**Peer Consulting Report**

**Board game rating**

**predictor**

**31 December 2023**

# Module 3 project of Ramona Herz & Lara Nonis

# Consultation report by: Patricio Becerra

# Content

[Content 1](#_Toc154565762)

[1 Project Summary 2](#_Toc154565763)

[2 Advantages 2](#_Toc154565764)

[3 Suggestions for Improvements 2](#_Toc154565765)

[4 Conclusions 2](#_Toc154565766)

[References 2](#_Toc154565767)

# 1 Project Summary

BoardGameGeek [1] is an online forum and games database for board game hobbyists. It holds information on over 125’000 tabletop games, including game data, reviews, images and videos. BoardGameGeek allows users to rate games on a scale of 1-10 and publishes a ranked list of board games based on these ratings. This ranking is a measure of the games’ popularity and is considered an important reference value among hobbyists (and thus a proxy for market success of a game).

The goal of the “Board game rating predictor” project is to develop an algorithm that can predict the BoardGameGeek rating of a new game or a game under development using a series of characteristics or features of the target game. The model was conceived as a tool for game developers to obtain an idea of the potential popularity or market value of their game before its launch, allowing them to fine-tune certain aspects of the game to improve its chances of success.

The project is split into two phases. Each phase will make use of different machine learning techniques:

* Phase I has the objective of finding the features of a game (as recorded in BoardGameGeek) that have the most effect on the game’s rating. These features could be e.g. number of players, mechanism of play, average play time. This phase will use unsupervised learning to find the features that most influence the target variable (rating).
* Phase II will use the information from Phase I to develop the predictor algorithm and provide a single predicted rating value for any game. This phase will use regression techniques (i.e., supervised learning) to find the rating prediction from multiple game features.

The game features include game ID, game name, minimal number of players, maximal number of players, rated weight, average rating, bgg rating, number of expansions, year published, designers, artists, publishers, mechanisms, categories, sculptors, game family and game subtype. The data is obtained via web scraping of the BoardGamesGeek platform through its API [2].

# 2 Advantages

# 3 Suggestions for Improvements

# 4 Conclusions

# References

1. S. Schaller, M. W. Buechi, B. Schuster, and F. S. Anselmetti (2023) Drilling into a deep buried valley (ICDP DOVE): a 252m long sediment succession from a glacial overdeepening in northwestern Switzerland. *Scientific Drilling* (2023), doi.org/10.5194/sd-32-27-2023
2. DOVE-Phase 1 Scientific Team, Anselmetti, F. S., Beraus, S., … , Schaller, S., et al. (2023): Drilling Overdeepened Alpine Valleys (DOVE) – Operational Report of Phase 1, (ICDP Operational Report), Potsdam : GFZ German Research Centre for Geosciences, 70 <https://doi.org/10.48440/ICDP.5068.001>
3. <https://www.icdp-online.org/projects/by-continent/europe/dove-switzerland/public-data>
4. S. Schaller (2023), ADS CAS Conceptual Design Report: Core and wireline log based lithoclasification of unconsolidated sediments from drill cores and boreholes.
5. Lee, A.-S., Enters, D., Huang, J.-J.S., Liou, S. Y. H. & Zolitschka, B. (2022) An automatic sediment-facies classification approach using machine learning and feature engineering. *Nature* *Commun. Earth Environ.* 3, 1–9, doi.org/10.1038/s43247-022-00631-2
6. Xing, Y.; Yang, H.; Yu, W., (2023) An Approach for the Classification of Rock Types Using Machine Learning of Core and Log Data. *Sustainability* , 15, 8868. [doi.org/10.3390/su15118868](https://doi.org/10.3390/su15118868)
7. Alzubaidi, F., Mostaghimi, P., Swietojanski, P., Clark, S. R., & Armstrong, R. T. (2021). Automated lithology classification from drill core images using convolutional neural networks. *Journal of Petroleum Science and Engineering*, doi.org/10.1016/j.petrol.2020.107933.
8. Lauper B., Zimmerli G. N., Jaeggi D., Deplazes G., Wohlwend S., Rempfer J., Foubert A. (2021) Quantification of Lithological Heterogeneity Within Opalinus Clay: Toward a Uniform Subfacies Classification Scheme Using a Novel Automated Core Image Recognition Tool. *Frontiers in Earth Science*, 9. doi.org/10.3389/feart.2021.645596
9. Jacq, K. et al. Sedimentary structure discrimination with hyperspectral imaging in sediment cores. *Sci. Total Environ.* 817, 152018 (2022).
10. Fu D., Su C., Wang W., Yuan R. (2022) Deep learning-based lithology classification of drill core images. *PLoS ONE* 17(7). doi. org/10.1371/journal.pone.0270826
11. Di Martino, A., Carlini, G., Castellani, G. et al. (2023). Sediment core analysis using artificial intelligence. *Nature Sci. Rep.* 13, 20409. doi.org/10.1038/s41598-023-47546-2
12. Polonia, A., et al. "Recognizing mega tsunamis in Mediterranean deep sea sediments based on the massive deposits of the 365 CE Crete event." *Scientific Reports* 12.1 (2022): 5253.
13. Morcillo, G., Faria, S. H., and Kipfstuhl, S. (2020). Unravelling Antarctica's Past through the Stratigraphy of a Deep Ice Core: an Image-Analysis Study of the EPICA-DML Line-Scan Images. *Quat. Int.* 566-567, 6–15. doi:10.1016/j.quaint.2020.07.011
14. Apollo to Artemis: Drilling on the Moon (2021). <https://www.nasa.gov/centers-and-facilities/kennedy/apollo-to-artemis-drilling-on-the-moon/>
15. Deep drilling for Mars (2023). <https://www.esa.int/ESA_Multimedia/Images/2023/03/Deep_drilling_for_Mars>
16. Becerra, P., I.B. Smith, et al. (2021), Past, Present and Future of Mars Polar Science: Outcomes and Outlook from the 7th International Conference on Mars Polar Science and Exploration, *The Planetary Science Journal*, 2:209, doi.org/10.3847/PSJ/ac19a5
17. Van Rossum, G. (2022). os—Miscellaneous operating system interfaces. <https://docs.python.org/3/library/os.html>
18. The Pandas Development Team. (2022). pandas-dev/pandas: Pandas (v1.5.1). Zenodo. <https://doi.org/10.5281/ZENODO.3509134>
19. Harris C.R., et al. (2020) “Array programming with NumPy,” *Nature*, vol. 585, no. 7825, Art. no. 7825, doi: 10.1038/s41586-020-2649-2.
20. P. Virtanen et al (2020) SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nat. Methods*, vol. 17, pp. 261–272. doi: 10.1038/s41592-019-0686-2.
21. S. Seabold and J. Perktold (2010) “statsmodels: Econometric and statistical modeling with python,” in *9th Python in Science Conference*. <https://www.statsmodels.org/stable/index.html>
22. F. Pedregosa et al. (2011) “Scikit-learn: Machine learning in Python,” J. Mach. Learn. Res., vol. 12, no. Oct, pp. 2825–2830, <https://scikit-learn.org/stable/about.html>
23. M. Abadi, A. Agarwal, et al. (2015) TensorFlow: Large-scale machine learning on heterogeneous systems,. Software available from tensorflow.org
24. S. van der Walt, J. L. Schönberger, et al. (2014) scikit-image: Image processing in Python. PeerJ 2:e453. <https://doi.org/10.7717/peerj.453>
25. Clark, J. et al. (2010). PIL : Python Imaging Library, <https://pypi.org/project/Pillow/>
26. OpenCV. (2015). Open Source Computer Vision Library. <https://opencv.org/>
27. Caswell, T. A., Lee, A., et al. (2022). matplotlib/matplotlib: REL: v3.6.1 (v3.6.1). Zenodo. <https://doi.org/10.5281/ZENODO.592536>
28. Y. Ikeda, “mpltern: Ternary plots with Matplotlib.” Zenodo, 2023. doi: <https://doi.org/10.5281/zenodo.3528354>.
29. Waskom, M. (2021). seaborn: Statistical data visualization. Journal of Open Source Software , 6 (60), 3021. <https://doi.org/10.21105/joss.03021>
30. Plotly Technologies Inc. (2015). Collaborative data science. Plotly Technologies Inc. <https://plot.ly>
31. Ronneberger, O., Fischer, P. & Brox, T. (2015) U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Interventio.* doi.org/10.1007/978-3-319-24574-4\_28.
32. Tan, M. & Le, Q. V. (2020). EfficientNet: Rethinking model scaling for convolutional neural networks. https://doi.org/10.48550/arXiv.1905.11946.
33. He, K., Zhang, X., Ren, S. & Sun, J. (2020). Deep residual learning for image recognition. doi.org/10.48550/arXiv.1512.03385.
34. Russakovsky, O. et al. (2015). ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* 115, 211–252.