Demand forecasting model for time-series pharmaceutical data utilizing shallow and deep neural network architectures

Shaik Salma Begum[1], Anusha S[2], Tejaswini TN[3], K Sai Krupa[4]
Department of Computer Science and Engineering, Presidency University

Abstract: Request determining may be a logical deliberate appraisal. This evaluation gages future needs for a key item. The show, the Request Figure Show (DFM) is key. It makes a difference pharmaceutical companies to find victory. They can do this within the worldwide showcase.

The inquire about paper's objective is to affirm different shallow and profound neural network methods. These strategies are utilized for foreseeing request. At that point suggestions are made. These suggestions are for deals and showcasing methodologies. They are based on slant and regular effects.

Eight item bunches are looked at. The items have diverse characteristics. Pharmaceutical items are in these bunches. For each bunch, a procedure is proposed. The prescient accuracy is calculated using root cruel squaredserror (RMSE) for DFMs.

Findings were made in this consider. The cruel RMSE esteem was 6.27. This was for neural network-based DFMs. These were for all medicate categories. This esteem was lower in shallow neural systems than in profound ones. Based on these findings, DFMs have value. They can precisely foresee future request. This is often particularly genuine for pharmaceutical items.

Keywords: Deep learning models, Demand forecasting Pharmaceuticalindustry Shallow neural network models

I. INTRODUCTION

The request estimating show gets to be fundamental. Apparatus is utilized by numerous worldwide pharmaceutical enterprises. It's got a double objective. Supply Ought to be coordinated. It ought to adjust with any rise or drop that will happen in request for the items. Too ,the stock ought to be as negligible as conceivable .On the off chance that that framework is deployed, it contributes to an productive supply chain administration. Within the conclusion, an advancement in in general benefit for the organization can happen .The structure Of the supply chain for pharmaceutical companies is complex .Now and then, certain parts are interlaced. These complexities ended up portion of various challenges. It's a challenge confronted by pharmaceutical companies all inclusive. The productive operation of supply chain administration is basic .It is important for any organization In essential service sector.

This is especially true for pharmaceutical industry. We've the current covid-19 pandemic to consider. Once thought as an Epidemic confined To China ,it has now become an Unplanned worldwide pandemic. This pandemic is One of A kind situation. Not experienced for last 100 years .It has Impacted supply and demand sides of even strong and well-known global pharmaceutical industry.

Sudden Outbreaks on one side have led to waves of demand increases While on the other hand lockdowns ,vaccine discoveries pill discoveries and medicine discoveries have led to unpredictability in supply of pharmaceutical products

Request Estimating Models (DFMs) are a Valuable apparatus for companies.. They offer assistance To stock or create the fundamental amount of pharmaceutical items. .Productive DFMs givesGreat bolster to pharmaceutical companies.. They can foresee request and make Suitable choices. Particularly in their generation and supply chain operations.

To be brief, highly effective DFMs can serve two purposes. They help surrmount the Complexity in the supply chain. They can also create a competitive edge .This edge is particularly useful for global pharmaceutical companies The vast majority Of pharmaceutical companies use machine learningbased models .They use these models to forecast the future demand Before it Actually arrives. The clear advantage of artificial neuralnetworks

Or data mining techniqques is the capacity They have the ability to produce high-quality predictions. Also ,they take lesser time to complete. The use of existing standard statistical Forecasting techniques In contrast, is unable to match.

It is virtually Impossible to achieve such Accuracy with The standard methods. In an article ,the authors stated that companies had a hard time delivering The right products .These companies couldn't timely provide a suitable amount Of products to their suppliers. The predominant cause For this was the companies' failure to recognize the quick changes in the level of demand. These companies also lacked awareness of Other competitors' demand for the supply Chain management.

Hence the writers Advised these bodies to employ a predictives tool. This tool should be used to Manage The demand forecasting and associated functions .This tool is beneficial since it can help companies to stay on Top of rapid demand changes And competitors' strategies.

Directly, the first well-known And prospering linear deciding time series models. Theysare Autoregresssive Integrated Moving Typical (ARIMA), Standard Autoregressive Arranges Moving Typical (SARIMA) and Autoregressive Moving Typical (ARMA). Be that because it may ,the estimate quality persevered Due to their assumption of clear exercises .Hence, this speeded up the change of machine learning methodologies.

We have information mining procedures and Fake Neural Systems (ANN) .They Are utilized for the time-series information estimating. Their tall forecast exactness has made these strategies exceptionally well known.

Growth of the ANN-based time-series forecasting model has been considerable. Also ,it has Opened the door for making accurate predictions .This is achieved by managing nonlinear input and output data. Linear forecasting methods are not ideal for modeling data with non-linear behavior. The above linear Forecast techniques are not well-suited For modeling data. This is especially the case if the data contains non-linear behavior.

There is A significant rise in The use of Ann for time-series forecasting. Its use increases the quality of prediction. several studies have suggested investigating the performance of ANN-based time-series forecasting. They suggest comparing these with common Forecasting methods. The studies show that there is no benchmark setting for the neural network .This Is relevant to The timeseries domain.

Parameter Setting is Purely dependent on the problem domain In this research, we utilize a variety of neural networks .These Include Probabilistic Neural Network (P_NN) ,Generalized Regression Neural Network GR_NN And Radial Basis Function Neural Network (RBF_NN) in addition to deep Learning Neural network. The deep Learning neural network used in this paper is Long-Term Short Memory Neural Network.

A comparative study is Presented in this paper .It is used to determine whether a shallow neural networkbased forecasting model is better We aim to determine if the model performs Better than the deep neural network This study is for time series pharmaceutical data.

In short, the paper's main contribution is summarized below.

• Developing a methodology for selection of ideal demand forecasting models is key. These models get used for pharmaceutical data . They use statistical and neural network models for Sales database. It’s a timeseries database.

• We need to investigate whether Different neural networks are effective .Special focus is on RBF_NN, P_NN, GR_NN ,LSTM and Stacked LSTM. These show more accurate ATC Thematic drug predictions. We compare these with non-neural network methods.

• Performance of ARIMA model Needs investigation. Shallow Neural Network and Deep Neural Network models also need to be examined .We'll look at their performance on time series dataset.

We split this research paper into 5 parts .These are Section 1 through Section 5.

Section 1 aims To Provide information on study's introduction. Section 2 focuses on review of related literature .The study fills research gaps.

Section 3 Gives short explanation on time-series forecasting methods. Section 4 presents the methodology proposed .It also includes experimental analysis.

At last We have Sect . 5. It sums up the present work. Sect. 5 also mentions possible future work.

II. REVIEW OF LITERATURE

In broad sense demand forecasting is major hurdle for big data application. It might use additional data context .Take For example value and veracity. It's For better prediction. The primary Focus For This research is demand forecasting In big data analytics .The setup discusses Integration of diverse data sources into Possible prediction process. However it will need data analyst to fulfil the requirement. The task also calls for adequate technical know-how. Financial resources are necessary too .They boost the technology.

Existing Literature clearly indicates insufficiency of statistical ways for optimization Of pharmaceutical assembly.. Likewise ,Production facility Faces a similar scenario. .Some researchers keep a constraint in their mind.. They have taken an improvised approach .The approach permits combination of simulation methods with data analysis. One key recommendation was that pharmaceutical unit should be optimized. The purpose of such optimization is performance enhancement.

In a separate study ,a Number of methods were used To predict energy use .These methods include Convolutional Neural Network (CNN) method, NN-based Genetic Algorithm, and NN-based Particle Swarm Optimization (PSO) methods. The results of simulation were clear. They showed That neural network approach outperformed the CNN approach for energy consumption prediction.

Authors have designed programs to predict short-run irrigation demand .These programs are based on Specific algorithms. Some of the algorithms are Genetic Algorithms (GA) ,Bayesian structure And Reinforcing Artificial Neural Networks (RANN). The programs have undergone experiments. It was discovered That the Artificial Neural Network model produced more precise results. It was more accurate compared to traditional algorithmic methods.

A different study examined Iran's yearly energy consumption. It utilized three variations of ARIMA model. These were combined with ANFIS model .The Research revealed that hybrids are more accurate than Individual models like Arima or ANFIS model. This was observed for the prediction of energy use.

Numerous researchers have tried generating close-to-perfect ,error-free models for forecasting. Curiously, a neuro-fuzzy approach surfaced in one study. It Was used To predict future demand .They employed past Sales data. ANFIS was put into use .Another feature was the use of ANN. This doubled as a 'neurofuzzy approach'. Despite its specificity, the Method shows its effectiveness.

State-of-the-art methods were studied. Also important forecasting topics regarding drug company demands were addressed. A notable study went over these aspects. They used three models For demand projections .They Finished that 'symbolic regression-based forecasting model' was the correct fit. Error rate was also observed to be less .It was noted to be less when compared to other models.

Some attempts were made to forecast Revenue streams for pharmaceutical companies. .One study focused on presenting a new approach for forecasting This looked at pharmaceutical companies' revenue estimates It implemented a technique using NN Specifically ,it used prediction methods involving time series datasets..

Because of COVID-19 pandemic ,there was a noticeable impact on supply chain functions. There Was a resulting short supply for ambulatory medicines .The pandemic had an interesting effect .It added to the demand for pharmaceutical products . It pointed out the reported shortages early In the pandemic. The Primary goal of this study was to find most efficient animal drugs. It was based on Drug distribution paths similar to those used for human$. This creates needs for investment in efficiency of Drug delivery systems. Also in health and feasibility of these delivery systems. The testing is critical. LSTM model was utilized in this study. The aim Was To improve predictive quality .The specific dataset used was timeseries dataset. This research proposed a demand forecasting strategy. It used multi-layer LSTM model. The model has a strong capability. It can predict exceptionally high or Low demand results .The researchers compared the LSTM model With CNN model. LSTM Model stood out .It has a more potent predictive performance.

Fuzzy Neural Network (FNN) was used in the theoretical development of forecasting methods in a separate study. FNN can Remove insignificant weights .This is Used to construct Fuzzy IF–THEN rules. These rules are used for additional learning. The results of FNN and ANN are combined. This is used to predict time-series data with more accuracy. Literature proposed various forecasting techniques .These included statistical and computational intelligence approaches. Among them are exponential smoothing ,KNN SVM, ARIMA RNN, And LSTM

Cognitive Architectures for forecasting methods were investigated in another research. Fuzzy Neural Networks (FNNs) and Artificial Neural Networks (ANNs) were among the cognitive architectures employed. The Fnns had a distinct advantage. They were able to delete insignificant weights. This enabled the construction of Fuzzy IF–THEN rules. These rules could be used to further learning .FNNs' Output was amalgamated with ANNs' results .This combination Was used for superior prediction of time-series data.

Various forecasting techniques are suggested In literature. These include: exponential smoothing, K-nearest Neighbors (KNN) ,Supporting Vector Machines (SVM), ARIMA, ANN, RNN and LSTM .They are considered for A Myriad of problem domains. The literature Points at statistical ,computational intelligence approaches for the same. The written work suggests the techniques for forecasting.

A method is used in the study .This method is for the development Of forecasting techniques. The method is a fuzzy neural network (FNN). Training FNN is beneficial. It discards the insignificant weights .Fnn uses them to develop Fuzzy IF-THEN rules. These rules are for further learning. The FNN's output is then pooled with ANN results. This combination boosts prediction Of time-series data.

Within the factual and computational insights approaches ,strategies exist. These strategies incorporate: exponential smoothing, K-nearest Neighbors (KNN), Supporting Vector Machines (SVM) ,ARIMA, ANN, RNN and LSTM. Writing has proposed them as determining procedures .They Aresput forward for a assortment Of issue spaces.

The study presented a new forecast technique. The Researchers applied it To dissect current sales data. They compared these findings to various machine learning models .The Goal was To ascertain future demands. Demand forecasting method revolved around LSTM. It grappled with e-commerce business' nonlinearities. Results Indicated LSTM's superiority over predictive univariate techniques . The study also proposed a hybrid approach.

Hybrid approach scrutinized lengthy memory time series. It incorporated Arfima and FFNN Authors stated that hybrid model outperformed standard approaches. Table 1 provides list of neuralnetwork-based predictive models. These models Are used for various demand forecasting operations. The demand forecasting task is extensive It involves detailed Examination of strategies .This included methods ,techniques and methodologies developed .These were also implemented in prior studies. The chosen model is the ANN-based demand forecasting model.

Model was deemed appropriate for time series dataset. This was after careful examination Of many methods and techniques. The neural Network utilizes Artificial neurons .These are used in complex deep learning networks. In this research ,we will develop ANN-based demand forecasting models. The models will Be for an Anatomical Therapeutic Chemical dataset. It is thematic for drug sales .The aim is to predict weekly demand accurately.

III. METHODS AND MATERIALS

Autoregressive Integrated Moving Average Model (ARIMA) is a Widely used forecasting model .It is lauded for its simplicity. ARIMA is also admired for its ability to generalize to non-stationary series. Detailed description of this model can be found. It can be discovered in [14, 23 ,35]. Here is a brief introduction to ARIMA.

3.1 Shallow neural network model: SNN

There are popular SNN models used for time-series data forecasting .Some Of them are ‘‘Radial Basis FunctionRBF) Neural Network (RBF_NN’’ and "Generalized Regression Neural Network (GR_NN)". Another popular model Is Probabilistic Neural Network (P_NN)". These models are used to build demand forecasting models for pharmaceutical products.

The SNN can be made up of an input layer (IL). .It can also have 1 or 2 hidden layers (HL).. The SNN has an output layer (OL) too.. The task it performs is regression. The target variable of the problem is continuous.

The Typical network configuration Diagram for SNN is shown in Fig .1. The Il has one Neuron for each predictor variable. The predictor variable are X1 ,X2 and so on. In HL the predictor variables values and its target values for each category are stored in each neuron. These categories are denoted as H1, H2 and so on.

When the input information are given to The HL it calculates the separated From the neurons brutal point for each test category. It vocations a appeared apportioned degree for this errand .The empty respect is given as the input to the RBF of HL. The resign of The HL is given to the Taking after layer pattern/summation layer or resign layer) for center of the road abandon

The decision Layer of P_NN/ GR_NN or output layer of RBF_NN is the last layer . This layer calculates the final output of the whole NN. The detailed descriptions of these neural Networks can be found below in Table 2.

3.2 The description of these shallow neural network models is provided below.

3.2.1 RBF_NN

The input layer Of RBF_NN receives the input data. Then it feeds it into hidden layer Of RBF_NN. The calculation takes place within the hidden layer. It uses the radial basis kernel function .The prediction task is performed in The output layer.

3.2.2 P_NN

The input layer of P_NN does computation. It computes euclidean distance from input-time series data. It does this to Training Input time-series data .The output is a vector. The vector's components tell How close the input is to Training input.

A second layer is there to compile the contributions for Each class of inputs.. There is also a denominator summation unit. .The result is used as predicted target value.. Lastly ,It handles task of prediction.

3.3 Deep Neural Network Model: LSTM

LSTM is made up of An input layer. It Consists of more hidden layers and an output layer .It is known as Recurrent Neural Networks (RNN). RNN has a deep learning architecture. This architecture holds the LSTM units in hidden layers. One major advantage of RNNs is this .They can be used for classification And regression tasks. The data used for this Can Be historical time points. The information can also be data.

A Single IL is part of the network .This IL leads into several HLs. It Also has a Single OL. The neurons found in the hidden layers are fully connected This includes memory cells It includes related gate units as well In the Figure 2 general architecture diagram of LSTM is shown.

An LSTM neural network has a cell .It has an input gate. It also has an output gate and a Forgot gate. The cell Remembers values .It creates its Net output as a vector of probabilities. This is done using the probability density function. The output layer selects the maximum probability neuron. It generates a 1. It is For the target classes .A 0 is for non-target classes.

3.2.3 GR_NN

The input layer of GRNN receives time series data It feeds the data into hidden Layer of GRNN .The calculation taking place within the hidden layer is using the radial Basis kernel function.

In summationpattern layer the denominator summation unit adds up the weight values of the hidden neurons.. The Numerator summation unit also adds up the weight values`. They are multiplied by actual target value for each hidden neuron. .The decision layer then divides the value. It accumulates in Numerator Summation unit It is Divided By the Value in the intervals

The input gate manages information flow into the cell . The Output gate manages flow of info out of the cell Forget gates manage the necessary or unnecessary information These functions are useful for predictions and classification

3.4 Performance metrics

The Root Mean Squared Error (RMSE) is A formula. It defines As square root. It is the square root Of sum of squared differences between Predicted sales values and actual sales values. We can see RMSE equated in Eq .5.

A smaller Rmse value Is better .It signifies the best model. There's one more Performance metric. It's called Normalized RMSE. This figure is calculated as RMSE of dataset divided by mean of dataset.

Finally ,we have the percentage of error (PE). Sum of the differences between forecasted sales values and the actual sales values Defines this .Then it is divided by sum of forecast sales values. It's multiplied By 100.

IV. PROPOSED METHODOLOGY

In the present article ,ARIMA is seen as central to our work. Neural network variants such as GR_NN, RBF_NN and P_NN are other models we use. LSTM and Stacked LSTM models are Also used for building DFMs. These are all constructed for pharmaceutical time-series sales data. The Graphical abstract of the proposed methodology is depicted in Fig .3 .Their performance has been measured. We used key performance indicators like RMSE and PE for comparisons.

Algorithm 1 shows the sequences Of the step .These are Used In the Neural Network-based Prediction Model for Demand Forecasting. The first step is initialization of model parameters and hyperparameters. This is For RBF_NN GR_NN ,P_NN, and LSTM models. The parameter of neural network model such as weights is initialized. Biases are also initialized in the parameters for the initial run of these models . Other hyper-parameters are initialized Randomly as well.

In this work dataset split Ratio is 70%:15%:15. It is maintained with 10 Kfold cro$s-validation. The term epoch represents the maximum number Of iterations .The objective of run is To minimize RMSE value of these models. The second step is To fix value for threshold (h). It denotes the RMSE value of the model .This is purely dependent on the dataset and userdefined.

Thirdly simulate the NN_net model with initial values.. Use the training dataset too.. After that evaluate performance.. If desired performance is reached then Stop the process .Stop the Training process .If otherwise ,repeat the training process. Use Different values of hyperparameters. These Are specified in Table 3. Termination criterion needs to be reached then.

Fourth stage is to build the NN_net model.. Use optimal Parameter values.. Next step Is to test NN_net model.. Use the testing data . After That estimate testing performance. Of the NN_net model That is. Following step Is to identify the best model .This is from the list of NN_models. A benchmark performance indicator is used .It is called RMSE.

Subsequently simulate best model. Use optimum neural Network configuration parameter values. This is for epidemiological data prediction .Finally Is to forecast values. These are for the subsequent ‘n’ time stamp.

4.1 Data description

We gathered weekly sales data. It is For pharmaceutical goods .The data was from the Kaggle website .The database included 600,000 Sales instances. These instances were for 57 drugs .It was during the period from 2014 to 2019. The database included attributes like date and time of sale. Brand name and amount of drugs were also included. In our research the demand for specific drug types was predicted. This was instead of predicting specific products. Therefore we classified the group of 57 drugs. They were organized In eight sections of anatomical treatment chemical (ATC) thematic analysis .The classification layout can be Seen in Fig. 4.

Sales Trends for all drug datasets categories are Presented in graphical form .Clear visualization is given in Fig. 5. This represents the Nature of the dataset used in our study. The actual dataset was used For implementing both shallow and deep neural network models. These models were used to predict demand.

Table 4 Shows the details of hyperparameters of demand forecasting models .We worked With models like ARIMA and RBF_NN. GR_NN and P_NN were also studied. Lastly we researched LSTM and stacked LSTM models as well.

Tables 5 and 6 illustrate RMSE value of the ARIMA model. .They also Show the Shallow neural network-based forecast model.. Additionally ,the deep neural network-based forecast model is presented..

Table 5 indicates all three shallow neural network models have similar values .These values are for all eight drug categories. The red highlighted values represent The lowest RMSE value. GR_NN-based DFM Appears to be more effective. It Works better for most drug types. The RMSE per drug is reduced by 5 to 23 %. ARIMA model for all drug categories is 15.02 .The overall percentage of prediction accuracy is 83.43%. This is For all drug categories using ARIMA model. If the model prediction accuracy is near 100% then it is considered the best model .In real-time datasets ,90% forecast accuracy is desirable.

Table 6 shows little difference in RMSE of LSTM Similar pattern was observed with Stacked LSTM for all drug types. For remaining RMSE value there was slight difference .Again values highlighted In red indicate lowest RMSE for Deep NN-based DFM. For M01AE and RO6 drugs Deep_NN-based DFMs worked better. This was in comparison To ShallowNN-based DFMS. Shallow_NN-based DFMS showed best performance for rest of drugs .Predictable accuracy levels above 90% were seen In these best models This is an inter`esting finding in current research.

LSTM And stacked LSTM showed almost Equal efficacy for all drug types .This Is except the M01AE drug. Another Interesting fact was shallow NN-based DFM showed better performance than deep NN-based DFM. It is because of the strong ability of shallow NN to tolerate noise and learn patterns from small databases. Lastly it is easy to create a model With Fewer parameters. Table 7 showed the PE values of shallow neural network-based demand.

This was against the deep neural network-based demand forecast model The PE values for The best model were shown. The range was from 0.2% To 13.57 .Therefore ,these DFMs could accurately predict ATC) thematic drug datasets.

DFMs' efficacy was tested with eight drug types. We used a benchmark performance indicator. Its name is RMSE .The performance was compared for these drug types. This comparison was shown graphically .You can see it in Fig .6. A comparison of minimum RMSE values Is shown in Figure 7. This comparison involves all Eight Drug types. Shallow NN and deep NN-based demand forecasting models were used. The findings are intriguing. They suggest The performance of shallow NN-based model was superior. This leads To The conclusion .This Shallow NN-based demand forecasting model is best for predicting demand for pharmaceutical products.

The performance of DFMs with eight drug types has been analyzed.. We used A benchmark performance indicator known as RMSE value Fig 6 features this graphically. .Figure 7 displays a comparison of minimum RMSE values.. These values are for all eight drug types .Shallow NN and Deep NN-based Demand forecasting models were used. The Results suggest an interesting pattern. It shows that demand forecasting model using Shallow Nn performed better. Therefore it can be concluded that the shallow NN-based demand forecasting model Worked best In predicting demand for pharmaceutical products

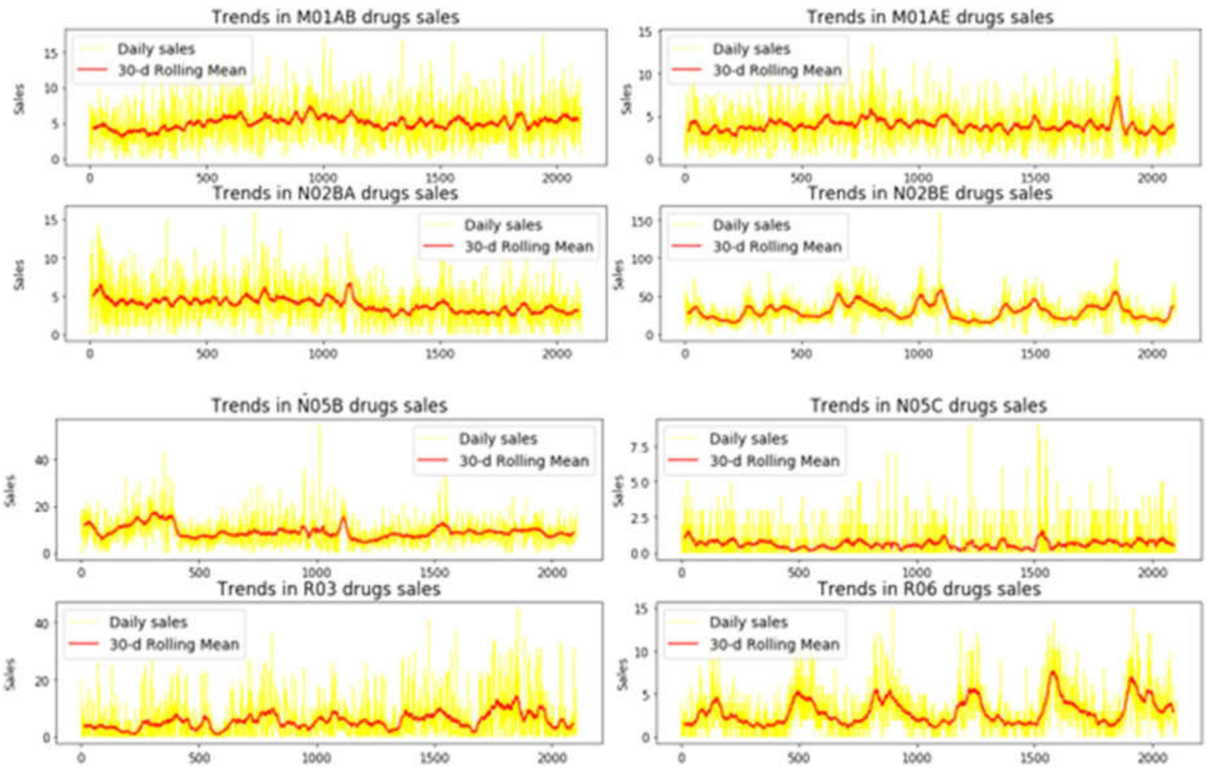
The performance of Deep Feedforward Networks ,DFMs was studied. This study was on eight drug types .A benchmark performance indicator was used. It's named RMSE value. Graphical illustration in Fig 6 shows the results .Figure 7 Makes A comparison. It compares minimum RMSE values for All eight Drug types. These values were Obtained with shallow NN And deep NN-based demand forecasting models.

Mean RMSE value of demand Forecasting models for eight drug types was displayed in Table 8. The value was found to be the same for shallow NN models. Models like GR_NN RBF_NN and P_NN showed identical RMSE values. These were for all eight categories of the drug .The conclusion was made that the shallow NN-based demand forecasting model is best. It is Best for Predicting demand for pharmaceutical products.

V. SUMMARY AND SCOPE

Main aim of demand forecasting models is laid out as follows:

- Accomplishment of intended production target
- Production stabilization against demand



RMSE for ARIMA and Shallow NN-DFM

The aim is to garner a shortlist of past data values .These would be combined using a mathematical formula. This formula aims to portray a forecast. This will then be available for use. The ARIMA Model is an exceptional tool for this. This is because it can handle both seasonal and non-seasonal data .Next ,the Shallow NN model will be used.

Shallow NN model is a simplification of its deep counterpart.. It has a single layer Of neurons.. These Perform a linear transformation on the input data.. Further ,activation function is Applied to This output .A single Output is Then produced. This simple Structure can make some tasks easier .These tasks include Image recognition and sentiment analysis.

One popular type of neural network is the Multilayer Perceptron (MLP). It has Several layers of neurons. The first layer analyses The inputs .It forms outputs for the next layer. This process continues through The network. The final result is the Output of the network.

In reinforcement learning ,the neural network is Used in a different way. It does not just create An Output based on an input .Instead it receives feedback as it interacts With an environment. This Feedback is then used to improve. Reinforcement learning Is Used in many tasks. Some of these include Game playing and robotics.

Deep neural networks are a recent advancement. Their complexity makes them very powerful .They are used in many applications today. These include self-driving Cars and music composition. However ,A large drawback is Its need for extensive data. This can Be a limitation in many fields.

- Future growth of sales
- Longstanding investment planning
- Budget preparation and Sales budget
- Control of raw materials

One clear thing from these Cases is That the shallow neural network models generally outperformed in terms of predictive accuracy. The most significant fact from this study is a deep neural network not always the best performance .Against this drug sales dataset It had lesser number of Samples For building the demand forecasting model. In short shallow neural networkbased time-series demand forecasting Models could offer potentially useful recommendations to pharmaceutical comppanies. The weekly sale data of ATC drug products has been useful .This identifies several days for implementation of special marketing campaigns. These campaigns Are to promote their sales.

VI. CONCLUSION AND FUTURE WORK

The primary Aim was to develop a method for forecasting demand. This method can help pharmaceutical companies produce or store Appropriate quantities of pharm products .Shallow and deep neural network models were leveraged. The models were used for forecasting for eight categories In the ATC thematic drug dataset.

The deep learning models utilized included LSTM And Stacked LSTM.. The shallow models were GRNN ,RBF_NN and P_NN.. The study Aimed to explore the predictability Of specific pharmacological agents.. But It was done Through the framework of the Atc categorization .This categorization was provided by WHO.

The data consisted of a time-series dataset spanning 1000 days .This data included the number Of times each drug was sold daily. Naturally it posed a significant challenge. This challenge was the identification Of a model that could capture the sequences temporal dynamics.

Three deep learning models were considered in the research. .These were LSTM ,Stacked LSTM and shallow neural network models.. In The LSTM model, the Input Layer consisted of 128 neurons.. The hidden layer contained 64 LSTM nodes.

The output layer was composed of 32 neurons.

The Stacked LSTM model was Used to model The data. The first LSTM Layer had 256 neurons. The second layer had 128 LSTM units. Thevlast layer had 32 neurons .All layers used the sigmoid activation function. P_NN shallow Neural network model was Also employed.

The data Was split into training And testing sets.. The training set consisted of 70% of the data. .It was used to train each of the models.. The remaining 30% of data was used to test the predictions of the trained models.

The performance of The models was assessed through the use of two evaluation metrics. These were root mean square error (RMSE) and Mean absolute percentage error (MAPE). The Rmse value for the LSTM model was 1.87. The MAPE value for the LSTM model was 3.12 .This is indicative Of its predictive power.

The RMSE value for the Stacked LSTM Model was slightly lower .It was 1.65. The MAPE value was also lower. It was 2.9. These scores demonstrate that The Stacked LSTM model performed better.

The RMSE value for the P_NN Architecture was 2.37. The Mape value for P_NN model was 7.2. These scores show that P_NN did not perform well. In comparison To the LSTM and Stacked LSTM models ,P_NN model had higher error .It Was noted as less accurate for The task of predicting drug demand.

Despite its shortcomings ,This work demonstrates potential for pharmaceutical companies .They can better predict Demand for pharm products The study showed limitations of certain models But it also showed promise for the use of deep learning models

This study may help Pharmaceutical companies .The companies might benefit from its findings. They could develop new models to forecast demand accurately. The proposed models might better represent the temporal dynamics of Drug sales.

The framework for Creating new predictive models is outlined. An even deeper study of Drug sales' temporal dynamics may be carried out. This Study could yield better results or the development of more sophisticated models .Better forecasting of pharmaceutical product demand Is anticipated.

Shallow Neural Networks (GRNN, RBF_NN and P_NN) and Deep Neural Network (LSTM and Stacked LSTM) models were used. These were Models for eight categories of the ATC thematic drug .In this study ,shallow neural network models have worked well for five drug categories out of eight .ARIMA model worked best for the remaining three drug categories. From the experimental study we know Single demand Forecasting model is not optimum for all drug categories. Shallow NN-based DFM did better for ATC thematic drug time-series dataset with mean RMSE value of 6.27. This was for all Eight drug types. Future Research work is needed to augment the predictive accuracy time series forecasting model.

REFERENCES

[1] Takahashi Y, Aida H ,Saito T (2000). Superiority of ARIMA model over f-ARIMA model. In: WCC 2000 - ICCT 2000 international conference on communication technology .Proceedings (Cat No00EX420). https://doi.org/10.1109/icct.2000.889171

[2] Mohammed NA ,Al-Bazi A (2021) .Adaptive Backpropagation algorithm for long-term electricity load forecasting. Neural Comput Appl. https://doi.org/10.1007/s00521-021-06384-x

[3] Livieris IE, Pintelas E ,Pintelas P (2020) .CNN–LSTM Model for Gold price time-series forecasting. Neural Comput Appl 32(23):17351–17360. https://doi.org/10.1007/s00521-020-04867x

[4] Oliveira DD, Rampinelli M ,Tozatto GZ, Andrea˜o RV, Mu¨ller SM (2021. Vehicular traffic flow forecasting using MLP and LSTM. Neural Comput Appl 33(24):17245–17256 .https://doi.org/101007/s00521-021-06315w

[5] Hamzac¸ebi C, Es HA ,C¸akmak R (2017). Forecasting Turkey's monthly electricity demand by Seasonal artificial neural network. Neural Comput Appl 31(7):2217–2231. https://doi.org/10.1007/s00521-017-3183-5

[6] Kanavos A, Kounelis F Iliadis L ,Makris C (2021) Deep learning models for forecasting aviation demand time series. Neural Comput Appl 33(23):16329–16343 .https:/doi.org/10.1007/s00521-021-06232-y

[7] Hu Z ,Zhao Q, Wang J (2018). .Prediction Model of worsted yarn quality.. It is based on a CNN–GRNN neural network. Neural Comput Appl 31(9):4551–4562. https://doi.org/10.1007/s00521-18-3723-7

[8] Merkuryeva G, Valberga A ,Smirnov A (2019) .Demand forecasting in pharmaceutical supply chains. This is a case study .Procedia Comput Sci 149:3–10. https://doi.org/10.1016/j.procs.2019.01. 100

[9] Thomson ME, Pollock AC ,O¨ nkal D, Go¨nu¨l MS (2019).. Combining forecasts: performance And coherence. .Int J Forecast 35(2):474–484. https://doi.org/10.1016/j.ijforecast.2018.10.006

[10] Nikolopoulos K, Buxton S, Khammash M ,Stern P (2016). .Forecasting branded and generic pharmaceuticals.. Int J Forecast 32(2):344–357. https://doi.org/10.1016/j.ijforecast.2015.08.001

[11] Siddiqui R, Azmat M, Ahmed S ,Kummer S (2021). A hybrid demand forecasting Model for greater forecasting accuracy. This is the case of the pharmaceutical industry. Supply Chain Forum Int J .https:/doi.org10.108016258312.2021.1967081

[12] Tayyab M, Zhou J, Zeng X ,Adnan R (2016). Discharge forecasting by applying artificial neural networks at Jinsha river basin, China. Eur Sci J ESJ 12(9):108 .https://doi.org/10.19044/esj.2016.v12n9p108

[13] Sun Y , Lang M, Wang D, Liu L (2014). It's A PSO-GRNN model For railway freight volume prediction. This empirical study is from China .J Ind Eng Manag. https://doi.org/10.3926/jiem.1007

[14] Ustaoglu B ,Cigizoglu HK, Karaca M (2008).Three Artificial neural network methods. They forecast daily mean, maximum and minimum temperature time series. Meteorol Appl 15(4):431–445. https://doi.org/10.1002/met.83

[15] Cifuentes J ,Marulanda G, Bello A, Reneses J (2020). Air temperature forecasting Occurs Using machine learning techniques .A detailed review describes this method. Energies 13(16):4215. https://doi.org/10.3390/en13164215

[16] Nayak SC ,Misra BB (2018). Estimating Stock closing indices occurs Using GA-weighted Condensed polynomial neural network. The method is inside Financ Innov .https://doi.org/10.1186/s40854-018-0104-2

[17] Li L, Xu Y ,Yan L et al (2020). A regional NWP's tropospheric delay inversion method is based on general Regression neural network model. Sensors 20:3167 .https://doi.org/10.3390/s20113167

[18] Ai S ,Chakravorty A, Rong C (2019). Household power demand prediction is done by evolutionary ensemble Neural network pool. They Utilize multiple network structures. Sensors 19:721 .https://doi.org/10.3390/s19030721

[19] Ng WW, Xu S ,Wang T et al (2020). Radial basis function neural network Is utilized with Localized stochastic-sensitive autoencoder. This Is For home-based activity recognition .Sensors 20:1479. https://doi.org/10.3390/s20051479

[20] Chao Z ,Kim H-J (2019) .Removal of computed tomography ring artifacts occurs via radial basis function artificial neural networks. Phys Med Biol 64:235015. https://doi.org/10.1088/1361-6560/ab5035

[21] Yu Q, Hou Z, Bu X ,Yu Q (2020) .An RBFNN-based data-driven predictive iterative learning control is used For nonaffine nonlinear systems. Ieee Trans Neural Netw Learn Syst 31:1170–1182. https://doi.org/10.1109/tnnls.2019.2919441

[22] Qiu R, Wang Y ,Wang D et al (2020). Water temperature Forecasting Is carried out based on modified ANN methods .Two cases of the Yangtze River in China are considered. Sci Total Environ 737:139729. https://doi.org/10.1016/j.scitotenv.2020.139729

[23] Foroughi M, Azqhandi MHA ,Kakhki S (2020). Bio-inspired adsorption takes place using Fe3O4–g–CN@PEI–b–CD nanocomposite High and fast adsorption of tetracycline occurs in aqueous media. This Is modeled by response surface methodology (RSM) .Boosted regression tree (BRT) and general Regression neural network (GRNN) are also used for modeling .J Hazard Mater 388:121769. https:/doiorg/10.1016/j.jhazmat.2019.121769

[24] Tkachenko R ,Izonin I, Kryvinska N et al (2020). They Propose an approach To increase prediction accuracy. This approach is for the recovery of missing IoT data .The approach is based on the GRNNSGTM ensemble .Sensors 20:2625. https//doi.org/10.3390s20092625

[25] Chu W-L, Lin C-J ,Kao K-C (2019) .Detection of a rotor malfunction and a ball-bearing system. They used Discrete Wavelet Transform integrated with Support Vector Machine. They used General Regression Neural Network. They also used visual dot patterns .Sensors 19:4806. https://doi.org/10.3390/s19214806

[26] Xie H, Li G, Zhao X ,Li F (2020). Forecasting of limb joint angles. They used multi-source signals. It was done by GS-GRNN for exoskeleton wearers. Sensors 20:1104 .https://doi.org/10.3390/s20041104.