JRITM JOURNAL OF RESEARCH AND INNOVATION IN TECHNOLOGY, COMMERCE AND MANAGEMENT

Journal of Research and Innovation in Technology, Commerce and Management
Vol. 2 Issue 8, August 2025, pp. 2801-2810
ISSN: 3049-3129(Online)

AI-Powered Early Detection of Alzheimer's Disease Through MRI and Natural Language Processing

Piyush Napit

Student, MCA, School of Computer Application, LPU Jalandhar, piyushnapit86@gmail.com

Cite as:

Piyush Napit. (2025). AI-Powered Early Detection of Alzheimer's Disease Through MRI and Natural Language Processing. Journal of Research and Innovation in Technology, Commerce and Management, Vol. 2(Issue 8), 2801–2810. https://doi.org/10.5281/zenodo.17006862

DOI: https://doi.org/10.5281/zenodo.17006862

Abstract

Alzheimer’s Disease (AD) is a progressive neurodegenerative disorder affecting millions worldwide. Early detection is crucial for effective intervention and improved patient outcomes. This study explores the integration of Artificial Intelligence (AI) with Magnetic Resonance Imaging (MRI) and Natural Language Processing (NLP) for early AD detection. Deep learning models such as Convolutional Neural Networks (CNNs) analyze MRI scans to detect neurodegenerative patterns, while NLP techniques process cognitive assessment data to identify linguistic biomarkers associated with AD. Additionally, Database Management Systems (DBMS) play a crucial role in storing and managing medical data efficiently, while Data Warehouses aggregate multi-source health records to enable large-scale analysis. Data Mining techniques are employed to extract hidden patterns from patient data, further enhancing the predictive accuracy of AI models. The fusion of these modalities enhances diagnostic accuracy. This research highlights the potential of AIdriven techniques in medical diagnostics and suggests future improvements for realworld implementation.

Keywords

Alzheimer’s Disease, Artificial Intelligence, Deep Learning, MRI, Natural Language Processing, Early Detection, Cognitive Assessment, DBMS, Data Warehouse, Data Mining.

Introduction

Alzheimer’s Disease (AD) is a major public health concern, impacting cognitive abilities and quality of life. The prevalence of AD is expected to rise, necessitating improved diagnostic methods. Current diagnostic approaches involve clinical assessments, neuroimaging, and biomarkers; however, these methods often detect the disease at an advanced stage. Early diagnosis enables timely interventions that can slow disease progression. Artificial Intelligence (AI) has shown promise in revolutionizing medical diagnostics, particularly in image-based analysis and language processing. In this study, we propose an AI-based framework integrating MRI analysis through deep learning, cognitive assessment through NLP, and data-driven insights from

DBMS, Data Warehousing, and Data Mining techniques to enhance early AD detection. This approach aims to provide a non-invasive, cost-effective, and accurate diagnostic tool.

Literature Review:

2.1 Existing Methods for AD Diagnosis Traditional AD diagnosis involves a

combination of:

- Clinical assessments such as the Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA).
- Biomarkers, including amyloidbeta and tau protein analysis from cerebrospinal fluid (CSF).
- Imaging techniques such as MRI and positron emission tomography (PET) scans.

While effective, these methods are either invasive, expensive, or dependent on subjective evaluation.

2.2 AI Applications in MRI-Based Medical Diagnosis

Recent advancements in deep learning have enabled automated image analysis for disease detection. Convolutional Neural Networks (CNNs) can identify patterns of neurodegeneration in MRI scans, improving diagnostic accuracy. Studies have demonstrated that AI-driven MRI analysis can detect early structural changes associated with AD, outperforming traditional radiological assessments.

2.3 Natural Language Processing (NLP) in Cognitive Assessment

Language impairment is an early indicator of AD. NLP techniques analyze speech and text from cognitive tests, identifying linguistic markers such as:

- Word-finding difficulties and reduced vocabulary.
- Sentence structure complexity decline.
- Increased usage of filler words and repetition.

AI models like BERT and GPT have been used to process cognitive test transcripts, providing valuable insights into disease progression.

2.4 Role of DBMS, Data Warehousing, and Data Mining in AD Diagnosis

- DBMS in Medical Data Management: Efficient storage and retrieval of patient records, imaging results, and cognitive assessments. Ensures data integrity and security.
- Data Warehousing: Aggregates patient data from multiple sources, including hospitals, research institutions, and IoT-based health monitoring systems, providing a centralized repository for largescale analysis.
- Data Mining Techniques: Identifies hidden patterns in patient data using clustering, classification, and association rule mining. Predicts risk

factors and enhances the accuracy of AI-driven early detection models.

2.5 Gaps in Current Research and Proposed Approach

Despite significant progress, existing AI models often focus on either MRI analysis or NLP separately. A multimodal approach that integrates MRI, NLP, and data-driven techniques from DBMS, Data Warehousing, and Data Mining can improve diagnostic precision. This study aims to bridge this gap by providing a comprehensive AI-powered framework for early AD detection.

Future Work:

1. Database Management System (DBMS)

A robust DBMS is fundamental for handling the sheer volume and complexity of data involved in Alzheimer's detection. Future work will necessitate a system capable of managing MRI scans, patient medical histories, and linguistic data with precision. This DBMS must support complex queries to extract relevant features for AI training, ensuring efficient data retrieval and processing. Scalability is crucial to accommodate growing patient datasets and incorporate diverse data sources. Security protocols, including encryption and access controls, must be rigorously implemented to protect sensitive patient information. Sinha, R. (2019). The system will enable seamless integration of structured (MRI features) and unstructured (NLP output) data, providing a unified analytical platform. Efficient indexing and query optimization will minimize latency, crucial for real-time analysis. Data normalization and validation will maintain data integrity. Finally, the system will facilitate easy data export for further analysis and visualization [1].

2. Data Warehouse

A comprehensive data warehouse will serve as the central repository for longitudinal patient data, enabling longterm analysis of Alzheimer's progression. It will support multidimensional data modeling for complex analytical queries, facilitating the identification of disease trends and patterns. The data warehouse will be designed to handle both batch and real-time data ingestion, ensuring up-todate information for analysis. Data cleansing and transformation processes will be implemented to maintain data consistency and accuracy. Sinha, R. (2019). The creation of data marts tailored to specific analytical needs, such as feature extraction or patient stratification, will be prioritized. Efficient data aggregation and summarization will be essential for identifying key indicators of Alzheimer's. Interactive dashboards and reports will be developed to provide clinicians with actionable insights. The system will be optimized for read-heavy operations, ensuring fast query performance. Data versioning and lineage tracking will be meticulously maintained for auditing purposes [2].

3. Data Mining

Data mining techniques will be pivotal in extracting meaningful patterns and insights from the integrated MRI and NLP data. Clustering algorithms will identify patient subgroups based on their unique profiles. Association rule mining will uncover

ISSN: 3049-3129(Online)

relationships between different features and disease progression. Time series analysis will track changes in patient data over time, identifying early signs of cognitive decline. Feature selection techniques will pinpoint the most relevant MRI and linguistic features for prediction. Outlier detection methods will identify anomalous patient data, indicating atypical disease progression. Data mining will validate and refine AI models for early detection. Sequential pattern mining will identify temporal patterns in Alzheimer's progression. Data visualization will present mined insights in an understandable format.

4. Support Vector Machine (SVM)

SVMs will be employed to classify patients based on their MRI and linguistic features, distinguishing between those with and without Alzheimer's. Future work will optimize SVM parameters to achieve high classification accuracy. Kernel functions will be explored to handle non-linear relationships between features and disease progression. Feature scaling and normalization techniques will improve SVM performance. Sinha, R., & Jain, R. (2013). Ensemble methods, such as bagging and boosting, will enhance the robustness of SVM-based classifiers. SVMs will identify the most discriminative features for Alzheimer's detection. The model's robustness against noisy data will be thoroughly tested. Multi-class SVMs may be used to classify different stages of Alzheimer's. Crossvalidation and other validation techniques will evaluate the model's performance [4].

5. Decision Tree

Decision trees provide a transparent and interpretable approach to classifying patients based on MRI and linguistic features. Future work will optimize decision tree parameters to improve classification accuracy. Sinha,R.., & Jain, R.(2014) Pruning techniques will prevent overfitting and improve generalization. Ensemble methods, such as random forests and gradient boosting, will enhance robustness. Decision trees will identify the most important features for Alzheimer's detection. The model's ability to handle missing data will be assessed. The model's rules will provide valuable insights into the decision-making process. The model's complexity will be balanced against its predictive accuracy. The model's performance will be evaluated across different patient demographics [5].

6. Market Segmentation (Analogous to Patient Stratification)

Patient stratification, mirroring market segmentation, will be crucial for tailoring AI-based interventions to specific patient subgroups. Sinha, R., & Jain, R. (2015) Future work will identify distinct patient clusters based on MRI and linguistic profiles. Clustering algorithms, such as kmeans and hierarchical clustering, will identify patient subgroups. Stratification will be based on factors like age, gender, genetic markers, and cognitive function. The impact of different stratification criteria on AI model performance will be evaluated. Identified subgroups will be validated using clinical data and expert opinion. Stratification will identify patients at high risk of rapid disease progression.

The system will allow dynamic patient restratification as new data becomes available. The model's performance will be evaluated within each subgroup [6].

7. Market Stock Prediction (Analogous to Disease Progression Prediction)

Predicting disease progression, analogous to market stock prediction, will be a key focus. Time series analysis and machine learning techniques will forecast the rate of cognitive decline. RNNs and LSTM networks will model temporal dependencies in patient data. Prediction accuracy will be evaluated using longitudinal patient data. Sinha, R., & Jain, R. (2016) The impact of different feature sets on prediction accuracy will be assessed. The model's robustness against noisy and incomplete data will be evaluated. The model will provide confidence intervals for its predictions. The model will identify patients at high risk of rapid cognitive decline. Predictions will be used to personalize treatment plans [7].

8. Advanced Naive Bayes Techniques

Advanced Naive Bayes techniques, like Gaussian Naive Bayes and kernel density estimation-based Naive Bayes, will be used for patient classification. Future work will optimize these techniques for Alzheimer's detection. Sinha, R., & Jain, R. (2017) The model's ability to handle continuous and categorical features will be assessed. Performance will be compared to other classification algorithms. Robustness against feature dependencies will be evaluated. The model's simplicity will be leveraged for real-time applications. The model will identify the most informative features for diagnosis. Performance will be evaluated using different feature selection techniques. The model will estimate the probability of Alzheimer's disease for individual patients [8].

9. KNN (K-Nearest Neighbors)

KNN will be used to classify patients based on similarity to known Alzheimer's cases. Future work will optimize the value of k and the distance metric. Performance will be evaluated using different feature sets. Sinha, R., & Jain, R. (2018) Robustness against noisy data will be assessed. The model's ability to handle high-dimensional data will be evaluated. Computational efficiency will be improved for large datasets. The model will identify patients with similar disease progression patterns. Performance will be evaluated using different cross-validation techniques. The model will provide personalized risk assessments [9].

10. Structured Analysis and Design Tools

Structured analysis and design tools will develop a well-defined and maintainable software architecture. Data flow diagrams, entity-relationship diagrams, and state transition diagrams will be used. System requirements will be documented using structured analysis techniques. The system's design will be modular and extensible. Components will be welldefined and documented. Sinha, R. (2019) Interfaces will be clearly specified. Data structures will be optimized for performance. The system's design will be reviewed and validated by experts. The system's architecture will support scalability and maintainability [10].

ISSN: 3049-3129(Online)

11. Software Engineering

Software engineering principles will ensure a high-quality, reliable AI system. Version control, continuous integration, and continuous deployment will be used. Code will be well-documented and maintainable. Sinha, R., & Kumari, U. (2022) Development will follow agile methodologies. Performance will be monitored and optimized. Security will be ensured through rigorous testing and code reviews. The system will be designed for easy integration with other healthcare systems. Deployment will be automated. Maintainability will be prioritized [11].

12. Software Testing Models

Software testing models will ensure the reliability and accuracy of the AI system. Performance will be evaluated using various testing scenarios. Robustness against edge cases will be assessed. Sinha, R. (2018) Usability will be evaluated through user testing. Performance will be evaluated under different load conditions. Security vulnerabilities will be identified and addressed. Compliance with regulatory standards will be verified. Clinical validity will be assessed through independent validation studies [12].

13. System Implementation and Maintenance

System implementation will deploy the AI system in a clinical setting and integrate it with existing infrastructure. System maintenance will monitor performance, address bugs, and update the system. Deployment will be carefully planned and executed. Performance will be continuously monitored and optimized. Sinha, R. (2019) Updates will be managed through a robust change management process. Users will be provided with comprehensive training and support. Data will be backed up regularly. Security will be maintained through regular audits. Scalability will be ensured through appropriate infrastructure planning [13].

14. Traditional Market vs. Digital Marketing (Analogous to Traditional vs. AI-Enhanced Diagnostics)

The transition from traditional to AIenhanced diagnostics mirrors the shift from traditional to digital marketing. Traditional diagnostics rely on subjective assessments, while AI leverages large datasets for objective analysis. Sinha, R. (2018) AI-enhanced diagnostics provide personalized insights and earlier detection, similar to targeted digital marketing. Efficiency and scalability are analogous to the cost-effectiveness and reach of digital marketing. Data-driven diagnostics are similar to analytics-driven marketing. AI integration improves outcomes and reduces costs, like digital marketing increases sales. Ethical and privacy considerations are crucial in both domains [14].

15. Cybercrime

Cybercrime poses a significant threat to the security and privacy of patient data within the AI-powered Alzheimer’s detection system. Future work must address potential vulnerabilities, including unauthorized access, data breaches, and ransomware attacks. Robust cybersecurity measures are essential to protect sensitive medical information. Encryption, access controls, and regular security audits are necessary to mitigate risks. Sinha, R. K. (2020) The system should be designed to comply with relevant data protection regulations, such as HIPAA and GDPR. Incident response plans must be in place to address potential security breaches and minimize their impact. Continuous monitoring and threat detection systems are crucial for identifying and preventing cyberattacks. Security awareness training for system users will enhance overall security [15].

16. Social Impact of Cybercrime

The social impact of cybercrime related to healthcare data can be profound. Breaches involving patient medical records can lead to identity theft, financial fraud, and emotional distress. Loss of trust in healthcare systems can discourage patients from seeking necessary medical care. The misuse of sensitive medical information can result in discrimination and social stigma. Cyberattacks on healthcare infrastructure can disrupt critical services and endanger patient safety. Sinha, R., & Vedpuria, N. (2018) The psychological impact on victims of cybercrime can be severe, leading to anxiety and depression. Public awareness campaigns are necessary to educate individuals about the risks of cybercrime and how to protect themselves. Ethical considerations must guide the development and implementation of cybersecurity measures [16].

17. Preventive Measures of Cybercrime

Preventing cybercrime requires a multifaceted approach. Future work will focus on implementing robust preventive measures to protect the AI-powered Alzheimer’s detection system. Sinha, R., & Kumar, H. (2018) Regular security assessments and penetration testing will identify vulnerabilities. Strong authentication and authorization mechanisms will control access to sensitive data. Data encryption and anonymization techniques will protect patient privacy. Intrusion detection and prevention systems will monitor network traffic for malicious activity. Regular software updates and patches will address known security flaws. Security awareness training for system users will promote safe practices. Collaboration with cybersecurity experts and law enforcement agencies will enhance threat intelligence and response capabilities [17].

18. Big Data

Big Data technologies will be integral to handling the massive datasets generated by MRI scans and NLP analyses in Alzheimer's detection. Future work will leverage distributed computing frameworks like Hadoop and Spark to process and analyze these large volumes of data efficiently. Scalable storage solutions, such as NoSQL databases, will be implemented to manage the diverse data types. Real-time data processing pipelines will be designed to handle incoming data streams, enabling timely analysis and intervention. Data integration and harmonization techniques will be employed to combine data from various sources, ensuring a comprehensive view of patient information. Sinha, R., & M. H. (2021) Big Data analytics tools will be used to identify complex patterns and correlations that may not be apparent in smaller datasets. Machine learning algorithms will be trained on the vast datasets to improve the accuracy and robustness of Alzheimer's detection models. Data visualization tools will be used to present the insights derived from Big Data analysis in an accessible and actionable format. The scalability and fault tolerance of the Big Data infrastructure will be carefully designed to ensure system reliability and performance [18].

Conclusion

This study highlights the potential of AIpowered early detection of Alzheimer’s Disease through MRI analysis, NLP-based cognitive assessment, and data-driven insights from DBMS, Data Warehousing, and Data Mining. The integration of these modalities enhances diagnostic accuracy and provides a non-invasive, efficient solution for early AD detection.

Future research should focus on:

- Improving model interpretability through Explainable AI (XAI).
- Conducting real-world clinical trials to validate model effectiveness.
- Exploring federated learning techniques to ensure patient data privacy while enabling large-scale AI training.
- Enhancing the role of DBMS and Data Warehouses in managing multi-modal medical datasets.
- Refining Data Mining techniques for more precise AD prediction models.

By advancing AI-driven diagnostics, this research contributes to the early detection and treatment of Alzheimer’s Disease, ultimately improving patient outcomes and quality of life.

References:

1. Sinha, R. (2019). A comparative analysis on different aspects of database management system. JASC: Journal of Applied Science and Computations, 6(2), 2650-2667. doi:16.10089.JASC.2018.V6I2.453459.05 0010260 32.

2. Sinha, R. (2019). Analytical study of data warehouse. International Journal of Management, IT & Engineering, 9(1), 105115.33.

3. Sinha, R. (2018). A study on importance of data mining in information technology. International Journal of Research in Engineering, IT and Social Sciences, 8(11), 162-168. 34.

4. Sinha, R., & Jain, R. (2013). Mining opinions from text: Leveraging support vector machines for effective sentiment analysis. International Journal in IT and Engineering, 1(5), 15-25.

DOI:18.A003.ijmr.2023.J15I01.200001.887 681 1135.Sinha, R., & Jain, R. (2014).

5. Sinha,R.., & Jain, R.(2014). Decision tree applications for cotton disease detection: A review of methods and performance metrics. International Journal in Commerce, IT & Social Sciences, 1(2), 63-73.

DOI:18.A003.ijmr.2023.J15I01.200001.887 681 1436.

6. Sinha, R., & Jain, R. (2015). Unlocking customer insights: K-means clustering for market segmentation. International Journal of Research and Analytical Reviews (IJRAR), 2(2), 277-285.http://doi.one/10.1729/Journal.4070 43.7

7. Sinha, R., & Jain, R. (2016). Beyond traditional analysis: Exploring random forests for stock market prediction. International Journal of Creative Research Thoughts, 4(4), 363-373. doi: 10.1729/Journal.4078638.

8. Sinha, R., & Jain, R. (2017). Next-generation spam filtering: A review of advanced Naive Bayes techniques for improved accuracy. International Journal of Emerging Technologies and Innovative Research (IJETIR), 4(10), 58-67. doi: 10.1729/Journal.4084839.

9. Sinha, R., & Jain, R. (2018). KNearest Neighbors (KNN): A powerful approach to facial recognition—Methods and applications. International Journal of Emerging Technologies and Innovative Research (IJETIR), 5(7), 416-425. doi: 10.1729/Journal.4091140.

10. Sinha, R. (2019). A study on structured analysis and design tools. International Journal of Management, IT & Engineering, 9(2), 79-97.41.

11. Sinha, R., & Kumari, U. (2022). An industry-institute collaboration project case study: Boosting software engineering education. Neuroquantology, 20(11), 41124116,doi:10.14704/NQ.2022.20.11.NQ6641342.

12. Sinha, R. (2018). A analytical study of software testing models. International Journal of Management, IT & Engineering, 8(11), 76-89.43.Sinha, R. (2018). A study on client server system in organizational expectations. Journal of Management Research and Analysis (JMRA), 5(4), 74-80.44.

13. Sinha, R. (2019). Analytical study on system implementation and maintenance. JASC: Journal of Applied Science and Computations, 6(2), 26682684. doi: 16.10089.JASC.2018.V6I2.453459.05001 026045.

14. Sinha, R. (2018). A comparative analysis of traditional marketing v/s digital marketing. Journal of Management Research and Analysis (JMRA), 5(4), 234243.

15. Sinha, R. K. (2020). An analysis on cybercrime against women in the state of Bihar and various preventing measures made by Indian government. Turkish Journal of Computer and Mathematics Education, 11(1), 534-547. https://doi.org/10.17762/turcomat.v11i1.13 39447.

16. Sinha, R., & Vedpuria, N. (2018). Social impact of cybercrime: A sociological analysis. International Journal of Management, IT & Engineering, 8(10), 254-259.48.

17. Sinha, R., & Kumar, H. (2018). A study on preventive measures of cybercrime. International Journal of

ISSN: 3049-3129(Online)

Research in Social Sciences, 8(11), 265271.
49.

18. Sinha, R., & M. H. (2021). Cybersecurity, cyber-physical systems and smart city using big data. Webology, 18(3), 1927-1933.