November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

AIDS's Drugs Quantification and Surveillance Using Deep Learning

Hadab Khalid Obayes[1], Nabeel Al – A'araji[2], Eman AL-Shamery[3]
[1]College of Information Technology, University of Babylon, Iraq
[2]Ministry of Higher Education, Iraq
[3]College of Information Technology, University of Babylon, Iraq
E-Mail: hedhab@uobabylon.edu.iq1
nhkaghed@itnet.uobabylon.edu.iq2
emanalshamery@itnet.uobabylon.edu.iq3

Article Info
Volume 81
Page Number: 282- 290
Publication Issue:
November-December 2019

Abstract: Consumption of medicines for a particular disease can be an indicator of the spread of the disease, as the increase in the consumption of medicines implies an increase in the incidence of the disease. Acquired immunodeficiency syndrome (AIDS) is a chronic, potentially life-threatening condition caused by the human immunodeficiency virus (HIV). AIDS is one of the deadliest diseases in human life. Therefore, monitoring the spread of AIDS through analyzing the consumption of its drugs and determining the places where the drugs are consumed geographically is an urgent necessity and brings useful information in the health sector. The main idea behind this paper is to employ a new approach of using deep learning as the main stage to predict the quantities of AIDS's drugs. Additionally, in the second stage the spatial concept is exploited to state the spread position of that disease. The deep neural network is a fully automated network that consists of a preprocessing layer, normalization layer and prediction layer depending on the state utilization drugs dataset of the USA for five consecutive years. Based on the results of the prediction process, the second stage represents the consumption of AIDS's drugs and produces a spatial map representing the disease Surveillance map. The results of the prediction process using the deep neural network are compared with the results of the linear regression method, as indicated by previous research. The deep network has given superior results by obtaining a very low value of (MSE).The results were encouraging and promising to rely on deep learning. The creation of a spatial map illustrated the spread of the disease and gave a clear vision based on the consumption of drugs.

Keywords:*Risk Management, process management etc.*

Article History
Article Received: *3 January 2019*
Revised: *25 March 2019*
Accepted: *28 July 2019*
Publication: *22 November 2019*

I. Introduction

Classical disease surveillance represent the main part of any public health file over many decades. Disease surveillance is vastly recognized as one of the most significant tools to evaluate, forecast, and lessen infectious disease spreading. classical disease surveillance depends on data aggregated by health institutions, and the data typically contains the information such as morbidity and mortality data, laboratory reports, individual casereports, field investigations, surveys, and demographic data (Salathé, 2016).

TEST Engineering & Management

November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

The development of healthcare and the medicinal industry most importantly resulted in the availability of data. Therefore, this field is considered to be a rich environment for different researches carried out scientifically. It carries several differences when compared to other sectors, as it is linked to the life of human beings. As well, it has a relatively higher priority because it is supposed to be within a high level of health services with no regards towards it's of cost, and therefore consumes a bigger amount of the specified budget for each country. Pharmaceuticals is probably one of the more influential fields of healthcare. Its industry depends on quantitative analyses of clinical studies and drug marketing(Taylor, 2017).

Pharmaceutical companies could also take benefit in healthcare Customer Relationship Management (**CRM**) and data mining. Pharmaceutical companies can specify their marketing target through following up which practitioners give out prescriptions for what drugs and for what purposes. This shows the least expensive or most effective treatment plan for a certain illness, as well as its contribution in identifying the practitioners who have practices with specific clinical trials like those who have a high percentage of patients that belong to a particular classification), and picture the course of an epidemic with the aim of supporting pharmaceutical salespersons, physicians, and patients (Boström, 2011).

A new research area that found its place in machine learning is deep learning, dating back to 2006. With the notion of a large-scale training data set like ImageNet, the issue of over-fitting that accompanies neural network training is partially solved. Furthermore, the capacity of a GPU chip in computer hardware integrates over hundreds of nuclear that provide huge computing energy to provide for the neural network. This enables the training of large-scale neural network. Simultaneously, considerable progress has been made in terms of models designed and the methods applied in neural network training.Deep learning aims to place machine learning as close to its original aim— namely Artificial Intelligence (AI) as possible (Lecun, Bengio, & Hinton, 2015). Conversely, the techniques of deep learning are applied to a large dataset. In modern computers, managing larger set of data could reduce the probability of over-fitting and increases the ease of accomplishment of complex computation in neural networks.

Analyzing drug utilization depends on drug prescriptions as well as its commerce or distribution in hospitals, pharmacies, and practitioners, as it is based on databases, electronic records and exploratory questions (Company, 1990). This process is adopted with several purposes, among which is the utilization of track after-sales services for medicinal products by several (substituting) brands for the same medical component. Therefore, substituted brands have exactly the same characteristics as the generic drug (Fy, Science, Effective, Office, & Drugs, 2017). It also detects and determines the use of over-the-counter drugs (Zhou, Sani, Lee, & Luo, 2016). The historical analysis of drug use of a long time span could support the planning and commerce of medicinal items through the discovery of consumption patterns and the prediction of production quantities needed for the next periods (Ghousi, Mehrani, & Momeni, 2012).

II. Related Works

This section will review the most recent prediction methods for drug utilization, where some of these methods are statistical ones whereas others used a machine learning method. In (Papana, Folinas, & Fotiadis, 2016) the authors have forecasted the quantities consumed and purchases for only one drug. The drug is RAPILYSIN LYPDINJ 2X1.16G/VIAL (RL). These drugs are used in the management of acute myocardial infarction (AMI). The data used in this work is collected from one hospital for the period from 2009 to 2011. This work presents a method of univariate time series analysis and forecasting. A drug purchase model and a drug consumption model were built, and the researchers obtained the best prediction result provided that the mean square error equals ($4.3245x10^6$, $4.1437x10^7$). As for the prediction of the demanding drug presented in (Ramos, Cubillas, & Feito, 2015), the researchers relied on spatial data in addition to demographic data and consumption data to predict

November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

the required quantities of the drug. A model was developed to predict the required quantities of the drug Salbutamol based on data from 2010 to 2014. Generalized Linear Models, Support Vector Machines with linear kernel and Support Vector Machines with Gaussian kernel are used as a model for prediction. The algorithm that gave the lowest ratio of line is Generalized Linear Models. The authors used the mean absolute error to evaluate the three models, and the error was (5.78, 6.33, 17.07) for Generalized Linear Models, Support Vector Machines with linear kernel and Support Vector Machines with Gaussian kernel respectively.

Monitoring the spread of disease by building an approved GIS map was presented in (Photis, 2016). In order to monitor the spread of diseases in Greece, the researcher devoted his work to analyze and evaluate the Greek Public Health Care System during the 1997 - 2007 decade, specifically for the health care supply, demand and service location quotients. Traditional statistical methods were used to compute Health care demand, and supply GIS was also used for the purpose of spatial visualization through mapping. The colour density is used to identify the quantities of supply and demand services applied in the polygons map of the country, which eventually provided another view for the evaluation of the health care.However, The work depended on traditional statistical methods only, as the data set was not used to predict future healthcare requirements. In the article (Kim, Won, Park, & Kang, 2015) the authors provided a method for predicting the quantities of drugs required based on consumption data as well as social networking data and wireless sensor data. Only three drugs were predicted. The researchers used consumption data as an endogenous variable resulting from time series and exogenous variables that come from topics trendanalysis of the Facebook, Twitter, and bloggers. Two prediction models have been tested in this article: anautoregressive model with endogenous variables only and an autoregressive model with both endogenous andexogenous variables**.** The results showed that the prediction model has improved after using the topics trendobtained from social networking as exogenous variables. As for the prediction of the demanding drug presented in (Ramos, Cubillas, & Feito, 2015), the researchers relied on spatial data in addition to demographic data and consumption data to predict the required quantities of the drug. A model was developed to predict the required quantities of the drug Salbutamol based on data from 2010 to 2014. Generalized Linear Models, Support Vector Machines with linear kernel and Support Vector Machines with Gaussian kernel are used as models for prediction. The algorithm that gave the lowest ratio of the line is Generalized Linear Models. The authors used the mean absolute error to evaluate the three models, and the error was (5.78, 6.33, 17.07) for Generalized Linear Models, Support Vector Machines with linear kernel and Support Vector Machines with Gaussian kernel respectively.

III. Theortical background

Drugs are the patient's essential necessity in order to be healthful, being them among the materials adopted by health experts or physicians in preventing disease or in therapy. Consequently, hospitals and/or pharmacies should guarantee providing the necessary drug at any times so they can be specified to patients when needed. Health care institutions face a challenge in providing required drugs and good services. As they control the provision of drugs and avoid harmful effects, inventory or stocking control is the main factor to keep the organization competitiveness in progress (Pratyaksa, Permanasari, Fauziati, & Fitriana, 2016). Drug quantification is a procedure followed to define the required quantity of a product or drug for procurement. Apart from the identifying quantity demands of drugs or certain items, the quantification process also defines the financial funds needed in order to purchased it as well. To achieve perfect drug quantification, different information is needed such as Essential Medicines List (EML), consumption of the drug, number of prescriptions, the upper and lower limits of stock levels, the frequency of stock-outs and how long the procurement cycle is (Iqbal, Geer, & Dar, 2018).

Deep learning can be defined in several ways, one of which is that it is a category of machine

November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

learning techniques which takes advantage of several layers of non-linear data performance for either the supervised or unsupervised reduction of data, extracting of features, transformation, as well as discovering and classifying patterns. Another way of defining this concept is that it is "a branch of machine learning field which depends on algorithms for representing multi-levels in order to model intricate links among pieces of information. "It is a recent field of Machine Learning research, that aims to move Machine Learning closer to its original target: Artificial Intelligence. This concept is mainly concerned with learning multiple levels of representing and abstracting used to make sense of data such as an image, sound, and text."(Deng & Yu, 2015). From what is mentioned above, the two common characteristics of the deep learning definitions are (1) having a basis of multiple computational layers or stages used to identify the dataset in an abstractive way and model nonlinear information process; (2) being used with supervised and unsupervised algorithm learning to extract features from the dataset. The application of this concept has been found in various sectors like image classification, drug discovery, and speech recognition (Minar & Naher, 2018).

Spatial visualizations: Spatial visualizations are one of the most popular design alternatives of information visualization, and geographic information system (GIS) is a main origin of inspiration for spatial visualization. The visualization of datasets which have spatial or geographical attributes can usually enhance our comprehension to the datasets (Yeh, 2008).

IV. The Proposed Methodology

This section describes the design and implementation of the proposed methodology using deep learning techniques for the prediction and surveillance of drugs and diseases and the description of these processes.Figure (1)shows the design and improvement of the deep neural network and Spatial visualizations.

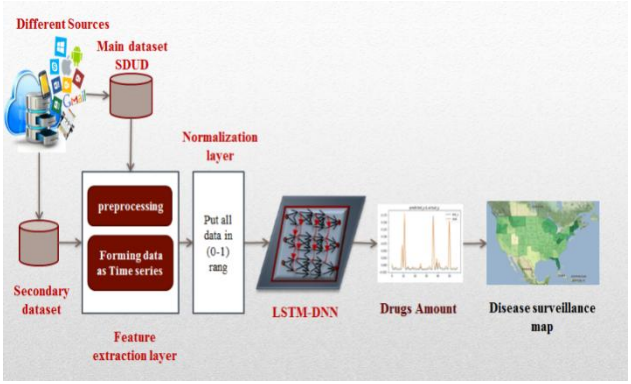

Figure (1) the proposed methodology

- **Data set**

In the USA, the data of the drug utilization is stored and organized as a dataset that is received from each state every year. State Drug Utilization Data (SDUD) are reported by states to cover outpatient drugs that are paid for by state Medicaid agencies since the start of the Medicaid Drug Rebate Program. The dataset includes many attributes which represent all information about the utilization of the drug. Datasets are available on Medicaid.gov: https://www.medicaid.gov. This data set represents the consumption data for a number of years and is constantly updated each year.

-Data Aggregation Stage

In the beginning of the work, the Data aggregation stage is applied. In this work, the data aggregation is a process in which information about disease (AIDS) and drugs brand name is gathered and expressed in a summary form. The purpose of aggregation is to get more information about AIDS disease based on its drugs. This stage depends on many websites specialized in the pharmacy field. The websites that are used in this work are www.drugs.com, www.medicen.com, and www.rxlist.com.The result from the data aggregation is the secondary dataset. The secondary dataset contains the list of drug brand names for the AIDS.

-features extraction layer

The preprocessing stage and forming data as time series stage. The work of these stages is (if- then layer, quarter collection, and year collection) to make all the feature extraction automatically. The

TEST Engineering & Management

November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

secondary dataset which resulting from the aggregation stage has been used to extract the drug brand name utilization information from the main dataset (SDUD). The input of the feature extraction layer is the raw data from the main dataset. The $x=(x_1,...x_n)$ is the input vector and $y=(y_1,...,y_m)$ is the output vector of the feature extraction layer.The aim of the deep neural network is to predict the future consumption of drugs. Since the prediction process depends heavily on the time data series, therefore, through the preprocessing layer, data will be formed in time series data. In this layer, the data is collected based on the quarter number. After collecting the data based on the quarter number, the data is then collected based on the year. The prediction process is related to two attributes: the number of prescriptions (P) and the number of units (U). The values of these attributes are grouped by season or quarter. Finally, the data is prepared to be in the acceptable form and passed to the next essential layer (Normalization layer).

Table (1) Example Of The Result feature extraction Layer

Drug	Year	Q1		Q2		Q3		Q4	
		P	U	P	U	P	U	P	U
Drug1	Year1	4844	328348	5214	357275	5554	381060	6198	419223
	Year2	6709	451579	7780	517104	8383	558720	6531	378085
	Year3	6516	378075	6973	402860	7525	436082	8750	508824
	Year4	108	6787	183	12149	123	8940	420	2675
	Year5	530	30090	487	28660	453	26438	489	28614
Drug2	Year1	50603	920025	53970	985180	53857	980567	61802	1137375
	Year2	68550	1263588	82698	1499839	93464	1695580	90602	1584597
	Year3	106058	1931853	112601	2046167	112050	2028865	117432	2095794
	Year4	14259	219687	2256	41340	16130	250117	16678	256141
	Year5	45446	799134	55871	1010911	51076	916817	63318	1074069

-Normalization layer

The Normalization layer aims to rescale real-valued numeric attributes within the range of 0 and 1 to be the input of the next layer (prediction layer). So before the data is entered to the network, all the data of attributes (the number of prescriptions (P) and the number of units (U)) must have the same scale such as between 0 and 1. This layer receives the output of the preprocessing layer which represents the vector that consists of eight inputs. Each drug has been represented by eight values; the number of prescriptions and required units for four quarters. According to the normalization layer, the input drugs vectors were scaled into the range of 0 and 1.The normalization layer is applying this formula.

$$mean_{normalized\ (x)} = \frac{x - mean(X)}{\max(X) - \min(X)} \qquad (2)$$

- **LSTM-DNN**: LSTM is a supervised network, a special kind of recurrent neural network, and has some internal contextual state cells that act as long-term or short-term memory cells. This network is used and trained for the purpose of predicting the quantities of drugs required for the next year. The output of the LSTM network is updated by the state of the memory cells. This property is very important when the prediction is the aim of the neural network in order to depend on the historical context of inputs, instead of the very last input only. Figure 2 shows the LSTM cell architecture.

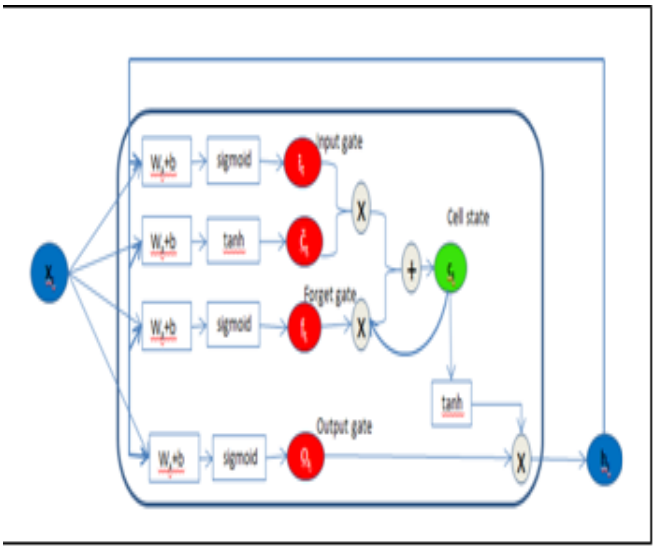

Figure 2 The LSTM Cell Architecture

A memory cell contains four units: the forget gate, the output gate, the input gate and a self-recurrent neuron. There is an internal state variable which is passed from one cell to another and modified by operation gates. The following algorithm represent the LSTM-DNN.

TEST Engineering & Management

November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

Algorithm IIofLSTM layer

D represents the number of memory (LSTM block)
S_j represents number of cell in bloch j
Input: $x \leftarrow [x_1, \dots, x_n], x_t \in \mathbb{R}^m$
Output : $h \leftarrow [h_1, \dots, h_n], h_t \in \mathbb{R}^p$

Given parameters $W_{fjm}, b_{fjm}, W_{cjm}, b_{cjm}, W_{ijm}, b_{ijm}, W_{ojm}, b_{ojm}$
Intilized $h_0, c_0 = \vec{0}$ *of length* p
for $\forall\, j \in D$ *do*
  for $\forall\, v \in S_j$ *do*
    $f_{tj} \leftarrow \sigma(W_{fjm} . [h_{t-1}, x_t] + b_f)$
    $C_{tj} \leftarrow \tanh(W_{cjm} . [h_{t-1}, x_t] + b_c)$
    $i_{tj} \leftarrow \sigma(W_{ijm} . [h_{t-1}, x_t] + b_i)$
    $C_{tj} \leftarrow f_{tj} * C_{tj-1} + i_{tj} * C_{tj}$ //Updat cell state//
    $O_{tj} \leftarrow \sigma(W_{ojm} . [h_{tj-1}, x_{tj}] + b_o)$//calculate Output
    $h_{tj} \leftarrow O_{tj} * \tanh C_{tj}$
    End for
  End for
End for
$e_k(t) \leftarrow t_k(t) - h_t(t)$
// Backward pass in LSTM //
for $\forall\, j \in D$ *do*
  for $\forall\, v \in S_j$ *do*
    $\Delta w_{ojm}(t) \leftarrow \propto \delta_{oj}(t) - x_m(t)$ \\ weight update of output gate\\
    $\Delta w_{ijm}(t) \leftarrow \alpha \sum_{v=1}^{S_j} e_{s_{c_j^v}}(t) \frac{\partial S_{c_j^v}(t)}{\partial w_{ijm}}$ \\ weight update of input gate\\
    $\Delta w_{fjm}(t) \leftarrow \propto \sum_{v=1}^{S_j} e_{s_{c_j^v}}(t) \frac{\partial S_{c_j^v}(t)}{\partial w_{fjm}}$ \\ weight update of forgate gate\\
    for $\forall\, v \in S_j$ *do*
      $\Delta w_{cjm}(t) \leftarrow \propto \sum_{v=1}^{S_j} e_{s_{c_j^v}}(t) \frac{\partial S_{c_j^v}(t)}{\partial w_{cjm}}$ \\ weight update of state cell\\
    End for
  End for
End for
Return $h \leftarrow [h_1, \dots, h_n], h_t \in \mathbb{R}^p$
End- algorithm

- **Evaluation**

MSE is the evaluation metric used to evaluate the prediction model performance. MSE (Mean Squared Error) demonstrates the extent to which the original is different from the predicted values that are estimated squaring the value for the average difference over the data set as a whole.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \qquad (1)$$

1- **Experimental results and discussion**

This section presents the results of the working stages of the proposed techniques as well as clarifies the comparative results on the dataset that are used in this study. The number of the AIDS's drugs are (41). The information of the drugs brand names extracted from the main dataset from 2012 to 2017 and the training data sperated as 80% training set, 10% validation set and 10% as test set. The data from 2012 to 2016 used as training data used to predict the quantities of drugs for 2017. ***The number of LSTM cells***; in the hidden layer, the empirical test is applying to choose the optimal number of LSTM cell; in our case the range between (5-20) cells. Table (1) shows the result of an empirical test, the best RMSE reached is (0.009) when the number of LSTM cell is (15) cell.

- **The Disease Surveillance Map**

A disease surveillance map is created based on the results of a deep neural network, where data for AIDS disease drugs are collected and linked to the spatial data provided by the dataset file DUSD. Through this map, it is possible to monitor the spread of the disease in all states. It provides information about the state where the disease is concentrated and can be an indicator of infection or wrong living habits. The Choropleth map is used to create a disease surveillance map. The required objects to create this map are the USA map used as a basemap (first layer), the loaded GEOjson file that contains the states polygon shapes as the second layer, and the third layer is the colour density that ranges fromtransparent to bold colour to represent the spread of the disease.

Table 4.6 Listed the Number of Cells and the RMSE

Number of cells	RMSE
5	0.010
10	0.015
15	**0.007**
20	0.012

Figure (3) shows the error decay with the number of cells where it reaches 0.015 with the number of cells 10.

November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

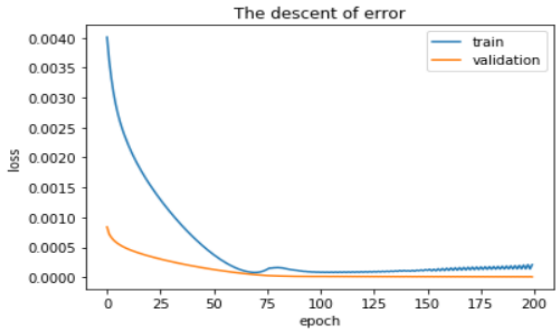

Figure (3) Loss Function of Training Set and Validation Set

Figure (4) shows the error decay with batch size where it reaches 0.007 with the number of cells 15.

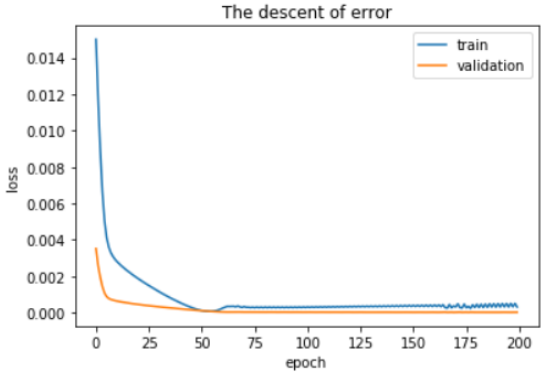

Figure (4) Loss Function of Training Set and Validation Set

While Figure (5) shows the test set and prediction results of the network of (2017) the prediction of the required quantities of drugs.The true convergence between the real data and the prediction result can be observed because the RMSE value is small (0.007).

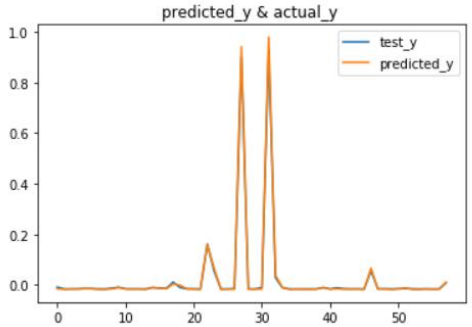

Figure (5) Prediction of Network and Test Set of 2017

Whereas when the Linear Regression method used on the data of the drug based on the five-year, the RMSE value that obtained is 0.094. Figure (6) shows the prediction resul by linear regression.

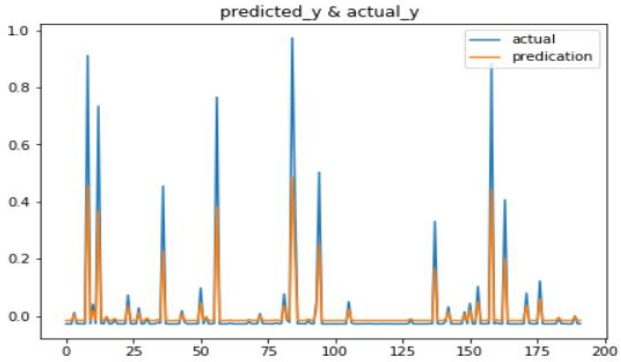

Figure (6) prediction result by linear regression

The goal of spatial visualization is to create maps to surveillance of the spread of diseases on the ground. This stage has employed the results from the previous stageprediction were used in addition to the spatial data in the data set SDUD. Two types of maps are created at this stage. The first is the disease surveillance map and the second is the drug consumption monitoring map.Figure (7) shows the ADIS surveillance map.

Figure (7) ADIS Surveillance Map

Figure (8) illustrates the monitoring drug map of the 'ISENTRESS', it is an ADIS drug using in the USA at 2017.

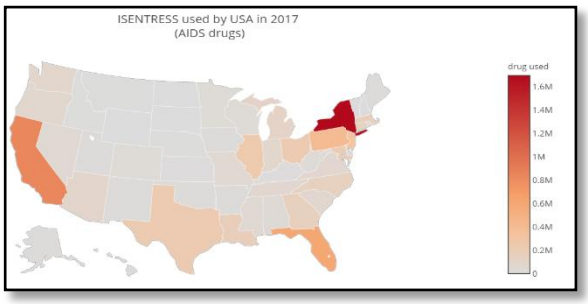

Figure (8) 'ISENTRESS' Monitoring Map

November-December 2019
ISSN: 0193-4120 Page No. 282 - 290

2- CONCLUSION

The pharmacy field is important nowadays because of its rapid development. The production and supply of drugs requires large amounts of money and because of the validity of the use of the drug within a limited period of time, the process of predicting the quantities to be provided in health institutions has become an urgent necessity. The availability of large data on the quantities spent for several years greatly facilitated the forecasting process. Therefore, the proposed methodology used the available data for the purpose of analysis and the most important characteristics that were discovered when implementing the proposed methodology through the results obtained are as follows:

-The aggregation step is applied at the beginning of this work, it provided important information about the disease and the drug brand name. This step affected the work positively, putting it in the right direction.

- The proposed method has provided the efficiency to predict the future needs of AIDS's drugs according to:
  - The proposed feature extraction layer has demonstrated its ability to extract useful features from the raw data.
  -The proposed deep neural network proved effective in feature extraction and prediction of the quantities of drugs, where the performance is enhanced by the network is fully automated.
- Spatial visualization: at this stage, a surveillance map for the spread of disease was established based on the consumption of drugs and a map showing the consumption of quantities of drugs by name of the drug brand. This stage provides a different vision for analyzing the data through which the quantities of the drugs started were linked to the places where these drugs dispensed.

REFERENCES

[1]. Boström, K. (2011). *Consumer behaviour of pharmacy customers Choice of pharmacy and over-the-counter medicines*.

[2]. Company, C. (1990). *Chapter 5 Application of Data Mining Techniques in Data Mining Needs of Global*. 138–180.

[3]. Deng, L., & Yu, D. (2015). *DEEP LEARNING: METHODS AND APPLICATIONS* (Journal of; W. 98052 Redmond, Ed.).

[4]. Fy, O. G. D., Science, R., Effective, I., Office, T., & Drugs, G. (2017). *FYs 2013 - 2017 Regulatory Science Report : Analysis of Generic Drug Utilization and Substitution*. (Figure 1), 1–10.

[5]. Ghousi, R., Mehrani, S., & Momeni, M. (2012). Application of Data Mining Techniques in Drug Consumption Forecasting to Help Pharmaceutical Industry Production Planning. *Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management*, 1162–1167.

[6]. Iqbal, M. J., Geer, M. I., & Dar, P. A. (2018). Evaluation of Medicines Forecasting and Quantification Practices in Various Evaluation of Medicines Forecasting and Quantification Practices in Various Public Sector Hospitals Using Indicator Based Assessment Tool. *Journal of Applied Pharmaceutical Science*, *7 (12)*(December 2017), 072–076. https://doi.org/10.7324/JAPS.2017.71209

[7]. Kim, W., Won, J. H., Park, S., & Kang, J. (2015). Demand Forecasting Models for Medicines through Wireless Sensor Networks Data and Topic Trend Analysis. *International Journal of Distributed Sensor Networks*, *2015*. https://doi.org/10.1155/2015/907169

[8]. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

[9]. Minar, M. R., & Naher, J. (2018). *Recent Advances in Deep Learning: An Overview*. *2006*, 1–31. https://doi.org/10.13140/RG.2.2.24831.10403

[10]. Papana, A., Folinas, D., & Fotiadis, A. (2016). Forecasting the consumption and the purchase of a drug. *INTERNATIONAL CONFERENCE ON SUPPLY CHAINS Functional*, *2*.

[11]. Photis, Y. N. (2016). Disease and Health Care Geographies : Mapping Trends and Patterns in a GIS Abstract. *IMedPub Journals*,

10(3:8), 1–8.

[12]. Pratyaksa, H., Permanasari, A. E., Fauziati, S., & Fitriana, I. (2016). ARIMA Implementation to Predict the Amount of Antiseptic Medicine Usage in Veterinary Hospital. *IBioMed*, (October), 5–6.

[13]. Ramos, M. I., Cubillas, J. J., & Feito, F. R. (2015). Improvement of the Prediction of Drugs Demand Using Spatial Data Mining Tools. *Springer Science+Business Media New York*, (November). https://doi.org/10.1007/s10916-015-0379-z

[14]. Salathé, M. (2016). Digital pharmacovigilance and disease surveillance: Combining traditional and big-data systems for better public health. *Journal of Infectious Diseases*, *214*(Suppl 4), S399–S403. https://doi.org/10.1093/infdis/jiw281

[15]. Taylor, D. (2017). *The Pharmaceutical Industry and the Future of Drug Development* (Pharmaceut; R. E. . Hester; & R. M. Harrison, Eds.). https://doi.org/10.1039/9781782622345-00001

[16]. Yeh, R. K. (2008). Visualization Techniques for Data Mining in Business Context : A Comparative Analysis. *International Journal of Engineering & Technology*, (2003), 310–320.

[17]. Zadeh, N. K., Sepehri, M. M., & Farvaresh, H. (2014). Intelligent Sales Prediction for Pharmaceutical Distribution Companies : A Data Mining Based Approach. *Hindawi Publishing Corporation*, *2014*.

[18]. Zhou, Y., Sani, N., Lee, C.-K., & Luo, J. (2016). Understanding Illicit Drug Use Behaviors by Mining Social Media. *2016 IEEE International Conference on Big Data (Big Data)*. Retrieved from https://arxiv.org/ftp/arxiv/papers/1604/1604.07096.pdf