Weekly Progress Report

This report outlines the progress made in the development of a deep learning-based **Brain Tumor Detection System** using MRI scans. The primary focus for this week was on dataset acquisition, exploratory data analysis (EDA), and initial preprocessing to establish a strong foundation for model training.

Brain Tumor Detection System using MRI scans.

Progress Summary:

2.1 Dataset Acquisition and Exploration

2.1.1 Library Importation

To ensure an efficient workflow, the following key libraries were utilized:

- NumPy & Pandas – For numerical computations and structured data handling.
- Matplotlib & Seaborn – For data visualization and statistical analysis.
- OpenCV – For image processing, including resizing and augmentation.
- TensorFlow & Keras – For deep learning model implementation and training.

2.1.2 Data Collection and Verification

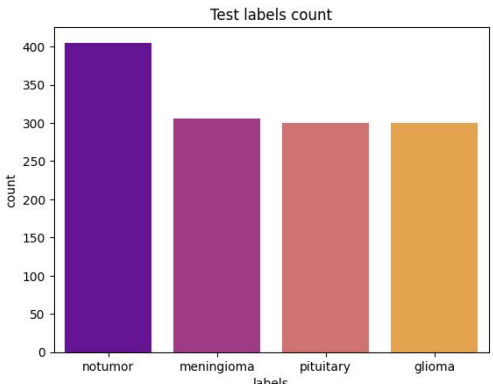
- The dataset consists of MRI scans categorized into "Tumor" and "No Tumor" classes.
- Image data was loaded using cv2.imread() and standardized to a fixed dimension (e.g., 128x128 pixels) to maintain uniformity.
- Initial inspections were conducted to ensure correct labeling and data integrity.

```
import os
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from matplotlib.image import imread
from PIL import Image

from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report,confusion_matrix

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense,
Dropout
```

2.2 Exploratory Data Analysis (EDA)

EDA was performed to understand the dataset structure and identify potential preprocessing requirements. Key activities included:

- **Class Distribution Analysis:** Verified the ratio of tumor vs. non-tumor samples to detect class imbalance.
- **Visual Inspection of MRI Scans:** Sample images were visualized using matplotlib to assess data quality and variability.
- **Pixel Intensity Distribution Analysis:** Histograms of grayscale pixel intensities were plotted to examine contrast and brightness levels across image



2.3 Data Preprocessing and Augmentation

To improve model performance and generalization, the following preprocessing techniques were applied:

- **Grayscale Conversion:** Images were converted to grayscale to reduce computational complexity while preserving essential structural information.
- **Normalization:** Pixel values were rescaled to the **[0,1]** range to facilitate stable training and faster convergence.
- **Data Augmentation:** Various transformation techniques were applied to artificially expand the dataset, including:
    - **Rotation:** Random rotations up to 30 degrees.
    - **Flipping:** Horizontal and vertical flips to increase diversity.
    - **Zooming and Shifting:** Introduced slight zoom-in/out effects to enhance model robustness.

CNN Model and Splitting

```
In [10]: train_df.shape[0]

Out[10]: 5712

In [11]: train_df, valid_df = train_test_split(train_df, test_size=0.2, random_sta
         te=42)
```

```
import random
random_indices = random.sample(range(len(test_df)), 4)
random_paths = test_df.iloc[random_indices]['filepaths'].values
actual_labels = test_df.iloc[random_indices]['labels'].values

plt.figure(figsize=(16, 4))

for i, img_path in enumerate(random_paths):
    img = load_img(img_path, target_size=(224, 224))
    img_array = img_to_array(img)
    img_array = np.expand_dims(img_array, axis=0)
    img_array /= 255.0
```

Challenges and Mitigation Strategies

3.1 Challenges Encountered

- **Class Imbalance:** A slight imbalance in the number of tumor vs. non-tumor images, which may affect model predictions.
- **Image Quality Variations:** Some MRI scans exhibit low contrast, potentially impacting feature extraction.

3.2 Mitigation Strategies

- **Data Augmentation:** Applied oversampling and synthetic transformations to balance classes.
- **Contrast Enhancement:** Implemented histogram equalization techniques to improve low-contrast images.

4. Conclusion

This week's progress successfully established the foundational components for developing an AI-powered **Brain Tumor Detection System**. Through comprehensive **dataset exploration, preprocessing, and augmentation**, we have prepared the data for subsequent deep learning model training. The upcoming week will focus on **model development and performance assessment** to validate the effectiveness of the approach.

This structured approach ensures a robust and reliable system for early and accurate brain tumor detection using MRI scans.