Project Report 1

Ethics Engineers

1. Will you be working in a group? If so, please state your group members so that we can create the student groups in Gradescope.
   1. Yannick Tanyi (me), Sana Fessuh, and Will Greenstone
2. Please state your initial brainstorming ideas for the question you plan to answer. Make sure to describe the connection to ethics, responsibility, privacy, or fairness. Remember, your answers should be BRIEF!
   1. State succinctly your the problem/question.
      1. Can user-centered debiasing strategies mitigate both popularity and exposure bias in Amazon's digital music recommendation system, especially for underrepresented user groups (e.g., those with low review frequency, fewer helpful votes, and lower engagement), while maintaining recommendation accuracy?
   2. What methodology you plan to use/learn to address the question?
      1. To address the question, we plan to use a combination of data analysis, machine learning, and debiasing techniques. Specifically, we will: 1. Data Collection and Preprocessing: We will collect data from Amazon’s digital music platform, focusing on user interaction metrics such as review frequency, helpful votes, and engagement. We'll preprocess this data to account for underrepresented groups. 2. Algorithm Auditing: We will audit existing recommendation algorithms to identify areas where popularity and exposure bias occur, using fairness metrics like exposure disparity and ranking parity. 3. Debiasing Techniques: We plan to implement user-centered debiasing strategies such as fairness-aware re-ranking and calibrated recommendations. These methods will focus on adjusting recommendations to give underrepresented users more exposure without sacrificing accuracy. 4. Evaluation: We will measure the effectiveness of these strategies by comparing recommendation accuracy (via precision and recall) and fairness (via exposure parity) before and after debiasing. In this process, we aim to learn and apply machine learning concepts, fairness in AI frameworks, and evaluation techniques for both bias and accuracy.
   3. What do you expect to find out?
      1. We expect to find that user-centered debiasing strategies can indeed mitigate popularity and exposure bias in Amazon's digital music recommendation system, particularly for underrepresented user groups. Specifically, we anticipate: 1. Increased Fairness: Underrepresented users, such as those with lower review frequency or engagement, will receive more equitable exposure in the recommendations, reducing the disparity caused by traditional popularity metrics. 2. Maintained Accuracy: While we aim to improve fairness, we expect that the accuracy of recommendations—measured by precision and recall—will remain high, showing that debiasing does not significantly compromise the system’s effectiveness in providing relevant suggestions. Overall, we hope to demonstrate that fairness and accuracy can coexist in recommendation systems when carefully designed debiasing strategies are applied.
   4. What type of deliverable do you plan on submitting (e.g., paper, software artifact, short movie, demonstrator, low-fidelity prototype, or any other deliverable you can justify)?
      1. Paper/powerpoint
3. Describe the next steps you plan to take to answer your question. If working in a group, make sure you roughly describe each member's plans. Remember, your answers should be BRIEF!
   1. The next steps we plan to take are to clean and prepare our Amazon data for analysis. We'll then run basic data visualizations to identify if there are specific user groups, such as those with low engagement or review frequency, that are underrepresented by the current recommendation algorithms. This will help us pinpoint areas where bias may exist and guide our subsequent analyses and debiasing strategies.
4. Github