Tampere University

Project work

COMP.SGN.120

Practical arrangements

- Two options:
  - **Option 1.** (default) consists of a programming task related to binary sound classification problem from scratch.
  - **Option 2.** (exceptional cases). You can propose your own course project work topic, related to audio. Accepted topics should be challenging, and after implementing the project, write a short report.

- Return project by Friday 15th December at 23:59

Research project from scratch

1. Data collection (weeks 1 and 2)
2. Feature extraction (week 3)
3. Define the model and split the data (Train, Validation and Test) (week 4)
4. Train the model and analyse the output (week 5)
5. Write report (week 6)

1. Data collection

Using your phone (annotate model, brand, year, microphone, …) record audio files of at least **5 seconds long**. (5-6s, not longer)

Classes related to vehicles, record as they pass by or idle, **not inside**:

- Tram
- Bus
- Car
- Motorcycle
- Truck

Make sure there is **no speech** recorded!

Choose **2 classes** and Collect **20-30 examples per class**

1. Data collection

Upload your recorded files to freesound.

- freesound.org, create username, upload and describe the files
- choose licence, you keep all rights to the audio.

Provide a list of your uploaded audio files (id, user)

The more you share, the more data you have to train your model!

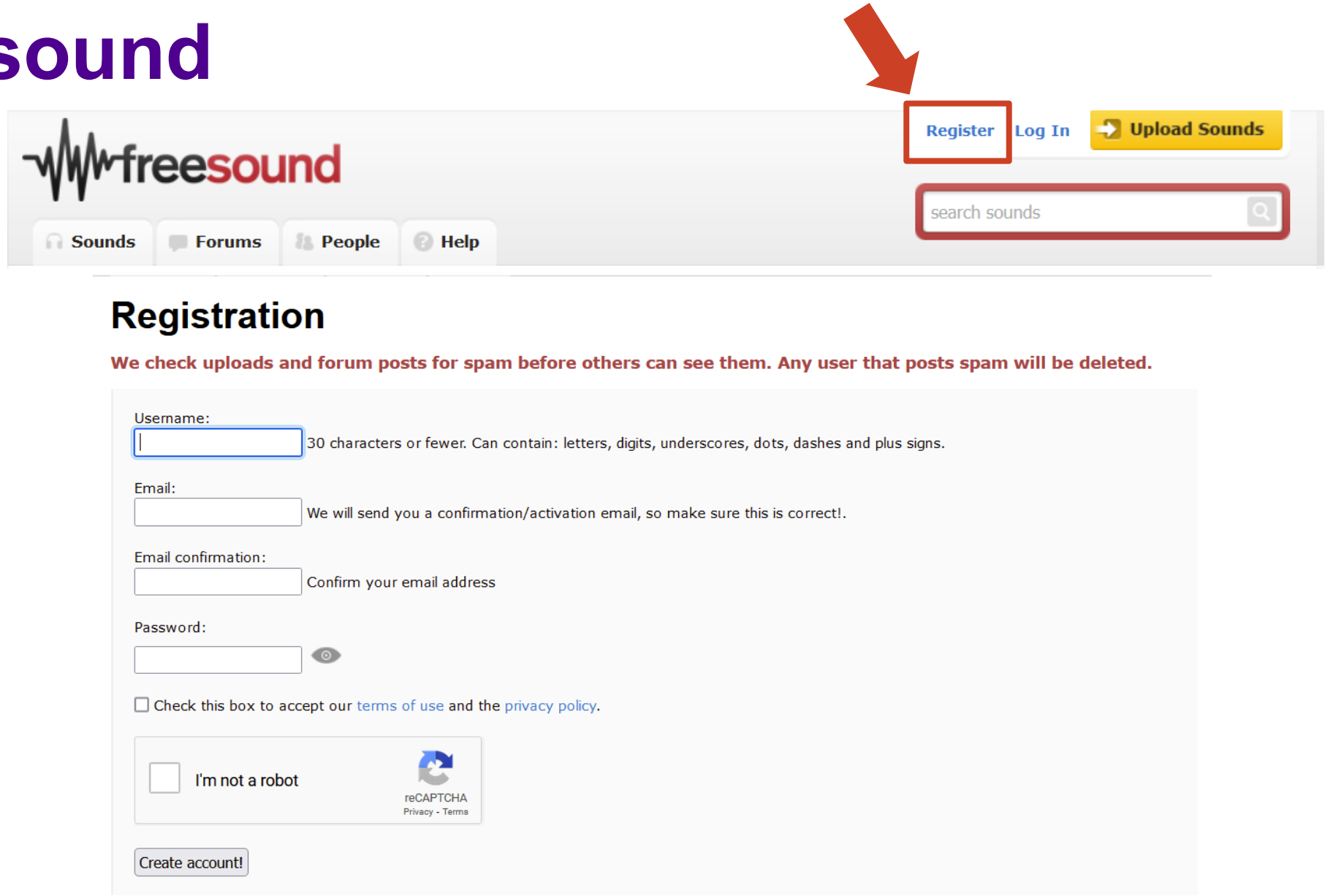
Data sharing

- Why?
  - Sharing is caring.
  - Open science is encouraged within the scientific community.
  - It allows to benefit from current scientific findings.
  - Speed up the transfer of knowledge

- How?
  - Collaborative database **Freesound** platform
  - https://freesound.org/

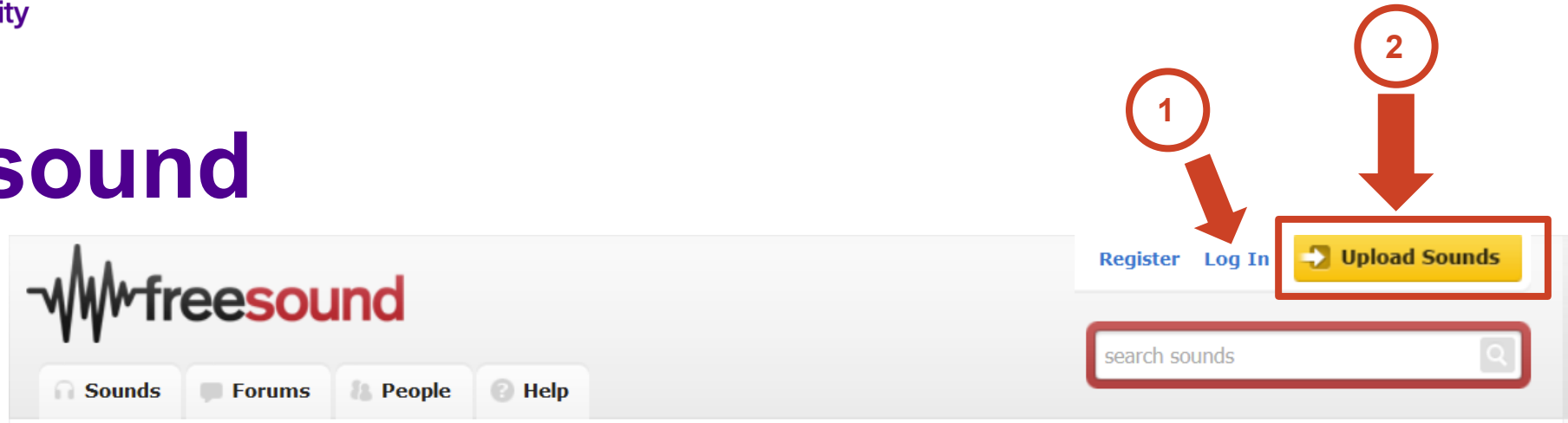
Tampere University

Freesound



Tampere University

Freesound



Please note before uploading..

Legality	Sounds only	Moderation	Formats
Only upload sounds created or recorded **by you**. *More details...*	Freesound is for **sounds** only. Self-created **music fragments/samples/loops** are also ok, but not complete songs or compositions.	All files are **moderated**. If you upload illegal files or files that don't follow the guidelines they will be removed.	We prefer **wav, aif and flac**, but we support ogg, m4a and mp3 too. For very large files please use some compressed format. Uploading more than 1GB at once will fail.

Licenses

- zero (cc0): http://creativecommons.org/publicdomain/zero/1.0/
  a. You can do whatever you want with the sound but you **can't claim authorship**!

     (so other people can do whatever they want with your uploaded sound…)

- attribution (by): http://creativecommons.org/licenses/by/4.0/
  a. Always **mention the original creators** of the sounds.

     (you will always be credited for the sound you upload)

- attribution noncommercial (by-nc): http://creativecommons.org/licenses/by-nc/4.0/
  a. You **can't earn any money** with the piece of work you create

     (others can use your sounds, but cannot sell it in any form)

We recommend to use **creative commons zero**

Describing sounds

Answer the following questions:

- **from where**: the source of the sound.
- **what**: what do I hear, try to describe the sound.
- **where**: if you sampled or recorded this sound, where did you do it?
- **method**: what gear did you use to sample the sound?
- purpose: why did you record this sound? what purpose can it fill?

Example: “car crash sound”

Description *"**Car Crash: breaking of glass, crushing metal on metal, burning rubber**. Hard impact sounds, glass shards falling to the floor. **Recorded with a SoundDevices recorder and stereo microphone (rode NT4)**. The sound was post-processed to make the impacts more dramatic (compression with waves C4). **Recorded on windy afternoon in Melbourne centre for our new movie The Car Crash**. The car we destroyed was a Fiat Panda."*

Dataset for the project

- you will choose **two classes** to work with, and use **all the data shared** by you and your colleagues, accessed through freesound
- each person makes a list of their own uploaded files

- you can also use files with same labels from different freesound users
  - not recorded in Tampere!
  - acoustic mismatch
  - incorrect labels?
  - different recording setup? e.g. more sounds, longer files, unusual environment

    -> may result in worse performance

2. Feature extraction

During the lectures we will study characteristics of the audio signals.

You will have to calculate some of those features from your selected audio files.

1. .wav conversion
2. Normalize
3. Feature extraction

3. Define Model

Simple binary classifier

- Support vector machine (SVM)
- Nearest Neighbor
- Logistic regression

(neural networks if you really really want)

Split the data: **disjoint users** for training, validation, test

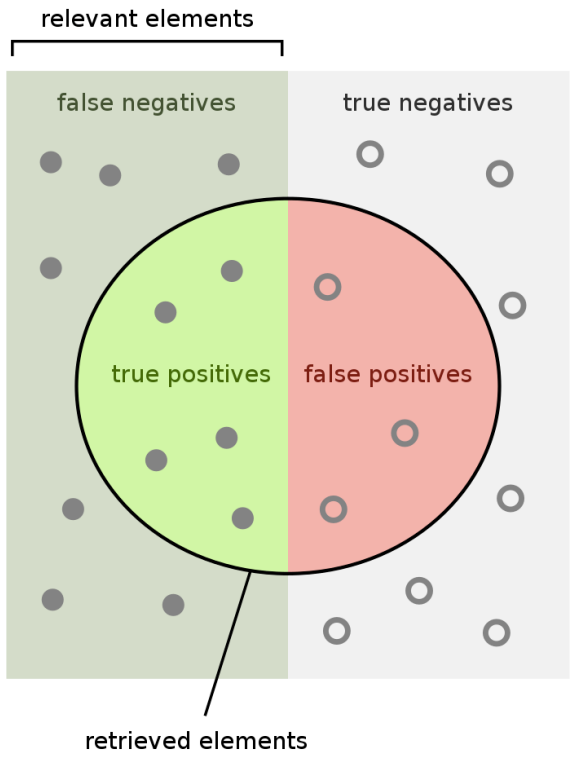
e.g. test data is your own, validation is one other user, train is everyone else

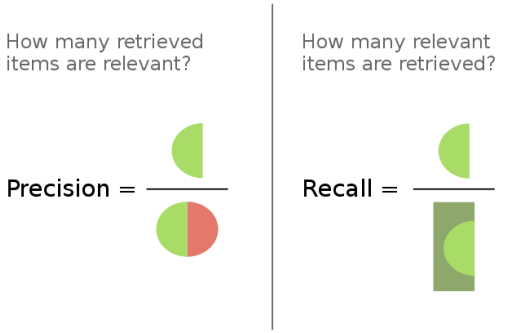
why?

4. Train the model

How to evaluate the model:

- **Accuracy**: ratio of correctly predicted observation to the total observations.
- **Precision**: ratio of correctly predicted positive observations to the total predicted positive observations.
- **Recall**: ratio of correctly predicted positive observations to the all the observations of that class.





5. Write report

The report has to include:

- Information from all the steps, from the data collection to train the model.
- Plots showing statistics of the data and results.
- Challenges encountered and how have been addressed.

Topic of your choice

- Signal processing methods (classical or more advanced ones)
  - Music signals, speech …
  - Something else?

Email to **archontis.politis@tuni.fi** describing the planned project, including:

- Problem to be solved
- Data to be used
- Method to be used
- Evaluation procedure

Instructions

1. Read scientific papers related to the topic.
2. Implement the algorithms needed using Python.
3. Evaluate and interpret the results.
4. Write a report about the results and your observations.

Return

1. Python Script → **project_family_names.py**
   - Return only a part of the test material to check the functionality of the algorithm (one signal) even though the evaluation results should be computed using more data
2. Report → **project_family_names.pdf**
   - **Introduction**: What problem is being solved in the project work? What assumptions were made?
   - **Data description**: how it has been collected? how many classes?
   - **Feature extraction**: which features were selected? Why?
   - **Model selection, data split**: How the problem has been solved? How the data has been split?
   - **Results**: what is evaluated, and how? How the model performed?
   - **Conclusions**: what trouble did you encounter? Why do you think you got those results? How could it be improved?
   - Include figures and diagram whenever descriptive.
   - Mention in the report **how the work was divided** in your group
   - **Length maximum 4 pages** (if exceed, strong and valid justification)

More information in Moodle, **PROJECT WORK** section