Capstone Project Proposal December 23, 2023

For my capstone project I will be creating a model for predicting brain tumors in humans using MRI (magnetic resonance imaging) scans. I will use a set of images from MRI scans to predict whether the scans show no tumor or one of three commonly occurring brain tumors. The three types of tumors are:

- Glioma - 33%[1] of all brain tumors are gliomas.
- Meningioma - 30% of all brain tumors are meningiomas.
- Pituitary - Most pituitary tumors are benign, but due to their location within the pituitary gland inside the brain these tumors can cause severe health issues including hormone irregularity which affects other parts of the body.

The reason for creating this predictive model is because there are a lot of abnormalities in the sizes and locations of brain tumors, making it really difficult to completely understand the nature of the tumor. Currently MRI scans are the best technique to detect brain tumors. The scans have to be examined by a specialist (ex. neurologist or radiologist). A well trained predictive model can help with three major issues:

1. By reducing the likelihood of human error via an additional review of a scan by an algorithm. The error rate due to human interpretation of radiological scans is roughly 30%, and is largely unchanged over the past 70 years. [2] Interruptions in the workplace contribute to human error, as does how well rested the person reading the scan results is and bias. Outsourcing of the interpretation of these scans may also be a factor.
2. The lack of an available specialist in certain geographic areas.
3. Quicker detection of tumors.

An additional benefit of integrating AI into the process could be the reduction in the time necessary to create MRI reports due to having less reliance on a qualified expert to interpret the results and generate the report. The generation of such a report is outside of the scope of this project.

Proper treatment, planning, and accurate diagnostics should be implemented to improve the life expectancy of patients. The 5-year survival rate for adults is approximately 35%, and can vary greatly with age. Time is of the essence in determining the health of a patient with potential brain tumors and starting any necessary treatment.

For this project I plan to use the Kaggle dataset found at: https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri/download?datasetVersionNumber=2

My approach will be a multiclass classification problem with four potential outputs. The four predicted outputs correspond to the three tumor types listed above, plus the possibility of no tumor being present in a scan. This will be a supervised learning project, and I plan to use deep learning (convolutional neural network) for image recognition. The position and size of a tumor will likely be key predictors.

---

[1] All referenced medical information taken from John Hopkins: https://www.hopkinsmedicine.org/health/conditions-and-diseases/

[2] https://www.rsna.org/news/2022/march/human-error-in-radiology

The final deliverable will be a web service with an exposed API. The modeling and deep learning will be done on my laptop, which has the following computing resources (Macbook Pro):

a. Processing power: CPU 16 cores
b. Memory: 48 GB
c. Specialized hardware: GPU 40 cores, Neural cores 16.