

# VISUM 2021 - 9th Edition

July 2021

# Contents

I	FAQ . . . . .	2
II	Initial Setup . . . . .	4
	II.1 Windows Users . . . . .	4
	II.2 Linux and Mac Users . . . . .	4
	II.3 All Users . . . . .	5
III	Opening and Editing Files . . . . .	5
	III.1 Windows Users . . . . .	5
	III.2 Linux and Mac Users . . . . .	6
	III.3 All Users - VS Code . . . . .	6
IV	Submissions . . . . .	6

# I FAQ

Q1. *Should/can I use my own computer to train my algorithms for the competition?*

A1. No, in order to ensure that every team has access to the same computing resources, each team will be assigned its own remote machine (from Google Cloud) with the following specifications:

- 5 vCPUs;
- 15 Gb RAM;
- 60 Gb disc;
- GPU NVIDIA Tesla T4;
- OS Ubuntu 18.04 LTS.

Q2. *How do I access my machine?*

A2. See Section II - “Initial Setup” and Section III - “Opening and Editing Files” below.

Q3. *Do I need to use any specific programming language and/or software to develop my solution?*

A3. Yes, you must use Python 3 to develop your solution. Basic knowledge of Unix shell commands will also be very useful. Although not strictly required, you may find the initial setup more straightforward if your personal computer runs a Linux distribution or Mac OS.

Q4. *Is the required software already installed in my remote machine?*

A4. Yes, CUDA, cuDNN, and Python 3.6 are already installed, as well as many useful Python libraries (e.g., Numpy, OpenCV, Scikit-Learn, TensorFlow, PyTorch, etc.).

Q5. *Can I install further Python packages?*

A5. Yes, you may install whichever packages you want. The recommended way is using pip (see Section III - “Opening and Editing Files” below). Please, do not use virtual environments (e.g., Conda).

Q6. *Can I work via Jupyter notebooks?*

A6. We do not recommend and strongly discourage the usage of Jupyter notebooks. These often lead to memory problems when multiple notebooks are running on the same machine.

Q7. *How do I submit my predictions?*

A7. You do not have to submit anything explicitly. Your code will be run by the competition organizers on the test data in a separate machine, in order to generate your predictions. Please note that you will not have access to the test data at any time during the competition. It is very important that you read Section IV - “Submissions”, where all instructions regarding this subject are provided.

Q8. *When and how is the leaderboard updated?*

A8. The daily leaderboard will be updated daily. Each day, at around 12:00 AM, your `test.py` will be copied (together with all additional files that you may have created and packages you may have installed) to a separate machine and will be run there. Therefore, it does not interfere with any jobs that you might be running in your machine. All groups may have an additional submission each day upon request in the submission’s platform. Note that the daily leaderboard will show results for a fraction of the test data only. The final leaderboard, computed using the whole test set, will be published on Friday, 9th July, in the morning. The leaderboard actualization will be published in our GitHub.

Q9. *Is there any maximum execution time that my inference script should respect?*

A9. Your `test.py` should run in a reasonable time: anything around one hour or less for generating the predictions for the whole training set should be fine.

Q10. *What is the competition deadline?*

A10. The competition is open until **9th July, 2:00AM**.

Q11. *I have read every section in this document and I still have questions. / I am having technical issues connecting to the machine, etc. Can anybody help me?*

A11. Sure, the VISUM project staff will be on Discord to help you. You can also contact the competition staff at any time via email:

- Eduardo – `eduardo.m.castro@inesctec.pt`;
- Wilson – `wilson.j.silva@inesctec.pt`;
- Isabel - `isabel.riotorto@inesctec.pt`;
- Mafalda – `mafalda.falcao@fe.up.pt`;
- Tiago – `tiago.f.goncalves@inesctec.pt`;
- Tomé – `tome.m.albuquerque@inesctec.pt`;
- Leonardo – `leonardo.g.capozzi@inesctec.pt`;

## II Initial Setup

Upon registration in the competition, your team will receive an e-mail with the IP address and the password of your machine. You will need these every time you need to establish a connection to your machine, so please keep them accessible throughout the competition.

You may change the password if you want to. However, please be aware that you and your team members share the same user account, so please share the new password with them. Your user account has no root privileges. Therefore, you will not be able to install new software (at least via apt) in your machine. You may install new Python packages using pip.

There is no graphical interface server running on the remote machines. Therefore, you will not be able to run any program that directly requires a GUI on your remote machine (e.g., image viewers, fancy text editors, etc.). You may, though, use any program running on your local machine to visualize and edit files on the remote machine. This is explained in Section III - “Opening and Editing Files”.

Below, you will find detailed steps on how to establish a connection to the remote machine and start working on the project.

### II.1 Windows Users

1. Unfortunately, in Windows, establishing an SSH connection is not as straightforward as in other OS. However, this can be done using third party applications like PuTTY (<https://www.putty.org/>). This tutorial (<https://mediateemple.net/community/products/dv/204404604/using-ssh-in-putty->) will guide you through all the required steps;
2. On Step 5 of the tutorial, use the provided IP address (e.g., 35.237.233.98) as the hostname, and leave the Port and Connection Type fields with their default values (22 and SSH, respectively);
3. On Step 8, type *visum* as the username;
4. On Step 9, type the password you have received;
5. After you have established the connection, please proceed to Section II.3 - “All Users” below.

### II.2 Linux and Mac Users

1. Open a terminal window;
2. Run the command: `ssh visum@<ip_address>` where <ip\_address> should be replaced with the IP address of your remote machine;

3. If you are asked if you want to proceed with the connection, type yes. Type the password you have received when you are asked for it;
4. After you have established the connection, please proceed to Section II.3 - “All Users” below.

## II.3 All Users

1. Now that you are connected to your remote machine, you are ready to start working on the project! The training set should be located at: `/home/master/dataset`;
2. Your machine will already have useful scripts. They may also be consulted here:  
`https://github.com/visum-summerschool/visum-competition2021`
3. You can run the first steps, shown in the README file of the repository to start working on the project as soon as possible.

## III Opening and Editing Files

As stated before, you may use any software installed in your personal computer to access, edit, and visualize any file stored in the remote machine. Hence, we suggest using SSHFS, which makes it easy to mount a remote directory into a local directory. Alternatively, you can also use Microsoft’s Visual Studio Code with the Remote SSH extension (see section Section III.3 - “All Users - VS Code”).

### III.1 Windows Users

1. Download and install WinFSP. You can download the installer of the latest stable version directly from **here**;
2. Download and install SSHFS-Win. You can download the installer of the latest stable version directly from **here**;
3. Open File Explorer, right-click on “This PC” and choose “Map network drive”;
4. Choose a drive to mount at;
5. In the “Folder” field, enter: `\\sshfs\visum@<ip_address>` replacing `<ip_address>` with the IP address of your remote machine;
6. Select the option “Connect with different credentials” and enter your remote machine’s password when asked for it;
7. Now, your remote machine’s home directory will be mounted at the Windows drive you have chosen. Confirm you have read and write access to the files.

### III.2 Linux and Mac Users

1. Open a terminal window;
2. Install SSHFS, using the appropriate command according to your OS:
  - **Ubuntu/Debian:** `sudo apt install sshfs`
  - **Mac OS:** `brew cask install osxfuse; brew install sshfs`
3. Create an empty directory in your computer, for instance: `mkdir ~/Desktop/visum;`
4. Mount the home directory of your remote machine at the directory you have just created:  
`sshfs visum@<ip.address>:/home/visum ~/Desktop/visum`  
... Replacing <ip.address> with the IP address of your remote machine;
5. Enter the password of your remote machine when asked for it;
6. Now, the home directory of your remote machine will be mounted at the directory you have chosen. Confirm that you have read and write access to the files.

### III.3 All Users - VS Code

Follow the instructions on <https://code.visualstudio.com/docs/remote/ssh>:

1. On step 1 under "Connect to a remote host" type *visum* as the username and the provided IP address (e.g., 35.237.233.98) as the hostname;
2. Type the password you have received when you are asked for it.

Note that all your team members share the same machine and its files, so be careful not to simultaneously access and modify the same file, as this can cause unintended behaviours.

## IV Submissions

**It is very important that you follow all the guidelines in this section. Otherwise, your model will not be evaluated and your submission will fail.**

As mentioned before, you shall not submit your test set predictions directly. Your code and model will be copied and run in a separate machine, together with all remaining files in your home directory (including Python packages installed via pip), to generate those predictions. For this process to run successfully, you must proceed as follows:

1. Implement your model's inference step in a file exactly at this location: `/home/visum/test.py`;
2. Your `test.py` should perform inference on the test data, which is located exactly at this location:  
`/home/master/dataset/test` (in your machine, this data is just the same as the training data);
3. Make sure that your `test.py` outputs a `.csv` file containing the model predictions exactly at this location:  
`/home/visum/preds.csv`;
4. Make sure that your `preds.csv` has the same format as the one generated by the baseline solution. Specifically, it should have a header and one line for each query after that:

```
,productid  
0,<product_id.0>  
1,<product_id.1>
```

You may verify the correctness of your file by running the `evaluate.py` script. For this reason, you should **not** modify the provided `evaluate.py`. Otherwise, you will not be able to confirm if your predictions are being generated correctly.

Note that leaderboards will be generated and provided daily. Such leaderboards are merely informative and calculated using only a small fraction of the test data. The final leaderboard will be obtained using the whole test set and will be the only one that the evaluation panel will take into account to decide the winner.

The winner will be determined by a jury including both VISUM and Farfetch members. The main criterion will be the highest accuracy. However, creativity, novelty, and the ability to communicate ideas will also be considered in the final decision. In this edition, by participating in our competition, one can win a **150€ Farfetch Voucher** and a **free registration for VISUM 2022**.