

2023 Introduction to Artificial Intelligence

Machine Learning Project



Schedule



12/6
2023



Competition start

12/20
2023



Competition ends

12/22
2023



Report due (eeclclass)

12/22
2023



Peer evaluation



Outline

- Introduction
- Kaggle
- Report

Introduction

Data

Data contents

- # examples: 500,000
- # attributes: 35
- # class (label 0/1): 2

Data

Data provided

- Training data (train.csv)
 - ▣ Data with labels, for model training
- Testing data (test.csv)
 - ▣ Data without labels, you need to predict the label of each data point
 - ▣ **Each data point has an id number**
 - ▣ The testing data is divided into two parts:
 - ▣ One for the public leaderboard
 - ▣ And the other for the private leaderboard
- Sample submission
 - ▣ The prediction you submit on Kaggle should follow the same format as the sample_submission.csv

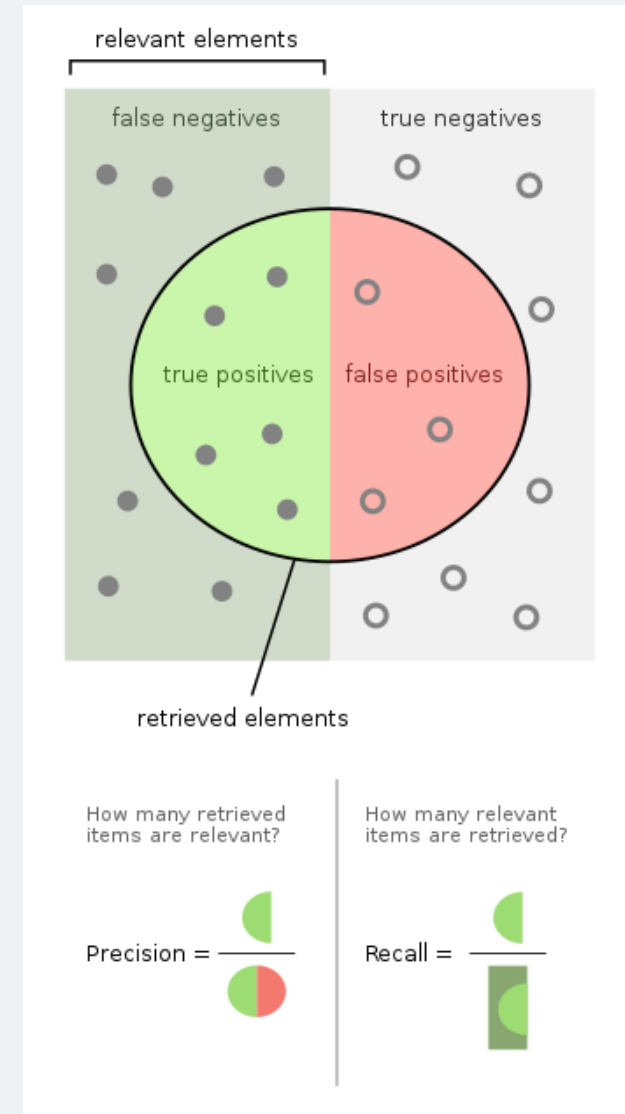
Performance indexes

F-score

- $Precision\ score = \frac{tp}{tp+fp}$
- $Recall\ score = \frac{tp}{tp+fn}$
- $F - score = 2 \cdot \frac{precision \cdot recall}{precision + recall}$

Macro F-score

- The metric used for Kaggle leaderboard
 - ▣ Can refer to the [code](#)





Kaggle

Join the Kaggle contest

Join through the Kaggle link

- You can get access to the contest if you've already filled out the team up [google form](#)
- The [Kaggle](#) link
 - ▣ Please contact TA 黃恩平 if you do not have access to the contest

Join the Kaggle contest

Accept the rules

- Click "I understand and Accept" to start competing on the "Rules" page

Rules

Read the following instructions carefully before entering this competition!

- You should finish your homework on your own.
- You should not modify your prediction files manually.
- Do not share codes or prediction files with any living creatures.
- Do NOT use any approaches to submit your results more than 5 times a day.
- Do NOT search or use additional data or pre-trained models.

Your grade on this hw will get deducted if you violate any of the above rules.

Be careful and wish you good luck!

For more detailed regulations of this course, please visit the course page.

If you have any questions, we encourage you to post the question on the discussion page.

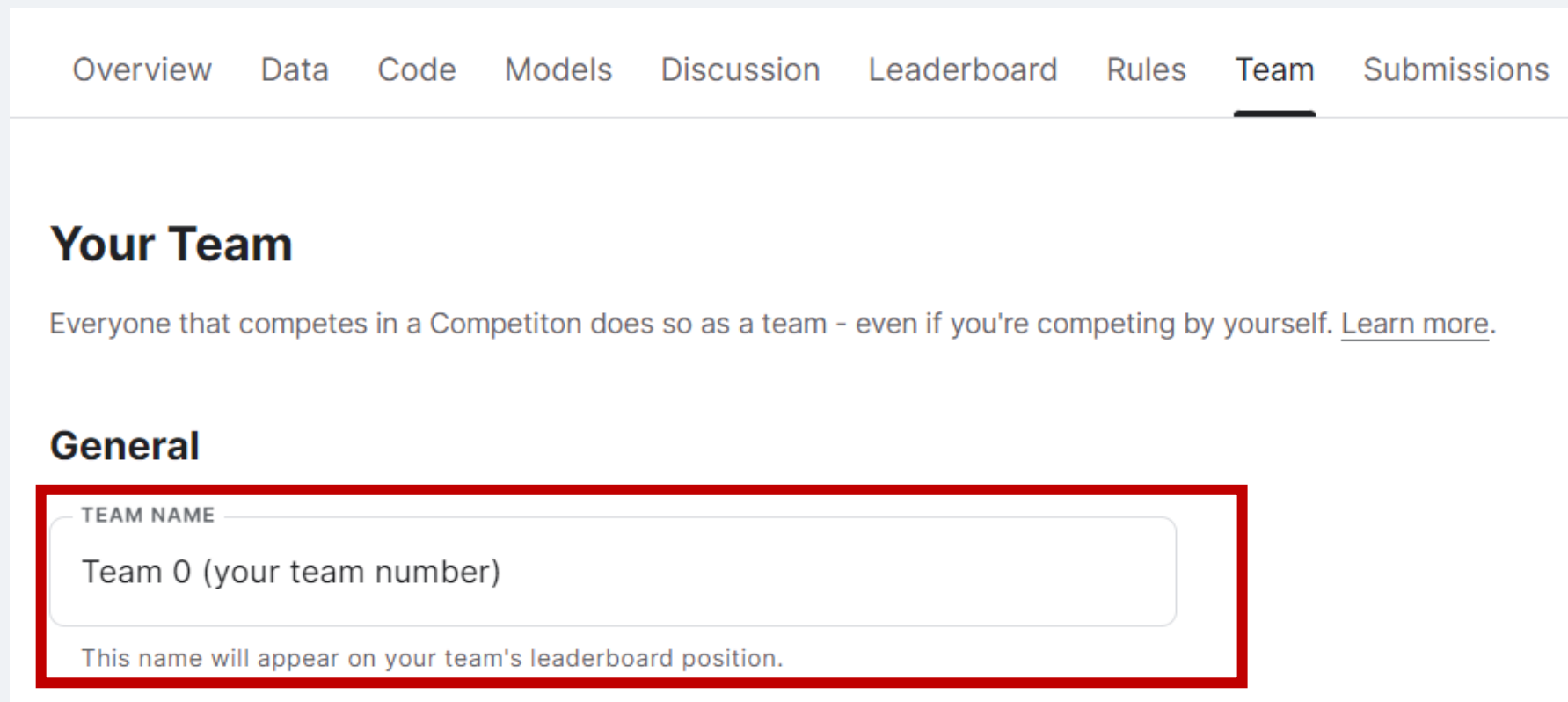
By clicking on "I Understand and Accept", you agree to be bound by the competition rules.

I Understand and Accept

Team up (Important!!)

Change your team's name

- See your team number on eeclass 小組專區



The screenshot shows the 'Team' page in eeclass. The navigation bar at the top includes links for Overview, Data, Code, Models, Discussion, Leaderboard, Rules, Team (which is underlined), and Submissions. The main content area is titled 'Your Team' and contains the text: 'Everyone that competes in a Competition does so as a team - even if you're competing by yourself. [Learn more.](#)'. Below this is a section titled 'General' which contains a form for changing the team name. The form has a label 'TEAM NAME' and a text input field containing 'Team 0 (your team number)'. A red rectangular box highlights the entire 'General' section, including the form and the explanatory text below it: 'This name will appear on your team's leaderboard position.'

Overview Data Code Models Discussion Leaderboard Rules Team Submissions

Your Team

Everyone that competes in a Competition does so as a team - even if you're competing by yourself. [Learn more.](#)

General

TEAM NAME

Team 0 (your team number)

This name will appear on your team's leaderboard position.


Team up (Important!!)

Invite your teammate(s)

- If you participate in this project by yourself, you don't need to do this step
- Search your teammates by their Kaggle account email
 - ▣ They must be added to the contest and "accept the rules" first.

Team Members


Your team can have a maximum of 10 members.



(You)
Team Leader

Send Invitation

You can invite 9 more people to join your team. They will need to have joined the competition.

 Search for the team name

Team up (Important!!)

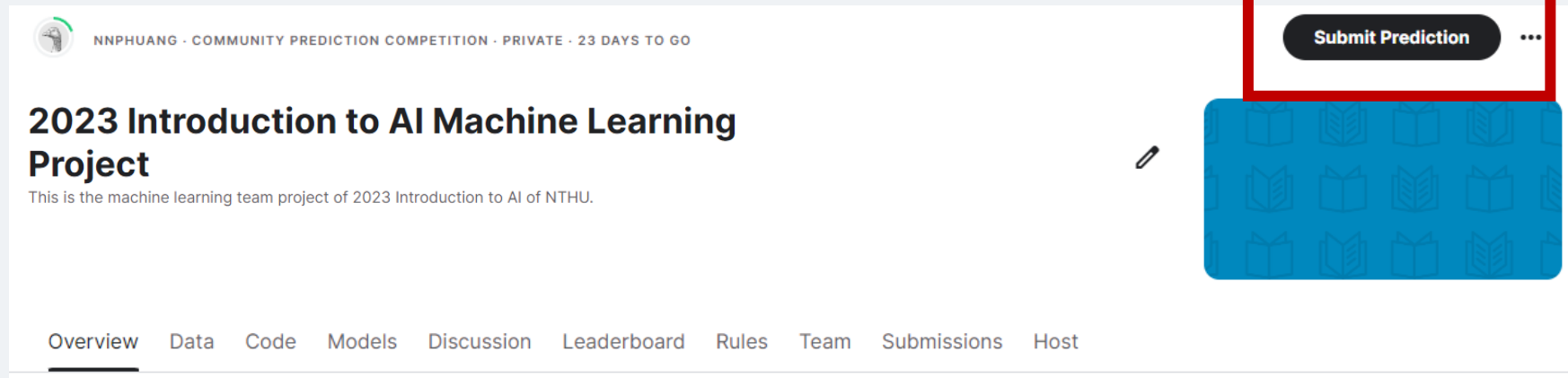
Anyone who doesn't follow these regulations
will be revoked from the contest!

You need to do these before you make any
submissions.

Kaggle submission

Public leaderboard submission

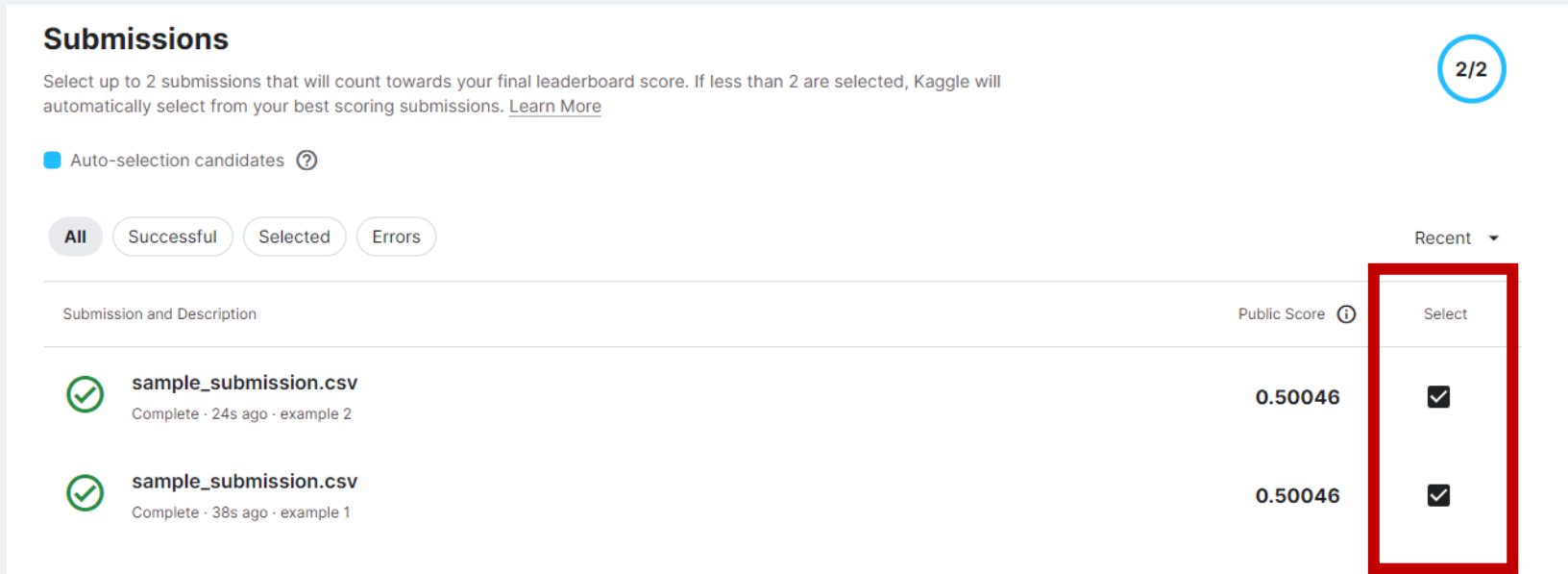
- You can submit up to 5 times per day as a team
- Your submission file must follow the format of sample_submission.csv





Kaggle submission

Private leaderboard submission

- Select 2 submissions as the private leaderboard submission
- The private leaderboard ranking will be released after the competition ends



The screenshot shows the 'Submissions' page on Kaggle. At the top, there's a header 'Submissions' with a subtext: 'Select up to 2 submissions that will count towards your final leaderboard score. If less than 2 are selected, Kaggle will automatically select from your best scoring submissions. [Learn More](#)'. A blue circle with '2/2' indicates that two submissions have been selected. Below this, there's a section for 'Auto-selection candidates' with a question mark icon. A filter bar shows 'All', 'Successful', 'Selected', and 'Errors' tabs, with 'Selected' being the active tab. A 'Recent' dropdown menu is also visible. The main table lists submissions with columns for 'Submission and Description', 'Public Score', and 'Select'. Two submissions are listed, both with a public score of 0.50046. The 'Select' column for both submissions shows a checked checkbox, indicating they are selected. A red rectangle highlights the 'Select' column and the checkboxes for both submissions.

Submission and Description	Public Score	Select
 sample_submission.csv Complete · 24s ago · example 2	0.50046	<input checked="" type="checkbox"/>
 sample_submission.csv Complete · 38s ago · example 1	0.50046	<input checked="" type="checkbox"/>

Report

Report

Things you should write in your report

- Explain your methods
- Describe how you improve the results

Format

- Save as a pdf file

Grading

Evaluation

- Report 60%
- Performance 60%
 - ▣ Private leader board F1 score 30%
 - ▣ Private leader board rank 30%
- Cooperation 30%
 - ▣ Teammates' evaluation by Google form. (The form will be announced later)
- No need to be stressed out about your grade 😊
 - ▣ If you aren't satisfied with your performance on the private leaderboard, you can still make up by writing an informative report.

Contact us if you have problems...

Kaggle discussion page

- We strongly encourage you to post your question here
- You can also search the competition forums in case a similar question has been answered

Email

- Questions regarding the dataset
 - ▣ 郭宥成 nelson891224@gmail.com
- Questions regarding Kaggle
 - ▣ 黃恩平 huangenping0512@gmail.com

