

SIOP 2024 Machine Learning FAQ

Version: Feb 29, 2024

Please email us your questions and we will update the FAQ.

- Sebastian Marin @ marin343@umn.edu
- Ivan Hernandez @ ivanhernandez@vt.edu
- Isaac Thompson @ thompsonisaacb@gmail.com

Table of contents

SIOP 2024 Machine Learning FAQ	1
Table of contents	1
Competition Related Questions	2
How does the competition work?	2
Are there any rules or requirements?	2
Do I need previous ML experience to enter the competition?	3
When will it start?	3
When will it end?	3
When can I sign up?	3
Who are the “winners” & who will get to present at the conference?	3
Who can I share the competition data with?	3
Can I add people to my team after I sign up?	4
Can team names be changed?	4
Can I remain anonymous?	4
Additional resources	4

Competition Related Questions

How does the competition work?

The competition is set up having a *training* dataset, a *development* dataset, and a *test* dataset. The *training* dataset is used to train models and the *development* dataset is used to adjust these models towards a generalizable solution. The *training* and *development* datasets were the basis for a public leaderboard shown throughout the entire competition, where all team rankings are shown. Near the end of the competition, the *test* dataset will be released for finalizing a generalizable model. This dataset is the basis of a private leaderboard, where a team's ranking is hidden from other teams. This leaderboard determines a team's best and final solution. After the submissions are closed, teams with the highest scores are the competition winners.

This year's competition focuses on developing Large Language Models (LLMs). Competitors will use LLMs of their choosing and design prompts to accomplish various tasks set out by benchmark datasets.

Given the advanced nature of LLMs, this year's competition will have slight changes. We will be providing minimal *training* and *development* data as compared to past competitions. However, the competition will keep its two phases, namely the development phase and the test phase. During the development ("dev") phase, teams will use the *training* and *development* datasets to refine their LLM techniques. Feedback and team rankings will be available on a public leaderboard during the dev phase. During the test phase, teams will tackle the *evaluation* dataset, featuring unseen cases to test the final performance for chosen LLMs.

Are there any rules or requirements?

The primary requirement is straightforward: if you rank as a top scorer in the final phase of the competition, you must submit **solutions that include LLMs** and **these solutions must be reproducible**. Your submission will qualify for consideration as a winning entry as long as it can be executed by the organizing committee and yields largely reproducible results. You will not qualify for consideration as a winning team otherwise. It is essential that the organizing committee can successfully execute your submission and achieve similar results.

Some things you might consider are:

- **Use of LLMs:** Since the use of LLMs is a requirement for *each* sub-challenge, ensure that the submitted work clearly integrates these models. To clarify, we are looking for a final solution where any kind of LLM could be used as long as it has been trained to receive a prompt (soft or hard prompts) and generate text. Predictions have to come from the LLM, but LLMs could be combined in a sequence or mixture, or chained. Check for proper implementation and usage of the LLMs in the solution. You should submit information around:
 - All models used and how they were used
 - Parameter settings for each model used
 - All prompts used (e.g., single prompts, prompt chains, soft and hard prompts)

- **Reproducibility of Results**: The results in your submission should be mostly reproducible. This means that when I run the code according to the provided instructions, anyone should obtain results that are similar to your best and final submission.
- **Code Quality and Documentation**: The code should be well-organized, clearly written, and thoroughly documented. This can include having clear comments, step-by-step instructions explaining how to run the code, a description of any dependencies or external libraries used, environment setup, and so on. Colab is likely the easiest way to share your final solution.

Do I need previous ML experience to enter the competition?

No previous experience is needed! Many of the methods are so new and there are no formal university classes on them yet, at least not in I-O Psychology. This is a learning opportunity for the field as a whole and requires no previous ML, big data, or data science experience.

When will it start?

The competition will begin on February 8th, which is when you will first have access to the *training* and *development* datasets.

When will it end?

The last day for submissions will be April 4th. Final submissions will be scored against the *evaluation* data set.

When can I sign up?

You can sign up any time before the last day of the competition. However, the sooner you sign up, the more time you will have to explore the data and experiment with different solutions.

Who are the “winners” & who will get to present at the conference?

Given the number of teams that have expressed interest in the past and the structure of past competitions, the **top 4 teams** will be invited to present their approach and methods at the competition’s special session at SIOP.

The special session will be held **Thursday, April 18, 2024 @ 9:00 AM - 10:20 AM in Montreux 2 Session C (Loc: Montreux 2/3)**

Who can I share the competition data with?

You can share the competition data with any member of your team. By participating in this competition, you agree to not share data outside your team. The data will be made public once the competition concludes.

Can I add people to my team after I sign up?

Yes, you can add members to your team on the EvalAI competition site. Follow the instructions on the Overview tab.

Can team names be changed?

Team names can be modified on the EvalAI website. Team names or names consisting of clever puns can be fun if you like.

Can I remain anonymous?

Yes, there is no risk to participating and just trying a few solutions out. You can remain anonymous as long as you like (unless you are a winner!).

Additional resources

For further reading and to deepen your understanding of LLMs, consider exploring [Using Large Language Models](#). While not required, this resource can enhance your understanding of LLMs and how to implement them.

Let's have some fun with this!