Week 3 Reflection

Adam Behun

3 Lessons and 1 Business Use case

**Classification problems in ML**

Classification problems try to predict the categorical class label of a new data point based on its independent variables. They are different from regression problems because regressions predict continuous (numerical) values, but classifications predict discrete (categorical) values. There are many types of classification problems, we have studied the binary classification problem where the algorithm picks the class label of a data point from 2 classes, zero and one respectively. The logic of 0 and 1 can be used for many use cases.

**Fitting an ML model**

We have learned that many problems (if not all depending on their complexity) can be transformed from numeric into categorical and so we can use either regression models or classification models based on their accuracy. We fit an ML model by trying a number of different models in an effort to have the highest accuracy 🡪 anything can be predicted but with low accuracy prediction is just a guess. Different algorithms use different techniques to show their accuracy.

**Logistic regression**

Logistic regression models the probability that a given input belongs to a class (0 or 1). Y-axis represents the predicted probability and the x-axis represent the independent variables. Y-axis only goes from 0 to 1 and we can adjust the cut-off based on the model's accuracy. We use logistic function to give the results boundaries from 0 to 1. The logistic regression graph shows an S-shaped curve (Age vs height - starts slow, then jumps, then slows down), that represents the relationship between the independent variables and the dependent variable. The function maps the predicted probability of a binary dependent variable given the values of the independent variables.

**How will this week’s topic help you make better business decisions?**

I can use different algorithm to predict the value of my dependent variable. I can also convert regression problems to classification problems and see which is performing better.