# Vehicle Trajectory Forecasting: A classification approach

## Comparison between LSTM and Transformers

# Contents

# Introduction

# Introduction

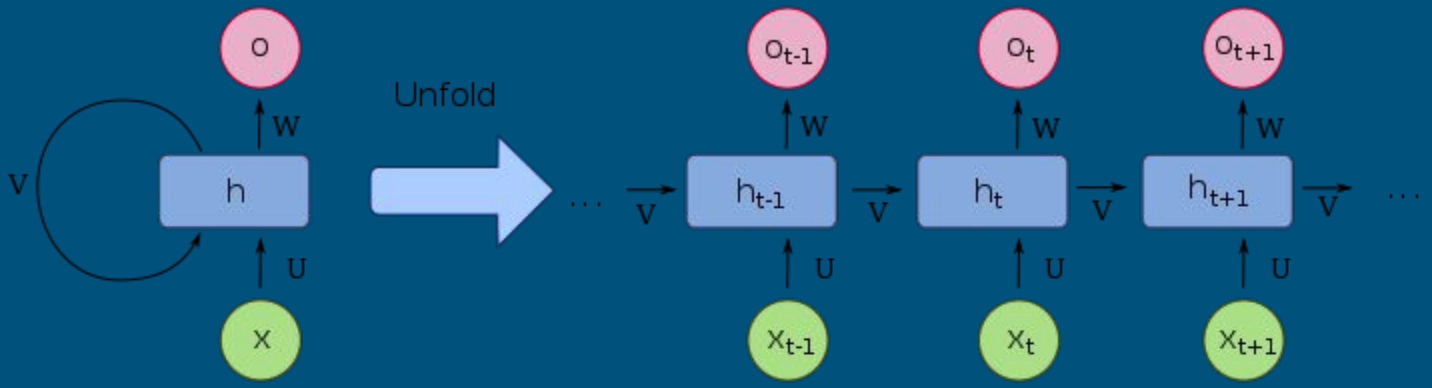
- Learning from a sequence.
- Memory.
- Attention is all you need (2018).

# Background

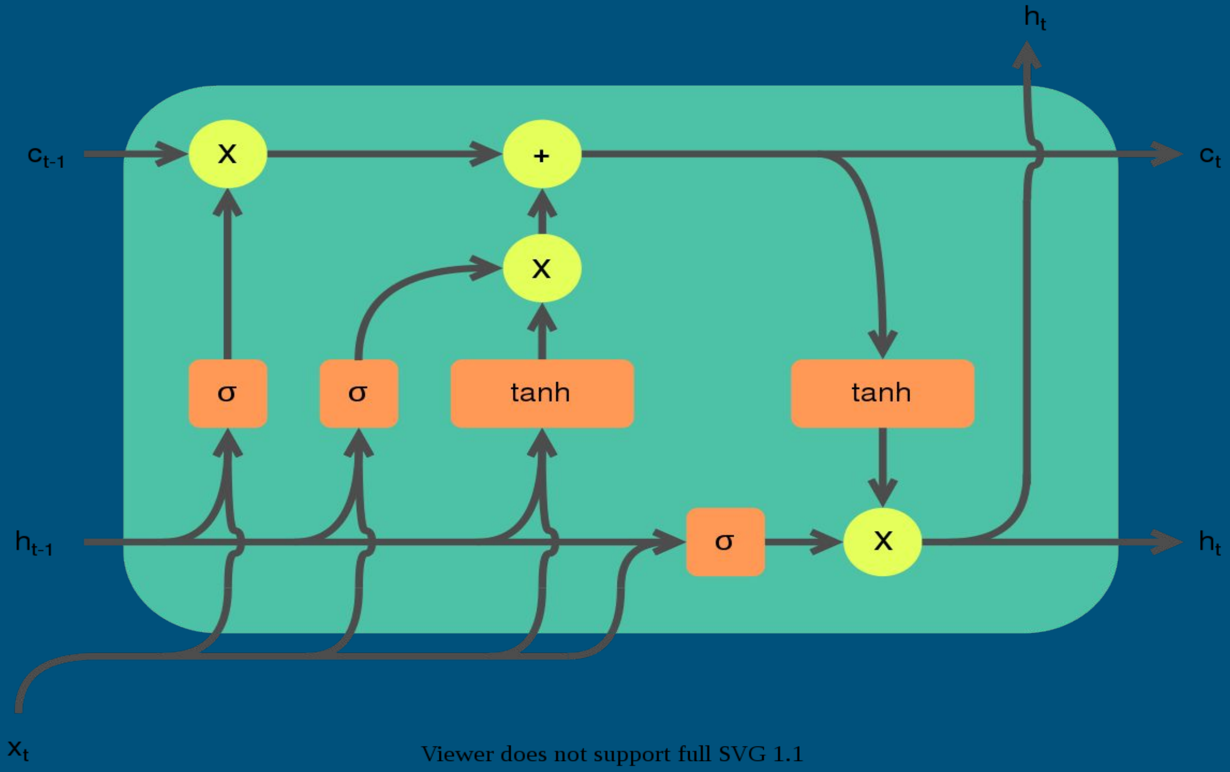
# Recurrent Neural Networks

- Memorizing sequence.
- Vanishing gradient.
- Long Short-Term Memory
- Complexity.

# RNN unfolded



# LSTM CELL




Viewer does not support full SVG 1.1

# Transformer Networks

- Attention
- Allows parallelization.

# Transformer architecture

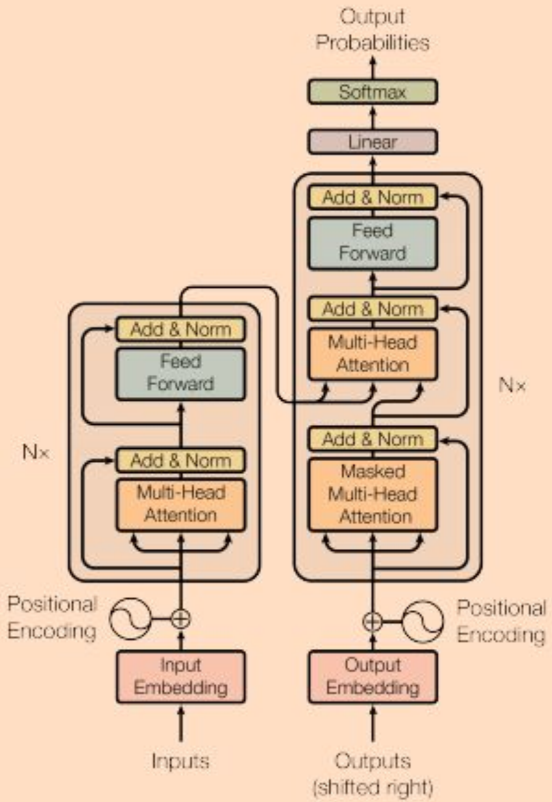
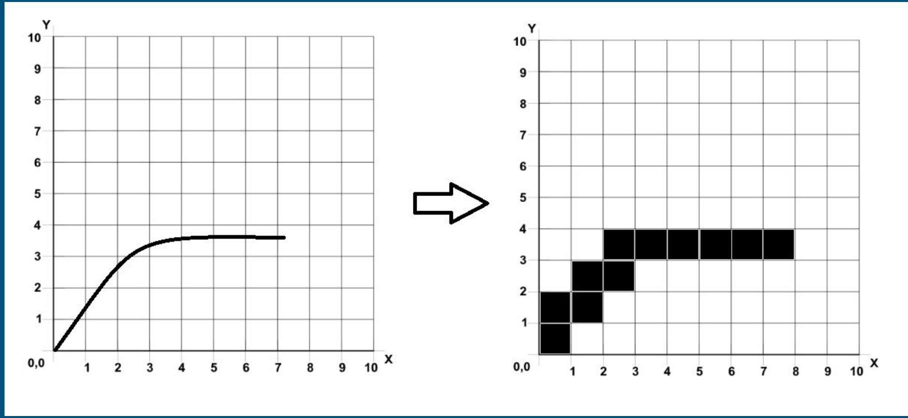



Figure 1: The Transformer - model architecture.

# Data description

- Vehicle movement in urban environments.
- Classification problem.
- 900 unique bins.
- Randomly rotating the trajectories.

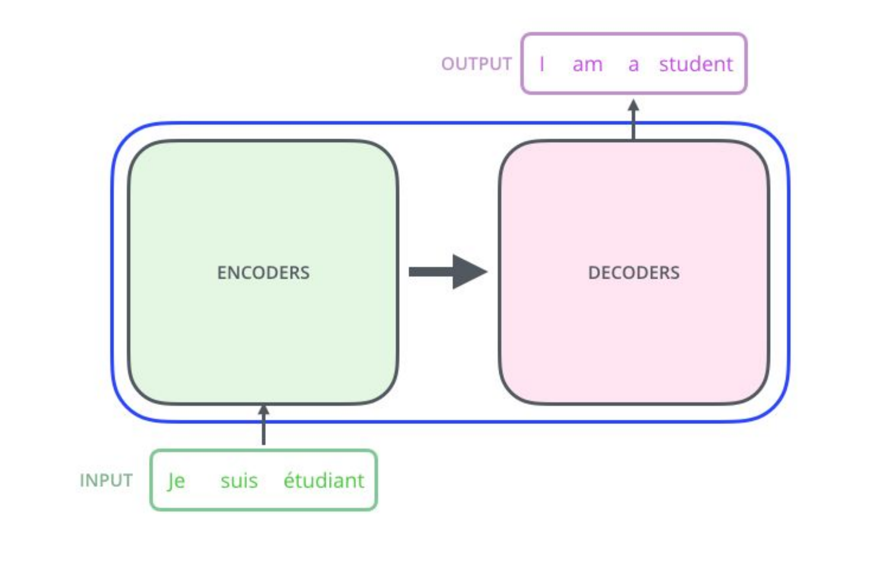


- Red : Training data
- Blue : Testing data

- Turns and roundabouts
- Split up into many sections.

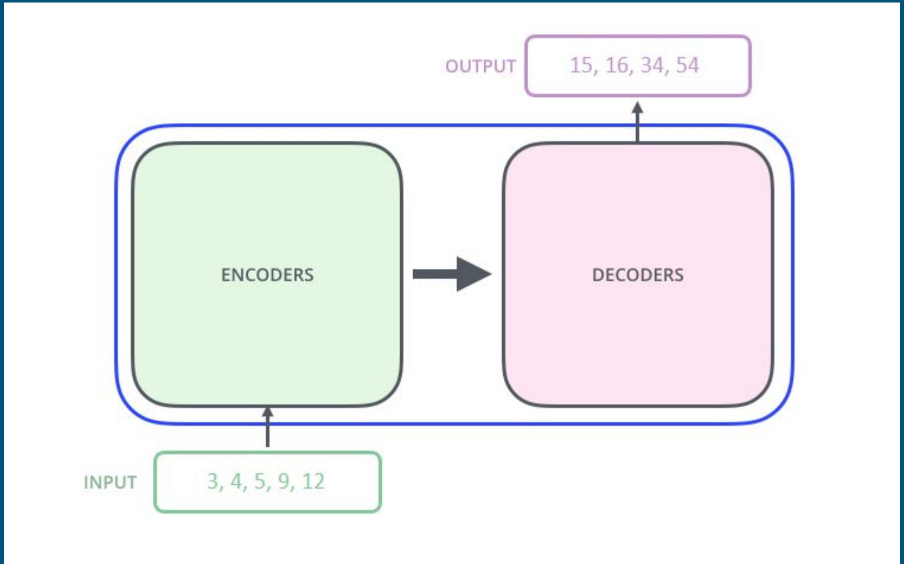
# Method: Input Output

- Modeled similarly to a translation task
- Machine Translation:
  - At each time step:
    - Assigns a probability for each word in the vocabulary
    - Predict the word with highest probability



# Method: Input Output

- Our method:
    - At each time step:
        - Assigns a probability for each bin in the possible bin space
        - Predict the bin position with highest probability

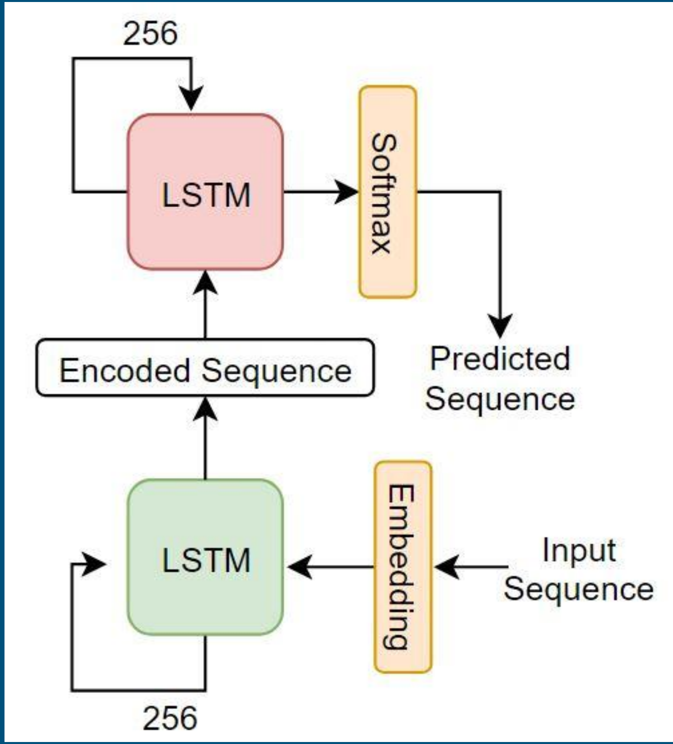


# Method: Transformer

- Implemented using PyTorch
- 256 dimensionality, 6 layers & 8 attention heads
- Cross-entropy loss
- Stochastic Gradient Descent with a decaying learning rate
- 20% dropout and early stopping

# Method: LSTM

- Implemented in Tensorflow using the Keras API
- Encoder and decoder both consist of 256 LSTM units
- Stochastic Gradient Descent with Momentum optimizer
- Categorical cross entropy loss
- To avoid overfitting dropout of 20% was used on the LSTM layers along with early stopping



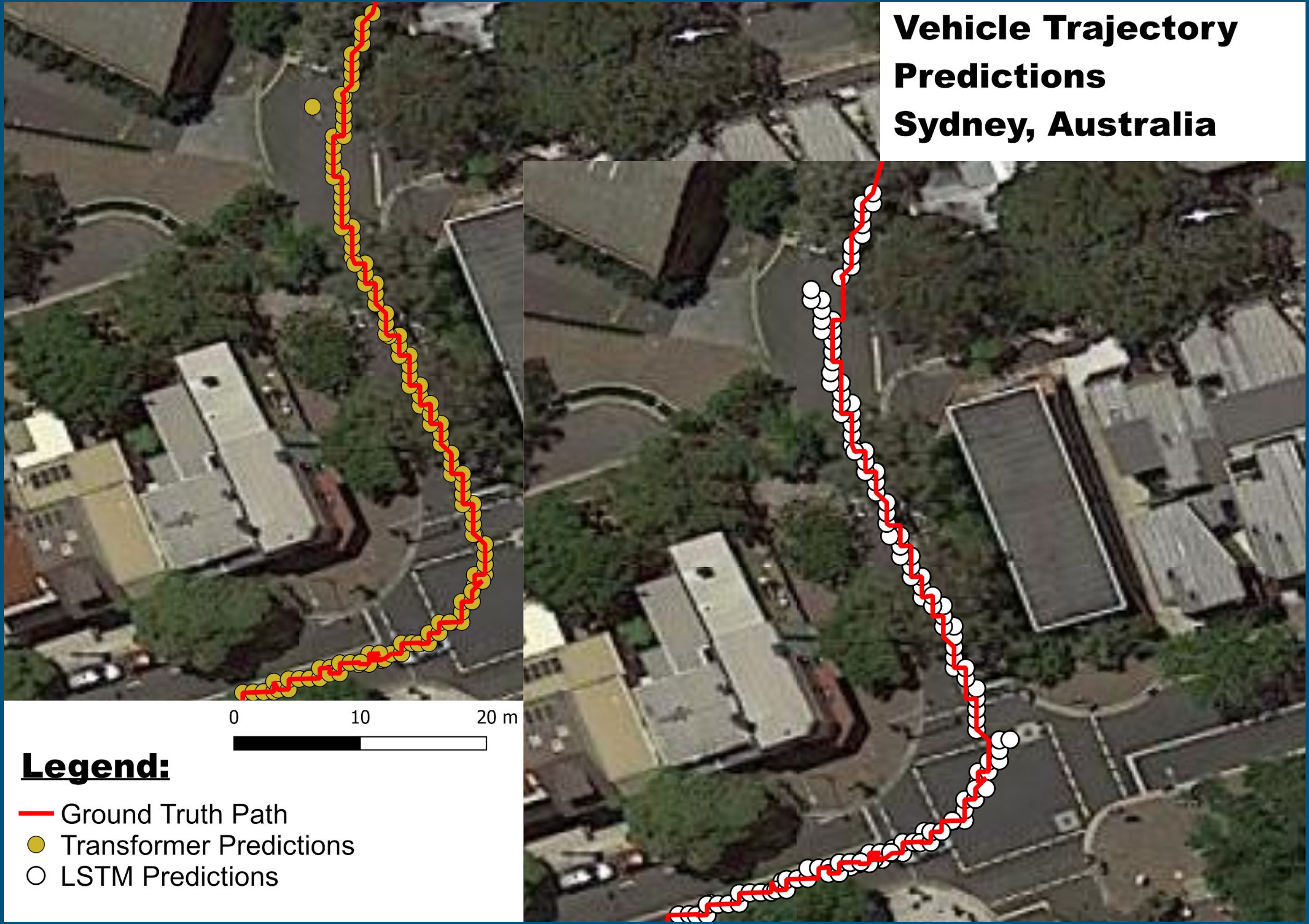
# Method: Evaluation Metrics

- Average Displacement Error (ADE):
  - The average of the error in meters between the ground truth and the predicted trajectory at every time step.
- Final Displacement Error (FDE):
  - The error in meters between the ground truth and the predicted trajectory at the last time step.
- Accuracy:
  - average of how many predictions are equal to ground truth
- Percentage of Perfect Predictions:
  - ratio of trajectories predicted with zero error

# Results

| Datasets (inputs, outputs) |
|---|
| 30, 70 |
| 11, 22 (every 3rd coordinate) |
| 30, 70 |
| 100, 100 |

| # of Bins | Model | Train loss | Validation loss | ADE (meters) | FDE (meters) | Bin Accuracy | % Perfect |
|---|---|---|---|---|---|---|---|
| 10,000 | Transformer | 1.06 | 1.008 | 0.1267 | 0.32 | 0.683 | 0.138 |
| 900 | Transformer | 0.16 | 0.135 | 0.0146 | 0.024 | 0.994 | 0.953 |
| 900 | Transformer | 0.12 | 0.094 | 0.0228 | 0.075 | 0.987 | 0.882 |
| 900 | Transformer | 0.17 | 0.482 | 0.1336 | 0.536 | 0.916 | 0.439 |
| | | | | | | | |
| 10,000 | LSTM | 1.754 | 2.205 | 1.2794 | 1.93 | 0.024 | ~0 |
| 900 | LSTM | 0.912 | 1.318 | 0.6456 | 0.925 | 0.386 | 0.023 |
| 900 | LSTM | 0.806 | 1.177 | 0.5909 | 0.993 | 0.444 | 0.031 |
| 900 | LSTM | 1.125 | 2.077 | 1.8461 | 2.558 | 0.331 | 0.075 |


Vehicle Trajectory
Predictions
Sydney, Australia
0
10
20 m
Legend:
Ground Truth Path
Transformer Predictions
LSTM Predictions

# Final Thoughts

- Transformer vs LSTM
  - More complex LSTM models / architectures
  - Regression vs classification
- Loss of information due to bins

# Questions?