IE 360 Term Project

June 1^{th}, 2021

1 Description

How many products will Trendyol sell each day? Forecasting the sales quantity of a product for an online retail is a challenging task. Forecasting heavily relies on the historical data and your modeling skills. In this project, you are expected to develop forecasting methods to predict the next day's sales quantity of the products in Table 1. Many decisions such as pricing, inventory placement and etc. require reliable forecasts and inaccurate business forecasts could result in actual or opportunity losses.

The data, covers daily sales of nine products of Trendyol and includes product, category level, brand and site level details. These are provided as explanatory variables but note that day of the week, special events and other can be important for forecasting. Site level data is provided for that purpose (i.e. you can identify Black Friday dates from the site traffic). Together, this robust dataset can be used to improve forecasting accuracy. You are also free to use any external data if you think it is useful (i.e. Google Trends).

Product Id	**Top Hier.**	**Bottom Hier.**	**Brand**
48740784	Dış Giyim	Mont	ALTINYILDIZ CLASSICS
73318567	Plaj Giyim	Bikini Üstü	TRENDYOLMİLLA
32737302	Plaj Giyim	Bikini Üstü	TRENDYOLMİLLA
31515569	Alt Giyim	Tayt	TRENDYOLMİLLA
6676673	Elektronik Aksesuarlar	Bluetooth Kulaklık	Xiaomi
7061886	Elektrikli Ev Aletleri	Dik Süpürge	Fakir
85004	Cilt Bakım	Yüz Temizleyici	La Roche Posay
4066298	Anne ve Bebek Bakım	Bebek Islak Mendil	Sleepy
32939029	Ağız Bakım	Şarj Edebilir Diş Fırçası	Oral-B

Table 1: Product characteristics

This project is organized as live competition in which you are expected to make submission everyday. For this purpose, we have built an application programming interface (API) so that you can programmatically get the data, manipulate it, build your forecast models and make submissions. The aim is to make you develop a forecasting product. Developed product can be used in various business areas, such as setting up appropriate inventory or service levels. Accurate forecasts help businesses in reducing waste and understanding the uncertainty and its risk implications.

Acknowledgements Additional thanks to Trendyol for providing the data for the project.

2 Evaluation

You are expected to provide your next day's (time $t+1$) predictions regularly before 13:00 every day (time t). You are allowed to make more than one submission, however we will only evaluate your latest submission before 13:00. Also, you need to make predictions for all of the products. System will not allow otherwise.

Each day, we will sort each group by the weighted mean absolute percentage error (WMAPE) of their predictions. Namely,

$$WMAPE_{day} = \frac{\sum_{p \in Products} |y_p - \hat{y_p}|}{\sum_{p \in Products} y_p}$$

where y_p is the real sales in the given day, and $\hat{y_p}$ is the prediction that you submitted. All of the groups will be sorted by the WMAPE of their latest submission. Then, according to your daily ranking, you will get your ranking points. For example, if there are 24 groups, the $1st$ group will get 24 points, $2nd$ will get 23... and so on. If you have not made any submissions, you will get 0 ranking points. Please, try to make the most of this experience!

You will have 300 requests/day as your total requests limit you can send to the system. After this point, you will be blocked for 24 hours. In this case, please contact us at mert.yuksekgonul@boun.edu.tr

3 Timeline

Beginning from **June 2, 2021**, you will be given your passwords and usernames for access to the system. This week, it will be the trial period. It will be the test period for the system. Take the trial phase to ask your questions, play around with the system and build your prediction pipeline. It will also be a chance for us to fix any unforeseen issues. We will proceed with the evaluations as usual, however we will restart all of the scores on **June 7, 2021**. Then, the actual project phase will begin. The submission system will be closed after **June 21, 2021**.

4 Deliverables

You will be presenting your work towards the end of the semester (to be announced later).

You are also expected to report your approach and findings as a document. Your report should have the following format:

1. Introduction: Problem description, summary of the proposed approach, descriptive analysis of the given data.
2. Related literature: Summarize relevant literature if there is any
3. Approach: Explain your approach to this problem.
4. Results: Provide your results and discussion.
5. Conclusions and Future Work: Summarize your findings and comments regarding your approach. What are possible extensions to have a better approach?
6. Code: Provide the Github link for your codes at the end of your report.

5 Grading

A group of at most three students will work on the project. Details of project grading is provided on Moodle. A slightly modified version of the grading template used for IE 492 (Graduation Project) course is going to be used. Current documents are almost 90% similar to the final one. The grades in the rubrics sum up to 90%. The remaining 10% will be determined from your group's performance in the competition.