☰ TASK 3

## Open-ended (graded)

You are in the group 👥 Importer consisting of 👤 bosliu (bosliu@student.ethz.ch (mailto://bosliu@student.ethz.ch)), 👤 hanwan (hanwan@student.ethz.ch (mailto://hanwan@student.ethz.ch)) and 👤 junzhe (junzhe@student.ethz.ch (mailto://junzhe@student.ethz.ch)).

📖 1. READ THE TASK DESCRIPTION

🖥 2. SUBMIT SOLUTIONS

✉ 3. HAND IN FINAL SOLUTION

📖 1. TASK DESCRIPTION

In this task, you will make decisions on food taste similarity based on images and human judgements.

We provide you with a dataset of images of 10.000 dishes, a sample of which is shown below.

We also provide you with a set of triplets (A, B, C) representing human annotations: the human annotator judged that the taste of dish A is more similar to the taste of dish B than to the taste of dish C. A sample of such triplets is shown below.

A   B   C

You task is to predict for unseen triplets (A, B, C) whether dish A is more similar in taste to B or C.

DATA DESCRIPTION

> Download handout (/task3/download)

In the handout for this project, you will find the the following files:

- **food.zip** - archive of the dish images
- **train_triplets.txt** - contains the training triplets. The entries of each triplet denote file names. For example, the triplet "00723 00478 02630" denotes that the dish in image "00723.jpg" is more similar in taste to the dish in image "00478.jpg" than to the dish in image "02630.jpg" according to a human annotator.
- **test_triplets.txt** - the triplets you should make predictions for
- **sample.txt** - a sample submission file

Your task is the following: for each triplet (A, B, C) in **test_triplets.txt** you should predict 0 or 1 as follows:

- 1 if the dish in image A.jpg is closer in taste to the dish in image B.jpg than to the dish in C.jpg
- 0 if the dish in image A.jpg is closer in taste to the dish in image C.jpg than to the dish in B.jpg

---

IMPORTANT

- The order for your predictions for the triplets should agree with the order of the triplets in test_triplets.txt and each prediction should be on a new line.
- You may (and probably should) use pretrained vision models for this task.

EVALUATION

Your submission will be compared to judgements produced by human annotators and the accuracy will be measured.

GRADING

We provide you with **one test set** for which you have to compute predictions. We have partitioned this test set into two parts (of the same size) and use it to compute a *public* and a *private* score for each submission. You only receive feedback about your performance on the public part in the form of the public score, while the private leaderboard remains secret. The purpose of this division is to prevent overfitting to the public score. Your model should generalize well to the private part of the test set. When handing in the task, you need to select which of your submissions will get graded and provide a short description of your approach. This has to be done **individually by each member** of the team. We will then compare your selected submission to our baseline. This project task is graded with either **pass (6.0), partial pass (4.0) or fail (2.0)**. To fully pass the project (grade: 6.0), you need to perform better than the baseline in both private and public score. If you only outperform the baseline in either the private or the public score, you will get a partial pass (grade: 4.0). In addition, for the pass/fail decision, we consider the code and the description of your solution that you submitted. The following **non-binding** guidance provides you with an idea on what is expected to pass the project: If you hand in a properly-written description, your source code is runnable and reproduces your predictions, and your submission performs better than the baseline, you can expect to have passed the assignment.

> ⚠ Make sure that you properly hand in the task, otherwise you may obtain zero points for this task.

FREQUENTLY ASKED QUESTIONS

◉ WHICH PROGRAMMING LANGUAGE AM I SUPPOSED TO USE? WHAT TOOLS AM I ALLOWED TO USE?

You are free to choose any programming language and use any software library. However, **we strongly encourage you to use Python**. You can use publicly available code, but you should specify the source as a comment in your code.

◉ AM I ALLOWED TO USE MODELS THAT WERE NOT TAUGHT IN THE CLASS?

Yes. Nevertheless, the baselines were designed to be solvable based on the material taught in the class up to the second week of each task.

◉ IN WHAT FORMAT SHOULD I SUBMIT THE CODE?

You can submit it as a single file (main.py, etc.; you can compress multiple files into a .zip) having max. size of 1 MB. If you submit a zip, please make sure to name your main file as *main.py* (possibly with other extension corresponding to your chosen programming language).

◉ WILL YOU CHECK / RUN MY CODE?

We will check your code and compare it with other submissions. We also reserve the right to run your code. Please make sure that your code is runnable and your predictions are reproducible (fix the random seeds, etc.). Provide a readme if necessary (e.g., for installing additional libraries).

◉ SHOULD I INCLUDE THE DATA IN THE SUBMISSION?

No. You can assume the data will be available under the path that you specify in your code.

◉ CAN YOU HELP ME SOLVE THE TASK? CAN YOU GIVE ME A HINT?

As the tasks are a graded part of the class, **we cannot help you solve them**. However, feel free to ask general questions about the course material during or after the exercise sessions.

◉ CAN YOU GIVE ME A DEADLINE EXTENSION?

> ⚠ We do not grant any deadline extensions!

◉ CAN I POST ON MOODLE AS SOON AS HAVE A QUESTION?

This is highly discouraged. Remember that collaboration with other teams is prohibited. Instead,

- Read the details of the task thoroughly.
- Review the frequently asked questions.
- If there is another team that solved the task, spend more time thinking.
- Discuss it with your team-mates.

⊙ WHEN WILL I RECEIVE THE PRIVATE SCORES? AND THE PROJECT GRADES?

We will publish the private scores, and corresponding grades before the exam the latest.