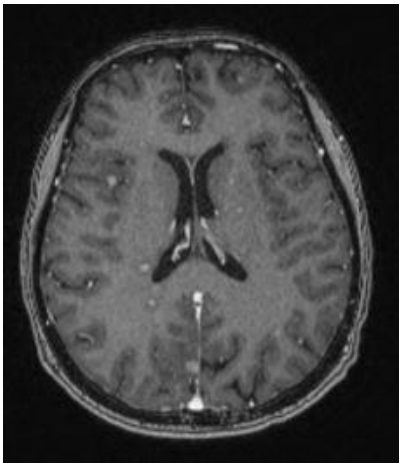
Data Appendix

MRI Brain Scan Dataset

The unit of observation for this dataset is an individual MRI brain scan image. The variables are the image data and the classification labels.

Yes Tumor



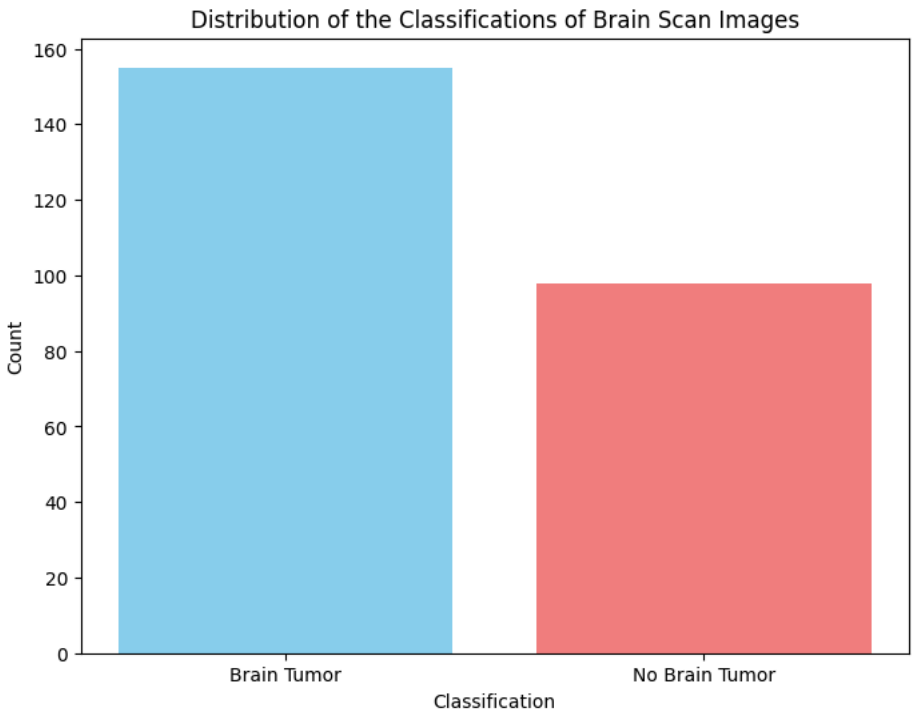
No Tumor

Image Data

This variable contains the pixel intensity values for black and white MRI scans of the brain. An MRI, or Magnetic Resonance Imaging, is a medical imaging technique that uses powerful magnets and radio waves to produce detailed images of the inside of the body. These images were originally from the brain_tumor_dataset.zip file and varied in size. This variable can be classified as a quantitative variable and each pixel can range from 0 to 255, where 0 is black and 255 is white. There are 253 observations for this variable and no observations are missing, which can be reported as 253(0). Preprocessing steps included resizing the images to be 244x244 pixels and normalizing the pixel values from the original range [0, 255] to [0, 1] in order to improve model performance.

Classification Labels

This variable indicates whether the MRI brain scan shows a brain tumor: a “yes” label indicates that there is a tumor and a “no” label indicates that there is not a tumor. On a brain MRI, a tumor typically appears as an abnormal mass or growth that stands out from the surrounding healthy brain tissue, showing up as a distinct area with different signal intensity depending on the type of tumor and the imaging sequence used [1]. The classification label is a categorical variable. Out of the 253 total image classification labels, there were 98 “no tumor” labels and 155 “yes tumor” labels. Since there were 253 observations for this variable and there were no missing observations, this can be reported as 253(0). No preprocessing was needed for this variable.



References

[1] American Health Imaging. “How Is an MRI Used to Detect Brain Tumors?” American Health Imaging, https://americanhealthimaging.com/how-is-an-mri-used-to-detect-brain-tumors/. Accessed 22 November 2024.