Universidad Autónoma del Estado de Morelos

Facultad de Contaduría Administración e Informática

Estimation model for arrival times in urban public transportation of Cuernavaca using artificial neural networks

Pedro Moreno

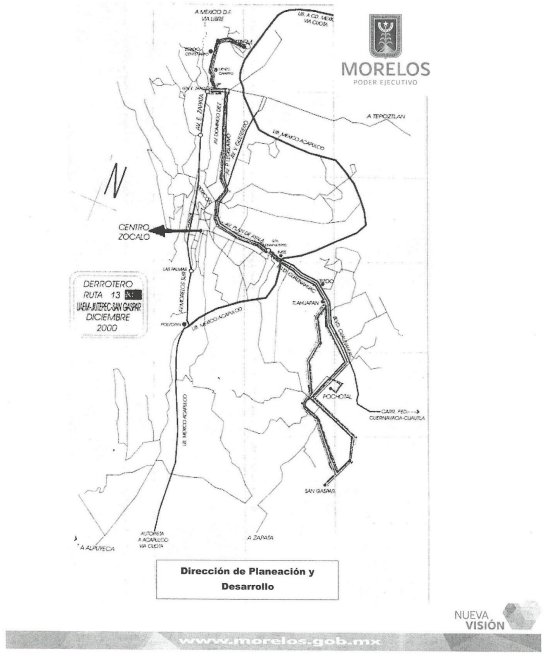
JUCC 2025

Introduction

- Importance of urban public transportation for sustainable cities
- Challenges: traffic congestion, unpredictable arrival times, lack of real-time data
- Need for intelligent systems to improve reliability and planning

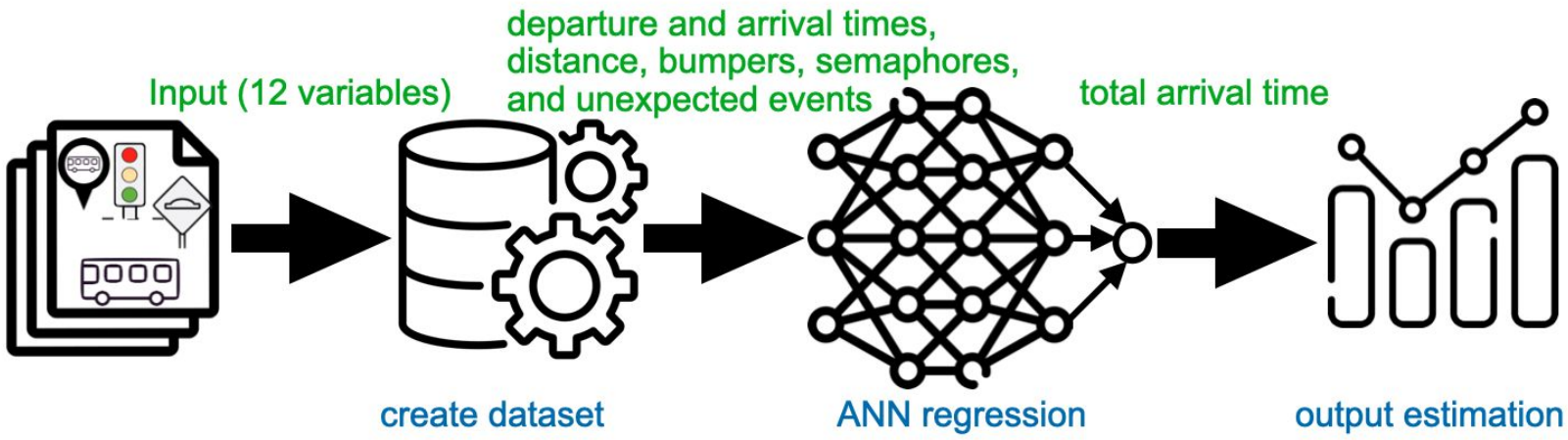
Problem Statement

- Inaccurate bus arrival times lead to:
  - Passenger dissatisfaction
  - Missed connections
  - Inefficient scheduling
- Route 13 in Cuernavaca-Jiutepec has 36 routes, 2,400+ buses, and high variability in service



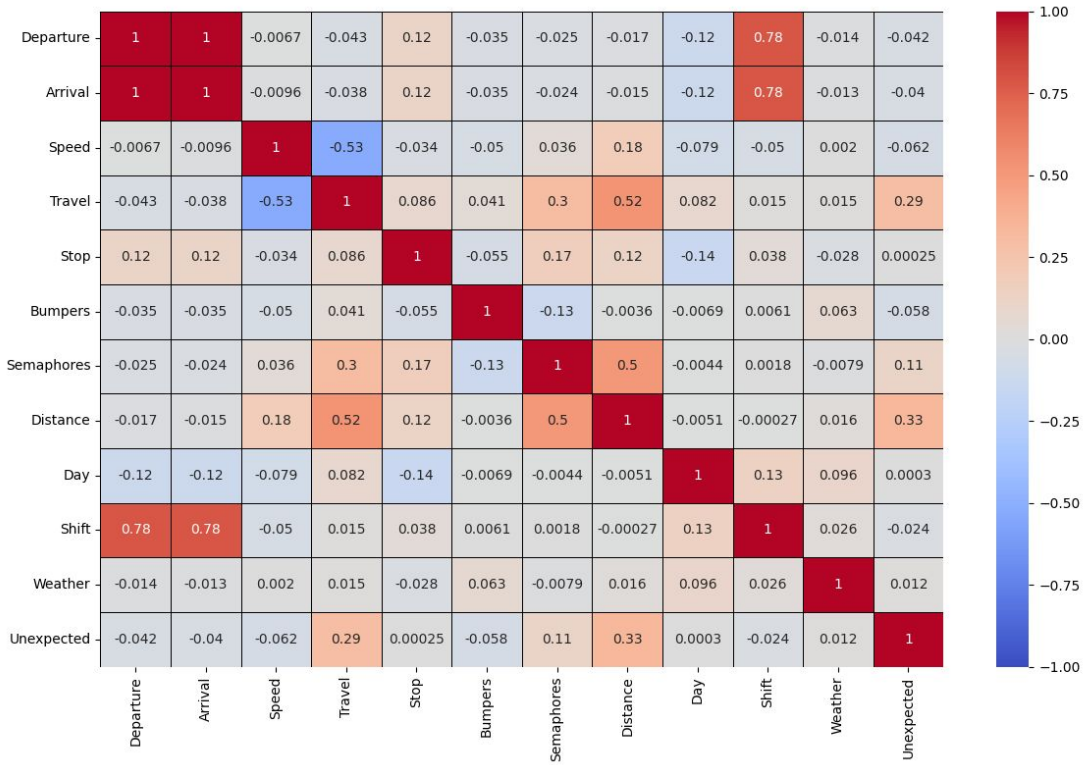
Proposed Method

- **Model:** Multilayer Perceptron ANN for regression
- **Input:** 12 features (e.g., speed, distance, traffic lights, weather)
- **Output:** Estimated travel time between any two bus stops
- **Activation:** ReLU in all layers



Data Collection & Preprocessing

- **Source:** 2,073 trip records from Route 13 (Oct 2024 – Feb 2025)
- **Features:** Departure/arrival times, speed, distance, events, etc.
- **Preprocessing:** Normalization, error correction, feature engineering
- **Correlation analysis** to identify key influencing variables



Case Study

- Route: UAEM to Walmart Jiutepec (~15.25 km)
- 37 stops (outbound), 41 stops (return)
- Data collected manually and via Google Maps

Summary statistics of ANN input variables

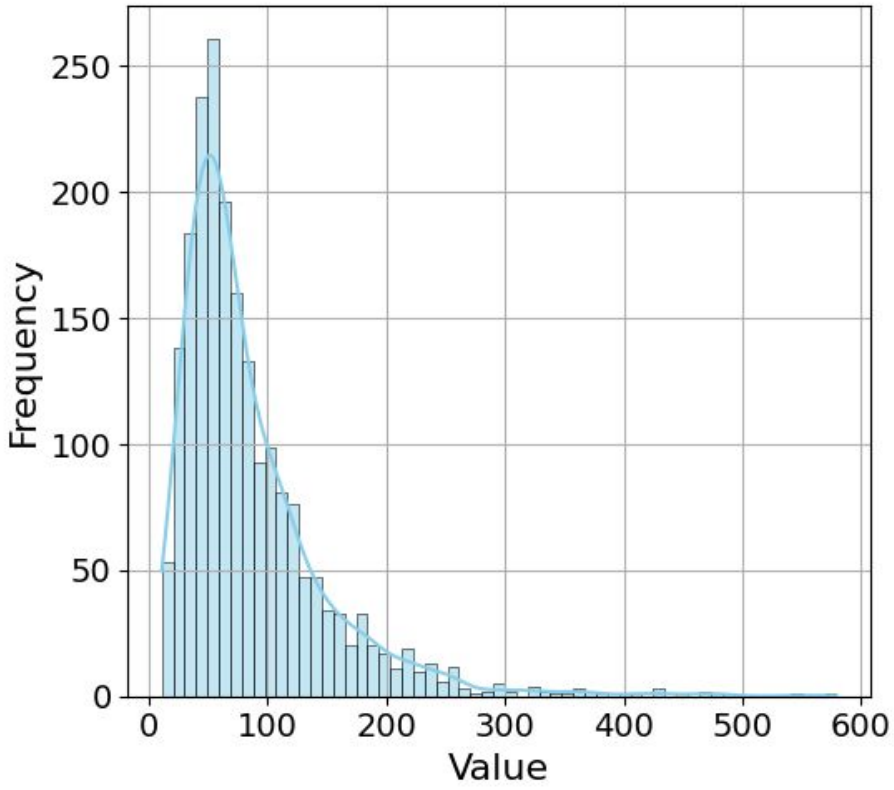
#	Variable	Mean	Median	Max	Min
1	Departure time (sec)	44 480.73	45 559.00	70 286.00	22 622.00
2	Arrival time (sec)	44 552.72	45 629.00	70 330.00	22 705.00
3	Distance traveled (m)	394.92	388.39	1 150.00	74.27
4	Travel time (sec)	71.98	54.00	557.00	7.00
5	Stop time (sec)	13.89	5.00	278.00	0.00
6	Real speed (km/h)	7.00	6.38	19.82	0.31
7	Bumpers	0.23	0.00	2.00	0.00
8	Semaphores	0.67	0.00	3.00	0.00
9	Day type (weekday/weekend)	0.70	1.00	1.00	0.00
10	Shift (morning/evening)	0.53	1.00	1.00	0.00
11	Weather (cold/warm)	0.04	0.00	1.00	0.00
12	Unexpected events	0.016	0.00	1.00	0.00

Hyperparameter Tuning

Method: Grid Search (GS)

Parameters Tuned:

- Batch size: 16–256
- Epochs: 50–1000
- Learning rate: 0.00001–0.01
- Layers: 3–4 hidden layers
- (e.g., 32-16-4, 64-6-16-6)
- Optimizers: Adam, RMSprop



Best Config: 3 layers, 32-16-4 neurons, 1000 epochs, Adam, LR=0.01

Experimental Results

Platform: Python, TensorFlow, Scikit-learn on ClusterUY (40-core Xeon, 128 GB RAM)

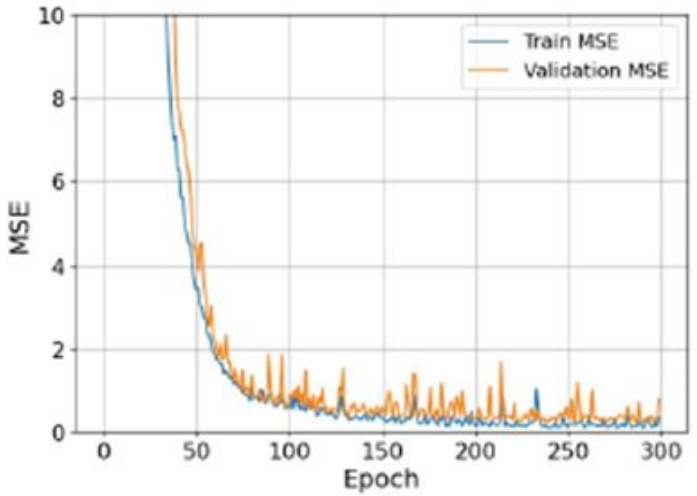
Metrics:

- MAE: 0.5951
- RMSE: 0.7932
- R^2: 0.9998

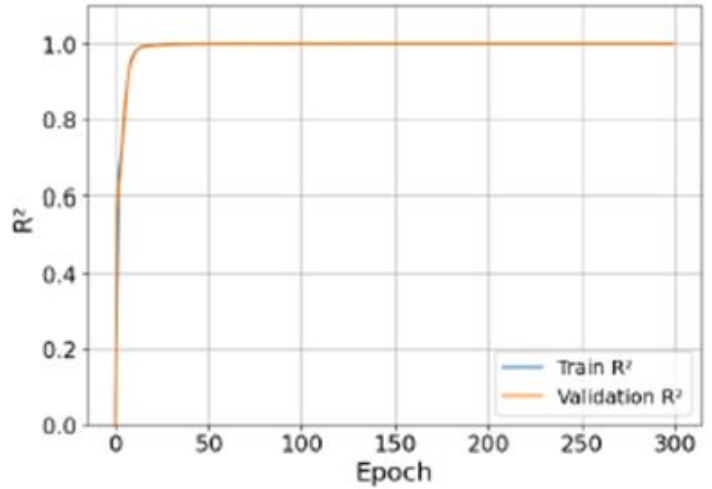
Comparison: ANN outperformed Multivariate Linear Regression by 14.99%

Results

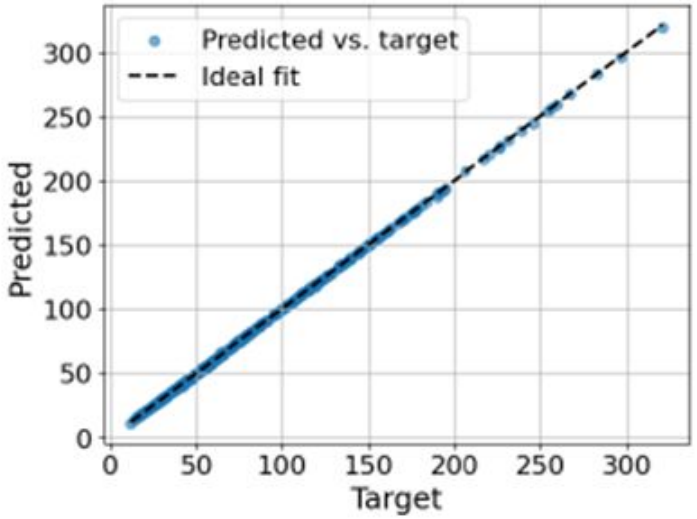
- High accuracy and generalization
- ANN model shows minimal overfitting
- Strong correlation between predicted and actual values
- Effective for ITS applications in medium-sized cities



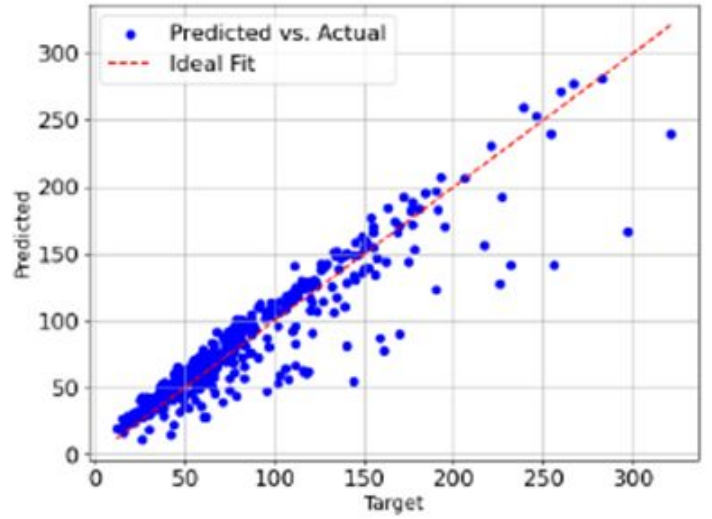
(a) MSE loss function



(b) R^2 function



(c) Predicted vs. target (ANN)



(d) Predicted vs. target (MLR)

Conclusions & Future Work

- ANN model effectively estimates bus arrival times
- Improves planning and passenger satisfaction

Future Work:

- Integrate GPS and passenger count data.
- Use synthetic data (GANs) for model enhancement.
- Expand to other routes and cities.

Thank you

Questions?