Thesis Proposal: A Comprehensive Study of Deep Learning Techniques for Supply Chain Demand Forecasting

Asef Shahriar

1

Khulna University of Engineering & Technology  
Department of Industrial Engineering and Management

Proposal for Thesis Research in Partial Fulfillment  
of the Requirements for the Degree of  
Bachelor of Science in Industrial and Production Engineering  
Course no. IPE 4000

Title: A Comprehensive Study of Deep Learning Techniques for Supply Chain Demand Forecasting

|  |  |
| --- | --- |
| Submitted by: | Asef Shahriar |
|  | ID-1611010 |
|  | Department of Industrial Engineering and Management |
|  | Khulna University of Engineering & Technology |

|  |  |
| --- | --- |
| Supervised by: | Md. Al Amin |
|  | Assistant Professor |
|  | Department of Industrial Engineering and Management |
|  | Khulna University of Engineering & Technology |

Signature of Author:  
Signature of Supervisor:

Date of Submission:

Expected Date of Completion: April 2021

1

|  |  |
| --- | --- |
| Nomenclature |  |
|  |  |
| ANN | Artificial Neural Network |
| ARIMA | Autoregressive Integrated Moving Average |
| BDA | Big Data Analytics |
| CNN | Convolutional Neural Network |
| DBN | Deep Belief Network |
| DL | Deep Learning |
| DNN | Deep Neural Network |
| EMD | Empirical Mode Decomposition |
| FCL-Net | Fusion Convolutional Long Short-term Memory |
| GA | Genetic Algorithm |
| IMF | Intrinsiic Mode Function |
| KNN | K-Nearest Neighbors |
| LSTM | Long Short-term Memory |
| MA | Moving Average |
| MAE | Mean Absolute Error |
| MAPE | Mean Absolte Percentage Error |
| ML | Machine Learning |
| MLR | Multiple Linear Regression |
| NN | Neural Network |
| RBM | Restricted Boltzman Machines |
| RF | Random Forest |
| RMSE | Root Mean Squared Error |
| RNN | Recurrent Neural Network |
| SC | Supply Chain |
| SCM | Supply Chain Management |
| ST-ResNet | Deep Spatio-Temporal Residual Networks |
| STL | Seasonal and Trend Decomposition uaing Loess |
| DADLM | Duo Attention Deep Learning Model |
| SVM | Support Vector Machine |
| SVR | Support Vector Regression |

A Comprehensive Study of Deep Learning Techniques for Supply Chain Demand Forecasting

# Overview of the Problem

1 Demand forecast has always been a vital part of supply chain. A forecasting model with adequate accuracy can increase the profit margin of the company. In a broader sense it can make the supply chain of the organization more responsive and resilient to failure. Traditional forecasting methods have been in practice for long time. These methods are based on simple arithmetic and of low accuracy. But seasonal variability, special days, e-commerce etc. have increased the complexity of demand forecasting now-a-days. To meet this requirement researchers have applied sophisticated time series and machine learning models. These models performed better to some extent than traditional regression models. Deep learning has emerged in recent time with great potential in almost every field. There deep learning techniques could be a better fit to address the issues in SC demand forecasting.

# Objectives

1 The objective of this research is to study and evaluate the performance of advanced deep learning techniques in forecasting demand with high variability and seasonal components.

# Literature Review

1 Deep learning techniques are being used for multitude of forecasting applications in recent times. used long short-term memory network, a variant of DL network, for demand forecasting in multi-channel retail. They combined random forest algorithm and deep learning in their work. Their findings shows that the proposed method can model complex relationship with elevated accuracy. In another research proposed a cross-temporal hierarchical framework integrating deep learning for supply chain demand forecasting. The resultant forecast is found to be coherent at all level of supply chain. implemented ML based forecasting techniques such as ANN, RNN, and SVM and benchmarked them with traditional methods such as exponential Theta model, MA, smoothing, linear regression etc. However their findings was not exciting as ML techniques did not outperform traditional methods. The SVM was found to be most accurate among the algorithms. Another study by used the identical algorithms in demand forecasting. They suggested RNN and SVM have superior performance than other techniques. But there was no statistically significant improvement over regression model(MLR). Many researchers applied ML techniques for demand forecasting outside of supply chain. used ML techniques to forecast tourism demand in Turkey. They mentioned ML models performed significantly better with the introduction of auxiliary variables in the dataset. built a end-to-end ML system using Apache Spark. Their primary focus was on demand forecasting in retail. Their model is superior in the sense that it includes distributed learning, evaluation and ensembling. used DL to forecast short-term passenger demand under on-demand ride services. They proposed a novel approach called FCL-Net. They concluded the model performed better than traditional time-series prediction methods and NN based algorithms(e.g, ANN and LSTM). Addition of exogenous variables reduced RMSE by 50.9% in the study.

Electronic power industries are using DL models for load demand forecasting extensively. presented an ensemble method integrating EMD with DL. EMD was used to decompose the demand series into some IMFs. Then DBN, a class of DNN is employed to model these IMFs. They also included RBMs in the DBN. The model was more attractive than nine other models studied in the research . Another study with electricity load application was done by . Their method support arbitrary time horizons and exhibits less than 2% error margin . Deep learning has the potential to pave the way of smartgrid technology. explored CNN for energy load forecasting at building level. They mentioned CNN outperformed support vector regression model. In conclusion they suggested more experiment on this topic. studied LSTM for electric load forecasting to provide better load scheduling and reduce unnecessary production. They fused GA in their model to find number of layers and optimal value of time lags. Their model was successful in capturing the features of the time series. As a result, reduced MAE and RMSE are observed in the forecasting. applied ML techniques to reduce instantaneous response to water demand by taking advantage of forecasting. They applied NN, RF, SVM and KNN on data collected from Portuguese water utilities. They commented that development and implementation of forecasting based response in the system can reduce cost by 18% or more . studied DL methods for forecasting monthly tourism demand in Macau. They stated that it has become quite difficult for existing models to accurately forecast data with the addition of great amount of search intensity indicators. Their first contribution is the development of a systemic conceptual model that makes use of all tourism demand factors without human intervention. Employing the attention score to portray the deep learning models is their second contribution. addressed two primary issues of tourism demand forecasting with DL: data inaccessibility and requirement of explanatory auxiliary variables. Therefore they proposed a decomposition model, STL-DADLM, rather than a complex over fitted model. utilized LSTM to forecast e-commerce demand of Walmart. They also developed a synthetic pre-processing unit to overcome challenges in e-commerce based forecasting. Their model achieved noteworthy improvement over state-of-the-art techniques in some categories of products. Big data analytics and machine learning techniques are widely developed for supply chain demand forecasting. NN, Regression, ARIMA, SVM, Decision Tree are the most popular methods for demands forecasting among the researchers as found by .They studied applications of BDA in demand forecasting in SCM. He shed light on the limitations of conventional methods and how BDA enables us to overcome these barriers . presented a DL model for urban taxi demand forecasting. They compared ST-ResNet and FLC-Net in their study. The study emphasized on right DNN structure and domain knowledge as most DNN models are superior than conventional ML techniques with proper design and tuning. applied pooling-based deep RNN to forecast uncertain household load. This novel approach surpassed ARIMA by a margin of 19.5%, conventional RNN by 6.5% and SVR by 13.1% . contrasted deep learning approach with conventional time series models for building level load forecasting. They found highest result with multi-step formulated CNN. It is also computationally less expensive due to less number of parameters in the CNN. Their CNN model provided 26.6% more accurate forecasts than seasonal ARIMAX. showed a very different approach by considering the forecasting problem as classification problem rather than a regression problem. In their study, ML based classification algorithms surpassed regression algorithms. Their main focus was on the influence of special days in demand variability. They concluded ML models are best suited for large-scale applications of demand forecasting alongside providing more accuracy.

# Methodology

1

## Data Collection/Synthesis

The research aim to apply the developed deep learning models on two different datasets: one real-world dataset and another one from simulation of a multi-step supply chain. The synthesized dataset will have significant variability and random noises. The real world dataset should come from a renowned supplier’s supply chain. Due to the implications connected with sharing industry data, the real world datasets may not be very effective in this study. As an alternative the synthesized datasets come handy to develop and evaluate the models.

For every ML/DL applications, several preprocessing steps are employed to process raw data. These methods include data cleaning, transforming operation, normalization, feature engineering etc. There is no single preprocessing strategy that works well on every applications. One has to find the best sequence of operations to suit his needs.

## Model Implementation

Generally, most SC forecasting problems are univariate. These univariate time series forecasting problem can be modeled as a supervised learning problems as follows:

Univariate time series datasets need restructuring to be modeled as a supervised learning problem. A well established method of restructuring is **windowing**. In this approach, fixed sized windows are created from previous time steps and the next immediate time step is considered as the output. These pairs of data are fed into the neural network for training. There exists variety of deep learning techniques for forecasting application.

The state-of-the-art neural network architectures include:

1. Artificial Neural Networks(ANNs)
2. Convolutional Neural Networks (CNNs)
3. Recurrent Neural Networks(RNNs)
4. Long-short Term Memory(LSTMs)

Many researchers have also applied hybrid-models for forecasting applications. Such models include:

1. CNN-RNN
2. CNN-LSTM

Applying these methods after decomposing the demand series into several IMFs is also a popular approach. Another approach is inclusion of auxiliary variables into the model to increase forecasting accuracy. Finally I intend to study the transfer learning approach on this regard. This approach would be novel in demand forecasting.

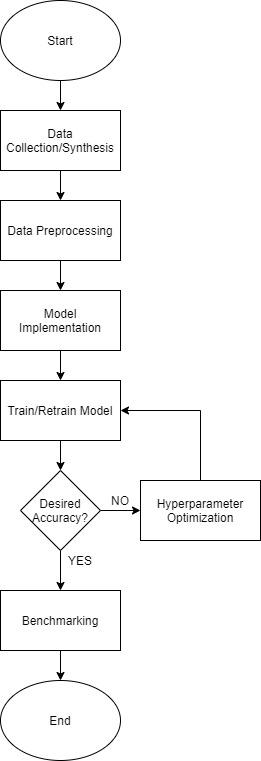
Deep learning models possess many hyper-parameters. Hence it is mandatory to tune those parameters to achieve best results from the models. Trying random values for the parameters and coarse to fine sampling scheme are widely used for optimization.

## Benchmarking

The thesis will include the benchmarks of costs and benefits of applying these models for supply chain demand forecasting. Specifically, I suggest the followings:  
Firstly, **RMSE** and **MAPE**. These will evaluate the forecasting error margins that the SC will have to endure.

Secondly, **data costs**. Generally DL models improve as the data volumes increases. But data collections and storage have costs connected to them. Thus the organizations have to face cost-accuracy trade-off.

Finally, **overall application impact**. For a new application like this a SC has to go through many challenges. At last these implications and expectancies will be discussed.



Research Methodology Flow Chart

# Expected Results

1

1. Reduced RMSE and MAPE
2. A model with low data and development costs. Especially, if the transfer learning approach show enough accuracy, the costs associated with data will be lesser than conventional models.
3. A feasible model considering costs and other challenges