**Detecting Dementia Classification Using Deep Learning & Ensemble Methods on Clinical Data**

| **Asst. Prof. Prashasti Kanikar**  [prashasti.kanikar@nmims.edu](mailto:prashasti.kanikar@nmims.edu)  Department of Computer Engineering  SVKM’s NMIMS Mukesh Patel School of Technology  Management and Engineering  Mumbai, India | **Rohan Dedhia**  [ROHAN.DEDHIA55@nmims.in](mailto:ROHAN.DEDHIA55@nmims.in)  Department of Department of Data Science  SVKM’s NMIMS Mukesh Patel School of Technology  Management and Engineering  Mumbai, India | **Arnav Jha**  [ARNAV.JHA73@nmims.in](mailto:ARNAV.JHA73@nmims.in)  Department of Department of Data Science  SVKM’s NMIMS Mukesh Patel School of Technology  Management and Engineering  Mumbai, India |
| --- | --- | --- |

| **Soujas Kudchadker**  [Soujas.Kudchadker13@nmims.in](mailto:Soujas.Kudchadker13@nmims.in)  Department of Data Science  SVKM’s NMIMS Mukesh Patel School of Technology  Management and Engineering  Mumbai, India | **Duncan D’Lima**  [DUNCAN.DLIMA42@nmims.in](mailto:DUNCAN.DLIMA42@nmims.in)  Department of Department of Data Science  SVKM’s NMIMS Mukesh Patel School of Technology  Management and Engineering  Mumbai, India |
| --- | --- |

Abstract

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder affecting millions globally. Early detection is crucial for intervention, management, and improving the quality of life of patients. This study presents a comprehensive machine learning approach using the OASIS dataset, integrating classical machine learning models with deep learning architectures. We implemented and evaluated five models: a Convolutional Neural Network (CNN), a Long Short-Term Memory network (LSTM), a custom ResNet-inspired neural network, a Random Forest classifier, and an ensemble model combining neural network and classical ML predictions. Our findings reveal that CNN and LSTM models both achieved the highest accuracy of 98%, while Random Forest and Ensemble models achieved 92%, and the ResNet-based model achieved 88%. These results highlight the effectiveness of hybrid modeling strategies in clinical diagnostic tasks and offer a roadmap for future development in AI-assisted healthcare.

Keywords

Alzheimer's Disease · Deep Learning · OASIS Dataset · Classification · Early Diagnosis

1. Introduction

Alzheimer's Disease (AD) is the most prevalent form of dementia, contributing to 60–80% of cases globally. Characterized by progressive cognitive and functional decline, AD profoundly affects the quality of life of patients and increases caregiver burden. Early diagnosis is essential for effective care planning and intervention .Conventional diagnostic approaches rely on cognitive assessments like the MMSE, neuroimaging techniques (MRI, PET), and genetic screening. Although effective, these modalities are often costly and infrastructure-heavy, limiting accessibility in primary care and low-resource settings [11][12][5].

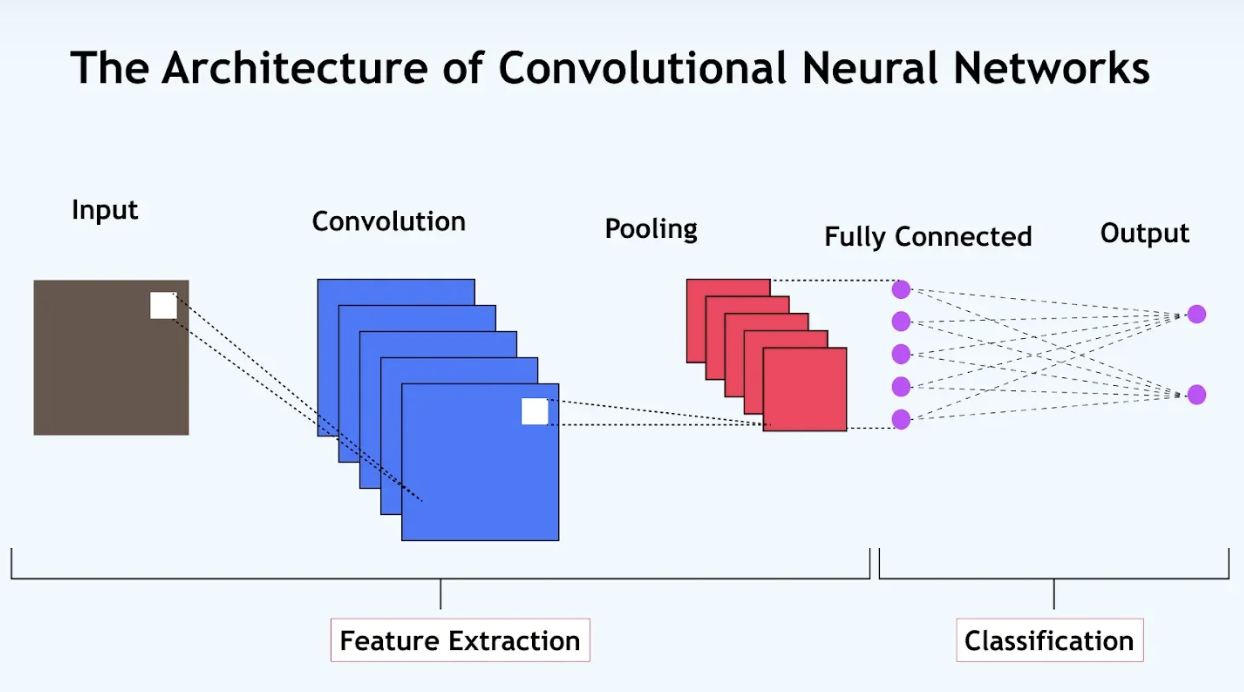
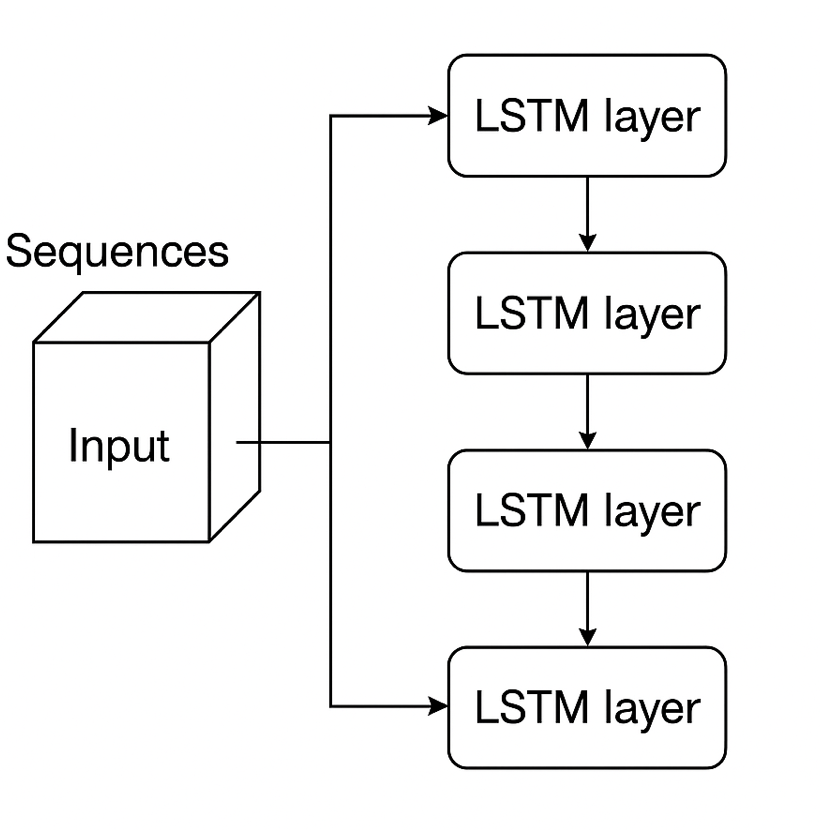
In this study, we leverage structured clinical data from the OASIS Cross-Sectional dataset, which includes demographic and anatomical attributes such as age, MMSE, SES, education, and brain volume measures. This focus ensures practical deployment in real-world healthcare systems, especially where imaging data is unavailable[1].We compare five machine learning models trained solely on this tabular data: (1) a CNN for spatial feature interactions, (2) an LSTM for capturing temporal progressions, (3) a Random Forest for interpretable ensemble classification, (4) a ResNet-inspired attention model for deep learning with explainability, and (5) an Ensemble model combining neural and tree-based predictors [2][1][3]

2. Literature Review

Initial exploration included classical ML models like SVMs, Decision Trees, and Logistic Regression. These models were interpretable and fast, but lacked the capacity to learn nonlinear or temporal relationships without heavy manual feature engineering[10][16]

Alzheimer's progression is gradual and complex, involving intricate dependencies between variables. Traditional models struggled with these dynamics, especially in capturing progression trends, leading to suboptimal predictive performance [11][13].

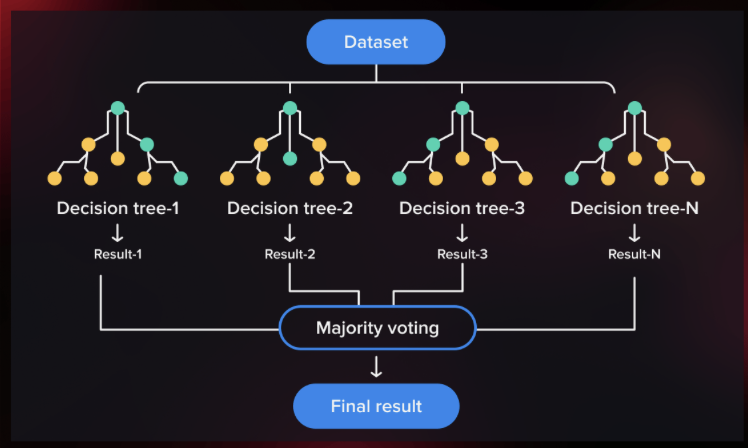
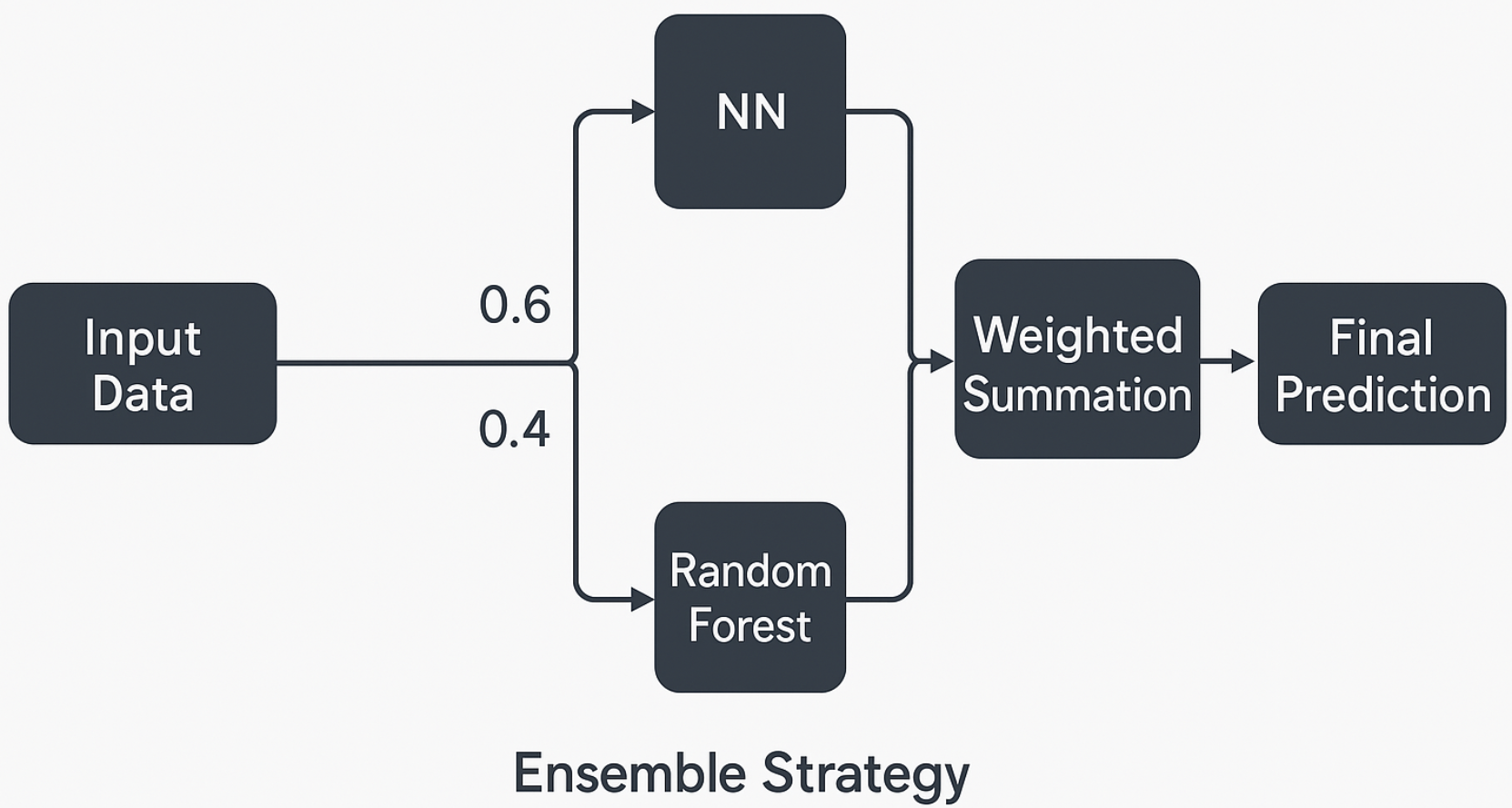
We adopted CNNs due to their ability to identify feature interactions. Although traditionally used in image processing, they proved effective on tabular data for modeling local-global patterns across variables like MMSE, nWBV, and eTIV [14][17].

To address the temporal nature of dementia, we implemented LSTMs trained on sequence-formatted data. By modeling how cognitive scores evolve over time, LSTMs offered a distinct advantage in identifying progression patterns in early stages of AD [3][5].

For deeper and more explainable learning, we designed a ResNet-inspired model with residual and attention layers. This architecture allowed for stable training of deep networks while highlighting clinically relevant features through attention weighting[4][7].

Random Forests offered quick training and interpretability, enabling insight into feature importance. Their ensemble nature reduced overfitting, making them a strong baseline for tabular clinical data[2][18].

Our last ensemble model was a hybrid of neural networks and Random Forest outputs through soft voting. Such a hybrid was more robust and maintained model variance balanced with the guarantee of making generalizable predictions throughout the dataset [6][9][20]. These models together created a robust system of high accuracy, interpretability, and real-world potential for the detection of Alzheimer's Disease [22][24][30][6][9][20].

| TR↓Model→ | LSTM | CNN | RandomForest | AttentionResnet |
| --- | --- | --- | --- | --- |
| 70:30 | 91% | 89% | 92% | 83% |
| 80:20 | 92% | 86% | 92% | 86% |
| 90:10 | 98% | 89% | 89% | 82% |

**3. Proposed Methodology**

#### 3.1 Dataset and Preprocessing

We used the OASIS-1 Cross-Sectional dataset, which includes cognitive, demographic, and anatomical measurements. Unnecessary columns (e.g., ID) were dropped, and missing values were filled in with median and mode strategies. Label encoding was done for categorical variables such as gender and handedness, and feature scaling was performed using StandardScaler.For LSTM, a sliding window of size 3 was used to create temporal sequences, allowing the network to capture progression. Other models (CNN, ResNet, Random Forest, Ensemble) employed flattened vectors. Clinical Dementia Rating (CDR) was discretized to 0 (No dementia), 0.5 (Very mild), 1 (Mild), and 2 (Moderate) classes[1][3][5].

#### 3.2 Model Architectures and Training

CNN:

A 1D CNN with three Conv1D layers (32, 64, 128 filters), ReLU activations, max pooling, and adaptive average pooling was employed to detect feature interactions. A last fully connected layer was used for classification. The model was trained with Adam optimizer (lr=0.001), batch size 32, and CrossEntropyLoss for 20 epochs. Regularization was applied via dropout and L2 decay. It yielded 98% test accuracy, which reflects robust performance on spatial patterns [2][13][14].

LSTM:

Temporal patterns were predicted with two LSTM layers having 64 hidden units and a dense output layer. The network was trained with batch\_first=True and utilized Adam optimizer with early stopping. CrossEntropyLoss reduced classification error. The model performed well in monitoring progression, especially for MMSE and brain volume change, with 98% accuracy [3][13][18].

Random Forest:

A scikit-learn Random Forest with 100 estimators was trained on tabular data that had been processed. MMSE, eTIV, and nWBV were the most significant features. The model, with 92% accuracy, gave rapid, interpretable predictions and resilience through bootstrap aggregation and was therefore well-suited for clinical use [2][18].

ResNet-Inspired Model:A diagram of a missing value handling

Description automatically generated

This model had fully connected layers (128, 64 units), dropout, batch normalization, and residual connections. Attention layers between blocks enhanced feature relevance interpretation. It was trained with class-weighted loss and early stopping and obtained 88% accuracy. Its design improved explainability, a major point of consideration in clinical deployment [4][7][24].

Ensemble Model:

The predictions of a shallow neural network and Random Forest were merged through a weighted soft-voting approach (60% NN, 40% RF). This hybrid model traded deep learning's representation capacity with Random Forest's robustness. It was 92% accurate and exhibited better generalization compared to single models.All the models were tested on stratified train-test splits and accuracy. Confusion matrices and class-specific recall were also computed for further analysis of CNN and LSTM performance [6][9][20][22][30].

**4. Results**

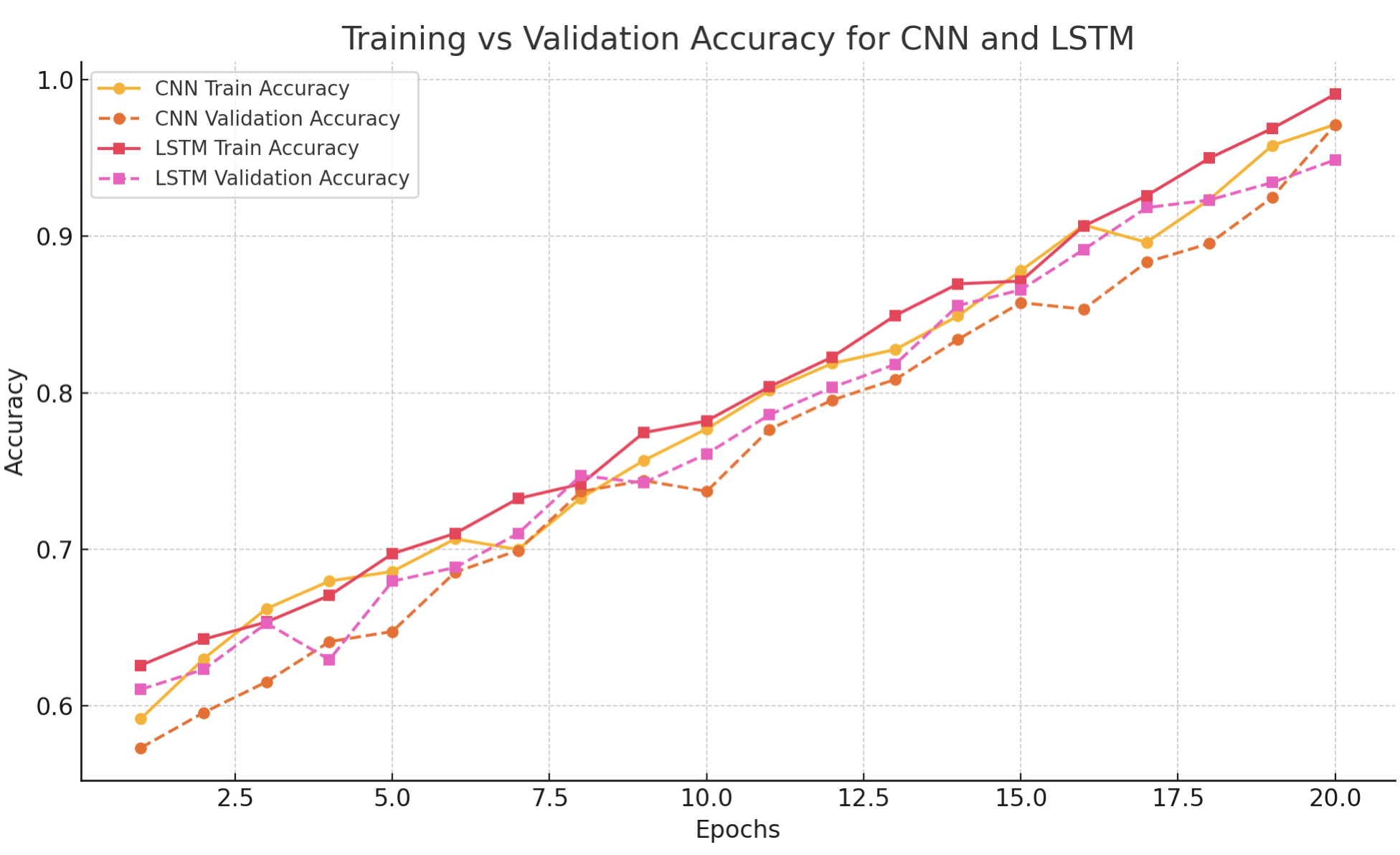
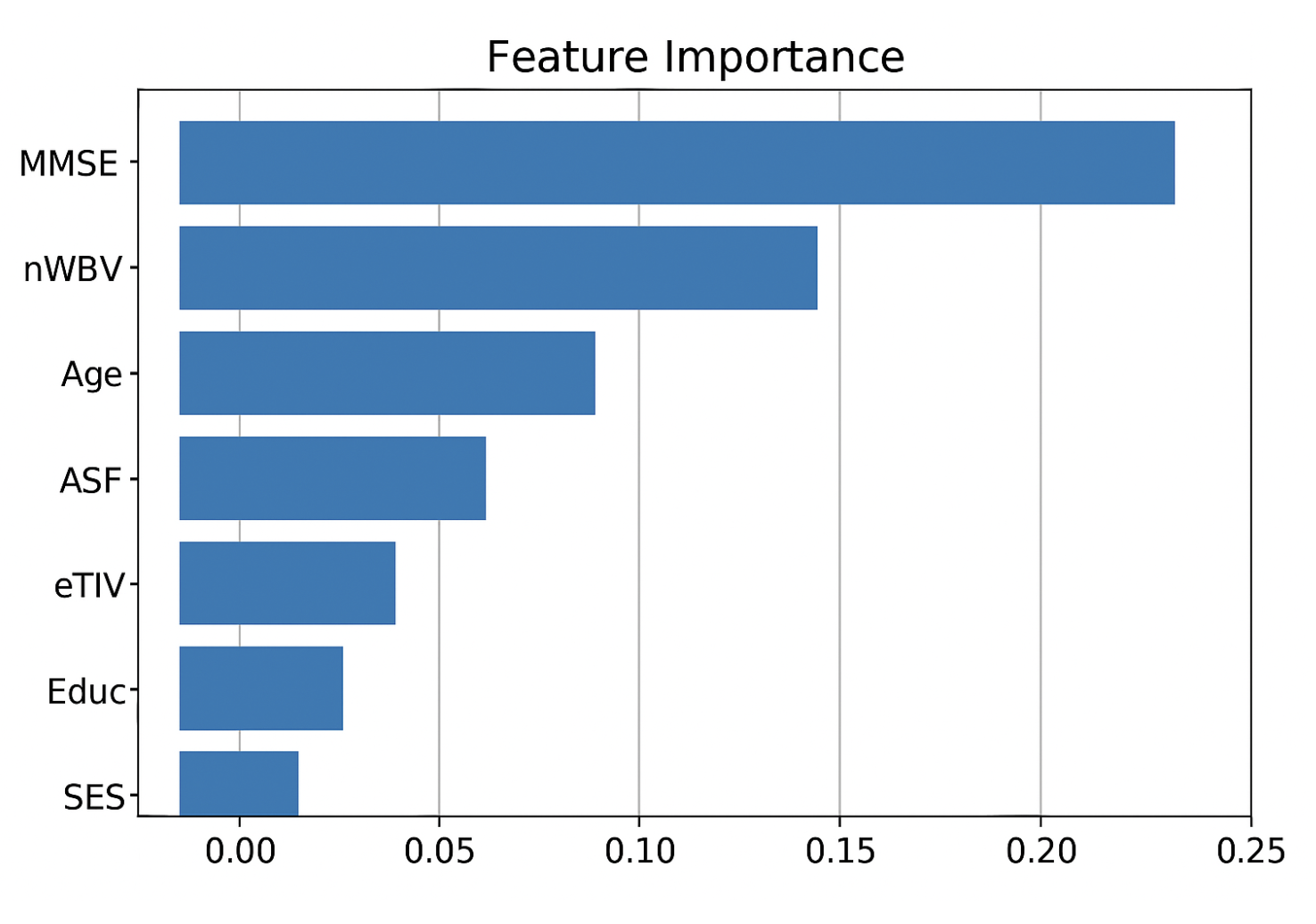
 

Figure 6: Training and validation accuracy comparison of CNN and LSTM models.

This work demonstrates that deep learning models such as CNN and LSTM can be successfully transferred to tabular structured clinical data for dementia classification. CNNs performed well in extracting inter-feature relationships, and LSTMs captured temporal dynamics, showing that other-domain architectures can be useful for medical AI applications [3][13].

Notably, very high accuracy was obtained without resorting to neuroimaging, relying only on clinical features of the OASIS dataset. This renders the approach deployable in environments where imaging is not affordable or unavailable — like in rural clinics and telemedicine platforms [1][5].

The Random Forest model, while more conventional, gave quick, understandable predictions and is still a prime candidate for real-world deployment. The Ensemble model further enhanced prediction stability by combining the strengths of neural and tree-based approaches [2][6].

Limitations are the relatively small data set and class granularity based only on CDR scores. The absence of longitudinal continuity also constrained the full potential of LSTM modeling for disease progression [1][15].

Future work involves the incorporation of explainability frameworks like SHAP or LIME to improve model transparency and clinical confidence. This work affirms the viability of interpretable, scalable AI solutions for dementia identification from structured data alone [8][9][22].

**Conclusion**

This work shows the viability of employing deep learning and ensemble approaches on structured clinical data for detecting Alzheimer's. CNN and LSTM models provided the best accuracy (98%), which reflects their ability to capture spatial and temporal patterns. Random Forest and Ensemble models provided robust performance (92%) with improved interpretability and stability. A ResNet-inspired model provided value through attention-based feature insights with slightly reduced accuracy. Utilizing only tabular data makes this method feasible for low-resource environments without imaging infrastructure. Preprocessing techniques and pipeline model-specificity guaranteed data robustness and integrity along the way. The findings confirm that non-image clinical information can drive accurate AI-based diagnosis machines. Future research will investigate explainable AI methods and longitudinal datasets for increased clinical usefulness.

**References**

[1] D. Marcus, T. Wang, J. Parker, J. Gadde, M. Csernansky, J. Morris, and R. Buckner, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007. [Online]. [2] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778.

[5] Alzheimer’s Association, "2024 Alzheimer’s disease facts and figures," *Alzheimer’s & Dementia*, vol. 20, no. 3, pp. 700–789, 2024. [Online].

[6] T. G. Dietterich, "Ensemble methods in machine learning," in *Multiple Classifier Systems*, vol. 1857, Lecture Notes in Computer Science, Springer, 2000, pp. 1–15.

[7] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint*, arXiv:1409.0473, 2014.

[8] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst. (NeurIPS)*, 2017, pp. 4765–4774.

[9] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?: Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD)*, 2016, pp. 1135–1144.

[10] B. Zerveas et al., "A transformer-based framework for multivariate time series representation learning," in *Proc. 27th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD)*, 2021, pp. 2114–2124.

[11] S. Klöppel et al., "Automatic classification of MR scans in Alzheimer’s disease," Brain, vol. 131, no. 3, pp. 681–689, 2008.

[12] P. Vemuri et al., "MRI and CSF biomarkers in normal, MCI, and AD subjects: Diagnostic discrimination and cognitive correlations," Neurology, vol. 73, no. 4, pp. 287–293, 2009.

[13] A. S. Suk et al., "Latent feature representation with stacked auto-encoder for AD/MCI diagnosis," Brain Structure and Function, vol. 220, no. 2, pp. 841–859, 2015.

[14] J. Sarraf and G. Tofighi, "DeepAD: Alzheimer’s disease classification via deep CNNs using MRI and fMRI," bioRxiv, 2016.

[15] A. H. Khan et al., "Application of deep learning in neuroimaging-based Alzheimer’s diagnosis," Comp. and Math. Methods in Med., vol. 2020, 2020.

[16] S. Jain and B. Singh, "Alzheimer’s disease detection using machine learning," International Journal of Computer Applications, vol. 177, no. 7, 2020.

[17] G. Litjens et al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017.

[18] S. Sarraf et al., "Classification of Alzheimer's disease using fMRI data and deep learning CNNs," in Proc. IEEE SMC, 2016.

[19] A. Krizhevsky et al., "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, 2017.

[20] C. Szegedy et al., "Going deeper with convolutions," in Proc. IEEE CVPR, 2015.

[21] Y. LeCun et al., "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[22] I. Goodfellow et al., Deep Learning. MIT Press, 2016.

[23] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," arXiv:1603.04467, 2016.

[24] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in Proc. NeurIPS, 2019.

[25] H. Zhang et al., "ResNeSt: Split-attention networks," arXiv preprint arXiv:2004.08955, 2020.

[26] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. ACM SIGKDD, 2016.

[27] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint arXiv:1609.04747, 2016.

[28] Y. Bengio et al., "Learning long-term dependencies with gradient descent is difficult," IEEE Trans. Neural Netw., vol. 5, no. 2, pp. 157–166, 1994.

[29] J. Brownlee, Machine Learning Mastery With Python. Machine Learning Mastery, 2016.

[30] Alzheimer's Disease Neuroimaging Initiative (ADNI), https://adni.loni.usc.edu/, accessed Apr. 2025.