Final Project Report

Timothy Sue (und2yw@virginia.edu)
December 13, 2023

1 OVERVIEW

As a part of the Time-Series Machine Learning team this fall, the primary responsibility was to complete and polish the existing paper, as well as submit it to the AAAI 2023 conference. Team members Md Khairul Islam, Kingsley Kim, Luke Benham, Zhengguang Wang, Ayush Karmacharya, and I (Timothy Sue) collaborated to finish and improve upon the existing draft of the Sensitivity Analysis Paper/project. Weekly paper readings were also held and attended, where existing related works were explored and discussed in detail to familiarize ourselves with various time-series models. Additionally, preparations for the financial aid project data set were taken in order to have maximum efficiency when the data becomes available.

2 TIMELINE MILESTONES

The first task that I was given was to familiarize myself with the basics of Pytorch and setting up local environments with GPU capabilities. Khairul assisted us by providing resources to help learn the basics of Pytorch as well as interpreting time-series models (GitHub repositories and related projects/works). This was useful and allowed us to have time to learn the functionality of a new interface which was very helpful. The existing website for the GPCE-COVID project was also merged and the interactive plot was integrated into the website that Team 4 developed in the summer. In addition to this, access to Rivanna HPC was given to us, as well as a hands-on tutorial which assisted in learning the basics of UVA's HPC. This helped me get a better understanding on how to use and utilize slurm scripts as well as the other features of Rivanna and the capabilities of it.

After having done these on boarding steps, the next step was to foray into reproducing the existing Sensitivity Analysis project as well as understanding the code/techniques used. Throughout this process, it became clear that the project focused on applying Sensitivity Analysis to COVID-19 Age Groups to determine the influence that age has on COVID-19 infections. Understanding this allowed us to have deeper knowledge on the results of the experiments, and the impacts that data had on the model's training and interpretation results. Reproducing the experiment locally was also very useful, as it consisted of tasks like training models using the data available on the Git repository (https://github.com/UVA-MLSys/COVID-19-age-groups).

The next step was to add more transformer-based time-series models, as well as a linear model (DLinear), in order to have a wider variety of models to be trained/experimented with. Khairul instructed each of us to handle a different model, and to train/test the model with multiple data sets (Total.csv, Top500.csv, etc) in order to prepare the models for interpretation. The model that I was in charge of was the normal Transformer model, which I used Rivanna slurm scripts to train/test the model using all the data sets available.

I also trained and recorded test results for a couple of other models, such as Autoformer, FEDformer and TimesNet. In addition to training/testing these models, there were also two interpretation files ran, one which interpreted the results with the ground truth using Morris Sensitivity, and another that interpreted the results without the ground truth. Both files were ran through slurm scripts and used the FEDformer model. After the model was trained and interpreted, the results were reported both in the Git repository's ReadMe.md as well as an excel sheet that contained data for all models. In the tables below, it shows the interpretation results that were recorded as well as the recorded training/testing results.

Method	**Comp.** (↑)		**Suff.** (↓)	
	MAE	MSE	MAE	MSE
Feature Ablation	4.91	8.64	9.53	10.5
Feature Permutation	4.00	7.08	8.00	8.28
Morris Sensitivity	**6.23**	9.39	**5.85**	**5.46**
Feature Occlusion	4.89	8.44	9.49	10.4
Augmented F.O.	4.18	7.66	7.96	8.09
Deep Lift	5.72	**9.54**	8.90	9.43
Integrated Gradients	5.52	9.09	9.25	10.2
Gradient Shape	4.78	8.17	8.04	8.27

Table 2.1: AOPCR results of the interpretation using the FEDformer model (without ground truth)

Model	**MAE**	**RMSE**	**RMSLE**	**R^2-score**
Autoformer	35.69	189.4	1.918	0.451
FEDformer	**30.19**	**182.2**	**1.467**	**0.481**
PatchTST	31.17	183.6	1.530	0.469
TimesNet	34.35	191.9	1.604	0.415
Transformer	34.818	192.720	1.601	0.410

Table 2.2: Test/Training performance of the deep learning models

After having completed the main interpretation results along with training the models, the paper was further written and developed to include the new features and interpretation results. Each of the team members went through the current draft of the paper, and commented on things that we should elaborate/add on. Furthermore, we were also given previous reviewer comments for an earlier draft of the paper submission (ICDH Conference). This allowed each of us to go through the paper and find if we have addressed the comments from past reviews, and if we hadn't, we were able to improve upon it; giving us the best chances for this submission.

Throughout this independent study, we also held weekly meetings where an individual would present a related time-series model/transformer paper. I presented the paper, "A Time Series is Worth 64 Words: Long-term Forecasting with Transformers" (from the ICLR 2023 conference). This paper was about a transformer-based time series forecasting model, called PatchTST. In this paper, they proposed that this new model, is able to outperform other state of the art transformer models in long-term forecasting accuracy. The model was able to do this through features such as channel-independence and patching (segmentation of time series into subseries level patches).

In the final portion of our project, many of our team members began to start researching how to improve the web page for the Sensitivity Analysis Project. This was brought to our attention by Khairul, who recommended that we start to research and create a new website to be more professional/formal. He mentioned that the GPCE-COVID website is a good reference, and also recommended we utilize JupyterBook. Moreover, each of our team members also made sure we were all able to reproduce the lab results discussed in our paper, and finished tasks in our Git to-do list relating to completing the web page and docs.

2.1 Sensitivity Analysis Project

As mentioned before, to get a basic understanding of the Sensitivity Analysis Project, we were given a week or so for on boarding, where we learned Pytorch basics, and were able to reproduce the project locally. These two things were particularly helpful in understanding the syntax of the project along with the library functions that it used. Our progress was relatively easy to keep track of, as we utilized the Github project to-do list which helped keep tasks organized. The Github repository's ReadMe.md also includes a detailed tutorial on how to reproduce the experiment, either using CPU or GPU capabilities. This allowed each of us to reproduce the experiment locally, to ensure that the results we calculated are accurate and correct. Below is a screenshot of our GitHub management and contributions/commits.

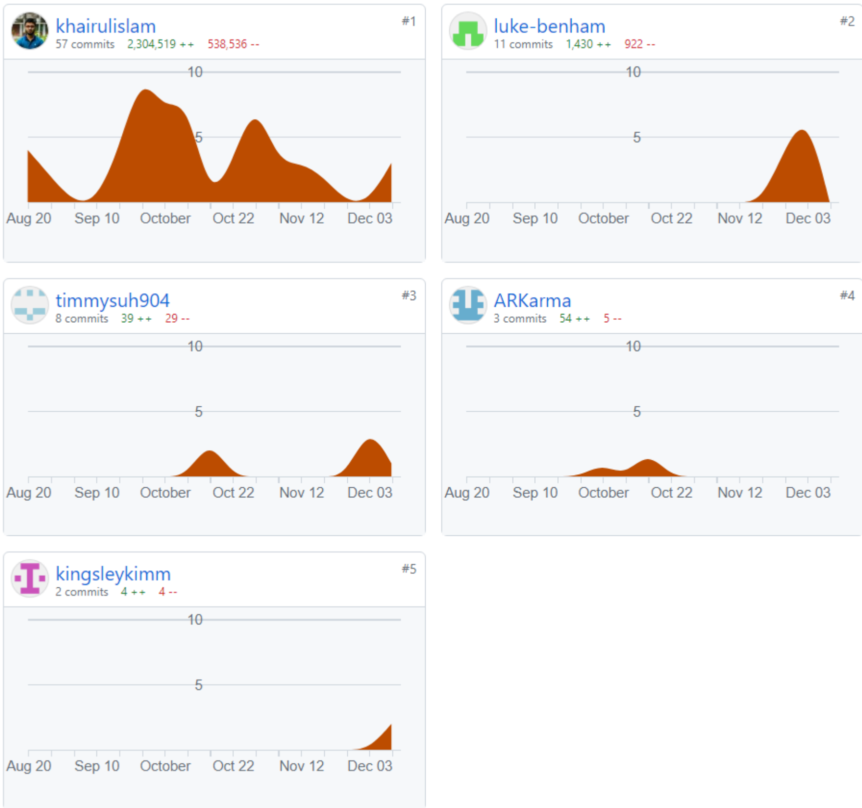

Figure 2.1: Our Teams Git Commit Graph

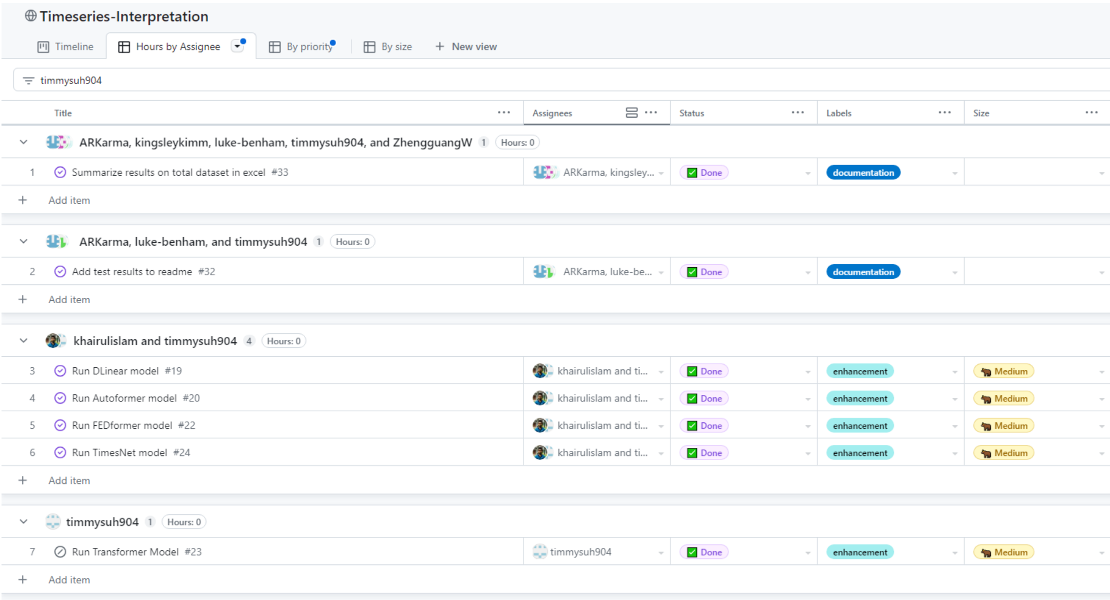

Figure 2.2: Individual completed tasks

After reproducing the results locally, our next task was to assist in writing the paper submission. The submission was mostly written by Khairul, as he had the best understanding of the material and experience in submitting projects to conferences. He did ask us to review the things that he had wrote, and comment on things that we could improve/change. I was mostly focused on the interpretation part of the submission, as that is what I mainly worked with and documented. Once our paper was roughly complete, we researched into integrating the web page with JupyterBook, and also on polishing the paper/writing our final reports.

2.2 Financial Aid Project

The Financial Aid Project, is a new project that relates to the use of AI/ML in relation to financial aid distribution. The data set is said to have sensitive information about students, so it was given to us on the Ivy Secure Environment. In order to access the data safely and securely, certain steps and regulations had to be taken, such as installing a high security VPN. Some team members and I have also sat in on some meetings relating to the Financial Aid project, where we got to introduce ourselves to the individuals in charge of the data and it's distribution.

Currently, since we have very recently gained access to the data set on Ivy, Ayush and I am currently going through the data and verifying that we were given the correct features. After this, our future plan is to work on the creating a model for this data set, as well as training/testing and interpreting the results.

2.3 Future Plans and Research

Currently, the Sensitivity Analysis project paper submission is complete and we are now polishing the project's Github web page along with preparing for the Financial Aid project. The paper should be submitted this week, as we are going to use the remaining time left to make improvements on clarity and other minor things. The web page for the Sensitivity Analysis project is also going well, as team members Zhengguang, Luke and Kingsley have prior experience working with JupyterBook. Once the work for the Sensitivity Analysis project is complete, the next task will be to focus on the Financial Aid data set. One thing that we will have to consider is what types of models can be used with the data set and how well we can fit it with our model. But, this should be easier once we can access the data and analyze it.