Spring 6-1-2021

Advancing The Ability To Predict Cognitive Decline and Alzheimer's Disease Based On Genetic Variants Beyond Amyloid-beta and Tau

Naveen Rawat
San Jose State University

Advancing The Ability To Predict Cognitive Decline and Alzheimer's Disease Based On Genetic Variants Beyond Amyloid-β and Tau

A Project

Presented to

The Faculty of the Department of Computer Science

San José State University

In Partial Fulfilment

Of the Requirements for the Degree

Master of Science

By

Naveen Rawat

May 2021

Naveen Rawat

has passed the defense for the project

Advancing Ability to Predict Cognitive Decline and Alzheimer's Disease Based on Genetic Variants Beyond Amyloid-b and Tau

DocuSigned by:
Leonard Wesley
D9F65983BBCE4B1...

5/25/2021

Advisor's Signature | Date

DocuSigned by:
Yulia Newton
65A3150EE8C7460...

5/25/2021

Committee Member's Signature | Date

DocuSigned by:
William Andreopoulos
3CFA6E56A8E74FF...

5/25/2021

Committee Member's Signature | Date

NOTE: The advisor should send the final report to the graduate coordinator so that the student can be cleared for graduation

The Designated Project Committee Approves the Project Titled

Advancing The Ability To Predict Cognitive Decline and Alzheimer's Disease Based On Genetic Variants Beyond Amyloid-b and Tau

by

Naveen Rawat

APPROVED FOR THE DEPARTMENT OF COMPUTER SCIENCE

SAN JOSÉ STATE UNIVERSITY

May 2021

Dr. Leonard Wesley, Associate Professor, Department of Computer Science

Dr. Yulia Newton, Instructor, Department of Computer Science

Dr. William Andreopoulos, Assistant Professor, Department of Computer Science

ABSTRACT

Advancing The Ability To Predict Cognitive Decline and Alzheimer's Disease Based On Genetic Variants Beyond Amyloid-b and Tau

By Naveen Rawat

A growing amount of neurodegenerative R&D is focused on identifying genomic-based explanations of AD that are beyond Amyloid-b and Tau. The proposed effort involves identifying some of the genomic variations, such as single nucleotide polymorphisms (SNPs), allele , chromosome, epigenetic contributors to MCI and AD that are beyond Aβ and Tau.

The project involves building a prediction model based on a support vector machine (SVM) classifier that takes into account the genomic variations and epigenetic factors to predict the early stage of mild cognitive impairment (MCI) and Alzheimer disease (AD). To achieve this, picking up important feature sets which will be input to the machine learning model were identified using statistical model tests. The data used in this research were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database / ADNI GO2 GWAS.

Future work may involve increase in sample size analyzed from ADNI DB, explore and analyze potential secondary effects/medical-conditions such as other diseases that might have influenced the observed results and separate out MCI from AD and further explore predictions and results.

Key Words: *SNP Name, Allele1 - Plus, Allele2 - Plus, Chromosome, SNP*

Acknowledgment

Here the proposal is to advance the ability to predict cognitive decline and Alzheimer's disease based on genetic variants beyond Amyloid-b and Tau.

I sincerely thank my advisor Dr. Wesley for giving me the opportunity to work and research in this field. Also I am greatly thankful to my committee members Dr. Newton and Dr. Andreopoulos for giving valuable feedback and suggestions during the tenure of this research. These esteemed computer science department faculty helped to shape up my research work with the most efficient, productive, and timely manner. With the aid of well established engineering tools, they guided me to narrow down on identifying some of the genomic variations, such as single nucleotide polymorphisms (SNPs), allele , chromosome, epigenetic contributors to MCI and AD. These genomic variations further were utilized in building prediction models for early onset of cognitive impairment detection.

They mentored me throughout the project, on building the background with literature reviews, helping me to find the right approach and method for my experimentations, evaluating the experimentation results, which I got statistically as well as through my machine learning classifiers, and finally helped me to conclude on my project work.

As a concluding remark, showing my utmost gratitude towards my committee members for guiding me around this experience, from start to till the end.

Table of Content

List of Figures

INTRODUCTION

Alzheimer's Disease (AD) is a type of dementia that is a neurovegetative disorder which impacted approximately 47 million or 0.6% of the global population in 2015 (Rawtaer et al., 2020). The number of AD diagnoses is projected to triple by 2050 (Rawtaer et al., 2020). More than 6 million Americans of any age have Alzheimer's (Lawlor, B. A. et al. ,1994). An anticipated 6.2 million Americans age sixty-five and older are residing with Alzheimer's dementia by 2021. Seventy- percentage are age seventy-five or older. One in nine human beings age sixty-five and older (11.3%) in America has Alzheimer's dementia (Dianxu Ren str al., 2020).

The global costs of dementia in 2015 were estimated at $818 billion, a 35.4% increase compared with 2010. Up to 1 in 5 of the community-dwelling older adults aged 65 years and above suffer from mild cognitive impairment (MCI). Between 10–15% of patients with MCI may develop dementia each year (Davis & Allen, 2013). In addition, a recent meta-analysis indicated that about 45% of MCI patients maintained stable cognitive ability, whereas 28% progressed to AD and 15% returned to normal status without recurrence (Hu et al., 2017). These are pandemic numbers and costs.

AD is an irreversible, fast-spreading brain disease that slowly destroys memory, critical thinking capabilities and, eventually, the capacity to perform simple tasks that are needed to maintain independence. Most human beings with the disorder signs and symptoms first seem to be in their mid-60s (Lawlor, B. A. et al. ,1994). MCI causes a mild, however great and measurable, decline in cognitive abilities, together with memory and questioning skills. An individual with MCI is at an accelerated danger of growing Alzheimer's or any other dementia.[1]

In recent years, there has been significant efforts to develop biomarkers that can help with the early detection of AD and MCI [2]. These efforts can be roughly partitioned into approaches that are (1) non-neuroimaging; (2) neuroimaging; (3) brain volume-based; and (4) genomic-epigenetic-based in nature. Example non-neuroimaging work to identify MCI and AD biomarkers typically involve evaluating metrics such as cerebrospinal fluid (CSF), positron emission tomography (PET), and plasma, β-amyloid, total tau (T-tau), and phosphorylated tau (Ptau) (Gao et al., 2021; Van et al., 2020). Additional non-neuroimaging metrics include differentially methylated positions (DMPs) as novel blood-based biomarkers of AD (Vasanthakumar et al., 2020). Example neuroimaging work to identify MCI and AD biomarkers involve evaluating metrics such as an analysis of MRI images (Stamate et al., 2020; C.B. Hall et al., 2009; Patnode et al., 2020). Example brain volume-based work to identify AD and MCI biomarkers include the work of Kotb's group, (Kotb et al., 2020).

The results of work presented here lies within the category of genomic-epigenetic-based approaches, and is beyond the well known Aβ and Tau based biomarkers which is the focus of groups such as Gao and Van (Gao et al., 2021; Van et al., 2020). Rather, related example work that falls within the category is that from Kim's group (Kim et al., 2020). That is, excessive levels of Aβ and tau are not sufficient to diagnose or explain all instances of AD. Previous work has reported that between 30% and 40% of normal individuals showed high levels of Aβ and tau (Bennett et al., 2006; Mintun et al., 2006).

In addition, despite high levels of tau showing a greater correlation with increased cognitive dysfunction than Aβ, taken together both have a relatively weak correlation with the degree of cognitive function. Furthermore, although tau accumulation showed a higher association with cognitive dysfunction than did Aβ, both pathogenic proteins demonstrated a weak to moderate association with the degree of cognitive function (Driscoll et al., 2011; Giannakopoulos et al., 2003).There must be, therefore, other pathogenic contributors, beyond Aβ and tau, that can contribute to the onset and progression of AD.

BACKGROUND

The current diagnostic methods and symptoms related to AD are presented in many concluded research studies that clearly identifies Brain Amyloid-b (Aβ), in elevated amounts found in brains. As per them Amyloid-b is the main component of deposits that are found in the brains of patients with cognitive impairment (CI) and Alzheimer's disease (AD) (Dianxu Ren et al. 2020; Lianne M. Reus et al. 2020; Zhu, Xc. et al. 2020).

However, relatively recent research has found that AD appears only if Aβ is followed by elevated levels of Tau (HangRai Kim et al. 2020). That is, AD does not appear in the absence of elevated Tau levels. (Rubinski et al. Alzheimer's Research & Therapy 202)

Genetic studies of Alzheimer's disease (AD) indicate that β-amyloid is important in the pathogenesis of the disease. However, amyloid-directed therapy usually does not slow down the development of patients with symptomatic diseases. (William J. Ray et al. Annu. Rev. Med. 2021. 72:15–28)

Despite many advances of the past two decades, the cause of many cognitive impairments (CI) and potent treatments remain puzzling to find. The hypothesis that amyloid- (A) peptides are the key causative agents of CI holds dominant among current researchers.

Nevertheless, many existing research studies show evidence that A peptides are certainly not to be the only factor in AD etiology. A view of AD pathogenesis that encompasses both the amyloid-dependent and amyloid-independent mechanisms will help fill the gaps in knowledge and reconcile the findings that cannot be explained solely by the amyloid hypothesis.

The search for purposeful CI and AD biomarkers remains wide and extensive. Many different academic as well as industry R&D sectors have and continue to make significant progress toward identifying CI and AD biomarkers.

One of many technical gaps that remains, despite successful previous work to date, is the identification and characterization of additional genomic and epigenetic factors that are correlated with CI and AD.

Having expanded knowledge about the genomic factors related to CI and AD is expected to be the basis of more robust biomarkers for these neurodegenerative diseases. Better and earlier detection of CI and AD can lead to earlier treatment, better outcomes, and lower costs associated with the diseases.

The research proposal here is to identify and characterize some of the genomic variations, such as single nucleotide polymorphisms (SNPs), allele, chromosome, epigenetic contributors to CI and AD that are beyond Aβ and Tau, which is present in most of the existing works.

DATA

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is an initiative of researchers who study data as they work to define the progression of Alzheimer's disease (AD). ADNI researchers collect, validate and utilize data, including MRI and PET images, genetics, cognitive test and other biomarkers as predictors of the cognitive impairment. Study resources and data from the North American ADNI study are available through website login of adni.loni.usc.edu. These data include Alzheimer's disease patients, mild cognitive impairment subjects, and elderly controls.

The datasets that are used for the project ADNI GO2_GWAS / Alzheimer's Disease Neuroimaging Initiative (Grand Opportunities) Genome-wide association study. The ADNI Genetics Core has released GWAS data for ADNI GO/2 participants which is available from the download section. Due to the large volume of data, the 750 individual subject files have been compressed into 15 zip files. In total around 735 subjects were analysed for this research project.

APPROACH AND METHOD

To develop any predictive model, the first step is the identifying strong features set or input variables for the machine learning model with which reduces cost and time complexity, also better model performance can be achieved.

Hence feature selection is an important criteria for success for any machine learning ML / SVM model. It can also be said as the process of reducing the number of input variables when developing a predictive model.

Statistical-based feature selection methods, for example Chi-Square and correlation analysis involve evaluating the relationship between each input variable and the output or the response variable using statistics and selecting those features that have the strongest relationship with the response variable. These Statistical-based methods are standard industry based practices and are fast and effective, also here the adoption of these mentioned statistical methods were governed by choice of the data sets used ADNI GO2 GWAS for both the input and response variables. As explained above, the Chi-Square test of independence and collreation test were used to decide if there's a considerable relationship among two specific input / output variables.

Furthermore, compare the frequency of each category of input variable with the type of the output variable. The statistical information is displayed in a contingency table, with each row representing the category of a variable and each column representing the category of the opposite variable. It can be said that researchers want to investigate the relationship between different category cognitive impairment Groups (CN 0, MCI 1, AD 2) versus genomic variations; Allele 1 & 2 Plus , chromosome, SNP and their portions in DNA sequence structure. A contingency table is a form of table in a matrix layout that presents the frequency distribution of the variables.

Then the next step involves building a predictive model primarily on a support vector machine classifier that takes into account the input features identified here based on genomic and epigenetic factors to train for predicting the early stage of MCI. More details about how support vector machine classification is performed can be found in Appendix A.

To generate training and test data sets for the SVM model the scikit Train-Test Split function is used. The train-test split function can be used to help buiuld classification or regression models which in turn and can be used as a basis for supervised learning algorithms. The procedure involves taking a dataset and dividing it into two subsets. In this project the 80: 20 ratio was used for train-test split evaluation. Common splits are 80% training data and 20% testing data, called simple hold-out splits, from sklearn.model_selection import train_test_split, X_train X_test y_train y_test = train_test_split (X, y, test_size=0.20, random_state=33)

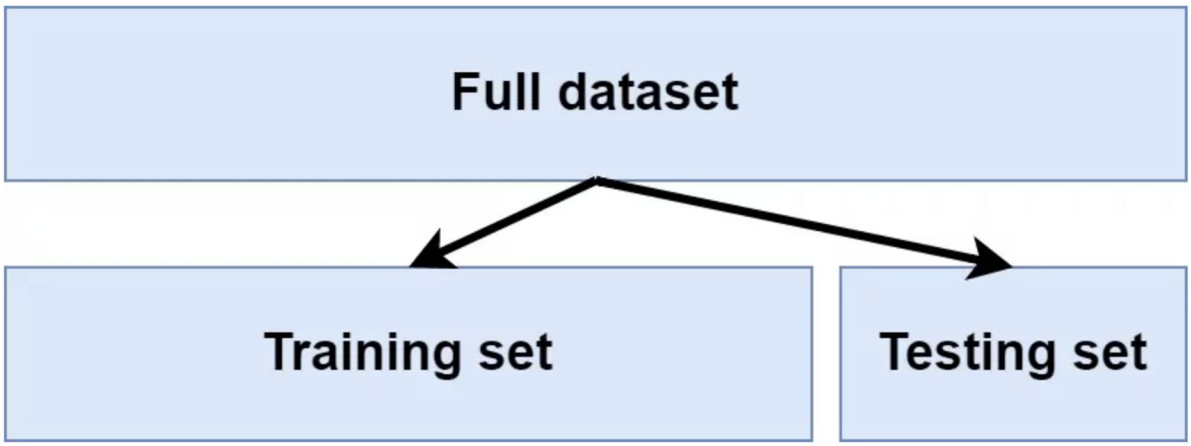

Figure 1. Train Test split function

EXPERIMENTAL METHOD

As per initial steps for the experimental process the clean data sets are the must prerequisite as it is directly responsible for the machine learning model, SVM binary classifier.

The pre-processing and cleaning stage for ADNA GWAS data files. Each file was in CSV format with a unique simple ID having around 75000 rows. Initially multiple attributes were considered; GC score, SNP index, SNP name, Allele 1 plus, Allele 2 plus, Chromosome, SNP position, GPA score, cluster separation, SNP , Allele 1 plus Vs Allele 2 plus,etc. Statistical analysis of the feature set further narrowed down to particular Chromosome type, SNP type & Positions.

['Chr'] == 4 & ['SNP'] == 13 ['SNP'] == 42

['Position'] == 21895517 or ['Position'] == 69033099 or

['Position'] == 116283010 or ['Position'] == 189690332

Some below substitutions were used here in the csv table as part of preprocessing datasets.

0 = control, 1 = MCI, and 2 = AD 2.

For the nucleotides in the Allele columns, let A=1, C=2, G=3, T=4 and for the SNP column an A/T = 14, an A/C =12, A/G=13,..., T/A=41, T/C=42 ... and so forth.

A Python-based support vector machine classified from scikit learn (CITE) was used for training and prediction of AD+MCI from control (CN) and non-control MCI /AD subjects). Generated a dataset that is separable and includes classes – so, in short, an easy and binary dataset.

Next created an SVM with a polynomial kernel to train a classifier, however not earlier than explaining the function of the kernel, also used C =1 and gamma to pass

a crucial part of SVMs. Finally post-processing helped in producing an accuracy confusion matrix as seen from high prediction model accuracy , lower counts on false positive , false negative results around the decision boundary of the model.

When using SVM, necessarily using one of the kernels: linear, polynomial or RBF=Radial Base Function (also called Gaussian Kernel). The larger the gamma, the narrower the gaussian "bell" is.

Split Train and Test data :: 80:20 ratio implemented by below function in code.
x_train,x_test,y_train,y_test=train_test_split(X1,y,test_size=0.2)

SVM with a polynomial kernel to train a classifier
classifier = svm.SVC(kernel='poly', C=1, gamma=0.01, class_weight='balanced')

Results will be verified by splitting the data-sets into training and test sets to get a more accurate assessment of SVM classifier's performance. Also SVM model binary outcome for predicting early onset of cognitive impairment will be compared with the actual ADNI sample dataset results to pick for any false positive and false negative based on confusion matrix.

RESULTS

In machine learning ML , functions are independent single variables or features that can be used as input to the system. In fact, when models make predictions, they use these attributes (features) to make predictions.

The class label is a discrete attribute, and you want to predict its value based on the values of other attributes. In this case, a controlled normal group CN or mild

cognitive impairment MCI /AD are binary class names. The intent is to analyze a function to calculate whether a person with a specific attribute value will be a candidate for MCI / AD or not.

Experimentation results were derived with the built SVM binary classifier. Having a polynomial kernel with C=1 and gamma=0.01, resulted in a prediction accuracy of 78.32 % or higher. Allele, Chromosome, SNP and SNP positions features were identified as input to the binary ML model predicting either cognitive impairments versus none for a certain control group.

How did you assess the quality and correctness of the obtained results?

To assess the quality and correctness of the obtained results , a two pronged approach is taken. First from statistical analysis features were identified using statistical models Contingency table, Chi square test, Pearson correlation test. The data used in this research were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database / ADNI GO2 GWAS.

Next best of the features narrowed down from above statistical analysis were fed into a predictive model based on a support vector machine (SVM) classifier that takes into account the genomic variations and epigenetic factors to predict the early stage of MCI/AD.

From these binary SVM classifiers an accuracy of around 78% were achieved to predict early stages of MCI / AD versus normal people.

Below examples of features extracted from genomic variations were fed into inputs to the SVM classifier.

['Chr'] == T

&

['SNP'] == AG ['SNP'] == TC

&

['Position'] == 21895517 or ['Position'] == 69033099

or

['Position'] == 116283010 or ['Position'] == 189690332

A small change of adding C and gamma value seems to be bumping Up accuracy considerably.

classifier = svm.SVC(kernel='poly', C=1, gamma=0.01, class_weight='balanced')

This comparatively higher accuracy for our experimental results.

EXPERIMANTAL RESULTS

As per observed experimental results statistical tools lead to better identification of feature sets used in machine learning models. Contingency table, Chi square test, correlation test, highlighted genomics variations with higher frequencies for example chromosome-type, allele or around SNP for a group.

Results observed from Statistical tools / Chi-square test and Pearson correlation **:**

1- Here MCI group has a high number of nucleotides in the Allele type T.

>>> print(contingency Table)

In statistics, a contingency table is a cross tabulation type of table in a matrix format that displays the frequency distribution between the variables. Here it's providing a picture of the interrelation between two variables that is the strong relationship between MCI and Allele1 plus type G.

Allele1 - Plus	1	2	3	4
Group				
0	475	380	452	943
1	1117	938	1255	**2465**
2	325	266	345	699

2- Similarly, the MCI group can observe a high number of Chromosome 4 genomic types.

Here as per contingency table it's providing a picture of the interrelation between two variables, that is a strong relationship between mild cognitive impairment MCI and Chromosome type 4 .

```
>>> print(contingencyTable)
Chr     2    3     4    12
Group
0      149  300  1651  150
1      385  770  4235  385
2      107  216  1203  109
```

3- MCI group has observed a high number of A/G genomic types for single nucleotide polymorphisms SNP.

Here as per contingency table it's providing a picture of the interrelation between two variables, that is a strong relationship between mild cognitive impairment MCI and SNP 13 or A/G .

```
>>> print(contingencyTable)
SNP     12    13    42   43
Group
0      300  1049   749  152
1      770  2694  1924  387
2      218   758   543  116
```

The chi-square test of independence further can be used to examine these above stated relationships. Based on Chi-square statistics & p-value one can conclude that a relationship exists between the categorical variables; that is the results are significant or not.

Here objectives involve identifying some of the genomic variations, such as single nucleotide polymorphisms (SNPs), allele , chromosome, epigenetic contributors to MCI and AD that are beyond that related to Aβ and Tau. These areas of interest were answered here with encouraging results showing positive prediction accuracy for selected genomic features and strong potential to expand into research scoping beyond Aβ and Tau. And can further expand into identifying more genomic variations for developmental prediction models based classifiers.

CONCLUSION AND DISCUSSION

The proposed effort involved identifying some of the genomic variations, such as single nucleotide polymorphisms (SNPs), allele , chromosome, epigenetic contributors to MCI and AD that are beyond that related to Aβ and Tau.

The project led to building a predictive model based on a support vector machine (SVM) classifier that takes into account the above mentioned genomic variations and epigenetic factors to predict the early stage of MCI/AD. Based on the results from the SVM binary classifier it is evident that using certain genetic variations can be used for predicting early onset of cognitive impairment diseases.

The research strategy used here can be helpful to engineer better mechanisms to enhance classification models for early stage detection of cognitive impairments.

FUTURE WORK

Explore the genes the SNPs are associated with to see if there are metabolic pathways involved or epigenetic relationships with other known causes or factors related to AD/MCI. Secondly, increase sample size analyzed from ADNI DB. Next explore and analyze potential secondary effects/medical-conditions such as other diseases that might have influenced the observed results. Separate out MCI from AD and explore predictions and results.

APPENDIX A

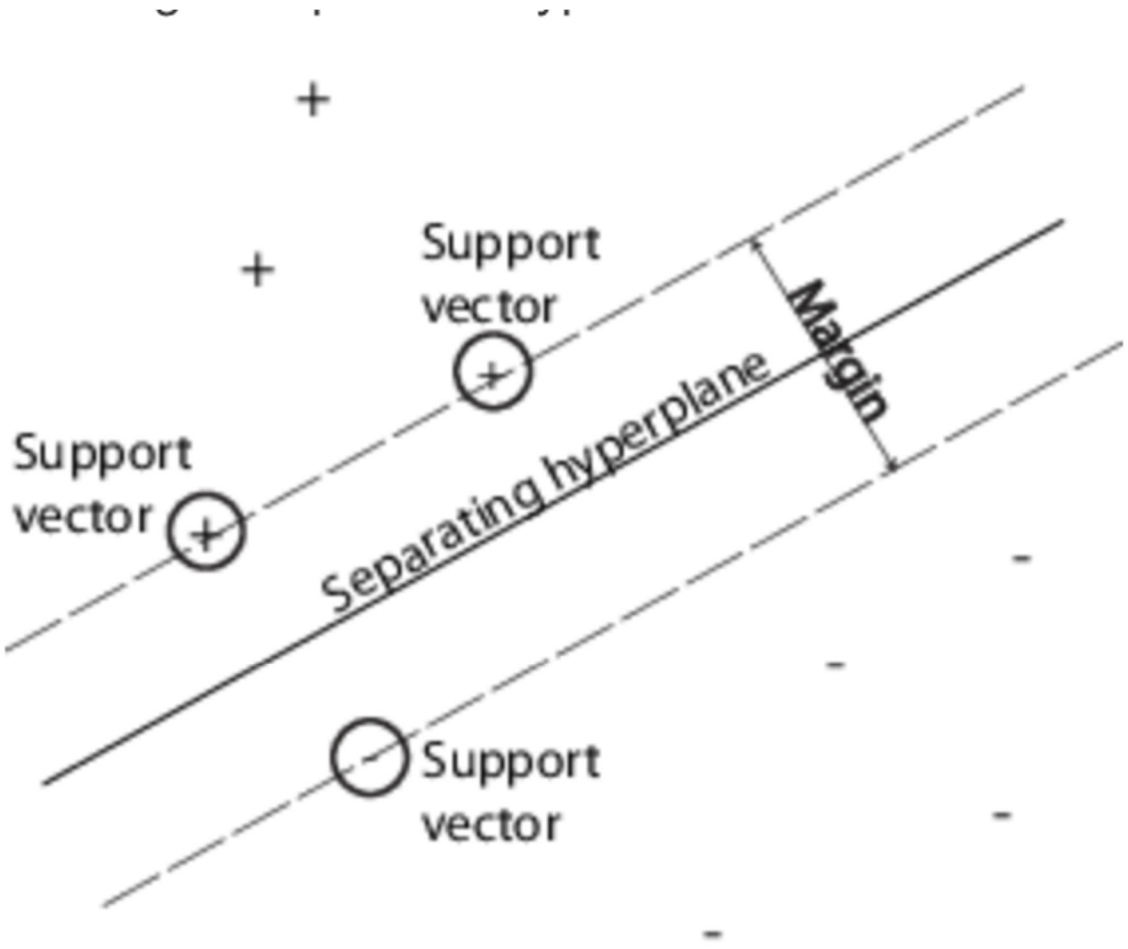

Fig 2 .Separable Data for SVM classifier

The above figure depicts the use of a support vector machine (SVM) when data has exactly two classes that are in here, two subjects control group (CN) Vs cognitive impairment (MCI/AD). An SVM classifies data by finding the best hyperplane that separates all data points of one class from those of the other class. The best

hyperplane for an SVM means the one with the distinct margin between the two classes. Margin means the width of the slab parallel to the hyperplane that has no internal data points.

As with any supervised learning model, first train a support vector machine, and then cross validate the classifier. Use the trained machine to predict new data. In addition, to obtain acceptable predictive accuracy, test various SVM kernel functions, and further tune the parameters of the kernel functions.

REFERENCES

Bennett D, Schneider J, Arvanitakis Z, et al. (2006) Neuropathology of older persons without cognitive impairment from two community-based studies. Neurology 2006;66:1837-1844.

C.B. Hall, R.B. Lipton, M. Sliwinski, M.J. Katz, C.A. Derby, & J. Verghese. (2009). Cognitive activities delay onset of memory decline in persons who develop dementia. *Neurology, 73*, 5

Davis, K. K., & Allen, J. K. (2013). Identifying cognitive impairment in heart failure: a review of screening measures. *Heart Lung, 42*(2), 92-97. doi:10.1016/j.hrtlng.2012.11.003

Dianxu Ren et al. (2020). The Effect of the APOE ε2ε4 Genotype on the Development of Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) in Non-Latino Whites. J Am Geriatr Soc. 2020 May; 68(5): 1044–1049. doi:10.1111/jgs.16337.

Driscoll I, troncoso J, (2011) Asymptomatic Alzheimer's disease: a prodrome or a state of resilience? Curr Alzheimer Res 2011;8:330-335.

Elizabeth E. Moore et al. (2020). Mild Cognitive Impairment Staging Yields Genetic Susceptibility, Biomarker, and Neuroimaging Differences. Frontiers in Ageing Neuroscience, Volume 12, Article 139

Giannakopoulos P, Herrmann F, Bussiere T, et al., (2003) Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. Neurology 2003;60:1495-1500

Guo, Y., Huang, Y.-Y., Shen, X.-N., Chen, S.-D., Hu, H., Wang, Z.-T., Tan, L., & Yu, J.-T. (2021). Characterization of Alzheimer's tau biomarker discordance using

plasma, CSF, and PET. *Alzheimer's Research & Therapy*, *13*(1), 1–12. https://doi.org/10.1186/s13195-021-00834-3

Haines, Jonathan L (2018). Alzheimer Disease: Perspectives from Epidemiology and Genetics. The Journal of Law, Medicine & Ethics, 46 (2018): 694-698

Hang-Rai Kim et al. (2020). Genetic variants beyond amyloid and tau associated with cognitive decline: A cohort study. American Academy of Neurology DOI: 10.1212/WNL.0000000000010724

Hu, C., Yu, D., Sun, X., Zhang, M., Wang, L., & Qin, H. (2017). The prevalence and progression of mild cognitive impairment among clinic and community populations: a systematic review and meta-analysis. *Int Psychogeriatr, 29*(10), 1595-1608. doi:10.1017/S1041610217000473

Kim, H. R., Lee, T., Choi, J. K., & Jeong, Y. (2020). Genetic variants beyond amyloid and tau associated with cognitive decline. *Neurology*, *95*(17), e2366–e2377. https://doi.org/10.1212/wnl.0000000000010724

Kotb, M. A., Kamal, A. M., Aldossary, N. M., Alsify, A. A., & Ahmed, Y. M. (2020). Value of magnetic resonance spectroscopy in geriatric patients with cognitive impairment. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery, 56*(1). doi:10.1186/s41983-020-0147-y

Lianne M. Reus et al. (2020). Degree of genetic liability for Alzheimer's disease associated with specific proteomic profiles in cerebrospinal fluid, Neurobiology of Aging, Volume 93, 2020, Pages 144.e1-144.e15.

Mintun M, Larossa G, Sheline Y, et al. (2006) [11C] PIB in a nondemented population: potential antecedent marker of Alzheimer disease. Neurology 2006;67:446-452.

Niaz, K., Shah, S.Z.A., Khan, F. et al. (2020) Ochratoxin A–induced genotoxic and epigenetic mechanisms lead to Alzheimer disease: its modulation with strategies. Environ Sci Pollut Res 27, 44673–44700. https://doi.org/10.1007/s11356-020-08991-y

Patnode, C. D., Perdue, L. A., Rossom, R. C., Rushkin, M. C., Redmond, N., Thomas, R. G., & Lin, J. S. (2020). Screening for Cognitive Impairment in Older Adults: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA, 323*(8), 764-785. doi:10.1001/jama.2019.22258

Rawtaer, I., Mahendran, R., Kua, E. H., Tan, H. P., Tan, H. X., Lee, T. S., & Ng, T. P. (2020). Early Detection of Mild Cognitive Impairment With In-Home Sensors to Monitor Behavior Patterns in Community-Dwelling Senior Citizens in Singapore: Cross-Sectional Feasibility Study. *J Med Internet Res, 22*(5), e16854. doi:10.2196/16854

Romero-Rosales B-L et al. (2020) Improving predictive models for Alzheimer's disease using GWAS data by incorporating misclassified samples modeling. PLoS ONE 15(4): e0232103. https://doi.org/10.1371/journal.pone.0232103

Rubinski et al. (2020). FDG-PET hypermetabolism is associated with higher tau-PET in mild cognitive impairment at low amyloid-PET levels. Alzheimer's Research & Therapy 12:133 https://doi.org/10.1186/s13195-020-00702-6

Sharmin Hossain et al. (2020) Longitudinal associations between dietary quality and Alzheimer's disease genetic risk on cognitive performance among African American adults. British Journal of Nutrition (2020), 124, 1264–1276 doi:10.1017/S0007114520001269.

Stamate, D., Smith, R., Tsygancov, R., Vorobev, R., Langham, J., Stahl, D., & Reeves, D. (2020). Applying Deep Learning to Predicting Dementia and Mild

Cognitive Impairment. In *Artificial Intelligence Applications and Innovations* (pp. 308-319).

Van Hulle, C., Jonaitis, E. M., Betthauser, T. J., Batrla, R., Wild, N., Kollmorgen, G., Andreasson, U., Okonkwo, O., Bendlin, B. B., Asthana, S., Carlsson, C. M., Johnson, S. C., Zetterberg, H., & Blennow, K. (2020). An examination of a novel multipanel of csf biomarkers in the alzheimer's disease clinical and pathological continuum. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. https://doi.org/10.1002/alz.12204

Vasanthakumar, A., Davis, J. W., Idler, K., Waring, J. F., Asque, E., Riley-Gillis, B., . . . Alzheimer's Disease Neuroimaging, I. (2020). Harnessing peripheral DNA methylation differences in the Alzheimer's Disease Neuroimaging Initiative (ADNI) to reveal novel biomarkers of disease. *Clin Epigenetics, 12*(1), 84. doi:10.1186/s13148-020-00864-y

Yu Guo et al. (2020). Discordant Alzheimer's neurodegenerative biomarkers and their clinical outcomes. Annals of Clinical and Translational Neurology published by Wiley Periodicals LLC on behalf of American Neurological Association

Zhu, Xc. et al. (2020). Impacts of CR1 genetic variants on cerebrospinal fluid and neuroimaging biomarkers in alzheimer's disease. BMC Med Genet 21, 181. https://doi.org/10.1186/s12881- 020-01114-x

Lawlor, B. A. et al.(1994). Clinical symptoms associated with age at onset in Alzheimer's disease. The American Journal of Psychiatry, 151(11), 1646–1649. https://doi.org/10.1176/ajp.151.11.1646