

**CSED 226**  
**Introduction to Data Analysis**  
**Competition: Rating Prediction**

# Overview

- The goal is to improve your model's performance **as much as possible**.
- You can improve your model's performance through **model selection and hyperparameter tuning**.
- You can submit your model's results to Kaggle and view your scores in real-time on the leaderboards.
  - You can only submit a **maximum of 20 per day**.

# Rules

- In this competition, you will use [Surprise](#), the python library for recommender systems.
- You must participate as an individual team in this competition.

# Evaluation

- Submissions are evaluated on the [Root Mean Square Error \(RMSE\)](#) between **the ground truth rating** and **your predicted rating**.
  - Please refer to the link.

# Dataset

- **Movie-Rating Prediction Dataset**

- The dataset consists of three columns.
  - ✓ user\_id, item\_id, rating
- You should train your model with **this interaction data**.

# Competition

- You will enter the competition individually through Kaggle.
- Kaggle Site: <https://www.kaggle.com>
- Competition URL: <https://www.kaggle.com/t/7c309521f9294bc4b5d42cabfdb32f31>

# Competition

Community Prediction Competition

**DataAnalysis-Classification**  
Plant species

Overview Data Notebooks Discussion **Leaderboard** Rules Team My Submissions **Submit Predictions**

① ② ③ ④ ⑤ ⑥

Your most recent submission

Name	Submitted	Wait time	Execution time	Score
sample2.csv	12 hours ago	0 seconds	0 seconds	0.80454

Complete

[Jump to your position on the leaderboard.](#)

Public Leaderboard Private Leaderboard

This leaderboard is calculated with approximately 50% of the test data.  
The final results will be based on the other 50%, so the final standings may be different.

[Raw Data](#) [Refresh](#)

#	Team Name	Notebook	Team Members	Score	Entries	Last
1	Team1			0.80757	2	12h

**Your Best Entry** ↑

Your submission scored 0.80454, which is not an improvement of your best score. Keep trying!

📍 Baseline 0.79040

## Kaggle-site

- ① Overview: Description of the task
- ② Data: Files to be used in the competition
- ③ Leaderboard: Current score updated in real time
- ④ Rules: What you must follow in the competition
- ⑤ My submissions: Files you submitted
- ⑥ Submit Predictions: Submission