Time Series Chains Forecasting using MLP

Frida Cantu, Sergio Valdez, Yuliana Jasso , Advisor: Dr. Li Zhang
Department of Computer Science, College of Engineering and Computer Science
The University of Texas Rio Grande Valley

INTRODUCTION

Time Series - a type of data that can be measured over time

Time Series Chains - an ordered set of subsequence patterns where each are similar to its previous but the first and last subsequences may be different by surprise.

Multilayer Perceptron (MLP) - A deep learning model that can do various deep learning tasks. In our case, we use MLP to learn the relationships between each subsequence in a chain in hopes of predicting the next subsequence.

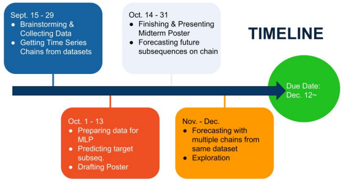
Motivation - Time series chains can tell yield important and interesting information. Our goal is to find a dataset with meaningful time series chains and predict future subsequences in a chain.

GOALS

- Finding meaningful chains in data.
- Predict future chains
- **Current goal**: having MLP successfully predict next subsequences in a known chain

PROJECT SUMMARY

Challenges/Struggles we faced - Time management, working with data with few points, translating data outputted from time series chain algorithm to MLP



METHODS

Time Series Chains

- Time series chains are created by identifying nearest neighbors for each subsequence on both sides.
- These nearest neighbor subsequences are stored in two data structures: **Left Matrix Profile** and **Right Matrix Profile**.
- Our primary emphasis lies in tracking the starting indexes of each subsequence within a chain.

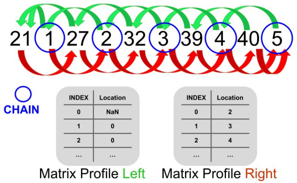
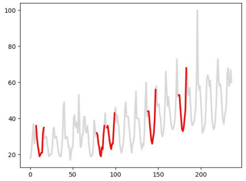


Figure shows toy example of how almost each number has a left and right nearest neighbor. Blue circle highlights which numbers are part of one chain.*

- For a subsequence to be considered part of a chain with length, $k > 1$, the left nearest neighbor of the right nearest neighbor of our current subsequence, $T(Ci,m)$, must be equal to $T(Ci,m)$.

$$TSC = \{T_{C1,m}, T_{C2,m}, \dots, T_{Ck,m}\}$$

$$LNN\big(RNN(T_{Ci,m})\big) = T_{Ci,m}$$



Subsequences from largest chain from one of the datasets we collected.

METHODS

Multilayer Perceptron

The general structure of an MLP can be described by the following equation:

$$\mathbf{f(x) = (T^{(L+1)} \circ \sigma \circ T^{(L)} \circ \cdots \circ \sigma \circ T^{(1)}) (x)}$$

The variable **T** describes the layers of the MLP. The structure of each layer of the MLP can be described with the following equation:

$$\mathbf{T^{(k)} (x) = W^{(k)}x+ b^{(k)}}$$

W represents a weight matrix and **b** represents a bias vector.

σ represents the ReLu activation function. The ReLu activation function can be described by the following equation:

$$\mathbf{\sigma(x) = max(0,x)}$$

The ReLu function will be applied to each element of the weight matrix **W**.

EXPERIMENT SETUP

Gathering Data - For our senior project, we gathered approximately 30 datasets from google trends, a website by Google that analyzes the popularity of top search queries in Google Search across various regions and languages. From those datasets, we chose the dataset, Bees All Time (trends ranging from 2004 to now).

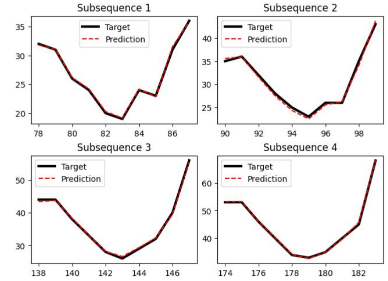
Getting Chains from Data - We utilized MATLAB to see if there were time series chains in the Bees All Time dataset. After we concluded that a time series chain exists, we obtained the time series chain starting indexes and sent them to be processed.

Data Processing - To process our data, we took our time series chain starting indexes and generated each subsequence. Then we separated the subsequences into a train subset and a target subset to prepare the data for the MLP model. In addition, we created a plot with each chain subsequence as well as a subplot, in which we normalized each chain, plotting the normalized chain on each subplot.

MLP and Results - After data processing, the train and target subsets becomes the input for the MLP model to predict the next subsequence.

RESULTS

After training the MLP with the processed data, we were able to successfully predict the next subsequences given the previous subsequence. After 3000 iterations we were able to achieve a MSE of 0.156.



NEXT STEPS

Our next steps include predicting future subsequences and finding meaning behind the time series chains. We would also like to work with larger data to have more subsequences to work with.

REFERENCES

Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., ... & Courville, A. (2019, May). On the spectral bias of neural networks. In *International Conference on Machine Learning* (pp. 5301-5310). PMLR.

Zhu, Y., Imamura, M., Nikovski, D., & Keogh, E.J. (2019). Introducing time series chains: a new primitive for time series data mining. *Knowledge and Information Systems, 60*, 1135-1161.