Purevmaa Damdinsuren

CPSC 4310/Spring 2020

WR5

**Part 0: Questions from articles**

1. *Write your understanding of the coefficient of determination R2*

R-squared error evaluates the performance of a regression model. It measures the strength of the relationship between your model and the dependent variable. Most people think R-squared value goes from 0 to 1, but actually it goes from negative infinity to 1. Higher R2 values represent smaller difference between observed values and predicted values.

1. *Can R2 value be negative? Provide one reason R2 results in a negative value.*

Yes, R2 can be negative. One of the reasons for negative R2 value is that chosen model does not follow the trend of the data.

1. *Are low R2 values always a problem? Provide your answer with a short explanation.*

No. Sometimes, model with low R2 value can be perfectly good model. “Some fields of study have an inherently greater amount of unexplainable variation. In these areas, your R2 values are bound to be lower.”

1. *Are high R2 values always great? Provide your answer with a short explanation.*

No. Sometimes, model with high R2 value can be overfitted or highly bias, which does not result good prediction on real-world data.

**Part 1: Questions from textbook chapter**

1. *Provide a short description of the fundamental idea behind Support Vector Machines?*

It is one of the most popular models in ML, and it is very powerful and versatile; capable of performing linear or nonlinear classification and regression. SVM is particularly fit well with classification of complex small-or medium-sized datasets.

1. *What is a support vector?*

Instances located on the edge of the decision boundary are called support vectors.

1. *Why is it important to scale the inputs when using SVMs?*

SVMs try to fit the largest possible space between the classes, so if the training set is not scaled, the SVM will tend to neglect small features.

**Part 2: Write one thing you really wish to take away from this course.**

From this course, I just want to get familiar with python tools to work with data and create simple, but useful models. However, I am not really interested in data science, but as far as my career concern Machine Learning is one of the fields, I want to get familiar with. That is why, getting to know what is available for me to conduct future research and ML models is my goal.