

An Ensembled Deep Learning Model for Mental Disorder Prediction

Prajna Paramita Debata¹, Midhunchakkaravarthy², Brojo Kishore Mishra³

¹ Department of Engineering and Science, Lincoln University College
Selangor, Malaysia; ² Faculty of AI Computing and Multimedia, Lincoln University College
Selangor, Malaysia; ³ Department of Computer Science and Engineering, NIST University
Berhampur, India

pdf.prajna@lincoln.edu.my

midhun@lincoln.edu.my

brojokishoremishra@gmail.com

Abstract: The detection of mental disorders is an evolving interdisciplinary domain that seeks to unravel the intricate neurobiological foundations of mental illnesses. With mental health challenges contributing increasingly to the global disease burden, the demand for accurate and reliable diagnostic systems has become more pressing than ever. Conditions such as schizophrenia and bipolar disorder, which impact a considerable share of the population, are particularly difficult to diagnose due to their overlapping genetic markers and clinical manifestations. To overcome these challenges, we introduce a deep learning-driven framework designed to pinpoint critical gene biomarkers using gene expression datasets from the Gene Expression Omnibus (GEO). The framework integrates the kernel-applied Fisher score (KFScore) for robust feature selection, and further Convolutional Neural Network (CNN) based RESNET is used to classify the schizophrenia and bipolar disorder data. The model's effectiveness is measured by its classification accuracy, the number of key genomic features it extracts, as well as its sensitivity and specificity.

Keywords: Mental disorder; Kernal Fisher Score; Deep learning; significant biomarker; Convolutional neural network

Introduction

The pharmacological management of schizophrenia and bipolar disorder is associated with considerable adverse effects. These include motor complications such as extrapyramidal reactions, tardive dyskinesia, and akathisia; endocrine, metabolic, and systemic disturbances such as hyperprolactinemia, weight gain, hyperlipidemia, hyperglycemia/diabetes, hypothyroidism (particularly with lithium), hepatic toxicity (e.g., valproate), and cognitive impairment; as well as other adverse outcomes including sedation, orthostatic hypotension, anticholinergic effects, cardiac conduction abnormalities, seizures, and, in rare but severe cases, neuroleptic malignant syndrome, agranulocytosis, myocarditis, and acute pancreatitis [1]. The frequency and severity of these side effects vary across drug classes—typical antipsychotics are more commonly associated with movement-related disorders, whereas atypical agents tend to induce metabolic and endocrine complications [2]. These risks strongly influence clinical decision-making and shape the individualized risk–benefit analysis in treatment planning.

Schizophrenia, which affects males and females at comparable rates, is characterized by hallucinations, delusions, cognitive disorganization, negative symptoms, and mood disturbances, all of which impair daily functioning and social relationships [3]. Its pathogenesis is believed to involve dysregulated neurotransmission, driven by an interplay of genetic and environmental factors [4].

Bipolar disorder, affecting approximately 2% of the global population, is a major psychiatric illness defined by alternating episodes of mania and depression [5]. Genetic predisposition is a key contributor, and differentiation from major depressive disorder can be challenging, often leading to delayed diagnosis, misdiagnosis, and heightened suicide risk [6][7]. Clinical features include mood instability, irritability, anxiety, and disturbances in sleep. Standard treatment strategies typically involve antipsychotics and mood stabilizers, including lithium and valproate.

The presented model for disease prediction and classification consists of in two parts.

Part I: A novel key gene extraction algorithm, Kernel Fisher Score (KFScore), is applied to identify the most significant genetic biomarkers.

Part II: A deep CNN (based on RESNET) is applied to classify the schizophrenia and bipolar disorder data.

Related work

Several representative approaches illustrate this process of mental disorder detection. Chen et al. [8] apply deep neural networks (DNNs) to genomic datasets to differentiate schizophrenia from bipolar disorder. Mashayekhi Shams & Jabbari et al. [9] introduce an end-to-end CNN–LSTM pipeline with 15- and 16-layer architectures for automated schizophrenia detection using neuroimaging data. Keshavan et al. [10] design a concise clinical scale that quantifies psychotic and affective symptomatology across illness trajectories. Arribas et al. [11] employ independent component analysis (ICA) and default mode network (DMN) measures to identify schizophrenia biomarkers. Karthik & Sudha [12] utilize deep neural networks for predicting both bipolar disorder and schizophrenia, while Fond et al. [13] apply CART decision trees for schizophrenia diagnosis. Trakaddis et al. [14] adopt gradient boosted trees with regularization to classify genomic schizophrenia profiles. Karrer et al. [15] incorporate meta-analytic cognitive priors to forecast cognitive outcomes in schizophrenia, and Talpalaru et al. [16] leverage logistic regression, support vector machines (SVM), and random forests to predict symptom severity. Finally, Salvador et al. [17] use Ridge Regression for schizophrenia classification. While prior studies have predominantly emphasized either deep learning or traditional machine learning paradigms, our work offers a direct comparative analysis of both approaches in the classification of bipolar disorder and schizophrenia. Furthermore, we propose a novel ensemble framework—KFScore CNN—which integrates a Kernel Fisher Score (KFScore) filter for effective feature extraction and CNN for classification.

Proposed Hybridized Approach

This study presents a two-phase ensemble framework aimed at selecting key genomic features and classifying disease-related data:

Figure 1 illustrates the overall workflow of the KFScore driven SC MBO CNN pipeline. Below is a detailed breakdown of the proposed methodology:

1. Data Preprocessing

In Table 1 shows the detailed elaboration of the dataset. Apply min–max normalization to both the bipolar disorder and schizophrenia datasets, followed by partitioning the normalized data into training and testing subsets.

2. Feature Selection

Employ the Kernel Fisher Score (KFScore) filter to extract the most discriminative genetic features.

3. Model Training and Optimization

Use the KFScore-selected features as input to a Convolutional Neural Network (CNN) classifier. Perform hyperparameter optimization during training.

4. Evaluation

Assess the performance of the optimized KFScore–CNN model on the test dataset, with classification accuracy as the primary evaluation metric.

5. Comparative Study

Compare the proposed KFScore–CNN framework against conventional machine learning classifiers to demonstrate its relative effectiveness

Table 1 The detailed elaboration of the dataset.

Data Source	Sample Nos.	Feature Nos.	Classes
GSE12649 [18]	102 no. of samples (33 bipolar disorder patients, 35 schizophrenia patients, 34 control samples)	22,283	Binary (0 means Control, 1 means Case)

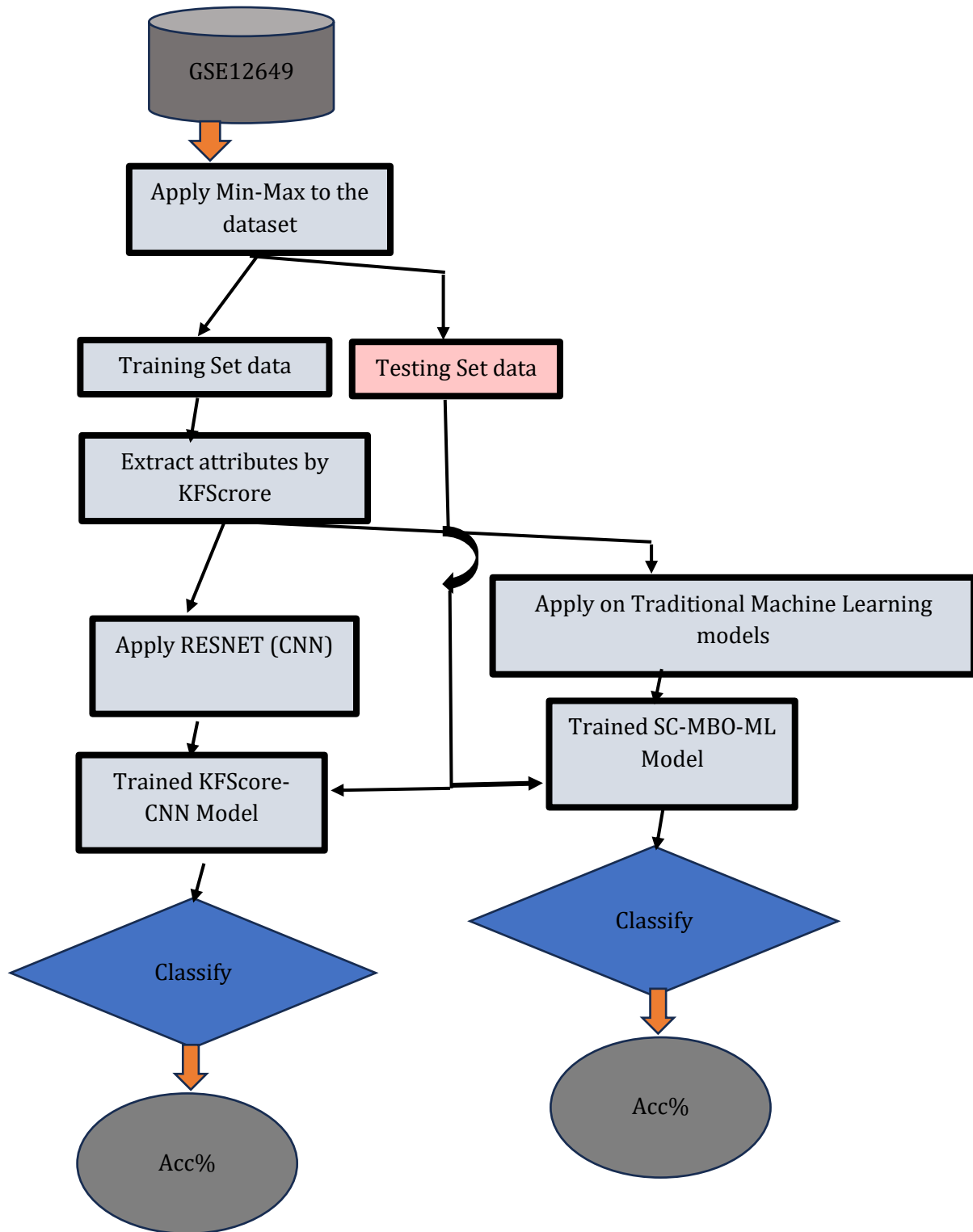


Figure 1. The overall architecture of the presented KFScore- CNN model.

Result Discussion (in GSE12649)

The kernel-based Fisher Score approach is applied to address this task, aiming to identify the most relevant and interdependent features from the input dataset. In each method, a score is assigned to every attribute, after which the attributes are ranked in descending order. The top-ranked features are then selected for further analysis. Table 2 shows the detailed comparison of the dataset's performance with proposed method and other ML algorithms.

Table 2 The detailed comparison of the dataset performance with proposed method and other ML algorithm.

Data Used	Methods used	Accuracy%	Sensitivity%	Specificity%
GSE12649	KFScore-CNN	89.12	91	88.38
	KFscore-KELM	82.15	82.10	80.76
	KFscore -ELM	80.19	80.88	81
	KFscore -MLP	76.27	77.27	75.10

Conclusions

In this study, we introduce an ensemble deep learning framework specifically developed for the GSE12649 dataset. Here, first, the KFScore algorithm is applied to extract the most informative genomic features. A key strength of KFScore lies in its ability to filter out irrelevant genes from the high-dimensional feature space by leveraging a kernel transformation. Then, a Convolutional Neural Network (CNN) based RESNET is used to classify the schizophrenia and bipolar disorder data. The effectiveness of the framework is assessed using multiple performance indicators, including classification accuracy, the number of significant genes selected, sensitivity, specificity. From the above experimental analysis, it is clear that the presented method outperforms the others.

References

1. S. Kiliçarslan and E. Dönmez, Improved multi-layer hybrid adaptive particle swarm optimization based artificial bee colony for optimizing feature selection and classification of microarray data. *Multimedia Tools and Applications*, 83(26), 67259-67281,2024.
2. S. Osama, H. Shaban and A. A. Ali, Gene reduction and machine learning algorithms for cancer classification based on microarray gene expression data: A comprehensive review. *Expert Systems with Applications*, 213, 118946, 2023.
3. F. Alharbi and A. Vakanski, Machine learning methods for cancer classification using gene expression data: A review. *Bioengineering*, 10(2), 173, 2023.

4. M. Mohammed, H. Mwambi, I. Mboya, M.K. Elbashir, and B. Omolo, A stacking ensemble deep learning approach to cancer type classification based on TCGA data. *Scientific reports*, 11(1), 15626, 2021.
5. L. Rukhsar, W.H. Bangyal, M. S. Ali Khan, A. A. Ag Ibrahim, K. Nisar, and D.B. Rawat, Analyzing RNA-seq gene expression data using deep learning approaches for cancer classification. *Applied Sciences*, 12(4), 1850, 2022.
6. A. A. Joshi and R. M. Aziz, Deep learning approach for brain tumor classification using metaheuristic optimization with gene expression data. *International Journal of Imaging Systems and Technology*, 34(2), e23007, 2024.
7. Y. Wang, Q. Chen, H. Shao, R. Zhang, and H. Shen, Generating bulk RNA-Seq gene expression data based on generative deep learning models and utilizing it for data augmentation. *Computers in Biology and Medicine*, 169, 107828, 2024.
8. X. Chen, et al. "Classification of Schizophrenia, Bipolar Disorder and Major Depressive Disorder with Comorbid Traits and Deep Learning Algorithms.", 2024.
9. A. M. Mashayekhi, and S. Jabbari. "A deep learning approach for diagnosis of schizophrenia disorder via data augmentation based on convolutional neural network and long short-term memory." *Biomedical Engineering Letters*, 1-13, 2024.
10. S. M. Keshavan, et al. "A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: the Schizo-Bipolar Scale." *Schizophrenia research* 133.1-3: 250-254, 2011.
11. I. J. Arribas, D. C. Vince. and T. Adali, "Automatic Bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from fMRI data." *IEEE Transactions on Biomedical Engineering* 57. 12: 2850-2860, 2010.
12. S. Karthik, and M. Sudha, "Predicting bipolar disorder and schizophrenia based on non-overlapping genetic phenotypes using deep neural network." *Evolutionary Intelligence* 14.2: 619-634, 2021.
13. G. Fond, et al. "Machine learning for predicting psychotic relapse at 2 years in schizophrenia in the national FACE-SZ cohort." *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 92: 8-18, 2019.
14. J. Y. Trakadis, et al. "Machine learning in schizophrenia genomics, a case-control study using 5,090 exomes." *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 180.2: 103-112, 2019.
15. M. T. Karrer, et al. "Brain-based ranking of cognitive domains to predict schizophrenia." *Human brain mapping* 40.15: 4487-4507, 2019.
16. A. Talpalaru, N. Bhagwat, GA. Devenyi, M. Lepage, MM. Chakravarty, Identifying schizophrenia subgroups using clustering and supervised learning. *Schizophr Res.*, 214:51-59, 2019.
17. R. Salvador, E. Canales-Rodríguez, A. Guerrero-Pedraza, et al. Multimodal integration of brain images for MRI based diagnosis in schizophrenia. *Front Neurosci*, 13:1203, 2019.
18. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12649>