
ISyE 6740 – Computational Data Analysis – Spring 2025

Final Report: Classification of Brain MRI Images

Team Member Names: Ryan Avera & Richard Soong

Project Title: Classification of Brain MRI Images

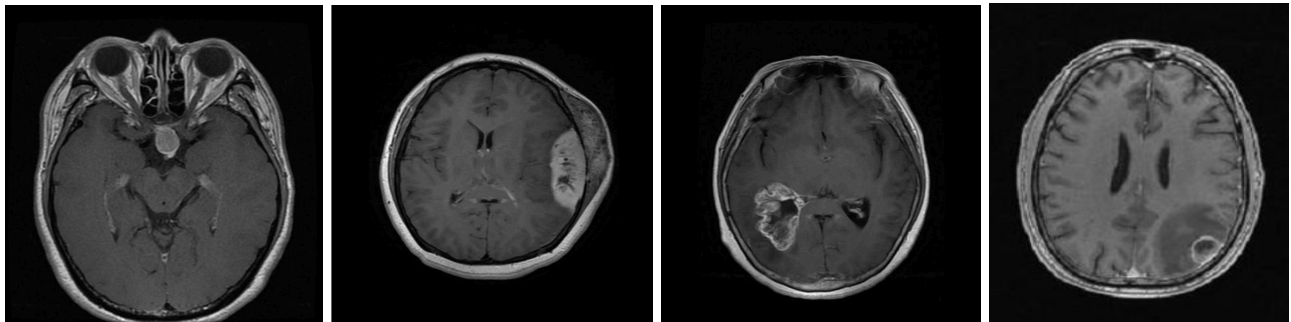
Data Source:

We used the “Brain tumor dataset” from Kaggle [1]. This dataset contains a combined 9000+ brain image MRIs across four types of brain types (three of which are abnormal), though we did not use all 9000 MRIs. Each brain type has many images taken from various angles, such as top-down and side views.

Problem Statement:

According to the Mayo Clinic, there are many different types of brain tumors including Gliomas, Meningiomas, Pituitary Tumors, and many more [2]. Our dataset focuses on the three types of tumors mentioned specifically along with brains that have no tumor. Though relatively rare, brain tumors can cause a number of issues within the brain and nervous system and, many times, can lead to death. Early-detection and classification are vital, so that proper treatment can begin promptly.

We want to explore various methods for classifying MRI images of brains based on their features. Our goal is to find a machine learning model that can, at a minimum, segment the brains that have tumors from those that do not and, ideally, classify the three types of tumors (and those brains with no tumor) into their respective categories. The main difference between the three types of tumor is the location within the brain. Gliomas tend to form within the brain tissue itself, while meningiomas form in the lining of the brain. Pituitary tumors are those that have developed on the pituitary gland. Classification of these types of tumors is critical because, according to Johns Hopkins Medicine, while most pituitary tumors and meningiomas are benign, gliomas tend to be malignant, meaning they grow rapidly and have a higher rate of mortality [3].



Methodology:

The dataset consists of brain MRI images sorted by tumor type or absence of a tumor. We used a subset (750 of each class) of the full dataset for model training/testing and a smaller subset for visualization tasks. Images were inconsistent sizes and dimensions, so preprocessing occurred while loading the images, including resizing images to 64x64 and removing an extra, unneeded dimension from some of the images. Images were also flattened to vector form as needed for some of the models. In addition to preprocessing the images themselves, we also had to create an array of labels. In the original dataset, grouped images of the same class were given in labeled folders, but individual images did not have labels.

After preprocessing, we split our data into standard 80% training and 20% testing sets and began training and comparing results of the different models using the test data.

We tried a number of different models as well as various classifications/labels:

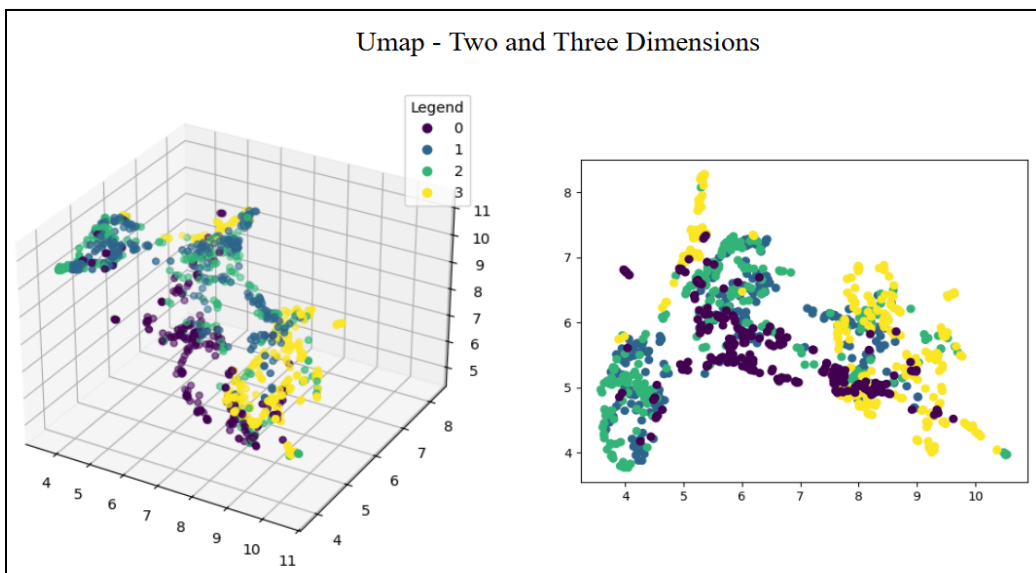
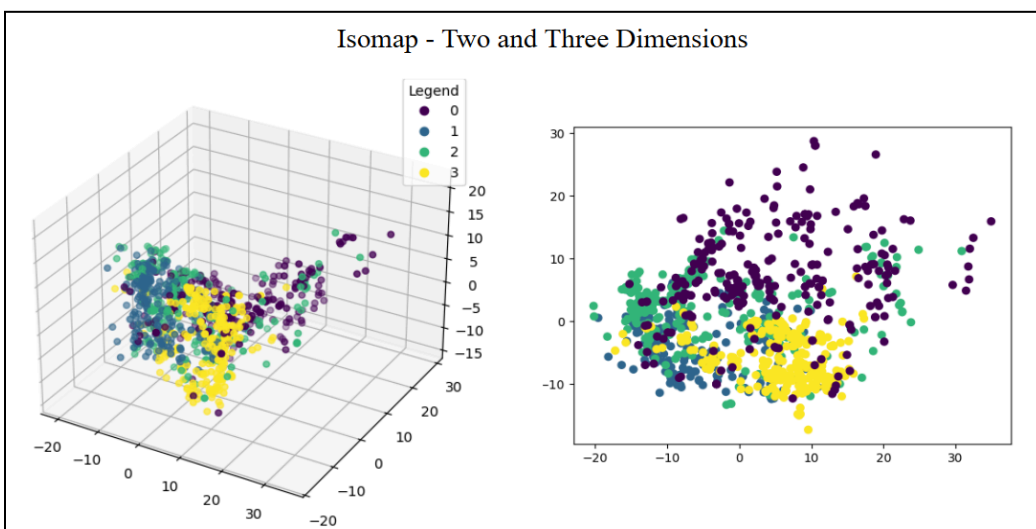
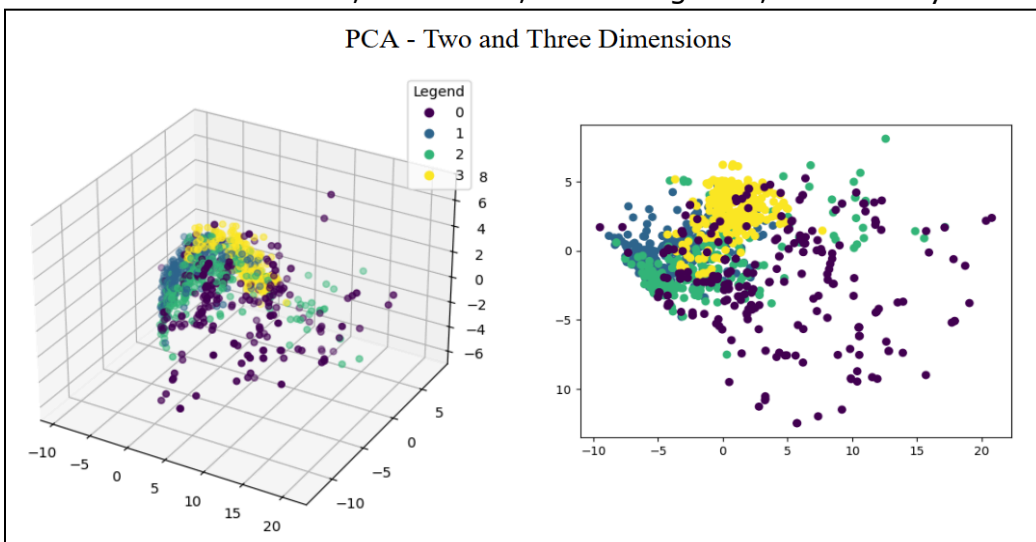
1. Dimensionality reduction techniques (PCA, Isomap, Umap) were used to aid visualization and try to uncover hidden patterns in the data that could be detected visually in a reduced dimension form.
2. In one approach we performed dimensionality reduction in combination with clustering to try to segment the images as if no labels had been provided. Here K-means was used to try and form clusters that would correspond to the correct label.
3. We also created labels for the data based on their given classification (0=No tumor, 1=Glioma, 2=Meningioma, 3=Pituitary tumor). We tried a number of classification models including K-nearest neighbors, multilayer perceptron, adaboost, and a convolutional neural network based on properties of TensorFlow image classification [4].
4. We also used the same classification models as above, but using only binary labels (0=No tumor, 1=Tumor).

Evaluation and Final Results:

Visualization with Dimensionality Reduction:

Using Isomap, Umap, and principal component analysis as a tool for dimensionality reduction does seemingly reveal some underlying structural pattern in the data. Groups are not well-defined, but we can see some clustering of similarly-labeled data points in the reduced-dimension forms. Brain images containing pituitary tumors and those images classified as having no tumor (labels = 3, 0 respectively) seem to form the most distinct groups in the plots below, which are sample scatter plots using two dimensions and three dimensions from the reduced dimension form given by PCA, Isomap (Isometric Mapping), and Umap (Uniform Manifold Approximation and Projection). Umap is a non-linear dimensionality reduction technique where points are weighted based on distance to a set number of nearest-neighbors. A weighted graph is created using these points and distance and Umap attempts recreating the graph in lower dimensional space penalizing points that were close in the high dimensional space, but are far in the lower dimensional space (and rewarding points that are close in both spaces).

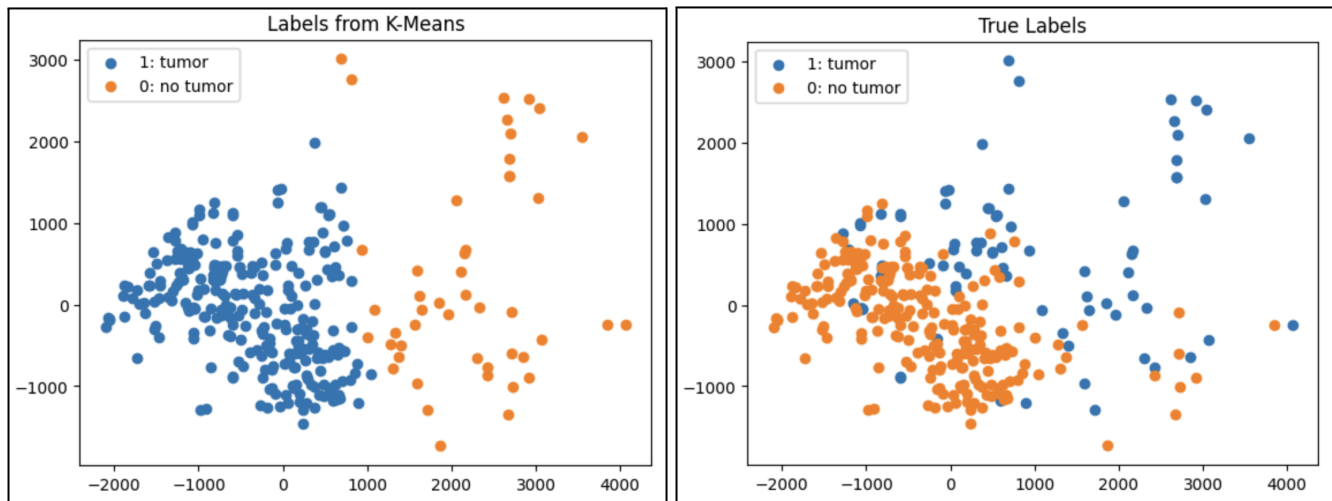
Reminder: 0=No tumor, 1=Glioma, 2=Meningioma, 3=Pituitary tumor



For both PCA and Isomap, in both 2-D and 3-D, the categories overlap a lot and are bunched very tightly together. Isomap seems to overlap a bit less than PCA, but not by much. On the other hand, Umap provides the clearest separation of classes. This is likely because of Umap's robustness in handling nonlinear data along with its innate ability to preserve local and global data structure [5].

Dimensionality Reduction + Binary Clustering:

We hoped to model and visualize the distinction between normal brains and tumorous brains. To do so, we simplified the complex brain MRI data using PCA, reducing it to 2 dimensions. We then modeled and implemented a K-Means clustering algorithm. We chose K-Means because we wanted to determine if providing unlabeled brain MRIs would be sufficient enough to classify brains.



From the above graphs, K-Means (right-side) seems to be particularly biased towards achieving a strict, clear-cut partition when labeling data points and creating clusters. This is due to how the decision boundaries for K-Means are inherently linear [6]. Unfortunately, the clusters for K-Means are in stark contrast to those of the true labels (left-side), which exhibit significant overlap between clusters. This indicates that K-Means clustering may not be entirely effective at distinguishing between brain types.

Binary Classification (Tumor or Not):

We tried a number of classification models to perform binary classification; labeling a given image as either positive or negative for tumors. Since the positive class had three times the number of samples when compared to the negative class, both accuracy and F1-scores were used to judge performance. These values can be seen below for each model: Naive Bayes, K-Nearest Neighbors, Multilayer Perceptron, and Adaboost.

	accuracy	f1_score
Perceptron	0.945455	0.962500
Adaboost	0.940909	0.959502
KNN	0.913636	0.940439
Naive Bayes	0.868182	0.910769

In terms of binary classification, it's evident that the multilayer perceptron model outperforms the other models, possessing both the highest accuracy and highest F1-score. This is likely due to the perceptron model's innate ability to progressively learn linearly separable patterns – continuously improving its accuracy [7]. Meanwhile, the Naive Bayes model seems to perform the worst. This may be because Naive Bayes uses a gaussian (normal) kernel whilst image data is often not normally distributed – especially with the various types of brains and brain MRIs present in this dataset.

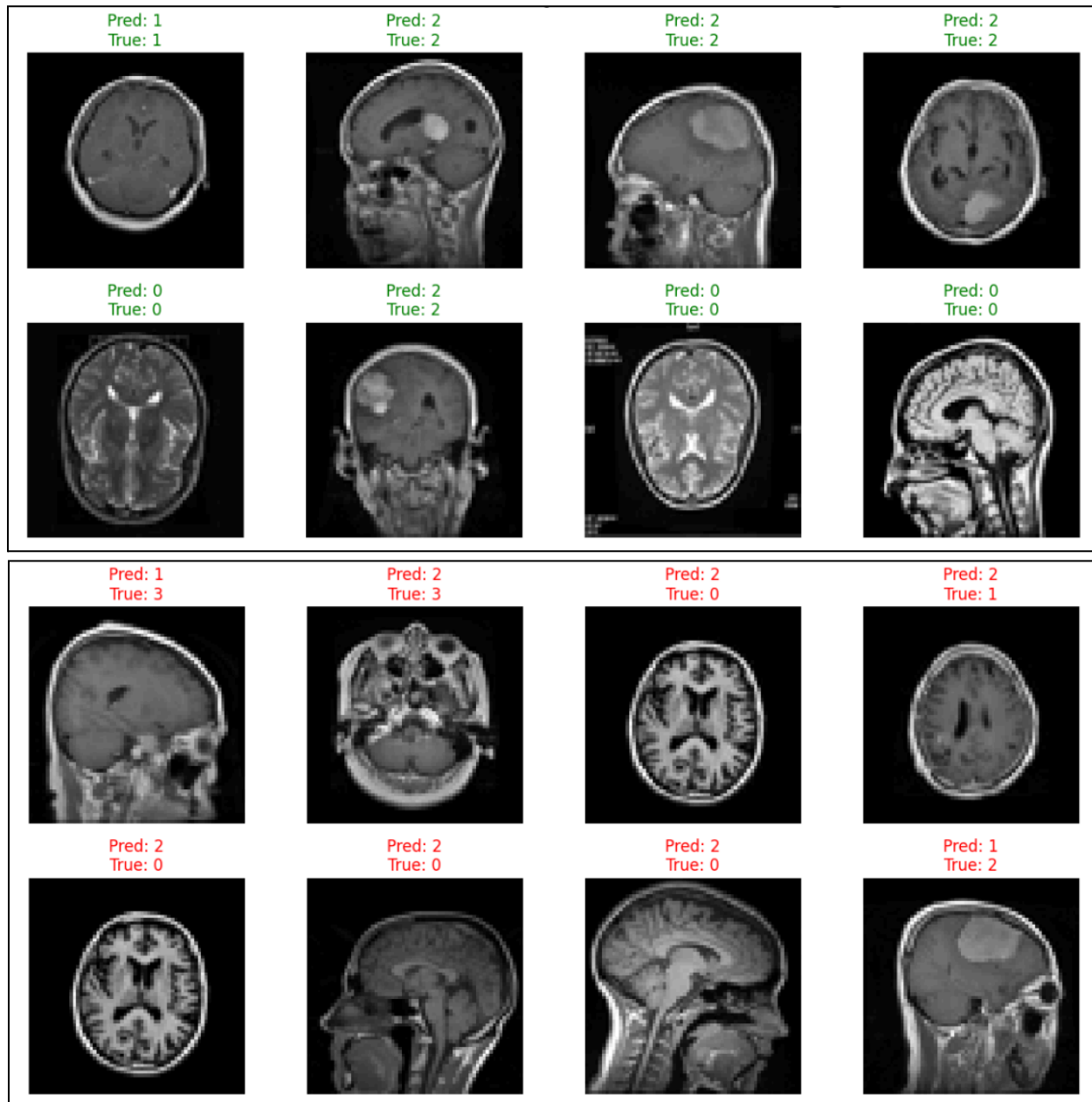
Classification by Tumor Type:

All four of the models above performed worse when trying to classify into the specific disease type (glioma, meningioma, pituitary tumors, and no tumor), with the multilayer perceptron achieving the greatest accuracy (~87%). While pretty good, we wanted to explore other options. We chose to build a relatively simple convolutional neural network to see if we could decrease misclassification rate achieved using the other four models. A convolutional neural network may combine parts/layers similar to those in a multilayer perceptron neural network with what are called convolutional layers along with other layers to reduce size and overfitting. In a convolutional layer, a small area (oftentimes called a filter) of each picture is “examined” and features are “learned”. This small area/filter is moved iteratively across the entire image to capture different features; simple features in early convolutional layers and then more complex features as more layers with filters of different sizes are added [8].

	accuracy
CNN	0.905000
Perceptron	0.863333
KNN	0.811667
Adaboost	0.781667
Naive Bayes	0.586667

Results show that the convolutional neural network performs the best for this image classification task. With further tuning of the model it may be possible to improve accuracy. With such great differences in orientation of the head in the MRI images, it may be worth exploring an ensemble method such that the data is first split based on perceived orientation and then a

neural network (or other classifier) applied for classification or using augmentation where multiple new images are created from a single image by reflecting, zooming, and rotating. which would account for some of the variations in orientation [4]. Examples of both correctly and incorrectly classified images are shown below.

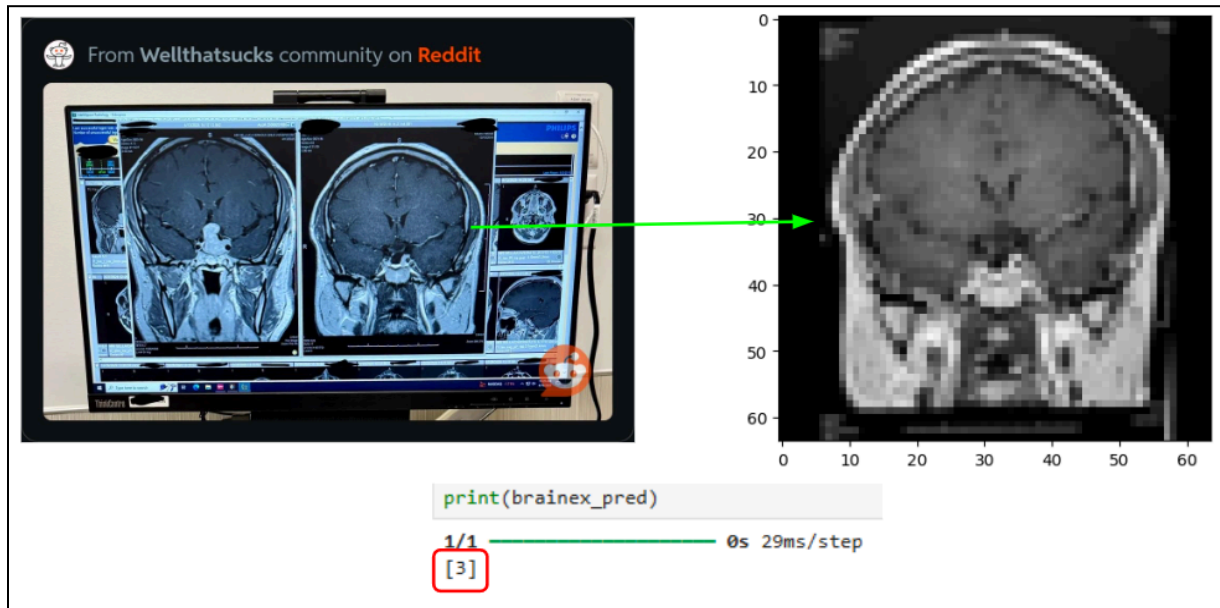


External Image Classification:

While working on the project, we came across a post on Reddit about a person whose family member received a brain MRI because of a possible growth. They had not received an official diagnosis at the time of image retrieval, but we thought it would be a great test for the trained model. We decided to process the image and feed it to the convolutional neural network to test the model on a piece of data from outside of our original data source. The original image was

cropped, converted to grayscale, pasted over a square black background, and then resized to the model's expected input size (64x64).

The image below shows the original source image as well as the image after preprocessing. It also shows that the predicted label given by the model is "3", which is the label for pituitary tumors as shown previously. This label is seemingly consistent with the location within the brain of the growth in the image when compared to other pituitary tumors.



Conclusion:

In this project, we performed a broad evaluation of multiple machine learning models including: Multilayer Perceptron, K-Nearest Neighbors, Naive Bayes, and Adaboost. These models were deployed for both binary and multi-class classifications on the brain MRI dataset. For four-class classification (by specific disease type, or lack thereof), we also included a convolutional neural network for comparison, which gave us superior classification performance due to its ability to extract features from image data. Overall, we achieved 95% accuracy for binary classification and 90% accuracy for multi-class classification.

To better understand the structure of the images and how they are related, dimensionality reduction techniques (Isomap, Umap, PCA) were used to aid in visualization. Additionally, k-means was implemented using a reduced-dimension form (given by principal component analysis) to explore how well clusters from the model compared to the true labels.

In terms of next steps, similar to how we classified a reddit brain MRI image, it might be insightful to run models like multilayer perceptron on additional external datasets, expanding the training and testing datasets to more MRI images. To further refine existing models, it might be worth applying regularization techniques like L2 regularization to convolution neural networks to lessen the chance of overfitting.

References:

- [1] Backert, Tom. "Brain tumor dataset." *Kaggle*, Updated Dec. 2024, <https://www.kaggle.com/datasets/tombackert/brain-tumor-mri-data>.
- [2] Mayo Clinic. "Brain tumor." *Mayo Clinic*, <https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084>.
- [3] Johns Hopkins. "Brain Tumor Types." *Johns Hopkins*, www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor/brain-tumor-types.
- [4] İlaslan, Düzgün. "Image Classification with TensorFlow." *Medium*, 23 Jul. 2023, <https://medium.com/@ilaslanduzgun/image-classification-with-tensorflow-a361c7b1eb05>.
- [5] Geelen, Peter. "Dimensionality Reduction and the advantage of non-linear, reproducible methods like UMAP." *Medium*, 8 Sep. 2023, <https://medium.com/aimonks/dimensionality-reduction-and-the-advantage-of-non-linear-reproducible-methods-like-umap-a75d8ad8b32b>.
- [6] CS 4/5780 Lecture Notes. "K-Means." *Cornell Computer Science Department*, Spr. 2022, <https://www.cs.cornell.edu/courses/cs4780/2022sp/notes/LectureNotes04.html>.
- [7] GeeksforGeeks. "What is Perceptron | The Simplest Artificial neural network." *GeeksforGeeks*, Updated 21 Oct. 2024, <https://www.geeksforgeeks.org/what-is-perceptron-the-simplest-artificial-neural-network/>.
- [8] IBM Think. "What are Convolutional Neural Networks?" *IBM*, <https://www.ibm.com/think/topics/convolutional-neural-networks>.