VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

SCHOOL OF INFORMATION TECHNOLOGY AND ENGINEERING

WINTER Semester –2022-23

B.Tech (IT) Capstone Project – Guide Interaction

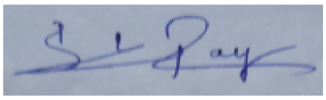
Date: 15-12-2022 to 16-12-2022

Register Number	19BIT0303
Student Name	SOUBHIK SINHA
Project Domain (Capstone Project)	MACHINE LEARNING
Project Title (Capstone Project)	ELECTRICTY DEMAND FORECASTING AND ANALYSIS USING TIME-SERIES ALGORITHMS
Keywords	Time-Series Algorithms , Forecasting , LSTM , RNN , Neural Network
Company Name & Address (For Off-campus students only)	N/A
External Mentor details (For Off-campus students only)	External Mentor name: N/A Designation: N/A Email: N/A Contact Number: N/A
Literature survey (10-15 references)	In [1] proposed a hybrid approach to long-term forecasting based on Time Series and Data Mining. The Support Vector Regression (SVR) approach is used to create a forecasting algorithm. The Particle Swarm Optimization (PSO) methodology is used to optimise the SVR algorithm's parameters as well as the size of the input samples. A hybrid forecasting technique is proposed for long-term annual electric peak load and total electric energy demand in order to reduce forecasting error. The Auto-Regressive Integrated Moving Average

	(ARIMA), Artificial Neural Network (ANN), and the suggested Support Vector Regression approach are the foundation of the proposed hybrid method. Based on the autocorrelation and partial autocorrelation of the original and differenced time series, the ARIMA method's parameters are established. In [2] used adaptive fourier decomposition to extract the fluctuation characteristics. The sub-series are run to measure and get rid of the seasonal pattern once the criterion of the linear and stationary sequence is met. The average periodicity length is quantified throughout the seasonal adjustment procedure. Additionally, the sine cosine optimization approach is used to choose the penalty and kernel parameters of the support vector machine in order to achieve the generalisation performance on actual electricity demand data. The results of the empirical study demonstrate that the proposed hybrid method's superior properties benefit from the impact of data pretreatment, and they also demonstrate that this hybrid modelling approach can produce promising prediction results while maintaining a manageable level of computational complexity. In [3] proposed a hybrid forecasting model to predict electricity consumption. The optimal smoothing parameters for the Holt-Winters exponential smoothing are chosen using the fruit fly optimization process. The forecasting effectiveness of the suggested model in comparison to competing approaches was thoroughly evaluated by the authors using power consumption data from a Chinese metropolis. The findings show that the suggested model, even with a limited number of training data, may significantly increase the forecast accuracy of monthly power use. The suggested model's calculation time is also the quickest among the benchmark hybrid algorithms that were tested. In [4] presented a version of the Multiple Aggregation Prediction Algorithm, a particular implementation of MTA, for high-frequency time series, to better manage the unfavourable effect of seasonality shrinkage that MTA implies and combine it with traditional cross-sectional hierarchical forecasting. An empirical evaluation of the effects of temporal aggregate inclusion in hierarchical forecasting is performed using real data from five bank branches. The authors demonstrate that the suggested MTA methodology exhibits higher accuracy, aggregate consistency, and dependable automated forecasting when used in conjunction with the best reconciliation technique. In [5] evaluated a number of machine learning algorithms to forecast the hourly power consumption at both the individual building level and the overall level. Understanding short-term dynamics requires forecasting at the hourly granularity, while the majority of neighbourhood scale studies are restricted to yearly, monthly, weekly, or daily data resolutions. Two years of data were utilised to train the

	model, and a third year of untrained data was used to make the forecast. For estimating the power consumption of specific buildings and groups of buildings, boosted-tree, random forest, SVM-linear, quadratic, cubic, and fine-Gaussian learning algorithms were all examined and put to the test. When criteria like computational time and error accuracy are examined, the findings indicated that boosted-tree, random forest, and ANN delivered the best results for prediction at hourly granularity. In [6], data were gathered from 100 civic public buildings for preliminary set reduction, and for nearly a year, data were gathered from a university lab building in Dalian to train and test deep neural networks. Both the short- and long-term forecasts for building energy use were tested. Comparisons were made between the predictions made by the deep neural network and those made by the back propagation neural network, the Elman neural network, and the fuzzy neural network. The outcomes demonstrate that the deep neural network and integrated rough set provided the most accurate results. The approach put forward in this study could offer a workable and precise answer for predicting building energy use. In [7] attempted to solve the issues, by creating a predictive model for energy usage in the cloud-based machine learning platform Microsoft Azure. For the algorithm of the prediction model, three methodologies—Support Vector Machine, Artificial Neural Network, and k-Nearest Neighbor—are suggested. The case study of two tenants from a business building in Malaysia focuses on real-world applicability. Before being utilised for model training and testing, the obtained data is analysed and pre-processed. RMSE, NRMSE, and MAPE measures are used to compare the effectiveness of each approach. The test demonstrates that the distribution of energy usage varies depending on the renter. In [8] employed hybrid models to examine how temperature and day type affect electricity demand, by utilizing data on Canadian power use. The point forecasts/predictions for the Ontario energy demand data are obtained using the hybrid neural network dynamic regression (NNDR) model. The experimental findings demonstrate the superiority of the proposed NNDR model over the Prophet and Seasonal Autoregressive Integrated Moving Average (DRSARIMA) errors models, two widely used dynamic regression models. Additionally, for the NNDR, DRSARIMA, and Prophet models, long-term point predictions and innovations are utilised to derive two classes of prediction intervals (PIs) utilising data-driven probabilistic innovation distribution and bootstrapping. In [9] utilized ensemble-based machine learning models for this task. In order to improve data quality for the forecasting job and enable the model to adjust itself in pandemic scenarios, lockdown temporal rules are added to the feature set. The short-term country-level demand

	forecast model is explored using a number of ensemble-based machine learning methods. For a probabilistic perspective, the quantile random forest regression is also used. The algorithms are taught to forecast Germany's national demand for case studies. The findings suggest that ensemble models, particularly boosting and bagging-boosting models, are capable of precise demand forecasting at the national level. Additionally, most of these models are resistant to missing pandemic policy data. The forecasting accuracy during a pandemic crisis is greatly increased by using the pandemic policy data as characteristics, though. For the aforementioned case study, the probabilistic quantile regression also showed great accuracy. In [10] anticipated power consumption using Artificial Neural Networks (ANN). With historical energy consumption statistics, educational data sets that account for social, economic, technological, and demographic characteristics were built in order to improve accuracy. The developing provinces of Düzce, Turkey, were subjected to the established forecasting methodology. The ANN-based technique was used to predict the region's 15-year electrical energy consumption. The outcomes have undergone a thorough analysis.
Problem Statement	The fluctuating demand of electricity has made the functioning and operations of powerplants in an ambivalent state. The deviations can result to much more than the already worst situation industries are dealing in the present – at one point only a single unit is working, and the very next day, all the available units have to be run at their max. capacity. This will eventually lead to early breakdown of machinery components – costing power corporations huge losses annually. Apart from the aforementioned, the society will not be catered with timely demands of electricity (especially in the chilling-cold months of december and hot months of May and June). This project aims to implement Time-Series algorithms to make use of past inputs to predict the future values – providing a chance to overcome the unplanned preparation needed for sudden power demands.
Meeting date & Time	15/12/2022 – 11:00 AM
Student Guide Interaction meeting	The points were discussed during the meeting • Study and Try to create a new architecture based on Short-Term Memory Algorithm. • Test and Compare the performance of ARIMA and other desired Time-Series algorithm(s) set for the project. • Attributes that can be included in the dataset. Rectifying the appropriate source of dataset (POSOCO Ltd., Contacting powerplant executives, Kaggle or any other Google data source). • Improvement of prediction performance based on proposed methodologies vs the existing ones • Output of LSTM – visualize and use w.r.t. to industry

Guide Name	Dr. BIMAL KUMAR RAY
Guide Signature	
Date	15/12/2022

References: (APA format)

References substantiating for framing the document.

(Minimum 10-15 references from reputed journals – IEEE, Elsevier)

1. Mohammad-Rasool Kazemzadeh, Ali Amjadian, Turaj Amraee. (2020). “A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting”. Energy, Volume 204, 117948, ISSN 0360-5442, https://doi.org/10.1016/j.energy.2020.117948.

2. Ranran Li, Ping Jiang, Hufang Yang, Chen Li. (2020). “A novel hybrid forecasting scheme for electricity demand time series”. Sustainable Cities and Society, Volume 55, 102036, ISSN 2210-6707, https://doi.org/10.1016/j.scs.2020.102036.

3. Weiheng Jiang, Xiaogang Wu, Yi Gong, Wanxin Yu, Xinhui Zhong. (2020). “Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption”. Energy, Volume 193, 116779, ISSN 0360-5442, https://doi.org/10.1016/j.energy.2019.116779.

4. Evangelos Spiliotis, Fotios Petropoulos, Nikolaos Kourentzes, Vassilios Assimakopoulos. (2020). “Cross-temporal aggregation: Improving the forecast accuracy of hierarchical electricity consumption”. Applied Energy, Volume 261, 114339, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2019.114339.

5. Shalika Walker, Waqas Khan, Katarina Katic, Wim Maassen, Wim Zeiler. (2020). “Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings”. Energy and Buildings, Volume 209, 109705, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2019.109705.

6. Lei Lei, Wei Chen, Bing Wu, Chao Chen, Wei Liu. (2021). “A building energy consumption prediction model based on rough set theory and deep learning algorithms”. Energy and Buildings, Volume 240, 110886, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2021.110886.

7. Mel Keytingan M. Shapi, Nor Azuana Ramli, Lilik J. Awalin. (2021). “Energy consumption prediction by using machine learning for smart building: Case study in Malaysia”. Developments in the Built Environment, Volume 5, 100037, ISSN 2666-1659, https://doi.org/10.1016/j.dibe.2020.100037.

8. S. Bowala, M. Makhan, Y. Liang, A. Thavaneswaran and S. S. Appadoo. (2022). "Superiority of the Neural Network Dynamic Regression Models for Ontario Electricity Demand Forecasting". IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 182-187, doi: 10.1109/CCECE49351.2022.9918212.

9. A. Arjomandi-Nezhad, A. Ahmadi, S. Taheri, M. Fotuhi-Firuzabad, M. Moeini-Aghtaie and M. Lehtonen. (2022). "Pandemic-Aware Day-Ahead Demand Forecasting Using Ensemble Learning". IEEE Access, vol. 10, pp. 7098-7106, 2022, doi: 10.1109/ACCESS.2022.3142351.

10. Y. E. UNUTMAZ, A. DEMİRCİ, S. M. Tercan and R. Yumurtaci. (2021). "Electrical Energy Demand Forecasting Using Artificial Neural Network". 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1-6, doi: 10.1109/HORA52670.2021.9461186.