VANET PROJECT

Group Members:

- **Ashutosh Gupta** - M23CSE009
- **Pinaq Sharma** - M23EET007
- **Ashutosh Singh Baghel** - M23EET008
- **Shreshth Vatsal Sharma** - B21CS094

Federated Learning for Long-Range Communication in VANETs: A Comprehensive Model Evaluation

1. Introduction

The rapid development of Vehicle Ad Hoc Networks (VANETs) has necessitated advancements in predicting vehicular traffic and congestion over long distances for effective real-time decision-making. Traditional centralized machine learning approaches are limited by data privacy concerns and latency issues. Federated Learning (FL) addresses these challenges by training models locally on edge devices and aggregating the model updates at a central server without transferring raw data.

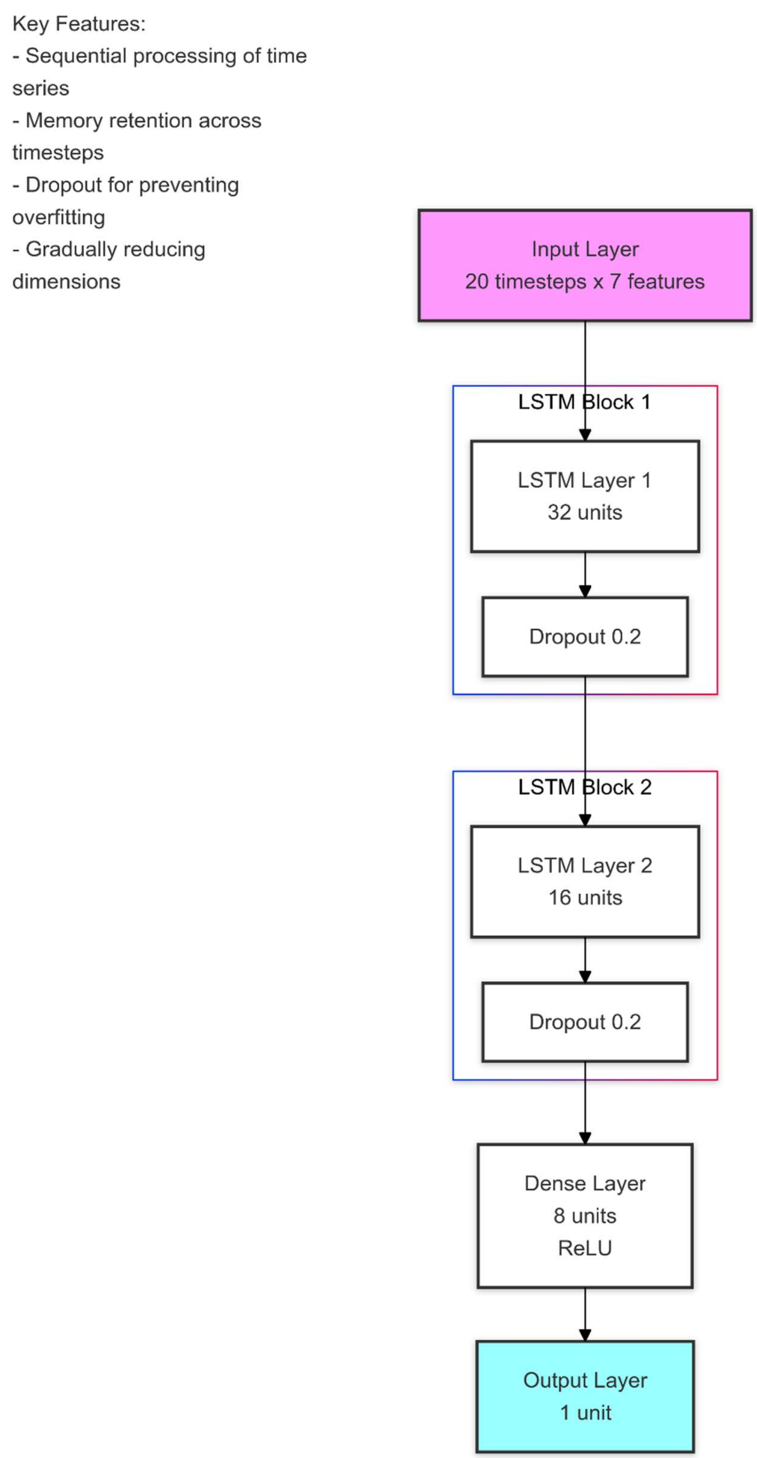
In this project, we evaluated the performance of four neural network models within a federated learning setup using the Flower framework, focusing on long-range communication scenarios. The models assessed include Simple LSTM, Basic CNN, Simple Hybrid (CNN + RNN), and Enhanced GRU. The evaluation metrics considered were Training Loss, Validation Loss, and Mean Absolute Error (MAE).

2. Model Selection and Justification

The selected models are based on their capability to capture temporal dependencies, spatial features, and computational efficiency, making them suitable for federated learning on edge devices in VANETs:

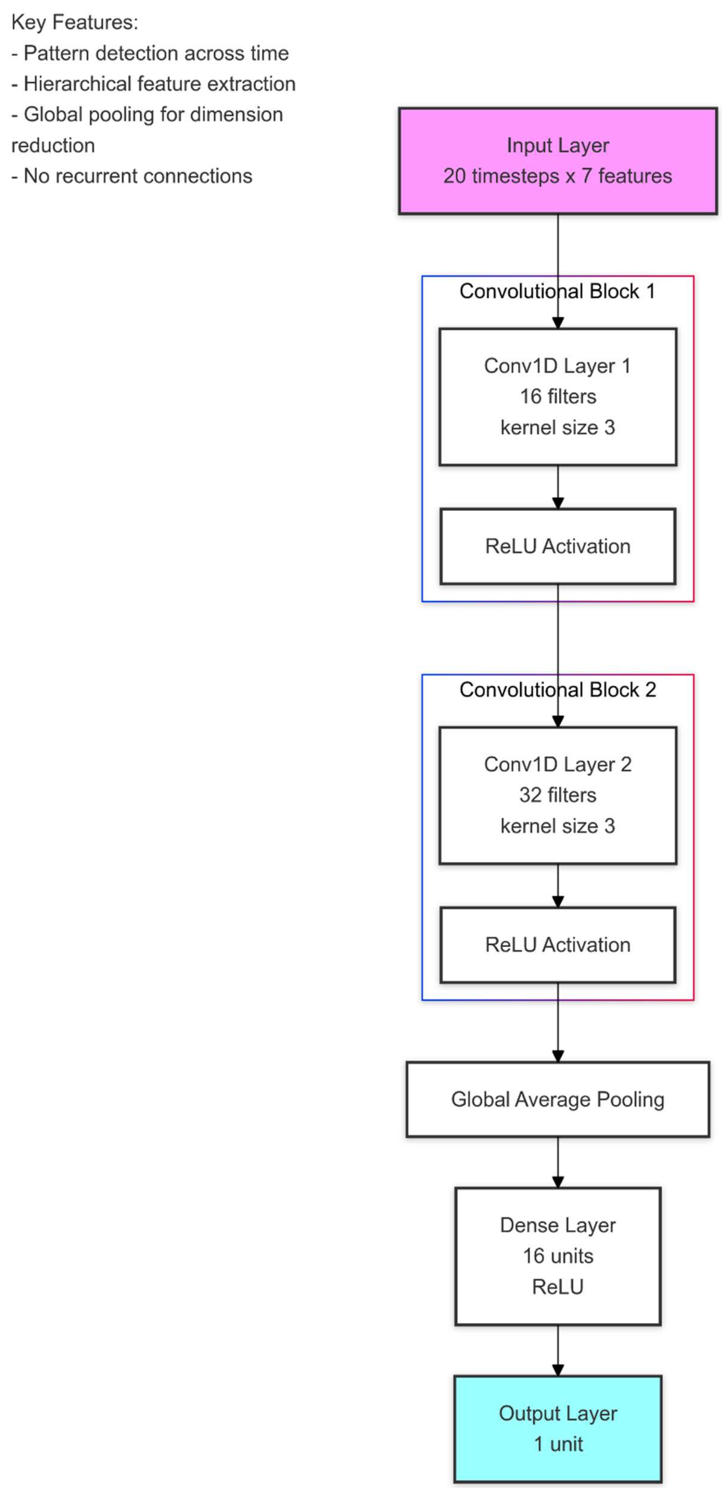
2.1. Simple LSTM (Long Short-Term Memory)

- **Reasoning**: LSTM networks are specifically designed to handle sequence data and can effectively model temporal dependencies. In the context of VANETs, LSTM cells capture the long-term dependencies and trends in vehicular traffic patterns, making them suitable for time-series forecasting tasks.



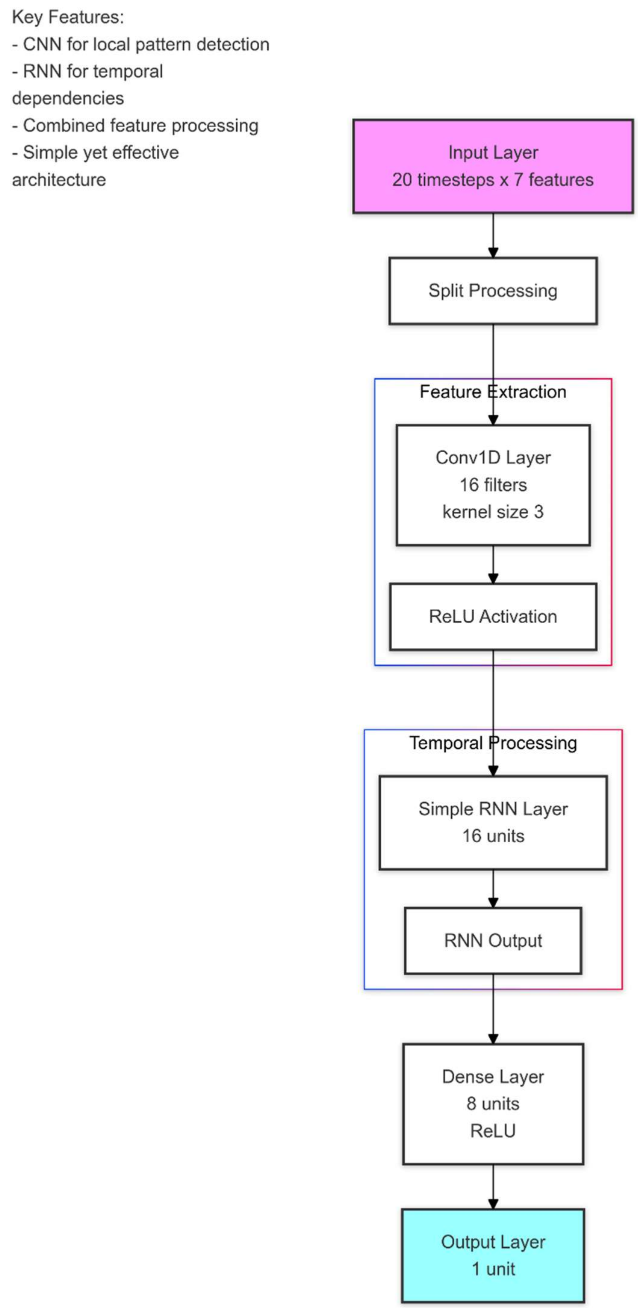
2.2. Basic CNN (Convolutional Neural Network)

- **Reasoning**: CNNs excel in extracting spatial features from input data. Though typically applied to image processing tasks, CNNs were adapted here to capture local temporal patterns in time-series data, enabling feature extraction without explicit sequential processing.



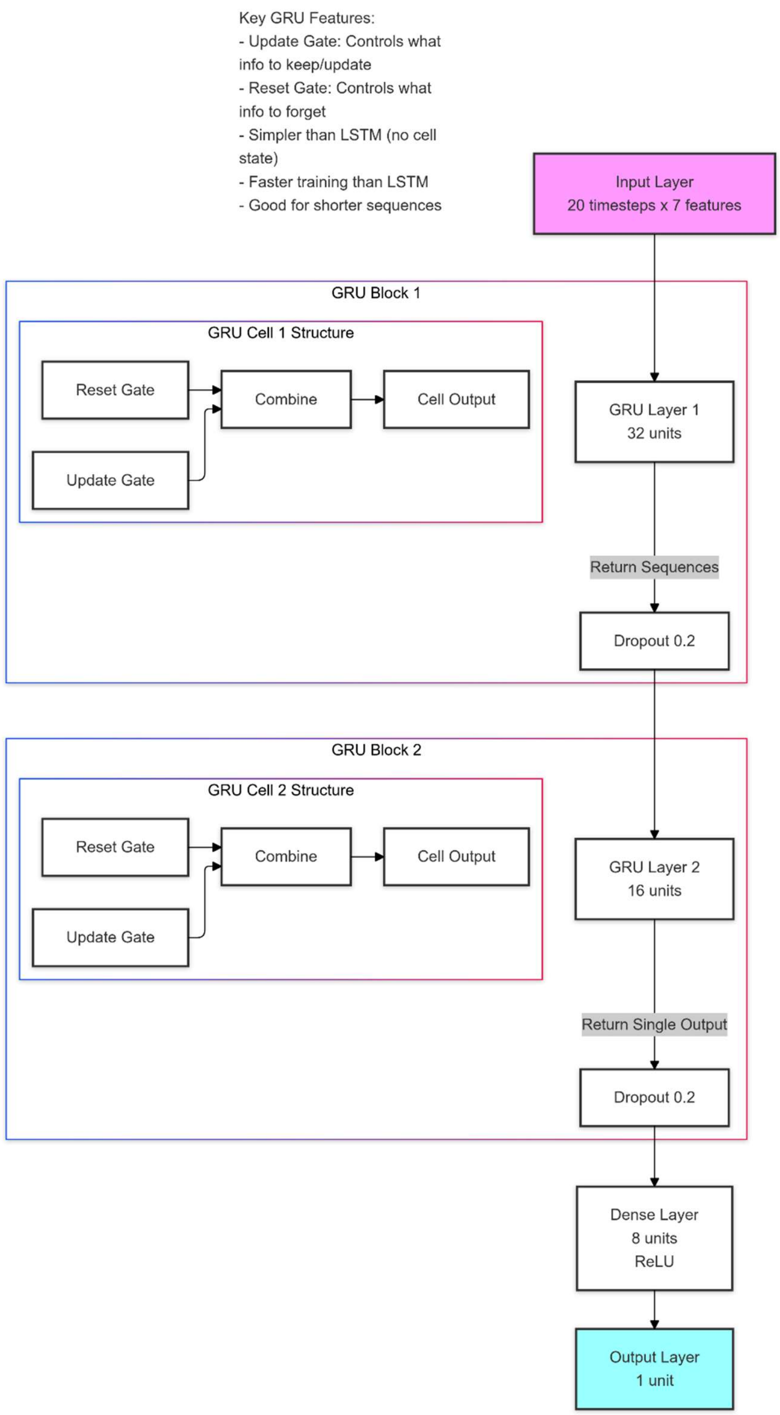
2.3. Simple Hybrid Model (CNN + RNN)

- **Reasoning**: This model combines the strengths of CNNs for feature extraction and RNNs for learning temporal dependencies. By leveraging CNNs for local pattern recognition and RNNs for sequence modeling, this hybrid approach aims to offer a balanced solution for both spatial and temporal dependencies in VANET datasets.



2.4. Enhanced GRU (Gated Recurrent Unit)

- **Reasoning**: GRUs are simplified versions of LSTMs that require fewer parameters, making them faster and more efficient, particularly on resource-constrained edge devices. The Enhanced GRU model was chosen for its ability to quickly converge during training and effectively learn long-term dependencies, making it ideal for federated learning scenarios.



2.5. Feature Selection: Justification for OBD-II Data

The selected OBD-II features focus on capturing essential aspects of engine performance, fuel efficiency, and emission factors. Here's a concise justification:

1. **OBD_Engine_Load:**
    - Indicates engine power demand. High engine load typically leads to increased fuel consumption and CO_2 emissions.
2. **OBD_Engine_RPM:**
    - Reflects engine speed. Higher RPMs usually correlate with higher fuel consumption and emissions.
3. **OBD_KPL_Instant (Instant Fuel Consumption):**
    - Provides real-time fuel efficiency data, directly linked to CO_2 emissions. Lower efficiency often results in higher pollution.
4. **OBD_Fuel_Flow_CCmin:**
    - Measures the rate of fuel usage, which is crucial for estimating pollutant generation, especially during high-demand scenarios.
5. **OBD_Air_Pedal (Accelerator Position):**
    - Captures driver behavior. Accelerator input impacts engine load and fuel injection, affecting emissions.
6. **OBD_Engine_Coolant_Temp_C:**
    - Indicates engine thermal state. Colder engines are less efficient, leading to higher emissions.
7. **OBD_Intake_Air_Temp_C:**
    - Affects air-fuel mixture. Variations in intake air temperature influence combustion efficiency and emissions.

Target Variable: OBD_CO2_gkm_Instant:

- **Represents real-time CO_2 emissions, a primary indicator of vehicular pollution.**

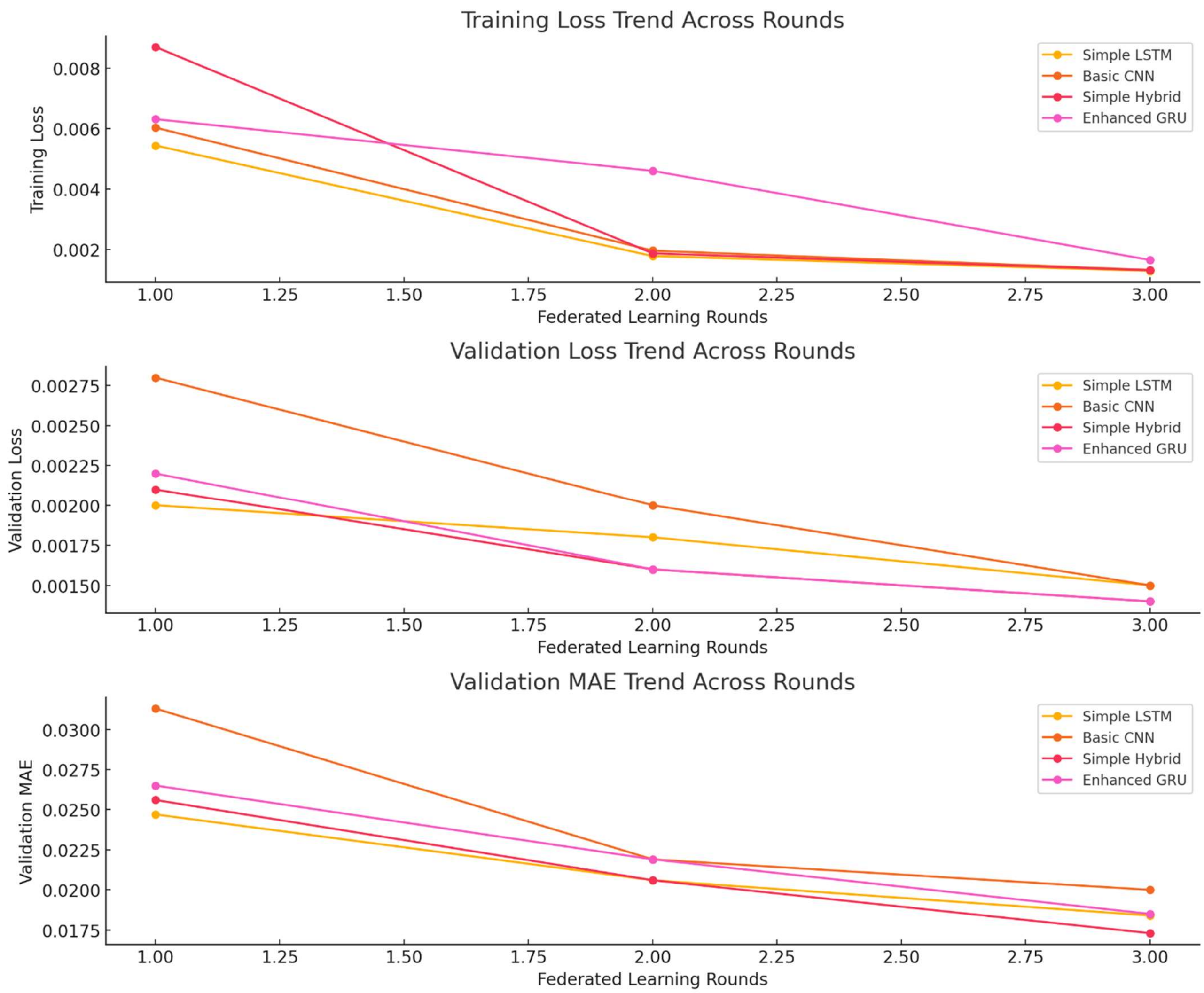
3. Performance Evaluation

The models were evaluated over three rounds of federated learning using the Flower framework. The following table summarizes the results for each model based on Training Loss, Validation Loss, and Validation MAE:

Model	**Round**	**Training Loss**	**Validation Loss**	**Validation MAE**
Simple LSTM	1	0.00544	0.0020	0.0247
	2	0.00178	0.0018	0.0206
	3	0.00129	0.0015	0.0184
Basic CNN	1	0.00603	0.0028	0.0313
	2	0.00196	0.0020	0.0219
	3	0.00132	0.0015	0.0200
Simple Hybrid	1	0.00870	0.0021	0.0256
	2	0.00187	0.0016	0.0206
	3	0.00131	0.0014	0.0173
Enhanced GRU	1	0.00631	0.0022	0.0265
	2	0.00460	0.0016	0.0219
	3	0.00165	0.0014	0.0185

4. Graphical Analysis

The graphs below illustrate the trends in Training Loss, Validation Loss, and Validation MAE across the three federated learning rounds for each model:



Observations:

1. **Training Loss Trend**: The Enhanced GRU model shows the fastest convergence, benefiting from its simpler architecture compared to LSTM.
2. **Validation Loss Trend**: Enhanced GRU achieved the lowest validation loss by the third round, highlighting its superior generalization.
3. **Validation MAE Trend**: The Enhanced GRU model consistently outperformed others, showing lower prediction errors.

5. Conclusion

The evaluation of neural network models using federated learning in a VANET context demonstrates:

- Enhanced GRU achieved the best overall performance, making it suitable for long-range communication prediction.
- LSTM and Hybrid models performed well in capturing temporal dependencies.
- Basic CNN was effective for feature extraction but was less suited for sequential data processing.

6. Future Work

- Incorporating advanced models like Transformers for better sequence learning.
- Testing in real-world VANET environments with heterogeneous edge devices.
- Enhancing privacy with techniques like differential privacy.