IRANIAN ECONOMIC REVIEW
Print ISSN: 1026-6542
Online ISSN: 2588-6096
https://ier.ut.ac.ir/
University of Tehran

Forecasting Gasoline Consumption in Iran using Deep Learning Approaches

Neda Bayat[a,*], Mansoor Davoodi[b], Ali Rezaei[c]

a. Department of Economics, Qazvin Branch, Islamic Azad University, Qazvin, Iran.
b. Department of Computer and Information Sciences, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran.
c. Faculty Statistics, Mathematics and Computer, Allameh Tabataba'i University,Tehran, Iran.
* Corresponding Author, E-mail: n.bayat@qiau.ac.ir

Article Info

Article Type: Research Article

Article History:
Received: 21 January 2021
Received in revised form: 15 May 2021
Accepted: 22 June 2021
Published online: 01 July 2023

Keywords:
Deep Learning,
Forecasting,
Gasoline Consumption.

JEL Classification:
Q47, C53, D83, C45.

ABSTRACT

Gasoline consumption is one of the challenging issues of energy management in Iran. The deficit of domestic production and the need for imports on one hand, and the impact of its consumption on macro-and micro-economic variables, on the other hand, cause gasoline consumption management has become more important. The more accurate, prediction of the trend of gasoline consumption is the more successful consumption management will be. Since gasoline consumption is affected by several parameters and factors, so, forecasting its consumption with high accuracy is difficult. In this paper, one recursive competitive learning method and two deep learning methods are utilized to provide more accurate forecasting of gasoline consumption. Due to the impact of gasoline consumption patterns on seasonal changes, climate and holidays, different periods are used for training the learning of these approaches, and their efficiency is compared in terms of the standard error metrics. The comparison results show the deep learning approaches and the training patterns with 12 months result in more accurate predictions. Finally, using the best approach and obtained setting, the gasoline consumption in Iran is predicted for the next years, which shows that gasoline consumption will grow 22 percent by 2027.

Cite this article: Bayat, N., Davoodi, M., & Rezaei, A. (2023). Forecasting Gasoline Consumption in Iran using Deep Learning Approaches. *Iranian Economic Review*, *27*(2), 347-376. DOI: https://doi.org/10.22059/ier. 2021.83902

 Publisher: University of Tehran Press.
DOI: https://doi.org/10.22059/ier.2021.83902

1. Introduction

In today's modern world, energy is one of the most indispensable elements of any economy. The transition from coal to oil resources and the emerge of industrialization and globalization have led to the exponential growth of energy consumption (Suganthi and Samuel, 2012). According to the statistics released by the International Energy Agency, within 10 years up to 2017, the average annual growth rate of energy consumption was 1.5 percent (IEA statistics, 2019). In the same year, EIA also predicts that global primary energy demand will grow at an annual rate of approximately one percent at least per year by 2040 (Kim and Cho, 2019). Some studies also have expressed that if current consumption patterns in the world continue, energy consumption will increase by about 50% by 2030 compared to 2012 (Suganthi and Samuel, 2012). Moreover, the main sectors of energy consumption will be the power and transportation sectors by 2030.

Population growth and economic development of countries are the main drivers of increasing energy demand in the world. Some studies declare that developing countries, especially in Asian regions like China and India will have the largest share of the world's energy demand soon. Because developed ones have both lower population growth rates and relatively mature levels of economy, which lead to no significant increase in energy consumption (Broadhead and Killmann, 2008). According to energy consumption, it is likely that demand management and energy supply will be taken into account as a challenge for countries shortly. For this reason, policymakers need to predict energy demand to ensure economic growth and country development so that can be effective in managing its supply and optimal resource allocation.

The transportation sector in countries has always been used as one of the indicators to evaluate industrial development. Among all types of energy, petroleum products have about 36.3% of Iran's total energy consumption in 2018, and the transportation sector with a share of

64.9% is the largest consumer of these products (Iran’s Energy Balance Sheet, 2018). In recent years, due to the development of communities and, consequently, the development of the transportation sector, gasoline has been considered as one of the important energy sources that always plays a serious role in countries’ economic growth (Azadeh et al., 2015). Today, the development of economies, population growth, car production rates, and an increase in car ownership, all are factors that lead to gasoline consumption increase. If this trend continues and there is no control over it, we expect to confront serious gasoline shortages in the near future and environmental problems.

Gasoline consumption is often considered as one of the countries’ development indicators by policymakers. But gasoline consumption policies in industrialized countries are usually different from oil-producing developing countries. For example, in industrialized countries such as the United States and Canada, policymakers pay special attention to gasoline demand, its pricing and control, while in oil-producing developing countries such as Iran and Saudi Arabia, pricing is less important than in industrialized countries, and policymakers prefer to provide inexpensive gasoline to achieve economic growth and social welfare (Azadeh et al., 2010).

According to the latest statistics published in Iran’s Energy Balance Sheet by the Ministry of Energy, primary energy production in 2018 was equal to 3060.6 million barrels of crude oil, of which 53.2% is devoted to crude oil, liquids, gas condensate and additives, 5.45% to natural gas, 0.3% to hydropower, wind, solar, 0.4% to nuclear energy, 0.3% to combustible renewable sources, and 0.3% is devoted to coal. Also, the total final energy consumption was 1363.5 million barrels, of crude oil, which compared to 2017, has increased by 0.8 percent. In recent decades, in developed countries, despite high per capita incomes and the possibility of having more diverse types of equipment and machinery, some measures have been taken to increase productivity, which has reduced per capita energy consumption. In contrast, Iran’s per capita final energy consumption is 1.8 times more than the world’s average per capita final consumption. Also, Iran's per capita final energy consumption in agriculture, household, commercial, public sectors, transportation, and industry is 3.3, 2.2, 1.5, and 1.5 times more than the world average, respectively.

Comparison of final per capita energy consumption by energy carriers with the global average shows that the per capita final consumption of natural gas is 6.6 and crude oil and petroleum products is 1.4 times the global average, which shows low productivity in operation and waste of energy in Iran. The highest share of petroleum products consumption is in gas oil and gasoline with 41.6% and 40.5%, respectively, and the lowest share is in liquid petroleum gas and kerosene with 5.2% and 0.6%. The transportation sector accounts for about 64.8% of the total oil product consumption. The main fuels consumed in Iran’s transportation sector are related to gasoline and gas oil, which with a share of 99.7% has the largest share of gasoline consumption in the country. Gasoline consumption in 2018, with a growth of 8% compared to 2017 reached 29.4464 million liters. A study of gasoline production and consumption statistics shows that gasoline consumption in Iran has always been higher than its production, so gasoline imports after a period of decline in the 1390s, has risen again and reached its highest level in 2018.

There are several reasons for the increase in gasoline imports in the recent decade such as insufficient growth of the country’s refinery production, implementation ceases of law enforcement on targeted subsidies in 2011 and 2014, termination of gasoline quotas in 2015 after 8 years and also growth cessation of alternative fuel share (CNG) from the light transport basket in the period of 2013 to 2017 .It increased by 4.3% to 12.63 million liters per day in 2017 (Iran Energy Balance Sheet, 2016). Funding the currency needed to import this level of gasoline, despite the sanctions and sharp rise in the exchange rate in recent years, has imposed an enormous cost on Iran's economy, which is making it even more urgent to manage gasoline demand. Accurate predictions on gasoline demand can be a great help to policymakers in managing gasoline demand and making the right strategic decisions.

So far, several studies have been conducted in different ways on the prediction of gasoline in the world and Iran. Conventional models require specific data processing methods to determine the linear relationships between dependent and independent variables. Multivariate calibration methods are usually used to maintain computational resources, reduce dimensions, and solve linearity problems in these models, which leads to the loss of some parts of the sample information. In recent years, new techniques have been introduced for predictive issues, such as economic variables as consumption. Artificial intelligence and machine learning are examples of these methods used to analyze and predict many variables, including gasoline consumption (Wang et al., 2020).

Deep learning methods as one of the new machine learning techniques have been extensively used by researchers in many fields. Unlike conventional learning methods, deep learning methods have a more complex architecture with more learning layers which makes them suitable for nonlinear and complex optimization problems (Kong et al., 2017). Deep learning methods can analyze and manage large amounts of information and complex issues over a reasonable time. The results of studies that have used these methods for prediction problems are various as they have used many different algorithms that do not necessarily mean the optimal solution (Wu et al., 2020). However, in recent years, the results of many studies that have used these methods to predict time series indicate that these methods have performed better than other traditional methods (Fu et al., 2016). Recursive neural networks with history, including LSTM and GRU, are very powerful examples of such methods. They have a recursive structure with memories to increase network performance in remembering the local and global patterns. These methods have been used in many real-world problems such as translation and processing of natural language, description of images, speech recognition, and prediction of economic variables, including variables related to energy (Kong et al., 2017).

This paper aims at utilizing deep learning approaches on the series of gasoline consumption in Iran and predicting it for future years. To this end, three well-known approaches *Recurrent Self-Organizing Map* (RSOM), *Long Short-Term Memory* (LSTM), and *Gated Recurrent Unit* (GRU) are customized and applied for the gasoline consumption in Iran's monthly gasoline consumption from 2002 up to 2017. Different settings and pattern lengths are examined in the approaches and the comparison results have been presented. Finally using the best-obtained setting and approach, future gasoline consumption is predicted.

2. Literature Review

Given the fact that the transportation sector in many developing countries including Iran, relies on gasoline due to poor public transportation, forecasting gasoline consumption plays a key role in planning its future demand. Numerous studies have been conducted in the world and Iran on gasoline forecasting. These studies can be divided into two general categories in terms of thematic classification: The main part of the study predicts gasoline consumption or models and estimates gasoline demand, then calculates price and income elasticities by focusing on the price and the consumption rate of gasoline and using the variables affecting gasoline demand. The results of this study mainly indicate that long-term elasticities are often greater than short-term, and income elasticities are in most cases greater than price elasticities (EIA, 2017).

The second group of studies is aimed not only at predicting gasoline consumption or modeling its demand, but also at examining the relationship between gasoline and various variables, including environmental variables such as air pollution and greenhouse gas emissions, or economic variables as economic growth, unemployment, welfare, and economic policies. These studies have often been considered since 2000 and different results have been obtained depending on the type of variables, the used models, samples, and the area of studies.

In gasoline modeling studies, three different approaches of time series, qualitative modeling, and neural networks have received more attention. Most studies have used time series modeling, combining auto-regressive (AR) models and moving averages, namely "Auto-Regressive Integrated Moving Average", which are mostly used to predict tax revenue based on gasoline by local governments and states. These models have been hailed by many researchers for their simplicity. One of the newest methods used to predict gasoline consumption in recent years is neural networks. These methods are data-driven and are based on mathematical modeling of how the human brain works and makes decisions, and are mainly used to determine nonlinear relationships between explanatory and dependent variables. Although, these models are not implemented based on predetermined economic and behavioral theories they are designed to discover how these theories work. These methods are still in their infancy, and researchers are looking to prove the superiority and differences of these methods compared to traditional methods (EIA Independent Statistics and Analysis, 2017).

Examples of this research include the study of Nasr et al. (2002) who used neural networks to forecast gasoline consumption in two separate models. The first model was uni-variable and it was based only on values of gasoline consumption; the second model was three variables and the previous values of gasoline consumption, price Also, the number of cars is considered as an explanatory variable. (Nasr et al., 2002). Rahimi-Ajdadi and Abbaspour-Gilandeh (2011) used artificial neural networks with different learning algorithms and multivariate regression to predict the fuel consumption of tractors. Wu and Liu (2011) also used the neural network model to predict the fuel type of cars by taking into account the variables of vehicle type, engine type, car weight, and transmission system. Predić et al. (2016) used an instantaneous neural network to predict the fuel consumption of passenger cars in the city of Nice and considered two variables as the average vehicle speed and urban traffic situation as the explanatory variable.

In recent years, the use of machine learning methods concerning gasoline studies has also attracted the attention of researchers, although the number of these studies is much more limited than other methods. Pelayo et al. (2016) tried to anticipate gasoline consumption by designing a pattern of customer feelings and emotions. First, they used statistical analysis to identify the characteristics of gas stations that customers tend to choose based on their mental criteria. then, with the help of supervised machine learning classification methods, they predicted the possibility of customers choosing gas stations.

Mustaffa and Yusof (2012) introduced a mutation optimizer based on Levy's Probably distribution in the *bee colony algorithm* and improved the algorithm’s performance by optimizing *Least Squares Support Vector Machine* learning parameters, and used the algorithm to predict gasoline. Finally, the results of this algorithm were compared with the results of some conventional algorithms such as Backpropagation Neural Network (BPNN), which indicated the superiority of this algorithm.

Numerous studies have been conducted on gasoline in Iran. Most of them have focused on predicting gasoline demand in the country or transportation sector, and often use similar explanatory variables to model and predict and identify the effect of these variables on gasoline consumption. These studies have mainly used annual data on macro variables, price, and gasoline consumption, to compare neural network models with the traditional regression models or estimate price and income elasticities. Table 1 summarizes some of the related studies. The difference between the current study and previous ones is customizing and applying the deep learning methods to predict gasoline as well as their comparison. Moreover, a recursive learning approach, a recurrent self–organizing map, has been applied to clarify and highlight the robustness of the deep learning approaches.

Table 1. Review of Some Gasoline Studies in Iran

Source	**Method**	**Data**	**explanatory variables**	**subject**	**Finings**
Azadeh et al. (2015)	support vector regression algorithm - time series	2009 - 2011	number of holidays per week, amount of transported freight, number of transported passengers	forecasting short-term gasoline consumption	SVR outperforms other intelligent tools as ANN
Azadeh et al. (2012)	Integrated Genetic Algorithm-Conventional Regression-Analysis	1987 - 2002	Price of gasoline, GDP, number of vehicles, and gasoline consumption in the previous periods	gasoline demand estimation	genetic algorithm has a much better-estimated value for gasoline consumption in Iran
Abrishami et al. (2010)	Hybrid intelligent with GARCH (1, 1)-GMDH & GARCH (1, 1)-Hybrid intelligent-GMDH	2004-2008	none	Forecasting Gasoline	Hybrid intelligence with GARCH (1, 1) forecast better than others
Azadeh et al. (2010)	an adaptive intelligent algorithm based on artificial neural network (ANN)-conventional regression	1992 – 2005	price, GDP, population, number of vehicles, gasoline demand in the last periods, the correlation coefficient between variables.	forecasting long-term gasoline demand in the USA, Canada, Japan, Kuwait and Iran	ANN provides far less error than regression
Kazemi et al. (2009)	multi-level ANN	1968-2006	GDP, the population, the total number of vehicles	Gasoline demand forecasting	validity of the model

Source	**Method**	**Data**	**explanatory variables**	**subject**	**Finings**
Kazemi et al. (2010)	hierarchical artificial neural networks (ANNs) model based on the supervised multi-layer perceptron (MLP), trained with the back-propagation (BP) algorithm	1968-2007	GDP, population, the total number of vehicles, transport energy demand in the last year	transport energy demand forecasting	hierarchical ANNs predictions show its superiority, estimating energy demand for the period 2008 to 2020
Assari et al. (2009)	particle swarm optimization (PSO) and genetic algorithm (GA	1981-2005	population, GDP, import, export, gasoline production, number of cars	estimating gasoline demand	PSO and GA models are in good agreement with the observed data
Asgharizadeh and Taghizadeh (2012)	hierarchical artificial neural network (ANN) based on supervised multi-layer perceptron (MLP), trained with back-propagation (BP) algorithm	1967-2008	gasoline demand, GDP, population, the total number of vehicles	forecasting gasoline demand in Tehran	validity of the model, estimating demand for the period between 2011 and 2030
Fania and Norouzib (2019)	multilayered perceptron neural network	2010-2018	fuel price, population, median household income, Gini coefficient, hybrid/gasoline cars ratio, the price index of goods and services, average vehicles lifetime	forecasting gasoline demand in the Transportation Sector	gasoline demand in Tehran’s transportation section will be increased by 2022

3. Time Series Prediction and Deep Learning

In this section, first, we present a formal definition of the time series prediction problem, then we explain how we utilize learning approaches for predicting gasoline consumption in our study.

3.1 Time Series Prediction

A time series is a set of serial observations such as $y_1, y_2, \dots, y_T$ in time order. y_T is the last available observation and the problem of estimating y_{T+1} is called the *time series prediction* problem. Note that, by applying y_{T+1} and repeating the prediction process, the next observations, $y_{T+2}, y_{T+3}, \dots$, are estimated as well. Generally, the prediction process is based on a crucial assumption, that is, y_{T+1} can be described by applying $m \geq 1$ previous consecutive observations $y_{T-m+1}, \dots, y_{T-1}, y_T$. So, the prediction process can be described as a nonlinear complex function P such as $y_{T+1} = P(y_{T-m+1}, \dots, y_{T-1}, y_T) + \varepsilon_{T+1}$, where ε_{T+1} is a random variable to consider noise and\or uncertainty at time $T+1$. For simplicity, let denote $X_t = \langle y_{t-m+1}, \dots, y_{t-1}, y_t \rangle$, for $t = m, m+1, \dots, T-1$, and call $s_t = (X_t, y_{t+1})$ as a *known sample*.

The problem of time series prediction has a bunch of applications in real-world problems such as the stock market, energy consumption, weather forecasting and many economic problems (Hussain et al., 2008; Sapankevych et al., 2009). Unfortunately, there is no explicit formulation to describe the function P, so, there are many studies on finding (or say *learning*) the pattern in time series (Hrasko et al., 2015; Lee et al., 2009). Because of the successful applications of artificial intelligence and particularly machine learning approaches in thinking and learning as humans, these types of approaches are appropriate tools to deal with time series problems. They are usually able to handle uncertainty and noisy data, able to understand trends and hidden patterns in time series without finding a strict mathematical relation between the samples, and able to remember the long history of data utilizing extra memories. However, these types of approaches need more resources time and space to train in comparison with classical time series prediction methods.

Recurrent neural networks are strong tools for time series analysis. Although the idea of the recurrent neural network was proposed in 1980, because of providing possible computational resources and GPU, recently they have been significantly applied to real-world problems. In the rest of this section, we briefly introduce three robust recurrent learning and deep learning approaches, *Recurrent Self-Organizing Map* (RSOM) (Varsta et al., 2001), *Long Short-Term Memory* (LSTM) (Hochreiter and Schmidhuber, 1997; Graves, 2013), and *Gated Recurrent Unit* (GRU) (Cho et al., 2014), that were used in this research to predict monthly gasoline consumption in Iran. All these three approaches are evolving variations of artificial neural networks, and they need to be trained using known samples $s_t = (X_t, y_{t+1})$ before asking a prediction query, i.e. $s_T = (X_T, ?)$. RSOM, LSTM and GRU are three recurrent learning approaches whose input samples at time t contain two parts; the current sample s_t, and the output of the network at time $t-1$. Indeed, the outcome of the recurrent networks at time t is affected by both current input and previous output. This is exactly the reason, RSOM, LSTM and GRU have been used here for gasoline consumption prediction.

3.2 Recurrent Self-Organizing Maps

Recurrent Self-Organizing Maps (RSOM) are competitive variations of neural networks. RSOSM is a set of neurons, called *competitive units* (CU), which are connected with a known structure. CUs are arranged in a two-dimensional grid and each of them has some neighbors. Let denote CUs by $\{w_1, w_2, w_3, \dots, w_n\}$. To apply RSOM to learn the patterns in a prediction function such as $y_{t+1} = P(y_{t-m+1}, \dots, y_{t-1}, y_t)$, it suffices to consider each w_i as a vector with size $m+1$, e.g., $w_i = \langle w_i^1, w_i^2, w_i^3, \dots, w_i^m, w_i^{m+1} \rangle$. Each first m

dimensions of the vector correspond to exactly one entry in the known sample $s_t = (X_t, y_{t+1})$, and w_i^{m+1} corresponds to the outcome of the prediction function which is y_{t+1}. All these values are randomly initialized at the first step when the map is constructed.

In the training phase of RSOM, the known samples are presented to the grid one by one. The idea behind training is that by presenting each sample, CUs try to get close to the sample by smoothly changing their corresponding values. Precisely, by presenting a sample $s_t = (X_t, y_{t+1})$, first the most similar unit to s_t is found. This can be simply done by finding the distance between each unit and s_t as follows:

$$dis(s_t, w_i) = \sqrt{\sum_{j=1}^{m} \left(y_{t-j+1}, w_i^j\right)^2} \quad (1)$$

Note that, the dimension $m+1$ is not considered to calculate the distance function. This is because of handling an unknown query sample $s_T = (X_T, ?)$ which will be explained later.

To apply the recurrent property, the following formula is used to find the *best matching unit b*.

$$z_i = (1-\alpha)\, z_i + \alpha\, dis(s_t, w_i), \quad (2)$$

$$z_b = min_{1 \le i \le n}\{\|z_i\|\}. \quad (3)$$

The parameter α is the *leaking rate* and is used to control the effect of the outcome from the previous step. In fact, z_i which is the outcome of i^{th} unit is updated after presenting each sample as equation (3). The *winner* unit with the minimum z_i denoted is called the *best matching unit.* In the next step, the dimension values of CUs are updated to get close to s_t. The general idea behind this step is the units with close distance to the winner unit, change more than the units far from the winner. The following equations explain this idea precisely.

$$w_i^j = w_i^j + \gamma(t)\, h_{ib}(t)\, dis(s_t, w_i), \text{ for } j = 1,2,\ldots,m+1, \quad (4)$$

where $\gamma(t) \in (0,1)$ is the *learning coefficient* to control the changing a unit to get closer to s_t. $h_{ib}(t)$ is the *excitation* of i^{th} competitive unit when the best matching unit is b. $h_{ib}(t)$ is related to the distance between i^{th} unit and the winner unit. $\gamma(0)$ usually set to 1 and while presenting the samples, it reduces smoothly.

The ordered known samples $s_m, s_{m+1}, \ldots, s_{T-1}$ are presented one by one, and CUs have been updated using the explained procedure. This training phase may repeat several times, each iteration is called one *epoch*. Repeating this process reduces the effect of random initialization of the CUs. At the end of the training phase, the trained map is a result and it will be able to predict the outcome of an unknown sample $s_T = (X_T, ?)$. In fact, the outcome of X_T is unknown and needs to be predicted. To this end, first, it is presented to the learning map, and then the best matching unit is found. Since the first m dimensions of the best matching unit are similar to X_T, it is naturally expected the $(m+1)-th$ dimension is close to the outcome of the sample as well.

RSOMs are less sensitive to uncertain or noisy data. Indeed, since they determine the calculation of the best matching unit based on all m dimensions, they can handle the uncertain data very well. Also, RSOMs can work with incomplete data by ignoring the missed dimensions for the incomplete data in the mentioned equations. Finally, an important advantage of RSOMs is projecting the $m-$dimensional samples to a two-dimensional map. Indeed, similar samples are projected to the same unit. This helps the network understand hidden patterns in the samples.

3.3 Long Short-Term Memory and Gated Recurrent Unit

As aforementioned, both Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are memory-based types of recurrent neural networks. Figure 1 illustrates the framework of a recurrent neural network. X_t is the input sample and h_t is the hidden state at time t. h_t plays the role of memory which is defined based on history h_{t-1} and input X_t as follows:

$h_t = f(W_1\, X_t + W_2\, h_{t-1} + b_h),$

where b_h is a constant value as a bias and f is an activation function *tangent hyperbolic* or RELU (graves, (2013)). The output o_t is also defined as follows:

$o_t = softmax(W_3\, h_t)$.

This general structure of the recurrent neural networks works based on gradient and backpropagation. That is using the partial derivative to reduce the output error, and greedily update the weights and parameters with respect to the current state. Thus, it suffers from the *vanishing gradient* (or *exploding gradient*) issue, that is, after some iterations of the network the gradient gets smaller and smaller and in some cases, it causes stopping learning of the network by preventing the updating of the weights. The inverse case may happen in an exploding gradient problem when the gradients get larger. LSTM and GRU are two efficient approaches to overcome these gradient issues which are introduced in the following.

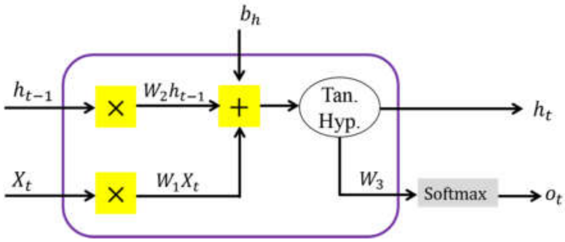

Figure 1. The Structural Framework Of A Recurrent Neural Network
Source: Research finding.

LSTM utilizes three gates to control and hold a balance between a new input sample and the information gathered and stored from the past samples in the network (Graves, 2013). The first gate is the *forget gate*, denoted by Γ_f, determines how much information from the previous time step is transferred to the next time step. The second one is the *update gate*, denoted by Γ_u. It determines whether the new information is used in the current time step or not. It also determines the weight of information if they would apply. Finally, the third gate is the *output gate* and denoted by Γ_o. It determines how much of the past information and the current information should be transferred to the next time step. Figure 2 illustrates the structural framework of LSTM networks, and the structure formulation of the gates is presented as follows:

$$\hat{C}_t = \tanh(W_c.[h_{t-1}, X_t] + b_C),$$

$$C_t = \Gamma_f.C_{t-1} + \Gamma_u.\hat{C}_t,$$

$$\Gamma_f = \sigma\left(W_f.[h_{t-1}, X_t] + b_f\right),$$

$$\Gamma_u = \sigma(W_u.[h_{t-1}, X_t] + b_u),$$

$$\Gamma_o = \sigma(W_o.[h_{t-1}, X_t] + b_o),$$

$$h_t = \Gamma_o.\tanh(C_t),$$

$$o_t = \text{softmax}(C_t),$$

where σ is the logistic sigmoid function.

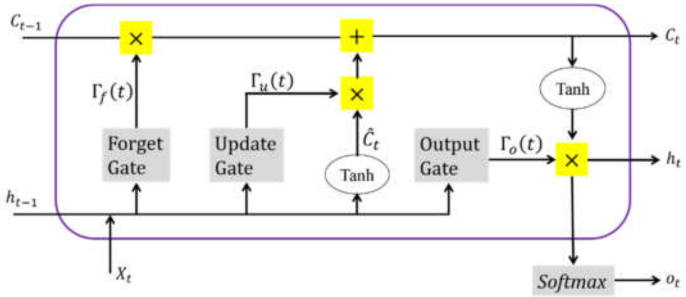

Figure 2. The Structural Framework of LSTM Networks
Source: Research finding.

GRU is a new simple variation of recurrent neural networks and has a similar structure to LSTM. It contains two gates called *update gate* Γ_u and *reset gate* Γ_r. The update gate operates as the update gate and forgets gate in LSTM and controls the tradeoff between the amount of information shod be thrown away and the amount of information that should be inserted into the history. The reset gate controls the amount of past information to forget before passing to the next time step. So, they can be formulated as follows:

$$\Gamma_u = \text{Sigmoid}(w_u[C_{t-1}, X_t] + b_u),$$

$$\Gamma_r = \text{Sigmoid}(w_r[C_{t-1}, X_t] + b_r),$$

$$\hat{C}_t = \tanh(w_c[\Gamma_r.C_{t-1}, X_t] + b_c),$$

$C_t = \Gamma_u . \hat{C}_t + (1 - \Gamma_u). C_t,$
$o_t = \text{softmax}(w_y . C_t + b_y),$

GRU uses fewer training parameters and memory. Consequently, it learns faster than LSTM. However, both LSTM and GRU networks are proper to work on the long sequence of the information. In the next section, the results of RSOM, LSTM and GRU for predicting gasoline consumption on the monthly data in Iran are presented.

4. Forecasting Results

This section presents the simulation results for the prediction of gasoline consumption in Iran. We coded the three learning approaches Recurrent Self-Organizing Map (RSOM), Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) using Python 3.7 programming language and implemented them on a PC with configuration CPU Intel(R) Core(TM) i5-2500k @ 3.30GHz and 4 GB RAM. The data used in this research was Iran’s monthly gasoline consumption from 2002 (1380 in the Solar calendar) up to 2017 (1396 in the Solar calendar) which was extracted from the reports of Iran's Energy Balance Sheet. All the available data was 192 monthly consumption which we separate into two *train samples* with a size 168 months and the *test samples* with a size 24 months.

Before using the time series data, first, we made it stationary by differencing a one-time step. Also, to reduce the sensitivity of the learning approaches to the scalability of the data, we normalized them in $[0, +1]$ at a preprocessing step. The preprocessed data is displayed in Figure 3.

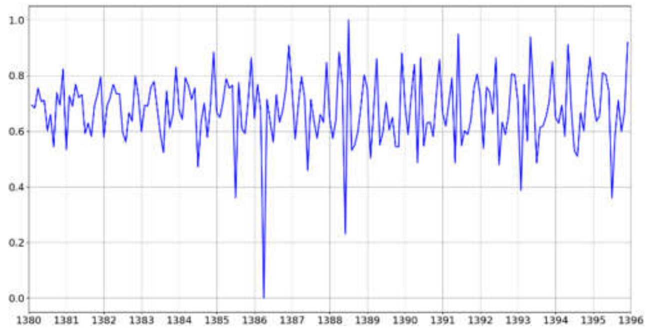

Figure 3. The Normalized and Stationery Monthly Iran's Gasoline Consumption
Source: Research finding.

We also used two standard error metrics to evaluate the learning algorithms, the *Root-Mean-Square Error* (RMSE) and the *Mean Absolute Percentage Error* (MAPE) which are formulated below.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{T}}, \tag{5}$$

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{y_t - \hat{y}_t}{y_t}\right|. \tag{6}$$

In these equations y_t is actual or measured output in period *t*, $\hat{y}_t$ is the predicted output, and T is the size of all observed samples. Since the gasoline consumption data is a seasonal series, we apply the learning algorithms for four different training patterns with size $m = 3, 6, 12$ and 24 months. Indeed, the parameter m is the number of consecutive previous seasons, used for predicting the gasoline consumption for the next season.

Since there are different learning parameters and coefficients in the learning approaches of RSOM, LSTM and GRU, we tried to set efficient values for them by trial and error approach. However, there is a tradeoff between the running time of the approaches and the prediction error. For the RSOM approach, we have used a grid with size 40×40, leaking rate $\alpha = 0.7$ and base learning rate $\gamma_0 = 1$. For LSTM and GRU approaches, we have default parameters in the library of *keras* at Python 3.7 with 2 hidden layers of 40 neurons, and

50 as the size of epochs. In each run of the approaches, we measured RMSE and MAPE in predictions of the algorithm on the 24 test samples. Finally, to ensure the results we ran the approaches 25 times and reported the best, the worst, the average and the standard deviation of the runs.

The prediction results of gasoline consumption for the last 24 months (months 169-192 in the available data) using pattern size $m = 3$ are shown in Table 2 and Figure 4. Figure 4 is the prediction data for the median results of RMSE. Note that, we ran each algorithm 25 independent times and reported the average results in the corresponding tables. However, in the figures, we depicted the median result (13^{th} best result out of 25 runs) in terms of RMSE. As can be seen. LSTM and GRU which are deep learning approaches obtained more precise results compared with RSOM.

Table 2. Comparison Results of RMSE and MAPE for Three Approaches RSOM, LSTM and GRU for Pattern Size m=3

	RMSE				MAPE			
	Min	**Max**	**Mean**	**Stdev.**	**Min**	**Max**	**Mean**	**Stdev.**
RSOM	5.910404	8.927331	6.928963	6.294258	0.427702	7.301528	8.413405	6.294258
GRU	3.869639	5.117954	4.298344	4.12426	0.326125	4.516787	5.326838	4.12426
LSTM	3.910419	5.413315	4.22059	4.089757	0.334523	4.433215	5.416821	4.089757

Source: Research finding.

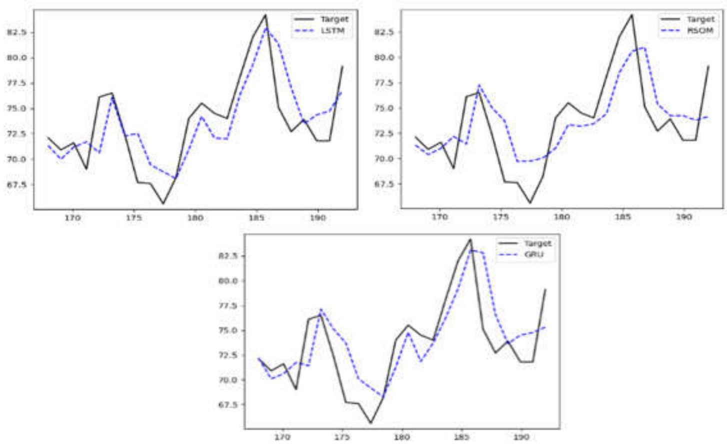

Figure 4. Prediction Results for Three Approaches of RSOM, LSTM, and GRU for Pattern size m=3.

Source: Research finding.

The prediction results for pattern size $m = 6$, $m = 12$, and $m = 24$ are presented in Table3 and Figure 5, Table 4 and Figure 6, and Table 5 and Figure 7, respectively.

Table 3. Comparison Results of RMSE and MAPE for Three Approaches RSOM, LSTM and GRU for Pattern Size $m = 6$

	RMSE				MAPE			
	Min	**Max**	**Mean**	**Stdev.**	**Min**	**Max**	**Mean**	**Stdev.**
RSOM	4.767876	6.830573	5.868465	0.407016	5.150814	7.15022	6.189264	0.423569
GRU	2.740791	4.61242	3.291588	0.399637	2.852976	4.790008	3.554744	0.381238
LSTM	2.935861	4.132814	3.147141	0.241833	3.17485	4.65224	3.387411	0.282936

Source: Research finding.

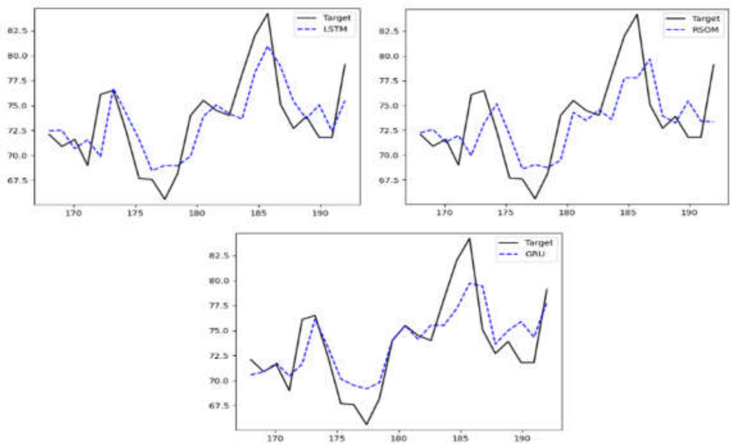

Figure 5. Prediction Results For Three Approaches Of RSOM, LSTM And GRU For Pattern Size $m = 6$.

Source: Research finding.

Table 4. Comparison Results of RMSE and MAPE for Three Approaches RSOM, LSTM and GRU for Pattern Size $m = 12$

	RMSE				MAPE			
	Min	**Max**	**Mean**	**Stdev.**	**Min**	**Max**	**Mean**	**Stdev.**
RSOM	4.009341	5.027954	4.412181	0.249776	4.337106	5.300941	4.664885	0.23006
GRU	2.256536	3.631777	2.846487	0.338666	2.208256	3.497033	2.75704	0.38211
LSTM	2.121221	3.776612	2.831736	0.419391	2.097501	3.717017	2.84112	0.374256

Source: Research finding.

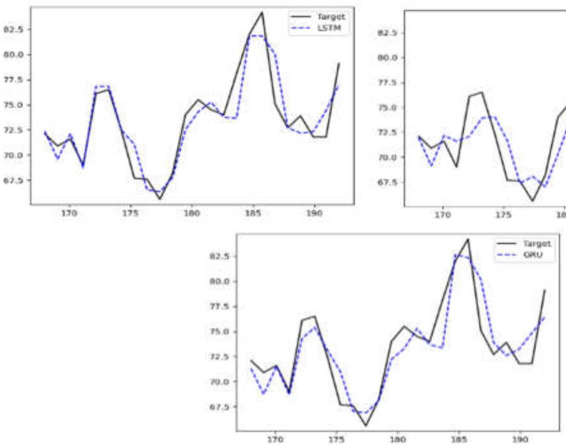

Figure 6. Prediction Results for Three Approaches RSOM, LSTM and GRU for Pattern Size $m = 12$.

Source: Research finding.

Table 5. Comparison Results of RMSE And MAPE for Three Approaches RSOM, LSTM and GRU For Pattern Size $m = 24$

	RMSE				MAPE			
	Min	**Max**	**Mean**	**Stdev.**	**Min**	**Max**	**Mean**	**Stdev.**
RSOM	4.356117	6.402825	4.980794	0.456191	4.903468	7.428351	5.584084	0.56188
GRU	2.824102	4.592952	3.423995	0.349356	3.115806	4.772153	3.827502	0.343904
LSTM	2.982085	4.348825	3.311752	0.337802	2.884763	4.38467	3.301966	0.358853

Source: Research finding.

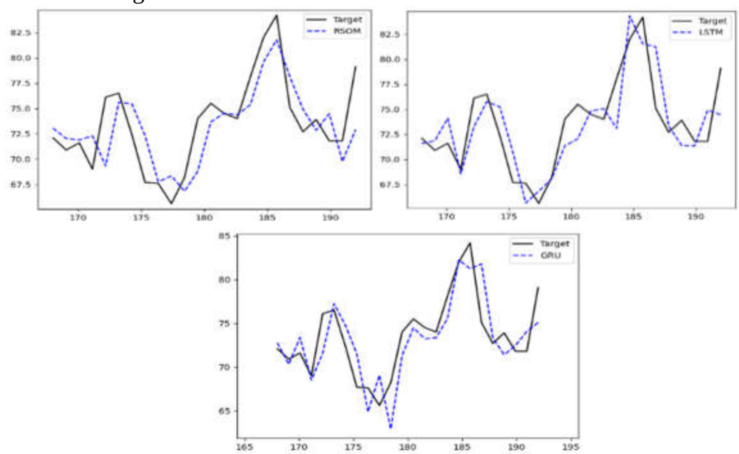

Figure 7. Prediction Results for Three Approaches of RSOM, LSTM and GRU for Pattern Size $m = 24$.

Source: Research finding.

The running time of each test for RSOM is about 34 seconds while for LSTM and GRU are about 72 and 63 seconds, respectively. So, as expected the deep learning approaches take more time to train the data. Of course, their accuracy is better than the learning approach. As it can be extracted from the reported results in the tables, in all the cases (different pattern's size m) the average accuracy of LSTM and GRU are better than RSOM's accuracy. However, there is no explicit comparison result between LSTM and GRU. Another important result refers to the learning pattern size, m. As it is shown, the results for $m = 12$ is better than the results for $m = 6$ and $m = 24$, and much better than for $m = 3$. These results hold for all three approaches.

Next, we use two approaches LSTM and GRU with the pattern size of $m = 12$, and predict future gasoline consumption. However, these approaches can simply predict any number of future years, we just predict the gasoline consumption for the years 2018−2025 (from 1397 up to 1404 in Solar colander). Note that, because of error propagation on the one hand, and changing the technology, economic and political circumstances, on the other hand, prediction for more future years may result in significant prediction error. Table 6 and Table 7 show the prediction for LSTM and GRU, respectively. Also, Figure 8 shows their diagram. It is worth mentioning that the difference between the diagrams is less than 2%, so, a trustable result may achieve using their average.

Table 6. Prediction Results of Iran Gasoline Consumption for 8 Years (1397-1404) Using LSTM with Pattern Size $m = 12$

	Month 1	**Month 2**	**Month 3**	**Month 4**	**Month 5**	**Month 6**	**Month 7**	**Month 8**	**Month 9**	**Month 10**	**Month 11**	**Month 12**
1397	78.6	81.3	80.4	78	84.1	80.9	85.6	83	84.9	77.5	76.1	86.1
1398	89	86.1	82.8	83.7	80.6	90.5	90.4	84	79.3	79.8	78.3	79.4
1399	88.7	83.1	85.9	83.6	86	92.1	88.5	89.4	82.7	81.6	81.6	83.2
1400	85	79.8	82.4	85.1	90.1	83.7	88.4	82.8	84.6	77.5	75.9	81.7
1401	87.2	86.1	84.6	85.6	86.2	80.8	93	91.4	82.4	83	81.1	83.6
1402	90.6	94.3	86.4	90	88.1	87.9	95.6	89	89.9	84.5	90.1	88.1
1403	96	97.1	85.8	89.7	89.6	93.5	96.4	96	93.3	89.8	91.3	88.4
1404	94.7	92.1	92.9	91.6	90	98.1	96.5	97.4	86.7	89.6	90.6	92.2

Source: Research finding.

Table 7. Prediction Results of Iran Gasoline Consumption for 8 Years (1397-1404) Using GRU with Pattern Size $m = 12$

	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12
1397	79.62	82.82	77.93	79.53	81.99	80.84	84.62	82.8	81.6	75.93	77.5	82.68
1398	87.52	84.5	83.67	81.59	79.15	89.51	92.3	84.63	82.51	79.68	79.6	81.4
1399	86.26	84.83	87.52	82.52	87.23	89.2	89.59	89.6	80.09	79.07	83.06	82.42
1400	85.57	81.21	80.47	85.43	90.33	85.28	88.28	84.93	83.3	76.3	75.83	82.44
1401	85.63	84.45	82.41	83.47	86.56	78.99	90.37	89.84	80.78	81.85	81.17	81.34
1402	90.12	89.51	83.87	87.11	87.06	89.95	93.72	89.16	86.06	83.24	86.86	86.21
1403	95.51	95.35	86.68	89.35	87.22	93.55	95	95.24	90.17	89.63	91.03	90.42
1404	92.26	90.93	91.11	88.95	92.19	95.55	94.65	95.15	86.98	88.25	91.15	92.08

Source: Research finding.

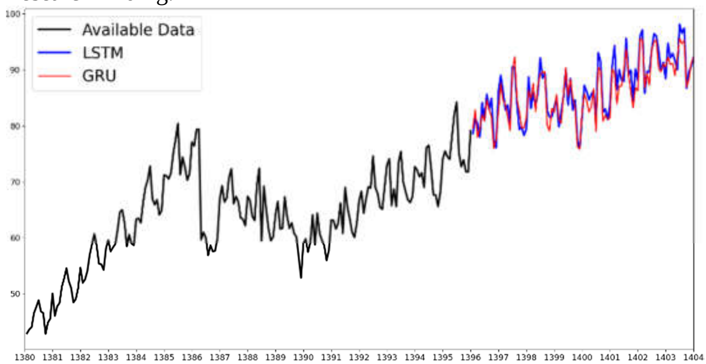

Figure 8. Prediction Results of Iran Gasoline Consumption for 8 Years (1397-1404) using LSTM and GRU with Pattern Size $m = 12$.

Source: Research finding.

Finally, we compare the forecasting performance of the three approaches by the DM test (Diebold and Mariano, 1995). We applied DM on both criteria RMSE and MAPE. The results are displayed in Table 8. The zero hypothesis points out there is no significant difference between the forecasted values. According to the DM test and the threshold value 1.96, the zero hypotheses cannot be rejected at the 5% level of significance for the pattern's sizes $m = 3$ and $m = 24$, which means, all three approaches have almost the same performance for such pattern sizes. However, for the pattern's sizes $m = 6$ and $m = 12$, LSTM and GRU have better performance and that means there is a significant difference between deep learning and RSOM. Further, LSTM outperforms GRU for the pattern size $m = 12$ at least in terms of the criterion MAPE.

Table 8. The DM Test with Criteria RMSE and MAPE. A, B and C are the Forecast Results Obtained by LSTM, GRU and RSOM, Respectively.

		$m = 3$			$m = 6$			$m = 12$			$m = 24$		
		(A,B)	**(A,C)**	**(B,C)**	**(A,B)**	**(A,C)**	**(B,C)**	**(A,B)**	**(A,C)**	**(B,C)**	**(A,B)**	**(A,C)**	**(B,C)**
RMSE	*DM*	-1.38	-1.34	-0.19	1.11	-2.28	-2.65	-1.80	-3.04	-2.53	-0.45	-0.67	-0.36
	p_val.	0.18	0.19	0.85	0.28	0.03	0.01	0.08	0.00	0.02	0.66	0.51	0.72
MAPE	*DM*	-1.00	-1.54	-0.94	1.10	-2.34	-2.43	-2.64	-3.04	-3.62	-0.45	-0.65	-0.24
	p_val	0.36	0.18	0.35	0.28	0.03	0.02	0.02	0.00	0.00	0.66	0.52	0.81

Source: Research finding.

4. Conclusion

Energy consumption management has become one of the priorities of policymakers, due to the essential role of energy in the economic development of countries and the increasing trend of its consumption. In Iran, gasoline is one of the major energy consumption items and is very important due to its 99.7% role in energy supply in the transportation sector. Gasoline imports have increased in recent years after a declining trend and reached its highest level in 2017 (1396 in the Solar calendar). Inconsistent policies that have led to inappropriate pricing, increased domestic car production, and also a serious weakness in the public transportation system are all reasons for the increase in gasoline consumption, which, in addition to insufficient domestic production, has led to an increase in gasoline imports. Import financing has doubled the importance of managing gasoline consumption due to rising exchange rates. Future consumption management is not possible without predicting its consumption trend.

There are several ways to predict economic variables, including gasoline consumption. Deep learning is one of the newest approaches to machine learning that, unlike conventional learning approaches, has a more complex architecture with more layers of learning and remembering global and local histories of data simultaneously. This powerful property makes deep learning approaches suitable to predict nonlinear and complex time series as well as to model the behavior of economic variables.

So far, several studies have been performed using traditional neural networks and classical methods to predict gasoline consumption in Iran, however, this study used two deep learning methods GRU, LSTM and the learning method RSOM to predict gasoline consumption and compare their efficiency with each other. Further, due to the possibility of influencing the gasoline consumption pattern from climate change and seasonal changes of different sizes of learning patterns $m = 3, 9,12$ and 24 months were used to train the models and the performance of each of them in terms of MAPE and RSME metric errors were compared. The results showed that, firstly, the LSTM and GRU deep learning methods performed better and $m = 12$ months resulted in the best predictor. Finally, by choosing the best method, the gasoline consumption was predicted until 2052 (1404 in the Solar calendar). The prediction showed that gasoline consumption will grow by about 22 percent.

If domestic production does not change and the current consumption pattern continues, the import of gasoline will inevitably increase. For this reason, policymakers are advised to make serious plans to build public transport infrastructure, increase domestic production, and improve consumption patterns. Also, considering the significant impact of gasoline import costs on microeconomic and macroeconomic variables, interested researchers are advised to study the effect of gasoline imports on household welfare, GDP, government budget, inflation and unemployment, and the impact of pricing policies.

References

Abrishami, H., Mehrara, M., Ahrari, M., & Varahrami, V. (2010). A

Hybrid Intelligent System for Forecasting Gasoline Price. *Iranian Economic Review*, *15*(27), 13-31.

Asgharizadeh, E., & Taghizadeh, M. R. (2012). A Hierarchical Artificial Neural Network For Gasoline Demand Forecast of Iran. *International Journal of Humanities*, *19*(1), 1-13.

Assari, M. R., Ghanbarzadeh, A., Assareh, E., & Behrang, M. A. (2009, June). Coal Demand Estimating In Iran Based on Socio-Economic Indicators Using Particle Swarm Optimisation and Genetic Algorithm. *2009 7th IEEE International Conference on Industrial Informatics*. Cardiff: IEEE.

Azadeh, A., Arab, R., & Behfard, S. (2010). An Adaptive Intelligent Algorithm for Forecasting Long Term Gasoline Demand Estimation: The Cases of USA, Canada, Japan, Kuwait and Iran. *Expert Systems with Applications*, *37*(12), 7427-7437.

Azadeh, A., Arab, R., & Behfard, S. (2010). An Adaptive Intelligent Algorithm for Forecasting Long Term Gasoline Demand Estimation: The Cases of USA, Canada, Japan, Kuwait and Iran. *Expert Systems with Applications*, *37*(12), 7427-7437.

Azadeh, A., Boskabadi, A., & Pashapour, S. (2015). A Unique Support Vector Regression for Improved Modelling and Forecasting of Short-Term Gasoline Consumption in Railway Systems. *International Journal of Services and Operations Management*, *21*(2), 217-237.

Azadeh, A., Boskabadi, A., & Pashapour, S. (2015). A Unique Support Vector Regression for Improved Modelling and Forecasting of Short-Term Gasoline Consumption in Railway Systems.

International Journal of Services and Operations Management, *21*(2), 217-237.

Azadeh, A., Mirjalili, M., Sheikhalishahi, M., & Nassiri, S. (2012). An Integrated Genetic Algorithm-Conventional Regression-Analysis Of Variance For Improvement Of Gasoline Demand Estimation. *International Journal of Industrial and Systems Engineering*, *11*(3), 205-224.

Broadhead, J., & Killmann, W. (2008). *Forests and Energy: Key Issues* (154). Rome: Food & Agriculture Org.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. *arXiv preprint arXiv*, Retrieved from https://arxiv.org/pdf/1406.1078.pdf?source=post_page

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, *13*, 253-63.

Fu, R., Zhang, Z., & Li, L. (2016). Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction. *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. Wuhan: IEEE.

Graves, A. (2013). Generating Sequences with Recurrent Neural Networks. *arXiv preprint arXiv*, Retrieved from https://arxiv.org/pdf/1308.0850

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735-1780.

Hrasko, R., Pacheco, A. G., & Krohling, R. A. (2015). Time Series Prediction Using Restricted Boltzmann Machines and Backpropagation. *Procedia Computer Science*, *55*, 990-999.

Hussain, A. J., Knowles, A., Lisboa, P. J., & El-Deredy, W. (2008). Financial Time Series Prediction Using Polynomial Pipelined Neural Networks. *Expert Systems with Applications*, *35*(3), 1186-1199.

Kazemi, A., Ganjavi, H. S., Menhaj, M., Mehregan, M., Taghizadeh, M., & Asl, A. F. (2009). A Multi-Level Artificial Neural Network for Gasoline Demand Forecasting of Iran. *2009 Second International Conference on Computer and Electrical Engineering*. Dubai: IEEE.

Kazemi, A., Shakouri, H. G., Menhaj, M. B., Mehregan, M. R., & Neshat, N. (2010). A Hierarchical Artificial Neural Network for Transport Energy Demand Forecast: Iran Case Study. *Neural Network World*, *20*(6), 761-772.

Kim, T. Y., & Cho, S. B. (2019). Predicting Residential Energy Consumption Using CNN-LSTM Neural Networks. *Energy*, *182*, 72-81.

Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Transactions on Smart Grid*, *10*(1), 841-851.

Lee, C. M., & Ko, C. N. (2009). Time Series Prediction Using RBF Neural Networks with A Nonlinear Time-Varying Evolution PSO Algorithm. *Neurocomputing*, *73*(1-3), 449-460.

Mustaffa, Z., & Yusof, Y. (2012). Lévy Mutation in Artificial Bee Colony Algorithm for Gasoline Price Prediction. *Knowledge*

Management International Conference (KMICe) 2012, Retrieved from https://repo.uum.edu.my/id/eprint/10963/1/CR180.pdf

Nasr, G. E., Badr, E. A., & Joun, C. (2002, May). Cross Entropy Error Function in Neural Networks: Forecasting Gasoline Demand. *FLAIRS Conference*, Retrieved from https://laur.lau.edu.lb:8443/xmlui/bitstream/handle/10725/6723/Cross.pdf?sequence=1

Pelayo, D. R., Pacheco, J. A. C., & Miralles-Pechuán, L. (2016). Emotion Analysis in Gasoline Consumption in Mexico using Machine Learning. *Advances in Computational Linguistics*, *8*, 1-20.

Predić, B., Madić, M., Roganović, M., Kovačević, M., & Stojanović, D. (2016). Prediction of Passenger Car Fuel Consumption Using Artificial Neural Network: A Case Study in The City of Niš. *Facta Universitatis, Series: Automatic Control and Robotics*, *1*(2), 105-116.

Rahimi-Ajdadi, F., & Abbaspour-Gilandeh, Y. (2011). Artificial Neural Network and Stepwise Multiple Range Regression Methods for Prediction of Tractor Fuel Consumption. *Measurement*, *44*(10), 2104-2111.

Report of Iran‘s Energy Balance Sheet. (2002-2018). Retrieved from https://www.iea.org/data

Sapankevych, N. I., & Sankar, R. (2009). Time Series Prediction Using Support Vector Machines: A Survey. *IEEE Computational Intelligence Magazine*, *4*(2), 24-38.

Suganthi, L., & Samuel, A. A. (2012). Energy Models for Demand Forecasting-A Review. *Renewable and Sustainable Energy Reviews*, *16*(2), 1223-1240.

Varsta, M., Heikkonen, J., Lampinen, J., & Millán, J. D. R. (2001). Temporal Kohonen Map and The Recurrent Self-Organizing Map: Analytical and Experimental Comparison. *Neural Processing Letters*, *13*(3), 237-251.

Wang, X., Yang, K., & Kalivas, J. H. (2020). Comparison of Extreme Learning Machine Models for Gasoline Octane Number Forecasting by Near-Infrared Spectra Analysis. *Optik*, *200*, 163325.

Wu, L. X., & Lee, S. J. (2020). A Deep Learning-Based Strategy to the Energy Management-Advice for Time-of-Use Rate of Household Electricity Consumption. *Journal of Internet Technology*, *21*(1), 305-311.