Universität St.Gallen
Institut für Wirtschaftsinformatik

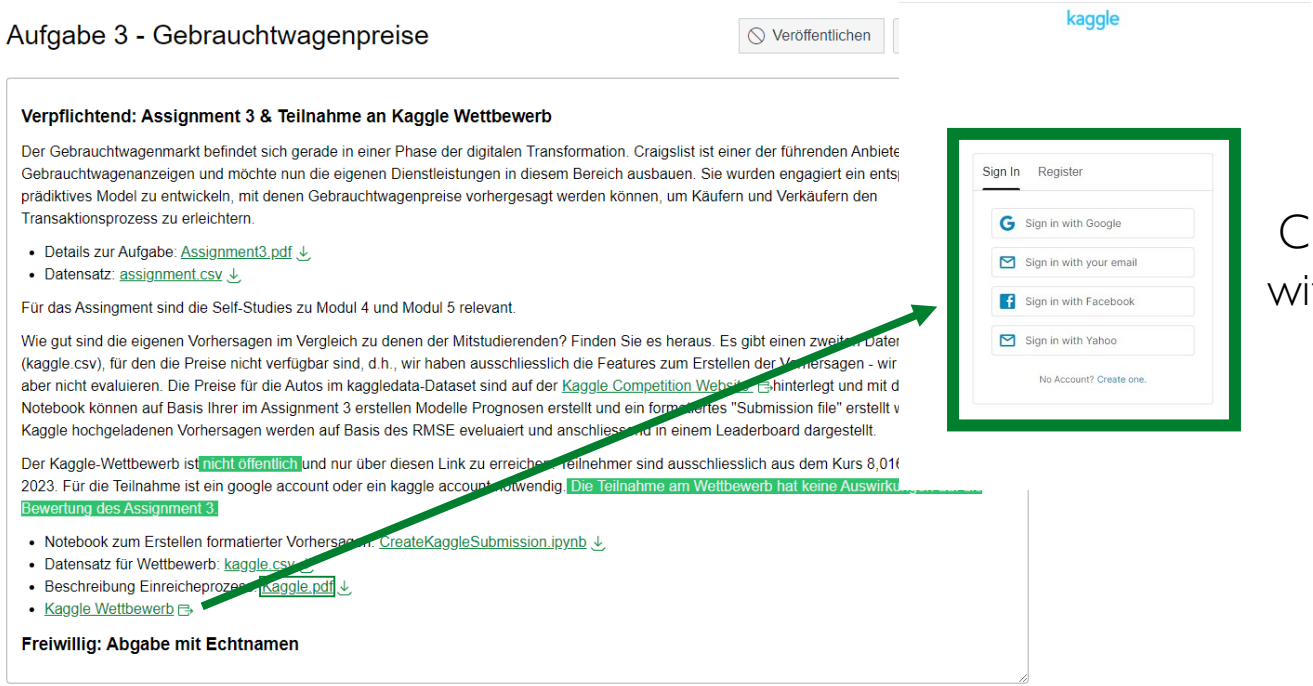
Submitting on Kaggle Competitions

Prof. Dr. Ivo Blohm (ivo.blohm@unisg.ch)

Associate Professor for Information Systems and Business Analytics

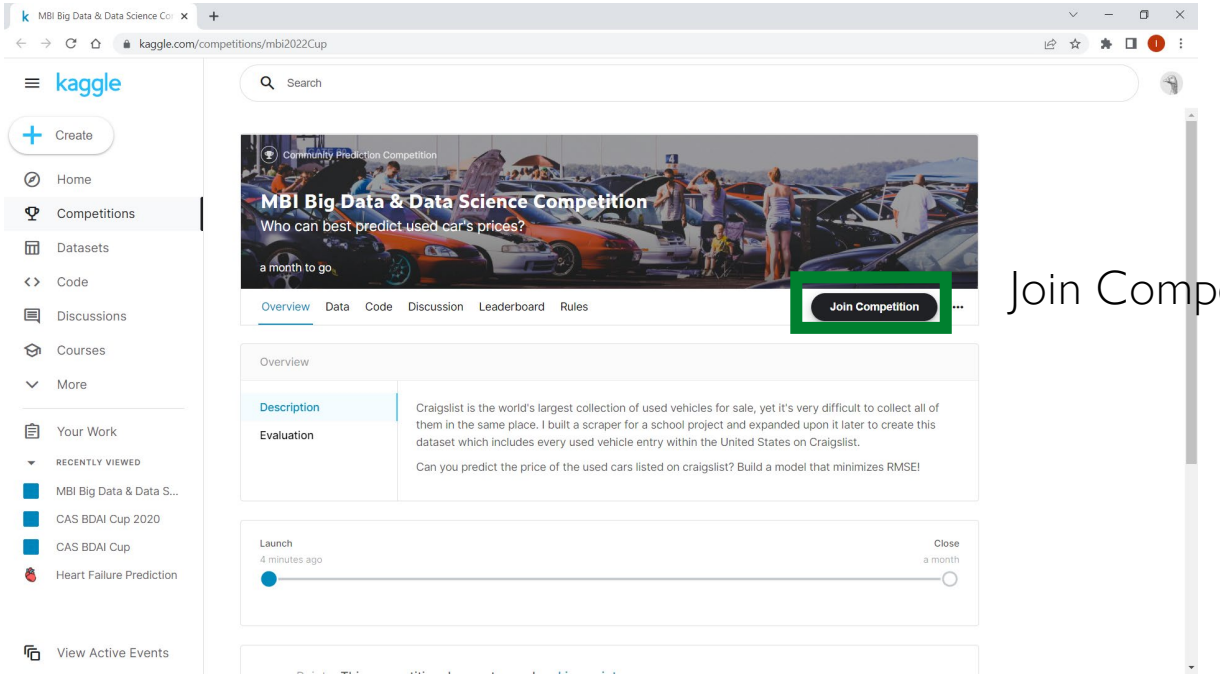
From insight to impact.

Competition Website



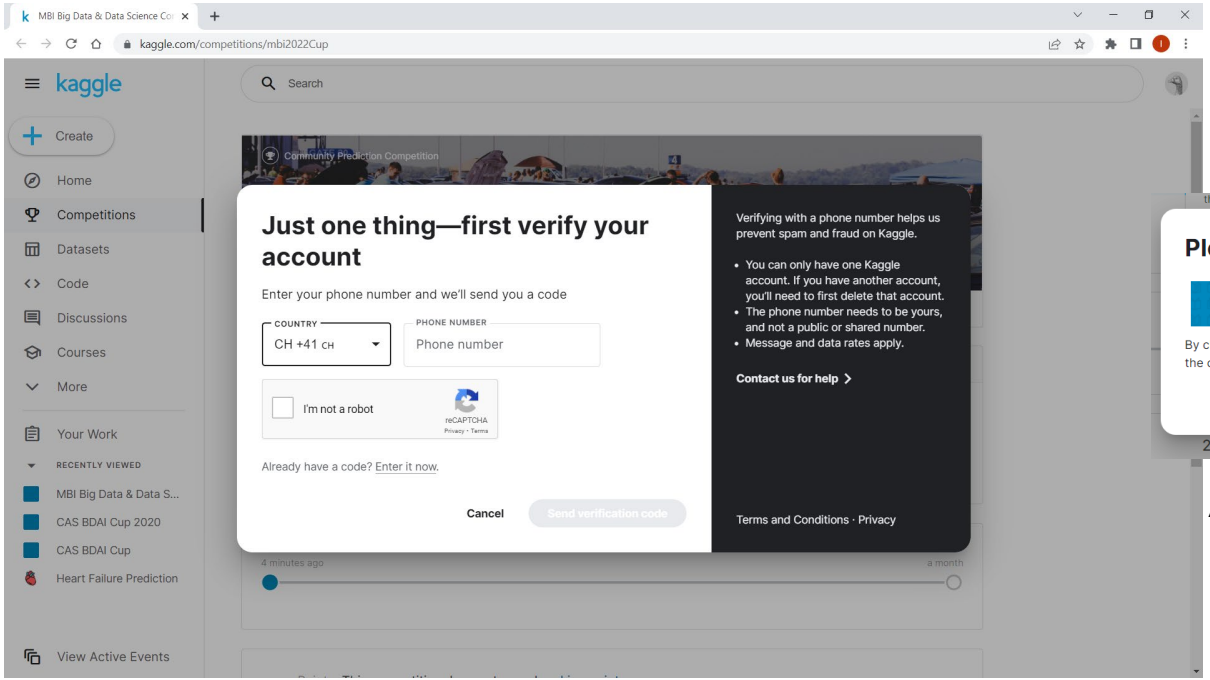
Create Login / Login with Google Account Or Email

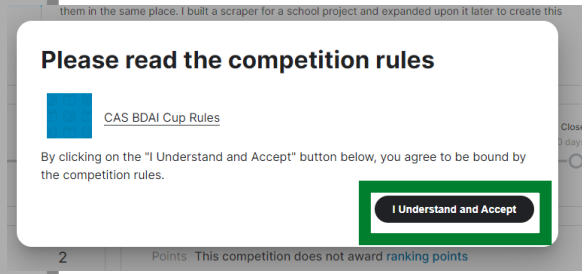
Competition Website



Join Competition

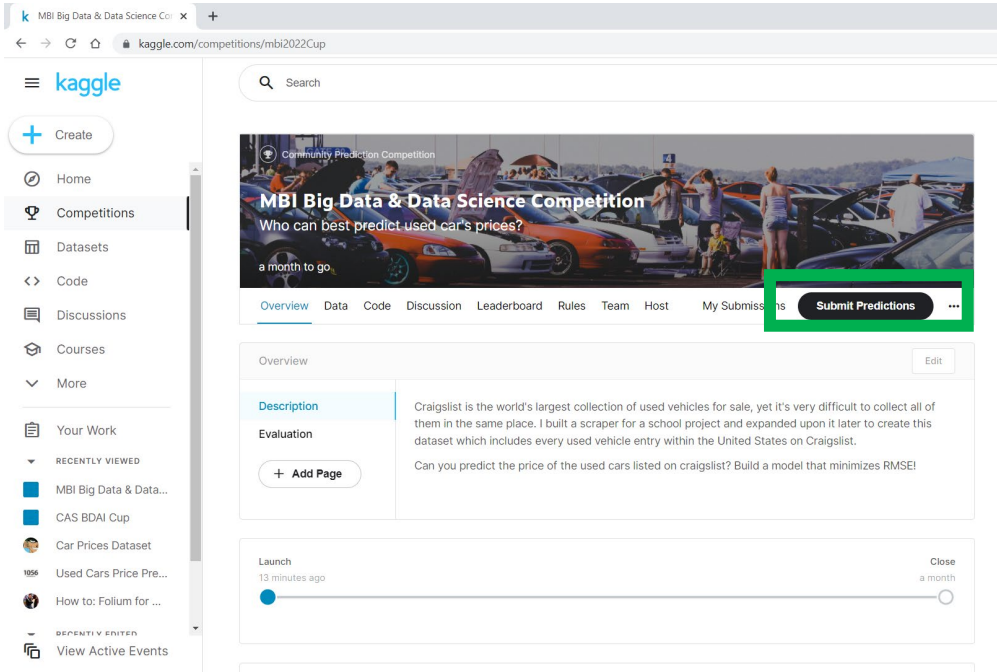
Verifiy your account





Accept Competition Rules

Make Submissions



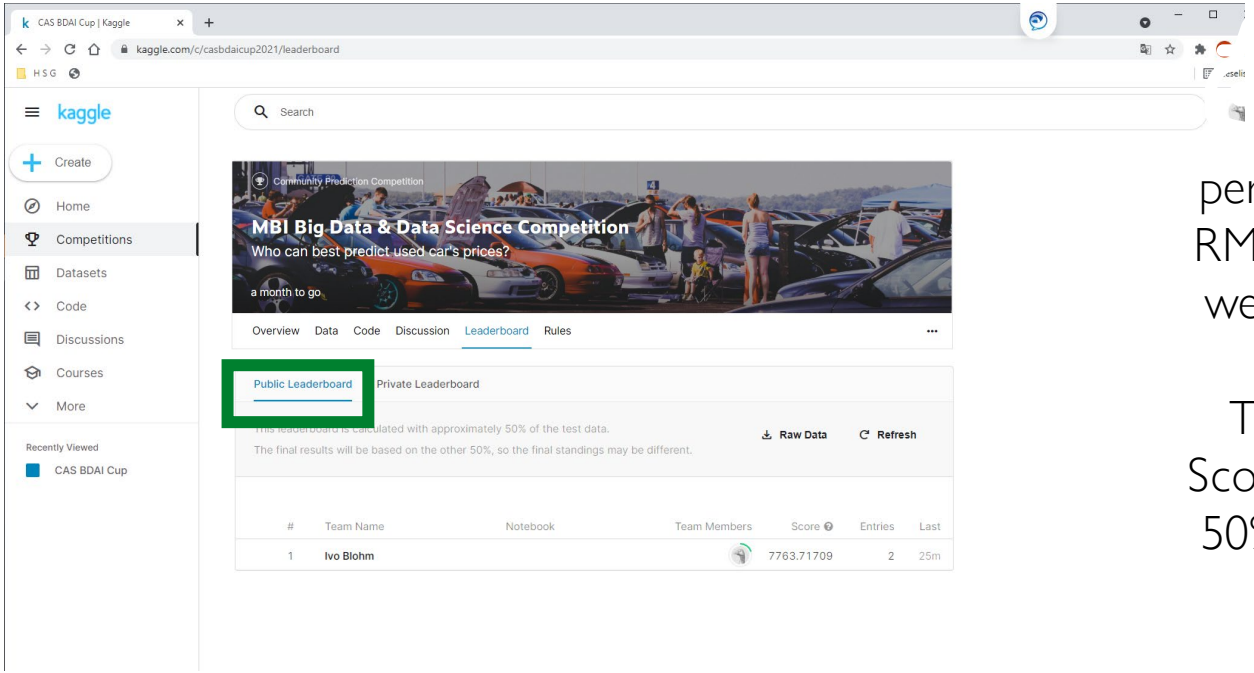
Take the CAS BDAI CUP Notebook

Develop your Model in Google Collab / Jupyter Notebook

Create Solution File and submit it here

10 submissions per day are allowed

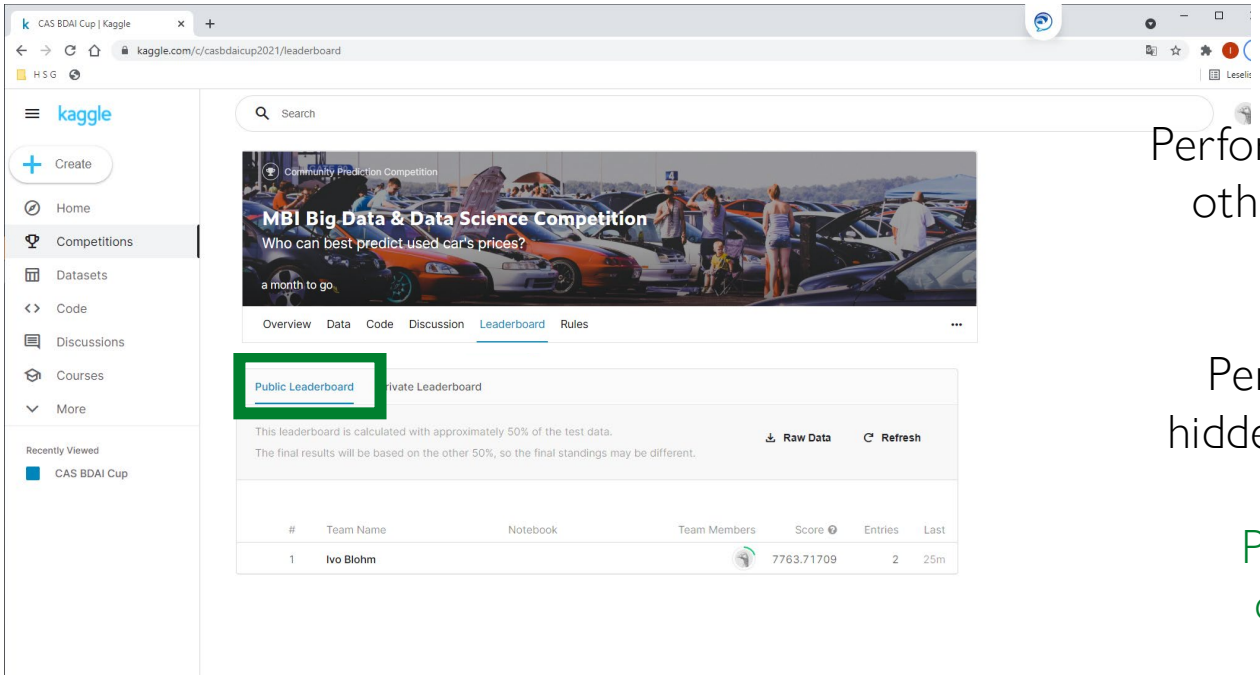
Compare your Performance – Public Leaderboard



After submitting a contribution, the performance get scored by RMSE and you can see how well your solution is doing

The Public Leaderboard Score is calculated based on 50% of the «ground truth» data

Compare your Performance – Private Leaderboard



Performance is scored on the other half of the «ground truth» data

Performance scores are hidden (also for submitters)

Private Leaderboard determines winner

Why so complicated? Target Leakage!

Presence of unexpected additional information in the training data, allowing a model to make unrealistically good predictions.

- Features in historic data that would not be available at the time of the prediction
- Duplicated instances of test data in training data
- Predictions get part of test data

Model looks good in development, but performs poor in the real-world

If contestants know their performance on the „evaluation data“ they would tailor their models to this specific data set!

HAPPY HACKING