**Homework 1: News Feed Recommendation**

**Objective:**

Given what you know at this point, how would you go about implementing your first recommendation system end-to-end? Specifically, think about recommending daily news via email. You have a list of users and their impressions, and your task is to design a simple news feed ranking system.

**Dataset:**

**MIND News Dataset:** <https://www.kaggle.com/datasets/arashnic/mind-news-dataset/data>

**Approach:**

Apply your collaborative filtering approach based on what we covered in class. You can use content-based, user-based, matrix factorization (MF), or even come up with your own idea. You can implement conventional methods, clustering, SVD, etc. Note that it doesn’t have to involve deep learning.

**Important Note:** Copying existing Kaggle solutions is not accepted and will not be graded.

**Evaluation:**

Evaluate your approach using the right metrics and submit a two-page short paper (maximum two pages). No code should be included in the report.

**Sample Papers on the MIND Dataset:**

* <https://aclanthology.org/2020.acl-main.331.pdf>
* <https://arxiv.org/pdf/2209.06131>

**Note:** The report should be in PDF format. It can be either double or single column. You may use [Overleaf](https://www.overleaf.com/) for typesetting if you wish.

**Grading Rubric:**

* **Approach (40 points)**
  + Example 1: Content-Based
  + Example 2: User-Based
  + Example 3: Clustering or SVD, etc
* **Results and Evaluation (40 points)**
  + Experimental Setup (10 points): How did you divide the data into training and testing sets? Did you follow the given setup and why?
  + Metrics (5 points) the reason you chose these metrics? Are there metrics reflective enough for the strength of recommendation model?
  + Results and Comparison (25 points) Please discuss the results and not just put numbers.
* **Executable Notebooks (20 points)**

**Submission:**

Submit a single ZIP file. Ensure that the ZIP file can be unzipped before submission. The submitted file should include:

* One PDF file containing your report.
* The code used for the project.