BLG202E Numerical Methods in Comp. Eng.

Spring 2024 - Term Project

Due: May 5, 2023

By turning in this assignment, I agree by the ITU honor code and declare that all of this is my own work.

Important Notes

- Upload your solutions through **Ninova**. Projects sent via e-mail and late submissions **will not be accepted**.

- Please make sure that you write your **full name** and student identification **number** to **every file** you submit.

- Cheating is highly discouraged. It will be punished by a negative grade. Also disciplinary actions will be taken. Please do your homework on your own. Team work is not allowed. Pattern of your **solutions must belong only to you**.

- All codes and reports will be run through **plagiarism check**s. Please **do not copy any text or code** from other sources.

- If you have any questions, please contact with **T.A. Kıymet Kaya** (kayak16@itu.edu.tr).

- Remember, there are only 10 types of people in the world – those who understand binary, and those who don't.

Project No 4: Item-based Filtering Enhanced by SVD

Please read the following prerequisites.

- Install a Conda environment if you do not have it already.

- Install Jupyter Notebook.

- The project should be done in Jupyter Notebook.

- Do not forget to format your code and leave comments for non-trivial sections.

- You are expected to use matplotlib, numpy / SciPy, json and pandas(for data reading).

Instructions.

- Re-implement and replicate the results of the following paper below on IMDB dataset.

    - **Applying SVD on item-based filtering.**
      In 5th International Conference on Intelligent Systems Design and Applications (ISDA'05) (pp. 464-469). IEEE.

    - **IMBD Dataset:**
      https://www.kaggle.com/datasets/ebiswas/imdb-review-dataset/data

- Method Implementation and Experimental Setup:

    - You need to implement SVD from scratch. You can't use built-in libraries or methods like 'numpy.linalg.svd', 'scipy.linalg.svd', 'sklearn.decomposition' etc.

    - You are expected to write code and get numerical results. Submissions without a working code will not be graded.

    - Please note that rating in the GroupLens data (in the paper) range from (1-5), while ratings in the IMDB dataset range from (1-10).

    - Include only the Users who have stated their opinions for at least 20 movies.

    - Observe the results for different active item's neighborhood= $\{20, 60, 80\}$.

    - Observe the results for different values of k, k= $\{2, 4, 6, 8, 10, 20\}$.

- Report:

    - Write a maximum of 3 pages report using IEEE Latex Template. State the problem, implementation details, dataset and experiments.

- Submission

    - Submit a zip file that includes your Python code, plots, and results until the deadline through Ninova.

    - Upload your project report until the deadline through Ninova.