Task 1

https://github.com/veeeel/Scholarships_Project.git

Task 2 - Business understanding

Identifying business goals

Facial emotion recognition is a rapidly evolving field within computer vision, offering significant value across industries ranging from digital marketing to mental health monitoring. This project focuses on Convolutional Neural Networks (CNN), to classify human facial expressions into distinct emotional categories such as anger, fear, joy, and surprise. Using a diverse dataset of facial images, the project aims to bridge the gap between raw pixel data and meaningful psychological interpretation, addressing the technical challenge of automating human emotional intelligence.

We started this project with the goal of deepening our understanding of image recognition. The primary goal of this project is to develop a machine learning model capable of accurately identifying emotions from images. Beyond simple classification, the project aims to address challenges in computer vision, such as varying head orientations and data noise. Our objectives are to construct CNN architecture that outperforms baseline dense neural networks in classification tasks. Implement data augmentation strategies, to ensure the model performs reliably on data that is not perfect and with at least seventy percent accuracy.

We define project success through both quantitative metrics and reliability. Specifically, the model should achieve validation accuracy exceeding 60-70%, demonstrating predictive capability far better than random chance. Furthermore, it must robustly handle augmented data—such as rotated images—and effectively distinguish between easily confused emotions like fear and surprise, as verified by confusion matrix analysis.

Assessing the situation

As our main resource, we use the “Human Face Emotions” dataset from Kaggle, and we also rely on the tools and knowledge we gained during the course, especially in machine learning and neural networks. For the technical side, we use Python, TensorFlow, Jupyter Notebook, and our personal computers for model training.

Our main requirements are that the model must be able to classify all five emotions, should reach accuracy significantly better than random guessing (at least around 60–70%), and must work on both original and augmented images. Our assumptions are that all photos in the dataset are correctly labeled (no misplaced images) and that TensorFlow can learn the relevant visual patterns with the amount of data and time we have. Our constraints are limited time and the fact that training may be slow depending on the computational power of our personal machines.

Our biggest risk is that the model might not reach the accuracy or speed we expect. To address this, we plan to modify our code and adjust hyperparameters. We will also test training on different computers to see what improves performance. If accuracy remains low, we will apply stronger data augmentation or simplify the model.

The main cost of this project is time. A lot of time was spent searching for a suitable dataset that would be both interesting and useful. Additional time was also spent working with the “Human Face Emotions” dataset to understand the different challenges within it, its limitations, and how we can apply our course knowledge effectively. The main benefit of the project is the opportunity to practice and strengthen our understanding of machine learning and neural networks. At the same time, we are creating a practical tool for human emotion recognition.

Defining data-mining goals

The goals we set for our model is to discover visual patterns in facial images that differentiate emotions. Since emotions are visible as specific changes in the eyes, eyebrows, mouth, we use a CNN that could compositionally learn these features starting from low level patterns building up to the entire expression. As part of the process we want to analyse which emotions does the model differentiate better and which ones it confuses.

We will consider the data-mining successful, if the model can detect meaningful structure in the data and reach an accuracy of at least 75%, showing clearly better performance than random guessing. It should be able to also identify emotions on images, different from the format it learned on, for example better resolution and colored images. The model should show somewhat consistent accuracy across emotions, but at the same time, if it makes confusions that make sense from human perspective, like mixing up fear and surprise, it shouldn't be rated lower because of that.

Task 3 - Data understanding

Gathering data

The foundation of this project is the human-face-emotions dataset, which serves as the source for training and validating our CNN based models. Data acquisition is automated using the Kaggle API. We utilize the kaggle.api.dataset_download_files method to fetch the samithsachidanandan/human-face-emotions dataset directly into our working environment. This ensures reproducibility and simplifies the pipeline for future updates or retraining. To ensure compatibility with our Convolutional Neural Network (CNN) architecture, the data must meet specific technical requirements. Images are processed as grayscale to focus on structural facial features rather than color information. All input images are resized to a standard 48x48 pixel resolution, this dimensionality reduction helps in managing computational resources while retaining essential facial landmarks for recognising emotions, we started off with 28*28 pixel images which lost important facial features. The dataset is publicly available and contains approximately 59 000 files. This volume is verified to be sufficient for training deep learning models without immediate overfitting. The data is split into training and validation subsets, with 80 to 20 ratio in order to allow for proper model evaluation.

Describing data

Our chosen dataset contains around 59,000 images representing five different emotions: surprise, happy, fear, sad, and angry. The images show emotions expressed in different forms — for example, happiness can appear as a wide open laughing face or just a small smile. The dataset also includes some group photos and colorful images, but the majority are black-and-white, single-person portraits. The images are organized into five folders, one for each emotion category. However, the number of images per class varies significantly — for example, the *happy* category contains over 18,000 images, while *fear* has around 9,000. Most files are stored in JPG format, and although there is some variation in resolution, the majority of images are approximately 48×48 pixels.

Exploring data

When looking at the data by categories, there is some imbalance in size, starting at around 8000 images in the smallest class, going up to around 18000 in largest. In order to not have this effect our model's predictions, we consider taking from each class the number of images that is present in the smallest one. In all classes images contain representation of different face angles. While most images have the face zoomed in around the same amount, there are a varying number of broader plan pictures in the classes as well. The class ‘happy’ especially contains multiple images, where the face is not even visible and the emotion is expressed by body and action. In this class are also some images with multiple people on them, which was not noticed in other classes at glance. With the classification of emotions comes also the question of ambiguity, as different people might set a class to some images that differs from how this dataset is constructed.

Planning the project

Make a detailed plan of your project with a list of tasks. There should be at least five tasks. Specify how many hours each team member will contribute to each task.

1. choose dataset
2. Set the goals for classification and analysis
3. verify data quality and preparing the data
4. Select modelling techniques
5. train model / choose the best performing model
6. Evaluate the results and make adjustments
7. Review the process - what could be improved / done differently
8. make poster - choose the appropriate visual representation for our analysis results

List the methods and tools that you plan to use. Add any comments about the tasks that you think are important to clarify.

convolutional neural network
confusion matrix