Jared Kelnhofer

May 5, 2020

Machine Learning Project 7

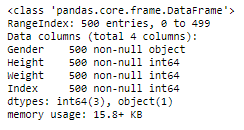
Using Weight, Gender, and Body Mass Index Data to Estimate Height

**Purpose**

In this project, I will attempt to use machine learning techniques to train a model that can accept 3 features of a person, (Weight, Body Mass Index (BMI,) and Gender) and give a reasonable estimate of that persons height. The challenge with this project is the fact that I have only 500 data instances to work with. This means that pre-processing and careful practices regarding my data are extremely important. One thing that I am excited to learn by doing this project is whether or not a neural network can discover a simple relationship, even with little training data. Because BMI is related to Weight and Height by a simple equation, my hope is that my model will discover this early on and be able to give very good answers despite my small amount of training data.

**Phase 1: Acquiring and Exploring the Data**

I started my project by using Pandas to import my data into a Dataframe. This allowed me to quickly and easily explore the data and check for missing values. Because I had so little data to worry about, I checked for missing values by eye in about a minute.



Here are the columns in my original DataFrame. All but one of the features, (Gender) turned out to be Integers. To handle the categorical Gender feature, I decided to go with one-hot encoding that was provided by the Pandas *get\_dummies()* method.