**M5 Forecasting - Accuracy**

Description:

Note: This is one of the two complementary competitions that together comprise the M5 forecasting challenge. Can you estimate, as precisely as possible, the point forecasts of the unit sales of various products sold in the USA by Walmart? If you are interested in estimating the uncertainty distribution of the realized values of the same series, be sure to check out its [*companion competition*](https://www.kaggle.com/c/m5-forecasting-uncertainty)

How much camping gear will one store sell each month in a year? To the uninitiated, calculating sales at this level may seem as difficult as predicting the weather. Both types of forecasting rely on science and historical data. While a wrong weather forecast may result in you carrying around an umbrella on a sunny day, inaccurate business forecasts could result in actual or opportunity losses. In this competition, in addition to traditional forecasting methods you’re also challenged to use machine learning to improve forecast accuracy.

The Makridakis Open Forecasting Center (MOFC) at the University of Nicosia conducts cutting-edge forecasting research and provides business forecast training. It helps companies achieve accurate predictions, estimate the levels of uncertainty, avoiding costly mistakes, and apply best forecasting practices. The MOFC is well known for its Makridakis Competitions, the first of which ran in the 1980s.

In this competition, the fifth iteration, you will use hierarchical sales data from Walmart, the world’s largest company by revenue, to forecast daily sales for the next 28 days. The data, covers stores in three US States (California, Texas, and Wisconsin) and includes item level, department, product categories, and store details. In addition, it has explanatory variables such as price, promotions, day of the week, and special events. Together, this robust dataset can be used to improve forecasting accuracy.

If successful, your work will continue to advance the theory and practice of forecasting. The methods used can be applied in various business areas, such as setting up appropriate inventory or service levels. Through its business support and training, the MOFC will help distribute the tools and knowledge so others can achieve more accurate and better calibrated forecasts, reduce waste and be able to appreciate uncertainty and its risk implications.

**Acknowledgements**  
Additional thanks go to other partner organizations and prize sponsors, National Technical University of Athens (NTUA), INSEAD, Google, Uber and IIF.

Evaluation:

This competition uses a **Weighted Root Mean Squared Scaled Error** (RMSSE). Extensive details about the metric, scaling, and weighting can be found in the [M5 Participants Guide](https://mofc.unic.ac.cy/m5-competition/).

## **Submission File**

Each row contains an id that is a concatenation of an item\_id and a store\_id, which is either validation (corresponding to the Public leaderboard), or evaluation (corresponding to the Private leaderboard). You are predicting 28 forecast days (F1-F28) of items sold for each row. For the validation rows, this corresponds to d\_1914 - d\_1941, and for the evaluation rows, this corresponds to d\_1942 - d\_1969. (Note: a month before the competition close, the ground truth for the validation rows will be provided.)

The files must have a header and should look like the following:

id,F1,...F28

HOBBIES\_1\_001\_CA\_1\_validation,0,...,2

HOBBIES\_1\_002\_CA\_1\_validation,2,...,11

...

HOBBIES\_1\_001\_CA\_1\_evaluation,3,...,7

HOBBIES\_1\_002\_CA\_1\_evaluation,1,...,4

Timeline :

* June 1, 2020 -  Full Training Labels Released. Participants will be provided with the actual values of the 28 days of data used for scoring performance (Public test set). The private test set will remain unchanged, and will still be used to determine the winners.
* June 23, 2020 - Entry deadline. You must accept the competition rules before this date in order to compete.
* June 23, 2020 - Team Merger deadline. This is the last day participants may join or merge teams.
* June 30, 2020 - Final submission deadline.

All deadlines are at 11:59 PM UTC on the corresponding day unless otherwise noted. The competition organizers reserve the right to update the contest timeline if they deem it necessary.

Prizes:

* 1st Place - $25,000
* 2nd Place - $10,000
* 3rd Place - $5,000
* 4th Place - $3,000
* 5th Place - $2,000

An additional $5,000 will be granted to the highest performing student team on the leaderboard at the end of the competition. A student team is one for which at least 50% of team members are current full-time students. If interested append \_STU to your team name.

Prizes will be distributed during the M5 Conference in December 2020, held in New York City, NY, USA.