**DATA ANALYTICS META-MODEL FOR IOT USING**

**HYBRID MACHINE LEARNING TECHNIQUES**

|  |  |  |
| --- | --- | --- |
| **Afreen** | **C. Sony** | **C. Kishore** |
| 21P61A0503 | 21P65A0540 | 21P61A0544 |
| Subhanafreen1011@gmail.com | Sonysonu6342@gmail.com | chinthakayalaKishore21@gmail.com |

**OBJECTIVE**

Develop a hybrid machine learning framework that integrates Federated Learning (FL) and Adaptive Moving Window Regression (AMWR) to enable robust prediction capabilities. The decentralized nature of FL ensures data privacy and security by keeping data local, preventing exposure to central servers. Meanwhile, AMWR enhances the model’s ability to adapt to dynamic IoT data streams by continuously adjusting to recent data, improving both prediction accuracy and adaptability over time. This combination offers a secure and efficient approach to handle real-time, privacy-sensitive predictions in IoT environments.

**ABSTRACT**

The growing volume of IoT data has raised concerns about data privacy and network costs, highlighting the need for innovative solutions in predictive analytics. A proposed hybrid machine learning framework integrates Adaptive Moving Window Regression (AMWR) and Federated Learning (FL) to analyze both historical and real-time dynamic IoT data. This framework seeks to overcome the limitations of current IoT data analysis methods by enhancing data privacy, lowering network costs, and enabling efficient training of Deep Neural Networks on devices. It aims to improve prediction accuracy, refine predictive precision for complex events, and facilitate more effective disease treatments without the need for invasive procedures.

**INTRODUCTION**

Internet of Things (IoT) devices generate vast amounts of data, which is essential for predictive analytics. However, traditional cloud-based data analysis encounters challenges related to privacy and network costs. Existing models include Cloud-Centric IoT Data Analysis, which relies on centralized cloud servers for processing and analysis but raises concerns about data privacy and high network costs. Edge Computing with Limited AI involves processing data locally on the IoT device or a nearby edge server, reducing reliance on the cloud. Centralized Predictive Models, commonly used in healthcare, train on large datasets at a single location but may struggle to adapt to the dynamic nature of IoT data streams.

**HARDWARE & SOFTWARE REQUIREMENTS**

**RAM:** 8GB  
**Processor:** Intel Pentium 4  
**Operating System:** Windows 10  
**Coding Language:** Python  
**IDE:** Google Colab

**EXISTING SYSTEM**

Cloud-Based Predictive Models involve sending IoT data to cloud servers for centralized processing, such as with Microsoft Azure IoT Hub, which connects, monitors, and manages IoT devices. These models offer high computational power and access to advanced machine learning algorithms, but they face limitations in terms of data privacy concerns and high network and storage costs.

Federated Learning (FL) Models take a decentralized approach, training models locally on IoT devices. A real-world example of this is Google’s Gboard, where FL is used for predictive text without transmitting user data to central servers. While FL enhances privacy and reduces data sharing, it faces challenges related to computational and communication limitations on resource-constrained devices.

Edge Computing Models process data locally on or near the IoT device rather than relying on central cloud servers. An example of this is a smart thermostat, which uses edge computing to adjust the temperature based on user movements and preferences in real-time, without sending data to the cloud. Edge computing reduces latency and supports real-time decision-making, but it is limited by the computational resources available on the device.

**PROPOSED SYSTEM**

The proposed system integrates Federated Learning (FL), Adaptive Moving Window Regression (AMWR), and Complex Event Processing (CEP) to enhance predictive analytics. FL trains machine learning models directly on local IoT devices, ensuring data privacy and reducing network costs. It employs optimization strategies such as structured pruning to reduce model size, weight quantization to compress model weights while maintaining accuracy, and selective updating to refresh only the active parts of the model. AMWR improves prediction accuracy by combining historical data with real-time updates, making it ideal for applications that require continuous adaptation to changing data patterns. CEP monitors real-time data streams to detect complex events and patterns, integrating with AMWR and FL to provide immediate event detection. While CEP handles real-time event identification, AMWR refines predictions with both historical and real-time data, and FL ensures secure and efficient model training, maintaining data privacy throughout.

**CONCLUSION**

In conclusion, the proposed hybrid framework combining Federated Learning (FL), Adaptive Moving Window Regression (AMWR), and Complex Event Processing (CEP) offers a robust solution for predictive analytics in IoT environments. FL ensures data privacy and reduces network costs by training models locally, while AMWR enhances prediction accuracy through continuous adaptation. CEP enables real-time event detection, facilitating immediate response to dynamic data streams. This system efficiently integrates historical and real-time data for precise predictions while maintaining security and privacy. Overall, it addresses the challenges of IoT data analysis by optimizing accuracy, privacy, and resource usage.

**REFERENCES**

Here are some potential references that can be cited for the concepts mentioned in the proposed system:

1. **McMahan, H. B., Moore, E., Ramage, D., & Hampson, S.** (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 1273-1282.
2. **Chen, T., & Guestrin, C.** (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
3. **Li, T., Sahu, A. K., Zaheer, M., & Talwalkar, A.** (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50-60.
4. **Srinivasan, S., & Xie, L.** (2020). Adaptive Moving Window Regression for real-time IoT data prediction. *Journal of Sensors and Actuator Networks*, 9(2), 24.
5. **Luckham, D. C.** (2002). The power of events: An introduction to complex event processing in distributed enterprise systems. *Addison-Wesley Longman Publishing Co.*
6. **Yin, X., Li, K., & Lin, X.** (2021). Edge computing for IoT data processing: A review. *IEEE Access*, 9, 145015-145032.

**Guide Name:** G. Arun Sir

**Guide Designation:** Associate Professor, Dept. of CSE

**Guide Signature:**

**Project Coordinator Signature:**

**HOD Signature:**