Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Computer Science and Statistics

MOdality and THought Unified - Mothú

Alanna O'Grady
14313298

April 10, 2019

A Masters Project submitted in partial fulfilment of the requirements for the degree of MAI (Computer Engineering)

Supervised by Professor Khurshid Ahmad

Declaration

I hereby declare that this project is entirely my own work and that it has not been submitted as an exercise for a degree at this or any other university.

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at http://www.tcd.ie/calendar.

I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write. anal-

ysis

Signed: ______________________ Date: ______________________

Summary

As sentiment analysis technology matures, the number applications to which it is applied continues to grow. One such application is Leadership Studies, through applying sentiment analysis to this area of research, an in depth revelation into the sentiments being felt by the leaders can be determined. It is important within Leadership Studies to understand the way in which a leader is communicating, as the sentiments which they express can be very influential on the mood and possibly the actions taken by their followers. As the majority of communication between humans is through non-verbal signals, it is particularly interesting to study leaders' non-verbal cues. In recent years, there has been great progress in the technology available for non-verbal sentiment analysis, although these systems remain underutilised due their poor usability. The aim of this project is to create a program which will automate the analysis of these sentiment results to reduce the burden of interpretation for the user.

The work within this project took place in four phases. The four phases consisted of:

- Data Collection
- Pre-Processing
- Post-Processing
- Data Analysis

As past leadership studies which have used sentiment analysis have focused mainly on male leaders, this research aimed to address this by expanding on prior studies by analysing the sentiments expressed by female leaders. The Data Collection phase consisted of impartially selecting the video footage in order to create a representative sample of the female leader population. The Pre-Processing phase involved manually editing the chosen sample videos to remove frames in which the subject was not centred within the frame, the subject was not looking towards the camera and in which another persons voice could be heard. The filtered sample data was then post-processed with the sentiment analysis systems FACET and openSMILE to get the facial expression and speech sentiment data, respectively. Within the Data Analysis phase, a program was created which would analyse the sentiment data of the subject to determine statistically significant results to help the user to understand the sentiment felt by the subject.

The analysis included in this program aims to reveal deeper insights into the sentiments felt by leaders. This program provided an overview of the sentiments conveyed by the leaders in both scenarios, through both non-verbal channels in the form of grouped barplots which will allow the user to easily understand the dominant sentiments felt by each leader in the different scenarios. This graphical representation will also allow the users to visually compare

the sentiments of each leader within each scenario, and from this create assumptions about the control of emotions by these women. Independence tests were used to determine a difference in the sentiments expressed in the emergency and the non-emergency scenarios. From this, an evaluation of the control which the leader exerts over their expression of emotions was performed to determine if their are any discrepancies within the leaders' behaviour. To verify the results of the independence test for the sentiments expressed through facial expressions, due to the fleeting nature of facial expressions, external methods were used. These external methods included independence tests using the audio sentiment results and the fused (facial-expression and speech) sentiment results. If the results from multiple modalities are in agreement, this suggests the results are reliable revelations into the behaviour of the leaders.

From the analysis performed, a deviance within the leaders' behaviour was determined from their facial-expressions. This was then validated with the chi-square test results from the audio and fused data as this confirmed that this change in behaviour was evident from multiple non-verbal channels. Based on the graphical representations of the sentiments expressed by the leaders, an increase in negative sentiments was evident in the emergency situations. From the evaluation of the sentiment analysis systems used within this project, FACET was found to perform correctly, whereas openSMILE was more prone to errors. FACET was deemed to be an accurate system, inaccuracies within its performance were generally found to be due to poor quality input data. openSMILE was found to perform less accurately, as there appeared to be a bias towards sadness and an unexpected positive correlation was observed between anger and happiness, revealing an issue within this system in differentiating sentiments of similar activation.

Based on the research conducted and the key findings, an insight into the leaders behaviour was sourced through unimodal and multimodal analysis. Although, for this technology to be used for practical applications, improvements must be made to the precision of the audio sentiment analysis for truly reliable results.

Abstract

Mothú translates to the Irish word for feeling, and it is this intrinsically complex cognitive process which we aim to understand through the program which was created as part of this research project. A large percentage of human communication consists of non-verbal communication, although in comparison with its verbal counterpart, non-verbal sentiment analysis remains a relatively new field. The non-verbal sentiment analysis systems currently available, although they are at the cutting edge of technology, they have not been widely adopted due to usability issues.

The program presented in this dissertation aims to analyse facial-expression and speech sentiment data in order to produce results which will allow the user to gain an insight into the sentiment felt by the subject, reducing the burden of interpretation. Previous studies which apply sentiment analysis techniques to the study of leaders have often focused on the male sex. This research aims to expand these prior studies by conducting an extensive investigation on the sentiment of a sample of female leaders. This study will consider the sample of leaders in both emergency and non-emergency situations to determine if a deviance within their expression of emotions can be determined.

From the statistical analysis performed and the graphical representations produced, a deviance within the behaviour of the leaders was found through unimodal analysis and then validated via multimodal analysis. As the validity of results presented in this project are a consequence of the accuracy of the sentiment analysis systems used, an evaluation of the performance of these systems was conducted. FACET, used for facial expression analysis, was found to be an accurate system whereas openSMILE used for speech analysis was found to have some inaccuracies. Therefore, the precision of the audio sentiment analysis must be improved for multimodal analysis such as this to be used for practical applications.

Acknowledgements

I would like to thank my supervisor Prof. Khurshid Ahmad for his support and encouragement during this research project. I would also like to acknowledge Prof. Ahmad's unpublished literature review (1) which was used as an aid during the writing portion of this project. I would also like to express my thanks to Prof. Ahmad's research group for the generosity they have shown me in terms of both advice and resources.

I would like to acknowledge the creators of both FACET and openSMILE not only for their invaluable sentiment analysis systems, but also for the extensive documentation provided.

Finally, I would like to thank my family for their enduring support and advice.

Contents

List of Figures

List of Tables

Listings

Nomenclature

r_{xy}	Pearson's Correlation Coefficient
χ^2	Chi-Squared Coefficient
AU	Facial Action Units
CERT	The Computer Expression Recognition Toolbox
FACET	The Facial Expression Analysis Engine
FACS	Facial Action Coding System
openEAR	Open-Source Emotion and Affect Recognition Toolkit
openSMILE	Open-Source Speech and Music Interpretation by Large-scale Extraction
SMILE	Speech and Music Interpretation by Large-scale Extraction
SNR	Signal-To-Noise Ratio
SVM	Support Vector Machine

1 Introduction

1.1 Project Outline

The Oxford English dictionary defines sentiment as a "view or opinion that is held or expressed", it is this which often governs the decisions made by humans. Due to this, sentiment analysis, the process of automatically detecting the sentiments being conveyed, is increasing in popularity as the availability of data continues to grow, a consequent of the increased use of social media and other websites as a means of broadcasting one's view to the world.

The behaviour expressed by leaders is a common area of study due to its' influential effects not only on their followers, but also on the public-opinion of the leader. Due to this, this project studies the sentiments expressed by leaders, using sentiment analysis to gain an insight into the behaviour of this sample of female leaders in various types of situations.

Non-Verbal Communication is one of the primary means of communication for humans, yet it still remains an exciting and dynamic area of research (6). As a result of this, the study presented will focus on sentiment analysis of the non-verbal signals expressed by the chosen leaders. The non-verbal channels being investigated as part of this study were visual (facial expressions) and audio (acoustic features of speech). A method of validating the sentiment results detected through facial expressions was required as facial-expressions are often short-lived, requiring external methods to be used. Audio sentiment data was chosen as the external validation, this was done through unimodal and multimodal analysis. A multimodal sentiment analysis allowed us to avail of the data from different channels, combining them to form a more accurate account of the sentiment expressed. Using facial expression and speech sentiment analysis was a natural choice for an investigation of sentiment expressed by non-verbal cues due to the efficient nature of these forms of analyses as well as the growth in the availability of other media such as videos and audio clips.

During this process, knowledge related to Non-Verbal Sentiment Analysis, and methods of extracting such data has been learnt. This research was an opportunity to create a program which automatically examines how facial-expressions and speech are impacted by sentiment through analysing the raw output from two major systems within this field: FACET and openSMILE. From this analysis, a deviance within the leaders behaviour was observed through the results of the Chi-Squared test performed on the facial-expression data. This result was then confirmed by the Chi-Square results for the speech and fused (facial-expression and speech) data. From evaluating the performance of the system used, FACET was found to be an accurate system, whereas openSMILE's performance often included inaccuracies. Therefore, until improvements are made to the speech sentiment analysis, multimodal analyses, such as this, will be deemed too unstable for implementation in real-world applications.

1.2 Research Question, Aim and Objectives

1.2.1 Research Question

It will be interesting to see women in key leadership positions, in a range of professions where it is important to not only be calm and rational but to be seen thus, both in emergency and non-emergency situations. The appearance matters in that there is a claim, perhaps with some justification, that sentiments can "leak", often through non-verbal signals. Leadership studies have stressed the importance of how a leader should be able to appear calm and confident at all times to assert his or her authority and empathy (7).

Leaders are individuals who make frequent public appearances, due to this, they are expected to behave in a "normal" way. As the sample under study consists of females who are leaders within their given fields, it is suggested that the sentiments they communicate are important to how these women are perceived. Often, the perception of their professional competency can be impacted by this expression of sentiment. Hence, these women are often well accustomed to regulating their sentiments in the presence of the public/media.

My Research Question:

From a combination of unimodal and multimodal sentiment analyses of this collection of female leaders in emergency and non-emergency scenarios, can deviance in their behaviour be determined automatically?

1.2.2 Aim and Objectives

The aim of this project is to investigate whether sentiment can be automatically detected from this sample through statistical analysis of the subject's facial-expression and acoustic speech data. The subjects chosen will be females who are considered leaders of their respective profession. Due to the importance of their position, how they communicate can have a high impact on various aspects of their lives (the organisation they work for, public opinion of the subject, etc).

Objectives:

- To become familiar with, and develop experience in sentiment detection methods, tools and techniques.
- To objectively source experimental data based on clearly defined criteria.
- To determine a method for automating sentiment detection by exploring whether or not we can seek confirmation of a signal of a sentiment in one modality (facial expression) by another (speech).

1.3 Technical Approach

The main focus of this research was to provide a comprehensive study of the expression of sentiments in female leaders. Due to this, it was of the utmost importance that the sample used was representative of the population. After sourcing the collection of videos, the data was filtered manually to ensure the data was of a high quality, which also met the criterion of both sentiment analysis systems.

The sample data was processed with FACET and openSMILE in order to determine the sentiments which were being expressed via the facial expression and acoustic non-verbal communication channels, respectively.

FACET is a facial expression analysis engine which has been developed by a team of specialist researchers led by Marian Bartlett, Javier Movellan, and Ian Fasel. This technology builds on over 20 years of machine-learning research, and has been inspired by the physiological and psychological system developed by Paul Ekman to taxonomise behaviour (the Facial Action Coding System) (8). openSMILE is a speech feature extractor which has been developed by Florian Eyben, Martin Wöllmer, and Björn Schuller. This system uses a combination of speech and music signal analysis to extract features from speech and due to this, it is widely used within the affective computing research community.

Unimodal analysis was performed on sentiment results from each modality. This allowed us to gain an insight as to the sentiments which the leaders were expressing primarily from that non-verbal communication channel. As part of this analysis, the occurrences of each sentiment were counted to determine the overall sentiment frequencies for each video. These sentiment frequencies were then used to determine overall sentiment for the scenarios in the form of data visualisation and chi-square test results.

A multimodal analysis was performed by aggregating the sentiment results from both modalities to determine the relationship between the sentiments expressed through each modality. This was also used as a method of validation of the sentiment detected through facial expressions. This data was analysed to determine how well these modalities supported each other and how often they agreed on the sentiment being expressed. A chi-square test was performed to determine if there was a significant difference in the sentiments expressed in emergencies and non-emergencies, the results of this test were then compared with the unimodal analysis to ensure the validity of these analyses.

1.4 Empirical Investigations & Evaluation Strategies

The results from the data analysis, specified above is discussed in the form of case studies. From presenting the results in this form, the outcomes for the empirical investigation done on the sentiment analysis results (the facial expression, speech and fused analysis results) are discussed in further detail.

As the accuracy of the results of this project are dependent on the accuracy of the sentiment analysis systems used, an evaluation of these systems were conducted. The main analysis consisted of determining the correlations between the sentiments which each system detected. The expectation was that for the system to be working accurately, there should be a strong negative correlation between positive sentiments (e.g. joy) and negative sentiments (e.g. anger and sadness). Another analysis was also conducted on FACET, this consisted of measuring the correlations between each sentiment and the action units which were active. The expectation in this case was that there should be a strong positive correlation between the sentiments and the Action Units from which it is derived.

1.5 Structure of Dissertation

Chapter 1 gives an outline of the project being presented, describing the research question it is addressing and the overall aim of this research. It also gives a brief description of the methods and evaluation strategies used as a part of this research.

Chapter 2 describes the motivation behind this research, including a critique of past studies in this area and the resulting research gap which this research aims to fill. There is an exploration into Leadership Studies and Sentiment Analysis, the primary areas of study within this project.

Chapter 3 gives a comprehensive description of the various stages within this project, including Data Collection, Pre-Processing (filtering), Post-Processing (sentiment analysis) and Data Analysis.

Chapter 4 presents the results of this research in the form of case studies. Both data visualisation and statistical measures were presented in an aim to highlight the importance of various aspects of the results and also to aid the interpretation of the results. The evaluation of the accuracy of the sentiment analysis systems and consequently the accuracy of the results obtained from them was also presented.

Chapter 5 comprises of a conclusion which has been drawn from the results gained during this project. Also discussed were challenges faced during the course of this research, and further work which could be done as a continuation of this research.

2 Motivation and Literature Review

2.1 Motivation

Sentiment is a feeling or view which is evoked in a person by a given subject. Non-verbal communication is a means by which humans communicate both sentiment and subliminal messages to each other. It has been argued by some scholars that non-verbal communication makes up 60-70 percent of how humans communicate with one another (6). Clearly this is a very important area of research regarding human behaviour, although compared with the field of verbal communication it is still in its infancy.

In leadership theory, it is thought that the positive sentiments expressed by a leader, can influence the behaviour of its followers through mood contagion (9) - this is a powerful attribute of a leader and for this reason, it has become an increasingly relevant topic in recent research relating to leadership. It will be fascinating to investigate not only the sentiments expressed by the leaders but also how they control this expression of sentiment in different scenarios.

The importance of sentiments conveyed by leaders and also that non-verbal communication research is in its early-stages has prompted the investigation conducted as part of this research project.

2.1.1 Non-verbal Communication

Non-verbal communication is a transmission of information by tactile, audio and visual channels. This a means of communication often used by humans which includes facial expressions, paralingual information and gestures/body movements. Charles Darwin was one of the first to investigate non-verbal communication in his study *The Expression of the sentiments in Man and Animals* (10) where he relates the non-verbal cues displayed by humans and animals to the sentiments/moods being felt. Work in this field was then continued by Ray Birdwhistell, whose work established the importance of non-verbal communication within human interaction. More recent studies are focused on affective computing, where artificial intelligence is used to recognise and process human affect states.

As the use of machines becomes evermore prevalent in everyday life, the need for systems which would enable human-machine interaction is evident. Systems which use verbal communication to allow this form of interaction are now readily available commercial products. Reis et al., (11), discuss the effectiveness of these systems by evaluating how well it interacts with carefully selected elderly persons - the results found that the intelligent personal assistants, such as Google Assistant[1] and Amazon Alexa[2], have scope to improve the effectiveness of the interaction to be a truly useful product for all. Although, if Ray Birdwhistell's assumption is correct (about non-verbal cues being the primary method of communication by humans), then these verbal communication systems are only the beginning of human-machine interaction. Research is currently being done to improve home robots to allow them to recognise sentiments from a human through facial expressions, although this work is still at an early stage as the only expression which the system recognises is a smile (12).

Human-machine interaction is only one example of a real-world use for sentiment analysis (which uses non-verbal communication), other uses for this innovative technology spans many industries. Non-verbal communication has particular uses within the domain of marketing analysis and brand strategy as non-verbal communication is an underused source of information in focus groups which can contain relevant data to the study at hand (e.g. how effective brand strategy is), but also aid in the decoding of verbal communication (13). Within the health-care industry, in the cases of communication disorders, such as Cerebral Palsy, there is a discrepancy between the pain recognised by a health-care professional and that felt by the patient (14), which indicates a need for a system which could automatically recognise sentiments through the non-verbal cues of the patient to ensure adequate pain-management by health-care professionals. Finance is an industry which has already adopted text-based sentiment analysis to analyse both the written text on a given subject, and the words spoken by an individual to predict changes on the market, although this method of analysis is considerably slower than its non-verbal counterparts. Research has been conducted, which relates the negative sentiments expressed through facial expressions by a CEO to spikes within the stock market (15) - indicating that this form of analysis would be particularly useful within this fast-paced industry. From the investigation of a small number of industries, there is a clear necessity for this technology within the future.

[1] https://assistant.google.com/

[2] http://alexa.amazon.com/spa/index.html.

2.1.2 Non-verbal Communication and Leaders

How charismatic a leader is perceived to be can be related to the amount of positive sentiments they express, which can also be associated with the positive sentiment aroused within their followers. Charismatic leaders have also been known to use negative sentiments to influence their followers in times of emergency (7, 16). Leaders can influence the affect state of their followers through a phenomenon called emotion/mood contagion, which can impact the perceived effectiveness of the leader (7). This contagion relies on non-verbal cues as the primary means of transfer (17), this phenomenon is dependent on the energy level with which the sentiment is expressed - as energy levels often correlate to an increased attentiveness of the viewer. So it can be assumed that leaders, particularly charismatic leaders, can influence the sentiments provoked within their followers - which begs the question, can leaders use this contagion to manipulate the sentiments of their followers and thereby, how they themselves are then perceived by their followers?

The term "emotional labour", is often applied to employees which are required to display the correct sentiment for their job, this could also be applied to a leader. Within the role as leader, both the perception of the leader (and the organisation they represent) and the impression this has on their followers may cause the leader to try and control the sentiments they express. This method of emotion regulation is generally used by leaders to manage emotion contagion within their followers, often enabling them to be more effective as a leader (18). Methods of emotion regulation include *Surface Acting* and *Deep Acting*. Surface Acting is the process by which an individual pretends to feel a particular sentiment (possibly suppressing their true sentiments), whereas Deep Acting is where an individual performs an action which allows them to change the sentiments they are actually feeling/displaying. As emotional labour is an important aspect of leadership, it has been found that leaders can be trained in these methods of emotion regulation which correlates to an improvement in leadership effectiveness (19). A method of analysing this expression of sentiment would enable the examination of a leader's emotion regulation to be done efficiently and accurately, giving us an insight into leadership behaviour in different scenarios.

In a study to differentiate the irrational and rational aspects of emotion in the decision making process, it was found that "emotions can reflect the functional rationality for decision making in times of uncertainty" (18). Given the significance of emotions in times of difficulty and the effect of emotion contagion on followers, the importance of the sentiment displayed by a leader at these difficult times is clear. Spyropoulou and Ahmad used FACET to detect sentiment through facial expression data of a set of male politicians in disaster-type scenarios and from this, calculated how much engagement, attention and sentiment was displayed by the politicians, (16) - this type of information could be useful to leaders wishing to communicate effectively with their audience during a disaster. This was an excellent example of investigating the sentiments which leaders display in emergency scenarios, although it lacked diversity within the sample investigated as it was limited to the male sex and to subjects of a single profession. What remains to be done is an extension of this research using a sample of female leaders of a range of professions in both emergency and non-emergency scenarios.

2.2 Sentiment Analysis

Sentiment Analysis, also known as Opinion Mining refers to the use of text processing, computational linguistics or biometrics to extract affect states (emotions) and from this, gain an insight into the subjective meaning being expressed. Sentiment Analysis is concerned with the 'opinion holder' and the 'sentiment' it has for a particular 'aspect' of a given 'entity' (20). Sentiment Analysis is used to analyse both verbal and non verbal communication in order to gain information relating to the emotional state of the subject under study.

Due to the accelerated growth in data, a sophisticated means of analysing this data must be determined which can provide an efficient and accurate mechanism of automatically processing this data. The aim of Automatic Sentiment Analysis is to computationally deduce the attitude or position a person has towards a particular entity.

As data has become available in different forms (such as text, image, video and voice-recordings), other methods for analysing this data was found. This has resulted in the various modalities for extracting sentiment from data (e.g. Text-based Sentiment Analysis, Audio-based Sentiment Analysis, etc.), which will be explored further in the following sections.

2.2.1 Text-based Sentiment Analysis

This method has been the primary focus within the field of Sentiment Analysis thus far. Research relating to the field of Sentiment Analysis began in the early 20th century, although the use of computer-based techniques for sentiment analysis began with the increased availability of text (on the web, etc.) (21). Research done in this topic mainly resulted in methods which used a rule-based classifier and a lexicon (a dictionary of words which are tagged to show that they relate to a particular sentiment or polarity) or data-driven methods which use large data-sets which are annotated for polarity (22).

The General Inquirer, developed in Harvard in 1961 (23), was one of the first systems used for content analysis. This system consists of a lexicon of words which are tagged to convey the psychological, sociological and part-of-speech (verb, noun, etc.) information which relates to this word. 83 tags are used to categorise the words within the dictionary, the tags consist of "legal", "academic" and "religious" to name but a few. This system is complex and rich in information although, due to its spreadsheet format it can easily be incorporated into computational applications. Stone and Hunt (23), discuss a method of incorporating the General Inquirer and the Hunt Concept Learner to produce a method for automatic theme analysis - an example of how this system has been used to further develop sentiment analysis systems.

Text-based Sentiment Analysis, although it is computationally less expensive than its counterparts, it has disadvantages which impacts significantly on the accuracy and efficiency of this technique. An issue which often arises for methods of text-based analysis is that words can convey different meaning depending on the context in which they are used. Developing a system which can convey the correct meaning in terms of the context is still being researched. The length of time which requires text media to be written, approved and then published is at least hours and generally days - due to this, text data can be insignificant at the time it is made available and therefore less useful than other forms of data (visual or audio), making it less suited to the fast-paced industries which would benefit from this technology.

2.2.2 Audio Sentiment Analysis

At the expense of some linguistic ambiguity, I will use acoustic speech sentiment and audio sentiment interchangeably. The increased availability of audio data along with the improvements in machine learning technology, have made the extraction and analysis of acoustic features from audio data possible. This analysis extracts and analyses acoustic features such as pause duration, pitch and voice intensity to detect emotion from the speaker. This type of Sentiment Analysis is often used in Multimodal Sentiment Analysis where the acoustic features of audio data can be used to disambiguate linguistic meaning (22).

For a given recording, from the audio signal, acoustic features such as pause duration, pitch, voice intensity and loudness are extracted. The data vectors are convolved with filters and become feature vectors, the input of the SVM. The SVM (which has been trained on datasets of audio recording, annotated with the sentiment it conveys) then classifies the sentiments detected.

Used in unimodal analysis, this form of sentiment analysis performs significantly worse than other unimodal analyses, text-based and visual-based (22). Although, as discussed in (24), pitch is an acoustic feature which when used alone, even without textual information (in unimodal analysis) can provide significant accuracy. It has been shown that audio-based sentiment analysis used in combination with other modalities can provide significant accuracy as it aids linguistic disambiguation (helps determine the meaning), the linguistic sparsity problem (as more sentiment data is available) and grounding (connecting it to the real-world environment) (22).

2.2.3 Visual Sentiment Analysis

Due to the increased availability of visual data (video and images), methods for accurately and efficiently analysing this data is becoming an increasingly popular topic within Sentiment Analysis research. Thanks to advances within Computer Vision, methods for processing this type of data are becoming increasingly more accurate and efficient - thereby making this a viable piece technology to incorporate into commercial products such as the Apple Animoji feature on the iPhone X. The Animoji feature uses Facial Expression-based Sentiment Analysis to detect the expression on the users face and alter the animation to correspond to this (Figure 2.1).

Figure 2.1: An image showing how the Animoji corresponds to the users facial expressions from (2). This Animoji technology uses the mechanisms included in FACET for sentiment recognition to mimic the expression of the user. This technology uses the images captured from the depth-sensing camera to analyse the active action units or muscles which have moved and dynamically mirrors this expression in the emoji.

Gesture-based Sentiment Analysis

The use of analysing body-gestures is a technique commonly used by humans, in the education and healthcare fields, to ascertain the affect state of their students and patients respectively (25). This mode exploits the suggestion that "the correlation existing between body movements and spoken user sentence(s) can be used to reveal user's emotions from gestures" (26). Developments in Machine Learning and Computer Vision technology have allowed this analysis to be computed automatically.

Body-gestures are generally used in combination with other modalities. Gentile et al. postulates that if there is a correlation between sentences and gestures, that there is also a correlation between the sentiment recognised via text-based analysis and the sentiment recognised from a gesture. Based on this assumption, a solution is described which uses a combination of sentences and the gestures which appear with them. This method uses a Naive Bayes classifier which is trained on a dataset of sentence-gesture pairs to compute gesture-emotion pairs, an emotion can then be recognised from a new gesture by finding the most similar gesture (nearest neighbour) within the gesture-emotion pair dataset from the training stage (26).

Body gestures often occur as a means of emphasising a particular part of speech - for this reason, it would be impractical to use this modality alone as occurrences of these gestures are sparse (when comparing to the frequency at which facial expressions occur). There is evidence that this mode of analysis can provide vital information related to the affect state of the subject being studied and that recognition of particular sentiments can perform better than other modalities (25). Although compared with other non-verbal communication channels, research into the association between sentiments and body-gestures is a comparatively new field, therefore technology in this area would not be mature enough to be of use in this project.

Facial Expression Sentiment Analysis

Improvements in computer vision technology allowed the Facial Action Coding System, FACS, to be automated. This enabled facial expression analysis to benefit from the years of research done by Eckman and Bartlett related to facial expression recognition. This form of sentiment analysis can be used for unimodal analysis (when used alone), as it has been found to have a high accuracy (22), although it is also used in multimodal analysis, in which its accuracy increases further.

Most Facial Expression-based Sentiment Analysis techniques consist of detecting a face within the frame, analysing the facial action units which are active over time, and from this, the sentiment being expressed can be determined. This method relies on SVMs which are trained on a dataset of facial expression data of which the corresponding affect state is known. New visual data is processed to get a feature vector (of data relating to facial features) which is then input into an SVM which determines the expression/s in the dataset which it is most similar - from this the corresponding sentiment is then determined.

According to Rosas et al., (22), Okao Vision, a rudimentary system for extracting facial expression data (which mainly recognises smiles) was found to be approximately as accurate as text-based analysis. From this, a conjecture can be made which suggests that with a more mature system (such as iMotions' FACET) facial expression-based analysis could surpass the performance of text-based analysis.

The efficiency of this form of Sentiment Analysis is greater than that of the other methods, particularly the text-based approach which often require analysis with large lexicons. This method of analysis is becoming increasingly important with the growth of video data on websites such as Youtube,[3] Vimeo[4] and Instagram[5] which have become increasingly popular as platforms for people upload videos voicing their opinion or view on a particular subject - this availability of data provides the opportunity for sentiment analysis to be applied to it in order to aid industries such as market research or business.

2.3 Technology Used

The systems being used for this research is the iMotions platform and openSMILE. The Emotient FACET engine within the iMotions platform will be used for facial expression analysis. The audio analysis will be done with openSMILE which will use the openEAR emotion recognition models to determine the sentiment. The following section will discuss the history and design of these technologies which will reveal to us the both the benefits and limitations of these systems.

2.3.1 iMotions

iMotions is a software platform which aids biometric research by providing resources such as facial expression analysis, eye tracking and EEG information (27), to name a few.

The module within the iMotions platform which was used for the post-processing of the data is FACET. FACET, the Facial Expression Analysis Engine, consists of facial recognition software created by Emotient, based on 20 years of research on facial expression/emotion recognition. FACET determines the facial expression of a given respondent in a frame-by-frame analysis in which the activity of action units of the face are used to determine the sentiment being expressed (28).

[3]https://www.youtube.com/
[4]https://vimeo.com/
[5]https://www.instagram.com/?hl=en

2.3.2 Evolution of FACS

The Facial Action Coding System, FACS, was developed in 1970s by Ekman and Friesan and was a method of analysing facial movement in terms of actions to determine the affect state (sentiment) being expressed (8). These facial action units were defined as the changes over a sequence of images of facial muscles which contract or relax during the expression of a given sentiment. 46 action units were defined, with the use of these, human experts (FACS coders) could decompose an expression into the action units which produced it (8). An example is shown where the action units corresponding to a particular affect state have have been labelled (Figure 2.2). Human FACS experts required at least 100 hours of training with the resulting rate of analysis being 1 minute of footage per hour (8), this expensive and exclusive process limited the accessibility of analysing of facial expressions. Thus prompting the automation of this complex and expert process, leading to the invention of CERT and more recently FACET - tools which have greatly increased the usability of this form of analysis.

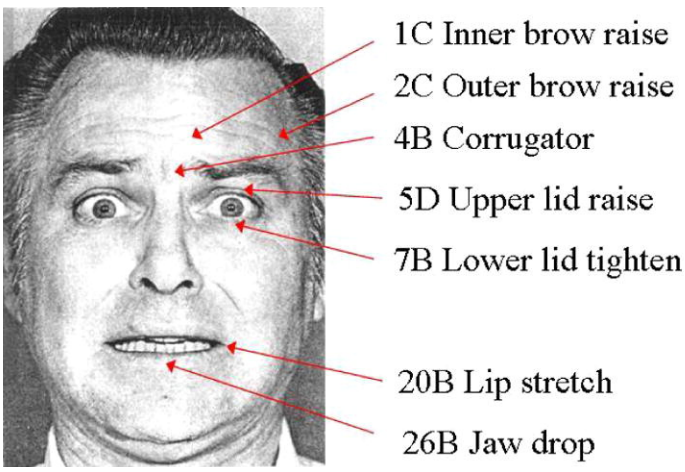

Figure 2.2: An example of FACS, indicating a subset of the defined action units, from (3).

Thanks to the advances in Computer Vision, the process of analysing facial action units was automated - which was the CERT system. The Computer Expression Recognition Toolbox (CERT) was a software which provided analysis of facial expression data outputting information related to the 19 Action Units, and the 6 basic emotions (happiness, sadness, disgust, surprise, fear) (29). CERT was freely available for academic purposes and was deployed by over 130 universities.

CERT's commercial successor was FACET (the Facial Expression Analysis Engine). FACET is a software that takes a video input of a subject and from this, will output the facial expression data relating to the sentiments they are expressing (in the form of Evaluations (positive, negative or neutral) which is derived from the 7 Affect States (basic emotions, e.g. joy, anger, surprise, fear, contempt, sadness, disgust) which are derived from 20

Action Units. By providing the 20 AU channels (Appendix A1.1 Table A1.1), it allows for more in-depth expression analysis to be conducted. Emotient, the start-up which created FACET was bought by Apple in 2016 (30), which resulted in the FACET SDK (Software Development Kit) being integrated into the iMotions biometric platform.

2.3.3 FACET

The process by which FACET operates is similar to that of its predecessor CERT, which consists of applying the following steps (29):

- Face Detection
- Feature Detection
- Face Registration
- Feature Extraction
- Action Unit Recognition and Intensity
- Basic Emotion Recognition and Pose Estimation

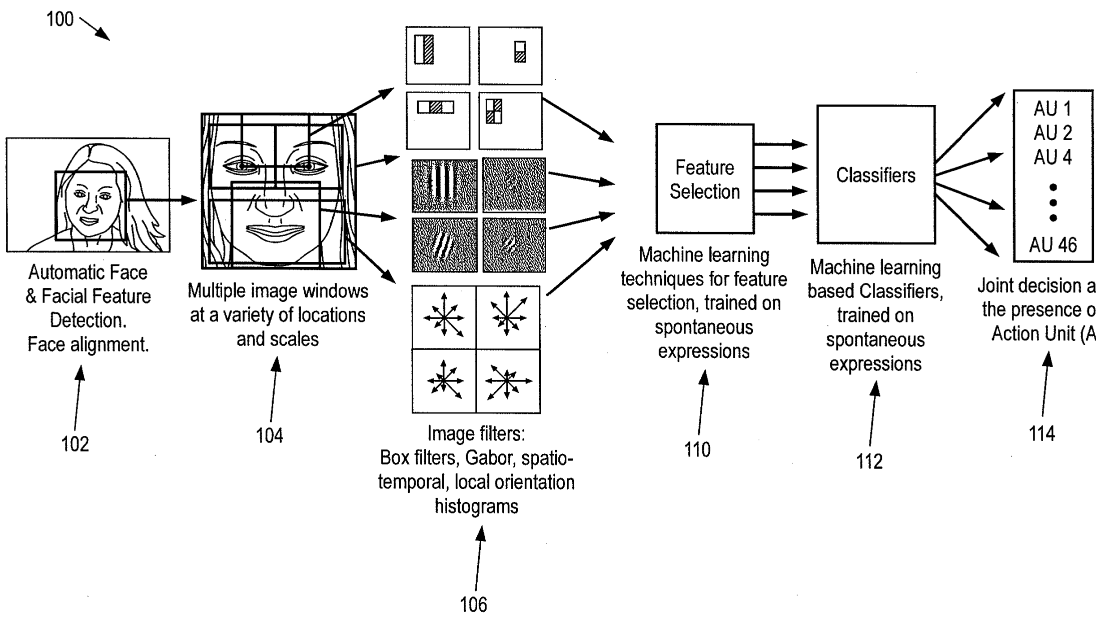

Figure 2.3: Automated facial expression analysis process used in CERT and FACET, from (4). The system initially detects the face within the current frame and then performs segmentation. These segments are then analysed further to detect the features within the face and their current location. The location of the features is then used to adjust the face patch to fit the face more accurately. The 2D array of this face patch is convolved with Gabor filters of varying scale and spatial frequency to give a single feature vector. This feature vector is input into classifiers to determine the active AUs and based on these results, determine the sentiments being conveyed.

Face Detection

The position of the face within the frame is found by applying a variation of the Viola Jones Cascaded Classifier Algorithm. The form of this algorithm used in both FACET and CERT uses GentleBoost as the boosting algorithm and WaldBoost for the automatic cascade threshold selection (29). This results in a face box appearing around the largest detected face which is segmented and used for further processing (Figure 2.3).

Feature Detection

The relevant facial landmarks (features, e.g. eyes, eye corners, brows, etc.) are detected using a feature-specific detector which was trained using GentleBoost. This detector then outputs the log-likelihood ratio of the feature in question being present at a given location (x, y). The log-likelihood ratio is combined with the prior locations of a given feature to give the posterior probability of a feature being located at (x, y) (29). The feature location estimates are then further refined using linear regression so that only the locations with a maximum likelihood estimate are used for further processing.

Face Registration

A face patch (internal face model) is scaled and its position adjusted using an affine warp to correspond with the detected face and the facial landmarks detected in the previous step (29) - the aim here is to minimise the difference between the detected face and the theoretical position of the face based on the features locations. This face patch is applied to the respondents face and adapts as the facial-expression changes. The face patch consists of less facial features than that actually found in the respondents face (i.e. using the brow corners rather than the entire brow arch), although it uses enough features to sufficiently detect activity in a given AU. The pixels of the resulting face patch are then extracted into a 2D array for for further processing (29).

Feature Extraction

The 2D array from the previous step is convolved using a Fast Fourier Transform with various complex Gabor filters to produce a single output feature vector. The biologically inspired Gabor filters remove the variability introduced in images by light and contrast to model the response properties of visual cortical cells (8) and therefore is particularly useful for texture analysis and discrimination. Gabor filters of different spatial frequencies and orientations are used. This multi-scale property is useful as it allows the results of the various frequencies and orientations to be combined to give more accurate results. Although it was found that the higher spatial frequencies are of particular importance in this process as they seem to contain more information than the lower frequencies (8).

Action Unit Recognition and Intensity

The feature vector is input into an SVM for each Action Unit which outputs the distance of the single feature vector to the SVM's separating hyperplane for that AU (29), this results in an intensity/evidence value of the activity for the given AU. The SVMs uses Supervised Learning which is a learning regiment used in Machine Learning which uses a large training set which consists of the inputs and the corresponding correct output with which the system is trained. The SVMs in CERT used a combination of well known facial expression datasets such as Cohn-Kanade (31) . Whereas FACET uses a training dataset acquired through crowd-sourcing, this approach was taken as it would likely create a training set consisting of "a large number of sample images of different people, in different illumination conditions, of various ethnicities and different ages, and with a range of facial artifacts" (32) - this diverse training set should allow the SVMs to perform more accurate recognition on subjects of varying genetic and biological factors.

Basic Emotion Recognition and 3D Head Pose Estimation

The AU channels, score vectors consisting of the facial action code score for each AU (evidence), are input into each emotion detector to compute the evidence of this emotion being present given the AUs which are active. The emotion detectors consist of multivariate logistic classifiers which have been trained using the computed AUs of the training data. This form of classifier is used when there is one dependent variable (emotion) and multiple independent variables (AU intensities) and the goal is to predict the probability of the dependant variable (emotion) based on the independent variables (AU intensities).

FACET can also be used to estimate the orientation of the subjects head. During the face registration stage, the face patch is passed through a range of classifiers to determine the ranges of yaw, pitch and roll. The output of the classifiers are combined with the location of the facial features and then linear regression is performed to refine the final values for yaw, pitch and roll. The head pose is an important aspect within visual sentiment analysis as these gestures can be used to determine how attentive the subject is (based on the yaw - whether the subject is facing towards or away) and also whether the subject is disagreeing (oscillating yaw - shaking there head).

2.3.4 Benefits and Limitations of FACET

When a system such as this is used for research, both the benefits and limitations must be considered as the accuracy of the system will be reflected in the results.

The development of FACET follows on the pioneering research of facial expression analysis conducted by Paul Ekman (FACS). This research lead onto revolutionary research in the area of automated sentiment recognition through behavioural analysis (analysis of facial expressions), which was developed by specialists within this domain and as a result of this, it is one of the most accurate and efficient systems available on the consumer market.

Limitations to the FACET engine are mainly concerned with the quality of the data on which the system is trained and its throughput rate. As the system uses Machine Learning technology, a very large training set is required to enable the system to learn how to recognise particular sentiments. The developers of FACET sourced the training data through crowd-sourcing in which the respondents were required to *act* a given sentiment (32). It is believed that this system will be able to recognise sentiments from the facial expression of an individual no matter the age, gender or race, although, given that the training data was sourced in this manner, it cannot be guaranteed that the training set is representative of the population. Recent advances in integrated circuit technology has allowed an increase in processing power, though the frame by frame processing power of approximately 30 frames/second is considerably lower than that of the human eye (which is believed to have the capacity to interpret up to 1000 frames/second (33)), which prompts us to question the accuracy of the system as compared with that of the human eye.

2.3.5 openSMILE and openEAR

openSMILE (Speech and Music Interpretation by Large-scale Extraction) is a flexible, modular system which unites features from both Automatic Speech Recognition (ASR) and Music Information Retrieval (MIR) to give rise to a robust feature extractor (34). Due to its modular architecture, openSMILE is modality independent, although it is primarily used for audio feature extraction. It has been extended to not only support audio feature extraction but also basic visual feature extraction. The SMILE feature extractor can be used with the openEAR (Emotion and Affect Recognition) toolkit models to source data about the sentiments of a person at a given time simply based on a sample of their speech.

openSMILE Architecture

The main component of the SMILE extractor being the Data Memory which enables the efficient incremental processing of data which is managed by the ring buffer. The Data Source writes the audio data (from a wave file or a sound card) to the Data Memory. Various Data Processors are then used to read and modify data to compute both the low-level descriptors (LLDs) and the functionals (which are commonly used for affect recognition) (5). The Data Sinks are used to write the features to files (CSV, WEKA) or to perform classification, such as sentiment recognition.

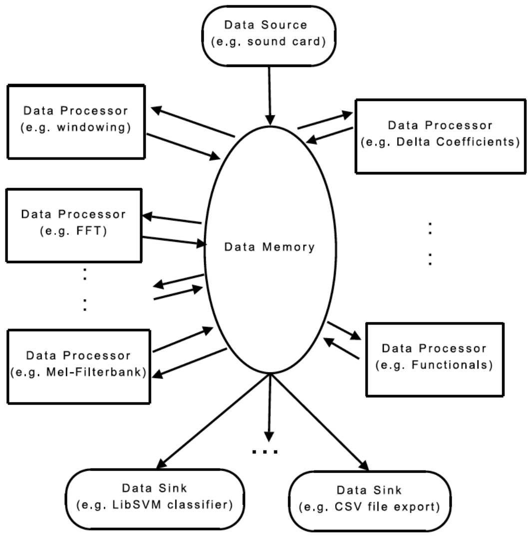

Figure 2.4: SMILE feature extractor component diagram, from (5).

2.3.6 openEAR

openEAR is an extension of the openSMILE framework which allows for automatic affect recognition from speech. There are three main components to openEAR, the core component being the SMILE extractor tool (described above), the support for classification modules in the feature extractor and the pre-trained emotion recognition models (also the ability to train your own models). The emotion recognition models are trained on various speech emotion databases, such as Berlin Speech Emotion Database, the eNTERFACE and the Audio Visual Interest Corpus, to name a few. In (5), a benchmark is performed which shows that the Berlin Speech Emotion Database performed better overall for affect recognition tasks. Due to the improved performance shown in the benchmark evaluation, the models which have been trained on the Berlin Speech Emotion Database will be used for this research project.

Berlin Speech Emotion Database

The Berlin Speech Emotion Database (EMO-DB), is an open-source database of emotional speech (in German). It consists of 10 sentences performed by 10 actors (5 female and 5 male) to convey seven different emotions (neutral, anger, fear, joy, sadness, disgust and boredom). Everyday sentences were used for the recording as it is "the natural form of speech under emotional arousal" (35). Professional actors were used as clear emotional expressions are rare in everyday life, but there is also possible ethical issues relating to recording a person's real emotional expressions.

2.3.7 Benefits and Limitations of openSMILE/openEAR

There are many benefits to using openSMILE/openEAR for research relating to sentiment as it is a renowned open-source system which has been greatly adopted by the computational paralinguistics research community. The modular architecture of openSMILE/openEAR allows for the this system to be used with other modalities such as visual data, ECG measurements, thus enabling this software to be a robust and versatile feature extractor. This system allows for the processing of live audio recordings or alternatively off-line processing (with pre-recorded data) - this flexibility enables this system to be used in a broader range of experiments.

Limitations to this system are mainly related to the databases which the sentiment recognition models have been trained on. These databases use professional actors to utter everyday sentences while *acting* a given sentiment. Due to this, there are concerns about the differences between the genuine expression of emotions and the artificial expression captured in these databases, although the databases which have been created using genuine expressions are usually private and limited in the emotions which they have captured (36). The video processing module within openSMILE is still at a relatively basic stage as it is generally used for classification of a subject (differentiating between gender and ethnicity). As this research relies on both audio and visual sentiment data, a more mature system, such as FACET (described in Section 2.3.3) is required for the facial expression sentiment analysis.

2.4 Summary

With the rise of human-machine technology, sentiment analysis has become an increasingly important area within computing. Improvements in computer vision and computational paralinguistics have allowed the accuracy of these systems to increase, causing many industries to gain interest in this technology. These system tend to be of particular interest to analysts (financial, political, etc.) who may be able to use the resulting sentiment data to make predictions.

As this technology can often be used for analysis and predictions, the study of leaders were considered as the analysis of the expression of sentiments shown by leaders is often of interest. From previous research conducted in this area, it was clear that there was a lack of research on the emotional expressions of female leaders as the previous research considered a sample which was confined to leaders of a single profession (16), indicating a gap within this research. The results from sentiment analysis systems are often large and complex, therefore a method of automatically determining the statistical significance of this data is needed to increase the adoption of this technology.

The lack of research conducted on non-verbal emotional expressions of female leaders and the improved accuracy of the sentiment analysis systems discussed in the previous sections, prompted the creation of a program which from performing data fusion, can automatically determine the statistical significance of the sentiment data. This program will be described in the following chapter, Chapter 3. I understand that sentiment in speech and sentiment in facial-expression are different modes of expressing sentiments. These two may be contemporaneous or one may lag whilst the other is leading. Moreover, there is a possibility that there may be no causal connection between the two. I am merely exploring the question of this relationship.

3 Method

In the previous chapter, the sentiment analysis systems used in this project have been explored and found to be the best technology available, although the usability of these systems could be improved upon. The program produced as part of this research project aims to improve the usability of these systems, using statistical techniques to reduce the burden of interpretation of the sentiment data for the user.

This program also allows us to test how well these systems, which were created in a lab environment, performs on real-world data. Behavioural analysis is a practice which began in the wild with Charles Darwin's study relating the emotions of man and animals to their non-verbal cues (10). The automation of behavioural analysis has brought this form of study to a laboratory setting, as these systems have been developed on carefully created and controlled data. This project will investigate the performance of these systems on non-laboratory data, bringing behavioural analysis back to operating in the wild.

In this chapter, the techniques used for the data collection, pre-processing, post-processing and data analysis are discussed at length, and the incentives for each technique is clarified.

The following diagram gives an overview of the technical approach used for this project:

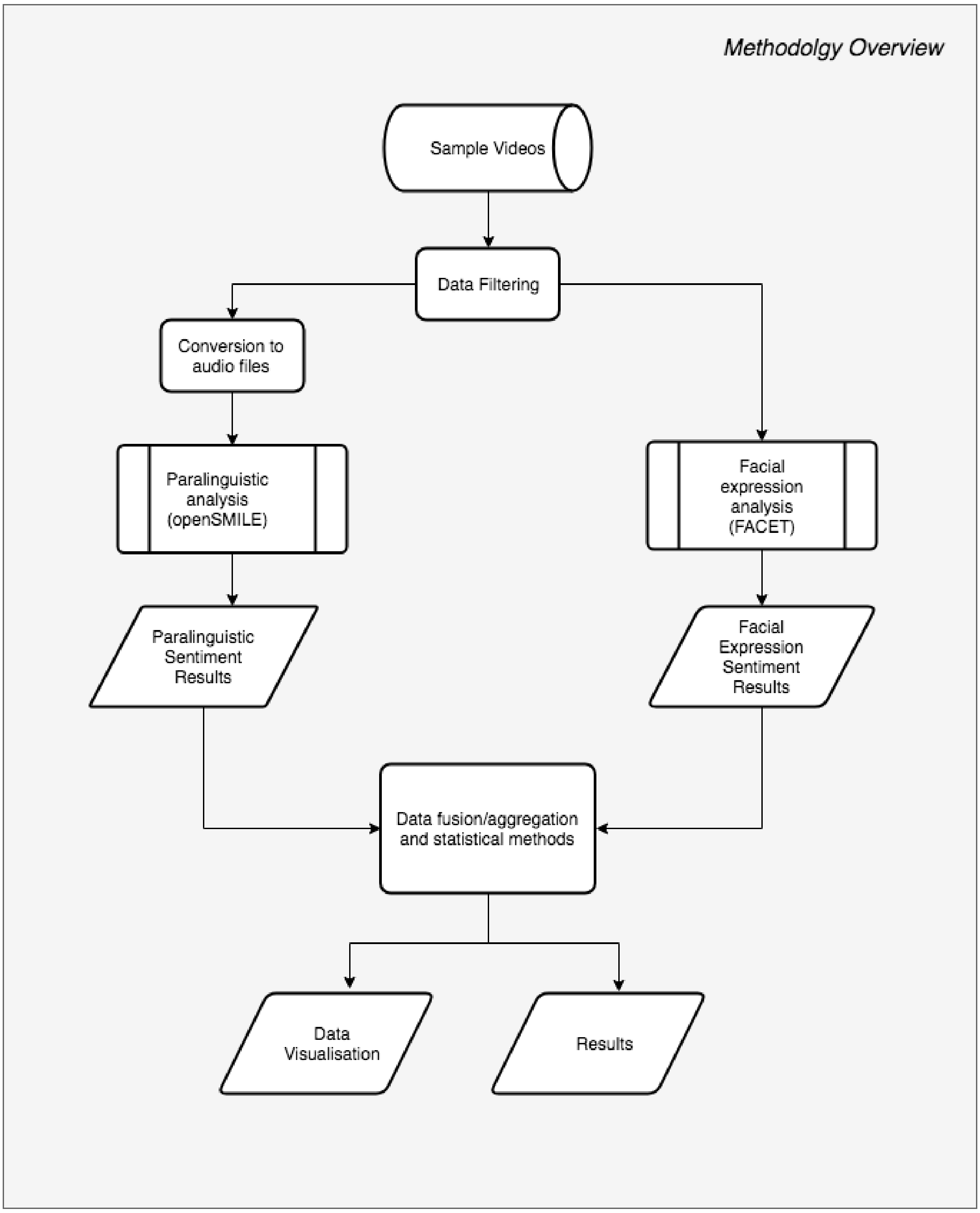

Figure 3.1: Overview of project method. The sample video footage of the chosen female leaders are filtered to remove frames which don't meet the criteria of FACET or openSMILE. The filtered data is then processed with FACET to output the raw facial sentiment data, the filtered data must be converted to a WAV file before processing with openSMILE can be performed. Unimodal and multimodal analyses are then performed on the resulting facial-expression and speech raw sentiment results to produce results in the form of statistical measures and graphical representations.

3.1 Data Collection

3.1.1 Sample Chosen

The method with which the training data for FACET was collected, consisted of crowd sourcing in which the respondents were asked to *act* a given affect state. Due to this, the population chosen for this investigation was leaders because of their ability to appear at ease in front of the media, and therefore they could be thought of as "semi-trained actors". The behaviour of a leader can be very influential, particularly in times of difficulty, for this reason, the sentiment displayed by the leaders will be considered in both emergency and non-emergency scenarios.

Within the field of facial-expression sentiment analysis, particularly that which focuses on the detection of sentiment within leaders as in (16), has mainly focused on male leaders. As a result, this study focuses on women who are leaders within their given profession. By focusing on female leaders, not only will this be extending previous work, but it will allow both sentiment analysis systems to be examined for gender bias within the matching algorithm.

The analysis for this project will mainly consist of Descriptive Statistics, of which further details can be found in Section 3.4. How accurately the Descriptive Statistics relate to the population relies heavily on the selection of the sample. Therefore, the selected sample must be representative of the population being studied. With this in mind, during the Data Collection stage, the aim was to create a diverse sample of female leaders, for this reason, the women chosen vary with regards to age, ethnicity and profession (Table 3.1). This sample consists of of four Americans (two American/Latina, one American/Ashkenazi Jew), two Europeans (one Irish, one British), and one Asian. This variance of race/nationality will allow us to evaluate the accuracy of the sentiment analysis systems on varying ethnicities. The chosen leaders' professions comprised of CEO, politician, judge and activist. This would allow us to investigate the sentiment shown by the leaders in different roles and possibly infer if there are particular behavioural tendencies in particular professions.

The sample consists of video footage of the chosen leaders in emergency and non-emergency scenarios, which was produced in a media setting where the lighting and position of the camera were controlled. The sample footage was sourced through Youtube [1] where the videos have been made available for public consumption.

[1] https://www.youtube.com/

Table 3.1: Review of samples chosen for this experiment. This consists of four American leaders (two American/Latina, one American/Ashkenazi Jew), two European leaders (one Irish, one British), and one Asian leader. The professions comprised of CEO, politician, judge and Human Rights Activist.

Name	Profession	Ethnicity	Date of birth	Emergency Situation	Non-emergency Situation
Sheryl Sandberg	COO Facebook	American/ Ashkenazi Jewish	28/08/1969	Death of Husband	Motivational talk
Heather Bresch	CEO Mylan	American	27/07/1969	EpiPen price controversy	Interview about Biosimilars
Mary Robinson	Politician	Irish	21/05/1944	1983 Abortion Referendum	Climate change talk
Sonia Sotomayor	Judge	American/Latina	25/06/1954	Interview about suitability for job	Interview about book
Malala Yousafzai	Human Rights Activist	Pakistani	12/07/1997	Speech on return to Pakistan	Talk at Harvard
Alexandria Ocasio-Cortez	Politician	British	13/10/1989	Interview about current state of political system	Chat Show Interview
Theresa May	British Prime Minister	American/Latina	01/10/1956	Press Conference about Brexit deal	New Year's Message

3.1.2 Criteria for collection

The FACET system has a criteria for which the data it analyses must follow. This criteria consisted of the chosen subject being the main focus of the frame, the scene being well lit, that the subject wasn't wearing a hat or sunglasses, that the subject was relatively static within the frame and that the angle of their face to the camera was no more than 10°(28). During the Data Collection stage, these criteria were kept in mind, although due to the uncontrolled nature of real-world data, pre-processing was conducted to ensure the data met the required criterion.

3.2 Pre-processing

The pre-processing of the data was a slow and thorough process as it required each video to be edited manually, filtering the data to ensure the criteria of FACET and openSMILE were met. The filtering was conducted frame by frame using the iMovie[2] software. Frames in which the subjects face was obstructed from view or facing away from the camera and also in which another persons voice could be heard (e.g. the interviewer) were removed.

[2]https://www.apple.com/lae/imovie/

3.3 Post-processing

3.3.1 Facial Expression Analysis

To get the sentiment results related to the facial expressions of each leader, FACET was used to post-process the sample collection. FACET is an emotion recognition system which uses Computer Vision technology to detect the face within the frame, the action units which are active and thereby the emotions which are being expressed. FACET analysed the subjects sentiments at a rate of approximately 30 frames per second (fps), this value could change depending on the quality of the video.

Upon completion, the raw sensor data for each video was exported as a tab delimited file. The senor data provides all results related to the sentiment analysis including the emotion evidences, the action unit evidences, the frame number and time in milliseconds, to name but a few. The tab delimited files were then read into the R program and stored in dataframes for further analysis.

3.3.2 Acoustic Speech Analysis

To get the acoustic sentiment results, the openSMILE speech feature extractor was used. To be processed with this system, the data had to be converted to audio files, segmented and processed with opensmile via the command-line. This process was automated by functions within the speechFunctions.R script within this program. The functions this script executes consists of converting the video files to 16 bit WAV files, segmenting the WAV files into 5 second chunks, creating a bash script which will process the segmented WAV files for each scenario with openSMILE, and then parse the ensuing openSMILE results file to determine the emotion probabilities. An example of the Bash script which was created and executed within this program can be found in Appendix A1.3.

The openSMILE configuration file used for processing was *emobase_live4_batch_single.conf* which depends on the pre-trained openEAR models. The model which was of particular interest in this project was *emodb.emobase.model*, which was trained on the Berlin Speech Emotion Database, commonly used in speech emotion recognition research, which gives the results in a machine-parsable format (Listing 3.1). Due to this, while parsing the openSMILE results file, pattern matching was used to find the data relating to this model and more complex regexes were used to source the emotion probabilities. A partial code listing for this process of parsing is shown below (Listing 3.2).

Listing 3.1: Partial listing of an openSMILE results file.

```
SMILE-RESULT::ORIGIN=libsvm::TYPE=classification::COMPONENT
   =emodbEmotion::VIDX=0::NAME=emodb_emotion::CATEGORY_IDX
   =1::CATEGORY=boredom::PROB=0;anger:0.045995::PROB=1;
   boredom:0.388477::PROB=2;disgust:0.236506::PROB=3;fear
   :0.031864::PROB=4;happiness:0.034387::PROB=5;neutral
   :0.062262::PROB=6;sadness:0.200510
```

Listing 3.2: Partial listing of code used to extract the emotion probabilities

```
if(str_detect(line, "emodbEmotion")) {
     anger_prob <- append(anger_prob, as.numeric(str_extract(line,
        regex("(?<=anger\\:)(\\d\\.\\d{1,6})"))) )
     boredom_prob <- append(boredom_prob, as.numeric(str_extract(line,
        regex("(?<=boredom\\:)(\\d\\.\\d{1,6})"))) )
     ...
   }
```

3.4 Data Analysis

3.4.1 Development of Program

The primary component of this research project is the program which was created to perform the analysis of the sentiment data. The aim when creating this program was to create an intelligent system which would automatically interpret this complex cognitive activity, thereby removing the complexity for the user of analysing sentiment data. This program will be used to gain insights to the sentiments being expressed by these leaders, particularly, whether a deviance in the expression of sentiments can be determined between the emergency and non-emergency situations.

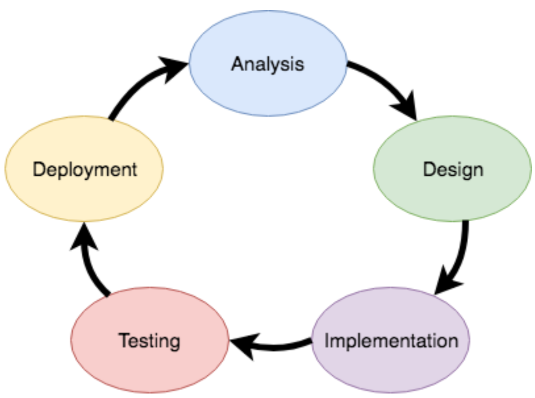

Figure 3.2: Diagram of software development lifecycle.

The process of creating this program followed the software development lifecycle, an integral method of software engineering. The predominant development style used was Agile as this is an efficient and adaptive form of development which can easily cope with changes to the initial scope of the project. The various stages of software development within this project consisted of: • Analysis • Design • Implementation • Testing • Deployment.

Analysis

The Analysis stage mainly comprised of research on the current state of the art in regards to Sentiment Analysis, particularly Multimodal Non-Verbal Sentiment Analysis. Another vital part was the requirement gathering, this involved learning both how to source the sentiment data using the chosen sentiment analysis systems (FACET and openSMILE) and the structure in which the results will be given. Learning how to use the sentiment analysis systems consisted of a training session with an iMotions Customer Success Manager in the case of FACET, and rigorous experimentation using documentation and online resources as a guide for openSMILE.

Design

Within the Design stage, the knowledge gained from the Analysis stage was used to greatly influence both the techniques used within the program which was to be created, but also the structure. From the research conducted on the state of the art, the statistical methods often used in this area of research was be determined, and based on this, the appropriate techniques were incorporated in the program. Knowledge of the format of the sentiment data which will be input into this program is important as it determines what procedures must be performed on this data before the analysis can take place, such as parsing and filtering the data to remove noise (thresholding). The program was then planned using flowcharts as this resulted in graphical representations of the overall process which aided in the evaluation and evolution of the design.

Implementation

The analysis program was created in the programming language R, a widely used language among statisticians and data scientists. The various statistical methods were implemented in a sequential fashion in which a full system test was conducted before another method was added. The flowcharts created in the previous stage were used as a visual aid both in the development and debugging of this program.

Testing and Deployment

Each time a new procedure was included in this program, rigorous testing was conducted to ensure that this procedure not only worked correctly, but also, that it did not have any adverse effects on other procedures. Upon completion of the Testing stage, if an issue within the code was exposed, the development cycle was iterated, returning to the Analysis stage. In the case that no issues were discovered in the Testing stage, the next stage of the development is the deployment of the program, or in the case of this research project, the program was used to source the final sentiment results (which are discussed in Chapter 4).

3.4.2 Facial Expression Analysis

FACET provides the sentiments and action unit results in the form of *evidences*. These evidences have a range of -8 to +8, for example, a sadness evidence of +2 means that the observed expression is (10^2) 100 times more likely to be classified as sadness (28). A threshold of 0.1 was used in the R program to determine if a sentiment was present as this gave a greater than equiprobable chance of the sentiment being present ($10^{0.1} \approx 1.259$). The chosen threshold for FACET is based on that we expect it to be able to recognise sentiments that vary slightly from that found in the training data. By choosing a threshold of 0.1, the program is accepting almost all cases in which FACET claims that a sentiment is present.

3.4.3 Speech Analysis

openSMILE provides the sentiment results in the form of probabilities (decimal form) which range from 0-1. The sentiment probabilities generated with the aforementioned model generally are within the 0-0.5 range. Due to this, the threshold used in this program to determine if an sentiment is present is 0.12.

3.4.4 Unimodal Analysis

The following methods were used to perform unimodal analysis on both the facial expression and speech sentiment data.

3.4.5 Thresholding & Signal-To-Noise Ratio (SNR)

As discussed in Sections 3.4.2 & 3.4.3, a given threshold was used to determine if the sentiment under consideration was present. This threshold was also used to determine the amount of noise within the resulting data from the sentiment analysis systems. As this research is concerned with the sentiments expressed through non-verbal channels, particularly visual and acoustic, which can be particularly noisy modalities. These systems learn based on the assumptions that expressions of a given sentiment will be similar to the training data, this introduces an inductive bias. Due to this, the threshold is used to get the number of values below this threshold (Rejected), where no signal was found (No Signal), the number of values above the threshold (Accepted) and the total number of values for that sentiment (Number of Values). These values were then used to compute the frequency at which that sentiment was expressed (Total Accepted %) and the Signal-To-Noise Ratio (SNR) for each sentiment (Signal-To-Noise Ratio %). The Total Accepted %, also known as the Emotion Frequency Percentages for each leader were derived from the Accepted and Number of Values parameters (Equation 1).

$$Emotion\ Frequency\ \% = \frac{Accepted}{Number\ of\ Values} \times 100 \tag{1}$$

The Signal-To-Noise Ratio was deduced by getting the total amount of noise within the signal (*Rejected* + *No Signal*) relative to the total signal (Equation 2). As the sentiment results under investigation are sourced through non-verbal cues, it is expected that there may be quite a high Signal-To-Noise ratio, which is why it was necessary to use a threshold for the sentiment values to filter out the noise.

$$Signal - To - Noise\ Ratio\ \% = \frac{Rejected + No\ Signal}{Number\ of\ Values} \times 100 \tag{2}$$

Data Visualisation

The Emotion Frequency Percentages which were determined in the thresholding process specified above, were used to create a Data Visualisation of the emotions detected in each leader, through each modality and for both scenarios. This visualisation consists of grouped barplots for all leaders within each scenario, this was performed on the data for each modality. From this, an overview of the sentiments expressed by all leaders in each scenario was provided. This visual aid could then be used to draw comparisons between the sentiments expressed in different scenarios, through the different non-verbal channels, and also between the different leaders. The resulting barplots are presented and discussed in further detail in Chapter 4.

Test of Statistical Independence

As a main goal of this project was to determine if a deviance in the emotions shown by the leaders could be deduced using data from different scenarios. The scenarios under consideration are the emergency scenario (one in which the leader is uneasy) and the non-emergency scenario (where the leader is more relaxed). The sentiments detected in the non-emergency scenario are interpreted as the "normal" or expected distribution of sentiments.

Pearson's Chi-Square Test (Equation 3) was used to determine the relationship between the sentiments conveyed in the different scenarios for each leader. The Emotion Frequency Percentages for the emergency and non-emergency scenarios were used as the observed and expected distributions for the Chi-Square Test, respectively (Equation 4). From this a probability, P-Value, was computed and if this was below the chosen significance value of 0.05, the Null Hypothesis given below would be rejected.

Null Hypothesis: The sentiments expressed in the emergency and non-emergency scenarios are significantly different.

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{3}$$

$$\chi^2 = \sum_{i=1}^{n} \frac{((emergency\ sentiment\ \%) - (non - emergency\ sentiment\ \%))^2}{non - emergency\ sentiment\ \%} \tag{4}$$

3.4.6 Multimodal Analysis

Data Fusion

The sentiments which are expressed via facial-expressions may differ to those expressed through acoustic speech features as there may be a delay between the two, although it is the relationship between these two modalities which will be investigated through the aggregation of these results. By performing this fusion, the hope was that the resulting sentiment data would be more truthful to the emotions actually being expressed by the subject.

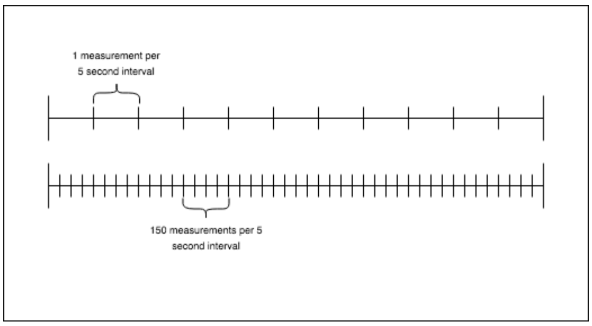

Figure 3.3: Conceptual diagram of differing sampling rates.

A complexity which had to be addressed for the fusion to be accomplished successfully was that the sentiment analysis systems used to get the sentiment results (FACET and openSMILE) used different sampling rates. Due to this, the number of measurements for each system differs. The resulting sensor data produced by FACET gives approximately 30 sentiment evidence values per second (as it generally processes 30 frames of video per second), whereas openSMILE sentiment results are given at a rate of 1 sentiment measurement per 5 seconds.

The fusion approaches which were attempted, although proved unusable due to the variance within the data include:

- Comparison of instantaneous values.
- Comparison of first moments.
- Comparison of facial expression first moments and the instantaneous speech values.

Comparison of instantaneous values - This involved comparing the speech values (which were given every 5 seconds with the facial expression emotion measurements at that time frame. This consisted of comparing every 150th facial expression measurement with the speech measurements. Upon analysing the resulting values, there didn't seem to be much of a correlation between the instantaneous values. This may be because the resulting speech data is an aggregation of the emotions expressed over that 5 second period.

Comparison of first moments - This comprised calculating the mean for each sentiment over the entire video, and comparing the resulting speech means and facial-expression means. In order for the resulting first moment values to be valid for comparison, there should be a low variance within the data. From studying the resulting data and its variance, it was found that the variance in this data would have caused the results of the statistics performed on it to be unsubstantiated.

Comparison of facial expression first moments and the instantaneous speech values - This included computing the mean of the facial expression for each 5 second segment - getting the means of every 150 sentiment measurements (30 *sentiment values per second* $\times$ 5 *seconds* $= 150$) and comparing these with the speech sentiment values. Although this approach reduces the effects of the variance within the data through considering smaller amounts of data more frequently, due to the transient nature of facial expressions, it was found to be unsuitable. By getting the mean for each sentiment over each 150 valued chunks of data, sentiment values which may have been quite high (confident in the expression of a given sentiment) were often lost - leading to very few values where the sentiment was over the given threshold to be considered present. This approach was not used because of the lack of information within its results.

Final Data Fusion Approach

Due to the variance within this data and the fleeting nature of facial expressions, it was necessary that a fusion method was created which could capture the emotion which was being expressed through the facial expressions. This method consisted of a **Comparison of max facial expression values and the instantaneous speech values**. For each chunk of 150 measurements for each sentiment, the max value over the chunks of data were determined. These max values were then compared with the speech sentiment values. This method reduced the effects of variance within the data through considering smaller chunks of data at regular intervals and is able to capture the sentiments detected within the facial expression through considering the max values.

Test of Independence

To validate the statistical independence results computed for the facial-expression data, the aggregation of the facial-expression and speech data are also statistically tested for independence. To truly determine whether a deviance within the leaders behaviour between the emergency and non-emergency scenarios was present, a chi-square test was performed on the fused data. The fused data was thresholded to get the percentage of speech data (within the fused data) which detected the sentiment, the percentage of facial-expression data (within the fused data) which detected the sentiment for each scenario. These values are then used as the inputs for a chi-square test (Equation 5), the result of which will reveal whether this deviation is present. It would be expected that these values should be similar to that computed in the unimodal analysis, although this result should be more representative of the true emotional state of the leader as it considers data from two non-verbal channels.

$$\chi^2 = \sum_{i=1}^{n} \frac{((fused\ emergency\ emotion\ \%) - (fused\ non-emergency\ emotion\ \%))^2}{fused\ non-emergency\ emotion\ \%} \quad (5)$$

Support between Modalities

As this project is concerned with the sentiment expressed through two non-verbal channels, Audio and Visual, it is important to determine how well the sentiment results detected from each channel support one another. These support measures are determined using the count of the audio and visual sentiment values being over the threshold (the number of times both modalities agreed that the given sentiment was present), and the count of each modalitiy's sentiment values being over the threshold (the number of times which each modality detected that the given sentiment was present). To determine how well the audio sentiment results are supported by the visual sentiment results, the ratio of both audio and visual agreeing on the presence of a sentiment and the count of when audio detected the sentiment was computed (Equation 6). A similar ratio is also computed to determine how well the audio data supports the visual data (Equation 7).

$$Audio\ supported\ by\ Visual = \frac{count\ of\ Audio\ and\ Visual\ over\ threshold}{count\ of\ Audio\ over\ threshold} \tag{6}$$

$$Visual\ supported\ by\ Audio = \frac{count\ of\ Audio\ and\ Visual\ over\ threshold}{count\ of\ Visual\ over\ threshold} \tag{7}$$

The results from this investigation can be used to reinforce the detection of particular sentiments, as if a given sentiment is detected by both modalities, it is highly likely that the given sentiment is being expressed by the subject. These results could also be used as a method of determining the accuracy of the sentiment analysis systems, if the support between the modalities seem irregular for particular sentiments, it could be suggestive of inaccuracies/biases within the sentiment analysis system for those particular sentiments.

3.4.7 Analysis of Accuracy of the System

The purpose of this analysis was to gain an insight into the accuracy of the systems. To perform this analysis, Pearson's Correlation Coefficient, defined in Equation 8, was used to determine the strength of the linear relationship between two factors. The correlation test was performed on the sentiments expressed and in the case of FACET, the correlation between the sentiments and the active AUs was also measured. The correlations were computed within the systemAnalysis.R script in which the relevant columns (emotions/AUs) within the raw sentiment data dataframes were the input data for the correlation matrix function in R (Listing 3.3).

Listing 3.3: Code used to compute the correlation matrix.

```
corMatrix <- round(cor(data, method = "pearson", use =
   "complete.obs"), 4)
```

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{8}$$

Correlation of Sentiments

The correlation of sentiments was performed on the results for all leaders in both emergency and non-emergency scenarios and for both systems (FACET and openSMILE). For each sentiment (emotion) it compares the evidences/probabilities with that of the other sentiments, for each frame/chunk - from this, the relationship between the sentiments for each video could be determined. The expectation was that there should be a strong negative correlation between Joy/Happiness (a principally pure sentiment) and sadness or anger (examples of negative sentiments).

Correlation of Sentiments and Action Units

A further analysis was conducted on FACET by computing the correlations between each sentiment and the action units, for each frame/chunk. The expectation in this case is that there should be a strong positive correlation between the sentiments and the AUs which are the criteria for the presence of that sentiment, specified in (37).

3.5 Summary of Methods Used

The method used for this project consisted of collecting and filtering the data to be an appropriate quality for the sentiment analysis systems. Facial-expression and speech sentiment results were sourced by processing the filtered sample data with FACET and openSMILE, respectively. A program was then created which performs data analysis on the aforementioned sentiment results to reduce the encumbrance of interpretation for the user.

The data analysis included in this program consisted of unimodal analysis, multimodal analysis and an assessment of the accuracy of the sentiment analysis systems used. The unimodal analysis consisted of investigating the sentiment results from each modality separately and using statistical methods such as the Chi-Square Test to determine if an anomaly within the behaviour of the chosen female leaders can be determined between emergency and non-emergency situations. The multimodal analysis consisted of fusing the sentiment data from both modalities to create a more robust model of the sentiments expressed and using this to determine if an anomaly can be determined within the leaders behaviour, but also to investigate how well the modalities support each other. The evaluation of the accuracy of the sentiment analysis systems mainly consisted of getting the correlations of the detected sentiments to see if this was consistent with our expectation of positive and negative sentiments being negatively correlated.

In the following chapter, the results of the data analysis performed in the program (which has been described above) in reference to the chosen female leader's expression of emotions shall be discussed.

4 Case Studies

In this chapter, the results from analysing the facial expression and speech (acoustic) sentiment data, which was performed using the program described in the Chapter 3, will be presented in the form of Case Studies.

The results of the unimodal analysis, in which the sentiment results from the modalities were analysed separately, are be discussed further in the Visual Sentiment Case Study and the Audio Sentiment Case Study. In these studies, the frequency of the sentiments expressed within the various scenarios by all the leaders is examined. The regularity of the sentiments expressed by these leaders is also investigated by comparing the distribution of sentiments expressed in both emergency and non-emergency situations.

The results from aggregating the facial-expression and speech sentiment data is presented and reviewed in the Multimodal Sentiment Case Study. This study discusses the support between the modalities, it also uses the fused data to examine the regularity of the sentiments expressed by the leaders in emergency and non-emergency situations.

The performance of the sentiment analysis systems used in this project were investigated. The validity of the results and the ensuing conclusions drawn are dependant on the accuracy of the system used to gain the raw sentiment outputs.

4.1 Unimodal Analysis

As this project is concerned with the expression of sentiment through facial-expression and speech (acoustic), it is important to investigate this behaviour not only on the aggregated data, but on each of the modalities results. From analysing the sentiment data of each modality separately, trends in behaviours which could be isolated to one non-verbal channel may be uncovered.

4.1.1 Visual Sentiment Case Study

This section discusses the sentiment results relating to facial-expressions for the given sample of leaders which was sourced using FACET.

The frequency of the sentiments expressed by the leaders, through their facial expressions, in different scenarios was presented in the figures below. Through visualising the sentiment results in this manner, the sentiments detected by the different leaders can be compared. These visualisations also allow trends and dissimilarities between the leader's emotional behaviour in the emergency and non-emergency scenarios to be observed.

From observing the frequencies of sentiments in the emergency (Figure 4.2) and non-emergency scenarios (Figure 4.1), an increase in negative sentiments (such as fear, disgust and sadness) can be observed in the emergency situation as compared to the non-emergency. Taking a closer look at the sentiments expressed by Mary Robinson, the expression of joy seems to remain constant - indicating significant control of this emotion. Whereas when examining the sentiments expressed by Heather Bresch, in both scenarios there a significant expression of negative emotions, particularly fear. This expression of negative sentiments in both scenarios, which spikes in the emergency scenario suggests that she has less control of the sentiments she expresses within the presence of media.

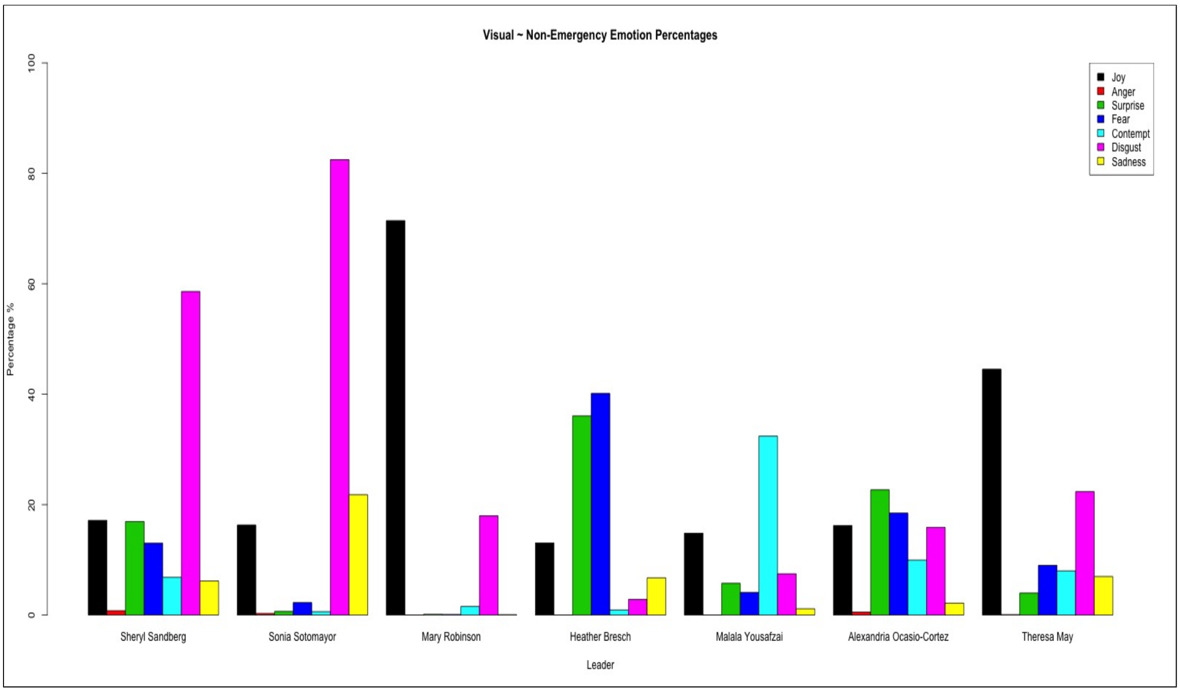

Figure 4.1: Emotion distribution during Non-Emergency scenarios from Facial Expressions. From left to right on the x-axis, the emotion frequencies for Sandberg, Sotomayor, Robinson, Bresch, Yousafzai, Ocasio-Cortez and May are graphed in the form of barplots.

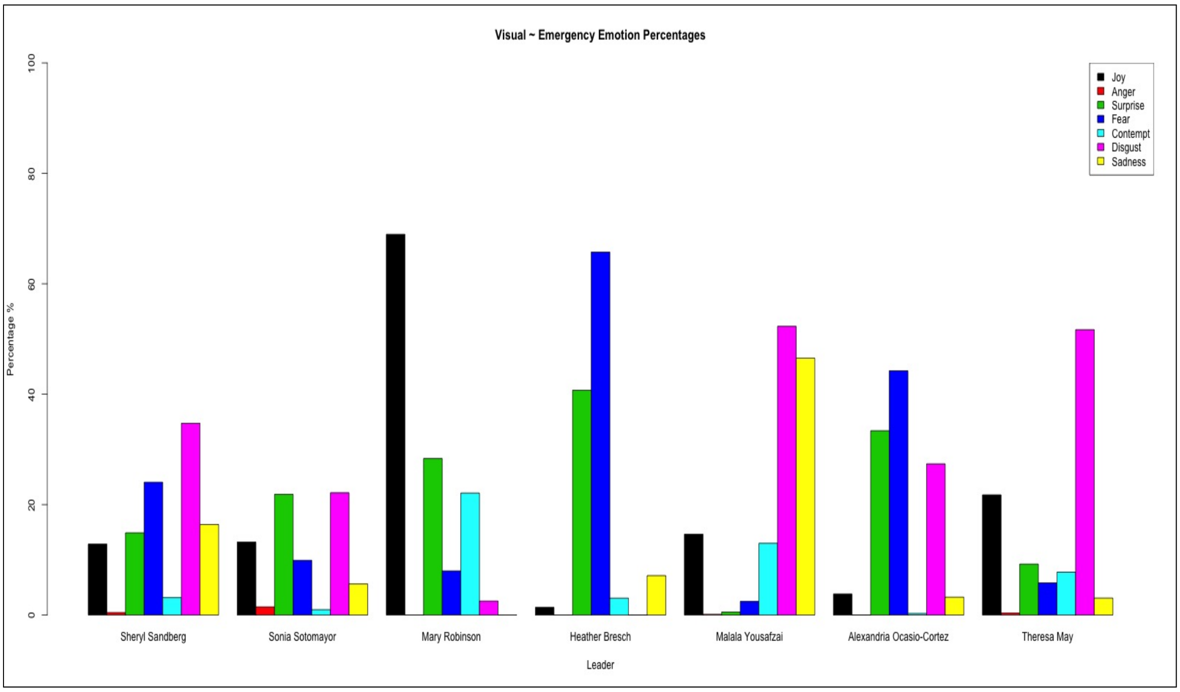

Figure 4.2: Emotion distribution during Emergency scenarios from Facial Expressions. From left to right on the x-axis, the emotion frequencies for Sandberg, Sotomayor, Robinson, Bresch, Yousafzai, Ocasio-Cortez and May are graphed in the form of barplots.

The purpose of the Chi-Square test was to determine if the expression of sentiments by these leaders would be constant in different scenarios. From comparing the distributions of sentiments for all the leaders in the emergency and non-emergency scenarios, the resulting p-values were all greater than the significance value of 0.05 leading us to accept the null hypothesis - confirming a significant difference in the sentiments expressed in both scenarios (Table 4.1). These results agree with what was suggested from the sentiment data visualisations discussed above, that generally the sentiments expressed varied significantly in both scenarios.

The results for both Heather Bresch and Alexandria Ocasio-Cortez are lightly lower, indicating that there may be slightly more similarities - from analysing Figure 4.1 and Figure 4.2 it was concluded that this may be due to the fact that both these women express the same emotions in both scenarios, but in differing scales.

Table 4.1: Chi-Square results determining the statistical significance between the emotions expressed by the leaders via **facial expressions** in an emergency and non-emergency scenario. Calculations show that the is a significant difference between the emotions shown in the different scenarios.

Leader	Scenario	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness	P-value
Sheryl Sandberg	Emergency	0.13	0.00	0.15	0.24	0.03	0.35	0.16	0.999
	Non-Emergency	0.17	0.01	0.17	0.13	0.07	0.59	0.06	
Sonia Sotomayor	Emergency	0.13	0.01	0.22	0.10	0.01	0.22	0.06	0.881
	Non-Emergency	0.16	0.00	0.01	0.02	0.01	0.82	0.22	
Mary Robinson	Emergency	0.69	0.00	0.28	0.08	0.22	0.03	0.00	0.960
	Non-Emergency	0.71	0.00	0.00	0.00	0.02	0.18	0.00	
Heather Bresch	Emergency	0.01	0.00	0.41	0.66	0.03	0.00	0.07	0.403
	Non-Emergency	0.13	0.00	0.36	0.40	0.01	0.03	0.07	
Malala Yousafzai	Emergency	0.15	0.00	0.01	0.02	0.13	0.52	0.47	0.951
	Non-Emergency	0.15	0.00	0.06	0.04	0.32	0.07	0.01	
Alexandria Ocasio-Cortez	Emergency	0.04	0.00	0.33	0.44	0.00	0.27	0.03	0.656
	Non-Emergency	0.16	0.01	0.23	0.18	0.10	0.16	0.02	
Theresa May	Emergency	0.22	0.00	0.09	0.06	0.08	0.52	0.03	0.998
	Non-Emergency	0.45	0.00	0.04	0.09	0.08	0.22	0.07	

4.1.2 Audio Sentiment Case Study

This section is discusses the sentiment results relating to acoustic features of speech for the given sample of leaders which was sourced using openSMILE.

From analysing the frequency of sentiments expressed by the leaders in the emergency (Figure 4.4) and non-emergency scenarios (Figure 4.3) as in the visual case study there is an increase in negative sentiments, such as disgust, anger and fear, in the emergency scenarios. From these plots, it is clear that there is a very high detection rate for boredom and sadness. A high frequency for boredom is not surprising as it is not one of the primary emotions but rather a composite emotion. Although the high frequency of sadness was unexpected. It may be perhaps that the leaders were expressing a lot of sadness through the way they were conveying their speech, although as this was seen for all leaders in both scenarios and it is significantly different to the sentiments detected within the visual case study - it leads us to question the results for sadness. Could this high frequency of sadness be due the model used having a bias towards sadness?

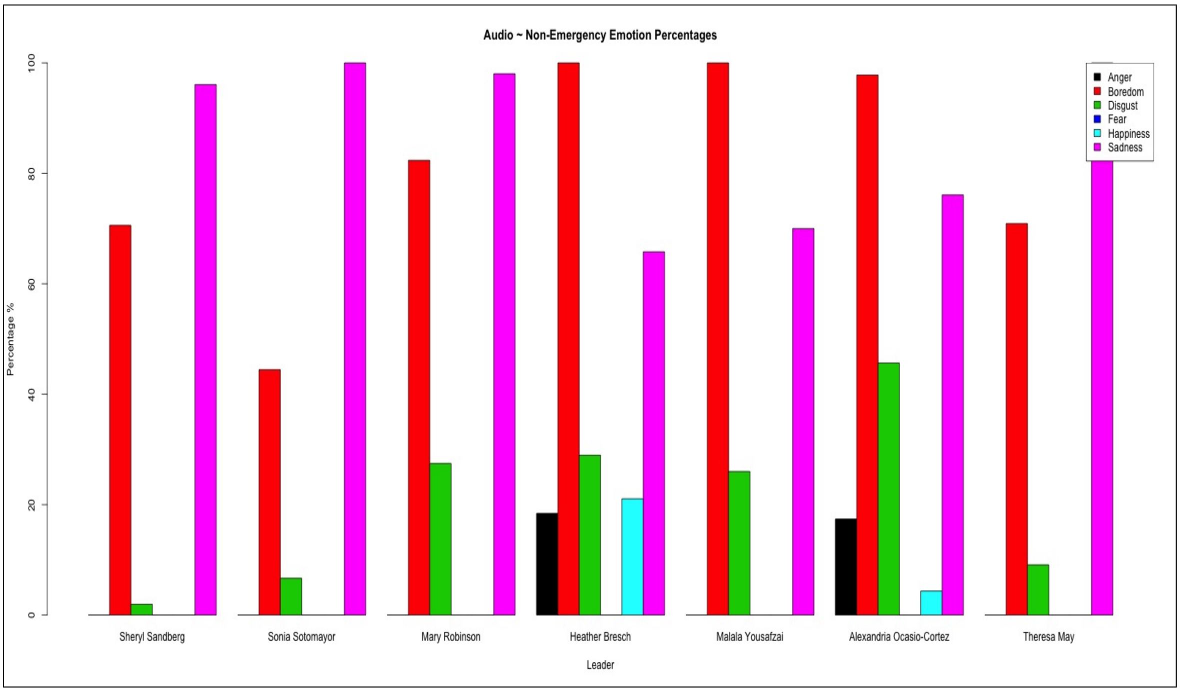

Figure 4.3: Emotion distribution during Non-Emergency scenarios from Speech. From left to right on the x-axis, the emotion frequencies for Sandberg, Sotomayor, Robinson, Bresch, Yousafzai, Ocasio-Cortez and May are graphed in the form of barplots.

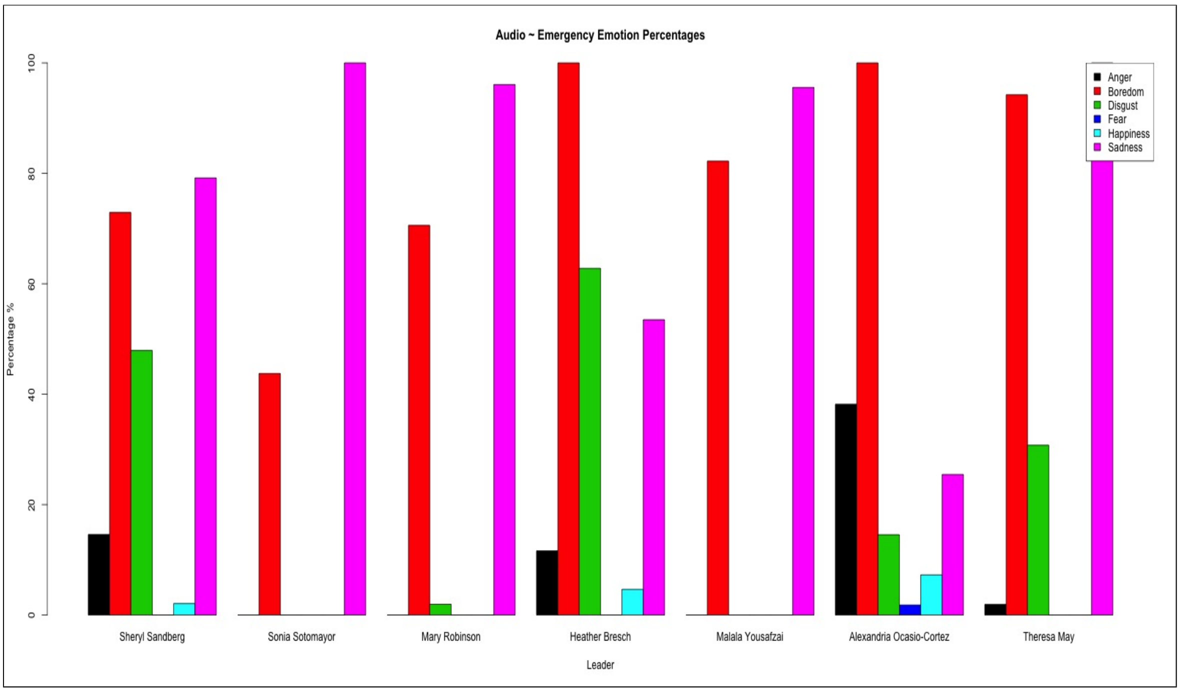

Figure 4.4: Emotion distribution during Emergency scenarios from Speech. From left to right on the x-axis, the emotion frequencies for Sandberg, Sotomayor, Robinson, Bresch, Yousafzai, Ocasio-Cortez and May are graphed in the form of barplots.

Table 4.2 shows the emotion percentages for each of the leaders in the emergency and non-emergency and the subsequent p-values. The p-values are over the significance value for all the leaders, therefore the null hypothesis is accepted, that there is a significant difference in the sentiments expressed in both scenarios. As discussed above, a significant increase in the negative sentiments was seen in the emergency scenario (Figure 4.4), this corresponds with the results of this chi-square test that all of the chosen leaders displayed significantly different sentiments in the different scenarios.

A slightly lower p-value was calculated for Mary Robinson's expression of sentiment in both scenarios which may be indicating that there are some similarities in the behaviour shown in both scenarios. From examining the emotions expressed by Mary Robinson in more depth, it can be seen that the same sentiments were detected in both scenarios, although the frequencies vary.

Table 4.2: Chi-Square results determining the statistical significance between the emotions expressed by the leaders via **acoustic features of speech** in an emergency and non-emergency scenario. Calculations show that the is a significant difference between the emotions shown in the different scenarios.

Leader	Scenario	Anger	Boredom	Disgust	Fear	Happiness	Sadness	P-value
Sheryl Sandberg	Emergency	0.15	0.73	0.48	0.00	0.02	0.79	0.986
	Non-Emergency	0.00	0.71	0.02	0.00	0.00	0.96	
Sonia Sotomayor	Emergency	0.00	0.44	0.00	0.00	0.00	1.00	0.999
	Non-Emergency	0.00	0.44	0.07	0.00	0.00	1.00	
Mary Robinson	Emergency	0.00	0.71	0.02	0.00	0.00	0.96	0.649
	Non-Emergency	0.00	0.82	0.27	0.00	0.00	0.98	
Heather Bresch	Emergency	0.12	1.00	0.63	0.00	0.05	0.53	0.975
	Non-Emergency	0.18	1.00	0.29	0.00	0.21	0.66	
Malala Yousafzai	Emergency	0.00	0.82	0.00	0.00	0.00	0.96	0.999
	Non-Emergency	0.00	1.00	0.26	0.00	0.00	0.70	
Alexandria Ocasio-Cortez	Emergency	0.38	1.00	0.15	0.02	0.07	0.25	0.874
	Non-Emergency	0.17	0.98	0.46	0.00	0.04	0.76	
Theresa May	Emergency	0.02	0.94	0.31	0.00	0.00	1.00	0.999
	Non-Emergency	0.00	0.71	0.09	0.00	0.00	1.00	

4.1.3 Comparison of Visual and Audio Case Studies

From the unimodal analysis performed on each modality, an increase in negative sentiments in the emergency scenario was evident in both cases, although the sentiment frequencies detected by each modality often differed. The results of Chi-Square tests, found that the sentiments expressed in the different scenarios were significantly different for each leader. As both modalities agreed on the deviance within the leaders' behaviour, the confidence of this result is strengthened.

The results of openSMILE revealed that a high frequency of sadness was observed in both scenarios by all leaders. This differed greatly from the results observed by FACET for the facial expressions, therefore this causes us to question the validity of the audio results. Although this may also be because the leaders are expressing sentiments differently through the different non-verbal channels, perhaps exerting more control over the sentiments expressed through a particular channel.

4.2 Multimodal Sentiment Case Study

Multimodal analysis was one of the primary focuses of this research. It was the hope that through aggregating the facial-expression and speech sentiment analysis, a more precise model would be formed of the emotional behaviour of these leaders. The results of this would also be compared with those computed in the Unimodal Analysis as a method of validating the facial-expression sentiment results. This analysis consisted of an independence test to determine if significantly different sentiments were being expressed by the leaders in the emergency and non-emergency scenarios, and also an evaluation of the support between the modalities.

The independence test consisted of a Chi-Square test which took the frequency of each sentiment of the leaders in the different situations and from this, determined if there was a significant difference in the sentiments being expressed. The results from this test will be compared with the independence test results from the unimodal analysis - in the case that there is an agreement between these results, the facial-expression results are validated and the confidence in this outcome is increased greatly.

Generally, it was found that for all sentiments expressed by each leader that the p-values were above the significance level, causing the the null hypothesis, that there was a significant difference in the expression of that sentiment between the different scenarios, to be accepted (Table 4.3). This result is in accordance with the independence results determined in the unimodal studies. Therefore a deviance within the leaders behaviour was successfully determined through the analysis of the leaders' non-verbal cues. There was an exception found, Mary Robinson's expression of fear in both scenarios and through both non-verbal channels remained constant. From investigating the sentiment frequencies determined for Mary Robinson by both modalities in the different situations (Tables 4.1 & 4.2), she appears to express very little fear through her speech or facial-expressions in either scenario - indicating great control over the expression of this sentiment. From analysing these results, it can be observed that the results for Alexandria Ocasio-Cortez, though it clearly signifies a difference between the behaviour in the two scenarios, they are less confident (as the results range from 0.5173 - 0.6359, whereas the the results of the other leaders are all above 0.7). This seems to be suggesting that there may be some similarities in her expression of emotion and from studying the graphical representations (Figures 4.1, 4.2, 4.3, 4.4) it can be noted that in both situations she is expressing the same emotions, although to different extents.

Table 4.3: Chi-Square Results for fused data determining the independence of the expression of the emotions in the non-emergency and emergency scenarios. Calculations show that the is a significant difference between the emotions shown in the different scenarios with Mary Robinson's expression of fear being the only exception.

Leader	Anger	Disgust	Fear	Happiness	Sadness
Sheryl Sandberg	0.8066	0.9154	0.8705	0.7035	0.8034
Sonia Sotomayor	0.7871	0.9834	0.4832	0.8378	0.8655
Mary Robinson	0.8253	0.7606	0.0001	0.9999	0.9124
Heather Bresch	0.9205	0.4207	0.9205	0.4317	0.9201
Malala Yousafzai	0.8614	0.7914	0.8157	0.7992	0.1438
Alexandria Ocasio-Cortez	0.5173	0.6220	0.6359	0.6117	0.6358
Theresa May	0.7945	0.9854	0.9100	0.9999	0.7933

As multimodal analysis was at the core of this research, it was important to determine how well these modalities agreed with the presence of a given sentiment. These results may reveal discrepancies between the modalities which may be due to the nature in which the sentiments are expressed, as some people can be more inclined to express sentiments through a particular non-verbal channel. These discrepancies may also be a consequence of inaccuracies within the sentiment analysis systems used. When there is a similar high level of support shown by both modalities, this indicates a reliable result.

Table 4.4 contains the results indicating how well the audio results for each sentiment by each leader in the different scenarios are supported by the visual data. This support was calculated by getting a ratio of the number of times in which the modalities agreed on the presence of a sentiment relative to the number of time this sentiment was detected from the audio data. The results showed very little support for the presence of anger, and relatively low support for fear and happiness. The support for sadness was generally quite high, indicating that the accuracy of openSMILE in regards to the detection of sadness is quite high. Although, the support for the detection of sadness for Mary Robinson are extremely low, showing discrepancies in openSMILE's detection of sadness. Although the detection of sadness is often found to be correct, the case with Mary Robinson clearly indicates an over-classification of this sentiment which may be due to a bias within the machine learning model used.

The results shown in Table 4.5 reveals how well the visual data is supported by the audio data. Sadness was found to be very well supported overall and again, emotions such as anger, happiness and fear are not as well supported

Table 4.4: Results pertaining to how well the fused audio results are supported by the visual results. Showing sadness to be generally well supported whereas the support for the other emotions was mush lower.

Leader	Scenario	Anger	Disgust	Fear	Happiness	Sadness
Sheryl Sandberg	Emergency	0.07	1.00	0.00	0.67	0.88
	Non-Emergency	0.00	1.00	1.00	1.00	0.71
Sonia Sotomayor	Emergency	0.00	0.00	1.00	1.00	0.73
	Non-Emergency	0.00	1.00	0.00	0.00	0.89
Mary Robinson	Emergency	0.00	0.64	0.83	1.00	0.05
	Non-Emergency	0.00	0.74	0.00	0.00	0.06
Heather Bresch	Emergency	0.00	0.02	1.00	0.00	0.96
	Non-Emergency	0.00	0.86	0.00	0.80	0.93
Malala Yousafzai	Emergency	0.25	1.00	0.50	1.00	1.00
	Non-Emergency	0.00	0.88	0.00	0.00	0.26
Alexandria Ocasio-Cortez	Emergency	0.03	1.00	0.00	0.67	0.63
	Non-Emergency	0.47	0.87	0.00	1.00	0.65
Theresa May	Emergency	1.00	1.00	0.00	0.00	0.51
	Non-Emergency	0.00	1.00	0.00	0.00	75.00

Based on the support results for both modalities, there are cases where similarly high support values were determined for both modes (this was usually for the detection of sadness and less often disgust). From the graphical representation of the emotions detected by the different modalities, it is clear that openSMILE has a preference towards detecting sadness whereas most other emotions were detected very little. Therefore, until the recognition of emotions (such as happiness, anger and fear) are improved, the audio results will continue to compromise the integrity of the multimodal sentiment analysis.

Table 4.5: Results pertaining to how well the fused visual data is supported by the audio data. Showing sadness to be generally well supported whereas the support for the other emotions was mush lower.

Leader	Scenario	Anger	Disgust	Fear	Happiness	Sadness
Sheryl Sandberg	Emergency	0.50	0.62	0.00	0.06	0.88
	Non-Emergency	0.00	0.16	0.02	0.02	1.00
Sonia Sotomayor	Emergency	0.00	0.00	0.02	0.02	1.00
	Non-Emergency	0.00	0.07	0.00	0.00	1.00
Mary Robinson	Emergency	0.00	0.33	0.44	0.44	1.00
	Non-Emergency	0.00	0.30	0.00	0.00	1.00
Heather Bresch	Emergency	0.00	1.00	0.02	0.00	0.68
	Non-Emergency	0.00	0.69	0.00	0.24	0.74
Malala Yousafzai	Emergency	0.25	0.02	0.10	0.10	1.00
	Non-Emergency	0.00	0.71	0.00	0.00	0.85
Alexandria Ocasio-Cortez	Emergency	1.00	0.38	0.00	0.24	0.39
	Non-Emergency	0.72	0.67	0.00	0.10	0.90
Theresa May	Emergency	0.17	0.44	0.00	0.00	1.00
	Non-Emergency	0.00	0.17	0.00	0.00	1.00

4.3 System Accuracy Evaluation

Another goal of this project was to evaluate the performance of these sentiment analysis systems. The accuracy of these systems substantiates the validity of the results presented in this project. This section analyses the performance of these systems during (FACET) and after post-processing (FACET & openSMILE).

4.3.1 Signal-To-Noise Ratio

While determining the accuracy of the sentiment analysis systems used, it was critical to evaluate the proportion of noise within the sentiment data. Appendix A1.4 contains the SNR results for all leaders, in each scenario, determined for each modality. When analysing the FACET sentiment results, a high SNR was determined, this is not surprising as non-verbal sentiment data is often more noisy than its verbal counterpart. When analysing openSMILE sentiment data, it was observed that it often had a high SNR, although these high values occurred less frequently for this modality. This is most likely because this modality uses a much lower sampling rate than FACET. A higher sampling rate causes the quantisation of noise to be distributed over a wider frequency and due to this, the SNR is generally higher. Due to the lower sampling rate of openSMILE, it does not face cases where no signal is found - in FACET this often occurs due to the quality of the video, or the position of the subject relative to the camera.

From examining the signal-to-noise ratio for each modality, it was found that sentiment detected through non-verbal signals generally contain quite a large quantity of noise, this is typical behaviour and is likely because "audio visual feature extraction is done automatically using methods that add additional noise" (38).

4.3.2 Analysis of FACET

As discussed in Chapter 2, the process by which FACET determines the sentiments being expressed by the subject is by first detecting the face within the frame, determining which action units are active and from this, classifying the sentiment being communicated. The accuracy of this system is therefore dependent on the training data on which the various SVMs used to detect active AUs and recognise the sentiments, and the efficiency of the face detection phase.

The criteria of FACET required that the input video consisted of frames in which the subject was the centre of the frame. Therefore, wideshots (in which the camera pans out) can then include faces other than the subject under study. In this case, FACET can fail to detect the face of the subject - due to this, the accuracy of this system is also dependent on the quality of the manual pre-processing conducted on the input beforehand.

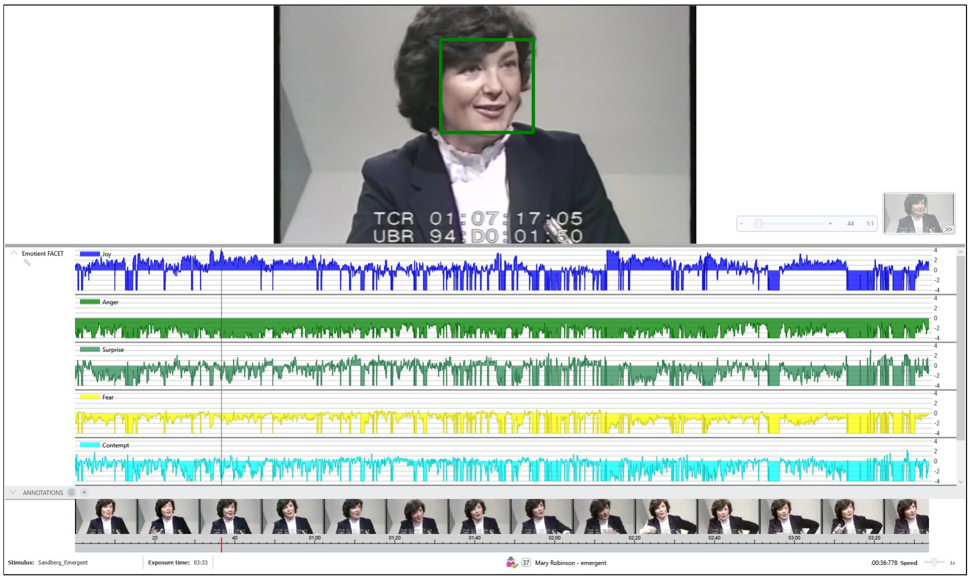

Figure 4.5: Screenshot of FACET in progress, detecting the face in the frame.

When FACET has successfully detected the face within the frame, it is bound by a green box as in Figure 4.5. The graphs below the frame of video shown depict the evidence values for the sentiments and action units, allowing the FACET results to be assessed with the corresponding video frame. From viewing the FACET results in this manner, times when the system failed to locate the face within the frame were noticed.

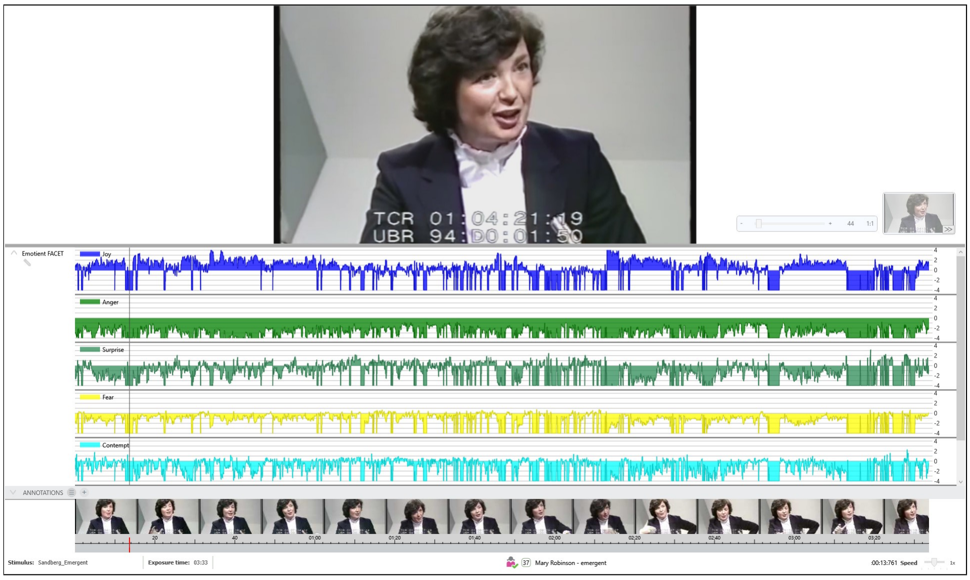

Figure 4.6: Screenshot of FACET in progress, not detecting the face in the frame.

Figure 4.6 corresponds to an instance in which the system failed to locate the face within the current frame although the subjects face was in a similar position as in Figure 4.5. These errors are often related to the quality of the video and/or the frequency of movement by the subject. Although overall, the face detection appeared to work quite well capturing the faces of the subjects at a range of angles and sizes.

Correlation of Sentiments

The correlation between the sentiments was conducted for each leader in each scenario and can be found in Appendix A1.5.1. Table 4.6 shows the sentiment correlations determined for Sheryl Sadnberg in an emergency scenario. It can be observed that there is a strong negative correlation between joy and anger/sadness - expected behaviour of an accurate systems as these emotions generally don't appear together. This negative correlation between the fundamental positive and negative sentiments implies an high level of accuracy in the performance of this system.

A strong correlation was noticed between sadness and contempt, and also between surprise and fear. This seems unusual, although according to Plutchik, and his wheel of emotions (a 3D model which defined the relations between various emotions) these emotions are closely related as they are in close proximity within the model (39). With this in mind, it is no longer surprising that there is a positive correlation between these sentiments, although this result does cause us to question the accuracy of their respective evidence results.

From analysing these sentiment correlations we have determined that FACET appears to perform at a high accuracy having a negative correlation between positive and negative sentiments, the instances of positive correlations relate to the highly complex relations between certain sentiments. These trends were representative of the sentiment correlations of all leaders within the sample.

Table 4.6: FACET sentiment correlation matrix for Sheryl Sandberg in an Emergency. This table shows a negative correlation between positive and negative sentiments, confirming the accuracy of this system.

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	0.081	-0.4253	0.091	0.2867	0.0707	-0.0758
Anger	0.081	1	-0.2947	-0.6876	0.2929	0.4412	0.0815
Surprise	-0.4253	-0.2947	1	0.1298	-0.2341	-0.5352	-0.1935
Fear	0.091	-0.6876	0.1298	1	-0.0513	-0.2739	0.057
Contempt	0.2867	0.2929	-0.2341	-0.0513	1	-0.0277	0.7512
Disgust	0.0707	0.4412	-0.5352	-0.2739	-0.0277	1	-0.1425
Sadness	-0.0758	0.0815	-0.1935	0.057	0.7512	-0.1425	1

Correlation of Sentiments and AUs

A further evaluation was conducted on FACET, which included computing the correlation between the sentiments which were detected and the action units which were deemed active at that point. It is expected that there should be a strong positive correlation between the sentiment detected and the action units which define them according to (37).

From observing the sentiment-action unit correlation for Sheryl Sandberg's expression of sentiments in an emergency-type scenario (Table 4.7), it is clear that action units often mingle and for this reason, their activity is not observed in isolation. From closely examining the correlation relating to joy which is defined by action units 6 and 12 (37, 40), positive correlations were also observed for action units 7, 9, 14, 20, 23, 25, 26 and 28 calling the accuracy of the system into question. The correlations computed for action units 6 and 12 are significantly higher than the other action units, confirming that this system performs at quite a high accuracy. Action units correspond to the contraction and relaxation of muscles within the face. Facial muscles, unlike other muscles within the body attach to other muscles or skin, rather than to the bone. Due to this, the contraction of one muscle or action unit may cause a change in the other facial muscles, causing this intermixing of action units.

Table 4.7: Sentiment - Action Unit correlations for Sheryl Sandberg in an Emergency. This table shows that there is some inter-mixing between the detection of AUs, although the AUs which define a given sentiment were found to be significantly higher than the others - showing FACET has perform accurately.

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
AU1	-0.3701	-0.5804	0.2313	0.3153	-0.4159	-0.4735	0.3578
AU2	-0.3424	-0.6633	0.307	0.3323	-0.3996	-0.3966	-0.0535
AU4	-0.3091	0.4923	-0.3261	-0.0124	-0.4429	0.1165	0.3651
AU5	-0.2424	-0.1825	0.4012	0.4816	-0.3655	-0.3179	-0.1682
AU6	0.4386	-0.6491	-0.6157	-0.0544	0.2676	0.1226	0.1172
AU7	0.2672	-0.0671	-0.3986	-0.1129	-0.0586	0.2647	0.0318
AU9	0.1801	0.2524	-0.4135	-0.3732	0.1288	0.631	-0.0036
AU10	-0.0083	-0.1704	-0.4205	-0.0401	-0.1247	0.7599	-0.2305
AU12	0.6494	-0.1882	-0.3433	0.2448	0.5719	-0.3281	-0.0013
AU14	0.0991	-0.0534	-0.6166	-0.0488	0.5754	-0.4706	0.2684
AU15	-0.147	0.0731	-0.3881	-0.0681	-0.1083	0.0672	0.0496
AU17	-0.0337	0.5659	-0.5915	-0.1764	0.4354	-0.1142	0.2506
AU18	-0.3918	0.6109	-0.2278	-0.3894	-0.4227	-0.2385	0.1589
AU20	0.135	-0.6052	-0.1608	0.4197	0.0746	-0.0665	-0.0835
AU23	0.138	0.7264	-0.4546	-0.0666	0.348	0.0845	0.136
AU24	-0.1088	0.4197	-0.53	-0.142	0.1874	-0.323	0.3306
AU25	0.1981	-0.4279	0.6003	0.2888	-0.1999	0.2063	-0.2296
AU26	0.2785	-0.5078	0.6958	0.1956	-5.00E-04	-0.1047	-0.1199
AU28	0.0184	-0.0371	-0.6345	0.0353	0.1233	-0.3525	0.2738
AU43	-0.3384	-0.7374	0.22	-0.3977	-0.3066	-0.0194	0.1376

4.3.3 Analysis of openSMILE

For openSMILE the audio data is read in and written to memory, various data processors compute the feature and functional values which are then input into the SVM which classifies the emotions being expressed. The results of this project are therefore dependent on the accuracy of this system and the data on which the SVM was trained.

Correlation of Sentiments

Table 4.8 contains the sentiment correlations calculated for the sentiments expressed by Sheryl Sandberg in an emergency through acoustic features within her speech. It is generally expected that positive sentiments such as joy should be negatively correlated with negative sentiments such as anger and sadness.

From examining the values within the table, a negative correlation was observed between happiness and sadness, as expected. In fact, sadness was found to be negatively correlated with all other emotions. This suggests that perhaps sadness is expressed acoustically in an isolated manner from other sentiments. Sadness arouses the parasympathetic nervous system causing a decrease in the heart rate resulting in speech which is slow, low-pitched and with low amount of higher frequency energy (36) - which tends to be dissimilar to the speech features aroused from other sentiments. Although, this result could also be indicative of a bias within the model used in openSMILE, towards the detection of sadness.

A positive correlation was observed between happiness and anger. This was unexpected, although according to Ayadi et al. happiness, anger and fear causes the sympathetic nervous system to be aroused causing an increased heart rate, which results in loud, fast speech with a higher frequency energy, a high average pitch and a wider pitch range (36). Classification of sentiments of similar activation (energy) levels such as happiness and anger depends on the valence dimension, for which there is no universal agreement on how to determine this through acoustic features. Therefore, these results confirm that further work must be done in the area of speech sentiment recognition to improve the classification of high frequency energy sentiments such happiness and anger.

From analysing the sentiment correlations from the openSMILE sentiment results, various trends were observed in the detection of sentiments which were also present in the results computed for the other leaders (Appendix A1.5.2). Based on these correlations it seems this system is classifying sadness very well as it was negatively correlated with all the other sentiments. This could be due to the fact that sadness is expressed very differently through speech than other sentiments, or possibly that the model is biased towards sadness. The positive correlation between happiness and anger shows inaccuracies in this system and emphasises that further work on the classification of sentiments of similar activation must be done.

Table 4.8: openSMILE sentiment correlation matrix for Sheryl Sandberg in an Emergency. Showing that sadness is negatively correlated with all other emotions (which could be indicative of a bias in the model) and also an unexpected positive correlation between happiness and anger indicating that this system has difficulties differentiating these sentiments.

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.1956	0.1297	0.6603	0.6851	-0.6013
Boredom	0.1956	1	0.1452	0.6969	0.594	-0.8443
Disgust	0.1297	0.1452	1	0.2631	0.1895	-0.4367
Fear	0.6603	0.6969	0.2631	1	0.9667	-0.9057
Happiness	0.6851	0.594	0.1895	0.9667	1	-0.8318
Sadness	-0.6013	-0.8443	-0.4367	-0.9057	-0.8318	1

4.4 Summary of Results

The results presented in this chapter revealed a deviance in the leaders' emotions which was determined by independence tests within the unimodal analyses (through facial expression and speech sentiment analysis) and was validated using the multimodal sentiment analysis. The independence tests consisted of chi-square tests in which the sentiment frequencies from the emergency and non-emergency scenarios were input to reveal whether a significant difference in the emotional behaviour of these women in these scenarios could be determined.

From testing the level of support between the modalities as part of the multimodal analysis, is was observed that the modalities often did not agree on this sentiment being expressed. The discrepancies between the facial-expression and speech results could be due to the different sampling rates used by the two systems, although the graphical representations of the sentiment results suggests a preference in openSMILE towards detecting sadness, which may be due to a bias within the model. Improvements must be made to the accuracy of the audio sentiment analysis system for multimodal analysis systems, such as this, to be used in a practical applications.

From the evaluation of the performance of the sentiment analysis systems used, FACET proved to operate correctly, whereas openSMILE appeared to have more inaccuracies. The inaccuracies which appeared in FACET were often due to the quality of the video or inherent complexities within emotion recognition (relationships between certain emotions and the intermingling of action units due to the connections between facial muscles). The inaccuracies within the performance of openSMILE generally concerned the detection of sentiments such as anger, happiness and fear. This may be because within speech affect recognition there are uncertainties related to the detection of sentiments of similar activation (such as happiness and anger).

5 Conclusion

5.1 Research conducted and challenges faced

Within leadership studies, techniques such as Sentiment Analysis are growing in popularity as they allow the user to gain insights into the emotional state and possibly the opinions held on a given subject through analysing how the leaders communicate non-verbally. The sentiment analysis systems available currently, although they use cutting edge machine learning, computer vision and signal processing technology, they remain underutilised due to usability issues.

This research aimed to improve the usability of these sentiment analysis systems to reduce the burden of interpretation on the user. In previous leadership studies which have utilised sentiment analysis techniques, the research has mainly focused on male leaders. Due to this, the chosen sample for this research consists of seven women who are considered leaders within their respective fields. This research aimed to expand the prior research done in this area by providing a comprehensive study of female leaders through analysing their non-verbal cues to see if a deviance within their behaviour within emergency and non-emergency scenarios can be determined. This research intended to interpret the results given by these systems through filtering and applying various analytical methods to gain statistically significant results in relation to the sentiment expressed by the leaders. The results are presented in both graphical representations of the overall sentiments felt by the leader and tabulated results in an effort to aid user comprehension.

This project consisted of four phases: • Data Collection • Pre-Processing • Post-Processing • Data Analysis. The Data Collection phase consisted of objectively gathering data from publicly available video footage of the chosen female leaders. The Pre-Processing phase was where the sample data was filtered to ensure it was of a suitable quality for sentiment analysis. The Post-Processing phase included the use of the FACET (facial-expression) and openSMILE (speech) sentiment analysis systems to gain the raw sentiment values from the aforementioned sample data. As the aim of this research was to aid the users understanding of these sentiment results, the Data Analysis Phase consisted of the creation of a program which would produce statistically significant results in relation to the sentiment of the leaders. The analysis performed within this program consisted of independence tests to determine deviations within the leaders behaviour between different scenarios, the production of graphical representations of the sentiment detected, data fusion and an evaluation of the performance of the sentiment analysis systems used. Independence tests were done using the sentiment frequencies of each leader in both scenarios, these values were input into chi-square tests which determined if there was a significant difference in the sentiments expressed in the different scenarios. A method to validate the statistical results relating to the facial expression sentiment data had to be determined. As facial expressions are often short-lived, an indirect method was needed to validate these results - using audio sentiment results. Multimodal results were used as a validation as the aggregation of the data from two non-verbal channels should be more indicative of the actual sentiments felt by the leader. A challenge faced during the Data Analysis phase was that different sampling rates were used by FACET and openSMILE. This added additional complexity to the data fusion process. To overcome this, various fusion approaches were attempted, but the final fusion approach consisted of getting the max values for every 150 FACET values (as 150 FACET values were equivalent to 1 openSMILE value) and fusing this with the sentiment values output from openSMILE.

By completing this research project an insight into non-verbal communication and it's importance in human communication was learnt. Concepts such as sentiment analysis and leadership studies were researched and critically evaluated. This research was an opportunity to work with big data, to automate the analysis of large amounts of sentiment data and from this derive useful insights into the leaders' behaviour.

5.1.1 Conclusions

From the graphical representations of the leaders' sentiments, an increase in negative sentiments was observed in the emergency situations - visually suggesting a significant difference in the behaviour of these women in this type of scenario. The facial expression chi-square test found that there was a significant difference in the sentiments expressed by the leaders in the emergency and non-emergency situations. This result was then validated externally using audio sentiment data in the unimodal and multimodal analyses by performing chi-square tests. The multimodal chi-square test was performed on each emotion of each leaders, this gave a more in depth analysis of their behaviour, revealing cases when a leader exerted control over selected emotions (as with Mary Robinson and her expression of fear), although overall, the leaders behaviour was deemed significantly different in the different type of scenarios. Based on these results, a deviance within the leaders behaviour was determined.

To verify the results determined within this project, an evaluation of the sentiment analysis systems was done. FACET's performance was deemed to be accurate based on the correlation of sentiments and the correlation between the sentiment detected and the action units which were active. Inaccuracies within its execution were viewed to be because of the quality of the video used or the position of the subject within the frame. openSMILE's performance was found to have some inaccuracies as it appeared to have a preference towards detecting sadness. This may be due to a bias within the training data of the machine learning model used. The sentiment correlations also revealed an unusual positive correlation between anger and happiness, it is likely this is due to an issue within speech affect recognition in differentiating sentiments with similar energy levels.

Although the program presented in this research was able to determine deviations within the leaders behaviour, at times the modalities disagreed with the presence of sentiments. Therefore, the precision of the audio sentiment analysis must be improved to gain more reliable multimodal results which could be used in real-world applications. Due to this, there are justifications for further work related to this research to be conducted.

5.2 Further Work

Upon completing this project, it was determined that further work could be conducted as a continuation of this research, this consisted of:

- Expanding the dataset to include multiple instances of each type of scenario.
- Using a sample of male and female leaders.
- Expanding the scope of project to include text analytics.
- Using a more diverse speech emotion dataset to create own emotion recognition model.
- Using Signal-To-Noise Ratio as a method of analysing the fused data.
- The use of human observers to create an inter-annotator agreement as to the sentiments being expressed.

The project presented in this dissertation consisted of an investigation of the sentiments expressed by female leaders in both emergency and non-emergency scenarios to determine if a "leakage" of emotion can be determined. If there was further time, this dataset could be expanded to include multiple instances of these these leaders in both types of scenarios, this would ensure the results would be truly representative of the population and not specific to the individual scenarios.

The sample could also be expanded further to investigate the expression of sentiment in both male and female leaders. By doing this, an investigation could be conducted to deduce if there are gender specific significant differences in the expression of sentiment in leaders.

The scope of this research could be expanded to perform text analytics on the spoken words of each leader in each scenario. By doing this, text-based sentiment analysis could be used to determine the sentiments expressed through the content of the leaders' speech. This would require the use of a speech-to-text software, to convert the spoken word to text. This would require a manual check of the resulting text to ensure the conversion has worked accurately. The converted text could then be processed with a text-based sentiment analysis system, to determine the sentiments being expressed. Further fusion could then be conducted with the non-verbal sentiment results presented in this project to determine if the sentiments expressed through verbal communication varies greatly with that conveyed through non-verbal communication.

The speech (acoustic) sentiment analysis conducted in this project was done so using an openEAR emotion recognition model which was trained with the Berlin Speech Dataset. The sentiment results gathered in this project shows a large detection of boredom and sadness in both scenarios for all leaders in the sample. This is an unexpected result and this leads us to question the accuracy of this model. If there was further time, an investigation into the available emotion speech datasets could be conducted to source a larger, more diverse dataset. This diverse dataset could then be used to train a new emotion recognition model which could hopefully provide more accurate speech sentiment results.

Through incorporating the Signal-To-Noise Ratio into the multimodal analysis, it provides another method of analysing the relationship between the modalities. In this method, one can observe if similar levels of noise are appearing at a given moment, which may indicate that the behaviour which the subject is presenting is unlike that found within the training data. This could also be useful for analysing cases in which the systems disagreed on the presence of a given emotion, if a high amount of noise is detected at this moment, the noise may be influencing the system causing unusual results.

Another method of analysing the performance of the sentiment analysis systems would be through the use of human observers. These human observers would be shown the filtered sample data and would generate there own judgements of the sentiments being expressed. All the judgements would then be aggregated (calculating the overlap between the agreeing judgements) to form an inter-annotator agreement. The inter-annotator agreement could then be used to evaluate the accuracy of the systems through computing how often the systems results correlate with the judgements of the annotators.

Bibliography

[1] Khurshid Ahmad. To Be Decided. Unpublished Literature Review, (no date).

[2] Sascha Segan. Apple iphone x sets the stage for full-screen ar. *PC Magazine*, pages 32 – 44, December 2017.

[3] Gwen C. Littlewort, Marian Stewart Bartlett, and Kang Lee. Automatic coding of facial expressions displayed during posed and genuine pain. *Image and Vision Computing*, 27(12):1797 – 1803, 2009. Visual and multimodal analysis of human spontaneous behaviour:.

[4] Marian Stewart Bartlett, Gwen Littlewort-Ford, Javier Movellan, Ian Fasel, and Mark Frank. Automated facial action coding system. US Patent 8798374 B2.

[5] F. Eyben, M. Wöllmer, and B. Schuller. Openear — introducing the munich open-source emotion and affect recognition toolkit. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–6. IEEE, Sep. 2009.

[6] Karen Anding Fontenot. Nonverbal communication and social cognition. *Salem Press Encyclopedia of Health*, 2018.

[7] Joyce E. Bono and Remus Ilies. Charisma, positive emotions and mood contagion. *The Leadership Quarterly*, 17(4):317 – 334, 2006.

[8] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski. Classifying facial actions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21 (10):974–989, Oct 1999.

[9] Joyce E. Bono and Remus Ilies. Charisma, positive emotions and mood contagion. *The Leadership Quarterly*, 17(4):317 – 334, 2006.

[10] Charles Darwin. *The Expression of the Emotions in Man and Animals*. London: Murray, London, 1872.

[11] A. Reis, D. Paulino, H. Paredes, I. Barroso, M. J. Monteiro, V. Rodrigues, and J. Barroso. Using intelligent personal assistants to assist the elderlies an evaluation of amazon alexa, google assistant, microsoft cortana, and apple siri. In *2018 2nd International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW)*, pages 1–5, June 2018.

[12] S. Shimizu, K. Shimada, and R. Murakami. Non-verbal communication-based emotion incitation robot. In *2018 IEEE 15th International Workshop on Advanced Motion Control (AMC)*, pages 338–341, March 2018.

[13] Alina Simona Tecău and Bianca Tescașiu. Nonverbal communication in the focus-group. *Bulletin of the Transilvania University of Brașov*, 8 (57)(2):119 – 124, 2015.

[14] Inmaculada Riquelme, Antonia Pades Jiménez, and Pedro Montoya. Parents and physiotherapists recognition of non-verbal communication of pain in individuals with cerebral palsy. *Health Communication*, 33(12):1448–1453, 2018.

[15] Luke Agnew. Sentiment analysis of facial expressions and asset price movement. Unpublished Bachelor's Dissertation, May 2018.

[16] Maria Spyropoulou and Khurshid Ahmad. Disaster-related public speeches: The role of emotions. *IEEE 2016 11th International Conference on Availability, Reliability and Security (ARES)*, page 800, 2016.

[17] Sigal G. Barsade. The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, 47(4):644 – 675, 2002.

[18] Neal M. Ashkanasy and Alana D. Dorris. Emotions in the workplace. *Annual Review of Organizational Psychology and Organizational Behavior*, 4:67 – 90, 2017.

[19] Daan van Knippenberg Peter J. Edelman. Training leader emotion regulation and leadership effectiveness. *Journal of Business*, 32(6):747 – 757, 2017.

[20] Mohammad Soleymani, David Garcia, Brendan Jou, Björn Schuller, Shih-Fu Chang, and Maja Pantic. A survey of multimodal sentiment analysis. *Image and Vision Computing*, 65:3 – 14, 2017. Multimodal Sentiment Analysis and Mining in the Wild Image and Vision Computing.

[21] Mika V. Mäntylä, Daniel Graziotin, and Miikka Kuutila. The evolution of sentiment analysis—a review of research topics, venues, and top cited papers. *Computer Science Review*, 27:16 – 32, 2018.

[22] V. Pérez Rosas, R. Mihalcea, and L. Morency. Multimodal sentiment analysis of spanish online videos. *IEEE Intelligent Systems*, 28(3):38–45, May 2013.

[23] Philip J Stone and Earl B Hunt. A computer approach to content analysis: studies using the general inquirer system. In *Proceedings of the May 21-23, 1963, spring joint computer conference*, pages 241–256. ACM, 1963.

[24] F. Mairesse, J. Polifroni, and G. Di Fabbrizio. Can prosody inform sentiment analysis? experiments on short spoken reviews. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5093–5096, March 2012.

[25] A. Kleinsmith and N. Bianchi-Berthouze. Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing*, 4(1):15–33, Jan 2013.

[26] V. Gentile, F. Milazzo, S. Sorce, A. Gentile, A. Augello, and G. Pilato. Body gestures and spoken sentences: A novel approach for revealing user's emotions. In *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*, pages 69–72, Jan 2017.

[27] iMotions. Biometric research, simplified, . `https://imotions.com/`. [Accessed: 29/11/2018].

[28] iMotions. Emotient facet module: Quick guide, . `https://help.imotions.com/hc/en-us/articles/203001451-Emotient-FACET-module-Quick-Guide?flash_digest=f6f1d9ec7b875f10b36400247625cdd51ed9acc4`. Requires iMotions account to access. [Accessed: 20/12/2018].

[29] Gwen Littlewort, Jacob Whitehill, Tingfan Wu, Ian Fasel, Mark Frank, Javier Movellan, and Marian Bartlett. The computer expression recognition toolbox (cert). *IEEE International Conference onFace and Gesture 2011, Automatic Face & Gesture Recognition and Workshops (FG 2011)*, page 298, 2011.

[30] Rolfe Winkler, Daisuke Wakabayashi, and Elizabeth Dwoskin. Apple buys artificial-intelligence startup emotient; emotient technology is used to assess emotions by reading facial expressions. *Wall Street Journal (Online)*, Jan 07 2016. `http://elib.tcd.ie/login?url=https://search.proquest.com/docview/1754234040?accountid=14404`. Last updated - 2017-11-22, [Accessed: 21/12/2018].

[31] T. Kanade, J. F. Cohn, and Yingli Tian. Comprehensive database for facial expression analysis. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pages 46–53, March 2000.

[32] Javier Movellan, Marian Stewart Bartlett, Ian Fasel, Gwen Ford Littlewort, Joshua Susskind, and Jacob Whitehill. Collection of machine learning training data for expression recognition.

[33] Notes on facial expression sentiment analysis. Unpublished notes, Personal communication. [Accessed: 15/01/2019].

[34] Florian Eyben, Felix Weninger, Florian Gross, and Björn Schuller. Recent developments in opensmile, the munich open-source multimedia feature extractor. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, pages 835–838, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2404-5.

[35] Felix Burkhardt, Astrid Paeschke, Melissa A Rolfes, Walter F. Sendlmeier, and Benjamin Weiss. A database of german emotional speech. In *INTERSPEECH*, 2005.

[36] Moataz El Ayadi, Mohamed S. Kamel, and Fakhri Karray. Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, 44(3):572 – 587, 2011. ISSN 0031-3203. doi: https://doi.org/10.1016/j.patcog.2010.09.020. URL http://www.sciencedirect.com/science/article/pii/S0031320310004619.

[37] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pages 94–101, June 2010. doi: 10.1109/CVPRW.2010.5543262.

[38] Minghai Chen, Sen Wang, Paul Pu Liang, Tadas Baltrušaitis, Amir Zadeh, and Louis-Philippe Morency. Multimodal sentiment analysis with word-level fusion and reinforcement learning. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, pages 163–171. ACM, 2017.

[39] Robert Plutchik. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 89(4):344–350, 2001. ISSN 00030996.

[40] iMotions. Facial action coding system (facs) – a visual guidebook, . https://imotions.com/blog/facial-action-coding-system/. [Accessed: 20/04/2018].

A1 Appendix

A1.1 19 action Units used in FACET

Table A1.1: Action Units used in FACET

Action Unit	Muscle/s Movement
AU1	Inner Brow Raiser Frontalis (Pars Medialis)
AU2	Outer Brow Raiser Frontalis (Pars Lateralis)
AU4	Brow Lowerer Depressor Glabellae, Depressor Supercilli; Corrugator
AU5	Upper Lid Raiser Levator Palpebrae Superioris
AU6	Cheek Raiser Orbicularis Oculi, Pars Orbitalis
AU7	Lid Tightener Orbicularis Oculi, Pars Palebralis
AU9	Nose Wrinkler Levator Labii Superioris Alaeque Nasi
AU10	Upper Lip Raiser Levator Labii Superioris aka Caput Infraorbitalis
AU12	Lip Corner Puller Zygomatic major
AU14	Dimpler Buccinator
AU15	Lip Corner Depressor Triangularis aka depressor anguli oris
AU17	Chin Raiser Mentalis
AU18	Lip Puckerer Incisivii Labii Superioris; Incisivii Labii I
AU20	Lip Stretcher Risorius
AU23	Lip Tightener Risorius
AU24	Lip Pressor Orbicularis Oris
AU25	Lips Part Depressor Labii, or Relaxation of Mentalis or Orbicularis Oris
AU26	Jaw Drop Masetter; Temporal and Internal Pterygoid Relaxed
AU28	Lip Suck Orbicularis Oris
AU43	Eye Closure; the levator palpebrae superioris muscle

A1.2 Video URLs

1. Sheryl Sandberg Emergent Scenario:

   Q&A: Facebook's Sheryl Sandberg

   `https://www.youtube.com/watch?v=HMO0hSogjhg`

2. Sheryl Sandberg Non-Emergent Scenario:

   Ideas for Tomorrow | Sheryl Sandberg, Chief Operating Officer of Facebook - Full Program

   `https://www.youtube.com/watch?v=YPhxCbRnLH8`

3. Sonia Sotomayor Emergent Scenario:

   Sonia Sotomayor: Supreme Court Nomination Hearings from PBS NewsHour and EMK Institute

   `https://www.youtube.com/watch?v=iMoatA0LWxI`

4. Sonia Sotomayor Non-Emergent Scenario:

   Sonia Sotomayor on Turning Pages: My Life Story and The Beloved World of Sonia Sotomayor

   `https://www.youtube.com/watch?v=GbPfucLRP0s`

5. Mary Robinson Emergent Scenario:

   William Binchy V. Mary Robinson (1983 Pro-Life Referendum)

   `https://www.youtube.com/watch?v=GLWnoQjTNiw`

6. Mary Robinson Non-Emergent Scenario:

   Everybody Matters: Climate Change and Human Rights

   `https://www.youtube.com/watch?v=n1YIedha8dM`

7. Heather Bresch Emergent Scenario:

   Rep. Chaffetz questions Mylan CEO Heather Bresch on EpiPen prices (C-SPAN)

   `https://www.youtube.com/watch?v=JokN5Aj0TV0`

8. Heather Bresch Non-Emergent Scenario:

   Mylan CEO on FDA-approved biosimilar, Trump's drug price plan

   `https://www.youtube.com/watch?v=PXI2YdQNnhU`

9. Malala Yousafzai Emergent Scenario:

   Malala Yousafzai speech in Islamabad

   https://www.youtube.com/watch?v=Ws6apTYM7nc

10. Malala Yousafzai Non-Emergent Scenario:

    A Conversation with Malala Yousafzai

    https://www.youtube.com/watch?v=EOLs8H7qrPE

11. Alexandria Ocasio-Cortez Emergent Scenario:

    Alexandria Ocasio-Cortez Talks to Glenn Greenwald About the Democratic Party and 2018 Midterms

    https://www.youtube.com/watch?v=zuoKLLLpiuE

12. Alexandria Ocasio-Cortez Non-Emergent Scenario:

    Congressional Candidate Alexandria Ocasio-Cortez on Healthcare, Education & Voting

    https://www.youtube.com/watch?v=LgvSTVw1Pw8

13. Theresa May Emergent Scenario:

    Theresa May gives press conference as she battles to save Brexit

    https://www.youtube.com/watch?v=zhJPRNMyFx4

14. Theresa May Non-Emergent Scenario:

    New Year 2018: Theresa May's Message

    https://www.youtube.com/watch?v=f205q1XQs78&t=50s

A1.3 Example Bash Script for Speech Processing

This is an example of the bash script which is created on-the-fly within the R program described in Chapter ??. Within the R program, the sample video was converted to a WAV file, this WAV file was then split into 5 second long segments which were named sequentially. The program then creates this script to iterate through each segment WAV file and process it with openSMILE. In the case that it is processing the first segment, it processes it with openSMILE and writes the emotion probabilities to a newly created results file, for each subsequent segment, the emotion results are appended to the aforementioned file.

Listing A1.1: Example bash script for speech processing

```
#!/bin/bash
cd /Users/alannaogrady/Downloads/opensmile-2.3.0
for i in $(seq -f "%03g" 0 43)
do
    echo "/out$i.wav"
    if [ $i -eq 0 ]
    then
        SMILExtract -C config/emobase_live4_batch_single.conf
        -I /NVC_Project/Data/bresch_emergent_1/out$i.wav >
        NVC_Project/AudioResults/bresch_emergent_1.txt
    else
        SMILExtract -C config/emobase_live4_batch_single.conf
        -I /NVC_Project/Data/bresch_emergent_1/out\$i.wav >>
        /NVC_Project/AudioResults/bresch_emergent_1.txt
    fi
done
```

A1.4 Unimodal Sentiment Results

A1.4.1 Facial-Expression Signal-To-Noise Ratio Data

Table A1.2: SNR Information for Sheryl Sandberg - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	47	47	47	47	47	47	47
Rejected	6267	7514	6283	6582	7054	3108	7105
Accepted	1307	60	1291	992	520	4466	469
Number of Values	7621	7621	7621	7621	7621	7621	7621
Total Accept %	17.1	0.8	16.9	13	6.8	58.6	6.2
Signal To Noise Ratio %	82.9	99.2	83.1	87	93.2	41.4	93.8

Table A1.3: SNR Information for Sheryl Sandberg - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	13	13	13	13	13	13	13
Rejected	6197	7081	6052	5399	6888	4638	5944
Accepted	916	32	1061	1714	225	2475	1169
Number of Values	7126	7126	7126	7126	7126	7126	7126
Total Accept %	12.9	0.4	14.9	24.1	3.2	34.7	16.4
Signal To Noise Ratio %	87.1	99.6	85.1	75.9	96.8	65.3	83.6

Table A1.4: SNR Information for Sonia Sotomayor - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	41	41	41	41	41	41	41
Rejected	5562	6635	6610	6503	6613	1132	5196
Accepted	1092	19	44	151	41	5522	1458
Number of Values	6695	6695	6695	6695	6695	6695	6695
Total Accept %	16.3	0.3	0.7	2.3	0.6	82.5	21.8
Signal To Noise Ratio %	83.7	99.7	99.3	97.7	99.4	17.5	78.2

Table A1.5: SNR Information for Sonia Sotomayor - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	0	0	0	0	0	0	0
Rejected	6166	7000	5552	6401	7036	5531	6705
Accepted	939	105	1553	704	69	1574	400
Number of Values	7105	7105	7105	7105	7105	7105	7105
Total Accept %	13.2	1.5	21.9	9.9	1	22.2	5.6
Signal To Noise Ratio %	86.8	98.5	78.1	90.1	99	77.8	94.4

Table A1.6: SNR Information for Mary Robinson - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	0	0	0	0	0	0	0
Rejected	2180	7631	7623	7625	7514	6263	7628
Accepted	5453	2	10	8	119	1370	5
Number of Values	7633	7633	7633	7633	7633	7633	7633
Total Accept %	71.4	0	0.1	0.1	1.6	17.9	0.1
Signal To Noise Ratio %	28.6	100	99.9	99.9	98.4	82.1	99.9

Table A1.7: SNR Information for Mary Robinson - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	872	872	872	872	872	872	872
Rejected	1116	5532	3716	5020	4117	5370	5530
Accepted	4416	0	1816	512	1415	162	2
Number of Values	6404	6404	6404	6404	6404	6404	6404
Total Accept %	69	0	28.4	8	22.1	2.5	0
Signal To Noise Ratio %	31	100	71.6	92	77.9	97.5	100

Table A1.8: SNR Information for Heather Bresch - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	0	0	0	0	0	0	0
Rejected	4915	5653	3614	3384	5601	5493	5273
Accepted	738	0	2039	2269	52	160	380
Number of Values	5653	5653	5653	5653	5653	5653	5653
Total Accept %	13.1	0	36.1	40.1	0.9	2.8	6.7
Signal To Noise Ratio %	86.9	100	63.9	59.9	99.1	97.2	93.3

Table A1.9: SNR Information for Heather Bresch - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	0	0	0	0	0	0	0
Rejected	6322	6411	3802	2196	6215	6410	5953
Accepted	89	0	2609	4215	196	1	458
Number of Values	6411	6411	6411	6411	6411	6411	6411
Total Accept %	1.4	0	40.7	65.7	3.1	0	7.1
Signal To Noise Ratio %	98.6	100	59.3	34.3	96.9	100	92.9

Table A1.10: SNR Information for Malala Yousafzai - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	45	45	45	45	45	45	45
Rejected	6318	7426	6996	7120	5005	6869	7341
Accepted	1108	0	430	306	2421	557	85
Number of Values	7471	7471	7471	7471	7471	7471	7471
Total Accept %	14.8	0	5.8	4.1	32.4	7.5	1.1
Signal To Noise Ratio %	85.2	100	94.2	95.9	67.6	92.5	98.9

Table A1.11: SNR Information for Malala Yousafzai - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	18	18	18	18	18	18	18
Rejected	4769	5583	5559	5452	4861	2658	2981
Accepted	821	7	31	138	729	2932	2609
Number of Values	5608	5608	5608	5608	5608	5608	5608
Total Accept %	14.6	0.1	0.6	2.5	13	52.3	46.5
Signal To Noise Ratio %	85.4	99.9	99.4	97.5	87	47.7	53.5

Table A1.12: SNR Information for Alexandria Ocasio-Cortez - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	622	622	622	622	622	622	622
Rejected	10708	12826	9833	10403	11554	10750	12607
Accepted	2190	72	3065	2495	1344	2148	291
Number of Values	13520	13520	13520	13520	13520	13520	13520
Total Accept %	16.2	0.5	22.7	18.5	9.9	15.9	2.2
Signal To Noise Ratio %	83.8	99.5	77.3	81.5	90.1	84.1	97.8

Table A1.13: SNR Information for Alexandria Ocasio-Cortez - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	0	0	0	0	0	0	0
Rejected	7791	8099	5396	4516	8077	5882	7838
Accepted	309	1	2704	3584	23	2218	262
Number of Values	8100	8100	8100	8100	8100	8100	8100
Total Accept %	3.8	0	33.4	44.2	0.3	27.4	3.2
Signal To Noise Ratio %	96.2	100	66.6	55.8	99.7	72.6	96.8

Table A1.14: SNR Information for Theresa May - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	31	31	31	31	31	31	31
Rejected	3721	6728	6463	6125	6192	5221	6261
Accepted	3012	5	270	608	541	1512	472
Number of Values	6764	6764	6764	6764	6764	6764	6764
Total Accept %	44.5	0.1	4	9	8	22.4	7
Signal To Noise Ratio %	55.5	99.9	96	91	92	77.6	93

Table A1.15: SNR Information for Theresa May - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
No Signal	67	67	67	67	67	67	67
Rejected	6022	7687	6998	7262	7113	3695	7477
Accepted	1694	29	718	454	603	4021	239
Number of Values	7783	7783	7783	7783	7783	7783	7783
Total Accept %	21.8	0.4	9.2	5.8	7.7	51.7	3.1
Signal To Noise Ratio %	78.2	99.6	90.8	94.2	92.3	48.3	96.9

A1.4.2 Speech Signal-To-Noise Ratio Data

Table A1.16: SNR Information for Sheryl Sandberg - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	50	14	43	51	51	0
Accepted	1	37	8	0	0	51
Number of Values	51	51	51	51	51	51
Total Accept %	2	72.5	15.7	0	0	100
Signal To Noise Ratio %	98	27.5	84.3	100	100	0

Table A1.17: SNR Information for Sheryl Sandberg - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	34	10	19	48	45	5
Accepted	14	38	29	0	3	43
Number of Values	48	48	48	48	48	48
Total Accept %	29.2	79.2	60.4	0	6.2	89.6
Signal To Noise Ratio %	70.8	20.8	39.6	100	93.8	10.4

Table A1.18: SNR Information for Sonia Sotomayor - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	45	22	42	45	45	0
Accepted	0	23	3	0	0	45
Number of Values	45	45	45	45	45	45
Total Accept %	0	51.1	6.7	0	0	100
Signal To Noise Ratio %	100	48.9	93.3	100	100	0

Table A1.19: SNR Information for Sonia Sotomayor - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	48	25	48	48	48	0
Accepted	0	23	0	0	0	48
Number of Values	48	48	48	48	48	48
Total Accept %	0	47.9	0	0	0	100
Signal To Noise Ratio %	100	52.1	100	100	100	0

Table A1.20: SNR Information for Mary Robinson - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	51	9	32	50	51	1
Accepted	0	42	19	1	0	50
Number of Values	51	51	51	51	51	51
Total Accept %	0	82.4	37.3	2	0	98
Signal To Noise Ratio %	100	17.6	62.7	98	100	2

Table A1.21: SNR Information for Mary Robinson - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	30	1	29	25	24	5
Accepted	13	42	14	18	19	38
Number of Values	43	43	43	43	43	43
Total Accept %	30.2	97.7	32.6	41.9	44.2	88.4
Signal To Noise Ratio %	69.8	2.3	67.4	58.1	55.8	11.6

Table A1.22: SNR Information for Heather Bresch - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	28	0	17	38	28	10
Accepted	10	38	21	0	10	28
Number of Values	38	38	38	38	38	38
Total Accept %	26.3	100	55.3	0	26.3	73.7
Signal To Noise Ratio %	73.7	0	44.7	100	73.7	26.3

Table A1.23: SNR Information for Heather Bresch - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	31	0	8	42	38	15
Accepted	12	43	35	1	5	28
Number of Values	43	43	43	43	43	43
Total Accept %	27.9	100	81.4	2.3	11.6	65.1
Signal To Noise Ratio %	72.1	0	18.6	97.7	88.4	34.9

Table A1.24: SNR Information for Malala Yousafzai - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	50	0	17	50	50	8
Accepted	0	50	33	0	0	42
Number of Values	50	50	50	50	50	50
Total Accept %	0	100	66	0	0	84
Signal To Noise Ratio %	100	0	34	100	100	16

Table A1.25: SNR Information for Malala Yousafzai - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	45	7	45	45	45	0
Accepted	0	38	0	0	0	45
Number of Values	45	45	45	45	45	45
Total Accept %	0	84.4	0	0	0	100
Signal To Noise Ratio %	100	15.6	100	100	100	0

Table A1.26: SNR Information for Alexandria Ocasio-Cortez - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	29	1	16	46	42	6
Accepted	17	45	30	0	4	40
Number of Values	46	46	46	46	46	46
Total Accept %	37	97.8	65.2	0	8.7	87
Signal To Noise Ratio %	63	2.2	34.8	100	91.3	13

Table A1.27: SNR Information for Alexandria Ocasio-Cortez - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	18	0	34	54	46	31
Accepted	37	55	21	1	9	24
Number of Values	55	55	55	55	55	55
Total Accept %	67.3	100	38.2	1.8	16.4	43.6
Signal To Noise Ratio %	32.7	0	61.8	98.2	83.6	56.4

Table A1.28: SNR Information for Theresa May - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	55	11	46	55	55	0
Accepted	0	44	9	0	0	55
Number of Values	55	55	55	55	55	55
Total Accept %	0	80	16.4	0	0	100
Signal To Noise Ratio %	100	20	83.6	100	100	0

Table A1.29: SNR Information for Theresa May - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
No Signal	0	0	0	0	0	0
Rejected	51	3	29	52	52	0
Accepted	1	49	23	0	0	52
Number of Values	52	52	52	52	52	52
Total Accept %	1.9	94.2	44.2	0	0	100
Signal To Noise Ratio %	98.1	5.8	55.8	100	100	0

A1.5 Emotion Correlations

A1.5.1 Facial Expression Emotion Correlations

Table A1.30: Sheryl Sandberg - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.0278	-0.2533	0.2904	0.1969	-0.3273	-0.0195
Anger	-0.0278	1	-0.0264	-0.1657	0.4828	0.0084	0.5947
Surprise	-0.2533	-0.0264	1	0.4555	0.0811	-0.502	0.1874
Fear	0.2904	-0.1657	0.4555	1	0.1177	-0.3554	0.2353
Contempt	0.1969	0.4828	0.0811	0.1177	1	-0.2341	0.7936
Disgust	-0.3273	0.0084	-0.502	-0.3554	-0.2341	1	-0.2224
Sadness	-0.0195	0.5947	0.1874	0.2353	0.7936	-0.2224	1

Table A1.31: Sheryl Sandberg - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	0.081	-0.4253	0.091	0.2867	0.0707	-0.0758
Anger	0.081	1	-0.2947	-0.6876	0.2929	0.4412	0.0815
Surprise	-0.4253	-0.2947	1	0.1298	-0.2341	-0.5352	-0.1935
Fear	0.091	-0.6876	0.1298	1	-0.0513	-0.2739	0.057
Contempt	0.2867	0.2929	-0.2341	-0.0513	1	-0.0277	0.7512
Disgust	0.0707	0.4412	-0.5352	-0.2739	-0.0277	1	-0.1425
Sadness	-0.0758	0.0815	-0.1935	0.057	0.7512	-0.1425	1

Table A1.32: Sonia Sotomayor - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.4419	0.2565	0.4921	0.191	-0.3561	-0.3719
Anger	-0.4419	1	0.1992	-0.2454	0.47	0.2206	0.5762
Surprise	0.2565	0.1992	1	0.5982	0.5292	-0.3729	0.309
Fear	0.4921	-0.2454	0.5982	1	0.3151	-0.2381	0.1109
Contempt	0.191	0.47	0.5292	0.3151	1	-0.2741	0.6816
Disgust	-0.3561	0.2206	-0.3729	-0.2381	-0.2741	1	0.0067
Sadness	-0.3719	0.5762	0.309	0.1109	0.6816	0.0067	1

Table A1.33: Sonia Sotomayor - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.4117	-0.4423	0.178	-0.1514	0.3342	-0.4328
Anger	-0.4117	1	-0.0211	-0.2695	0.3303	0.1208	0.3547
Surprise	-0.4423	-0.0211	1	0.2967	-0.0591	-0.2373	0.0135
Fear	0.178	-0.2695	0.2967	1	0.2601	-0.2644	0.2221
Contempt	-0.1514	0.3303	-0.0591	0.2601	1	-0.3054	0.8025
Disgust	0.3342	0.1208	-0.2373	-0.2644	-0.3054	1	-0.4184
Sadness	-0.4328	0.3547	0.0135	0.2221	0.8025	-0.4184	1

Table A1.34: Mary Robinson - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.7605	-0.4343	-0.1848	-0.6679	-0.8858	-0.6544
Anger	-0.7605	1	0.3776	0.0984	0.6968	0.8119	0.7622
Surprise	-0.4343	0.3776	1	0.6848	0.5865	0.4496	0.4266
Fear	-0.1848	0.0984	0.6848	1	0.3523	0.1428	0.2514
Contempt	-0.6679	0.6968	0.5865	0.3523	1	0.7284	0.8834
Disgust	-0.8858	0.8119	0.4496	0.1428	0.7284	1	0.7237
Sadness	-0.6544	0.7622	0.4266	0.2514	0.8834	0.7237	1

Table A1.35: Mary Robinson - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.6467	-0.734	-0.4897	-0.5653	-0.6937	-0.6602
Anger	-0.6467	1	0.5531	0.4741	0.668	0.8689	0.8151
Surprise	-0.734	0.5531	1	0.7053	0.4283	0.6746	0.4874
Fear	-0.4897	0.4741	0.7053	1	0.3226	0.5341	0.4324
Contempt	-0.5653	0.668	0.4283	0.3226	1	0.5932	0.821
Disgust	-0.6937	0.8689	0.6746	0.5341	0.5932	1	0.7552
Sadness	-0.6602	0.8151	0.4874	0.4324	0.821	0.7552	1

Table A1.36: Heather Bresch - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.2074	-0.4005	0.4474	0.0437	0.32	-0.3906
Anger	-0.2074	1	0.2528	-0.307	0.51	0.3583	0.4838
Surprise	-0.4005	0.2528	1	-0.1387	-0.2819	-0.3288	-0.2017
Fear	0.4474	-0.307	-0.1387	1	-0.1005	0.1481	-0.2797
Contempt	0.0437	0.51	-0.2819	-0.1005	1	0.2338	0.8231
Disgust	0.32	0.3583	-0.3288	0.1481	0.2338	1	0.219
Sadness	-0.3906	0.4838	-0.2017	-0.2797	0.8231	0.219	1

Table A1.37: Heather Bresch - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.2151	-0.1931	0.5948	0.2945	0.3355	-0.4198
Anger	-0.2151	1	-0.0722	-0.3619	0.4343	0.3043	0.4853
Surprise	-0.1931	-0.0722	1	-0.2804	-0.6859	0.0364	-0.4599
Fear	0.5948	-0.3619	-0.2804	1	0.051	0.1556	-0.391
Contempt	0.2945	0.4343	-0.6859	0.051	1	0.2133	0.6354
Disgust	0.3355	0.3043	0.0364	0.1556	0.2133	1	0.0335
Sadness	-0.4198	0.4853	-0.4599	-0.391	0.6354	0.0335	1

Table A1.38: Malala Yousafzai - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.1171	0.0171	0.5928	0.639	-0.1537	0.1839
Anger	-0.1171	1	0.4085	0.2476	-0.1193	0.0038	0.257
Surprise	0.0171	0.4085	1	0.5265	-0.2246	-0.2455	0.1113
Fear	0.5928	0.2476	0.5265	1	0.3412	-0.195	0.2502
Contempt	0.639	-0.1193	-0.2246	0.3412	1	-0.304	0.4093
Disgust	-0.1537	0.0038	-0.2455	-0.195	-0.304	1	-0.0476
Sadness	0.1839	0.257	0.1113	0.2502	0.4093	-0.0476	1

Table A1.39: Malala Yousafzai - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.2218	-0.1126	-0.1135	0.5172	-0.1561	-0.1469
Anger	-0.2218	1	0.3141	0.4015	0.0754	0.3229	-0.3149
Surprise	-0.1126	0.3141	1	0.6261	0.0992	-0.3202	-0.2782
Fear	-0.1135	0.4015	0.6261	1	-0.1704	0.0616	-0.2725
Contempt	0.5172	0.0754	0.0992	-0.1704	1	-0.4968	0.205
Disgust	-0.1561	0.3229	-0.3202	0.0616	-0.4968	1	-0.1835
Sadness	-0.1469	-0.3149	-0.2782	-0.2725	0.205	-0.1835	1

Table A1.40: Alexandria Ocasio-Cortez - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	0.0473	-0.3144	0.1712	-0.1846	0.4275	-0.3082
Anger	0.0473	1	0.3279	0.2534	0.6688	0.3607	0.7048
Surprise	-0.3144	0.3279	1	0.5665	0.3049	-0.412	0.3505
Fear	0.1712	0.2534	0.5665	1	0.2615	-0.0985	0.2431
Contempt	-0.1846	0.6688	0.3049	0.2615	1	0.07	0.9093
Disgust	0.4275	0.3607	-0.412	-0.0985	0.07	1	0.0142
Sadness	-0.3082	0.7048	0.3505	0.2431	0.9093	0.0142	1

Table A1.41: Alexandria Ocasio-Cortez - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.0117	-0.5597	-0.1599	-0.0581	0.5971	-0.2565
Anger	-0.0117	1	-0.0176	-0.3106	0.5056	0.3207	0.3875
Surprise	-0.5597	-0.0176	1	0.4421	-0.1347	-0.5889	-0.0781
Fear	-0.1599	-0.3106	0.4421	1	-0.407	-0.1956	-0.3471
Contempt	-0.0581	0.5056	-0.1347	-0.407	1	0.0534	0.9097
Disgust	0.5971	0.3207	-0.5889	-0.1956	0.0534	1	-0.1031
Sadness	-0.2565	0.3875	-0.0781	-0.3471	0.9097	-0.1031	1

Table A1.42: Theresa May - Non-Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.3025	-0.433	0.1112	0.0842	0.0425	-0.296
Anger	-0.3025	1	0.1264	-0.518	0.4201	0.4754	0.3362
Surprise	-0.433	0.1264	1	0.0922	0.1974	-0.2542	0.4576
Fear	0.1112	-0.518	0.0922	1	-0.0711	-0.1529	0.1099
Contempt	0.0842	0.4201	0.1974	-0.0711	1	-0.0727	0.8088
Disgust	0.0425	0.4754	-0.2542	-0.1529	-0.0727	1	-0.1324
Sadness	-0.296	0.3362	0.4576	0.1099	0.8088	-0.1324	1

Table A1.43: Theresa May - Emergency

	Joy	Anger	Surprise	Fear	Contempt	Disgust	Sadness
Joy	1	-0.0715	0.1661	0.3889	0.4951	-0.0873	0.1579
Anger	-0.0715	1	0.0704	-0.1666	0.2772	0.3124	0.1802
Surprise	0.1661	0.0704	1	0.2057	0.3286	-0.6081	0.2288
Fear	0.3889	-0.1666	0.2057	1	0.0671	-0.0314	0.148
Contempt	0.4951	0.2772	0.3286	0.0671	1	-0.1943	0.8068
Disgust	-0.0873	0.3124	-0.6081	-0.0314	-0.1943	1	-0.091
Sadness	0.1579	0.1802	0.2288	0.148	0.8068	-0.091	1

A1.5.2 Speech Emotion Correlations

Table A1.44: Sheryl Sandberg - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.777	0.6964	0.9129	0.9777	-0.8465
Boredom	0.777	1	0.7289	0.8877	0.8485	-0.9868
Disgust	0.6964	0.7289	1	0.7509	0.7276	-0.8017
Fear	0.9129	0.8877	0.7509	1	0.9443	-0.9299
Happiness	0.9777	0.8485	0.7276	0.9443	1	-0.9057
Sadness	-0.8465	-0.9868	-0.8017	-0.9299	-0.9057	1

Table A1.45: Sheryl Sandberg - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.1956	0.1297	0.6603	0.6851	-0.6013
Boredom	0.1956	1	0.1452	0.6969	0.594	-0.8443
Disgust	0.1297	0.1452	1	0.2631	0.1895	-0.4367
Fear	0.6603	0.6969	0.2631	1	0.9667	-0.9057
Happiness	0.6851	0.594	0.1895	0.9667	1	-0.8318
Sadness	-0.6013	-0.8443	-0.4367	-0.9057	-0.8318	1

Table A1.46: Sonia Sotomayor - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.8387	0.9063	0.9592	0.9815	-0.9182
Boredom	0.8387	1	0.6996	0.9015	0.8875	-0.9816
Disgust	0.9063	0.6996	1	0.8326	0.8853	-0.8198
Fear	0.9592	0.9015	0.8326	1	0.9737	-0.9504
Happiness	0.9815	0.8875	0.8853	0.9737	1	-0.9507
Sadness	-0.9182	-0.9816	-0.8198	-0.9504	-0.9507	1

Table A1.47: Sonia Sotomayor - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.8317	0.9235	0.8746	0.9298	-0.8815
Boredom	0.8317	1	0.8852	0.8449	0.925	-0.9929
Disgust	0.9235	0.8852	1	0.8178	0.904	-0.9191
Fear	0.8746	0.8449	0.8178	1	0.9732	-0.8917
Happiness	0.9298	0.925	0.904	0.9732	1	-0.9611
Sadness	-0.8815	-0.9929	-0.9191	-0.8917	-0.9611	1

Table A1.48: Mary Robinson - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.7433	0.7196	0.7366	0.8896	-0.8504
Boredom	0.7433	1	0.5706	0.643	0.8111	-0.9728
Disgust	0.7196	0.5706	1	0.6009	0.6826	-0.7124
Fear	0.7366	0.643	0.6009	1	0.7656	-0.748
Happiness	0.8896	0.8111	0.6826	0.7656	1	-0.8966
Sadness	-0.8504	-0.9728	-0.7124	-0.748	-0.8966	1

Table A1.49: Mary Robinson - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	-0.033	-0.3677	0.6599	0.8682	-0.7134
Boredom	-0.033	1	0.2163	-0.2743	0.1133	-0.5861
Disgust	-0.3677	0.2163	1	-0.6296	-0.4487	0.2084
Fear	0.6599	-0.2743	-0.6296	1	0.7407	-0.5343
Happiness	0.8682	0.1133	-0.4487	0.7407	1	-0.8346
Sadness	-0.7134	-0.5861	0.2084	-0.5343	-0.8346	1

Table A1.50: Heather Bresch - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	-0.3555	-0.4032	0.5376	0.5214	-0.4457
Boredom	-0.3555	1	-0.147	-0.603	-0.5952	-0.2834
Disgust	-0.4032	-0.147	1	-0.3614	-0.5827	0.7866
Fear	0.5376	-0.603	-0.3614	1	0.7426	-0.3216
Happiness	0.5214	-0.5952	-0.5827	0.7426	1	-0.5315
Sadness	-0.4457	-0.2834	0.7866	-0.3216	-0.5315	1

Table A1.51: Heather Bresch -Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.2274	-0.6086	0.8345	0.9235	-0.65
Boredom	0.2274	1	-0.528	0.181	0.063	-0.8209
Disgust	-0.6086	-0.528	1	-0.5726	-0.5531	0.5091
Fear	0.8345	0.181	-0.5726	1	0.8173	-0.5557
Happiness	0.9235	0.063	-0.5531	0.8173	1	-0.5312
Sadness	-0.65	-0.8209	0.5091	-0.5557	-0.5312	1

Table A1.52: Malala Yousafzai - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.4752	0.1503	0.5174	0.9163	-0.7636
Boredom	0.4752	1	0.1272	0.4276	0.5496	-0.8761
Disgust	0.1503	0.1272	1	-0.0589	0.057	-0.2091
Fear	0.5174	0.4276	-0.0589	1	0.6297	-0.6133
Happiness	0.9163	0.5496	0.057	0.6297	1	-0.8502
Sadness	-0.7636	-0.8761	-0.2091	-0.6133	-0.8502	1

Table A1.53: Malala Yousafzai - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.9506	0.9773	0.9517	0.9936	-0.9734
Boredom	0.9506	1	0.9338	0.9542	0.95	-0.9929
Disgust	0.9773	0.9338	1	0.9289	0.9777	-0.9639
Fear	0.9517	0.9542	0.9289	1	0.965	-0.961
Happiness	0.9936	0.95	0.9777	0.965	1	-0.9757
Sadness	-0.9734	-0.9929	-0.9639	-0.961	-0.9757	1

Table A1.54: Alexandria Ocasio-Cortez - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.0375	-0.167	0.8124	0.9451	-0.6086
Boredom	0.0375	1	-0.4385	0.3348	0.0645	-0.795
Disgust	-0.167	-0.4385	1	-0.3062	-0.2567	0.3722
Fear	0.8124	0.3348	-0.3062	1	0.8464	-0.7853
Happiness	0.9451	0.0645	-0.2567	0.8464	1	-0.6173
Sadness	-0.6086	-0.795	0.3722	-0.7853	-0.6173	1

Table A1.55: Alexandria Ocasio-Cortez - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.1236	-0.4146	-0.2193	0.7483	-0.5528
Boredom	0.1236	1	0.2188	-0.6775	-0.3019	-0.2809
Disgust	-0.4146	0.2188	1	-0.5083	-0.6768	0.6427
Fear	-0.2193	-0.6775	-0.5083	1	0.1953	-0.2424
Happiness	0.7483	-0.3019	-0.6768	0.1953	1	-0.5991
Sadness	-0.5528	-0.2809	0.6427	-0.2424	-0.5991	1

Table A1.56: Theresa May - Non-Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.8124	0.6935	0.9214	0.9623	-0.884
Boredom	0.8124	1	0.6329	0.873	0.8774	-0.9825
Disgust	0.6935	0.6329	1	0.663	0.6976	-0.7417
Fear	0.9214	0.873	0.663	1	0.9513	-0.9209
Happiness	0.9623	0.8774	0.6976	0.9513	1	-0.9327
Sadness	-0.884	-0.9825	-0.7417	-0.9209	-0.9327	1

Table A1.57: Theresa May - Emergency

	Anger	Boredom	Disgust	Fear	Happiness	Sadness
Anger	1	0.3036	0.511	0.5897	0.8695	-0.6076
Boredom	0.3036	1	-0.1455	0.5699	0.5412	-0.9085
Disgust	0.511	-0.1455	1	0.0996	0.4786	-0.2367
Fear	0.5897	0.5699	0.0996	1	0.6869	-0.7098
Happiness	0.8695	0.5412	0.4786	0.6869	1	-0.8003
Sadness	-0.6076	-0.9085	-0.2367	-0.7098	-0.8003	1