Project Laboratory : Traffic Prediction

Optimizing Transport Systems Using Sensor Data

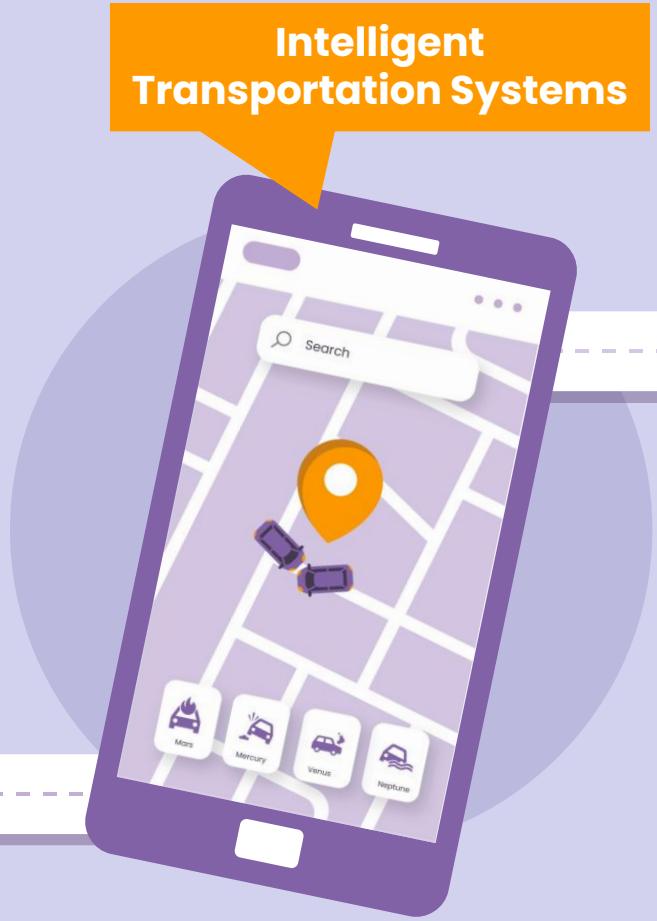
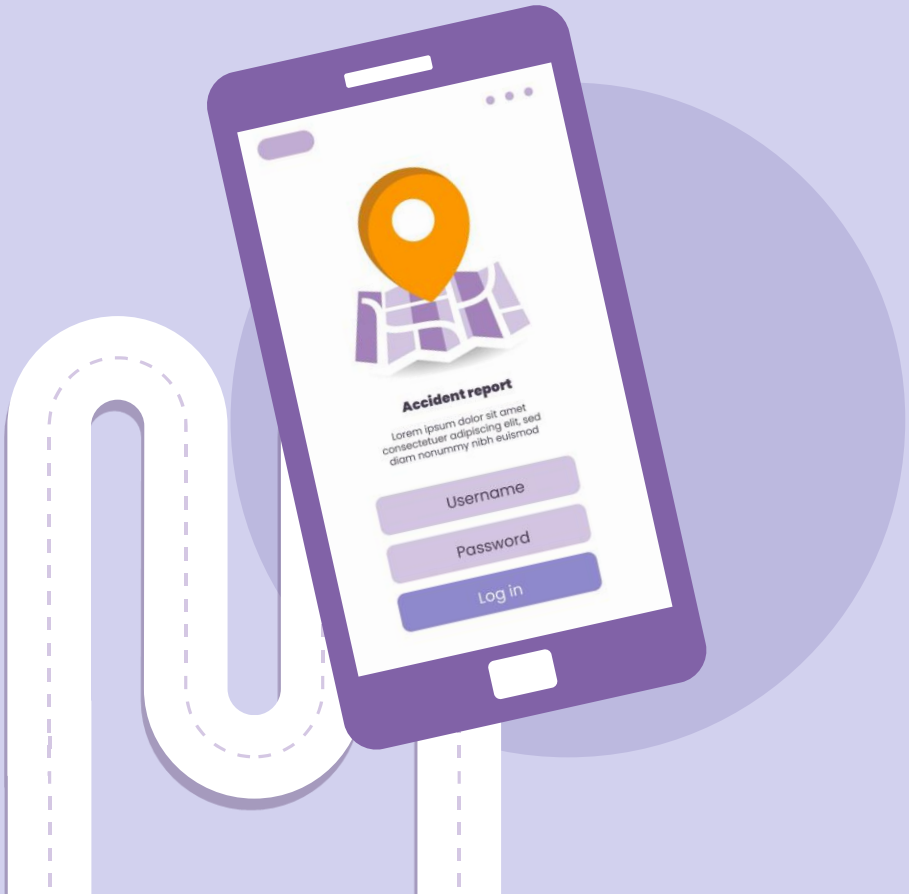


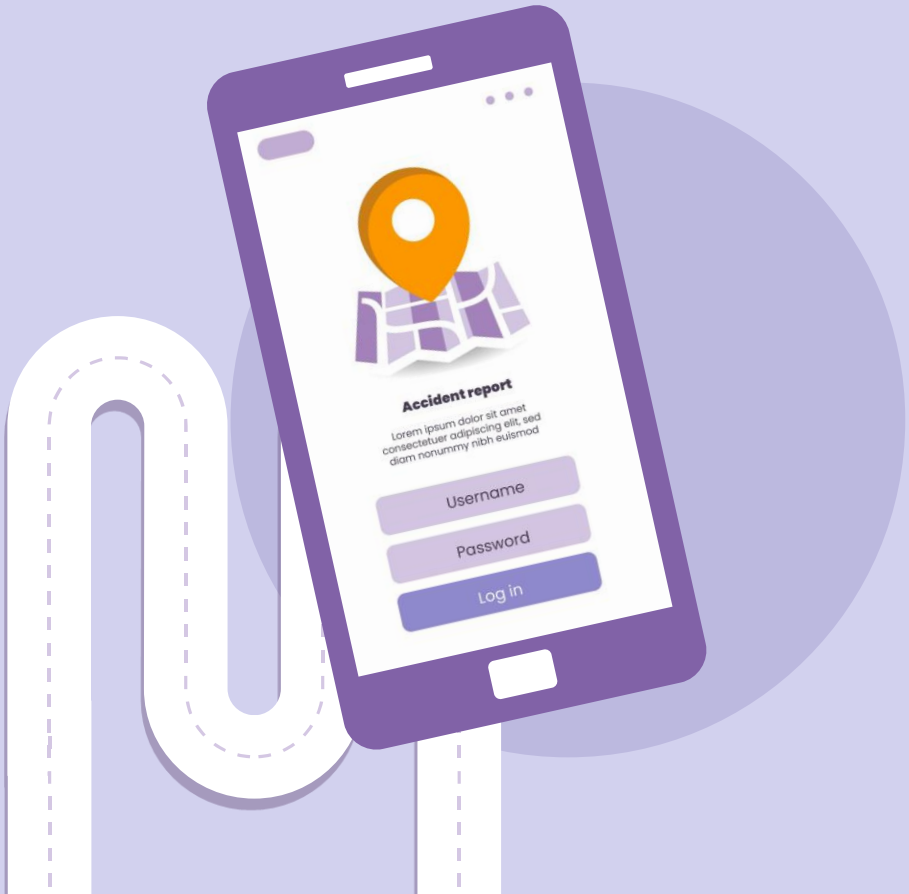
Table of contents



01

Overview

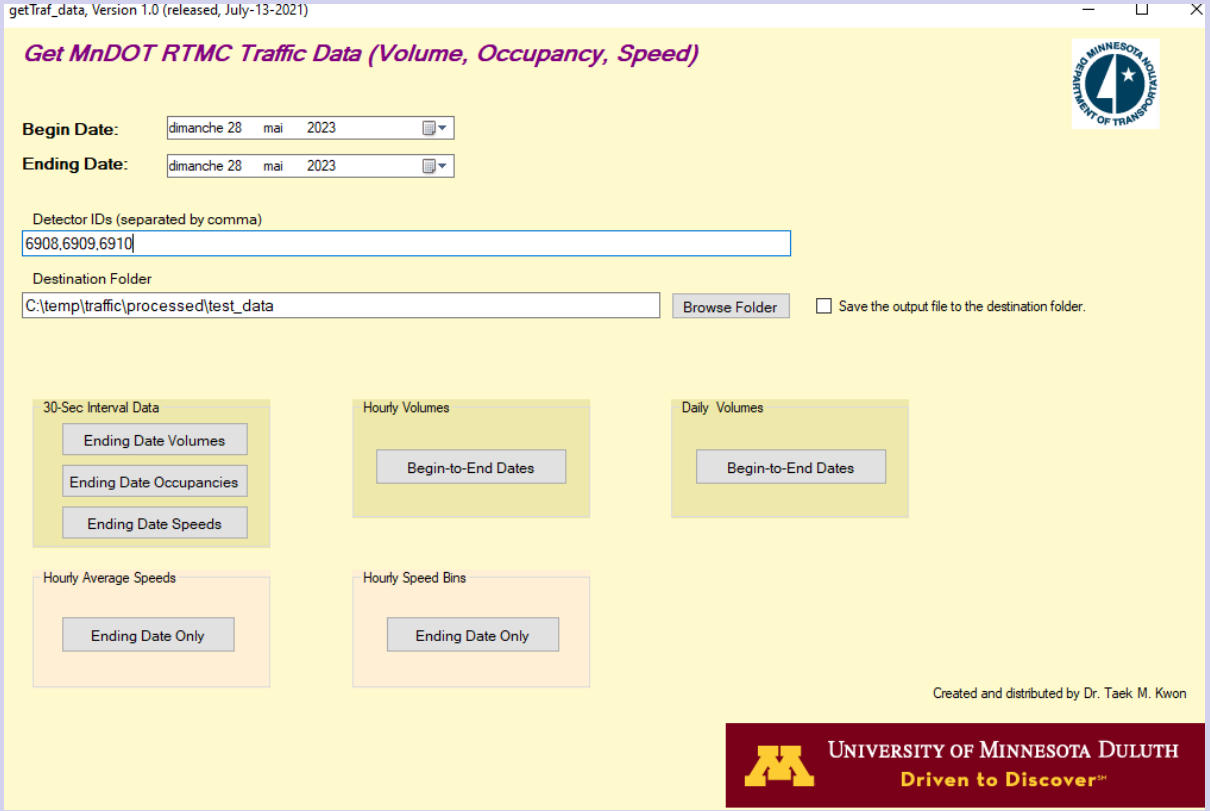
Predicting the state of traffic is a fundamental and challenging task.



02

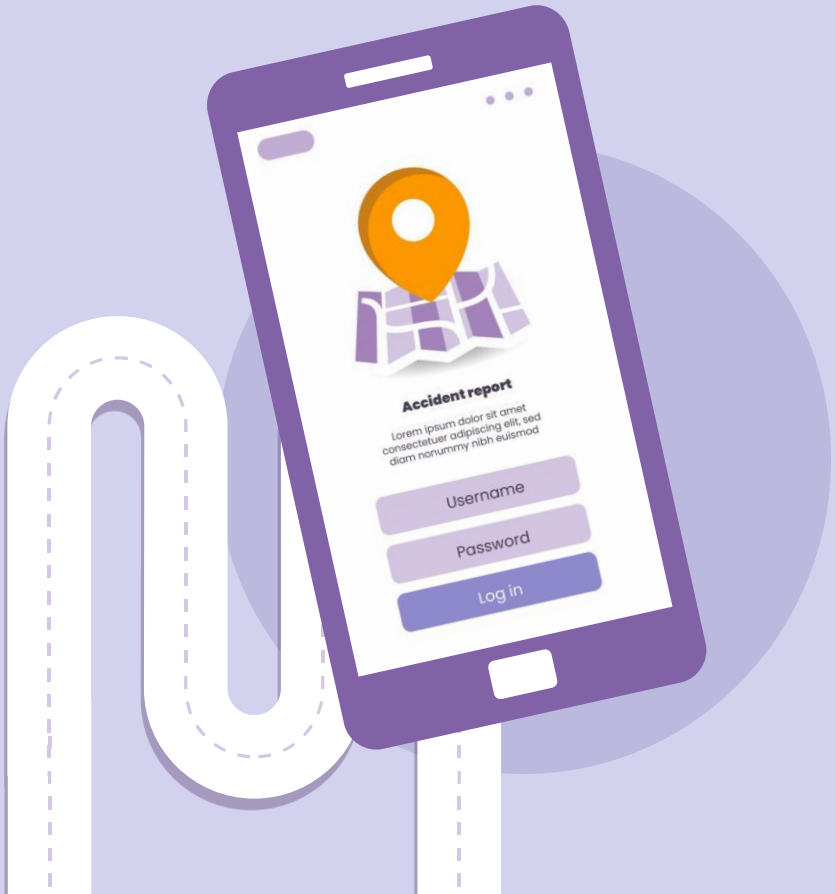
Data Collection

getTraf_data



getTraf_data

	A	B	C	D	E	F
1	date	hou	6908	6909	6910	Total Vol
2	01/01/2022	0	259	288	83	630
3	01/01/2022	1	223	296	77	596
4	01/01/2022	2	135	243	84	462
5	01/01/2022	3	103	175	46	324
6	01/01/2022	4	96	95	7	198
7	01/01/2022	5	69	111	17	197
8	01/01/2022	6	116	138	17	271
9	01/01/2022	7	181	205	50	436
10	01/01/2022	8	237	274	72	583
2898	01/05/2022	16	1060	1410	1168	3638
2899	01/05/2022	17	940	1233	840	3013
2900	01/05/2022	18	837	1175	701	2713
2901	01/05/2022	19	722	959	536	2217
2902	01/05/2022	20	583	791	376	1750
2903	01/05/2022	21	503	575	240	1318
2904	01/05/2022	22	394	447	143	984
2905	01/05/2022	23	260	341	105	706



03

Data Pre-Processing

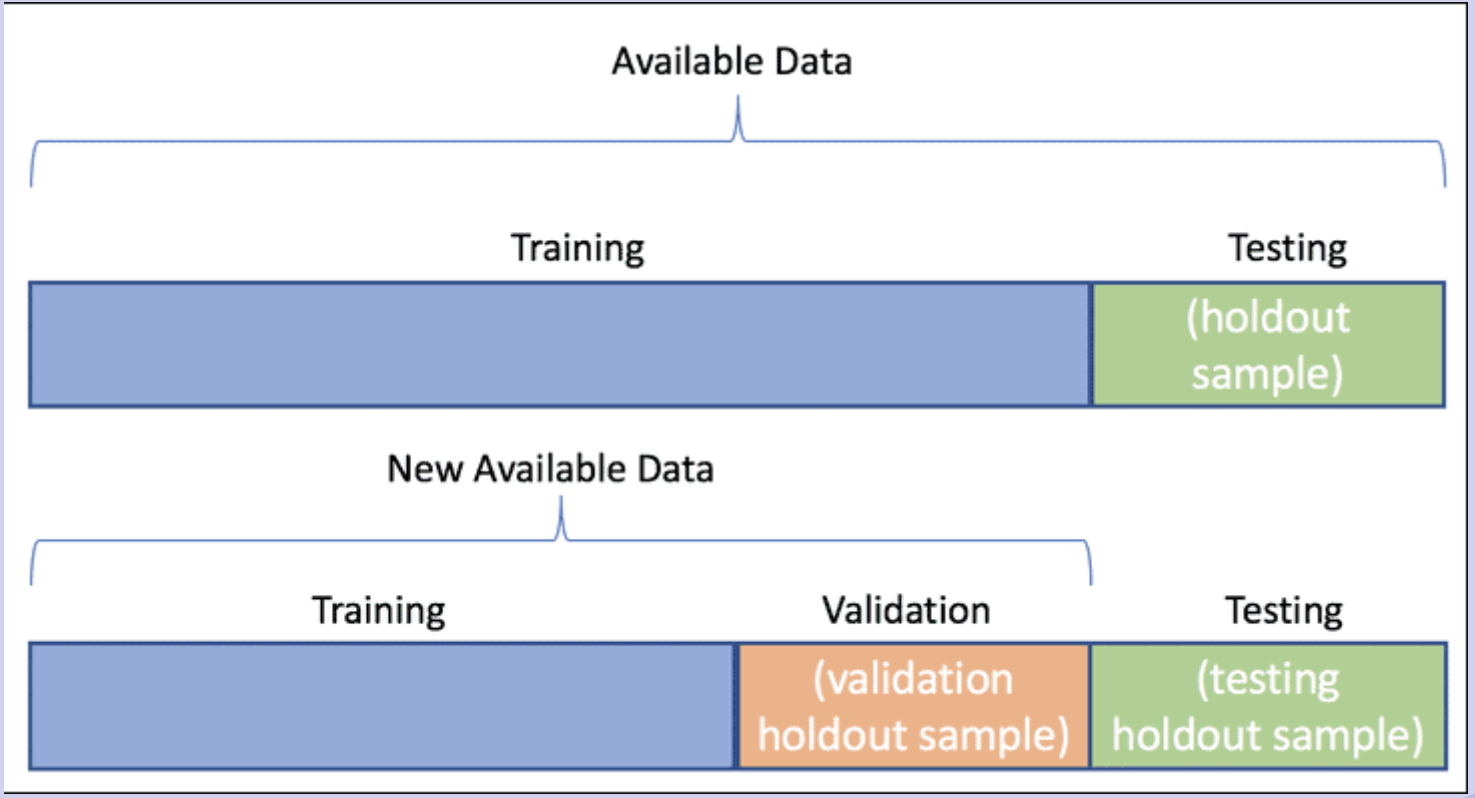
Handling missing values

Mean imputation

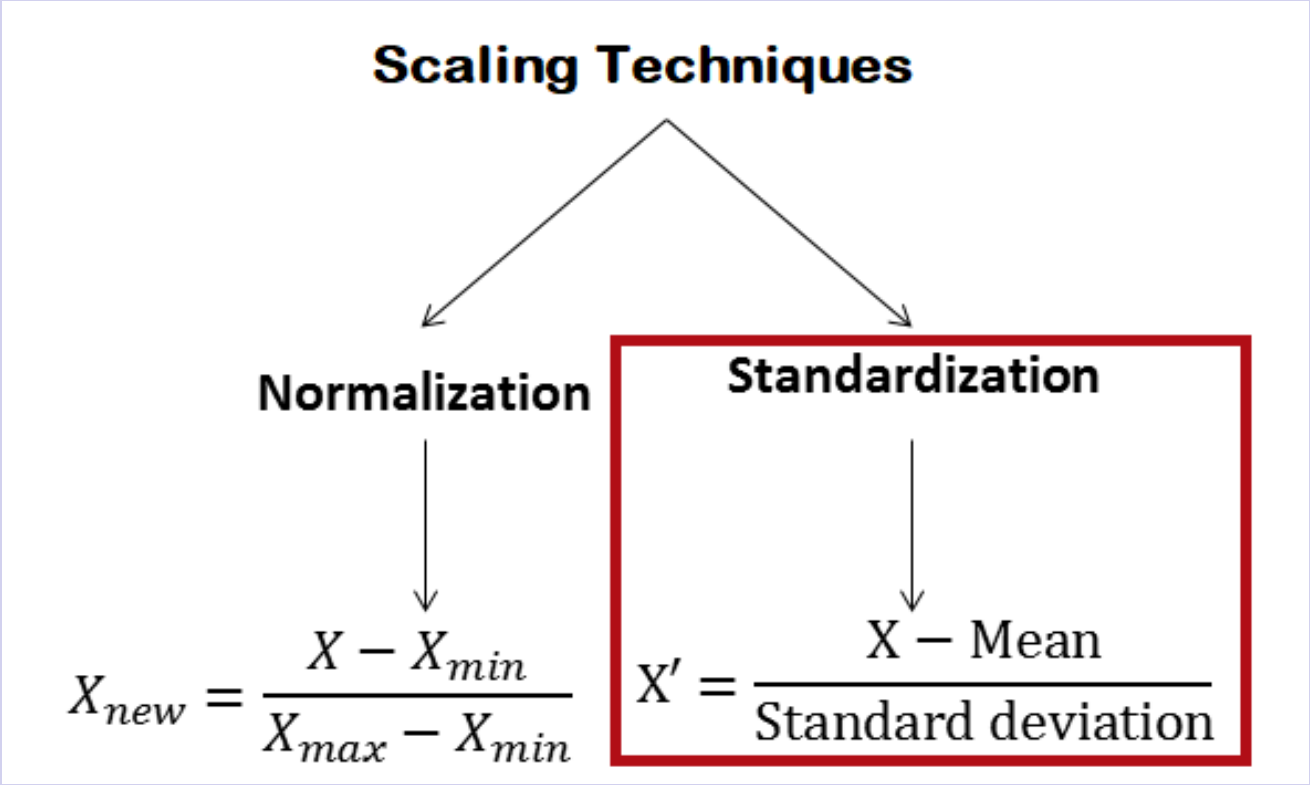
Forward fill

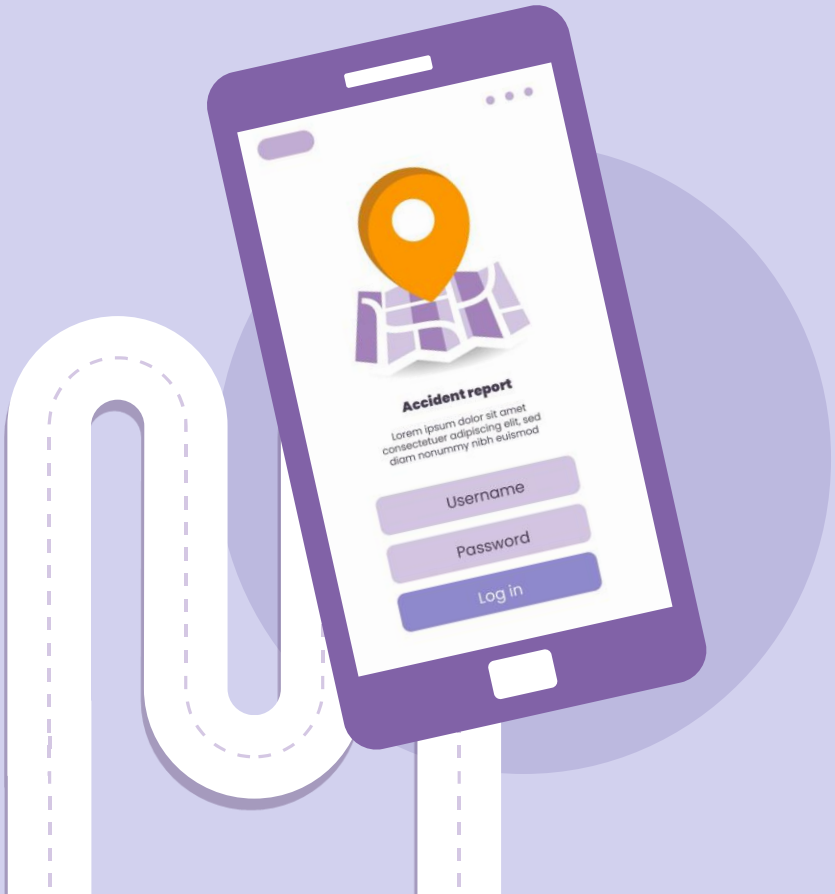
Backward fill

Split the dataset



Feature Scaling

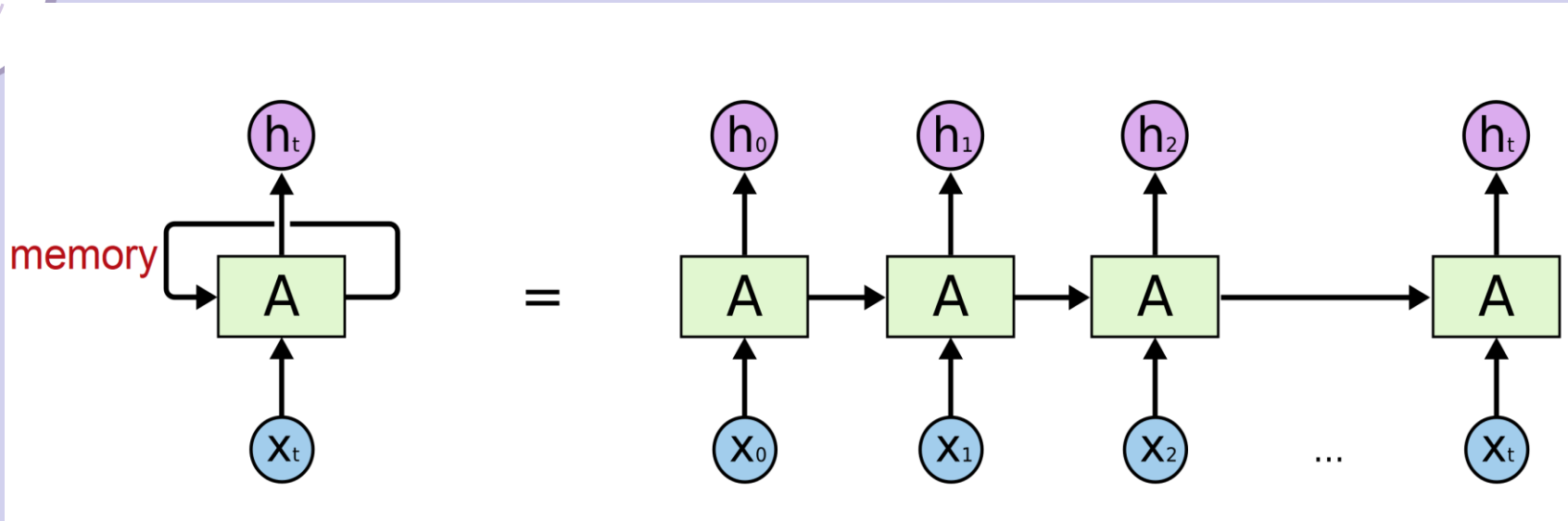




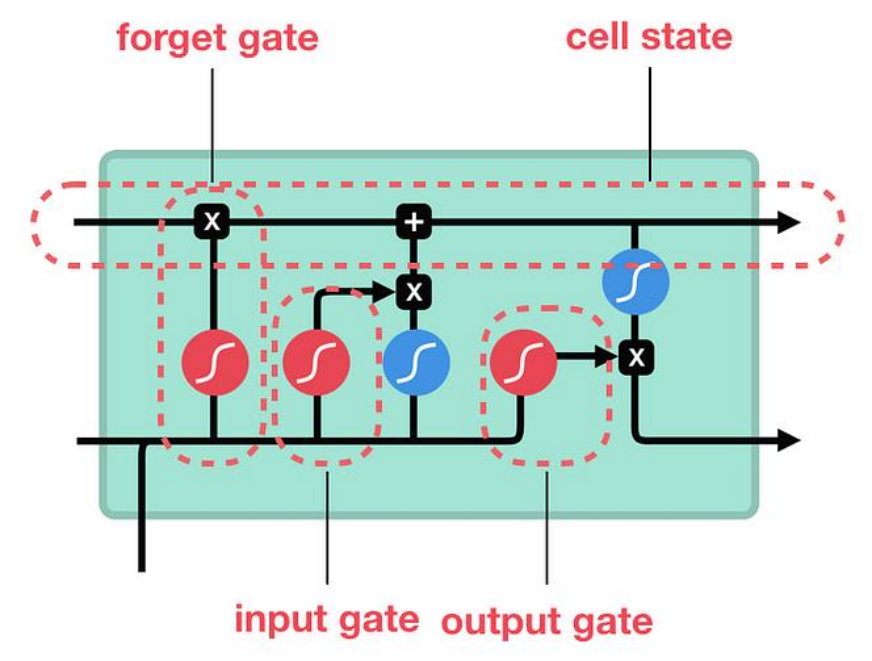
04

Model Architecture: RNN & LSTM

RNN Model



LSTM Model


forget gate
cell state
input gate
output gate


sigmoid
tanh
pointwise multiplication
pointwise addition
vector concatenation

Impact of RNN and LSTM on the model

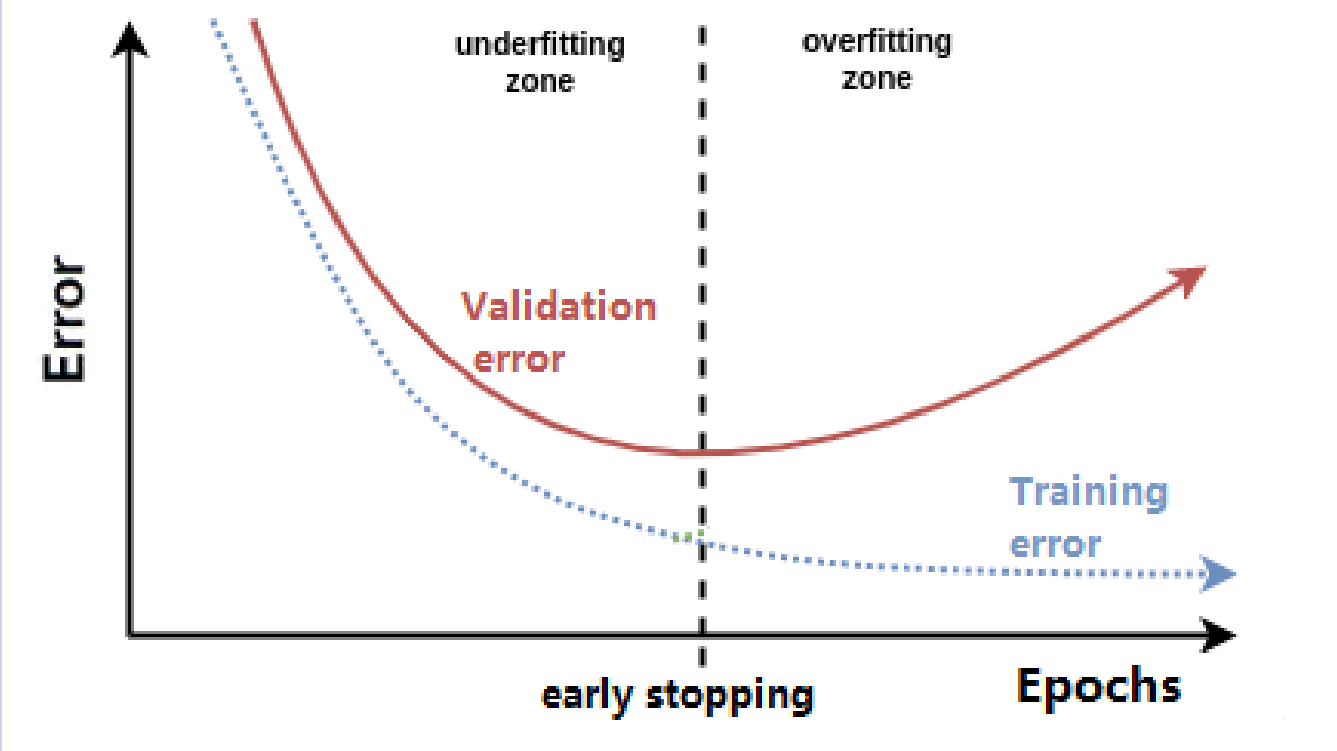
RNN

- ✔ Lower risk of overfitting
- 🚫 Less robust to noise.

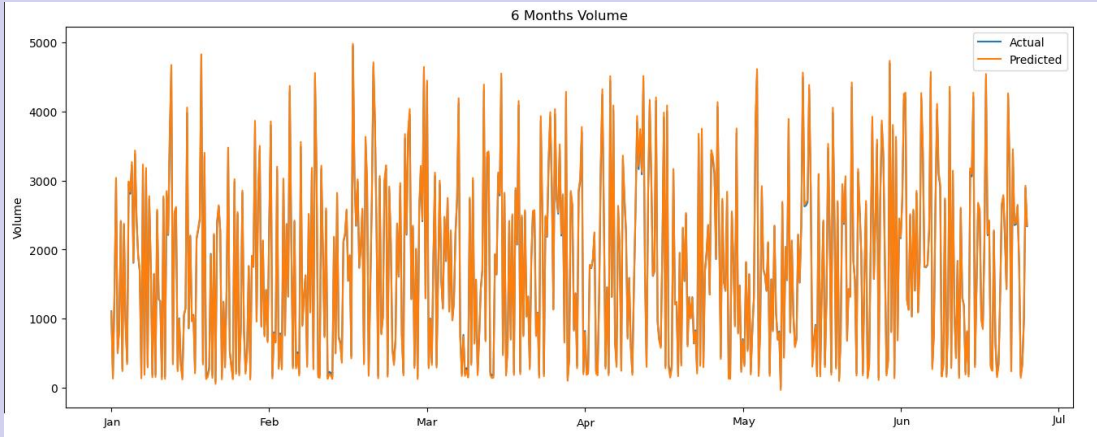
LSTM

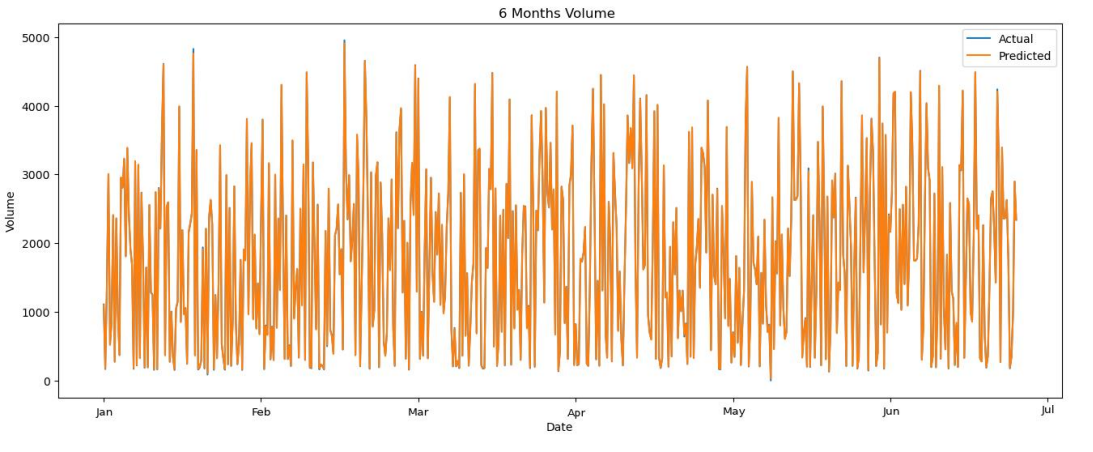
- ✔ Better performance on long sequences.
- 🚫 Slower to train and run.

Early stopping



Results


6 Months Volume
Actual
Predicted
Volume
5000
4000
3000
2000
1000
0
Jan
Feb
Mar
Apr
May
Jun
Jul


6 Months Volume
Actual
Predicted
Volume
5000
4000
3000
2000
1000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Date

Evaluation :

```
# Calculate MAPE for training and test sets
mape_train = mean_absolute_percentage_error(y_train, y_train_pred)
mape_test = mean_absolute_percentage_error(y_test, y_test_pred)

print(f'Training MAPE: {mape_train}%')
print(f'Test MAPE: {mape_test}%')
```

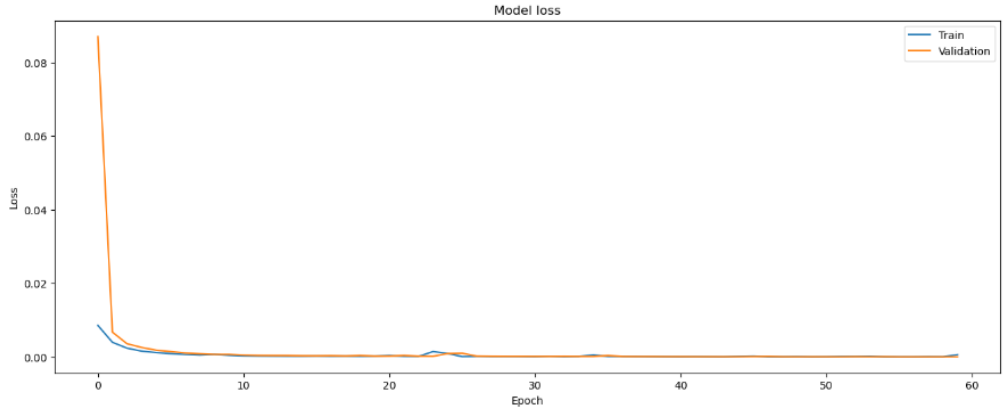
```
Training MAPE: 5.508993370855774%
Test MAPE: 4.227828255798661%
```

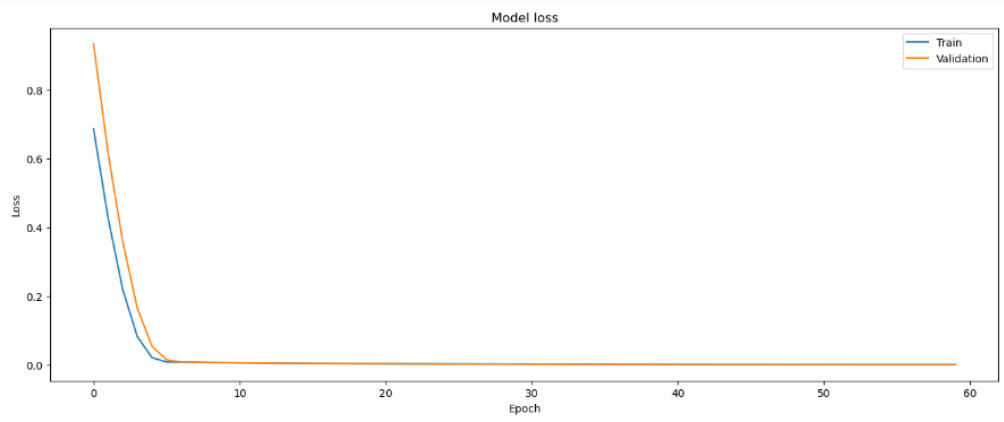
```
# Calculate MAPE for training and test sets
mape_train = mean_absolute_percentage_error(y_train, y_train_pred)
mape_test = mean_absolute_percentage_error(y_test, y_test_pred)

print(f'Training MAPE: {mape_train}%')
print(f'Test MAPE: {mape_test}%')
```

```
Training MAPE: 2.6603522927608365%
Test MAPE: 1.9339970744402364%
```

Training and Validation loss





Conclusion:

"Our research shows that using sensor data and LSTM networks can indeed be a powerful tool for predicting future traffic states"

Potential business applications for our model

- Smart Traffic Apps
- Autonomous Vehicles
- Taxi Companies.

Thank you for your time and attention