



Facial Expression Recognition

Milestone 1

Malak Gaballa - 900201683

Masa Tantawy - 900201312

CSCE 4604 - Advanced Machine Learning
Dr. Moustafa Youssef

Table of contents



01

**Proposal
Summary**



02

**Progress
Report**



03

**Timeline &
Member
Contribution**



Proposal Summary



Human communication is a very complex process. **Facial expressions recognition (FER)** from images of human faces has a diverse range of applications including in the fields of robotics or human-computer interaction systems.

Yet, this can be challenging as facial expressions vary tremendously across individuals due to factors such as human age, gender, or race can be obstacles or even other factors such as background, sunglasses, or scarves can hinder the process of accurate recognition of facial expressions.



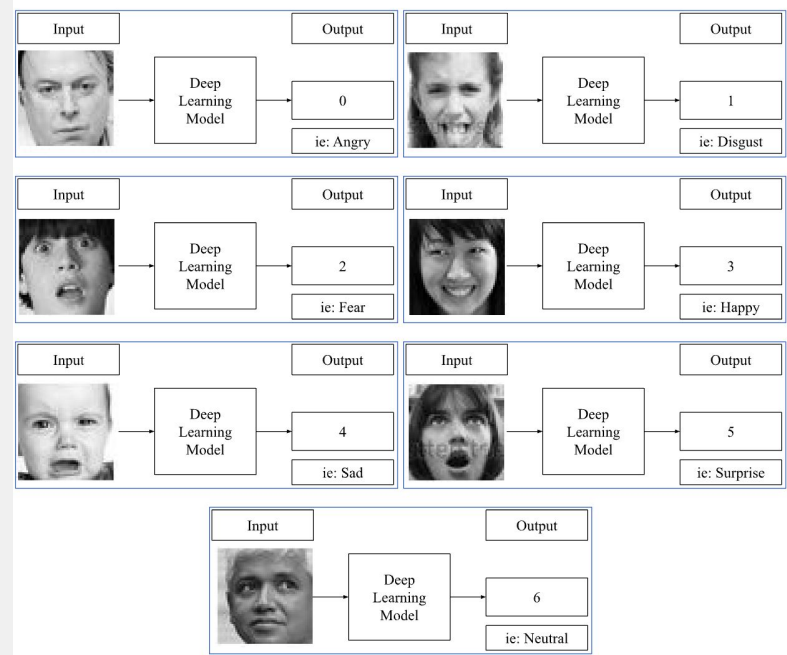
Problem Statement

Given images of human faces showing different expressions, the model should be able to **categorise each image into one of 7 categories, each representing a facial expression**. These are: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral).

Model input: image - vector of pixels for a 48x48 pixel grayscale image,

Model output: A number from 0 to 6 which indicates the facial expression illustrated in the image.

To evaluate the model effectiveness, we opt for the **weighted accuracy metric**, which accounts for class imbalance in the data.

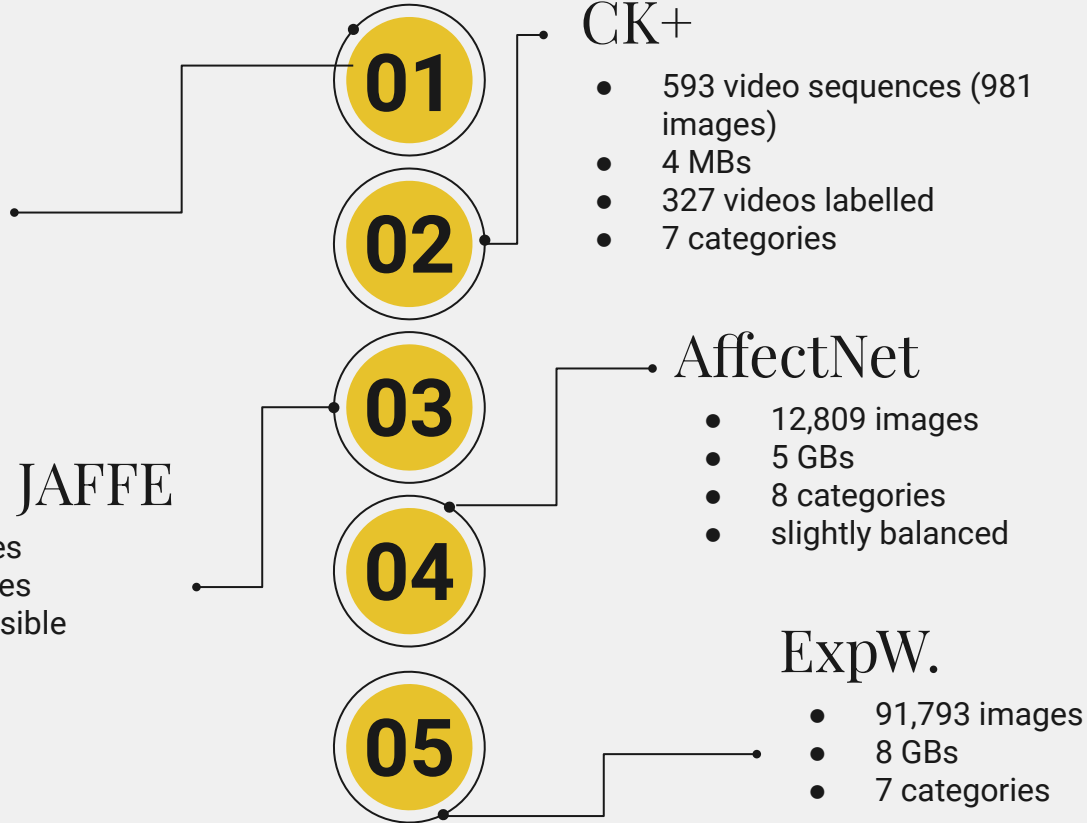


Available Datasets

FER2013

- 35,887 facial grayscale images (48x48 pixels)
- 63 MBs
- 7 categories (**highly imbalanced**)
- Has a test-train split

→ **SELECTED DATASET**



Related Work (SOTA)

Model	Datasets	Accuracy (%)
Ensemble ResMaskingNet with 6 other CNNs	FER2013 and VEMO	76.82
Residual Masking Network	FER2013 and VEMO	74.14
EmoNeXt	FER2013	76.12
Segmentation VGG-19	FER2013 and CK+	75.97
Ensemble of 7 Models	FER2013, CK+ , JAFFE	75.8
Local Learning Deep + BOVW	FER2013	75.42
	FER+	87.76
	AffectNet 8-way classification	59.58
	AffectNet 7-way classification.	63.31

Related Work (SOTA)

Model	Datasets	Accuracy (%)
VGG,Res-Net, and Inception	FER2013	75.2
LHC-Net	FER2013	74.42
VGGNet	FER2013	73.28
CNN Hyperparameter Optimization	FER2013	72.16
Ad-Corre	FER2013, AffectNet , RAF-DB	72.03
DeepEmotion	FER2013, CK+, FERG, JAFFE	70.02

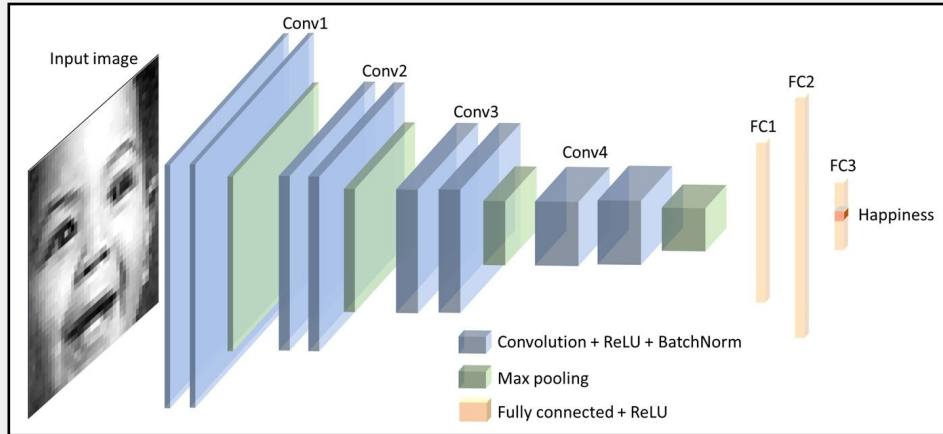
The VGGNet model, short for Visual Geometry Group Network

- *Research Paper:* [Facial Emotion Recognition: State of the Art Performance on FER2013](#)
- *Repository:* [Github link](#)
- *Frameworks:* PyTorch

Baseline Model – VGGNet

A classical convolutional neural network architecture used in large-scale image processing & pattern recognition.

The network consists of 4 convolutional stages and 3 fully connected layers. The convolutional stages are responsible for feature extraction, dimension reduction, and non-linearity. The fully connected layers are trained to classify the inputs as described by extracted features.

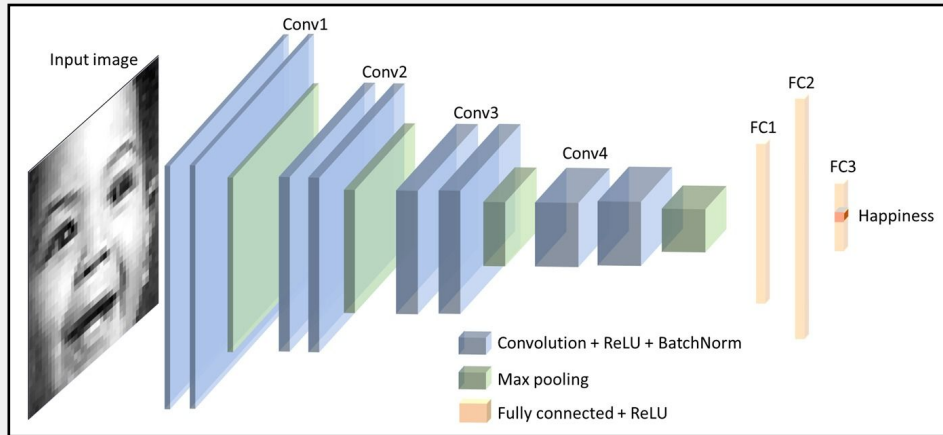


- **Each convolutional stage:** 2 convolutional blocks & a max-pooling layer.
- **Convolution block:** consists of a convolutional layer, a ReLU activation, and a batch normalization layer. Batch normalization is used to speed up the learning process, reduce the internal covariance shift, and prevent gradient vanishing or explosion.

Baseline Model – VGGNet

A classical convolutional neural network architecture used in large-scale image processing & pattern recognition.

The network consists of 4 convolutional stages and 3 fully connected layers. The convolutional stages are responsible for feature extraction, dimension reduction, and non-linearity. The fully connected layers are trained to classify the inputs as described by extracted features.



- The first two fully connected layers are followed by a ReLU activation. The third fully connected layer is for classification.

Proposed Updates



Hyperparameters Tuning

- Number of epochs,
- Regularisation, ...

Data Imbalance Handling

- Oversampling: ROS & SMOTE
- ~~Undersampling~~: RUS & Tomeklinks
- Smote + Tomek & Smote + ENN

Model Ensemble

To enhance model performance and interpretability

Data Augmentation

- Adding auxiliary datasets to train the model
- Image manipulation: such as mirroring/reflecting them, adding background noise, or other appropriate approaches.

Real-time App (TENTATIVE)



To afford high generalisability on the Egyptian/Arab race



Progress Report

On deploying the baseline and the proposed solutions



Milestone 1 Progress



Baseline Model

Trying to deploy the baseline model through multiple approaches

Created a custom function to evaluate the performance of the model



Proposed Solutions

Worked on first 2 elements of the proposed solutions.

HYPERPARAMETERS: Number of Epochs

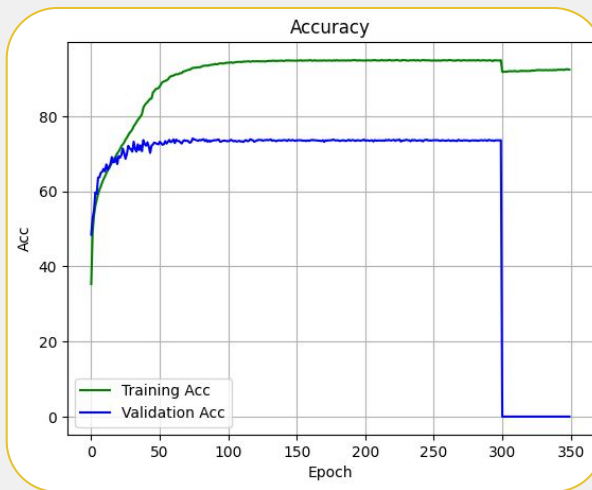
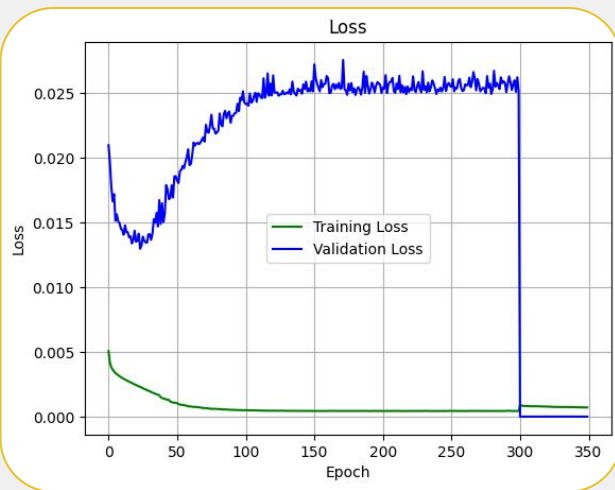
DATA IMBALANCE: created new balanced datasets

```
graph TD; BM[Baseline Model] --> HP[HYPERPARAMETERS: Number of Epochs]; PS[Proposed Solutions] --> DI[DATA IMBALANCE: created new balanced datasets]; PS --> HP;
```

Baseline Model 1

This model is based in **PYTORCH** model but we were unable to modify the hyperparameters.

- Epochs = 350
 - Top-1 Accuracy : 73.27%
 - Top-2 Accuracy : 86.45%



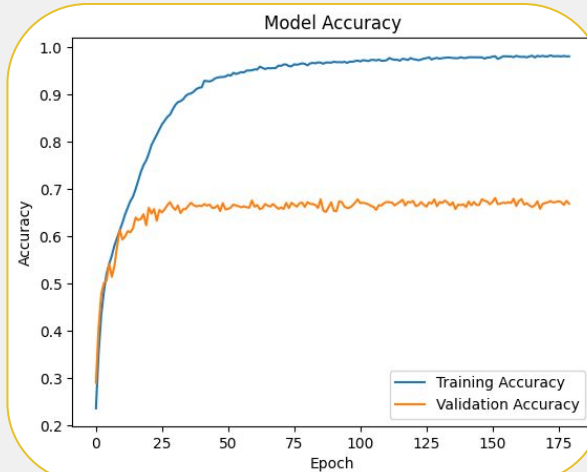
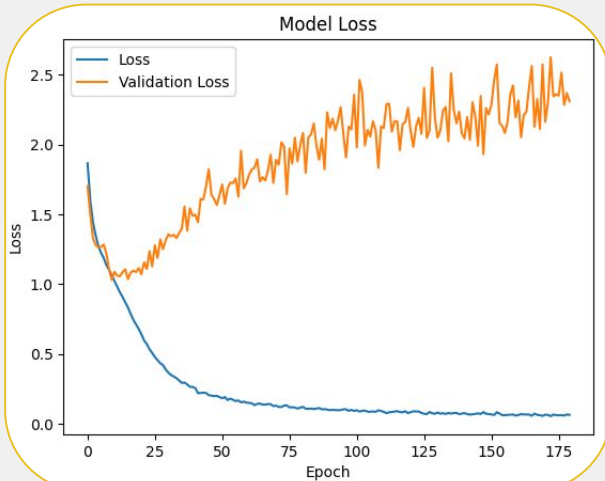
Confusion Matrix:

[[308	3	34	20	49	9	44]
[13	36	2	1	3	1	0]
[38	0	284	11	84	32	47]
[11	0	10	797	12	15	50]
[56	0	54	19	425	8	91]
[7	0	26	16	7	351	8]
[31	0	21	39	78	9	429]]

Baseline Model 2

This is a trained model in **KERAS TENSORFLOW**

- Epochs = 180 instead of 350 due to GPU limit
 - Top-1 Accuracy : 65.76%
 - Top-2 Accuracy : 79.91%
 - Top-3 Accuracy : 88.49%



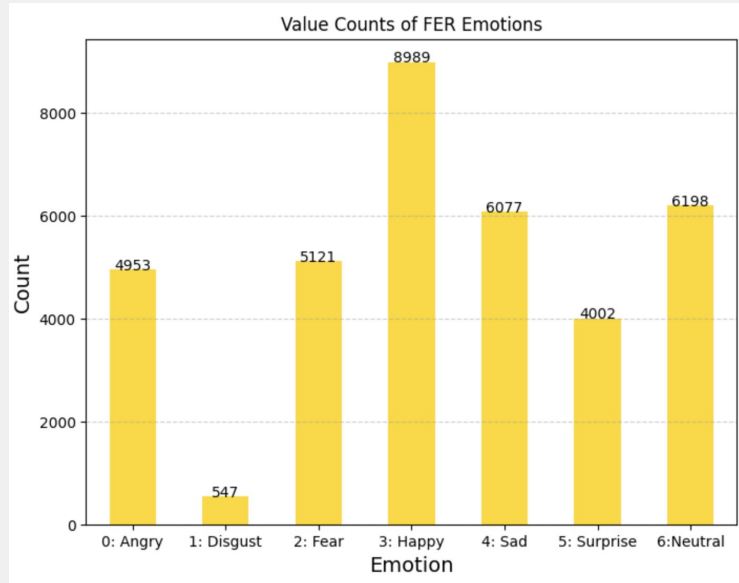
Confusion Matrix

	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
Angry	276	6	53	23	58	5	46
Disgust	14	35	2	0	4	0	1
Fear	45	0	248	10	119	20	54
Happy	19	0	29	736	33	15	63
Sad	66	4	73	26	380	8	96
Surprise	13	0	44	22	10	314	12
Neutral	36	0	42	41	114	3	371

Y-axis: true label
X-axis: predicted label

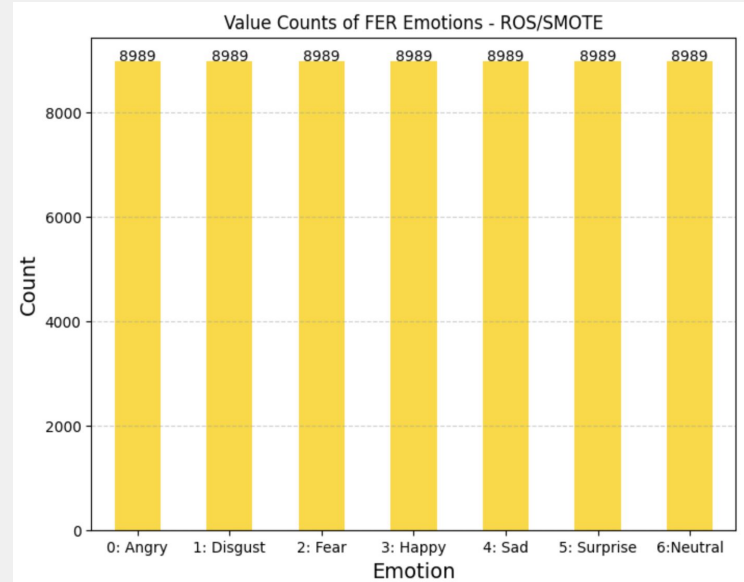
Data imbalance

Original Data



Highly imbalanced
A total of 35,887 images

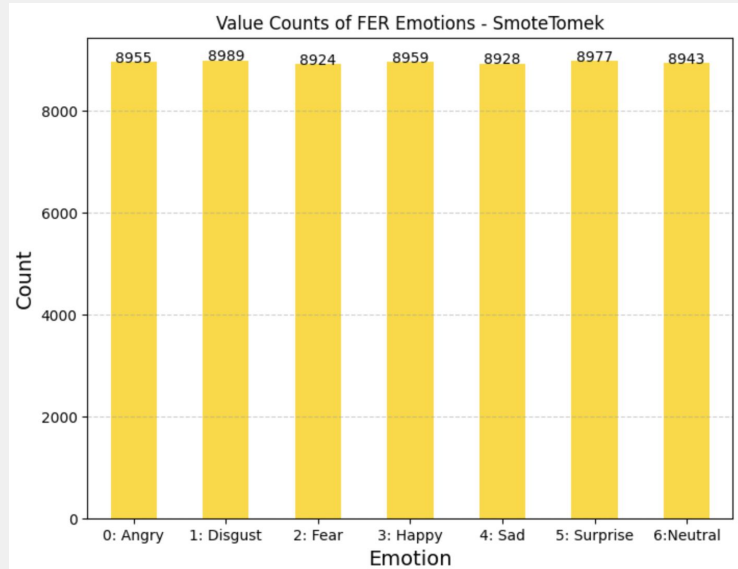
Balanced Data



ROS and SMOTE
A total of 62,923 images

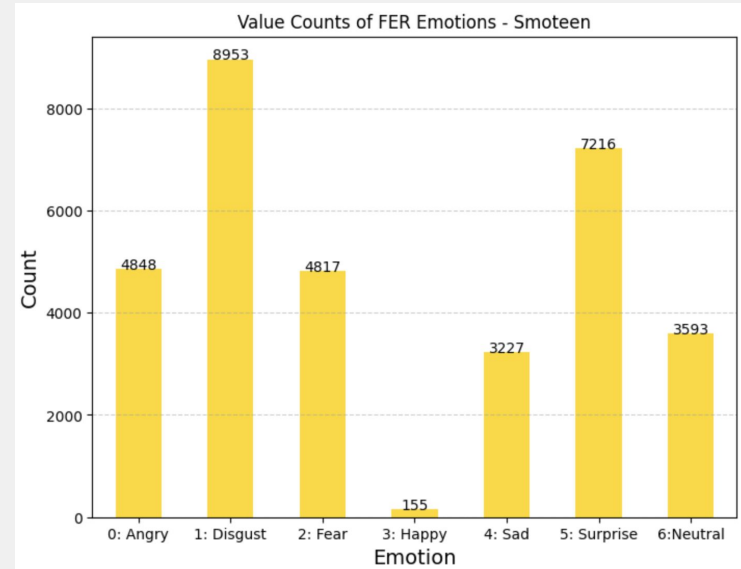
Data imbalance

Balanced Data



SmoteTomek
A total of 62,675 images

Balanced Data



Smoteen
A total of 32,809 images

Data imbalance

Original Data



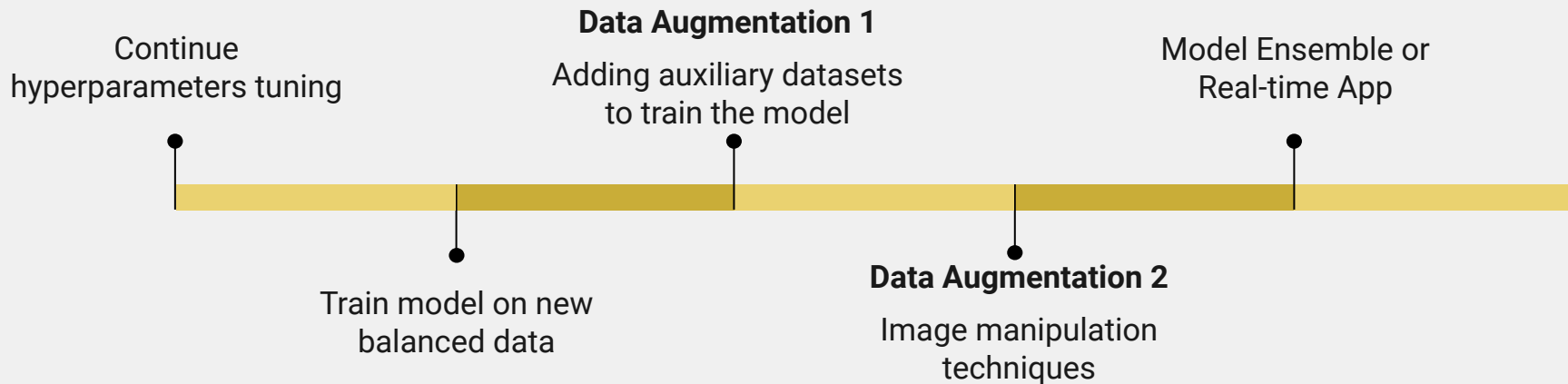
Highly imbalanced
Angry, Angry, Fear, Sad

Balanced Data



SMOTE, SmoteTomek, Smoteen
All neutral

Timeline & Member Contribution



Note: The timeline is tentative and subject to adjustments depending on progress / time availability.

Malak

- Deploying baseline model
- Model Evaluation Function

Masa

- Proposal Summary & Slides
- Data Imbalance Handling



Thanks!

Facial Expression Recognition Milestone 1

Malak Gaballa - 900201683

Masa Tantawy - 900201312



CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**