledge of breast cancer features. Subsequent improvements, especially the investigation of a voting classifier, augment predictive powers, solidifying the project's dedication to innovation in breast cancer prognosis. The adoption of an intuitive Flask framework enables effortless access and rapid decision-making, hence enhancing patient care.

The initiative offers significant insights on algorithmic appropriateness and highlights the necessity of adaptability and personalization in medical predictive modeling, therefore advancing breast cancer diagnosis and therapy.

6. FUTURE SCOPE

The scope of the plant for the proposed "SVM tool core for breast cancer survival estimates contains a wide selection of clinical, imaging and molecular properties. This includes patient demographics including age, gender and ethnicity; Tumor properties in the form of size, degree and phase; Treatment history; And molecular indicators such gene expression patterns and copy number variants. In addition, depiction properties increased from complete sliding paintings, which provides spatial insight into tumoran atomy and microement. The Utility nucleus of SVM architecture facilitates the integration of complex links and interactions between many properties, so the model improves the future of capacity. The proposed SVM tool core uses different properties that reflect different dimensions of biology and patient properties of breast cancer, which provide accurate and individual prognosis of the results of the existence of breast cancer.

REFERENCES

[1] G. M. Clark, "Do we really need prognostic factors for breast can cer?," Breast Cancer Res. Treat., vol. 30, no. 2, pp. 117–126, 1994.

[2] L. R. Martin, S. L. Williams, K. B. Haskard, and M. R. Dimatteo, "The challenge of patient adherence," Therapeutics Clin. Risk Man age., vol. 1, no. 3, pp. 189–199, Sep. 2005.

[3] C. Curtis et al., "The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel

subgroups," Nature, vol. 486, no. 7403, pp. 346–352, Apr. 2012.

[4] K. Tomczak, P. Czerwinska, and M. Wiznerowicz, "Review the cancer genome atlas (TCGA): An immeasurable source of knowl edge," Wspo»czesna Onkologia, vol. 1A, pp. 68–77, 2015. [Online].Available:http://www.termedia.pl/doi/10.5114/wo.2014.47136

[5] D. Delen, G. Walker, and A. Kadam, "Predicting breast cancer survivability: A comparison of three data mining methods," Artif. Intell. Med., vol. 34, no. 2, pp. 113–127, Jun. 2005. [Online].Available:https://linkinghub.elsevier.com/retrieve/ pii/S0933365704001010

[6] K. Polyak, "Heterogeneity in breast cancer," J. Clin. Investigation, vol. 121, no. 10, pp. 3786–3788, Oct. 2011. [Online]. Available: http://www.jci.org/articles/view/60534

[7] Z. Obermeyer and E. J. Emanuel, "Predicting the future– Big Data, machine learning, and clinical medicine," New England J. Med., vol. 375, no. 13, pp. 1216–1219, Sep. 2016. [Online]. Avail able: http://www.nejm.org/doi/10.1056/NEJMp1606181

[8] G.Viswanath "Improved Light GBM Model Performance Analysis and Comparison for Coronary Heart Disease Prediction", in International Journal of Information Technology and Computer Engineering, 2024, Vol.12, pp.658-672.

[9]D. M. Abd El-Rehim et al., "High-throughput protein expression analysis using tissue microarray technology of a large well-char acterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses," Int. J. Can cer, vol. 116, no. 3, pp. 340–350, Sep. 2005.

[10] G.Swapna, "A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification" Journal of Computer Science, Vol.19, 2023, pp.1203-1211, DOI: https://doi.org/10.3844/jcssp.2023.1203.1211

[11] X. Xu, Y. Zhang, L. Zou, M. Wang, and A. Li, "A gene signa ture for breast cancer prognosis using support vector machine," in Proc. 5th Int. Conf. BioMed. Eng. Inform., 2012, pp. 928–931.

[12] Viswanath Gudditi, "A Smart Recommendation System for Medicine using Intelligent NLP Techniques", 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), 2022, pp.1081-1084.

[13] C. Nguyen, Y. Wang, and H. N. Nguyen, "Random forest classi f ier combined with feature selection for breast cancer diagnosis and prognostic," J. Biomed. Sci. Eng., vol. 06, no. 05, pp. 551–560, 2013. [Online]. Available: http://www.scirp.org/journal/doi.aspx?DOI¼10.4236/jbise.2013.65070

[14] G. Viswanath "Machine-Learning-Based Cloud Intrusion Detection", in International Journal of Mechanical Engineering Research and Technology, 2024, Vol.16, pp 38-52.

[15] M. Khademi and N.S. Nedialkov, "Probabilistic graphical models and deep belief networks for prognosis of breast cancer," in Proc. IEEE 14th Int. Conf. Mach. Learn. Appl., 2015, pp. 727–732.

[16] G.Swapna & K Bhaskar "Early-Stage Autism Spectrum Disorder Detection Using Machine Learning "International Journal of HRM and Organizational Behavior, pp269-283, DOI: https://ijhrmob.org/index.php/ijhrmob/article/view/242

[17] H. Wang, F. Xing, H. Su, A. Stromberg, and L. Yang, "Novel image markers for non-small cell lung cancer classification and survival prediction," BMC Bioinf., vol. 15, no. 1, Dec. 2014, Art. no. 310.[Online]. Available: https://bmcbioinformatics. biomedcentral.com/articles/10.1186/1471–2105-15-310

[18] G.Swapna & K Bhaskar"Malaria Diagnosis Using Double Hidden Layer Extreme Learning Machine Algorithm With Cnn Feature Extraction And Parasite Inflator" International Journal of Information Technology and Computer Engineering, Vol. 12 , 2024, pp536-547,DOI: https://ijitce.org/index.php/ijitce/article/view/704

[19] K.-H. Yu et al., "Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features," Nat. Commun., vol. 7, no. 1, Nov. 2016, Art. no. 12474. [Online].Available:http://www.nature.com/articles/ncomms12474

[20] G. Viswanath "Multiple Cancer Types Classified Using CTMRI Images Based On Learning Without Forgetting Powered Deep Learning Models", in International Journal of HRM and Organizational Behavior, 2024, Vol.12, pp 243-253.

[21] S. J. Giri, P. Dutta, P. Halani, and S. Saha, "MultiPredGO: Deep multi-modal protein function prediction by amalgamating protein structure, sequence, and interaction information," IEEE J. Biomed. Health Inform., vol. 25, no. 5, pp. 1832–1838, May 2021.

[22] G.Viswanath & Dr.G.Swapna "Diabetes Diagnosis Using Machine Learning with Cloud Security" in Cuestiones de Fisioterapia, 2025, Vol.54, No.2, pp 417-431.

[23] N. AryaandS. Saha,"Multi-modal classification for human breast cancer prognosis prediction: Proposal of deep-learning based stacked ensemble model," IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 19, no. 2, pp. 1032–1041, Mar./Apr. 2022.

[24] G.Viswanath & Dr.G.Swapna "Health Prediction Using Machine Learning with Drive HQ Cloud Security" in Frontiersin Health Informatics, 2024. Vol. 13, No. 8, pp. 2755-2761.

[25] D. Sun, A. Li, B. Tang, and M. Wang, "Integrating genomic data and pathological images to effectively predict breast cancer clini cal outcome," Comput. Methods Prog. Biomed., vol. 161, pp. 45–53, Jul. 2018.

[26] B. Tang, A. Li, B. Li, and M. Wang, "CapSurv: Capsule network for survival analysis with whole slide pathological images," IEEE Access, vol. 7, pp. 26 022–26 030, 2019.

[27] N. Arya and S. Saha, "Generative incomplete multi-view progno sis predictor for breast cancer:GIMPP," IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 19, no. 4, pp. 2252–2263, Jul./Aug. 2022.

[28] Y.-F. Zhu, L.-F. Tian, Z.-Y. Mao, and L. Wei, "Mixtures of kernels for SVM modeling," in Proc. Int. Conf. Natural Comput., 2005, pp. 601–607.