IT3389 – Applied AI Project

Diploma in Applied AI and Analytics (DAAA)

School of Information Technology

<u>13 February 2024</u>

Nasrullah, Jia Jun, Skye

Table of Contents

Final Report

Module Group: 2 and 3

Project Group Number: Group 4

Team Name: Group 4

Background:

Singapore's growing elderly population brings with it a significant rise in dementia cases, posing a major challenge for accurate diagnosis, effective care, and maintaining the quality of life for patients and their families. Currently, the absence of a cure for Alzheimer's and the often subtle onset of symptoms lead to delayed diagnoses, exacerbating the difficulties in managing this condition.

Recognizing this critical need, we, a specialized team in AI and healthcare technology, have been entrusted by Synapxe, a leading health-tech agency in Singapore, and SingHealth to develop a comprehensive solution. Our project focuses on two key areas: early detection and effective home care support for stabilized dementia patients.

Early detection is crucial for managing the progression of dementia and improving patient outcomes. We aim to equip both patients and caregivers with tools for continuous monitoring, empowering them to identify potential changes and seek timely intervention. Our solution will also address the critical need for improved home care support, particularly for stabilized patients where family members often lack the required knowledge and skills. Our approach leverages the power of artificial intelligence, incorporating MLOps for efficient deployment and management of machine learning models. Additionally, we will adhere to the CRISP-DM methodology to ensure a structured and transparent development process for our AI solution.

This project goes beyond just addressing the immediate needs of dementia patients and their loved ones. It aligns with Singapore's broader healthcare objectives, serving as a potential model for improved dementia care within the community. We are committed to delivering a solution that not only benefits individuals but also contributes to the nation's healthcare aspirations.

Objectives:

MRI Image Analysis (Jia Jun):

The objective of implementing MRI Image Analysis is grounded in addressing critical gaps in the healthcare sector, particularly in the realm of diagnostic accuracy and efficiency. The need arises from the increasing complexity of medical conditions that demand precise and early detection for effective treatment. By harnessing advanced machine learning algorithms, MRI Image Analysis holds the potential to revolutionize diagnostic processes, providing clinicians with a more nuanced understanding of intricate patterns within medical images. This technology can significantly enhance the speed and accuracy of detecting abnormalities, enabling early intervention and personalized treatment plans. Moreover, the implementation of MRI Image Analysis can contribute to reducing the workload on healthcare professionals, streamlining the diagnostic workflow, and ultimately improving patient outcomes. The integration of artificial intelligence in this context not only augments the capabilities of medical practitioners but also paves the way for more efficient and cost-effective healthcare solutions. Achieving this objective involves meticulous data preprocessing, feature extraction, and the deployment of specialized neural networks tailored for medical image analysis, thereby ensuring a robust and reliable framework for advancing healthcare diagnostics.

Dementia Risk Prediction (Nas):

The objective of implementing Dementia Risk Prediction through machine learning is to enhance patient care and optimize treatment strategies in the healthcare sector. This initiative is crucial due to the growing complexity of patient profiles, which requires personalized and informed prediction decisions.

By leveraging a comprehensive dataset comprising patient demographics, medical history, lifestyle factors, and relevant health indicators, the Dementia Risk Prediction model aims to analyze and identify patterns associated with the risk of developing dementia. The goal is to provide healthcare professionals with accurate predictions, enabling early intervention and personalized care plans for patients at risk. This approach not only facilitates proactive management of dementia but also contributes to reducing healthcare costs and improving patient outcomes.

Through the integration of machine learning algorithms, this objective seeks to leverage data-driven insights for more effective risk assessment and intervention strategies. Implementation involves rigorous model training, validation, and continuous refinement to ensure the reliability and adaptability of the Dementia Risk Prediction system in diverse healthcare settings. This initiative represents a significant advancement in proactive healthcare management, enhancing patient well-being and fostering a more efficient healthcare system.

Prescription Recommendation (Skye):

The objective of implementing Prescription Recommendation through machine learning is to optimize treatment strategies and enhance patient care in the healthcare sector. This initiative is essential due to the increasing complexity of patient profiles, necessitating personalized and informed prescription decisions. By leveraging a diverse dataset comprising patient demographics, medical history, lifestyle factors, and relevant health indicators, the Prescription Recommendation model aims to analyze and identify patterns associated with effective treatment outcomes.

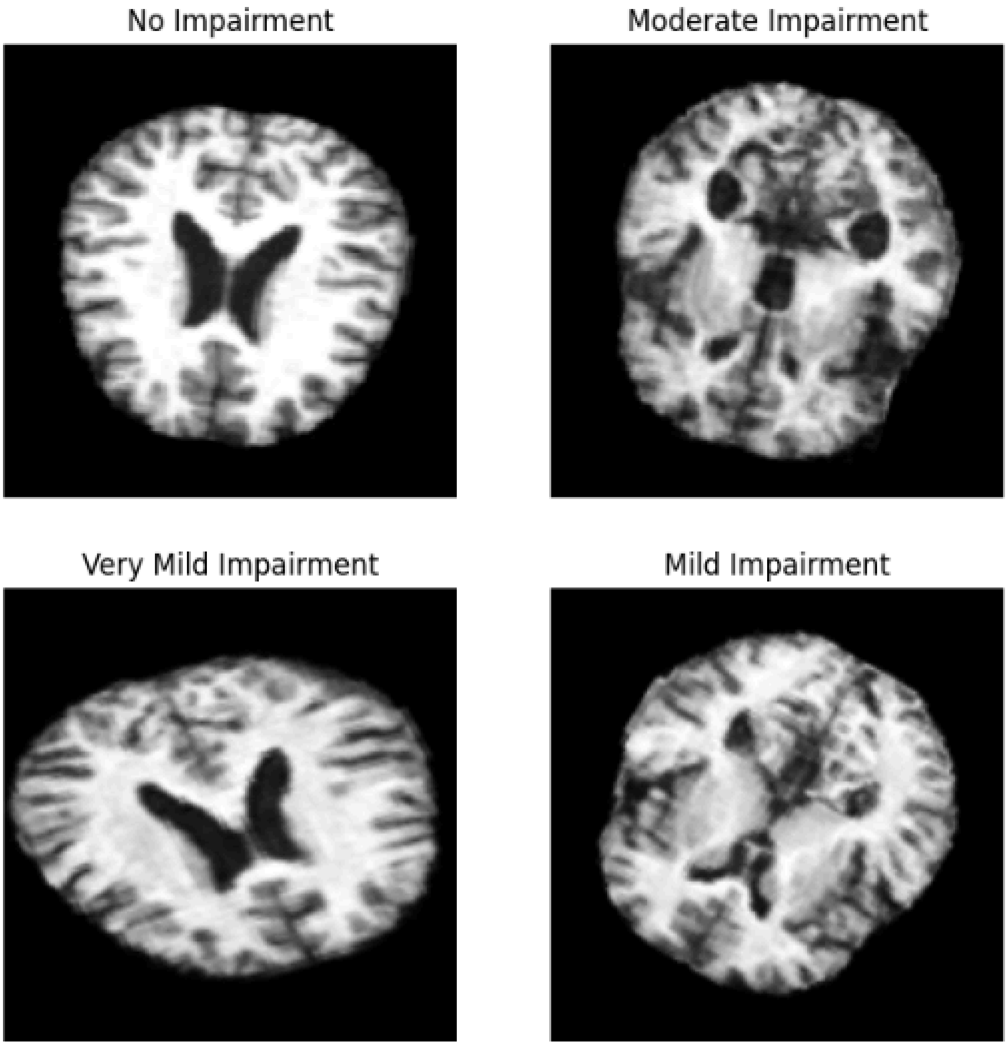
The goal is to provide healthcare professionals with tailored prescription recommendations, improving the precision and relevance of medication choices for individual patients. This approach streamlines the prescription decision-making process and contributes to minimizing adverse drug interactions and potential side effects. Through the integration of machine learning algorithms, this objective harnesses data-driven insights for more effective and personalized treatment plans.

Implementation involves rigorous model training, validation, and continuous refinement to ensure the adaptability and reliability of the Prescription Recommendation system in diverse healthcare scenarios. This initiative represents a pivotal step towards optimizing healthcare practices, promoting patient-centric care, and fostering a more efficient and responsive healthcare system.

Solution & Insights

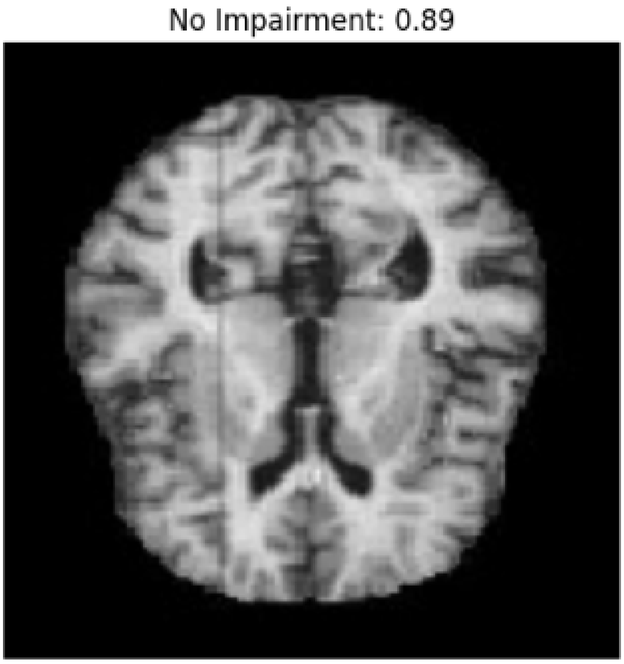
MRI Image Analysis Solution (Jia Jun):

The implementation of MRI Image Analysis offers a transformative solution to the challenges in healthcare diagnostics, particularly in the field of medical imaging. By employing state-of-the-art machine learning algorithms, our solution aims to enhance the accuracy and efficiency of detecting subtle abnormalities within MRI scans. The utilization of convolutional neural networks (CNNs) allows for a comprehensive analysis of intricate structures, enabling early detection of medical conditions. This approach not only expedites the diagnostic process but also facilitates timely intervention and personalized treatment plans. Furthermore, the integration of MLOps and adherence to the CRISP-DM methodology ensures the robust deployment and management of machine learning models, fostering a scalable and sustainable solution. The impact of our MRI Image Analysis solution extends beyond improved diagnostic capabilities, promising to reduce the burden on healthcare professionals, enhance patient outcomes, and contribute to more effective and cost-efficient healthcare practices.



These are some sample data from the dataset.

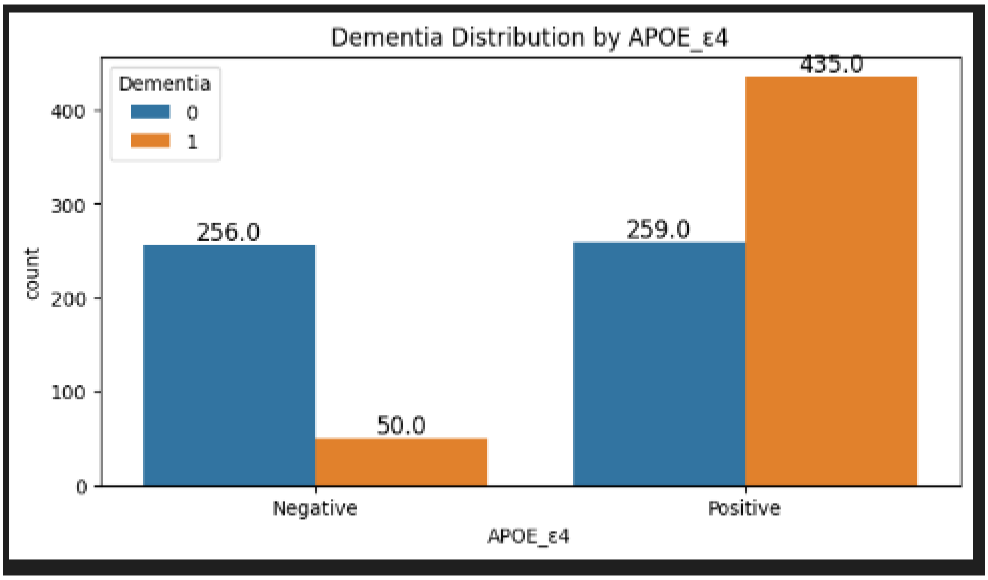
```
1/1 [==============================] - 0s 48ms/step
```



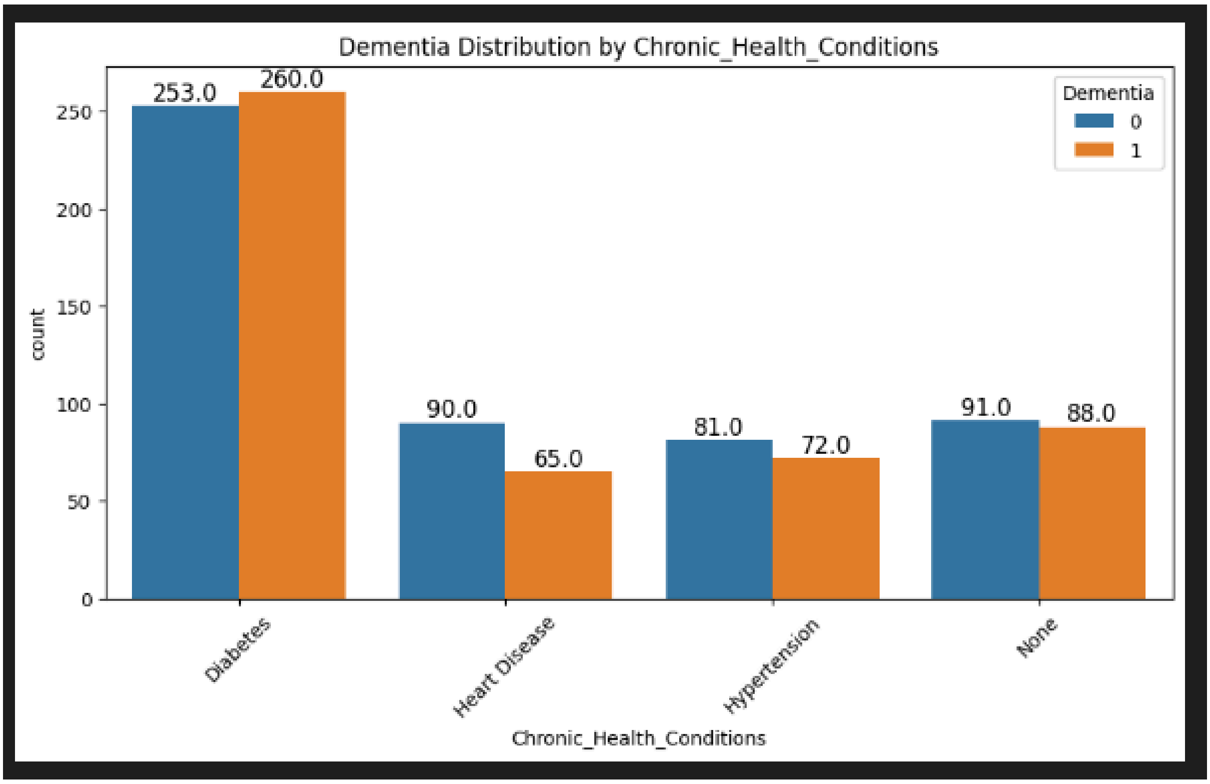
After creating a model to detect MRI scans, this is the final result giving a percentage of the predicted class and the label for the predicted class.

Dementia Prediction (Nas):

Our aim with Dementia Risk Prediction is to enhance patient care by leveraging machine learning to optimize treatment strategies. With patient profiles becoming increasingly complex, personalized prediction decisions are essential. By analyzing comprehensive datasets including demographics, medical history, and lifestyle factors, our model identifies patterns associated with dementia risk. The goal is to provide accurate predictions for early intervention and personalized care plans, ultimately reducing healthcare costs and improving patient outcomes. Through rigorous model training and validation, we ensure the reliability and adaptability of our system in diverse healthcare settings.

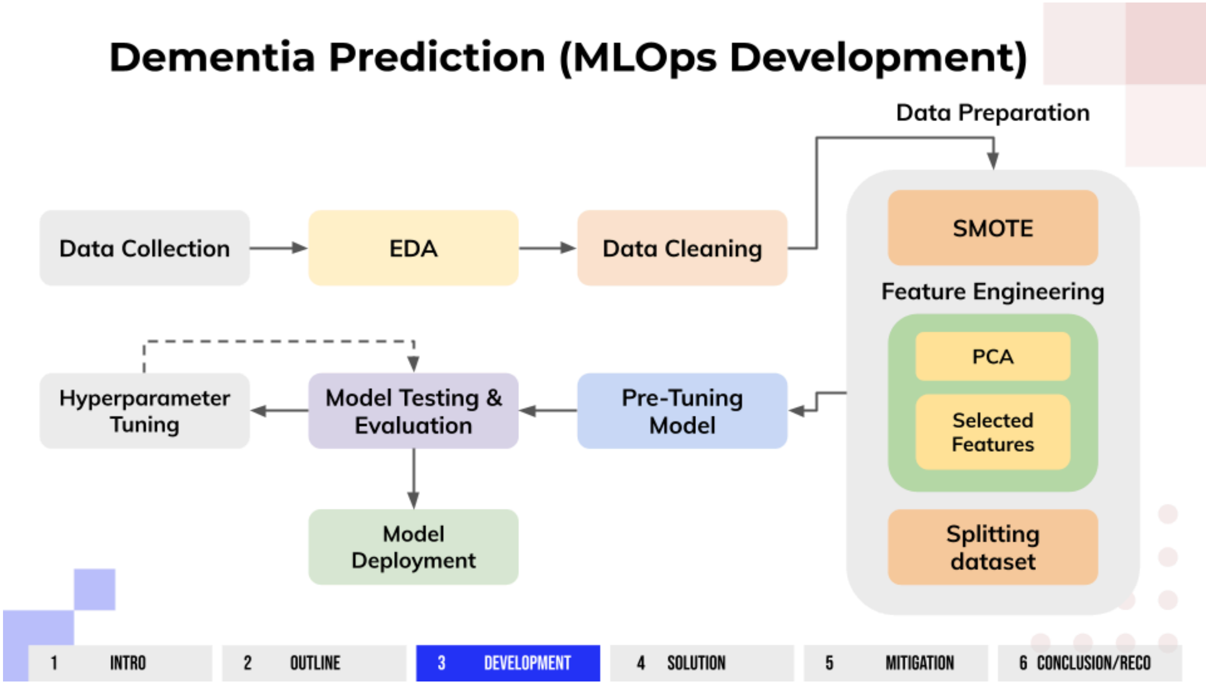


Patients who have the gene APOE_E4 are prone to get a dementia diagnosis later in life.



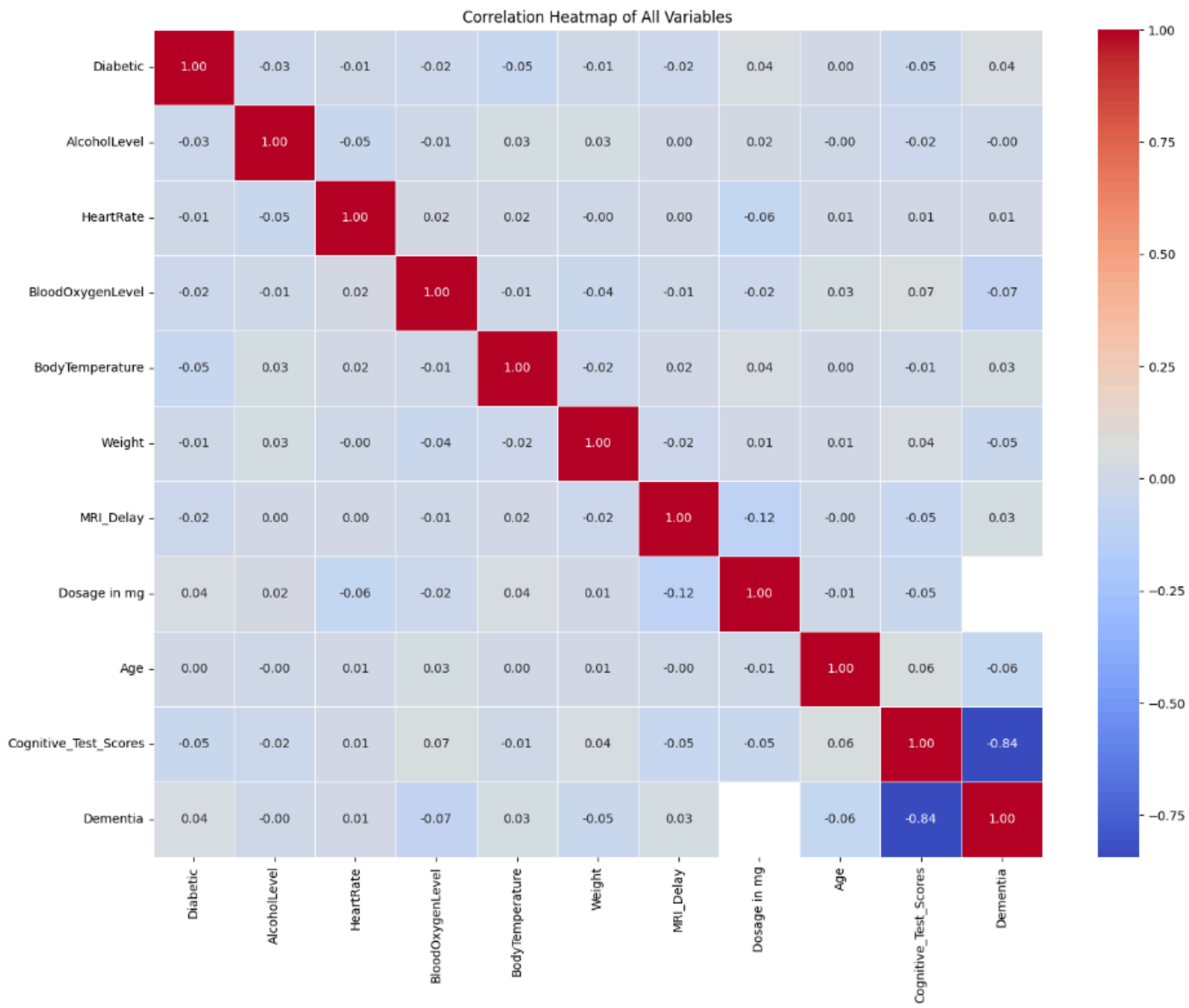
Patients who are diagnosed with Diabetes are also are prone to get dementia.

Implemented Streamlined MLOps Workflow:



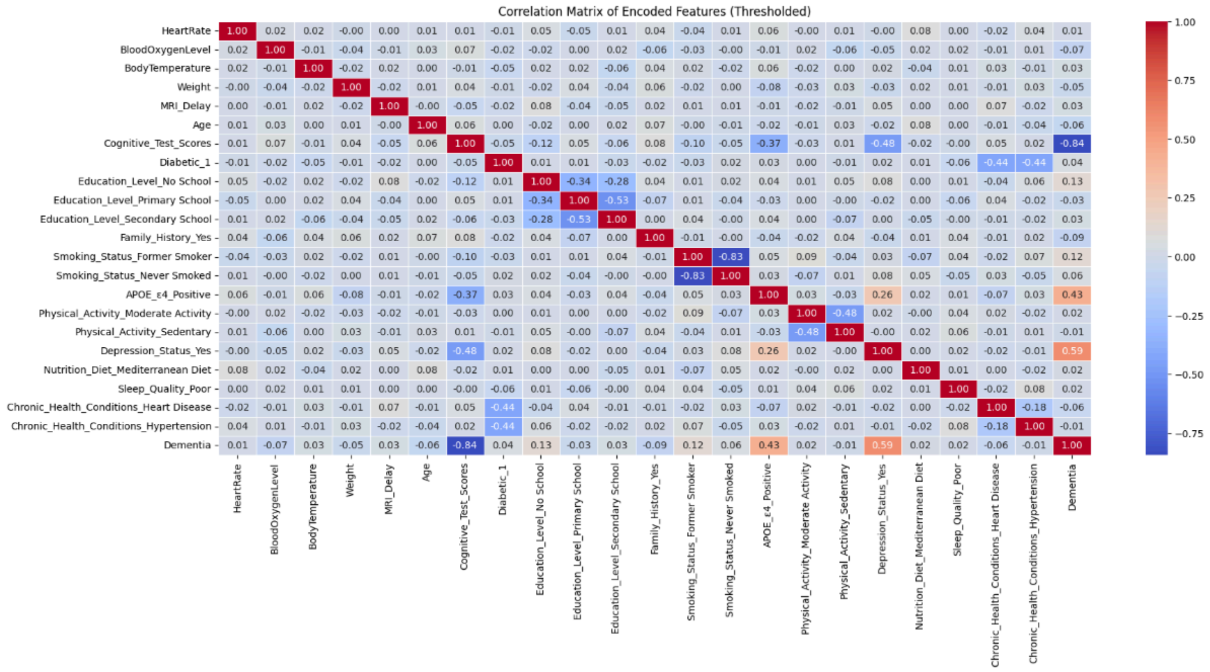
Before one-hot encoding:

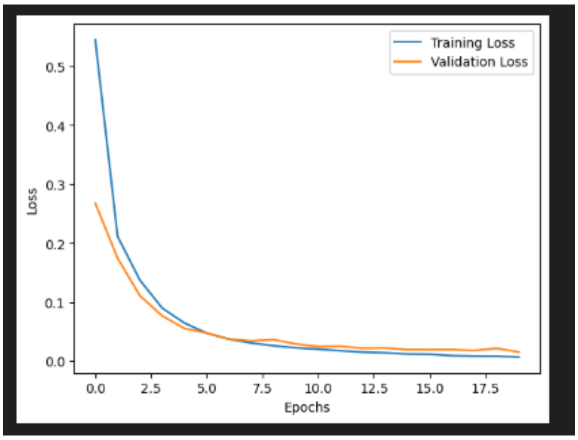
We can see that in this feature correlation heatmap, variables like [Cognitive_Test_Scores] is highly correlated with the target variable [Dementia].



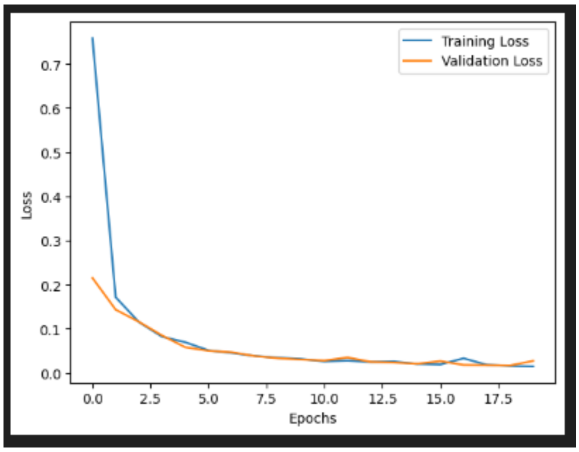
After one-hot encoding:

We can see many features that we created from the categorical variables, and we can see many more features that is highly correlated with the target variable [Dementia]. Variables like [Cognitive_Test_Scores], [APOE_E4_Positive] and [Depression_Status_Yes], are highly correlated to the target variable.





Best Epoch: 16
Validation Accuracy at Best Epoch: 0.9944
PCA Model Performance:
Data Accuracy: 0.9944

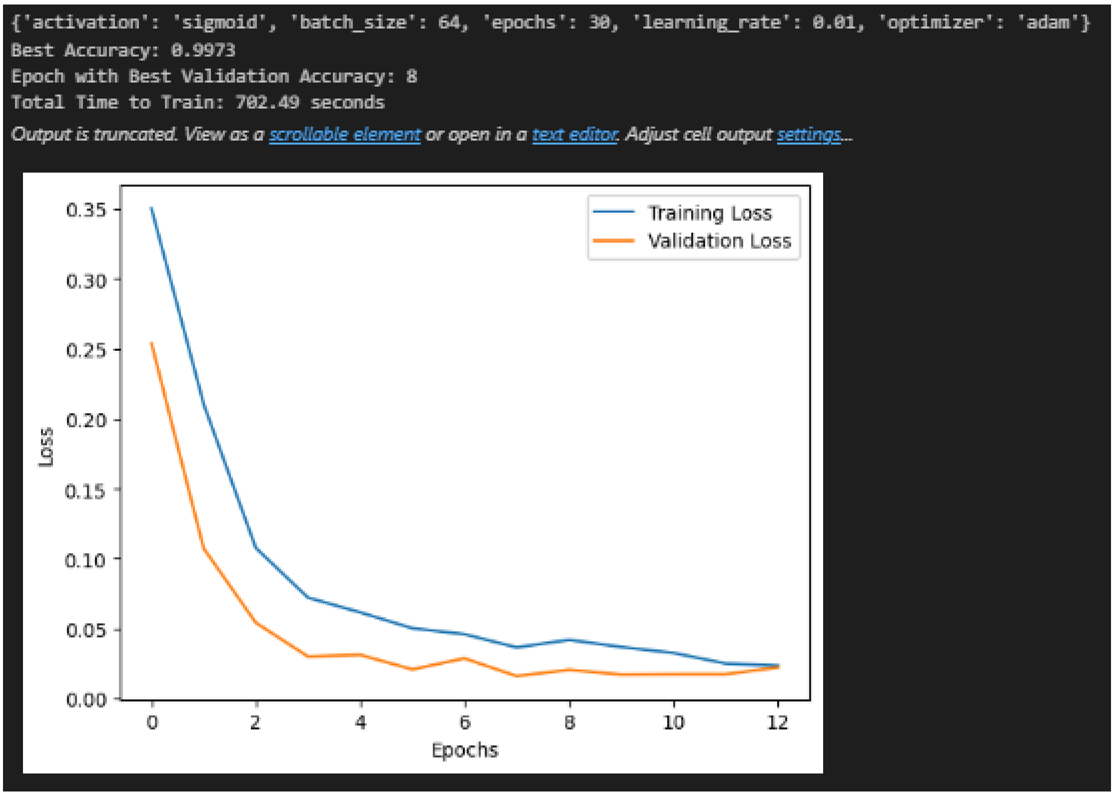


Best Epoch: 18
Validation Accuracy at Best Epoch: 0.9938
Selected Features Model Performance:
Data Accuracy: 0.9938

Before hyper parameter tuning, I trained the data on a simple deep learning model. It will return me the best model with the best results which is the PCA Model.

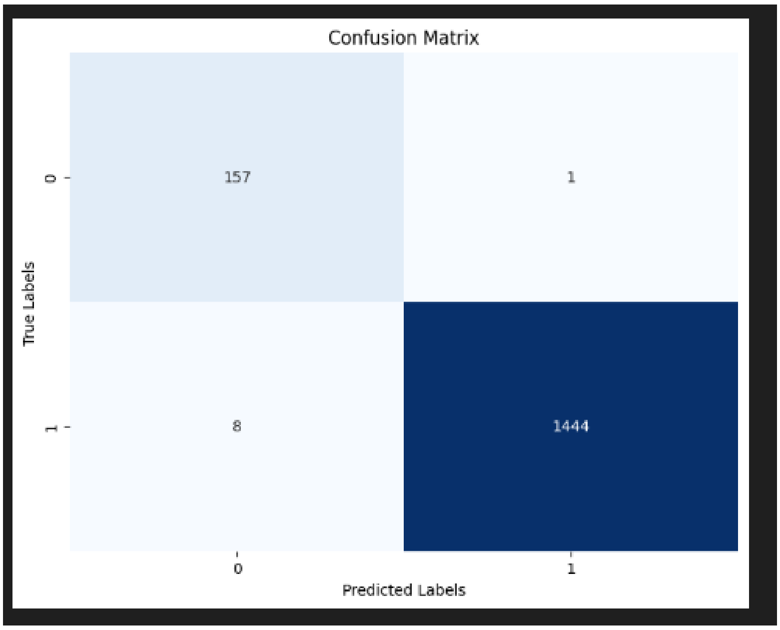
Hyperparameter Tuning

I automated the hyperparameter tuning using GridSearch and RandomSearchCV such that it returns me the best parameters to use. This model also reduce overfitting and displays the perfect fit on epoch 8.

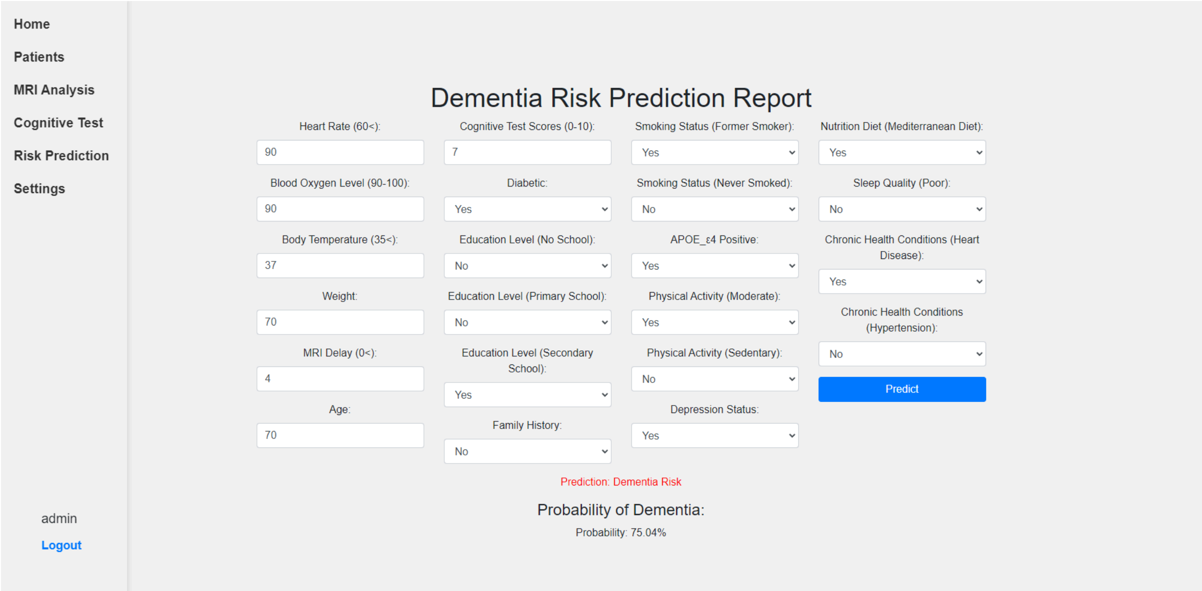


Correlation matrix:

Correct Predictions: 1600 out of 1610



Deployed to Web App:



Prescription Recommendation (Skye):

Our goal with Prescription Recommendation is to optimize treatment strategies and enhance patient care using machine learning. With patient profiles becoming more complex, personalized prescription decisions are crucial. By analyzing diverse datasets, including demographics and medical history, our model identifies patterns associated with effective treatment outcomes. This streamlines the prescription process, minimizes adverse drug interactions, and improves medication choices for patients. Through continuous refinement, we ensure the adaptability and reliability of our system in diverse healthcare scenarios, promoting patient-centric care and a more efficient healthcare system.

Supporting Data:

Data Source:

MRI Image Detection:

https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images

Prescription Recommendation:

https://www.kaggle.com/datasets/kaggler2412/dementia-patient-health-and-prescriptions-dataset

Content of data:

- **MRI Image Detection:**

  Our dataset consists of MRI images categorized into four classes: Mild Impairment, Moderate Impairment, No Impairment, and Very Mild Impairment. The goal is to develop a highly accurate model for predicting the stages of Alzheimer's disease. By leveraging this diverse dataset, our aim is to enable AI models to effectively classify different stages of Alzheimer's based on MRI imagery, leading to more precise and early diagnoses and ultimately improving healthcare outcomes for affected individuals.

- **Prescription Recommendation:**

  This dataset includes a wide range of health-related features, providing insights into factors influencing dementia presence and progression. It covers binary indicators for diabetes and dementia status, continuous measures like alcohol consumption and cognitive test scores, demographic aspects, genetic information (APOE ε4 allele), lifestyle factors, medication history, and chronic health conditions. Sourced from reputable channels and healthcare professionals, this dataset is a valuable resource for in-depth analysis, allowing researchers to explore complex relationships surrounding dementia onset and progression.

Conclusion & Recommendations

In our AI system's development, ethics and transparency have been paramount. We've meticulously sourced data from credible sources and conducted proactive risk assessments, documenting potential ethical challenges comprehensively. Our commitment to transparency extends to clear communication about our system's objectives, methodologies, and implications with stakeholders.

Recommendation (Nas)

To ensure fairness and mitigate biases, we've rigorously examined our data and model development processes, implementing techniques for bias reduction. Ethical considerations guide our decision-making, prioritizing equitable outcomes. Going forward, we'll maintain our ethical focus, monitoring system performance and addressing potential challenges to uphold high ethical standards and deliver impactful, ethical AI solutions in healthcare.

Recommendation (Jia Jun)

These are some recommendations for further training as my tuned model did not perform better.

Longer Training: Training the tuned model for more epochs may help in improving its performance. It's essential to monitor the model for overfitting and adjust the training process accordingly.

Regularization Techniques: Implementing dropout, L2 regularization, or exploring different architectures might prevent overfitting and improve model generalization.

Data Augmentation: To enhance the model's ability to generalize and learn from a more diverse dataset, data augmentation strategies could be beneficial.

Early Stopping: Utilize early stopping to halt training when the validation performance no longer improves, preventing overfitting and saving computational resources.

Given the constraints of time, our training was limited. Future work should consider extending the training duration while closely monitoring for improvements in performance metrics.

Recommendation (Skye)

Model Tuning: As I did not really manage to tune my model, the advantage of model tuning lies in its capability to optimize and enhance the performance of machine learning models, allowing for improved accuracy and generalization on diverse datasets.

Cross-Validation: Implement cross-validation to get a more robust estimate of my model’s performance.