AITA: A new framework for Trading Forward Testing with an Artificial Intelligence Engine

Ivan Letteri[1]

[1] University of L'Aquila, Italy - Department of Information Engineering, Computer Science and Mathematics

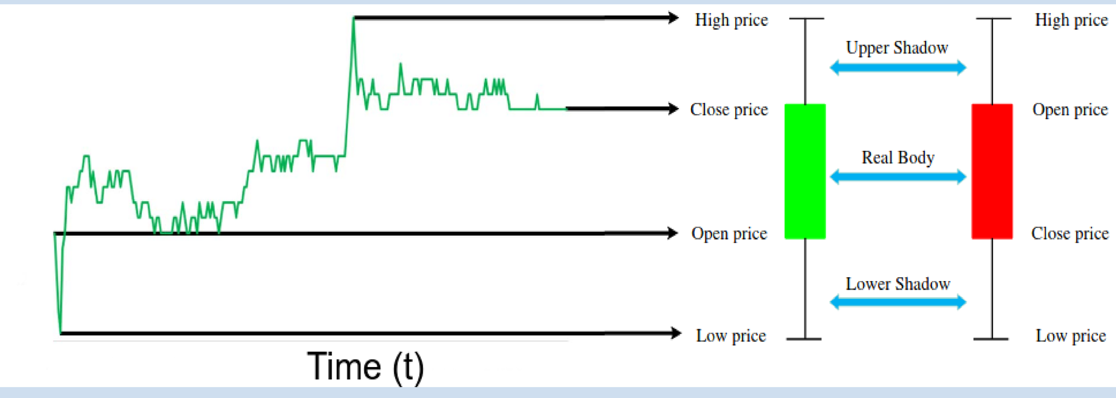
UNIVERSITÀ DEGLI STUDI DELL'AQUILA

Abstract

In general, traders test their trading strategies by applying them to historical market data (backtesting) and then reapply those that have made the most profit on that past data. We propose AITA (*Artificial Intelligence Trading Assistant*), our framework for generating trading systems by applying the new trading strategy called DNN-forwardtesting by determining the best strategy based on the prediction issued by a deep neural network. The experiment with AITA involves the use of 10 stocks that are first filtered according to their volatility using the *Kmean++* model. Having determined the assets with average volatility, we use this historical data to train a deep feed-forward neural network to predict price trends over the next 30 days of the open stock market. The trading system, created by AITA, calculates the most effective technical indicator by applying it to the DNN forecasts to generate the trading strategy.

AITA features

1. Price Action - AITA uses the analysis of asset price history series, defined as OHLC, i.e., the opening, higher, lowest and closing prices of an asset, typically represented with candlesticks charts. For each timeframe t, the OHLC of an asset is represented as a 4-dimensional vector $X_t = (x_t^{(o)}, x_t^{(h)}, x_t^{(l)}, x_t^{(c)})^T$, where $x_t^{(l)} > 0$, $x_t^{(l)} < x_t^{(h)}$ and $x_t^{(o)}, x_t^{(c)} \in [x_t^{(l)}, x_t^{(h)}]$.



2. Volatility Estimators - AITA framework is designed to use the PK, GK, RS, YZ Historical Volatility measures to create a dataset.

- The *Parkinson* (PK) estimator incorporates the stock's daily *high* and *low* prices and is calculated as follow:

$$PK = \sqrt{\frac{1}{4N ln_{CR(n)} 2} \sum_{i=1}^{N} (ln_{CR(n)} \frac{x_t^{(h)}}{x_t^{(l)}})^2}.$$

PK provides completely separate information from using time-based sampling such as closing prices, but it cannot handle trends and jumps, so it systematically underestimates the volatility.

- The *Garman-Klass* (GK) estimator incorporates open, low, high, and close prices and is calculated as follows:

$$GK = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} (ln_{CR(n)} \frac{x_t^{(h)}}{x_t^{(l)}})^2 - \frac{1}{N} \sum_{i=1}^{N} (2 ln_{CR(n)} 2 - 1)(ln \frac{x_t^{(c)}}{x_t^{(o)}})^2}.$$

This method is robust for opening jumps in price and trend movements. However, it takes into account not only the price at the beginning and end of the time interval but also intraday price extremums.

- The *Rogers-Satchell* (RS) estimator, unlike the PK and GK estimators, RS incorporates drift term (mean return not equal to zero). It is calculated as follows:

$$RS = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (ln_{CR(n)} (\frac{x_t^{(h)}}{x_t^{(c)}}) ln_{CR(n)} (\frac{x_t^{(h)}}{x_t^{(o)}}) + ln_{CR(n)} (\frac{x_t^{(l)}}{x_t^{(c)}}) ln_{CR(n)} (\frac{x_t^{(l)}}{x_t^{(o)}})}.$$

However, it does not take into account price movements between trading sessions, and underestimate volatility since price jumps periodically occur between sessions.

- The *Yang-Zhang* (YZ) estimator Yang and Zhang [2000] incorporates open, low, high, and close prices, and is calculated as follows:

$$\sigma^2_{OpentoCloseVol} = \frac{1}{N-1} \sum_{i=1}^{N} (ln_{CR(n)} \frac{x_t^{(c)}}{x_t^{(o)}} - ln_{CR(n)} \frac{x_t^{(c)}}{x_t^{(o)}})^2$$

$$\sigma^2_{OvernightVol} = \frac{1}{N-1} \sum_{i=1}^{N} (ln_{CR(n)} \frac{x_t^{(o)}}{x_{t-1}^{(c)}} - ln_{CR(n)} \frac{x_t^{(o)}}{x_{t-1}^{(c)}})^2$$

$$YZ = \sqrt{\sigma^2_{OvernightVol} + k\sigma^2_{OpentoCloseVol} + (1-k)\sigma^2_{RS}}$$

where $k = \frac{0.34}{1.34 + \frac{N+1}{N-1}}$.

3. Stocks Evaluated

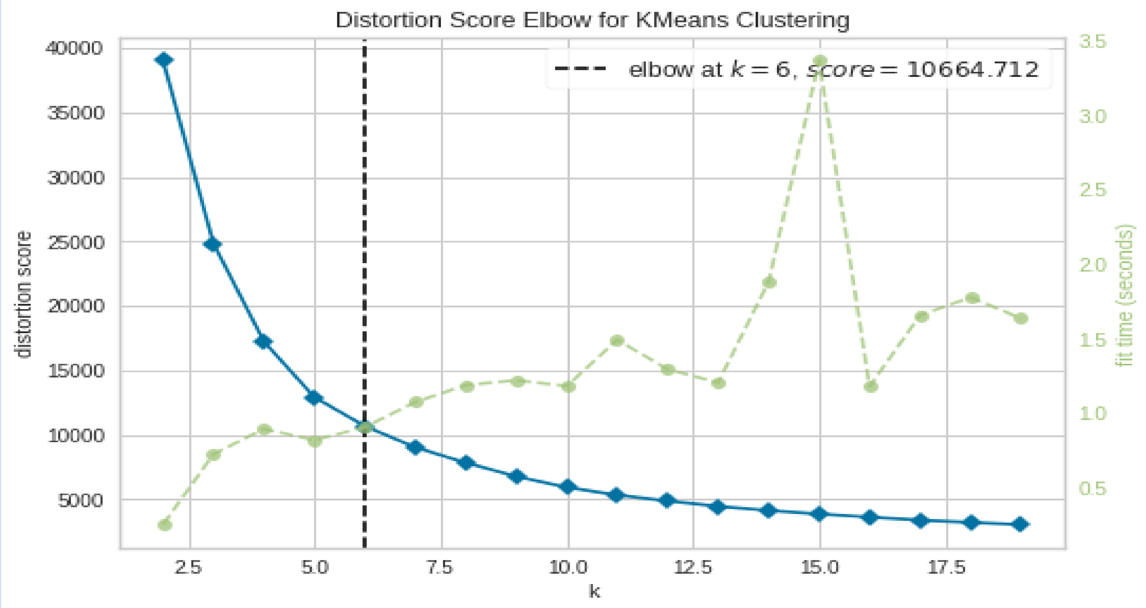
Ticker	Company	Market
CSGKF	Credit Suisse Group AG	Other OTC
EOG	*EOG Resources, Inc.*	*NYSE*
META	Meta Platforms, Inc.	Nasdaq GS
NKE	NIKE, Inc.	NYSE
DIS	Walt Disney Co.	NYSE
PG	Procter & Gamble Co.	NYSE
QQQ	Invesco QQQ Trust	Nasdaq GM
IBM	Business Machines Corp.	NYSE
ANF	*Abercrombie & Fitch Co.*	*NYSE*
CS	Credit Suisse Group AG	NYSE

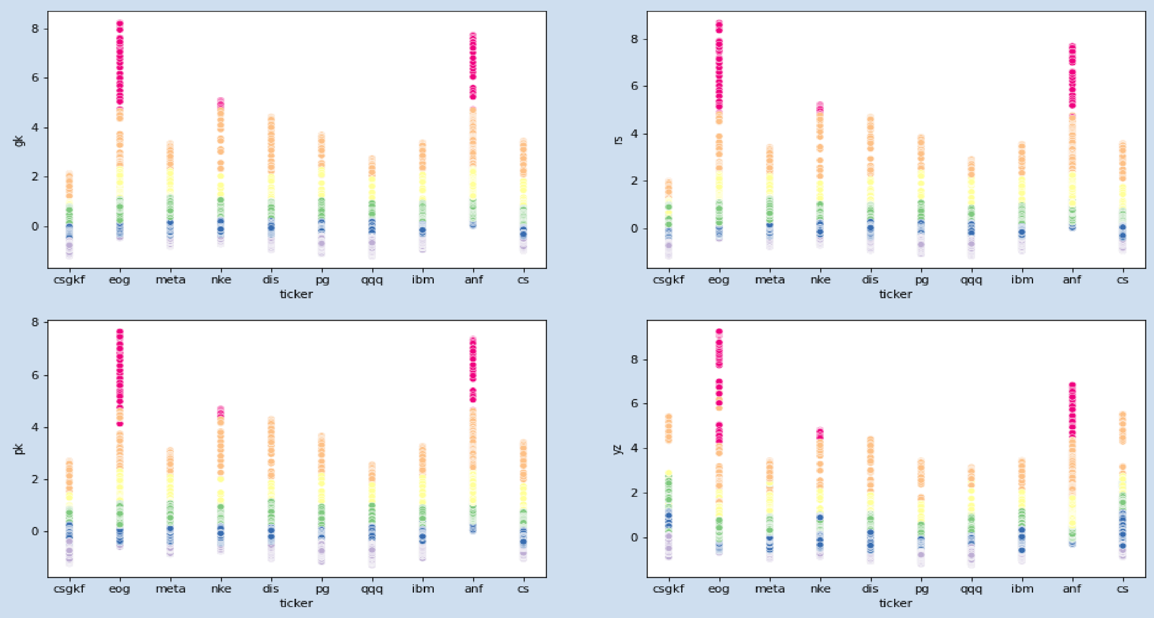
Metrics

(i) *Max drawdown* $max_{\tau \in (0,t)}[max_{t \in (0,\tau)} \frac{n_t - n_\tau}{n_t}]$. (ii) *Sharpe ratio (SR)* $\frac{\mathbb{E}[r]}{[r]}$. (iii) *Sortino ratio (SoR)* $\frac{\mathbb{E}[r]}{DD}$. (iv) *Calmar ratio (CR)* $\frac{\mathbb{E}[r]}{MDD}$.
To check the goodness of trades: *Total Returns* $R_k(t)$ for each stock $(k = 1, ..., p)$ in the time interval $(t = 1, ..., n)$, where $TR = R_k(t) = \frac{Z_k(t+\Delta t) - Z_k(t)}{Z_k(t)}$, and analysing the standardized returns $r_k = (R_k - \mu_k)/\sigma_k$, with $(k = 1, ..., p)$, where σ_k is the std dev of R_k, e μ_k denote the average overtime.

The Experiment

Dataset is pre-processed: (i) Compute k-means++ clustering for different values of k. In our case, we varied k from 2 to 20 clusters. (ii) For each k, is calculated the total within-cluster sum of squares (ws). (iii) Plot the curve of wss according to the number of clusters k. (iv) Find the location of a bend (knee) in the plot, which is generally considered as indicator of the appropriate number of clusters, and the best clustering in our experiments is for $k = 6$. ANF and EOG have all the types of observations spread over all the k-means++ clusters for each historical Volatility estimator considered.





ARIMA model as benchmark. the predictions on the closing prices made with such auto-selected optimal ARIMA model for k = 30 days following the training timespan, from October 30, 2011, to October 16, 2021. **DNN model.** The NN performs better is the MLP when its input layer is fed with the $t = 5$ previous values. The resulting optimal geometry has two hidden layers composed of $10 * t$ and $5 * t$ neurons, respectively, with dropout of 20% on each of the two internal layers. To introduce non-linearity between layers, we used ReLU as the activation function, and to estimate the network learning performance we use the L1loss function. The optimisation algorithm used to minimise such loss function is the adaptive moment (Adam).

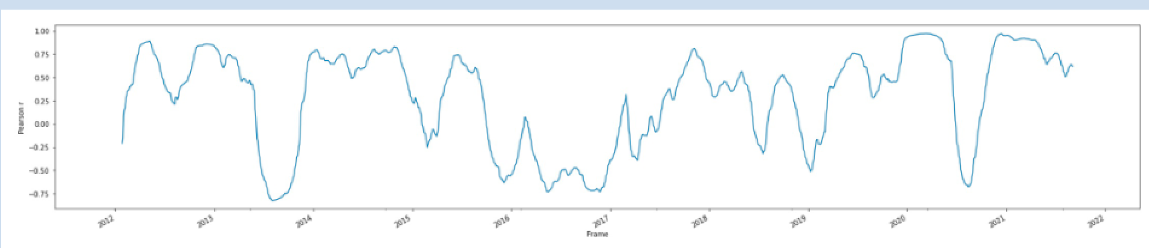
ARIMA					
	MSE	RMSE	MAE	MAPE	EVS
ANF	25.49	5.05	3.86	0.09	-0.02
EOG	56.23	7.50	5.42	0.06	-3.94

DNN					
	MSE	RMSE	MAE	MAPE	EVS
ANF	1.75	1.32	1.07	0.02	0.91
EOG	2.39	1.55	1.23	0.01	0.7

Tables reports ARIMA errors (very high) and DNN errors acceptable.

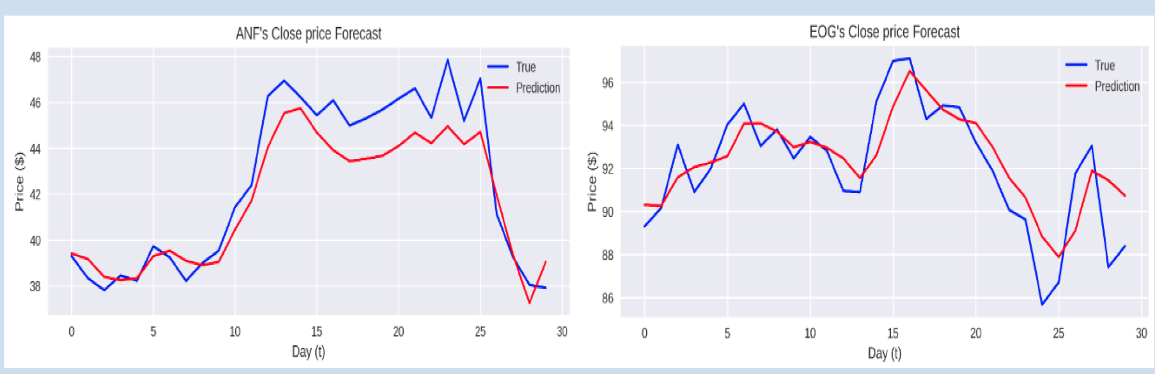
Forecasting Methods

AITA framework also proves that the price time series corresponding to the selected assets are completely uncorrelated, i.e., ANF and EOG do not influence each other. Therefore, we evaluated the synchrony between the two financial assets using (i) the Pearson coefficient. (ii) Dynamic Time Warping.



ANF	
Entry	$((x^{(l)} < TEMA^{(l)}) \vee (x^{(h)} < TEMA^{(h)})) \wedge ((x^{(c)} < TEMA^{(c)}) \vee (x^{(o)} < TEMA^{(o)}))$
Exit	$((x^{(l)} > TEMA^{(l)}) \vee (x^{(h)} > TEMA^{(h)})) \wedge ((x^{(c)} > TEMA^{(c)}) \vee (x^{(o)} > TEMA^{(o)}))$
EOG	
Entry	$(+DI > -DI) \wedge (ADX > 25)$
Exit	$(-DI > +DI) \wedge (ADX > 25)$

Best Techinical Indicators. The AITA (algorithmic) trading strategy is encoded in a set of entry and exit trading rules which are in turn based on the value of a single indicator chosen from a set of twelve common technical indicators, i.e., SMA, EMA, MACD, BBs, ADX, ...
In this case **DNN win on ARIMA.** So, AITA forecasts OHLC prices using ML models.



Results

With backtesting would choose ADX for the EOG share, as with our forward testing technique. However, the TEMA indicator would not be chosen for the ANF share. Indeed, the most promising indicator would be RSI. However, if applied to the future, it would result in a loss of 1.16%

	#Trades	TR ($)	ShR	SoR	CaR
ANF	3	6.126	2.194	3.340	12.403
EOG	3	1.374	1.253	2.556	5.814

	#Trades	TR ($)	ShR	SoR	CaR
ANF	1	-1.168	0.119	0.158	-0.935