Least Squares, Most Gains

Alban Brüchig, Osman Örnek, Filip Mikovíny

Machine Learning for Behavioral Data (MLBD) Course, EPFL, 2025

1 INTRODUCTION + RESEARCH QUESTIONS

We are interested in observing different learning patterns and subgroups of students among the GoGymi student community.

With this in mind we chose to research *whether*:

- **RQ1:** Does the *AI model / Gymitrainer-usage links to increased quiz grades?*
- **RQ2:** Do *student clusters based on essay skill scores form a relationship with mathematics and text quiz scores?*
- **RQ3:** *Is student activity an effective predictor of whether a student will disengage from an educational application or not?*

Motivation: infer on how and which students learn best (or the least) and why (not).
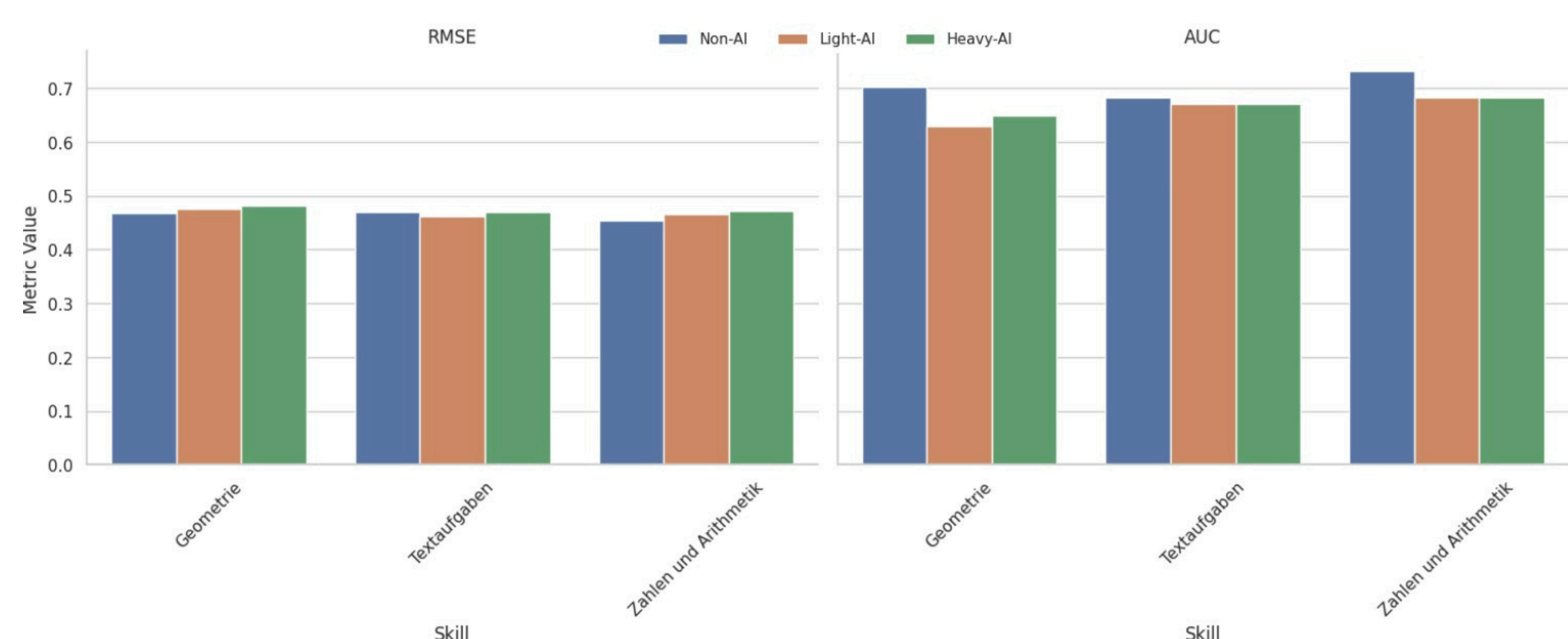
3 VISUALIZATION OF RESULTS

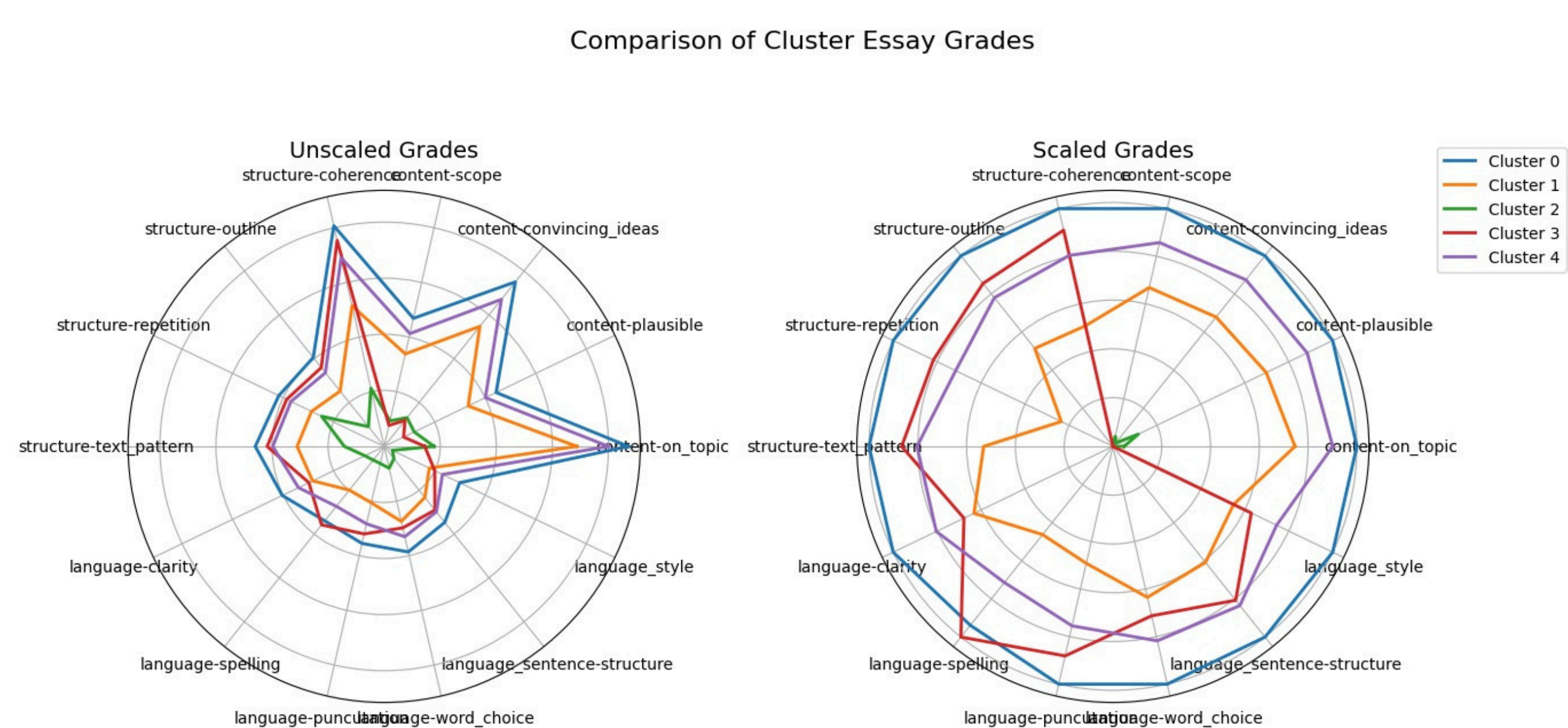**Figure 1:** RMSE and AUC across AI Usage Groups and Math Skills

Figure 2: Mean Score across Essay Skills by Cluster, Unscaled and Scaled
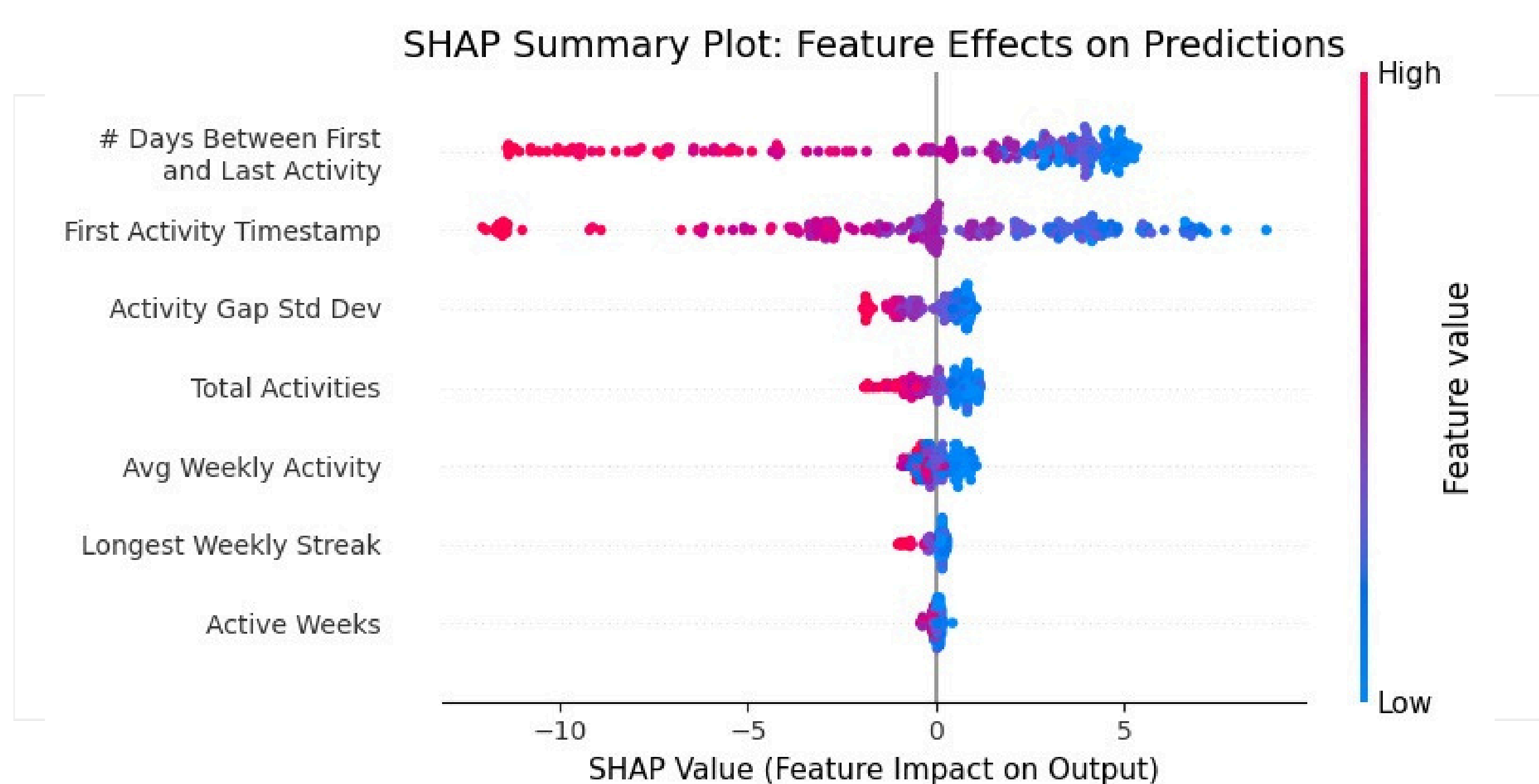

Figure 3: Feature Effects on Predictions on Churn for GoGymi

2 METHODOLOGY

We classified a user's AI usage into one of 3 groups by its count of messages with Gymitrainer, and developed multiple numerical metrics to quantify a user's activity.

For **RQ1** we used BKT to predict a pass or fail for a math quiz by tracking a user's grade evolution over time.

For **RQ2** we clustered students using PCA/K-means for essay scores by skill and measured for each student-cluster average success in mathematics and text quizzes for students of the cluster.

For **RQ3** we used XGBoost classification to predict students' churn/drop-out rate defined to circa a month before the exam based on their activity.

4 DISCUSSION ON RESULTS

- **Figure 1** shows the RMSE and AUC across GoGymi usage groups, by math skill. We notice that they are quasi-equal across usage levels, indicating that GoGymi is, as far as we have data, not correlated to better math grades over time. That is, students are not improving their math skills from using Gymitrainer.
- We notice on **Figure 2** from splitting in 4 clusters that each cluster can be ordered to be better or worse than another, across all essay skills. This means that users are in general (i.e, when looking at a small number of clusters) distributed across their overall level at essays rather than their level at an individual skill, indicating low relevance of skills when looking at global essays performance.
- We also see by that students that belong to an outer cluster (respectively an inner cluster) show better (respectively worse) average performance in math and text assignments. We deduce from **Figure 3** that an early GoGymi first- use date is by far our best predictor for an engaged user, that is, a user that has used GoGymi less than approx. a month before the exam date.

5 CONCLUSIONS/KEY TAKEAWAYS

1. Somehow, as far as we have data, Gymitrainer is not helpful to students' math skills over time.

2. Students are well-divided by overall essay level, indicating that they improve(d) these skills simultaneously.

3. Users that registered a long time ago are the most likely by far to use GoGymi to train for their entry exam, indicating trust and belief in the app.