CM3070 Computer Science Final Project

Final Report

Feelings-Score

Project template **CM3015 Machine Learning and Neural Networks**, titled "**Deep Learning on a Public Dataset"**

By: Deon Robert de Swardt
Student: 200218533

Table of Contents

List of Figures

Vision Statement

The vision of creating the Feeling Index Score is to provide a comprehensive and standardised measure that accurately captures and quantifies individuals' emotional and sentiment responses during video calls. Combining facial emotional recognition and sentiment analysis, the Feeling Index Score offers valuable insights into the intensity and frequency of different emotional states and sentiment categories.

This feelings score will serve as a reliable tool for understanding and evaluating the emotional dynamics of individuals, enabling better communication, empathy, and decision-making in various domains such as mental health, human-computer interaction, and social sciences.

Ultimately, the Feeling Index Score enhances our understanding of human emotions and sentiments in digital communication, promoting more meaningful and impactful interactions.

Introduction

This project aims to build an emotional index score used to ascertain the outcomes of meetings or sessions with online users within the video conferencing space. The scope of the project is to use sentiment analysis combined with emotional prediction to get to a composite score.

The project idea came to light while on a video call interviewing a potential candidate; he was saying all the right things, but his expressions were saying something else sometimes.

The possibility of measuring a candidate and assigning a score based on the interview's success was considered, with the intention of developing interaction ratings.

The latest trend is people working more and more remotely from different locations, which has increased the demand for video conferencing. Only a few visible statistics are being created based on meeting interactions. These meetings can range from professional consultations to employee meetings or even personal video calls.

Through this project, the aim is to create a plug-in for video conferencing tools that offers the aforementioned score, enabling users to assess and measure the performance of their meetings.

The baseline for the product will be established, and the conceptual feasibility of this idea will be demonstrated. This will involve creating a model for emotional classification and sentiment analysis using deep learning models, which will be trained accordingly. The resulting outcomes of these models will be combined to generate the index score.

There is significant motivation for this project, with the primary focus being on the usability of the project/product. The term "project" and "product" will be used interchangeably, considering its ultimate transformation into a product. This project can be used in multiple industries and will significantly benefit if used correctly.

Similar projects and products

There are a couple of similar products that do analyse emotions but do not create a score using sentiment analysis as well.

iMotion - human behaviour research software

iMotions specialises in providing technology and solutions for consumer insights research. iMotions' platform enables researchers to conduct studies using a range of biometric sensors, eye tracking devices, facial expression analysis, and surveys [5]

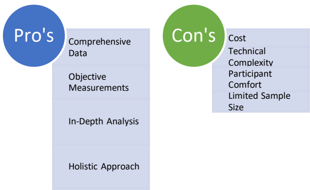


Figure 1 iMotions Pro's and Con's

EmoReact: A Multimodal Approach and Dataset for Recognising Emotional Responses in Children [13]

The research paper introduces the EmoReact dataset, which captures multimodal data to recognise emotional responses in children. The proposed approach combines feature extraction and machine learning techniques to analyse the data and classify emotional responses. The research contributes to emotion recognition in children and has implications for various applications related to emotional understanding and development in this population.

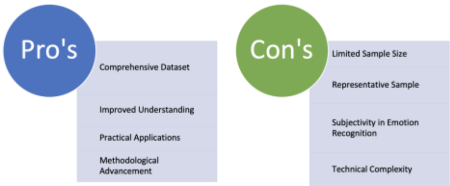


Figure 2 EmoReact Pro's and Con's

NoldusHub: is an all-in-one platform for human behaviour studies designed to combine multiple measurements and easily achieve reliable, high-quality data and direct insights. [14]

The Noldus Observer software allows researchers to capture, record, and analyse behavioural data from different sources. It supports video recording and analysis, enabling researchers to observe and code behaviours systematically and reliably. The software also provides tools for synchronisation with physiological data, eye tracking, and other biometric measurements.

Project Ethics

Commented [RdS1]: Rewrite the below into what is going to be applicable of army project

When conducting a project involving public data, it is crucial to prioritise and uphold ethical considerations throughout the process.
To ensure the project's success, data must be collected from users engaging in video calls, and consent must be obtained to utilise the recorded data from these calls.

Informed Consent:

For this project, obtaining informed consent from all participants is crucial. The Consent Form, provided in **Appendix A**, clearly outlines the project's purpose, procedures, potential risks, and benefits. Participants will have a comprehensive understanding of their rights, including the right to withdraw from the project at any time without facing any negative consequences. [16]

Confidentiality and Anonymity:

Strict measures will be implemented to ensure the confidentiality and privacy of the research group participants. Personal information will be protected. Honouring participants' preferences regarding the use and storage of their data is crucial.

Voluntary Participation:

Participation in the research project will be voluntary, without coercion or pressure. Participants will be free to choose whether or not to take part, and their decision will not result in any negative consequences or discrimination.

Data Protection and Security:

Participants' data will be handled with the utmost care, adhering to relevant data protection regulations. Appropriate security measures will be implemented to prevent unauthorised access, use, or disclosure of sensitive information.

Respect for Diverse Perspectives:

Recognising and respecting the diverse backgrounds, beliefs, and perspectives of the research group participants is crucial. It is essential to represent their views and experiences accurately without bias.

Beneficence and Minimising Harm:

Efforts will be made to maximise potential benefits while minimising potential risks and harm to participants. Throughout the project, participants' well-being and emotional welfare will be monitored, and necessary support or referrals will be provided in case of any adverse effects during and after the video conference call.

Transparency and Openness:

Maintaining transparency throughout the research project is vital. Clear communication regarding the objectives, methods, and expected outcomes will be promptly provided to participants and communicated in the consent form. Any questions or concerns they may have will be addressed promptly.

Literature review

Title: Development of Strategies for Automatic Facial Feature Extraction and Emotion Recognition

Introduction:

Facial feature extraction and emotion recognition have gained significant attention in computer vision and affective computing. This literature review aims to explore the research paper titled "Development of Strategies for Automatic Facial Feature Extraction and Emotion Recognition" by David Restrepo and Alejandro Gómez [15]. The authors present their work on advancing the development of strategies for automatic facial feature extraction and emotion recognition. The review will provide an overview of the paper's objectives, methodologies, and critical findings while highlighting relevant related works.

Summary of the Research Paper

The research paper focused on developing automatic facial feature extraction and emotion recognition strategies. The authors propose a novel approach combining computer vision and machine learning techniques to extract facial features and classify emotions.

The methodology involves face detection, feature extraction, and emotion classification. The authors also address challenges such as variations in facial expressions, lighting conditions, and occlusions.

Facial Feature Extraction Techniques

The paper discusses various facial feature extraction techniques, such as geometric and appearance-based approaches. Geometric-based methods utilise facial landmarks or critical points to represent the face, while appearance-based methods use texture or appearance descriptors. The authors present a comprehensive analysis of these techniques and evaluate their effectiveness in terms of accuracy and computational complexity.

Several related studies have explored facial feature extraction techniques. For example, Zhang et al. [20] proposed a method based on facial landmarks to extract geometric features. Li et al. [11] introduced a hybrid approach combining geometric and appearance-based methods for improved feature extraction.

Emotion Recognition:

The authors explored different emotion recognition approaches, including traditional machine learning methods and deep learning-based approaches. Traditional methods involve extracting handcrafted features and classification using classifiers such as support vector machines (SVM) or random forests. Deep learning approaches, on the other hand, utilise convolutional neural networks (CNN) to learn features directly from raw facial images. The authors compare the performance of these techniques and discuss their advantages and limitations.

Several notable studies have contributed to emotion recognition research. For instance, Jung et al. [6] proposed a deep learning-based approach using facial expression images for emotion classification. Zhao et al. [21] investigated the use of handcrafted features and SVM for emotion recognition from facial images.

Discussion and Analysis:

The literature review section of the paper provides a comprehensive overview of the existing research in the field of automatic facial feature extraction and emotion recognition.

The authors discuss relevant studies contributing to developing facial feature extraction techniques and emotion classification models. They highlight the advancements made in the field and identify areas for further improvement.

Evaluation and Results:

The research paper presented experimental results to evaluate the proposed automatic facial feature extraction and emotion recognition strategies.

The authors conducted experiments on publicly available datasets and reported performance metrics such as accuracy, precision, recall, and F1-score.

The results demonstrate the effectiveness of the proposed approach compared to existing methods, highlighting the potential for practical applications in real-world scenarios.

The authors validated their approach using widely used datasets such as the Cohn-Kanade [12] and the Extended Cohn-Kanade [7] databases.

Critical Analysis:

The authors discuss various techniques for facial feature extraction, including geometric-based and appearance-based methods, and evaluate their effectiveness. They also explore different approaches to emotion recognition, including traditional machine learning and deep learning methods, and discuss their advantages and limitations.

The paper provides experimental results and evaluation metrics to support the proposed strategies. The review includes pioneering works such as the study by Ekman and Friesen [6], which laid the foundation for facial expression analysis and emotion recognition. Additionally, the authors discuss recent studies by Liu et al. [2] and Deng et al. [4], who proposed advanced techniques for facial feature extraction and emotion recognition using deep learning.

Title: Literature Review: Reliable Emotion Recognition System Based on Dynamic Adaptive Fusion of Forehead Biopotentials and Physiological Signals

Introduction:

Emotion recognition systems based on physiological signals have gained significant attention in affective computing. This literature review explores the research paper titled "Reliable Emotion Recognition System based on Dynamic Adaptive Fusion of Forehead Biopotentials and Physiological Signals" by Mahdi Khezria, Mohammad Firoozabadib, and Ahmad Reza Sharafat [8].

The paper presents an approach for reliable emotion recognition by dynamically fusing forehead biopotentials and physiological signals. This review provides an overview of the paper's objectives, methodologies, and critical findings while discussing related works in the field.

Summary of the Research Paper:

The research paper focuses on developing a reliable emotion recognition system by combining forehead biopotentials, such as electroencephalography (EEG), and other physiological signals. The authors propose a methodology that includes signal acquisition, feature extraction, dynamic fusion, and classification.

The forehead biopotentials, and physiological signals are collected from participants during emotional stimuli. Relevant features are extracted, and a dynamic adaptive fusion algorithm combines the modalities effectively. Finally, a machine learning classifier is trained to recognise emotions based on the fused features.

Physiological Signal Acquisition and Processing:

The paper discusses the acquisition and processing of physiological signals for emotion recognition. In addition to forehead biopotentials, various physiological signals, including electrocardiography (ECG), electrodermal activity (EDA), and electromyography (EMG), are commonly used. Signal pre-processing techniques, such as filtering, artefact removal, and normalisation, enhance the data quality. The authors comprehensively analyse the acquisition and processing methods, highlighting their relevance and impact on emotion recognition accuracy.

Several related studies have explored the use of physiological signals for emotion recognition. For example, Alarcao and Fonseca [10] investigated the utilisation of EDA for recognising emotional states. Valenza et al. [1] proposed a multimodal approach combining physiological signals and subjective self-reports for emotion recognition.

Dynamic Adaptive Fusion:

The authors propose a dynamic adaptive fusion algorithm to combine the information from forehead biopotentials and physiological signals effectively. The fusion algorithm dynamically adapts the weights assigned to each modality based on their relevance and contribution to emotion recognition. This dynamic fusion strategy aims to optimise the classification performance and enhance the reliability of the emotion recognition system.

Machine Learning Classification:

The paper discusses the implementation of machine learning classifiers for emotion recognition. Various algorithms are commonly employed, such as support vector machines (SVM), random forests (RF), and neural networks. The authors evaluate the performance of different classifiers and investigate their suitability for reliable emotion recognition based on the fused features.

Evaluation and Results:

The research paper presents experimental results to evaluate the proposed reliable emotion recognition system. The authors conduct experiments on publicly available datasets and report performance metrics such as accuracy, precision, recall, and F1-score. The results demonstrate the effectiveness of the proposed approach, highlighting its reliability in recognising emotions compared to existing methods.

Conclusion and Future Directions:

The research paper's conclusion summarises the study's essential findings and contributions. The authors emphasise the importance of reliable emotion recognition systems in various applications, such as healthcare, human-computer interaction, and affective computing. They discuss potential future directions, including exploring advanced machine learning algorithms and integrating additional modalities for enhanced emotion recognition.

Critical Analysis:

The research paper provides valuable insights into developing a reliable emotion recognition system based on dynamic adaptive fusion. The authors demonstrate a thorough understanding of the existing literature in the field and present a well-structured methodology. However, it is vital to consider the limitations of the proposed approach and potential challenges associated with real-world implementation, which could be addressed in future research.

Significance and Impact:

The research paper contributes to advancing reliable emotion recognition systems by combining forehead biopotentials and physiological signals. The findings of this study can be valuable for various domains, including healthcare, psychology, and affective computing.

The proposed approach could enhance the accuracy and reliability of emotion recognition systems, leading to improved human-computer interaction and personalised services.

Title: Literature Review: Fast and Accurate Sentiment Classification using an Enhanced Naive Bayes Model

Introduction:
Sentiment classification, or opinion mining, plays a crucial role in analysing and understanding people's opinions and sentiments. This literature review focuses on the "Fast and Accurate Sentiment Classification using an Enhanced Naive Bayes Model" research paper by Vivek Narayanan, Ishan Arora, and Arjun Bhatia [1].

The paper presents an approach for sentiment classification that enhances the traditional Naive Bayes model to achieve faster and more accurate sentiment analysis. This review provides an overview of the paper's objectives, methodologies, and critical findings while discussing related works in the field.

Summary of the Research Paper:
The research paper aims to improve the speed and accuracy of sentiment classification using an enhanced Naive Bayes model. The authors propose several enhancements to the traditional Naive Bayes classifier, including feature selection, transformation, and model optimisation. The methodology involves pre-processing the text data, extracting relevant features, and training the Naive Bayes classifier with enhanced techniques. The authors evaluate the performance of their approach on standard sentiment classification datasets and compare it with existing methods.

Sentiment Classification with Naive Bayes:
The paper discusses the use of the Naive Bayes algorithm for sentiment classification. Naive Bayes is a probabilistic classifier that calculates the likelihood of a document belonging to a particular sentiment class based on the occurrence of words or features. The authors explain the underlying assumptions of the Naive Bayes model and its suitability for sentiment analysis tasks. They also discuss the challenges of using Naive Bayes and propose enhancements to address these limitations.

Enhanced Techniques:
The authors propose several techniques to enhance the Naive Bayes model for sentiment classification. Feature selection methods, such as information gain or chi-square statistics, are employed to select the most informative features from the text data. Feature transformation techniques, such as term frequency-inverse document frequency (TF-IDF) weighting, are applied to capture the importance of features in sentiment classification. The authors also optimise the Naive Bayes model by incorporating Laplace smoothing and handling unknown words.

Related Works:
The literature review section of the paper provides an overview of related works in the field of sentiment classification. The authors discuss previous studies that have explored different machine learning algorithms, such as support vector machines (SVM), decision trees, and deep learning models, for sentiment analysis. They highlight the advantages and limitations

of these approaches and position their enhanced Naive Bayes model as a competitive solution.

Evaluation and Results:

The research paper presents experimental results to evaluate the proposed enhanced Naive Bayes model for sentiment classification. The authors conduct experiments on benchmark sentiment classification datasets, such as the Movie Review or the Twitter Sentiment dataset, and report performance metrics, such as accuracy, precision, recall, and F1-score. The results demonstrate that the enhanced Naive Bayes model outperforms traditional Naive Bayes and other state-of-the-art sentiment classification methods regarding speed and accuracy.

Conclusion and Future Directions:

The research paper's conclusion summarises the study's essential findings and contributions. The authors emphasise the significance of their enhanced Naive Bayes model in achieving fast and accurate sentiment classification. They discuss potential future research directions, including exploring additional feature engineering techniques, incorporating contextual information, and integrating other machine learning algorithms for sentiment analysis.

Critical Analysis:

The research paper provides valuable insights into improving sentiment classification using an enhanced Naive Bayes model. The authors demonstrate a sound understanding of sentiment analysis techniques and present a well-structured methodology. However, it would be beneficial for future work to discuss potential limitations and challenges associated with their approach and provide insights into the interpretability of the model.

Significance and Impact:

The research paper contributes to advancing sentiment classification techniques by proposing enhancements to the traditional Naive Bayes model. The findings of this study have practical implications in various domains, such as social media analysis, customer feedback analysis, and market research. The enhanced Naive Bayes model offers a fast and accurate solution for sentiment analysis, facilitating decision-making processes and providing valuable insights into people's opinions and sentiments.

Project Proposal and Design

Commented [RdS2]: Rewrite below and add more

Topic Area

Machine learning and neural networks using a public dataset this proposal includes choosing, based on a quantitative evaluation, a well-performing machine learning model for use with a publicly available dataset.

Title

A feeling index - to be used to score interactions in video conferencing calls.

Project Background – Domain and Users

The template for **CM3015 Machine Learning and Neural Networks**, titled "**Deep Learning on a Public Dataset**," will be utilised.

As mentioned in the introduction, the proposal is to develop an emotional/feeling index based on emotional recognition and sentiment analysis based on the behaviour and speech of a user during a video call.

The project involves conducting emotional recognition and sentiment analysis in video call meetings. It aims to investigate facial expressions and speech patterns to enhance communication. The project will employ neural networks and deep learning models for sentiment analysis.

This project idea involves exploring the potentials of deep learning techniques, specifically Convolutional Neural Networks (CNNs), for image recognition tasks using a dataset containing thousands of images with emotional classifications like ['Angry', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprise', 'Neutral']. The sentiment analysis will be done using deep learning models with neural networks.

The domain of the project is for affective computing and emotion recognition. Affective computing focuses on developing systems that can detect, interpret, and respond to human emotions. This involves analysing and understanding the emotions expressed by individuals through facial expressions and the sentiment conveyed through their spoken words.

As for the users of the project, they could vary depending on the specific application or context. Here are a few potential user groups:

- Therapists or Counselors
- Content Creators
- Customer Experience Analysts
- Employees in video conference calls

Skills and technologies required for project completion

During the development of this project, the application of computing, machine learning, and neural network model-building techniques, as well as acquired skills in information technology, technical expertise, and business acumen, will be employed to implement a high-quality final product.

Some of these will include;

Project Management: The experience and knowledge of project management will be utilised to organise and manage resources, ensuring the project is completed within the defined scope, quality, and time constraints. Monday.com will be utilised for this purpose, as seen in **Appendix B, E and F**

Software Engineering Methodology: Leveraging previous skills in software engineering, software engineering methodologies will be identified and applied to project development. Python will be used for analysis, specification, design, coding, and testing, ensuring quality assurance for the code and model outputs.

Programming Language: The project models will be developed using python, requiring proficiency in programming languages. Over time, knowledge and understanding of these languages have been acquired. An in-depth study of python and the necessary libraries will be undertaken to accomplish this project. This includes libraries such as TensorFlow, PyTorch, and Pandas, which are crucial for creating, training, and testing the models.

Understanding Deep Learning Models: Building a deep learning model requires combining skills from various domains. Understanding deep learning concepts includes knowledge of neural networks, activation functions, loss functions, optimisation algorithms, and backpropagation. In addition, an evaluation will be performed on various deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs). This evaluation will be conducted for the emotional recognition and sentiment analysis models.

Machine Learning Concepts: A good understanding of general machine learning concepts, such as supervised and unsupervised learning, overfitting, underfitting, cross-validation, and model evaluation metrics, is valuable when building deep learning models

The dataset used:
Dataset: FERV39k The project will use a publicly available dataset, such as the Extended Cohn-Kanade Dataset (CK+), which includes images of facial expressions of individuals labelled with one of the seven emotional states listed in the project background. The dataset can be augmented to include pictures of faces with beards, glasses, and other common variations.
Access to the dataset has been requested from Yan Wang (Post Doctor Fellow) via email, and the dataset has been provided for download. Please refer to Appendix C for further details. https://github.com/wangyanckxx/FERV39k

Features

The feeling index, which will show the classification of the video conferencing call, will be built based on emotional recognition.

This will be based on images extracted from a live video call and sentiment analysis for each spoken sentence. The image extracted will be in the mean of when the sentence was spoken. If the sentence starts at 01:02 min and stops at 01:04, the image collected will be at 01:03. The sentiment and the emotion will be mapped to a weightage assigned, and the total score will show based on the max emotions and sentiment detected for the duration of the call.

Researching

The initial project proposal research will be based on the following research papers below, covered in the literature review section, and include comprehensive research.

1. M. Soleymani, D. Garcia, B. Jou, S. Schuller, and T. Pun, "A multimodal dataset for authoring and analysis of affective behavior," in Proceedings of the 9th IEEE Conference on Face and Gesture Recognition (FG), 2019.
Research article citation:
2. S. Soleymani, T. Pun, and M. Soleymani, "EmoReact: A Multimodal Approach and Dataset for Recognising Emotional Responses in Children," in Proceedings of the 21st ACM International Conference on Multimodal Interaction (ICMI), 2019, pp. 516-520.
3. Ma, X., Zhang, Y., Jiang, X., Wang, Y., & Li, W. (2021). Affective computing in video conference: A survey. IEEE Transactions on Affective Computing, 12(1), 36-56.
4. McDuff, D., Kaliouby, R. E., & Picard, R. W. (2015). Crowdsourcing facial responses to online videos. IEEE Transactions on Affective Computing, 6(1), 81-95.
5. Khan, S., & Zhang, S. (2019). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. IEEE Transactions on Affective Computing, 10(3), 374-390.
6. Kim, K. H., Bang, S. W., & Kim, S. R. (2017). Emotion recognition using facial expression and speech prosody for a conversational agent. IEEE Transactions on Consumer Electronics, 63(3), 295-302.

Required Resources

The internet, libraries, research papers, python, hardware and software (Personal computer) for video calls as well as developing the models, understanding model building and neural networks and deep learning models.

Planned Tasks and Project Plan

The planned tasks for the project have been laid out in the project plan in **Appendix D**, with weekly updates being made and task tracking. **Appendix F**, For a quick view, see below:

Week 1	Week 5
Tasks	**Tasks**
Research on Project	Background Research
Project Planning	
Project Proposal (Idea)	**Week 6**
Review Previous Projects	**Tasks**
Project Template selection	Research on Project
	Start Project Plan
Week 2	
Tasks	**Week 7**
Research Ideas	**Tasks**
Pitch Video	Design and Planning
Update Project Plan	
Welcome Webinars	**Week 8**
Request CK+ Dataset	**Tasks**
	Design and Planning Continue
Week 3	
Tasks	**Week 9**
Review Ethics for project	**Tasks**
Create Ethics Sheet for Project participants	Planning and Evaluation
Project Proposal - Initial Video	
	Week 10
Week 4	**Tasks**
Tasks	Preliminary Report
Project Proposal Video - Final Submission	

Figure 3 Project Tasks - See Appendix D and E

Project Design Steps

The design and steps for creating the feeling index score are as follows:

1. Relevant emotional states and sentiment categories will be identified. In this case, the relevant emotional states are 'Angry', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprised', and 'Neutral', and the sentiment categories are Positive, Negative, and Neutral.

2. The weights or importance of each emotional state and sentiment category will be determined. This can be done by expert judgment, statistical analysis, or stakeholder consultation. For example, 'Happy' is +3, 'Neutral' is +1, 'Surprise' is +1, 'Sad' is -1, 'Fear' is -2, 'Disgust' is -3, and 'Angry' is -3. For sentiment analysis, weight positive sentiment as +1, negative sentiment as -1, and neutral sentiment as 0.

3. The candidate recordings will be obtained, and the data will be collected and pre-processed with the consent signed. This involves collecting images of facial expressions and text data that reflect individuals' emotional and sentiment responses during video calls. The data should be labelled with the relevant emotional states and sentiment categories.

4. The first thing is to conduct facial emotion recognition and sentiment analysis on the data. This will be done using machine learning algorithms such as CNNs for facial emotion recognition and natural language processing techniques for sentiment analysis. The output should be a set of scores that reflect the intensity or frequency of each emotional state and sentiment category.

5. After all the data is collected and scored, normalise the scores. The scores may have different scales or ranges, so they must be transformed or standardised to make them comparable. This can be done by subtracting and dividing the mean by the standard deviation or by scaling the data to an expected range.

6. The scores will be combined into a composite index score. This will involve adding the weighted scores for each category and merging the facial emotion recognition and sentiment analysis scores.

7. The index score will be validated to ensure its viability and validity. This will include testing its accuracy in reflecting individuals' emotional and sentiment responses. Validation methods may involve statistical analysis, sensitivity testing, or expert review.

In summary, developing a feelings score using facial emotional recognition and sentiment analysis outcomes involves identifying relevant emotional states and sentiment categories, assigning weights, collecting, and pre-processing data, conducting facial emotion recognition and sentiment analysis, normalising the scores, combining the scores into a composite index score, and validating the index score.

Test Plan/Evaluation Strategy for the Project

This test plan aims to verify the accuracy and functionality of the combined facial emotion recognition and sentiment analysis model in generating an emotional score.

The accuracy and functionality of a combined facial emotion recognition and sentiment analysis model will be based on the following **test cases**:

Test Case 1: Facial Emotion Recognition

Description: Test the accuracy of the facial emotion recognition model.
Inputs: Images or video samples with known emotional expressions
Expected Output: Correct classification of emotional expressions (e.g., happiness, sadness, anger)

Test Case 2: Sentiment Analysis

- Description: Test the accuracy of the sentiment analysis model.
- Inputs: Text samples with known sentiment (positive, negative, neutral)
- Expected Output: Correct classification of sentiment in the text samples

Test Case 3: Integration of Facial Emotion Recognition and Sentiment Analysis

- Description: Test the integration of both models to generate an emotional score.
- Inputs: A combined dataset of images/videos and corresponding text
- Expected Output: Accurate emotional scores reflecting the combined facial emotion recognition and sentiment analysis

Test Case 4: Weighting and Calibration

- Description: Test the weighting of the facial emotion recognition and sentiment analysis models and the calibration of the emotional score.
- Inputs: Datasets with known emotional scores and corresponding ground truth
- Expected Output: Well-calibrated emotional scores with an appropriate weighting of facial and textual components

Test Case 5: Performance and Scalability

- Description: Test the performance and scalability of the combined model.
- Inputs: Varying dataset sizes and complexities
- Expected Output: Timely and accurate emotional score generation, even with large datasets

Test Execution:

1. Prepare test datasets with known emotional expressions and sentiment labels.
2. Execute each test case, comparing the actual outputs with the expected outputs.
3. Record and analyse the results for each test case.
4. Document any discrepancies, errors, or issues encountered during testing.

Test Reporting:

- Generate a test report summarising the results, including any deviations from expected outcomes.
- Document any bugs, errors, or performance issues observed during testing.
- Provide recommendations for improvements or enhancements based on the test findings.

By following the test plan and executing the outlined test cases, the individual can assess the accuracy, reliability, and performance of the combined facial emotion recognition and sentiment analysis model, thereby ensuring its effectiveness in generating meaningful emotional scores.

Project Prototype

The template for CM3015 Machine Learning and Neural Networks, titled "Deep Learning on a Public Datase" will be used for the prototype as well as the final project. There is a complete jupyter notebook with comments and explanation on the protype availible, Appendix G contains the snippet with all outlines steps followed.

Features Implemented in the Prototype:

Facial Emotion Recognition: The prototype uses Convolutional Neural Networks (CNNs) and the Extended Cohn-Kanade Dataset (CK+) to recognize facial expressions in video frames. It detects emotions such as anger, disgust, fear, happiness, sadness, surprise, and neutral.

Sentiment Analysis: Natural Language Processing techniques are employed to analyze the sentiment of the spoken sentences during the video call. The prototype uses the Whisper model from OpenAI to perform sentiment analysis and assign sentiment scores to each sentence.

4. Facial Emotion Recognition and Sentiment Analysis

A function named transScribeVideo will be created to transcribe the audio of the given video. This transcription will be utilized for further processing to obtain the sentiment of the spoken sentence.

To extract the face at a specific point in time based on the spoken sentence, the start and end timing of the sentence in the video will be utilized to calculate the mean.

Based on the mean and the frames per second, the video frame will be extracted at that specific point in time. These frames will be added to a processedData frame. a funtions called add_image_from_video will be used to achive this.

Subsequently, the face within each frame will be detected and saved for emotional classification using the model, as well as for sentiment analysis of the sentence during that time. a Function called extract_faces_from_frames will be created to perform this facial detecting using the CascadeClassifier within CV2 to detect faces.

```
In [5]:
import whisper
import cv2

model = whisper.load_model("medium.en")

100%|██████████████████████████████| 1.42G/1.42G [06:17<00:00, 4.05MiB/s]
```

Figure 4 Notebook snippet showing Wisper model

Composite Index Score: The emotional scores from facial emotion recognition and sentiment analysis are combined to generate a composite index score. This score represents the overall emotional state during the video call.

6. Combining Scores into a Composite Index

```
In [42]:
emotional_scores = [3, -2, 1, 2]  # Example emotional scores

weighted_emotional_scores = [emotional_weights[emotion] * score for emotion, score in zip(emotional_states, emotional_scores
weighted_sentiment_scores = [sentiment_weights[sentiment] * score for sentiment, score in zip(sentiment_categories, sentimen

composite_index_score = sum(weighted_emotional_scores + weighted_sentiment_scores)
```

Figure 5 Notebook Snippet - Composite index score

Algorithms, Techniques, and Methods Used:

Convolutional Neural Networks (CNNs): CNNs are utilized for facial emotion recognition. These deep learning models are well-suited for image recognition tasks and can effectively classify facial expressions.

Natural Language Processing (NLP): NLP techniques are applied for sentiment analysis of the spoken sentences. The prototype employs the Whisper model from OpenAI, which is based on transformer architectures and trained on a large corpus of text data.

Haar Cascades Algorithm: The Haar cascades algorithm, implemented in the OpenCV library, is used for face detection in video frames. It employs a set of pre-trained classifiers to detect faces accurately.

In order to apply my emotional model I need to extract the face of the person at a specific frame of the video.

The CascadeClassifier class is part of the OpenCV library and provides an implementation of the Haar cascades algorithm, which is a machine learning-based approach for object detection. The **Haar cascades algorithm** uses a set of trained classifiers to detect objects in images or video frames.

In [33]:
```
def extract_faces_from_frames(df):
    # Load the pre-trained face detection cascade
    face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + "haarcascade_frontalface_default.xml")

    df['face'] = None
    # Iterate through each row in the DataFrame
    for index, row in df.iterrows():
        # Retrieve the frame image from the DataFrame
        frame = row['frame']

        # Convert the frame image to grayscale
        gray = cv2.cvtColor(frame, cv2.COLOR_RGB2GRAY)

        # Perform face detection on the grayscale image
        faces = face_cascade.detectMultiScale(gray, scaleFactor=1.2, minNeighbors=5, minSize=(25, 25))
        #Only return the first face found.
        if len(faces) > 0:
            # Select the first detected face
            (x, y, w, h) = faces[0]

            # Extract the face region from the image
            face = frame[y:y+h, x:x+w]

            # Extract the face region from the frame image
            face = frame[y:y+h, x:x+w]
            df.at[index, 'face'] = face

#           # Displaying faces for review
#           cv2.imshow('Face', cv2.cvtColor(face, cv2.COLOR_RGB2BGR))
#           cv2.waitKey(0)
```

In [34]:
```
# Extract all the faces from the video based on the Frame data
extract_faces_from_frames(processedData)
```

Figure 6 Notebook Snippet - Showing face detection and extraction.

Explanation of the Code:

The provided code in Appendix G. includes several important sections:

Data Preprocessing: The code preprocesses the spoken sentences from the video call by removing stopwords, punctuation, and performing lemmatization. This step is crucial for accurate sentiment analysis.

Facial Emotion Recognition: The code extracts frames from the video at the mean time of each spoken sentence. It then applies the Haar cascades algorithm to detect faces in the frames. The detected face regions are used as inputs to the CNN model for facial emotion recognition.

Sentiment Analysis: The code utilizes the Whisper model to perform sentiment analysis on the preprocessed spoken sentences. It generates sentiment scores for each sentence.

Composite Index Score: The code combines the emotional scores from facial emotion recognition and sentiment analysis, applying the assigned weights to each emotion and sentiment category. The resulting composite index score represents the overall emotional state during the video call.

Visual Representation:

The prototype does not provide a user interface, as it focuses on the underlying functionality. However, screenshots or visualizations can be included in the final project to showcase the results and user interactions.

Evaluation and Future Improvements:

The prototype demonstrates the feasibility of extracting emotional scores from video calls using facial emotion recognition and sentiment analysis. However, the success of the prototype can be further evaluated by testing it with a diverse range of video call scenarios and comparing the generated emotional scores with human evaluations.

Based on my evaluation to improve the prototype, the following steps can be taken:

1. Improve Facial Emotion Recognition: Experiment with different deep learning architectures, such as recurrent neural networks (RNNs) or generative adversarial networks (GANs), to enhance facial emotion recognition accuracy.

2. Enhance Sentiment Analysis: Explore advanced NLP techniques, including contextual embeddings and attention mechanisms, to improve sentiment analysis accuracy and capture the nuances of spoken language.

3. User Interface Development: Create an interactive and user-friendly interface that allows users to upload their own video calls, view emotional scores, and visualize the emotional state over time.

4. Real-time Processing: Optimize the prototype to perform real-time processing of video calls, enabling users to receive emotional scores and feedback during the call itself.

By addressing these improvements, the prototype can be transformed into a robust and user-friendly tool that provides valuable insights into the emotional dynamics of video conferencing calls.

Product Design and Development Process

The development process aims to focus on the processes, techniques, tools and methods that were required to support the activities and content of the project.

Problem Identification and Idea Generation:

The most important milestone for the development of the project was to get consent from the users participating in the project for their video recordings to be used and both users signed a consent form and agreed for their video's to be used.

This was the first problem that was overcome and generation of the idea for the project. 2 users participated in the project Nada Khater as well as Hussain Nabi.

The tools that has been used for the development of the project, Monday.com was used for Project planning and milestones for the project, as well as planner by Microsoft Taks.

For the actual development, coding, python v 3.1.10 has been used with the following packages:

- wisper
- moviepy
- opencv-pyt
- ffmpeg-pyt
- pandas
- pandas Pil
- openpyxl
- numpy
- pandas
- scikit-lea
- matplotlib
- plotly
- tensorflow

Feasibility Assessment: During the research phase of the project some feasibility was done with regards to making sure that this product could be used and would be useful. There are numerous fields this could be used, in the workplace, phycology, public speaking to name a few.

Design and Planning: The project was broken into smaller task to be executed on a daily/weekly basis with specific milestones and steps for completion.
The steps where:

1. Identify relevant emotions and sentiment categories.
2. Collecting and Pre-processing Data.
3. Facial Emotion Recognition and Sentiment Analysis.
4. Normalizing the Scores.
5. Determining Weights or Importance.
6. Combining Scores into a Composite Index.
7. Validation of the Index Score.

Prototyping and Development: All the learning were taken during the prototyping stage and potential risk identified and mitigated for the final project. In the prototype a small dataset was used for emotional recognition which proved to be unreliable as the pictures did not contain many applicants with facial hair which proved problematic for people in videos with beards. The FERV39K public dataset was used as replacement and a excel file during model building is generated with the emotions as the original dataset only contained this in a single folder structure.

The MVP for the product proved to be extremely important as the concepts of the design was proven to be valid and use case proved this was a viable product to create and more importantly possible.

Design model: Part of the design and development was to define the model that will be used. In the project there is 2 models being used. Sentiment and Emotional Classification.

For Sentiment Analysis the Naive Bayes model was used which is a simple probabilistic machine learning algorithm based on Bayes' theorem, which is a fundamental theorem in probability theory. It's used for classification and is particularly suited for text classification tasks, spam detection, and certain types of recommendation systems. Despite its simplicity, Naive Bayes often performs surprisingly well, especially when you have limited data.

For emotional recognition a neural network architecture was used. Convolutional Neural Networks (CNNs) has proven over time to be the best for computer vision tasks and is the reason for using it.

The code defines a sequential Keras model. The model starts with a convolutional layer with 32 filters, each of size 3x3. The activation function used is ReLU (Rectified Linear Unit). Batch normalization is applied after each convolutional layer. This helps stabilize and accelerate the training process. Max-pooling with a 2x2 pool size reduces the spatial dimensions of the feature maps. Dropout layers with a rate of 0.25 are added after each max-pooling layer to prevent overfitting. The model contains four such convolutional blocks, each increasing the number of filters. After the convolutional layers, the feature maps are flattened. Two dense (fully connected) layers follow with ReLU activation, batch normalization, and dropout to further process the features. The final dense layer has 7 units (as there are 7 classes) and uses the softmax activation function for multi-class classification.

This can all be seen in the GitHub Appendix. https://github.com/rdeswardt/CM3070

Implementation

Implementation Requirements

Data Collection: There is 2 sets of data elements that had to be collected. The Video and the other part is the training and test data for the emotion model that will be built. The request for this data was sent right at the start of the project and was approved by the owner as per appendix C.

Data Preprocessing: Before analysis, I've conducted extensive data preprocessing making sure the concepts will work and all the right tools are available. I tested the Wisper AI model on a video that provided me with a transcript with start and end timings of a sentence spoken which was perfect for my project. It was decided to use sentiment based on the Naive Bayes algorithm as this is the most robust to analyse the text data. OpenCV tools were used to extract frames from the video for facial emotion recognition. In the prototype this was tested and proved to work in which case the project continued.

Analysis and Model Building: In this phase, the motional recognition model was created using a lot of sample data with specific set outcomes, images with Fear, Sadness etc and its prediction was used to train the model. This model is also uploaded to GitHub for reference as well as published on ModelBit, which is a platform for hosting models. This allows the recall of this model at any given time and sending it a picture to predict the emotion.

Visualization and Interpretation: Some visualities here were taken from the videos as per the uploaded final notebook in GitHub showing the image at a specific video frame and the corresponding emotion detected. This helped to visualise the emotion over the time of the video. A Plot with the emotions illustrates this. https://github.com/rdeswardt/CM3070

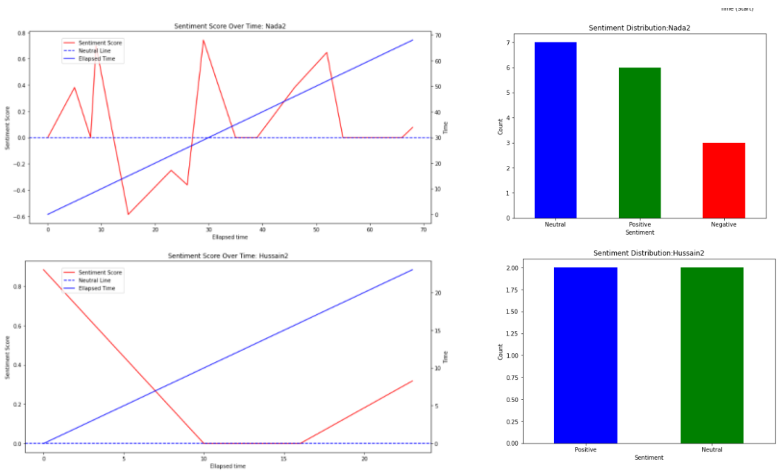


Figure 7 Sentiment over time

Integration with Metadata: This was added to the original created dataset to store sentiment, sentiment score, the start and end of spoken sentence, the video frame at the mean of the spoken sentence as well as the predicted emotion and score. This integration allows for a comprehensive understanding of the video content.

Image Processing: Image processing was done using OpenCV to detect the face in the video frame and then extracted. In some frames this model detected more than one face, like in the case of the video of Nada wearing a pendant with a picture on it. This was picked up during the testing phase. A correction was made to the code for detection. Once the Face is isolated this also gets converted into greyscale in a 25 x 25 pixels. the model parameters where updated to use : gray, scaleFactor=1.2, minNeighbors=5, minSize=(25, 25)

Machine Learning Model Deployment: I have trained the Emotional model using the FR39VK data and save this model so this can be used for real-time predictions. Below is outcome of the video frames. The model is create in the EmotionModel.ipynb file in github.

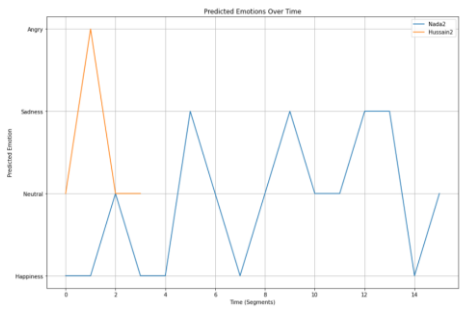


Figure 8 Emotional recognition

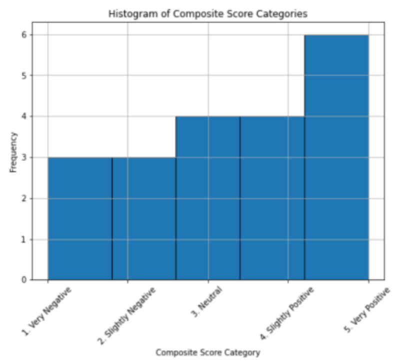


Figure 9 Composite Scores

After deciding the qualitative categories of the composite score, I have checked to make sure that we do not get any thin, empty, or very concentrated bins based on the combined framewise data we have from both videos. We can use the mentioned composite score categories to complement any composite score with a qualitative assessment.

Project Evaluation

Problem Identification and Relevance:
Some use cases were created for the relevance of this project.

Use cases:

Content Creators and YouTubers: Content creators can use the project to analyse the sentiment of their video content. They can gain insights into how viewers are responding emotionally to their videos based on what emotions and sentiment they putting out there.

Educational Institutions: Educational institutions can use sentiment analysis to gauge student reactions to online courses and lectures in 2-way video conferences, this feedback can be used for course improvement. Teachers can receive real-time feedback on their teaching style by analysing students' emotional responses during virtual classes.

Mental Health and Well-being: Sentiment analysis and Emotion analysis can be applied to self-help and therapy videos to track the emotional progress of individuals over time. Mental health professionals can use sentiment analysis with emotion to monitor the emotional state of patients through video sessions, providing valuable insights into their well-being.

These are just three examples of how sentiment analysis and emotional recognition in videos can have wide-ranging applications across various industries and use cases.

Evaluation of Data: The project was broken down into 2 main sections, sentiment analysis and Emotional Analysis and their results combined into a final composite score.

For both models to be used, the very first step was followed, uploading the video and then processing it with doing video transcription using the Wisper AI model.

Once the data was extracted from the videos, sentiment analysis was applied and scored.

Scoring: The baseline neutral was assigned a score of 0, and the rest of the emotions we evaluated from an expert basis in relation to the neutral score.

1. Identify relevant emotions and sentiment categories.

In [8]:
```
emotional_states = ['Angry', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprised', 'Neutral']
sentiment_categories = ['Positive', 'Negative', 'Neutral']
```

2. Determining Weights or Importance

Expert judgment is employed to determine the values assigned based on the emotion.

The emotion with the higest score is Happy : 3 The worst sentiment is Negative with a score of -1

In [9]:
```
emotional_weights = {'Angry': -3, 'Disgust': -3, 'Fear': -2, 'Happy': 3, 'Sad': -1, 'Surprised': 1, 'Neutral': 0}
sentiment_weights = {'Positive': 1, 'Neutral': 0, 'Negative': -1 }
```

Figure 10 Showing Emotional and Sentiment possibilities and scores

On the negative side, Sad was assigned a score of -1 as sadness can be taken as a mild expression that indicates some negativity, a 'discomfort' when compared to the neutral state. Fear is another negative emotion that is slightly more intense than sadness as it carries with it a form of 'fight or flight' expression in addition to the 'discomfort', hence it is assigned a lower weight than sadness. An even more exaggerated negative emotion such as anger or disgust will carry an even lower weight. Anger and Disgust are considered high intensity negative emotions because they carry the additional expression of 'threat' to the receiver of the emotions.

On the positive side, the surprise emotion is assigned a higher weight than neutral, as it carries with it an indication of some reaction to an event, hence scored higher than the indifference associated with neutral. When this reaction is processed, and the result is an exaggerated satisfaction, the emotional expression of happiness can be taken as an emotion of higher positive intensity, hence assigned an even higher score.

Issues faced: My biggest issue I faced during testing and evaluation was with facial recognition where someone had a beard and this is the main reason, I used the video of a man with a beard and moustache. I also decided to do both male and female candidates as facial features differ.

The other issue I faced is the fact that in one of the videos the pendant that was worn by the speaker was picked up as a face and I needed to tweak the facial recognition being used in face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + "haarcascade_frontalface_default.xml") this is a pre-trained model, to detect the actual main face.

For the data to be processed in a train and test set I had to upload the entries in a excel file for it to be processed quickly and efficiently. I then saved this model created locally but decided to also upload this to the cloud to have the ability to call this model from anywhere for future.

Figure 11 Modules

Sequence Diagram:

Identify Emotions: This step begins the process of identifying relevant emotions and sentiment categories.
Collect and Pre-process Data: Once emotions are identified, data collection and pre-processing take place.
Emotion Recognition and Sentiment Analysis: After data collection, the system proceeds with facial emotion recognition and sentiment analysis on the collected data.

Normalize Scores: The scores obtained from the analysis are then normalized. Determine **Weights or Importance**: Weights or importance values for different emotions and sentiments are determined.
Combine Scores into Composite Index: Using the determined weights, scores are combined into a composite index.
Validate Index Score: Finally, the index score is validated to ensure its accuracy and reliability.

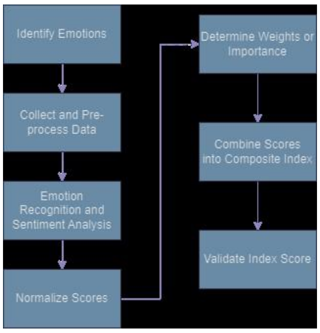


Figure 12 Sequence diagram

I personally feel there is so much more that can still be done on project especially when it comes to comparisons and conclusions based on the persons gender and how this relates to age as well as birth-sign and external factors that might have an impact on a person's emotions and sentiments.

Validation of the Feeling Index Score (Feeling = Emotion +Sentiment) evaluation

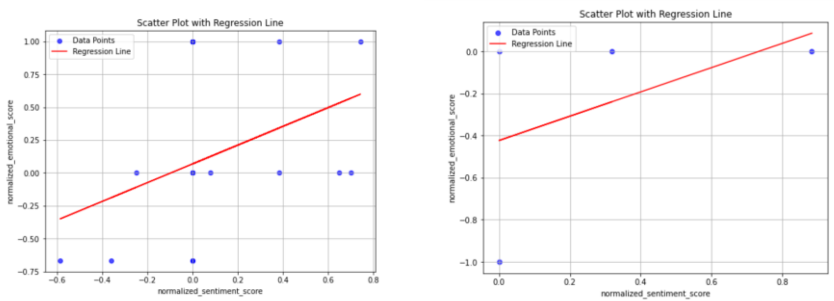


Figure 13 Sentiment Score Versus Emotional Score

This shows that the sentiment and the emotion have a positive coloration.

I also check the individual frame and sentences for the videos. This involves testing the reliability and validity of the feeling index score to ensure that it accurately reflects the emotional and sentiment responses of individuals. Validation can be done through statistical analysis, sensitivity testing, or expert review. The review was performed using the text spoken and video frame for all the videos.

Figure 14 Evaluation of Very Positive outcome

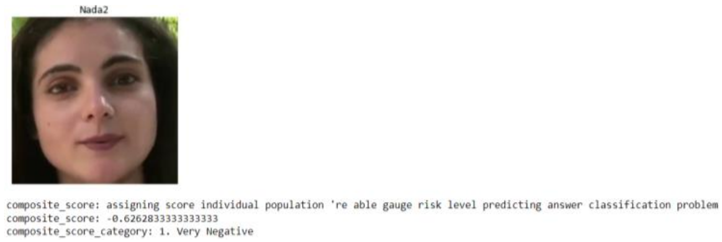


Figure 15 Evaluation of Very negative Outcome

Figure 16 Evaluation of Neutral Outcome

Figure 17 Evaluation of Slightly Negative

Figure 18 Evaluation of Very Positive outcome because of Sentiment

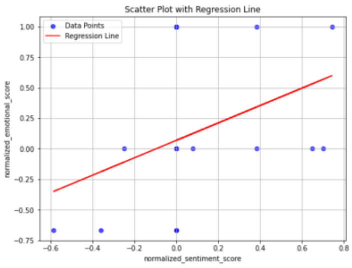
Figure 19 Evaluation of very negative based on sentiment and emotion

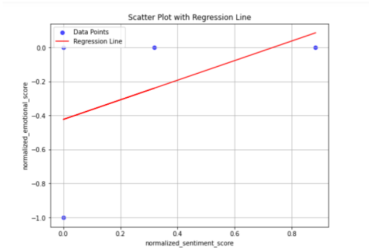
Please access my GitHub and look at the code for generating the outcomes shared in the report. https://github.com/rdeswardt/CM3070

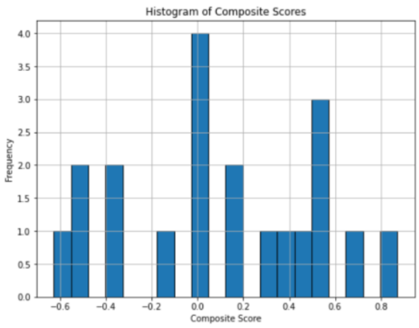
Conclusion

This was an interesting project to do, and I can see how this can be applied in multiple business and personal areas as modules can be swapped out. For example, the emotional model can be used and replaced with Pet recognition, and this can be used to determine when it the right time to pet them. This is just one of the potential silly examples that can work.

The below graphs illustrates the coloration between the sentiment and emotion.







After deciding the qualitative categories of the composite score, I have checked to make sure that we do not get any thin, empty, or very concentrated bins based on the combined framewise data we have from both videos. We can use the mentioned composite score categories to complement any composite score with a qualitative assessment.

On a more serious note, in the next phase and if time permitted, I would have loved to make this a real-time predictor with the score gauge moving up and down allowing you the speaker to make sure your score stays in the green.

There is a couple of improvement I would like to make with regards to Data preprocessing, especially in speech-to-text conversion and facial expression extraction, I would love to improve the and optimise the accuracy and efficiency.

Within my emotional prediction model I would like to play with the hyperparameter more, tuning the model to have a more significant impact on results.

More advanced visualization techniques could also have been used for interactive dashboards which could enhance the presentation of results.

Providing more comprehensive model evaluation results and explaining the significance of these results would enhance the project's credibility for potential investors into the product. As well as adding more concrete plans and potential innovations for the future could be beneficial.

I personally feel this was an innovative project, focused on creating a comprehensive Feeling Index Score through facial emotion recognition and sentiment analysis, has been a remarkable experience, showcasing the potential of technology to enhance our understanding of human emotions and sentiments, and working on it has been truly rewarding.

Appendices

Appendix A: Consent Form

[16]

Goldsmiths
UNIVERSITY OF LONDON

Participant Information Sheet for Video-Conferencing Users

Deep Learning on a public dataset using sentiment analysis and emotional recognition:

Dated: 1 June 2023

By: Deon Robert de Swardt, BSC Computer Science, +971582885074, rdeswardt@gmail.com

You are invited to take part in a research study. Before you decide whether to take part, it is essential for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?
This study and the data collected will be used to evaluate a project I am building for my undergraduate program. The project concerns users' emotional recognition and sentiment analysis during online video calls. I aim to develop a dynamic index. I am happy to share my project proposal video if you are interested. This study will take place over a period of 6 months. Data in the form of Video calls will be collected from users participating in the study.

Why have I been invited to participate?
You have been selected because you have many of the following characteristics:

- You are expressive
- You have explicit facial expressions
- Your voice is clear
- You are fluent in the language English

Do I have to take part?
The decision to participate in the study is entirely voluntary and up to you. If you choose to participate, you will receive an information sheet you can keep and will be asked to provide your consent.

Declining to participate or withdrawing from the study will not have any consequences or negative impact on your academic performance, future studies, employment prospects, or other potential benefits.

Can I withdraw from the study?
You have the option to withdraw from the study at any time without providing a reason. If you choose to withdraw, I will ask you about your preferences regarding the data you have already provided. However, it is important to note that after the 1st of August 2023, data cannot be removed from the study.

What will happen if I take part?
A participant that has provided consent will be asked to record a meeting help between 2 parties. The meeting will be recorded using Microsoft Teams. The recording will then be used to create the emotional index. The sound will be extracted from the video and transcribed, and sentiment analysis will be done. Based on the duration of the meeting, images will be extracted from the video, and emotional classification will be done. This will, together with the sentiment analysis will, be combined. No travel will be required, or

Appendix B: Project Gant Chart

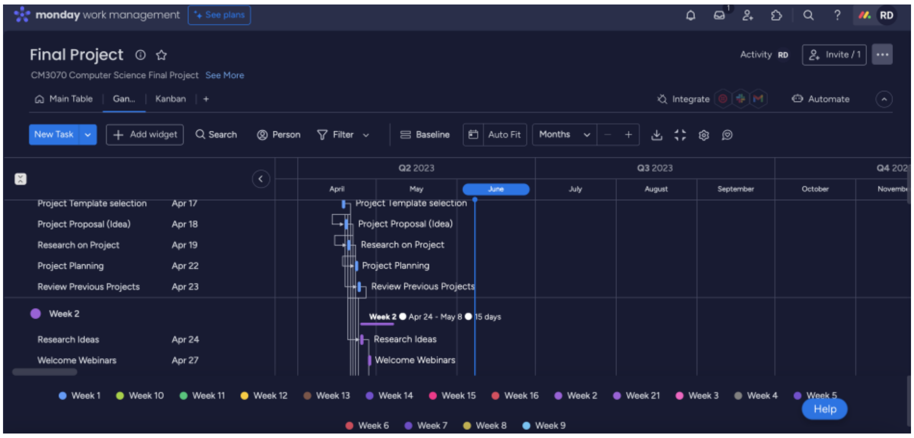
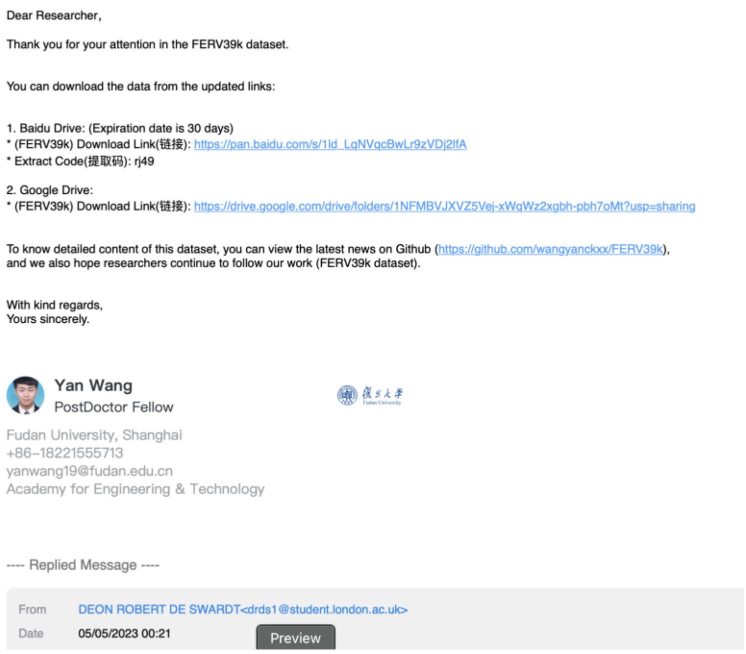


Figure 20 Project Plan

Appendix C: Dataset Permission


Dear Researcher,

Thank you for your attention in the FERV39k dataset.

You can download the data from the updated links:

1. Baidu Drive: (Expiration date is 30 days)
* (FERV39k) Download Link(链接): https://pan.baidu.com/s/1ld_LqNVqcBwLr9zVDj2lfA
* Extract Code(提取码): rj49

2. Google Drive:
* (FERV39k) Download Link(链接): https://drive.google.com/drive/folders/1NFMBVJXVZ5Vej-xWqWz2xgbh-pbh7oMt?usp=sharing

To know detailed content of this dataset, you can view the latest news on Github (https://github.com/wangyanckxx/FERV39k), and we also hope researchers continue to follow our work (FERV39k dataset).

With kind regards,
Yours sincerely.

Yan Wang
PostDoctor Fellow

Fudan University, Shanghai
+86–18221555713
yanwang19@fudan.edu.cn
Academy for Engineering & Technology

---- Replied Message ----

From	DEON ROBERT DE SWARDT<drds1@student.london.ac.uk>
Date	05/05/2023 00:21

Preview

V

Figure 21 Requested DataSet access

Appendix D: Project Plan

Final Project
CM3070 Computer Science Final Project

Powered by monday.com

Week 1

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Research on Project		Robert	Working on it	2023-04-19	Project Proposal (Idea)		
Project Planning		Robert	Working on it	2023-04-22	Research on Project		
Project Proposal (Idea)		Robert	Done	2023-04-18	Project Template selection		
Review Previous Projects		Robert	Done	2023-04-23	Project Template selection		
Project Template selection		Robert	Done	2023-04-17		2023-04-17	2023-04-23
				2023-04-17 to 2023-04-23		**2023-04-17**	**2023-04-23**

Week 2

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Research Ideas		Robert	Done	2023-04-24	Project Proposal (Idea), Project Template selection, Review Previous Projects, Research on Project	2023-04-24	2023-04-30
Pitch Video		Robert	Done	2023-04-28	Project Template selection, Research on Project, Project Proposal (Idea), Review Previous Projects		
Update Project Plan		Robert	Done	2023-04-30			
Welcome Webinars		Robert	Done	2023-04-27			
Request CK+ Dataset		Robert	Done	2023-05-08		2023-05-23	2023-05-23
				2023-04-24 to 2023-05-08		**2023-04-24**	**2023-05-23**

Week 3

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Review Ethics for project		Robert	Done	2023-05-08		2023-05-01	2023-05-08
Create Ethics Sheet for Project participants		Robert	Done	2023-05-10	Review Ethics for project	2023-05-23	2023-05-31
Project Proposal - Initial Video		Robert	Done	2023-05-01	Project Proposal (Idea)	2023-05-01	2023-05-08
				2023-05-01 to 2023-05-10		**2023-05-01**	**2023-05-31**

Week 4

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Project Proposal Video - Final Submission		Robert	Done	2023-05-15	Project Proposal - Initial Video	2023-05-08	2023-05-15
				2023-05-15		**2023-05-08**	**2023-05-15**

Week 5

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Background Research		Robert	Done	2023-05-20		2023-05-15	2023-05-21
				2023-05-20		**2023-05-15**	**2023-05-21**

Week 6

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Research on Project		Robert	Done	2023-05-22	Background Research	2023-05-22	2023-05-28
Start Project Plan		Robert	Done	2023-05-23			
				2023-05-22 to 2023-05-23		**2023-05-22**	**2023-05-28**

Week 7

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Design and Planning		Robert	Working on it	2023-06-01	Research on Project, Create Ethics Sheet for Project participants, Review Ethics for project, Background Research, Research on Project	2023-05-29	2023-06-04
				2023-06-01		**2023-05-29**	**2023-06-04**

Week 8

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Design and Planning Continue		Robert	Working on it	2023-06-05		2023-06-05	2023-06-11
				2023-06-05		**2023-06-05**	**2023-06-11**

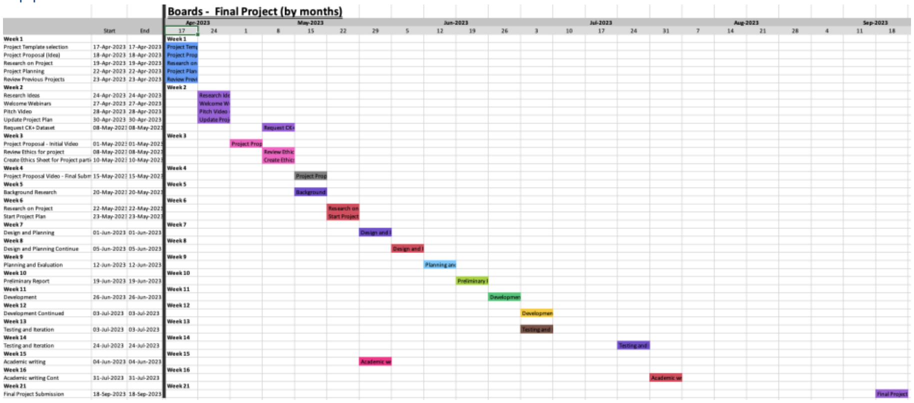
Week 9

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Planning and Evaluation		Robert	Working on it	2023-06-12		2023-06-12	2023-06-18
				2023-06-12		**2023-06-12**	**2023-06-18**

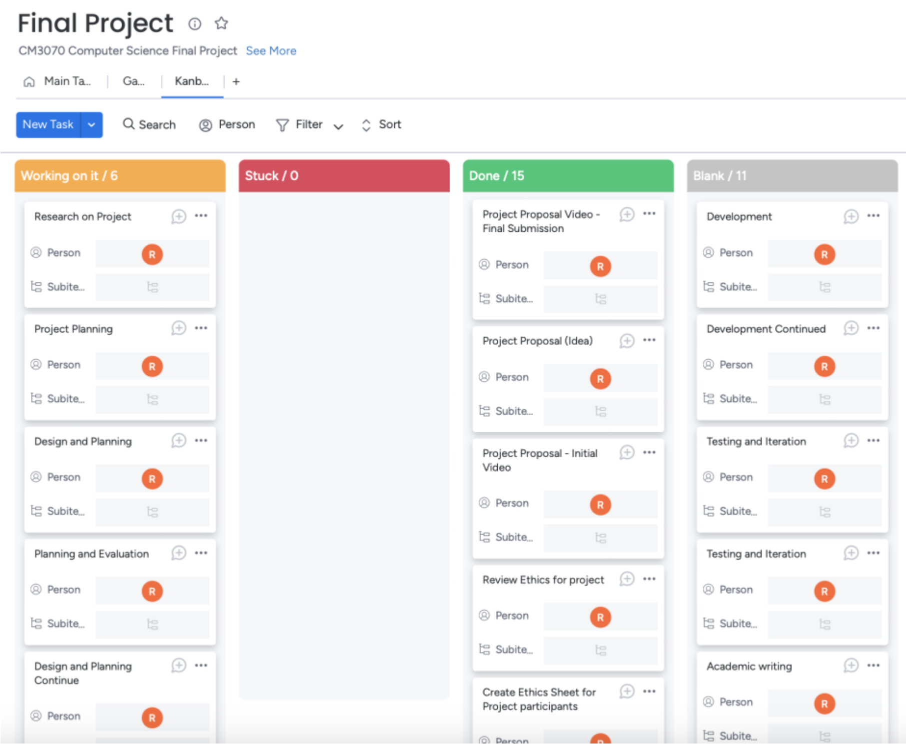
Week 10

Name	Subitems	Person	Status	Date	Dependency	Timeline - Start	Timeline - End
Preliminary Report		Robert	Working on it	2023-06-19		2023-06-19	2023-06-25
				2023-06-19		**2023-06-19**	**2023-06-25**

Appendix E: Timeline



Appendix F: Kanban View



Appendix G: Python Jupyter Notebook

jupyter RND Last Checkpoint: 14 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help | Not Trusted | Python 3 (ipykernel) | Logout

Prototype and R&D for the Feelings Score Development

In order to create the prototype, a couple of requirements are needed first.

The decision has been made to use **Wisper** from OpenAI to perform the sentiment analysis on a video call.

The **OpenCV: Open Source Computer Vision Library** will also be utilized for video manipulation and extracting the images of the subjects in the video.

ffmpeg-python serves as a Python wrapper for the FFmpeg multimedia framework. FFmpeg, an open-source software suite, is capable of handling multimedia data, including audio and video, which is required for this project.

For the prototype, Video's that have been published by individuals known to the project creator will be used. The subjects in the video have completed the consent form as per **Appendix A**.

A link to the consent form can be found in the reference section of the report.

To achieve the steps outlined in the project design section I have created the sections as below:

1. Identify relevant emotions and sentiment categories.
2. Determining Weights or Importance.
3. Collecting and Pre-processing Data.
4. Facial Emotion Recognition and Sentiment Analysis.
5. Normalizing the Scores.
6. Combining Scores into a Composite Index.
7. Validation of the Index Score.

The following libraries will be installed and used:

In [121]:
```
!pip install wisper
```

```
Requirement already satisfied: wisper in /opt/anaconda3/lib/python3.9/site-packages (1.0.0)
Requirement already satisfied: requests in /opt/anaconda3/lib/python3.9/site-packages (from wisper) (2.26.0)
Requirement already satisfied: psutil in /opt/anaconda3/lib/python3.9/site-packages (from wisper) (5.8.0)
Requirement already satisfied: werkzeug in /opt/anaconda3/lib/python3.9/site-packages (from wisper) (2.0.2)
Requirement already satisfied: boto3 in /opt/anaconda3/lib/python3.9/site-packages (from wisper) (1.24.2)
```

1. Identify relevant emotions and sentiment categories.

In [3]:
```
emotional_states = ['Angry', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprised', 'Neutral']
sentiment_categories = ['Positive', 'Negative', 'Neutral']
```

2. Determining Weights or Importance

Expert judgment is employed to determine the values assigned based on the emotion.

The emotion with the higest score is Happy : 3 The worst sentiment is Negative with a score of -1

In [4]:
```
emotional_weights = {'Angry': -3, 'Disgust': -3, 'Fear': -2, 'Happy': 3, 'Sad': -1, 'Surprised': 1, 'Neutral': 1}
sentiment_weights = {'Positive': 1, 'Neutral': 0, 'Negative': -1 }
```

3. Collecting and Pre-processing Data.

For the prototype, recorded videos of short duration will be utilized to demonstrate the functionality and accuracy of the approach in determining sentiment and facial recognition.

The extraction of audio from the video and the application of facial recognition will be employed to capture snapshots of the subjects in the video and perform emotional recognition.

Wisper will be employed for video transcription, with each spoken sentence being assigned a timestamp indicating the start and end time within the video. The average timestamp will be used to extract the subject's facial expression at the specified time and analyze it for emotion, assigning a relevant score.

The Open CV library will also be utilized, known for its speed and efficiency, making it a popular choice for real-time applications and performance-critical tasks.

In order to use the wisper openAI model you have to have ffmpeg instaled:

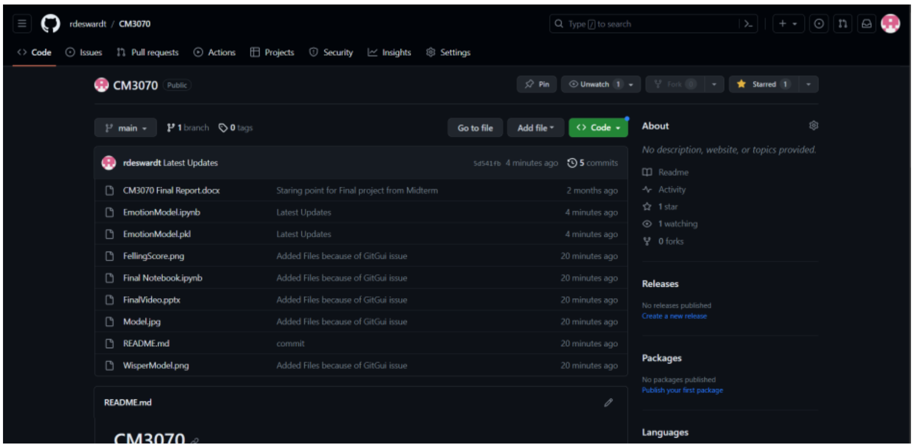
Linux: sudo apt update && sudo apt install ffmpeg

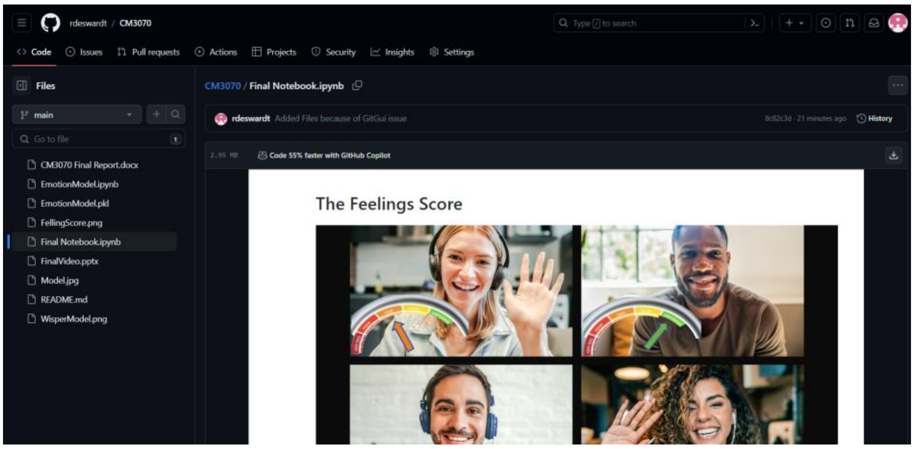
MacOS: brew install ffmpeg

Windows: chco install ffmpeg

ffmpeg guide

Appendix H : GitHub





References

[1] CHANEL, G., KRONEGG, J., GRANDJEAN, D. AND PUN, T. 2006. Emotion Assessment: Arousal Evaluation Using EEG's and Peripheral Physiological Signals. 530. .

[2] CLARK, E.A., KESSINGER, J., DUNCAN, S.E., BELL, M.A., LAHNE, J., GALLAGHER, D.L. AND O'KEEFE, S.F. 2020. The Facial Action Coding System for Characterization of Human Affective Response to Consumer Product-Based Stimuli: A Systematic Review. *Frontiers in psychology* 11, 920. https://search.proquest.com/docview/2412984813.

[3] DOMÍNGUEZ-JIMÉNEZ, J.A., CAMPO-LANDINES, K.C., MARTÍNEZ-SANTOS, J.C., DELAHOZ, E.J. AND CONTRERAS-ORTIZ, S.H. 2020. A machine learning model for emotion recognition from physiological signals. *Biomedical signal processing and control* 55, 101646. https://dx.doi.org/10.1016/j.bspc.2019.101646.

[4] GU, J., WANG, Z., KUEN, J., MA, L., SHAHROUDY, A., SHUAI, B., LIU, T., WANG, X., WANG, G., CAI, J. AND CHEN, T. 2018. Recent advances in convolutional neural networks. *Pattern recognition* 77, 354-377. https://dx.doi.org/10.1016/j.patcog.2017.10.013.

[5] HTTPS://IMOTIONS.COM/ABOUT-US/. About Us - iMotions. 2023, .

[6] JUNG, H., LEE, S., YIM, J., PARK, S. AND KIM, J. Dec 01, 2015. Joint Fine-Tuning in Deep Neural Networks for Facial Expression Recognition. In Anonymous IEEE, , 2983-2991.

[7] KANADE, T., COHN, J.F. AND YINGLI TIAN. 2000. Comprehensive database for facial expression analysis. In Anonymous IEEE, , 46-53.

[8] KHEZRI, M., FIROOZABADI, M. AND SHARAFAT, A.R. 2015. Reliable emotion recognition system based on dynamic adaptive fusion of forehead biopotentials and physiological signals. *Computer methods and programs in biomedicine* 122, 149-164. https://www.clinicalkey.es/playcontent/1-s2.0-S0169260715001959.

[9] KOELSTRA, S., MUHL, C., SOLEYMANI, M., JONG-SEOK LEE, YAZDANI, A., EBRAHIMI, T., PUN, T., NIJHOLT, A. AND PATRAS, I. 2012. DEAP: A Database for Emotion Analysis ;Using Physiological Signals. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING* 3, 18-31. https://ieeexplore.ieee.org/document/5871728.

[10] KOELSTRA, S., MUHL, C., SOLEYMANI, M., JONG-SEOK LEE, YAZDANI, A., EBRAHIMI, T., PUN, T., NIJHOLT, A. AND PATRAS, I. 2012. DEAP: A Database for Emotion Analysis ;Using Physiological Signals. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING* 3, 18-31. https://ieeexplore.ieee.org/document/5871728.

[11] LIU, Q., YANG, J., DENG, J. AND ZHANG, K. 2017. Robust facial landmark tracking via cascade regression. *Pattern recognition* 66, 53-62. https://dx.doi.org/10.1016/j.patcog.2016.12.024.

[12] LUCEY, P., COHN, J.F., KANADE, T., SARAGIH, J., AMBADAR, Z. AND MATTHEWS, I. 2010. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. .

[13] NOJAVANASGHARI, B., BALTRUŠAITIS, T., HUGHES, C.E. AND MORENCY, L. Oct 31, 2016. EmoReact: a multimodal approach and dataset for recognizing emotional responses in children. In Anonymous ACM, , 137-144.

[14] NOLDUSHUB. Noldus. 2023, .

[15] RESTREPO, D. AND GOMEZ, A. Oct 2017. Short Research Advanced Project: Development of strategies for automatic facial feature extraction and emotion recognition. In Anonymous IEEE, , 1-6.

[16] ROBERT DE SWARDT. 2023. Consent Form. 2023, .

[17] SHU, L., XIE, J., YANG, M., LI, Z., LI, Z., LIAO, D., XU, X. AND YANG, X. 2018. A Review of Emotion Recognition Using Physiological Signals. *Sensors (Basel, Switzerland)* 18, 2074. https://www.ncbi.nlm.nih.gov/pubmed/29958457.

[18] WANG, F., WU, S., ZHANG, W., XU, Z., ZHANG, Y., WU, C. AND COLEMAN, S. 2020. Emotion recognition with convolutional neural network and EEG-based EFDMs. *Neuropsychologia* 146, 107506. https://dx.doi.org/10.1016/j.neuropsychologia.2020.107506.

[19] WANG, Z., WANG, Y., ZHANG, J., HU, C., YIN, Z. AND SONG, Y. 2022. Spatial-Temporal Feature Fusion Neural Network for EEG-Based Emotion Recognition. *IEEE transactions on instrumentation and measurement* 71, 1-12. https://ieeexplore.ieee.org/document/9751142.

[20] ZHANG, Z., LUO, P., LOY, C.C. AND TANG, X. 2014. Facial Landmark Detection by Deep Multi-task Learning. 94. .

[21] ZHAO, X. AND ZHANG, S. 2011. Facial Expression Recognition Based on Local Binary Patterns and Kernel Discriminant Isomap. *Sensors (Basel, Switzerland)* 11, 9573-9588. https://www.ncbi.nlm.nih.gov/pubmed/22163713.