SMART LABORATORY REAL-TIME MONITORING AND PREDICTION

SALAVA V SATYANARAYANA
Associate Professor, Department of EEE, Hyderabad institute of technology and management, Gowdavally, Medchal-Malkajgiri, Hyderabad, Telangana 500043
9828568689
Pshve2011@gmail.com
AKHILSAI VODNALA
Department of EEE, Hyderabad institute of technology and management, Gowdavally, Medchal-Malkajgiri, Hyderabad, Telangana 500043
akhilvodnala30@gmail.com
RAPELLI SANDEEP
Department of EEE, Hyderabad institute of technology and management, Gowdavally, Medchal-Malkajgiri, Hyderabad, Telangana 500043
Sandeeprapelli767@gmail.com
MD.WASIM AKHRAM ALI
Department of EEE, Hyderabad institute of technology and management, Gowdavally, Medchal-Malkajgiri, Hyderabad, Telangana 500043
wasimakhramali@gmail.com
S.SANDEEP REDDY
Department of EEE, Hyderabad institute of technology and management, Gowdavally, Medchal-Malkajgiri, Hyderabad, Telangana 500043
Sandeepreddysaddala@gmail.com
PILLALAMARRI MADHAVI
Assistant Professor, Department of EEE, Hyderabad Institute of Technology and Management, Gowdavally, Medchal-Malkajgiri, Hyderabad, Telangana 500043
madhavip.eee@hitam.org

Abstract—This project emphasizes renewable energy, predictive maintenance, and insulation health monitoring. It examines solar energy generation trends based on various states of India to identify regional variations and efficiency. It also entails forecasting the lifespan of electric motors based on performance data, minimizing unexpected failures and maintenance expenses. Moreover, dielectric oil testing is performed to determine the quality and aging of transformer oil for safe and reliable transformer operation. This collaborative research seeks to improve the sustainability, reliability, and efficiency of electrical systems in India.[1].

Keywords— *Solar Energy Forecasting, Motor Life Prediction, Machine Learning, Renewable Energy, Predictive Maintenance*

I. INTRODUCTION

Power systems are the core of everyday life and industrial operations, and their efficiency, safety, and reliability are the central concerns of electrical engineering. The three topics mentioned above in this project are critical areas: solar energy production, motor life prediction, and transformer oil testing. We have compared solar energy generation in different states of India to determine the geographical variations in solar power generation. This comparison assists in determining the high-potential solar energy states and suggests ways to enhance the utilization of renewable energy.

To forecast the motor lifespan, we gathered operating conditions like temperature, voltage (V), current (I), and frequency (Hz) of operational motors. We applied supervised machine learning techniques to estimate how much longer the motors can operate effectively. The model forecasts the motor lifespan so that industries can plan maintenance ahead and reduce the risk of abrupt motor failure and enhance efficiency[2].

II. PROPOSED METHODOLOGY

A. Block Diagram

The block diagram shows the fundamental steps in a supervised machine learning prediction model. Initially, data collection acquires the necessary information required for predictions. In the second step, data preprocessing involves cleaning and organizing the data to ensure it is ready for analysis.

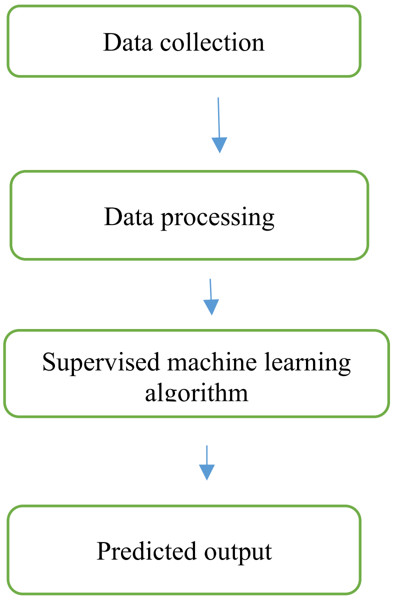

Fig 2.1 Basic Block diagram of forecasting data

B. PREDICTION OF SOLAR ENERGY GENERATION IN VARIOUS STATES OF INDIA

The "Prediction of Solar Energy Generation in Different States of India" project is to predict solar power generation trends in different states of India based on data science models. State-specific predictions are required because climatic, geographical, and infrastructural conditions vary. Polynomial regression with Ridge regularization, XGBoost and supervector machines is utilized in this project to predict solar energy production between 2017 and 2022. 2017-2021 data is used as the model base, 2021-2022 data to test the model, and for predicting the 2022-2023 output.

A.Code Implementation

To predict for the year 2021-22:

A.Polynomial + Ridge regression

```
import pandas as pd
import numpy as np
from sklearn.linear_model import Ridge
from sklearn.preprocessing import PolynomialFeatures
from sklearn.pipeline import make_pipeline
from sklearn.metrics import mean_absolute_error, r2_score
import matplotlib.pyplot as plt
# Load data
df = pd.read_csv("/content/Generation.csv")  # Replace with your CSV filename
df.columns = df.columns.str.strip()
# Define input and target columns
year_cols = ['2017-18', '2018-19', '2019-20', '2020-21']
target_col = '2021-22'
X_years = np.array([2017, 2018, 2019, 2020]).reshape(-1, 1)
# Initialize lists
predicted = []
actual = []
states = []
# Model pipeline: PolynomialFeatures + Ridge Regression
model = make_pipeline(PolynomialFeatures(degree=3), Ridge(alpha=30))
```

```
# Fit and predict per state
for _, row in df.iterrows():
    y = row[year_cols].astype(float).values
    model.fit(X_years, y)
    y_pred = model.predict([[2021]])[0]
    predicted.append(y_pred)
    actual.append(float(row[target_col]))
    states.append(row['State/UT'])
# Compile results
results = pd.DataFrame({
    'State/UT': states,
    'Actual 2021-22': actual,
    'Predicted 2021-22': predicted,
})
# Metrics
mae = mean_absolute_error(actual, predicted)
r2 = r2_score(actual, predicted)
# Save to CSV
results.to_csv("improved_prediction_2021.csv", index=False)
# Print table and metrics
print("\n📄 Improved Prediction Table for 2021–22:\n")
print(results.to_string(index=False))
print(f"\n📈 MAE: {mae:.2f}")
print(f"📈 R² Score: {r2:.4f}")
```

B. Prediction of Lifespan of Single-Phase Induction Motor

This project focuses on predicting the remaining useful life (RUL) of a single-phase induction motor based on real-time electrical and ambient data.

a. Type of learning approach

The target variable "motor life remaining" was used to train regression models. Key features included voltage, current, power, frequency, temperature, and speed. Three models were evaluated: Random Forest, XG Boost, and Light GBM.

b. Code Implementation

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
from sklearn.metrics import mean_absolute_error, r2_score
# Load dataset
file_path = "/content/NewData.csv"  # Update path if needed
df = pd.read_csv(file_path)
# Convert timestamp to datetime
df["timestamp"] = pd.to_datetime(df["timestamp"])
# Rename incorrect column name if exists
df.rename(columns={"temparature": "temperature"}, inplace=True)
# Encode categorical columns: 'status', 'anomaly', and 'load_level'
label_encoders = {}
```

```
categorical_columns = ["status", "anomaly", "load_level"]  # Add 'load_level' for encoding
for col in categorical_columns:
    if col in df.columns:
        label_encoders[col] = LabelEncoder()
        df[col] = label_encoders[col].fit_transform(df[col])  # Convert categorical to numeric
# Define feature set (now including 'load_level')
features = ["voltage", "current", "power", "frequency", "temperature", "speed", "anomaly", "load_level"]
features = [col for col in features if col in df.columns]  # Ensure all columns exist in the dataset
    df["motor_life_remaining"] = np.abs(np.random.normal(loc=500, scale=100, size=len(df)))  # Random life
hours
# Drop missing values
df = df.dropna()
# Train-Test Split
X_train, X_test, y_train, y_test = train_test_split(df[features], df["motor_life_remaining"], test_size=0.2,
random_state=42)
# Train the Random Forest Model
model = RandomForestRegressor(n_estimators=300, max_depth=15, random_state=42)
model.fit(X_train, y_train)
```

III. RESULT &DISCUSSIONS

A. PREDICTION OF SOLAR ENERGY GENERATION IN VARIOUS STATES OF INDIA

To predict for the year 2021-22:

1. Polynomial + Ridge regression

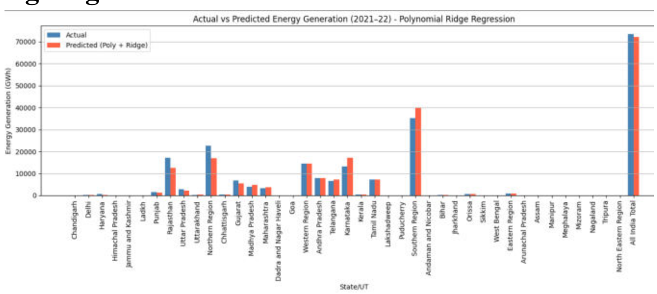

Figure 7 Actual vs Predicted Energy Generation (2021–22) – Polynomial Ridge Regression

In the bar graph "Actual vs Predicted Energy Generation (2021–22) – Polynomial Ridge Regression" the predicted values derived from a Polynomial Ridge Regression model are contrasted with the actual energy generation in gigawatt-hours (GWh) for different states of India. The red bars show the predicted values, and the blue bars show the actual energy generation.

2. XGBoost

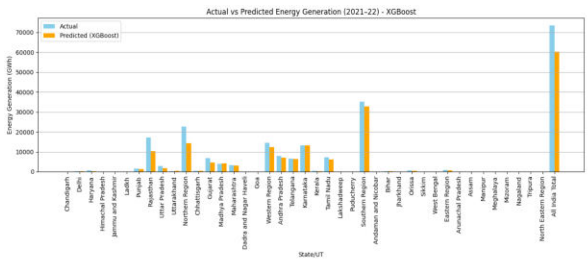

Figure 8 Actual vs. Predicted Energy Generation (2021–22) – XGBoost

"Actual vs. Predicted Energy Generation (2021–22) – XGBoost" compares the actual and predicted energy generation values across several states of india.The XGBoost model's predicted values are displayed in orange bar, while the actual values are represented as blueb.

3. SVR

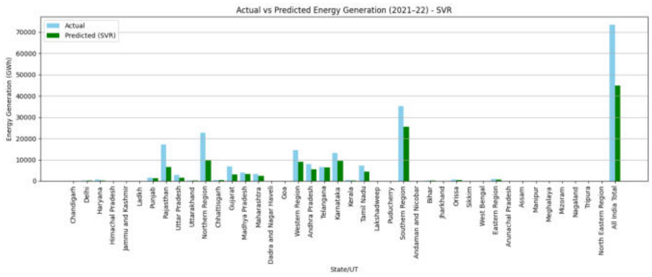

Figure 9 Actual vs Predicted Energy Generation (2021–22) – SVR

"Actual vs Predicted Energy Generation (2021–22) – SVR" shows the actual and predicted energy generation for states of india based on Support Vector Regression (SVR). The light blue bars indicate actual values, and the dark green bars indicate predictions by the SVR model. Unlike XGBoost, SVR shows evident underestimation in most of the areas, particularly in highly energy-generating states like Gujarat, Maharashtra, Tamil Nadu, and the "All India Total." The projected bars are even much lower compared to actual ones in some cases, showing SVR could not efficiently model the data distribution lying behind, particularly for values with greater magnitudes.

Accuracy and nature of the model

Model	R^2 Score	performance
Polynomial + Ridge (deg=4)	0.9858	Best fit, very high accuracy, sensitive to overfitting.
XGBoost Regressor	0.9552	Good accuracy, better generalization.
SVR (RBF Kernel)	0.8161	Lower accuracy; not optimal for dataset.

Fig 12 Nature of the data of soalr energy generation

Tabular & Time Series-Like: Each row = a state's yearly energy data (2017–2021).
Small Time Horizon: Only 4 time points used to predict the 5th (2021–22).
Heterogeneous Scale: Large variation in values across states (some near 0, others >10,000).
Skewed Target: Big states dominate; small ones risk poor model fit.

B. Prediction of Lifespan of Single-Phase Induction Motor

1. Residuals vs Predicted Values Plot Using Random Forest Regression

The **Residuals vs Predicted (Random Forest)** plot shows the predicted motor life on the X-axis and the residuals (actual minus predicted values) on the Y-axis. The residuals are scattered between -300 and +300, indicating inconsistent accuracy. This suggests the model captures general trends but lacks precision in individual predictions, highlighting the need for further tuning or more advanced models to improve performance

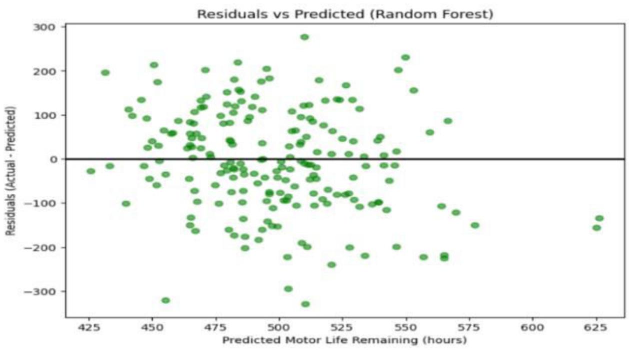

Fig 13 The plot of residual vs predicted values using random forest

2.Residual Analysis of LightGBM Model: Predicted vs Actual

Predicted vs Actual. This plot shows the difference between actual and predicted motor life values using the Light GBM model. Most data points are close to the zero line, indicating accurate predictions for the majority of cases. A few outliers appear at higher predicted values, where the model's error increases slightly. Despite this, the overall spread of residuals is small, suggesting the model is generally stable and reliable across a wide range of motor conditions.

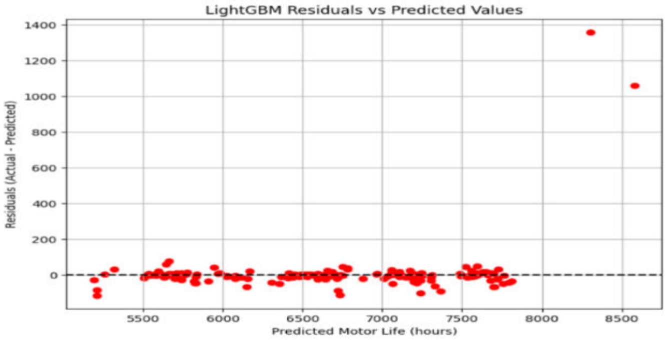

Fig 14 The plot of predicted vs actual values using lightGBM

3. **Residual Analysis of XG Boost Model: Predicted vs Actual Range of predicted values.**

Predicted vs Actual Range of predicted values. The graph titled "Residuals vs Predicted (XGBoost)" illustrates the distribution of residuals against the predicted motor life remaining using the XGBoost model**. The x-axis denotes the predicted motor life remaining in hours, while the y-axis represents the residuals (actual value minus predicted value).** Most of the data points are tightly clustered around the zero line, suggesting that the model predictions are quite accurate for a majority of the instances. However, a few outliers can be observed with higher residuals, indicating some extreme deviations where the model made significant prediction errors.

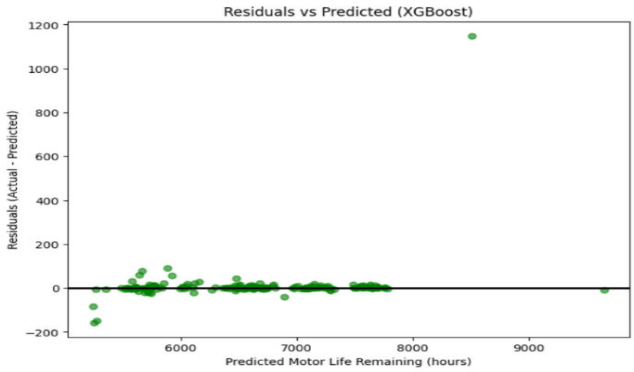

Fig 15 The plot of residual vs predicted values using XGboost

4. **Predicted Performance Metrics for Motor Life Span Models**

Algorithm	R^2 Score	Performance
Random Forest	0.05	Poor performance
XG Boost	1	Overfitting.
Light GBM	0.97	Great performance.

Table 16 Model Evaluation for Motor Life Prediction

IV. CONCLUSION

The predictive analysis performed in both projects demonstrates the transformative role of machine learning in energy systems and equipment maintenance. For energy generation, ensemble regression models provided accurate year-wise forecasts, essential for planning and resource management. For motor lifespan prediction, the models enabled data-driven maintenance strategies that reduce downtime and improve safety. Collectively, these initiatives illustrate the potential of predictive analytics in informing smart decision-making and process effectiveness.

REFERENCES

[1] S.V Satyayana and P. Madhavi, "Big Data Analytics for Electrical Systems using Machine Learning Algorithms," 2023 IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), Mysuru, India, 2023,pp. 1-4, Doi: 10.1109/TEMSMET56707.2023.10149919.

[2] P. Madhavi and S. V. Satyanarayana, "Data Analysis of Electrical Systems Using Machine Learning Algorithms," 2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR), Hyderabad, India, 2022, pp. 1-4, Doi: 10.1109/ICAITPR51569.2022.9844178.

[3] Xinwei Zhang, Mingming Zhang, Quanyi Hu. "Research on the Application of Artificial Intelligence in Operation and Maintenance for Power Equipment," IOP Conf. Ser.: Earth Environ. Sci. 617 (2020) 012001. DOI: 10.1088/1755-1315/617/1/012001

[4] Buddhahai, B., Wongseree, W., & Rakkwamsuk, P. "An Energy Prediction Approach for a Nonintrusive Load Monitoring in Home Appliances," IEEE Transactions on Consumer Electronics, DOI: 10.1109/TCE.2019.2956638

[5] Prince Waqas Khan, Yung-Cheol Byun, "Genetic Algorithm Based Optimized Feature Engineering and Hybrid Machine Learning for Effective Energy Consumption Prediction," IEEE Access, vol. 8, 2020, DOI: 10.1109/ACCESS.2020.3034101

[6] Selahattin Kosunalp, "A New Energy Prediction Algorithm for Energy-Harvesting Wireless Sensor Networks With Q-Learning," IEEE Access, vol. 4, 2016, DOI: 10.1109/ACCESS.2016.2606541

[7] S. Fallah, M. Ganjkhani, S. Shamshirband, and K.-W. Chau, ''Computational intelligence on short-term load forecasting: A methodological overview,'' Energies, vol. 12, no. 3, p. 393, Jan. 2019.

[8] A. Ghasempour and J. Lou, ''Advanced metering infrastructure in smartgrid: Requirements, challenges, architectures, technologies, and optimizations,'' in Smart Grids: Emerging Technologies, Challenges and Future Directions. Hauppauge, NY, USA: Nova Science, 2017.

[9] M. R. Haq and Z. Ni, ''A new hybrid model for short-term electricity load forecasting,'' IEEE Access, vol. 7, pp. 125413–125423, 2019.

[10] A. A. Mamun, M. Sohel, N. Mohammad, M. S. Haque Sunny, D. R. Dipta,and E. Hossain, ''A comprehensive review of the load forecasting techniques using single and hybrid predictive models,'' IEEE Access, vol. 8,pp. 134911–134939, 2020.

[11] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[12] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.