

Image Content Analysis on Instagram

Winter Term 2019/2020

2. Meeting

University of Freiburg, Germany

Bernhard Lutz, Tano Müller
Chair for Information Systems Research



UNI
FREIBURG

■ **Schedule**

- Meeting 1: (October 31, 2019 4:15pm, PC Pool 9a, Werthmannstr. 4)
 - Introduction
 - Explanation of Labeling Process
- Meeting 2 (December 3, 2019 4:15pm, PC Pool 1, Werthmannstr. 4):
 - Discussion of unclear background labels
 - Explanation of tasks, guidelines code
- **January 15, 2020** 11:59pm: Submission of model code
- Meeting 3 (January 30, 2020 3:15pm, PC Pool 3, Werthmannstr. 4):
 - Presentations, 10min each
 - Results of model performance
- February 25, 2020 11:59pm: Submission of report (5-10 pages)

- **Grading:** 1/3 model code, 1/3 presentation, 1/3 written report

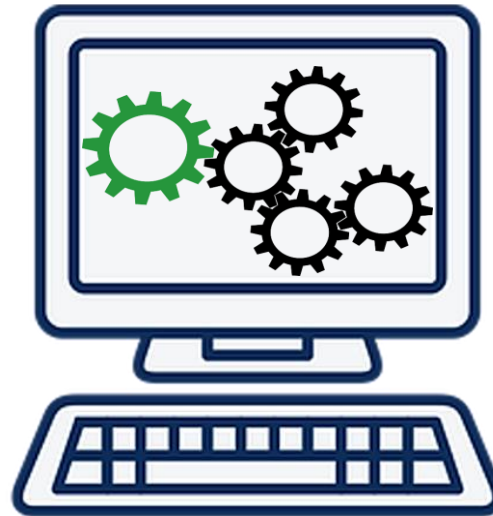
Goal of this seminar: Predict attributes of Instagram images



1: Select an image



2: Model makes
prediction



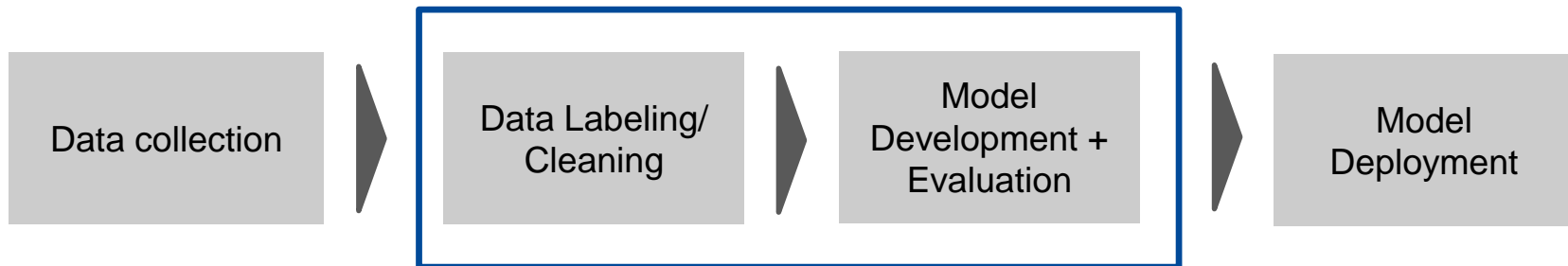
3: Obtain
attributes

- 1 million likes
- no advertisement
- human in focus
- a bit of nudity

Your tasks



- Label images of dataset
- Build a model to **predict content** of images or **number of likes**
- Challenge against the other students
- Present your model
- Write a report



Focus of the seminar

Step 1: Data collection



100K Likes

50K Likes

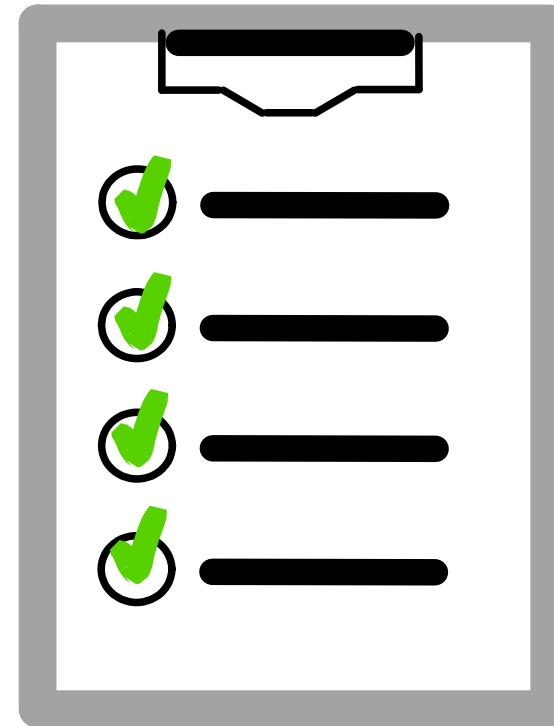
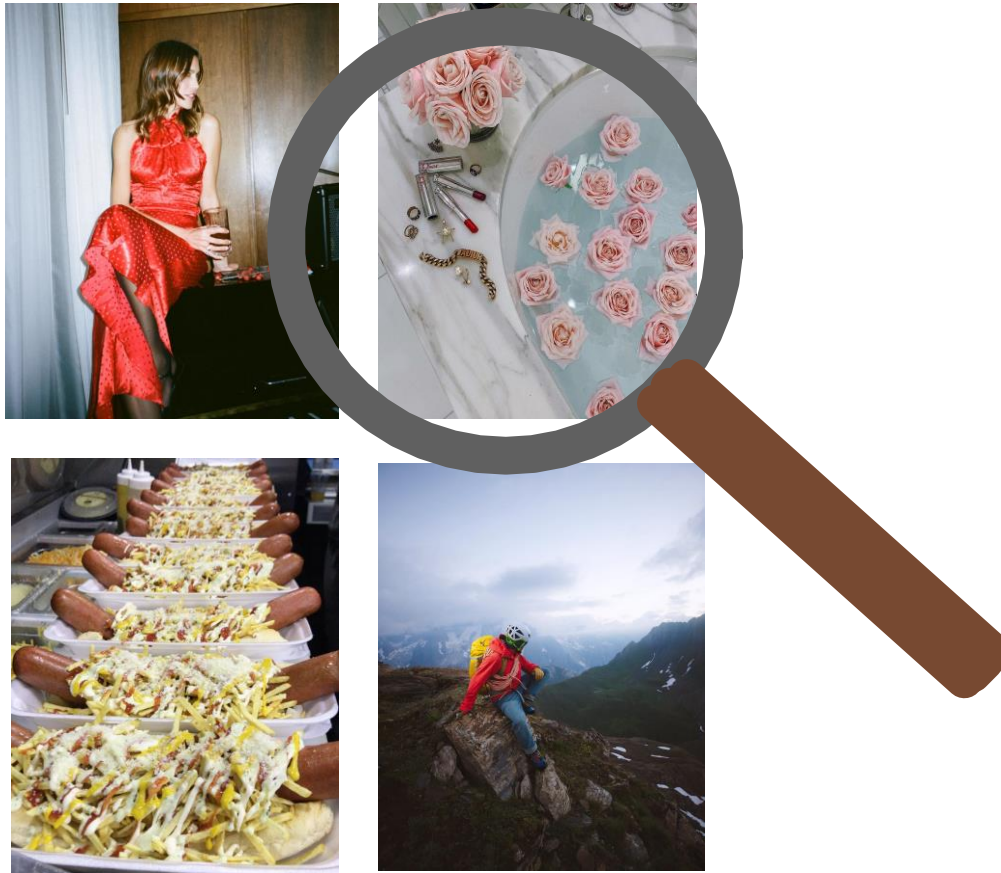
10M Likes

20K Likes

500 Likes

We already did
this

Step 2: Labeling of dataset

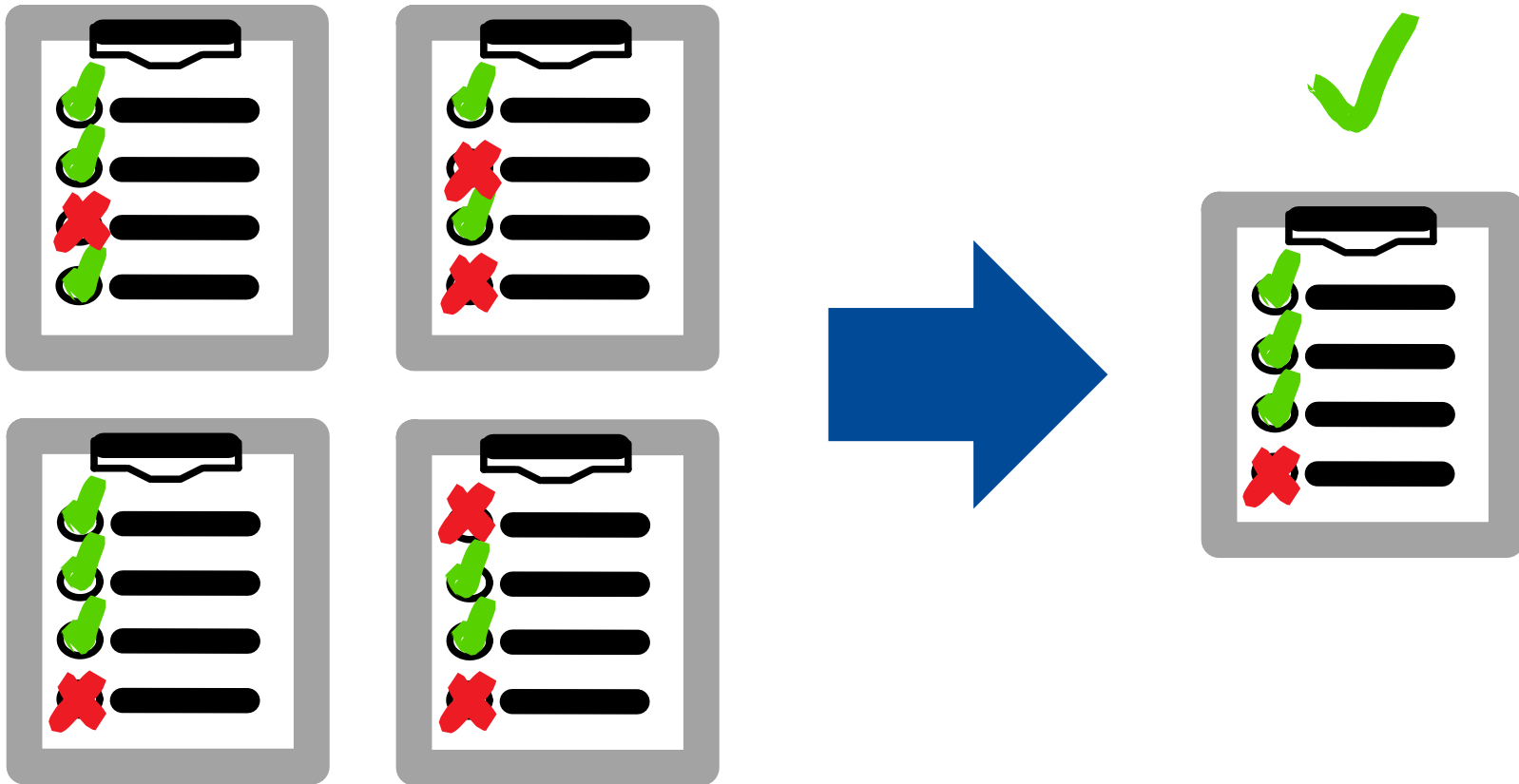


You did this
last month

Step 3: Aggregation of labels



We aggregate the ordinal features by the median, but **this does not work for the background**



- Let us discuss 154 images, where there was no majority for a background...

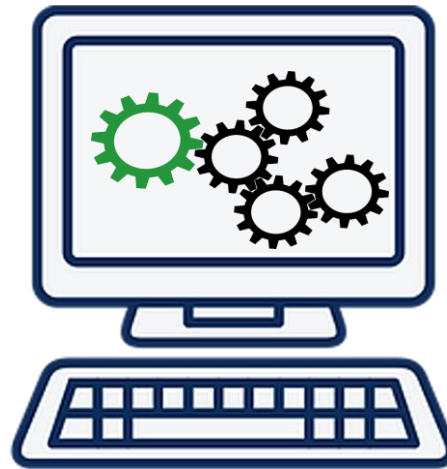
Task A: Predict image content



New image



Model



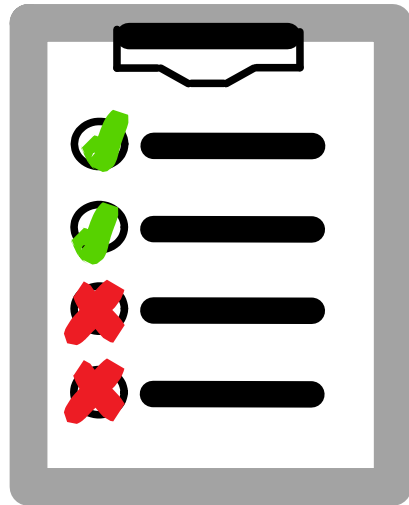
Predict image features

- human in focus
- no advertisement
- no nudity

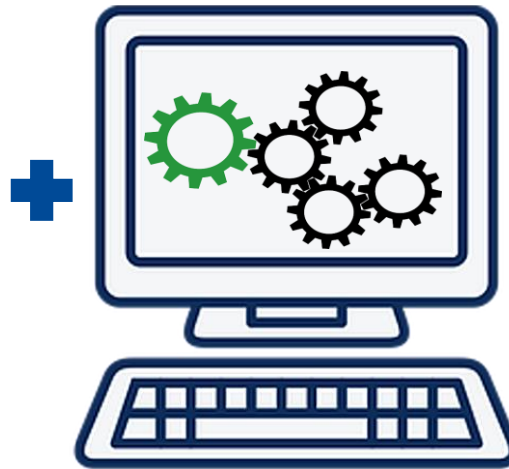
Task B: Predict number of likes



List of features



Model



Predict likes



- **Your code must be organized as follows:**
 - You need one file called Train.R or Train.py that we can call from RStudio or from the command line. This file contains the code to fit and save your model
 - You need to provide another file Evaluate.R or Evaluate.py that loads your model and makes predictions for unseen data
 - We will provide templates containing the files and the code to load the datasets. **DO NOT CHANGE paths or names of functions!**

Next Steps



- We assign task A or task B **according to your skills**
- The task assignment and the datasets will be provided **by the end of the week**
- Submission of model code on January 15, 2020 via email:
 - Task A: tano.mueller@is.uni-freiburg.de
 - Task B: bernhard.lutz@is.uni-freiburg.de