Mohammed Uddin, Jason A. Charles Nov. 7th, 2023

CISC 3440

Check In 1

1. Brain Tumor MRI Dataset <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset/data>

**Problem Description:** Identify certain brain tumors including glioma, meningioma, and pituitary by learning from a dataset of brain images. Model should be able to correctly do what radiologists do daily.

**Why is it worth Solving?** It will reduce the cost of tumor identification. It can also reduce the cost of getting MRIs, as a good working model will lead to doctors and technicians working less on identifying a healthy brain or not and more on administering MRIs. It can speed up the process in tumor identification as models work much faster than humans.

**What types of machine learning systems might you use to solve it?** Supervised Classifiers. Regression model won’t work, as there aren’t continuous values to be predicted. It would work similar to the iris or mnist model as the model will be checking a brain scan for certain features.

**Inputs:** The Model takes in a jpg image

**Outputs:** Classify an image as one of 4 classes: no tumor, glioma, meningioma, or pituitary

1. Heart Disease Prediction

<https://www.kaggle.com/datasets/rishidamarla/heart-disease-prediction/data>

**Problem Description:** Doctors and mainly their assistants spend hours trying to put the profile of a patient together. Sometimes the patient profile raises alarms to the doctor because the information they gathered on their patient may indicate a high risk of heart disease. Usually there are common habits and symptoms of patients with and without heart disease. A model should be able to process some info about a patient like any person and be able to predict whether or not the patient will have heart disease or not.

**Why is it worth Solving?** By developing a good model that can predict heart disease, It will save time for healthcare providers. They won’t have to spend potentially minutes or hours going through data and trying to remember what numbers or symptoms should raise an alarm and potentially making an error.

**What types of machine learning systems might you use to solve it?** A supervised classifier is most appropriate. As the classifier goes through the features of one column it should get closer and closer to the conclusion that a patient has heart disease or not

**Inputs:** Age, Sex, Chest pain type, BP, Cholesterol, FBS over 120, EKG results, Max HR, Exercise angina, ST depression, Slope of ST, Number of vessels fluro, Thallium

**Outputs:** Heart Disease presence

1. Heart Attack Risk Prediction Dataset

<https://www.kaggle.com/datasets/iamsouravbanerjee/heart-attack-prediction-dataset>

**Problem Description:** Many things can be recorded from a patient that could indicate if an individual is at risk of a heart attack or not. Information ranging from age, cholesterol, having diabetes, to living in a certain country can all have an impact on a person’s risk of getting a heart attack. There are many things to remember when it comes to flagging bad values for certain features, a person can easily forget but a model doesn’t need to rely on memory the way a person does.

**Why is it worth Solving?** It can be annoying, long, and very error prone when a person is going through patient data. Reading numbers may not be a physician or their assistant’s favorite job. It would make their jobs much easier if they could simply input data in a certain format and just rely on a good model to read through the data and make a prediction

**What types of machine learning systems might you use to solve it?** A supervised classifier. It seems most appropriate since we are just using a bunch of features to predict a binary outcome

**Inputs:**Patient ID, Age, Sex ,Cholesterol, Blood Pressure, Heart Rate, Diabetes, Family History, Smoking, Obesity, Alcohol Consumption, Exercise Hours Per Week, Diet, Previous Heart Problems, Medication Use, Stress Level, Sedentary Hours Per Day, Income, BMI, Triglycerides, Physical Activity Days Per Week, Sleep Hours Per Day, Country, Continent, Hemisphere

**Outputs:** Heart Attack Risk