

Extracting Radiomic Features from Brain MRI Scans for Tumor Classification

Daniel-Luke Isemonger

Interest: 10/10

1. This project aims to classify brain tumors using an ensemble of machine learning models trained on radiomic features extracted from MRI scans. The models will include kNN, logistic regression, naive Bayes, random forest and gradient boosting. Each simple model will be trained separately, then their predictions will be aggregated via techniques like weighted averaging, stacking or majority voting to produce the final classification. I believe that using an ensemble will leverage the strengths of different algorithms to improve prediction accuracy.
2. The dataset will consist of 2D MRI scan slices of patient brains and corresponding diagnostic data. Proposed Features and example row:

Feature	Representation	Example Value
Age	Integer	62
Gender	Binary	1 (Male)
Tumor Location	Categorical	Frontal Lobe
Tumor Area	Continuous	52.3 cm ²
Tumor Texture	Continuous	0.73
Tumor Intensity Mean	Continuous	198
GLCM Contrast	Continuous	0.082
Target	Categorical	Glioblastoma

3. The dataset will be gathered from Kaggle (<https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>), The Center for Biomedical Image Computing and Analytics (<https://www.med.upenn.edu/cbica/>), and Figshare (https://figshare.com/articles/dataset/brain_tumor_dataset/1512427). The scans will then be preprocessed to extract the relevant feature data into a .csv or similar file.

