Application of Hidden Markov Model in Crude Oil Price Forecasting

Abdul Talib Bon
Faculty of Technology Management, Business and
Entrepreneurship University Tun Hussein Onn Malaysia
Parit Raja, Batu Pahat
86400, Johor, Malaysia
talib@uthm.edu.my

Nuhu Isah
Faculty of Technology Management, Business and
Entrepreneurship University Tun Hussein Onn Malaysia
Parit Raja, Batu Pahat
86400, Johor, Malaysia
nuhuisah33g@gmail.com

Abstract

Forecasting of future crude oil price in one of the most challenging problem in forecasting technique due to its haphazardness, non-linearity and complexity, many models are available to predict these non-linear and complexity. Recently, many researchers developed models based on Artificial Neural Network (ANN), Support Vector Machine (SVM), Fuzzy Logic (FL) and Moving Average (MA). In this paper, we developed a crude oil price forecasting model based on Hidden Markov Model due to its proven fittingness for modeling vigorous systems and pattern classification. We apply the HMM methodology to forecast the crude oil price from 1996 to 2015 using available past datasets from West Texas Intermediate. We used Matlab software in training and validating the data. The results obtained using Hidden Markov Model is inspiring because of the little error given. Lastly, we recommend development of fusion models by employing Hidden Markov Model with other lenient computing models.

Keywords
Crude Oil Price, Hidden Markov Model, Forward-Backward Algorithm, Baum-Welch Algorithm

1. Introduction

Crude oil is a critical and unavoidable product around the world. It is essential part of the financial development of the developing and developed countries in the study of Benería and Floro (2015). Moreover, political occasions, climate, unpredictability, multifaceted nature in crude oil market are a few determinants of unrefined petroleum price Kilian and Murphy (2014). The impact of unrefined petroleum price unpredictability will colossally influences significant number of market practitioners which have direct impact on the monetary value. In this manner, to diminish the hurtful impact of crude oil price varieties, it is imperative to forecast the variability of the crude oil price.

In spite of the fact that it could be wrangled about that the time of raw petroleum is going to be over on the planet. Nevertheless, some inquires about have contended that worldwide request of unrefined petroleum would increment for the long haul in spite of the way that crude oil demand from organization for economic co-operation and development countries (OECD) have declined, however, the need for crude oil has expanded due to the expansion in requests of non OECD nations, particularly China Kaygusuz, K. (2012). In addition, the way that considerable measure of crude oil comes from the temperamental Middle East implies that there are conceivable outcomes of crude oil price instabilities Middle East Research Institute (2015), accordingly, forecasting crude oil price is extremely basic for financial improvement of numerous countries according to Kaygusuz (2012). In this paper we exhibit Hidden Markov Model for crude oil price forecasting.

According to Agnolucci (2009), there were a significant number of researches with respect to the Box-Jenkins models and GARCH models in crude oil price forecasting. Crude oil price unpredictability has multiplier as well as topsy-turvy consequences for the economy by Basher and Sadorsky (2006). A little change in crude oil price massively influence monetary exercises of a country while any progressions in financial exercises in a country affect crude oil price. Along these lines, crude oil price forecasting have vast macroeconomic impacts on an economy. Sadorsky (2006) found that the out-of-test estimates of a solitary condition GARCH model are better than those of state space, vector autoregression and bivariate GARCH models in forecasting crude oil price. GARCH model is capable to have a superior execution than the inferred instability models regarding prescient accuracy by Kang and Yoon (2013).

Be that as it may, crude oil price forecasting is among the prime issues of artificial intelligent (AI) community following the analysis of Ahmadi et al. (2013). The aim of forecasting research has been generally beyond the ability of outdated AI research which has mainly focused on developing intelligent systems that are supposed to emulate human intelligence by Kang and Yoon (2013) and Ahmadi et al. (2013). The nature of crude oil price is complexity, nonlinearity and additionally volatility by Yu et al. (2014). The rate of price swings in such arrangement relies upon many causes, for example, financial, political and social etc. Yu et al. (2014). Thusly, to create AI frameworks for this kind of forecasting needs an iterative procedure of uncovering learning and change of the framework through information mining, information designing, information driven displaying, and experimentation. Crude oil is a fundamental piece of the worldwide economy by Bon and Isah (2016). A shakiness in unrefined petroleum price influences our individual and corporate money and related exercises, and the financial prosperity of any given country by Bon and Isah (2016). A splendid forecasting model for unrefined petroleum price is to a great degree required. A considerable number of researches have been undertaking and is keeping on finding a perfect forecasting model for raw petroleum price by Ahmadi et al. (2013) and Yu et al. (2014). A large portion of the forecasting research has utilized time series techniques like ARIMA model, GARCH model and multiple regressions by Chatfield (2016). As of late, various unrefined petroleum price forecasting methods in view of AI, including artificial neural network, fuzzy logic and support vector machines have been proposed by Chatfield (2016). Nonetheless, the vast majority of them have their own weaknesses by Bon and Isah (2016). ANN has a few issues on account of its chosen auxiliary design by Teo et al. (2015). Some researchers have used fuzzy systems in forecasting, but in order to develop a model to forecast crude oil price behavior using fuzzy logic, you need to have some background skilled knowledge about it because of its complexity by Li et al. (2015). In this paper, we used Hidden Markov Model (HMM) to forecast crude oil price. The HMMs have been widely used in different areas such as speech recognition, electrical signal's forecasting and image processing by Farhadi et al. (2015)etc. Hidden Markov Model is used here to forecast and or model crude oil price.

2 Methodology

In this section, we discussed HMM methodology along with the relevant processes to build, initialize and train the HMM.

2.1 Hidden Markov Model

A hidden Markov model (HMM) is a statistical model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. An HMM can be presented as the simplest dynamic Bayesian network by Wilson and Bobick (1999). The HMM represents the overall process behavior in terms of movement between states and describes the inherent variations of the observations within a state. In fact, HMM is a stochastic sequences as Markov chains where the states are not directly observed but are associated with a probability function by Patterson et al. (2009). HMM is characterized by the following:

i. The number of states in the model,
ii. The number of observation symbols,
iii. The state transition probabilities. Probabilities to move from state i to state j, i.e $P(q_j \mid q_i)$,
iv. The observation emission probability distribution that characterize each state q_i, i.e. $p(x \mid q)$,
v. The initial state distribution.

Thus having a Markov chain with transition matrix A, observation emission matrix B and prior state probability π, we can build the Markov model $\lambda = (A, B, \pi)$ by Smyth (1995). By training the model and making adjustments to the values of A, B and π, we can compute $P(O \mid \lambda)$, the probability of the observation sequence $O(=O_1O_2\dots O_T)$ given the model λ, using the equation

$$P(O / \lambda) = \sum_{allQ} P(O / Q\lambda) P(Q / \lambda)$$

$Q = states\ sequence\ q1, q2, ..., qT$ in the markov model.

To construct a HMM for a given problem, we must build an initial HMM and then train the HMM suitably to solve the problem. The following sections explain the operations to initialize and train the HMM.

2.2 Initialization

To construct an HMM we assume initial values of A, B and π. In this study, we generate the initial values of A and π randomly. During implementation of HMM, however, the choice of the model (spherically), choice of the number of states and observation symbol (discrete) become a tedious task by Krogh et al. (2001). For the present application of crude oil price forecasting, we constructed HMM by varying the parameters, and observed a better performance resulting from the left-right HMM. Since the choice of good initial estimates of the parameters of bqT (O) densities is essential for rapid and proper convergence of the re-estimation formulae, we used the segmental k-means training procedure for estimating the parameter values by Krogh et al. (2001).

2.2 Forward-Backward Algorithm

To calculate the value of $P(O|\lambda)$, it requires 2TNT calculations, since at every time slot t = 1,2, …. T, there are N possible states which can be reached. So there are NT possible state sequences and for each such state sequence about 2T calculations are required for each term. This is a complex task to do. Luckily, there is an effective procedure called Forward- Backward algorithm, which can be used to compute the value of $P(O|\lambda)$. If the number of states are N and the number of observations in a sequence is M, we assume that the state transition probability from state i to j is a_{ij} (i, j = 1, …..,N), and the observation emission probability from state i is $b_i(O_t)$, t = 1 ,….., M. The two probabilities given satisfied the probability law below

$$\sum_j aij = 1,$$

where, $a_{ij} \geq 0$ and $b_j(O_t) \geq 0$

Thus, we defined the forward probability iteratively as follows:

$$\pi_1(i) = \pi_i b_i(O_i)$$

$$\alpha_{t+1}(j) = \left[\sum_{j=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1})$$

Similarly we defined the backward probabilities as:

$$\beta_T(i) = 1$$

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$$

Similarly, we defined two more probabilities γ and ξ which we used in the learning of constraints of HMM. The equations used to calculate these two probabilities are:

$$\gamma\ (i) = \frac{\alpha_t(i)\beta_t(i)}{P(O/\lambda)} = \frac{\alpha_t(t)\beta_t(i)}{\sum \alpha_{t\, i=1}(i)\beta_t(i)}$$

$$\xi(i, j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O/\lambda)}$$

2.3 Baum-Welch Algorithms

A number of algorithms are available to adjust the constraints or parameters of the HMM with observation sequences. Among them are Baum-Welch (BW) algorithm by Yu and Siskind (2013) that is widely used. In BW algorithm, the parameters of an HMM are trained to maximize the sum of the log- likelihood of each training sequences. Baum-Welch algorithm is used to re-estimates the probabilities π, $\hat{a}$ and $\hat{b}$ using the values of γ and ξ. These are

estimates are $\hat{\pi}_i = \gamma_1(i)$

2.4 Forecasting Process

To forecast crude oil price using HMM, the continuous signal formed by the number of predictors and the variable of interest (for example, for a d-dimensional data, if the predictor variables are: X_1, X_2, ….., X_d and the variable of interest is y then a signal will be formed using values of X_1, X_2, X_3, ….., X_d and y) is input into a HMM for training the HMM. After the HMM is trained, this becomes an expert HMM that is well suited to the training data signals. Then, for any continuous signal of the size of training data signal, the HMM produces the probability which depicts how well the data signal matches with the trained HMM. In other word, it can be explained that this probability gives us a measurement of closeness of the data signal (in terms of fluctuations with respect to time) with the set of training data signals. This value being a small number we take the log of probability, which is called the likelihood value. After having a trained HMM we produce likelihood values for each of the training data signals. The average of the produced likelihood values is calculated. We refer to this average likelihood value as η.

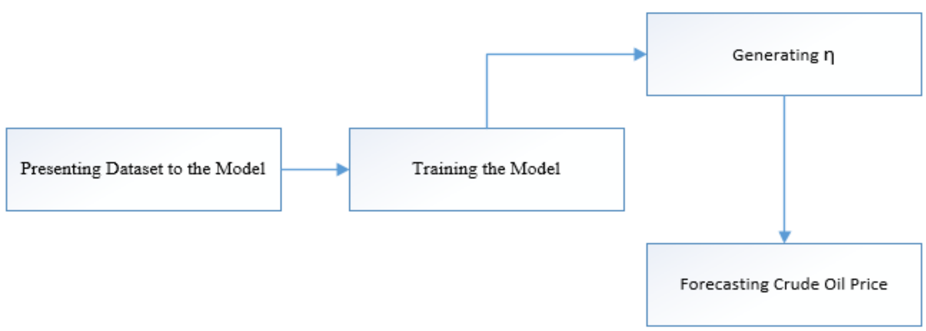

Figure 1: Steps in the Forecasting Process

For any unknown data signal wherein the values of predictors are known but the value of variable of interest is unknown, it is assumed that if the value of unknown parameter would have been known then the signal formed could produce the same likelihood value as that of the average likelihood value η. So by taking into the consideration the aforementioned assumption we estimate the exact value of the unknown parameter via an iterative process, such that the signal formed produces a likelihood value equal to the average likelihood value η .The value of unknown parameter thus found is then the forecast value for the given values of the predictor variables. Figure 1 shows the steps of the forecasting process.

3 Result

The data used in this study was collected from West Texas Intermediate for the period 1996 to 2015.We selected 70% of the data for training and reserved the remaining 30% of the data for testing. To evaluate the forecast efficiency of the model, we used the first feature from the test dataset as inputs variables for the trained HMM and the last remaining feature to compare with the obtained forecast.

3.1 Choosing the Model Parameters

In building the HMM, we choose 2 states HMM and mixture observation probability, and the left-right HMM process for the crude oil price forecasting. The initial value *A* and π were chosen randomly and then normalized to satisfy the following conditions:

$$\sum_{i=1}^{N} \pi_i = 1 \text{ and } \sum_{j}^{N} a_{ij} = 1$$

Where, *N* = number of states

A three-dimensional Gaussian distribution was used to determine the observation probability by Gelman et al. (2014).

$$b_j(O) = \sum C_{jm} \xi \left[O, \mu_{jm}, U_{jm} \right]$$

Where, C_{jm} = mixture coefficient for the m-th mixture in

3.2 Test Result

Figure 2 displays the actual and forecasted crude oil prices from 1996 to 2015. The performance of the model is evaluated by calculating the coefficient of determination, R^2 and the mean relative error by Zeng et al. (2014). R^2 close to 1 indicates that the forecasted price closely follows the fluctuations in the actual price index. A positive mean relative error indicates that the model tends to underestimate the index.

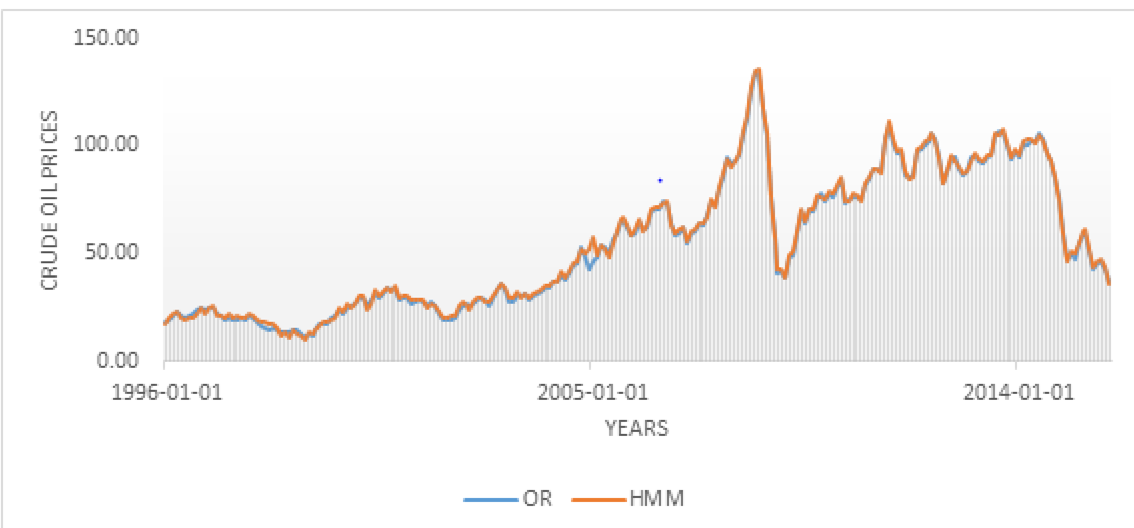

Figure 2. Actual and Forecasted Movement of Crude Oil Price
from 1996 to 2015 Using Hidden Markov Model

The above figure shows the forecasted value of crude oil price together with the actual crude oil price. Blue line show the flow of actual crude oil price while the red line indicates the forecasted value of crude oil price using hidden markove model. In this model, the R^2 is 0.93112, it shows that there is string relationships between the actual crude oil price and the forecasted crude oil price.

Acknowledgements
I would like to thank UTHM, ORICC for supporting and sponsoring this research paper under the grant no. U731. My appreciation goes to my supervisor for his guidance throughout this research paper.

References
Benería, L., Berik, G., & Floro, M. (2015). Gender, development and globalization: economics as if all people mattered. Routledge.

Agnolucci, P. (2009). Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models. Energy Economics, 31(2), 316-321.

Ahmadi, M. A., Ebadi, M., Shokrollahi, A., & Majidi,

S. M. J. (2013). Evolving artificial neural network and imperialist competitive algorithm for prediction oil flow rate of the reservoir. Applied Soft Computing, 13(2), 1085-1098.

Basher, S. A., & Sadorsky, P. (2006). Oil price risk and emerging stock markets. Global Finance Journal, 17(2), 224 251.

Bon, A. T., & Isah, N. (2016, November). Hidden Markov Model and Forward-Backward Algorithm in Crude Oil Price Forecasting. In IOP Conference Series: Materials Science and Engineering (Vol. 160, No. 1, p. 012067). IOP Publishing.

Chatfield, C. (2016). The analysis of time series: an introduction. CRC press.

Farhadi, H., AmirHaeri, M., & Khansari, M. (2015). Alert correlation and prediction using data mining and HMM. The ISC International Journal of Information Security, 3(2).

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). Bayesian data analysis (Vol. 2). Boca Raton, FL, USA: Chapman & Hall/CRC.

Kang, S. H., & Yoon, S. M. (2013). Modeling and forecasting the volatility of petroleum futures prices. Energy Economics, 36, 354-362.

Kaygusuz, K. (2012). Energy for sustainable development: A case of developing countries. Renewable and Sustainable Energy Reviews, 16(2), 1116-1126.

Kilian, L., & Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. Journal of Applied Econometrics, 29(3), 454-478.

Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. Journal of molecular biology, 305(3), 567-580.

Li, H., Pan, Y., & Zhou, Q. (2015). Filter design for interval type-2 fuzzy systems with D stability constraints under a unified frame. IEEE Transactions on Fuzzy Systems, 23(3), 719-725.

Patterson, T. A., Basson, M., Bravington, M. V., & Gunn, J. S. (2009). Classifying movement behaviour in relation to environmental conditions using hidden Markov models. Journal of Animal Ecology, 78(6), 1113-1123.

Smyth, P. J. (1995). U.S. Patent No. 5,465,321.

Washington, DC: U.S. Patent and Trademark Office.

Teo, T. T., Logenthiran, T., & Woo, W. L. (2015, November). Forecasting of photovoltaic power using extreme learning machine. In Smart Grid Technologies-Asia (ISGT ASIA), 2015 IEEE Innovative (pp. 1-6). IEEE.

Wilson, A. D., & Bobick, A. F. (1999). Parametric hidden markov models for gesture recognition. IEEE transactions on pattern analysis and machine intelligence, 21(9), 884-900.

Yu, H., & Siskind, J. M. (2013). Grounded Language Learning from Video Described with Sentences. In ACL (1) (pp. 53-63).

Yu, L., Zhao, Y., & Tang, L. (2014). A compressed sensing based AI learning paradigm for crude oil price forecasting. Energy Economics, 46, 236-245.

Zeng, K., Gao, S., Mozafari, B., & Zaniolo, C. (2014, June). The analytical bootstrap: a new method for fast error estimation in approximate query processing. In Proceedings of the 2014 ACM SIGMOD international conference on Management of data (pp. 277-288). ACM.

Biographies

Dr. Abdul Talib Bon is Professor of Technology Management in the Department of Production and Operations Management at the Universiti Tun Hussein Onn Malaysia. He has a PhD in Computer Science, which he obtained from the Universite de La Rochelle, France. His doctoral thesis was on topic Process Quality Improvement on Beltline Moulding Manufacturing. He studied Business Administration in the Universiti Kebangsaan Malaysia for which he was awarded the MBA. He's bachelor degree and diploma in Mechanical Engineering which his obtained from the Universiti Teknologi Malaysia. He received his postgraduate certificate in Mechatronics and Robotics from Carlisle, United Kingdom. He had published more 150 International Proceedings and International Journals and 8 books. His research interests include manufacturing, forecasting, simulation, optimization, TQM and Green Supply Chain. He is a member of IEOM, IIE, IIF, TAM, MIM and council member's of MSORSM.

Nuhu Isah is from Katsina state and was born in 1985 at Katsina state of Nigeria. He attended Musa Yar'adua Qur'anic Model Primary School, then to Sir Usman Nagogo College of Arabic and Islamic Studies and finished his secondary school at Kiddies International College Katsina, Katsina State of Nigeria. He also attended Ahmadu Bello University Zaria, Nigeria where he obtained B.Sc. Business Administration in 2011. He started working with Unity Bank Nigeria Plc from 2011 to 2013 where he worked as recruitment assistant officer in Human Capital Management Department and then transferred to branch as teller in Operations Department. Thereafter, he applied and later gains admission for Master's Degree in Technology Management at University Tun Hussein Onn Malaysia (UTHM). After his MSc in 2015, he immediately continues his PhD in the same university at the same faculty. The author published 1 book and more than 15 International Proceedings and Journals.