
A DEEP LEARNING BASED METHODS FOR FORECASTING GOLD PRICE WITH RESPECT TO PANDEMICS

Mahtab Mohtasham Khani

Institute of Informatics
Istanbul Technical University
Istanbul, Turkey, 34467
khani17@itu.edu.tr

Sahand Vahidnia

School of Engineering and IT
University of New South Wales
Canberra, Australia, 2612
s.vahidnia@unsw.edu.au

Alireza Abbasi

School of Engineering and IT
University of New South Wales
Canberra, Australia, 2612
a.abbasi@unsw.edu.au

November 19, 2020

ABSTRACT

The spread of COVID-19 in the world had a devastating impact on the world economy, trade relations, and globalization. As the pandemic advances and new potential pandemics are on the horizon, a precise analysis of recent fluctuations of trade becomes necessary for international decisions and controlling the world in similar crisis. The COVID-19 pandemic made a new pattern of trade in the world and affected how businesses work and trade with each other. It means that every potential pandemic or any unprecedented event in the world can change the market rules. This research develops a novel model to have a proper estimation of the stock market values with respect to COVID-19 dataset using long short-term memory networks (LSTM). The nature of the features in each pandemic is totally different, thus, prediction results for a pandemic by a specific model cannot be applied to other pandemics. Hence, recognising and extracting the features which affect the pandemic is in the highest priorities. In this study, we develop a framework, providing a better understanding of the features and feature selection. This study is based on a preliminary analysis of such features for enhancing forecasting models' performance against fluctuations in the market. Our forecasts are based on the market value data and COVID-19 pandemic daily time-series data (i.e. the number of new cases). In this study, we selected Gold price as a base for our forecasting task which can be replaced by any other markets. We have applied Convolutional Neural Networks (CNN) LSTM, Vector Out-put Sequence LSTM, Bidirectional LSTM, and Encoder-Decoder LSTM on our dataset and our results achieved an MSE of $6.0e-4$, $8.0e-4$, and $2.0e-3$ on the validation set respectfully for one day, two days, and 30 days predictions in advance which is outperforming other proposed method in the literature.

Keywords COVID-19 · Economy · Deep Learning

1 Introduction

The novel Coronavirus (SARS-CoV-2) disease identified as COVID-19 has been initiated in Wuhan, China and had a quick global spread. This disease first led to a big human health problems in all countries and many cascading financial problems resulted from social distancing and travel restrictions. Governments proceeded with specific procedures to control the speed of spreading the disease such as canceling flights, locking down their national and state borders, preventing most of exports and imports, and shutting down some businesses which lead to a huge economic shock to the world. In this situation, trading which is an indivisible and crucial part of today's life can become a potential of spreading the virus in many ways like associated group of people in this process. Traveling has high risk of spreading the disease and most of the trades are based on traveling [1]. On the other hand, without trading economy will collapse and the effects of such economic devastation will remain for a long time. Consequently, the pandemic has damaged the global economy by creating problems in the world supply chain [2]. For example, equity markets in the EU and US dropped by as much as 30% [3]. Generally, there was a dramatic shock to global trading activities during COVID-19 pandemic such as increasing the demand for essential goods such as medical products and food, as well as a sharp

decrease in the prices of some products like oil, or the collapse of some airlines declaring bankruptcy with hopes to resume operations after the the end of the outbreak [4]. According to a recent review [5] trade implications of the COVID-19 pandemic that China and the rest of the world follow a new pattern which leads some economies to win and some to loose.

Understanding the economic implications and impact of the pandemic on the economy is a way to minimize the impact and support the economic decision makers on their trade choices. Conventional methods such as tacking economic factors of companies and the markets, like statement of cash flow and balance sheet, or technical analysis to understand the market trend based on the volume and price, have been in place for many years and predictions have been made for decades using these methods. With the emergence and advancements in machine learning and deep learning, new methods have emerged to better understand the patterns and make more accurate predictions. These data-driven methods require access to large datasets, which have been available during the recent years, and has paved the ground for the research in this direction [6]. The nature of the features of each pandemic is different, hence, one result cannot be applied to other cases. Hence, in this study we develop a framework facilitating the overall procedure of time series analysis and predictive robust models in various cases.

In this work, we use LSTM for time series analysis in order to better understand the patterns in the sequences and improve the prediction performance. LSTMs are basically artificial recurrent neural networks (RNNs) that can process the entire sequence of data. LSTM was first investigated by Hochreiter [7] to deal with the problem of vanishing or exploding the gradients of RNNs. LSTM Networks are a solution for extracting the pattern of the input data-set which spans over a sequence of time which can be responsible for COVID-19 as a nonlinear feature over a sequence of time.

The main objective of this research is to analyze and explore the implications of COVID-19 on the economy considering short-term implications and patterns of pandemics on world trade and market sector, by studying the case of Gold price on the stock market, based on the COVID-19 time series and eleven sectors of the market during the recent years and months. We study the correlation between different sectors and COVID-19 new cases for distinguishing the effective features in market prediction. This paper contributes to market prediction accuracy by developing an accurate forecasting model on the stock market values with respect to the COVID-19 dataset, using LSTM. Our forecasts are based on the market value data extracted from Yahoo! Finance and COVID-19 pandemic daily time-series data (i.e. the number of new cases). We applied our method for 1, 2 and 30 days in advance predictions, concerning COVID-19 data. The performance of our proposed method has been examined and validated using Mean Square Error (MSE) on the validation dataset, by comparing to the most recent developments in the field.

2 Literature Review

2.1 Economy and Pandemics

Human beings have faced many pandemics throughout the history such as SARS Epidemic [8], Middle East respiratory syndrome (MERS) [9], West African Ebola epidemic [10], and H1N1 Swine Flu pandemic [11]. The pandemics are known to have significant effects on economy, some of which may persist for decades after the pandemic [12]. There are many studies trying to understand the pandemic and its effects on trade. Some of these studies are purely economic and are not relying on machine learning for predictions. Machine learning is a branch of analytic predictions in which has lower reliance on human expert supervision for understanding and analyzing the reasons and causes [13]. Machine learning algorithms doesn't count out the need for human expert analyzes, meanwhile they are one of the best solutions for extracting the underlying patterns [14]. The studies in this section analyze the economical, social, and even political aspects of the markets to understand and explain the market. These studies provide insights into the economic point of view of the pandemics and helps to better understand the variables. In a review paper by Barua [5], likely trade implications of the COVID-19 pandemic have been found to provide a better understanding of the impacts of COVID-19 on coming days. This study investigates a standard trade framework and then proposes a theoretical mapping that depicts the progress of trade implications. The study then reviews some real-life evidence for testing his map and concluded with clues for controlling pandemic situations. The study divides the effects of a pandemic into different short and long term stages, and it suggests reducing reliance on specific countries to reduce the economic impact. In another survey about macroeconomic implications of COVID-19, the effects of negative supply shocks on shortage of demands is discussed [15]. They argue the future of economic shocks of COVID-19 as well as presenting the theory of Keynesian supply shock, which is about larger changes in aggregate demand than the shocks, in comparison to standard supply shocks which concluded in a good understanding about availability or lack of some goods in the pandemic situation. The Center of Economic Policy Research (CERF), which is a network of economists mostly from European universities, has asserted that "The virus is likely to be as 'contagious' economically as it is medically" [16]. In the past recessions, global trade has slowed faster than global growth. This study [16] also discussed about the demand shock on trade as well as the supply-sides of this virus. This study is an analysis across the COVID-19 and trade which

emphasized on the danger of permanent collapse on trade system. Also the optimal lock-down policies for minimizing the output costs of the lock-down was researched by [17]. The study uses a linear economy model to formalize the planner's dynamic control problem and identifies the features to measure the optimal intensity, shape, and time that this lock-down will last. Gormsen et al.[3] studies the impacts of Coronavirus on stock prices and the growth expectation. They show how effective are news, events, and the data on dividend futures on forecasting and analysing the drop and growth of the economic over time. They've studied the expected growth in US and EU S&P 500 and Euro Stoxx 50 and estimated that Expected growth over the next year is down by 2.6% both in the both US and the EU.

2.2 Stock Market Prediction

The studies on market prediction and the surrounding literature can be generally divided in two categories: (1) the studies which predict the market using social networks analysis; and (2) the studies of time series analysis. The former category comprises the methods like the diffusion models which are useful for minding lots of latent information and market predicting. Li et al. [18] investigates the research methods and techniques on diffusion models which could facilitate the prediction of social network influences or predicting the trade future leads. In another study, [19] proposed a system for detecting influenza epidemics using Twitter data. They extracted the tweets that mention actual influenza patients using the support vector machine (SVM) based classifier for separating the negative and positive tweets.

The second category, which are more in terms of quantity, mainly focus on methodologies to predict various financial time series. LSTM networks are always one of the best solutions for a sequence of data and single data points [20]. Hence, Fischer and Krauss propose and use the use of LSTMs in financial time series. The feature space in their study is the standardized daily stock market return of specific markets and days. They provide a comparison of LSTM, random forest, a standard deep net, and logistic regression and conclude that LSTM outperforms all other methods. In a more recent study, Livieris et al. [21] implement a CNN-LSTM model to predict gold price time-series and its fluctuations. The paper asserts that the combination of LSTM layers with some other convolutional layers increases in the forecasting performance. They do not consider other variables and the pandemic, which could be limitation of this study.

In a review paper on time series classification [22], it has been emphasized that LSTM networks are the most efficient way for time series classification. Among the usages of LSTMs, time series forecasting problems are considered to reduce the complication of study in comparison to the other traditional ones [23]. This study [23] suggest to use traditional methods in non-complicated tasks and to use LSTMs as substitute methodology. They conclude that LSTM has a great promise and opportunity for applying to the problem of time series forecasting. In another study [24] LSTM is used in order to predict the China's stock market return and has reported the power of LSTM in stock market prediction. Besides, Persio and Honchar [25] investigated a similar problem by applying artificial neural network architectures to predict trend movement on stock market based on past returns. According to their results, in a comparison of multi-layer perceptron (MLP), convolutional neural networks, LSTMs for feature extractions, and the combination of wavelet transform and CNN, the latter one has achieved the best results.

In a study surveying time series forecasting of COVID-19 transmission in Canada [26], LSTMs are used to predict the possible stopping time of COVID-19 in Canada and around the world. They propose multi step LSTM method and predicted 2, 4, 6, 8, 10, 12 and 14th day for two successive days. The study proposes the use of bidirectional LSTM for the forecasting model. LSTM also used as a data driven estimation method in India for predicting the spread of COVID-19 and the effects of preventing protocols [27]. In this study they predicted the number of COVID-19 cases of India 30 days ahead.

The prediction of number of cases in a pandemic has demonstrated to be somehow predictable based on the historical data. However, the same is not true for market data, as it has far more fluctuations. Additionally, as the studies suggest, the world pandemic situations would make the market movements more chaotic and unpredictable. As a results, we aim to adopt LSTM,as one of the best forecasting methods, to develop a predictive model for stock market prediction using LSTM-series by considering COVID-19 data-set for market prediction for the firs time. This will establish a guideline for future works of this category.

3 Methodology

Our method comprises four steps. First,our data which is the overall market sectors and tickers are analyzed, compiled and extracted. Then, relevant features are processed and extracted. Later, LSTM models are trained and tuned on the aforementioned features to create forecasting models with a range of hyper parameters, steps and history periods. Finally,the method and the results are evaluated and validated to be used for future forecasting tasks.

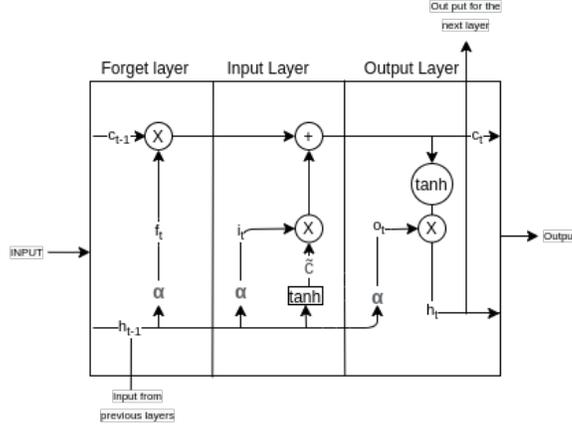


Figure 1: LSTM internal architecture which consist of: h_{t-1} for the output from last LSTM unit, c_{t-1} for memory of the last LSTM unit and c_t for candidates of cell state at time t Two α s to represent non-linearity as a sigmoid layers, \tanh : represent \tanh layer Vector operations: \times expresses scaling the information, and $+$ expresses adding information

3.1 LSTM Models

Originally, LSTM was investigated in 1991 by Hochreiter [7] to deal with the problem of vanishing or exploding gradients, which are very common in RNNs. Generally, RNNs are good at handling the sequence dependencies. LSTMs are a type of RNNs which are better suited for larger architectures and more capable at extracting patterns from large sequences of datasets. LSTMs are also known to have better response for non-linearity[28]. The COVID-19 and market time series data show a non-linear behaviour which motivates the application of LSTM in this research.

As illustrated in Figure 1, each LSTM unit consists of three gates: a forget gate which remembers the values over arbitrary time intervals, and two other gates to regulate information into and out of the cell, which are called input and output gates. Each LSTM cell maintains a cell state vector and at each time step the next LSTM can choose to read from it, write to it or reset the cell. These gates gives the ability of controlling the process of memorizing to LSTM, and therefore, it can avoid long-term dependency [29] which is a key factor of solving problems related to COVID-19 with short historical dataset. The parameters of the gates are expressed in Eq. 1, where α expresses the sigmoid function, w_x expresses the weights for the neurons of gate x , h_{t-1} expresses the output from last LSTM unit, x_t expresses current input, and b_x expresses the biases for the gate x .

$$\begin{aligned}
 i_t &= \alpha(w_i[h_{t-1}, x_t] + b_i) && \text{for input gate} \\
 f_t &= \alpha(w_f[h_{t-1}, x_t] + b_f) && \text{for forget gate} \\
 o_t &= \alpha(w_o[h_{t-1}, x_t] + b_o) && \text{for output gate}
 \end{aligned} \tag{1}$$

In this study, we conduct both single-step and multi-step analysis, predict gold prices at least a day ahead. In order to have better sensitivity analysis, we also employ both multi-variate and uni-variate approaches to demonstrate the effectiveness of other variables. There are numerous methods and approaches involving LSTMs to tackle time series forecasting analysis which depends on the dataset and the task. At the following sections these approaches will be discussed and established to pave the ground for comparison and model selection.

3.1.1 Single Step LSTM

To predict the gold price a day in advance, single-step feed forward stacked LSTM networks are used. As mentioned earlier, a series of hyper-parameters and input variables are tested to better understand the effect of feature space on the prediction error. The overall structure of the LSTM networks employed at this section is depicted in Figure 2. Generally, the higher number of LSTM cells within a layer would allow us to have a longer memory. This means that for longer historic days, we can grow the width of LSTM network and vice versa to have an optimal fit. The activation layers in all architectures are Rectified Linear Unit (ReLU) (Eq. 2) and the optimizer of choice for LSTM networks are usually

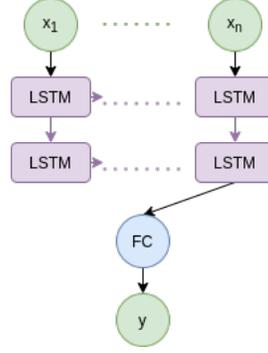


Figure 2: LSTM network architecture. X is input, Y is output, with stacked LSTM cells.

ADAM, as opposed to Stochastic Gradient Descent (SGD) which are usually known for robust optimization. After tuning numerous hyper-parameters, the selection of top models have been presented at table 3.

$$ReLU(x) = \max(0.0, x) \quad (2)$$

3.1.2 Multi-Step LSTM

Forecasting for more than one day or step, makes it a multi-step forecasting problem. The methods for addressing the multi-step forecasting problem can be categorized under vector-output sequence prediction approach and encoder-decoder approach. The Encoder-Decoder approach in addition to the vector output sequence prediction methods are the main focuses of this study. To validate the results and have a comparison to other suggested methods in the literature, the results will also be compared to Bidirectional LSTM and CNN-LSTM.

The output for a multi-step forecasting LSTM can be a vector sequence. This can be achieved by simply adding n-output neurons to a simple vanilla LSTM network. Hence, the overall architecture of multi-step vector output approach is almost identical to Figure 2.

Encoder-Decoder as it is explained by its name, predicts by encoding the inputs and then decoding the output. This approach is used for multi-step time series forecasting[30]. The model was designed to solve sequence to sequence problems like natural language processing[31], text translation, and answering to textual questions. Encoder-Decoder also known to yield good results for image classification, image to text, movement classification, and describing images by text tasks[32]. Encoder-Decoder approach in LSTM can have many different implementations, suiting different workloads. In general, the overall architecture of experimented Encoder-Decoder models in this study can be illustrated as Figure 3. As illustrated, the model takes and encodes the inputs, then repeats the final state of the encoding layer for all time steps. The decoder comprises at least an LSTM layer and time distributed dense layer to provide the output of desired shape and structure.

CNN-LSTMs are the combination of CNNs and LSTMs[6] that are common to be utilized in computer vision problems[33]. CNN-LSTMs are also Encoder-decoder based approaches, where the encoding happens in the CNN section. They have been utilized in various tasks in the literature like caption generation[32] and prediction of gold prices [21]. Thus, CNN-LSTM models have also been experimented in this study and the results are discussed in the following section.

Bidirectional LSTM is inspired by Bidirectional Recurrent Neural Networks [34]. The network learns the sequences both from forward and backward and then concatenates all the data for prediction. Bidirectional LSTM networks can be more beneficial than unidirectional ones in terms of results[35].

Similar to the single-step LSTM approaches, ReLU activation and ADAM optimizer has been used in all architectures. At the following section we discuss the evaluation metrics, the data required for this analysis and forecasting models, and the result of the analysis.

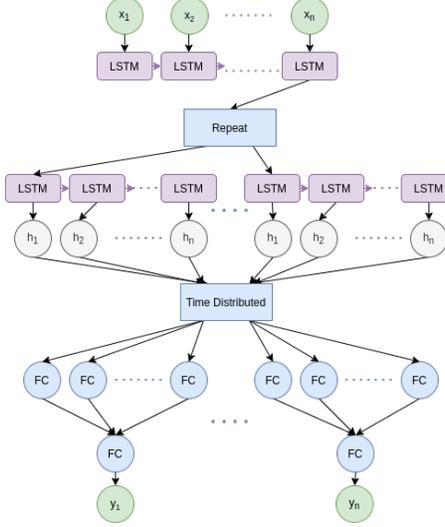


Figure 3: Encoder-Decoder LSTM network architecture. X is input, Y is output, and h is hidden state of the LSTM cells.

3.2 Validation and Evaluation

For validation purpose we separate a 90 recent days period of our dataset as validation set, and validated the model on the aforementioned period. In order to accommodate the randomness factor, the tests have been repeated for a number of times and only the top-performing seeds and training instances have been referred to.

There are various metrics to calculate the loss in regression and prediction tasks. Root Mean Square Error (RMSE), Mean Square Error (MSE), Mean Absolute Error (MAE), Mean Squared Logarithmic Error (MSLE) are all techniques to find the difference between the predicted value and the actual value. However, in this study the models are optimized on the MSE values. Hence, the comparison results will favor this metric. This effects both best-model and training checkpoint selection, and the comparison of different LSTM methodologies and models based on validation errors. However, to better understand the results, we take advantage of all of them which are respectfully defined by Eq. 3, where n is for the number of predictions and y_i is the ground truth of i instance and the \hat{y}_i is the predicted results of them.

$$\begin{aligned}
 RMSE &= \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}} & MSE &= \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\
 MAE &= \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| & MSLE &= \sqrt{\frac{1}{n} \sum_{i=1}^n \log(y_i - \hat{y}_i)^2}
 \end{aligned} \tag{3}$$

3.3 The Data

There are thousands of publicly traded stocks around the world and every one of them can be categorized as a member of the 11 major market sectors [36], including Financial, Utilities, Consumer Discretionary, Consumer Staples, Energy, Healthcare, Industrial, Technology, Telecom, Materials, and Real Estate. These 11 sectors are responding to the key areas of the economy and all the companies in each sector share the same broad focus. The list of corresponding tickers of the 11 market sectors have been collected from ETFdb¹. As the list of the market sectors is long and the market values and volume of the tickers vary drastically, only top 10 tickers with the highest values for each sector has been selected. The market data in this study is gathered for this selected top 10 sector tickers for the past five years from 07/2015 to 07/2020 from Yahoo! Finance, which also covers the recent global COVID-19 pandemic period. On the other hand, we also need COVID-19 pandemic data (including newly infected and total infections) to incorporate in our model and it has been collected from "JHU CSSE COVID-19 Data" daily time series. As the market data has been collected from 30-07-2015 to 30-07-2020 and COVID-19 data starts from 22-01-2020, the COVID-19 values prior to 22-01-2020 has been set to zero to match the dimensionality of market dates.

¹ETFdb

3.4 Understanding the Feature Space

The COVID-19 time-series data comprises both the world and USA data separately. To best utilize this data, an aggregation of these cases have also been recorded and added to the feature space. Then the new cases have been calculated as the difference of the daily cases, yielding six features as follows: Us-new, Us-all, World-wide-new (except US), World-wide-all(except US), Total-new, Total-all. Finally, the case numbers are normalized before feeding to the neural networks, making the feature space even better suited for the task. The stock market data has many missing rows as a result of market closures for holidays and weekends. Hence, data was padded to interpolate the missing data. The reason to prefer padding to other interpolation techniques is that it's intuitive to refer to the final exchange rates and values as the current one. To obtain sector data, the selected ticker data for market close rate, volume and daily average rates are acquired and calculated. Later, the mean of the corresponding values are taken as the overall sector values and then normalized to better fit the neural networks. To better understand the feature space and the relationships among the features, figure 4 illustrates the technology sector symbol average vs. the COVID-19 cases, that hints at a correlation in the data. Hence, a correlation analysis is carried out to better understand these relationships among COVID-19 and the market. This analysis provides further insights to the strength of relationships among the variables and parameters to be used in our model, helping us to confidently include important variables in the feature space and eliminate less important ones.

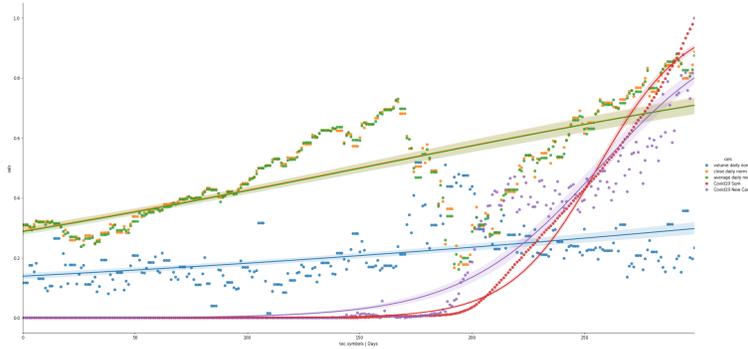


Figure 4: Technology Symbols and COVID-19 time series in 300 days until 03 Jul 2020.

To prepare the data for analysis, the volume and the average of daily sectors values are calculated. Later, the results are normalized and then the logistic regression model of the values are calculated. As the prediction goal and case of this study is to forecast gold price, the feature space should be prepared. To better understand the feature space, the correlation coefficients are calculated against daily gold price (as the possible dependent variable) within the last 300 days period. A strong correlation can be subjective and vary from one study to other[37], but in this study $r = \pm 0.4$ has been selected as the correlation coefficient threshold to eliminate weak correlations. As shown in 1, only some sectors including total new COVID-19 cases in the world, Consumer-staples (closing price), and Technology (closing price) have statistically significant correlations (significance of over 95% and $r > 0.4$ or $r < -0.4$) with gold price. Hence, all the remaining sectors values with non-significant correlations are eliminated to finalize the feature space. Table 1 shows the correlation coefficients across the sectors. In Table 1 the correlations for both 'Close-gold' and 'Average-gold' are presented to illustrate the difference between the correlations of closing price of the market and the daily average prices. Overall, we can see similar results on both.

As shown in Table 1 new COVID-19 cases have stronger correlation (r value) with the market data. On the other hand, it is also observable that daily market average value (normalized) has stronger correlation with COVID-19 pandemic than the market volume (normalized). It should also be reminded that correlation does not necessarily result in causation, yet the strong correlations can bear latent underlying connections, relationships or meanings. For instance, the energy sector has the strongest (negative) correlation coefficient to the new cases. As the cases rise, the energy sector market value falls. This is very sensible, as can be observed from the recent drop in fuel prices, that hints at causality. The same also applies to the industrial sector, which comes second after the energy sector. Financial sector also has a very strong correlation coefficient, with -0.954, which comes third in this table. This indicates that the financial sector has also been hit hard by the pandemic at similar levels.

Table 1: Correlation analysis of all sectors vs. Gold.

Variables	r (Close-gold)	r (Average-gold)
Consumer-staples(Close)	0.817	0.817
Consumer-staple(Avg.)	0.817	0.818
Volume energy	0.438	0.440
Financial(Close)	0.468	0.469
Financial(Avg.)	0.466	0.467
Healthcare(Close)	0.721	0.722
Healthcare(Avg.)	0.722	0.722
Materials(Close)	0.600	0.600
Materials(Avg.)	0.598	0.599
Real-estate(Close)	0.492	0.493
Real-estate(Avg.)	0.495	0.496
Tec-symbols(Close)	0.777	0.778
Tec-symbols(Avg.)	0.775	0.777
Telecom(Close)	0.675	0.675
Telecom(Avg.)	0.675	0.675
Utilities(Close)	0.710	0.711
Utilities(Avg.)	0.720	0.711
Gold(Close)	1.0	0.999
Gold(Avg.)	0.999	1.0
Volume-food	0.4665	-0.682
Covid-all-total-cases(norm)	0.877	0.885
Covid-all-new-cases(norm)	0.881	0.886
Covid-us-total-cases(norm)	0.889	0.898
Covid-us-new-cases(norm)	0.858	0.860
Covid-ww-total-cases(norm)	0.873	0.881
Covid-ww-new-cases(norm)	0.879	0.884
Covid-all-total-cases	0.877	0.885
Covid-all-new-cases	0.881	0.886
Covid-us-total-cases	0.889	0.898
Covid-us-new-cases	0.858	0.860
Covid-ww-total-cases	0.873	0.881
Covid-ww-new-cases	0.879	0.884

The p-values for the corresponding correlations in this table, which already have been filtered, are below $1.59e - 7$.

“ww” notion in COVID-19 cases indicates world-wide cases except the United States.

“US” indicates the United States cases.

“norm” indicates normalized data.

“all” denotes the aggregated cases of the United States and the rest of the world.

“Close” indicates the mean closing price of sector symbols.

“Avg.” indicates the mean daily-average price of sector symbols.

4 Results

The experiments in this study have been implemented in Python 3.8. Keras[38] has been used for deep learning implementations and Scikit-learn[39] has been used for some loss evaluations.

Time series analysis have 1,2 and 30 days of future data (single and multi-step) predictions. The history and the range of the data to infer from has also been a variable tuned in this study for both single and multi-step predictions. The number of historical days to tune are 5, 9, 15, 22, and 30 for the 1 and 2 day predictions, and 45 and 60 historical days for 30 days prediction models. All models have been trained for 400 epochs, with best model checkpoints.

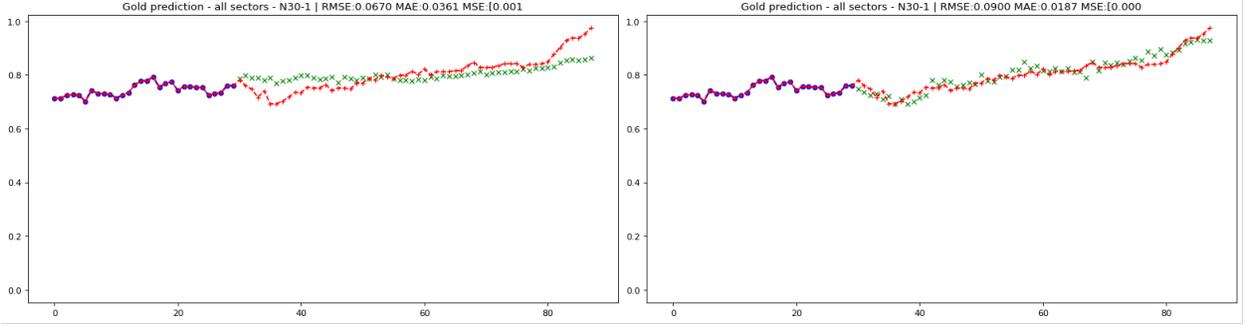


Figure 5: Single-step time series prediction. Figure on the left (a) has been trained on the market data without COVID-19 time series. Figure on the right (b) has been trained on all features including COVID-19 time series data. Red markings are ground truth validation points and green marking are the predictions. The model makes prediction based on 30 days of historical data. The architecture is: LSTM(300)-LSTM(300)-Dense(100)-Dense(1).

4.1 The Feature Space

An objective in this study is to see the effect of COVID-19 pandemic data in the market future prediction error. Hence, the models are trained with varying features to see the influence of COVID-19. In this study as presented at the results Table 3, the inclusion of COVID-19 data in feature space would increase the model’s performance for 1 and 2 days in advance predictions. Taking into account that the variation in feature space might require architectural modifications and tuning, the tests and tuning were carried out on three different feature space of uni-variate gold data, multi-variate data with COVID-19 and multi-variate data without COVID-19. We also notice that as opposed to many other studies which implement uni-variate methods, implementation of multi-variate approach and taking advantage of related variables in the market has improved the performance of the models by reducing their errors. As discussed in *Understanding the Feature Space* section, the features have been selected based on their correlation coefficients with the target market value.

4.2 Single-step

The best results were achieved with 30 days of historical data as sequence inputs for the 1 day predictions, and the LSTM network with architecture of the winner model is mentioned on Table 3. It is worth mentioning that the output of the stacked LSTM layers were taken and the hidden states were ignored, as it was noticed that having the hidden states does not improve the performance. A sample of this experiment has been recorded on the Table 3. The statement regarding the improved predictions using COVID-19 data stands true in single-step models as illustrated in Figure 5. We can see that the predictions can follow the trend much faster when we incorporate and enrich the market data features with COVID-19 data, visible on Figure 5 (b). Error values are also self explanatory in this comparison, where the model on Figure 5 (b) has MSE of 0.00053 and the model on Figure 5 (a) stops at 0.00197.

4.3 Multiple-step

As discussed previously in *Multi-Step LSTM* section, various multi-step models are tested in this study, including vector-output sequence prediction and encoder-decoder. According to our experiments and after many different runs due to the stochastic and random behaviour, we determined that the winner method is vector output sequence prediction with 22 days of history, predicting two days in advance. The MSE of the model on validation data has been recorded as 0.00080. Encoder-decoder approach also performed very good in comparison to the rest of the approaches. Using the encoder-decoder approach, the MSE of 0.00091 was hit on the validation set. We noticed that simple stacked LSTM networks would outperform other methods for 1 and 2 days of prediction in advance, as probably it requires a less complicated approach for this kind of prediction tasks and over-complicating the approach does not necessarily increase the performance.

As it can be seen in Table 3 COVID-19’s data has a noticeable effect on the results. Generally, the results of the experiments without COVID-19 as demonstrated in3 do not shown any improvement. Besides, to calculate the effects of uni-variate data-sets, we predict some of the experiments with only the Gold-close data-set and our winner architecture and the results didn’t show any improvement in the performance. As it can be seen in Table 3 the error rate with uni-variate data-set is increasing totally.

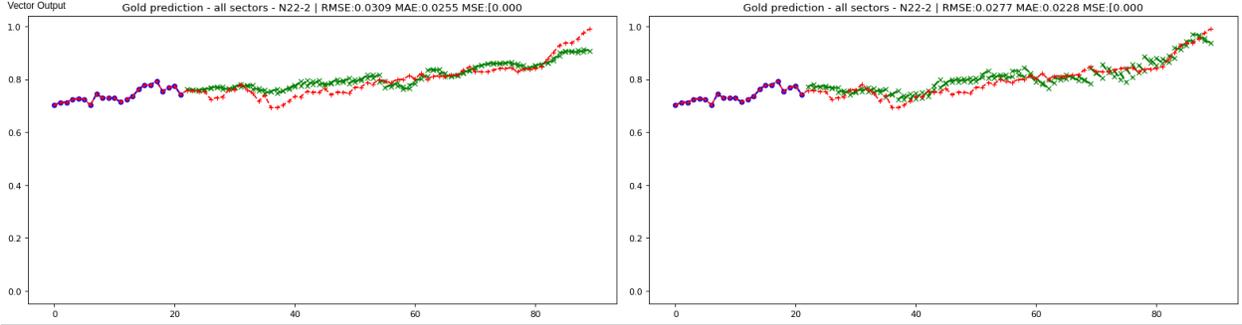


Figure 6: Multi-step time series prediction. Figure on the left (a) has been trained on the market data without COVID-19 time series. Figure on the right (b) has been trained on all features including COVID-19 time series data. Red markings are ground truth validation points and green marking are the predictions. The model makes prediction based on 22 days of historical data. The architecture is: LSTM(200)-LSTM(200)-Dropout-Dense(2).

Finally, 30 days prediction of stock market value sequence was carried out, that are presented in the Table 2. The history periods of 45 days and 60 days are compared as the best resulted experiments in this category. The MSE of the best model is 0.0021 which indicates the model has converged and reached a low error value. As illustrated, the vector output model has achieved the best results in comparison to encoder-decoder, bidirectional, and CNN-LSTM models. These results are in line with 2-days prediction models, hinting at the capability of vector output model, given correct architecture and hyper-parameters.

Table 2: MSE error rates for 30 days prediction in advance with some of the best architectures.

Type	History	Network	MSE
E-D	60	L500-L100-d100	0.0217
BD	45	BDL300-L100-d100	0.0024
CNN	45	C100-C200-pool-L100-d100	0.0061
SE	60	L250-L250-d30	0.0021

History column shows the number of days that predictions are based on.

L denotes LSTM layers

d denotes Dense layers

C denotes Convolution layers

BDL denotes bidirectional LSTM layers

pool denotes Max-pooling layers

E-D denotes encoder-decoder model

BD denotes bidirectional model

CNN denotes CNN-LSTM model

SE denotes vector output sequence prediction model

4.4 Comparison of the Proposed Method

Here we discuss and compare our proposed methods' results with two other proposed models. The first study proposes a time-series forecasting prediction for the price of gold, by a CNN-LSTM model [21] that using uni-variate data and single step prediction (single-day forecast) which is shown in Table 3 (marked by * on row 27). Changing the uni-variate to multi variate model using the same overall architecture improves the results (specifically MSE), as illustrated at the next row. Also, comparing these results with our best architectures (i.e. vector-output sequence) with the same features, we notice that regarding the prediction of Gold price with this much fluctuations, our models yield better results in terms of prediction error (lower MSE).

The second study [26] surveyed the prediction of COVID-19 in Canada Using LSTM Networks for 14 days in advance and used uni-variate dataset and Bidirectional LSTM (model '28, td14' in Table 3). It was noticed that the method could not be fitted to this problem efficiently and the errors are higher than other methods. For a more accurate comparison, we also predicted 14 days with Encoder-Decoder architecture and uni-variate dataset (model 'L200-r-L200-td100-td1'),

Table 3: Table of results

#	Steps	Variable	Method	History	Model	RMSE	MSE	MAE	MSLE	R2	ACC
01	1	Multi		5	L100-L100-d100	0.0251	0.00063	0.01774	0.00018	0.85759	0.97803
02		Multi		15	L100-L100-d100	0.0662	0.00439	0.04733	0.00128	0.02247	0.94534
03		Multi		22	L100-L100-d100	0.0290	0.00084	0.02168	0.00024	0.76198	0.97438
04		Multi		30	L100-L100-d100	0.0246	0.00060	0.02072	0.00017	0.79249	0.97592
05		Multi		9	L100-L100-drop-d100-d1	0.0781	0.00067	0.02066	0.00020	0.82069	0.94616
06		Multi		30	L300-L300-d100-d1	0.0900	0.00053	0.01870	0.00016	0.80063	0.91020
07		COVID less		30	L300-L300-d100-d1	0.0020	0.00197	0.03611	0.00059	0.50048	0.93532
08	2	Multi	SE	5	L200-L200-drop-d2	0.0439	0.00192	0.03294	0.00057	0.54020	0.95891
09		Multi		15	L200-L200-drop-d2	0.0426	0.00174	0.03738	0.00052	0.58592	0.95340
10		Multi		22	L200-L200-drop-d2	0.0311	0.00096	0.02539	0.00028	0.70735	0.97065
11		Multi		22	L200-L200-drop-d2	0.0283	0.00080	0.02334	0.00024	0.75998	0.97118
12		Uni		22	L200-L200-drop-d2	0.0983	0.00479	0.09456	0.00274	0.51678	0.87320
13		COVID less		22	L200-L200-drop-d2	0.0287	0.00082	0.02410	0.00024	0.75448	0.97081
14		Multi		30	L200-L200-drop-d2	0.0414	0.00171	0.03342	0.00048	0.36443	0.96104
15		Multi		30	L200-L300-drop-d2	0.0423	0.00178	0.02961	0.00050	0.34163	0.96603
16	2	Multi	E-D	5	L200-r-L200-td100-td1	0.0525	0.00275	0.04250	0.00082	0.32775	0.94633
17		Multi		15	L200-r-L200-td100-td1	0.0303	0.00091	0.02290	0.00027	0.78431	0.97180
18		COVID less		15	L200-r-L200-td100-td1	0.0317	0.00100	0.02456	0.00030	0.76589	0.96948
19		Uni		15	L200-r-L200-td100-td1	0.0745	0.00554	0.05769	0.00165	-0.30800	0.83155
20		Multi		15	L200-r-L200-drop-td100-td1	0.0386	0.00148	0.03276	0.00045	0.64927	0.95866
21		Multi		22	L200-r-L200-td100-td1	0.0424	0.00180	0.03723	0.00052	0.10212	0.96339
22		Multi		30	L100-r-L100-drop-td200-td1	0.0500	0.00250	0.03244	0.00070	0.42893	0.95437
23		Multi		30	L200-r-L200-drop-td100-td1	0.0348	0.00121	0.02520	0.00033	0.55437	0.97100
24	14	Multi		14	L200-r-L200-td100-td1	0.0938	0.00280	0.04374	0.00088	-0.10390	0.94558
25	14	Multi		15	L200-r-L200-td100-td1	0.0523	0.00270	0.04312	0.00084	-0.09150	0.94591
26	14	Uni		14	L200-r-L200-td100-td1	0.0513	0.00263	0.04029	0.00080	0.07561	0.95002
27	*1	Uni	CNN	9	C64-C128-Max128-L200-td32-td1	0.0762	0.00581	0.05752	0.00172	-0.31150	0.93195
28	1	Multi	CNN	9	C64-C128-Max128-r-L200-td32-td1	0.0768	0.00278	0.03868	0.00080	0.37183	0.93054
29	2	Uni	CNN	9	C64-C128-Max128-L200-td32-td1	0.0532	0.00283	0.03821	0.00083	0.32742	0.95390
30	2	Multi	CNN	9	C64-C128-Max128-L200-td32-td1	0.0663	0.00430	0.04930	0.00130	-0.03590	0.94067
31	**14	Uni	BD	14	BDL28, td1	0.46173	0.54144	0.4371	0.09122	-87.826	0.45193
32	14	Multi	BD	14	BDL28, td1	0.0337	0.25308	0.02388	0.00034	0.57346	0.96969
33	14	Multi	BD	14	BDL100,td1	0.0436	0.25430	0.03513	0.00057	0.33846	0.95614
34	14	Multi	BD	14	BDL100,td28,td1	0.0459	0.25460	0.03723	0.00064	0.26752	0.95343

The column "Step" indicates the number of steps ahead (days) predicted that are 1, 2, and 14 days in this table.

The column "Variable" shows different variables used in feature space (i.e Uni indicates the dataset which only includes Gold, Multi includes all the dataset, and COVID less includes all the financial, variables without COVID-19 data).

History column shows the number of days that predictions are based on.

L denotes LSTM layers

d denotes Dense layers

C denotes Convolution layers

BDL denotes bidirectional LSTM layers

pool denotes Max-pooling layers

td denotes time distributed dense

pool denotes Max-pooling layers

E-D denotes encoder-decoder model

BD denotes bidirectional model

CNN denotes CNN-ISTM model

SE denotes vector output sequence prediction model

* based on [21]

** based on [26]

and as the results reflect, our proposed model outperforms the approach. Besides, we further modified their model with changing the dataset from uni-variate to multi-variate ('BDL28, td1' on rows 31 and 32) and as it can be seen the results improved with the reducing the errors (considering all four errors). Changing the architecture, yet preserving the bidirectional LSTM approach improves the errors, but fails to outperform the rest of the approaches. We can see in 3 on rows 24 to 26 that using Encoder-Decoder approach improves the errors drastically. Hence, we can conclude that tuning and adapting bidirectional and CNN-LSTM approaches for this task would be much harder and even impossible. Finally, we can assert that the multivariate dataset affects the final results positively in almost every architecture, yielding lower errors.

5 Conclusion and Future Work

In this study, we developed a set of prediction models for financial markets based on various features (including COVID-19 and different analysis of market). We further investigated the effects of these features, specifically the COVID-19 pandemic, on the predictive models. The paper focuses on three aspects: (1) days to predict, (2) machine learning approaches, and (3) the feature space.

The results and experiments showed that although vanilla stacked LSTMs might seem simple, yet they are very powerful at prediction compared to Encoder-Decoder, Bidirectional, and CNN-LSTM approaches in time series prediction problems. In this study we achieve the MSE of $5e - 4$ for single-step and $8e - 4$ for multi-step forecasting for two days of prediction in advance as illustrated in Table 3.

We also showed the power other financial indicators (i.e. market data like oil, technology symbols, etc.) and non-financial parameters (e.g. COVID-19 new cases), in prediction of a specific market. Correlation analysis were utilized in this study to establish the connection among markets and the features in general. According to the results, we can assert that having a dataset of other pandemics and the variations of the economy in that situation in the history can boost the results. Although the models can achieve very low errors, it should also be acknowledged that markets rely on many variables like geopolitical decisions and escalations that can result in unpredictable movements. This is a limitation of machine learning at market forecasting, as designing an AI to comprehend every variable is not easily achievable. As a future work, understanding the semantics and context of social networks can also be considered as a potential feature in predictive models.

Conflict of interest

The authors declare that they have no conflict of interest.

Authors' contributions

Authors Mahtab Mohtasham Khani and Sahand Vahidnia have had equal contributions to this work.

References

- [1] Mary E Wilson. Travel and the emergence of infectious diseases. *Emerging infectious diseases*, 1(2):39, 1995.
- [2] Richard Baldwin. The greater trade collapse of 2020: Learnings from the 2008-09 great trade collapse. <https://voxeu.org/article/greater-trade-collapse-2020>, 2020.
- [3] Niels Joachim Gormsen and Ralph S. J. Koijen. Coronavirus: Impact on Stock Prices and Growth Expectations. *SSRN Electronic Journal*, pages 1–27, 2020.
- [4] DAVID SLOTNICK. Many of the world's airlines could be bankrupt by may because of the covid-19 crisis, according to an aviation consultancy. these airlines have already collapsed because of the pandemic. <https://www.businessinsider.com.au/coronavirus-airlines-that-failed-bankrupt-covid19-pandemic-2020-3?r=US&IR=T>, May 2020.
- [5] Suborna Barua. Covid-19 pandemic and world trade: Some analytical notes. *Available at SSRN 3577627*, 2020.
- [6] Mahtab Mohtasham Khani, Sahand Vahidnia, Leila Ghasemzadeh, Y Eren Ozturk, Mustafa Yuvalaklioglu, Selim Akin, and Nazim Kemal Ure. Deep-learning-based crack detection with applications for the structural health monitoring of gas turbines. *STRUCTURAL HEALTH MONITORING-AN INTERNATIONAL JOURNAL*, 2019.
- [7] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- [8] QN Chen, YW Liu, and FL Wang. A chronicle on the sars epidemic. *CHINESE LAW AND GOVERNMENT*, 36(4):12–15, 2003.
- [9] Raoul J de Groot, Susan C Baker, Ralph S Baric, Caroline S Brown, Christian Drosten, Luis Enjuanes, Ron AM Fouchier, Monica Galiano, Alexander E Gorbalenya, Ziad A Memish, et al. Commentary: Middle east respiratory syndrome coronavirus (mers-cov): announcement of the coronavirus study group. *Journal of virology*, 87(14):7790–7792, 2013.
- [10] Jolie Kaner and Sarah Schaack. Understanding ebola: the 2014 epidemic. *Globalization and health*, 12(1):1–7, 2016.
- [11] Centers for Disease Control, Prevention, et al. The 2009 h1n1 pandemic: summary highlights, april 2009-april 2010. *Official Online Article Published by the Centers for Disease Control*, 4, 2010.
- [12] Òscar Jordà, Sanjay R Singh, and Alan M Taylor. Longer-run economic consequences of pandemics. Technical report, National Bureau of Economic Research, 2020.
- [13] Christa Pettie. *Modeling Syndromic Surveillance and Outbreaks in Subpopulations*. PhD thesis, 2020.

- [14] R Shouval, O Bondi, H Mishan, A Shimoni, R Unger, and Arnon Nagler. Application of machine learning algorithms for clinical predictive modeling: a data-mining approach in sct. *Bone marrow transplantation*, 49(3):332–337, 2014.
- [15] Veronica Guerrieri, Guido Lorenzoni, Ludwig Straub, and Iván Werning. Macroeconomic implications of covid-19: Can negative supply shocks cause demand shortages? Technical report, National Bureau of Economic Research, 2020.
- [16] Richard Baldwin and Beatrice Weder di Mauro. *Economics in the Time of COVID-19*. 2020.
- [17] Fernando E Alvarez, David Argente, and Francesco Lippi. A simple planning problem for covid-19 lockdown. Working Paper 26981, National Bureau of Economic Research, April 2020.
- [18] Mei Li, Xiang Wang, Kai Gao, and Shanshan Zhang. A survey on information diffusion in online social networks: Models and methods. *Information (Switzerland)*, 8(4), 2017.
- [19] Eiji Aramaki, Sachiko Maskawa, and Mizuki Morita. Twitter catches the flu: detecting influenza epidemics using twitter. In *Proceedings of the conference on empirical methods in natural language processing*, pages 1568–1576. Association for Computational Linguistics, 2011.
- [20] Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018.
- [21] Ioannis E Livieris, Emmanuel Pintelas, and Panagiotis Pintelas. A cnn-lstm model for gold price time-series forecasting. *Neural Computing and Applications*, pages 1–10, 2020.
- [22] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, 33(4):917–963, 2019.
- [23] Felix A Gers, Douglas Eck, and Jürgen Schmidhuber. Applying lstm to time series predictable through time-window approaches. In *Neural Nets WIRN Vietri-01*, pages 193–200. Springer, 2002.
- [24] K. Chen, Y. Zhou, and F. Dai. A lstm-based method for stock returns prediction: A case study of china stock market. In *2015 IEEE International Conference on Big Data (Big Data)*, pages 2823–2824, 2015.
- [25] Luca Di Persio and Oleksandr Honchar. Artificial neural networks approach to the forecast of stock market price movements. *International Journal of Economics and Management Systems*, 1, 2016.
- [26] Vinay Kumar Reddy Chimmula and Lei Zhang. Time series forecasting of covid-19 transmission in canada using lstm networks. *Chaos, Solitons & Fractals*, page 109864, 2020.
- [27] Anuradha Tomar and Neeraj Gupta. Prediction for the spread of covid-19 in india and effectiveness of preventive measures. *Science of The Total Environment*, page 138762, 2020.
- [28] Erik Marchi, Fabio Vesperini, Felix Weninger, Florian Eyben, Stefano Squartini, and Björn Schuller. Non-linear prediction with lstm recurrent neural networks for acoustic novelty detection. In *2015 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2015.
- [29] Xuan Hien Le, Hung Viet Ho, Giha Lee, and Sungho Jung. Application of Long Short-Term Memory (LSTM) neural network for flood forecasting. *Water (Switzerland)*, 11(7), 2019.
- [30] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.
- [31] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.
- [32] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3156–3164, 2015.
- [33] Jeffrey Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, and Trevor Darrell. Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2625–2634, 2015.
- [34] Mike Schuster and Kuldip K Paliwal. Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11):2673–2681, 1997.
- [35] Alex Graves and Jürgen Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, 18(5-6):602–610, 2005.
- [36] etfdb. The 11 sectors of the stock market. <https://etfdb.com/etf-education/what-is-an-in-kind-redemption-for-etfs/>, 2020.

- [37] Sahand Vahidnia, Ömer Özgür Tanrıöver, and I Askerzade. An early phase software project risk assessment support method for emergent software organizations. *Int J Advan Comp Sci Appl*, 8(5):105–118, 2017.
- [38] François Chollet et al. Keras. <https://keras.io>, 2015.
- [39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.