Deep Learning based Crude Oil Price Forecasting Model for Non-Linear Dynamics Modeling

[1]Ramesh Mocharla, [2]L Prasanna Lakshmi, [3]Lenkapally Mahesh
Assistance Professor, Department Of CSE, Sree dattha institute of engineering and science

ABSTRACT

With the popularity of the deep learning model in the engineering fields, it has attracted significant research interests in the economic and finance fields. In this project, we use the deep learning model to capture the unknown complex nonlinear characteristics of the crude oil price movement. We further propose a new hybrid crude oil price forecasting model based on the deep learning model. Using the proposed model, major crude oil price movement is analyzed and modeled. The performance of the proposed model is evaluated using the price data in the WTI crude oil markets. The empirical results show that the proposed model achieves the improved forecasting accuracy.

Keywords: Crude oil price, deep learning, Non-linear dynamics.

1. INTRODUCTION

The crude oil price movements are subject to diverse influencing factors. The dynamic complicated interactions among these factors result in the mysterious behavior of the crude oil price movement, whose characterization and prediction remained one of the most interesting and intriguing research issues in the economic and financial analysis field. Recently numerous empirical studies have revealed the nonlinear nature of economic and financial data, where traditional methods such as linear prediction methods are not able to analyze the complex nonlinear dynamics involved [1]. [2] used the Qual VAR model to model the nonlinear autocorrelation characteristics of WTI crude oil price changes and forecast its future movement. They found that the Qual VAR model outperforms the benchmark Random Walk and VAR model. [3] proposed a Hidden Markov Model (HMM) with threshold effect to model hidden factors influencing the crude oil price movement. They demonstrated that the proposed models outperform the ARMA model, in h-day ahead forecasting exercise. On the other hand, the Artificial Intelligence (AI) and Machine Learning (ML) based approaches received more and more research interests. [4] used the Recurrent Neural Network (RNN) to forecast the crude oil indices. [5] proposed a Genetic Algorithm optimized Neural Network model to forecast the crude oil price fluctuations. They showed this evolutionary neural network model brings statistically significant performance improvement. [6] provided a comprehensive survey on the AI and ML based crude oil forecasting models.

The deep learning model is a new artificial intelligence paradigm developed beyond the neural network. It has become a popular phenomenon in the computer science and engineering fields such as image recognition, text classification and speech recognition recently. For example, [7] proposed HMM-DNN model for speech synthesis method. In addition to the HMM-DNN model, the CNN model and the RNN model are also applied to the construction of speech recognition models. However, the deep learning model has only been introduced to the financial field quite recently. There are some very early exploratory attempts. For example, [8] proposed a deep learning hierarchical decision model and used it to construct new portfolio with the desired level of performance. Empirical studies show that the proposed model achieves the superior performance. [9] predicted the real estate price by combining the Boltzmann machine with the genetic algorithm. [10] applied the convolution neural network to the biological aspects of DNA-protein transcription prediction. [11] used the deep belief network model to predict the short-term power load in Macedonia from 2008 to 2014. The empirical results showed that the deep belief network model has obvious advantages compared with the traditional forecasting models.

In this project, we propose a new Deep learning-based hybrid crude oil price forecasting methodology to model the nonlinear dynamics involved in the crude oil price movement and forecast its future movement at higher level of accuracy. Considering both the linear and non-linear characteristics of historical data, we integrate the prediction results of the ARMA model and the prediction results of the deep learning model. Empirical studies have been conducted using the major crude oil prices to evaluate the performance of the proposed model. The superior performance of the crude oil price forecasting model using the deep belief network and recurrent neural network provide the empirical evidence that the market is inefficient in the regional and submarkets. Work in this paper shows methodologically the merit of the deep learning model in tracking and capturing the nonlinearity and dynamic crude oil price movement. When analytic solutions are lacking, it would provide the best approximation to the nonlinear dynamics in the crude oil price movement. It contributes to the literature by providing a new methodology on how the deep learning model can be used to improve the crude oil forecasting accuracy.

2. LITERATURE SURVEY

Monthly Energy Review; U.S. Energy Information Administration. Available online: https://www.eia.gov/ totalenergy/data/monthly/index.php (accessed on 10 February 2020).

Note 1. Merchandise Trade Value. Imports data presented are based on the customs values. Those values do not include insurance and freight and are consequently lower than the cost, insurance, and freight (CIF) values, which are also reported by the Bureau of the Census. All exports data, and imports data through 1980, are on a free alongside ship (f.a.s.) basis. “Balance” is exports minus imports; a positive balance indicates a surplus trade value and a negative balance indicates a deficit trade value. “Energy” includes mineral fuels, lubricants, and related material. “Non-Energy Balance” and “Total Merchandise” include foreign exports (i.e., re-exports) and nonmonetary gold and U.S. Department of Defense Grant- Aid shipments. The “Non-Energy Balance” is calculated by subtracting the “Energy” from the “Total Merchandise Balance.” “Imports” consist of government and nongovernment shipments of merchandise into the 50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and the U.S. Foreign Trade Zones. They reflect the total arrival from foreign countries of merchandise that immediately entered consumption channels, warehouses, the Foreign Trade Zones, or the Strategic Petroleum Reserve. They exclude shipments between the United States, Puerto Rico, and U.S. possessions, shipments to U.S. Armed Forces and diplomatic missions abroad for their own use, U.S. goods returned to the United States by its Armed Forces, and in-transit shipments. Note 2. Non-Combustion Use of Fossil Fuels. Most fossil fuels consumed in the United States and elsewhere are combusted to produce heat and power. However, some are used directly for non-combustion use as construction materials, chemical feedstocks, lubricants, solvents, and waxes. For example, coal tars from coal coke manufacturing are used as feedstock in the chemical industry, for metallurgical work, and in anti-dandruff shampoos; natural gas is used to make nitrogenous fertilizers and as chemical feedstocks; asphalt and road oil are used for roofing and paving; hydrocarbon gas liquids are used to create intermediate products that are used in making plastics; lubricants, including motor oil and greases, are used in vehicles and various industrial processes; petrochemical feedstocks are used to make plastics, synthetic fabrics, and related products.

[2] Saggu, A.; Anukoonwattaka, W. Commodity Price Crash: Risks to Exports and Economic Growth in Asia-Pacific LDCs and LLDCs. United Nations ESCAP Trade Insights 2015, 6, 2617542.

This issue of the Trade Insights series identifies Asia-Pacific LDCs and LLDCs with export-portfolios and economies which are at greatest risk from the recent collapse in global commodity prices. Asia-Pacific LDCs and LLDCs account for less than 2% of global commodity exports and just 7% of Asia-Pacific commodity exports; however many these economies have export-portfolios which are highly concentrated in one or two major commodities: mainly crude oil, natural gas, aluminum, iron ore/steel, cotton and copper. This note finds that economic growth is at significant risk from changes in commodity prices across many Asia-Pacific LDCs and LLDCs, particularly in fuel-exporting economies and metal and mineral exporting economies.

[3] Chatzis, P.S.; Siakoulis, V.; Petropoulos, A.; Stavroulakis, E.; Vlachogiannakis, N. Forecasting stock market crisis events using deep and statistical machine learning techniques. Expert Syst. Appl. 2018, 112, 353–371.

Predictions of stock and foreign exchange (Forex) have always been a hot and profitable area of study. Deep learning applications have been proven to yield better accuracy and return in the field of financial prediction and forecasting. In this survey, we selected papers from the Digital Bibliography & Library Project (DBLP) database for comparison and analysis. We classified papers according to different deep learning methods, which included Convolutional neural network (CNN); Long Short-Term Memory (LSTM); Deep neural network (DNN); Recurrent Neural Network (RNN); Reinforcement Learning; and other deep learning methods such as Hybrid Attention Networks (HAN), self-paced learning mechanism (NLP), and Wavenet. Furthermore, this paper reviews the dataset, variable, model, and results of each article. The survey used presents the results through the most used performance metrics: Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Square Error (MSE), accuracy, Sharpe ratio, and return rate. We identified that recent models combining LSTM with other methods, for example, DNN, are widely researched. Reinforcement learning and other deep learning methods yielded great returns and performances. We conclude that, in recent years, the trend of using deep-learning-based methods for financial modeling is rising exponentially.

[4] Koch, N.; Fuss, S.; Grosjean, G.; Edenhofer, O. Causes of the EU ETS price drop: Recession, CDM, renewable policies or a bit of everything?–New evidence. Energy Policy 2014, 73, 676–685

This paper examines the usefulness of asset prices in predicting the beginnings of recessions in the G-7 countries. It finds that equity/house price drops have a substantial marginal effect on the likelihood of a new recession. Increased market uncertainty, which is a second-moment variable associated with equity price changes, is also a useful predictor of new recessions in these countries. These findings are robust to the inclusion of the term spread and oil prices. The new recession forecasting performance of our baseline model is superior to that of a similar model estimated over all recession and expansion periods, highlighting a difference between the probabilities of a new recession versus a continuing recession.

[5] Baumeister, C.; Kilian, L. Forecasting the real price of oil in a changing world: A forecast combination approach. J. Bus. Econ. Stat. 2015, 33, 338–351.

The U.S. Energy Information Administration (EIA) regularly publishes monthly and quarterly forecasts of the price of crude oil for horizons up to 2 years, which are widely used by practitioners. Traditionally, such out-of-sample forecasts have been largely judgmental, making them difficult to replicate and justify. An alternative is the use of real-time econometric oil price forecasting models. We investigate the merits of constructing combinations of six such models. Forecast combinations have received little attention in the oil price forecasting literature to date. We demonstrate that over the last 20 years suitably constructed real-time forecast combinations would have been systematically more accurate than the no-change forecast at horizons up to 6 quarters or 18 months. The MSPE reductions may be as high as 12% and directional accuracy as high as 72%. The gains in accuracy are robust over time. In contrast, the EIA oil price forecasts not only tend to be less accurate than no-change forecasts, but are much less accurate than our preferred forecast combination. Moreover, including EIA forecasts in the forecast combination systematically lowers the accuracy of the combination forecast. We conclude that suitably constructed forecast combinations should replace traditional judgmental forecasts of the price of oil.

3. PROPOSED SYSTEM

Long short-term memory (LSTM) RNN in TensorFlow

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. It was proposed in 1997 by **Sepp Hochreiter** and **Jurgen schmidhuber**. Unlike standard feed-forward neural networks, LSTM has feedback connections. It can process not only single data points (such as images) but also entire sequences of data (such as speech or video).

For example, LSTM is an application to tasks such as unsegmented, **connected handwriting recognition,** or **speech recognition**.

A general **LSTM** unit is composed of a cell, an input gate, an output gate, and a forget gate. The cell remembers values over arbitrary time intervals, and three gates regulate the flow of information into and out of the cell. LSTM is well-suited to classify, process, and predict the time series given of unknown duration.

Long Short- Term Memory (LSTM) networks are a modified version of recurrent neural networks, which makes it easier to remember past data in memory.

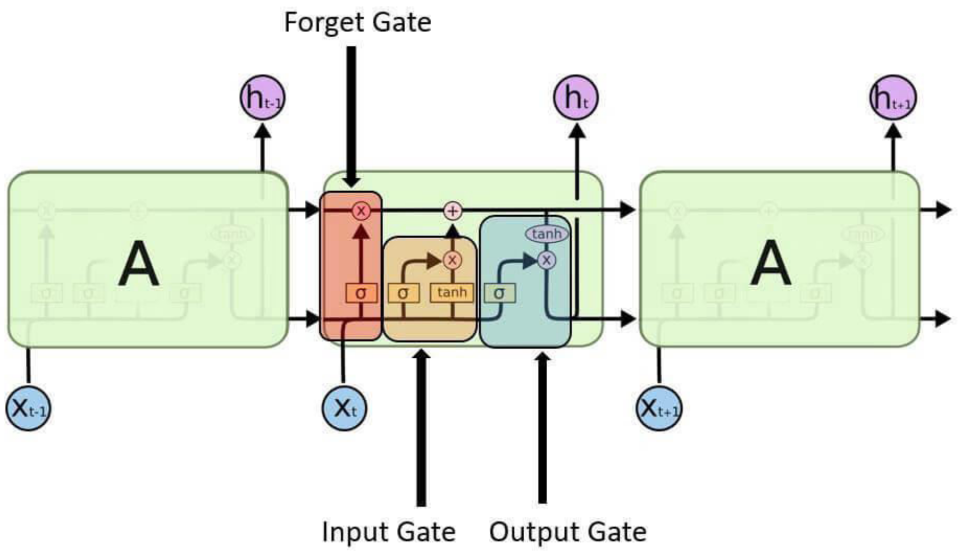


Fig. 1: LSTM cell.

1. Input gate- It discover which value from input should be used to modify the memory. **Sigmoid** function decides which values to let through 0 or 1. And **tanh** function gives weightage to the values which are passed, deciding their level of importance ranging from **-1** to **1**.

$$i_t = \sigma(W_i.[h_t - 1, x_t] + bi)$$
$$C_t = \tanh\,(W_C.[h_t - 1, x_t] + b_C)$$

2. Forget gate- It discover the details to be discarded from the block. A sigmoid function decides it. It looks at the previous state **(ht-1)** and the content input (Xt) and outputs a number between 0(omit this) and 1(keep this) for each number in the cell state **Ct-1**.

$$f_t = \sigma(W_f.[h_t - 1, x_t] + b_f)$$

3. Output gate- The input and the memory of the block are used to decide the output. Sigmoid function decides which values to let through 0 or 1. And **tanh** function decides which values to let through 0, 1. And tanh function gives weightage to the values which are passed, deciding their level of importance ranging from **-1 to 1** and multiplied with an output of **sigmoid**.

$$O_t = \sigma(W_o[h_t - 1, x_t] + b_o$$
$$h_t = o_t * \tanh(C_t)$$

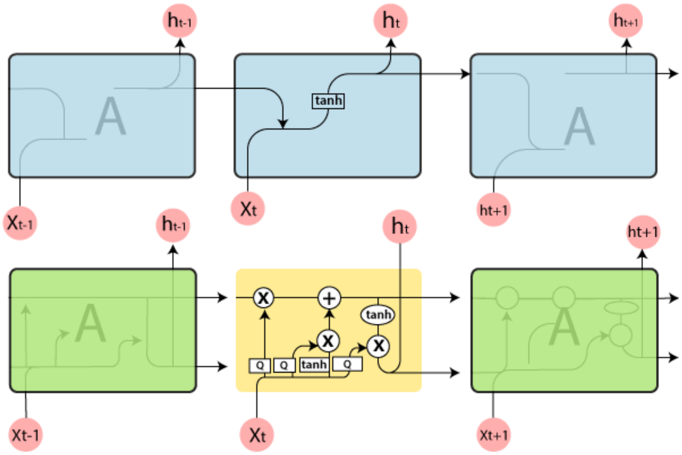


Fig. 2: LSTM forward cells.

It represents a full RNN cell that takes the current input of the sequence xi, and outputs the current hidden state, hi, passing this to the next RNN cell for our input sequence. The inside of an LSTM cell is a lot more complicated than a traditional RNN cell, while the conventional RNN cell has a single "internal layer" acting on the current state (ht-1) and input (xt).

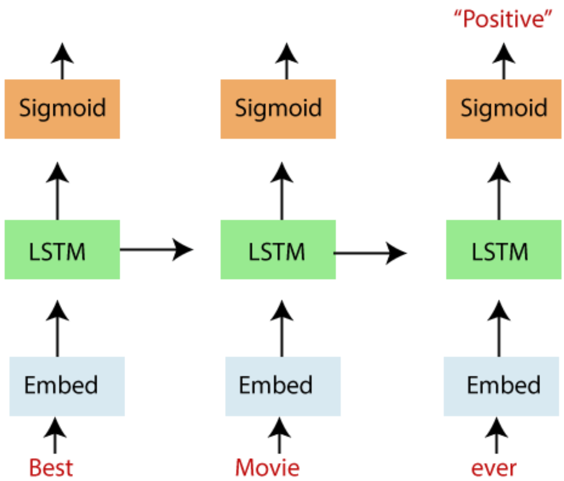


Fig. 3: Output generation using LSTM.

In the above diagram, we see an "unrolled" LSTM network with an embedding layer, a subsequent **LSTM** layer, and a sigmoid activation function. We recognize that our inputs, in this case, words in a movie review, are input sequentially.

The words are inputted into an embedding lookup. In most cases, when working with a corpus of text data, the size of the vocabulary is unusually large.

This is a multidimensional, distributed representation of words in a vector space. These embeddings can be learned using other deep learning techniques like **word2vec**, we can train the model in an end-to-end fashion to determine the embedding as we teach.

ISSN PRINT 2319 1775 Online 2320 7876
Research Paper , UGC CARE Listed (Group -I) Journal Volume 11, Iss 13, 2022

These embeddings are then inputted into our **LSTM layer**, where the output is fed to a sigmoid output layer and the **LSTM cell** for the next word in our sequence.

Algorithm:

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data (such as speech or video). For example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition, speech recognition and anomaly detection in network traffic or IDSs (intrusion detection systems).

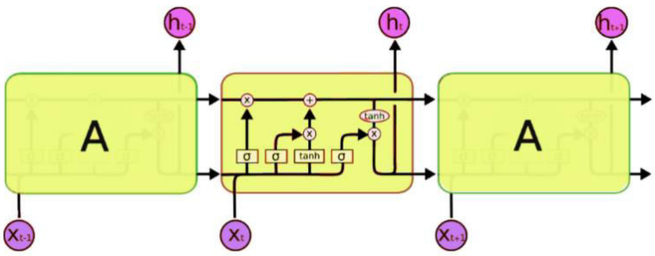
A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.

LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. LSTMs were developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs. Relative insensitivity to gap length is an advantage of LSTM over RNNs, hidden Markov models and other sequence learning methods in numerous applications

Training:

An RNN using LSTM units can be trained in a supervised fashion, on a set of training sequences, using an optimization algorithm, like gradient descent, combined with backpropagation through time to compute the gradients needed during the optimization process, in order to change each weight of the LSTM network in proportion to the derivative of the error (at the output layer of the LSTM network) with respect to corresponding weight.

A problem with using gradient descent for standard RNNs is that error gradients vanish exponentially quickly with the size of the time lag between important events. However, with LSTM units, when error values are back-propagated from the output layer, the error remains in the LSTM unit's cell. This "error carousel" continuously feeds error back to each of the LSTM unit's gates, until they learn to cut off the value.



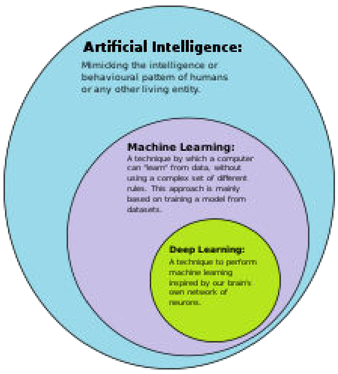
Deep Neural Networks

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.

Deep-learning architectures such as deep neural networks, deep belief networks, deep reinforcement learning, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.

Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems. ANNs have various differences from biological brains. Specifically, artificial neural networks tend to be static and symbolic, while the biological brain of most living organisms is dynamic (plastic) and analogue.

The adjective "deep" in deep learning refers to the use of multiple layers in the network. Early work showed that a linear perceptron cannot be a universal classifier, but that a network with a nonpolynomial activation function with one hidden layer of unbounded width can. Deep learning is a modern variation which is concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically informed connectionist models, for the sake of efficiency, trainability and understandability, whence the "structured" part.



Deep learning creates many layers of neurons, attempting to learn structured representation of big data, layer by layer.

Architecture of the network

Network models

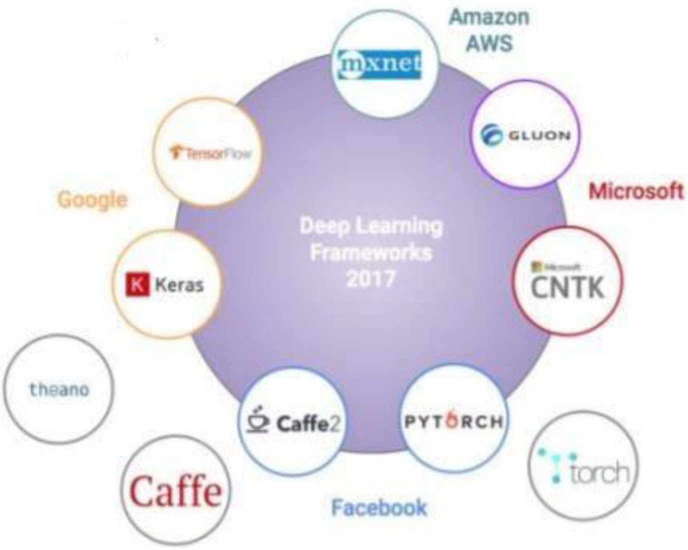
- Deep neural networks are mathematical models of intelligence designed to mimic human brains.
- Network models define a set of network layers and how they interact.
- Questions to answer while designing a network models include: –

Which layer type to use? – How many neurons to use in each layer? – How are layers arranged? – And more

- There are many standard CNN models available today which work great for many standard problems. Examples being AlexNet, GoogleNet, Inception-ResNet, VGG, etc

Deep learning frameworks

- Building a deep learning solution is a big challenge because of its complexity.
- Frameworks are tools to ease the building of deep learning solutions.
- Frameworks offer a higher level of abstraction and simplify potentially difficult programming tasks.

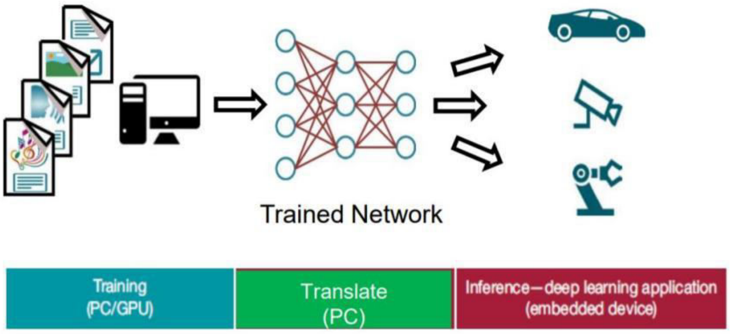


Popular Frameworks

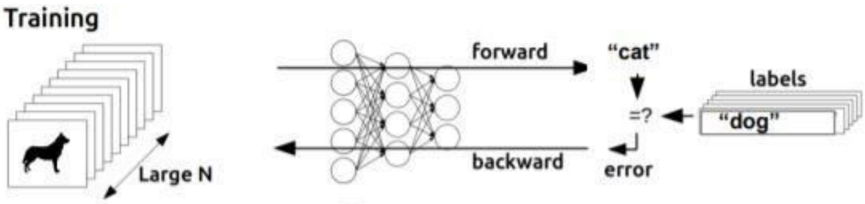
- TensorFlow: – Developed by Google – The most used deep learning framework – Based on Github stars and forks and Stack Overflow activity
- Caffe: – Developed by Berkeley Vision and Learning Center (BVLC) – Popular for CNN modeling (imaging/computer vision applications) and its Model Zoo (a selection of pre-trained networks) Next to all these frameworks, there are also interfaces that are wrapped around one or multiple frameworks. The most wellknown and widely-used interface for deep learning today is Keras. Keras is a high-level deep learning API, written in Python.

Deep learning development flow

1. Selection of a framework for development
2. Selecting labeled data set of classes to train the network upon
3. Designing initial network model
4. Training the network
5. Saving the parameters and architecture in a binary file
6. Inference



Training data:



Introduction to deep learning summary

- What is deep learning? Artificial intelligence, or AI, is an umbrella term for any computer program that does something smart. Machine learning is a subset of AI and Deep Learning is subset of Machine Learning.
- Deep learning has its roots in neural networks.
- Neural networks are sets of algorithms, modeled loosely after the human brain, that are designed to recognize patterns.
- Speaking deep learning: Network types, nodes, layers, development frameworks and network models.
- Deep learning solution development flow
- Application spaces

CNN

Convolution leverages three important ideas that can help improve a machine learning system: sparse interactions, parameter sharing and equivalent representations. Moreover, convolution provides a means for working with inputs of variable size. Traditional neural network layers use matrix multiplication by a matrix of parameters with a separate parameter describing the interaction between each input unit and each output unit. This means every output unit interacts with every input unit. Convolutional networks, however, typically have sparse interactions This is accomplished by making the kernel smaller than the input. For example, when processing an image, the input image might have thousands or millions of pixels, but we can detect small, meaningful features such as edges with kernels that occupy only tens or hundreds of pixels.

Store fewer parameters, which both reduces the memory requirements of the model and improves its statistical efficiency. It also means that computing the output requires fewer operations. These improvements in efficiency are usually quite large For many practical applications, it is possible to obtain good performance on the machine learning task while keeping several orders of magnitude less data. This allows the network to efficiently describe complicated interactions between many variables by constructing such interactions from simple building blocks that each describe only sparse interactions.

The Future of Deep Learning

The performance of deep learning systems can often be dramatically improved by simply scaling them up. With a lot more data and a lot more computation, they generally work a lot better. The language model GPT-318 with 175 billion parameters (which is still tiny compared with the number of synapses in the human brain) generates noticeably better text than GPT-2 with only 1.5 billion parameters. The chatbots Meena and BlenderBot also keep improving as they get bigger. Enormous effort is now going into scaling up and it will improve existing systems a lot, but there are fundamental deficiencies of current deep learning that cannot be overcome by scaling alone, as discussed here.

Comparing human learning abilities with current AI suggests several directions for improvement:

1. Supervised learning requires too much labeled data and model-free reinforcement learning requires far too many trials. Humans seem to be able to generalize well with far less experience.
2. Current systems are not as robust to changes in distribution as humans, who can quickly adapt to such changes with very few examples.
3. Current deep learning is most successful at perception tasks and generally what are called system 1 tasks. Using deep learning for system 2 tasks that require a deliberate sequence of steps is an exciting area that is still in its infancy. What needs to be improved. From the early days, theoreticians of machine learning have focused on the id assumption, which states that the test cases are expected to come from the same distribution as the training examples. Unfortunately, this is not a realistic assumption in the real world: just consider the non-stationarities due to actions of various agents changing the world, or the gradually expanding mental horizon of a learning agent which always has more to learn and discover. As a practical consequence, the performance of today's best AI systems tends to take a hit when they go from the lab to the field.

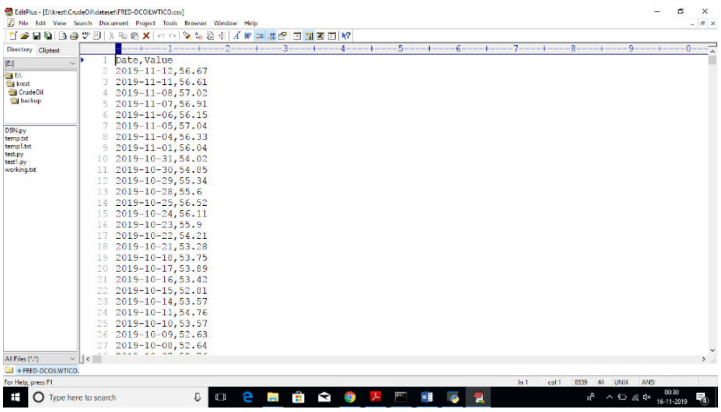
Our desire to achieve greater robustness when confronted with changes in distribution (called out-of-distribution generalization) is a special case of the more general objective of reducing sample complexity (the number of examples needed to generalize well) when faced with a new task—as in transfer learning and lifelong learning or simply with a change in distribution or in the relationship between states of the world and rewards. Current supervised learning systems require many more examples than humans (when having to learn a new task) and the situation is even worse for model-free reinforcement learning since each rewarded trial provides less information about the task than each labeled example. It has already been noted that humans can generalize in a way that is different and more powerful than ordinary id generalization: we can correctly interpret novel combinations of existing concepts, even if those combinations are extremely unlikely under our training distribution, so long as they respect high-level syntactic and semantic patterns we have already learned. Recent studies help us clarify how different neural net architectures fare in terms of this systematic generalization ability.

4. RESULTS AND DISCUSSIONS

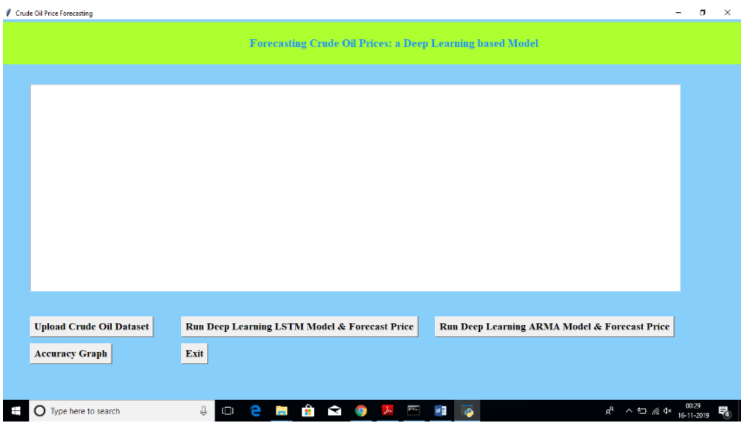
This project employs 2 deep learning algorithms such as LSTM (Long Short-Term Memory Network), and ARMA to forecast crude oil prices. LSTM algorithm is a deep learning famous algorithms used to train and predict any kind of data such as voice recognition, image classification or data classification, this algorithm consist of three layers called input, output, forget layer and at input layer algorithm will read features from dataset and find out best features and saved that best features in output layer and this process continue to filter features between input and output layer and if best output found then assign it current output and old output will be assigned to forgot layer. Final filtered features will be used to train model. While prediction algorithm accept test data and then apply that test data on train model to predict accurate class to which this test data belongs.

ARMA algorithm will also use for same training and testing model but its accuracy of prediction is less compare to LSTM.

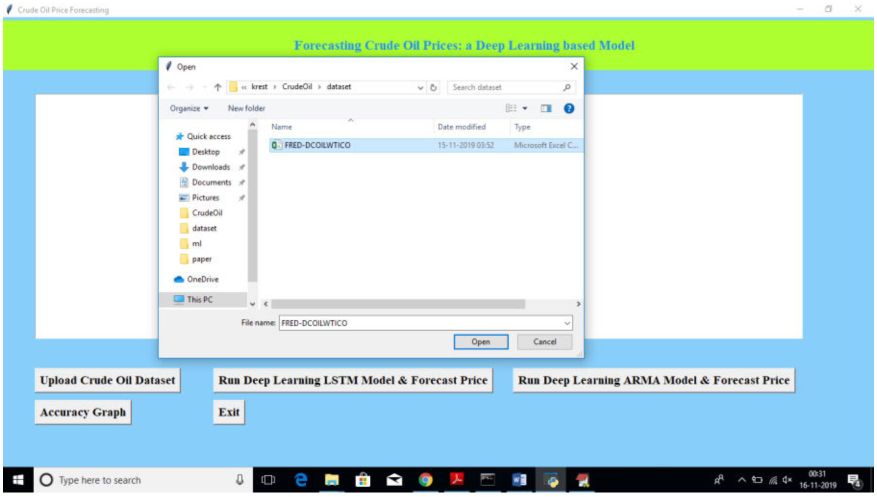
In this project we are using historical crude oil prices from QUANDL website as dataset and this dataset saved inside 'dataset' folder. This dataset contains two values such as DATE and price value. See below screen of dataset values

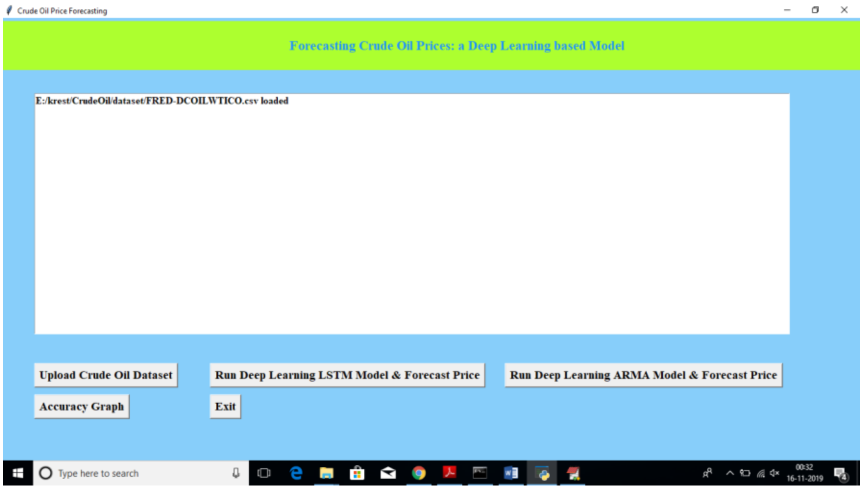


Screen shots

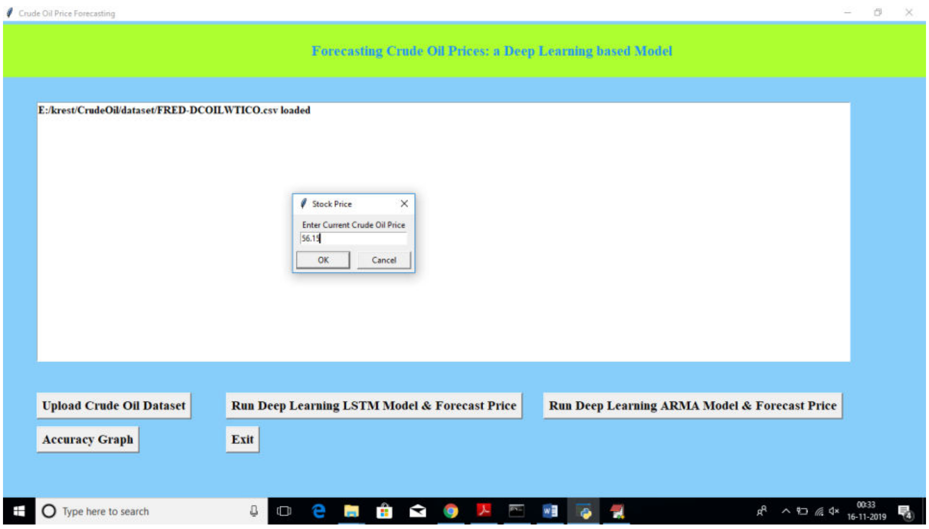


In above screen click on 'Upload Crude Oil Dataset' button to upload dataset

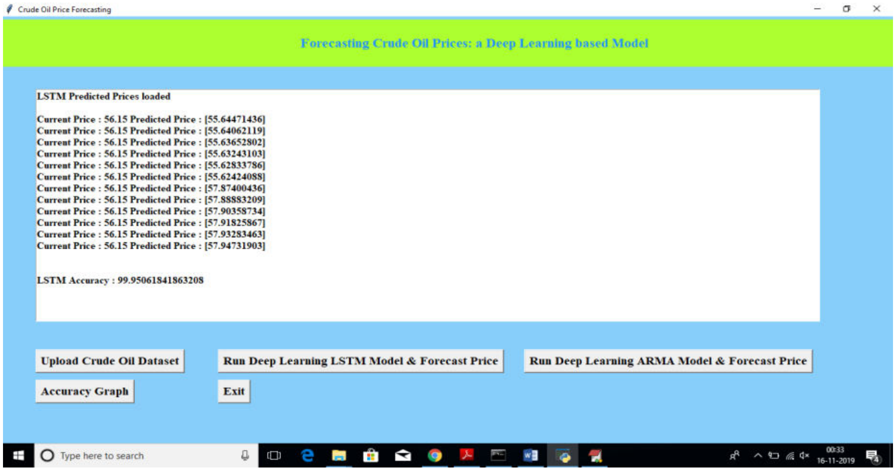




After uploading dataset click on 'Run Deep Learning LSTM Model & Forecast Price' button to allow LSTM to read dataset and then generate model. Here we need to enter current oil price in next dialog box to predict future prices

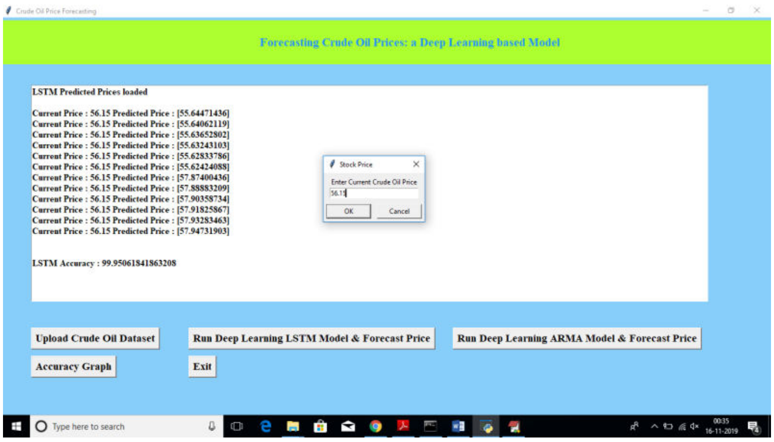


In above screen I entered current price as '56.15' and then we will get 10 new predicted prices

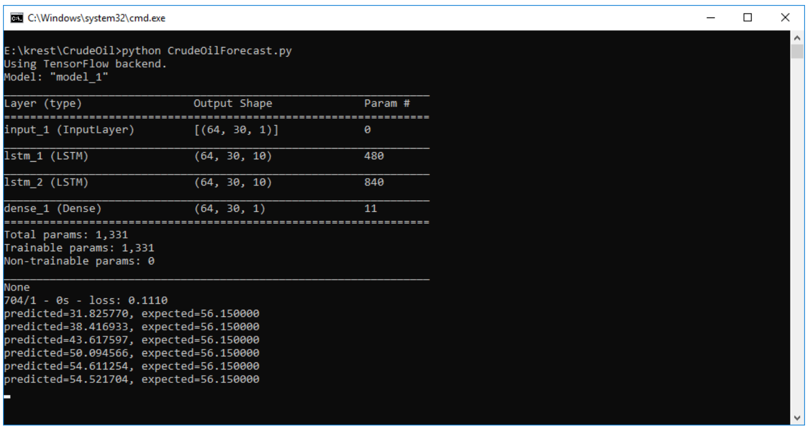


In above screen we can see current price and new best possible forecast prices. We will get LSTM prediction accuracy also.

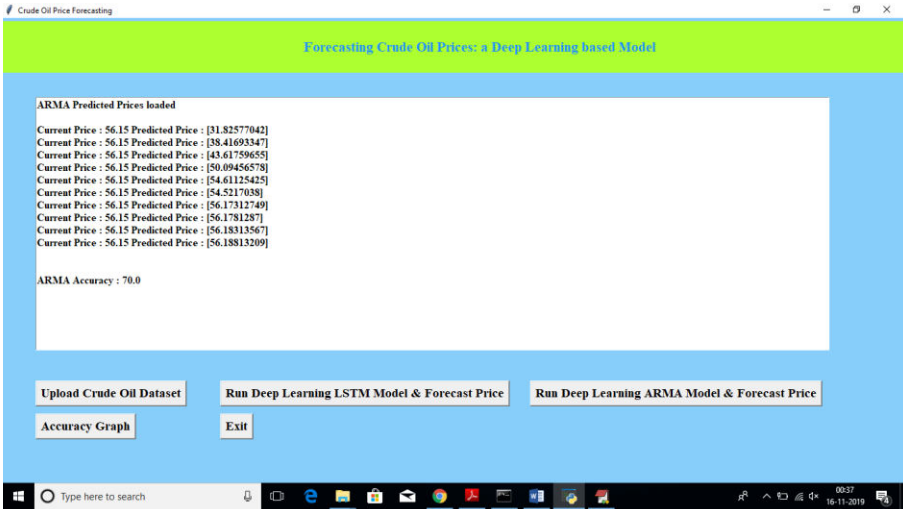
Now click on 'Run Deep Learning RAM Model & Forecast Price' button to allow application to read dataset and to generate model. Here also we need to enter current price to get future prices.



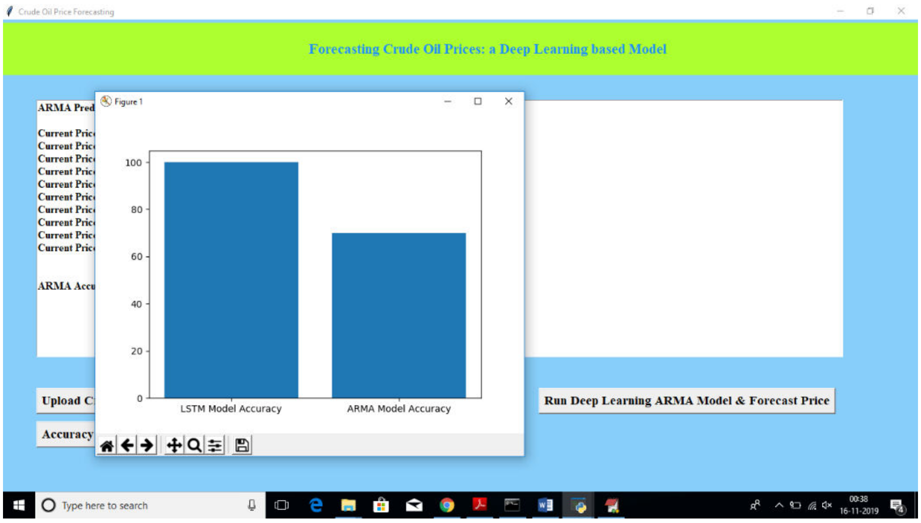
In above screen for ARMA algorithm also I gave current price as 56.15 and now below are the predicted prices. In below black console also we can see algorithms process details



Below are predicted prices with ARMA



In above screen we can see predicted prices from ARMA and we can also see its prediction accuracy. Now click on “Accuracy Graph’ button to to see accuracy in graph



In above graph x-axis represents algorithm name and y-axis represents accuracy. From above graph we can say LSTM perform better than ARMA

5. CONCLUSION AND FUTURE ENHANCEMENT

In this paper, we apply the emerging deep learning model to the crude oil price forecast. More specifically we identified two particular deep learning models, i.e. the deep belief network and the recurrent neural network, to be useful in modeling the nonlinear dynamics in the crude oil price movement. We construct a hybrid model that combines the forecasts from ARMA model as well as the forecasts from the deep learning models. We have conducted the empirical studies using the representative WTI crude oil market. We found the introduction of Deep Learning model in the crude oil price models lead to the improved forecasting accuracy. Work in this paper implies that there is exploitable forecasting opportunity in the crude oil price movement. More accurate modeling of the nonlinear dynamics in the crude oil price movement is critical to the further understanding of the determinant underlying the crude oil price movement. In the meantime, we found that the performance of the deep learning model is very sensitive to the parameters. Increasing model complexity with more hidden layers and hidden neurons may not necessarily lead to higher level of nonlinear modeling accuracy. This performance constraint may be attributed to the limited types of deep learning model attempted. It merits further research in constructing some innovative forecasting models based on different types of deep learning models.

REFERENCES

[1] F. Shen, J. Chao, J. Zhao, Forecasting exchange rate using deep belief networks and conjugate gradient method, Neurocomputing 167 (2015) 243 – 253.

[2] R. Gupta, M. Wohar, Forecasting oil and stock returns with a qual var using over 150 years off data, Energy Economics 62 (2017) 181–186.

[3] D. M. Zhu, W. K. Ching, R. J. Elliott, T. K. Siu, L. M. Zhang, Hidden markov models with threshold effects and their applications to oil price forecasting, Journal of Industrial and Management Optimization 13 (2) (2017) 757–773.

[4] J. Wang, J. Wang, Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations, Energy 102 (2016) 365–374.

[5] H. Chiroma, S. Abdulkareem, T. Herawan, Evolutionary neural network model for west texas intermediate crude oil price prediction, Applied Energy 142 (2015) 266 – 273.

[6] H. Chiroma, S. Abdul-kareem, A. S. M. Noor, A. I. Abubakar, N. S. Safa, L. Shuib, M. F. Hamza, A. Y. Gital, T. Herawan, A review on artificial intelligence methodologies for the forecasting of crude oil price, Intelligent Automation and Soft Computing 22 (3) (2016) 449–462.

[7] Z. H. Ling, L. Deng, D. Yu, Modeling spectral envelopes using restricted boltzmann machines and deep belief networks for statistical parametric speech synthesis, IEEE Transactions on Audio, Speech, and Language Processing 21 (10) (2013) 2129–2139.

[8] J. B. Heaton, N. G. Polson, J. H. Witte, Deep learning for finance: deep portfolios, Applied Stochastic Models in Business and Industry 33 (1) (2017) 3–12.

[9] M. H. Rafiei, H. Adeli, A novel machine learning model for estimation of sale prices of real estate units, Journal of Construction Engineering and Management 142 (2) (2016) 04015066.

[10] H. Y. Zeng, M. D. Edwards, G. Liu, D. K. Gifford, Convolutional neural network architectures for predicting dna-protein binding, Bioinformatics 32 (12) (2016) 121–127.

[11] A. Dedinec, S. Filiposka, A. Dedinec, L. Kocarev, Deep belief network based electricity load forecasting: An analysis of macedonian case, Energy 115, Part 3 (2016) 1688 – 1700, sustainable Development of Energy, Water and Environment Systems.

[12] G. E. Hinton, R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (5786) (2006) 504.

[13] Y. Bengio, Learning Deep Architectures for AI, Now Publishers, 2009. [14] H. Lee, R. Grosse, R. Ranganath, A. Y. Ng, Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations, in: International Conference on Machine Learning, 2009, pp. 609 – 616.

[14] G. E. Hinton, S. Osindero, Y. W. Teh, A fast learning algorithm for deep belief nets, Neural Computation 18 (7) (2014) 1527–1554.

[15] I. Arel, D. C. Rose, T. P. Karnowski, Deep machine learning - a new frontier in artificial intelligence research [research frontier], IEEE Computational Intelligence Magazine 5 (4) (2010) 13–18.

[16] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation 9 (8) (1997) 1735.

[17] F. A. Gers, J. Schmidhuber, F. Cummins, Learning to forget: Continual prediction with lstm. neural computation 12(10): 2451-2471, Neural Computation 12 (10) (2000) 2451–2471