Project Proposal:

Alzheimer's Disease Image Classification using CNN Models

Business Understanding

1. Problem Statement:

The objective of this project is to develop a deep learning model using Convolutional Neural Networks (CNN) to accurately classify brain MRI images of Alzheimer's disease. We have four classes to identify: MildDemented, ModerateDemented, NonDemented, and VeryMildDemented. By accurately identifying these classes, we can assist in the early detection and diagnosis of Alzheimer's disease, which is a neurodegenerative disorder that affects millions of people worldwide.

2. Why this topic?

Alzheimer's disease is a significant and increasing public health crisis with a huge impact on patients, families, and healthcare systems. Early detection of Alzheimer's can lead to better control and treatment, potentially slowing down its progression and enhancing the quality of life for affected individuals.

3. Target Audience:

The project applies to the healthcare and medical research domain. The target audience for this project includes healthcare professionals, neurologists, medical researchers, and data scientists working on Alzheimer's disease-related studies. Additionally, medical institutions and hospitals looking to incorporate advanced technology for early diagnosis of Alzheimer's could also benefit from the outcomes of this project.

4. Impact on the Real World:

If successfully developed and deployed, the CNN model for Alzheimer's disease identification could have several real-world impacts:

- ***Early Detection***: Early detection allows for intervention and treatment, potentially slowing down the advancement of the disease and improving patient outcomes.
- ***Reduced Healthcare Costs:*** Early detection and accurate diagnosis can lead to more targeted treatment plans, potentially reducing long-term healthcare costs.
- ***Research Advancement:*** The model can assist researchers by providing a tool for large-scale screening and analysis, leading to a deeper understanding of Alzheimer's disease and potential areas for further research.

5. Pre-existing Projects/Research:

The field of Alzheimer's disease identification using image classification has seen significant research and progress in recent years. Several papers and projects have explored the use of CNN models and other deep-learning techniques for this purpose. Relevant work includes:

- Raza N, Naseer A, Tamoor M, Zafar K. Alzheimer Disease Classification through Transfer Learning Approach. Diagnostics (Basel). 2023 Feb 20;13(4):801. doi: 10.3390/diagnostics13040801. PMID: 36832292; PMCID: PMC9955379.

6. Motivation:

As a data science student, the motivation for this project is to apply groundbreaking deep learning techniques to address a crucial real-world problem. Alzheimer's disease is devastating, and the possibility to contribute to early detection and diagnosis through image classification is both intellectually thrilling and socially impactful. By leveraging TensorFlow and Keras to develop CNN models, I aim to showcase the power of AI and data-driven approaches in the medical domain.

Data Understanding

1. Source of Raw Data:

The raw data for this project was obtained from /huggingface.co. Hugging Face is a popular platform for accessing and sharing various natural language processing (NLP) and computer vision datasets, including datasets for image classification tasks. Hugging Face datasets are typically publicly available for research and academic purposes.

2. Building on Previous Work:

Previous work on the same dataset exists, but this project will be from scratch, it will be up to me, to make use of it.

Data Preparation

1. Form of Data Storage:

The data is stored in the form of JPEG (.jpg) images. Each image represents a brain MRI scan, and they are organized into folders based on their respective classes (MildDemented, ModerateDemented, NonDemented, and VeryMildDemented).

2. Preprocessing Steps:

- ***Normalization:*** Scaling the pixel values to a range between 0 and 1 to improve model performance.
- ***Data Augmentation:*** Introducing variations in the data by applying transformations like rotation, and flipping, increasing the diversity of the dataset, and improving model generalization.
- ***Train-Test Split:*** Dividing the dataset into training and testing sets to evaluate the model's performance on unseen data.

Modeling

1. Modeling Techniques:

For the image classification problem of identifying Alzheimer's disease stages based on brain MRI images, we will use Convolutional Neural Networks (CNNs).

2. Target Variable:

The classes represent different stages of Alzheimer's disease: MildDemented, ModerateDemented, NonDemented, and VeryMildDemented.

Evaluation

1. Metris:

The Metris used to evaluate this model will be Accuracy and AUC scores.

2. Comparison with Baseline:

The final results will include a comparison of the developed CNN model with the baseline model. This comparison will highlight the improvements achieved through the use of advanced CNN architectures.

Tools/Methodologies

1. Python Libraries:

- NumPy
- Os
- TensorFlow
- Keras
- Sklearn

2. Analysis Environment:

The analysis will be performed using Python programming language and its libraries on the local machine