Verification and Validation Report: TTE RecSys

Yinying Huo

April 2, 2025

1 Revision History

Date	Version	Notes
Apr. 2 2025	1.0	First draft

2 Symbols, Abbreviations and Acronyms

symbol	description
T	Test

[symbols, abbreviations or acronyms – you can reference the SRS tables if needed —SS]

Contents

List of Tables

List of Figures

This document includes the results of VnV plan.

3 Functional Requirements Evaluation

The functional requirements are tested using both system tests and unit tests.

3.1 Dataset

To ensure the project receives a valid dataset for training:

Manually, a Jupyter Notebook (data_preprocessing.ipynb) is used to merge the raw data (three CSV files) into a single CSV file.

Then, the automated system test (test_data_validation.py) runs each time the dataset is updated before training starts to ensure that the input CSV file contains all the required columns.

Additionally, an automated unit test (test_data_processing.py) runs on each push to verify the correctness of the data processor, which generates new features for users and items.

All tests have passed successfully.

3.2 Model Training Convergence

The system test (test_model_convergence.py) checks whether the training loss decreases as training progresses.

This test has passed successfully.

3.3 Model storage

The system test (test_model_storage.py) runs each time model training is completed. The test ensures that the trained model and computed embeddings are correctly stored in the 'output' folder.

This test has passed successfully.

3.4 Embedding generation

The unit test 'test_embedding_generation.py' will check the correctness of embedding generation. This test has passed successfully.

4 Nonfunctional Requirements Evaluation

4.1 Usability

4.2 Reliability

4.3 Protability

5 Unit Testing

6 Changes Due to Testing

[This section should highlight how feedback from the users and from the supervisor (when one exists) shaped the final product. In particular the feedback from the Rev 0 demo to the supervisor (or to potential users) should be highlighted. —SS]

7 Automated Testing

8 Trace to Requirements

9 Trace to Modules

10 Code Coverage Metrics

References

Author Author. System requirements specification. https://github.com/..., 2019.

Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. The faiss library, 2024. URL https://arxiv.org/abs/2401.08281.

Carlo Ghezzi, Mehdi Jazayeri, and Dino Mandrioli. *Fundamentals of Software Engineering*. Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 2003.

Daniel M. Hoffman and Paul A. Strooper. *Software Design, Automated Testing, and Maintenance: A Practical Approach*. International Thomson Computer Press, New York, NY, USA, 1995. URL http://citeseer.ist.psu.edu/428727.html.

Yinying Huo. Software requirements specification for tte recsys, 2025. URL https://github.com/V-AS/Two-tower-recommender-system/blob/main/docs/SRS/SRS.pdf. Accessed: February 18, 2025.

David L. Parnas. On the criteria to be used in decomposing systems into modules. *Comm. ACM*, 15(2):1053–1058, December 1972.

David L. Parnas. Designing software for ease of extension and contraction. In *ICSE '78: Proceedings of the 3rd international conference on Software engineering*, pages 264–277, Piscataway, NJ, USA, 1978. IEEE Press. ISBN none.

David L. Parnas and P.C. Clements. A rational design process: How and why to fake it. *IEEE Transactions on Software Engineering*, 12(2):251–257, February 1986.

D.L. Parnas, P.C. Clement, and D. M. Weiss. The modular structure of complex systems. In *International Conference on Software Engineering*, pages 408–419, 1984.

James Robertson and Suzanne Robertson. *Volere Requirements Specification Template*. Atlantic Systems Guild Limited, 16 edition, 2012.

W. Spencer Smith. Systematic development of requirements documentation for general purpose scientific computing software. In *Proceedings of the 14th IEEE International Requirements Engineering Conference, RE 2006*, pages 209–218, Minneapolis / St. Paul, Minnesota, 2006. URL http://www.ifi.unizh.ch/req/events/RE06/.

W. Spencer Smith and Nirmitha Koothoor. A document-driven method for certifying scientific computing software for use in nuclear safety analysis. *Nuclear Engineering and Technology*, 48(2):404–418, April 2016. ISSN

1738-5733. doi: http://dx.doi.org/10.1016/j.net.2015.11.008. URL http://www.sciencedirect.com/science/article/pii/S1738573315002582.

W. Spencer Smith and Lei Lai. A new requirements template for scientific computing. In J. Ralyté, P. Ågerfalk, and N. Kraiem, editors, *Proceedings of the First International Workshop on Situational Requirements Engineering Processes – Methods, Techniques and Tools to Support Situation-Specific Requirements Engineering Processes, SREP'05*, pages 107–121, Paris, France, 2005. In conjunction with 13th IEEE International Requirements Engineering Conference.

W. Spencer Smith, Lei Lai, and Ridha Khedri. Requirements analysis for engineering computation: A systematic approach for improving software reliability. *Reliable Computing, Special Issue on Reliable Engineering Computation*, 13(1):83–107, February 2007.

W. Spencer Smith, John McCutchan, and Jacques Carette. Commonality analysis of families of physical models for use in scientific computing. In *Proceedings of the First International Workshop on Software Engineering for Computational Science and Engineering (SECSE 2008)*, Leipzig, Germany, May 2008. In conjunction with the 30th International Conference on Software Engineering (ICSE). URL http://www.cse.msstate.edu/~SECSE08/schedule.htm. 8 pp.

W. Spencer Smith, John McCutchan, and Jacques Carette. Commonality analysis for a family of material models. Technical Report CAS-17-01-SS, McMaster University, Department of Computing and Software, 2017.

Wikipedia. Dot product — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Dot%20product&oldid=1268352915, 2025a. [Online; accessed 17-February-2025].

Wikipedia. Gradient descent — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Gradient%20descent&oldid=1274464503, 2025b. [Online; accessed 17-February-2025].

Wikipedia. Mean squared error — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Mean%20squared%20error&oldid=1276072434, 2025c. [Online; accessed 17-February-2025].

Wikipedia. Stochastic gradient descent — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Stochastic%20gradient%20descent&oldid=1265558819, 2025d. [Online; accessed 17-February-2025].

Xinyang Yi, Ji Yang, Lichan Hong, Derek Zhiyuan Cheng, Lukasz Heldt, Aditee Ajit Kumthekar, Zhe Zhao, Li Wei, and Ed Chi, editors. *Sampling-Bias-Corrected Neural Modeling for Large Corpus Item Recommendations*, 2019.

Appendix — Reflection

The information in this section will be used to evaluate the team members on the graduate attribute of Reflection.

1. What went well while writing this deliverable?

2. What pain points did you experience during this deliverable, and how did you resolve them?

3. Which parts of this document stemmed from speaking to your client(s) or a proxy (e.g. your peers)? Which ones were not, and why?

4. In what ways was the Verification and Validation (VnV) Plan different from the activities that were actually conducted for VnV? If there were differences, what changes required the modification in the plan? Why did these changes occur? Would you be able to anticipate these changes in future projects? If there weren't any differences, how was your team able to clearly predict a feasible amount of effort and the right tasks needed to build the evidence that demonstrates the required quality? (It is expected that most teams will have had to deviate from their original VnV Plan.)