SDAIA
الهيئة السعودية للبيانات والذكاء الاصطناعي
Saudi Data & AI Authority

أكاديمية طويق
TUWAIQ ACADEMY

Weekly Mini Project:

T5

Project Members and roles :

- **Sarah Aljuwayr : choosing an existing dataset and project presentation, train the models**
- **Aliah Alotaibi : project report, EDA, Preprocessing**
- **Najla Aldhubaib: particbited in report, models evaluation**

Project Title:	**Traffic Volum Prediction using Time Series**

Description

Project Overview:

The project aims to recognize images of cyclists and determine whether they are wearing a helmet or not from the Traffic Violations Prediction dataset. The project will explore different aspects of deep learning, including model design, hyperparameter tuning, overfitting handling, feature transformation, and visualization techniques.

Dataset: Traffic Volum Prediction from UC Irvine

Download Dataset Link: https://archive.ics.uci.edu/dataset/492/metro+interstate+traffic+volume

Tasks:

1-Data Loading and Preprocessing: Download the Traffic Volum Prediction dataset and load it into google colab environment (Python). Preprocessed the data by 1-missing values: We had
missing values in one column and this column was dropped.
2-duplicates:After checking duplicates we recognized that we have duplicates and where handled using (drop_duplicates). 3-Normalize: we choose to apply Min-Max Scaling: Rescale the data to a fixed range (usually 0 to 1). 4- Applying label encoder on categoricl columns.

2-Design and implement RNN, LSTM and GRU:

Design and implement the RNN, LSTM and GRU models using (Keras)TensorFlow. The breakdown of the Recurrent Neural Network (RNN) model architecture contains : input layer, 3 RNN layers, 3 dropout layers, dense layer. Also, for LSTM model: input layer, 2 LSTM layers, dense layer. Finally, GRU model: input layer , 2 GRU layers, 2 dense layers. Group members chose number of target size,batch size, neurons per layer, and epochs number because it is sutiable for the dataset and after run the code multiple times.

3-Splitting the data :

- S plitting the data into train and test. Applying MinMaxScaler .
- Function (create_dataset) was used to prepare the data. Reshape the data for LSTM input

4-Training the Model and Evaluation:

Train and evaluate the model's performance using: mean squared error.

Conclusion

After applying three models, LSTM has the best performance based on model evaluation.