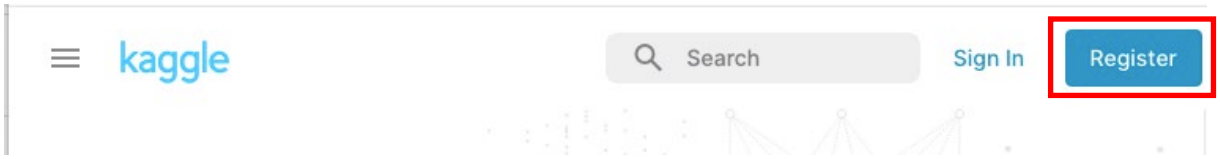
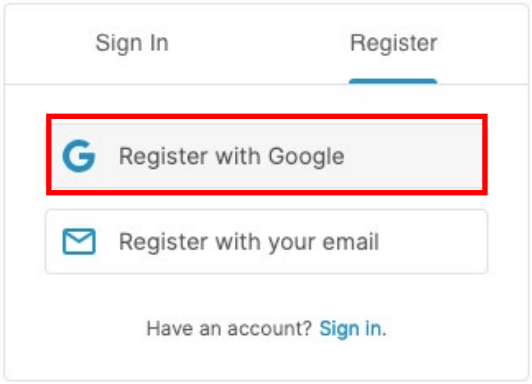
# Instructions of using Kaggle

## 1. Registration

The first step for using Kaggle is creating an account. To do so, you can access the Kaggle homepage and click on the register option at top right corner of the screen.




Please use the Register with Google option and use your student.unimelb.edu.au email address to make an account.

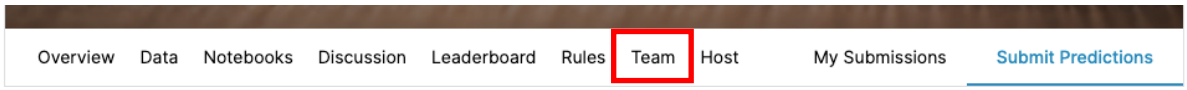



**PLEASE ONLY USE YOUR STUDENT ID AS YOUR TEAM NAME.**

## For group submissions please use BOTH Student IDs (e.g. 12345 & 12354)

> ***NOTE**: We will only consider submissions under the correct Student ID. All the other submissions are considered fake and will be ignored.*

If you made a mistake, you could update your TEAM NAME, going to "TEAM" tab → General → TEAM NAME.

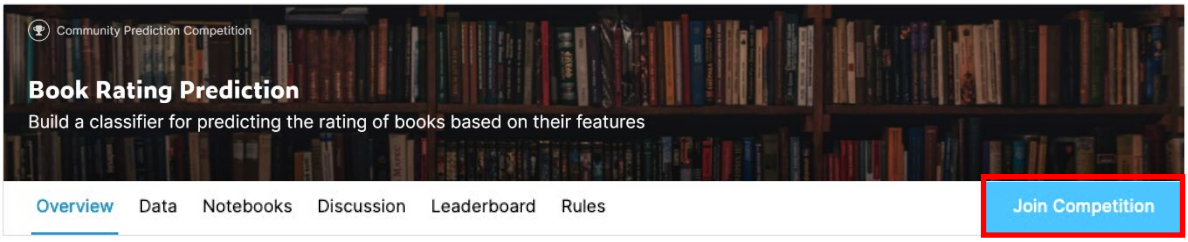


## 2. Competition

The COMP30027 Project 2 is a *private* competition so only people who have access to this link can participate.

*Link:* https://www.kaggle.com/t/a7661361a7174e4e87f08fb2403ae5c6

After accessing the competition page, you need to "Join" the competition by clicking on the option on the top-right corner and accepting the rules.

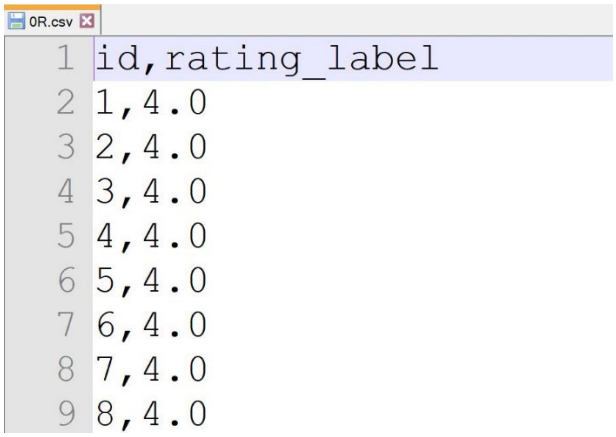



Your prediction file needs to be in *.csv format*.

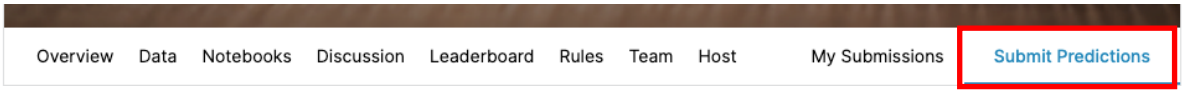
This .csv file should have exactly **two** columns.

- **First column** should be named: *id*. This column should include a sequence (series) from 1 to 5766 (indicating the sequence of the instances in the book_rating_test file).
- **Second column** should be named: *rating_label*. This column should include the predicted label (3.0, 4.0, or 5.0) for the book rating. These predictions are the output of your model for the instances in the book_rating_test file.

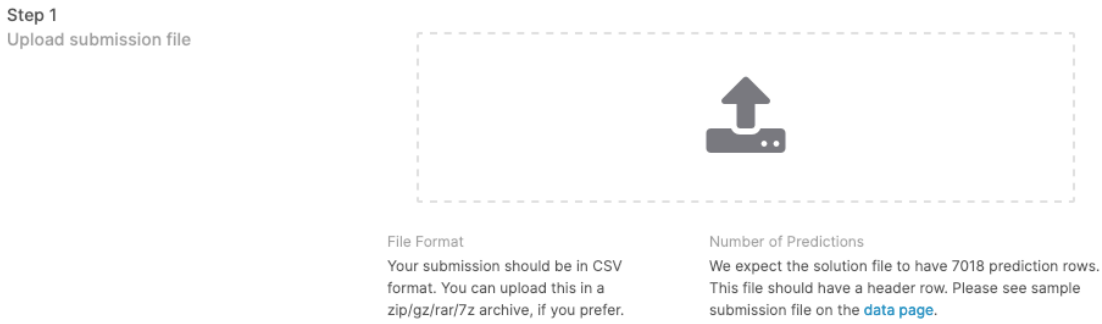
Your .csv file should have exactly 5767 rows. *First* row including the *header* row: {id, rating_label} and the rest of *5766 rows* should include the content.



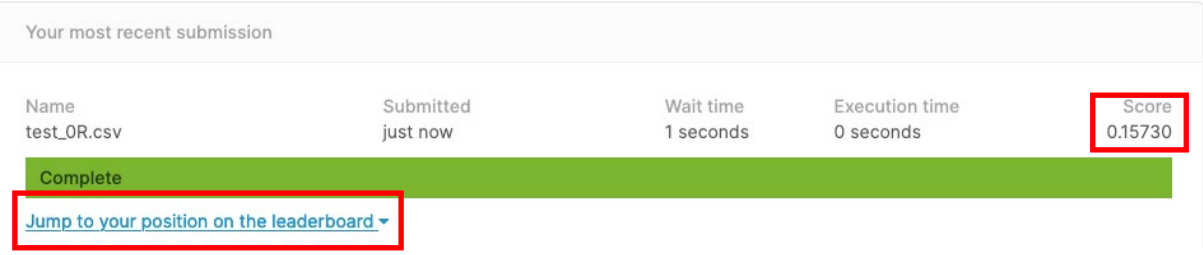

After that you would be able to "Submit Predictions" using the provided option.




If your prediction file has the correct format (2 columns, 5767 rows, *correct* header and *correct* id-s) it will be loaded in Kaggle *Leader Board* successfully.




After a successful submission, Kaggle will give you a score (the accuracy of your test data predictions using 50% of the data). And you can also find the ranking of your results using the *public* leaderboard. After competition closes, public 50% test scores will be replaced with the private leaderboard with 100% test data.




***NOTE**: We are checking your prediction accuracy results on 100% of the data using the private Leader Board.*

It is because we do not want you to try and improve your rank just by *overfitting* your results for the test data (using excessive try and error submissions on Kaggle).

You can only submit up to 8 predictions on each day. It is important to keep in mind that we are NOT marking the accuracy of your model, but we are assessing your ability and skills in developing and analysing of a logical argument about the given task, using different Machine Learning methods.

Prior to competition closes, you may select a final submission out of the ones submitted previously – by default the submission with highest public leader board score is selected by Kaggle.