# Learning Journal draft

In the first two chapters of "Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow" and during my first project in the course, my primary focus was on data cleaning and conversion. I mainly learned about the statistical effects of different data transformation and the mathematical background, which I see as a foundational and important asset.

It was very insightful to learn and differentiate between the different types of machine learning, and to overview the relevant concepts that are put to use when training different models.

This ignited my motivation to create a machine learning project (assignment 3.1) from scratch. Since we began with linear regression models, I chose to construct a linear regression model independently. However, I didn't want to use pre-existing data; instead, I desired a dataset that I could consistently update and extend. My interest in real-time financial data led me to build a Linear Regression model with the scikit library and yfinance. I aimed to predict the price of Bitcoin by assessing its correlation with other assets in the market.  
  
To utilize the data consistently, I developed a database that stores the predictions, allowing me to create a trading bot that can make informed decisions based on the model's predictions. I also calculated the correlation between assets over various time periods, dynamically implementing a time range that adapts as I retrain the model. To diversify the assets, I included those with positive correlations, such as SPY to Bitcoin, those with negative correlations, such as Gold to Bitcoin, and a broad index of bonds. I utilized as many parameters as possible in the model to account for biased predictions due to the differences in base values of the assets, and converted the dataset to work on percentage change of the assets values.

For my second project (assignment 3.2), I constructed a basic neural network that predicted Bitcoin's price based on sequential analysis. I utilized historical data from Yahoo Finance via the yfinance library to focus specifically on BTC. I designed and trained an LSTM-based neural network model to predict future BTC prices based on past data. The model's efficiency was evaluated on a test set, and I created functions to determine the optimal training sequence length and visualize the model's predictions. The ultimate goal of this project was to develop a model that can accurately predict BTC prices based on historical data, potentially allowing for more informed cryptocurrency trading decisions. Both projects are available for review in my assignments repository as Assignments 3.1 and 3.2, respectively, and will be part of a trading bot, as I plan to adapt the projects to work more efficiently with the knowledge that I gain from this course.

As a third project (assignment 3.3), I used to create a linear regression model and a random forest to predict the probability of survival for passengers of the titanic. Firstly, I analyzed the dataset, whereas already some clear patterns emerged, converted the data, and selected a couple features for the models. I was quite happy with my third project, although I tried to keep it simple, as it was mainly a review of last week’s lesson.