HIGH SCHOOL EDITION
Journal of Student Research
Volume 11 Issue 3 (2022)

Machine Learning in Alzheimer's Disease: Prognostic Prediction via Neuroimaging and Numerical Data

Ydathip Phetchrungruengphol[1], Phasit Thanitkul[1], Chavakorn Arunkunarax[1], Thanapol Wongtharua[1], Kunapas Sumpunwetchakul[1], Jetnipat Kongsirirungruang[1], Niracha Janavatara[1], Witawin Sittisirinukul[1], Chutikarn Kanchana-art[1], Pingpan Krutdumrongchai[1], Tansiri Praditphonlert[1] and Kanokkorn Pitayakornpakdee[1#]

[1]The Newton Sixth Form School, Thailand
[#]Advisor

ABSTRACT

Alzheimer's disease (AD) is often detected too late or inaccurately in clinical practice. Therefore, improvement in the current methods of AD detection will provide opportunities for early intervention, symptomatic treatment, and, overall, better quality of patients' and their caregivers' lives. The paper is an in-depth study of how functional brain imaging and support vector machine (SVM) could be utilized to detect the risk for AD which works by assessing and plotting data into multi-dimensional graphs for various results. It aims to identify patients in presymptomatic stages for early treatment to delay or prevent progressive cognitive decline and disease. With knowledge of machine learning, our medical tool is a breakthrough in the methodology of AD detection. The future of our tool requires a substantive amount of brain scan data for the machine learning algorithm to produce reliable results, so further research in this field of study is crucial and strongly encouraged.

Introduction

Dementia, with over 55 million people in the world affected, is currently the seventh leading cause of death among all diseases and one of the major causes of disability and dependency among elderlies globally (WHO, 2021). Being the most common cause of dementia, AD disrupts patients' behavioral abilities and interferes with one's daily life and activities. There are a total of 7 stages of AD - divided into the early and late stages. In the early stages, toxic chemical changes take place in the brain. Beta amyloid plaques (also known as neuritic plaques, Aβ plaques, or senile plaques) are formed and tau tangles, due to abnormal accumulation of proteins, are manufactured. This causes healthy neurons to experience functional and communication decline and die. Initial damage occurs in the hippocampus and the entorhinal cortex, both of which are essential parts of the brain in forming memories, thus symptoms of early-stage AD patients being the deterioration in the patients to perform cognitive functions that require thought or the ability to retain memory. As more neurons die, other parts of the brain are affected and begin to shrink. By the final stage of AD, damage is widespread, and the mass of brain tissue decreases significantly until the patient is no longer able to perform basic cognitive or physical functions and gradually passes away (NIA, 2021). In both of these stages, harm is already done to the patients' cognitive functions, thus diagnosis should be established during the preclinical stage before it accumulates to the point where it could be exhibited in the form of symptoms and the effects would be irreversible.

Nowadays, the diagnosis of Alzheimer's is done using combined data from the CSF by performing a lumbar puncture combined with brain scans (CT, MRI, PET). Such methods are proven to be insufficient, with morphological changes only becoming visible for diagnosis when the damage has already been done. Therefore, by the time said drugs can be widely used in clinical practice, it will be too late and the drugs that are used to treat severe AD cannot

do much to delay the inevitable. However, if accurate diagnosis can be issued earlier, people would be able to receive proper treatment quicker, increasing the chances of recovery. In effect, an optimal solution would be to use artificial intelligence (AI) as a tool to detect morphological changes in scans of the early stages of AD that the human eye may not be able to recognize.

Machine learning (ML), a branch of AI, is a method for extracting patterns through the use of raw data through algorithms (IBM Cloud Education, 2020). AI application in the medical field has been increasingly experimented with in the past decade, producing invaluable results for the healthcare field and patients of tomorrow. A successful case of using ML algorithm as a detection tool to analyze patterns of pulmonary diseases, of which has been proven to double the accuracy of diagnoses in comparison to those done by human medical professionals (Topalovic M., 2019). The concept can be applied to detect early signs of AD from several scans of the brain and as well. By inputting several parameters that are early indicators of the disease, the algorithm will be able to detect small unseen differences through training with a large amount of corresponding data.

Our research aims to solve the problem of insufficient diagnosis strategies for AD through the use of ML. If our hypothesis is proven to be successful, countless lives will be saved annually - with our methodology aiming to get patients the treatment that they need before it develops to the point where it may cause them life-long harm, likely even resulting in death. Similar prototypes can also be adopted and used to diagnose other neurodegenerative diseases or any disease in general to similarly prevent the problem before it is able to manifest, given sufficient data and the right parameters to train the algorithm.

Literature Review

Biomarkers of beta-amyloid (Aβ) are used as one of the criteria for the early stages of Alzheimer disease (AD) and are increasingly used in clinical trials. This stresses the need for reliable and available biomarkers of brain Aβ pathology. Four modalities have been established: CSF Aβ42/40, Florbetapir-F18 PET, amyloid PET, and tau PET. In which, CSF Aβ42/40 and Florbetapir-F18 PET correlate highly with findings that will be discussed further. In our project, we intend to examine these biomarkers using machine learning to detect the possibility of developing AD in early stages.

Machine learning (ML) is a type of artificial intelligence (AI) that enables softwares to accurately predict outcomes without being explicitly programmed to do so. It is incredibly complex and its operation varies depending on the task and the algorithm used to accomplish it. However, at its core, an ML model looks at data, identifies patterns, and uses those insights to complete assigned tasks. Any task that relies upon a set of data points or rules can be automated using ML.

ML is more accurate than human detection because it detects details which humans cannot see. It works by inputting a database of images and finding similarities between them to detect even the smallest changes, pixel by pixel. This helps decrease the amount of time doctors spend diagnosing images and provide a more detailed report to patients. In 2021, ML has been used to diagnose lung cancer and the results demonstrate the potential to help prevent lung cancer deaths through early detection: a ML model was more accurate for early diagnosis than standard eligibility criteria for screening (Am J Respir Crit Care Med, 2021). We aim to use ML, specifically supervised learning, to achieve similar results as the case above with the diagnosis of AD.

Supervised learning, a category of ML, involves training the predictive model through classification algorithms, by inputting a labeled set of data and telling the computer which category the piece of data belongs to. Then, the model can be used to classify the input data. To complete such algorithms, the model requires support vector machines, neural networks, linear and logistic regression, random forest, and classification trees.

Biomarkers of AD Used in This Study

Tau PET

Chemical changes, namely hyperphosphorylation, occur in tau protein in AD. Under normal circumstances, tau is contained intracellularly and regulates microtubule polymerization. However, when hyperphosphorylated, tau threads pair up into paired helical filaments (PHFs) and form extremely insoluble self-aggregates (NFTs), which disrupt synaptic function and precipitate cell death. These filaments tangle together, causing disintegration of microtubules and collapse of the neuron's transport system. Neurofibrillary tangles (NFTs) made of tau are the pathologic hallmark of a group of neurodegenerative disorders known as tauopathies (Rowley et al., 2020).

The ideal PET tau tracer would have a stronger selective binding potential to PHFs and phosphorylated tau over β-amyloid (β-amyloid has higher concentration in diseased brains), high permeability of blood–brain barrier, low metabolism, and low non-tau (non-target) binding to other CNS receptors and tissues (e.g. white matter). Worse memory performance was associated with greater PET tau binding in the entorhinal cortex.

PET tau imaging enables the longitudinal assessment of the spatial pattern of tau deposition and its relation to amyloid-β pathology and neurodegeneration. Recently, several tau PET tracers including T807, THK-5117, and PBB3 have been developed and succeeded in imaging neurofibrillary pathology. For use of tau PET as a biomarker of tau pathology in AD, PET tracers should have high affinity to PHF-tau and high selectivity for tau over amyloid-β and other protein deposits (Okamura et al., 2014).

The accumulation of pathologic tau is more closely related to functional and structural deterioration in the AD spectrum than β-amyloid (Neurol, 2015).

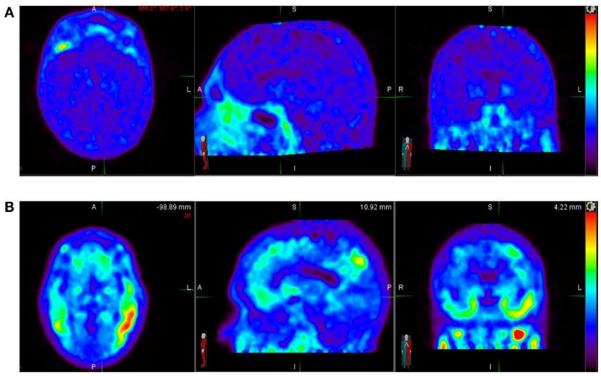

Figure 1. (A) PET Scan

Normal 18F-AV-1451 PET study illustrating a tau-PET scan from an elderly cognitively normal subject (MMSE score of 30). No abnormal uptake is demonstrated above diffuse background activity. **(B)** Abnormal 18F-AV-1451 PET study illustrating a tau-PET scan of images of a subject with mild Alzheimer's disease (MMSE score of 22). Unlike an individual with normal cognition, there is abnormally increased radiotracer uptake distributed in the temporal lobes bilaterally.

Amyloid PET

Beta-amyloid is produced when a much larger protein referred to as the amyloid precursor protein (APP) is broken down. APP is composed of 771 amino acids and is split by two enzymes to produce beta-amyloid. Amyloid beta peptide (Aβ) is produced through the proteolytic processing of a transmembrane protein, APP, by β- and γ-secretases

(Sally Robertson, B.Sc, n.d.).

Beta-amyloid 42 is especially toxic, as abnormal levels of this naturally occurring protein clump together to form plaques that collect between neurons and disrupt cell function. Amyloid PET results do not establish a diagnosis of AD. It is an adjunct to other diagnostic evaluations. PET scan is for additional information to help clarify an otherwise unclear diagnosis.

Initially, 2-Deoxy-2-[18F]fluorodeoxyglucose ([18F]FDG), a glucose analog, was utilized for PET brain imaging for the management of AD. Healthy brain cells avidly take up the substance, as they highly metabolize glucose, but the substance is relatively reduced in uptake in the temporo-parietal cortical regions that are affected by AD.

Amyloid precursors for PET imaging such as the short radioactive half-life (20 minutes) Pittsburgh compound B ([11C]PiB), by crossing the blood brain barrier (BBB) and selectively binding to beta amyloid plaques (Aβ), are able to provide a virtual ante mortem histopathological portrait of the brain.

More specifically, [18F] FDG PET is used to differentiate AD from other clinical diagnoses by normalizing the uptake intensity to the mean metabolic rate for glucose utilization in the whole brain (CMRglc) or the cerebellar glucose consumption, as these areas allow for accurate distinction of AD by being maximally stable in subjects but minimally affected by external stimuli (Lois et al., 2019).

Florbetapir-F 18 PET

Florbetapir(^{18}F)((E)-4-(2-(6-(2-(2-(2[18F]fluoroethoxy)ethoxy)ethoxy)pyridine-3-yl)vinyl)-N-methylbenzamine)), also known as ^{18}F-AV-45, is a molecular imaging agent that binds to Aβ aggregates which is intended to be used with PET scan of the brain via intravenous injection ("Amyvid", n.d.).

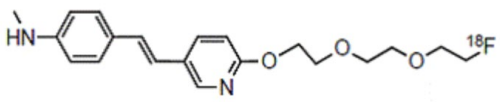

Figure 2. (E)-4-(2-(6-(2-(2-(2[18F] fluoroethoxy)ethoxy) ethoxy) pyridine-3-yl)vinyl)-N-methylbenzamine)

This is to estimate the Aβ neuritic plaque density in patients being evaluated for AD ("Florbetapir (18F)", n.d.). The use of this radiopharmaceutical compound has been backed up by numerous studies to be safe and display strong correlation between ^{18}F-florbetapir PET images and postmortem assessment of Aβ deposition (Okamura &Yanai, 2010) (Cortes-Blanco, 2014).

^{18}F-florbetapir PET scans are binarily classified as either negative (Aβ-) or positive (Aβ+) for each region of the brain and shown in terms of regional-to-whole cerebellar standard uptake value color scale (SUVR).

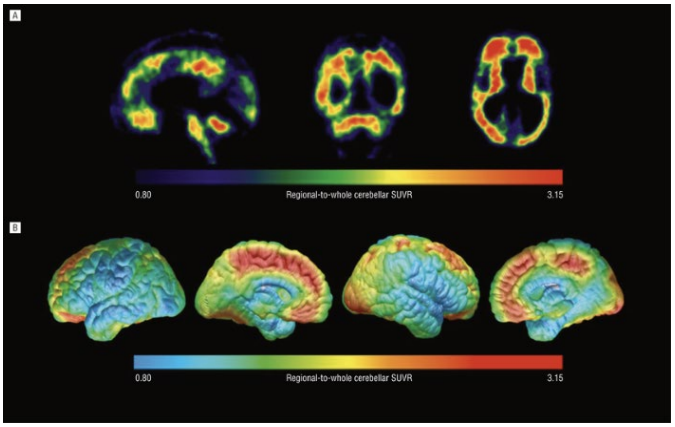

Figure 3. Example of ^{18}F-florbetapir PET scan on a regional-to-whole SUVR color scale

A negative ^{18}F-florbetapir scan indicates very little to no neuritic plaques, which is inconsistent with AD diagnosis and partially rules out that a patient's cognitive impairment is due to AD. Although a positive ^{18}F-florbetapir scan can translate to moderate to abundant amounts of amyloid neuritic plaques, it does not confirm the diagnosis of AD or any other cognitive disorder (Yang et al, 1970).

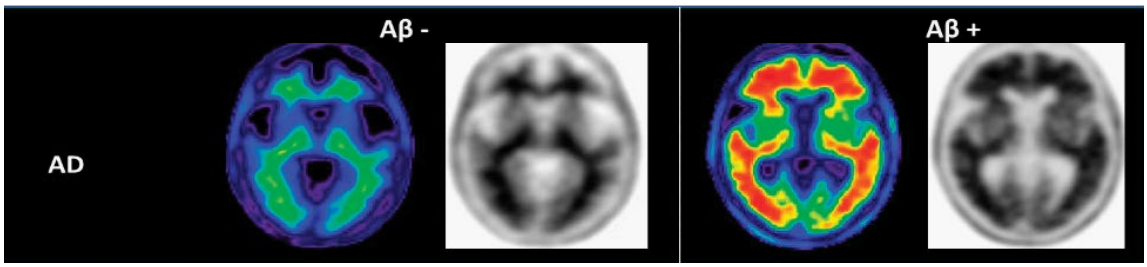

Figure 4. shows Aβ- and Aβ+ ^{18}F scans of AD patients

These ^{18}F-florbetapir PET scans are usually visually assessed by trained sets of human eyes. This can be both time consuming and incredibly difficult, especially for images with low levels of Aβ deposition. Over the past decade, studies have applied machine learning and deep learning to improve the accuracy of classifying these images and their results have been promising (de Vries et al., 2020).

CSF Aβ42/40 ratio

Aβ40 is a 40 amino acid proteolytic product from the amyloid precursor protein (APP). In addition, it is a biomarker correlating with Alzheimer's disease. In healthy states, Aβ40 is the most abundant form of the amyloid peptides in both cerebrospinal fluid (CSF) and plasma (10 to 20 times higher than Aβ42). The only difference between Aβ40 and Aβ42 is that two additional free carboxyl groups reside on Aβ42.

Aβ42 is a γ-secretase modulator, which can be used for Alzheimer's detection ("Aβ42-IN-2", 2022). It is significantly more neurotoxic than Aβ40 and forms amyloid fibrils much more rapidly than Aβ40 (*J. Am. Chem. Soc.* 2009). Recent studies suggest that a decrease in the ratio of Aβ42/Aβ40 may indicate AD progression (Aβ40 Assays, 2022).

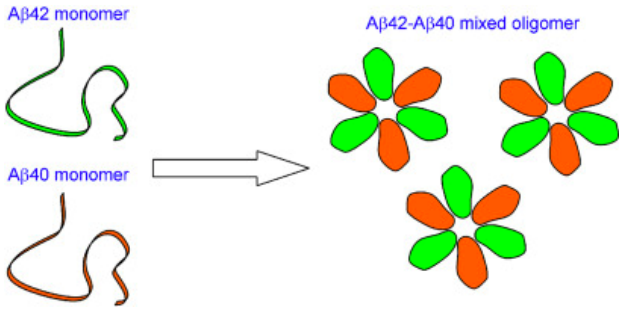

Figure 5. Diagrams of mixed oligomers

Aβ40 and Aβ42 interact strongly, existing as mixed species labeled Aβ40/Aβ42 monomer mixtures according to Figure 4 diagram of mixed oligomers. The concentration ratio of Aβ42 to Aβ40 (Aβ42/40 ratio) has been suggested to be superior to the concentration of either Aβ42 or Aβ40 alone (Risto Cukalevski et al., 2015) (Oskar Hansson et al., 2019).

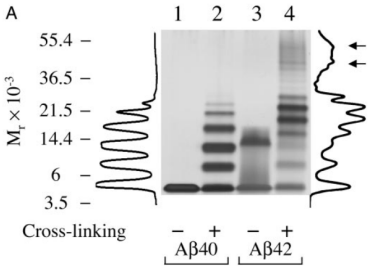

Figure 6. Cross-linked PICUP Scanner

1= non crossed Aβ40; 2=crossed Aβ40; 3= non crossed Aβ42; 4= crossed Aβ42
Arrows next to the intensity profile of cross-linked Aβ42 indicate intensity maxima corresponding to the PICUP scanner. This scanner works by the PICUP stabilizing oligomer populations by covalent cross-linking and when combined with fractionation methods, such as sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) or size-exclusion chromatography (SEC). PICUP provides snapshots of the oligomer size distributions that existed before cross-linking. Hence, PICUP enables visualization and quantitative analysis of metastable protein populations and can be used to monitor assembly and decipher relationships between sequence modifications and oligomerization.

Findings from a recent CSF research indicates plasma Aβ42 and Aβ42/40 were lower in individuals who were Aβ positive compared with those who were Aβ negative whereas there were no differences in the levels of Aβ40 (Janelidze, 2021).

Correlation between CSF Aβ42/40 and ^{18}F-florbetapir PET

Alone, ^{18}F-florbetapir PET scan has not yet displayed accurate usefulness in studies, but combining CSF Aβ42/40 ratio has shown an established correlation that is proven valuable to the accuracy of AD diagnosis. Reduced CSF Aβ42/40 and increased ^{18}F-florbetapir (PET) uptake reflects brain Aβ accumulation which indicate AD progression (Mattsson et al., 2014).

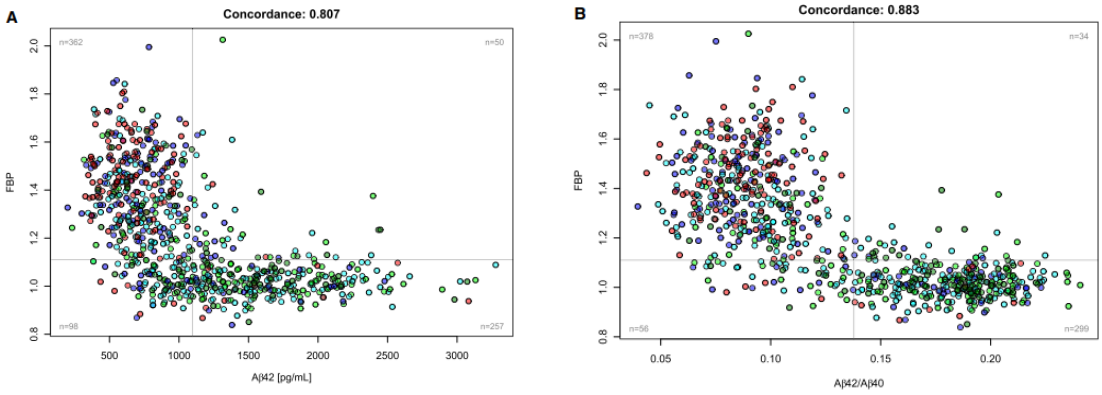

Figure 7 and 8. Relationship between CSF biomarkers and Cortical Florbetapir

Based on the analysis of the relationships between CSF biomarkers and cortical ^{18}F-florbetapir SUVRs shown above, the concordance for Aβ42 and ^{18}F-florbetapir PET was 81% and increased to 88% for the CSF Aβ42/40 ratio.

Application of biomarkers in the diagnosis of Alzheimer's Disease (AD)

After obtaining all the values, we can train the ML model to detect abnormalities in the biomarkers and determine the risk of a person having AD based on the information provided.

PET scans

The regional-to whole cerebellar standard uptake value color scale (SUVR) is used to classify the density of different biomarkers from PET scans to determine the prospect of a patient's symptoms deriving from AD in accordance to its consistency in AD diagnosis. Regions on a PET scan that appear red indicate a concentrated density of biomarkers and inversely appear black otherwise (in the case where there is a low density of biomarkers). A higher concentration of beta amyloid neuritic plaques in ^{18}F-florbetapir PET scans and more tau protein binding are trends usually observed in AD patients that can be visually assessed through PET scans using the aforementioned SUVR scale. Identification in the lack or surplus in accumulation of said biomarkers could thus help eliminate AD from the potential diseases that may be the cause of the patient's cognitive abnormality, or in effect also help identify AD as its cause (Yang et al, 1970).

Aβ42 and Aβ42/40 ratio

In regards to determining abnormality, we will use a cutoff ratio of Aβ42 and Aβ42/40 plasma and put the value into the machine to determine its degree of abnormality or lack thereof. A lower ratio than the cutoff will indicate Aβ accumulation in the brain which suggests AD progression.

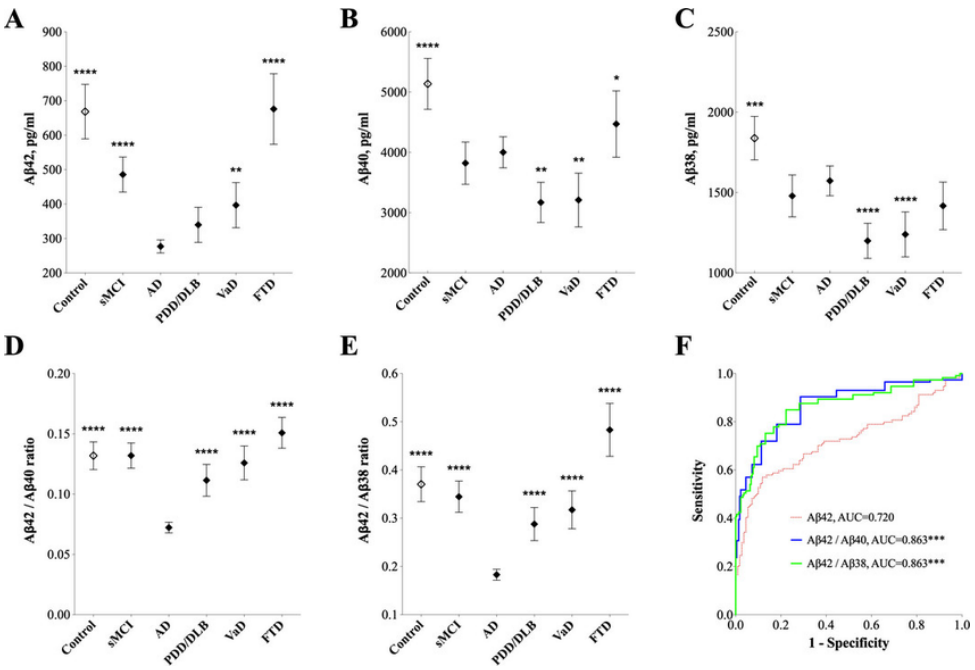

Figure 9. Comparison of ratio biomarkers in control groups and those in patients with AD

If we succeed, we would be able to train a machine learning algorithm to assess a patient's likelihood of contracting AD by visually processing scans of biological markers to match the results in accordance to its consistency in AD diagnosis; especially in scans where there are lower levels of deposition of biomarkers. By its nature, PET scans are incredibly time consuming and challenging to evaluate even to trained eyes. In addition to the fact that the trends of each biomarker could only suggest chances of a patient contracting AD, diagnosis by utilizing the trends and correlations of a combination of biomarkers and a tool that would be able to appraise data sets efficiently would substantially facilitate if not improve the current methods of AD diagnosis in the medical field. With the biomarkers mentioned above combined with machine learning, we will be able to detect early stages of AD and raise the quality of life of patients with sooner treatment.

Methods

Though there is presently no available medication for the treatment of Alzheimer's Disease (AD), we utilize computing technology and various resources: machine learning (ML) model by Python and supervised learning algorithm by Support Vector Machine (SVM) to bring forth accurate and comprehensive diagnosis to improve the quality and productivity of healthcare. Proposed below are the materials and methods of a promising solution for the management of AD.

PET scans

A positron emission tomography (PET) scan is an imaging test which can be used as a tool to diagnose chances of a patient being affected by neurological diseases, using radioactive drugs such as Flortaucipir (18F), Fludeoxyglucose (18F) and ^{18}F-florbetapir. The drug is injected into your body, where it attaches to the plaques. PET will then make β-amyloid neuritic plaques light up on a tomography scan in colors corresponding to their density on the standardized

uptake value ratio (SUVR) scale in the cortex and fatty membrane surrounding nerve cells of the brain with the Florbetapir. The same applies to Tau tangle forms when the ligand is Flortaucipir, and Aβ42 accumulation in amyloid scans when the radiotracer is Fludeoxyglucose (red in areas with high density and black otherwise). PET scans can be used to diagnose patients in the early stages of AD.

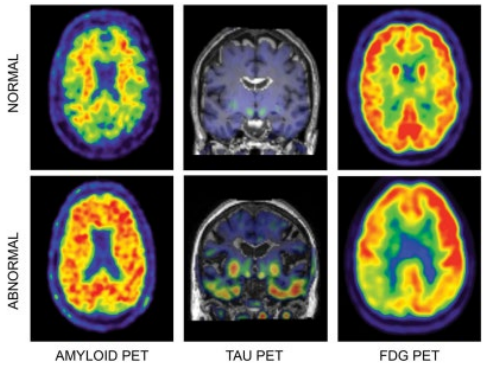

Figure 10. Example of positive and negative amyloid, tau and PET scans

Cerebrospinal Fluid Aβ40/42

In AD, there are two types of proteins that can be crucial in Alzheimer's diagnosis, Aβ40 and Aβ42. Under normal circumstances, there are about 10-20x more Aβ40 compared to Aβ42. Aβ42 in humans are known to be toxic, they form amyloid fibrils at a rapid pace, causing disruption in the brain's function. In patients with AD, the ratio of Aβ42/40 is lower than those who do not have the disease and this ratio between the two proteins has been proven to be more effective than the levels of each individual protein and these values can contribute to the diagnosis of Alzheimer's disease.

Known Correlations

Tau PET scans and beta amyloid protein correlations

The APOEε4 allele- an established genetic risk factor for AD and amyloid-β is found to be related to increased tau load. One *APOEε4* allele and amyloid-β is found in areas with higher vulnerability to AD and more a higher density of tau, while the interaction between amyloid-β and two *APOEε4* alleles is shown to lead to a more widespread pattern of tau aggregation. A higher amyloid-β load detectable by both CSF and PET is required before substantial tau deposition can be observed. Compared to participants with abnormal Aβ levels on both PET and CSF, the CSF+/PET− group has a distinctly better prognosis (Reimand et al., 2020).

Proof that Florbetapir actually predicts the accumulation of beta amyloid in the brain

Christopher M. Clark's research shows that ^{18}F-florbetapir PET image visualizations and the mean estimate of cortical uptake correlates with the abundance and amount of beta amyloid in the brain. This research is done by performing PET scans on 35 individuals who are expected to pass away soon and use it to compare to the amount of beta amyloid shown in their autopsy. This further confirms the legitimacy of ^{18}F-florbetapir in finding telltale signs of AD (Jama, 2011).

^{18}F-florbetapir and CSF Aβ42/40 ratio correlation

As stated before in the literature review, ^{18}F-florbetapir alone cannot be used as an indicator of AD. However, studies have shown that by combining both increased ^{18}F-florbetapir color uptake results with reduced CSF Aβ42/40 ratio can indicate brain Aβ formation and progression of AD.

Criterias and Requirements for Subject Population

In AD research, patients are usually recruited from clinical practice, memory clinics or nursing homes. For the purpose of this interdisciplinary research, the diagnostic criteria of our research are as follows. Firstly, we expect there to be a wide variety of ethnicities in our subject population to allow for an unbiased and holistic correlation study. This can include Caucasians, Asians, Hispanics, African Americans, Latinos, Indigenous groups and many others. Secondly, AD family history of subjects will be taken into account, ensuring that at least one preceding family member was diagnosed with AD. The subject population's age range must also fall within the range of 65 to 85 years old. This is because 65 years of age is often considered as the universal threshold to differentiate between younger-onset and older-onset AD (“What is Alzheimer's?”, n.d.,). Gender balance within the subject population must be equally distributed between males and females to ensure fairness when observing correlations and formulating conclusions.

Furthermore, subject groups will be required at least ten years of schooling, as studies indicate how people with low education levels seem to lead unhealthier lifestyles contributing to dementia or AD (American Academy of Neurology, 2007). Participants must fall within the moderate rating from the health status captured by a poly-environmental risk score called the LIBRA index. Measures of self-reported household wealth of the sample must be from middle to high tertiles, calculated by summing wealth from the total value of a respondent’s home, physical wealth, business assets, and financial assets. Research results reveal that increased dementia risk is associated with an increase in LIBRA scores, and lower wealth was associated with an increased risk of dementia (Deckers, 2019).

In addition to the selection criteria aforementioned, the presence of Apolipoprotein E4 (APOε4 gene) must also be noted in the population of samples. Numerous studies have shown that the allele is associated with greater risks of AD, along with significantly earlier onset of both AD and Parkinson’s Disease (PD) (Presence of an APOE4 allele…, n.d.). The subject must also have no traumatic brain injury (TBI) history to minimize the external factors influencing the results of this study. Varying CERAD scores, Braak scores, and NIA-Reagan criterias are to be expected within the sample population. CERAD is a 4-point scale designed to quantify the primary cognitive deficits in AD using neuritic plaques, and is often widely used to evaluate cognitive decline and progression of dementia (Fillenbaum et al., 2008). Agreement between the NIA-Reagan and CERAD criteria is moderate as research shows that the lower percentage of the NIA-Reagan criteria measured brings forth greater likelihood of AD. Whereas, lower Braak scores show lower risks (low likelihood scores ranged from I – IV; intermediate likelihood II to V). CERAD criteria revealed 41% as no AD, 13% as possible, 36% as probable and 11% as definite AD (Mufson et al., 2016). Collectively, varying values for these tests in the subject population will allow for definitive correlation to be deduced.

Classifying Data

The team’s software will require four scans as input - Amyloid PET, ^{18}F-florbetapir PET, PET tau, and CSF Aβ42/40. The prior three are graphical data, while the latter is numerical. Each set of scans is to come with additional information on their sex, age, race and the physicians’ diagnosis — whether they have AD, if so, the current stage between early and late. The graphics and numerical data from each patient will then be used to construct and train a machine learning model.

Required Python Libraries

The machine learning model requires the following Python libraries:

- os
- matplotlib
- matplotlib.pyplot
- IPython.display
- pandas
- NumPy
- PIL
- scikit-learn
- Skimage

Data Preprocessing

Before constructing and training the model using the graphical data, amyloid PET, tau PET and ^{18}F-florbetapir PET scans, all images must be resized so that they all have the same dimensions. Furthermore, all resized scans are to be converted into hexadecimal and stored in 3-dimensional NumPy arrays containing RGB values of each pixel so that the computer can process them as numerical data.

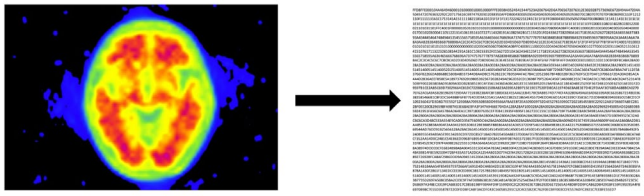

Figure 11. Partial conversion of a PET scan into hexadecimals

Model Construction and Training

The computer will plot each three-dimensional array into a 3-dimensional graph, where the x-axis, y-axis, and z-axis of the graph are the red, green, and blue values of the pixel respectively, and CSF-Aβ 42/40 numerical data into a 1-dimensional graph. Then, the program uses a kernel function to transform all the graphs into 4 multi-dimensional graphs.

The algorithm's goal is to construct two support vector classifiers from support vectors. These support vector classifiers are the maximum margin hyperplanes that best separate the data into three groups: no AD, early AD, and late AD. A support vector classifier is always one dimension lower than the dimension of the graph. Meanwhile, the support vectors are data points that are closer to the hyperplanes and influence the position and orientation of the hyperplanes, as shown in examples in Figure 3 and 4.

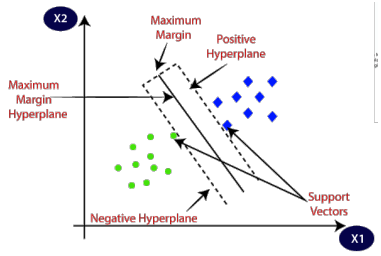

Figure 12. 2-dimensional SVM

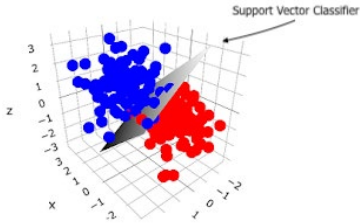

Figure 13. 3-dimensional SVM

The classified training data is used to train the model. The larger the size of the dataset, the more accurate the support vector classifier is built (Shanmukh, 2021; Joshstarmer, 2019; Morris, 2018).

Model Testing

After obtaining a trained model using SVM, example data values will be inputted to test if it gets classified correctly. The test on a 2D plane can be done by having a set of data and plotting a new one in while using the margins to see if it is on the right side of the threshold. On 3D planes, a support vector classifier, which is a plane, is used to classify which side is the point on, as shown in example in Figure 5.

Expected Output

When using the model, the input data will be plotted into the graphs and the program will calculate what class each of the 4 parameters belongs to. Each parameter will be classified into following classes: no AD, early stage AD, or late stage AD. Thus, there will be a total of 4 results.

Summary

In conclusion, our software will mark a major breakthrough in the future of AD diagnosis, potentially advancing its treatment. With relevant data and technology, patients will be able to assess their likelihood of developing AD with data from machine learning. However, further research in this field of study is crucial and strongly encouraged as AD is mostly detected at a point far from return since it is still extremely challenging for current technology to detect abnormalities until the brain scan samples are clear enough. Although there may not be any existing cures for Alzheimer's, sooner detection will prompt patients to undertake treatment at earlier stages. This could largely benefit the quality of life of both the patient and the caregivers, as they would be aware of what the disease entails and prepare themselves physically, emotionally, and financially. The future of our tool requires a greater amount of data.

Discussion

The purpose of this study was to evaluate the practicality of a machine learning model in the early detection of Alzheimer's disease (AD). In which, we analyzed how PET and CSF scans, together, could potentially provide patients greater accuracy in the prognosis and diagnosis of the disease. It was hypothesized that through implementing data of biomarkers and Support Vector Machines (SVM), contriving the machine would effectively enhance the efficiency of modern diagnostic methods for AD.

Our study consists of detecting abnormalities in the brain and comparing it to a database and giving feedback according to what was matched. We decided to conduct this study after realizing that there is a pattern in an Alzeimer's disease patient's brain scan image which could allow for earlier detection and prevention. Our approach is similar to

Europe's hospital AI analysis which matches their scans to a database. The results obtained imply that, in theory, the study could also be used to match brain scans for Alzeimer's disease. The limitation of this method is the database since the program compares images with preexisting data. Therefore, without sufficient images in the database for comparison, there is a high chance that the AI will match the wrong disease for patients. Our research could potentially be followed up by changing the detecting parameters into somewhere else in the body that is not the brain to detect other types of abnormalities.

A multidimensional panel of preclinical AD biomarkers presents the best chance for a diagnosis and prediction of progression to AD. As proposed earlier, a combination of multiple sets of information is more reliable to predict AD onset; by abnormalities in AD biomarkers (e.g., CSF, Aβ42/40 ratios, PET scans, etc.). Preclinical AD diagnostic algorithms have been proposed based on this combinatorial approach of neuroimaging, genetic testing, and CSF biomarker tests. Incorporation of biomarkers and genetic information into the preclinical AD diagnostic scheme moreover permits prediction of the in vivo physiological changes occurring in the brain before a clinical AD diagnosis. Together with advanced technology of machine learning from the suggested algorithm mentioned above, advancements and more precise predictions can be retrieved from the greater collection of biomarkers datas, leading to more reliability in measuring and tracking different disease mechanisms, cellular alterations, and disturbances in pathways "in vivo", and will pave the way for the discovery of novel drug targets.

Conclusion

Alzheimer's Disease (AD) is the most common neurological disease, with over 55 million people currently suffering from its effects. These neurodegenerative diseases affect the quality of life severely, on both the patient themselves and the people around them. Currently, there is no complete cure for these diseases, especially when by the time a proper diagnosis can be established, it would be well past the point of no return due to the vague signs and symptoms presented in early-stage patients, often going unnoticed by the human eye.

To address this issue, the team proposes to develop a machine learning model to diagnose Alzheimer's disease using scans of Amyloid PET, Florbetapir-18 PET, PET tau and Amyloid-beta 42/4 CSF results to train up. The model uses Support Vector Machine to classify each value of scans between the negative, early, or late stages of AD.

In the application of our proposal, the team aims to gather data from different parts of the world. This will increase the overall accuracy and decrease the duration of diagnosis. The model, with all four biomarkers included, will efficiently reduce doctors' workload, and increase the accuracy of diagnosis. Patients will be aware of the possibilities ahead of time and take precautions in case of future deterioration, allowing them a higher chance of recovery.

However, lack of PET scanners as well as CSF in Thailand results in scarcity of data for constructing and training and testing the model. Furthermore, existing samples cannot be altered to create more samples as it will affect the accuracy of the prediction. Therefore, if we are to develop the model using the current amount of data, the software might not be accurate enough to be used.

Thus, we expect to develop this model in the future when CSF and PET scanners are more common and data becomes abundant in Thailand or if we are able to collect sufficient data from abroad.

Limitations

The limitations that may come with the application of our method of diagnosis mainly lies with the availability of data. Due to the method relying on multiple sets of data, majority of which are types of PET scans, and said types of scans usually being not readily available in addition to being an extremely expensive form of imaging at that can pose to be a challenge, potentially resulting in the accessibility of our proposed method to be be limited- at least until PET scans would become more on hand to the general population. Furthermore, while the use of a ML algorithm trained up using Supervised Learning as a tool for detection could certainly be an advantage, this only came to be with the

copious amounts of data which hones this accuracy, thus also being a potential avenue for a shortcoming in the case where there is lack of said data, affecting the accuracy of the established prediction. Finally, while with our study we hope to be able to increase the precision and hasten the establishment of AD diagnosis, the knowledge of neuropathological processes in AD and normal aging are not entirely understood especially in elderly populations where similar clinical pictures might have different neurobiological backgrounds. Thus, we would at best, be able to come to a conclusion at the theoretical possibility of a patient contracting AD, and not a concrete ultimatum as one would desire.

Acknowledgements

First and foremost, we would like to express our utmost appreciation and gratitude towards Dr. Teerakiat Jareonsettasin, co-founder of The Newton Sixth Form School and former Minister of Education of Thailand. Other than his inspiring role as our Headmaster, his close guidance and tutelage has been an indispensable influence in refining a critical scholarly mindset imperative to flourish in the world of academia. Without such profound insight and grounding hands on approach in implementing his ideas in the school, there would have been no research class, the birth of this group, nor would have this study been conducted and completed. Furthermore, we would also like to thank Kanokkorn Pitayakornpakdee, our research mentor for teaching us the basics of doing a literature review and offering valuable guidance throughout. Her advice has been invaluable in completion of this project.

We are particularly obliged for the assistance given by Dr. Dran Rukarcheep, Dr. Yossamon Jiratthawong and Akarat Hanmatheekuna who, with their extensive knowledge in the field, enlightened us with the possibilities and the limitations of our project. Without their guidance, we would not have been as successful as we have been in our work.

Finally, we would like to thank our parents for supporting us along the way and supporting our dreams of becoming involved in the fields of medicine, computer science, and engineering. Without their financial and emotional support, we would not be the people we are today, owing them our eternal gratitude and debt.

References

Aβ40 assays. Quanterix. (2022). https://www.quanterix.com/simoa-assay-kits/ab40/

Aβ42-in-2: AB42 inhibitor: Medchemexpress. MedchemExpress.com. (n.d.). https://www.medchemexpress.com/a%CE%B242-in-2.html

Amyloid β protein: AB40 inhibits AB42 ... - pubs.acs.org. (n.d.).https://doi.org/10.1021/ja8092604

Applications of amyloid, TAU, and neuroinflammation pet ... (n.d.). https://doi.org/10.1002/hbm.24782

BioMed Central. (2018). Secondary prevention of Alzheimer's dementia: neuroimaging contributions - Alzheimer's Research & Therapy. https://doi.org/10.1186/s13195-018-0438-z

Brain : a journal of neurology, 2017 Mar ; Relationships between flortaucipir PET tau binding and amyloid burden, clinical diagnosis, age and cognition. https://doi.org/10.1093%2Fbrain%2Faww334

Brain scans. Stanford Health Care (SHC) - Stanford Medical Center. (2017). https://stanfordhealthcare.org/medical-conditions/brain-and-nerves/dementia/diagnosis/brain-scans.html

Campbell MR;Ashrafzadeh-Kian S;Petersen RC;Mielke MM;Syrjanen JA;van Harten AC;Lowe VJ;Jack CR;Bornhorst JA;Algeciras-Schimnich A; (n.d.). *P-Tau/AB42 and AB42/40 ratios in CSF are equally predictive of amyloid PET status*. Alzheimer's & dementia (Amsterdam, Netherlands). https://doi.org/10.1002/dad2.12190

Cortés-Blanco A;Prieto-Yerro C;Martínez-Lázaro R;Zamora J;Jiménez-Huete A;Haberkamp M;Pohly J;Enzmann H;Zinserling J;Strassmann V;Broich K; (n.d.). *Florbetapir (18F) for brain amyloid positron emission tomography: Highlights on the European Marketing Approval*. Alzheimer's & dementia: the journal of the Alzheimer's Association. https://doi.org/10.1016/j.jalz.2013.09.007

Cukalevski, R., Yang, X., Meisl, G., Weininger, U., Bernfur, K., Frohm, B., Knowles, T. P. J., & Linse, S. (2015). *The AB40 and AB42 peptides self-assemble into separate homomolecular fibrils in binary mixtures but cross-react during primary nucleation*. Chemical Science. https://doi.org/10.1039/c4sc02517b

Deckers, K., Cadar, D., van Boxtel, M. P. J., Verhey, F. R. J., Steptoe, A., & Köhler, S. (2019). Modifiable risk factors explain socioeconomic inequalities in dementia risk: Evidence from a population-based prospective cohort study. Journal of Alzheimer's Disease. https://doi.org/10.3233/JAD-190541

Doraiswamy, P., Sperling, R., Johnson, K., Reiman, E., Wong, T., Sabbagh, M., Sadowsky, C., Fleisher, A., Carpenter, A., Joshi, A., Lu, M., Grundman, M., Mintun, M., Skovronsky, D., Pontecorvo, M., Ranjan Marwan Geoffrey Richard F Mildred V Beth Emmie Gus Duara Sabbagh Lawrence Ahern Holub Farmer Safirste, Duara, R., Ahern, G. L., Holub, R., … Williamson, J. (1970). *[pdf] florbetapir F 18 amyloid PET and 36-month cognitive decline:a prospective Multicenter Study: Semantic scholar*. undefined. https://doi.org/10.1038/mp.2014.9

Fillenbaum, G. G., van Belle, G., Morris, J. C., Mohs, R. C., Mirra, S. S., Davis, P. C., Tariot, P. N., Silverman, J. M., Clark, C. M., Welsh-Bohmer, K. A., & Heyman, A. (2008). Consortium to establish a registry for Alzheimer's disease (CERAD): The First Twenty Years. Alzheimer's & dementia : the journal of the Alzheimer's Association. https://doi.org/10.1016%2Fj.jalz.2007.08.005

Hansson, O., Lehmann, S., Otto, M., Zetterberg, H., & Lewczuk, P. (2019). *Advantages and disadvantages of the use of the CSF amyloid β (AB) 42/40 ratio in the diagnosis of alzheimer's disease - alzheimer's research & therapy*. BioMed Central. https://doi.org/10.1186/s13195-019-0485-0

JAMA. (2011). Use of florbetapir-PET for imaging beta-amyloid pathology. https://doi.org/10.1001%2Fjama.2010.2008

James, O. G., Doraiswamy, P. M., & Borges-Neto, S. (n.d.). *PET imaging of tau pathology in alzheimer's disease and Tauopathies*. Frontiers. https://doi.org/10.3389/fneur.2015.00038

K;, O. N. Y. (n.d.). *Florbetapir (18F), a pet imaging agent that binds to amyloid plaques for the potential detection of alzheimer's disease*. IDrugs : the investigational drugs journal. https://pubmed.ncbi.nlm.nih.gov/21154149/

Lois, C., Gonzalez, I., Johnson, K. A., & Price, J. C. (2019). *PET imaging of TAU Protein Targets: A Methodology Perspective*. Brain imaging and behavior. https://doi.org/10.1007%2Fs11682-018-9847-7

Lundeen, T. F., Author AffiliationsFrom the Department of Medical Imaging, A, W., VL, V., Al, E., JC, M., J, S., WE, K., I, D., CA, M., MA, M., KA, J., S, M., CC, R., Jr, L. R., Healthcare, G. E., Lilly, E., Piramal, DF, W., … Sugi, M. D. (2018). Signs and Artifacts in Amyloid PET. https://doi.org/10.1148/rg.2018180160

Machine learning. International Business Machines (IBM Cloud Education). (2020). https://www.ibm.com/cloud/learn/machine-learning

Mattsson, N., Insel, P. S., Landau, S., Jagust, W., Donohue, M., Shaw, L. M., Trojanowski, J. Q., Zetterberg, H., Blennow, K., Weiner, M., & Alzheimer's Disease Neuroimaging Initiative. (2014). *Diagnostic accuracy of CSF AB42 and Florbetapir Pet for Alzheimer's disease*. Annals of clinical and translational neurology. https://doi.org/10.1002%2Facn3.81

Mufson, E. J., Malek-Ahmadi, M., Perez, S. E., & Chen, K. (2016). Braak staging, plaque pathology, and APOE status in elderly persons without cognitive impairment. Neurobiology of aging. https://doi.org/10.1016%2Fj.neurobiolaging.2015.10.012

National Institute on Aging (NIA). (2021). *What Is Dementia? Symptoms, Types, And Diagnosis.* https://www.nia.nih.gov/health/what-is-dementia

Neurology. (2020). Association of amyloid-β CSF/PET discordance and tau load 5 years later. https://doi.org/10.1212/WNL.0000000000010739

Okamura N;Harada R;Furumoto S;Arai H;Yanai K;Kudo Y; (n.d.). *Tau Pet Imaging in Alzheimer's disease*. Current neurology and neuroscience reports. https://doi.org/10.1007/s11910-014-0500-6

OUP Academic. (2017). Association between tau deposition and antecedent amyloid-β accumulation rates in normal and early symptomatic individuals. https://doi.org/10.1093/brain/awx046

Pankratz N;Byder L;Halter C;Rudolph A;Shults CW;Conneally PM;Foroud T;Nichols WC; (n.d.). Presence of an APOE4 allele results in significantly earlier onset of parkinson's disease and a higher risk with dementia. Movement disorders : official journal of the Movement Disorder Society.https://doi.org/10.1002/mds.20663

Pontecorvo MJ;Devous MD;Navitsky M;Lu M;Salloway S;Schaerf FW;Jennings D;Arora AK;McGeehan A;Lim NC;Xiong H;Joshi AD;Siderowf A;Mintun MA; ; (n.d.). *Relationships between Flortaucipir Pet Tau binding and amyloid burden, clinical diagnosis, age and cognition*. Brain : a journal of neurology. https://doi.org/10.1093/brain/aww334

Presence of an APOE4 allele results in significantly ... (n.d.). https://doi.org/10.1002/mds.20663

Robertson, S. (2018). *What are amyloid plaques?* News. https://www.news-medical.net/health/What-are-Amyloid-Plaques.aspx#:~:text=Amyloid%20plaques%20form%20when%20pieces,enzymes%20to%20produce%20beta%2D amyloid

Rpubs. (N.D.). RPubs - Image classification using SVM. https://rpubs.com/Sharon_1684/454441

Sauvée M;DidierLaurent G;Latarche C;Escanyé MC;Olivier JL;Malaplate-Armand C; (n.d.). *Additional use of AB_{42}/AB_{40} ratio with cerebrospinal fluid biomarkers P-Tau and AB_{42} increases the level of evidence of alzheimer's disease pathophysiological process in routine practice*. Journal of Alzheimer's disease : JAD. https://doi.org/10.3233/JAD-131838

ScienceDaily. (2007). Low education level linked to Alzheimer's, study shows. ScienceDaily. https://www.sciencedaily.com/releases/2007/10/071001172855.htm

Shanmukh, V. (2021). *Image classification using machine learning-support vector machine(SVM)*. Medium.https://medium.com/analytics-vidhya/image-classification-using-machine-learning-support-vector-machine-svm-dc7a0ec92e01

Shorena Janelidze, P. D. (2021). *Head-to-head comparison of 8 plasma amyloid-β 42/40 assays in Alzheimer disease*. JAMA Neurology. https://doi.org/10.1001/jamaneurol.2021.3180

The First Twenty Years. Alzheimer's & dementia : the journal of the Alzheimer's Association. https://doi.org/10.1016%2Fj.jalz.2007.08.005

JAMA. (2011). Use of florbetapir-PET for imaging beta-amyloid pathology. https://doi.org/10.1001/jama.2010.2008

Therriault, J., Benedet, A. L., Pascoal, T. A., Mathotaarachchi, S., Savard, M., Chamoun, M., Thomas, E., Kang, M. S., Lussier, F., Tissot, C., Soucy, J.-P., Massarweh, G., Rej, S., Saha-Chaudhuri, P., Poirier, J., Gauthier, S., & Rosa-Neto, P. (2020). *APOEε4 potentiates the relationship between amyloid-β and tau pathologies*. Nature News. https://doi.org/10.1038/s41380-020-0688-6

Topalovic, M., Burgel, P., Das, N., Daenen, M. (2019). *Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests.* https://www.researchgate.net/publication/334587622_Artificial_intelligence_outperforms_pulmonologists_in_the_interpretation_of_pulmonary_function_tests

What is Alzheimer's? (n.d.). https://www.alz.org/alzheimers-dementia/what-is-alzheimers#:~:text=Alzheimer's%20is%20not%20a%20normal,to%20as%20early%2Donset%20Alzheimer's

World Health Organization. (2021). *Dementia*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/dementia

Yang, L., Author AffiliationsFrom the Division of Medical Imaging Products (L.Y., Steele, N. Z. R., Others, M. W. W. and, O'Brien, P. L. A. and K. L., L. J. Abu-Raddad and Others, Others, E. T. and, & Others, N. A. and. (1970). *Brain amyloid imaging - FDA approval of Florbetapir F18 Injection: Nejm*. New England Journal of Medicine. https://doi.org/10.1056/nejmp1208061

YouTube. (2019). *Support Vector Machines Part 1 (of 3): Main Ideas!!! YouTube*. https://www.youtube.com/watch?v=efR1C6CvhmE