Machine Learning Documentation

Sunpreet Singh Chahal

Table of Contents:

1. Foreword (Pg. 3)
2. Introduction (Pg. 4 – Pg. X)
3. Machine Learning Dictionary (Pg. X – Pg. X)

Foreword:

After spending some time on learning a handful of machine learning models, I realised that I had made progress on learning the code for each of the models, however, I made minimal progress on understanding the models well enough to where I could converse the results the model had produces.

The end-goal of the document is to gain a deeper and more well-rounded understanding of machine learning.

*This document hasn’t been reviewed to be used as a professional reference, more so, just personal journaling of what I’ve learned so-far and for personal reference.*

Introduction:

In this section cover what machine learning is…

**\*\*What machine learning is\*\* / \*\*It’s applications\*\* / \*\*Future of ML\*\* / \*\*Why it is important for my career (as a data scientist)\*\***

The purpose of this document is to go over:

1. Every model (as I learn) in detail; what the model is, the purpose of said model and where it could be applied in a real-life scenario.
2. To understand key information that the models present; P-value, R-Squared, MAPE, MSE, F-Statistic etc. Rather than glossing over these variables, but to understand them and be able to extract *that* much more value from the model.
3. To cover the mathematical theory behind machine learning – my mathematical is quite poor, however, with the support of my colleagues I’ve gathered a range of helpful resources to help me built relevant mathematical knowledge regarding machine learning. This is very important, as understanding the theory behind this topic will allow me to **\*\* INCLUDE ANSWER SENT TO SHANGHUA/BRUNO/LEO – IN SNIP OR NOTEBOOK REGARDING THIS DOCUMENT (AT THE BOTTOM) \*\***

General ML Knowledge:

In this section cover topics such as different libraries used within machine learning

**\*\*SciKit / TensorFlow / NLTK / Statsmodel etc..\*\***

Machine learning, though computer-based, may suffer from human behaviours…

**\*\*Selection bias\*\* / \*\*Undercoverage bias\*\* / \*\*Survivorship bias\*\* / \*\*Different biases\*\***

ML Model Dictionary:

In this section I’ll cover important values that can be extracted from models.

Having a good understanding of:

* How to calculate each value
* What each of these values represents
* The importance of each value
* The use case of each value – where it’s appropriate to use

Combined, will 1) Allow me to extract further value from each of my models, and 2) Gain deeper insight on each model’s capabilities.

Machine Learning Models:

In this section I will cover various machine learning models. Each model will have comprehensive notes about; what the model is, what the model does and what its use case is.

There are three types of machine learning models: Supervised, Unsupervised & Reinforcement Learning. Detailed notes on each class of model can be found in their respective sections.

Diagram

Description automatically generated

**Supervised Learning Models:**

A model that uses labelled data to train the machine. The output(s) are known, the machine just needs to class the input accordingly.

For example, our data has 2 labels/classes; ‘cats’ and ‘dogs’, with various features such as ‘bark’ (1/0) and ‘purrs (1/0). We use a classification model to see what class our input animal is.

1. The animal barks – the model will predict the animal to be a dog.
2. The animal purrs – the model will predict the animal to be a cat.

**Linear Regression – Single Variable:**

Mathematical Theory Behind ML

**Linear Algebra:**

xxx