Author: Vijeet Yarlagadda

Brain Tumor Scan Classification using Computer Vision and Neural Networks

Business Understanding

- What problem are you trying to solve, or what question are you trying to answer?

The question that this project aims to answer is "Can you use a neural network to aid in the diagnosis of brain tumors?"

- What industry/realm/domain does this apply to?

This project applies to the medical research industry.

- What is the motivation behind your project? (Saying you needed to do a capstone project for flatiron is not an appropriate motivation)

According to VeryWell Health, brain tumors are the second most occurring tumors found in children. Brain tumors can cause extremely adverse effects, and affect thousands of families every year. If there is a possibility that we can aid in the diagnosis of this disease, I believe that exploring that possibility is not only a matter of curiosity, but one of responsibility.

Data Understanding

- What data will you collect?

I will collect brain scan images of patients with different types of brain tumors.

- Is there a plan for how to get the data (API request, direct download, etc.)?

I will use the following dataset from Kaggle. (Brain Tumor MRI Images 44 Classes | Kaggle)

- What are the features you'll be using in your model?

I will be using all of the images which are 630px*630px.

Data Preparation

- What kind of preprocessing steps do you foresee (encoding, matrix transformations, etc.)?

Since this is a multiclass classification problem, I will be using encoding to transform the text based classes into integer based classes using dummy variables. I will also be using image preprocessing tools from the Keras library to convert the images into tensors to be used with the model.

- What are some of the cleaning/pre-processing challenges for this data?

A challenge that I will have to deal with is the fact that there are 3 different types of scans amongst the data (T1, T1C+, and T2). Dealing with these different types of scans may require using different models for each type of scan.

Modeling

- What modeling techniques are most appropriate for your problem?
  I believe that transfer learning will be the most appropriate technique for this problem.
- What is your target variable? (remember - we require that you answer/solve a supervised problem for the capstone, thus you will need a target)
  My target variable will be the class of tumor, which represents the type of tumor shown in the image.
- Is this a regression or classification problem?
  This is a classification problem.

Evaluation

- What metrics will you use to determine success (MAE, RMSE, etc.)?
  I will use accuracy as a metric of success.

Tools/Methodologies

- What modeling algorithms are you planning to use (i.e., decision trees, random forests, etc.)?
  I am planning to use a Convolutional Neural Network.