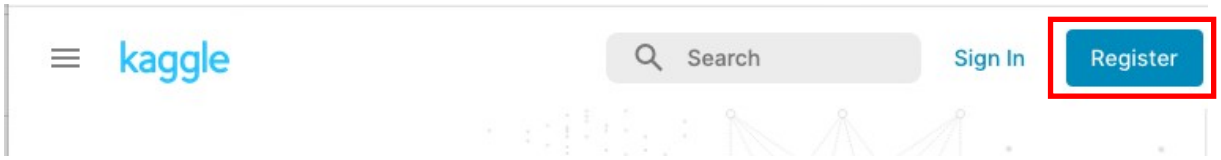
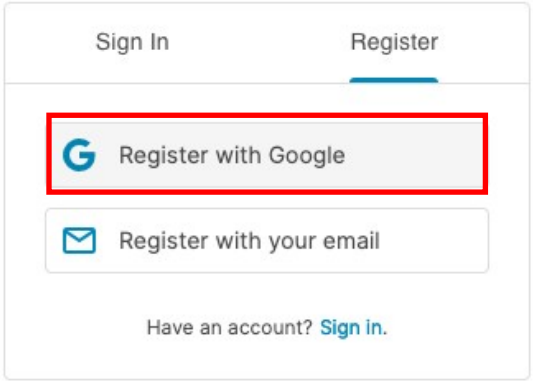
Instructions for Using Kaggle

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.

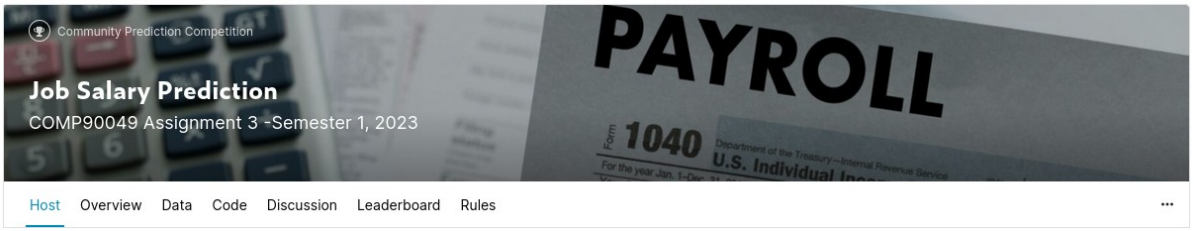


PLEASE ONLY USE YOUR STUDENT ID AS YOUR TEAM NAME.

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

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

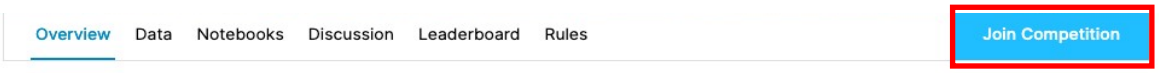
2. Competition



The COMP90049 2023 SM1 Assignment3 is a *private* competition so only people who have access to this link can participate.

https://www.kaggle.com/t/00fd833c73c1472eb3b257ad49f138dc

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.



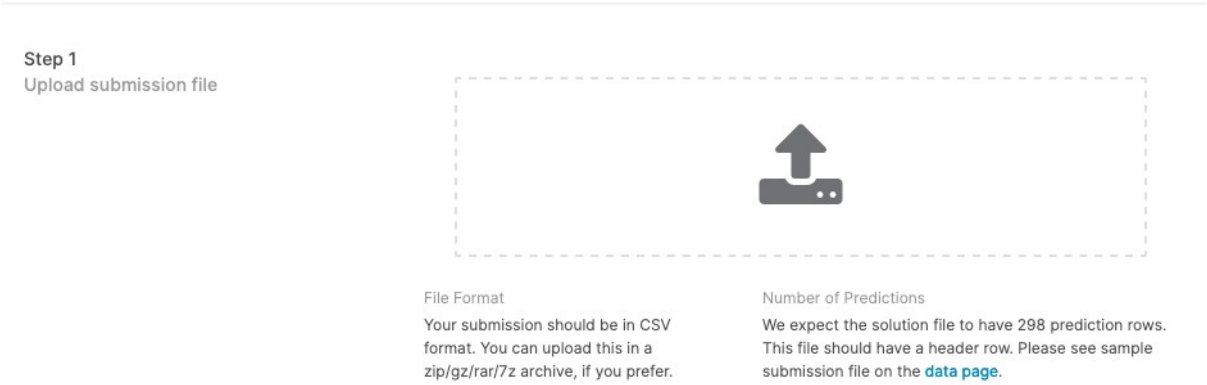
Your prediction file needs to be in **.csv format** (*.csv* stands for Comma Separated Values file)

Your CSV file should exactly **two** columns and **1739** rows. Your file should have a header row: {'job_id', 'salary_bin'} and 1738 prediction rows.

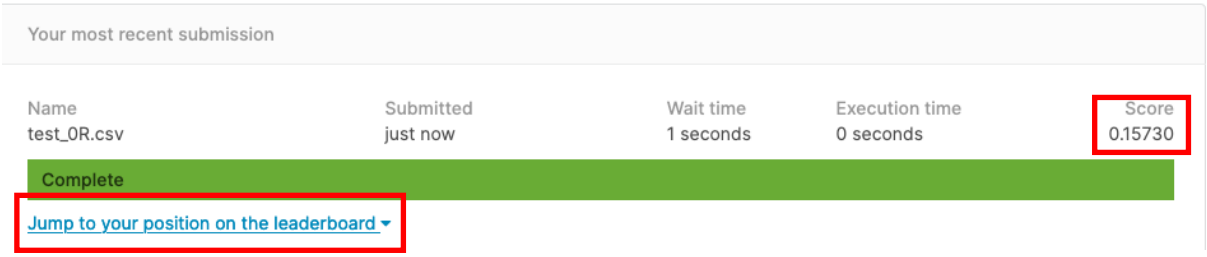
	A	B
1	job_id	salary_bin
2	JOB-2019-0016087	9
3	JOB-2019-0008055	9
4	JOB-2019-0001658	9
5	JOB-2019-0016777	9
6	JOB-2019-0017600	9
7	JOB-2019-0000279	9
8	JOB-2019-0010945	9
9	JOB-2019-0010096	9
10	JOB-2019-0017154	9
11	JOB-2019-0008605	9

Given this correct format, you will be able to "Submit Predictions" using the provided option.

If your prediction file has the correct format (2 columns, 1739 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* leader board.



***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, we are assessing your ability and skills in developing and analyse of a logical argument about the problem of tweet sentiment classification, using different Machine Learning methods.

Prior to competition close, 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.