Abstract

A review of the representative models of machine learning research applied to the foreign exchange rate and stock price prediction problem is conducted. The article is organized as follows: The first section provides a context on the definitions and importance of foreign exchange rate and stock markets. The second section reviews machine learning models for financial prediction focusing on neural networks, SVM and evolutionary methods. Lastly, the third section draws some conclusions.

Key words: Stock price prediction, machine learning, artificial neural networks, kernel machines and evolutionary methods.

1. Introduction

One of the main topics of interest in the field of finance is the value of the so-called financial assets. Different from real assets, financial assets are not directly used in the production of goods. They are claims on real assets representing the most developed stage of an economic system, which implements an explicit separation between property and usage of means of production. Financial assets play a
central role on the economic flow of resources. They allow scarce resources to easily migrate from less to better economic productive activities.

Financial assets can be divided in terms of the claim they encapsulate. Some of the essential financial assets include stocks and currency pairs. Stocks entitle their owners to receive residual economic profit from business entities called companies. Stocks derive their value from the potential cash flow generated by the means of production owned by the corresponding company. As an example, Microsoft, Google or Intel stocks, give the right to the owner of claiming the economic benefit produced by those companies. In contrast, bonds are considered debt securities. Bonds represent a credit stake over the business entity issuing them. Bondholders have the right to claim from the debtor the total lent resources plus an opportunity cost recognition commonly called interest payments. The main difference between stocks and bonds is the residual feature of stocks. Bonds mainly derive their value from debtor’s probability of default because bondholders do not participate on issuer’s profit generation capacity.

On the other hand, a currency pair refers to the relative value of one currency related to another currency. Before 1970, currency pairs (FX) markets were underdeveloped due to governments’ restriction and a fully negotiated monetary order established in 1944, the Bretton Woods. After Bretton Wood’s failure in 1971, main currencies moved to a flotation standard. Foreign exchange rates were not defined by well-established international agreements but by market forces. Thus, conversion rates started to be settled based on relative purchasing power of both currencies, commercial issues, government budget deficit levels, economic productivity or simply, demand and supply of corresponding currencies. According to the BIS (Bank for International Settlement), main FX markets are EUR/USD and USD/JPY, which represent 42% of the US$4 trillion daily transaction volume.

1.1 Predictability on Foreign Exchange Rates and Stock Prices

Several empirical studies have confirmed considerable predictability levels on the foreign exchange and stock markets. For example, [33] showed that there exists evidence of price prediction on short horizons. This prediction can be connected to the type of strategy used by agents extrapolating in future prices.

On this context, [19], [18], [20], [21], [3] and [44] showed that investors tended to use chartist strategies for prediction at short periods of time and fundamentalist for long ones. As observed by [18] p. 264, «It may be that each respondent is thinking to himself or herself, I know that in the long run the exchange rate must return to the equilibrium level dictated by fundamentals. But in the short run, I will ride the current trend a little longer. I only have to be careful to watch for the turning point and to get out of the market before everyone else does». These findings are still valid today.

Extra predictability is also supported by performance of institutional investors specialized on exploiting market opportunities. The Hedge Fund Weighted Composite Index

---

1 The Bretton Woods agreement promoted free trade and new economic consensus based on fixed exchange rates between developed countries' currencies and between the dollar and gold. The dollar stood as the new reserve value currency. The Bretton Wood agreement collapsed in 1971 when the US started to finance the Vietnam War printing dollars with no gold backup.

2 EUR,USD, JPY stands for Euro,Dollar and Japanese Yen respectively.

3 Chartism states that historical prices and their transformation can be an informative source for predicting future price performance. On the other hand, fundamentalism looks at company’s or real asset’s performance to predict future financial asset prices.
Javier Sandoval

(HFRIFWI) summarizes the performance of the hedge fund industry, private, actively managed institutional investors that exploit sophisticated strategies. As shown in Figure 1, from 1992 to 2010, the index has increased 530%. In contrast, the S&P500 has gained, including dividends, 310%. This extra return cannot be explained by high-risk exposure. The standard deviation of HFRIFWI’s annual returns has been 12.08% compared to 20.04% on the S&P500 index. Observing the maximum drop-back over a year on both indexes, 12% and 20% respectively, further supports this conclusion.


Studies in machine learning are concerned with the ability to learn patterns, adapt behaviors and make intelligent decisions that are not explicitly programmed in computer systems. These systems usually learn from historical or online data that is presented to them. As mentioned by [39], much of the classic, knowledge-based work in artificial intelligence does not appear in the recent work of financial price prediction applications. Pure knowledge-based models applied to financial price prediction were rapidly abandoned and replaced by machine learning, data driven techniques. In the following subsections a review of machine learning techniques applied to the financial asset price prediction domain is described.

The review is focused on three main subareas: neural networks, Support Vector Machines (SVM) and evolutionary methods. Science Direct was used as the main source of information sorted by relevance. Studies are presented following a chronological timeline that shows how techniques have evolved in terms of complexity and input datasets.

2.1. Artificial Neural Networks (ANNs)

One of the first models applied to price prediction in computational intelligence were Artificial Neural Networks (ANNs). ANNs are a family of mathematical methods inspired by the function of human neural system, and the way neurons interact with each other. First publications on application of ANNs to price prediction can be traced back to 1988 to the work in [47] which focused on ANNs for predicting the daily value of IBM stock based on historical prices. Although their results were not astonishing, yet it is considered
a seminal paper for a list of countless studies devoted to the topic of market predictability using non-traditional techniques.

The ability to predict the daily values of the S&P500 using ANNs was explored in [45], focusing on a market index. This article gave evidence of how a combination of rule-based expert system techniques and an ANN outperformed a simple passive strategy. Prediction in the short term of other financial assets has been the subject for different ANN models. In [23] the ANN’s ability to forecast the S&P500 index was studied again and they also tested ANN’s predictions on Gold futures. The latter article not only used historical prices as input data but also included open interest patterns, which were thought to incorporate the beliefs of a majority of the traders in the corresponding market.

Before 1997, articles on ANN applications to financial forecasting showed primal frameworks such as Multilayer Perceptron (MLP) and mostly used historical prices on US stock markets. After 1997, researchers began to explore new sets of input data and more sophisticated frameworks.

With regard to foreign exchange markets, several ANNs models have been proposed. In [22] a model based on ANNs is described to predict four major exchange rates, which included the Dutch Guilder. Over the sample data, the paper reported unfavorable in-sample fits or forecast performance. On the contrary, in [25] evidence was found of superior forecasting performance of ANNs over Random Walk predictions on major foreign exchange markets. This kind of contradictory conclusions were reached every time simple ANN models fed with historical prices, were constructed.

Researchers moved forward and put together ANNs and technical analysis indicators as in [41]. That paper abandoned simple historical price datasets and implemented an MLP-ANN towards prediction of major foreign exchange rates including the CHF/USD and JPY/USD, using typical inputs from chartist. Prediction power was compared to ARIMA’s performance. The authors found that on almost every market the ANN outperforms the ARIMA benchmark. Other articles which combined technical indicators and ANNs are [5] and [4].

Input data was also extended to cover international stock markets others than the US. For example, [16] implemented an ANN-type model in order to forecast the general index of the Madrid Stock Market. Authors found that developed trading rules had a superior performance than a buy-and-hold strategy especially on «bear» and «stable» market episodes. Another study on international markets was that of [36] which used accounting ratios to predict stock prices on the Canadian market.

Other studies on ANNs applied to price prediction combine Radial Basis Function Neural Networks (RBF) with dimensionality reduction techniques [1] and local linear wavelet neural network (LLWNN) together with an estimation of distribution algorithm (EDA) [14].

2.2. Kernel Machines

Following the definition of [12] (page 655), kernels methods can be regarded as machine learning techniques which are «kernelised» versions of other fundamental machine learning methods, that is, learning algorithms incorporating a positive semidefinite kernel function representing similarities between input vectors in a transformed space.
In finance and especially in price prediction, research has focused on the seminal and most widely known kernel machine: the Support Vector Machines (SVMs). One of the main drawbacks observed on implementing SVM models in price prediction is that financial price series are highly noisy and non-stationary. The latter means that the pattern relating input and output variables changes over time and since the SVM assumes the data is explained by a static prediction rule, it may be unsuitable for an accurate financial forecast.

Before 2000 there is no evidence to our knowledge of applications of SVMs to the price prediction problem. One of the first contributions was [11]. This paper proposed to overcome the over-fitting problem usually found on some ANNs, with the use of SVMs. Compared to a MLP trained with the back-propagation algorithm over historical and chartist indicators of the S&P 500, results showed that an SVM obtained better predictions mainly due to SVM’s empirical risk minimization principle which minimize an upper bound on the generalization error rather than on the training error.

Extending their previous results, [43] presented the same year a model that combines an SVM with a self-organized feature map (SOM). First, the SOM was used to partition the input data space into disjoint regions. Then, multiple SVMs were trained within the partitioned regions. The models were tested on prices from five future contracts. The authors reported a higher performance when compared to predictions from a single SVM. These two works marked the beginning of multiple publications devoted to SVMs applied to price prediction.

In order to elude the static assumption of the prediction rule in the SVM, [42] and [10] suggested the implementation of an Adaptive SVM model. Such extension has been motivated by noting that in non-stationary financial prices, most recent data provides the most relevant information. Therefore, the authors proposed to give higher weights to more recent observations and to allow the SVM parameters (regularization coefficient and kernel parameter) to change over time as new observations are given to the model. They found better performances when compared to a back-propagation MLP and a regularized RBF ANN using historical prices and chartist information for prediction on the S&P500.

Although for general-purpose forecasting, SVM algorithms implemented before 1999 were generally slower than ANNs with similar generalization performance ([24], p. 345), significant improvements were achieved in the following years.

A comparative SVM vs. ANN study [26] over the NIKKEI225 series was conducted on their ability to forecast the series direction over a week time interval. It was showed that in terms of the hit ratio the SVM model outperformed the back-propagation ANN as well as a random walk and a linear discriminant analysis model. Other papers that show evidence of superior performance of SVM models over ANNs are [27] and [40], the latter regarding prediction in an electricity market.

Other comparative study worth to mention is [29] which examined SVM performance compared to a back-propagation ANN and a case-based reasoning (CBR) model. In such study SVM provided better predictions when tested on the daily Korea composite stock price index (KOSPI).

It is interesting to mention that despite of the promising results of SVMs in finance forecasting, other studies have suggested that some knowledge extraction techniques
may improve performance of these models, as a pre-processing step for input data. A good example of this approach is given in [28] which proposed combining a data-driven technique, i.e. SVMs, together with parametric models such as integrated moving average (ARIMA) and vector autoregressive (VAR) models in order to perform input variable selection. These findings showed that filtering relevant input variables with the parametric model helped improve the prediction power of the SVM.

Another example of an SVM extension using variable selection is developed in [31]. That paper used a filtering and wrapping method called F-score and Supported Sequential Forward Search (FSSFS) for feature selection. Again, better performance was obtained as a result of coupling the SVM with variable selection, in this case for the prediction of the daily value of the NASDAQ index using as raw inputs a set of technical analysis variables.

Filtering and wrapping is only one of a number of data pre-processing options proposed in the literature. A second alternative was dimensionality reduction using well-known algorithms such as independent component analysis (ICA) as in [35] which focused on reducing noise in input data before calibrating the SVM model. Specifically, the authors proposed a two stage modeling approach that firstly implemented an ICA algorithm in order to generate independent components from input variables. ICA-SVM was tested on Nikkei 225 and TAIEX indexes obtaining a better performance than a regular SVM and a random walk model.

So far all described SVM models have focused on historical prices, technical analysis and other macroeconomic variables. To the best of our knowledge, the only paper experimenting with the limit order book as an input set for an SVM aimed to predict the market’s directional changes is [17]. That study proposes the implementation of multiple kernel learning methods (SimpleMKL and LPBoostMKL) to train a multiclass SVM model in order to predict EUR/USD changes based on cumulative volumes shown in the limit order book. The authors found promising results compared to a regular SVM’s prediction.

2.3. Evolutionary Methods

As defined by ([8], page 37), evolutionary methods comprise a series of algorithms inspired by the theory of the evolution of the species, including Genetic Algorithms (GA), Genetic Programming (GP) and Grammatical Evolution (GE). Application of these models to prediction of financial assets has been conducted over two main avenues: direct evolution of trading systems aimed to exploit price prediction capabilities and optimization of ANNs topologies or parameters with evolutionary techniques.

One of the first works to propose an evolutionary technique to tackle the price prediction problem was [15]. In that work, the author proposed an ANN model calibrated with a GA using the maximum profitability/maximum drawdown ratio as a performance variable to be optimized in a training dataset. Other study using GA to train ANNs can be found in [46].

On the other hand, there is no clear evidence of superior performance of GP models when applied to the price prediction domain. One of the first works in this topic was [13] which studied the Efficient Market Hypothesis (EMH, as defined in section 1). The authors used GP to evolve math expressions representing trading strategies in order to predict

---

4 The limit Order Book stores limit orders, orders which have specific execution prices.
returns on the S&P500 and the TAIEX (a Taiwan stock exchange index). The paper concluded with no indication of a superior performance of the GP-based model compared to a buy-and-hold strategy. Other works suggesting the same assertion are [2] where technical rules on S&P500 values from 1928 to 1995 were evolved using GP, and [38] which studied Canadian individual stocks using historical prices and transaction volumes.

In contrast, the GP-based model for foreign exchange markets reported in [32] showed that GP is able to find better strategies than buy-and-hold. They also gave evidence that the superior return was not due to a greater exposure to risk. This claim is also supported by [6] where the proposed fitness function considers consistency of performance and takes into account the transaction cost effect. Similar conclusions were drawn recently in [34].

A different form of evolutionary technique is grammatical evolution (GE). GE as defined by [8] (page 73), is a grammar-based form of GP which extends the biological analogy by considering a distinction between the genotype and phenotype. Studies using GE in the price prediction domain were presented for the first time in [37]. That paper proposed discovering useful technical trading techniques with GE using data from the FTSE 100 index for the period 26/4/1984 to 4/12/1997. The results presented showed great potential of GE based models. GE models were also tested in foreign exchange series in the same paper.

As a final note, [9] presented a GE model tested over the GBP/USD daily series from 1993 to 1997. Technical indicators were evolved using a grammar-defined model. Experimental results showed that the GE model was able to beat buy-and-hold strategies even on out of sample data using a fitness function, which emphasized excess returns and the maximum drawdown of each trading strategy. Another study specialized on GE focused on exchange rate prediction was [7].

3. Conclusions

Research on computational intelligence applied to financial price prediction can be traced from the last part of the 80’s. First works proposed ANNs as the main modeling technique and linear models as benchmarks. In 2000, SVM began to supersede ANNs. This trend is mainly explained by superior generalization and parameter estimation properties in SVMs. In contrast, evolutionary methods have transversally accompanied ANN and SVM methods focusing on tuning previous models and trying to evolve trading rules independently. Based on the reviewed literature, there is a tendency of dominance for SVMs over ANNs when compared in terms of performance and calibration feasibility.

Our review revealed that technical analysis has been mainly used as the input domain for the majority of models of financial price prediction. If one regards technical analysis variables representing transformations in assets price, then it can be stated that historical transaction prices have been the main source of information for the machine learning price prediction models reviewed above. Other sources of information used in these studies have included trading volume and order book variables. Because of information limitations, limit order features are mainly confined to best bid and offer and fixed interval liquidity levels. Despite the richness of information that can be found in the order book, no other elements have been used for the purpose of machine learning financial prices prediction to the best of our knowledge.
Another finding of this review is that research on price prediction has focused on developed markets. In fact, we only found studies about financial market data from the USA, Europe, and Southeast Asia including Japan. This geographical bias might be due to the availability of information and the degree of maturity of local financial markets.

As a concluding remark, two paths for future extensions are observed. Based on input data, there exist opportunities on applying standard computational techniques on the financial price prediction problem in developing economics such as Brazil, Colombia, and Mexico. It will be relevant and interesting to test standard findings such as dominance relationships in new datasets. On the other hand, this review also notices the absence of expert knowledge on studied works. One alternative will be to extend mentioned models in order to include this new element. Other alternative is to focus on new techniques that naturally incorporate the concept of expert knowledge, such as Bayesian Networks.

References