A Robust Deep Capsule Network for IoT-Enabled Cloud-Based Predictive Analytics

Mrs.M. Nandhiya[1], Dr.C.R Durga devi[2]

[1]*PhD Research Scholar, Department of Computer Science, NGM College, Pollachi, Tamil Nadu, India.*

[2]*Assistant Professor, Department of Information Technology, NGM College, Pollachi, Tamil Nadu, India.*

***Abstract*—Cloud computing offers scalable resources such as storage, networking, analytics, and software services through the internet. Predictive analytics in cloud platforms is widely used to forecast future events from historical data. Existing prediction models often suffer from lower accuracy and high computational complexity when handling large and diverse datasets. This paper presents an IoT-enabled Robust Autoencoded Deep Capsule Network for efficient predictive analytics in cloud environments. The proposed framework integrates preprocessing, feature extraction, classification, and optimization into a unified architecture. IoT sensors continuously gather environmental data and transfer it to the cloud platform for analysis. Missing values and abnormal records are corrected during preprocessing using polynomial imputation and modified Z-score analysis. A robust autoencoder identifies the most relevant features, reducing dimensionality and improving efficiency. Finally, the capsule network applies similarity-based classification to generate accurate prediction results. Experimental evaluation using the Global Air Pollution Dataset demonstrates that the proposed framework achieves superior prediction accuracy, reduced error rate, and lower prediction time compared with existing deep learning approaches.**

***Index Terms*—IoT, Cloud Computing, Predictive Analytics, Deep Capsule Network, Autoencoder, Feature Selection, Air Pollution Prediction, Machine Learning**

I. INTRODUCTION

Cloud computing has become an important platform for delivering scalable computing services over the internet. Organizations rely on cloud infrastructure to process and store large amounts of data without maintaining physical resources. The integration of Internet of Things devices with cloud computing has enabled real-time monitoring and intelligent decision making in healthcare, agriculture, environmental monitoring, and smart cities.

Recent studies have applied machine learning and deep learning methods for predictive analytics. Models such as LSTM, BiLSTM, CNN, Transformer, and hybrid neural networks have improved forecasting performance. However, many of these methods require extensive training time and fail to manage high-dimensional datasets effectively. Some techniques also struggle to capture spatial and temporal relationships among features.

To address these limitations, this work proposes a Polynomial Regressive Robust Autoencoded Deep Capsule Network (PRRA-CapsNet). The proposed model improves prediction accuracy while minimizing computational complexity. The architecture combines preprocessing, feature selection, classification, and optimization techniques into a single framework for cloud-based predictive analytics.

II. RELATED WORKS

Several researchers have developed advanced approaches for air pollution prediction and predictive analytics. Transformer-GRU models were introduced to improve forecasting performance, but they faced scalability issues with large datasets. Spatial-temporal feature fusion networks improved pollutant prediction but required extensive parameter tuning. Deep learning frameworks based on CNN, LSTM, GAN, and attention mechanisms have also been explored for environmental monitoring.

Autoencoder-based methods were proposed to reduce dimensionality and enhance feature extraction. Although these techniques improved learning efficiency, they often increased training complexity. Hybrid forecasting models combined machine learning and statistical methods to improve prediction accuracy, but many failed to reduce prediction time and error rates simultaneously.

The review of earlier studies shows that most existing models either suffer from high computational cost or reduced generalization capability. Therefore, a lightweight and accurate prediction framework is necessary for cloud environments handling large-scale IoT data.

III. PROPOSED METHODOLOGY

The PRRA-CapsNet framework is designed to perform efficient predictive analytics in cloud computing environments. The architecture consists of four major stages: data acquisition, preprocessing, feature selection, and classification.

Initially, IoT sensors collect environmental information from multiple locations. These data samples are transmitted to the cloud server for storage and analysis. The collected dataset is then passed to the preprocessing stage. During preprocessing, missing values are estimated using polynomial predictive mean imputation. Outlier samples are detected and removed through the modified Z-score technique. This process improves the quality of the dataset and reduces noise.

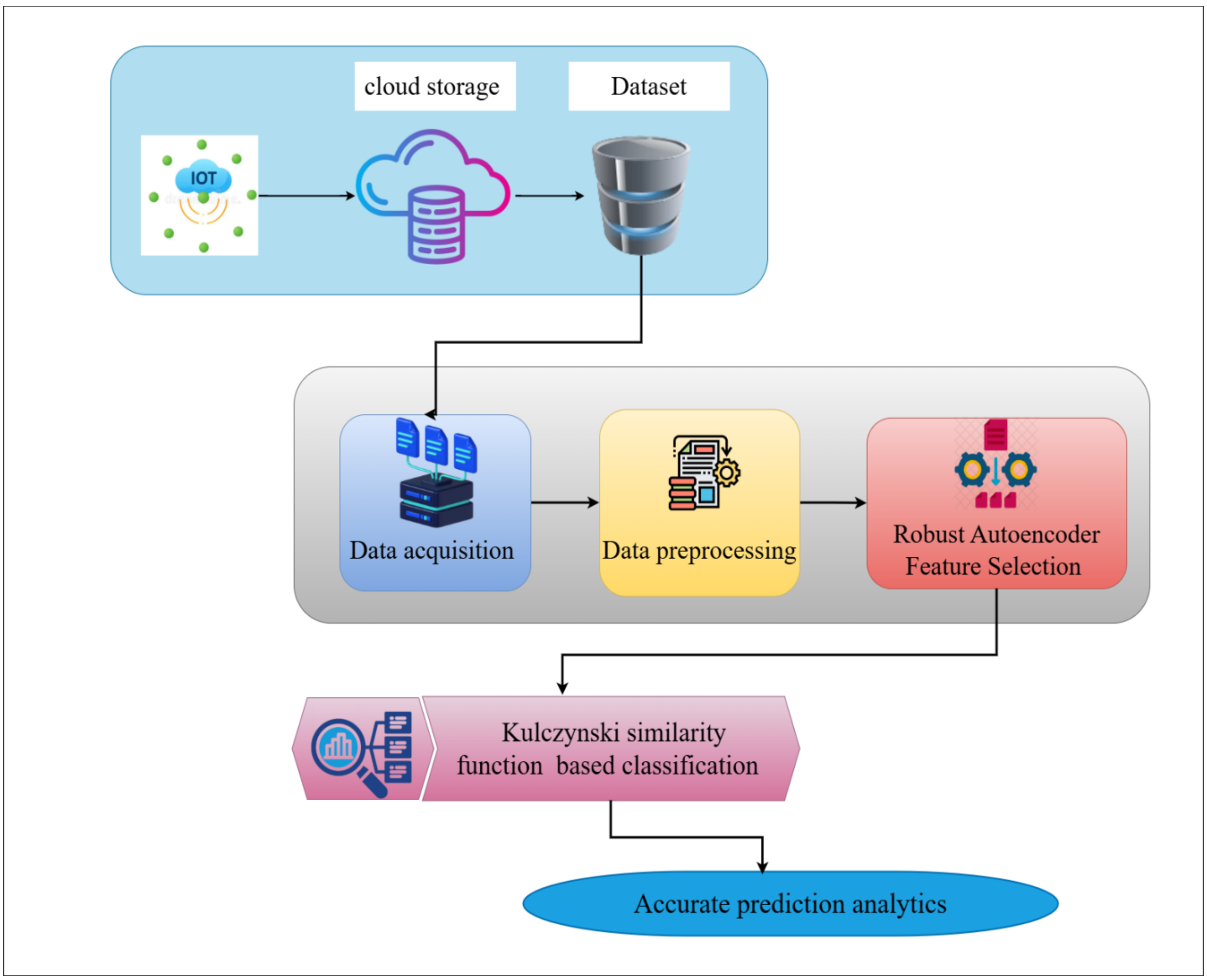

Figure 1. Architecture of the proposed PRRA-CapsNet model

After preprocessing, a robust autoencoder performs feature selection. The encoder transforms the high-dimensional input into a compact latent representation, while the decoder reconstructs the original data. The autoencoder identifies relevant features and removes redundant attributes. This dimensionality reduction process decreases computational overhead and enhances prediction performance.

The selected features are then provided to the class capsule layer. Capsule networks capture spatial relationships between features more effectively than traditional neural networks. The proposed model employs the Kulczynski similarity function to measure similarity between training and testing samples for accurate classification. To minimize classification error, the network parameters are optimized using a wild horse herd optimization algorithm. The optimization process iteratively updates weight values to achieve better prediction accuracy and lower false prediction rates.

IV. EXPERIMENTAL SETUP

The proposed PRRA-CapsNet model was implemented using Python programming language. Experiments were conducted on the Global Air Pollution Dataset obtained from the Kaggle repository. The dataset contains 23,463 instances with 12 attributes related to air quality, including AQI value, ozone level, nitrogen dioxide concentration, carbon monoxide, and particulate matter.

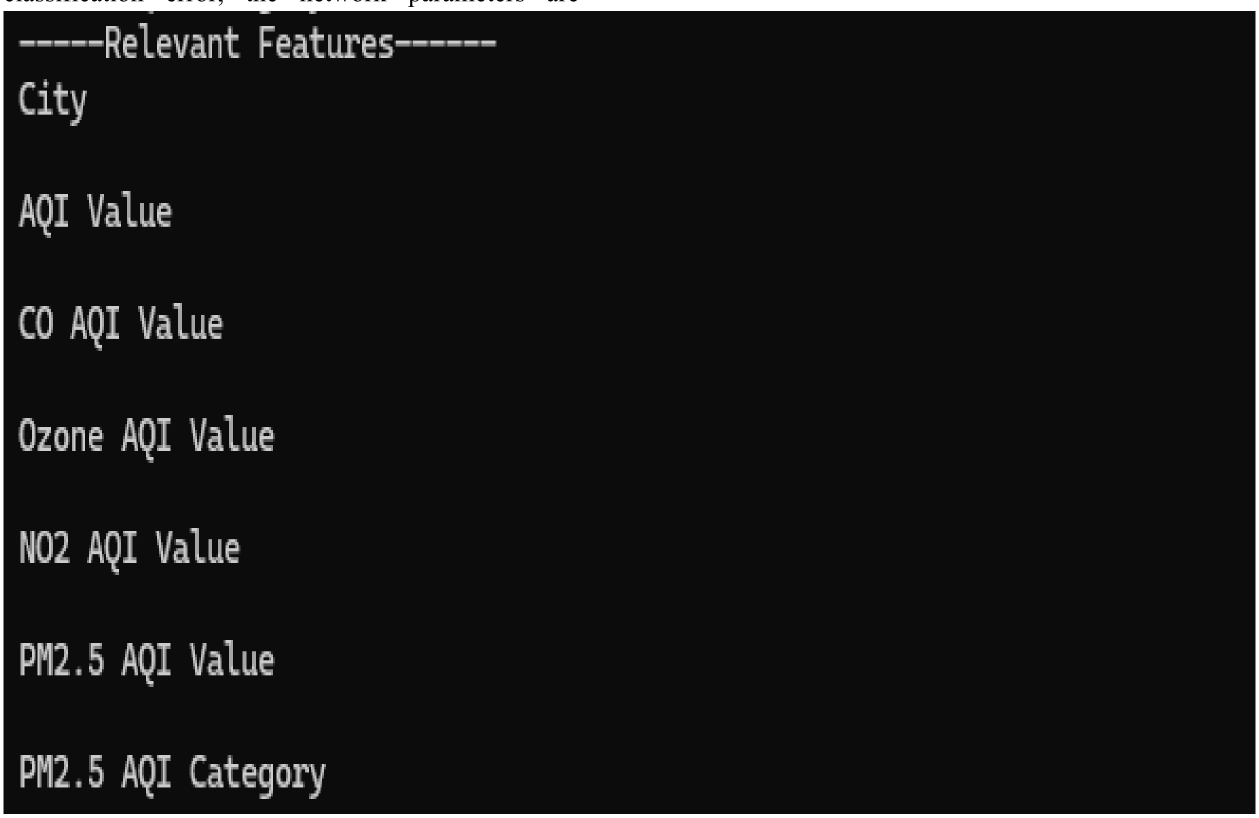

Figure 7 selected features

Figure 7 illustrates the feature selection outcomes which identify seven relevant attributes to perform air pollution prediction.

In the final phase, the model carries out air pollution prediction by applying Kulczynski similarity function implemented into the deep capsule network. Using the seven selected features, the PRRA-CapsNet model is accurately predicting the air pollution based on air quality index is good, moderate and unhealthy, very unhealthy and hazardous. The categories Unhealthy, Very Unhealthy, and Hazardous indicate that the air pollution levels are high and potentially harmful to health.

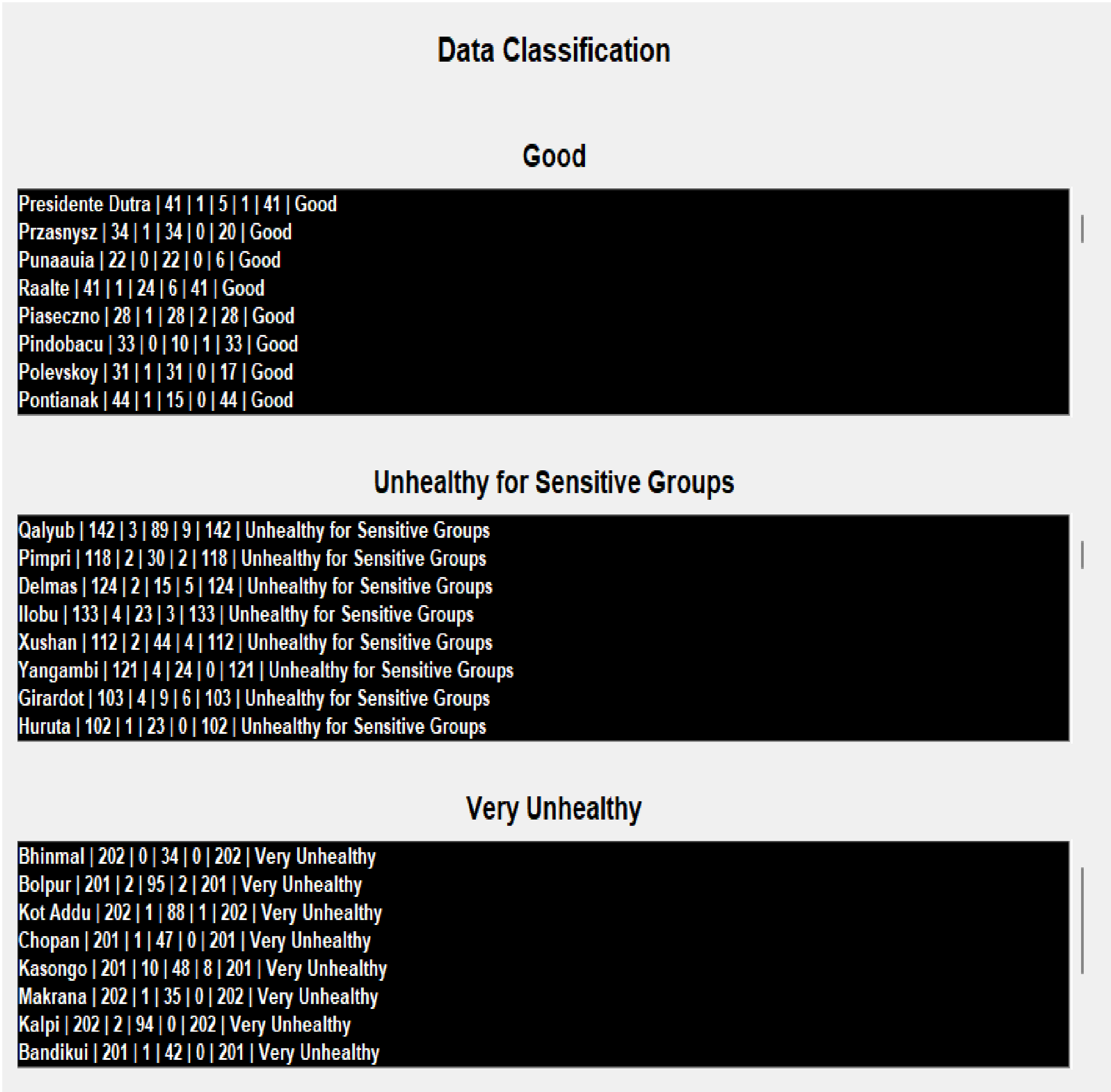

Figure 8: Prediction Outcomes

V. RESULTS AND DISCUSSION

Experimental results indicate that the proposed PRRA-CapsNet framework outperforms existing deep learning methods in predictive analytics tasks. The model achieved higher prediction accuracy because of efficient preprocessing and similarity-based classification.

The robust auto encoder improved feature extraction by removing irrelevant attributes and reducing dimensionality. As a result, computational complexity and prediction time were significantly reduced. The capsule network effectively captured feature relationships, leading to improved classification accuracy.

The optimization strategy based on wild horse herd behavior minimized prediction error and improved convergence. Compared with TRANS_GRU and MSTVFFN models, the proposed approach demonstrated better precision, recall, and F1-score. In addition, the framework achieved lower RMSE and MAE values, confirming its reliability for cloud-based predictive analytics applications.

VI. CONCLUSION

This paper introduced a robust IoT-enabled predictive analytics framework called PRRA-CapsNet for cloud computing environments. The model integrates preprocessing, robust autoencoder feature selection, deep capsule classification, and optimization into a unified architecture. By combining these techniques, the framework improves prediction accuracy while reducing computational complexity and prediction time.

Experimental analysis using the Global Air Pollution Dataset confirmed that the proposed model provides superior performance compared with existing deep learning approaches. Future work can focus on integrating real-time big data analytics and advanced optimization strategies for smart city and environmental monitoring applications.

REFERENCES

[1] Sudhir Kumar et al., "Optimizing Air Pollution Forecasting Models Through Knowledge Distillation," IEEE Access, 2025.
[2] Xinmeng Zhou et al., "A Multiscale Spatial-Temporal Variable Feature Fusion Network," Environment International, 2025.
[3] Gang Chen et al., "Hybrid Deep Learning Air Pollution Prediction," Scientific Reports, 2025.
[4] Sabina-Cristiana Necula et al., "Advanced Hybrid Models for Air Pollution Forecasting," Electronics, 2025.
[5] Cog Cao et al., "Physics-Based Machine Learning for Urban Air Pollution Prediction," Environmental Research Communications, 2025.

Authors Profile

M.Nandhiya received her B.Sc Information Technology from NGM College, Pollachi. She had her Master of Computer Applications, from SVS Insitute of Technology under Anna University, Chennai. She holds her M.Phil degree in Computer Science at NGM College, Pollachi. She had 2 years 7 months experience in the field of teaching. She is presently working as a Assistant professor in NGM College, Pollachi. Her research interest includes Data Mining, Big Data Analytics. Now she is pursing her PhD degree at NGM College, Pollachi.

Dr. C.R. Durga devi, Assistant Professor in the Department of Information Technology at Nallamuthu Gounder Mahalingam College, Pollachi, is an accomplished academic with over 18 years of teaching and research experience. She holds a Ph.D. in Computer Science (2022) from Bharathiar University and has qualified both NET (2013) and SET (2012). Her research interests include cloud computing, load balancing, big data analytics, and optimization algorithms, with 4 Scopus indexed publications, 10 international journal papers, and 20 conference presentations. Dedicated to continuous learning, she has completed 25 FDPs, 12 workshops, and industry trainings with Wipro on Java Full Stack and Cloud Computing. A recognized Ph.D. supervisor, she contributes as a doctoral committee member and evaluator while also serving as Criterion Manager for NAAC, member of the Institution's Innovation Council, and organizer of innovation-focused initiatives. She has been recognized as a Wipro Certified Faculty (2022) and Innovation Ambassador (MIC, 2024–25), and actively engages in academic responsibilities such as external examiner, question paper setter, curriculum designer, and editor. Frequently invited as a resource person, she exemplifies academic excellence, research leadership, and a strong commitment to advancing higher education in Information Technology.