ANNs using Pytorch (Regression) – NYC Taxi Fares

Data:

In this playground competition, hosted in partnership with Google Cloud and Coursera, you are tasked with predicting the fare amount (inclusive of tolls) for a taxi ride in New York City given the pickup and dropoff locations. While you can get a basic estimate based on just the distance between the two points, this will result in an RMSE of $5-$8, depending on the model used (see the starter code for an example of this approach in Kernels). Your challenge is to do better than using Machine Learning techniques!

Data reference: New York City Taxi Fare Prediction | Kaggle

Attributes:

- pickup_datetime - timestamp value indicating when the taxi ride started.
- pickup_longitude - float for longitude coordinate of where the taxi ride started.
- pickup_latitude - float for latitude coordinate of where the taxi ride started.
- dropoff_longitude - float for longitude coordinate of where the taxi ride ended.
- dropoff_latitude - float for latitude coordinate of where the taxi ride ended.
- passenger_count - integer indicating the number of passengers in the taxi ride.

Key asks:

- You can get a basic estimate based on just the distance between the two points, this will result in an RMSE of $5-$8, depending on the model used. Your challenge is to do better than these using Machine Learning techniques!