

You are a data scientist tasked with developing a machine learning model to classify brain MRI images into three tumor types: glioma, meningioma, and pituitary tumor. You have been provided with the "Brain Tumor Classification (MRI)" dataset from Kaggle, which contains MRI images along with labels indicating the type of tumor present. Your goal is to build, train, and evaluate an SVM-based multiclass classification model for this medical imaging task.

1. Data Exploration and Visualization:

- Visualize the distribution of the three tumor classes in the dataset. Are the classes balanced or imbalanced?
- Plot the distribution of patient ages in the dataset. Are there any trends or patterns related to age and the presence of specific tumor types?

2. Data Preprocessing:

- Split the dataset into training and testing sets (e.g., 80% training and 20% testing).
- Apply data augmentation techniques to increase the size of the training dataset. Visualize some of the augmented images.

3. Model Building and Training:

- Extract relevant features from the MRI images. Visualize a few examples of these extracted features.
- Train an SVM classifier on the training data for multiclass classification. Visualize the decision boundaries of the SVM classifier along with the support vectors for each class.

4. Model Evaluation:

- Evaluate the SVM model's performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC for each class. Visualize the confusion matrix.
- Plot the Receiver Operating Characteristic (ROC) curves for each class and calculate the Area Under the Curve (AUC) to assess the model's discrimination ability.

5. Hyperparameter Tuning:

- Experiment with different kernel functions (e.g., linear, polynomial, RBF) and SVM hyperparameters. Visualize how changes in hyperparameters impact the model's performance.

6. Model Interpretation:

- Determine the importance of extracted features in classifying the different tumor types. Visualize the feature importances.

7. Prediction:

- Make predictions on the testing dataset and visualize some of the MRI images with their predicted labels.
- Identify and visualize a few examples of misclassified MRI images. Explore common patterns among these misclassifications.