Comparing Pure Stock Portfolio with Stock and Crypto-currency mixed Portfolio through LSTM to Compare & Analyze Investment Opportunities for Portfolio Performance Measurement

Hasan M Sami[1] and S.M. Arifuzzaman[1]*

[1]School of Business, Canadian University of Bangladesh (CUB), Dhaka-1212, Bangladesh.

*Correspondence: arif@cub.edu.bd (S.M. Arifuzzaman, Associate Professor and Head, School of Business, Canadian University of Bangladesh (CUB), Dhaka-1212, Bangladesh).

ABSTRACT

LSTM (Long Short-Term Memory) has revolutionized the approach to time series prediction many folds due to its appropriate capability to forecast through Non-Linear forecasting methods. It's observed that RNN has the capability to similarly think through given enough training in accordance to desired functionality models. Due to the Gated Structure referring to storing relevant information's and forgetting the irrelevant information's LSTM made revolutionary accomplishments towards non linear forecasting that is dependent on human like behavior. In this research we have focused on making comparison between two different portfolio's which will depend upon LSTM's future forecasting capability in terms of predicting the best possible output which gets illustrated through Portfolio Optimization principles.

Keywords: LSTM, RNN, Gates, Markowitz portfolio Theory, Performance, and Non-linear price prediction.

INTRODUCTION:

LSTM has the capability to solve time series processes through by feed forward network using permanent size time windows (Gers *et al.,* 2002). The task primarily associated with LSTM applies within the boundary of training under a fixed data series and this training fore casting refers to the future point of interest. It's seen that LSTM have predicted human motion and their predictiveness with high degree of accuracy (Alahi *et al.,* 2016) in terms of their behavior effecting regarding choices for motion in crowded spaces. Alahi *et al.* have successfully implemented the behavioral track in terms of intelligent tracking systems that simu-lates with human behavioral and decision pattern in relation to smart environments. Wang *et al.* have suggested that the risk of making investment is directly correlated with less financial knowledge and better financial decisions always leads to better and profitable financial decisions (Wang *et al.,* 2011).

Evidential perception suggests that in USA and Germany, investors are more inclined towards their local investments rather than foreign mixed investment portfolios (Kilka *et al.,* 2000). Our research paper would focus upon the enlightened view of statistically controlled processes through nonlinear approaches for portfolio asset choices and time series-based decision making. It's found that crypto currencies remain an important investment alternative (David *et al.,* 2017) since its birth with bit coin in 2009. The multifold growth of various crypto currencies has led investors to undertake crypto currencies as strong investment alternative (David *et al.,* 2017). According to Tobias,

it's observed that in the world of portfolio-based investment introducing crypto currencies in reference to a long only strategy with mixture and effective understanding of all portfolio performance optimizing parameters would result to create the best possible portfolio returns (Tobias, 2019). In our research we would thus make comparisons between pure stock portfolios under the pretense of selecting them with good financial standpoints and would compare the mixture of stocks with crypto currencies through performance optimizations for financial performance betterment. It's found that there are evidential circumstances that suggests crypto currency-based portfolio to be extremely good performers (Jiang & Liang, 2017). In relation to nonlinear forecasting and time series prediction accuracy it's found that Deep reinforcement learning plays an effective role in portfolio management.

Due to high amount investments being poured into crypto currencies because of its growing popularity and ROI, a highly volatile behavior is observed (Fruehwirt *et al.,* 2020). Our paper would thus try to reduce the risk variance of investment by mixing up the finances with stock with proper financial indications for long only gains rather than pure statistical evidences for portfolio investment on which we would rely completely for the financial contributions of crypto currencies. Thus, this research process would describe the following events:

1. Creating a pure Stock Portfolio
   a) Each stock would get selected under the pretense of solid financial ratio as evidence for investment as suggestion.
   b) Using LSTM for nonlinear forecasting processes suggests the best stocks for performing well in long only investment method for best portfolio performance under the judgment of best stocks to select under the boundary of a specific time period.
2. The best selected stocks will be mixed best possible crypto currencies which will perform best under the pretense of LSTM for that specific time period where the Stock with crypto currency combined portfolio would be compared with pure stock only portfolio in reference to financial performance or financial gains.

Literature Review

In financial markets, the most important factors remained as financial literacy and risk perception in terms of selecting assets for investment purposes (Aren *et al.,* 2016). It's found that as financial literacy proceeds thus better human decision occurs. As Human Like thinking capability determines the overall investment perspective of market behavior (Caparrelli *et al.,* 2004), it's better to simulate the future asset pricing forecast in association with training asset choice option through biased factor of non-linear forecasting principle based on LSTM. It's seen that in relation to learning the pattern of choices based on various situations LSTM makes the most effective choice of holistic feature selection. It's based on dilated causal convolution network which aims to learn various features which are completely trait dependent and multi-dimensionality reduction associates with multi-conditional data hence the learning with due training makes the nonlinear prediction effective (Shen *et al.,* 2020). Its further being evaluated under the circumstances that multi range and multi-level features are completely under predictive accuracy using fixed time intervals (Shen *et al.,* 2020 and Alahi *et al.,* 2016). For commodity pricing, LSTM also made effective predictions for Gold pricing (Livieris *et al.,* 2020). The forecasting model will associate with verified by linear regression by major global indices (Cao *et al.,* 2018; Livieris *et al.,* 2020). As the gated structure supports the feed forward network hence the networking model can learn appropriately and can make accurate predictions (Gers *et al.,* 2002). It's seen that LSTM is capable of effectively predicting the Bit coin price while learning and training was the only feature for making accurate predictions (Wu *et al.,* 2018). It's found that as crypto currencies indicate no evidence associated with financial performance indicators hence their fall and gain is still more volatile in comparison effective stocks (Fruehwirt *et al.,* 2020). Hence our decision is focused to associate stocks with crypto currencies to increase the portfolio performance but at the same time would try to reduce the mixed portfolio variance. Furthermore, to conduct the selection process, we would associate with favorable financial ratios as indicative benchmarks. Since the height (Chen 1981 & Beaver 1966) of decision-making regarding selection of assets and for portfolio performance optimization

proposed by (Henry Markowitz, 1952; Markowitz Theory - Jorion P, 1992), it's seen that financial ratios play an important role for asset selections. It's observed that LSTM offers better predictability than ARIMA (Siami *et al.,* 2018). Our predictability feature is dependent on various factors aligned with financial dependency of the assets. This research would focus, how we would make selections of assets using K Nearest Neighbor algorithm (Chen *et al.,* 2017). It's found that EPS ratios plays a significant role in regards to making selections of assets for investments (Lin *et al.,* 2011). Husna & Satria in (2019) have found that Current ratio plays an effective role to define the pre-sent liquidity condition in order to proof their investors its financial capability. Referring to such current ratio status, it's seen that throughout human decision of investment history, it's always better to make investments in a liquid company be-cause that refers to more or easier way out in case of financial distress or easy ROI (Amihud & Mendelson, 1991).

In process of selecting the effective Crypto currencies, our method of justification is laid upon with the usual alliance of volume model choice (Capon *et al.,* 1996) which works perfectly well with mutual funds. In association with non linear methods it's seen that LSTM makes accurate and affirmative predictions for future aspects of any growing models (Livieris *et al.,* 2020). For Fruit Fly Algorithm process LSTM makes effecttive predictions regarding location attributes (Peng *et al.,* 2020). Hence, this paper focuses entirely on LSTM for asset pricing related to future forecasting and accuracy.

Theoretical Framework

K Nearest Neighbor Algorithm - It's seen that K nearest neighbor algorithm has been highly successful to locate the best group of employees in order to select through defined quality settings (Sami HM, 2021). This research would also impose a set of quality factors those are considered as benchmarks for identifying the best possible stocks for investment.

Point of Interest:

EPS - CFI (Corporate Finance Institution) and major accounting standards of GAAP suggests that, EPS is explained as the ratio of Net Income against the Weighted Average Outstanding Shares.

Current Ratio - Current Ratio shows if an organization is financially solvent or not. In reality it's a standard benchmark to suggest, how fast the company can pay off its current debts with the usage of its current assets.

The exquisite point among the various values of both EPS & Current Ratios and for Crypto currencies, the specific volume amount that suggests the best theoretical practice in lieu to accounting standard and popularity means (Leong *et al.,* 2009 ; Arel *et al.,* 2016; Husna and Satria, 2019). It's found that volumetric suggestion as the basis for mutual fund investment (Capon *et al.,* 2016).

Benchmarks
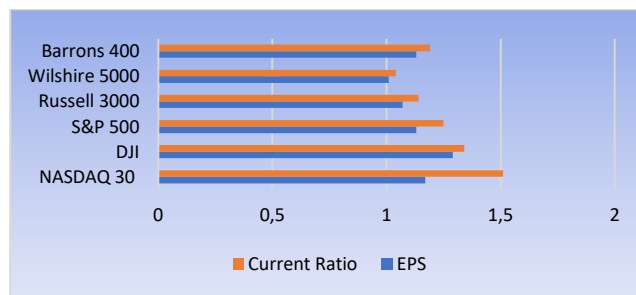
In reference to evidential presence through accounting standard guidelines it found that major stock indexes have most EPS and Current ratios illustrated as greater than 1 and for in between 1 & s respectively. Leong et al mentioned along with supportive proofs of accounting standards from GARP it's seen that most of the major stock market index suggests to more than 1 EPS as a benchmark standard and for Current Ratio it's always best to be between 1 and 2. This research paper thus would associate with comprising these as benchmark standards for selecting stocks under the benchmark associations of NASDAQ 30, DJI 30, S&P 500 and etc.

Volume & Affordability of Sales - Stoics have found an immediate correlation between anti persistence with volumetric changes associated with price associations of major currencies (Stosic *et al.,* 2019). Stosic *et al.* have shown that with directional improvement in volume there is clear evidence in changes of prices in crypto currencies. Our research focuses exclusively to associate volumetric sales of crypto currencies as the basic measure for choice. Novy Marx mentions that if we make selection of stocks based on its quality attribute then with quality increment, we can expect to increase the payment for attaining a premium stock (Novy Marx, 2013). The simple method would suggest relying on average quality stocks which are available at discounted prices and predicted through market determined movement flows.

LSTM (Long Short-Term Memory) - LSTM is a neural network structure which feeds forward the exclusive rights to predict a phenomenon of correct prediction which is primarily based on artificial recurrent neural network. It gains the capability to process sequences of data through its network structure. The entire common LSTM unit gets composed of a cell, an input Gate, an output gate and a forget gate (Jiang *et al.,* 2017);
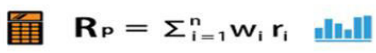
a) Cell Structure remembers the valued structures rather than the arbitrary time intervals.
b) Gates regulate the flow of information in and out of the cell.

LSTM is organized to achieve its best lapses due to its solution by vanishing gradient problem. The solution process gets allowed by the flow of information unchanged through unchanging gradient LSTM units (Wu *et al.,* 2018). In case of non-linear time series forecasting which involves vanilla RNN through back propagation, because of computational process involvement and finite precision numbers (Round off Errors), RNN's keep feeding the gradients as vanished or explosion based on no change (tend to zero) or infinite change (tend to infinity) respectively. Generally, LSTM contains a forget gate that allows to sequentially ignore the error values that are back propagated from the output layer and slowly with repetition would cut them off through the feed forward neural network (Siami *et al.,* 2019). The recurrence would only allow LSTM to train for those gradients for which the weight

updates seemed effective for future value references (Alahi *et al.,* 2016). Finally, the effective weight update references and gradient growths suggests the factor of propagation that associates with time series forecast for any value based on training and testing values in terms of accuracy.

Portfolio Optimization - Portfolio is the mix of different assets with different proportions as investment for one specific entity. It's observed that, the best mix of assets associates the best profitability given the lowest risk level in comparison where the situation of holding only one asset as the measure of investment for earning would define the risk return parity. It's generally a selection process that uses the right investment tools in the correct proportions to generate optimum returns with a balance of risk and return measure associated with the investments (*Bauman et al.,* 2012).

Portfolio Return Formula

$$R_P = \Sigma_{i=1}^{n} w_i\, r_i$$

The overall return of a portfolio is directly dependent on the weight allocation of investments in each portfolio assets and the return of each of these assets defined along the risk factor. As Markowitz Portfolio Theory directly relates to return and risk as proportional to each other, as the probability of risk reduces with introduction of less returning assets into the portfolio, hence the ideal return point should be the actual mix of return and risk after which each additional unit of return would bring more than one additional unit of risk (Layard *et al.,* 2008). Hence the effective return of a portfolio is possible when we attain the highest Sharpe ratio. Sharpe ratio is the measure of return and risk. If a Sharpe ratio remains higher than 1 that means return of the portfolio is greater than the risk associated with price variations (Wang *et al.,* 2011).

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p}$$

R_p = Return of portfolio
R_f = Risk-Free rate
σ_p = Standard deviation of portfolio's excess return

Research Process - Process 1:
a) Selecting a random sample of 100 stocks as investment assets in accordance to various online stock suggestion platforms such as:
www.fknol.com; www.businessinsider.com; www.investing.com; and www.fools.com

b) Applying K Nearest Algorithm in order to select stocks within the given criteria of preferential EPS and Current Ratio suggested through effect-tive researches.

c) Applying LSTM as nonlinear forecasting method for each of the selected stocks for a given time range as training purposes priory and for testing purposes later so the effective forecasting quality of LSTM in terms of accuracy could be suggested.

d) Associate the positive returning stocks with their LSTM forecasted prices in the PyPortfolioOpt for portfolio optimization.

Process 2:

a) Select all the suggested crypto currencies for investment purposes from
https://finance.yahoo.com/cryptocurrencies/
In reference to their historical price data and from market suggestions such as from -
www.nasdaq.com; www.money.usnews.com
www.analyticsinsight.net; www.capital.com
www.investopedia.com

b) Apply LSTM to forecast the future prices of the crypto currencies.

c) Mix the winner stocks for a given time period with Long only gaining crypto currencies in order to build a portfolio.

Compare the result of Process 1 and Process 2 in reference to Sharpe Ratio value and ROI within the context of gains for both of these portfolio's, in order to determine if mixing Crypto currencies in portfolio would make investment more profitable.

METHODOLOGY:

In order to select the stocks based on effective financial ratio as boundary line for criteria's, it's more important to identify the standard ratio for which the stocks could have better or higher priced future value. According to Chen & Shimerda it found that the financial ratio values should have following limits:

EPS >>>> AS High as Possible
1.0 < Current Ratio <= 2.0

In the raw analysis process xi and yi could be considered as reference values for EPS and Current ratio respectively. Through K Nearest neighbor clustering process, this research specifically focuses towards selecting those assets which are effectively closest to

the targeted point of reference suggested by the expert researches and standard accounting guidelines.

Ux is the average point for EPS
Uy is the average point for Current Ratio

Arg-Min (SQRT‖ xi – Ux‖^2 + ‖ yi – Uy‖ ^ 2)=

Minimum Point of Difference
Hence the best asset chosen will have the least difference between the intended asset features and target assets.

Min (ABS (K (Xi, Yi) – K (Mod X, Mod Y)))
Where, ABS refers to absolute value. K (xi, yi) refers to the respective weight attributes of EPS and Current Ratios for selected assets. K stands for the values referred with each variable.

```
for i in asset:
        If
            EPS >>>> As High as Possible && Current Ratio of
Asset (K): (1 < K <= 2) Accept
        Else:
                Reject
```

The process of asset selection on the basis of evaluated criteria is thus completed. It's seen that after the selection process we need to get assured which asset would probably make the most benefit in terms of future investment. In order to make this process successful we have selected the process of non-linear time series price forecast in association with LSTM which has the capability to remember the necessary information for forecasting but would ignore irrelevant information for forecasting purpose.

Selection of Crypto currencies - Crypto currency selection process would also get associated with two criteria's –

a) The crypto currency requires to have high degree sales i.e. the crypto currencies will have high volume

b) The crypto currency needs to have a cheaper contribution margin. The reason behind this is to make the contribution amount being made available to most investors.

Hence our selection process would comprise under the same principle of K Nearest Neighbor Clustering.

Ai = Volume of Crypto currency
Bi = Price of Crypto currency

Ua = Highest Volume of Crypto currency in terms of Sale

Ub = Lowest Price of Crypto currency

MIN (ABS ∥ Ua – Ai ∥) = Selection of Crypto currencies with lowest possible difference in volumes for Crypto currency……………………(1)

MIN (ABS ∥ Ub – bi∥) = Selection of Crypto currencies with lowest possible price… (2)

Selecting the common among them will suggest the best possible crypto currency to invest in for optimized portfolio performance.

Results (Equation (1)) ^ Results (Equation (2))
The final results will be the appropriate crypto currencies which are required for portfolio optimization processes.

LSTM (Long Short-Term Memory) - LSTM is a specialized RNN (Recurrent Neural Network) that has the capability to remember the relevant and required information's rather than irrelevant information. Generally, in the process of RNN it starts from activation function. Activation Function is method that defines how the weighted sum of input gets transformed into an output from a node or nodes of layer in the network. a<t> refers to the activation function. The output function refers to the output of the previous time step based on the input x<t>and activation function a<t> based on the x<t-1> and a<t-1> as input and activation function values of previous time instance. y<t>refers to the output based on the input at that specific time step x<t>.In order to effectively run the activation functions to generate output, its highly important for all the following alternatives to function properly.

$$W_{ax}, W_{aa}, W_{ya}, b_a, b_y$$

All of these factors have effective affect in the RNN structure through which the following results are generated:

$$a^{<t>} = g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a)$$

In the activation function, the bias of the activation function is getting adjoined with the weight allocation of Wax for input and Waa for activation function. Similarly for the output activity:

$$y^{<t>} = g_2(W_{ya}a^{<t>} + b_y)$$

The resultant activation is associated with the bias factor of output and Wya as the allocated weights for

output. It relates with bias and market factor for price prediction g1, g2 remains the activation functions that take effect one after another. All of these functions make the recurrent network to be effective. Hence the complete RNN unit is defined as
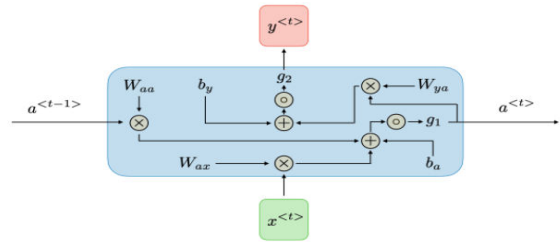
Where each time step motion allows to effectively train based on the given information thus generated from the network. Each of these steps thus illustrates the favorable fact of previous asset price and its relativity with the new asset price.
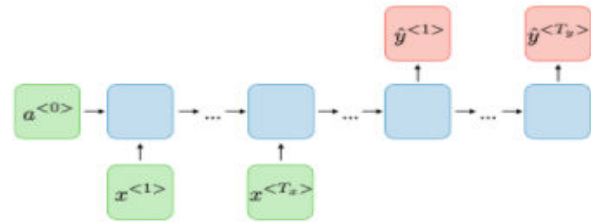
As many instances would be included in this feed forward neural network system in order to train the information's to gather relevant information in response to situational advancements. The RNN architecture preferentially allows the training process to be completely dependent on biased conditional approaches. The bias is the principal component in price prediction because each stock is relative market through beta but also each stock has own performance capability (Stosic *et al.,* 2019). In the process of LSTM, there are various gated structures like the following:

$$\Gamma_u \ \Gamma_r \ \Gamma_f \ \Gamma_o$$

Γ Represent the notation of gates in the RNN process. The above notations represent update gate, relevance gate, forget gate and output gate respectively. Thus the gates allow remembering features for predicting the price or cutoff the features for predicting the price. The following function explains the association of gates:

$$\Gamma = \sigma(W x^{<t>} + U a^{<t-1>} + b)$$

In order to comply, the specifics of gated structure, sigmoid function is being shows here but the relu and soft max functions could be integrated in this equation

as well with respect demand of situation. Each function has its own benefits and applications towards predicted suggestions.

$$g(z) = \frac{1}{1 + e^{-z}}$$

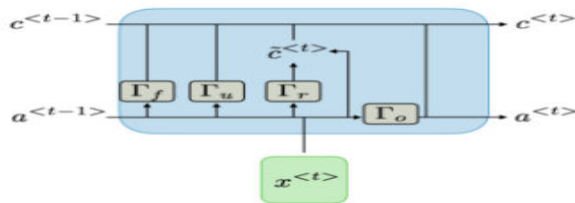Generally sigmoid function refers to g(z) as an output existent from [0,1]

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

Tanh inadvertently refers to the valued answer of g (z) from [-1, 1]. Finally, relu

$$g(z) = \max(0, z)$$

Refers to the maximum value between 0 and z that gets attained due to recurrent neural network process.

Application through LSTM

In association with RNN, LSTM is observed to indulge various gates within the processing unit. Thus, the target of LSTM lies on to the prospects of effective evaluation and training revolution that assists;

a) Proper number of trainings that guides to project that intended value for future forecasting
b) As RNN experiences both exploding gradient and vanishing gradient problem hence the output values that would assist the activation functions for future values hence through RNN clipped methods are emphasized to deliberately control the exploding gradient problem.
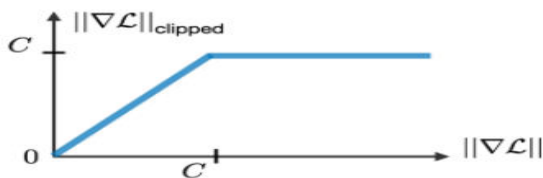
Through the clipping process, RNN thus eliminates the risk of infinity by maintaining a parallelism through boundary line. Although RNN effectively solves the exploding gradient problem but it's unable to solve the vanishing gradient problem. If we carefully look at the gated structure of RNN unit in LSTM, we can observe that there are four gates present in the unit. In the primary output function of c~<t> Relevance gates are

associated with previous activation functions and present input. The situational bias is added with it to make the primary output effective in association with the activation functions, which could be either of softmax, relu, tanh and sigmoid etc.

$$\tanh(W_c[\Gamma_r \star a^{<t-1>}, x^{<t>}] + b_c)$$

In case of final output function $c<t>$, the primary output associated with the update gate structure and the forget gate model will be added with previous output in order to determine if any required information's are supposed to be remembered or not to make the prediction process much efficient. This gated feature thus makes the elimination of vanishing gradient problem.

$$\Gamma_u \star \tilde{c}^{<t>} + \Gamma_f \star c^{<t-1>}$$

Furthermore, this process makes the activation function for the next step highly efficient, as with each step of final output the activation function is remembering which are relevant information's those are required for prediction process.

$$\Gamma_o \star c^{<t>}$$

Relevance with Financial Forecasting - As financial forecasting is highly dependent on nonlinear motions, hence LSTM makes much accurate predictions rather than linear prediction methods of Linear regression and ARIMA process (Abediyi *et al.,* 2013). Each step of pricing is denoted against a fixed level of previous level prices as reference. For the purpose of research process, we have indulged the number of observations steps as -

Product	Data Sample Size	Repetition Steps
Stock	Large	60
Stock	Small	20
Crypto-Currency	Large	60
Crypto-Currency	Small	20

Epochs for Loss Reduction - In RNN, the loss function is denoted under the jurisdiction of loss occurrence through training process and when the previous step feeds irrelevant information for forward step. Here we can see that among the loss occurrences if the primary output commences less loss than the final output, then the previous information should be chosen in the feed forward process.

$$\mathcal{L}(\hat{y}, y) = \sum_{t=1}^{T_y} \mathcal{L}(\hat{y}^{<t>}, y^{<t>})$$

LSTM preferably sustains to perform the exact step of loss reduction through its gated structure process, hence with effective training it's seen that the loss in re-ference to predictions has reduced. Throughout our extended Epoch calculation, it's seen that with favorable observations, epoch of 75 has made effective predictions for the process to be successful.

Back propagation Method - In order to make the financial price forecasting issue suggestively good with reference to time-based propagations, the derivative of loss L with respect to matrix W needs to be updated in reference to each time steps. It's found that through such process, associated weights which will applied for forecasting the price will get update din each step as the relevant information is being updated and forwarded through the activation steps.

$$\frac{\partial \mathcal{L}^{(T)}}{\partial W} = \sum_{t=1}^{T} \frac{\partial \mathcal{L}^{(T)}}{\partial W}\bigg|_{(t)}$$

It's seen that through observation steps and feed forwarding process of LSTM the loss function sustainably gets reduced throughout for large datasets much favorably. In case of smaller datasets, if we practice taking the similar number of training steps then the loss doesn't get reduced. But moreover, with shorter datasets and shorter training steps its seen that the losses are reduced favorably.

Portfolio Optimization - A Portfolio is a mix of assets, which if combined properly considering;

 a) Proper Time for Investment
 b) Proper Amount for Overall Investment
 c) Investing the right amount in each asset

Would bring the best possible return of the investment made in the financial markets.

Hence expe cted portfolio return will be defined as
E (Pr) = Wai * Rai + ……………….. + Wan * Ran

Here i refer to all the assets present in the portfolio and n refers to the final asset number in the portfolio. W represents the weight or how much of investment contribution did we make in the portfolio. R stands for the return of each individual respective asset which are present in the portfolio. In association with getting a return from the portfolio, the reason behind reducing the risk-free earning is that throughout all the investment processes the risk-free earnings are already

making an acceptable profit, hence the actual profit accrues only when return of the portfolio is greater than the risk-free rate and the return difference is greater than portfolio risk.

Sharpe Ratio > 1 ………… Acceptable
Sharpe Ratio <= 1 ……….. Unacceptable

In relationship to selecting crypto currencies, it found that the selection criteria were highly dependent on volumetric sales basis. The secondary idea lies within the context of cheaper contribution margin. In the boundary of these the selected assets were thus used for financial performance and optimized performance.

RESULTS AND ANALYSIS:

Among the online portal of suggestions for the purpose of research we have collected 100 stocks which have been considered by the investment portals for investment. After conducting the K nearest Neighbor clustering, it found that a sample of 42 stocks based on Higher EPS values and Higher Current Ratio has been found out. With the list of assets, which are required to be best possible mix hence within the boundary of volumetric suggestions, it's been figured out that the below list maintains to make the best possible list of asset selection. The secondary step would be to make selections of crypto currencies those would make effective contributions in the portfolio in reference to overall portfolio return but less risk occurrence. Hence under the effective criteria based on higher volume but less price several crypto currencies are chosen.
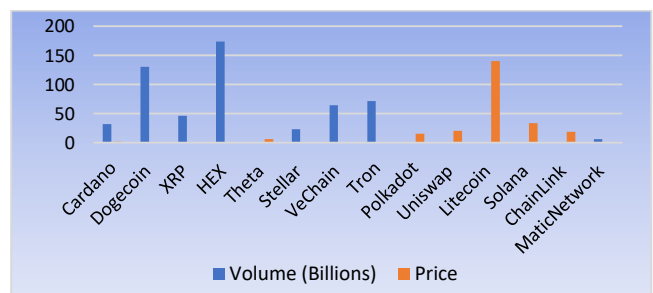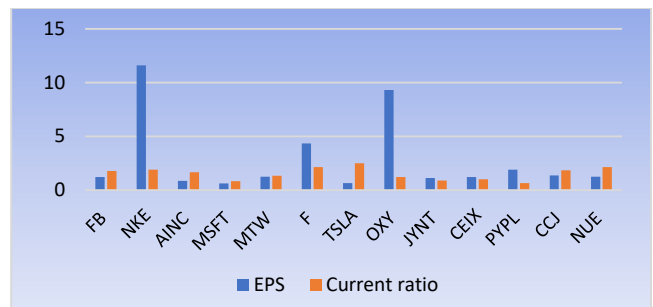
In this research process we have seen that in case of crypto currency investment we can see the pattern that with the increase in prices the number of volumes of crypto currencies. In order to forecast the future prices for both stocks and crypto currencies, it found that LSTM has made significant effectivity in relation to error calculations for the forecasting process when they are effectively undergone through a specific number of epoch steps. In our research process in order to comply a standard fashion of similar number epoch practices to reduce the loss calculation, the LSTM process has been run through various epoch steps.

The above chart is the average representation of loss accumulation that accrues through LSTM training process. It's seen that nearly after the 50-60th Epoch the loss count basically becomes null, hence to make the process of calculation faster the total number of epochs are 75 for both stocks and crypto currencies.
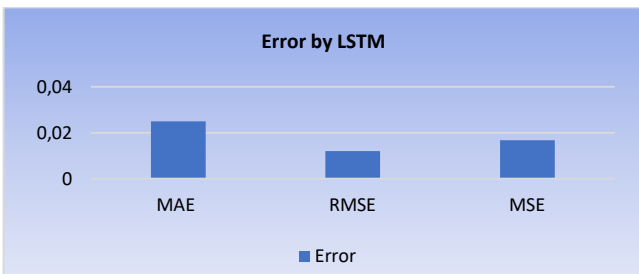
In order to check the accuracy by forecasting for each single stocks and crypto currencies, the accuracy results for the stocks and crypto currencies are explained below:
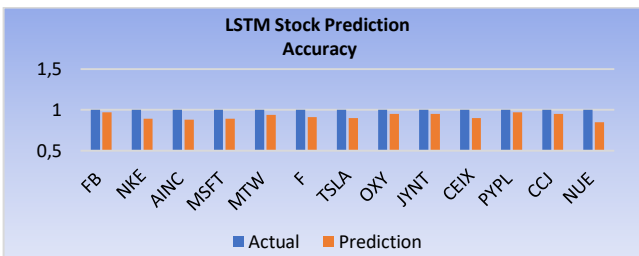
For crypto currencies LSTM makes effective predictions as well. Even though this research has made various comparisons against other price forecasting

methods like ARIMA, SVR, Random forest Regression, Decision Tree Regression and Adaboost Method. But with various research analyses it could be well established that LSTM by far makes the most effective forecasting.

Portfolio Optimization Process - Using Portfolio Optimizer, the simple stock portfolio made following returns. Such returns were possible as LSTM made effective suggestions regarding these stocks and their future price forecasts.

```
from pypfopt.efficient_frontier import EfficientFrontier
from pypfopt import risk_models
from pypfopt import expected_returns

#Portfolio Optimization
#Calculate the Expected Returns
#Calculate the annualized Sample Covariance Matrix of Asset Returns
average = expected_returns.mean_historical_return(df)
s = risk_models.sample_cov(df)

#Optimize for Maximum Sharpe Ratio
ef_line = EfficientFrontier(average, s)
weights = ef_line.max_sharpe()
cleaned_weights = ef_line.clean_weights()
print(cleaned_weights)
ef_line.portfolio_performance(verbose = True)

OrderedDict([('AINC ', 0.04521), ('F', 0.0), ('MTW ', 0.02052), ('TSLA ', 0.0), ('OXY ', 0.00931), ('JYNT ', 0.36875), ('CEIX
', 0.0), ('PYPL ', 0.0), ('CCJ', 0.0), ('NUE ', 0.22714), ('FB', 0.0), ('NKE ', 0.0), ('MSFT ', 0.32908)])
Expected annual return: 437.1%
Annual volatility: 25.2%
Sharpe Ratio: 17.29

(4.370631854712281, 0.2516671286500466, 17.287247158773813)
```

```
portfolio_return_pct = (((sell_ainc + sell_mtw + sell_oxy + sell_jynt + sell_nue + sell_msft +
                        fund_remaining) / (buy_ainc + buy_mtw + buy_oxy + buy_jynt + buy_nue + buy_msft)) - 1) * 100

portfolio_return_pct
```

```
116.38849232100364
```

In relation with effectively predicted prices, the real prices are compared in terms of allocations of stocks and actual gain of the stocks in comparison to predicted stock gains. It's seen that predicted portfolio is 95.72% accurate in terms forecasted price and in respect with allocation of resources. The predicted portfolio with the usage of LSTM has successfully identified the best possible assets but the allocation of the assets was a bit different for real portfolio and predicted portfolio. The predicted portfolio and real portfolio have a short difference of 17.29 and 17.63 respectively as the new Sharperatio's.

```
real_portfolio_gain = 2810.55
forecasted_portfolio_gain = portfolio_gain
effective_prediction = forecasted_portfolio_gain / real_portfolio_gain
effective_prediction
```

```
0.9572218960701637
```

As crypto currencies are operating always hence the best possible representation of prices are considered for that specific period. This research focused to represent the market performance of mixed portfolio and pure stock portfolio in a highly bullish market situation considering the period of January February and March of 2021 as the quarter year for portfolio performance testing. As the general market was highly profitable, this research process was deliberately created in order to check if statistical and nonlinear forecasting method

could make the investment process more effective than the general market return scenario. In reference to mixed portfolio scenario, it seems that due heavy fund allocation in crypto currency, and serious profit allocation assigned through the introduction crypto currencies as alternate investments into the portfolio. It's observed that with the introduction of crypto currency the investment amount shift towards crypto currency was huge due to profit accumulation.

```
from pypfopt.efficient_frontier import EfficientFrontier
from pypfopt import risk_models
from pypfopt import expected_returns

#Portfolio Optimization
#Calculate the Expected Returns
#Calculate the annualized Sample Covariance Matrix of Asset Returns
average = expected_returns.mean_historical_return(df)
s = risk_models.sample_cov(df)

#Optimize for Maximum Sharpe Ratio
ef_line = EfficientFrontier(average, s)
weights = ef_line.max_sharpe()
cleaned_weights = ef_line.clean_weights()
print(cleaned_weights)
ef_line.portfolio_performance(verbose = True)

OrderedDict([('AINC ', 0.13957), ('F', 0.0), ('MTW ', 0.00934), ('TSLA ', 0.03774), ('OXY ', 0.00609), ('JYNT ', 0.0), ('CEIX
', 0.0), ('FYPL ', 0.0), ('CCJ ', 0.07601), ('NUE ', 0.05019), ('FB', 0.19452), ('NKE ', 0.0), ('MSFT ', 0.13774), ('xrp ', 0.
0), ('dogecoin ', 0.02748), ('cardano ', 0.32131), ('stellar ', 0.0), ('theta ', 0.0), ('vet ', 0.0), ('trx ', 0.0)])
Expected annual return: 206856.9%
Annual volatility: 55.9%
Sharpe Ratio: 3697.71

(2068.5694095707076, 0.5594140667596029, 3697.707176998153)
```

In this mixed portfolio the Sharpe ratio is more than 3500 as each crypto currency gained 300% during this quarter yearly period. It's also seen that not only crypto currencies own the portfolio but also stocks own portfolio share as it would definably reduce the portfolio variance, hence although we can see that MSFT was present for both pure stock portfolio and mixed portfolio but the quarter yearly jump is less than 6%.

```
#Get the discrete allocation for each pair of Stock
from pypfopt.discrete_allocation import DiscreteAllocation, get_latest_prices

latest_prices = get_latest_prices(df)
weights = cleaned_weights
da = DiscreteAllocation(weights, latest_prices, total_portfolio_value = 5000)

allocation , leftover = da.lp_portfolio()
print('Discrete Allocation: ', allocation)
print('Funds Remaining: ${:.2f}'.format(leftover))

Discrete Allocation:  {'AINC ': 84, 'MTW ': 2, 'OXY ': 1, 'CCJ': 24, 'NUE ': 3, 'FB': 4, 'MSFT ': 3, 'dogecoin ': 2849, 'cardan
o ': 1227}
Funds Remaining: $0.01

portfolio_return_pct = (((sell_ainc + sell_mtw + sell_oxy + sell_ccj + sell_nue + sell_msft + fund_remaining + sell_doge + sell_
                        sell_fb) / (buy_ainc + buy_mtw + buy_oxy + buy_ccj + buy_nue + buy_msft + buy_cardano + buy_doge + buy_
portfolio_return_pct

97.62856864274558

real_portfolio_gain = 2510.55
forecasted_portfolio_gain = portfolio_gain
effective_prediction = forecasted_portfolio_gain / real_portfolio_gain
effective_prediction

0.979965159427172
```

Drawbacks & Further Improvements - It's seen that the overall initialization of stock and crypto currency selection is purely based on financial ratios and market factor indicators but this research requires more data considering there are nearly 10 more financial ratio's which are important for stock classification and selection process. There are various statistical indicators like EMA, MACD and Bolinger bands as practical de facto indicators for investment but this process purely got depended on financial ratios and that also only in EPS and Current ratio for investment purpose which seems a narrow section for research preview. In order to make the Crypto currency an important investment alternative we need to start making effective

selection criteria rather than simple price and volumetric indication. Moreover, it's found that only LSTM has been used in this research paper to illustrate regarding price forecasting process on the basis of suggestions other researches. More research is suggestive regarding considering LSTM as the best method and its proof should effective in the process of investment-based forecasting.

CONCLUSION:

It's seen that for pure stock portfolio the profit earnings are bit more than the mixed portfolio although the observed Sharpe ratio is way higher for mixed portfolio rather than pure stock portfolio. So, in terms profit expectation we should consider making investment only in stocks but in order to make the investment safer and much prediction oriented with less volatility expectation we should think about making investments in mixed portfolio containing crypto currencies and stocks.

ACKNOWLEDGEMENT:

CONFLICTS OF INTEREST:

The authors declared no possible conflicts of the interest with respect to the research, authorship and publication of this article.

REFERENCES:

1) Alahi, A., Goel, K., Ramanathan, V., (2016). Social LSTM: Human Trajectory Prediction in Crowded Spaces. IEEE *Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/CVPR.2016.110

2) Amihud, Y., & Mendelson, H. Husna, A., & Satria, I. (2019). Effects of Return on Asset, Debt to Asset Ratio, Current Ratio, Firm Size, and Dividend Payout Ratio on FIirm Value. *International Journal of Economics and Financial Issues,* **9**(5), 50–54.

3) Amihud Y, and Mendelson H. (1991). Liquidity, Asset Prices and Financial Policy, *Financial Analysts Journal,* **47**(6), 56–66. https://doi.org/10.2469/faj.v47.n6.56

4) Aren, S., & Zengin, A. N. (2016). Influence of Financial Literacy and Risk Perception on Choice of Investment. *Procedia - Social and Behavioral Sciences,* **235**, 656–663. https://doi.org/10.1016/j.sbspro.2016.11.047

5) Baumann, P., & Trautmann, N. (2012). Portfolio-optimization models for small investors. *Mathematical Methods of Operations Research,* **77**(3), 345–356. https://doi.org/10.1007/s00186-012-0408-3

6) Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *J. of Accounting Research,* **4**, 71. https://doi.org/10.2307/2490171

7) Burggraf T. (2019). Risk-Based Portfolio Opti-mization in the Crypto currency World, SSRN, 25th September, 2019. https://doi.org/10.2139/ssrn.3454764

8) Cao, J., Li, Z., & Li, J. (2018). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and Its Applications*, **519**; 127-139. https://doi.org/10.1016/j.physa.2018.11.061

9) Caparrelli, F., D'Arcangelis, A. M., & Cassuto, A. (2004). Herding in the Italian Stock Market: A Case of Behavioral Finance. *J. of Behavioral* Finance, **5**(4), 222–230.

10) Capon, N., Fitzsimons, G. J., & Alan Prince, R. (1996). An individual level analysis of the mutual fund investment decision. *Journal of Financial Services Research,* **10**(1), 59–82. https://doi.org/10.1007/BF00120146

11) Chen, K. H., & Shimerda, T. A. (1981). An Empirical Analysis of Useful Financial Ratios. *Financial Management,* **10**(1), 51. https://doi.org/10.12691/jfa-2-1-2

12) Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. *Expert Systems with Applications*, **80**, 340–355. https://doi.org/10.1016/j.eswa.2017.02.044

13) Fruehwirt, W., Hochfilzer, L., and Roberts, S. (2020). Cumulation, crash, coherency: a Crypto currency Bubble Wavelet analysis. *Finance Research Letters,* 101668.

14) Gers, F. A., Eck, D., & Schmidhuber, J. (2002). Applying LSTM to Time Series Predictable through Time-Window Approaches. *Neural Nets WIRN Vietri*-01, 193–200. https://doi.org/10.1007/978-1-4471-0219-9_20

15) Jiang, Z., & Liang, J. (2017). Cryptocurrency portfolio management with deep reinforcement learning. Intelligent Systems Conference (IntelliSys), Reinforcement learning, *Intelligent Systems Conference (IntelliSys).* https://arxiv.org/abs/1612.01277

16) Jorion, P. (1992). Portfolio Optimization in Practice. *Financial Analysts Journal,* **48**(1), 68–74. https://doi.org/10.2469/faj.v48.n1.68

17) Kilka, M., & Weber, M. (2000). Home Bias in International Stock Return Expectations. *J. of Psychol. and Finan. Markets,* **1**(3-4), 176–192. https://doi.org/10.1207/S15327760JPFM0134_3

18) KuoChuen, D. L., Guo, L., & Wang, Y. (2017). Crypto currency: A New Investment Opportunity? *The J. of Alternative Investments,* **20**(3), 16–40. https://doi.org/10.2139/ssrn.2994097

19) Layard, R., Mayraz, G., & Nickell, S. (2008). The marginal utility of income. *J. of Public Economics,* **92**(8-9), 1846–1857. https://doi.org/10.1016/j.jpubeco.2008.01.007

20) Leong, K., Pagani, M., & Zaima, J. K. (2009). Portfolio strategies using EVA, earnings ratio or book to market. *Review of Accounting and Finance,* **8**(1), 76–86. https://ideas.repec.org/a/eme/rafpps/v8y2009i1p76-86.html

21) Lin, F., Liang, D., & Chen, E. (2011). Financial ratio selection for business crisis prediction. *Exp Syst with Applications*, **38**(12), 15094–15102. https://doi.org/10.1016/j.eswa.2011.05.035

22) Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Comp and Applications*, **32**, pages: 17351–17360. https://doi.org/10.1007/s00521-020-04867-x

23) Novy R Marx, (2013). The Quality Dimension of Value Investing. National Bureau of Economic Research, Simon Graduate School of Business, University of Rochester.

24) Peng, L., Zhu, Q., Lv, S.-X., & Wang, L. (2020). Effective long short-term memory with

fruit fly optimization algorithm for time series forecasting. *Soft Computing*; 15059–15079. https://doi.org/10.1007/s00500-020-04855-2

25) Sami HM. (2021). Optimizing organizational overall performance, the use of quantitative choice of HR in carrier quarter enterprise of Bangladesh, *Can. J. Bus. Inf. Stud.*, **3**(3), 49-59. https://doi.org/10.34104/cjbis.021.049059

26) Shen, Z., Zhang, Y., Lu, J., Xu, J., & Xiao, G. (2019). A novel Time Series Forecasting Model with Deep Learning. *Neuro computing*. https://doi.org/10.1016/j.neucom.2018.12.084

27) Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series. 2018 17th IEEE *International Conference on Machine Learning and Applications (ICMLA)*.

28) Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The Performance of LSTM and BiLSTM in Forecasting Time Series. IEEE International Conference on Big Data, 9-12 Dec. 2019. https://doi.org/10.1109/BigData47090.2019.9005997

29) Stosic, D., Stosic, D., & Stosic, T. (2019). Multi-fractal behavior of price and volume changes in the crypto currency market. *Physica A: Statistical Mechanics and Its Applications, Fractals,* **28**(05), 2050076. https://doi.org/10.1142/S0218348X20500760

30) Wang, M., Keller, C., & Siegrist, M. (2011). The Less You Know; the More You Are Afraid of - A Survey on Risk Perceptions of Investment Products. *J. of Behavioral Finance,* **12**(1), 9–19. https://doi.org/10.1080/15427560.2011.548760

31) Wu, C.-H., Lu, C.-C., Ma, Y.-F., & Lu, R.-S. (2018). A New Forecasting Framework for Bitcoin Price with LSTM. *IEEE International Conference on Data Mining Workshops*. https://doi.org/10.1109/BigData47090.2019.9005997