UtilitasMathematica

ISSN 0315-3681 Volume 122, 2025

Predicting Smart Grid Stability Using Advanced Machine Learning Techniques

Chintala Deepak Srinivas [1], M Kavitha [2]

[1] Student, [2] Associate Professor

Department of CSE

Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.

Deepaksrinivas249@gmail.com , mkavita@kluniversity.in

Abstract: An electric lattice incorporates transformers, era centers, communication joins, control stations, along with wholesalers, all of which together permit for the exchange of power from control stations to commercial and private customers . Customary electrical systems cannot predict quick shifts in electricity demanded by users and often lack adequate flexibility and robustness. This limitation has overwhelmingly driven onward the transition into modern smart grids—new electrical systems adequately designed for being self-healing, durable, and adaptive for constantly evolving client needs. Machine learning has emerged as an influential tool in the enhancement of grid stability via dealing with the challenges posed by dynamically shifting demands. This transformation decidedly minimizes the risk of breakdowns, ensuring a reliably more efficient grid system. In this analysis, advanced machine learning methods, inclusive of the CatBoost framework, were used in terms of grid stability prediction with a Kaggle-hosted public dataset. The experiments were conducted in a largely Python-based simulation environment. They leveraged thorough preprocessing, careful feature scaling, and refined predictive modelling techniques. The prior CatBoost-based model achieved a great accuracy of 98.96% during load stability prediction. This accurate forecasting of power demand notably reduces the likelihood of grid failure, improving the stability, reliability, as well as the robustness of the smart grid system.

Keyword: *Classification, Kaggle, Machine learning, Smart grid stability, load balancing, prediction model.*

1 Introduction:

The change of ordinary electrical frameworks into progressed keen networks marks a noteworthy advancement in control conveyance frameworks. Standard electric systems work unidirectionally, in this way constraining control transport to a singular course and causing huge vitality wastage amid conveyance. By presenting capabilities with respect to two-way control transmission, shrewd lattices handle those limits, fortified by means of control frameworks that are progressed, in conjunction with communication apparatuses for two-way utilize, as well as data frameworks that are intelligent. These particular systems speak to progressed advances, such as Supervisory Control sand Data Acquisition (SCADA) frameworks, Phasor Data Concentrators (PDCs), and Phasor Measurement Units (PMUs). These components authorize administrators to watch and after that evaluate lattice dauntlessness in real-time. This moves forward both unwavering quality coupled with useful competence. Savvy lattices speak to progressed qualities such as sharp computerized meters for the outfitting of prompt client input, the discovery of systemic absconds, as well as the invigorating of shopper information within the center of disruptions. Supplemental switches ease the transmission of power within the center of squeezing circumstances, whereas electrochemical frameworks hoard lavish vitality to fulfill request amid deficiencies. These increases upscale network competence and, besides, handle the things displayed by energetic power requests. These moreover concern, in conjunction with those things, decentralized control era.

The consideration with clever calculations, progressed by means of machine learning, has changed network steadiness administration. The aptitude of machine learning toward generalization, coupled nearby adjustment inside significantly energetic situations, positions it as a profitable instrument in estimating power request as well as inside keeping up framework steadiness. These calculations use wide-ranging datasets, as well as those produced by modern frameworks, to figure designs. They too refine asset dispersion to secure harmony between vitality creation and utilization.

The move from routine control lattices to keen frameworks offers both benefits and impediments. Major obstacles include taking care of complicated communication and electrical frameworks while adapting to fluctuating requests, costs, and vitality quality. Cleverly frameworks, fueled by machine learning, offer a vigorous arrangement to these challenges by empowering real-time decision-making and versatility.

This consider centers on foreseeing savvy framework soundness utilizing progressed machine learning procedures. The essential goals incorporate creating predictive models to preserve lattice harmony in powerfully changing environments and utilizing these expectations to avoid framework disappointments by proactively overseeing control dissemination. The comes about of this investigate give basic experiences into upgrading the soundness, unwavering quality, and vigor of savvy frameworks, clearing the way for maintainable and productive vitality frameworks.

2. Review of Literature :

The developing appropriation of savvy frameworks is making unused openings for investigating the application of manufactured insights in real-time steadiness expectation. Different considers have inspected the utilize of machine learning methods to address distinctive challenges related with the execution and operation of canny systems . A nitty gritty dialog of these approaches is given in Subsection 2.1.

2.1. Related Work

The expanding appropriation of shrewd lattices has opened roads for leveraging progressed machine learning strategies to address different challenges in framework steadiness and administration.

Alazab et al. [3] presented the Multidirectional Long Short-Term Memory (MLSTM) show for determining savvy framework soundness and illustrated its predominance over conventional approaches such as Long Short-Term Memory (LSTM), Repetitive Neural Systems (RNNs), and Gated Repetitive Units (GRUs).

Hafeez et al. [12] emphasized the significance of precise vitality utilization determining in savvy lattices, especially given the non-linear nature of vitality request designs. Their think about utilized a profound learning demonstrate with the ReLU enactment work, prepared on U.S. control lattice information, and assessed its execution utilizing measurements such as fluctuation, relationship, and merging rate. Reenactment comes about highlighted the model's exactness and proficiency in short-term stack determining.

Gorzaczany et al. [11] proposed a fluffy inference-based framework for exact estimating of framework solidness in decentralized shrewd lattices, adjusting interpretability and precision.

Yu et al. [28] tended to the instabilities in vitality request expectation utilizing measurable modeling and back vector machines (SVMs). Their think about depended on real-world power meter information from Stanford College, utilizing Gaussian dissemination for information guess to improve determining precision.

Ma et al. [16] presented a software-based Lattice Soundness Mindfulness Framework that screens and analyzes real-time network soundness, centering on small-signal, voltage, and temporal soundness through a secluded system.

Ahmed et al. [2] investigated the integration of keen networks with renewable vitality sources as a feasible arrangement to rising vitality requests. By joining machine learning into Gaussian Prepare Relapse, they proposed an vitality administration demonstrate that beated molecule swarm optimization and hereditary algorithm-based approaches.

Wang et al. [27] utilized Bayesian-optimized LightGBM for framework soundness evaluation, accomplishing productive and fast expectations for high-dimensional datasets.

Rai and De [23] created a stack estimating demonstrate based on real-time information from savvy meters, illustrating the prevalent execution of back vector relapse in terms of Cruel Outright Blunder (MAE) and Root Cruel Square Blunder (RMSE).

Krc et al. [14] highlighted the significance of stack adaptability in savvy frameworks and proposed a hub characterization demonstrate utilizing counterfeit neural systems, parameter space investigation, and clustering strategies.

Panda and Das [20] explored decentralized control frameworks utilizing relapse investigation and highlight positioning, whereas Ghosh and Kole [10] and Bashir et al. [6] compared different machine learning calculations for lattice solidness forecast. Their comes about recognized the XGBoost classifier and Choice Tree classifier as best entertainers, accomplishing correctnesses of 97.5% and past.

Malbasa et al. [17] proposed a machine learning-based show for voltage soundness expectation in transmission frameworks, leveraging dynamic learning to move forward the dataset and recognizing irregular timberland as the foremost compelling classifier.

2.2. Research Gap

Despite advancements in power systems, traditional energy generation resources remain limited, and the rising demand for electricity necessitates a shift towards more efficient and reliable systems. While smart grids address this challenge, much of the research has focused on hardware and communication domains, with limited exploration of artificial intelligence (AI) for grid management. Accurate demand forecasting through AI has the potential to enhance grid stability and improve consumer satisfaction by minimizing transmission losses and ensuring reliable power delivery.

2.3. Motivation

The recognized ask around hole have spurred the makers to form a machine learning-based appear for foreseeing framework stack . By leveraging AI, this think about points to address crest stack circumstances viably, guaranteeing framework steadiness and operational productivity. The proposed technique, point by point in Area 3, gives a system for coordination progressed AI strategies into savvy framework operations to meet energetic vitality requests and make strides in general framework execution.

3. Methodology

3.1. Dataset Overview

The proposed demonstrate utilizes the freely accessible Keen Framework Soundness Dataset, initially compiled by Vadim Arzamasov and facilitated on Kaggle. This dataset contains 60,000 occurrences, each portrayed by 14 qualities, with 12 input highlights impacting network solidness. The target variable categorizes the grid's state as either steady or unsteady, serving as the premise for the prescient assignment.

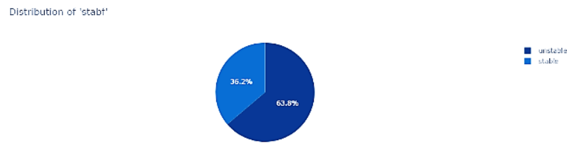

Figure 1 Distribution of stable and unstable

A detailed breakdown of the dataset features is provided below:

- To begin with Four Highlights: Talk to the reaction times of organize individuals, with values expanding between 5 and 10 [16] .
- Another Four Highlights: Show control generation and utilization for organize members. Positive values speak to generation, whereas negative values indicate utilization. Their extend is from -2.0 to -0.5.
- Ensuing Four Highlights: Capture the cost flexibility coefficients for organize members, labeled as gama1, gama2, gama3, and gama4. These coefficients extend from 0.05 to 1.00.
- The dataset incorporates two subordinate factors, cut and stabf, which are computed from the beginning 12 input highlights. The wound variable speaks to a nonstop solidness score measuring the grid's operational state, whereas stabf gives a twofold classification—labeling the framework as either 'stable' or 'unstable'—based on a limit connected to the solidness score.

A heatmap representation of the trait connections is outlined in Figure 1. The heatmap, which is determined from a relationship framework, highlights the degree of affiliation between factors within the dataset. The solid interconnects among the highlights demonstrate their appropriateness for foreseeing lattice soundness [1].

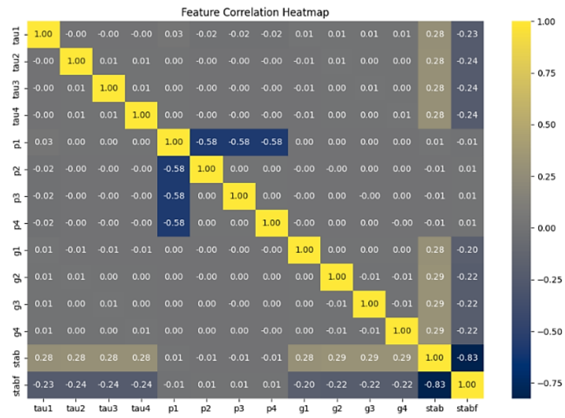

Figure 2 Correlation heatmap of smart grid dataset features.

3.2 Dataset Pre-Processing

The dataset utilized in this consider was falsely delivered , guaranteeing the nonattendance of lost values. As all highlights within the dataset are pertinent to the assignment, include choice is superfluous. In any case, due to the numerical nature of the dataset, appropriate encoding of highlights is fundamental. The dataset was in this way part into two subsets: 70% for preparing and 30% for testing, guaranteeing an fitting adjust for demonstrate assessment and approval.

Points of interest of the classifiers utilized for preparing the model are examined within the taking after subsection (3.3).

3.3 Classifier Details

The taking after classifiers were utilized for preparing and testing the show, each chosen for its special qualities in prescient examination:

Logistic Regression: A broadly utilized classification calculation that predicts the likelihood of a subordinate variable based on one or more free factors. It yields probabilities, allotting a course name of or 1 based on a predefined limit [13].

Random Forest: An ensemble-based directed learning classifier that totals the yields of numerous free choice trees. Each tree autonomously classifies the input, and the ultimate yield is decided by the lion's share vote (for classification) or the normal forecast (for relapse). The execution of a Arbitrary Timberland classifier depends on components such as the number of choice trees and the highlights chosen for part at each hub [25].

Decision Tree: A directed learning strategy that structures information into a tree-like show, comprising choice hubs and leaf hubs. Choice hubs speak to the choice of highlights, whereas leaf hubs indicate the anticipated results. This strategy is basic, interpretable, and broadly pertinent [15].

K-Nearest Neighbors (KNN): KNN can be a non-parametric calculation that classifies data centers based on their region to the closest kkk neighbors. It depends on a separate metric (e.g., Euclidean or Manhattan separate) to decide course names, making it straightforward and viable for different classification assignments.

Artificial Neural Networks (ANN): ANN imitates the working of the human brain by utilizing interconnected layers of hubs (neurons) to memorize complex designs in information. It comprises an input layer, one or more covered up layers, and an yield layer. Each neuron applies an enactment work to the weighted whole of its inputs, empowering the show to capture non-linear connections and perform exceedingly exact forecasts. ANNs are especially successful for dealing with high-dimensional and non-linear information.

Bagging Classifiers: Bagging, or bootstrap , is an outfit strategy that trains numerous base classifiers on arbitrary subsets of the preparing information. The ultimate expectations are made by amassing person yields through averaging (for relapse) or lion's share voting (for classification). In this think about, Support Vector Classification (SVC) was utilized as the base classifier for stowing [21].

CatBoost: CatBoost, short for "Categorical Boosting," could be a slope boosting calculation particularly planned to handle categorical information successfully without requiring broad preprocessing, such as one-hot encoding. It leverages requested boosting to maintain a strategic distance from overfitting and applies effective encoding strategies for categorical highlights. CatBoost is vigorous to hyperparameter tuning and gives great execution on both classification and relapse. assignments, particularly in datasets with a blend of numerical and categorical highlights.

4. Result and Analysis:

As detailed in Section 3.3, seven distinct machine learning algorithms—Decision Tree, Logistic Regression, Artificial Neural Networks (ANN), Random Forest, K-Nearest Neighbors (KNN), Bagging Classifier, and CatBoost Classifier—were employed to predict smart grid stability. The experiments were conducted utilizing Python 3.7 on a Windows 10 system which installed with an Intel i3 processor and 8 GB of RAM. Model performance was assessed using standard evaluation metrics such as Accuracy, Specificity, Sensitivity (Recall), Area Under the Curve (AUC), and the F1-score.

Precision measures the proportion of positive class predictions that are true positives, while Sensitivity (Recall) calculates the extent of genuine positive cases accurately recognized. The F1-score gives a adjusted degree of Recall and Precision. Accuracy reflects the classifier's overall correctness, and AUC quantifies the model's capacity to recognize between steady and unsteady lattice conditions. A better AUC esteem shows superior show execution in isolating the two classes.

Classifier / Metrics	**Sensitivity (Recall)**	**F1-Score**	**Specificity**	**Accuracy**	**Area Under Curve (AUC)**
Logistic Regression	0.686	0.722	0.881	0.811	0.889

ISSN 0315-3681 Volume 122, 2025

Random Forest	0.858	0.895	0.967	0.928	0.884
Decision Tree	0.827	0.842	0.827	0.889	0.911
ANN	0.684	0.943	0.787	0.909	0.807
KNN	0.782	0.848	0.965	0.911	0.900
Bagging Classifier	0.819	0.858	0.690	0.9739	0.953
CatBoost Classifier	0.980	0.985	0.990	0.989	0.987

Table 1 presents the results for each machine learning model.

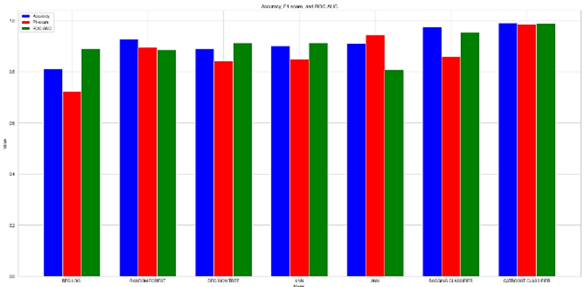

Figure 3 Comparison of machine learning techniques for predicting smart grid stability.

The Logistic Regression model achieved an accuracy of **81.16%**, a F1-score of **72.27%**, Sensitivity of **68.67%**, Specificity of **88.10%**, and an AUC score of **88.97%**, making it the least effective among the models.

The Decision Tree classifier performed better with **88.91% accuracy**, **84.22% F1-score**, and an AUC score of **91.19%**. KNN yielded an AUC score of **91.19%,** an accuracy of **90.02%** and a F1-score of **84.87%**.The ANN model recorded an accuracy of **91.00%**, a F1-score of **94.31%**, but a relatively lower AUC score of **80.75%**.

Random Forest outperformed many models with **92.81%** accuracy, **89.51%** F1-score, Sensitivity of **85.80%**, Specificity of **96.70%**, and an AUC score of **88.44%**. However, the Bagging Classifier delivered superior performance with an accuracy of **97.39%**, F1-score of **85.88%**, Sensitivity of **81.96%**, Specificity of **69.07%**, and an AUC score of **95.34%**.

The best results were obtained with the CatBoost Classifier, which exhibited **98.96%** of **accuracy**, F1-score of **98.55%**, Recall of **98.04%**, Precision of **99.06%**, and an AUC score of **98.76%**.

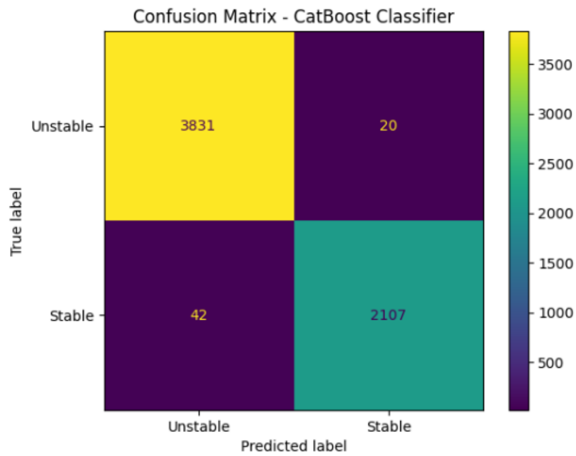

Figure 4 Confusion matrix for CatBoost Classifer

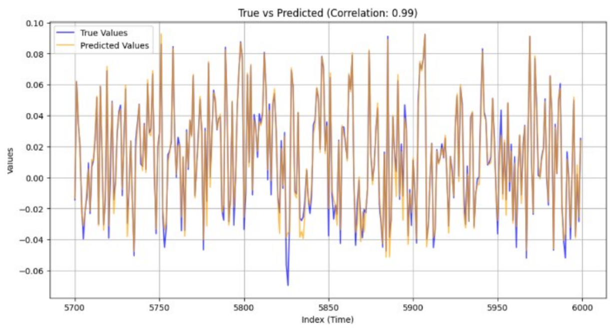

Figure 5 Comparison of True and Predicted values

The performance comparison plot in highlights the superiority of the Bagging Classifier and CatBoost Classifier over other models. The CatBoost Classifier demonstrated well-calibrated probabilities in reliability diagrams aligning closely with the diagonal identity line, indicating better probabilistic predictions compared to the other models.

5. Conclusion

The smart grid represents a transformative shift from traditional electromechanical power management to an advanced, electronically controlled infrastructure. With the increasing global population and economic expansion, electricity demand has surged, making efficient power distribution essential to minimize losses. Smart grids offer a solution by enhancing energy management and reducing inefficiencies in electricity delivery. Machine learning and computational intelligence plays a significant part in progressing request estimating and guaranteeing grid stability.

This research assessed multiple machine learning algorithms for forecasting smart grid stability and determined that the CatBoost classifier delivered the highest performance. The CatBoost classifier achieved the highest accuracy of **98.97%**, demonstrating superior performance over other models. The high AUC score further validates its ability to distinguish between stable and unstable grid conditions.

Future studies should aim to improve the robustness of the proposed model by incorporating supplementary techniques and investigating cutting-edge deep learning approaches. Expanding the dataset and incorporating real-time data processing could further improve smart grid resilience and efficiency.

References

[1] Shamik Tiwari., Anurag Jain., Kusum Yadav., Rabie Ramadan.,"Machine Learning-Based Model for Prediction of Power Consumption in Smart Grid," The International Arab Journal of Information Technology, Vol. 19, No. 3, May 2022

[2] Ahmed W., Ansari H., Khan B., Ullah Z., Ali S., Mehmood C., Qureshi M., Hussain I., Jawad M., Khan M., and Ullah A., "Machine Learning Based Energy Management Model for Smart Grid and Renewable Energy Districts," IEEE Access, vol. 8, pp. 185059-185078, 2020.

[3] Alazab M., Khan S., Krishnan S., Pham Q., Reddy M., and Gadekallu T., "A Multidirectional LSTM Model for Predicting The Stability of A Smart Grid," IEEE Access, vol. 8, pp. 85454 85463, 2020.

[4] Arif A., Javaid N., Anwar M., Naeem A., Gul H., and Fareed S., "Electricity Load and Price Forecasting Using Machine Learning Algorithms in Smart Grid: A Survey," in Proceedings of 34th International Conference Information on Advanced Networking and Applications, Caserta, pp. 471-483, 2020.

[5] Arzamasov V., Böhm K., and Jochem P., "Towards Concise Models of Grid Stability," in Proceedings of IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, Aalborg, pp. 1-6, 2018.

[6] Bashir A., Khan S., Prabadevi B., Deepa N., Alnumay W., Gadekallu T., and Maddikunta P., "Comparative Analysis of Machine Learning Algorithms for Prediction of Smart Grid Stability," International Transactions on Electrical Energy Systems, vol. 31, no. 9, pp. 1 23, 2021.

[7] Devi S. and Radhika Y., "A Survey on Machine Learning and Statistical Techniques in Bankruptcy Prediction," International Journal of Machine Learning and Computing, vol. 8, no. 2, pp.133-139, 2018.

[8] Fang X., Misra S., Xue G., and Yang D., "Smart Grid—The New and Improved Power Grid: A Survey," IEEE Communications Surveys and Tutorials, vol. 14, no. 4, pp. 944-980, 2011.

[9] Ghojogh B. and Crowley M., "Linear and Quadratic Discriminant Analysis: Tutorial," arXiv preprint arXiv:1906.02590, 2019.

[10] Ghosh A. and Kole A., "A Comparative Analysis of Enhanced Machine Learning Algorithms for Smart Grid Stability Prediction," TechRxiv, 2021.

[11] Gorzałczany M., Piekoszewski J., and Rudziński F., "A Modern Data-Mining Approach Based on Genetically Optimized Fuzzy Systems for Interpretable and Accurate Smart-Grid Stability Prediction," Energies, vol. 13, no. 10, pp. 25-59, 2020.

[12] Hafeez G., Alimgeer K., Wadud Z., Shafiq Z., Khan M., Khan I., Khan F., and Derhab A., "A Novel Accurate and Fast Converging Deep Learning-Based Model for Electrical Energy

Consumption Forecasting in A Smart Grid," Energies, vol. 13, no. 9, pp. 22-44, 2020.

[13] Jain A., Tiwari S., and Sapra V., "Two-Phase Heart Disease Diagnosis System Using Deep Learning," International Journal of Control and Automatio, vol. 12, no. 5, pp. 558-573, 2019.

[14] Krč R., Kratochvílová M., Podroužek J., Apeltauer T., Stupka V., and Pitner T., "Machine Learning-Based Node Characterization for Smart Grid Demand Response Flexibility Assessment," Sustainability, vol. 13, no. 5, pp. 29-54, 2021.

[15] Lamba R., Gulati T., Dhlan K. and Jain A., "A Systematic Approach to Diagnose Parkinson's Disease through Kinematic Features Extracted from Handwritten Drawings," Journal of Reliable Intelligent Environments, vol. 7, no. 3, pp. 253-262, 2021.

[16] Ma J., Feuerborn S., Black C., and Venkatasubramanian V., "A Comprehensive Software Suite for Power Grid Stability Monitoring Based on Synchrophasor Measurements," in Proceedings of IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, Washington, pp. 1-5, 2017.

[17] Malbasa V., Zheng C., Chen P., Popovic T., and Kezunovic M., "Voltage Stability Prediction Using Active Machine Learning," IEEE Transactions on Smart Grid, vol. 8, no. 6, pp. 3117-3124, 2017.

[18] Miraftabzadeh S., Foiadelli F., Longo M., and Pasetti M., "A Survey of Machine Learning Applications For Power System Analytics," in Proceedings of International Conference on Environment and Electrical Engineering and IEEE Industrial and Commercial Power Systems Europe, Genova, pp. 1-5, 2019.

[19] Moldovan D. and Salomie I., "Detection of Sources of Instability in Smart Grids Using Machine Learning Techniques," in Proceedings of 15th International Conference on Intelligent Computer Communication and Processing, Cluj Napoca, pp. 175-182, 2019.

[20] Panda D. and Das S., "Regression Analysis of Grid Stability under Decentralized Control," in Proceedings of International Conference on Engineering, Science, and Applications, pp. 1-6, Tokyo, 2019.

[21] Parseh M., Rahmanimanesh M., and Keshavarzi P., "Persian Handwritten Digit Recognition Using Combination of Convolutional Neural Network and Support Vector Machine Methods," The International Arab Journal of Information Technology, vol. 17, no. 4, pp. 572-578, 2020.

[22] Qinyu B., Lin Y., Jiayi M., Tianqi L., Xiaotian Z., and Youyin W., "Analysis of Influence with Connected wind Farm Power Changing and Improvement Strategies on Grid Voltage Stability," in Proceedings of China International Conference on Electricity Distribution, Tianjin, pp. 2029-2033, 2018.

[23] Rai S. and De M., "Analysis of Classical and Machine Learning Based Short-Term and Mid Term Load Forecasting for Smart Grid," International Journal of Sustainable Energy, vol. 40, no. 9, pp. 821-839, 2021.

[24] Rani P., Kumar R., Ahmed N., and Jain A., "A Decision Support System for Heart Disease Prediction Based Upon Machine Learning," Journal of Reliable Intelligent Environments, vol. 7, no. 3, pp. 263-275, 2021.

[25] Rani P., Kumar R., and Jain A., "HIOC: A Hybrid Imputation Method to Predict Missing Values in Medical Datasets," International Journal of Intelligent Computing and Cybernetics, vol. 14, no. 4, pp. 598-616, 2021.

[26] Sharma S., Challa R., and Kumar R., "An Ensemble-based Supervised Machine Learning Framework for Android Ransomware Detection," The International Arab Journal of Information Technology, vol. 18, no. 3A, pp. 422-429, 2021.

[27] Wang R., Liu Y., Ye X., Tang Q., Gou J., Huang, M., and Wen Y., "Power System Transient Stability Assessment Based on Bayesian Optimized Lightgbm," in Proceedings of IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2), Changsha, pp. 263 268, 2019.

[28] Yu W., An D., Griffith D., Yang Q., and Xu G., "Towards Statistical Modeling and Machine Learning Based Energy Usage Forecasting in Smart Grid," ACM SIGAPP Applied Computing Review, vol. 15, no. 1, pp. 6-16, 2015.