

NTUST, CSIE

Machine Learning (CS5087701), Fall 2025

Final Project

Candidate Topics: Data-oriented

You can choose from one of the following datasets and test a few machine learning algorithms on it.

- I. *ECML/PKDD 15: Taxi Trajectory Prediction (I)*
<https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i>
- II. *ECML/PKDD 15: Taxi Trip Time Prediction (II)*
<https://www.kaggle.com/c/pkdd-15-taxi-trip-time-prediction-ii>
- III. *Sentiment Analysis*
<https://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences>
http://ai.stanford.edu/~amaas/data/sentiment/aclImdb_v1.tar.gz
- IV. *Real and Fake Face Detection*
<https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection/data>
- V. *Huge Stock Price Data: Intraday Minute Bar*
<https://www.kaggle.com/datasets/arashnic/stock-data-intraday-minute-bar/data>
- VI. *Global Fishing Watch data*
<https://globalfishingwatch.org/datasets-and-code/>
- VII. *Static IR Drop Estimation Using Machine Learning*
<https://github.com/ASU-VDA-Lab/ML-for-IR-drop?tab=readme-ov-file>
- VIII. *MacroRank*
<https://drive.google.com/drive/folders/1TKHLMwHAMXxGo2zsbVSbO51Qv1HiKc8O>
- IX. *ImageNette dataset*
<https://github.com/fastai/imagenette>