Implementation of Machine Learning Technique for Efficient Power Distribution & Utilization in Energy Sector of Pakistan

U. Bashir[1], M. Ali[2], A. Ali[3]

[1,2]Energy Engineering Department, University of Engineering and Technology, Taxila, Pakistan
[3]Water & Power Development Authority (WAPDA), Islamabad, Pakistan

[1]umair.bashir95@gmail.com

Abstract- Due to Pakistan's energy problems, including inefficient power distribution, frequent power outages, and rising demand that causes huge load shedding, this study looks into how machine learning methods can be used to improve power distribution for reduction in load shedding, also it gives us insight of generation fuel cost with respect to GDP of Pakistan using MATLAB. In this research paper we use data of Pakistan energy sector of 144 month from January 2010 to December 2021. With the help of trend analysis of data, we train our model and implement it for reduction of load shedding. For per unit cost relation with GDP we use MATLAB analysis. With this analysis we can easily see which fuel type is feasible for our GDP. The study concludes with a findings and recommendations, providing a valuable contribution to the field and offering practical insights for policymakers, industry professionals, and researchers alike.

Keywords- Machine Learning (ML), Mega Watt (MW), Mega Watthour (MWh), Deep Learning (DL), Distribution Companies (DISCOs), Central Power Purchasing Agency (CPPA), Water and Power Development Authority (WAPDA), Autoregressive Integrated Moving Average (ARIMA Model)

I. INTRODUCTION

An essential asset for modern economies is energy. Consequently, the technologies employed in its production and utilisation have gained prominence. Securing a pristine, economically efficient, and reliable energy provision in various manifestations is inextricably interconnected considering both environmentally friendly development and energy security. Having said that, attaining this objective would need a maximum allocation of renewable energy within the entirety of the energy composition. Meanwhile, the exponential surge in energy consumption is compelling legislators and planners to address this matter.

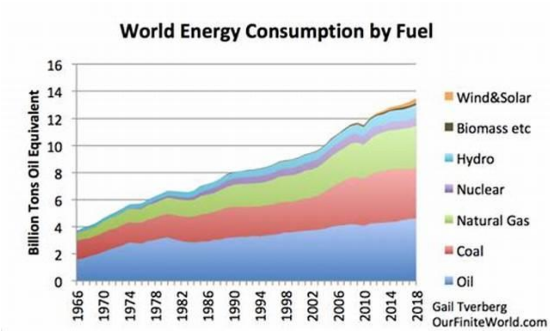

Figure 1 World Energy Consumption by Fuel

According to reports, between 1990 and 2019, the world's primary energy supply grew by 66%, while at the same time, power consumption soared to almost 130%[1].
Fossil fuels, a major contributor to the climate-altering greenhouse gas (GHG) effect, provide the vast majority of the world's energy. The two predominant fossil fuels, petroleum and coal, provide around 57% of the global energy provision[2].
Developing countries are lucky to own sufficient renewable energy resources., but poor economic conditions, political instability and inequitable energy planning are the main and most daunting challenges they face in meeting energy demand. Under such circumstances, these nations may benefit from depending on both conformist energy planning and conventional energy supplies. Therefore, in order to achieve sustainable growth and energy security, energy planning and regulations that take into account the best possible penetration of renewables in the entire energy mix are essential.[2]

1.1 Pakistan's Energy Sector

Comparable to several emerging countries, Pakistan has been encountering significant energy issues for over two decades. The shortage of power in the country has resulted in a reduction of production and even drove some businesses to shut down, resulting in worsening unemployment. The nation's limited natural gas and oil supplies are depleting rapidly, while the production of coal and other sources of clean energy requires a significant amount of time. With a population of over 200 million and significant expansion over the past 20 years, Pakistan has an abundance of renewable energy resources; yet poor planning and administration have resulted in a disastrous energy crisis[3]. Consequently, the country utilizes a mere 452 kilowatt-hours of power per capita, which is almost a quarter of the global average use. Due to excessive heat and humidity during the summer months in Pakistan's lower regions, there is a surge in demand for energy beyond generation and transmission capacity. As a result, load shedding occurs, with metropolitan regions experiencing it for 8–12 hours and rural areas for 12–18 hours[4]. The energy scarcity has resulted in the closure of several industries or the adverse effects on their industrial output, so diminishing the living conditions of numerous households and exerting a substantial detrimental influence on the economy. The nation has experienced a decline of almost 10% in its overall GDP expansion. If Pakistan's energy deficit is not addressed quickly, it may pose a danger to the nation's energy security[5].

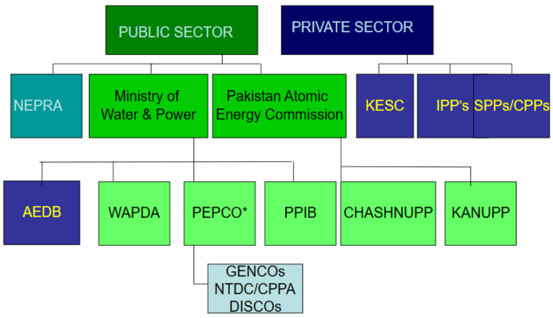

Figure 2 Players in Power Sector of Pakistan

In order to satisfy the growing demand for energy, governments and academic institutions worldwide are engaged in extensive energy planning exercises. Energy modelling is being used to assist in the development of energy policy. These initiatives facilitate the review of policy alternatives across multiple tiers and in response to diverse scenarios. Researchers have also focused a lot of effort on applying energy modelling as the best and most appropriate way to conduct energy planning and policy analysis[6].

It is also important to note that Pakistan has abundant renewable energy resources at its disposal. It wasn't until recently that a scientific estimate of these resources was attempted. Multiple attempts in renewable energy modelling are necessary to evaluate these renewable energy sources' realistic potential and their long-term diffusion in order to harness them.

1.1.1 Pakistan's Energy Crisis

Pakistan has not generated enough energy over the past several decades to fulfil its increasing demand, which has put the nation's power sector in risk. Pakistan relies on imported gas and oil for its energy needs[7].

The Asian Development Bank proclaimed in a 2019 white paper that Pakistan is a country that doesn't have sufficient energy security. In addition to Pakistan, a great number of other countries, even wealthy ones, are dealing with energy insecurity[1]. In 2021, a circular debt of PKR 2.8 trillion, representing 52% of total debt, will have resulted from the incapacity of distribution sector; operating companies (DISCOs) to collect revenue and achieve high recovery rates, thereby placing Pakistan's distribution sector under a financial crisis[8].

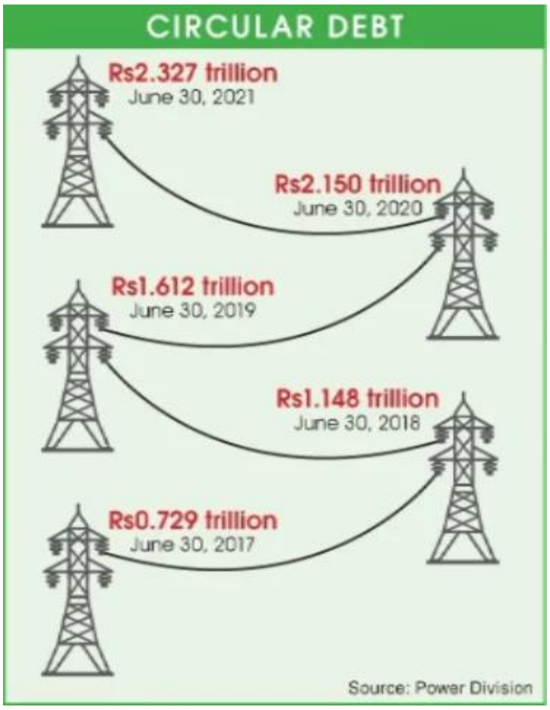

Figure 3 Circular Debt of Pakistan

Inefficiencies in the nation's power system, such as feeder overloads, overloaded transformers, and poor line conditions, have led to frequent power outages and an unstable power supply. Population expansion has increased demand for electrical connections, but bureaucratic culture, corruption, and technological barriers have made it difficult for DISCOs to provide connections in the time frame given. In order to enhance the performance of DISCOs and regulate the supply-demand gap, load shedding has been implemented; DISCOs with lower rates of electricity theft, line losses, and recovery rates experience shorter outages[9].

Over the past decade, Pakistan has been faced with energy crises as a result of inadequate energy planning and governance concerns. Energy-related issues in the nation have resulted in the closure of thousands of businesses, impacting both industrial output and poor survival for thousands of people[3]. With an estimated cost of 10% of GDP over the previous five years, the energy crisis has been a significant drag on the economy and a substantial growth hindrance. If Pakistan's energy problem is not addressed promptly and effectively at both the operational and strategic levels, it has the potential to evolve into a significant national security concern[10].
Only 2–4% of the energy mix comes from these renewable sources, which have not yet reached their full potential[11].

1.2 Alternative and Renewable Energy Policy 2019 Pakistan

The primary objective of the 2019 policy was to facilitate and support the nation's advancement in the utilisation of sustainable energy sources. To suit domestic demand, Pakistan generates a very tiny percentage of its total oil supply. Domestic oil production is subject to financial, technological, and technical limitations. The latest available data indicates that the expenditure on oil imports witnessed a significant surge of 95.9% over the period from July to April in the fiscal year 2022, escalating from US$8.69 billion to US$17.03 billion[3].
Energy is becoming more expensive due to rising global oil prices and the steep devaluation of the Pakistani rupee. This puts pressure on the country's export industry and worsens its trade deficit. In a similar vein, the importation of LNG (liquefied natural gas) had a substantial surge of 82.90% over the period spanning from July to April in FY2022, in contrast to a more modest 39.86% rise in LPG imports.
Pakistan is one of the countries that produces a growing amount of power from nuclear technology. Between July and March of FY2022, the gross capacity of nuclear power reactors rose by 39% to 3,530 MW, providing 12,885 million units of electricity to the national grid[11].

1.3 Power Distribution System of Pakistan

At 132 kV and lower, DISCOs are in charge of the management and upkeep of the transmission and distribution infrastructure[9].
The sufficiency of three key components is the primary determinant of power delivery across Discos' networks:

1. Distribution transformers.
2. Power Transformers.
3. 11 kV feeders.

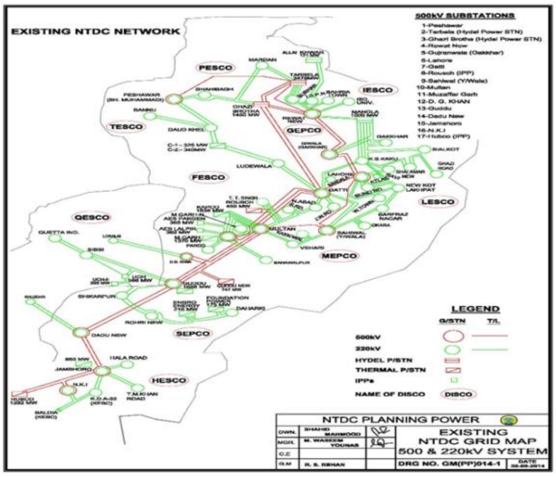

Figure 4 Power Distribution Network of Pakistan

1.4 Machine Learning

Machine learning is a branch of artificial intelligence (AI) that uses trained algorithms on data sets to create self-learning models that can categorise data and forecast events without human intervention. In modern times, the use of machine learning extends to a wide array of commercial efforts, including tasks such as language translation, forecasting fluctuations in the stock market, and providing clients with personalized product suggestions based on their past acquisitions[12].
Due to the extensive use of machine learning in modern life for the purposes of artificial intelligence, the terms "machine learning" and "artificial intelligence" are commonly utilized interchangeably[13]. However, the two words are essentially distinct. Machine learning, as contrast to artificial intelligence (AI), specifically pertains to the utilisation of algorithms and datasets in pursuit of the overarching objective of developing machines possessing cognitive capabilities like to those exhibited by humans.
We can apply machine learning to examine both historical and real-time data since we have access to data from a wide range of sources. In order to anticipate an optimal merit, machine learning algorithms are also more adept at accounting for all the numerous aspects that affect the price, including as demand, weather, energy availability from several sources, past usage, etc[13]. This enables you to choose your source of power with greater knowledge. Since it can be challenging to ensure the supply of electricity from various sources, this is particularly useful in markets where there is a lot of renewable energy, like wind[14].

II. RELATED WORKS, MOTIVATION AND CONTRIBUTION

Ensuring that power generation and demand are always in balance is a significant priority in the management of power systems. Machine learning

has the potential to enhance the longevity and balance of renewable energy systems. [15]
The models are categorised into two groups: knowledge-driven and data-driven. These models are used to anticipate and assist energy systems. Knowledge-based approaches are often created using a thorough comprehension of systems and their mechanics. As a result, they rely heavily on experience and information, which might lead to a higher likelihood of errors. Nevertheless, data-driven methods, particularly those utilising artificial intelligence, have gained significant attention due to their ability to operate without prior knowledge or complex analytical models. These methods can be effectively employed with suitable data and a general understanding of the system. Today, the issue of insufficient data is nearly resolved, which may result in increased interest and focus in this area. Thus, the latest and most useful models that rely on data and artificial intelligence, such as machine learning and deep learning, have gained immense popularity and are extensively utilised[16].
Typically, when it comes to matters concerning machine learning, we come across two different viewpoints. If we want to include the temporal aspect of the data in the problem, we need additionally look at it from the standpoint of temporal Series (TS). The relevant section provides known application strategies for TS. In forecasting vital parameters, it is necessary to include the time dimension. To produce accurate results, the model should integrate both traditional and fundamental machine learning methods as well as time series algorithms. When it comes to renewable systems, it's vital to note that as renewable energy becomes more popular, it's becoming more and more crucial to include time-varying input data in energy system optimisation studies. In recent years, time-series aggregation has been popular as a way to simplify models and tackle this difficulty[17].
Optimisation plays a crucial role in the design, analysis, control, and operation of real-world systems. Optimisation also includes the task of determining the best suitable objective, variables, and limitations. The objective is to choose a model that provides valuable understanding of the current practical issues and creates a scalable algorithm that can discover a proven optimum (or nearly optimal) solution within an acceptable timeframe. ML has been influenced by recent advancements in contemporary optimisation[18].
The findings indicate that the learning algorithm based on artificial neural networks is both more precise and computationally economical compared to conventional approaches in expressing the performance of optimisation. Overall, the results indicate that teaching-learning approaches are more effective than methods like PSO when it comes to achieving optimal energy output[19].

Every algorithm performs better in specific situations; thus, we should search for the optimal method based on accuracy and error rate for the desired application. For instance, in the area of predicting global solar radiation, four algorithms were studied: SVM, KNN, DL, and ANN. The study's findings indicate that the ANN method demonstrated high accuracy.

III. RESEARCH METHODOLOGY

Research is focused on identifying a solution to a pressing issue for society or a commercial entity. Applied research aims to solve practical problems, while basic research focuses on expanding scientific knowledge with broad applicability. Real data from reliable and official sources was used for research. The annual reports and energy data files were used for data mining. It also talks about almost every part of Pakistan's current power economy, including demand, load, and generation. Proposed Methodology for research is explained in flow chart as illustrated in;

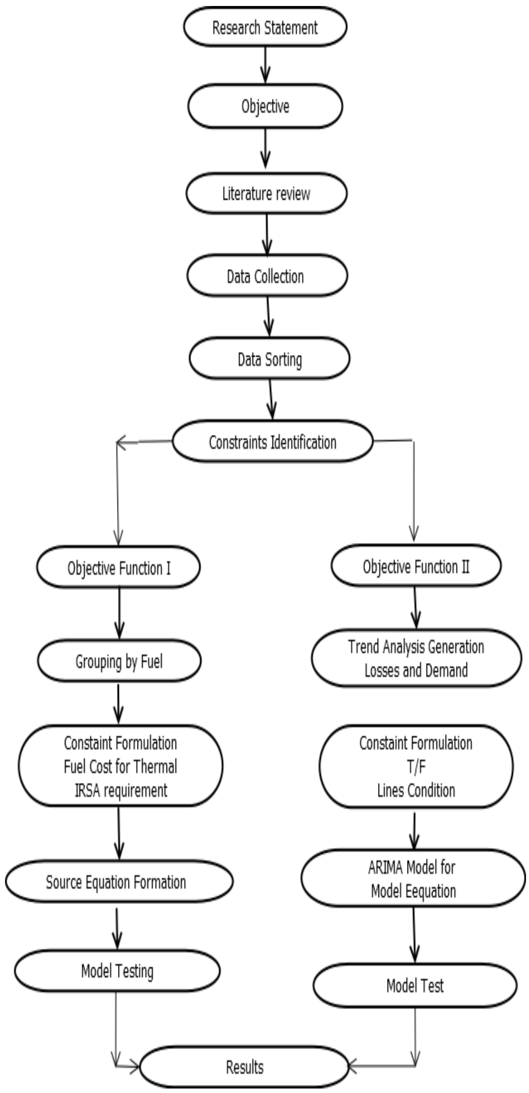

Figure 5 Flow Diagram for Methodology

Objective Function-I

Data on energy generation from various sources including Thermal, Hydel, Nuclear, Solar, Bagasse, and Wind has been collected over a specific duration to assess their respective impacts on GDP while considering the per unit production cost of each source. This comparative analysis aims to evaluate the efficiency and economic viability of different energy sources in relation to their contribution to GDP growth. By examining the percentage share of energy production from each source and analysing their associated production costs, insights can be gained into the optimal utilization of resources to drive economic development. This analysis will provide valuable information for policymakers, energy planners, and stakeholders to make informed decisions regarding energy policy, investment priorities, and sustainable development strategies aimed at fostering economic growth while ensuring energy security and environmental sustainability. Data of energy generation and GDP are correlated, and best fitting model has been adopted by comparing coefficient of correlation.

Following Equation has been developed for defining relation between GDP and Total generation from 2007 to 2021.

$$y = 2.955 \times 10^{-29}x^7 + 2.412 \times 10^{-23}x^6 - 8.406 \times 10^{-18}x^5 + 1.622 \times 10^{-12}x^4 - 1.871 \times 10^{-7}x^3 + 0.0129x^2 - 493.3x + 8.051 \times 10^6$$

Where y represents GDP and x represents total generation.

For finding out trend of GDP with hydel generation, following equation has been established.

$$y = 0.003254x^7 - 0.6204x^6 + 50.53x^5 - 2278x^4 + 6.143 \times 10^4x^3 - 9.9 \times 10^5x^2 + 8.831 \times 10^6x - 3.363 \times 10^7$$

Where y represents GDP and x represents percentage contribution of hydel in total power generation. Another equation showing relation of GDP with percentage of thermal generation has also been found, and is given by:

$$y = -0.001033x^4 + 0.3962x^3 - 51.18x^2 + 2765x - 5.384 \times 10^4$$

Where y represents GDP and x represents percentage contribution of thermal in total power generation.

Objective Function-II

Data on energy generation, load, energy demand by DISCOs, over loaded feeders, over loaded and transformer has been collected over a specific duration to assess their respective impacts on load shedding and losses. This comparative analysis aims to evaluate the efficiency and economic viability of our distribution system. By examining the trends of our energy data over the last 144 months and analysing their correlation with losses This analysis will provide valuable information for policymakers, energy planners, and stakeholders to make informed decisions regarding energy policy, investment priorities, and sustainable development strategies aimed at fostering economic growth while ensuring energy security and environmental sustainability.

Data attached at end of this research work in appendix. This study is aimed at finding relation between distribution system over load, losses and load shedding.

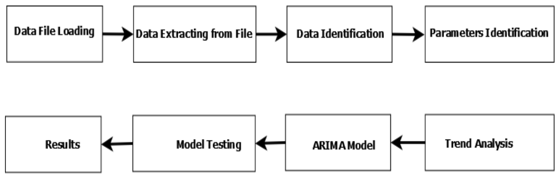

Figure 6 Block diagram of ML program

For this purpose, we have tried to find correlation between electricity production, losses and load shedding data of energy Jan 2010 to Dec 2021. To analyse our data which is monthly data spanning a 12-year period, employing machine learning technique offers a robust methodology for uncovering underlying trends. Below are trends over given span.

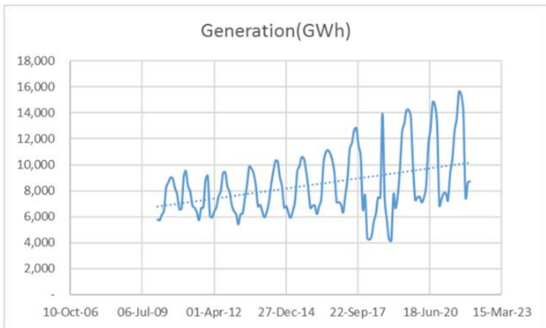

Figure 7 Energy Generation Trend

Below, the trend plot illustrating these dynamics is presented in the figure;

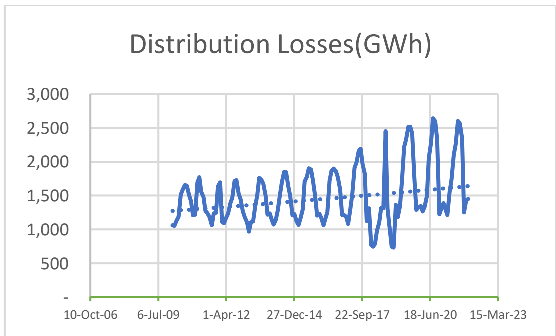

Figure 8 Trend of Distribution Losses

Data is provided by National Power Control Centre. Analysing our 12-year monthly dataset by using machine learning techniques for revealing hidden patterns. This approach involves thorough data preprocessing to ensure accuracy and completeness, followed by feature extraction to effectively capture temporal information. Presented below are the trends observed over the specified time span.

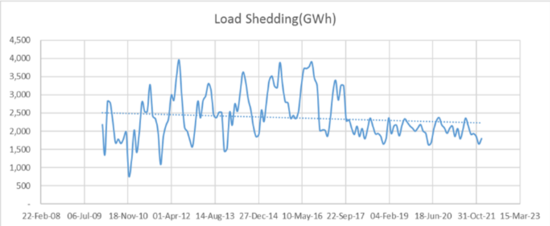

Figure 9 Load shedding Trend

Model Development

As we analysis the trends of our energy generation, distribution losses, energy supplied to DISCOs, distribution losses, over loading of transformers and load shedding over the past 144 months. As we are satisfied with our model trend analysis, we deploy it as a predictive tool that can be integrated into existing systems used by power utility companies and government agencies for load shedding management. This involves implementing the model in a production environment and providing documentation and support for its ongoing use. Now we know trends of all data sets we have now we will use ARIMA model to design our model. Below is analysis of our data on the basis of previous trends;

Stationarity Check for Data:

ADF Statistic: -1.5518165872844143

p-value: 0.5077406613076128

Critical Values:

1%: -3.481

5%: -2.884

10%: -2.579

Data is non-stationary.

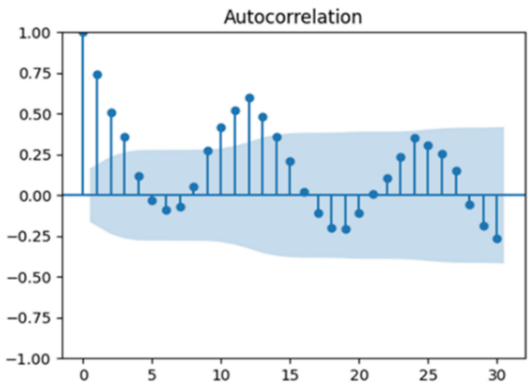

Figure 10 Autocorrelation Graph

The autocorrelation graph illustrates the correlation between a time series and its lagged values. Suggesting a strong positive autocorrelation with a lag of one year.

The model equation is;

Forecast_t = -0.06 + -0.06 * Forecast_t-1 + -0.12 * Forecast_t-2 + 0.15 * Forecast_t-3 + -0.21 * Forecast_t-4 + -0.18 * Forecast_t-5

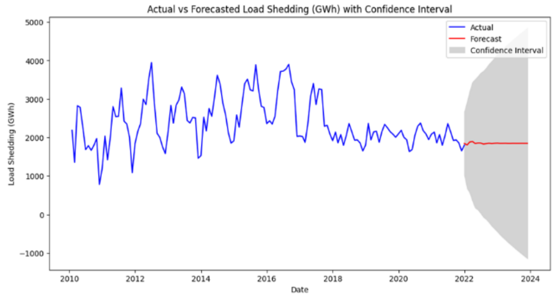

Figure 11 Forecast Load Shedding Using ARIMA Model

An accuracy improvement of 0.1038 indicates that the ARIMA model has reduced the Mean Absolute Error (MAE) by approximately 10.38% compared to the baseline model.

IV. RESULTS AND DISCUSSION

Firstly, we will cover the objective function-I. Discuss it's results which we get using MATLAB.

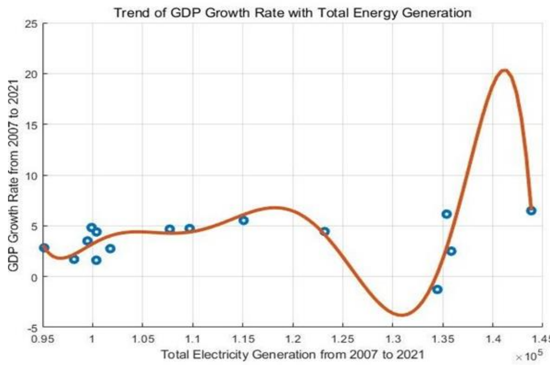

Figure 12 Trend of GDP with total Generation in GWH

From Figure 11; Once the minimum threshold of electricity is met, further increases in electrical generation lead to proportional increases in GDP, reflecting a linear relationship between the two variables. However, as electrical generation continues to increase beyond a certain point, the additional economic benefits become less definite, resulting in the sine wave pattern.

From our Figure 12; we clearly see that, the initial rise in GDP with low hydropower generation indicates that hydropower is a critical factor driving economic growth in the region. However, as hydropower generation increases beyond a certain point, it reaches the capacity limits of existing infrastructure or natural resources. Trend of GDP with Percentage of Thermal in Total Power Generation in GWH

After plotting the graph of GDP vs Percentage of thermal in total power generation in Figure 13; we can see the initial brick rise in GDP could represent a period of significant infrastructure investment in thermal power generation.

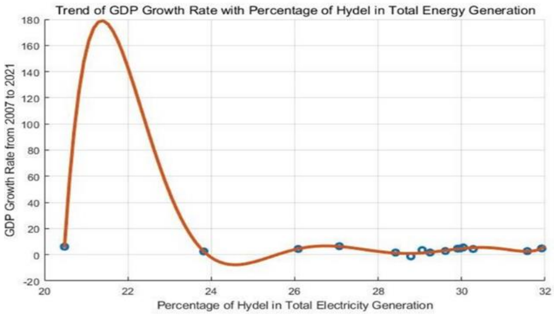

Figure 13 Trend of GDP with Percentage of Hydel in Total Power Generation in GWH

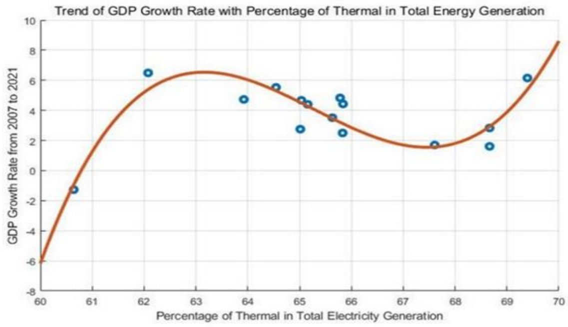

Figure 14 Trend of GDP with Percentage of Thermal in Total Power Generation in GWH

The negative value at the beginning of the curve could indicate a period of disinvestment or economic downturn preceding the investment phase. As thermal power infrastructure is developed and brought online, GDP experiences a sharp increase.

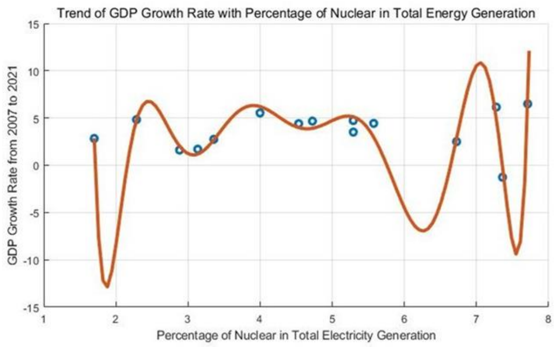

Figure 15 Trend of GDP with Percentage of Nuclear in Total Power Generation in GWH

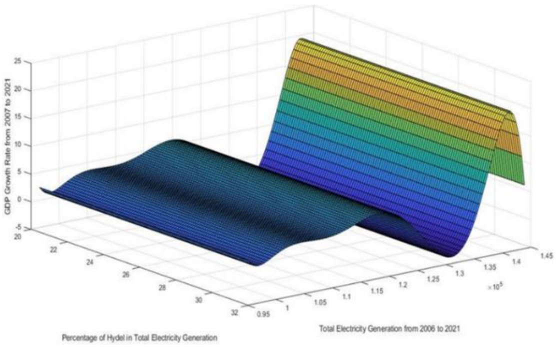

Figure 16 Trend of GDP with Total Generation and Percentage of Hydel Power in GWH

As the order of the equation is increased, R2 value keeps on increasing but as suggested by the above figures, there is no specific trend followed by the graph and determining relation between the two quantities becomes difficult. In this analysis, relation of per unit cost on GDP growth rate is estimated through its variation with type of source as the contribution of type of source has great impact on determining per unit cost. The percentage of hydel power is related to GDP growth rate with an equation of 7th having R2 value of 0.8, which shows strong positive correlation but still order of equation makes it difficult to determine exact trend both the quantities follow. Similarly, thermal production seems to show a better correlation with 4th order equation having R2 value of 0.76.

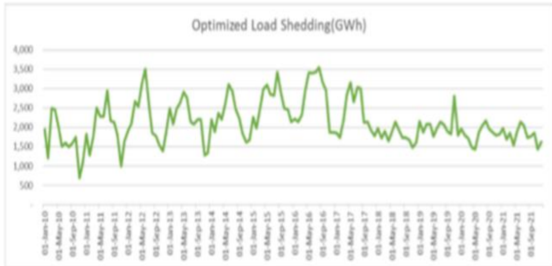

Figure 17 Optimized Load shedding after applying model

From Figure 16; we clearly see that; after applying our AI model loadshedding is reduced by 9%.

V. Conclusions and Future Recommending

The model presented in this thesis have the following benefits for power producers, energy-users and the government:

Surety of a reliable power supply for the consumers. The available domestic-power-plants can also be optimized. It is advisable to optimise each energy source separately for plant capacity assessment, focusing on upgrading and overhauling plants to maximise their capabilities and ensuring they are readily available for the load dispatch centre.

It needs special attention that as Pakistan is considering coal as an alternative for energy modification. The country aims to develop ultra-supercritical plants to increase the amount of power generation from coal, emphasising the need of efficiency in fuel consumption.

Assess grid resilience with machine learning. Evaluate the grid's ability to withstand disruptions and identify vulnerabilities. Implement resilience-enhancing measures to minimize losses during outages.

Address power quality issues with machine learning. Detect and mitigate voltage harmonics, disturbances, and other quality issues to reduce equipment damage and energy losses.

ISSN:1813-1786 (Print) 2313-7770 (Online)

Implement real-time voltage control systems. Utilize machine learning algorithms to adjust transformer taps and capacitor banks to maintain voltage within specified limits. This enhances power quality and reduces losses.

Optimize renewable energy integration with machine learning. Predict renewable energy generation patterns and use machine learning to enhance the utilization of energy storage systems. This minimizes curtailment and improves grid efficiency.

Use reinforcement learning for grid optimization. Apply machine learning algorithms to find the most efficient grid configurations and operation strategies, minimizing losses and improving overall performance.

With the rise of decentralized energy resources like solar and wind, machine learning will be used to optimize the integration of these sources, reducing losses and improving grid stability

Governments and regulatory bodies are likely to encourage the adoption of machine learning technologies through policies and incentives to improve grid efficiency and reduce losses.

Collaboration between utilities, technology providers, and research institutions will accelerate the development and implementation of machine learning solutions, fostering innovation in the sector.

REFERENCES

[1] ‘2019 Asian Development Bank Annual Report’, Asian Development Bank, May 2020. doi: 10.22617/FLS200124.

[2] ‘Alternative Energy Development Board’. [Online]. Available: https://mail.aedb.org/

[3] ‘pakistan-energy-sector’. [Online]. Available: https://www.paradigmshift.com.pk/pakistan-energy-sector/

[4] ‘CTBCM Detailed Design Report’. [Online]. Available: https://nepra.org.pk/Admission%20Notices/2020/08%20Aug/CTBCM/CTBCM%20Detailed%20Design%20Report.pdf

[5] ‘Tribal Areas Electric Supply Company (TESCO)’. [Online]. Available: https://www.tesco.org.pk/

[6] ‘National Electric Power Regulatory Authority’. [Online]. Available: https://nepra.org.pk/

[7] ‘Survey of Pakistan’. [Online]. Available: http://surveyofpakistan.gov.pk/

[8] ‘IPPs’. [Online]. Available: https://dailytimes.com.pk/1131871/the-1990s-and-the-ipps/

[9] ‘WAPDA Website’. [Online]. Available: https://www.wapda.gov.pk/

[10] ‘NTDC’. [Online]. Available: https://www.ntdc.gov.pk/the-company

[11] ‘Energy policy 2019’. [Online]. Available: https://nepra.org.pk/Policies/ARE_Policy_2019_-_Gazette_Notified.pdf

[12] E. Barbierato and A. Gatti, ‘The Challenges of Machine Learning: A Critical Review’, *Electronics*, vol. 13, no. 2, p. 416, Jan. 2024, doi: 10.3390/electronics13020416.

[13] T. Ahmad and H. Chen, ‘Utility companies strategy for short-term energy demand forecasting using machine learning based models’, *Sustainable Cities and Society*, vol. 39, pp. 401–417, May 2018, doi: 10.1016/j.scs.2018.03.002.

[14] M. Alizamir, S. Kim, O. Kisi, and M. Zounemat-Kermani, ‘A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: Case studies of the USA and Turkey regions’, *Energy*, vol. 197, p. 117239, Apr. 2020, doi: 10.1016/j.energy.2020.117239.

[15] T. Ahmad, R. Madonski, D. Zhang, C. Huang, and A. Mujeeb, ‘Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm’, *Renewable and Sustainable Energy Reviews*, vol. 160, p. 112128, May 2022, doi: 10.1016/j.rser.2022.112128.

[16] M. M. Forootan, I. Larki, R. Zahedi, and A. Ahmadi, ‘Machine Learning and Deep Learning in Energy Systems: A Review’, *Sustainability*, vol. 14, no. 8, p. 4832, Apr. 2022, doi: 10.3390/su14084832.

[17] H. Teichgraeber and A. R. Brandt, ‘Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities’, *Renewable and Sustainable Energy Reviews*, vol. 157, p. 111984, Apr. 2022, doi: 10.1016/j.rser.2021.111984.

[18] R. Zahedi, A. Ahmadi, R. Eskandarpanah, and M. Akbari, ‘Evaluation of Resources and Potential Measurement of Wind Energy to Determine the Spatial Priorities for the Construction of Wind-Driven Power Plants in Damghan City’, *13*, vol. 11, no. 1, pp. 1–22, Feb. 2022, doi: 10.18488/13.v11i1.2928.

[19] Y. Zhou, S. Zheng, and G. Zhang, ‘Artificial neural network based multivariable optimization of a hybrid system integrated with phase change materials, active cooling and hybrid ventilations’, *Energy Conversion and Management*, vol. 197, p. 111859, Oct. 2019, doi: 10.1016/j.enconman.2019.111859.