

Artificial Counselor System For Stock Investment

Hadi NekoeiQachkanloo*, Benyamin Ghojogh*, Ali Saheb Pasand*, Mark Crowley

Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada

{hnekoeiq, bghojogh, ali.sahebpasand, mcrowley}@uwaterloo.ca

Abstract

This paper proposes a novel trading system which plays the role of an artificial counselor for stock investment. In this paper, the stock future prices (technical features) are predicted using Support Vector Regression. Thereafter, the predicted prices are used to recommend which portions of the budget an investor should invest in different existing stocks to have an optimum expected profit considering their level of risk tolerance. Two different methods are used for suggesting best portions, which are Markowitz portfolio theory and fuzzy investment counselor. The first approach is an optimization-based method which considers merely technical features, while the second approach is based on Fuzzy Logic taking into account both technical and fundamental features of the stock market. The experimental results on New York Stock Exchange (NYSE) show the effectiveness of the proposed system.

Introduction

Consider the situation of an investor, with an existing budget, who wants to know how to divide the budget between several existing stocks. The fraction of budget invested in stock i is a weight denoted by w_i in range $[0, 1]$. The problem is to design an artificial financial counselor system for suggesting the optimum weights for investing in the stocks. In order to find the optimum weights, the next day prices of the stocks need to be estimated. Therefore, this paper firstly tries to predict the future prices of stocks based on their previous behaviour which can be modeled by time series. The primary goal of this project is to find the best weights according to the predicted prices and fluctuations of every stock as well as the risk tolerance of the investor. The overall structure of the proposed system is depicted in Fig. 1. In general, there are n stocks each of which has five time series of technical features and several time series of fundamental features. Two indices, namely Average Directional Index (ADX) and Parabolic Stop and Reverse Index (SAR), are calculated from the technical features of stocks. The time series are first preprocessed then a time series predictor forecasts the future prices. The variances (risks) of time series and the future prices go to the next stage computing the optimum weights based on risk tolerance of the investor. That

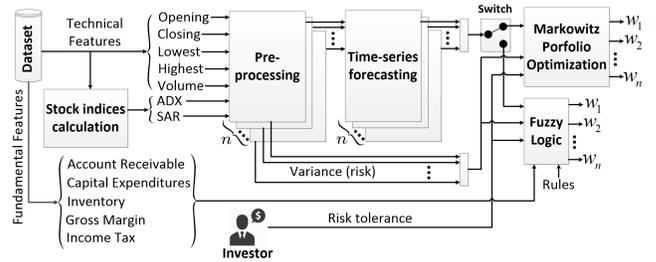


Figure 1: Overall structure of the proposed system

stage can be done in two approaches which are Markowitz portfolio theory (Bodie, Kane, and Marcus 2014) (an optimization problem with the purpose of maximizing the profit) and fuzzy investment counselor (fuzzy logic). The latter considers both technical and fundamental features providing opportunity for other sentimental features like an expert broker.

Two approaches are commonly used to analyze the financial behaviour of the markets, (I) technical analysis, and (II) fundamental analysis (Cavalcante et al. 2016; Atsalakis and Valavanis 2009b). The former considers technical attributes which are the history of raw prices of stocks. Some economists believe that all the information needed to predict financial behaviour of the market exist in the prices. Fundamental analysis considers the economic factors, such as liabilities, size of the firm, expenses, assets, revenue, etc, affecting the economic movements of market (Vanstone and Finnie 2009; Abarbanell and Bushee 1997).

There have been developed different machine learning methods for time series prediction (Makridakis, Spiliotis, and Assimakopoulos 2018; Cavalcante et al. 2016). Two main methods are commonly used for stock time series prediction (Cavalcante et al. 2016), which are Multi-Layer Perceptron (MLP) (e.g., Ican and Çelik 2017; Kayal 2010) and Support Vector Regression (SVR) (e.g., (Cao and Tay 2003; Chen 2010)). In this paper, the new version of the New York Stock Exchange (NYSE) dataset (<https://www.kaggle.com/dgawlik/nyse/data>), including both technical and fundamental features, is utilized. In the literature, there are several works attempting to forecast the previous versions of the NYSE dataset, such as (Atsalakis and Valavanis 2009a; Halliday 2004) and the new version of it (Song 2018). These works mainly focus on trend prediction

*The first three authors contributed equally to this work.

of stocks. For instance, a neuro-fuzzy controller is proposed in (Atsalakis and Valavanis 2009a) for price prediction. The thesis (Song 2018) is one of the very recent works on this dataset. Our work has a comparative performance with it in terms of hit rate. A more important advantage of our system against these works is proposing a complete trading system which recommends the optimum weights of investments as well as predicting the stock exchange, while those works merely attempt to forecast the stocks. Our main contributions are threefold: (I) We propose a complete, novel trading system which not only predicts the future stock prices but also suggests the best fractions of budget for investments. According to (Cavalcante et al. 2016), there is no such thing as a complete and well-established system proposed in the literature. (II) We consider both technical and fundamental attributes of the stocks. (III) We use two perspectives, i.e., optimization and artificial intelligence, to address the trading problem.

Methodology & Implementation

Preprocessing

Three types of preprocessing will be applied on the dataset. The first preprocessing for time series prediction is moving average which smooths the small fluctuations of stock helping to better capture the trend (Cavalcante et al. 2016). We use moving average on technical features, i.e., opening, closing, lowest, and highest prices, and volume time series, $s'(t) = \frac{1}{\Delta_{ma}} \sum_{k=0}^{\Delta_{ma}-1} s(t-k)$, where $s(t)$ and $s'(t)$ denote the data and average data over the period Δ_{ma} days for day t , respectively. The Δ_{ma} is 50 in this work to capture both short-term and long-term trend changes (Fletcher 2012).

The second preprocessing for time series prediction is Z-score normalization which removes the mean and scales the variance to unit. Every stock has several time series, which are opening, closing, lowest, and highest prices, volume, and the ADX and SAR indices. In each series, let $\hat{s}(t)$ denote the normalized data in day t . We define Δ_p as the size of normalization window. The normalization is $\hat{s}(t) = (s'(t) - \overline{s'(t)}) / \text{std}(s'(t))$, where $\overline{s'(t)}$, and $\text{std}(s'(t))$ are mean and standard deviation of averaged time series in range $[t - \Delta_p, t]$, respectively. Note that after time series forecasting, data is denormalized by reversing the transformation.

Furthermore note that, in economics, the profit rate (relative change of budget) is mostly used rather than the raw budget. Therefore, for the input of weight suggestion modules, the averaged time series is converted to profit rates by $r_i(t) = (s'(t+1) - s'(t)) / s'(t)$, for every day indexed by t in stock i . Henceforth, $r_i(t)$ is denoted by r_i for simplicity, where r_i means the return of the target future day. Note that the return implies the meaning of profit rate in this work.

Prediction of Future Prices

It has been shown that usage of stock indices is helpful in prediction of individual stock prices (Song 2018). In this work, two well-known stock indices, Average Directional Index (ADX) and parabolic Stop and Reverse (SAR) (Wilder 1978), are calculated and used for prediction as well as the

technical features. In the following, calculation of these two indices are explained and the prediction is detailed.

Average Directional Index The ADX index measures the strength of stock trend regardless of its direction (Wilder 1978). For its calculation, first True Range (TR) is found by $\text{TR}(t) = \max(H(t) - L(t), |H(t) - C(t-1)|, |L(t) - C(t-1)|)$, where H , L , and C denote the highest, lowest, and closing prices of a stock respectively. The smoothed TR (STR) is $\text{STR}(t) =$

$$\begin{cases} \text{invalid} & \text{if } t \leq \tau \\ (1/\tau) \sum_{i=1}^{\tau} \text{TR}(i) & \text{if } t = \tau + 1 \\ \frac{1}{\tau} [(\tau - 1)\text{STR}(t-1) + \text{TR}(t)] & \text{otherwise,} \end{cases} \quad (1)$$

where τ is the smoothing period and is set to 14 according to (Wilder 1978). The Plus Directional Movement (PDM) and Minus Directional Movement (MDM) measure trend direction over time and are calculated as $\text{PDM} = \max(H(t) - H(t-1), 0)$ and $\text{MDM} = \max(L(t-1) - L(t), 0)$. The smoothed PDM (SPDM) and smoothed MDM (SMDM) are calculated using a similar approach to Eq. (1) but with replacing TR with PDM and MDM, respectively. The Smoothed Plus Directional Indicator (SPDI) and Smoothed Minus Directional Indicator (SMDI) are then calculated as $\text{SPDI}(t) = 100 \times \frac{\text{SPDM}(t)}{\text{STR}(t)}$ and $\text{SMDI}(t) = 100 \times \frac{\text{SMDM}(t)}{\text{STR}(t)}$. Afterwards, the Directional Index (DX) is found as $\text{DX} = 100 \times \frac{|\text{SPDI}(t) - \text{SMDI}(t)|}{\text{SPDI}(t) + \text{SMDI}(t)}$. The Average DX (ADX) is obtained using a similar approach to Eq. (1) by replacing TR with DX and τ with 2τ since $t \leq \tau$ is already invalid for calculating SPDM and SMDM.

Parabolic Stop and Reverse Index The SAR index captures the overall trend of stock over the time. It includes information of stop/reverse points where the trend is changed from uptrend (UT) to downtrend (DT) or vice versa (Wilder 1978). The SAR index is $\text{SAR}(t) =$

$$\begin{cases} \text{invalid} & \text{if } t < 4 \\ \min\{L(t-i) | i = 0, \dots, 3\} & \text{if } t = 4, \text{ UT} \\ \max\{H(t-i) | i = 0, \dots, 3\} & \text{if } t = 4, \text{ DT} \\ \text{SAR}(t-1) + \text{AF}(t-1) \times \\ (\text{EP}(t-1) - \text{SAR}(t-1)) & \text{otherwise,} \end{cases} \quad (2)$$

where EP is the Extreme Point and found by $\text{EP}(t) =$

$$\begin{cases} \text{invalid} & \text{if } t < 4 \\ \max\{H(t-i) | i = 0, \dots, 3\} & \text{if } t \geq 4, \text{ UT} \\ \min\{L(t-i) | i = 0, \dots, 3\} & \text{if } t \geq 4, \text{ DT,} \end{cases} \quad (3)$$

and AF is the Acceleration Factor and starts from 0.02 and is increased by a step of 0.02 whenever the EP changes. Note that the AF is saturated to 0.2 if it reaches 0.2 and it is reset to 0.02 when the trend changes from UT to DT or vice versa.

Prediction To have a visual sense of technical features (opening, closing, highest, and lowest prices) and the SAR, SPDI, SMDI, and ADX indices, see Fig. 2. This figure depicts the candle plot of a stock in a time span. The candle plot is a useful visualization of the opening, closing, highest, and lowest prices as well as the trend of stock (green for

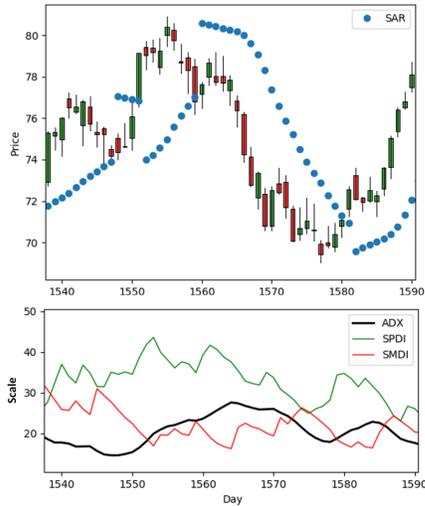


Figure 2: The candle plot of a stock and its indices.

increasing and red for decreasing). The end points of every candle show the highest and lowest prices of the day and the middle points determine the opening and closing prices. The SAR index is also shown in Fig. 2. It can be seen that this index is showing the trend of the stock over time. The SPDI and SMDI indices from which ADX index is found are also depicted. The ADX index is capturing the strength of the stock trend. For example in days 1550 to 1570, it is showing that stock trend is strong because the difference of SPDI and SMDI is noticeable.

The future prices of stocks are to be predicted using a time series forecasting method. In this work, SVR is used for prediction of stock time series. The utilized kernel function is the radial basis function $k(\mathbf{x}_n, \mathbf{x}_m) = \exp(-\gamma \|\mathbf{x}_n - \mathbf{x}_m\|_2^2)$, where γ , which is found to be 0.001 by validation, determines the bias of trained model, and \mathbf{x}_n and \mathbf{x}_m are training and testing data, respectively. The C variable, controlling the amount of penalizing slack variables, is found to be 1000 by validation. In order to prepare the training and testing data, the technical features (opening, closing, highest, and lowest prices and the volume) as well as the ADX and SAR indices throughout the days within a window with size Δ_p are concatenated to form a vector. We remove the first 2τ days from the training data because of invalidity of indices. In this work, the highest price of stocks are predicted and used for suggestion of weights, as explained in the next section because the highest price has a significant impact on investors to invest in a stock or not.

Suggestion of Weights

Markowitz Portfolio Theory The first approach for suggesting investment weights is using Markowitz portfolio theory (Bodie, Kane, and Marcus 2014). The Markowitz portfolio theory can be formalized as a regularized quadratic op-

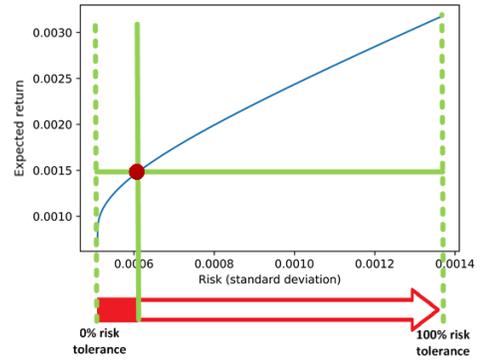


Figure 3: The frontier curve of Portfolio theory.

timization problem:

$$\begin{aligned} & \text{minimize}_{\mathbf{w}} \quad (1/\mu)\mathbf{w}^\top \mathbf{S} \mathbf{w} - \mathbf{w}^\top \mathbf{r} \\ & \text{subject to} \quad \mathbf{w} \succeq 0, \quad \mathbf{w}^\top \mathbf{1} = 1, \end{aligned} \quad (4)$$

where $\mathbf{w} = [w_1, \dots, w_n]^\top$ is the vector of weights whose i -th element determines the portion of budget to be invested in stock i . Vector $\mathbf{r} = [\mathbb{E}(r_1), \dots, \mathbb{E}(r_n)]^\top = [\tilde{r}_1, \dots, \tilde{r}_n]^\top$ includes the expected value of returns of all the stocks where \tilde{r}_i denotes the predicted return. The matrix S is the covariance matrix whose element (i, j) is $\mathbb{E}[(r_i - \bar{r}_i)(r_j - \bar{r}_j)]$ where \bar{r}_i denotes the mean of r_i in a window of size Δ_c . The cost function in Eq. (4) consists of two parts, i.e., $\mathbf{w}^\top \mathbf{S} \mathbf{w}$ and $\mathbf{w}^\top \mathbf{r}$. The former represents the overall variance (risk) of investment and the latter is the overall expected return of investment. The goal of portfolio theory is to maximize the overall expected return of investment while minimizing the overall risk. Note that the weights are positive and they should sum up to one, $\sum_{i=1}^n w_i = 1$. The μ in Eq. (4) is the regularization parameter. The higher the μ , the less the variance is penalized and thus the higher risk is allowed to take for investment.

We solved this quadratic optimization problem using Python software for convex optimization (CVXOPT) (<https://cvxopt.org/index.html>). The value of $\mu > 0$ is swept from a very small number in order to obtain the frontier curve of portfolio theory (Bodie, Kane, and Marcus 2014) which is a curve of expected return versus standard deviation (risk) of investment (see Fig. 3). This sweeping can be stopped whenever the curve reaches the maximum possible risk of investment which is equal to the risk of the stock having maximum variance. In other words, whenever $\mathbf{w}^\top \mathbf{S} \mathbf{w} = \max(\text{diag}(S))$.

The maximum risk tolerance of the investor, denoted by $\eta \in [0, 1]$, is translated to the range between minimum and maximum standard deviations (risks) of the generated frontier curve of portfolio (see Fig. 3). The point on the curve having the largest expected return up to that risk level is considered as the optimum portfolio point corresponding to a specific \mathbf{w} . For calculating S , the Δ_c , which is 100 in this work, should be large enough to capture the true risk of the stock but not so large as to include old and invalid risks.

Fuzzy Investment Counselor In this work, a Fuzzy Investment Counselor (FIC) is proposed to model the be-

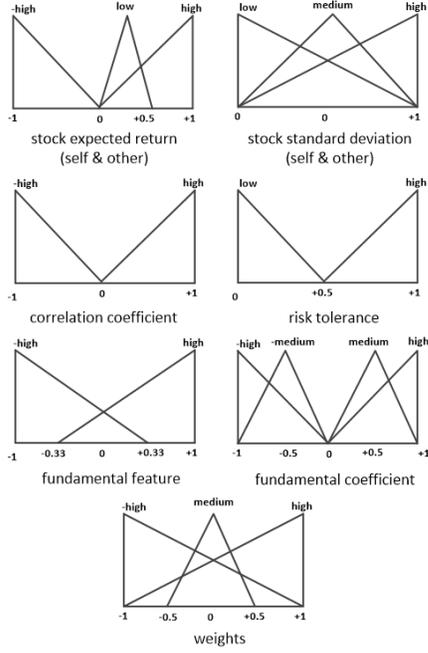


Figure 4: Membership functions of fuzzy variables.

haviour of an expert investor under a rationality assumption according to predefined rules. Note that this module takes both technical and fundamental attributes into account while portfolio theory is merely based on technical features.

Parts of FIC: The FIC consists of two parts, i.e., technical and fundamental parts. The technical part includes two separate fuzzy systems. In the technical part, the first fuzzy system, called **self-stock system**, considers each stock individually, while the second fuzzy system, **pairwise-stocks system**, considers the effect of all the other stocks on every stock. The fundamental part considers fundamental features of each stock.

Settings of Fuzzy Logic: The Mamdani fuzzy model (Karray and De Silva 2004) is used for the FIC. The membership functions for fuzzy technical and fundamental variables are shown in Fig. 4. The singleton and centroid methods are used for fuzzification and defuzzification, respectively. Note that fuzzification converts crisp variables to fuzzy qualitative values, and defuzzification does the reverse (Klir and Yuan 1995). The fuzzy rules of both parts are illustrated using tree structures in Fig. 5. Every rule in the illustrated tree is a path from root to a leaf where fuzzy T-norm is applied on the non-leaf nodes and the leaf is the output of the rule. To better understand these trees, two sample rules from Fig. 5b are explained here: “IF $E_{other}(r)$ is HIGH and ρ is HIGH and σ_{other} is MEDIUM and η is LOW / HIGH, THEN w is MEDIUM / HIGH.” The intuition of this rule is that if the correlation of two stocks is high and the return of one of them is going up with medium risk, it is better to invest in the other stock. However, it depends on the risk tolerance of the investor to give a large or medium weight to investing in that stock. The standard operators (minimum and maximum respectively for T-norm and S-norm) are used in this work.

Algorithm 1 Technical Analysis in FIC

- 1: **Self-stock Fuzzy system:**
 - 2: **for** i from 1 to n **do** ▷ iterate on self stock
 - 3: $\mathbb{E}'(r_i)$ and $\sigma'_i \leftarrow$ Normalize $\mathbb{E}(r_i)$ and σ_i
 - 4: Do fuzzification for $\mathbb{E}'(r_i)$, σ'_i , and η for self-stock fuzzy system
 - 5: Apply fuzzy rules for self-stock fuzzy system
 - 6: $w_i^{t,s} \leftarrow$ Do defuzzification for w for self-stock fuzzy system
 - 7: **Pairwise-stocks Fuzzy system:**
 - 8: **for** i from 1 to n **do** ▷ iterate on self stock
 - 9: **for** j from 1 to n excluding i **do** ▷ iterate on other stock
 - 10: $\mathbb{E}'(r_j)$ and $\sigma'_j \leftarrow$ Normalize $\mathbb{E}(r_j)$ and σ_j
 - 11: Do fuzzification for $\mathbb{E}'(r_j)$, σ'_j , $\rho_{i,j}$, and η for pairwise-stocks fuzzy system
 - 12: Apply fuzzy rules for pairwise-stocks fuzzy system
 - 13: $w_{i,j}^{t,p} \leftarrow$ Do defuzzification for w for pairwise-stocks fuzzy system
 - 14: $w_i^{t,p} \leftarrow \sum_{j=1}^{n-1} \rho_{i,j} w_{i,j}^{t,p}$
 - 15: $w_i^t \leftarrow (\eta \times w_i^{t,p}) + w_i^{t,s}$ ▷ Fusion of Self-Stock and Pairwise-Stock
 - 16: $\mathbf{w}^t \leftarrow [w_1^t, \dots, w_n^t]^\top$
 - 17: $\mathbf{w}^t \leftarrow \mathbf{w}^t / \sum_{i=1}^n w_i^t$
-

Algorithm 2 Fundamental Analysis in FIC

- 1: $\mathbf{c}_f \leftarrow [0.3, 0.15, -0.4, -0.5, -0.9]^T$
 - 2: **for** i from 1 to n **do** ▷ iterate on stock
 - 3: **for** k from 1 to n_f **do** ▷ iterate on fundamental features
 - 4: $f'_{i,k} \leftarrow$ Normalize $f_{i,k}$
 - 5: Do fuzzification for $f'_{i,k}$ and \mathbf{c}_f
 - 6: Apply fuzzy rules
 - 7: $w_i^f \leftarrow$ Do defuzzification for w
 - 8: $\mathbf{w}^f \leftarrow [w_1^f, \dots, w_n^f]^\top$
 - 9: $\mathbf{w}^f \leftarrow \mathbf{w}^f / \sum_{i=1}^n w_i^f$
-

Self-Stock Fuzzy System: As seen in Algorithm 1, the self-stock fuzzy system in the technical analysis part of FIC iterates on the stocks and for each of them, it suggests the weights ($w_i^{t,s}$ for stock i) according to its expected return $\mathbb{E}(r_i)$, risk σ_i (standard deviation), and risk tolerance η (see Fig. 5a). In order to prepare the inputs for the suitable ranges of defined membership functions, the self-stock expected return and self-stock standard deviation are respectively normalized to $\mathbb{E}'(r_i) \leftarrow \mathbb{E}(r_i) / (0.001 + \sum_{i=1}^n |\mathbb{E}(r_i)|)$ and $\sigma'_i \leftarrow \sigma_i / ((\sum_{i=1}^n \sigma_i)(0.0001 + \sigma_{\text{scaling}}))$ where $\sigma_{\text{scaling}} = \eta + ((1 - \eta) \times \max_{i=\{1, \dots, N\}} \sigma_i)$. This σ_{scaling} is a function of η to emphasize the risk tolerance of the investor.

Pairwise-Stocks Fuzzy System: In the pairwise fuzzy approach (see Algorithm 1), every stock is found to be better or worse than the other stocks to invest in. For every stock i , we iterate on all other stocks, indexed by j , and

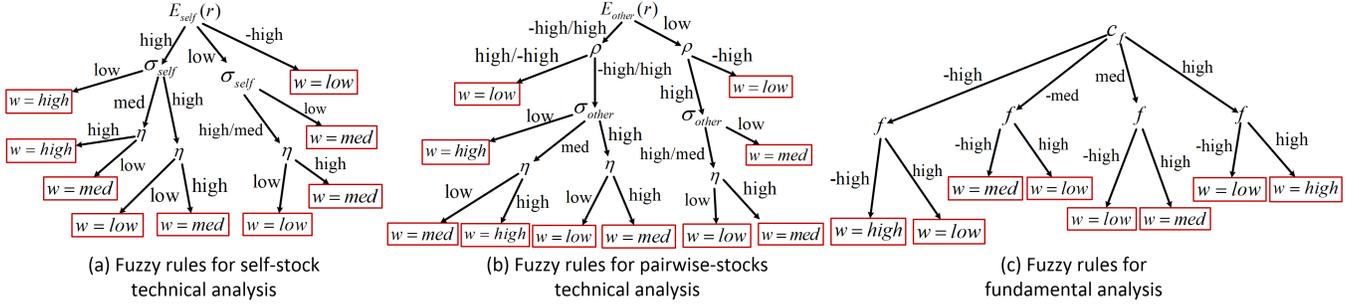


Figure 5: Fuzzy rules for (a) self-stock technical analysis, (b) pairwise-stocks technical analysis, and (c) fundamental analysis.

according to the rules of pairwise-stocks system (see Fig. 5b), the weights $w_{i,j}^{t,p}$ are suggested. The rules are applied on the fuzzified values of other stocks' expected return $\mathbb{E}(r_j)$, risk σ_j , risk tolerance η , and mutual correlation coefficient, $\rho_{i,j} = S(i,j)/\sigma_i\sigma_j$, where $S(i,j)$ is the element (i,j) of covariance matrix. The expected return and risk of other stocks should also be normalized as explained before. Finally, for every stock, $n-1$ pairwise weights are obtained and they should be fused. For this fusion for stock i , the $w_{i,j}^{t,p}$ weights are summed up while weighted by their $\rho_{i,j}$ which determines their impact $w_i^{t,p} = \sum_{j=1}^{n-1} \rho_{i,j} w_{i,j}^{t,p}$.

Fusion of Self-Stock and Pairwise-Stocks Fuzzy Systems: Finally, the total technical weight is $w_i^t = (\eta \times w_i^{t,p}) + w_i^{t,s}$ where the influence of $w_i^{t,p}$ is ignored when η is small because when the risk tolerance is low, self-stock variations should be more valued. The vector of total technical weights, denoted by $\mathbf{w}^t = [w_1^t, \dots, w_n^t]^\top$, is then normalized so that the weights sum up to one, $\mathbf{w}^t \leftarrow \mathbf{w}^t / \sum_{i=1}^n w_i^t$.

Fundamental Analysis in FIC: For fundamental analysis in FIC, n_f (here five) fundamental features, i.e., **account receivable, capital expenditure, inventory, gross margin, and income tax**, are selected to be used in this work from the dataset. We define fundamental coefficients, denoted by \mathbf{c}_f , for the utilized fundamental features. This coefficient is in the range $[-1, 1]$ and determines the positive or negative impact of a fundamental feature on the stock's expected return. Considering the analysis reported in (Abarbanell and Bushee 1997), which reports the relative impacts of the mentioned fundamental features on the expected return of the stock, we define $\mathbf{c}_f = [0.3, 0.15, -0.4, -0.5, -0.9]^\top$ as a vector of fundamental coefficients of the used fundamental features with the same mentioned order. As seen in Algorithm 2, for every stock, separate fuzzy rules are used for different fundamental features (see Fig. 5c). For every stock i , the fundamental features, indexed by k and denoted by $f_{i,k}$, are normalized with summation of their absolute values to be appropriate for the defined membership function, $f_{i,k}' \leftarrow f_{i,k} / (0.001 + \sum_{k=1}^{n_f} |f_{i,k}|)$. The obtained vector of fundamental weights $\mathbf{w}^f = [w_1^f, \dots, w_n^f]^\top$ is finally normalized to sum up to one, $\mathbf{w}^f \leftarrow \mathbf{w}^f / \sum_{i=1}^n w_i^f$.

Combining Technical and Fundamental Analysis: Eventually, for stock i from technical weight w_i^t and fundamental weight w_i^f , the overall fuzzy weight is calculated

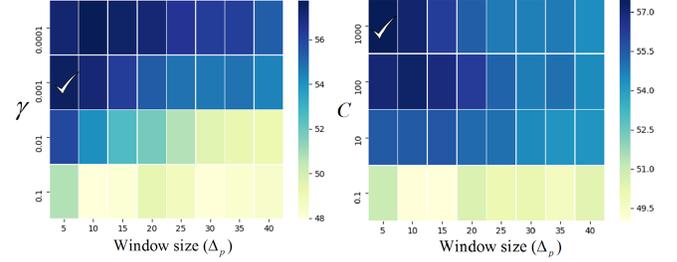


Figure 6: Sweeping γ and C versus Δ_p

as $w_i = \alpha w_i^f + w_i^t$ where $\alpha = (n_f + \mathbf{c}_f^\top \mathbf{f}_i') / 2n_f$ and \mathbf{f}_i' is a vector containing the normalized fundamental features of i -th stock. Note that $\mathbf{c}_f^\top \mathbf{f}_i'$ is in range $[-n_f, n_f]$ and thus the range of α is $[0, 1]$. This α determines the importance of fundamental analysis against the technical analysis. Finally, the overall fuzzy weights $\mathbf{w} = [w_1, \dots, w_n]^\top$ are normalized so that they sum up to one, $\mathbf{w} \leftarrow \mathbf{w} / \sum_{i=1}^n w_i$.

Experimental Results

Utilized Dataset The utilized dataset in this work is the New York Stock Exchange (NYSE) dataset. This dataset contains information on 140 stocks from years 2010 to 2016. We selected 25 well-known stocks for evaluating the prediction and comparison to the literature. Most of the stocks are selected inspired by (Song 2018) although our algorithm can be applied on any number of stocks. The 25 selected stocks are AAPL (Apple), AIG (American International Group), AMZN (Amazon), BA (Boeing), CAT (Caterpillar), COF (Capital One), EBAY, F (Ford), FDX (FedEx), GE (General Electric), GM (General Motors), GOOG (Google Alphabet Inc.), HD (The Home Depot), IBM, JNJ (Johnson & Johnson), JPM (JPMorgan Chase & Co.), KO (Coca-Cola), MSFT (Microsoft), NKE (Nike), ORCL (Oracle), PEP (PepsiCo), T (AT&T), WMT (Walmart), XOM (ExxonMobil), and XRX (Xerox). The used information of this dataset in this work includes: (I) *Prices*: The prices (in US dollar) represented as time series, which are opening (the morning price), closing (the evening price), lowest, and highest prices as well as volume (the number of traded stocks in the day). For training data, the first 80% of the data (almost years 2010 to 2014) is used for training and the rest (almost 2015 and 2016) for test. For tuning the parameters, the last 20% of the

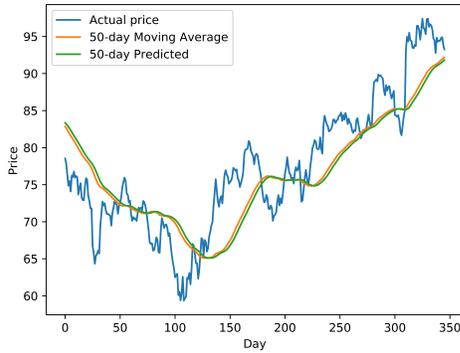


Figure 7: Forecasting the averaged time series of highest price in CAT stock.

training set is held out for validation. (II) *Fundamental Attributes*: It contains some numeric metrics showing the annual company performance such as account receivable, capital expenditure, inventory, gross margin, and income tax. Features for years 2013 to 2015 are used as fundamental inputs to FIC.

Results for Price Prediction The proposed framework includes three parameters Δ_p , C , and γ . To tune these parameters, different values, i.e., $\Delta_p = \{5, 10, \dots, 40\}$, $C = \{0.1, 10, 100, 1000\}$, and $\gamma = \{0.1, 0.01, 0.001, 0.0001\}$ were tested on the validation set. Figure 6 illustrates the average hit rates over the 25 stocks on the validation set. Note that the Hit Rate (HR) is formulated as $HR = \frac{1}{m} \sum_{t=1}^m [\text{sign}(\tilde{r}(t)) \stackrel{?}{=} \text{sign}(r(t))]$ where m is the number of days of prediction. The HR assesses how well the trend of price is predicted. As can be seen in Fig. 6, the best obtained parameters are $C = 1000$, $\Delta_p = 5$, and $\gamma = 0.001$. As expected, a large C is penalizing slack variables for better prediction, and small γ avoids a biased model. A small window size also makes sense according to (Song 2018).

We evaluated the predictions using standard methods from the literature, which are Hit Rate (HR), Mean Absolute Error ($MAE = (1/m) \sum_{t=1}^m |s(t) - \tilde{s}(t)|$), and Root Mean Square Error ($RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^m (s(t) - \tilde{s}(t))^2}$) where $\tilde{s}(t)$ denotes the predicted time series. To have a better comparative sense of these metrics, we also report Mean Absolute Percentage Error ($MAPE = (100/m) \sum_{t=1}^m |(s(t) - \tilde{s}(t))/s(t)|$). The results on the test set are reported for the 25 stocks in Table 1. We evaluated price prediction for two cases, i.e., Not Smoothed Prices (NSP) and Smoothed Prices (SP) with moving average, reported in this table. The MAE, RMSE, and MAPE errors of our work on different stocks are relatively small. Moreover, the hit rates are fairly high in almost all the stocks showing the fine performance of SVR in predicting the time series and its trend. The hit rates of SP are all high enough showing the effectiveness of our algorithm. Table 1 also compares the stock prediction of the proposed framework with a very recent work for Trend Prediction (TP) (Song 2018). As reported in this table, the NSP

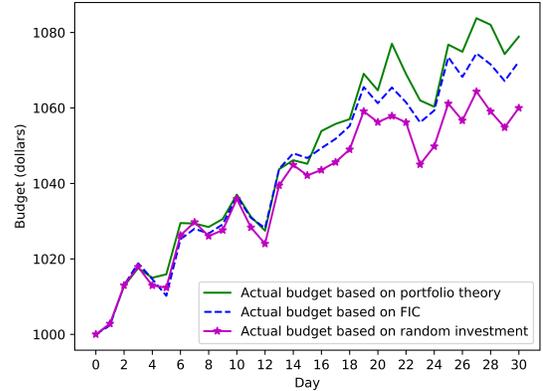


Figure 8: Actual and predicted budgets for a 30-day period.

outperforms TP or has comparative result with it in stocks AAPL, AMZN, BA, COF, GOOG, NKE, ORCL, and T. It is noteworthy that TP (Song 2018) merely predicts +1 for uptrend and -1 for downtrend and measures the hit rate based on that. However, NSP predicts the actual price and our hit rate is based on the predicted price (and not the trend) which is more difficult.

In order to capture a visual sense of the performance of price forecasting, see Fig. 7 which shows that the predicted averaged time series is following the actual averaged series closely.

Results for Suggestion of Weights The predicted averaged prices, i.e., SP, are used as input to weight suggestion part. In order to illustrate the performance of the weight suggestion modules, suppose that the initial budget of the investor, denoted by B , is \$1000. The proposed weight suggestion methods, which are based on portfolio theory or fuzzy logic (FIC), are tested on this budget invested in the 20 stocks (all except the stocks GE, GOOG, ORCL, and JNJ because of lack of fundamental features in the dataset and the stock GM because of its incomplete time series of technical features in the dataset).

For every day of a period of time, the optimum weights $\{w_i\}_{i=1}^n$ are found, from which the actual future return r and the expected future return $\mathbb{E}(r)$ are obtained as $\sum_{i=1}^n w_i r_i$ and $\sum_{i=1}^n w_i \mathbb{E}(r_i)$, respectively. Then the actual future budget (B) is updated based on the obtained r as $B \leftarrow B \times (1 + r)$. Note that, according to rationality, if the expected return is negative, we exit from stock market for that future day. This procedure is performed for 30 days, as an example, to compare and analyze the actual remained budgets for three different methods of weight suggestion which are portfolio theory, FIC, and random weights. This example 30-day period includes days 1537 to 1567 of dataset although can be any time span. The actual budgets, are illustrated in Fig. 8 for $\eta = 30\%$. On the 30th day, the actual budget is updated to \$1078.88, \$1072.26, and \$1059.98 using portfolio theory, FIC, and random weighting, respectively. As seen in this figure, the actual budget based on portfolio theory is higher than the actual budget based on fuzzy logic and

		AAPL	AIG	AMZN	BA	CAT	COF	EBAY	F	FDX	GE	GM	GOOG	HD
NSP	HR	61.16%	55.36%	57.68%	60.00%	60.29%	58.26%	56.81%	61.45%	58.26%	53.33%	61.46%	57.97%	56.52%
	MAE	1.058	0.410	7.973	1.310	0.848	0.757	0.291	0.131	1.517	0.228	0.332	6.676	1.086
	RMSE	1.483	0.710	11.085	1.748	1.150	1.018	0.474	0.190	2.227	0.320	0.455	9.707	1.478
	MAPE	0.99%	0.84%	1.23%	0.97%	1.12%	1.04%	1.07%	1.00%	0.97%	0.78%	1.04%	0.92%	0.85%
SP	HR	91.59%	92.17%	93.33%	85.50%	88.69%	95.65%	89.56%	86.08%	93.04%	92.17%	83.76%	89.70%	92.75%
TP (Song 2018)	HR	57.58%	58.79%	56.21%	60.15%	×	57.12%	59.70%	×	59.55%	61.21%	×	58.33%	59.70%
		IBM	JNJ	JPM	KO	MSFT	NKE	ORCL	PEP	T	WMT	XOM	XRX	
NSP	HR	55.94%	52.75%	56.23%	54.49%	53.62%	60.00%	58.26%	57.68%	57.68%	56.52%	54.49%	60.00%	
	MAE	1.178	0.613	0.598	0.236	0.461	0.598	0.286	0.579	0.229	0.502	0.669	0.110	
	RMSE	1.681	0.889	0.879	0.333	0.708	0.834	0.408	0.802	0.313	0.799	0.885	0.160	
	MAPE	0.80%	0.56%	0.92%	0.55%	0.87%	1.02%	0.74%	0.57%	0.61%	0.75%	0.80%	1.10%	
SP	HR	92.46%	93.04%	94.49%	90.14%	89.27%	90.14%	93.91%	88.98%	93.91%	94.49%	93.62%	85.50%	
TP (Song 2018)	HR	×	56.52%	61.67%	59.24%	59.09%	60.61%	58.94%	59.39%	58.64%	60.91%	60.45%	×	

Table 1: Evaluation of our work on forecasting technical features and comparing to related work. The × symbol means that the related work does not have result on that stock.

random investment because it is based on optimization. The FIC, which considers both technical and fundamental features, results in invested budget very close to result of portfolio theory; suggesting that the FIC closely models the optimized behaviour of portfolio theory. In conclusion, as Fig. 8 shows, both the proposed weight suggestion methods are better than random investment.

Conclusion and Future Direction

In this paper, a novel end-to-end artificial investment counselor was demonstrated which first forecasts the time series of prices in several stocks and then suggests the optimum portions of the budget to be invested in the stocks to have the best possible return in the future. The FIC, in contrast to portfolio theory, gives us the opportunity to consider even more features, such as news or sentimental attributes (e.g., social media impact (Yang, Mo, and Liu 2015)), similar to an expert broker. Some future work can be considering more professional properties in finance, e.g., short selling and Elliott waves to have a more complete trading system.

References

- Abarbanell, J. S., and Bushee, B. J. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* 35(1):1–24.
- Atsalakis, G. S., and Valavanis, K. P. 2009a. Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert systems with Applications* 36(7):10696–10707.
- Atsalakis, G. S., and Valavanis, K. P. 2009b. Surveying stock market forecasting techniques—part ii: Soft computing methods. *Expert Systems with Applications* 36(3):5932–5941.
- Bodie, Z.; Kane, A.; and Marcus, A. J. 2014. *Investments*. McGraw-Hill Education.
- Cao, L.-J., and Tay, F. E. H. 2003. Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on neural networks* 14(6):1506–1518.
- Cavalcante, R. C.; Brasileiro, R. C.; Souza, V. L.; Nobrega, J. P.; and Oliveira, A. L. 2016. Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications* 55:194–211.
- Chen, J. 2010. Svm application of financial time series forecasting using empirical technical indicators. In *Information Networking and Automation (ICINA), 2010 International Conference on*, volume 1, V1–77. IEEE.
- Fletcher, T. 2012. *Machine learning for financial market prediction*. Ph.D. Dissertation, UCL (University College London).
- Halliday, R. 2004. Equity trend prediction with neural networks. *Research Letters in the Information and Mathematical Sciences* 6:15–29.
- Ican, Ö., and Çelik, T. B. 2017. Stock market prediction performance of neural networks: A literature review. *International Journal of Economics and Finance* 9(11):100–108.
- Karray, F. O., and De Silva, C. 2004. *Soft computing and intelligent systems design: theory, tools, and applications*. Pearson Education.
- Kayal, A. 2010. A neural networks filtering mechanism for foreign exchange trading signals. In *Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference on*, volume 3, 159–167. IEEE.
- Klir, G., and Yuan, B. 1995. *Fuzzy sets and fuzzy logic: Theory and Application*. Prentice hall New Jersey.
- Makridakis, S.; Spiliotis, E.; and Assimakopoulos, V. 2018. Statistical and machine learning forecasting methods: Concerns and ways forward. *PloS one* 13(3):e0194889.
- Song, Y. 2018. Stock trend prediction: Based on machine learning methods. Master’s thesis, University of California, Los Angeles (UCLA).
- Vanstone, B., and Finnie, G. 2009. An empirical methodology for developing stockmarket trading systems using artificial neural networks. *Expert systems with applications* 36(3):6668–6680.
- Wilder, J. W. 1978. *New concepts in technical trading systems*. Trend Research.
- Yang, S. Y.; Mo, S. Y. K.; and Liu, A. 2015. Twitter financial community sentiment and its predictive relationship to stock market movement. *Quantitative Finance* 15(10):1637–1656.