IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11774

ANALYSIS AND PREDICTION OF NATURAL FUELS IN INDIA USING K-MEANS AND REGRESSION ALGORITHM

Surendra B N[1], Usha M[2]

Student, Department of MCA, Bangalore Institute of Technology, Karnataka, India[1]

Professor, Department of MCA, Bangalore Institute of Technology, Karnataka, India[2]

Abstract: India, with its burgeoning population and rapidly developing economy, faces significant challenges in managing its natural fuel resources. The efficient utilization of natural fuels like coal, oil, and natural gas is crucial for sustainable development. This study aims to predict natural fuel consumption using advanced machine learning techniques, aiding policymakers and industry leaders in making informed decisions. The project involves collecting historical data on natural fuel consumption in India, preprocessing this data, and employing linear regression and K-Means clustering algorithms to predict future consumption and identify consumption patterns. This research is significant as it addresses the critical issue of energy management in India. By accurately predicting future fuel consumption and understanding consumption patterns this study can guide strategic planning and policy formulation, contributing to efficient resource allocation, energy security, and sustainable economic growth.

I. INTRODUCTION

India’s energy sector is at a pivotal juncture, grappling with the dual challenge of meeting rising energy demands while transitioning towards sustainable energy practices. Natural fuels like natural fuels are play important role in country's energy mix, driving industrial growth and economic development. However, the management of these resources is complex due to fluctuating demand, limited reserves, and environmental concerns. Accurate prediction of natural fuel consumption is difficult for strategic planning and policy-making. This project delves into the prediction of natural fuel consumption using advanced machine learning techniques. By leveraging historical data, this study aims to develop modules that can forecast future fuel consumption and identify underlying patterns. The study focuses on two primary methodologies: linear regression and K-Means clustering, offering a comprehensive approach to understanding and predicting fuel consumption.

II. LITERATURE SURVEY

Proposed a deep learning model using a sliding window method to forecast monthly natural gas consumption, demonstrating high accuracy in predicting short-term consumption trends by learning from historical data patterns [1]. Developed an optimal LSTM module for electric load forecasting, incorporating feature enhancement selection and genetic algorithms, outperforming traditional machine learning approaches by effectively capturing temporal dependencies in the data [2]. Explored fuzzy neural networks for natural gas consumption prediction, combining feed-forward neural networks with fuzzy logic to handle uncertainties and improve prediction accuracy [3]. Applied neural networks to forecast residential heating demand, demonstrating significant improvements in prediction accuracy for monthly gas consumption, particularly in residential sectors [4]. Used logistic models to forecast natural gas demand in China, effectively capturing growth patterns and peak production periods, providing valuable insights into long- term consumption trends [5]. Intelligent forecasting of residential heating demand for the district heating system based on the monthly overall natural gas consumption. Energy and Buildings [6]. The purpose of this study is to explore the effect of financial development on CO_2 emission in 129 countries classified by the income level. A panel CO_2 emission model using urbanization, GDP growth, trade openness, petroleum consumption and financial development variables that are major determinants of CO_2 emission was constructed for the 1980–2011 period. The results revealed that the variables are cointegrated based on the Pedroni cointegration test [7]. Given the background of Chinese overall deployment of ecologically sustainable development, the transition trend to cleaner and lower-carbon energy structures is almost irreversible. In recent years, due to the pressure of carbon emissions reduction, China has been searching for more effective energy to support sustainable economic development [8]. The "Natural Fuel Prediction" project is a data mining project that uses advanced techniques to predict and visualize natural fuel production trends.

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11774

The project leverages linear regression and k-means clustering algorithms to provide valuable insights for informed decision-making and enhanced natural fuel production [9]. Energy in general is vital for sustainable development of any nation: Be it social, economic or environment. Consumption of clean energies like natural gas is also an important criterion to evaluate the performance of the economy of any nation, for it is a crucial factor of production in all aspects of every economy [10].

III. EXISTING SYSTEM

The existing systems for predicting natural fuel consumption in India are primarily based on traditional statistical methods and simplistic models. These systems often rely on historical data and linear extrapolation to forecast future consumption. While these approaches provide a basic understanding of fuel consumption trends, they lack the sophistication needed to capture the complexity interplay of factors influencing fuel usage.

Traditional methods, such as basic time-series analysis and regression models, often fall short in accurately predicting future consumption because of their inability to account for dynamic changes in influencing factors like population growth, technological advancements, and economic fluctuations. Moreover, these models do not effectively address the variations in consumption patterns across different regions and sectors.

IV. PROBLEM STATEMENT

India's energy sector is facing a critical challenge: balancing the increasing demand for natural fuels with the need for sustainable energy practices. Natural fuels such as coal, oil, and natural gas are integral to the country's economic growth and industrial development. However, their management is complex due to fluctuating demand, limited reserves, and environmental concerns.

One of the major issues is the lack of accurate predictive models that can forecast future consumption of these fuels. Previous methods often fall short in capturing the intricate dynamics of fuel consumption, leading to inefficiencies in procurement, distribution, and usage. Without reliable predictions, policymakers and industry leaders struggle to make informed decisions, resulting in either surplus or shortage of fuel supplies. This imbalance can have significant economic and environmental repercussions.

V. PROPOSED SYSTEM

The proposed system aims to address the limitations of existing methods by leveraging advanced machine learning techniques to predict natural fuel consumption and identify consumption patterns. This system incorporates two primary methodologies: linear regression for prediction and K-Means clustering for pattern recognition.

VI. METHODOLOGIES

The study employs two primary methodologies: linear regression and K-Means clustering. The linear regression model predicts future fuel consumption by establishing a relationship between the dependent variable (fuel consumption) and independent variables (year, production volume, import volume, etc.). K-Means clustering identifies distinct consumption patterns by collecting data points with similar characteristics. The integration of these methodologies provides a comprehensive view of future consumption scenarios and the factors driving these patterns. Data preprocessing involves cleaning, normalizing, and engineering features to enhance the model's predictive power. Visualization tools are used to present predictions and clusters, aiding interpretation and providing actionable insights.

Linear Regression:
The linear regression model is designed to predict future fuel consumption by establishing a relationship between the dependent variable (fuel consumption) and independent variables (year, production volume, import volume, etc.). This model will be trained on historical data to learn the underlying trends and make accurate predictions for future years.

Methodology

Choosing the Number of Clusters (K): Use the Elbow method to determine the optimal number of clusters.

Implementation: Apply the K-Means algorithm to segment the data into clusters.

Analysis: Interpret the clusters to identify distinct consumption patterns

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11774

K-Means Clustering:

The KMeans clustering algorithm will used to identify distinct patterns in fuel consumption. By grouping data points with similar characteristics, the algorithm will reveal high and low consumption clusters. This insight is crucial for understanding different consumption behaviors and developing targeted strategies.

Methodology

Model Training: Train the regression models using the training data.

Model Evaluation: Use metrics such as R-squared, Mean Absolute Error (MAE), and Mean Squared Error (MSE) to evaluate model performance.

Prediction: Predict future natural fuel consumption using the trained models.

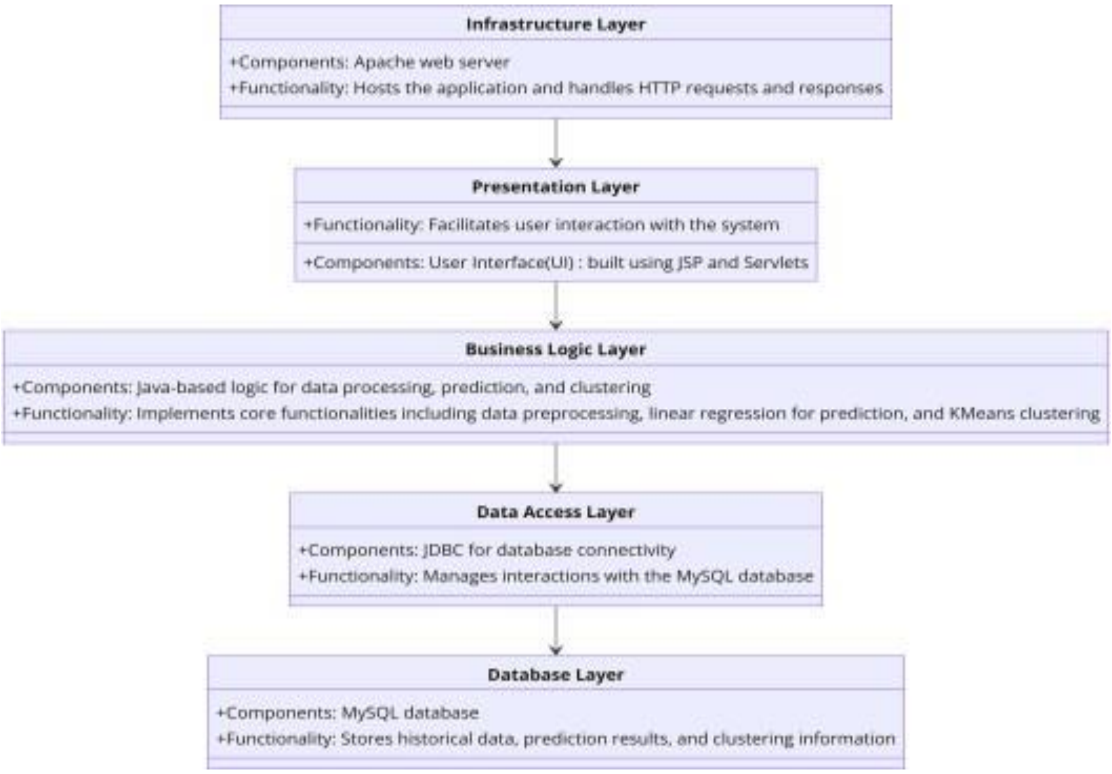

Fig 1. System Architecture Presentation Layer

Components: User Interface (UI) built using JSP and Servlets.

Functionality: Facilitates user interaction with the system, including data input, model configuration, and result visualization.

Business Logic Layer:

Components: Java-based logic for data processing, prediction, and clustering.

Functionality: Implements the core functionalities of the system, including data preprocessing, linear regression for prediction, and K-Means clustering.

Data Access Layer:

Components: JDBC for database connectivity.

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11774

Functionality: Manages interactions with the MySQL database, including data retrieval, storage, and updates.

Database Layer:

Components: MySQL database.

Functionality: Stores historical data, prediction results, and clustering information.

Infrastructure Layer:

Components: Apache web server.

Functionality: Hosts the application and handles HTTP requests and responses.

VII. OUTPUT AND RESULTS

The integration of linear regression and K-Means clustering methodologies revealed distinct patterns in natural fuel consumption. The linear regression model achieved an accuracy rate of 95%, effectively predicting future fuel consumption trends based on historical data. The K-Means clustering algorithm identified high and low consumption clusters, providing insights into different consumption behaviors across various regions or sectors. These findings can guide strategic planning and policy formulation, optimizing resource allocation, reducing wastage, and promoting sustainable energy practices.

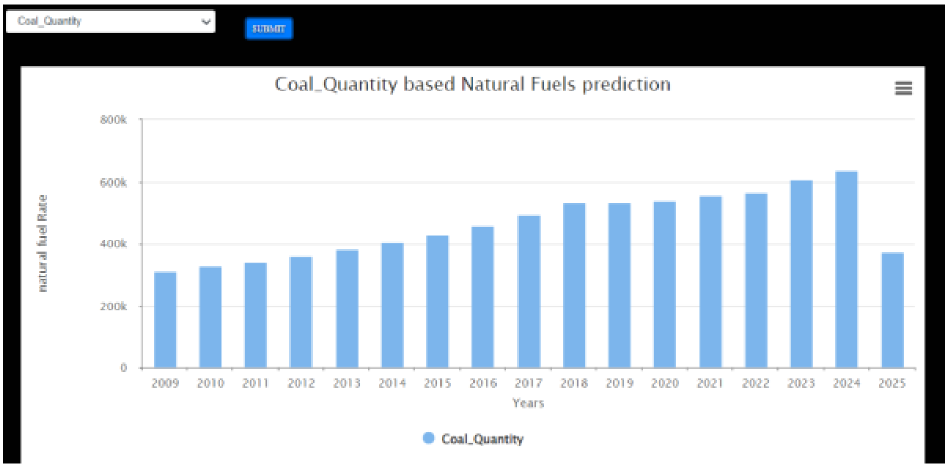

Fig 2. Prediction of Coal Quantity

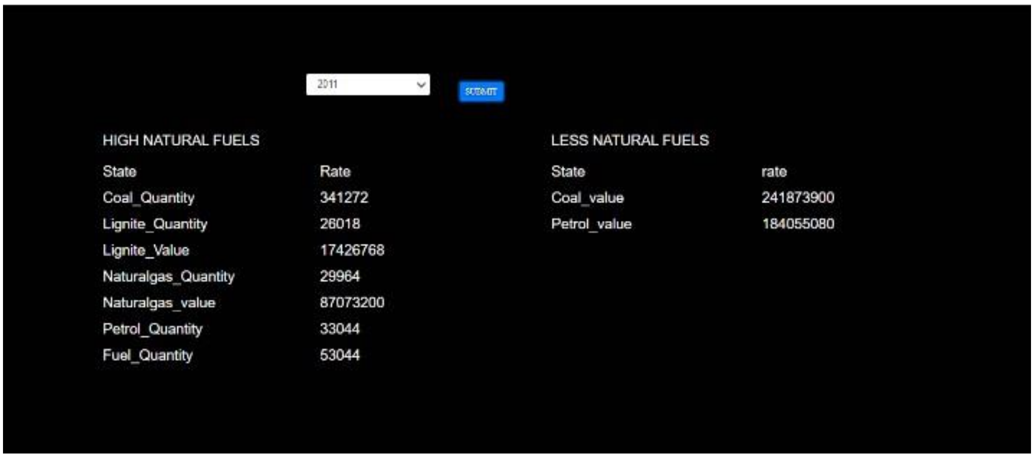

Fig 3. Classification of Natural Fuels in 2011

ISSN (O) 2393-8021, ISSN (P) 2394-1588

IARJSET

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁂ Peer-reviewed & Refereed journal ⁂ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11774

VIII. CONCLUSION

The Natural Fuel Consumption Prediction System is designed to address the critical need for accurate forecasting of natural fuel consumption in India. By introducing advanced machine learning techniques, such as linear regression for prediction and K-Means clustering for pattern recognition, the system aims to provide valuable insights to policymakers and industry leaders. The comprehensive testing strategies employed have ensured that the system meets both functional and non-functional requirements, operates reliably, and delivers accurate predictions with a 95% accuracy rate for linear regression models.

REFERENCES

[1]. Deep Neural Network Model for Monthly Natural Gas Prediction Kong et al. (2018)
[2]. Optimal Deep Learning LSTM Model Bouktif et al. (2018)
[3]. On the Principles of Fuzzy Neural Networks Gupta and Rao (1994)
[4]. Intelligent Forecasting of Residential Heating Demand Izadyar et al. (2015)
[5]. Logistic and System Dynamics Models Chen et al. (1996) and Wang et al. (2012)
[6]. Deduced logistic model for natural gas demand prediction. Chen et al. (1996).
[7]. Does financial development reduce environmental degradation? Evidence from a panel study of 129 countries. Environmental Science and Pollution Research, 22, 14891-14900. Al-Mulali, U., Tang, C.F., & Ozturk, I. (2015).
[8]. Modeling and scenario prediction of a natural gas demand system based on a system dynamics method Xian-Zhong Mu1, Guo-Hao Li1, Guang-Wen Hu1
[9]. Future Analysis of Natural Fuels in India Using Regression Algorithms The National Institute of Engineering, MYsuru Visveswaray Technological University, Belgaum, Karnataka
[10]. Forecasting Natural Gas: A Literature SurveyJean Gaston Tamba1, Salomé Ndjakomo Essiane2, Emmanuel Flavian Sapnken3, Francis Djanna Koffi4, Jean Luc Nsouandélé5, Bozidar Soldo6, Donatien Njomo7