Practical Sessions:

Detailed

Julio C. S. Jacques Junior

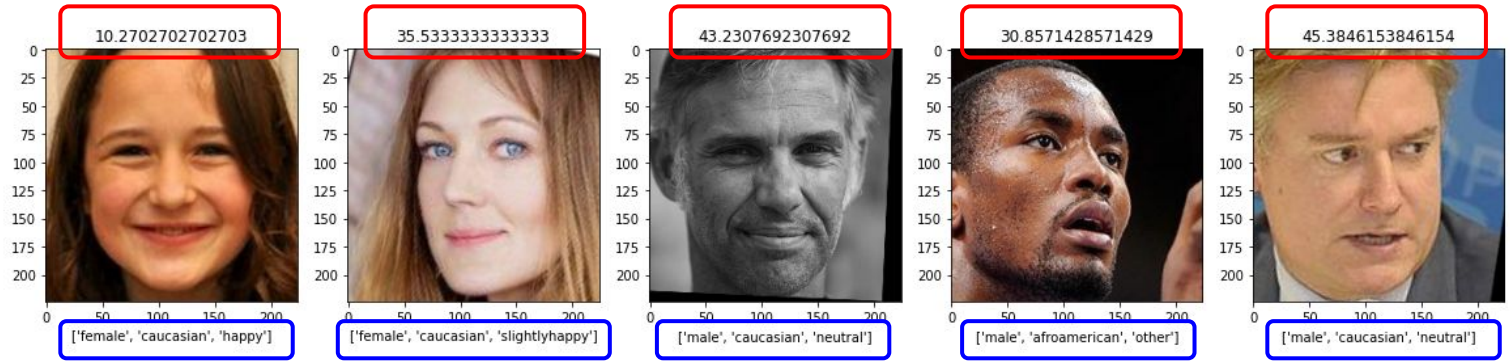
Summary

- The problem
- The database
- Our goal
- The dynamics of the sessions
- Tasks and deliverables
- Evaluation

Problem: Automatic Age Perception

- You will need to solve a **regression** problem
- Given a face image, regress the **perceived age**

Age labels

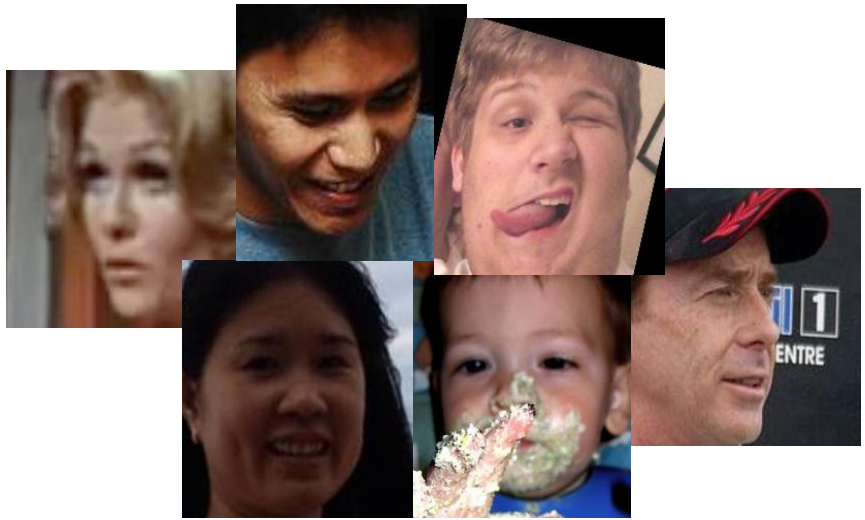


Metadata

Gender (male / female)
Ethnicity (asian / afroamerican / caucasian)
Facial expression (neutral / slightly-happy / happy / other)

Problem: Automatic Age Perception

- **It looks simple but** several challenges are involved
  - Pose variation
  - Different image qualities
  - Different illumination conditions
  - Occlusions, etc



Dataset: Appa-Real Age Dataset

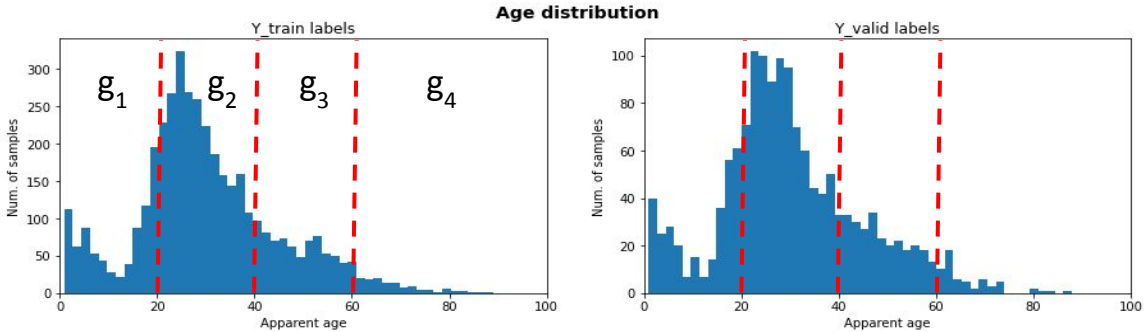
- The data is divided in:
  - **Train** (4065 images),
  - **Validation** (1482 images) and
  - **Test** (1978 images) set

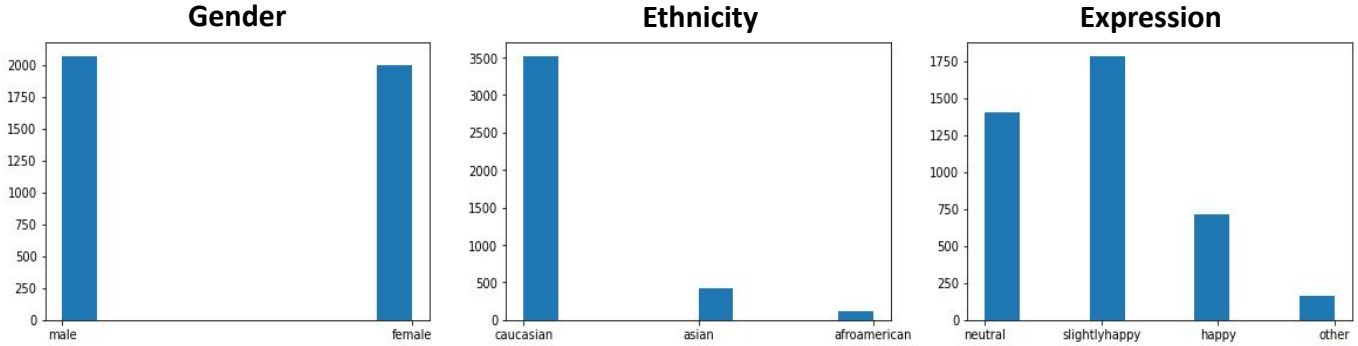
- Matadata is also provided:
  - **Gender**: male / female
  - **Ethnicity**: asian / afroamerican / caucasian
  - **Facial expression**: neutral / slightly-happy / happy / other

http://chalearnlap.cvc.uab.es/challenge/13/track/13/description/

- **Dataset is biased** *w.r.t* different attributes

Training data distribution: Age and Metadata



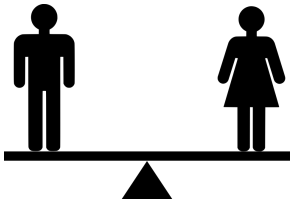


Goal: maximize accuracy & minimize the bias scores

- Low Mean Absolute Error (MAE)
- Low Bias scores
  - **Gender** bias (2 groups)
  - **Age** bias (4 age groups)
  - **Ethnicity** bias (3 groups)
  - **Facial expression** bias (4 groups)

➔ **Bias metric goal:** for each attribute (A), minimize the MAE (M_A) difference among different groups (N)

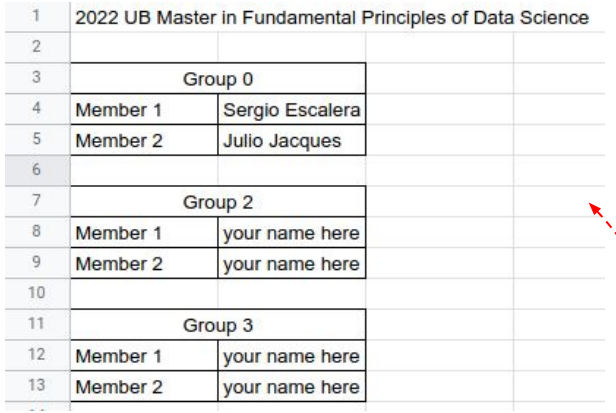
```
for i in range(0,N):
  for j in range(1,N):
    if ( j > i ):
       bias.append( abs( M_A(i) - M_A(j) ))
print("Bias(A) = " np.mean(bias) )
```



Ideally, the method should predict with similar accuracy for all different subgroups

Working in **pairs**

- Our proposal is to work in **groups of 2**
  - Stimulate collaborative work
  - Receive quick feedback



1	2022 UB Master in Fundamental Principles of Data Science	
2		
3	Group 0	
4	Member 1	Sergio Escalera
5	Member 2	Julio Jacques
6		
7	Group 2	
8	Member 1	your name here
9	Member 2	your name here
10		
11	Group 3	
12	Member 1	your name here
13	Member 2	your name here

- **Groups should be defined ASAP, as the Tasks (and deliverables) will be defined by the and of this class;**
  - Please, include the information about your group in the **shared doc** (as illustrated above) → https://docs.google.com/spreadsheets/d/1sIvpdqqCNol2y8oGDg83GDz6L7QgunF0eQCiOjUU8oc/edit?usp=sharing

Dynamics and Details

- After you have defined your group, you can start working on the practical exercises.
- Next, I will comment about
  - The **dynamics** of the practical sessions
  - The **starting-kit**
  - The **metric** used to evaluate the proposed solution
  - The **Tasks**, the additional **(optional) exercise** and **Deliverables**
  - The **way the deliverables will be evaluated**

Workflow

- **Colab** "Starting kit"
- Make predictions
- **Submission file** (.zip)

Submitting your predictions

Receive Real-time Feedback

Colab File + Report:
- Deliverable 1
- Deliverable 2
- Extra

Results

#	User	Entries	Date of Last Entry	Ranking ▲	Gender (bias) ▲	Expression (bias) ▲	Ethnicity (bias) ▲	Age (bias) ▲	MAE ▲
1	[redacted]	1	02/01/21	1.200000	0.293883 (1)	1.027788 (2)	2.130759 (1)	7.460989 (1)	10.975767 (1)
2	[redacted]	1	02/01/21	1.800000	0.628002 (2)	0.819692 (1)	2.447512 (2)	8.988890 (2)	11.141265 (2)

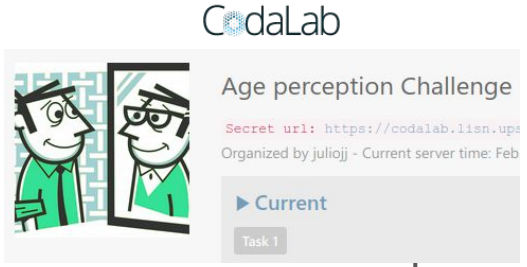
Ranking: *average rank position*

analysed **biases**

"**Accuracy**"

Colab

- **Colab** "Starting kit"
- Make predictions
- **Submission file** (.zip)



Submitting your predictions

Receive Real-time Feedback

Colab File + Report:
- Deliverable 1
- Deliverable 2
- Extra

Results

#	User	Entries	Date of Last Entry	Ranking ▲	Gender (bias) ▲	Expression (bias) ▲	Ethnicity (bias) ▲	Age (bias) ▲	MAE ▲
1	[redacted]	1	02/01/21	1.200000	0.293883 (1)	1.027783 (2)	2.130759 (1)	7.460989 (1)	10.975767 (1)
2	[redacted]	1	02/01/21	1.800000	0.628002 (2)	0.819692 (1)	2.447512 (2)	8.988890 (2)	11.141265 (2)

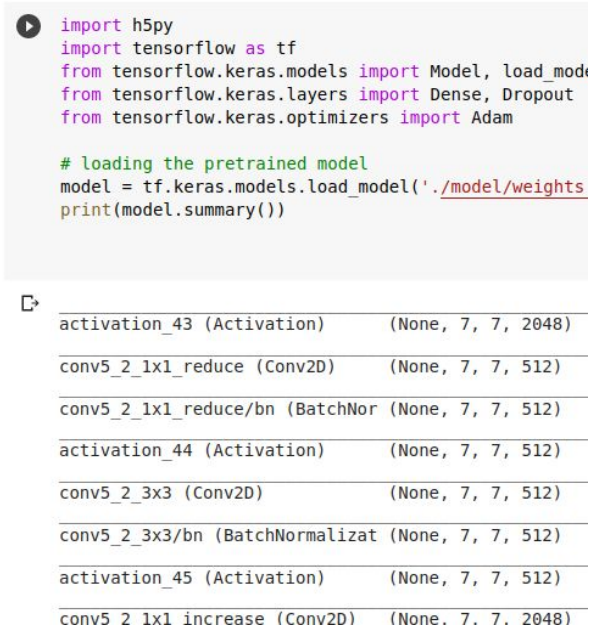
analysed **biases**

"**Accuracy**"

Ranking: *average rank position*

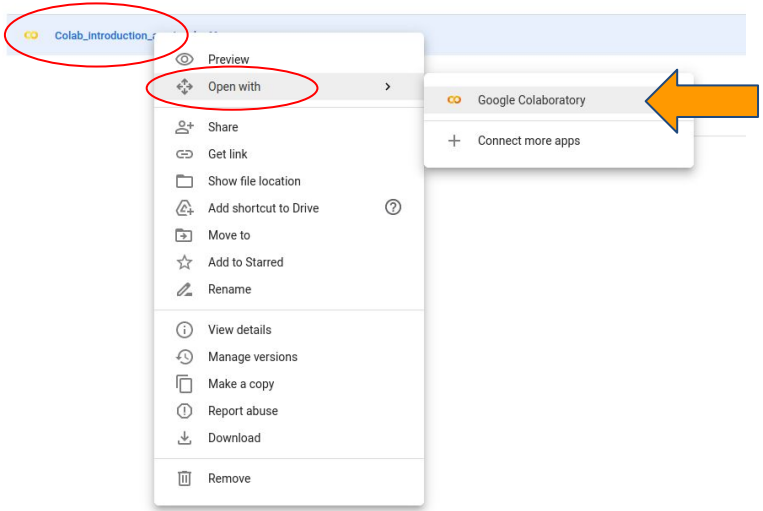
Colab

- Allow you to use CPU/GPU units on the cloud (GPU: *not unlimited*)
- We have prepared a **jupyter notebook** where you can:
  - Get introduced to the problem **progressively**
  - **Download** the data (train/valid/test)
  - **Visualize** the data/metadata
  - Run **baseline** methods (*code available*)
  - **Train / Load** pre-trained models
- Edit / **adapt / improve** the baseline methods



“Hello Colab”

- Upload the provided **“.ipynb”** file to your Drive
- Open the file with **“Google Collaboratory”** colab



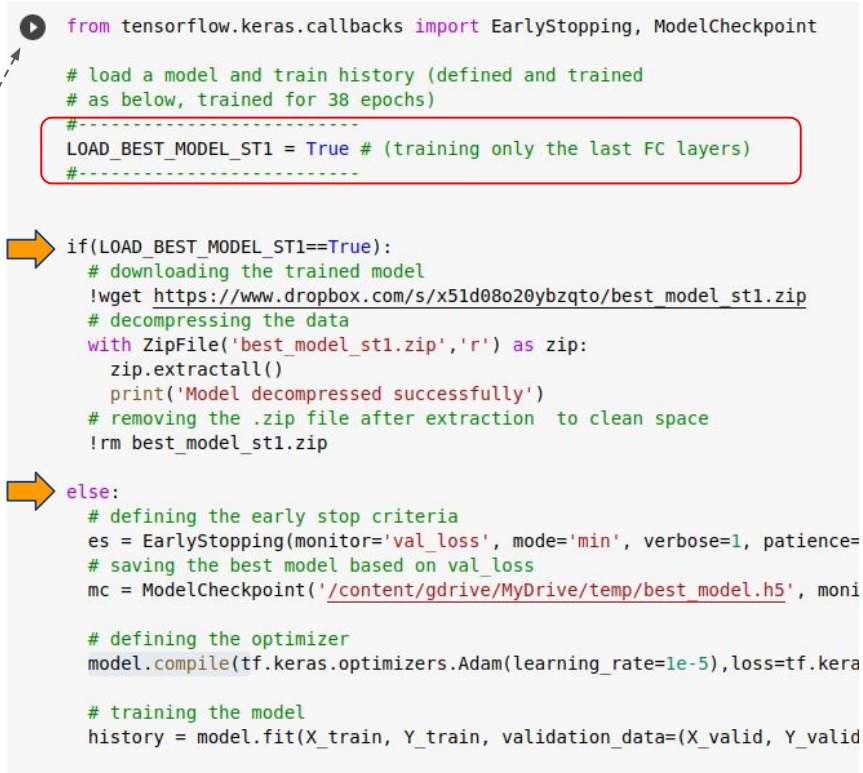
“Hello Colab”

- Data loading
- Visualization
- Modeling
- Training (*stop & continue*)
- Evaluation

Recommendation:

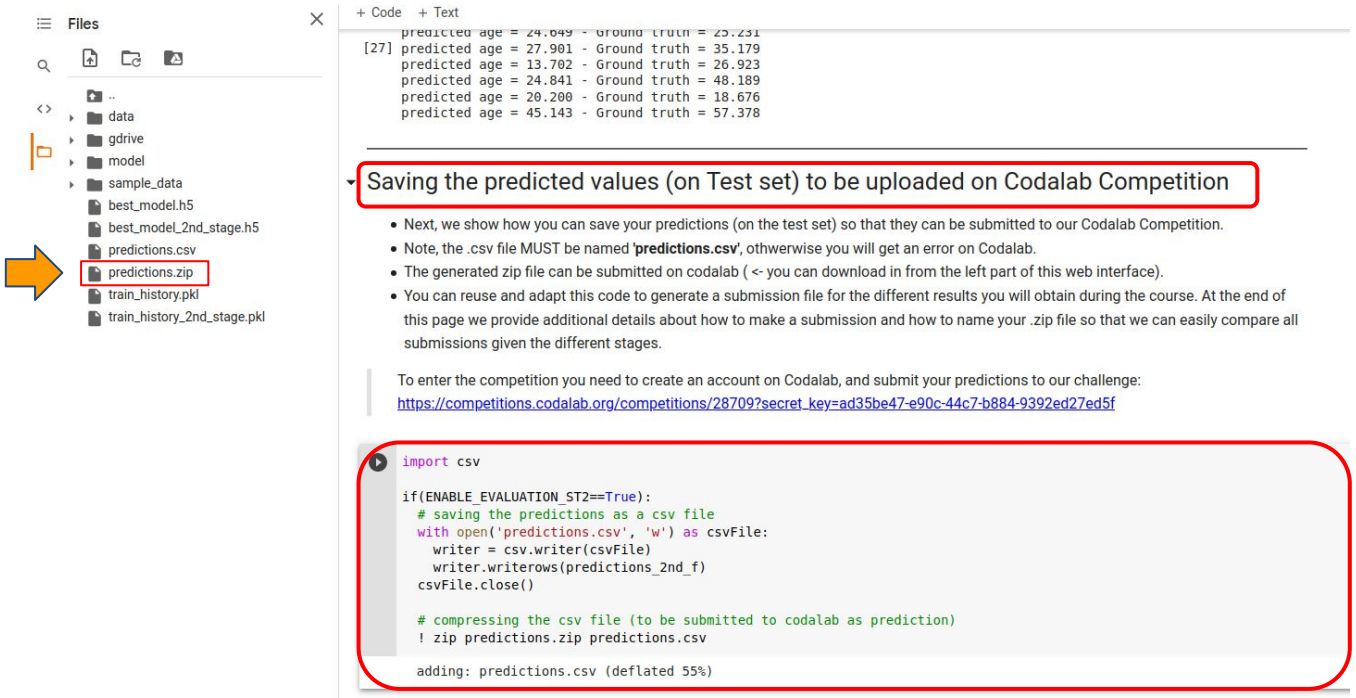
A) Press **“Play”** and get used with everything

B) Edit → Improve



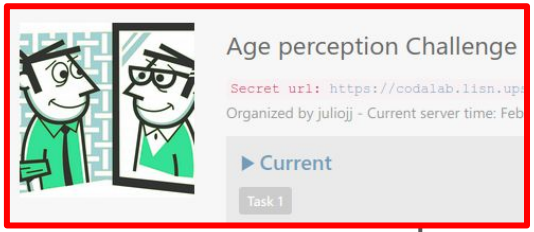
```
from tensorflow.keras.callbacks import EarlyStopping, ModelCheckpoint

# load a model and train history (defined and trained
# as below, trained for 38 epochs)
#--------------------------
LOAD_BEST_MODEL_ST1 = True # (training only the last FC layers)
#--------------------------


if(LOAD_BEST_MODEL_ST1==True):
  # downloading the trained model
  !wget https://www.dropbox.com/s/x51d08o20ybzqto/best_model_st1.zip
  # decompressing the data
  with ZipFile('best_model_st1.zip','r') as zip:
    zip.extractall()
    print('Model decompressed successfully')
  # removing the .zip file after extraction  to clean space
  !rm best_model_st1.zip

else:
  # defining the early stop criteria
  es = EarlyStopping(monitor='val_loss', mode='min', verbose=1, patience=
  # saving the best model based on val_loss
  mc = ModelCheckpoint('/content/gdrive/MyDrive/temp/best_model.h5', moni

  # defining the optimizer
  model.compile(tf.keras.optimizers.Adam(learning_rate=1e-5),loss=tf.kera

  # training the model
  history = model.fit(X_train, Y_train, validation_data=(X_valid, Y_valid
```

Submission file: Colab → Codalab



Workflow

- **Colab** "Starting kit"
- Make predictions
- **Submission file** (.zip)



Submitting your predictions

Receive Real-time Feedback

Colab File + Report:
- Deliverable 1
- Deliverable 2
- Extra

Results

#	User	Entries	Date of Last Entry	Ranking ▲	Gender (bias) ▲	Expression (bias) ▲	Ethnicity (bias) ▲	Age (bias) ▲	MAE ▲
1	[redacted]	1	02/01/21	1.200000	0.293883 (1)	1.027788 (2)	2.130759 (1)	7.460989 (1)	10.975767 (1)
2	[redacted]	1	02/01/21	1.800000	0.628002 (2)	0.819692 (1)	2.447512 (2)	8.988890 (2)	11.141265 (2)

Ranking: *average rank position*

analysed **biases**

"**Accuracy**"

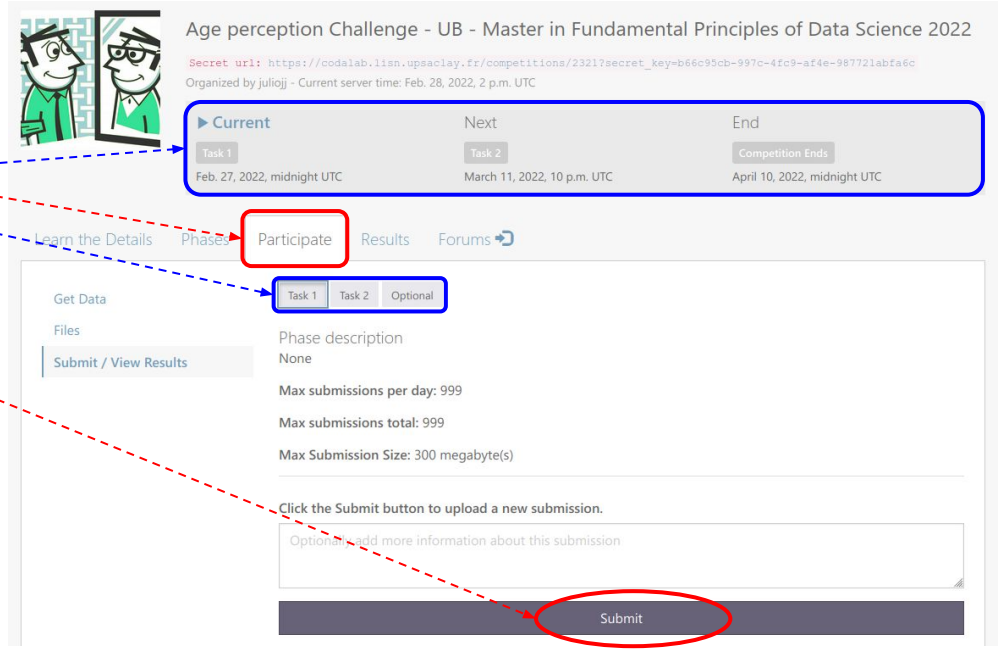
Codalab: **main goal → motivation**

1. **Motivate you** to improve your method and results
    a. Compared to your previous submissions
    b. Compared to your colleagues
2. Simulate a real scenario in research (to **motivate you**)
3. Have fun while learning new skills (to **motivate you**)

IMPORTANT: The Rank position on Codalab **won't be considered** for the evaluation!

Codalab: **Submitting your results**

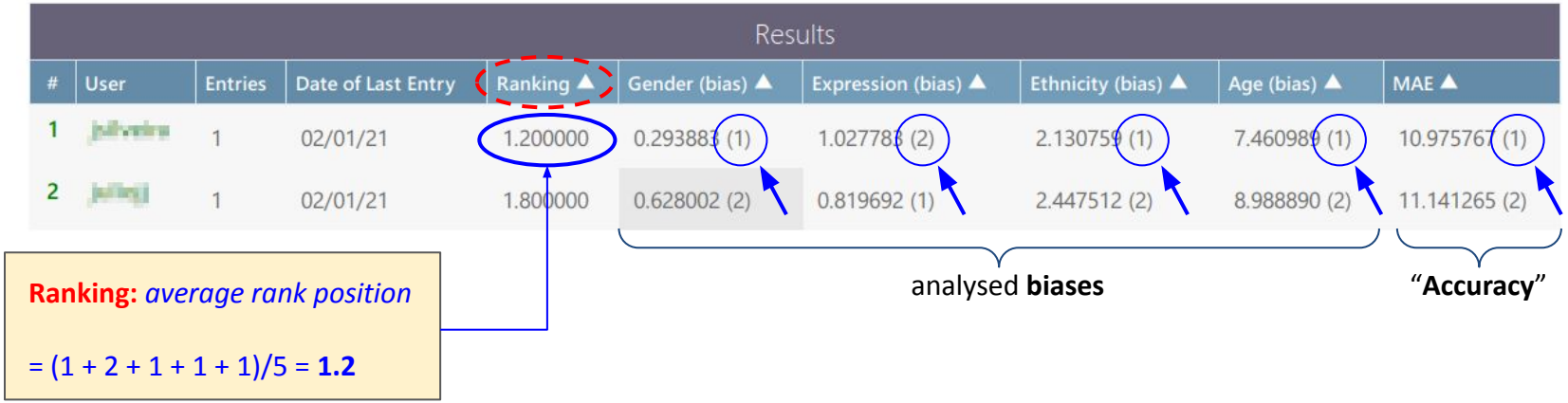
1. **Register** on Codalab: https://codalab.lisn.upsaclay.fr
2. Access our Challenge
3. **Participate**
   (*submission **deadlines***)
4. **Submit** your file



Challenge link:

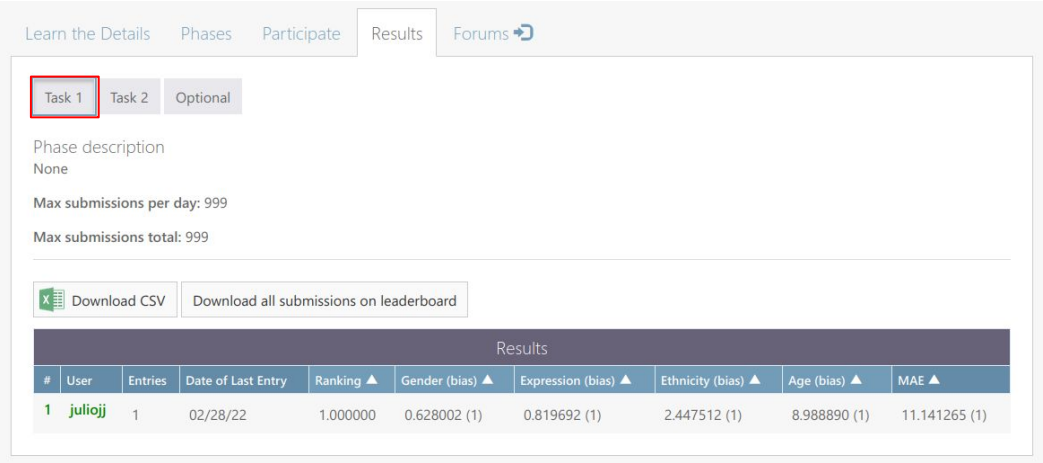
https://codalab.lisn.upsaclay.fr/competitions/2321?secret_key=b66c95cb-997c-4fc9-af4e-987721abfa6c

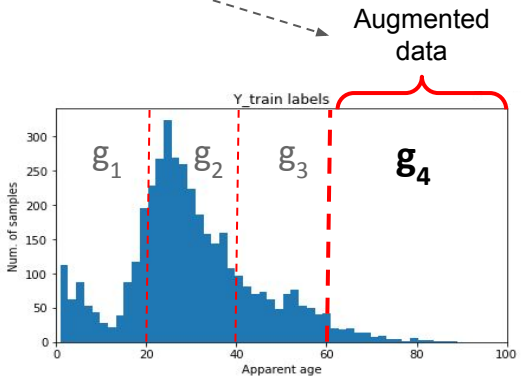
Codalab: **Real-time feedback** on the Leaderboard



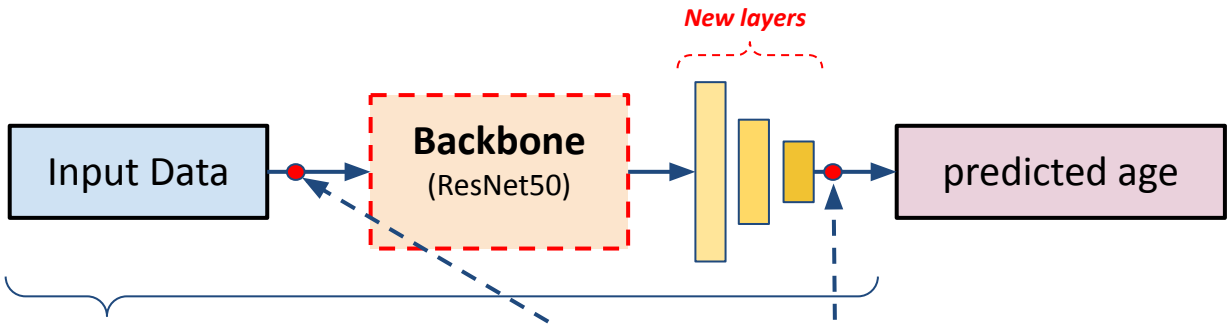
Baselines

- For each phase (Task) of the challenge, you will find different baseline results… they are there to motivate you to improve and beat them :)
  - **Task 1:** *simple method from the starting-kit (with no data augmentation / custom loss)*
  - **Task 2:** *simple method from the starting-kit (basic data augmentation only)*
  - **Extra:** *winning approach from a past curse.*





Starting kit: **General Working Plan**



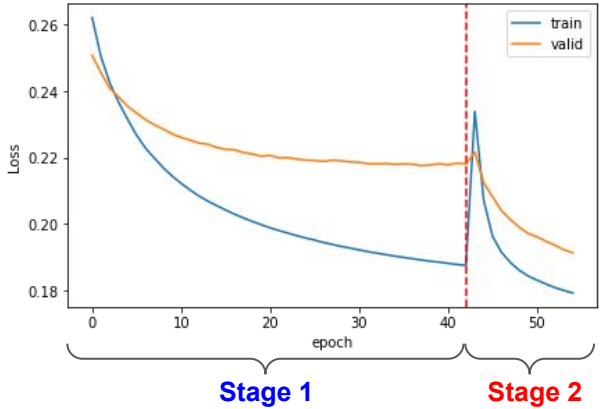
Task 1: play with hyperparameters, model definition, learning strategy, and **with data augmentation**.

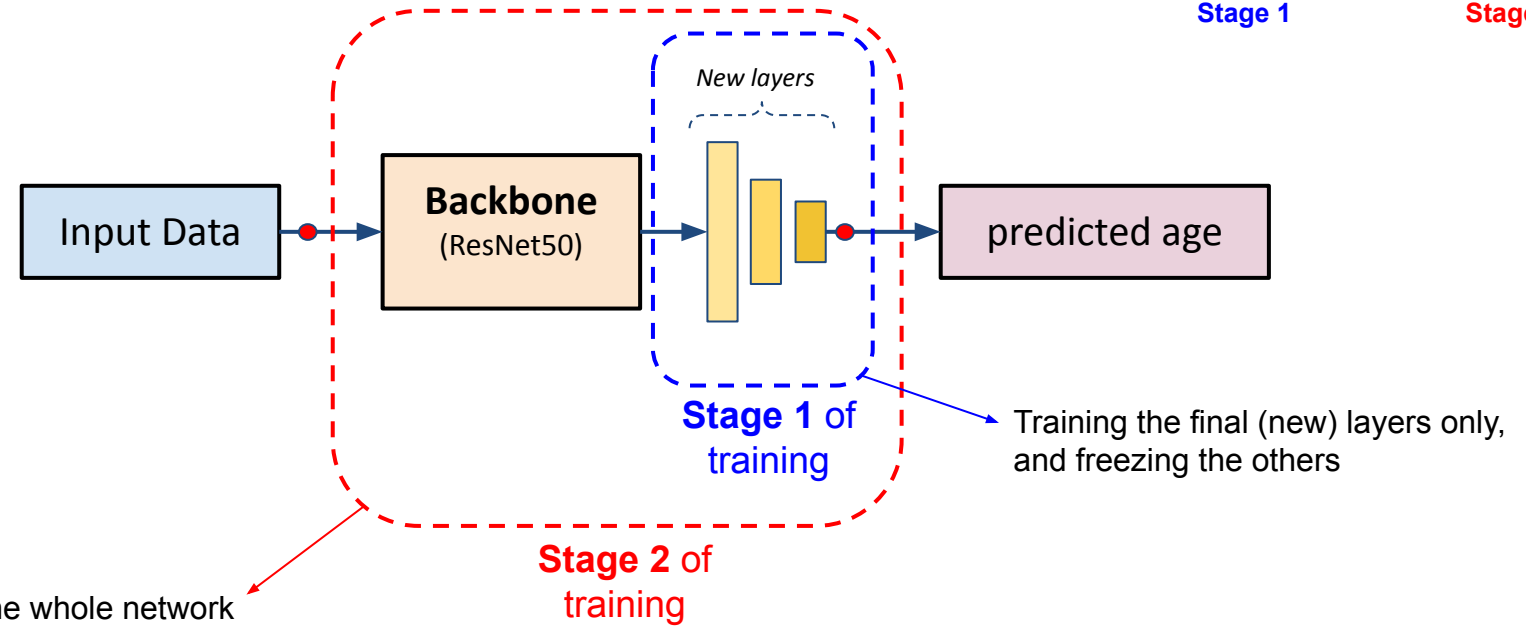
Task 2: fix the previous model and play with hyperparameters, learning strategy, **custom loss / weighted samples** (**without data augmentation**)

Extra: all together + **"surprise us"**

Starting kit: **Training Strategy**

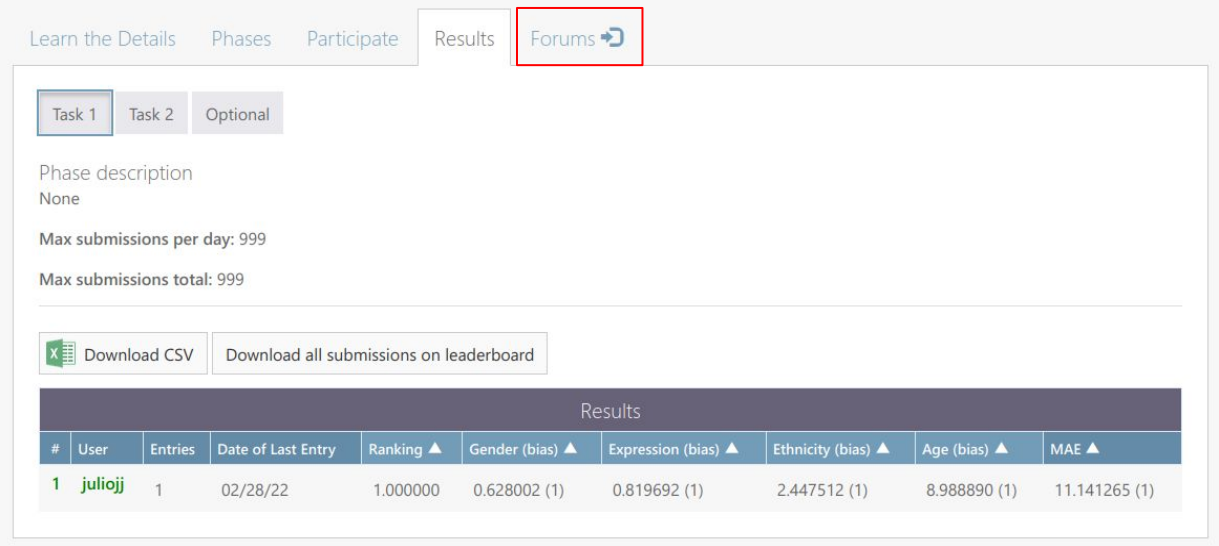
- You are free to employ any training strategy you want





Forums

- You can use the Forum to **exchange experiences**, ask **questions** and report any **problem**.



Workflow

- **Colab** "Starting kit"
- Make predictions
- **Submission file** (.zip)

Submitting your predictions

Receive Real-time Feedback

Colab File + Report:
- Deliverable 1
- Deliverable 2
- Extra

Results

#	User	Entries	Date of Last Entry	Ranking ▲	Gender (bias) ▲	Expression (bias) ▲	Ethnicity (bias) ▲	Age (bias) ▲	MAE ▲
1	[redacted]	1	02/01/21	1.200000	0.293883 (1)	1.027783 (2)	2.130759 (1)	7.460989 (1)	10.975767 (1)
2	[redacted]	1	02/01/21	1.800000	0.628002 (2)	0.819692 (1)	2.447512 (2)	8.988890 (2)	11.141265 (2)

Ranking: *average rank position*

analysed **biases**

"**Accuracy**"

Colab file + Report

- For each deliverable,
  - ***Task 1***,
  - ***Task 2***, and
  - *Extra exercise (optional)*
- **Your group** will have to deliver
  - ***Colab file*** (**well documented** and with **clean code/results**, saved as **.ipynb**)
  - **Report document** (saved as **.pdf**)
- ***How to share these files?***
  - *Zip both files → your_names-Task_ID.zip*
  - *Sent it to juliojj@gmail.com before each deliverable deadline*
    - *Subject: 2022 UB Master / Deliverable Task X (where X can be 1, 2 or Extra)*

Report document Template

HEADER

1. SUMMARY OF CONTRIBUTIONS

2. EXPERIMENTAL SETUP & DISCUSSION OF THE RESULTS

3. FINAL REMARKS

OPT 2 – COMPUTER VISION (2022 / UB)
REPORT **Task 1** (Task 2, Final Project, or Extra exercise)

Group members:
full name (1), <email_1@domain>, Codalab_user_1
full name (2), <email_2@domain>, Codalab_user_2

1. SUMMARY OF CONTRIBUTIONS
Summarize the strategy you followed in this section. Describe the backbone you have selected (with some justification, if possible), the changes you have performed in the model if any, the different models and/or hyperparameters you have evaluated, what kind of attributes (age, gender, ethnicity, etc) you have prioritized during your experiments, etc.

2. EXPERIMENTAL SETUP & DISCUSSION OF THE RESULTS
Describe in detail the different experiments you have defined/performed and present the different results you have obtained. You can use Tables to compare the different results and experiments (e.g., progressively), images or graphs to illustrate whatever you need. Next, we illustrate how to include tables and images in your report.

Table 1: illustrative Table, hypothetically comparing the results obtained for model X and Y.

Model	Learning rate	Training strategy	Gender bias	Expression bias	Ethnicity bias	Age bias	MAE
X	1e-5	1	0.628002	**0.819692**	2.447512	8.988890	11.141265
Y	1e-4	2	**0.293883**	1.027783	**2.130759**	**7.460989**	**10.975767**

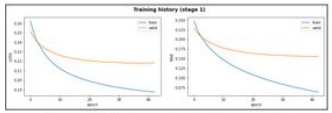


Figure 1: illustrative example of training curve.

Don't forget to discuss the results. For instance, "as it can be seen in Table 1, method Y obtained overall better accuracy and lower bias scores, except for the case of Expressions. This may be explained by the fact that…", or "Figure 1 illustrates the training curves of both stages. As it can be observed…" Don't forget that the main goal is to maximize accuracy and minimize the bias scores.

3. FINAL REMARKS
Draw your final remarks, conclusions and findings. For instance, you can comment why you believed method Y worked better than method X for the problem at hand (and goal), or what you believe could make a difference, as suggestions for future work, etc.

|---- The Report Document is limited to 4 Pages -----

What should be in the Report?

1. SUMMARY OF CONTRIBUTIONS

- Summarize the **strategy you followed** in this section.
- Describe and present the **backbone you have selected,** or your new model if you have implemented a new model from scratch. Justify your decision.
- The **changes you have performed in the backbone model** if any
- The different **models and/or hyperparameters you have evaluated**
- **What kind of attributes** (age, gender, ethnicity, etc) **you have prioritized** during your experiments, etc. Justify your decision.
- **Compared with the starting-kit, what are the main differences of your solution** (+ any possible advantage or disadvantage?)

What should be in the Report?

2. EXPERIMENTAL SETUP & DISCUSSION OF THE RESULTS

- Describe in detail the different **experiments you have defined** and **present the different results you have obtained**. Justify your decisions (e.c., why have you defined such experiments? What was your objective?)
  - For instance, an experiment evaluating the best learning rate with all other parameters and methods fixed; or an experiment comparing method X versus method Y; or any other strategy;
- You can use **Tables** to **illustrate the different results and experiments** (e.g., **progressively**), **Images** or **Graphs**.

What should be in the Report?

2. EXPERIMENTAL SETUP & DISCUSSION OF THE RESULTS

- **Discuss and analyse the results**.
  - For instance, *what could be the reason model Y obtained overall better results?* or why the bias score of attribute X did not decrease as much as the one evaluated for attribute Y? Be creative: **think about interesting experiments, analysis and discussions.**
- Don't forget that the main **goal** is to **maximize** accuracy and **minimize** the bias scores.
- Do not show images and/or tables without discussing or commenting on them.

What should be in the Report?

3. FINAL REMARKS

- **Draw your final remarks, conclusions and findings**.
  - For instance, you can comment *what worked best/worst* in your experiments, and *what you believe could make a difference*, as suggestions for **future work**, etc.

In summary, **we expect** to receive a **clear and good discussion**, with **Tables** and **visualizations** (training curves, augmented data examples, etc) **+ "surprise us"**

Don't forget that **the report document complements the Colab code**. That is, **we also expect a clear and well documented code**, where we can check the defined experiments, models, training strategy and obtained results.

Evaluation

- **Colab File + Report: List of items** and ***achievement levels*** **+ Creativity**
  - Task goal (e.g., data augmentation) will have high weight

		Level of achievement		
✅	Played with **hyperparameters**?	❌ Low	❌ Mid	✅ High
✅	Played with different **backbones** (optional)?	❌ Low	❌ Mid	✅ High
✅	Played with the **layers** of the Net?	❌ Low	❌ Mid	✅ High
✅	Performed data **augmentation**?	❌ Low	❌ Mid	✅ High
✅	**Presentation** of the results	❌ Low	❌ Mid	✅ High
✅	**Analysis/discussion** of the results	❌ Low	❌ Mid	✅ High
✅	...	❌ ...	❌ ...	✅ ...

IMPORTANT: The Rank position on Codalab **won't be considered** for the evaluation!

Schedule

- Mar 1st: **Detailed Practical Class**
- Mar 11th: **Deliver Task 1**
- Mar 18th: **Control Session** → **Feedback Task 1 + Questions about Task 2**
- Apr 1st: **Deliver Task 2**
- Apr 10th: Deliver Extra exercise ***(optional)***
- Apr 29th: Release of (practical) grades

Task objectives

- Task 1 and 2 are already solved in the starting-kit
- We expect you to **go beyond** the provided code
  - Improve and/or acquire new skills
  - Get a better understanding about different problems associated to
    - Computer vision
    - Deep learning
    - Fairness and bias mitigation methods

Task 1: **data augmentation**

- Play with the model architecture (e.g., included/removed some layers) or implement a new model from scratch
- Test and evaluate different backbones and/or losses can be a plus
- Play with hyperparameters (e.g., learning rate, batch size, num. of epochs)
- Play with the training strategy (going beyond the starting kit can be a plus)
- **Intelligent data augmentation**
- Submit the results on codalab
- **Deliver a well documented and clean code**
- **Deliver a clear report with strong analysis and discussion**

Task 2: **custom loss** (without data augmentation)

- Fix the model used in Task 1 (not mandatory but can help when comparing task 1 vs task 2)
- Play with hyperparameters (e.g., learning rate, batch size, num. of epochs)
- Play with the training strategy (going beyond the starting kit can be a plus)
- **Define a custom loss (e.g., weighted samples)**
- Submit the results on codalab
- **Deliver a well documented and clean code**
- **Deliver a clear report with strong analysis and discussion**
    - Here it is expected that you can compare and discuss the results obtained for task 1 vs task 2.

Extra (optional) Task

- **Exploit your creativity as much as you can (“surprise us”)**
- Play with the training strategy / backbones / hyperparameters
- Intelligent data augmentation
- Custom loss
- Submit the results on codalab
- **Deliver a well documented and clean code**
- **Deliver a clear report with strong analysis and discussion**
    - Here it is expected that you can compare and discuss the results obtained for task 1 vs task 2 vs the extra task

Codalab submissions: **best practices**

- People can have different usernames that are difficult to recognize the user, like “***abc2021”***
    - **Inform the codalab username** on deliverables

Colab demo