Recommending Boulder Problems with Graph Neural Networks

Žan Mervič, Tadej Jeršin, Vid Čabrian
Slovenian
Večna pot 113, Ljubljana Slovenia
zm3587@student.uni-lj.si,
tj90781@student.uni-lj.si,
vc76033@student.uni-lj.si

1 Dataset

We used publicly available data from the MoonBoard mobile application and its official website, https://moonboard.com/. The MoonBoard platform hosts a large collection of user-generated climbing problems along with extensive user interaction data.

Since no standardized dataset existed for this domain, we collected the data ourselves through web scraping. The resulting dataset includes entities such as users, climbing problems, and their relationships all of which could naturally be represented as a graph. The data collection process focused solely on aggregating and structuring publicly available information, no private or sensitive user data were collected.

1.1 Dataset Description, Task, and Metrics

The MoonBoard dataset consists of climbing problems, users and their interactions. Each climbing problem includes the following attributes: grade, rating, number of sends, setter hash, and the set of holds defining the problem.

Each user is represented with attributes such as ranking, highest grade, number of problems sent, and a dictionary of problems they had climbed, including their ratings and comments. Example entries for both users and problems are shown below.

```
"47C5D8F8-7E58-4480-A6": {
  "ranking": 15704,
  "highest_grade": "6C",
  "problems_sent": 17,
  "problems": {
      "PREMIERA": {
          "grade": "6A+",
          "rating": 4.0,
          "date": "2025-02-08",
          "comment": "Flashed"
      }
  }
}
```

```
"PIZZA DECHU": {
  "grade": "7A",
  "rating": 4.0,
  "num_sends": 35,
  "setter_hash": "01dfa...",
  "holds": ["3G", "4I", "6E", ...]
}
```

We will model the dataset as a bipartite graph, where nodes represent users and climbing problems, and edges represented user–problem interactions. Additional attributes such as grades, ratings, and send counts are stored as node or edge features.

The primary task is to build a recommendation system that predicts which climbing problems a user is most likely to attempt or enjoy, based on historical user–problem interactions and problem similarity. We will model this as a link prediction problem. To evaluate the system, we will use Recall@k, which

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

measures the proportion of relevant routes that appear among the top k recommendations for each user. Recall@k is particularly suitable in this context because climbing users typically try only a small subset of available routes, and it emphasizes whether the system successfully surfaces the most relevant options, rather than penalizing the model for ordering less relevant routes beyond the top recommendations.

1.2 Why Did We Choose This Dataset?

We chose the MoonBoard dataset because it represents a unique, real-world domain that combines aspects of social interaction, user preferences, and spatial configuration, all of which can be effectively modeled as a graph. The data naturally forms a bipartite structure between users and climbing problems, making it well-suited for applying graph neural network (GNN) techniques to learn user–problem relationships.

Additionally, the MoonBoard platform is a vibrant and active global climbing community where users continuously generate new problems and record ascents. This dynamic and diverse data source provides an excellent opportunity to study recommendation strategies in a niche but meaningful context.

2 Graph ML Techniques

We plan to use LightGCN[1], PinSAGE[3], and GFormer[2] for recommending Moonboard climbing routes. All three models are appropriate for this task because the data naturally forms a user-route interaction graph, making graph-based recommendation methods ideal. These models were specifically designed for recommender systems as they leverage the graph structure, interaction patterns, and node features (with the exception of LightGCN) to generate high-quality personalized route recommendations.

2.1 LightGCN

LightGCN is a simplified graph convolutional network for collaborative filtering that removes nonlinear transformations and feature projections, focusing solely on neighborhood aggregation. Its simplicity makes it efficient, stable, and highly suitable for sparse interaction data such as MoonBoard ascents and ratings. We will use LightGCN as a strong baseline to model general user–problem relationships.

2.2 PinSAGE

PinSAGE extends graph convolutional methods to web-scale recommendation systems by combining random-walk-based neighborhood sampling with feature aggregation. It effectively captures graph connectivity while remaining scalable and robust. We plan to apply PinSAGE to model higher-order user–problem relationships and incorporate side information such as user ratings, comments and hold configurations.

2.3 GFormer

GFormer is a recent graph transformer model that applies attention mechanisms to learn both local and global dependencies in user–item graphs. It integrates self-supervised learning through masked graph autoencoding, improving robustness to noise and data sparsity. We intend to evaluate GFormer as an advanced model for capturing complex interaction patterns and user preferences in the MoonBoard dataset.

References

[1] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, pages 639–648, 2020.

[2] Chaoliu Li, Lianghao Xia, Xubin Ren, Yaowen Ye, Yong Xu, and Chao Huang. Graph transformer for recommendation. In *Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval*, pages 1680–1689, 2023.

[3] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L. Hamilton, and Jure Leskovec. Graph convolutional neural networks for web-scale recommender systems. *CoRR*, abs/1806.01973, 2018. URL http://arxiv.org/abs/1806.01973.

MLG:
Machine Learning with Graphs (*Strojno učenje na grafih*)

Assignment: Project Proposal

Submission time: 15.32 **and date:** 21.10.2025

Submission Fill in and include this cover sheet with each of your assignments. It is an honor code violation to write down the wrong date and/or time. Assignments are due at 9:00am and should be submitted through Gradescope and eUcilnica. Students should check Piazza for submission details.

Late Periods Each student will have a total of *two* free late periods. *Late period expires the morning on the day before the next class.* (Assignments are usually due on Fridays, which means the late period expires on the following Tuesday at 9:00am.) Once these late periods are exhausted, any assignments turned in late will be penalized 50% per late period. However, no assignment will be accepted more than *one* late period after its due date.

Honor Code We strongly encourage students to form study groups. Students may discuss and work on assignments in groups. However, each student must write down their solutions independently, i.e., each student must understand the solution well enough in order to reconstruct it by him/herself. Students should clearly mention the names of all the other students who were part of their study group. Using code or solutions obtained from the web (ChatGPT, GitHub, Google, previous years etc.) is considered an honor code violation. We check all submissions for plagiarism. We take the honor code very seriously and expect students to do the same.

Name: Tadej Jeršin

Email: tj90781@student.uni-lj.si **SID:** 63240451

Study group: DS 2.year

I acknowledge and accept the Honor Code.

(Signed)

MLG:
Machine Learning with Graphs (*Strojno učenje na grafih*)

Assignment: Project Proposal

Submission time: 15.32 **and date:** 21.10.2025

Submission Fill in and include this cover sheet with each of your assignments. It is an honor code violation to write down the wrong date and/or time. Assignments are due at 9:00am and should be submitted through Gradescope and eUcilnica. Students should check Piazza for submission details.

Late Periods Each student will have a total of *two* free late periods. *Late period expires the morning on the day before the next class.* (Assignments are usually due on Fridays, which means the late period expires on the following Tuesday at 9:00am.) Once these late periods are exhausted, any assignments turned in late will be penalized 50% per late period. However, no assignment will be accepted more than *one* late period after its due date.

Honor Code We strongly encourage students to form study groups. Students may discuss and work on assignments in groups. However, each student must write down their solutions independently, i.e., each student must understand the solution well enough in order to reconstruct it by him/herself. Students should clearly mention the names of all the other students who were part of their study group. Using code or solutions obtained from the web (ChatGPT, GitHub, Google, previous years etc.) is considered an honor code violation. We check all submissions for plagiarism. We take the honor code very seriously and expect students to do the same.

Name: Žan Mervič

Email: zm3587@student.uni-lj.si **SID:** 63200190

Study group: DS 2.year

I acknowledge and accept the Honor Code.

(Signed)

MLG:
Machine Learning with Graphs (*Strojno učenje na grafih*)

Assignment: Project Proposal

Submission time: 15.32 **and date:** 21.10.2025

Submission Fill in and include this cover sheet with each of your assignments. It is an honor code violation to write down the wrong date and/or time. Assignments are due at 9:00am and should be submitted through Gradescope and eUcilnica. Students should check Piazza for submission details.

Late Periods Each student will have a total of *two* free late periods. *Late period expires the morning on the day before the next class.* (Assignments are usually due on Fridays, which means the late period expires on the following Tuesday at 9:00am.) Once these late periods are exhausted, any assignments turned in late will be penalized 50% per late period. However, no assignment will be accepted more than *one* late period after its due date.

Honor Code We strongly encourage students to form study groups. Students may discuss and work on assignments in groups. However, each student must write down their solutions independently, i.e., each student must understand the solution well enough in order to reconstruct it by him/herself. Students should clearly mention the names of all the other students who were part of their study group. Using code or solutions obtained from the web (ChatGPT, GitHub, Google, previous years etc.) is considered an honor code violation. We check all submissions for plagiarism. We take the honor code very seriously and expect students to do the same.

Name: Vid Čabrian

Email: vc76033@student.uni-lj.si **SID:** 63240450

Study group: DS 2. year

I acknowledge and accept the Honor Code.

(Signed)