ADiag: Graph Neural Network Based Diagnosis of Alzheimer's Disease

Project ID: TMED004

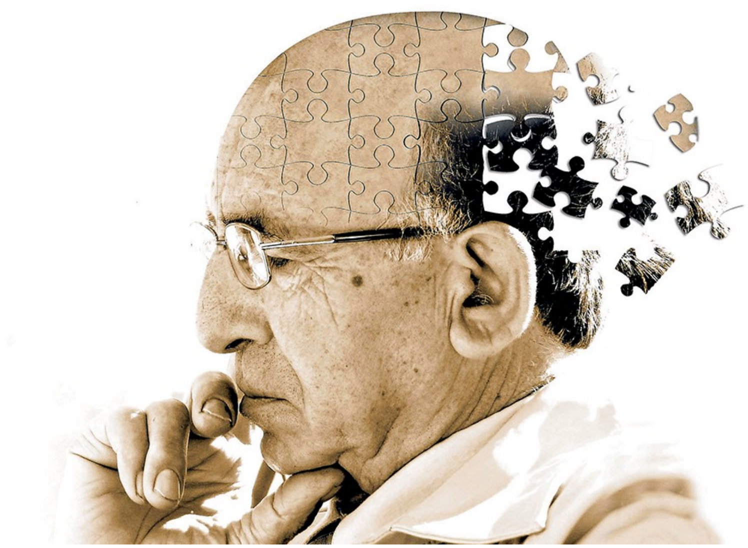
Vishnu Ram Sampathkumar

Class 12

National Public School, Indiranagar

Bengaluru, Karnataka

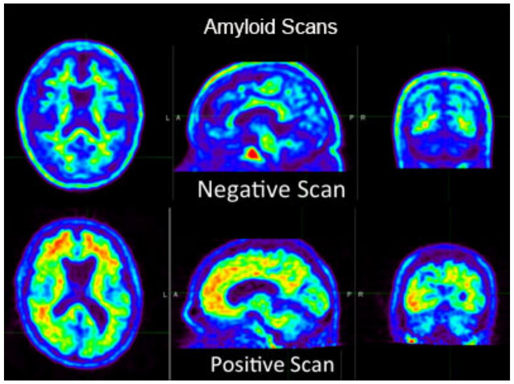
INDIA



Introduction - Problem Statement

1. **Alzheimer's Disease (AD)**
   - AD is a progressive **neurodegenerative disease** and the most common form of **dementia -** affects more than **50 million people** worldwide[1]
   - **No cures for AD, only treatment of symptoms**
   - In **advanced stages**, complications from severe loss of brain function — such as malnutrition, dehydration or infection— **result in death**
   - Global **economic burden: US$ 800B+** spent on medical + social care for AD and related dementia[1]
2. **Mild Cognitive Impairment (MCI)**
   - Measurable decline in cognitive abilities beyond the expected decline of normal aging
   - Person with MCI is at an increased risk of developing AD or any other form of dementia
   - Sometimes, MCI reverts to normal cognition or remains stable

- **Current QUALITATIVE** clinical diagnosis of AD through **MMSE and CDR Tests**
  - Highly variable as it **depends on clinician's competence**
  - **>25% chance of misdiagnosis**
- **Current QUANTITATIVE Diagnosis - PET Imaging**
  - Low specificity
  - Prohibitively expensive
- **Impact of high misdiagnosis rates**
  - Quality of life affected as symptom relief therapy not given
  - Promising clinical trials showing discouraging results as patients wrongly classified as AD



Amyloid PET Imaging
Source: UCSF Medicine

[1] Anders Wimo, Maëlenn Guerchet, Gemma-Claire Ali, Yu-Tzu Wu, A. Matthew Prina, Bengt Winblad, Linus Jönsson, Zhaorui Liu, and Martin Prince, "**The worldwide costs of dementia 2015 and comparisons with 2010,**" https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5232417/

Introduction - My Goal

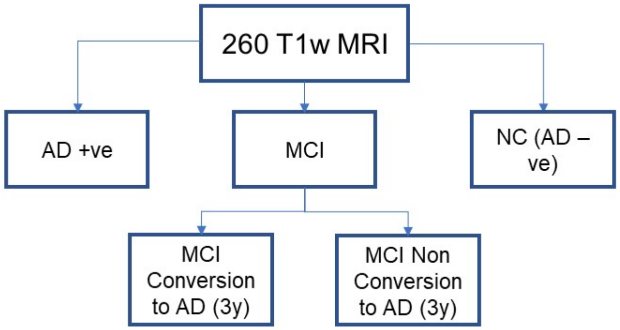
- Provide a **novel tool** - **ADiag** - to **help clinicians** so they can **quickly, quantitatively and accurately** diagnose AD and MCI with early signs of cortical atrophy patterns
- Use **Graph Theory** and **Deep Learning Architecture** to build this diagnostic model
- Achieve **accuracies > 80%** in classifying brain images quantitatively as:
  - AD Positive
  - AD Negative
  - MCI Conversion to AD in 3 years (MCIc)
  - MCI Non-Conversion to AD in 3 years (MCInc)
- **Use cortical thickness as imaging biomarker**
  - An excellent biomarker for diagnosis with a high specificity as AD/MCI pathology shows distinct regional pattern of cortical atrophy

Methodology: Data Acquisition and Pre-Processing

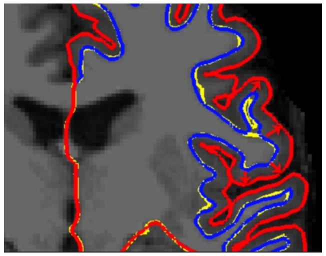
- **MR Image Acquisition**
  - **Dataset:** 75 NC (Controls), 68 MCIc, 45 MCInc and 72 AD T1w image scans sourced from Alzheimer's Disease Neuroimaging Initiative (ADNI) database after receiving prior permission through NIH grant application
  - Age group of subjects is 42 to 95 years
- **Conditions for selection**
  - To simulate biological realism, few subjects with conditions such as alcoholism and depression were selected

Group	MMSE	CDR
MCI	20-26	0.5
AD	<24	>0.5
NC	24-30	0

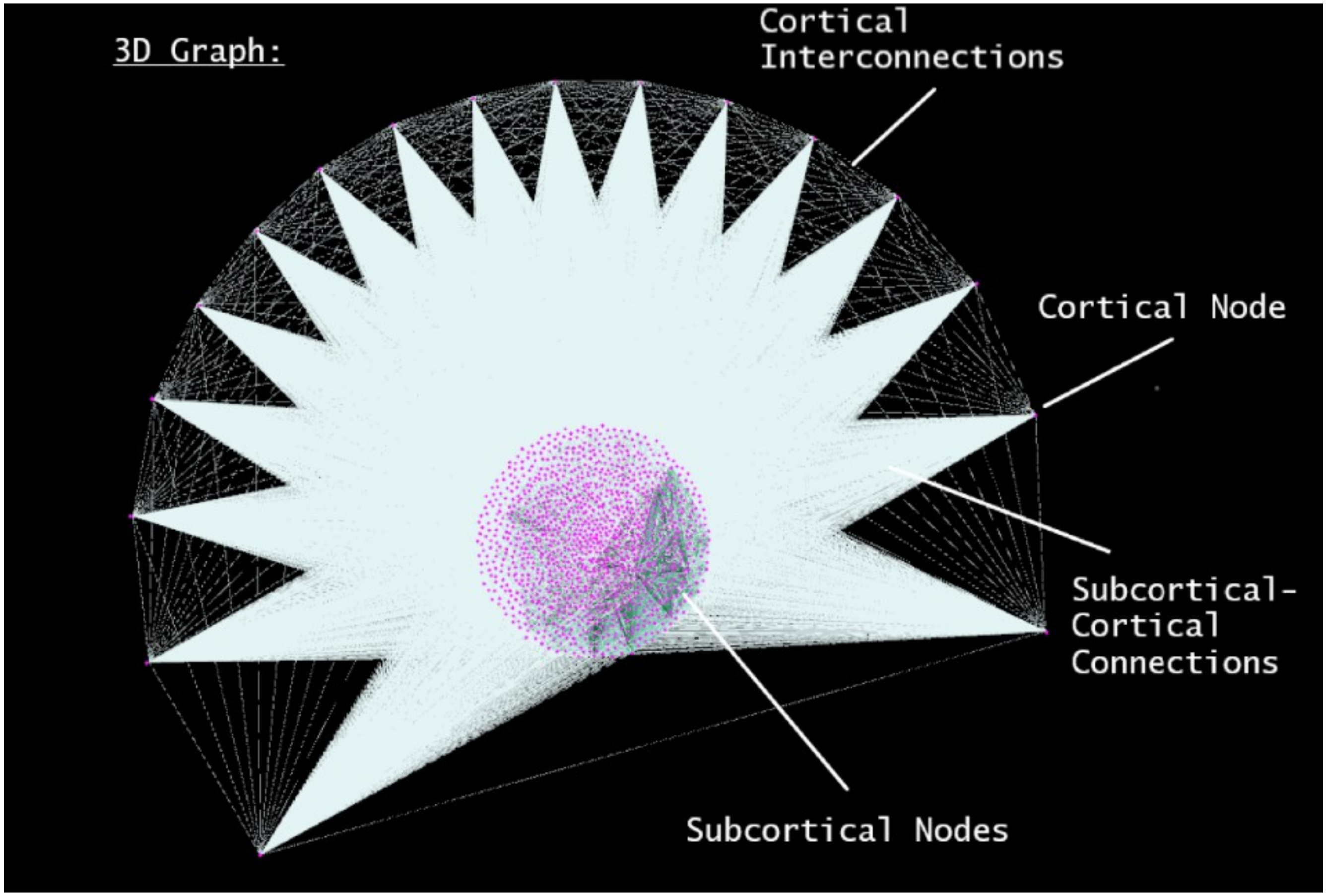
  - Three year time period considered to verify whether MCI converts to AD
- **Thickness features extracted from graphs via FreeSurfer software**
- **Graynet software used to model thickness features into Graphs** (series of nodes and edges)
  - Edge weights based on thickness differences between connected nodes
  - Each scan yielded 1162 nodes and 674,541 edges



Dataset Composition


FreeSurfer thickness extraction

Methodology: Overview of 3D Graph

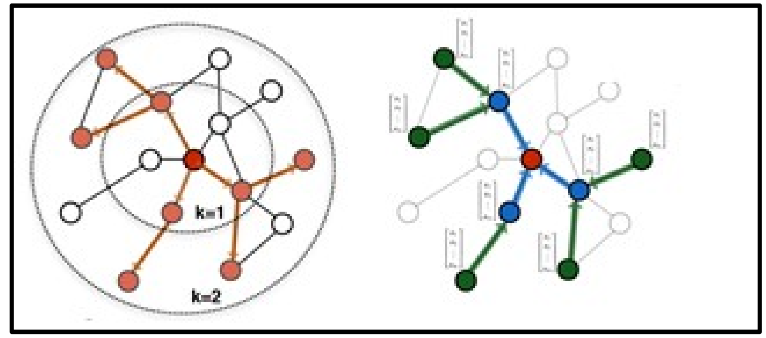


Methodology: Deploying Graph Neural Network

GraphSAGE layer

$$\mathbf{h}^k_{\mathcal{N}(v)} \leftarrow \text{AGGREGATE}_k(\{\mathbf{h}^{k-1}_u, \forall u \in \mathcal{N}(v)\})$$
$$\mathbf{h}^k_v \leftarrow \sigma\left(\mathbf{W}^k \cdot \text{CONCAT}(\mathbf{h}^{k-1}_v, \mathbf{h}^k_{\mathcal{N}(v)})\right)$$

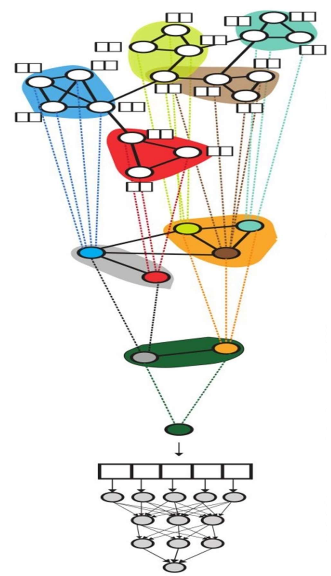
- Responsible for **aggregating information** from all nodes in a graph's neighbourhood
- Each node **is simultaneously enriched** with information from neighbourhood
- Extremely **relevant for brain graphs** → neighbourhoods in a graph can be compared to lobes of a brain



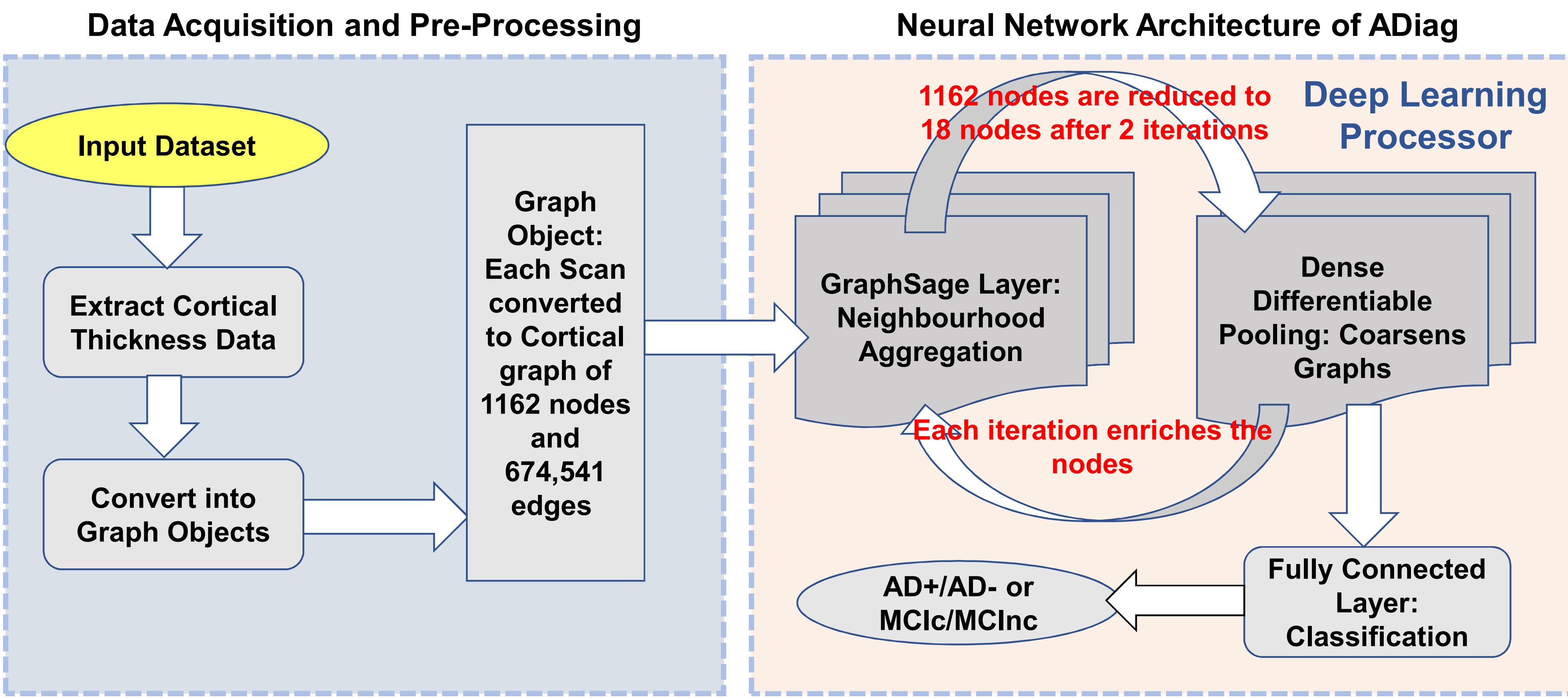
GraphSAGE: Neighbourhood Aggregation

Dense Differentiable Pooling

- Responsible for **coarsening**/reducing size of graph
- Generates **assignment tensor** which decides how many nodes to cluster together based on GraphSAGE output
- Extremely relevant for whole graph classification as opposed to node classification

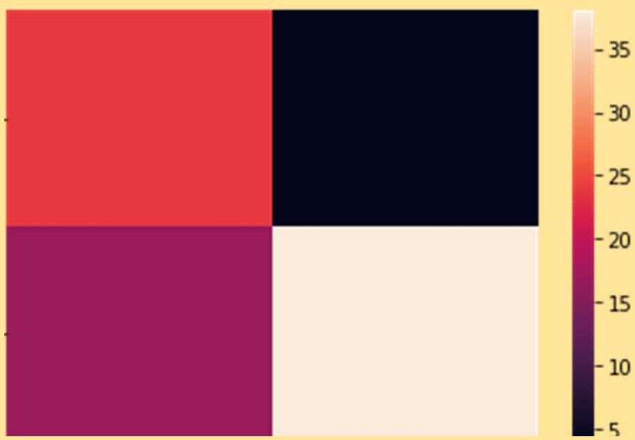

DDP: Reduces Node Number

ADiag Design and Methodology

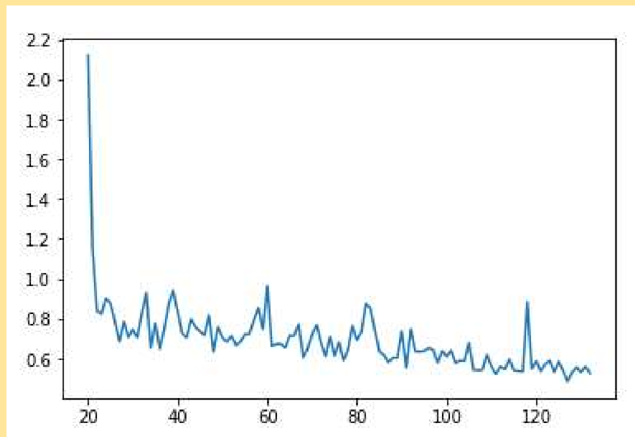


Data Analysis and Results – AD v. NC

Confusion Matrix

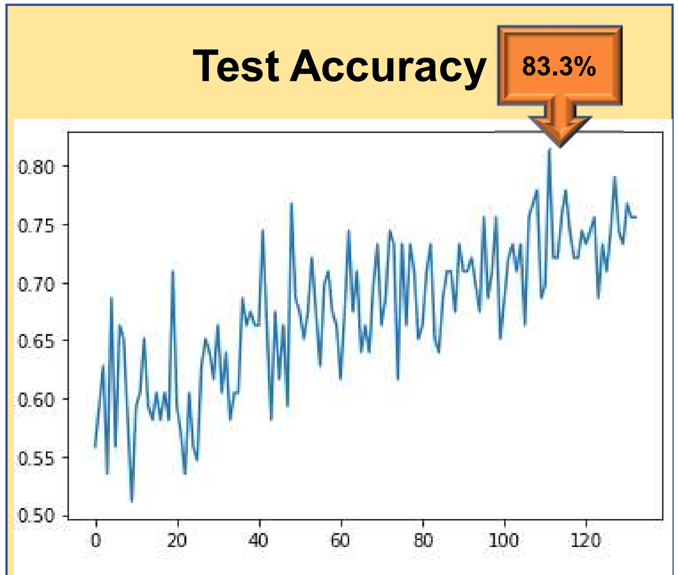


Training Loss v. # of Epochs



# of Epochs

Test Accuracy

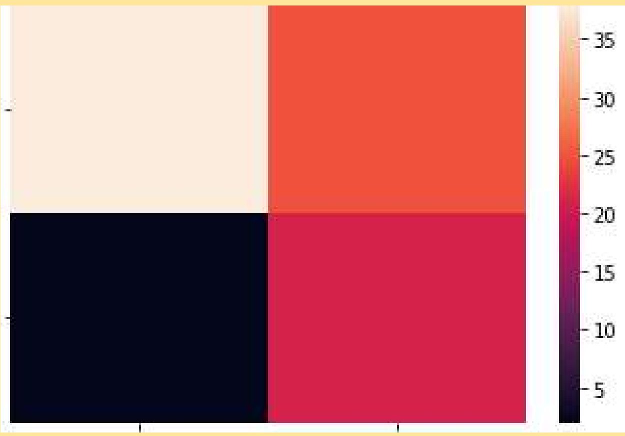


# of Epochs

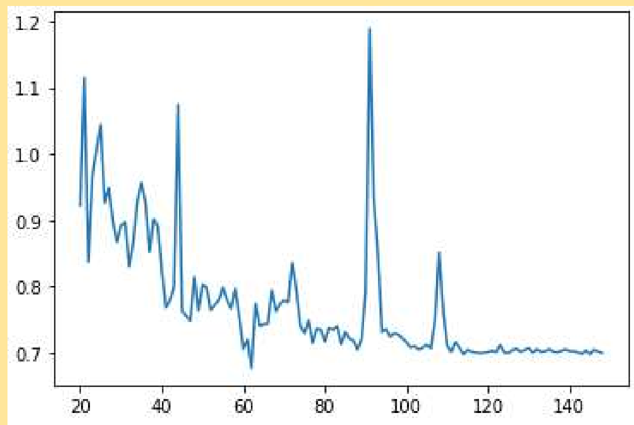
- **Accuracy is 83.3%**
- **Training optimized with Learning Rate Optimization, K-Fold Cross Validation**
- **Specificity: 85.7%; Sensitivity: 70.4%**

Data Analysis and Results – MCIc v. MCInc

Confusion Matrix

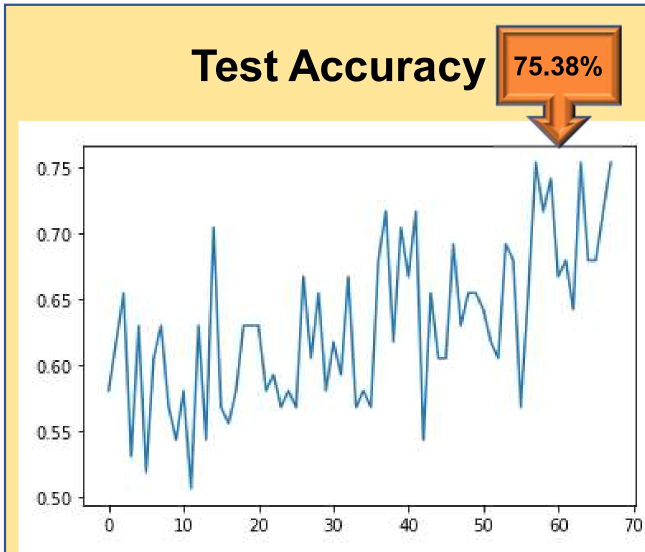


Training Loss v. # of Epochs



# of Epochs

Test Accuracy



# of Epochs

- **Accuracy is 75.38%**
- **Training optimized with Learning Rate Optimization, K-Fold Cross Validation**
- **Specificity: 80.2%; Sensitivity: 68.6%**

Discussion

Category	ADiag: GNN	CDR and MMSE	ThickNet Graph Learning	PET Imaging
Accuracy (AD v. NC)	83.3% from 147 samples **** Estimated 95 % accuracy with 250 scans**	Less than 75%	89% from 509 samples	N/A
Accuracy (MCIc v. MCInc)	75.38%	N/A	64.5%	N/A
Feasibility	High: based on widespread T1w MRI	High: based on written/oral exam	High: based on widespread T1w MRI	Moderate: dependant on sparse PET scan
Expense	Low: ~ $700	Extremely Low	Low: ~ $700	High: ~ $6000
Effect of Data	Accuracy Scales with Data	N/A	Accuracy does not scale with data	N/A

- I attempted to use gene expression as a secondary variable along with cortical thickness. Using a PCA map, I found a low correlation of gene expression values with the patients' condition and hence, abandoned it as a secondary variable.
- Initially the dataset was from the Open Access Series of Imaging Studies (OASIS-3) database which had AD and NC scans only. When I expanded the project scope to include MCI patients' data, I had to source the data from the ADNI database.

Conclusions

- **ADiag is a novel, quantitative, low-cost diagnostic tool that diagnoses AD and MCI**
- **Clinicians can use ADiag to diagnose AD with higher certainty than qualitative diagnosis (83% v 75% or less)**
- **Clinicians can predict with 75.4% accuracy whether the MCI patient can progress to AD in three years**
  - I have achieved my goal of creating such a model with an accuracy of 83.3% and 75.38% for AD v. NC and MCIc v. MCInc, respectively
  - I have also proved my hypothesis that cortical thickness is a powerful biomarker to diagnose Alzheimer's Disease
  - ADiag is one of two GNN-based models for AD and MCI diagnostics: paper at **https://arxiv.org/pdf/2101.02870.pdf**
- **Future goals and objectives**
  - Doubling dataset size: this will increase accuracy to approximately 95%
  - Validation of ADiag model at National Institute of Mental Health and Neuro Sciences (NIMHANS), Bengaluru; incorporate Indian dataset
  - Include PET data to access uptake features

References

[1] E. Nichols, C. E. Szoeke, S. E. Vollset, N. Abbasi, F. Abd-Allah, J. Abdela, M. T. E. Aichour, R. O. Akinyemi, F. Alahdab, S. W. Asgedom *et al.*, "Global, regional, and national burden of alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 1, pp. 88–106, 2019.

[2] S. Balsis, J. F. Benge, D. A. Lowe, L. Geraci, and R. S. Doody, "How do scores on the adas-cog, mmse, and cdr-sob correspond?" *The Clinical Neuropsychologist*, vol. 29, no. 7, pp. 1002–1009, 2015.

[3] J. Warren, J. Schott, N. Fox, M. Thom, T. Revesz, J. Holton, F. Scaravilli, D. Thomas, G. Plant, P. Rudge *et al.*, "Brain biopsy in dementia," *Brain*, vol. 128, no. 9, pp. 2016–2025, 2005.

[4] P. J. LaMontagne, T. L. Benzinger, J. C. Morris, S. Keefe, R. Hornbeck, C. Xiong, E. Grant, J. Hassenstab, K. Moulder, A. Vlassenko *et al.*, "Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease," *medRxiv*, 2019.

[5] O. Querbes, F. Aubry, J. Pariente, J.-A. Lotterie, J.-F. Démonet, V. Duret, M. Puel, I. Berry, J.-C. Fort, P. Celsis *et al.*, "Early diagnosis of alzheimer's disease using cortical thickness: impact of cognitive reserve," *Brain*, vol. 132, no. 8, pp. 2036–2047, 2009.

[6] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[7] P. R. Raamana and S. C. Strother, "Graynet: single-subject morphometric networks for neuroscience connectivity applications," *Journal of Open Source Software*, vol. 3, no. 30, p. 924, 2018.

[8] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in neural information processing systems*, 2019, pp. 8026–8037.

[9] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Advances in neural information processing systems*, 2017, pp. 1024–1034.

[10] Z. Ying, J. You, C. Morris, X. Ren, W. Hamilton, and J. Leskovec, "Hierarchical graph representation learning with differentiable pooling," in *Advances in neural information processing systems*, 2018, pp. 4800–4810.