Surrogate Loss Learning for Dynamic Time Warping(DTW)

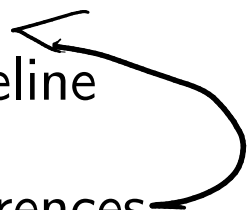
Supervisor: ~~Vijaya Yalavarthi~~ Prof. Schmidt-Thieme

Tanvi Joglekar

Univesity of Hildesheim

Agenda

1. Motivation
2. Literature Review
3. Research Idea
4. Data Foundation
5. Timeline
6. References



Motivation

- **Dynamic Time Warping(DTW)** is a technique used to quantify the similarity between two varying time sequences.
- DTW and its variants like soft-DTW, Shape and Time Distortion Loss have bounded efficiency and performance.
- This is because of their hyper-parameter dependency and quadratic complexity.
- Surrogate Losses are a superior option.

Motivation

Research has manifested that a neural network can approximate a desired loss.

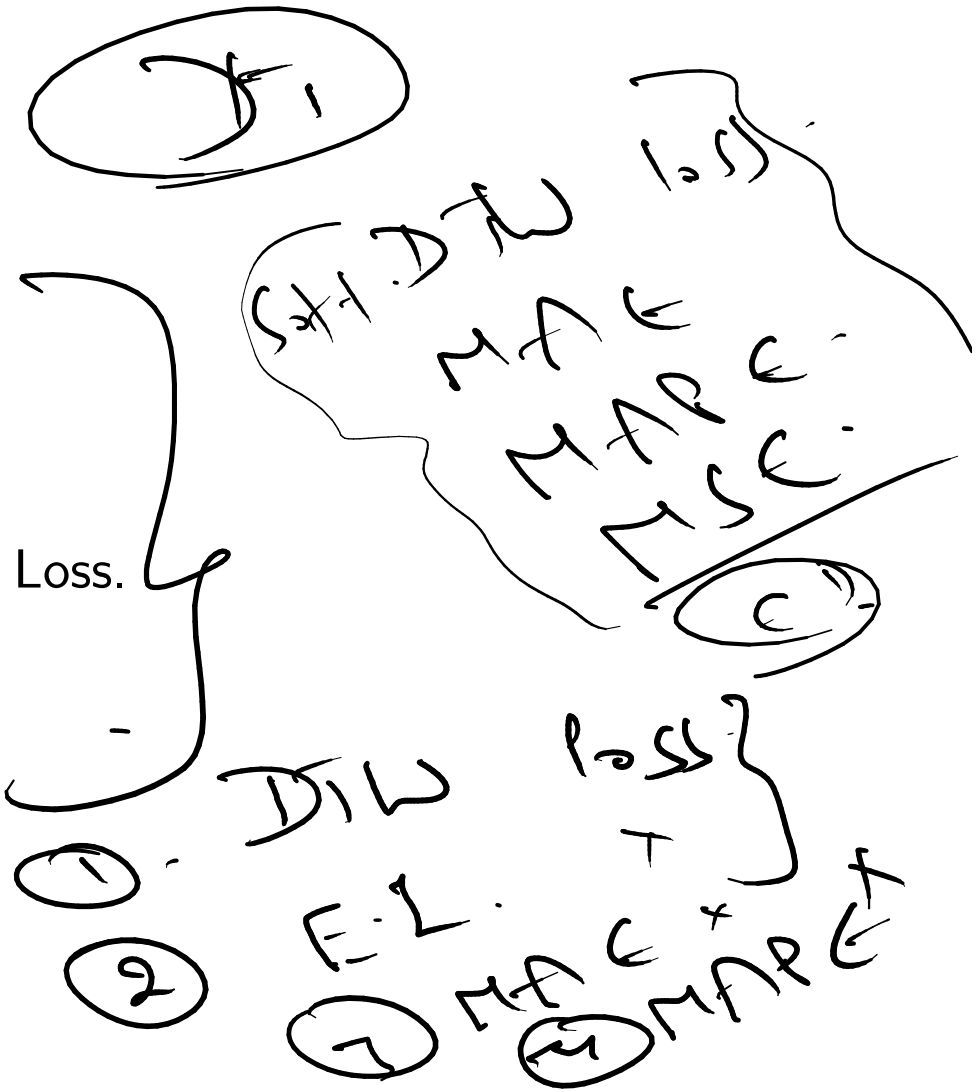
Goal: Develop a surrogate loss function for Shape and Time Distortion Loss(DILATE) as a meta-level neural network.

- Surrogate Loss network will be trained before forecasting model.

Literature Review

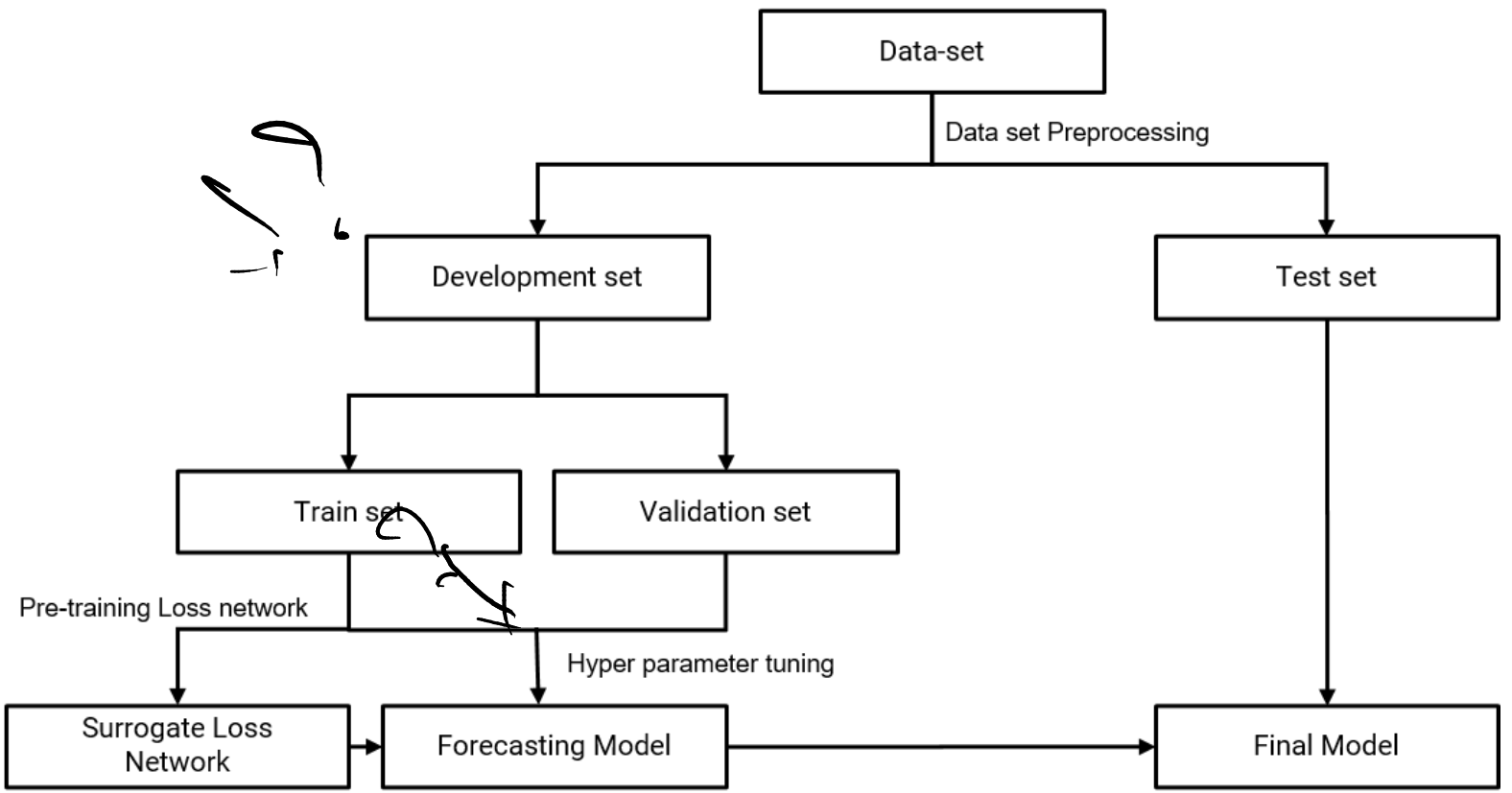
Baselines are bifurcated as:

- DTW based:
  - soft-DTW
  - Shape and Time Distortion Loss.
- Surrogate Loss Learning for:
  - MAE
  - MAPE
  - MSE



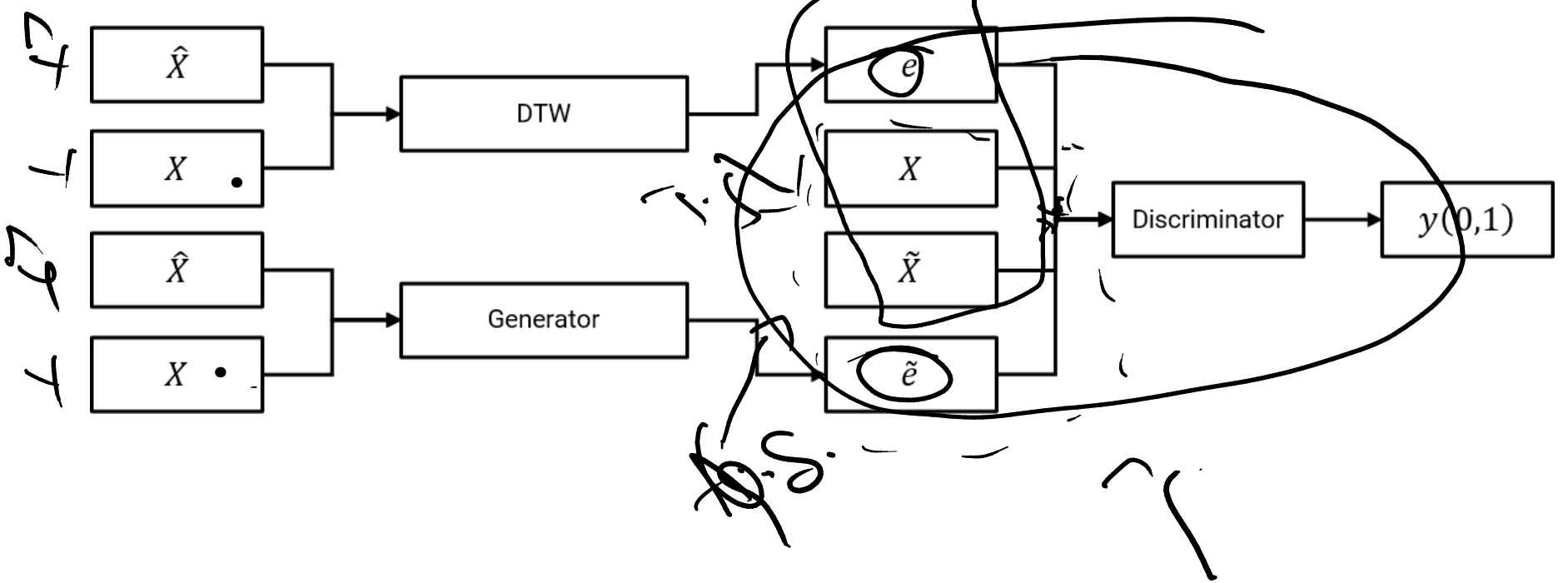
Research Idea

Overall Process:



Research Idea

Surrogate Loss Network- GAN:

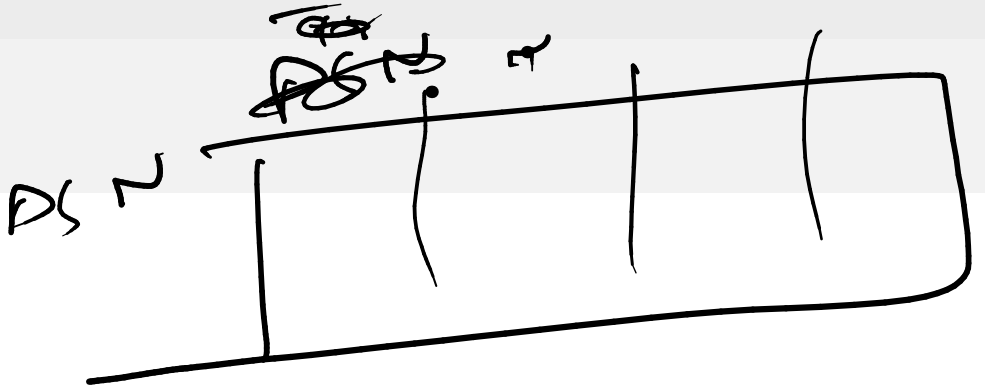


Data Foundation

Public Data-sets:

- Traffic:
  - Hourly Traffic Data from California State Department.
  - Data Tenure of 2015-2016.
  - Size: 17,544
- Electricity:
  - Electricity Load Diagrams Dataset, containing the electricity consumption every 15 minutes.
  - Data Tenure of 2011-2014.
  - Size: 26,304
- Retail:
  - Favorita Dataset is a combination of metadata for different products and the stores, sampled at the daily level.
  - Data Tenure for 2012.

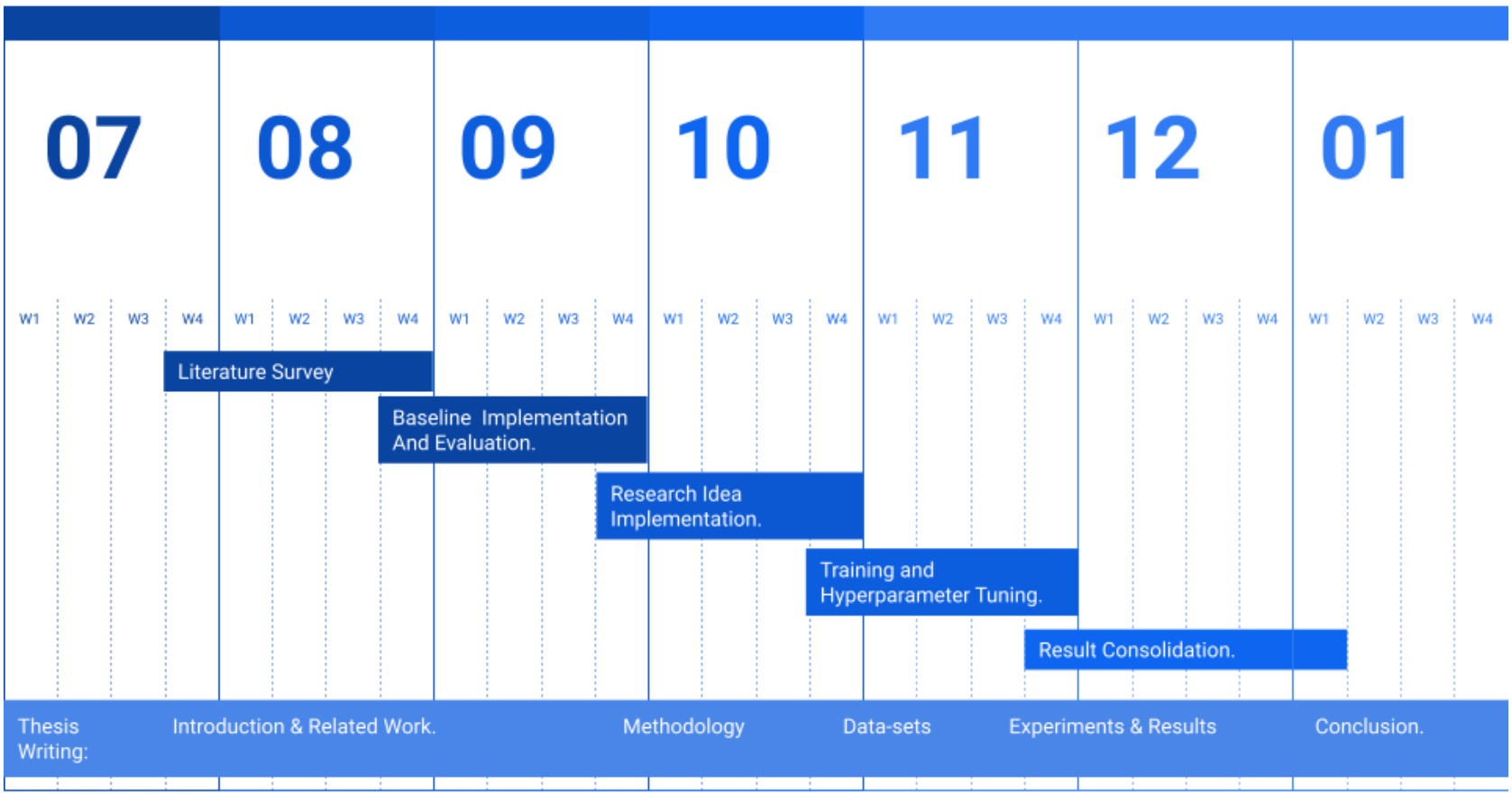
Data Foundation



- Complete Journey:
  - A record set of household transactions of about 2500 households that are frequent shoppers at a retailer.
  - Data Tenure for 2 years.
  - Size: 300 million

Internal Data-set(From MunichRe):
SME sensor data-set
OR
RBB data-set. (Confirming it with the supervisor at the Company.)

Timeline


07
08
09
10
11
12
01
W1 W2 W3 W4 W1 W2 W3 W4 W1 W2 W3 W4 W1 W2 W3 W4 W1 W2 W3 W4 W1 W2 W3 W4 W1 W2 W3 W4
Literature Survey
Baseline Implementation And Evaluation.
Research Idea Implementation.
Training and Hyperparameter Tuning.
Result Consolidation.
Thesis Writing:
Introduction & Related Work.
Methodology
Data-sets
Experiments & Results
Conclusion.

References

- Cuturi, Marco, and Mathieu BlondelCuturi, Marco, and Mathieu Blondel. Soft-DTW: a differentiable loss function for time-series, International Conference on Machine Learning. PMLR, 2017.
- Guen, Vincent Le, and Nicolas Thome. Shape and Time distortion loss for training deep time series forecasting models, arXiv:1909.09020(2019).
- Grabocka, Josif, Randolf Scholz, and Lars Schmidt-Thieme. Learning surrogate losses, arXiv:1905.10108 (2019).
- Lamb Alex et al, Professor Forcing: A New Algorithm for Training Recurrent Networks 2016(https://arxiv.org/abs/1610.09038)
- Electricity Data-set: https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014
- Traffic Data-set: http://pems.dot.ca.gov/

References

- Complete Journey Data-set: https://www.dunnhumby.com/source-files/
- Favorita Data-set: https://www.kaggle.com/c/favorita-grocery-sales-forecasting/data