

# Project Proposal Report: Eric Larsen

**Tentative project title:** Temporal Projection of Grocery Sales

**Type:** Research

**Student name:** Eric Larsen

## Project summary:

For this project I will be participating in the Kaggle competition for temporal prediction of grocery sales shown at the URL <https://www.kaggle.com/competitions/store-sales-time-series-forecasting/overview>. The goal of this project is to learn how to use LSTM, RNN, GRU, and Transformer Neural Networks to predict a time series dependent variable for the future. This competition will be used to understand the ability of Neural Networks to predict time series information. The goal is to understand what Neural network structure performs the best and how they can be applied to other problems. This project will be inputting temporal data from a large Ecuadorian-based grocery retailer to predict the unit sales of items sold at this grocery retailers stores. The dataset is given on Kaggle and is readily available. The input parameters are of range 178 for Store Identification, family of product, and holiday information to map to a prediction output of approx. 33 outputs for each store id with there being 54 stores, leading to a total output of about 1800 values. Therefore, with so many values to predict this project is focused on temporal prediction accuracy of different time series network structures. Understanding the overall performance of these different time series networks then lends itself to further applications where time series predictions are necessary.

## References:

Similar projects of temporal forecasting for systems can be found in the following papers.

Tang, X., Yao, H., Sun, Y., Aggarwal, C., Mitra, P., & Wang, S. (2020). Joint Modeling of Local and Global Temporal Dynamics for Multivariate Time Series Forecasting with Missing Values. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04), 5956-5963. <https://doi.org/10.1609/aaai.v34i04.6056>

Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J., & Li, Z. (2018). Deep Multi-View Spatial-Temporal Network for Taxi Demand Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1). <https://doi.org/10.1609/aaai.v32i1.11836>

Wu, Y., Ni, J., Cheng, W., Zong, B., Song, D., Chen, Z., Liu, Y., Zhang, X., Chen, H., & Davidson, S. B. (2021). Dynamic Gaussian Mixture based Deep Generative Model For Robust Forecasting on Sparse Multivariate Time Series. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1), 651-659. <https://doi.org/10.1609/aaai.v35i1.16145>

Zhang, C., Song, D., Chen, Y., Feng, X., Lumezanu, C., Cheng, W., Ni, J., Zong, B., Chen, H., & Chawla, N. V. (2019). A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 1409-1416. <https://doi.org/10.1609/aaai.v33i01.33011409>

## Availability of dataset/code:

The data set is listed on Kaggle and is readily available. It is a pool of temporal information for the aforementioned Ecuador Grocery chain. This data is already partitioned into Train and Test data sets with examples given in their provided .zip file.