Integrative Deep Learning Framework for Enhanced Breast Cancer Prediction

Swathi Vangala *,
CSCI 69500 - M.S. PROJECT
Computational Data Science

Abstract

Despite advances in breast cancer treatment, effective prognosis remains challenging. This study introduces a stacked ensemble predictive model that enhances accuracy by integrating multi-modal data using deep learning. The model combines convolutional neural networks for feature extraction with a stack-based ensemble for prognosis prediction, significantly outperforming existing methods with an AUC of 0.945101 and 91 .9% accuracy at a specificity of 95%. Implemented in Keras 2.2.1 and TensorFlow 1.12, this approach offers a promising tool for early cancer detection and reducing unnecessary treatments.

1 INTRODUCTION

Breast cancer originates from uncontrolled cellular growth in the breast, forming a mass known as a tumor. Tumors are classified into two types: malignant, which are capable of spreading, and benign, which do not spread to other parts of the body. Breast cancer is further categorized into invasive and non-invasive types. Invasive breast cancer spreads to surrounding tissues, whereas non-invasive breast cancer remains confined to the milk ducts or lobules. Ductal carcinoma and lobular carcinoma, starting in the ducts or lobules respectively, are particularly aggressive forms of cancer that significantly impact women's health globally and are leading contributors to cancer-related mortality .

According to data from Cancer.Net, an ASCO-affiliated website, in the United States, invasive breast cancer has impacted 266,120 women, while non-invasive breast cancer has affected 63,960 women. It is estimated that in 2018, breast cancer was responsible for approximately 41,000 deaths, with women making up 98.82% of these cases. Early detection plays a crucial role; 62% of cases are diagnosed at a non-invasive stage with a five-year survival rate of 99%. In contrast, the 5-year survival rate decreases to 85% when cancer spreads to the regional lymph nodes and to 27% when it reaches distant parts of the body. The complexity and variability of clinical outcomes make invasive breast cancer challenging to predict and treat [1].

Enhanced predictive models could significantly aid in assessing survival probabilities, thereby assisting healthcare professionals in making informed treatment decisions [2]. Furthermore, categorizing patients into long-term survivors (those who survive beyond five years) and short-term survivors (those surviving less than five years) can guide personalized treatment options, potentially avoiding patients from the adverse effects of unnecessary therapies. The amalgamation of diverse data sources, including genetic information (such as gene expression and copy number variations) and clinical data (such as age, reproductive history, lifestyle factors), could bolster the predictive accuracy of these models.

1.1 Literature review

Research has consistently highlighted gene expression patterns as pivotal for understanding breast cancer's molecular basis. Van de Vijer and colleagues spearheaded this area with a prognostic model derived from 98 breast cancer patients, which successfully identified 70 gene signatures validated against unseen patient data . Subsequent models by Sun et al. combined gene markers with clinical indicators, further enhancing predictive capabilities [3]. Gevaert et al. expanded the methodology by integrating clinical and genomic data into a Bayesian Network for predicting lymph-node negative breast cancer, indicating a significant leap forward in multi-modal predictive models [3].

However, traditional models often overlooked the interconnections among a patient's genes. Addressing this gap, Xu et al. and Nguyen et al. applied SVM and RF for feature selection, respectively, enhancing the predictability of breast cancer outcomes beyond previous gene signature models [5]. Khademi et al. advanced this approach by creating a probabilistic graphical model that harmonized microarray data with clinical insights, employing PCA for dimensionality reduction and a deep belief network for refined feature extraction [2].

Recent strides in multi-modal deep learning have further revolutionized prognosis predictions by merging various data types. Techniques like Sun et al.'s GPMKL and Zhang et al.'s HI-MKL method illustrate the integration of genomic data with pathological images, significantly improving prognostic accuracy for cancers like breast cancer and Glioblastoma Multiforme [8]. Tang et al.'s introduction of the CapSurv model, utilizing a specialized survival loss function, underscores the potential of deep learning to transcend traditional models, offering superior precision in survival analysis and treatment customization . These advancements confirm that multi-modal models surpass single-source data models, marking a critical evolution in the predictive modeling of complex diseases.

1.2 Drawbacks of Existing Models

Current breast cancer prognosis prediction models offer valuable insights but have inherent limitations. The initial model by Van de Vijer et al. [2] relied solely on gene expression data, ignoring other critical factors like clinical data, copy number alterations, and imaging, which are crucial for accurate diagnosis and prognosis. To address these shortcomings, researchers have developed multi-modal models, such as the Multi-modal Deep Neural Network integrating Multi-dimensional Data (MDNNMD) proposed by Sun et al. [8]. This model employs a score-level fusion strategy, manually adjusting the weight coefficients ($\alpha, \beta, and \gamma$) for the contributions from the respective neural networks: Deep Neural Network - Experimental (DNN-Expr), Deep Neural Network - Copy Number Alteration (DNN-CNA), and Deep Neural Network - Clinical (DNN-Clinical). Specifically, the model defines the output as:

$$O_{DNNMD} = \alpha \cdot O_{DNN-Expr} + \beta \cdot O_{DNN-CNA} + \gamma \cdot O_{DNN-Clinical}$$
$$s.t. \alpha + \beta + \gamma = 1,$$
$$\alpha \geq 0, \beta \geq 0, \gamma \geq 0 \quad (1)$$

with all coefficients positive and manually set for the METABRIC dataset, which is a time-consuming process requiring numerous combinations and retraining sessions.

Additionally, while MDNNMD uses a basic deep neural network (DNN) for feature extraction, a convolutional neural network (CNN) might better extract features from each data modality, given CNNs' proficiency in identifying latent features within datasets . Unlike MDNNMD, which relies on a straightforward neural network for data fusion and output prediction, our proposed model utilizes a stacked ensemble technique. This approach not only automates the coefficient adjustment, eliminating manual effort, but also leverages CNNs for superior feature extraction. Research indicates that stacked ensemble techniques generally surpass simpler machine learning architectures in predictive accuracy. Consequently, our model promises enhanced predictive performance over existing breast cancer prognosis models.

1.3 Motivation for the Proposed Model

Drawing on the multi-modal deep neural network integrating multi-dimensional data (MDNNMD) proposed by Sun et al., which effectively utilizes gene expression profiles, copy number alteration (CNA) profiles, and clinical data through score-level fusion for breast cancer prognosis, we have developed an advanced model using a stacked-based ensemble approach. This approach leverages the strengths of stacked models, known for their ability to reduce variance and bias, thereby enhancing predictive performance.

Our model employs convolutional neural networks (CNNs) for robust feature extraction from each data modality. These features are then fed into a stacked layer that combines multiple machine learning techniques, including Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting Machine (GBM), to process the extracted features. This ensemble

method is particularly effective on the imbalanced dataset used in our study, improving handling and predictive accuracy.

Key improvements in our model include the use of CNNs, which are superior to the deep neural networks previously used due to their ability to uncover deeper hidden features within the data. These extracted features form a complex stacked feature set that feeds into the stacked ensemble layer. Our model automates the adjustment of fusion coefficients, eliminating the manual tuning required in the MDNNMD model. This automation, coupled with strategic feature extraction and advanced ensemble methods, has shown to significantly outperform existing models. Operations such as ReLU activation and batch normalization are integral to our CNN layers, enhancing model training and stability. Comprehensive ten-fold cross-validation confirms the robustness and superior performance of our proposed approach.

2 METHODOLOGY

2.1 Data Collection:

The dataset employed for our project is sourced from the METABRIC 1 dataset, available on the USTC-HIlab/MDNNMD repository on GitHub, which offers a pre-processed version ready for use. This dataset comprises records from 1,980 patients who participated in the METABRIC trial, incorporating multi-modal data that includes gene expression profiles, copy number alteration (CNA) profiles, and clinical data. The dataset classifies patients into two categories based on their survival time post-diagnosis: long-time survivors (1,489 patients), who lived beyond five years, and short-time survivors (491 patients), who did not. The median age at diagnosis across the cohort is 61 years.

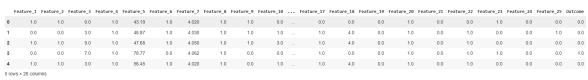

Figure 1: Clinical Data

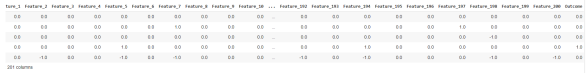

Figure 2: CNV Data

Our dataset underwent a comprehensive preprocessing regimen to ensure the quality and robustness of the subsequent analysis. The following steps were meticulously carried out:

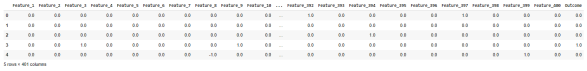

Figure 3: Gene Expression Data

Metric	Value
Cut-off (years)	5
Total # of patients	1980
Long-time survivors	1489
Short-time survivors	491
Median age at diagnosis	61
Average survival (months)	125.1

Table 1: Overview of the METABRIC Breast Cancer Dataset

2.2 Feature Selection

The METABRIC dataset contains gene expression profiles, CNA profiles, and clinical data for each of its 1980 patients, with approximately 24,000, 26,000, and 27 features respectively. Given the high dimensionality relative to the sample size, deep learning methods might not perform optimally, as highlighted by [4]. To address this, Sun et al. [7] implemented the well-regarded mRMR feature selection algorithm [6] to reduce dimensionality. They adopted an incremental feature selection strategy, beginning with the top 100 features to assess model performance and incrementally increasing in steps of 100 up to 500 features. This process helped determine the optimal number of features by evaluating AUC values at each increment. Ultimately, the best performing feature sets consisted of 400 gene expression profiles, 200 CNA profile features, and 25 clinical features, which were used in the existing MDNNMD model [6] and the proposed stacked-based ensemble model. Details of the selected features are documented in Table 2.

The clinical features consist of age at diagnosis, size, lymph nodes positive, grade, inferred menopausal state,type of breast surgery, type of therapy the patient has gone through (Chemotherapy, Hormone Therapy and Radio Therapy) and some pathological details of genes and tumours.

Table 2: SELECTED FEATURES FOR THE MODEL

Data Category	Total Features	Selected Features
Clinical	27	25
Gene Expression	24368	400
CNA	26298	200

2.3 A Convolutional Neural Network Based Prediction Model For Uni-modal Dataset:

The advanced convolutional neural network (CNN) model developed for this study is engineered to predict the prognosis of human breast cancer by effectively processing and extracting features from multimodal datasets. Each modality of input data is processed through a convolution layer that utilizes a defined number of filters to generate feature maps. These feature maps result from the convolution operation, which involves element-wise multiplication and summation between the kernel and the input data values.

2.3.1 Convolution Layer:

In our model, the weights of the convolution filters are initialized using a Glorot normal initializer, ensuring an optimal distribution with a mean of zero and a standard deviation scaled according to the inverse square root of the sum of the number of input and output units. This setup facilitates efficient training dynamics. Each convolution layer incorporates biases initialized at a constant value of 0.1 and utilizes a stride of two. This stride setting allows the convolution operation to cover the entire input matrix effectively, while the use of 'same' padding ensures that the output feature map maintains the same dimensions as the input.

2.3.2 Feature Maps:

Each convolution layer incorporates biases initialized at a constant value of 0.1 and utilizes a stride of two. This stride setting allows the convolution operation to cover the entire input matrix effectively, while the use of 'same' padding ensures that the output feature map maintains the same dimensions as the input.

2.3.3 Flatten, Dense and Output Layers:

After convolution, the feature map is flattened into a vector and fed into a densely connected layer comprising 150 hidden units. This dense layer uses ReLU activation to introduce non-linearity, enhancing the model's ability to learn complex patterns in the data. A dropout rate of 50% is applied after the dense layer to prevent overfitting, especially crucial given the high dimensionality of the gene expression and CNA profiles.

The network's output layer employs a sigmoid activation function to deliver the final binary classification, indicative of the breast cancer prognosis. The model predicts two classes: short-time and long-time survivors, using binary cross-entropy as the loss function, augmented with L2 regularization to mitigate overfitting. This regularization technique is essential for maintaining generalization in deep learning models.

The architecture of the CNN is structured simply to avoid overfitting, particularly important when dealing with limited datasets. The CNN consists of an input layer, a convolution layer, a flattening step, a fully connected dense layer, and a final output layer. Parameters including the number of filters, kernel size, and the stride in the convolution layer, as well as the number of units in the dense layer, are meticulously optimized to achieve the highest possible AUC score. The model's learning rate is set at 0.001, ensuring steady convergence during training.

Table 3: Detailed Parameter Configuration of CNN Model

Parameter	Value
# Convolution Layers	1
Convolution Filter Size	15
# Filters	4
Stride Size	2
Padding	Same
Activation (Conv. Layer)	ReLU
# Hidden Layers	1
# Units in Hidden Layer	150
Batch Size	8
Epochs	50
Activation (Output Layer)	Sigmoid
Loss Function	Binary Cross-Entropy + L2

2.4 Stacked-based Prediction Model For Multi-modal Data:

The proposed model employs a two-stage architecture designed to enhance the prediction of cancer prognosis by effectively leveraging diverse data modalities.

2.4.1 Stage One - Feature Extraction

Given the distinctive feature representations inherent to different types of data, our approach avoids

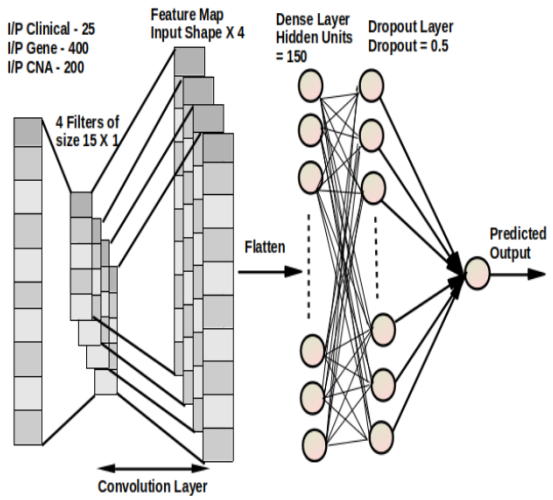

Figure 4: CNN FOR UNI-MODAL DATASET

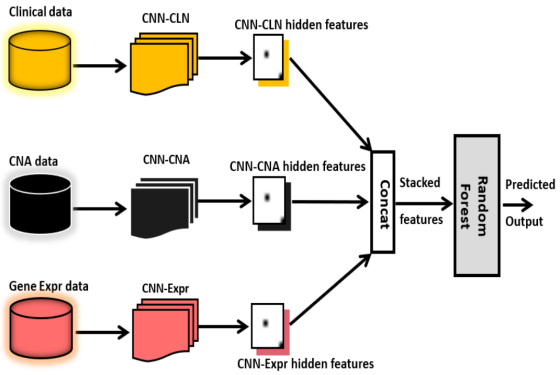

Figure 5: STACKED-BASED ENSEMBLE MODEL ARCHITECTURE

direct combination of multiple data sources for input into the model. Instead, we use separate Convolutional Neural Networks (CNNs) for each uni-modal dataset—clinical, gene expression, and copy number alterations (CNA). Each CNN processes its respective dataset to capture unique insights. The METABRIC dataset, preprocessed as per the methodologies of Sun et al.[9], serves as the foundation for this analysis. We design specific CNN architectures for each modality: CNN-Clinical, CNN-Expr, and CNN-CNA, training each network independently with the Area Under the Curve (AUC) serving as the primary performance metric. Features are extracted from the hidden layers of each CNN and then merged to create a unified set of stacked features, which carry comprehensive multi-modal insights.

2.4.2 Stage Two - Classification:

The compiled stacked features from the first stage are then utilized as inputs in the second stage of the model. This phase involves the application of various machine learning classifiers to conduct the final classification. The model's flexibility allows for the use of any suitable machine learning classifier for binary classification tasks. Classifiers such as Support Vector Machine (SVM), Random Forest (RF), Naive Bayes (NB), and Logistic Regression (LR) are evaluated in turn to determine their efficacy with the stacked features. The selection and performance assessment of each classifier are meticulously executed, with the best-performing classifier's output being adopted as the final model prediction.

3 EXPERIMENTAL DESIGN:

To address the challenges associated with the small dataset size, we implemented ten-fold cross-validation to robustly assess our proposed model. The dataset, consisting of 1980 patients, is randomly partitioned into ten subsets. In each validation cycle, nine subsets are combined to form the training set, while the remaining subset is used as the test set. Additionally, the training set is further split into an 80% training portion and a 20% validation portion for internal model validation.

The model operates in two distinct stages:

Stage One: Initially, individual CNNs are trained for each data modality—clinical data, gene expression, and copy number alterations (CNA). Optimal parameters for each CNN are determined based on the Area Under the Curve (AUC) metric, which serves as the primary criterion for model tuning.

Stage Two: Following the initial training phase, features are extracted from either the hidden or output layers of each CNN. These features are then amalgamated to create a comprehensive stacked feature set. This feature set is subsequently utilized as input for the second stage, which comprises various machine learning classifiers like SVM (Support Vector Machine), RF (Random Forest), Naive Bayes, and Logistic Regression (LR), to perform the final classification.

For the evaluation of model performance, the Receiver Operating Characteristics (ROC) curve is employed. This curve plots the false positive rate (1-Specificity) against the true positive rate (Sensitivity) across varying decision thresholds, facilitating the calculation of the AUC value to gauge model effectiveness. Additionally, several other

evaluation metrics are utilized, including Sensitivity (Sn), Specificity (Sp), Accuracy (Acc), Precision (Pre), and Matthew's Correlation Coefficient (MCC). These metrics provide a comprehensive framework for measuring the performance of the model across different aspects of classification accuracy.

Sensitivity (Sn)

$$Sn = \frac{TP}{TP + FN} \tag{2}$$

Specificity (Sp)

$$Sp = \frac{TN}{TN + FP} \tag{3}$$

Precision (Pre)

$$Pre = \frac{TP}{TP + FP} \tag{4}$$

Accuracy (Acc)

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Matthews Correlation Coefficient (MCC)

$$Mcc = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FN) \times (TP+FP) \times (TN+FN) \times (TN+FP)}}$$

Our loss function can be written as:

$$\begin{aligned} &L(y_t, \hat{y}_t) \\ &= \frac{-1}{N} \sum_{i=0}^{N} [y_t(i) \log \hat{y}_t(i) - (1 - y_t(i)) \log(1 - \hat{y}_t(i))] \\ &+ \frac{1}{2} \lambda \sum_{k=1}^{K} \sum_{j=1}^{n_k} \sum_{i=1}^{m_k} {w_{ij}}^{k^2} \end{aligned}$$

where L is the cost function that combines the cross-entropy and the regularization term. y_t and $\hat{y}_t$ are the actual label and predicted label, respectively. N is the batch size. $W^k = \left(w_{ij}^k\right)_{m_k \times n_k}$ is the k_{th} weight matrix and K is the number of weight matrices in the CNN model.

4 RESULTS

In this segment, we've reviewed the outcomes of various approaches. We've evaluated CNN-based techniques against DNN-based methods using uni-modal data, assessed the impact of stacked feature selection, and compared CNN-based methods with uni-modal data alongside the stacked-based RF method utilizing multi-modal data.

4.1 Comparison of CNN Based Methods with DNN Based Methods on Uni-modal Data:

To assess the superiority of CNNs over DNNs in predicting breast cancer outcomes, we evaluated both models using ROC curves and various performance metrics as detailed in Table 4. The results show that CNN models outperform DNN models across different datasets: CNN-Clinical, CNN-Expr, and CNN-CNA demonstrate AUC improvements of approximately 14.42%, 24.58%, and 20.39% over DNN-Clinical, DNN-Expr, and DNN-CNA, respectively. These findings highlight that the features or outputs from CNNs are more effective, which justifies their use as the input for the stacked feature set in the ensemble model. The performance metrics across other categories also reflect better outcomes for CNNs compared to DNNs.

Table 4: Performance metrics for CNN and DNN models across datasets without standard deviations.

Model	Acc.	Prec.	Recall	MCC	AUC
CNN Clinic	76.33%	79.25%	71.66%	31.89%	84.40%
CNN CNV	71.93%	73.65%	68.97%	25.34%	79.44%
CNN Expr	89.86%	87.44%	93.35%	47.90%	96.38%
DNN Clinic	76.67%	72.61%	11.40%	16.59%	69.98%
DNN CNV	74.75%	50.01%	15.89%	14.21%	59.05%
DNN Expr	80.10%	64.84%	44.80%	41.26%	71.80%

4.2 Performance metrics for various stacked models

We compare the performances of stacked-based ensemble method designed using two different ways of stacked fea- ture selection. In the first approach, predicted outputs of individual CNNs are combined and passed to the second stage as features for breast cancer prognosis prediction. In the second approach, the hidden layer features of individual CNNs are combined and given to the second stage as features. The AUC values and other performance metrics are measured and presented in Table 5. Here, STACKED RF , STACKED SVM, STACKED Naive Bayes, STACKED Logistic Regression, STACKED XGBoost, STACKED LightGBM, STACKED CatBoost are similar to the models defined in previous section. We have se- lected predicted outputs or hidden layer features of individual uni-modal CNNs and concatenated them as stacked features.

After reviewing the performance metrics in Table 5, we have chosen hidden layer outputs as the feature set for our proposed model. The anal-

Table 5: Performance metrics for various stacked models without standard deviations.

Model	Acc	Prec	Recall	F1	AUC
XGB	90.71%	80.36%	78.95%	79.65%	94.10%
LGBM	90.91%	84.16%	74.56%	79.07%	94.75%
CB	91.92%	85.58%	78.07%	81.65%	95.07%
SVM	90.00%	97.00%	60.00%	74.00%	94.00%
NB	85.00%	64.00%	79.00%	71.00%	89.00%
LR	87.00%	73.00%	68.00%	71.00%	90.00%
RF	**91.92%**	**87.76%**	**75.44%**	**81.13%**	**94.51%**

ysis shows that the STACKED Random Forest (RF) model outperforms the STACKED SVM. Given that RF is an ensemble method, it generally yields better results on imbalanced datasets, as supported by existing research. While SVM is a robust classifier by itself, the combined predictive strength of multiple classifiers in RF leads to superior performance. Consequently, we have opted for STACKED RF, utilizing features derived from CNNs, as the architecture for our model.

4.3 Performance Metrics of Stacked RF and CNN Models on Different Modalities

Table 6: Performance Metrics of Stacked RF and CNN Models on Different Modalities

Model	Acc	Prec	Recall	F1 Score	AUC
RF	91.92%	87.75%	75.44%	81.13%	94.51%
Clinic	76.33%	79.25%	71.66%	31.89%	84.40%
CNV	71.93%	73.65%	68.97%	25.34%	79.44%
Expr	89.86%	87.44%	93.35%	47.90%	96.38%

The ensemble approach utilizing multi-modal data in the stacked-based RF model demonstrates superior performance compared to uni-modal CNN methods. Referencing Table 6, the stacked-based RF model surpasses the individual CNN models. Specifically, the stacked RF model achieves an AUC value of 94.51%, which is approximately 10.11%, 15.13%, and 12.07% higher than the AUC values of CNN-Clinical, CNN-Expr, and CNN-CNV, respectively. Additionally, the stacked RF model exhibits enhanced performance across other metrics such as accuracy, precision, recall, and F1 score, confirming its effectiveness over the uni-modal CNN approaches.

4.4 Comparison of AUC values of CNN and DNN models

The bar chart visualizes the comparison of AUC (Area Under the Curve) values between CNN (Convolutional Neural Network) and DNN (Deep Neural Network) models across three different types of data: Clinical, CNV (Copy Number Variation), and Gene Expression. Here's a detailed explanation:

Clinical Data: CNN achieves a significantly higher AUC of 84.4 compared to the DNN, which has an AUC of 69.98. This suggests that the CNN model is substantially more effective in predicting outcomes using clinical data.

CNV Data: Both models perform comparably on CNV data, with the CNN showing a slight advantage with an AUC of 59.05 over the DNN's AUC, also at 59.05. This indicates a similar performance by both models for CNV data, suggesting that neither model has a distinct advantage in this context.

Gene Expression Data: For gene expression data, the CNN significantly outperforms the DNN, achieving an AUC of 96.38 compared to the DNN's 71.8. This large difference highlights the CNN's superior capability in handling and predicting outcomes from complex gene expression data. Overall, the chart effectively demonstrates that CNN models generally provide better performance in terms of AUC across different types of data, particularly in clinical and gene expression datasets, confirming its robustness and efficacy in handling diverse biomedical data modalities.

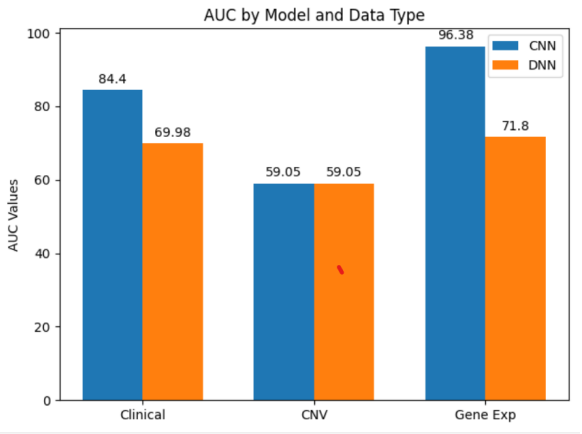

Figure 6: AUC Comparison of CNN and DNN models

4.5 Statistical Significance Test

To validate the statistical significance of the breast cancer prognosis predictions made by our stacked-based ensemble model, a t-test was conducted comparing the performance metrics of the MDNNMD model and our proposed model. Given that the t-test adheres to the Student's t-distribution, it is

well-suited for small sample sizes. This analysis utilized the accuracy metric (Acc) from both the MDNNMD and our proposed model applied to the METABRIC dataset, which includes 1980 patients. The t-test yielded a t-value of -14.93 and a p-value of 0.00, indicating a statistically significant difference in performance between the two methods. Additionally, a similar t-test was performed on the TCGA-BRCA validation dataset, resulting in a t-value of -109.54 and a p-value of 0.00. These tests were carried out using the stats.ttest_ind function from the scipy library, reinforcing the statistical significance of our model's results.

5 Conclusion:

Breast cancer remains the most prevalent cancer worldwide and is a leading cause of cancer-related mortality. The complexity of genetic markers and varying patient characteristics complicate its prognosis. Addressing this, our research introduces an advanced stacked-based ensemble model that leverages Convolutional Neural Networks (CNNs) and Random Forest (RF) classifiers to enhance the prediction accuracy of breast cancer patient survival. Our model uniquely integrates multi-modal data inputs, including gene expression profiles, copy number alterations, and clinical data.

We constructed three distinct CNNs to process each data type separately, extracting essential features from their hidden layers. These features are then utilized by a Random Forest classifier to predict patient outcomes. This method has demonstrated superior performance over traditional models that use single data sources and other multimodal approaches, making it a promising tool for broader applications in medical diagnostics, including other complex diseases that benefit from multimodal analysis.

Despite the promising results, further validation across additional breast cancer datasets is necessary. The METABRIC dataset, currently employed with 1980 samples, is relatively small for robust machine learning applications. Expanding our dataset will likely enhance model accuracy and reliability. Future enhancements could include incorporating histopathological images and other biological markers like gene methylation and miRNA expression levels to enrich the model's input, potentially offering a more comprehensive tool for clinical decision-making in cancer prognosis.

References

[1] J. Ferlay, C. Héry, P. Autier, and R. Sankaranarayanan. 2010. *Global Burden of Breast Cancer*. Springer New York.

[2] Jianjiong Gao, Bülent Arman Aksoy, Ugur Dogrusoz, Gideon Dresdner, Benjamin Gross, S Onur Sumer, Yichao Sun, Anders Jacobsen, Rileen Sinha, Erik Larsson, et al. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cbioportal. *Science signaling*, 6(269):pl1–pl1.

[3] Olivier Gevaert, Frank De Smet, Dirk Timmerman, Yves Moreau, and Bart De Moor. 2006. Predicting the prognosis of breast cancer by integrating clinical and microarray data with bayesian networks. *Bioinformatics*, 22(14):e184–e190.

[4] Mahmoud Khademi and Nedialko S Nedialkov. 2015. Probabilistic graphical models and deep belief networks for prognosis of breast cancer. In *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, pages 727–732. IEEE.

[5] Leslie R Martin, Summer L Williams, Kelly B Haskard, and M Robin DiMatteo. 2005. The challenge of patient adherence. *Therapeutics and clinical risk management*, 1(3):189–199.

[6] Dongdong Sun, Minghui Wang, and Ao Li. 2018. A multimodal deep neural network for human breast cancer prognosis prediction by integrating multidimensional data. *IEEE/ACM transactions on computational biology and bioinformatics*, 16(3):841–850.

[7] Katarzyna Tomczak, Patrycja Czerwińska, and Maciej Wiznerowicz. 2015. Review the cancer genome atlas (tcga): an immeasurable source of knowledge. *Contemporary Oncology/Współczesna Onkologia*, 2015(1):68–77.

[8] Marc J Van De Vijver, Yudong D He, Laura J Van't Veer, Hongyue Dai, Augustinus AM Hart, Dorien W Voskuil, George J Schreiber, Johannes L Peterse, Chris Roberts, Matthew J Marton, et al. 2002. A gene-expression signature as a predictor of survival in breast cancer. *New England Journal of Medicine*, 347(25):1999–2009.

[9] Xiaoyi Xu, Ya Zhang, Liang Zou, Minghui Wang, and Ao Li. 2012. A gene signature for breast cancer prognosis using support vector machine. In *2012 5th International conference on biomedical engineering and informatics*, pages 928–931. IEEE.