# Swiss equity market forecasting

Some of the major sub-fields of finance is trading; a highly profitable activity loaded with historical data. Accordingly, this thesis seeks to explore and make use of Computer Science Technologies, Machine Learning approaches and sequences of historical time series data in order to achieve the ability to spot opportunities in the Swiss equity market and implement orders for the profit making.

## Introduction

The following project will seek to explore two main final products: The first one would be a Tkinter based dashboard showcasing the latest financial updates of companies to buy/short-sell its stocks. The second product would be a standard or a benchmark within the algorithms of the machine learning process that would potentially spot the overbought/oversold stocks in the Swiss equity market by learning from the historical prices and the technical indicators. The algorithm with the best metrics results will be used to trade automatically Swiss companies stocks.

## Data gathering and stocks trading

It is safe to say that financial data is very expensive to retrieve through the platforms. In this context, Alphavantage.io and Quandl.com are almost-free platforms. They provide a python API and a restful API to retrieve the historical data necessary for the technical indicators and for the various transactions of prices. The API was used to regain the free available Swiss stocks. The Swiss stocks list were recovered through Six Group (SWX Swiss Exchange platform).

## Data pipeline

A pipeline-like architecture is established to ensure the appropriate work of the project, data is retrieved from restful APIs and python APIs and consequently stored in a so-called Data Lakes under pickle format. Accordingly, another module is called to recover this raw data for processing, cleaning, aggregating and finally persisting it under CSV format for scalability and reusability.

A part of the cleaned data is handled by a module for plotting which will help in the procedure of highlighting the financial aspects of the traded companies.

Last but not least, the final module takes the processed data and passes it to multiple machine learning algorithms in order for it to learn trends from historical data that which are used, in this context, thanks to four classification algorithms , Multi-Layer Perceptrons, Bayesian Networks, Decision Trees and Random Forests.

## User Interface

The user interface is a jupyter-notebook which uses ipywidgets to enable the user to interact with the program and choose the traded stock that plots on a tkinter based GUI. Therefore, the user can provide its keys for Alpaca platform to trade with his portfolio.
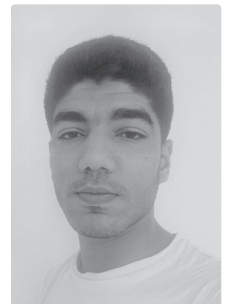
Chadi Taieb

## Results and conclusion

The accuracy differs a bit among classification models when it comes to decide whether to buy or short-sell stocks, however, The overall average accuracy is 65% on test sets. Testing in real environment may take up to 6 months in order to observe the results on the long run. The results are encouraging for further development yet human intervention is essential to overcome market constraints mostly with different market time-zones, unavailable shares, portfolio management etc.
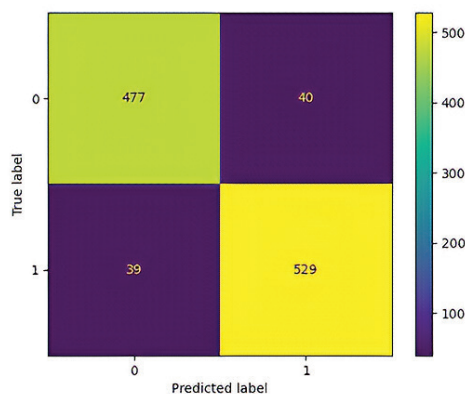


Confusion Matrix for SENS stock using Random Forests model on the whole dataset