IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510142

Early Detection of Alzheimer's Disease with Blood Plasma Proteins Using Support Vector Machines

Sharanabasavaraddi B[1], Spandana KP[2], Yashaswini LM[3], Madhuri JK[4]

Assistant Professor, Department of Computer Science, S.J.M. Institute of Technology, Chitradurga, Karnataka, India[1]

B.E. Students, Department of Computer Science and Engineering, S.J.M. Institute of Technology, Chitradurga, Karnataka, India[2,3,4]

Abstract: Alzheimer's disease (AD) is a progressive neurodegenerative disorder that affects memory, cognitive abilities, and daily functioning. Early identification of the disease is important for timely clinical assessment and patient management. This paper presents a machine-learning-based framework for classifying Alzheimer's disease using biomarker-related data and supervised learning algorithms, with particular emphasis on Support Vector Machine (SVM) and Logistic Regression (LR). The proposed workflow consists of dataset acquisition, data preprocessing, handling missing values, encoding categorical attributes, splitting the dataset into training and testing subsets, model training, and performance evaluation. The models are evaluated using standard classification metrics, including accuracy, precision, recall, sensitivity, and confusion matrix analysis. A web-based interface can make the trained model accessible to users by providing a structured input and prediction workflow. The framework aims to investigate the usefulness of machine learning in supporting the classification of Alzheimer's disease and to compare the performance of the selected algorithms. However, the reliability of the results depends on the quality, representativeness, and correct identification of the dataset. The proposed system is intended as a research prototype for decision support and not as a replacement for professional clinical diagnosis.

Keywords: Alzheimer's disease, Blood plasma proteins, Biomarkers, Machine learning, Support Vector Machine, Logistic Regression, Early detection, Classification, Predictive analytics, Healthcare decision support.

I. INTRODUCTION

Alzheimer's disease is one of the most common causes of dementia and is associated with progressive deterioration in memory, thinking, and other cognitive functions. As the disease advances, individuals may experience increasing difficulty performing routine activities and maintaining independence. The growing impact of dementia creates a need for reliable methods that can assist in identifying individuals who may require further clinical evaluation.

The proposed framework includes data preprocessing, training and testing, model classification, and performance evaluation. Its objective is to establish a structured and reproducible approach to machine-learning-based Alzheimer's disease classification and to explore the potential value of biomarker-related data in supporting further research.

II. RELATED WORK

Research on Alzheimer's disease detection has explored different data sources and computational approaches, including linguistic features, neuroimaging, electroencephalography, clinical variables, and blood-based biomarkers. These approaches demonstrate the broad potential of computational methods for analysing disease-related patterns, while also highlighting challenges involving data quality, feature selection, class imbalance, and generalization

III. PROBLEM STATEMENT AND OBJECTIVES

Alzheimer's disease classification is challenging because disease-related patterns may be difficult to identify from individual observations or multiple clinical measurements. Manual interpretation of complex datasets can also be time-consuming. Machine-learning methods offer a way to analyse input features systematically and identify patterns associated with predefined classes.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510142

The problem addressed in this study is to develop and evaluate a supervised machine-learning framework for classifying Alzheimer's disease using the selected dataset and to compare SVM and Logistic Regression based on appropriate classification metrics. The dataset and its target labels must be clearly documented to establish whether the study specifically evaluates blood plasma protein biomarkers.

IV. PROPOSED SYSTEM ARCHITECTURE

The dataset is loaded into the system and processed to address missing values and convert categorical attributes into a suitable numerical representation. The processed features and labels are divided into training and testing subsets. SVM and Logistic Regression are trained using the training data, and their predictions on the held-out test data are evaluated using the selected metrics. The results can then be presented through the application's interface.
For a valid experiment, preprocessing operations such as imputation and scaling must be fitted using the training data only and then applied to the test data. This prevents information leakage from the test set.

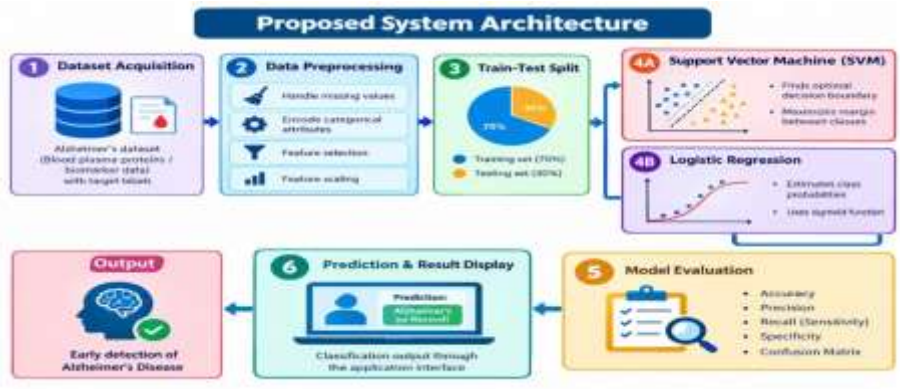


Fig. 1. Proposed System Architecture for Early Detection of Alzheimer's Disease Using Blood Plasma Proteins and Machine Learning.

V. METHODOLOGY

The methodology consists of six stages: dataset selection, preprocessing, data splitting, classifier training, performance evaluation, and result presentation. In the present study, we analyzed written samples obtained from Greek native speakers diagnosed with Alzheimer's in mild and moderate stages and from age matched cognitively normal controls (NC). We adopted a computational approach for the comparison of morph syntactic complexity and lexical variety in the samples. We used text classification approaches to assign the samples to one of the two groups. The classifiers were tested using various features: morph-syntactic and lexical characteristics. Degenerative conditions, such as Alzheimer's disease (henceforth AD) are commonly associated with deficits across a range of subcomponents of linguistic competence.

1 Dataset selection

The study uses an existing Alzheimer's disease dataset containing input features and corresponding class labels. For a paper specifically focused on blood plasma proteins, the dataset must contain documented plasma protein measurements and appropriate diagnostic labels.

2 Data preprocessing

Preprocessing prepares the dataset for model training and reduces errors caused by inconsistent or incomplete data. The following operations are considered

3 Data splitting

The project report specifies a 70:30 split, with 70% of the observations allocated to training and 30% reserved for testing. The training set is used to fit the classifiers, while the test set measures their performance on unseen observations

4 Support Vector Machine

Support Vector Machine is a supervised learning algorithm that identifies a decision boundary separating classes. It

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

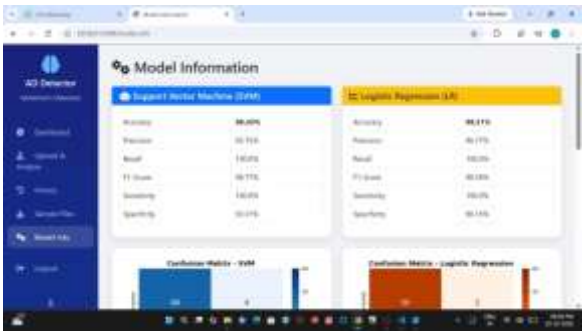
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

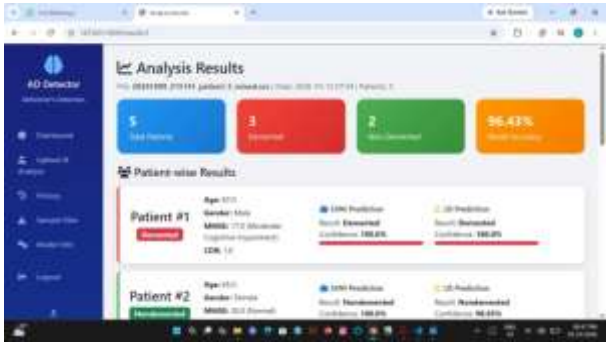
DOI: 10.17148/IJARCCE.2026.1510142

aims to maximise the margin between the boundary and the closest training observations, known as support vectors

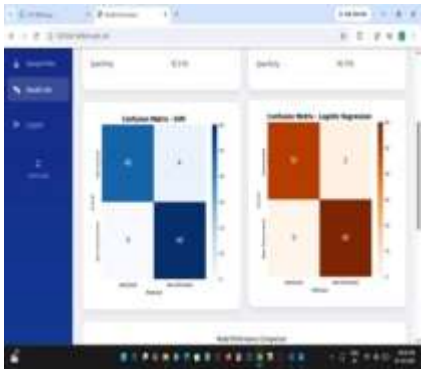
VI. RESULTS AND DISCUSSION

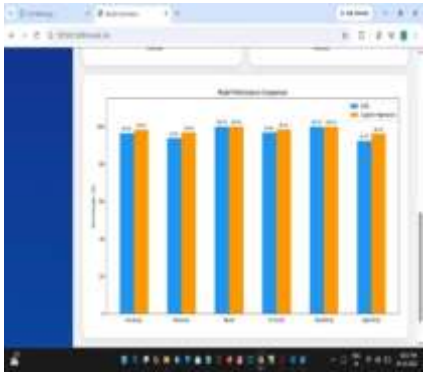
The project report identifies accuracy, precision, recall, sensitivity, and confusion matrix as evaluation measures. However, the numerical values needed to report the final experimental findings are not established in the extracted report text. The following table should therefore be populated using the actual output from your completed model.











VII. CONCLUSION AND FUTURE WORK

This paper presents a machine-learning framework for Alzheimer's disease classification using SVM and Logistic Regression. The workflow includes data preprocessing, training and testing partitioning, supervised model development, and evaluation using standard classification metrics. A software interface can provide a practical way to demonstrate the prediction workflow.

The framework establishes a structured approach to comparing classification algorithms on the selected dataset. Its effectiveness must be determined from the actual test-set results, rather than assumed in advance. Furthermore, the reliability and applicability of the findings depend on the dataset's provenance, feature definitions, sample size, and experimental design.

The proposed system should be regarded as a research prototype intended to support investigation of machine-learning-based classification. Additional validation is necessary before considering any clinical application.

ACKNOWLEDGMENT

The authors express their gratitude to the management, Principal, Department of Computer Science and Engineering, project coordinator, and faculty members of S.J.M. Institute of Technology, Chitradurga, for their guidance and support during the project. (Please add the project guide's name here if required by your college.)

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026
DOI: 10.17148/IJARCCE.2026.1510142

REFERENCES

[1] Alzheimer's Association, "2018 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 14, no. 3, pp. 367–429, 2018.
[2] M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, and M. Karagiannidou, World Alzheimer Report 2016: Improving Healthcare for People Living with Dementia: Coverage, Quality and Costs Now and in the Future. London, U.K.: Alzheimer's Disease International, 2016.
[3] B. Dubois et al., "Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria," Alzheimer's & Dementia, vol. 12, no. 3, pp. 292–323, 2016.
[4] M. S. Albert et al., "The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's & Dementia, vol. 7, no. 3, pp. 270–279, 2011.
[5] G. M. McKhann et al., "The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's & Dementia, vol. 7, no. 3, pp. 263–269, 2011.
[6] R. A. Sperling et al., "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's & Dementia, vol. 7, no. 3, pp. 280–292, 2011.

 This work is licensed under a Creative Commons Attribution 4.0 International License