

# A Big-Data Approach to Investigating the Mechanisms of Dementia-Related Neurodegenerative Disorders

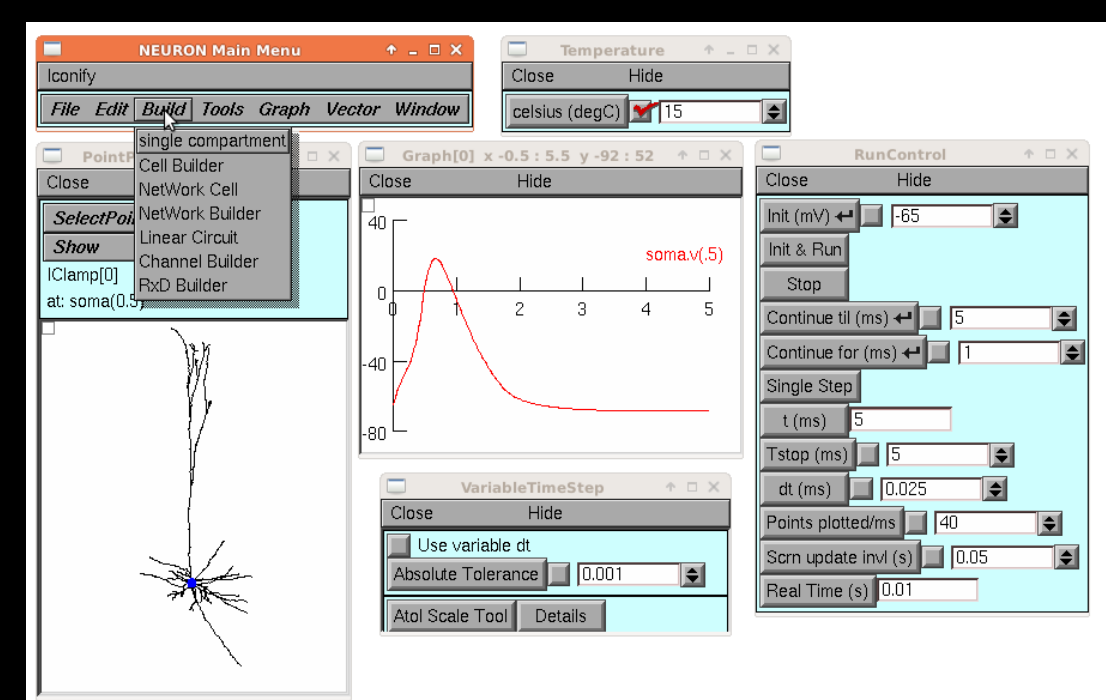
Christopher J. Manning, Andra G. Armstrong Ph.D.

## Objective

Design a research poster that showcases your main 1-2 results and is:

- ☐ To use machine learning techniques to investigate Alzheimer's data
- ☐ Understand the values of specific dementia associated factors/predictors

## Background



- Use of online databases have made it easier for collaboration in research efforts
- Machine learning is becoming a larger part of data exploration, the neuroscience field being one of these places
- Coded neural simulators such as the ones from the Thousand Brain Project (Monty) or the NEURON simulation by Yale and Duke Universities also allow for the production of simulated data when human data is unreliable

## Methods

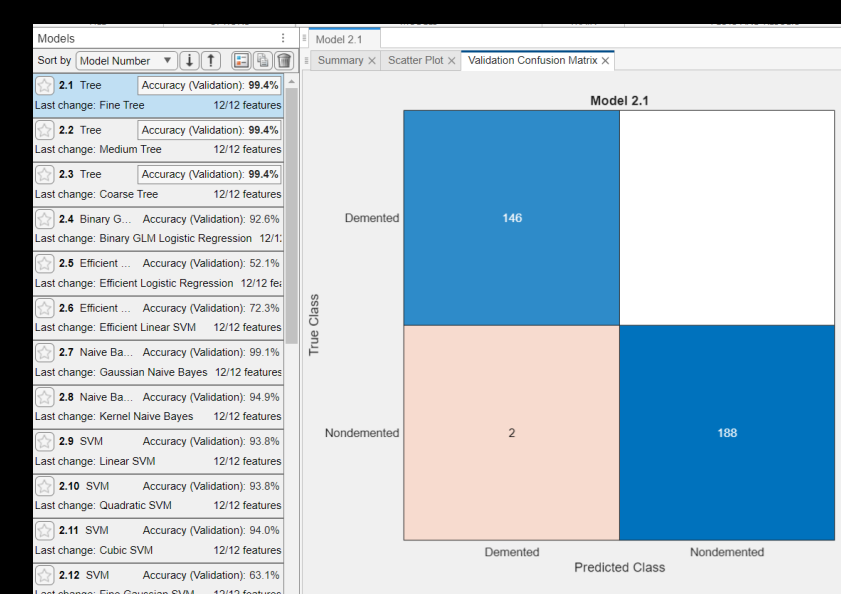
This research project used one dataset containing various dementia and participant demographic measures

### Classification:

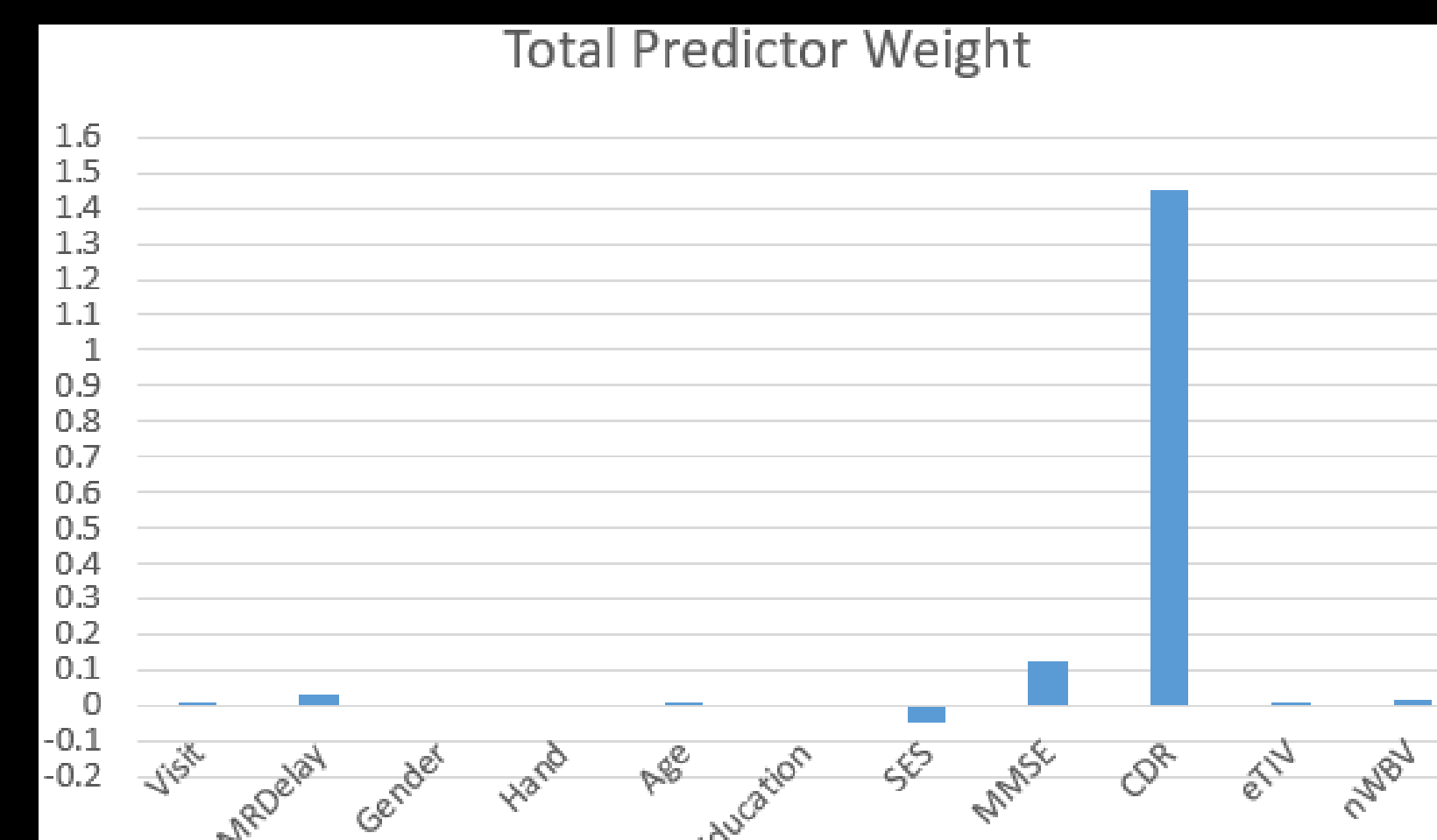
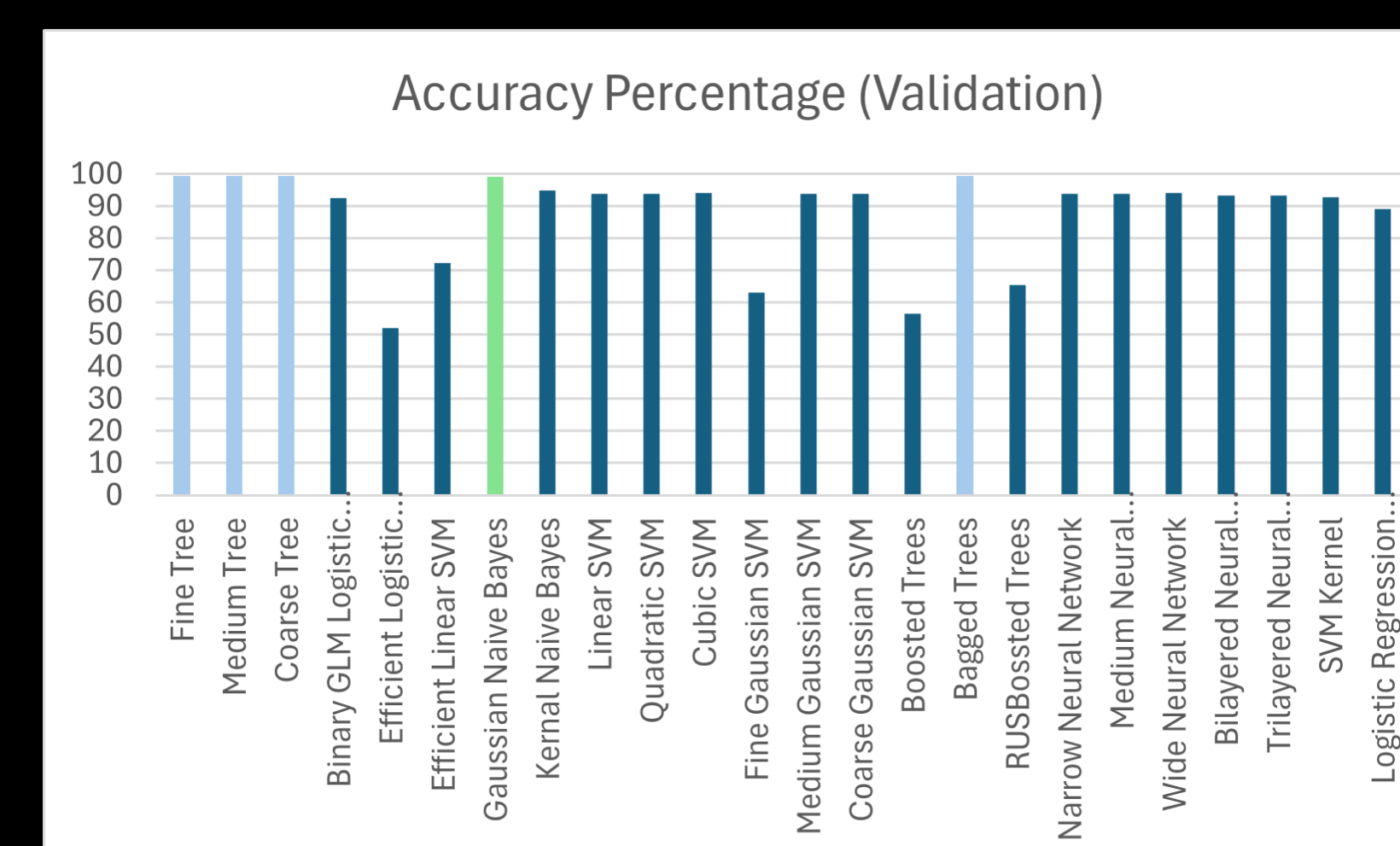
- Using models to separate data points into distinguished predicted classes
- Different modes of classification work different:
  - Decision Trees:** Data is split into subsets based on selected variables, refining the variables in too their correct category
  - Gaussian Naive Bayes:** Classification based on the Naive Bayes Theorem where continuous attributes follow a Gaussian Distribution

### Regression:

- Finding model or function that distinguishes the data into real values, not classes
- The prediction is a continuous target variable that is found through independent features of the data

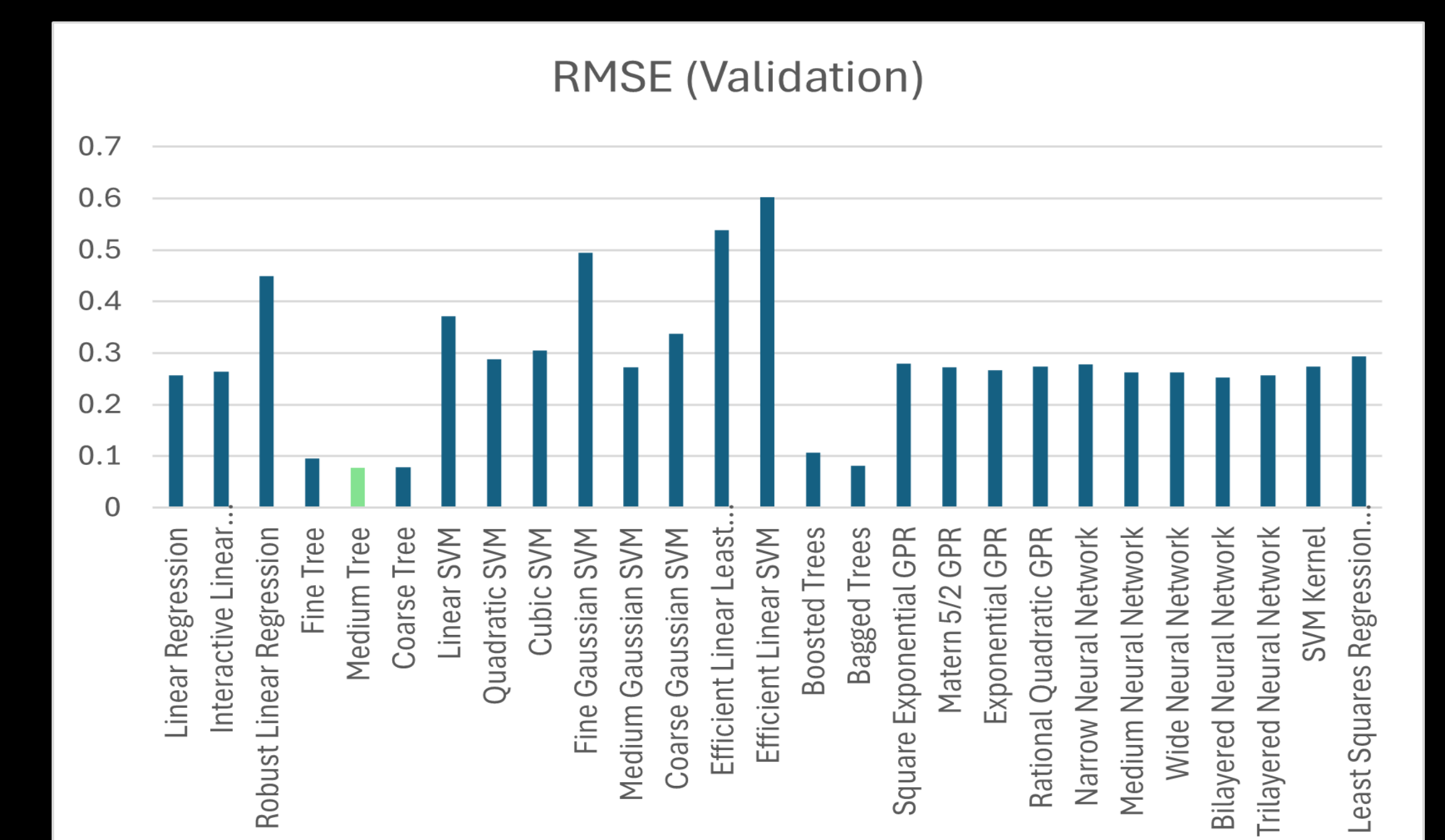


## Results



- This graph plots the accuracies of the different classifiers used to evaluate the data

- This graphs show the weight each predictor has on the final classification
  - The main three are CDR, MMSE, and SES



- This graph is of the validation scores of the Regression models
  - The RMSE gives the average distance between the predicted values given by the model and the actual values from the data

## Implications / Future Directions / Challenges

### Implications:

- Gives further evidence backing a new and more efficient way for Medical Practitioners to diagnose Alzheimer's Disease
- Allows for further understanding of Alzheimer's as a disease

### Future Directions:

- Due to some of the outlined challenges with acquiring data plans to work on a model to simulate Alzheimer's (specifically synaptic degeneration) and its dementia symptoms in being worked on using one of the mention neural networks

### Challenges:

- Some of the data sets looks at had to much missing data to use
- Accessing data is a challenge as a student research without the credentials or the amount of funding it takes to get some of the data sets