MACHINE LEARNING BASED FORECASTING IN POWER SYSTEM

PARUL

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

AUGUST 2023

MACHINE LEARNING BASED FORECASTING IN POWER SYSTEM

by

PARUL

Department of Electrical Engineering

Submitted

in fulfilment of the requirements of the degree of Doctor of Philosophy
to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

AUGUST 2023

Certificate

This is to certify that the thesis entitled **"Machine learning based forecasting in power system"**, submitted by **Parul** to the Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy** in 2023, is a record of the original, bonafide research work carried out by her under our supervision and guidance. The thesis has reached the standards fulfilling the requirements of the regulations related to the award of the degree.

The results contained in this thesis have not been submitted in part or in full to any other University or Institute for the award of any degree or diploma to the best of our knowledge.

Prof. B.K Panigrahi
Department of Electrical Engineering,
Indian Institute of Technology, Delhi
New Delhi, India

Prof. P.N Suganthan
KINDI Center for Computing Research
College of Engineering
Qatar University,Doha,Qatar

Acknowledgements

First and foremost, I am extremely grateful to my supervisors, Prof. B.K Panigrahi and Prof. P.N Suganthan (KINDI Center for Computing Research, Qatar University) for their invaluable advice and continuous support during my Ph.D. Their plentiful experience and immense knowledge have encouraged me in my academic research. I am grateful to Prof. Panigrahi for providing me the confidence to work on fundamental research challenges throughout the Ph.D. journey. Working with him was so comforting because I knew he would assist me in all the ways both scientifically and personally. Thank you again Prof. Panigrahi for your magnificent support. I am really thankful to Prof.P.N Suganthan for his expertise in the field of evolutionary algorithms that helped me a lot during my research. His critical inputs on my papers were really helpful, that gave different perspective and added immense value in my publications.

I would also like to thank Prof.Abbas khosravi from Deakin University, Australia (through project collaboration) , he always encouraged me with his motivational words and having belief in me that my work can be published in good IEEE transactions. Through his motivation and support of his student Seyed Mohammad Jafar Jalali, new ideas in the field of deep neuroevolution could be explored and more research was conducted in this area.

Finally, I would like to thank my family, my parents, my in-laws, and my daughters, without their support, my Ph.D. wasn't possible. I would like to express my sincere gratitude to my husband, Munish Arora who has equally supported me in this Ph.D. journey. Right from the beginning, he was the one who encouraged me to pursue the Ph.D., and over time, whether it was staying alone to study or travelling daily, or shifting house, he always supported me. Even after the birth of my daughter (Shivya Arora) in 2019, he was the one who was taking care of my daughter along with his job, so that I could do my research work. During the COVID times in 2020, when situations changed, then we were working in shifts as sometimes he used to take care of our daughter and then do his office work, and other times, I used to take care of our daughter and then do mine research work. Special thanks to my father

(Ashwani Mahajan), who pushed me for completing my thesis and supported me wherever I required help for managing things at home and work place and looking after my younger daughter (Aavya Arora). So, I would like to say that it was my entire family's support that helped me in completing my research work.

Parul Arora
Department of Electrical Engineering,
Indian Institute of Technology, Delhi
New Delhi, India

Abstract

In the past few years, the power system planning and scheduling systems have been improving day by day, and for that purpose, the need for accurate forecasts is increasing. Time-series forecasts are required for demand, electricity price, and wind and solar forecasting. The forecasts can be provided through statistical, machine learning, or hybrid methods. The Machine learning (especially the neural network) models have achieved outstanding performance in the areas like image classification, natural language processing, and time-series forecasting. Nonetheless, the architecture of neural networks significantly impacts their overall performance. Concerning the design of the most cutting-edge NNs, they are generally hand-crafted by experts in that domain, which requires heavy trial and error efforts. On the other hand, it becomes a challenging task for users who are not experts in the NN domain to choose the best architecture for their application. Also, there is another issue that the basic NNs only sometimes suit the particular application. Therefore, several strategies have been explored for automating the designing of the NNs, like evolutionary algorithms (EAs), bayesian optimization and reinforcement learning (RL), and concatenating with some other statistical methods. Among these methods, the evolutionary algorithms have gained the maximum interest due to their less complexity and easy integration for automating the design of the NNs. Deep Neuroevolution is one such efficient class of models referring to the automated procedure and training of NNs using EAs. This thesis first describes the comprehensive analysis, survey, and evaluation of the current state-of-the-art studies on employing statistical and machine learning algorithms along with evolutionary algorithms. Then, the major issue of designing optimal probabilistic models through neural networks is being addressed along with the designing of neuroevolution models on two different case studies: one on wind forecasting and the other one on load forecasting. The experimental results show that the deep neuro-evolution models designed in this research work perform better than the other state-of-the-art algorithms. In this thesis, deep learning-based novel algorithms are proposed in each chapter. As in Chapter 3, neural grasshopper optimized DeepAr is proposed in which the evolutionary algorithm is used for the optimization of the hyperparameters of the DeepAr and this model is tested on the GEFCom-14 and AEMO datasets and the prediction interval coverage probability (PICP) and Pinball loss (PL) for the two datasets are $[0.902, 0.320]$ and

$[0.933, 1.4885]$, respectively. In Chapter 4, a hybrid model based on combining a deep neural network and state-space model is proposed, which utilizes the basic features of both models. This model has been tested on univariate as well as multivariate loads, and this model can be employed in areas where less data is available to the places where extensive data is available. Whereas chapter 5 proposes a novel neuroevolution algorithm based on CNN for handling the uncertainty associated with load forecasting. Also, the mean scaled interval score metric is used to evaluate the forecasts.

सार

पिछले कुछ वर्षों में, बिजली प्रणाली योजना और शेड्यूलिंग प्रणालियों में दिन-ब-दिन सुधार हो रहा है, और उस उद्देश्य के लिए, सटीक पूर्वानुमानों की आवश्यकता बढ़ती जा रही है। मांग, बिजली की कीमत और पवन और सौर पूर्वानुमान के लिए समय-श्रृंखला पूर्वानुमान की आवश्यकता होती है। पूर्वानुमान सांख्यिकीय, मशीन लर्निंग या हाइब्रिड तरीकों के माध्यम से प्रदान किए जा सकते हैं। मशीन लर्निंग (विशेषकर तंत्रिका नेटवर्क) मॉडल ने छवि वर्गीकरण, प्राकृतिक भाषा प्रसंस्करण और समय-श्रृंखला पूर्वानुमान जैसे क्षेत्रों में उत्कृष्ट प्रदर्शन हासिल किया है। बहरहाल, तंत्रिका नेटवर्क की वास्तुकला उनके समग्र प्रदर्शन पर महत्वपूर्ण प्रभाव डालती है। सबसे अत्याधुनिक एनएन के डिजाइन के संबंध में, वे आम तौर पर उस क्षेत्र के विशेषज्ञों द्वारा हाथ से तैयार किए जाते हैं, जिसके लिए भारी परीक्षण और त्रुटि प्रयासों की आवश्यकता होती है।

दूसरी ओर, जो उपयोगकर्ता एनएन डोमेन के विशेषज्ञ नहीं हैं उनके लिए अपने एप्लिकेशन के लिए सर्वोत्तम आर्किटेक्चर चुनना एक चुनौतीपूर्ण कार्य बन जाता है। इसके अलावा, एक और मुद्दा यह है कि बुनियादी एनएन केवल कभी-कभी विशेष एप्लिकेशन के लिए उपयुक्त होते हैं। इसलिए, एनएन की डिजाइनिंग को स्वचालित करने के लिए कई रणनीतियों का पता लगाया गया है, जैसे विकासवादी एल्गोरिदम (ईएएस), बायेसियन अनुकूलन और सुदृढीकरण सीखना (आरएल), और कुछ अन्य सांख्यिकीय तरीकों के साथ संयोजन। इन विधियों के बीच, एनएन के डिज़ाइन को स्वचालित करने के लिए विकासवादी एल्गोरिदम ने अपनी कम जटिलता और आसान एकीकरण के कारण सबसे अधिक रुचि प्राप्त की है। डीप न्यूरोएवोल्यूशन मॉडलों का एक ऐसा कुशल वर्ग है जो ईएएस का उपयोग करके एनएन की स्वचालित प्रक्रिया और प्रशिक्षण का संदर्भ देता है। यह थीसिस सबसे पहले विकासवादी एल्गोरिदम के साथ-साथ सांख्यिकीय और मशीन लर्निंग एल्गोरिदम को नियोजित करने पर वर्तमान अत्याधुनिक अध्ययनों के व्यापक विश्लेषण, सर्वेक्षण और मूल्यांकन का वर्णन करती है। फिर, तंत्रिका नेटवर्क के माध्यम से इष्टतम संभाव्य मॉडल को डिजाइन करने के प्रमुख मुद्दे को दो अलग-अलग केस अध्ययनों पर न्यूरोएवोल्यूशन मॉडल के डिजाइन के साथ संबोधित किया जा रहा है: एक पवन पूर्वानुमान पर और दूसरा लोड पूर्वानुमान पर। प्रायोगिक परिणाम बताते हैं कि इस शोध कार्य में डिज़ाइन किए गए गहन न्यूरो-इवोल्यूशन मॉडल अन्य अत्याधुनिक एल्गोरिदम की तुलना में बेहतर प्रदर्शन करते हैं।

इस थीसिस में, प्रत्येक अध्याय में गहन शिक्षण-आधारित उपन्यास एल्गोरिदम प्रस्तावित हैं। जैसा कि अध्याय 3 में है, तंत्रिका टिड्डी अनुकूलित डीपएआर प्रस्तावित है जिसमें डीपएआर के हाइपरपैरामीटर के अनुकूलन के लिए विकासवादी एल्गोरिदम का उपयोग किया जाता है और इस मॉडल का परीक्षण

GEFCom-14 और AEMO डेटासेट और भविष्यवाणी अंतराल कवरेज संभावना (PICP) पर किया जाता है। दो डेटासेट के लिए पिनबॉल हानि (पीएल) क्रमशः 0.902, 0.320 और 0.933, 1.4885 है। अध्याय 4 में, एक गहरे तंत्रिका नेटवर्क और राज्य-अंतरिक्ष मॉडल के संयोजन पर आधारित एक हाइब्रिड मॉडल प्रस्तावित है, जो दोनों मॉडलों की बुनियादी सुविधाओं का उपयोग करता है। इस मॉडल का परीक्षण यूनीवेरिएट और मल्टीवेरिएट लोड पर किया गया है, और इस मॉडल को उन क्षेत्रों में नियोजित किया जा सकता है जहां कम डेटा उपलब्ध है, जहां व्यापक डेटा उपलब्ध है। जबकि अध्याय 5 लोड पूर्वानुमान से जुड़ी अनिश्चितता से निपटने के लिए सीएनएन पर आधारित एक उपन्यास न्यूरोएवोल्यूशन एल्गोरिदम का प्रस्ताव करता है। साथ ही, पूर्वानुमानों का मूल्यांकन करने के लिए माध्य स्केल अंतराल स्कोर मीट्रिक का उपयोग किया जाता है।

Contents

List of Figures

List of Tables

Abbreviations

AACE	Average Absolute Coverage Error
ACO	Ant Colony Optimization
AEMO	Australian Energy Market Operator
AIC	Akaike Information Criterion
AI	Artificial Intelligence
ALO	Ant Lion Optimizer
ANN	Artificial Neural Networks
ARCH	Autoregressive Conditional Heteroskedasticity
ARIMA	Auto-Regressive Integrated Moving Average
BBO	Bio-geography Based Optimizer
CDF	Cumulative Distribution Function
CNN	Convolutional Neural Network
CRPS	Continous Ranked Probability Score
CSS	Charged System Search
DFFNN	Deep Feed Forward Neural Network
DL	Deep Learning
DMDNN	Deep Mixture Density Neural Network
DNE	Deep Neuro Evolution
DNN	Deep Neural Network
DeepAR	Deep Auto-Regressive
EA	Evolutionary Algorithm
EGWO	Enhanced Grey Wolf Optimiser
ESM	Error Feedback Stochastic Modelling
ETS	Exponential Smoothing
FPA	Flower Pollination Algorithm
GAN	Generative Adversarial Network

GARCH	Generalized AutoRegressive Conditional Heteroskedasticity
GA	Genetic Algorithm
GBM	Gradient Boosting Machines
GBRT	Gradient Boosting Regression Tree
GEFCom	Global Energy Forecasting Competition
GOA	Grasshopper Optimisation Algorithm
GPU	Graphical Processing Units
GRU	Gated Recurrent Unit
GWO	Grey Wolf Optimizer
HELM	Hysteretic Extreme Learning Machine
ISN	Improved Seasonal Naive
ISSM	Innovation State-Space Models
IS	Interval Score
LSTM	Long Short-Term Memory
LUBE	Lower Upper Bound Estimation
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MFO	Moth Flame Optimizer
MIS	Mean Interval Score
MLR	Multiple Linear Regression
ML	Machine Learning
MQR	Multivariate Quantile Regression
MSE	Mean Square Error
MSIS	Mean Scaled Interval Score
MVO	Multi Verse Optimizer
ND	Normal Deviation
NE	Neuro Evolution
NGOA	Neural Grasshopper Optimisation Algorithm
NN	Neural Network
NPTS	Non-Parametric Time Series
NRMSE	Normalised Root Mean Square Error
NWP	Numerical Weather Prediction
OBL	Opposition Based Learning
PCFM	Proposed Combined Forecasting Model
PICP	Prediction Interval Coverage Probability

PI	Prediction Interval
PLF	Probabilistic Load Forecasting
PL	Pinball Loss
PSO	Particle Swarm Optimization
PWF	Probabilistic Wind Forecasting
QRF	Quantile Regression Forest
RBFNN	Radial Basis Function Neural Network
RBM	Restricted Boltzmann Machine
RF	Random Forest
RL	Reinforcement Learning
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
ReLU	Rectified Linear Unit
SES	Simple Exponential Smoothing
SGD	Stochastic Gradient Descent
SHADE	Success History Based Adaptive Differential Evolution
SSA	Salp Swarm Algorithm
SSE	Sum of Squared Errors
SSM	State-Space Model
SS	State-Space
SVM	Support Vector Machine
TAR	Autoregressive Conditional Heteroskedasticity
VAR	Value at Risk
VECM	Vector Error Correction Model
VRAE	Variational Recurrent Auto Encoders
WCA	Water Cycle Algorithm
WOA	Whale Optimization Algorithm
WS	Wind Speed
XGBoost	Extreme Gradient Boosting