Detecting Preclinical Cognitive Change

The Clock Drawing Test for Early Detection of Alzheimer’s

Kelly “Scott” Sims

Table of Contents

The Problem

In the year 2015 in the USA, there were 5.3 million cases of Alzheimer's that contributed to an exponentially inflated medical cost of more than $200 Billion a year. By 2050, the number of cases is expected to triple. And these statistics, that are only attributed to cases in the USA, are expected to inflate medical costs to over $1 Trillion.

The issue with early detection of Alzheimer's is that the majority of techniques are either outrageously expensive, extremely invasive, or both. These methods include P.E.T scans, Synaptic Dysfunction tests, Spinal Taps, and MRI's. Another test that is neither expense nor invasive is a simple cognition test. The following illustration from NYU Langone Medical Center highlights the issue with current cognitive screenings however.

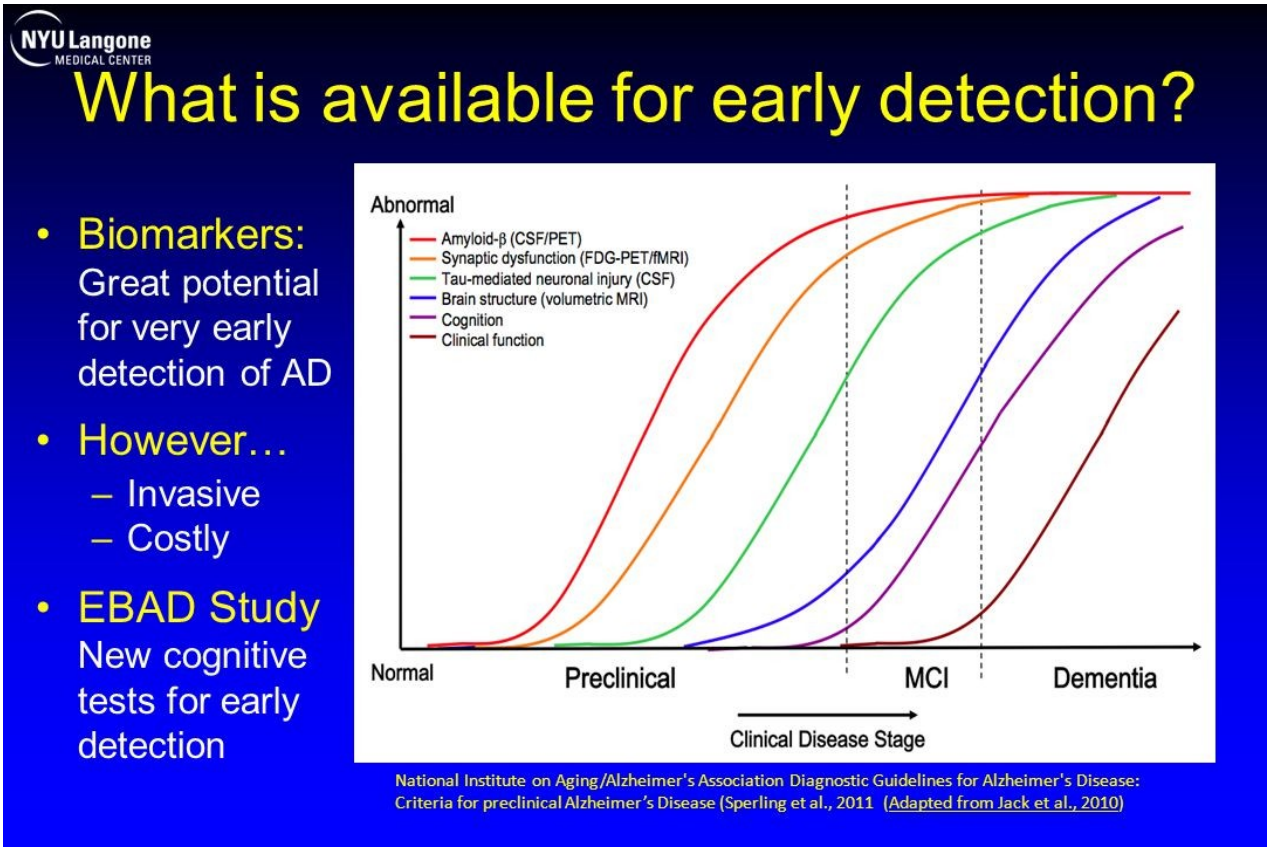

Illustration 1: Ferris, Steven Ph. D. "Diagnostic Guidelines for Alzheimer's Disease". Sperling et al., 2011

The x-axis is increasing severity of the disease. The y-axis is brain function ranging from Normal to Abnormal. Each test's slope shows how likely it is to detect signs of dementia. The first 4 are extremely painful, invasive and expensive. Whereas stated earlier, Cognition (purple), is neither. The problem with it however, is that it can't detect any signs of dementia until it is pretty much too late.

Cognitive Screenings

The cognition tests are actually very simple in nature. First, a patient is given a blank sheet of paper and asked to draw a clock showing 10 minutes after 11. This is called a “command clock”. Next, they’re given a sheet of paper with a clock already drawn on it and they’re asked to copy the clock. This is called the “copy clock”. These two simple tests challenge both spatial perception (copying the clock) and normal cognitive processes such as understanding the language, imagining a clock and drawing one (command clock). The great thing about these tests is that they’re fast, easy, non-invasive, and cheap. The downside is that they’re slow to score, where scoring is subjective. And they don’t sufficiently detect a deficiency until a patient is severely impaired. So how can we improve this test to detect early signs of dementia?

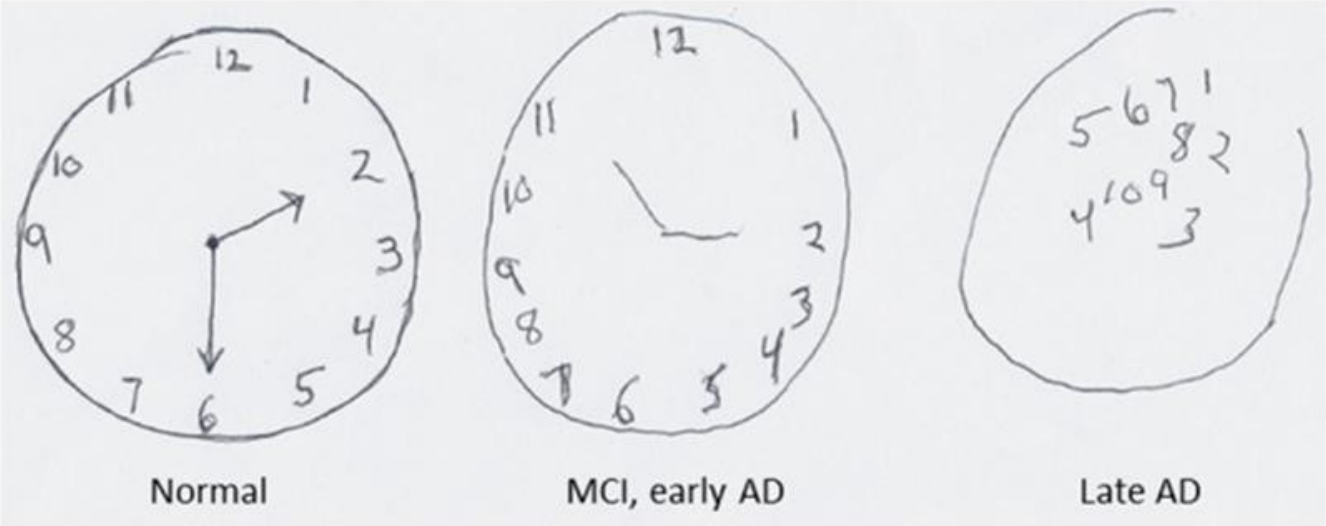

Illustration 2: Image: Mark P. Mattson / Frontiers in Neuroscience

The Data

Since 2005 there have been 5300 clock tests performed using a special method. A labeled data set of 4200 patients, both healthy and impaired, has been curated from these special tests. All data is in the form of hand drawn clocks. The clocks were drawn using a ball point pen with a camera embedded in it. The paper had specially designed microdots on it so the pen could always track its movements with respect to the paper. It measured location 75 times a second and is extremely precise. This is in direct contrast to how the test was originally performed. Using plain paper and pen, the only data collected during the session was the final drawing with no insight or data of what transpired while the patient was drawing the clocks.

Initial Thoughts

From the method of data collection, I can brainstorm the following initial quantitative variables.

- Time to draw a clock showing 10 minutes past 11
- Time to draw the boundary of the clock
- Time to draw each number
- Amount of pause between drawing the next part of the clock
- Location of each number with respect to each other (change in angle)
- Order in which each part of the clock was drawn
- How close of resemblance the drawn clock is to the provided clock

These can be potential features or foundations for new features via feature engineering. Some look more predictive than others So let's engineer some features and implement some models

Feature Engineering and Modeling

Distribution Modeling

From both drawn images, an analysis can be performed between how well a patient drew a clock from memory (the command clock) vs simply copying the provided clock. From the center of the clock, symmetry can quantified by measuring angle offsets of each number. In perfect symmetry, the angle between 2 numbers should be about 1/12th of 360 degrees. From this exercise of symmetry, 4 different distributions can be compiled

1. distribution of angles from non impaired patients on the "command clock"
2. distribution of angles from non impaired patients on the "copy clock"
3. distribution of angles from impaired patients on the "command clock"
4. distribution of angles from impaired patients on the "copy clock"

Arguably, the distributions from the impaired patients would be more skewed. Therefore, the key statistics for modeling would be the median and standard deviation of each distribution. From these distributions, we can create 4 new features – the Z-score for each patient from each distribution. Z-score is simply how many standard deviations a value is from the center of the distribution. Being closer to the center of the distribution means a higher probability of being a significant value of the

group being model by the distribution. By being closer to the center, a patient will have a lower Z-score.

Features	Description
Z-score Command Clock Non Impaired	Patient's Z-score for median angle between each number in a clock with respect to the Non Impaired Command Clock Distribution
Z-score Copy Clock Non Impaired	Patient's Z-score for median angle between each number in a clock with respect to the Non Impaired Copy Clock Distribution
Z-score Command Clock Impaired Distribution	Patient's Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution
Z-score Copy Clock Impaired Distribution	Patient's Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution

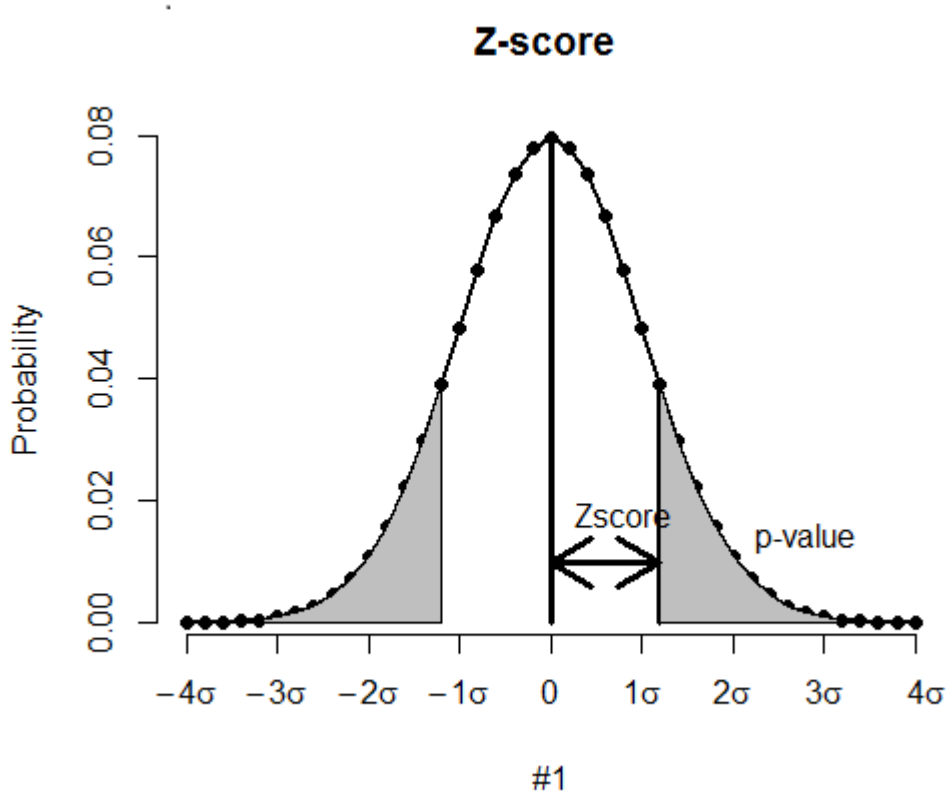

Illustration 3: Relation of Z-Score to Probability

Outlier Detection as a Feature

The analysis of the final images is good, but not great. Remember with these cognitive tests that they're not very sensitive to potential indicators. They're only good classifiers after heavyset dementia has started to kick in. Because of this, we need to draw insights not from just the final images, but the processes that take place while the images are being drawn.

Since every stroke and movement is monitored with a camera that measures position with respect to time, we can quantify the "pause" from one motion to the next. E.g. a patient has drawn the diameter of the clock and the first number of the face, but then significantly pauses for some duration before drawing the next piece of the clock. This could be a potential indicator of a mental lapse. We measure the potential lapse as so:

- The patient begins drawing the first piece of the clock
- The moment the pen either stops or is lifted up, that is time T=0
- The moment the pen is pressed down and begins drawing the next piece of the clock, that is time T = deltaT
- Continue measure deltaT from one piece of the clock being drawn to the next.

Presumably for a non impaired patient, deltaT would be nearly constant during the whole drawing process (the patient draws 12 on the clock, then it takes 3 milliseconds for the patient to start drawing a 1, then it takes 3 milliseconds for the patient to start drawing a 2 after drawing the 1, etc. etc.). For an impaired person, there would probably be one or multiple "outliers" in deltaT during their session. We then count the number of outliers for each patient. This is a new feature.

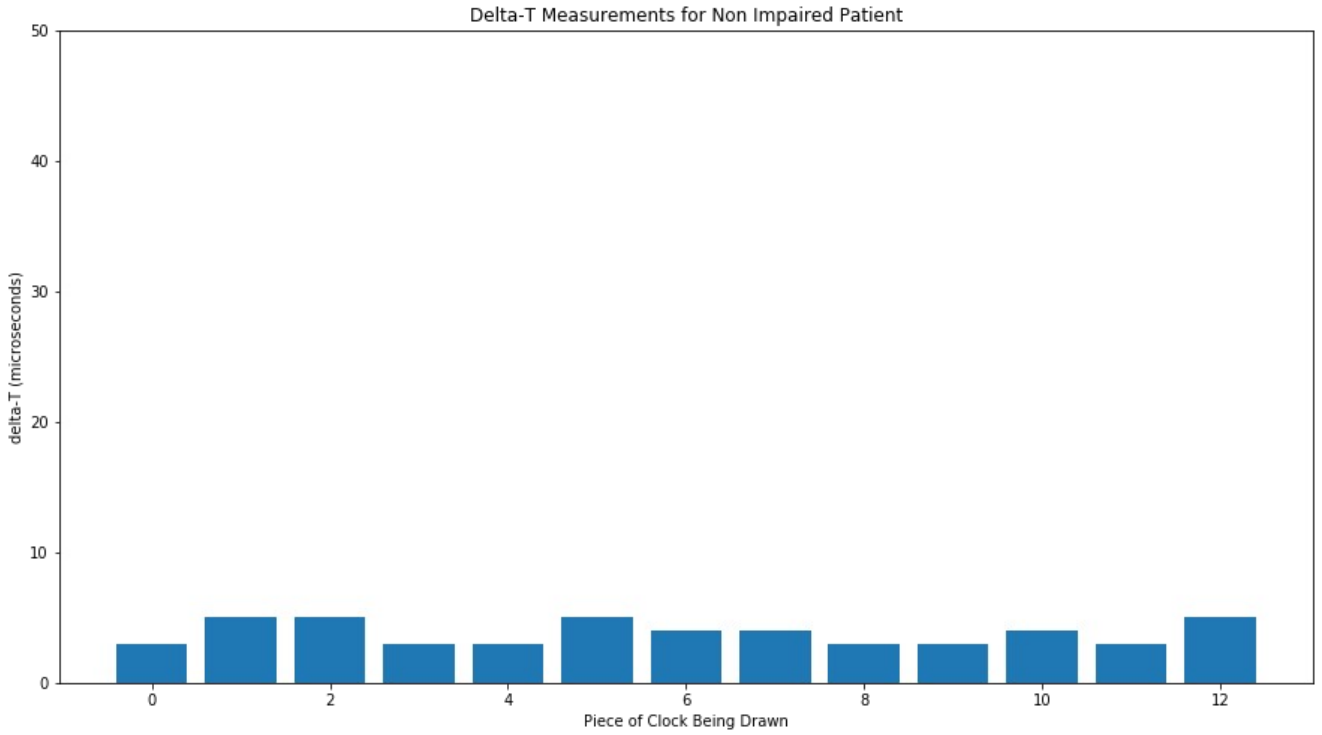

Illustration 4: Theoretical Delta-T for a Non Impaired Patient. We would expect no outliers

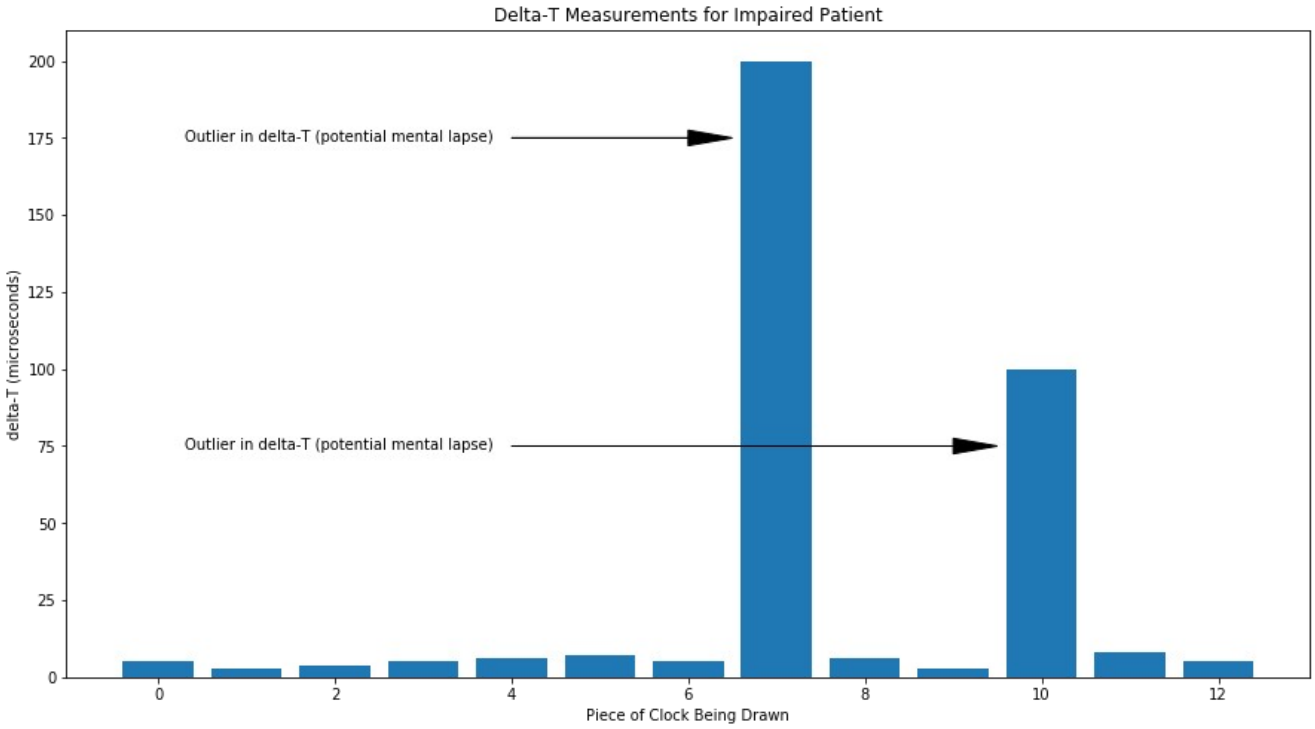

Illustration 5: Theoretical Delta-T for an Impaired Patient. Outliers detected which could be potential indicators of mental lapses

Features	Description
# of Delta-T outliers	Number of outliers present in the Delta-T distribution
Z-score Command Clock Non Impaired	Patient's Z-score for median angle between each number in a clock with respect to the Non Impaired Command Clock Distribution
Z-score Copy Clock Non Impaired	Patient's Z-score for median angle between each number in a clock with respect to the Non Impaired Copy Clock Distribution
Z-score Command Clock Impaired Distribution	Patient's Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution
Z-score Copy Clock Impaired Distribution	Patient's Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution

Modeling Process Order for Abnormality

The order in which a normal person completes a task is usually systematic and structured. One might argue that someone with no abnormalities would either complete the clock by drawing the circle, then

filling in the numbers in order (e.g. 1,2,3,4,5….12), or they would draw the perimeter numbers first, (12, 6, 3, and 9) then fill in the numbers between, in order to facilitate symmetry. But because we wouldn't want to introduce bias in our thinking, process order is something we could model in order to classify a "normal" vs "abnormal" process order. The output of this model would be another feature in the ultimate model.

Let's make the standardization rule of 14 processes to draw the clock. Drawing the circle is 1 process, then drawing each of the 12 numbers would be process numbers 2 – 13, starting with drawing 12 and filling in the numbers clockwise, followed by drawing the hands as process number 14. We could then train a Bayesian Model to give the probability of process ordering to be normal or abnormal.

The feature inputs would be the order in which a patient completed the clock. E.g. a patient that draws the circle first, then the perimeter numbers followed by the filled in numbers, and finally the hands would be encoded for input into Bayesian model as {1, 2, 5, 8, 11, 3, 4, 6, 7, 9, 10, 12, 13, 14}. Since Bayesian statistics operates on joint probabilities, the presence of the same numbers, regardless of order would net the same output from the model. Because of this, each input will be multiplied by a weight. These weights will be decreasing with process ordering. Put in other terms, processes completed first will be multiplied by a higher weight. The output of the Bayesian model wouldn't be 0 (normal) or 1 (abnormal), but it would be probabilities of being 1. The lower the probability, the stronger the chance of the process being normal.

Features	Description
Process Ordering Probability	Output from a Bayesian Model that gives the probability of process ordering be abnormal.
# of Delta-T outliers	Number of outliers present in the Delta-T distribution
Z-score Command Clock Non Impaired	Patient's Z-score for median angle between each number in a clock with respect to the Non Impaired Command Clock Distribution
Z-score Copy Clock Non Impaired	Patient's Z-score for median angle between each number in a clock with respect to the Non Impaired Copy Clock Distribution
Z-score Command Clock Impaired Distribution	Patient's Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution
Z-score Copy Clock Impaired Distribution	Patient's Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution

Descriptive Statistics as Features

For final features to aggregate, we can just add the initial variables I outlined in the first section as descriptive features

Features	Description
Accuracy of “Copy Clock” to Original Clock	Quality score of the “copy” clock on a scale of 5 1 – Not Even Close 2 – Major Flaws 3 – Minor Flaws 4 – Nearly Perfect 5 - Perfect
Time to Draw all the Numbers	Self Explanatory
Time to Draw the Boundary of the Clock	Self Explanatory
Time to draw a clock showing 10 past 11	Self Explanatory
Process Ordering Probability	Output from a Bayesian Model that gives the probability of process ordering be abnormal.
# of Delta-T outliers	Number of outliers present in the Delta-T distribution
Z-score Command Clock Non Impaired	Patient’s Z-score for median angle between each number in a clock with respect to the Non Impaired Command Clock Distribution
Z-score Copy Clock Non Impaired	Patient’s Z-score for median angle between each number in a clock with respect to the Non Impaired Copy Clock Distribution
Z-score Command Clock Impaired Distribution	Patient’s Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution
Z-score Copy Clock Impaired Distribution	Patient’s Z-score for median angle between each number in a clock with respect to the Impaired Command Clock Distribution

FINAL MODEL

With our newly collected and generated features, we can incorporate a “Majority Voting” Ensemble method for early detection of dementia. This simply means that various types of classification models are used simultaneously. The output from each model is tallied, and the response that appears the most is the final prediction. The voters for this analysis are:

- Logistic Regression
- Random Forest
- Naive Bayes
- SVM
- k-Means Clustering

As an example output of the voters, considering the following

Logistic Regression = 1

Random Forest = 0

Naive Bayes = 1

SVM = 1

k-Means Cluster = 1

Since there are 4 positive classifications to 1 negative, the final prediction is positive for early onset detection of dementia.

Conclusion

The great thing about this whole process is that with every patient that takes the test, the distributions get updated and more representative of the population. This in turn makes the model more predictive with each new patient.