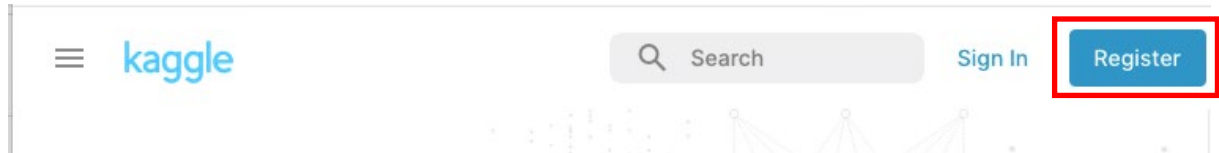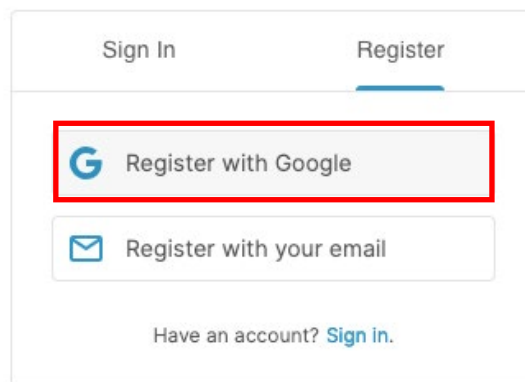# Instructions for Using Kaggle

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.

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

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

If you made a mistake, you could update your TEAM NAME, in your Kaggle profile.
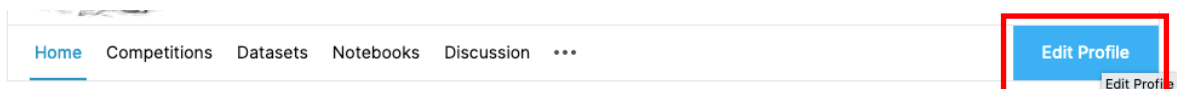
2. Competition

The COMP30027 2021 Assignment2 is a *private* competition so only people who have access to this link can participate.

Link: https://www.kaggle.com/t/179624c0610641f888093168f5b81ff5

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.
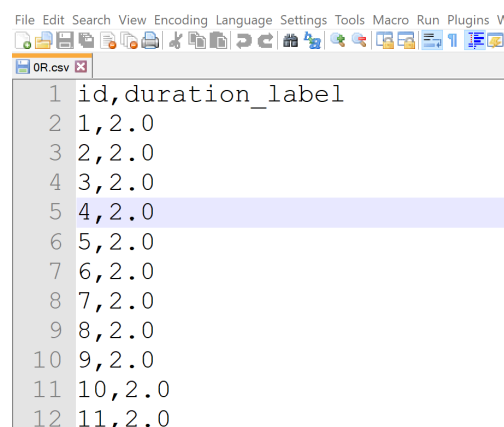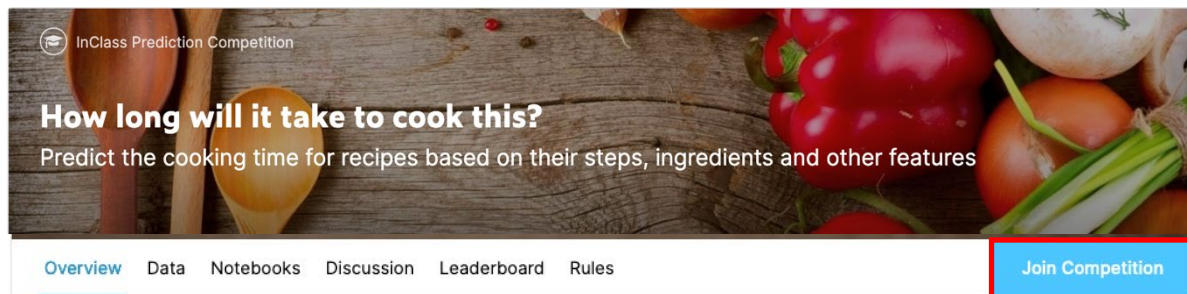

Your prediction file needs to be in *.csv format*.

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

- **First column** should be named: *id*. This column should include a sequence (series) from 1 to 10000 (indicating the sequence of the instances in the recipe_test datasets.
- **Second column** should be named: *duration_label*. This column should include the predicted label (1.0, 2.0, or 3.0) for the cooking duration. These predictions are the output of your model for the instances in the recipe_test datasets.

Your .csv file should exactly have **10001** rows. *First* row including the *header* row: {id, duration_label} and the rest of *10000 rows* should include the number of the instance and your rating prediction {1.0, 2.0 or 3.0}.

After that you would be able to "Submit Predictions" using the provided option.
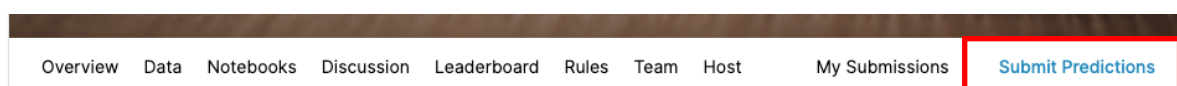
If your prediction file has the correct format (2 columns, 10001 rows, *correct* header and *correct* id-s) it will be loaded in Kaggle *Leader Board* successfully.

After a successful submission, Kaggle will give you a score (the accuracy of your test data predictions using 30% of the data). And you can also find the ranking of your results using the *public* leaderboard. After competition closes, public 30% test scores will be replaced with the private leaderboard with 100% test.

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.