Steps:

1. In adni\_dataset\_original.csv Replace bl->0 (1763 instances), m->space(9601 instances) under VISCODE column and saved it to adni\_dataset1\_replaceVISCODE.csv
2. Replace EMCI and LMCI by MCI, SMC by NL under DX attribute and saved it to adni\_dataset1.csv
3. Remove the rows with all the attribute values as “NA”. Dataset from 11364 rows down to 8036 rows and saved it to adni\_dataset2\_removeRowsWithNA.csv
4. Run PEMM algorithm in Rstudio to find missing values with adni\_dataset2\_removeRowsWithNA.csv as the input and store the findings in adni\_dataset3\_fromPEMM.csv
5. Copy DX values from adni\_dataset2\_removeRowsWithNA.csv to adni\_dataset3\_fromPEMM.csv and save it as adni\_dataset3\_fromPEMM\_addedDX.csv
6. Aggregate the data as per RID.(Run aggregate.R) and store it in fromPEMM\_after\_mean.csv and aggregate\_temp.csv
7. Delete the extra column of ID’s in the fromPEMM\_after\_mean.csv and aggregate\_temp.csv
8. Sort aggregate\_temp.csv in asc order by RID. Then make rows unique using the advanced filter option in excel.
9. Copy the DX in aggregate\_temp.csv to fromPEMM\_after\_mean.csv and save it as fromPEMM\_after\_mean\_addDX.csv (MCI = 862, AD = 351, NL = 523, Total = 1736)
10. Run forestFeatureSelection.py for selecting the necessary attributes from a number of them.
11. Run compareClassifier.py and choose your classifier.

MySQL

1. <http://localhost/phpmyadmin/>
2. Db: alzDetect
3. Tables: patientReadings………contains noncognitivemarkers