Article:
Li, S., Liu, K. and Yang, P. orcid.org/0000-0002-8553-7127 (Accepted: 2025) An interpretable deep learning approach for Alzheimer's disease diagnosis using gene expression data. IEEE/ACM Transactions on Computational Biology and Bioinformatics. ISSN 1545-5963 (In Press)

An Interpretable Deep Learning Approach for Alzheimer's Disease Diagnosis Using Gene Expression Data

***Abstract*—With the global ageing population, the diagnosis of Alzheimer's disease (AD) has become an urgent public health priority. Gene expression techniques offer the advantages of being less invasive and cost-effective, but their high dimensionality and small sample sizes make them prone to the curse of dimensionality in AD diagnosis. This study proposes a novel interpretable deep learning approach to address these challenges. We introduce a shallow sparse autoencoder for dimensionality reduction and combine it with XGBoost for classification, achieving an Area Under the Receiver Operating Characteristic curve (AUROC) of up to 95.13%. Additionally, we develop a fast, low-cost feature selection algorithm that dynamically adjusts feature elimination to enhance model efficiency. Comprehensive cross-dataset evaluation demonstrates the model's strong generalisation performance on the public datasets: Alzheimer's Disease Neuroimaging Initiative (ADNI), AddNeuroMed1 (ANM1), and ANM2. Our method also provides biological interpretability through enrichment analysis, offering insights into the mechanisms underlying AD and potential therapeutic targets. This makes our approach a promising tool for early, accurate diagnosis and clinical application.**

***Index Terms*—feature selection, dimensionality reduction, enrichment analysis, gene expression, alzheimer's disease, deep learning**

I. Introduction

ALZHEIMER'S disease (AD) is a progressive neurodegenerative disorder that affects millions of people around the world. It is characterised by gradual loss of memory and cognitive functions, leading to a decline in the ability to perform daily activities. From 2000 to 2019, the recorded deaths attributed to Alzheimer's disease increased by 145%, more than doubling in number [1]. The World Health Organization's report indicates that currently, there are more than 50 million individuals globally who have dementia, and this number is projected to nearly triple by the year 2050 [2]. In contrast, deaths from the leading cause of death, which is heart disease, decreased by 7.3% [1]. This indicates that as the population ages, Alzheimer's disease has become a more prevalent cause of death. The early stage of the disease presents a crucial opportunity to implement interventions aimed at modifying and preventing the progression of the disease, achieving maximum effectiveness. Despite significant advances in understanding the pathology of AD, early diagnosis remains a challenge. This is largely due to the complex nature of the disease, which involves multiple genetic and environmental factors.

As researchers delve deeper into Alzheimer's disease, while clinical core based on impairment of episodic memory are currently the main diagnostic criterion, other biomarkers are gradually being introduced into the AD diagnostic process as well. In a 2007 study, Dubois et al. recommended that in addition to the clinical core of early and significant episodic memory impairment, the NINCDS-ADRDA and DSM-IV-TR criteria, which are the most popular diagnostic criteria, should also take into account at least one or more biomarkers that have been shown to be effective in the diagnosis of AD, such as magnetic resonance imaging (MRI), positron emission tomography (PET) and cerebrospinal fluid (CSF) [3]. MRI, FDG-PET, amyloid PET, and CSF biomarkers can detect early brain and body changes related to Alzheimer's disease. MRI shows brain tissue shrinkage in the medial temporal lobe. FDG-PET shows brain cell glucose use. Amyloid PET shows amyloid plaque accumulation in the brain. CSF biomarkers show Aβ and tau protein levels and ratios in the cerebrospinal fluid. These proteins and plaques are signs of Alzheimer's disease. Due to the complex pathology of Alzheimer's disease, multi-omics data-based AD diagnostic studies [4]–[7] have become a hot topic in recent years. However, each of these diagnostic methods has some drawbacks, such as cognitive scales that rely on subjective diagnosis by clinicians, CSF being an invasive approach, and MRI and PET being expensive. The World Health Organisation (WHO) projects that the population of individuals aged 80 years or older will triple by 2050, reaching 426 million [8]. Among them, two-thirds are expected to reside in lower- and middle-income nations [8]. Therefore, a more affordable and less invasive method of objective diagnosis is needed to make early diagnosis of Alzheimer's disease widely available.

Gene expression testing represents a more affordable and cost-effective approach to the early detection of Alzheimer's disease (AD) compared to neuroimaging techniques. While gene expression testing typically costs under £100, neuroimaging tests such as magnetic resonance imaging (MRI) range from £500 to £1,500, and positron emission tomography (PET) tests can cost £1,400 to over £4,900. The lower cost of gene expression testing makes it a more accessible diagnostic option, particularly in resource-limited settings, and its ability to identify early biomarkers of AD has significant potential to promote early diagnosis. Early detection facilitates timely intervention, which is critical for improving patient outcomes and potentially slowing the progression of the disease. Recent research has highlighted the potential of gene expression data in improving the accuracy of AD diagnosis [9]–[12]. Gene expression is the process by which genes are transcribed to produce active proteins. Gene expression data, also known as transcriptomics data, refers to data reflecting mRNA abundance based on DNA microarray experiments. This information can be used to identify patterns and correlations that may be indicative of disease states. However, the high

dimensionality and complexity of gene expression data present significant challenges. Traditional statistical methods often lack the ability to capture complex patterns and interactions among genes. Therefore, if conventional statistical methods or machine learning algorithms are applied directly, they often encounter the 'curse of dimensionality', particularly in contexts where the number of features (e.g., genes) far exceeds the number of samples. While this issue is more pronounced in single-cell gene expression data due to the interplay of cells, genes, individuals, and time points, it remains relevant in bulk gene expression analysis for Alzheimer's disease (AD). Genome-wide expression profiling for AD often involves tens of thousands of genes, necessitating dimensionality reduction techniques to identify informative biomarkers, improve model performance, and enhance interpretability. It means that as the number of features increases, the performance and interpretative capability of the model may instead decrease. Furthermore, these methods often require a priori knowledge of the disease, which may not always be available.

In the context of processing high-dimensional data, both deep learning and traditional machine learning methods offer distinct advantages and limitations. Deep learning techniques, such as neural networks and their variants (e.g., convolutional neural networks and recurrent neural networks), excel in dimensionality reduction due to their ability to automatically learn complex representations and features from raw data [13], [14]. These methods leverage hierarchical architectures to capture intricate patterns in high-dimensional spaces, often resulting in superior performance for tasks such as image and speech recognition [15]. However, this performance comes at the cost of interpretability, as deep learning models are often described as "black boxes" due to their complex and opaque internal mechanisms [16], [17]. In contrast, traditional machine learning methods, such as linear regression, support vector machines (SVMs), and decision trees, offer greater interpretability and transparency [18], [19]. These models allow for a clearer understanding of how input features influence predictions, which is crucial in domains requiring explanations for decision-making, such as healthcare and finance [20]. Nonetheless, traditional methods may struggle with high-dimensional data due to their limited capacity to capture complex relationships, often leading to lower predictive performance compared to their deep learning counterparts [21], [22] In recent years, for the high-dimensional small sample size problem in the field of AD diagnosis, there have been research attempts to solve it by combining the high performance of deep learning with the high interpretability of traditional machine learning or statistical learning [23]. The choice between deep learning and traditional machine learning approaches involves a trade-off between performance and interpretability, with each method offering unique benefits suited to different types of problems and data characteristics.

The main contributions and novelties of this work are summarised as follows:

1) **Proposed an interpretable deep learning framework for Alzheimer's disease diagnosis:** A shallow sparse autoencoder was developed to extract biologically interpretable high-level features from gene expression data. This approach demonstrated superior diagnostic performance compared to traditional deep learning models, achieving an Area Under the Receiver Operating Characteristic curve (AUROC) of up to 95.13%. The interpretability of the extracted features enhances the model's potential utility in clinical and research contexts.
2) **Designed a novel, computationally efficient feature selection algorithm for gene expression data:** A fast and low-cost feature selection algorithm was introduced, capable of dynamically adjusting the number of eliminated features to efficiently identify high-weight features. This method reduces computational overhead while maintaining diagnostic accuracy, making it particularly suitable for processing large-scale datasets.
3) **Conducted extensive cross-dataset generalisation analysis:** The proposed framework was rigorously evaluated for cross-dataset generalisability. Trained on the ADNI gene expression dataset, the model achieved strong classification performance on external datasets, including ANM1 and ANM2, demonstrating its adaptability and broad applicability to different gene expression datasets in Alzheimer's research.

The rest of the paper is organised as follows: the 'METHODS' section describes the details of the dataset and an overview of the methodology, the 'EXPERIMENTAL RESULTS AND ANALYSIS' section presents and discusses the results of the experiments, and the 'CONCLUSION' section summarises the outcomes and points out the future work.

II. Related Work

The use of gene expression data for disease diagnosis has been extensively studied in recent years [24], [25]. Various feature selection methods have been proposed to identify relevant genes from high-dimensional gene expression data. Booij et al. conducted a study to develop a blood-based gene expression test for the early detection of Alzheimer's Disease [26]. They utilized oligonucleotide microarray analysis on blood samples from 94 AD patients and 94 healthy controls, employing a Jackknife gene selection method and Partial Least Square Regression (PLSR) to create a disease classifier algorithm. This algorithm, based on 1239 probes, achieved an accuracy of 87%, sensitivity of 84%, and specificity of 91%. Lunnon et al. (2013) proposed a blood gene expression marker for early diagnosis of Alzheimer's Disease (AD) using data from HT-12v3 BeadChips [27]. They developed an AD diagnostic classifier in a training cohort of 78 AD and 78 control blood samples, achieving 75% accuracy in a validation group. The classifier was compared with structural MRI measures, showing 70% accuracy for gene expression versus 85% for MRI. The study highlighted the potential of blood expression markers to detect AD earlier in the prodromal phase. Li et al. conducted a comprehensive analysis to identify differentially expressed genes (DEGs) in blood and brain tissues of Alzheimer's Disease (AD) patients [28]. They utilized microarray gene expression profiles from large datasets, applying the *limma* R package for DEG identification. For feature selection, they employed the Least Absolute Shrinkage

and Selection Operator (LASSO) regression, combined with Support Vector Machine (SVM), Random Forest (RF), and logistic Ridge Regression (RR) models. The study revealed significant overlaps in DEGs between blood and brain tissues. However, they may encounter challenges in high-dimensional and complex gene expression datasets, as evidenced by the results of this study. The average AUC values for the ADNI, ANM1, and ANM2 datasets were 0.657, 0.874, and 0.804, respectively, with further performance degradation observed during external validation across datasets. These findings suggest potential overfitting due to the curse of dimensionality, which can limit generalisability.

Deep learning techniques have shown promise in handling high-dimensional data and capturing complex patterns. Ahmed et al. explore various deep learning algorithms for the classification of gene expression data, which is crucial in bioinformatics, particularly for cancer classification [29]. The study evaluates the performance of Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), and an improved DNN with a preprocessing technique to handle overfitting. The improved DNN incorporates Dropout to enhance accuracy. The authors also discuss several feature selection methods, including Sequential Random k-Nearest Neighbours (SRKNN), Single Sequential Feature Selection (SSFS), Incremental Wrapper-based feature subset selection with Markov Blanket (IWSSMB), and a hybrid genetic algorithm and learning automata (GALA). While these methods show promising results, they often suffer from high computational complexity and sensitivity to noisy data, which can impact the robustness and generalizability of the models. Xie et al. developed a regression-based predictive model using a MultiLayer Perceptron and Stacked Denoising Auto-encoder (MLP-SAE) to assess the impact of genetic variants on gene expression [30]. The model was trained with a stacked denoising auto-encoder for feature selection and a multilayer perceptron framework for backpropagation, incorporating dropout to prevent overfitting. The results demonstrated that the MLP-SAE model with dropout outperformed other models such as Lasso and Random Forests. However, the study noted that the high-dimensional nature of genomic data and the low signal-to-noise ratio posed significant challenges, potentially limiting the model's ability to identify trans associations and necessitating further improvements. Dincer et al. (2020) introduced the Adversarial Deconfounding AutoEncoder (AD-AE) to address the challenge of disentangling confounders from true biological signals in gene expression data [31]. The AD-AE model comprises two neural networks: an autoencoder to generate embeddings that reconstruct original measurements and an adversary trained to predict confounders from these embeddings. By jointly training these networks, the model aims to produce embeddings that encode significant biological information while excluding confounding signals. However, the method has limitations, including potential overfitting due to the unregularized autoencoder and the complexity of training adversarial networks, which may require substantial computational resources and careful tuning of hyperparameters. However, while these methods offer good performance in dimensionality reduction, they often lack interpretability, which is crucial in medical applications for understanding disease mechanisms and making informed clinical decisions.

Enrichment analysis using KEGG and GO has been widely used to interpret the biological relevance of selected features [32]. These analyses provide insights into the biological processes, molecular functions, and cellular components associated with the selected genes, thereby validating their relevance to the disease under study.

In this paper, we build upon these previous works by proposing a novel approach that integrates deep learning techniques with traditional statistical methods for feature selection from gene expression data. Our approach aims to leverage the strengths of both techniques to improve AD diagnosis accuracy while maintaining interpretability.

III. Methods

The proposed method is divided into four steps: preprocessing the data, reducing the dimensionality of the sparse autoencoder, XGBoost classification and interpretability analysis. In the data pre-processing, we constructed a dataset based on ADNI by selecting probes common to the ANM1 and ANM2 datasets. The vectorized probe data was then fed into the sparse autoencoder as input for feature selection. At the same time, dimensionality reduction is achieved by limiting the number of nodes in the hidden layer of the sparse autoencoder. Hence, the nodes in the hidden layer of the sparse autoencoder were used as selected features and input to the XGboost classifier, which was trained and performs the classification task. Finally, we performed an interpretability analysis. We first ranked the features in terms of importance using XGBoost, and then filtered out the high weight nodes and high weight probes, which were used for enrichment analysis to verify the interpretability of the extracted probes. The general framework is shown in Fig. 1.

A. Data Pre-Processing

The experiments in this study used peripheral blood gene expression data. We introduced the gene expression dataset from ADNI to train and validate our feature selection and classification model. In addition to the data in ADNI, we also used gene expression data from AddNeuroMed1 (ANM1) and AddNeuroMed2 (ANM2) to validate the generalisability of our model across databases. To classify participants, the samples from the three databases were classified using Mini Mental State Examination (MMSE) as diagnostic criteria. MMSE is a joint effort of the National Institute of Neurological and Communicative Disorders and Stroke (NINDS) and the Alzheimer's Disease and Related Disorders Research Association (ADRDA). MMSE is a measure of general cognitive status that includes 30 areas of ability, including memory, orientation, comprehension, attention, reading, writing, learning, etc. In this study we included 744 samples from ADNI (containing 246 NC, 382 MCI and 116 AD), 329 samples from ANM1 (containing 104 NC, 80 MCI and 145 AD) and 382 samples from ANM2 (containing 134 NC, 109 MCI and 139 AD). Details of the dataset are given in Table I.

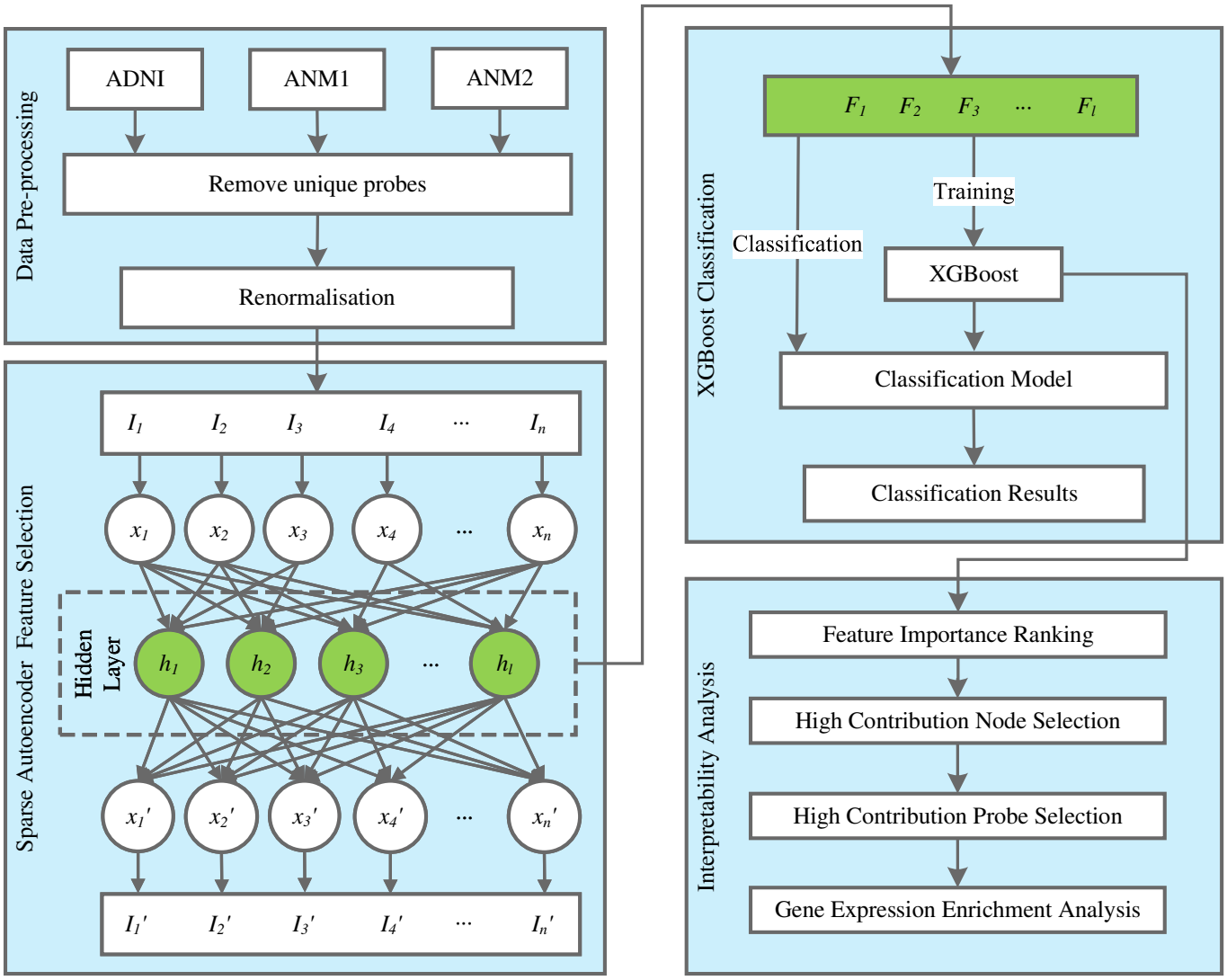


Fig. 1: The Architecture of the proposed Sparse Autoencoder-XGBoost model for AD Classification. Step 1: Select the probes common to the three datasets to build our dataset; Step 2: Use the sparse autoencoder to reduce the dimensionality and extract the features of the data; Step 3: Use the hidden layer nodes of the sparse autoencoder as features, and use XGBoost to classify these features. Step 4: Select high contribution probes using feature importance ranking and feature selection methods, and perform enrichment analysis on the probes to validate the interpretability of the proposed method.

TABLE I: Details of participants from ADNI, ANM1 and ANM2

	ADNI			ANM1			ANM2		
	NC	AD	MCI	NC	AD	MCI	NC	AD	MCI
Number of Cases	246	116	382	104	145	80	134	139	109
Gender, % Males	47.6%	63.8%	56.5%	37.1%	30.0%	47.3%	31.5%	36.4%	32.8%
Age	76.2±6.5	77.3±7.7	72.9±8.0	73.7±7.5	76.0±6.6	74.9±5.3	75.7±6.1	78.6±5.4	77.2±3.2
MMSE	29.1±1.2	21.3±4.4	28.1±1.7	29.1±1.1	20.9±4.7	26.8±1.7	28.4±1.7	19.9±4.6	28.1±1.1

Regarding the collection chip of gene expression data, ADNI uses Affymetrix Human Genome U219, while ANM1 and ANM2 use Illumina HumanHT-12 Expression BeadChip v3 and v4, respectively. As the genes and probes targeted by the Affymetrix Human Genome U219 and the Illumina HumanHT-12 Expression BeadChip are not identical, we performed data pre-processing on the three datasets to enable controlled experiments between the datasets. The first step was to remove probes that were not common to the three datasets: we removed genes that were unique to each of the three datasets, leaving 14,498 genes, which reduced the number of probes from 49,386 to 38,947 for ADNI, from 48,804 to 29,485 for ANM1, and from 47,231 to 20,177 for ANM2. We selected the probes common to all three datasets, so the final number of probes selected was 16,482. The second step is to normalise the three datasets. Although

all three datasets provided normalised probe data, the median RNA expression values for ADNI, ANM1 and ANM2 were calculated to be 3.897, 7.584 and 6.154 respectively. This indicated that the gene expression intensity of ADNI dataset is significantly lower than that of ANM1 and ANM2. To ensure the quality of the cross-dataset experiments by mitigate the influence of batch effect and difference between the datasets, we renormalised the three datasets by performing Robust Multi-chip Average (RMA) [33].

B. Shallow Sparse Autoencoder for Dimensionality Reduction

The autoencoder (AE) is an unsupervised learning model based on artificial neural networks [34], designed to extract hidden features from input data and efficiently reconstruct the original data. One of its key advantages is the ability to learn both linear and nonlinear features, making it particularly well-suited for complex data. By employing deep neural network structures and nonlinear activation functions, autoencoders excel at capturing intricate patterns and relationships in data.

Gene expression data is characterised by a complex non-linear relationship among genes, influenced by transcription factors, pathway memberships, and other biological properties. Traditional linear dimensionality reduction methods, such as principal component analysis (PCA), rely on linear mappings and often fail to capture these nonlinear associations. In contrast, autoencoders are capable of learning nonlinear manifolds, enabling them to better model the intricate relationships inherent in gene expression data.

An autoencoder typically consists of two main stages: encoding and decoding. Each stage uses specific activation functions tailored to the task. The encoding stage is represented as follows:

$$h(x) = s(W_1 x + b_1), \tag{1}$$

where x is the input vector, $h(x)$ represents the activations in the hidden layer, W_1 is the weight matrix, b_1 is the bias vector, and s denotes the sigmoid activation function.

The decoding stage is expressed as:

$$x' = s(W_2 x + b_2), \tag{2}$$

where x' is the reconstructed output vector, W_2 represents the weight matrix connecting the hidden and output layers, and b_2 is the bias vector for the output layer.

The training of an autoencoder requires a loss function to evaluate its performance. The loss function typically includes a reconstruction error term, which measures the mean squared error (MSE) between the reconstructed and original inputs, and an L_2 regularisation term to mitigate overfitting:

$$J(w,b) = \frac{1}{N}\sum_{n=1}^{N}(x_n - \widehat{x}_n)^2 + \frac{\lambda}{2}\sum_{l=1}^{L}\sum_{j=1}^{N}\sum_{i=1}^{k}\left(w_{ji}^{(l)}\right)^2, \tag{3}$$

where L is the number of hidden layers, N is the number of samples, k is the number of variables in the dataset, and $w_{ji}^{(l)}$ represents the weights in the hidden layers.

Sparse autoencoders extend standard autoencoders by incorporating a sparsity constraint on the hidden layer neurons [35]. This constraint ensures that only a limited number of neurons are activated, enhancing feature extraction and noise immunity. Given input x, the average activation of hidden neuron j is calculated as:

$$\widehat{\rho}_i = \frac{1}{n}\sum_{j=1}^{n} h_i(x_j), \tag{4}$$

where n is the number of training samples, and x_j represents the j-th sample. The sparsity penalty is defined using the Kullback-Leibler (KL) divergence:

$$\sum_{j=1}^{D^{(1)}} KL(\rho \parallel \hat{\rho}_j) = \sum_{j=1}^{D^{(1)}}\left(\rho \log\frac{\rho}{\hat{\rho}_j} + (1-\rho)\log\frac{1-\rho}{1-\hat{\rho}_j}\right), \tag{5}$$

where $D^{(1)}$ denotes the number of neurons in the hidden layer, ρ is the desired sparsity level, and $\hat{\rho}_j$ is the actual average activation. When ρ and $\hat{\rho}_j$ are similar, the penalty approaches zero. Incorporating this term into the loss function results in the sparse autoencoder loss function:

$$J_{SAE}(w,b) = J(w,b) + \beta\sum_{i=1}^{D^{(1)}} KL(\rho \parallel \hat{\rho}_i), \tag{6}$$

where β controls the weight of the sparsity penalty.

While deep autoencoders with multiple hidden layers can extract higher-level features, only the first hidden layer directly relates to the original features, limiting their interpretability. Therefore, this study employs a shallow sparse autoencoder with a single hidden layer to ensure interpretability while achieving dimensionality reduction.

C. XGBoost for Multi-Class Classification

XGBoost is a machine learning system based on boosted trees proposed by Chen et al. [36] based on the work of gradient boosting algorithm (GBDT). The algorithm consists of a collection of iterative residual trees, i.e., the Nth decision tree learns the residuals of the previous N-1 trees, and the predicted outputs of each tree are summed up to be the final output of the sample. At the same time, the splitting strategy adopted by XGBoost in constructing residual trees can be used to evaluate the importance of features by metrics. It has been proved that XGBoost achieves excellent results compared with other classifiers in classifying small samples and unbalanced data.

Assume that in the sample dataset $D = (x_i, y_i)$, x_i is the feature data of the *i*-th sample, and y_i is the label output value. XGBoost consists of K CART trees, which assign scores to each leaf node, and finally the predicted scores of each CART are summed up to obtain the final total score, which is evaluated by K additive functions as shown in the following:

$$\widehat{y}_\iota = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in F \tag{7}$$

where f_k denotes the independent tree structure with leaf node weights, and $f_k(x_i)$ denotes the weight value of the i-th sample x_i that falls on a leaf node in the *k*-th tree. F is the overall

space of the K trees. To optimise the function objective $Obj(\theta)$ is:

$$Obj(\theta) = \sum_{i}^{n} l(y_i, \widehat{y}_\iota) + \sum_{k}^{K} \Omega(f_k) \tag{8}$$

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{9}$$

where $\iota(\cdot)$ is the differential loss function,which measures the error between the true value and the predicted value of the model. $\Omega(\cdot)$ is the regularisation term, which represents the complexity of each CART tree, T represents the number of leaf nodes in each CART tree, and w_j represents the fraction of each leaf node, and in this way it is used to constrain the objective function to prevent overfitting. γ and λ are the constants that control the degree of regularisation of the constants. Since the algorithm uses additive training to generate trees one by one, then for round t the predicted value $\widehat{y}_\iota$ and the loss function $Obj(\theta)$ can be expressed as:

$$\widehat{y}_\iota^{(t)} = \Sigma_{k=1}^{t} f_k(x_\iota) = \widehat{y}_\iota^{(t-1)} + f_\iota(x_\iota) \tag{10}$$

$$Obj(\theta)^{(t)} = \Sigma_i^n \, l(y_i, \widehat{y}_l^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{11}$$

Gradient boosting decision tree using first order derivative information in optimisation. While XGBoost transforms $\iota(\cdot)$ using second order Taylor's formula for faster convergence of the objective function. The second order Taylor expansion is given as:

$$f(x + \Delta x) \cong f(x) + f'(x)\Delta x + \tfrac{1}{2} f''(x)\Delta x^2 \tag{12}$$

Let $f_t(x_i)$ be Δx in Taylor's formula, and a Taylor second order expansion of the loss function $Obj(\theta)^{(t)}$ has:

$$\begin{aligned} Obj(\theta)^{(t)} \cong \Sigma_i^n (l(y_i, \widehat{y_i}^{(t-1)}) + g_i f_t(x_i) \\ + h_i f_t^2(x_i)) + \Omega(f_t) \end{aligned} \tag{13}$$

where $g_i$and h_i denote the first and second order derivatives of $l(y_i, \widehat{y_i}^{(t-1)})$ with respect to $\widehat{y_i}^{(t-1)}$. The function $l(y_i, \widehat{y_i}^{(t-1)})$ in round t can be treated as a constant term. Hence, substituting Eq. 9 into Eq. 13, the following equation is obtained:

$$\begin{aligned} Obj(\theta)^{(t)} = \Sigma_i^n \left(g_i f_t(x_i) + h_i f_t^2(x_i)\right) \\ + \ \gamma T + \frac{1}{2}\lambda \Sigma_{j=1}^{T} w_j^2 \end{aligned} \tag{14}$$

Let $w \in R^T$, w be the sequence of weights of the leaf nodes, $q : R^d \to \{1, 2, ...T\}$, q be the tree structure. Therefore $q(x)$ is denoted as the position of the sample x falling in the leaf node. $f_t(x_i)$ can be expressed by the following equation:

$$f_t(x_i) = w_{q(x)}, w \in R^T, q : R^d \to \{1, 2, ...T\} \tag{15}$$

Convert the loss function for traversing the sample data to a loss function for traversing the leaf nodes:

$$Obj(\theta)^{(t)} = \Sigma_{j=1}^{T} \left(G_j w_j + \tfrac{1}{2}(H_j + \lambda) w_j^2\right) + \gamma T \tag{16}$$

where I_j is the set of samples belonging to the leaf node j,and subsequently its derivative on $Obj(\theta)^{(t)}$ yields the extreme points with extreme values of:

$$w_j^* = -\frac{G_j}{H_j + \lambda} \tag{17}$$

$$Obj^* = -\frac{1}{2}\Sigma_{j=1}^{T} \frac{{c_j}^2}{H_j + \lambda} + \gamma T \tag{18}$$

XGboost uses equation 19 to evaluate whether or not a node splits, and ultimately determine the structure of the tree:

$$Gain = \frac{1}{2}\left[\frac{{G_L}^2}{H_L + \lambda} + \frac{{G_R}^2}{H_R + \lambda} - \frac{(G_R + G_L)^2}{H_R + H_L + \lambda}\right] - \gamma \tag{19}$$

The number of times a feature acts as a split node throughout the construction of the model is $weight$ and the average gain of the feature as a split node is $gain$:

$$gain = \sum_{Gain} /weight \tag{20}$$

D. Fast Recursive Feature Elimination for Feature Selection

Medical tasks place more emphasis on interpretability than traditional machine learning tasks. This requires not only generating classification results, but also interpreting these results in a biologically meaningful way. Since the sparse autoencoder has only the first layer of nodes directly connected to the probes, and the deeper nodes represent a high-dimensional mapping of the upper features, which poses a challenge in recognising the relationship between the representations and the probes, the dimensionality of the features was not reduced in this study by using a stacked autoencoder. Nonetheless, the dimensionality of the features after dimensionality reduction using a sparse autoencoder with a single hidden layer, while allowing the classifier to achieve optimal classification, is still too high for interpretable biological analysis. Therefore, to achieve interpretable analyses, we need to perform further feature selection on these features.

One advantage of using the gradient boosting algorithm is that after the boosting tree has been created, the importance score of each feature can be obtained relatively for effective feature selection. The importance score, in general, measures the value of features in the model for boosting decision tree construction. The more a feature is used for node segmentation in the model, the higher its relative importance. Feature importance is obtained by calculating and ranking each feature in the sample dataset. The importance of a feature is calculated in the decision tree by the amount of each feature's split-point improvement performance measure (typically the Gini index). The larger a feature's performance measure for split-point improvement, i.e., the closer it is to the root node, the larger its importance weight is. Also, the more features are selected by more boosting trees the higher the importance degree. Finally, the results of a feature in all the boosting trees are weighted and summed and then averaged to get the importance score.

Inspired by the Recursive Feature Selection (RFE) feature selection algorithm proposed by Guyon et al [37], we proposed an improved version of RFE optimised for high-dimensional data called Fast-RFE. Recursive feature selection is a feature selection algorithm based on the importance of model features, eliminating a number of end features at each iteration based on the feature importance ranking, and using the dataset containing the retained features as the training samples for

the next round until the features are reduced to a certain dimension. RFE is effective for small-sample classification tasks, but its computational complexity and costs are high when the feature dimensions are high.

We proposed Fast Recursive Feature Selection to accommodate the high-dimensional nature of the gene expression data classification task. The XGBoost model can be used to rank feature importance by the feature importance metric, and in this work, the normalised weight score of $gain$ in Eq. 20 is used as the metric for feature importance ranking. The Fast-RFE algorithm, as presented in Algorithm 1, employs a two-phase approach to identify significant features efficiently. Initially, features are sorted based on their gain values, and the mean (μ) and standard deviation (σ) of these gains are calculated. The algorithm begins with an initial threshold $\delta = \mu$ and iteratively refines it. In the first phase, features with gain exceeding δ are selected to construct an XGBoost classifier. If the resulting model's accuracy decrease is less than 0.01 compared to the original model, δ is incremented by σ. This process continues until a significant accuracy drop is observed, establishing an interval $[a, b]$ for further refinement. The second phase employs a binary search within $[a, b]$ to optimise the feature selection. At each iteration, features with $|w| > \delta$, where $\delta = (a+b)/2$, are chosen to train an XGBoost classifier. Based on the model's performance, either a or b is updated to δ. This binary search persists until the interval $[a, b]$ contains only one feature, at which point features with $|w| > a$ are selected as the optimal feature subset.

Algorithm 1 Fast Recursive Feature Selection

```
Sort features by gain values
Calculate μ and σ of all features' gain values
Set initial threshold δ = μ
a ← 0, b ← ∞
while true do
    Select features with gain > δ
    Build XGBoost classifier with selected features
    Calculate classification accuracy
    if accuracy drop < 0.01 compared to original model then
        a ← δ
        δ ← δ + σ
    else
        b ← δ
        break
    end if
end while
while a < b do
    δ ← (a + b)/2
    Select nodes with |w| > δ
    Train XGBoost classifier with selected features
    Calculate classification accuracy
    if accuracy drop < 0.01 compared to original model then
        a ← δ
    else
        b ← δ
    end if
    if only 1 node satisfies a < |w| < b then
        Select features with |w| > a to construct the optimal feature subset
        break
    end if
end while
```

IV. Experimental Results and analysis

To validate the effectiveness of our proposed approach in classifying Alzheimer's disease and to verify the biological significance of the extracted features. First, we used feature extraction algorithms and classifiers that have previously been shown to be effective for AD classification in the literature as a control group to demonstrate the effectiveness of our proposed method. The detailed flow of the control experiment is shown in the Fig. 2. We introduced feature extraction algorithms and classifiers that have previously been shown to be effective in the literature as a control group to demonstrate the effectiveness of our proposed method. Three feature extraction algorithms are included: Principal Component Analysis (PCA) Least Absolute Shrinkage and Selection Operator (LASSO) regression and Differential Gene Expression Analysis (DGE). Four classifiers are included: Support Vector Machine (SVM) Random Forest (RF), L1 regularisation Logistic Regression (L1-LR) and Deep Neural Network (DNN). Then, using the feature selection method proposed in Section III-D, we selected the high contributing nodes from the hidden layer of the sparse autoencoder. Further, we filtered out the high contributing gene expression probes from the high contributing nodes. Last but not least, Gene Ontology (GO) biological enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis uncovered the biological significance of the high contributing probes.

A. Evaluation Matrixes for Classification

When evaluating the classification performance of a classifier, common metrics that can be used include: Accuracy, Precision, Sensitivity, Specificity and Receiver Operating Characteristic (ROC). Accuracy can be expressed as the ratio of the number of samples correctly classified by the classifier to the total number of samples. Precision can be expressed as the ratio of the number of positive samples correctly classified by the classifier to the number of all positive samples predicted by the classifier. Sensitivity, also known as the True Positive Rate (TPR), is the ratio of the number of predicted results for positive samples to the actual number of positive samples. Specificity, also known as the False Positive Rate (FPR), is the ratio of the number of results that were incorrectly predicted as positive samples but were actually negative samples to the actual number of negative samples.The ROC (Receiver Operating Feature) curve is used to evaluate the performance of the classifier, the horizontal axis indicates the proportion of negative samples that are incorrectly classified as positive samples, and the vertical axis indicates the proportion of positive

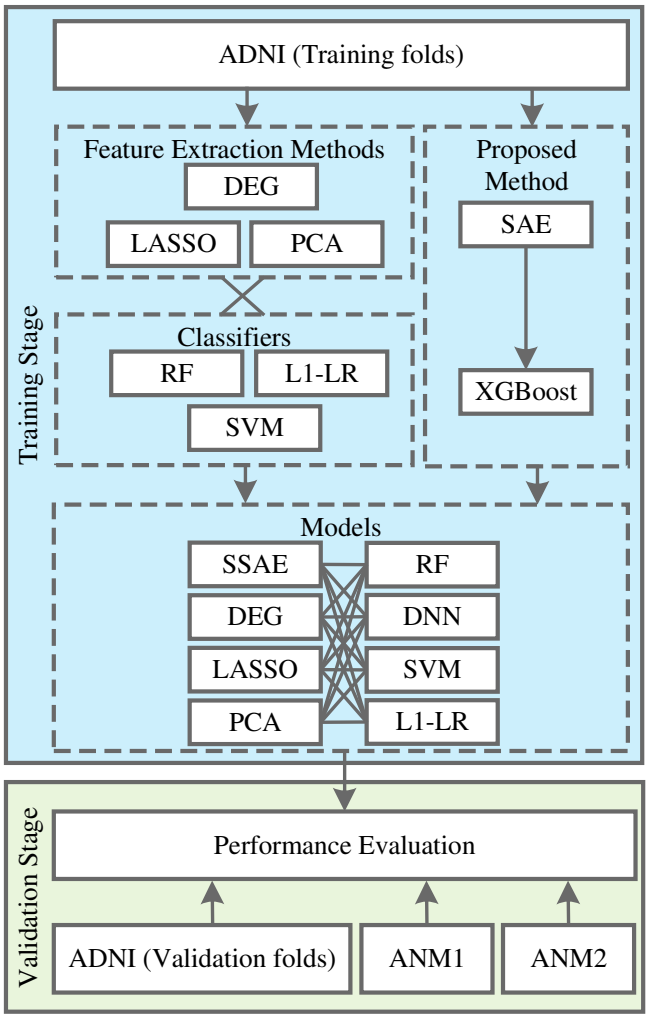


Fig. 2: The pipeline of the control experiment. We utilise three of the most commonly employed feature selection methods in the field of gene expression research, namely DEG, LASSO, and PCA, in conjunction with four commonly used classifiers, namely RF, SVM, and L1-LR. The various feature selection algorithms were combined with the classifiers one by one in order to train the model, and their performance was compared with that of the proposed algorithms using SSAE as the feature selection method and XGBoost as the classifier. The performance on three datasets of each model was then compared with that of the proposed algorithm using SSAE as the feature selection method and XGBoost as the classifier. DEG, Differently expressed gene; LASSO, Least absolute shrinkage and selection operator; PCA, principal component analysis; RF, random forest; SVM, support vector machine; L1-LR, L1-regularised LR; SSAE, Sparsed Autoencoder; XG, XGboost.

samples that are correctly classified as positive samples. The AUC value is the area under the ROC curve, which ranges from 0 to 1. The closer the AUC is to 1, the better the performance of the classifier. The traditional ROC curve is for a binary classification approach, whereas this experiment needs to evaluate the multiclass-classification tasks for AD, MCI, and NC. For ROC to be used in this study, we plotted ROC curves for each of the three classification tasks for each experiment and then averaged the three to obtain the ROC curve for the given task.

TABLE II: Parameters of the Sparse Autoencoder

Parameter	Value
Active Function	Sigmoid
Batch Size	256
Decay	0.99
Sparsity Parameter	0.001
Penalty factor	1.00

TABLE III: Parameters of XGBoost

Parameter	Value
Booster	gbree
eta	0.003
max depth	6
gamma	0.3
objective	multi:softprob
subsample	1
num class	3

B. Optimal Hyperparameter Selection for the Models

The experimental environment of this study is: Intel i9-13900k, Nvidia RTX 3090, 64gb Ram, windows 11. In order to obtain appropriate hyperparameters for SSAE, we use 10-fold cross-validation to train each combination of hyperparameters of the model, while avoiding over-training that leads to over-fitting. Finally, the hyperparameter combination that minimises the average reconstruction error is selected. For the XGBoost classifier, we randomly divided the ADNI dataset into training and testing sets with a ratio of 7:3. Then we optimise the model hyperparameters by learning curve and grid search. The model hyperparameters are also adjusted to avoid overfitting by 10-fold cross-validation on the training set. The parameter settings for the Sparse Autoencoder and XGBoost are shown in Table II and Table III respectively.

C. Comparison of Different Models

We use the preprocessed ADNI dataset to train SSAE and XGBoost for classification tasks on AD, MCI and NC. The experiments are compared with three dimensionality reduction algorithms and four classifiers commonly used in the gene expression field. Fig. 3 shows the comparison of Average Area Under the Receiver Operating Characteristic (AUROC) curve. As shown in the Fig. 3, the proposed method outperforms other comparative methods.

D. Analysis of Bioinformatics Interpretability

In order to obtain biologically significant gene expression probes, we firstly used the proposed Fast-RFE algorithm to extract high weight nodes from the high weight features obtained by SSAE, and then we used the algorithm to extract high weight probes from the high weight nodes.

As shown in Fig. 4, the gain value of the nodes is heavily clustered around 0, and the contribution of these nodes to the

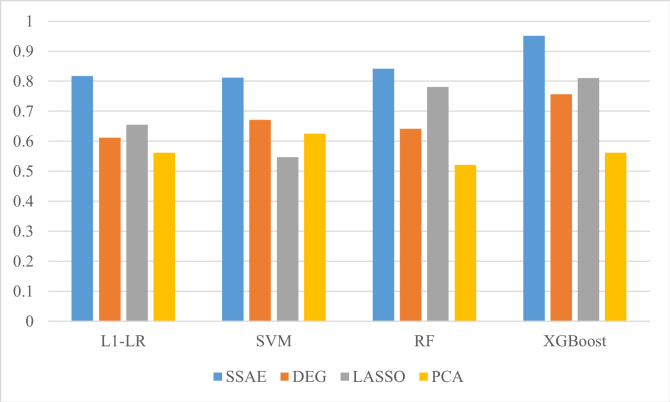


Fig. 3: Average AUROC of all comparison and the proposed model.

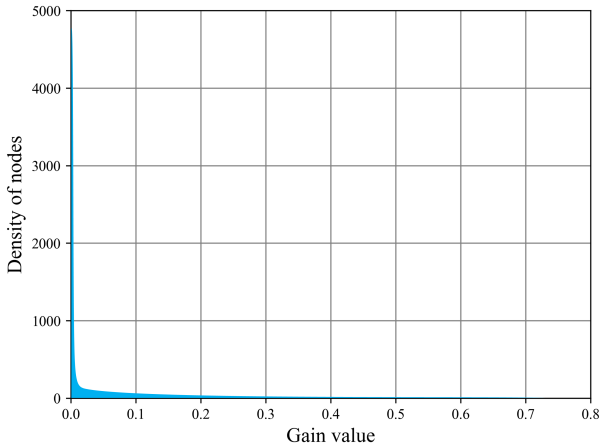


Fig. 4: Histogram of the normalised gain score of the nodes (features)

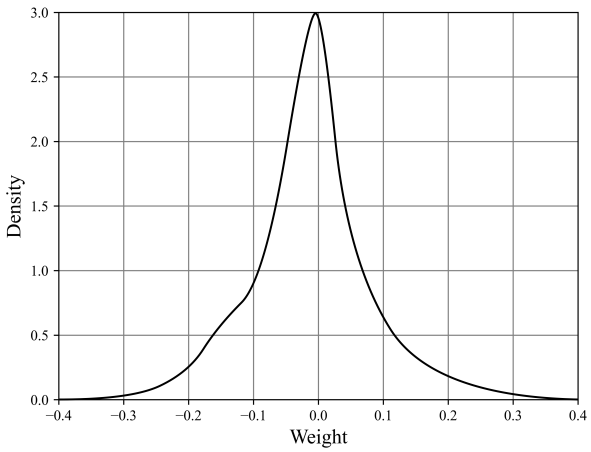


Fig. 5: Distribution of probe weight density curve inside node 2

classification is negligible. Further, as shown in Fig. 5, we plot the weight density curve of the nodes, which shows that the

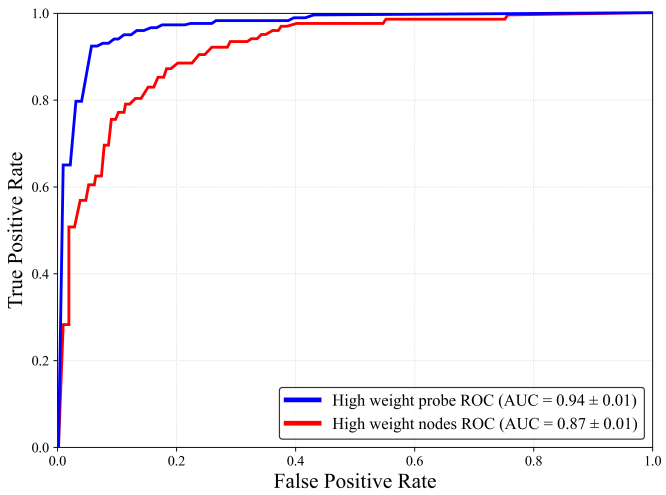


Fig. 6: ROC curves of high contributing nodes and high contributing probes

TABLE IV: Evaluation of the model's performance with different features (%) for classifying between CN, AD, and MCI

Features	Accuracy	Specificity	Sensitivity	Precision
Nodes	**93.23**	**94.51**	**93.12**	**90.51**
Probes	85.18	86.10	85.12	81.48

weights of all probes on a single node approximately follow a normal distribution with mean 0. In other words, in a node, the weights of all probes on a single node are distributed in the same way as the weights of all probes on a single node. That is to say, in a node, there are always a large number of probes whose weights are enriched around 0, and the influence of these probes on the value of the node is very small, while the probes distributed at the two ends of the weight density curve affect the value of the node to a great extent, and these probes with larger weights are called high weight probes.

We finally selected 37 high contributing nodes from 5000 nodes, and then filtered 4790 high weight probes from these 37 high weight nodes. We constructed XGBoost classifiers using high weighted nodes and high weighted probes respectively. Table IV shows the performance metrics of the two classifiers after cross-validation, and Fig. 6 plots their ROC curves. Combining the classification performance of Table IV and Fig. 6, the effect of the classifier constructed by the high weight probe is only slightly decreased compared with that of the high weight node, which indicates that the feature nodes have largely retained the information of the high weight nodes, and further proves that the selection of the high weight probe is effective. At the same time, it can be seen that the classifier constructed by the feature nodes is stronger than the high weight probe in every aspect, which may be due to a series of nonlinear transformations performed by the feature nodes on the high weight probe to improve its feature expressiveness, and thus it is more suitable for the AD classification problem. Compared with the original 5000 nodes classification, the accuracy decreases less than 2%, which shows that the feature nodes can represent the original nodes well. In conclusion,

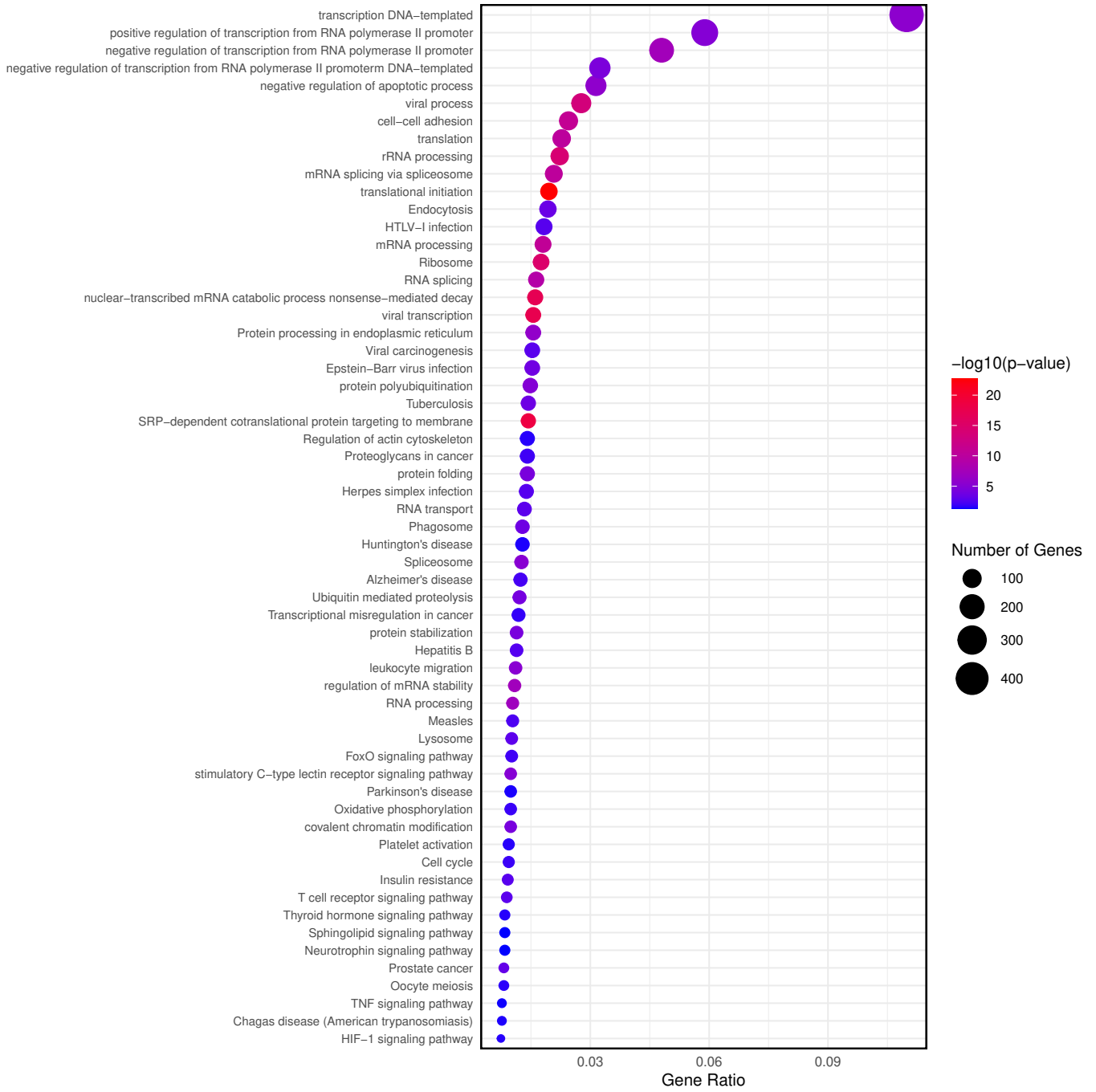


Fig. 7: Enriched Pathways dot plot. This dot plot illustrates the results of the gene enrichment pathway analysis, with each dot representing an enriched pathway. The horizontal axis displays the gene ratio, calculated as the number of genes in a pathway divided by the total number of genes analysed, where a higher ratio indicates a greater proportion of genes from the dataset present in that pathway. The size of each dot corresponds to the number of genes from the dataset found in the pathway, with larger dots signifying pathways containing more genes. The colour of the dots represents the statistical significance of the enrichment, shown as -log10(p-value), transitioning from blue (less significant) to red (more significant).

from the ADNI dataset alone, the feature nodes we constructed can significantly enhance the AD classification effect.

We then performed GO bioprocess enrichment and KEGG pathway enrichment analyses on the high weight probes and plotted the Enriched Pathways dot plot, in order to find out which biochemical pathways with relevant biological functions are more likely to be distributed in the two types of data. We used $p \leq 0.05$ as the significance threshold to screen for significant GO entries or KEGG pathways. As shown in Fig. 7, according to the GO enrichment results, the high weight probes were not directly enriched in biological processes related to Alzheimer's disease, but as shown by the KEGG

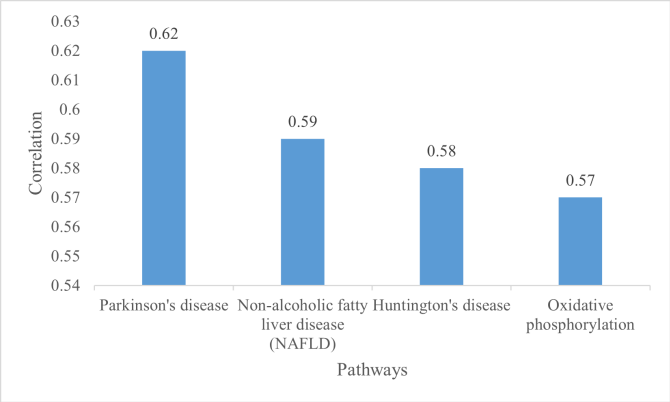


Fig. 8: AD-associated pathways in high weight probes

enrichment results, the high weight probes were significantly enriched in three metabolic pathways, namely Alzheimer's disease, Parkinson's disease and Huntington's disease. The results demonstrate that high weight probes exhibit superior biological performance in Alzheimer's disease.

To further validate the biological interpretability of the selected high weighted probes, we used the KEGG Cluster tool provided by Database for Annotation, Visualisation and Integrated Discovery (DAVID) to cluster enriched pathways with similar genes in AD-related pathways. Fig. 8 show that there is a strong correlation between Alzheimer's disease, Parkinson's disease and Huntington's disease, which are neurological disorders and may share similar metabolic processes. In addition, we found that Non-alcoholic fatty liver disease (NAFLD) and oxidative phosphorylation pathways are also strongly associated with Alzheimer's disease, which is consistent with some of the current findings. Liver disease has been reported to be a risk factor for cognitive decline in the elderly [38], and that NAFLD accelerated the emergence of AD pathology in a rat model of AD [39]. From these results, it is clear that high weighted probes have better biological interpretability in Alzheimer's disease.

E. Evaluation of Model Generalisability Performance

In order to validate the generalisability of the proposed method and to test whether it can be generalised to other Alzheimer's disease gene expression datasets, we carried out this work using the gene expression datasets of ANM1 and ANM2. We first preprocessed all the datasets using the method in Chapter III, and then used SSAE and XGBoost to extract and classify features from the data of ANM1 and ANM2, and plotted the ROC curves respectively. From the results, it can be seen that the classification performance of the proposed method on ANM1 and ANM2 datasets is only slightly degraded compared with ADNI, and the feature nodes have proved to be effective for the AD classification problem, considering that the high weight probes on ANM1 and ANM2 datasets are partially missing.

V. Conclusion

In conclusion, this study presents a novel and interpretable deep learning approach for the diagnosis of Alzheimer's disease using gene expression data. By employing a shallow sparse autoencoder, our model achieves high diagnostic accuracy, with an AUROC of up to 95.13%, while extracting biologically interpretable features. Additionally, we developed a fast and low-computational-cost feature selection algorithm, capable of dynamically adjusting feature elimination, further enhancing the model's efficiency. Our comprehensive experimental analysis demonstrates the model's strong cross-dataset generalisation, achieving consistent performance on the ANM1 and ANM2 datasets, which supports its broader applicability to diverse gene expression datasets.

Our method offers a promising solution for early, non-invasive diagnosis of Alzheimer's disease, with clinical applications that could enhance patient outcomes by enabling timely intervention. The combination of deep learning and traditional machine learning techniques not only boosts model performance but also ensures interpretability—critical in building trust in clinical decision-making.

Looking ahead, further improvements, such as incorporating advanced techniques like multimodal data fusion, can enhance both the model's accuracy and its ability to unravel complex disease processes. While this study focuses on gene expression data, we recognise the importance of integrating multi-omics data, such as proteomics and genomics, to provide a more comprehensive understanding of Alzheimer's disease. Future work will explore the integration of multi-omics data, such as proteomics and genomics, to provide a more comprehensive understanding of Alzheimer's disease.

Acknowledgments

We gratefully acknowledge the use of datasets provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) and AddNeuroMed. The ADNI data used in this study were obtained from the ADNI database (adni.loni.usc.edu), which is supported by the National Institute on Aging (NIA) and the National Institute of Biomedical Imaging and Bioengineering (NIBIB). The AddNeuroMed data were provided by the AddNeuroMed consortium, supported by the European Commission under the Sixth Framework Programme (FP6). We extend our sincere thanks to the participants and researchers involved in these initiatives for their valuable contributions to advancing Alzheimer's disease research.

References

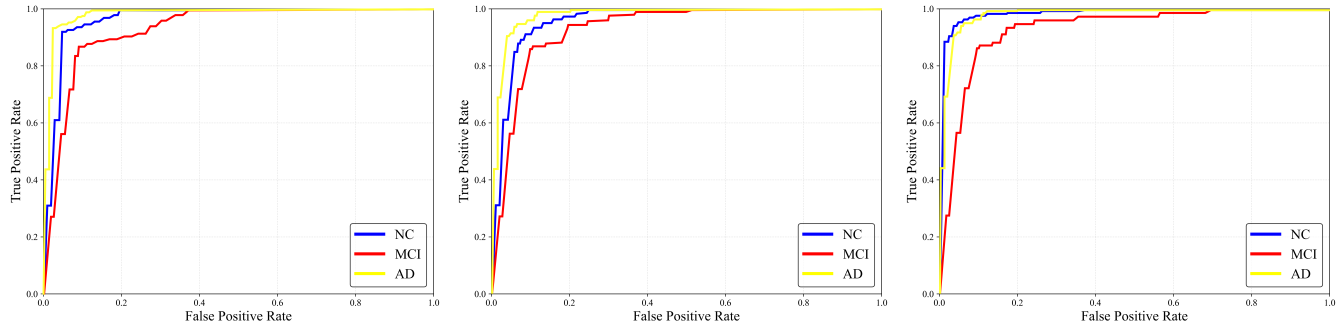
[1] "2023 alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 19, no. 4, pp. 1598–1695, 2023. [Online]. Available: https://alz-journals.onlinelibrary.wiley.com/doi/abs/10.1002/alz.13016

[2] W. H. Organization, *Dementia*. World Health Organization, 2021.

[3] B. Dubois, H. H. Feldman, C. Jacova, S. T. DeKosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, G. Jicha *et al.*, "Research criteria for the diagnosis of alzheimer's disease: revising the nincds–adrda criteria," *The Lancet Neurology*, vol. 6, no. 8, pp. 734–746, 2007.

[4] P. Kodam, R. Sai Swaroop, S. S. Pradhan, V. Sivaramakrishnan, and R. Vadrevu, "Integrated multi-omics analysis of alzheimer's disease shows molecular signatures associated with disease progression and potential therapeutic targets," *Scientific reports*, vol. 13, no. 1, p. 3695, 2023.

[5] M. V. Fernandez, M. Liu, A. Beric, M. Johnson, A. Cetin, M. Patel, J. Budde, P. Kohlfeld, K. Bergmann, J. Lowery *et al.*, "Genetic and multi-omic resources for alzheimer disease and related dementia from the knight alzheimer disease research center," *Scientific data*, vol. 11, no. 1, p. 768, 2024.



(a) ADNI (b) ANM1 (c) ANM2

Fig. 9: ROC curves for different datasets

[6] G. A. Cary, J. C. Wiley, J. Gockley, S. Keegan, S. S. Amirtha Ganesh, L. Heath, R. R. Butler III, L. M. Mangravite, B. A. Logsdon, F. M. Longo *et al.*, "Genetic and multi-omic risk assessment of alzheimer's disease implicates core associated biological domains," *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, vol. 10, no. 2, p. e12461, 2024.

[7] Z. Abbas, H. Tayara, and K. T. Chong, "Alzheimer's disease prediction based on continuous feature representation using multi-omics data integration," *Chemometrics and Intelligent Laboratory Systems*, vol. 223, p. 104536, 2022.

[8] W. H. Organization, *Ageing and health*. World Health Organization, 2022.

[9] N. Ertekin-Taner, "Gene expression endophenotypes: a novel approach for gene discovery in alzheimer's disease," *Molecular neurodegeneration*, vol. 6, pp. 1–14, 2011.

[10] A. Grubman, G. Chew, J. F. Ouyang, G. Sun, X. Y. Choo, C. McLean, R. K. Simmons, S. Buckberry, D. B. Vargas-Landin, D. Poppe *et al.*, "A single-cell atlas of entorhinal cortex from individuals with alzheimer's disease reveals cell-type-specific gene expression regulation," *Nature neuroscience*, vol. 22, no. 12, pp. 2087–2097, 2019.

[11] X. Sun, G. He, H. Qing, W. Zhou, F. Dobie, F. Cai, M. Staufenbiel, L. E. Huang, and W. Song, "Hypoxia facilitates alzheimer's disease pathogenesis by up-regulating bace1 gene expression," *Proceedings of the National Academy of Sciences*, vol. 103, no. 49, pp. 18 727–18 732, 2006.

[12] S. A. Semick, R. A. Bharadwaj, L. Collado-Torres, R. Tao, J. H. Shin, A. Deep-Soboslay, J. R. Weiss, D. R. Weinberger, T. M. Hyde, J. E. Kleinman *et al.*, "Integrated dna methylation and gene expression profiling across multiple brain regions implicate novel genes in alzheimer's disease," *Acta neuropathologica*, vol. 137, pp. 557–569, 2019.

[13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. [Online]. Available: https://www.nature.com/articles/nature14539

[14] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608014002135

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097–1105. [Online]. Available: https://papers.nips.cc/paper/2012/hash/6c1d0de8b8d5d90c77b0ef6e9b2201d4-Abstract.html

[16] L. H. Gilpin, M. Sandhu, and S. Jain, "Explaining explanations: An overview of interpretability of machine learning," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–14. [Online]. Available: https://dl.acm.org/doi/10.1145/3173574.3173580

[17] M. T. Ribeiro, S. Singh, and C. Guestrin, ""why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144. [Online]. Available: https://dl.acm.org/doi/10.1145/2939672.2939778

[18] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available: https://link.springer.com/article/10.1023/A:1010933404324

[19] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995. [Online]. Available: https://link.springer.com/article/10.1007/BF00994018

[20] R. Caruana, J. Gehrke, E. Koch, and P. Koch, "Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission," in *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1721–1730. [Online]. Available: https://dl.acm.org/doi/10.1145/2783258.2783347

[21] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2009. [Online]. Available: https://web.stanford.edu/~hastie/ElemStatLearn/

[22] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1995. [Online]. Available: https://link.springer.com/book/10.1007/978-1-4757-2440-0

[23] A. S. Alatrany, W. Khan, A. Hussain, H. Kolivand, and D. Al-Jumeily, "An explainable machine learning approach for alzheimer's disease classification," *Scientific Reports*, vol. 14, no. 1, p. 2637, 2024.

[24] S. Liu, C. Xu, Y. Zhang, J. Liu, B. Yu, X. Liu, and M. Dehmer, "Feature selection of gene expression data for cancer classification using double rbf-kernels," *BMC bioinformatics*, vol. 19, no. 1, pp. 1–14, 2018.

[25] Uzma, F. Al-Obeidat, A. Tubaishat, B. Shah, and Z. Halim, "Gene encoder: a feature selection technique through unsupervised deep learning-based clustering for large gene expression data," *Neural Computing and Applications*, pp. 1–23, 2020.

[26] B. B. Booij, T. Lindahl, P. Wetterberg, N. V. Skaane, S. Sæbø, G. Feten, P. D. Rye, L. I. Kristiansen, N. Hagen, M. Jensen *et al.*, "A gene expression pattern in blood for the early detection of alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 23, no. 1, pp. 109–119, 2011.

[27] K. Lunnon, M. Sattlecker, S. J. Furney, G. Coppola, A. Simmons, P. Proitsi, M. K. Lupton, A. Lourdusamy, C. Johnston, H. Soininen *et al.*, "A blood gene expression marker of early alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 33, no. 3, pp. 737–753, 2013.

[28] X. Li, H. Wang, J. Long, G. Pan, T. He, O. Anichtchik, R. Belshaw, D. Albani, P. Edison, E. K. Green *et al.*, "Systematic analysis and biomarker study for alzheimer's disease," *Scientific reports*, vol. 8, no. 1, p. 17394, 2018.

[29] O. Ahmed and A. Brifcani, "Gene expression classification based on deep learning," in *2019 4th Scientific International Conference Najaf (SICN)*. IEEE, 2019, pp. 145–149.

[30] R. Xie, J. Wen, A. Quitadamo, J. Cheng, and X. Shi, "A deep auto-encoder model for gene expression prediction," *BMC genomics*, vol. 18, pp. 39–49, 2017.

[31] A. B. Dincer, J. D. Janizek, and S.-I. Lee, "Adversarial deconfounding autoencoder for learning robust gene expression embeddings," *Bioinformatics*, vol. 36, no. Supplement_2, pp. i573–i582, 2020.

[32] K.-S. Hung, C.-C. Hsiao, T.-W. Pai, C.-H. Hu, W.-S. Tzou, W.-D. Wang, and Y.-R. Chen, "Functional enrichment analysis based on long noncoding rna associations," *BMC Systems Biology*, vol. 12, no. 4, pp. 109–118, 2018.

[33] R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed, "Exploration, normalization, and summaries of high density oligonucleotide array probe level data," *Biostatistics*, vol. 4, no. 2, pp. 249–264, 2003.

[34] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning repre-

sentations by back-propagating errors," *nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[35] A. Ng *et al.*, "Sparse autoencoder," *CS294A Lecture notes*, vol. 72, no. 2011, pp. 1–19, 2011.

[36] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[37] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine learning*, vol. 46, pp. 389–422, 2002.

[38] Y. Yilmaz and O. Ozdogan, "Liver disease as a risk factor for cognitive decline and dementia: an under-recognized issue," *Hepatology*, vol. 49, no. 2, pp. 698–698, 2009.

[39] D.-G. Kim, A. Krenz, L. E. Toussaint, K. J. Maurer, S.-A. Robinson, A. Yan, L. Torres, and M. S. Bynoe, "Non-alcoholic fatty liver disease induces signs of alzheimer's disease (ad) in wild-type mice and accelerates pathological signs of ad in an ad model," *Journal of neuroinflammation*, vol. 13, pp. 1–18, 2016.