Opleiding Informatica

Using machine learning to predict price trends

in foreign exchange rates

Mark Meijhuis

Supervisors:
Jan N. van Rijn and Holger Hoos

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

12/08/2020

Abstract

Much research has been conducted in foreign exchange price trend prediction in the machine learning industry. Many different models have been used, among which classification models using technical indicators of historical data. In this thesis, we try to expand on this previous research by varying the training time, combining technical indicators of currency pairs and backtesting on daily data of 30 currency pairs from 2012 up until and including 2018. We use random forest and gradient boosting classifiers, in addition to state-of-the-art automated machine learning model generation to train and test our data. We found that some models could be improved by adding technical indicators of other currencies, but the question remains if this can be used to generate consistent improvement. From the backtesting experiments we observed that the mean annualized return on investment is 13.71% over all currency pairs over the years 2012 up until and including 2018, with several noticeable high annualized ROI values of above 30%.

Contents

1 Introduction

1.1 The situation

The foreign exchange, or forex, market is a highly volatile market and the largest trading market in the world. In the past, this market was only used by central banks, commercial banks and hedge funds for currency trading in order to make a profit [24]. As a result of the development of the internet, many more players have entered the market and it is now possible to trade on this market as an individual. Predicting the future trend is key in making good investments and, ultimately, good profits.

Forex trend prediction has been done with many different prediction models, such as neural networks, genetic algorithms, regression models and classification models [1]. In many of these models, technical analysis was used. Technical analysis produces technical indicators which are calculations done on historical data, such as price and volume of assets being traded. Traders use these technical indicators to discover investment opportunities.

Classification models can use these technical indicators to predict whether the trend of the price of a particular currency pair will rise or fall [3]. Schut (2019) has already done previous research on classifying an uptrend or downtrend [23]. Schut produced random forest, gradient boosting and classification models with hyperparameter optimization, algorithm selection, or both. The research has shown that the gradient boosting classifier outperformed the random forest classifier using default hyperparameters. The random forest classifier outperforms the gradient boosting classifier when optimizing the hyperparameters. This research showed that it is possible to significantly outperform the baselines using just the technical indicators. Optimizing the data that is used as input for the model, in both training and testing data, has not been explored yet.

1.2 Contributions

In this thesis, we will expand on the technique used by Schut. We will do this by varying the training period in years and training a prediction model. Currency pairs are correlated [21], and to try and exploit this we will combine the technical indicators of currency pairs to observe if this improves the accuracy of the model. Finally, we will use our classification model with a trading strategy to determine the extent to which we can use this model for real-life trading scenarios.

1.3 Research question

In this work we focused on technical analysis. More precisely, we trained classifiers with technical indicators calculated on certain time points with the aim of predicting price fluctuations in the next daily interval. We looked for ways to improve the accuracy of these predictions and backtest a trading strategy with the produced models. The research question that try to answer is *'What is the effect of applying technical indicators of multiple foreign exchange currency pairs to predict the future price of a single currency pair and selecting different training periods to gain the highest accuracy in price prediction and subsequent trading profits?'*

1.4 Thesis overview

Section 2 includes the definitions and discusses related work; Section 3 describes the experiments and their outcome; Section 6 concludes. Finally, we give a recommendation for future research in 7.

2 Related work and background

2.1 Foreign exchange

The foreign exchange market is a global decentralized market for trading currencies. The foreign exchange rate of two currencies is determined by the value of the base currency divided by the value of the quote or counter currency. With an exchange rate higher than 1, the base currency is valued higher than the quote currency. If the exchange rate is lower than 1, then the base pair is valued lower than the quote pair. For example, if the EUR/USD exchange rate is 1.13, then one could purchase 1.13 US Dollars in exchange for 1 Euro. Exchange rates float freely against one another, which means they are in constant fluctuation. The foreign exchange market is an imperfect market, which means that buyers and sellers can influence prices and trading can earn profits. Open 24 hours a day for five days a week [24], this market has many opportunities for algorithmic currency trading robots.

2.2 Definitions

This subsection describes the most important definitions used in this thesis.

2.2.1 Currency pairs and major currency pairs

Any currency pair is an exchange rate between a base currency and a quote currency. There are several currency pairs, although the major ones in term of the number of daily transactions are Euro vs US Dollar (EUR/USD), Australian Dollar vs. US Dollar (AUD/USD), Great Britain Pound vs. US Dollar (GBP/USD), US Dollar vs. Canadian Dollar (USD/CAD), US Dollar vs. Swiss Franc (USD/CHF) and US Dollar vs. Japanese Yen (USD/JPY) [24].

2.2.2 Open, close, low and high

Each time interval, which can be of arbitrary length such as a minute, an hour or a day, has 4 indicators which illustrates the price movement of a currency pair over time. We use this data to determine price movements for model training and for backtesting trades using these models. These indicators are:

- **Open**: the open price is the price of the currency pair when this interval *starts*.
- **Close**: the closing price is the price of the currency pair when this interval *ends*.
- **Low**: the low price is the lowest price at which the currency pair is traded in this interval.
- **High**: the high price is the highest price at which the currency pair is traded in this interval.

It is important to realize that the open price of an interval is not necessarily equal to the closing price of the next interval, because time might have passed between the intervals due to closing times of the market.

2.2.3 Spread and commission

The spread is the difference between the bid and the ask price, i.e. the best potential price at which a currency be bought or sold. The difference between these prices is profit for the market maker. Spread is not the only expense when trading currencies, as commission is paid to the broker for making a trade. Commission is either a fixed dollar amount or a percentage of the price at which the currencies are traded.

2.2.4 Volume

In financial markets, the *volume* of a financial instrument is the amount of a security that was traded during a particular time interval. Since foreign exchange is a decentralized market, there is no way of measuring the volume of the currencies traded. The volume for traded currency pairs is measured by counting tick movements, i.e. the number of times the price moves up or down in any given time interval. It is known that the volume of financial transactions in traded currencies is much higher than that of stocks [21].

2.2.5 Volatility

The *volatility* of a financial instrument refers to how frequent and how drastic the price increases or decreases in any given time interval.

2.2.6 Leverage

Leverage is the use of borrowed funds to trade more units of currency to increase potential profits. This is, however, a double-edged sword, as the risk in terms of losses increases likewise. Using leverage is also called margin trading.

2.3 Types of analysis

A key point in trading currencies and making a profit is the analysis of price trends and the prediction of the exchange rate in the future. This can be done with fundamental analysis and technical analysis.

2.3.1 Fundamental Analysis

Fundamental analysis encompasses the analysis of the overall state of the economy on factors that could influence the value of a currency. For example, one could look at economic indicators, government policy, societal and other factors within a business cycle framework. News headlines or articles are analysed using term weighting or sentiment analysis. Previous research has shown that 30% of daily price fluctuations in exchange rates are the result of macro news [7]. Every hour, approximately 300 foreign exchange related news items were published between 2013-2017 on Reuters Business and Financial News for the major currency pairs [4]. Nassirtoussi *et al.* (2015) argue that it takes 1 hour to reflect impactful news on the price of foreign exchange pairs [11].

2.3.2 Technical Analysis

The aim of technical analysis is identifying trading opportunities by analyzing statistical trends gathered from historical trading activity, namely the price and the volume that was traded at certain points in time [1]. Technical indicators are calculated based on previous intervals for a certain interval to determine the current market trend, including support and resistance areas, while others are focused on determining the strength of a trend and the likelihood of its continuation. Technical analysis is based on the principle that history might repeat itself [7]. Previous studies have indicated that technical analysis gives a higher prediction accuracy than fundamental analysis [15].

2.4 Related work

Due to the size and the volatility of the foreign exchange market, predicting the future price can be a daunting task. It has even been suggested that the market follows a random walk, there is however substantial evidence that this is not the case [6, 8]. Several reports of substantially higher than 50% accuracy models and substantial returns using these models can be seen as empirical evidence that the market does not follow a random walk [9]. A trader looking for a profitable trading strategy would aim to find the most accurate prediction models that have the best return on investment (ROI) with a low drawdown, which is the highest peak-to-trough decline in investment portfolio. There are several approaches to finding such a model. It is widely understood that no model can significantly outperform alternative models in a real-world trading strategy [19], so we have to look at multiple approaches to get an idea of the efficacy of different methods. There seems to be a general agreement that models trained using technical indicators are well-suited for financial data modelling and forecasting [9].

2.4.1 Technical indicators classification

Using a classification model using technical indicators for time series reduces the prediction model into simple, independent classification problems as opposed to predicting future trends while taking the recent history of the trend in mind. Maggini *et al.* (1997) stated that the rules for predicting the future trend this way are constantly changing for financial time series. They suggest that there is an inherent difficulty to predicting future trends using technical indicators, because the rules for classification seem to be constantly changing for financial time series. They even suggest that there are contradictory instances of the classification in the training sets [16]. Tsang and Park (1999) generated an annual ROI of over 40% with a C4.5 decision tree model using technical indicators [14].

2.4.2 Correlation between currencies

Correlation between foreign exchange pairs was demonstrated [21]. This observation can be explained by the relative values which can be reflected by a single pair in the currency pairs. For example, it is logical for the EUR/USD currency pair to be correlated with the GBP/USD currency pair due to the relative value of the US Dollar. So-called 'arbitrage' can occur when these related currency pairs are misaligned and these can be exploited to make a profit [21]. Petropoulos (2017) used genetic algorithms combined with multiple machine learning techniques to analyze the correlation between currency pairs and generate a trading system that resulted in an annualized ROI of 17.4% using leverage [19].

2.4.3 Genetic algorithms

Genetic algorithms are algorithms that use Darwinian concepts such as reproduction, mutation, recombination and selection to improve the algorithms accuracy [1]. Using genetic algorithms, a Naive Bayes classifier and technical indicators, Abreu *et al.* (2018) produced a model with 54.95% accuracy and a subsequent annualized ROI of 10.29% [1].

2.4.4 Neural networks and deep learning

LSTM-based neural networks are recurrent neural networks, meaning they use previous data to perform their prediction, combined with long-term dependency learning [21]. Rundo (2019) showed that using arbitrage in the triangular currency pairs USD/EUR, EUR/GBP and GBP/USD can yield a 98.23% ROI from 2014 to 2018 with a 85% accurate trend prediction on minute data [21]. Kuroda (2017) used neural networks and genetic algorithms to achieve a 15.97% annualized ROI on the USD/JPY currency pair [13]. Evans *et al.* (2013) used neural networks with genetic algorithms to achieve a model with 72.5% prediction accuracy and a mean annualized ROI of 27.83% on currency pairs GBP/USD, EUR/GBP and EUR/USD using leverage [8]. Kondratenko and Kuperin (2003) claimed to to get about an 80% accuracy in trend prediction for foreign exchange [12].

3 Methodology

This section provides the source of the data, the techniques used to manipulate and analyse the data and the models used to generate predictions.

3.1 Data

The foreign exchange rates of 30 currency pairs were extracted from the Dukascopy database for the years 2012 up to and including 2018. The technical indicators used for classification were later added using the Technical Analysis Library in Python [18]. The data is normalized to the average between bid and ask price, with an extra column to supply the spread on that particular data point. The data used is purely day-to-day open, close, high and low prices, because an inherent property of financial data is that the higher the frequency of data collection, the worse the signal/noise ratio [12]. On the other hand, the more accurate the forecast, the higher its practical value. Relatively frequent trading is a must for our data, so we want to minimize the signal/noise ratio while maintaining the accuracy of the forecast. When taking this into consideration, daily data seems the best suited for our practical needs. The currency pairs used are displayed in Table 1.

Currency pair	Description
AUD/CAD	Australian Dollar vs. Canadian Dollar
AUD/CHF	Australian Dollar vs. Swiss Franc
AUD/JPY	Australian Dollar vs. Japanese Yen
AUD/NZD	Australian Dollar vs. New Zealand Dollar
AUD/SGD	Australian Dollar vs. Singapore Dollar
AUD/USD	Australian Dollar vs. US Dollar
CAD/CHF	Canadian Dollar vs. Swiss Franc
CAD/JPY	Canadian Dollar vs. Japanese Yen
CHF/JPY	Swiss Franc vs. Japanese Yen
CHF/SGD	Swiss Franc vs. Singapore Dollar
EUR/AUD	Euro vs. Australian Dollar
EUR/CAD	Euro vs. Canadian Dollar
EUR/CHF	Euro vs. Swiss Franc
EUR/DKK	Euro vs. Danish Krone
EUR/GBP	Euro vs. Pound Sterling
EUR/HKD	Euro vs. Hong Kong Dollar
EUR/JPY	Euro vs. Japanese Yen
EUR/NOK	Euro vs. Norwegian Krone
EUR/NZD	Euro vs. New Zealand Dollar
EUR/PLN	Euro vs. Polish Zloty
EUR/SEK	Euro vs. Swedish Krona
EUR/GSD	Euro vs. Singapore Dollar
EUR/TRY	Euro vs. Turkish Lira
EUR/USD	Euro vs. US Dollar
GBP/USD	Pound Sterling vs. US Dollar
NZD/USD	New Zealand Dollar vs. US Dollar
USD/CAD	US Dollar vs. Canadian Dollar
USD/CHF	US Dollar vs. Swiss Franc
USD/DKK	US Dollar vs. Danish Krone
USD/JPY	US Dollar vs. Japanese Yen

Table 1: Currencies used in trend prediction, currency combination and backtesting

3.2 Trend prediction and technical indicators

Trend prediction is a central theme in this thesis. The target for the models and the features used as training data are described here.

3.2.1 Technical indicators

The technical indicators described in previous work [23] and the Technical Analysis Library in Python [18] add 86 features to the data. Some technical indicators require a time window of 51 intervals to calculate the value of the technical indicator, which generates some rows with empty values for these indicators. These rows with empty values are removed, leaving only the rows with all the technical indicator values.

3.2.2 Target

The closing price is used as a target for the machine learning models. A binary variable is generated for each row of data, which indicates whether the closing price of the next daily interval will be higher or lower than the closing price of the current data row. A rising trend in the next interval is indicated as *true*, while a falling trend is indicated as *false*.

3.2.3 Models

The models used for trend prediction are the Random Forest (RF) classifier, Gradient Boosting Classifier (GBC) from the widely-used Scikit-Learn Python library. In addition, we have used the Auto-Sklearn classifier to automate model optimization with techniques such as model selection, hyperparameter optimization and feature selection [17]. This model can be used as a drop-in replacement for scikit-learn estimators, such as the random forest and gradient boosting classifiers.

3.3 Variable training time

State of the art techniques do not use a system that focuses on validation but only use a training and a testing split [1]. Due to indications in the literature that the training period may have an effect on trend prediction accuracy [16], we used different training periods with daily data from 2012 up to and including 2017, while using the daily data of 2018 as testing data. The scikit-learn gradient boosting classifier is used to generate the accuracy values. To test whether one training period was significantly better than another, the Friedman test was used.

3.4 Neural network

To replicate and verify promising results in the past on foreign exchange price prediction [12, 21, 13, 8], we created a recurrent neural network (RNN) based on the model that was used by Kondratenko and Kuperin (2003) [12]. We used the percentual change in the open, close, high and low price, in addition to the percentual change in volume and a percentual change in moving average of the closing price with a time window of 20 days. Sequences of 20 days were used as input to the neural network. A LSTM architecture was used in addition to an Adam optimizer with a learning rate of 0.003. Daily data of 2012 until and including 2018 was used with a 85% training split and a 15% testing split. 200 epochs were used to acquire the final accuracy. The trend of the next daily interval was predicted. The model was created in Python with the Keras framework.

There was too little time for optimizing the neural network, so the results are shown in the appendix A. There is no further elaboration on these results.

3.5 Combining technical indicators of currency pairs

Due to correlation between the price trends of currency pairs [21], it is possible that the accuracy of a machine learning model using technical indicators might be improved by combining the technical indicators of multiple currency pairs. We used the low, high, open and closing prices in addition to the volume and the technical indicators tied to these metrics and merged these with the exact same data of another currency pair on the corresponding date.

We used the gradient boosting classifier to generate an accuracy for each currency pair combination. This generates a matrix of values, which indicates for each currency pair if the accuracy has increased, decreased, or whether the change is statistically insignificant. For each entry, we generated 10 accuracy values and compared them using the Wilcoxon signed-rank test [5]. We generated the matrices for 2017 and 2018. We took the best increase in accuracy from the previous year and used as advice for the subsequent year to test if the improvement is consistent.

To draw a conclusion whether this technique can improve our models in real-life situations, we used the Auto-Sklearn classifier to test EUR/CHF, the currency pair for which the gradient boosting classifier performed best. We did the same for EUR/SGD, which was the currency pair on which the gradient boosting classifier performed worst. The baseline accuracy values can be seen in Figure 1 in Section 4.1. The data was trained from the 1st January 2012 up until 30th June 2018 and tested on the months July, August and September. The same was done for the months October, November and December. The Wilcoxon signed-rank test was used to test whether a model was significantly superior than an alternative model [5].

3.6 Backtesting

Backtesting is the process of trading with a prediction model and a strategy on historical data. For this work, we used the backtesting framework Backtrader [20]. We tested on the daily data of a certain whole year, while using he other years between 2012 and 2018 as training data for the gradient boosting classifier. We gave the Backtrader framework 100.000 units of currency to trade with. No leverage was used.

3.6.1 Strategy

When the classifier predicted a *rise* in the exchange rate, a percentual stake of the starting cash was used to buy the base currency. For example, in the EUR/USD pair the base currency would be the Euro. If the classifier predicted a *fall* in price, the base currency would be sold with a stake of 20% of the starting cash. This stake has been determined experimentally, since a higher stake seemed to result in a higher potential ROI, but also potentially higher losses. A stake of 20% therefore seemed the most suitable when finding a balance between risk and ROI. Before buying or selling, any open position would be closed. The commission per trade was the spread of the corresponding date.

3.6.2 Output metrics

We measured several metrics to quantify the backtesting results. The main result is the return on investment (ROI), which is the ratio between net profit and cost of investment, measured in percentage.

Another metric is the volatility, which is measured with the average true range (ATR), an indicator of volatility [10]. It is measured by calculating the relative difference between the closing prices of two sequential days and averaging this difference over the entire year.

To determine whether there is a correlation between metrics, we use the Spearman's rank correlation test [22]. This is a test that assesses how well the relationship between two variables can be described using a monotonic function. This test results in a correlation coefficient and a p-value to determine significance.

4 Results

In this section, we discuss the results of the experiments described in Section 3.

4.1 Baseline results

To make sense of any result of our experiments, we have to look at the baseline accuracy of the classifiers.

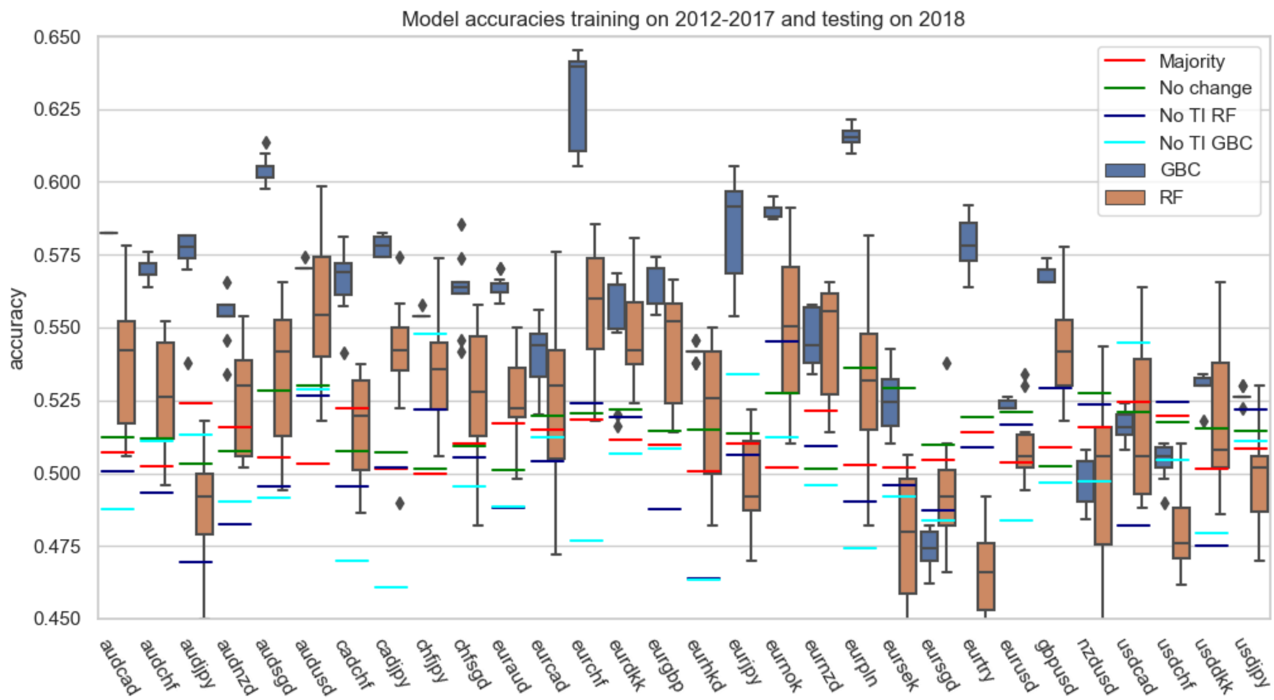

Figure 1: Baseline results of basic currency pair data with technical indicators

Figure 1 shows the accuracy of the random forest and gradient boosting classifiers on the currency pairs in Table 1, along with the baseline classifier accuracy values. The *majority* classifier assumes that the prediction, either a rising or falling trend, that has had the most instances in the past will be the next prediction. The *no change* classifier assumes that the prediction of the preceding day will be the next correct prediction. The *no TI* uses the random forest and gradient boosting classifiers to predict the trend without adding the technical indicator values to the data. The figure shows that the gradient boosting classifier generally performs better than the random forest classifier, with EUR/CHF having the highest accuracy and EUR/SGD having the lowest accuracy. In 27 of the 30 currency pairs, the gradient boosting classifier has a higher average prediction accuracy than the random forest classifier.

4.2 Variable training time

To determine whether a certain training period would be beneficial for the prediction accuracy on testing data, we measured performance of classifiers trained on different time periods with the gradient boosting classifier and comparing the testing accuracies. The results can be seen in Table 2. In the column headers, we can see what time period was trained on. We did all testing on the daily data of 2018. For some currency pairs, such as CHF/JPY and NZD/USD, a shorter training time period seems to be beneficial when we look at the higher accuracy values at the right hand side of the table. For other currency pairs, such as AUD/SGD and AUD/USD, a longer training time period seems to be beneficial when we look at the higher accuracy values at the left hand side of the table. When we run the Friedman test with each column as a distinct group, we find a p-value of 0.098, meaning that one group is not consistently performing better than the other groups. This is reflected in the means of the columns, which have accuracy values that are a short distance from each other in terms of accuracy values.

Currency pair	2012-2017	2013-2017	2014-2017	2015-2017	2016-2017	2017-2017
AUD/CAD	0.67	0.63	0.62	0.65	0.69	0.56
AUD/CHF	0.61	0.61	0.54	0.51	0.62	0.63
AUD/JPY	0.57	0.58	0.49	0.52	0.48	0.38
AUD/NZD	0.58	0.63	0.60	0.63	0.52	0.54
AUD/SGD	0.60	0.70	0.55	0.59	0.54	0.54
AUD/USD	0.63	0.68	0.61	0.62	0.59	0.62
CAD/CHF	0.51	0.53	0.53	0.50	0.49	0.48
CAD/JPY	0.47	0.54	0.49	0.51	0.50	0.50
CHF/JPY	0.53	0.50	0.48	0.49	0.58	0.61
CHF/SGD	0.58	0.51	0.59	0.55	0.55	0.56
EUR/AUD	0.52	0.54	0.55	0.56	0.51	0.61
EUR/CAD	0.57	0.60	0.51	0.53	0.53	0.44
EUR/CHF	0.59	0.60	0.62	0.58	0.59	0.63
EUR/DKK	0.47	0.46	0.50	0.45	0.51	0.50
EUR/GBP	0.65	0.67	0.65	0.64	0.53	0.54
EUR/HKD	0.52	0.57	0.54	0.52	0.54	0.51
EUR/JPY	0.50	0.52	0.61	0.54	0.52	0.53
EUR/NOK	0.50	0.55	0.62	0.51	0.53	0.55
EUR/NZD	0.62	0.58	0.52	0.50	0.54	0.45
EUR/PLN	0.53	0.53	0.62	0.64	0.53	0.57
EUR/SEK	0.48	0.50	0.45	0.43	0.41	0.47
EUR/SGD	0.53	0.53	0.47	0.51	0.53	0.54
EUR/TRY	0.59	0.51	0.52	0.52	0.47	0.43
EUR/USD	0.48	0.53	0.56	0.49	0.53	0.57
GBP/USD	0.55	0.51	0.55	0.50	0.47	0.50
NZD/USD	0.54	0.52	0.48	0.57	0.59	0.54
USD/CAD	0.50	0.58	0.56	0.51	0.53	0.49
USD/CHF	0.49	0.52	0.50	0.55	0.53	0.46
USD/DKK	0.55	0.59	0.53	0.55	0.47	0.50
USD/JPY	0.55	0.52	0.54	0.56	0.55	0.45
Mean	0.55	0.56	0.55	0.54	0.53	0.52

Table 2: Testing accuracies on daily data in 2018 with training data January of the first year to December of the second year

4.3 Combining technical indicators of currency pairs

To improve the accuracy of the models, we tested whether adding the technical indicators of a different currency pair had a positive or negative effect. We compared the resulting accuracy with the baseline accuracy, which is the currency pair data with only its own technical indicator values. The Wilcoxon signed-rank test was used to compare 10 accuracy values. If the result was not significant, a '/' was inserted in the corresponding cell. Table 3 shows a subset of the resulting matrix from the daily data of 2018 to give an impression of the resulting data.

	AUD/CAD	AUD/CHF	AUD/JPY	AUD/NZD	AUD/SGD	AUD/USD
AUD/CAD		-1.93	/	/	1.85	-1.77
AUD/CHF	/		-1.76	-3.44	0.72	-3.39
AUD/JPY	/	-4.34		/	-6.06	/
AUD/NZD	-4.86	/	/		/	/
AUD/SGD	-3.61	-5.36	-2.23	-2.63		-2.4
AUD/USD	2.17	/	-0.8	0.88	/	
CAD/CHF	-2.45	-5.79	-3.53	1.17	-5.2	-3.34
CAD/JPY	-3.78	/	2.65	2.41	/	-1.69
...	...	...	...	...	...	...
Max	6.75	5.32	2.65	4.52	3.75	6.63
Best	USD/JPY	EUR/DKK	CAD/JPY	EUR/DKK	EUR/SGD	USD/CAD
Median	-2.73	-3.01	-2.16	-1.45	-2.50	-1.35
Mean	-1.96	-2.55	-2.20	-0.99	-2.49	-0.39

Table 3: Subset of matrix with accuracy values of combined technical indicators on daily data of 2018.

In the lower rows, we can see the maximum improvement that was acquired, the currency whose added technical indicators made that improvement possible, and what the mean and median accuracies over that column are. The column headers are the currency pairs which target is predicted.

These matrices were produced for the years 2017 and 2018. We analysed the resulting matrices of 2017 and 2018 on increased accuracies (positive effects) and decreased accuracies (negative effects). The result is shown in Table 4.

	2017	2018
Positive effect	219	180
Negative effect	489	411
No significant difference	162	279

Table 4: Testing accuracies on daily data in 2018 with training data January of the first year to December of the second year

The 'Best' row in Table 3 can be interpreted as the best currency pair to combine with the base currency pair to get the best improvement. To measure whether this currency pair can be used to improve our models in the next year, we used it as advice and tested how many times the model would actually be improved. This result can be seen in Table 5. The 'Accuracy' row is the number of positive and neutral effects divided by the total number of currency pairs.

	Instances
Positive effect	8
Negative effect	12
No significant difference	10
Accuracy	0.6

Table 5: Effects of using the advised currency pair to combine with the base currency pair in the previous year.

4.3.1 Auto-Sklearn

The previous results have all been produced by using the gradient boosting classifier. To better understand whether this technique is effective or not, we used the Auto-Sklearn classifier on the EUR/CHF and EUR/SGD currency pairs. This model automatically applies optimizations, such as feature selection and hyperparameter optimization, to acquire a better performing model.

In Table 6 and Table 7 we can see the base accuracy without combined technical indicators. This accuracy is a result of training from January 2012 up until June 2018 and tested on July, August and September 2018. In Table 6 we can see that the base accuracy of the EUR/CHF currency pair is 55%. The technical indicators of the currency pairs AUD/USD, EUR/NZD and USD/DKK improved the accuracy significantly in the months July, August and September 2018 (a). When tested on October, November and December 2018 (b), these combined technical indicators seemed to significantly decrease the accuracy.

(a) July, August and September

Base	Combined	Accuracy	P-value
EUR/CHF	-	0.55	-
EUR/CHF	AUD/USD	0.57	0.02
EUR/CHF	EUR/NZD	0.62	<0.01
EUR/CHF	USD/DKK	0.6	<0.0

(b) October, November and December

Base	Combined	Accuracy	P-value
EUR/CHF	-	0.64	-
EUR/CHF	AUD/USD	0.62	0.05
EUR/CHF	EUR/NZD	0.61	<0.01
EUR/CHF	USD/DKK	0.62	0.11

Table 6: Accuracy of EUR/CHF in the indicated months with significant improvements from July, August and September.

In Table 7 we can see that the base accuracy of the EUR/SGD currency pair is 64%. Only the technical indicators of AUD/USD seemed to significantly improve the model in the months July, August and September to 69%. Upon testing in October, November and December we found that the combined technical indicators significantly reduced the base accuracy from 55% to 50%.

(a) July, August and September

Base	Combined	Accuracy	P-value
EUR/SGD	-	0.64	-
EUR/CHF	AUD/USD	0.69	0.05

(b) October, November and December

Base	Combined	Accuracy	P-value
EUR/SGD	-	0.55	-
EUR/SGD	AUD/USD	0.5	<0.01

Table 7: Accuracy of EUR/SGD in the indicated months with significant improvements from July, August and September.

5 Backtesting

In this section we discuss the results of backtesting in terms of return on investment (ROI) for the currencies in different years and the correlation between the ROI and a variety of variables.

5.1 Basic results

The results in Figure 2 show the return on investment in the year 2018 with the investment strategy described in Section 2 and the baselines. We used the gradient boosting classifier trained on the years 2012 up until and including 2017.

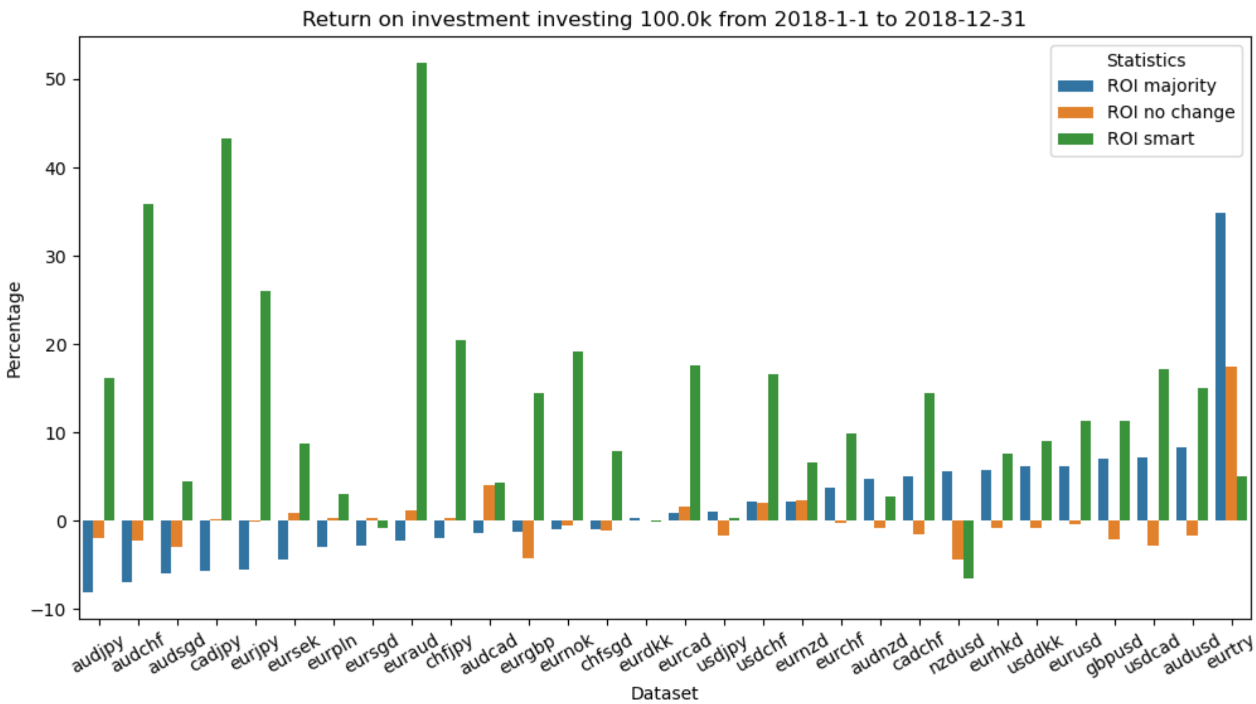

Figure 2: Return on investment with the investment strategy described in 3.6.1 and baselines.

The results show certain high ROIs with our investment strategy, such as the currency pairs CAD/JPY and EUR/AUD with ROIs of 43.2% and 51.85% respectively. There are relatively few losses with our investment strategy. The baselines seem to result in a ROI between -10% and 10%, with EUR/TRY as the only currency pair with a significant exception.

Research has been conducted on investment techniques and prediction models with very good results, but a ROI of 52% is significantly higher than other results found in the literature. To investigate this further, we will look at the ROIs of the different currency pairs in the years 2012 up until 2018 and see if we can find useful trends in the data and the results.

5.2 Consistency of returns

In Table 8 we can see how good or bad our model has performed in terms of return on investment in different years. We can see that the currency pair EUR/AUD performed the best and EUR/DKK performed the worst. On average, we would expect to find an average annualized return of 13.17%. We can see that some models perform very well in some years and very poor in other years, such as USD/DKK, which had a ROI of 29.38% in 2017 and a ROI of -33.05% in 2014.

The Friedman test over the columns 2012-2018 results in a p-value of <0.001, meaning that these years are significantly different and some years produced better results than others in terms of ROI. The Friedman test over the rows EUR/AUD - EUR/DKK also results in a p-value of <0.001, meaning that this model of investing has worked better on some models, such as EUR/AUD, than other models, such as EUR/DKK.

Currency pair	2012	2013	2014	2015	2016	2017	2018	Mean
EUR/AUD	4.88	18.28	14.25	48.02	28.22	14.51	51.85	25.72
AUD/USD	25.73	23.83	18.09	29.29	41.8	21.03	14.97	24.96
CHF/JPY	6.3	34.98	-3.81	60.33	24.33	16.65	20.38	22.74
AUD/CHF	36.56	14.44	9.82	34.18	6.35	17.19	35.85	22.06
CAD/JPY	10.06	16.92	12.59	5.49	21.24	34.55	43.24	20.58
USD/JPY	7.51	17.47	19.05	64.14	13.35	21.37	0.31	20.46
EUR/JPY	13.26	24.26	22.26	12.72	17.59	26.32	25.95	20.34
GBP/USD	7.69	33.02	6.65	14.81	25.11	34.44	11.34	19.01
AUD/NZD	44.77	10.32	11.92	17.9	23.64	11.97	2.81	17.62
AUD/CAD	11.49	10.88	15.04	20.26	30.86	24.33	4.27	16.73
EUR/USD	12.75	11.41	-4.6	44.51	17.23	21.18	11.28	16.25
EUR/CAD	6.07	0.17	5.39	29.07	30.01	24.49	17.64	16.12
USD/CAD	3.53	6.38	6.19	25.72	36.21	16.19	17.1	15.9
USD/CHF	9	-4.65	14.8	41.51	17.94	13.47	16.61	15.53
EUR/GBP	2.47	4	11.53	8.44	25.64	26.2	14.4	13.24
EUR/NZD	-7.8	15.33	9.45	17.86	19.34	20.71	6.65	11.65
CAD/CHF	-6.31	18.96	5.39	22.54	2.6	17.83	14.39	10.77
EUR/HKD	11.64	0.05	-0.36	31.18	9.36	15.94	7.61	10.77
AUD/JPY	9.46	29.9	18.26	14.5	-30.14	15.96	16.15	10.58
EUR/NOK	-0.67	-2.14	14.6	11.26	21.85	9.08	19.1	10.44
EUR/SEK	3.9	9.61	12.78	14.63	3.19	8.29	8.69	8.73
AUD/SGD	17.67	-9.33	15.5	10.47	7.16	14.4	4.47	8.62
EUR/SGD	11.6	-0.96	5.03	13.93	10.18	12.93	-0.86	7.41
USD/DKK	2.77	4.07	-33.05	25.15	14.25	29.38	8.98	7.36
CHF/SGD	5.2	-2.08	-5.1	26.19	-0.68	9.52	7.83	5.84
EUR/TRY	0.62	-14.42	-2.66	14.19	21.8	9.79	5.09	4.92
NZD/USD	-2.99	9.83	42.54	-5.12	-0.99	-2.7	-6.5	4.87
EUR/PLN	4.72	6.54	-3.89	-6.81	23.11	-0.67	2.97	3.71
EUR/CHF	-11.11	6.99	0.96	-2.27	4.71	8.86	9.88	2.57
EUR/DKK	-0.95	-0.35	-0.3	-0.14	-0.37	-0.13	-0.14	-0.34
Mean	7.99	9.79	7.94	21.47	15.5	16.44	13.08	13.17

Table 8: Return on investment with the strategy as described in 3.6.1 tested on different years, sorted by average ROI.

5.3 ROI vs. Accuracy

If we are able to perfectly predict the future price of a financial instrument, we would expect to easily make money trading that asset. Therefore, we would expect a model with a higher prediction accuracy to have a positive effect of the resulting return on investment. In Figure 3 we can see a plot of the accuracy of the model used by the backtesting framework, along with the resulting ROI after executing the investment strategy on 2018.

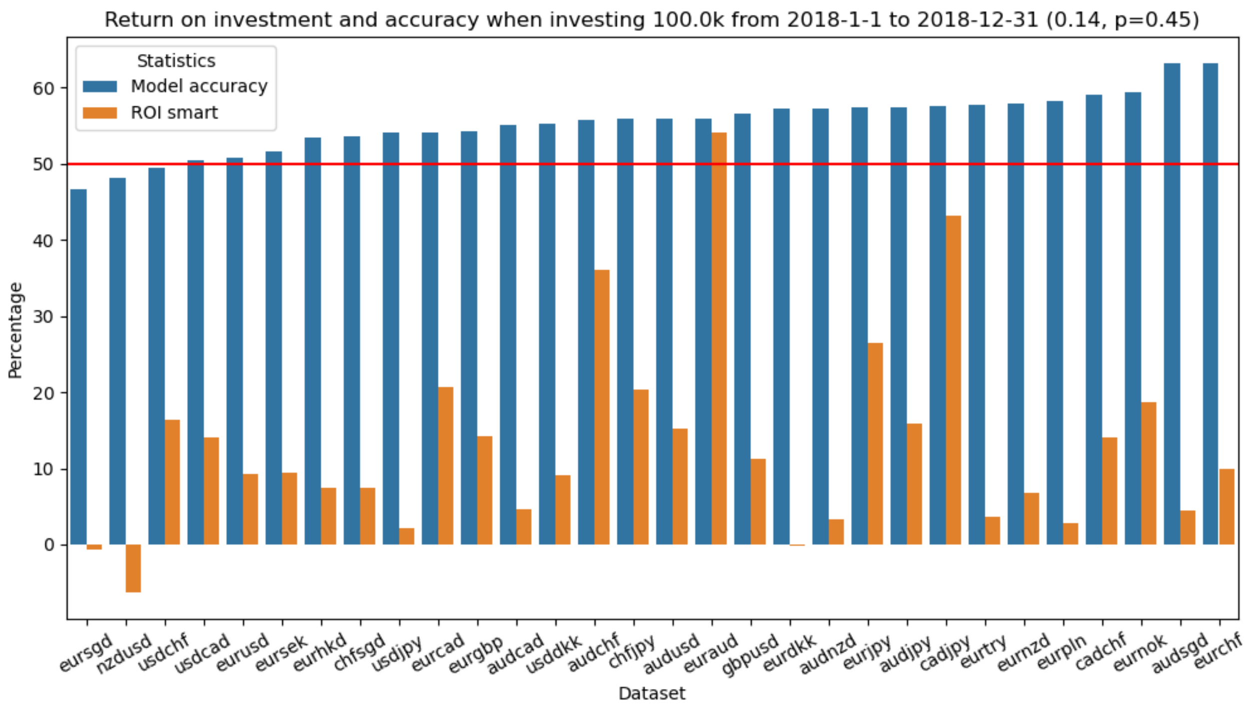

Figure 3: ROI and model accuracies when backtesting on 2018.

The Spearman's rank correlation coefficient is 0.14 with a p-value of 0.45, meaning there is no relation between the accuracy of the model and the resulting ROI. Table 9 shows whether there is a significant correlation between model accuracy and ROI in other years. Some years have a significant correlation, but this correlation seems to be mostly attributed to models with an accuracy lower than 50% making a loss.

	2012	2013	2014	2015	2016	2017	2018
Correlation	0.03	0.67	0.43	0.05	0.24	0.37	0.14
P-value	0.87	<0.01	0.02	0.77	0.21	0.05	0.45

Table 9: Correlation between model accuracy and resulting annualized ROI.

5.4 ROI vs. Volatility

Volatility is a measurement which represents how much a financial instrument swings in price. A volatile financial asset has the inherent property of being risky to trade with, but having a higher potential return than other financial assets. We used the Spearman's rank correlation test to test whether there is a significant correlation between ROI and volatility. The results of this measurement over 2017 are shown in Figure 4. The values of volatility are in arbitrary units.

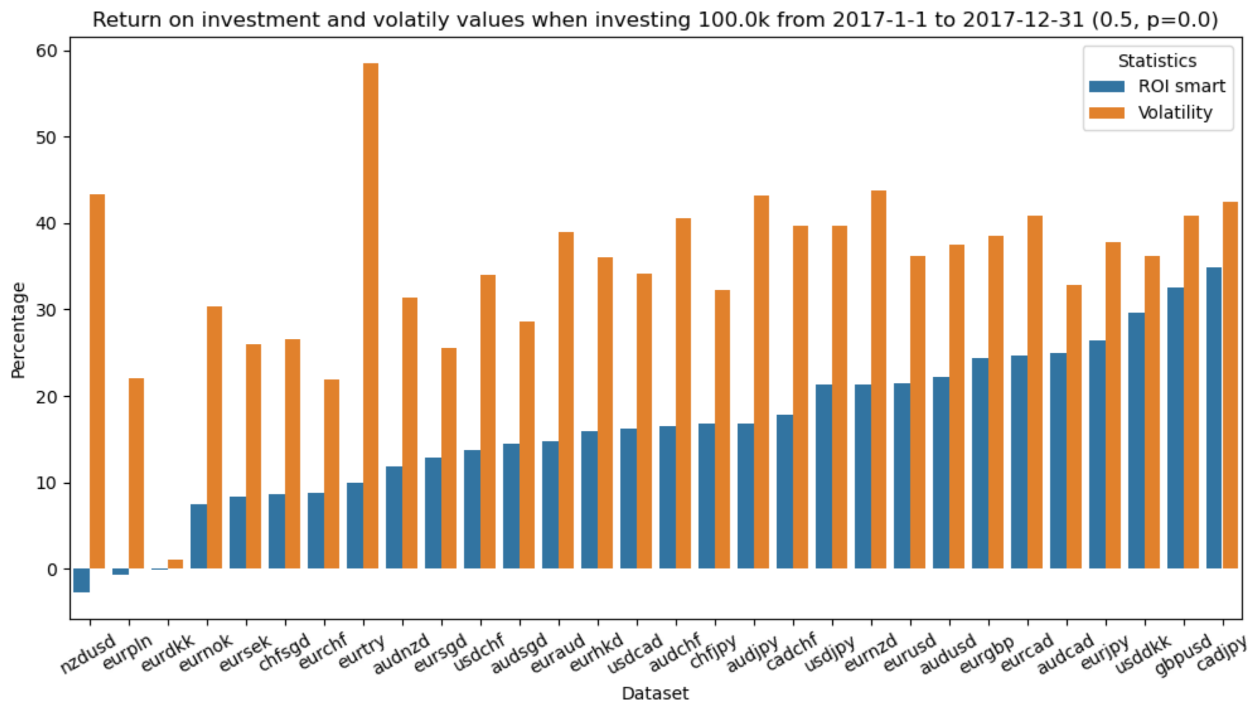

Figure 4: ROI and volatility when backtesting on 2017.

The Spearman's rank correlation coefficient is 0.5 with a p-value of <0.01, meaning that there is a significant correlation between the volatility and the ROI of the currency pairs in 2017. The correlation coefficient and corresponding p-value of this relationship in other years can be seen in Table 4.

	2012	2013	2014	2015	2016	2017	2018
Correlation	0.21	0.53	0.54	0.35	0.21	0.5	0.34
P-value	0.26	<0.01	<0.01	0.06	0.27	<0.01	0.06

Table 10: Correlation between volatility and annualized ROI.

6 Conclusions

In this section we will try to draw conclusions from the results described in Section 4 and Section 5. In the baseline results, we can see that adding technical indicators to the data increases the accuracy of the models, as the random forest and gradient boosting classifiers seem to outperform the baselines. This was also shown in previous work with a similar setup [23]. Generally we can see that the gradient boosting classifier seems to outperform random forest classifier.

6.1 Variable training time

In the results in Section 4.2 we can see that some models perform better when using less years of data into the past as training data. Some models seem to perform better if we use more years of data into the past as training data. The result of the Friedman test over all the accuracies of different training periods, we can see that the resulting p-value is 0.098. We can conclude that there is no optimal training period that makes the models for all currency pairs improve. We might be able to draw conclusions when we look at individual currency pairs.

6.2 Combining technical indicators of currency pairs

We tried combining technical indicators of currency pairs to predict a single trend one day into the future. In the results in Section 4.3 we can see that we would expect to see a negative effect two times more often than we would see a positive effect. If we use the advice generated by the experiments for 2017 and use it to test on 2018, we would see either an improvement of the model or no significant improvement for 60% of the cases. This does, however, include the insignificant model accuracy changes, which would have no ultimate effect on the accuracy. In 40% of the cases, we would see a negative effect, which could have a significant impact on our return on investment if we would use it in a backtesting strategy. This is because losses are probable when the accuracy of the model is below 50%.

We used the Auto-Sklearn classifier to make sure we are using an optimal model in terms of model optimizations such as feature selection and hyperparameter optimizations. When using the advice from July, August and September 2018 on October, November and December 2018 we only observed a significant decrease in all the models for both the EUR/CHF and the EUR/SGD currency pair.

From the previous experiments there does not seem to be a consistent method to improve the base prediction model of just a single currency pair with its technical indicators to predict the trend of the next day.

6.3 Backtesting

When backtesting our strategy on a certain year between 2012 and 2018, we used the other years as training data. The backtesting results are promising, with a mean ROI of 13.71% over all currency pairs in all years. There were better performers, such as EUR/AUD with a mean ROI of 25.72% over all years, and worse performers, such as EUR/DKK with a mean performance of -0.34%. We can see many different ROI values for individual currency pairs and for different years, and there does not seem to be an obvious consistency on within those results. We have found that a better prediction model does not necessarily lead to a higher return on investment, which would mean that there are other factors at play, which we have not yet been measured. It could, for example, be possible that the shape of the price trend over a certain year is an important factor for the resulting ROI.

We have found that there are years where there was a significant correlation between ROI and volatility, with a positive correlation coefficient. This would indicate that a higher volatility would lead to a higher ROI. This is, however, not the case for every year on which we backtested, so the relationship is not conclusive.

We can say that a mean ROI of 13.71% is a high enough percentage for practical trading and more research into improving this ROI could lead to more promising results.

6.4 Answering the research question

In this work, we wanted to answer the research question *'What is the effect of applying technical indicators of multiple foreign exchange currency pairs to predict the future price of a single currency pair and selecting different training periods to gain the highest accuracy in price prediction and subsequent trading profits?'*. The base performance we wanted to improve was shown by the work of Schut (2019) [23]. We tried to draw a conclusion on how we can optimize the input data and training period by varying the years into the past we used for training and by combining technical indicators of other currency pairs to exploit potential correlation. It seems that varying the training time might be beneficial for some currency pairs. It seems that combining currency pair technical indicators to improve model prediction accuracy is not practical in our experimental setup. To draw definitive conclusions on what could be beneficial for improving this prediction model, future research might want to focus on a single currency pair or a triangle of currency pairs, e.g. EUR/USD, USD/GBP and GBP/EUR.

7 Future research

Future research on this topic could include improving the prediction model or improving the backtesting strategy by improving the trading strategy or finding correlations that could predict a higher ROI based on certain price movements.

There are many promising models that can predict price movements, such as neural networks [13, 8], using genetic algorithms for feature selection [1] or using arbitrage between currency pairs [21]. The properties that are used in these different techniques can be used to create a model with a higher prediction accuracy. In Section 4.3 we have seen that more data does not necessarily lead to an improvement in prediction accuracy. There are other models that are correlated with the price trend of currency pair prices, such as the gold price, which is correlated with the Australian dollar [2]. Models could be trained on certain parts of multiple data sources to create a model that uses a voting system to predict the trend price, which could lead to a higher accuracy.

In addition to using different or multiple data sources with a variety of prediction models, fundamental analysis could be incorporated into the prediction model. With over 200 financial news articles per hour on each individual major currency, such as the Euro, the US Dollar and the Japanese Yen, there is enough data available to experiment with it being incorporated in this prediction model [4].

In terms of backtesting, not many trading strategies have been tested on this data with this prediction model. No leverage was incorporated in the strategy. Also, the maximum drawdown has not been analysed while using this trading strategy, so the results might underestimate the risk involved with this type of investment. However, the results are consistent over multiple currency pairs so this way of automatic trading seems promising for the future.

References

[1] G. Abreu, R. Neves, and N. Horta. Currency exchange prediction using machine learning, genetic algorithms and technical analysis. 05 2018.

[2] N. Apergis and D. Papoulakos. The australian dollar and gold prices. *The Open Economics Journal*, 6:1–10, 01 2013.

[3] A. Baasher and M. Fakhr. Forex trend classification using machine learning techniques. *Proceedings of the 11th WSEAS International Conference on Applied Computer Science*, pages 41–47, 01 2011.

[4] D. Chen, S. Ma, K. Harimoto, R. Bao, Q. Su, and X. Sun. Group, extract and aggregate: Summarizing a large amount of finance news for forex movement prediction. pages 41–50, Nov. 2019.

[5] J. Demsar. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7:1–30, 01 2006.

[6] T. Doan and A. Lo. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1:41–66, 02 1988.

[7] . L. R. Evans, M.D. How is macro news transmitted to exchange rates? *Journal of Financial Economics*, 88:26–50, 2008.

[8] C. Evans, K. Pappas, and F. Xhafa. Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation. *Advances in mobile, ubiquitous and cognitive computing" in the Mathematical and Computer Modelling Journal*, 58, 09 2013.

[9] E. A. Gerlein, T. Mcginnity, A. Belatreche, and S. Coleman. Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems with Applications*, 54, 02 2016.

[10] A. Hayes. Average true range - atr definition. https://www.investopedia.com/terms/a/atr.asp, 2019 (accessed August 8, 2020).

[11] A. Khadjeh Nassirtoussi, S. Aghabozorgi, T. Wah, and D. Ngo. Text mining of news-headlines for forex market prediction: A multi-layer dimension reduction algorithm with semantics and sentiment. *Expert Systems with Applications*, 42:306–324, 01 2015.

[12] V. Kondratenko and Y. Kuprin. Using recurrent neural networks to forecasting of forex. *arXiv.org, Quantitative Finance Papers*, 04 2003.

[13] K. Kuroda. Predicting optimal trading actions using a genetic algorithm and ensemble method. *Intelligent Information Management*, 09:229–235, 01 2017.

[14] J. Li and E. Tsang. Investment decision making using fgp: A case study. 2:1259 Vol. 2, 02 1999.

[15] I. R. Luciana Abednego*, Cecilia Esti Nugraheni. Forex trading robot with technical and fundamental analysis. *Journal of Computers*, 13:1089–1097, 2018.

[16] M. Maggini, C. L. Giles, and B. Horne. Financial time series forecasting using k-nearest neighbors classification. pages 169–181, 1997.

[17] S. F. M. L. Matthias Feurer, Katharina Eggensperger and F. Hutter. Automated machine learning with scikit-learn. https://github.com/automl/auto-sklearn.

[18] D. L. Padial. Technical analysis library in python. https://github.com/bukosabino/ta.

[19] A. Petropoulos, S. Chatzis, V. Siakoulis, and N. Vlachogiannakis. A stacked generalization system for automated forex portfolio trading. *Expert Systems with Applications*, 90, 08 2017.

[20] D. Rodriguez. A feature-rich python framework for backtesting and trading. https://www.backtrader.com/, 2020 (accessed July 29, 2020).

[21] F. Rundo. Deep lstm with reinforcement learning layer for financial trend prediction in fx high frequency trading systems. *Applied Sciences*, 9:4460, 10 2019.

[22] P. Schober, C. Boer, and L. Schwarte. Correlation coefficients: Appropriate use and interpretation. *Anesthesia Analgesia*, 126:1, 02 2018.

[23] F. Schut. Machine learning and technical analysis for foreign exchange data with automated trading, 2019.

[24] H. Talebi, W. Hoang, and M. Gavrilova. Multi-scale foreign exchange rates ensemble for classification of trends in forex market. *Procedia Computer Science*, 29, 06 2014.

A Neural network

The resulting accuracy values from the neural network described in Section 3.4 are described in Table 11. The mean accuracy of the neural network shows that the neural network does not improve over the classifier baselines described in Section 4.1.

Currency pair	Accuracy
AUD/CAD	0.5
AUD/CHF	0.49
AUD/JPY	0.48
AUD/NZD	0.5
AUD/SGD	0.5
AUD/USD	0.55
CAD/CHF	0.49
CAD/JPY	0.54
CHF/JPY	0.52
CHF/SGD	0.52
EUR/AUD	0.52
EUR/CAD	0.53
EUR/CHF	0.52
EUR/DKK	0.52
EUR/GBP	0.52
EUR/JPY	0.5
EUR/NOK	0.54
EUR/NZD	0.5
EUR/PLN	0.56
EUR/SEK	0.52
EUR/SGD	0.5
EUR/TRY	0.5
EUR/USD	0.52
GBP/USD	0.54
NZD/USD	0.52
USD/CAD	0.47
USD/CHF	0.49
USD/DKK	0.52
USD/JPY	0.5
Mean	0.51
Max	0.56
Min	0.47

Table 11: Currencies used in trend prediction, currency combination and backtesting