De La Salle University

Recognizing Reader's Affect Using EEG Data

A Thesis Document
Presented to
the Faculty of the College of Computer Studies
De La Salle University Manila

In Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Computer Science

by

KALAW, Kristine Ma. Dominique F.

Ethel Chua Joy ONG
Adviser

April 19, 2017

APPROVAL SHEET

This project paper hereto entitled:

Recognizing Reader's Affect Using EEG Data

prepared and submitted by **Kristine Ma. Dominique F. Kalaw** in partial fulfilment of the requirements for the degree of **Master of Science in Computer Science** has been examined and is recommended for acceptance and approval for ORAL EXAMINATION.

Ms. Ethel Chua Joy Ong
Adviser

Approved by the Committee on Oral Examination with a grade of PASSED on April 18, 2017.

Dr. Judith Azcarraga
Chair

Dr. Nelson Marcos
Member

Ms. Ethel Chua Joy Ong
Member

Accepted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science.

Dr. Merlin Teodosia C. Suarez
Dean
College of Computer Studies

The thesis entitled

Recognizing Reader's Affect Using EEG Data

after having been recommended and reviewed, is hereby approved by the Software Technology Department, College of Computer Studies, De La Salle University:

RACabredo

Dr. Rafael A. Cabredo
Chairperson
Software Technology Department

Dr. Merlin Teodosia Suarez
Dean
College of Computer Studies

April 19, 2017
Date

4/19/17
Date

Acknowledgments

I thank my thesis adviser, Ms. Ethel Ong, for sparking my interest in research as well as for challenging me intellectually. Her words "*Follow your interests*" stuck with me while I was reflecting on the options for my thesis topic. This passion prompted me to consult with her because her research area in computational storytelling align with my interest in reading fiction novels. Her steady guidance helped me commit to my passion and own this research.

Next, I am grateful to Dr. Shirley Lua for her recommendations to this research, and Dr. Judith Azcarraga and Mr. Roy Navea for their expertise regarding EEG processing. Much gratitude also goes to my past professors in the various colleges of the university, for sparing me a part of their class time to present my work to their students and solicit participants in the data collection. Without their participation and input, the experiments would not have been conducted successfully.

Much appreciation also goes to my friends and batch mates who are my sounding boards for my ideas. Special thanks also to Sharmaine and Bianca, not only for sharing and discussing ideas with me, but also for assisting me in some of the more tedious tasks of this research.

I am likewise grateful to Tita Pet for proofreading my work. I appreciate all her comments and suggestions in improving the documentation of this thesis.

More importantly, I acknowledge the Department of Science and Technology - Science Education Institute (DOST-SEI) through its Engineering Research and Development for Technology (ERDT) program. I am indebted to them for their support and for funding this research.

Finally, thank you very much to my parents, sisters, and family for their unfailing support and continuous encouragement in my studies.

Abstract

Emotion or affect is known to play vital roles in rational and intelligent behavior, such as cognition and decision-making. Detecting or recognizing affect can be done by analyzing physiological data or a combination of various physiological data. The current work presents a study on brainwaves or EEG signals, which are examples of physiological data, and their association to emotions while a person is reading literary fiction, an unexplored domain. EEG data from 32 participants were collected while they were reading a short story. These EEG signals were collected with the use of an Emotiv Insight EEG headset, attached to the head of each participant while reading the story segments presented via the developed data collector tool. After which, features were extracted and different datasets were built according to sex, reading preference, and reading frequency profiles. Decision Trees were used to establish baseline performance results, and these were able to classify the Hourglass of Emotion model and Emotions of Literary Response models. Support Vector Machines and Multilayer Perceptrons were trained on the same datasets to see if there is an increase in performance. Results show that they indeed yielded better performance results than DT, however, only by a small degree. Principal Component Analysis was used as an approach for feature selection, and results show comparable performance as opposed to using the base feature set of all EEG features with an averaged ±5 margin of error.

Keywords: affective computing, affect recognition, EEG, digital signal processing, machine learning

Contents

List of Figures

List of Tables

Listings

Chapter 1

Research Description

This chapter introduces the research undertaken in the field of Affective Computing. It is divided into four sections which discuss the research problem, the research objectives, the scope and limitations, as well as the significance of the study.

1.1 Overview of the Current State of Technology

Although emotion theorists still do not have a clear definition as to what emotions are (Kleinginna Jr. & Kleinginna, 1981), many studies suggest that *emotions* or *affect* play vital roles in rational and intelligent behavior such as cognition and decision making (Bechara, Damasio, & Damasio, 2000; Schwarz, 2000). Emotions interplay with rational thinking in ways that are imperceptible but important for intelligent functioning. Humans already observe emotional intelligence (EI), which is a subset of social intelligence that involves the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions (Salovey & Mayer, 1990). Ciarrochi et al. (2000) present a critical evaluation of the importance of EI versus the intelligence quotient (IQ). They conclude that while a number of people claim that EI is more important than IQ, there is little scientific evidence to support the said claim. However, using the Multi-factor Emotional Intelligence Scale (MEIS), they were able to justify some of the claims of EI like being an important determinant of success in relationships and careers. Note that the same stimulus creates different emotions in different individuals, and the same individual may express different emotions in response to the same stimulus, at different times.

With regard to affective computing (which relates to, arises from, or influences emotions), it is simply about causing a computer to be capable of emotional intelligence. Imagine a computer that is aware of a person's psychological well-being, that is his mental or emotional state. Theoretically, its responses would adjust accordingly depending on the person's current mental state. Thus, affective interactions with computers can easily and immediately give direct feedback as opposed to human interactions (Picard, 1997). However, before computers can give feedback, it has to detect a person's psychological state first. This is where emotion recognition comes in.

Humans usually perceive emotions by facial or vocal expressions. However, there are other physiological cues such as knowing that a person is nervous because their hand is clammy when held (relating to skin conductance) or determining that a person is excited when their pulse is felt (relating to heart rate). Hence, people naturally read many physiological signals of emotions (Picard, 2000). Translating this to computers, emotion recognition involves pattern recognition of physiological data. This proves to be a challenge because, again, there is no clear definition for what emotions are, and there is still the issue of whether there are distinct physiological patterns that are associated with each emotion (Cacioppo & Tassinary, 1990). Regardless of the difficulty, there have been various studies in emotion detection and recognition for various domains and motivations. Such detection is possible through analysis of facial and/or vocal expressions (Zeng et al., 2006; Zeng, Pantic, Roisman, & Huang, 2007), a combination of physiological data (e.g., heart rate, skin conductance, muscle tension) (Picard, Vyzas, & Healey, 2001), or even something as unusual as mouse-click behavior (J. Azcarraga & Suarez, 2012).

Electroencephalography (EEG) is the recording of the brain's electrical activity and it is one way to look into brain functions in real time (Rossetti & Laureys, 2015). The brain's electrical activity or brainwaves is another example of physiological data. EEG is a noninvasive procedure that can be recorded digitally with commercial portable devices such as Emotiv EPOC/EPOC+[1], Muse[2], or iBrain[3]. In the field of medicine, EEG evaluation plays critical roles in accurate patient diagnosis, seizure detection, evaluation of patients with consciousness disorders, and even sleeping disorders. Whereas in the field of affective computing, EEG serves as a standalone or an additional modality for affect detection and recognition while

[1]Emotiv, `https://emotiv.com/`
[2]InteraXon, `http://www.choosemuse.com/`
[3]Neurovigil, `http://www.neurovigil.com/`

doing an activity, such as the study by J. Azcarraga and Suarez (2012) where they used EEG coupled with mouse-click behavior to predict academic emotions (confidence, excitement, frustration, and interest) of students while solving varying difficulty levels of math problems.

Reading literary texts or fiction is not simply an activity but rather an experience that is never the same from one reading to the next (Tompkins, 1980). It is a pleasurable activity when the reader's imagination is engaged in an active and creative way (Iser, 1972). Reading non-fiction differs from reading fiction. Usually, when one reads non-fiction such as academic texts or news articles, the goal is to be informed. Whereas when one reads fiction such as novels, short stories, or poems, the goal is to be entertained and moved (experience a variety of emotions) (Mar, Oatley, Djikic, & Mullin, 2011). Reader-response criticism (a school of literary theory that focuses on the *reader*, the *reading process*, and *response*, rather than the literary text itself) is said to have started with I.A. Richard's discussion of emotional response (Tompkins, 1980). Emotions arising from reading the literary text are called *narrative emotions* (Mar et al., 2011).

There have been studies establishing the association of brainwave patterns and emotions while being engaged in an activity. Yazdani et al. (2012) made use of EEG and various physiological data to classify the valance, arousal, and preference of the participants on music videos. Nie et al. (2011) attempted to find the relationship of EEG and human emotions while watching movies. Lastly, Lin et al. (2010) applied machine learning algorithms to categorize EEG dynamics according to emotional states while listening to music. There have also been empirical works that have established the relation between reading and emotions or emotional response in areas of culture, media, and arts. Cupchik et al. (1998) showed how different literary texts elicit different emotional responses (emotions of identification vs. remembered emotions). Miall and Kuiken (1994) tested how stylistic variations in the literary text affects the response of people in terms of reading time, strikingness, and affect. However, there is no current work that has studied brainwave patterns and their association to affect while a person is reading literary fiction.

1.2 Research Objectives

1.2.1 General Objective

To build an affect model that maps the EEG signals collected from readers (while they are reading stories) to specific emotions.

1.2.2 Specific Objectives

1. To review the approaches, methodologies, and experiments of existing affect detection or recognition studies that use EEG data;
2. To identify different emotions that can be elicited from the readers as they read the stories;
3. To determine which elements of a story affect the reader's emotional state;
4. To build a corpus of EEG signals;
5. To implement machine learning algorithms for classifying the emotion based on the EEG signals; and
6. To define evaluation metrics for assessing the performance of the model.

1.3 Scope and Limitations of the Research

The research will focus on EEG data as its modality for emotion detection and recognition. A review of existing affect detection or recognition studies that use EEG data is needed to determine what the approaches and techniques are in tackling this area of research. It includes the review of how the data are collected and prepared, how tests and experiments are conducted, what machine learning algorithms and features are used to train their models. Aside from emotion detection of recognition studies, a review of researches establishing the association of affect and EEG is also needed.

To associate the brainwave patterns to specific emotions, the research must identify these specific emotions first. This will entail a review of different emotion models and determine which of them is appropriate for reader affect. The basis for *appropriateness* is dependent on the goal of the research and its participants. The academic emotion model is an example of a model that may not be aligned to the goal of this research. Ekman's model of six basic emotions (happiness, sadness, fear, anger, disgust, and surprise) may not be enough for the participants to fully articulate their emotional state.

The study will attempt to determine which element of the story triggered the reader to evoke a particular emotion. The story elements are limited to character traits and behavior, the reader's empathy to the character/s, the story plot or casual chain of story events, and lexical choices and sentence structure.

For the computer to be able to associate patterns of brainwaves to specific affect, a data corpus of EEG signals must be built. This will be collected from participants aged 18 and above of diverse demographics. (The rationale behind this is that the selected short stories have appeared in must-read lists for *high school* and *college students*. Having said that, a younger age group may have difficulty in understanding the material itself, and thus, may lead to negative emotions such as confusion and frustration. Refer to Section 3.4.3 for further explanation.) Data collection will follow the methodology of the studies mentioned in Section 2.2 while the presentation of the stimulus is similar to the experiments in the studies in Section 2.3. The chosen literary fiction are short stories because they can be read in a single sitting. The selected short stories are *The Veldt* by Ray Bradbury, *Man from the South* by Roald Dahl, and *The Fisherman and the Jinni* from *One Thousand and One Nights*. Since this proposed study involves the participation of human subjects, informed consent forms were given to them following the ethical research guidelines of the university (refer to Appendix A and B).

Weka (Hall et al., 2009) and RapidMiner (Hofmann & Klinkenberg, 2013) are examples of some of the tools being considered to perform the machine learning tasks. Both are open source tools for data science tasks.

The usual metrics for supervised machine learning tasks, which are precision, recall, accuracy, and f-measure, will be used to evaluate the performance of the affect model. Unsupervised machine learning metrics will still be determined.

1.4 Significance of the Research

Affective computing is a field that is a combination of various existing technologies and computer science concepts like wearable computing, digital signal processing, computer vision, pattern recognition, machine learning, and human-computer interaction (HCI). This study can contribute to the feasibility and application of these technologies and concepts on an unexplored application domain. Furthermore, this study provides a basis of the brainwave activity of people's emotions while they are reading, compared to their self-reported emotions. The methodology on data collection, data preparation, preprocessing, and classification will also be useful to future studies relating to EEG.

The results of this study can provide more understanding of emotions, which can be further utilized by affect-aware systems such as intelligent tutoring systems (ITS) (Anderson, Boyle, & Reiser, 1985) or embodied conversational agents (ECA) (Cassell, 2000). These kinds of systems replicate face-to-face human interactions with, for example, a digital avatar. Since these systems serve as stand-ins for humans, it should at least on a certain level, feel like having a natural interaction with a human, or teacher/tutor in the case of an ITS. For example, we have an ITS/ECA that helps a child practice reading. In reality, effective teachers gauge certain cues from their students and respond accordingly. If the system detects distress or frustration from the child based on his or her brainwaves, it can react appropriately by suggesting to have a break. If the example system is more advanced, then it could determine where the distress or frustration is coming from, for example, if the child's cause of frustration is the difficulty of understanding the words, then the system can replace those words with easier ones.

Aside from the field of computer science, the contributions of this work may be informative to affective science and psychology as well. Various iterations of experiments can be done depending on research objectives such as in-depth comparative studies (per demographic profile, e.g. per age group, per gender) of affective states while reading. Emotion detection can also be an indicator for behavior prediction, most especially when monitoring certain emotional states that could lead to potentially harmful or dangerous behaviors.

From a business perspective, the contributions of this work may also be of use to product and market analyses. Given a product description, commercial script, or narrative of the same nature, the findings of this research could give insights to the public's or audience's response to the narrative. By knowing their target audience's likes or dislikes, companies can efficiently create more effective products or services. On the other side of this perspective, that is, the target audience of the business companies, knowing the emotions or the response of other buyers can provide them more information on the product. This is why some people check the rating and reviews of a items or services before actually buying. However, not everyone rates and reviews items or services after they purchase it. If they do, it is not as detailed as some would have preferred. So if there is some way to refine the rating and reviewing system to be automatic or semi-automatic and to present the detected emotions in an intuitive and objective manner, then the consumers could make informed decisions and know that the product or service is truly worth their money. Another example can be considering emotions as tags for content media retrieval such as looking for book recommendations that make one feel happy.

1.5 Research Methodology

This section lists and discusses the phases and activities that are performed to accomplish the research project. The phases occur in a sequential manner unless otherwise stated. They are also revisited when new limitations are discovered.

1.5.1 Concept Formulation and Literature Review

This phase was concerned with the formulation of the thesis, particularly the identification of the research problem and its objectives, as well as defining its scope and limitations. A review of related literature on emotions, brainwaves, and reading fiction was done to understand how these concepts correlate with each other. The methodologies and approaches presented in some of these studies were considered to see if they are applicable to this research. The tools that are used are also learned in this phase in addition to canvassing and procuring the needed equipment. This phase is important because it builds the background knowledge needed before proceeding to the next phases.

1.5.2 Development of the Data Collector Tool

In this phase, the tool to be used for data collection was developed. Prior to the development, the pre-selected short stories were manually split into segments while still retaining meaning and coherence per segment (see Appendix C for a sample of the short story segments used). The data collector tool was based on the system used by J. Azcarraga and Suarez (2012). The software tool output of this phase is relevant to the next phase.

1.5.3 Data Collection and Dataset Building

The data for training and evaluation of the affect model were collected in this phase. At least 30 participants have read the segmented short stories while an EEG sensor was attached to them. Their EEG recordings and self-assessments were gathered via the data collector tool described in the previous section. This phase may took up most of the duration of the research because each session per participant per story lasted up to an hour and a half. Data preparation and preprocessing were also involved here.

1.5.4 Training and Evaluation

This phase is involved with the application of the to-be-determined machine learning techniques as well as feature extraction and selection. Experiments include various combinations of machine learning techniques and feature sets applied to the data sets. The results of each model will be evaluated according to the metrics set in Section 1.3. This phase will also occur more than once as more data are added and new machine learning algorithms are discovered.

1.5.5 Documentation

The documentation phase is necessary all throughout the duration of the research. Taking down notes on the observations and findings during the other research activities and review of related literature, writing related technical documents, and the research paper itself fall under this phase.

1.5.6 Calendar of Activities

Table 1.1 shows the Gantt chart of the activities. Each asterisk represents approximately 1 week's worth of activity.

Table 1.1: Timetable of activities.

Activities	Jan (2016)	Feb	Mar	Apr	May	Jun	Jul	Aug
Concept Formulation and Review of Related Literature	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆__				
Development of the Data Collector Tool			__⋆⋆	⋆⋆__	⋆⋆__			
Data Collection and Dataset Building					__⋆⋆	⋆⋆⋆⋆	⋆⋆__	
Training and Evaluation							⋆⋆⋆⋆	⋆⋆⋆_
Documentation	__⋆⋆	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆__	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆_

Activities	Sep	Oct	Nov	Dec	Jan (2017)	Feb	Mar	Apr
Data Collection and Dataset Building			__⋆⋆	⋆___	⋆⋆⋆⋆	⋆⋆⋆⋆		
Training and Evaluation	_⋆⋆⋆	⋆⋆⋆⋆			⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆__	
Documentation	_⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆___	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆___

Chapter 2

Review of Related Literature

This chapter discusses the related studies to the research being undertaken. It presents how emotions are correlated to computing, brainwaves, and reading fiction.

2.1 Emotions and Computing

The field of Affective Computing was pioneered by Picard in 1995 when she presented her ideas on the feasibility of affect-aware computers and their possible implications and applications (which eventually was published into a book in 1997). *Affective computing*, defined as computing that relates to, arises from, or influences emotions, is an aggregation of already existing technologies and computer science concepts such as wearable computing, digital signal processing, computer vision, pattern recognition, machine learning, and human-computer interaction. However, before Picard presented her ideas and theoretical applications on the field, she first justified as to why emotions should be imbued to computers (Picard, 1997).

Humans usually read and express emotions through face and voice. However, there are other physiological cues such as feeling clammy hands (relating to skin conductance) to mean that a person is nervous or having a fast heart rate to mean that a person is excited. Hence, people naturally read many physiological signals of emotions (Picard, 2000). Translating this to computers, emotion recognition involves pattern recognition of physiological cues. This proves to be a challenge because there is still the issue of whether there are distinct physiological patterns that are associated with each emotion (Cacioppo & Tassinary, 1990). A fast heart rate may imply that a person is either nervous or excited. So there must be other indicators that differentiate the two.

For an affect recognition system to be accurate, it needs to combine multiple kinds of physiological signals from the user as well as information about the user's context, situation, goals, and preferences (Picard, 1997). Picard (2000) attempted to classify a single person's physiological patterns of over a period of several weeks into eight distinct emotions. They used Clynes and Menuhin's set of emotions for their classification. The emotion set was chosen not because it is the best for HCI but rather because it has an already existing method for eliciting emotions (via *Sentograph*, also developed by Clynes and Menuhin). This set includes *no emotion (neutral)*, *anger*, *hate*, *grief*, *platonic love*, *romantic love*, *joy*, *reverence* with high/low arousal and positive/negative valence taken into consideration. Results show an 81% recognition accuracy on all eight classes of emotions.

Picard (2003) wrote a paper addressing most of the concerns and challenges in affective computing. One of the issues she pointed out was that her experiment (Picard, 2000) was forced to eight choices only. However, naturally, we know that each emotion has its own varying degrees of intensity, thus, it is not only limited to eight. Picard gives an analogy to treat emotions like the weather. To quantify the weather, there are various sensors that measure temperature, pressure, humidity, etc. With emotions, it is building the sensors for the equivalents of temperature, pressure, etc. Like weather, the prediction may not be perfectly reliable, however, at least having a little foresight is useful.

Another issue Picard has addressed is with gathering *good* affect data. The first concern here are the sensors. Sensors are typically expensive, invasive, and/or obtrusive. There can be difficulty in gathering accurate physiological data due to technical factors such as where the sensors are applied or how much gel is used for the electrode. However, there are now advances in wearable technology that seamlessly integrate these sensors to what humans usually wear such as the smart watch. Apart from the sensors, another challenge she presented regarding affect data is the ground truth to compare the classifications against. An outsider cannot objectively say that the subject is feeling this certain emotion. Only the subject can know what emotional state he is in. If the outsider explicitly asks what the subject is feeling, that may also compromise the ground truth depending on how comfortable the subject is with expressing his feelings, how aware he is to his feelings, or if the subject becomes irritated with the constant asking of how he feels. Thus, Picard set the following factors on obtaining good affect data as shown in Table 2.1. Underlined are the best conditions for gathering genuine affect data. The problem with it is that it's being opportunistic and quite impossible to obtain as it breaches some privacy and ethical issues (Picard, 2000).

Table 2.1: Five factors on obtaining *good* affect data.

Subject-elicited vs. event-elicited	Does the subject purposefully elicit emotion or is it elicited by a stimulus or situation outside the subject's efforts?
Lab setting vs. real-world	Is the subject in a lab or in a special room that is not their usual environment?
Expression vs. feeling	Is the emphasis on external expression or on internal feeling?
Open-recording vs. hidden-recording	Does the subject know that anything is being recorded?
Emotion-purpose vs. other-purpose	Does the subject know that the experiment is about emotion?

2.2 Emotions and Brainwaves

J. Azcarraga and Suarez (2012) made use of EEG data coupled with mouse-click behavior to classify the academic emotions of the subjects. They limited the academic emotions to *confidence*, *excitement*, *frustration*, and *interest* with the intensity taken into consideration. They used an Emotiv EPOC sensor, which has 14 channels, to record the EEG data and took note of the number of mouse clicks, of each click's duration, and of each mouse movement. They first established the baseline EEG recording *resting state* by allowing the participant to relax for 3 minutes. The subjects were tasked to solve four algebra equations of varying levels of difficulty for 15 minutes. A self-assessment window will pop up every 2 minutes and the participant will tag each academic emotion with intensity from 1 to 100 using their own observation module.

They identified 17 features from the prepared data sets (14 EEG channels and 3 mouse behavior information) and used Multilayer Perceptrons (MLP) and Support Vector Machines (SVM), with a 10-fold cross-validation, as classification algorithms for their success in general approximation. Their metrics used for assessing performance were precision, accuracy, and f-measure. Issues encountered in data preparation were with cleaning the data and balancing the dataset. They had to manually synchronize the EEG recordings to the self-assessments. In addition, noise artifacts in the EEG which are not concerned with affective states were removed (i.e. blinking, hand movements), only 16 out of 25 data from the students were found useful given the conditions they set for balancing the data set. Six different datasets were prepared based on the percentage of feature outliers. A feature is considered an outlier if it exceeds 1 standard deviation away from the mean of that particular feature of that particular subject. Their results show that as the number of outlier features increases, prediction accuracy also increases. The combination of brainwave and mouse behavior data yielded better results than brainwave or mouse behavior alone. Note, however, that brainwave data performed better than mouse behavior data. On all cases of their experiments, MLP algorithm performed better than SVM. Table 2.2 shows a summary of their prediction accuracy results.

Table 2.2: Summary of MLP accuracy results of Azcarraga and Suarez (2012).

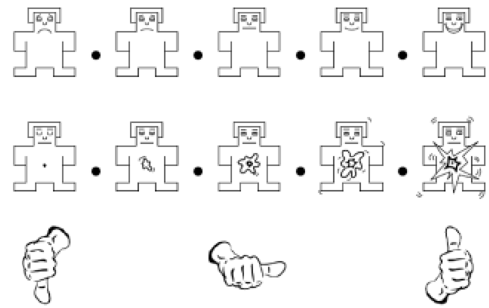
Dataset	**Brainwaves**	**Mouse**	**Brainwaves + Mouse**
0	54.66	32.26	61.04
10	63.74	38.9	69.8
25	75.27	45.11	78.58
33	74.92	45.46	80.69
50	83.65	43.85	88.56
60	88.33	48.79	92.27

An extension of their work shows that by implementing gender-specific classifiers, the classification performances significantly improves from a baseline accuracy of 58.70% to 65.33% for the male participants and 70.50% for the female participants (A. Azcarraga, Talavera, & Azcarraga, 2016).

Yazdani et al. (2012) combined EEG signals and various physiological signals to classify the emotional state of a person while they are watching music videos. Based on the valence-arousal scale proposed by Russell (1980), they classified the EEG and physiological signals into high/low arousal and positive/negative valence. They also added an additional like/dislike classification to see the preferences of their participants. They recorded the EEG and physiological signals using a

BioSemi ActiveTwo system[4]. The EEG recording was composed of 32 channels and the accompanying physiological signals were the following: Galvanic Skin Response (GSR), respiration, skin temperature, blood volume pulse by plethysmograph, EMGs of of zygomaticus major and trapezius muscles (2 channels each), and a 4-channel electrooculogram (EOG). Figure 2.1a shows a participant with the EEG and physiological sensors attached to him.

(a) A participant with the EEG and physiological sensors attached to him.



(b) Images used for self-assessment of (from top to bottom) arousal, valence, and preference.

Figure 2.1: Experimental set-up of Yazdani et al. (2012)

Their experiment started with having the participant relax for 2 minutes; a fixation cross was displayed on the screen. This recording served as their baseline. Twenty music video clips were presented at random in separate runs. Each run consists of a 5-second baseline recording (display of the fixation cross), 2-minute display of the music video, and the participant's self-assessment for arousal, valence, and preference. Figure 2.1b shows the symbols they used for the self-assessment, which was based on the Self-Assessment Manikins (SAM) by Morris (1995).

Table 2.3: Summary of accuracy results of Yazdani et al. (2012).

	Arousal	**Valence**	**Preference**
EEG Data	69.58%	73.66%	70.25%
Physiological Data	55.7%	54%	66.6%

Preprocessing includes filtering, downsampling, and artifact removal to ensure artifact-free signals for all channels. They introduced a novel approach for feature extraction, which was the use of the relative wavelet energies (RWE) of each channel and the RWE of symmetrical channel pairs as extracted features. SVM

[4]BioSemi, `http://biosemi.com/`

classifier with radial basis function kernels and a 20-fold cross validation were used in the classification. Table 2.3 shows the averaged accuracy of their results. Despite the higher accuracy rate for EEG data, their performance results are incomparable because they used different window lengths.

An additional experiment they explored was the feasibility of having a general-purpose affect recognition system. To test this, they trained a classifier on the physiological data of the participants and employed a leave-one-participant-out cross-validation. Their conclusion was that accuracy varies among participants, indicating that classification is participant-dependent. Further investigation needs to be done to build a general classifier for all participants.

Nie et al. (2011) used EEG signals to classify *positive* and *negative* emotions while watching movie clips. Their rationale for choosing only two classes is because emotions come in mixed forms, hence, if the subject could not properly distinguish his emotions, he could at least identify whether it is a positive or a negative emotion. They used a 62-channel electrode cap to record the EEG with a 32-bit level at a sampling rate of $1000Hz$.

Their experimental protocol is as follows: For each session, one movie clip is played. Five seconds was allocated to indicate the start of the session followed by a 4-minute showing of the movie clip. After the viewing, there was 45 seconds of self-assessment and then a 15-second rest before proceeding to the next session. They did not indicate whether they have a baseline recording of the participants before the experiment. For self-assessment, they used the SAM by Bradley and Lang (1994).

Data preparation and cleaning include the removal of the artifacts not related to the emotional states and the extraction of the five frequency bands. To further remove noise, they smoothed the features by applying a linear dynamic system (LDS) approach for each band. For feature selection, they calculated the correlation coefficients between features and labels for each channel and band on the training set. They ranked them and obtained the top 50 and top 100 features. This is done to find subject-independent features. For classification, they used SVM with linear kernel to train the data for each band. This was accomplished with a 7:3 ratio of training and testing data. Lastly, they employed another SVM for all of them. This was checked with a 10-fold cross-validation. Table 2.4 shows the summary of accuracy results of their classification.

Lin et al. (2010) applied machine learning algorithms to associate EEG patterns

Table 2.4: Summary of accuracy results of Nie et al. (2011).

Subject	Delta	Theta	Alpha	Beta	Gamma	All
1	68.23	66.21	92.38	82.83	100	99.63
2	57.58	87.09	86.30	72.86	73.65	81.95
3	91.30	77.74	84.06	85.82	91.20	87.16
4	38.80	74.28	65.63	100	88.47	91.13
5	45.81	74.19	90.48	77.42	62.26	82.90
6	71.24	84.65	94.35	82.71	89.82	82.39
Average	62.16	77.36	85.53	83.61	84.23	87.53

to self-reported emotional states while listening to music. The goal of their research was to identify emotion-specific EEG features as well as explore the efficiency of two classifiers, MLP and SVM. They used a 32-channel EEG module by Neuroscan, Inc. to record the EEG signals. Their emotion model is based on the four emotional states following the 2D arousal valence-arousal scale by Russell (1980): *joy* (positive valence, high arousal), *anger* (negative valence, high arousal), *sadness* (negative valence, low arousal), and *pleasure* (positive valence, low arousal). These emotional states were recorded with the use of FEELTRACE by Cowie et al. (2000).

From the recorded data, they have identified four feature sets. First is the individual spectral power from the 30 scalp electrodes, namely Fp1, Fp2, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, O1, Oz, and O2. The first feature set was named PSD30 for *power spectrum density for all 30 channels*. The next two feature sets involve the symmetric electrode pairs, namely Fp1-Fp2, F7-F8, F3-F4, FT7-FT8, FC3-FC4, T7-T8, P7-P8, C3-C4, TP7-TP8, CP3-CP4, P3-P4, and O1-O2. Asymmetry indexes were calculated in two ways. DASM12 (*differential asymmetry of 12 electrode pairs*) was computed by power subtraction (i.e. power of C3 - power of C4) whereas RASM12 (*rational asymmetry of 12 electrode pairs*) was computed by power division (i.e. power of C3 / power of C4). The last feature set is the PSD24 which is the *power spectrum density of 24 channels*. The PSD24 is a subset of PSD30 wherein the channels along the midline were not included (Fz, FCz, Cz, CPz, Pz, and Oz).

MLP and SVM classifiers were trained and evaluated with a 10-fold cross validation. Table 2.5 and Table 2.6 show the results of their classification. DASM12 gave the best performance results among the frequency bands and general classifier of the frequency bands for SVM and MLP, with SVM having a marginally higher result. Since DASM12 gave the best results, they applied an F-score index with a leave-N-feature-out scheme to rank the features across all frequency bands. They obtained the top 30 subject-independent features and applied SVM, obtaining a result of approximately 74% accuracy.

Table 2.5: Summary of MLP accuracy results of Lin et al. (2010).

Feature Type	**EEG Frequency Band**					
	Delta	**Theta**	**Alpha**	**Beta**	**Gamma**	**All**
DASM12	**63.93**	**63.67**	**64.07**	55.71	53.24	**81.52**
RASM12	48.54	50.69	55.40	48.21	44.82	65.33
PSD24	49.20	52.10	57.79	53.20	54.46	75.66
PSD30	52.12	55.61	61.89	**57.92**	**58.04**	79.54

Table 2.6: Summary of SVM accuracy results of Lin et al. (2010).

Feature Type	**EEG Frequency Band**					
	Delta	**Theta**	**Alpha**	**Beta**	**Gamma**	**All**
DASM12	**69.91**	**68.27**	**66.94**	**58.83**	57.35	**82.29**
RASM12	50.91	51.39	56.95	50.29	47.61	65.81
PSD24	51.02	53.27	54.61	55.42	56.80	69.54
PSD30	53.38	55.61	56.64	58.71	**59.54**	71.15

Following Picard's five factors on obtaining good affect data, it is observed that the experimental setup of J. Azcarraga and Suarez (2012), Yazdani et al. (2012), Nie et al. (2011), and Lin et al. (2010) is *event-elicited*, conducted in a *lab setting*, concerned with *feeling*, is *open-recorded*, and is *emotion-purpose*. Table 2.7 shows a summary of EEG-based affect recognition studies while doing an activity.

Table 2.7: Summary of EEG-based affect recognition studies.

Reference	EEG Recording	Emotion Model	Self-Assessment Scheme	Classification Algorithm	Features	Results
Azcarraga & Suarez (2012)	14-channel Emotiv EPOC sensor	Confidence, interest, excitement, frustration	Own observation tool	MLP (10-fold cross validation)	Power spectrum density of 14 channels	54-88%
Yazdani et al. (2012)	32-channel EEG electrodes recorded via Biosemi ActiveTwo system at 512Hz sampling rate	Arousal, valence, like/dislike	Morris's SAM for arousal and valence; thumbs up and down symbols for like/dislike	SVM with radial basis function kernel (20-fold cross validation)	Relative wavelet energies of each electrode together with RWE of symmetrical electrode pairs	69.58% (valence), 73.66% (arousal), 70.25% (like/dislike)
Nie et al. (2011)	62-channel electrode cap recorded with 32-bit level at 1000Hz sampling rate	Positive, negative	Bradley's SAM	Linear SVM (7:3 ratio)	Top 100 and top 50 subject independent features obtained through the reduction of the original features by correlation coefficients	89.22% (Top 100), 84.94% (Top 50)
Lin et al. (2010)	32-channel module by Neuroscan, Inc. at 500Hz sampling rate	Joy, anger, sadness, pleasure	FEELTRACE	MLP, SVM with radial basis function kernel (10-fold cross validation)	Power spectrum density of 30 channels, differential asymmetry of 12 electrode pairs, rational asymmetry of 12 electrode pairs, power spectrum density of 24 channels	65-81% (MLP), 65-82% (SVM)

2.3 Emotions and Reading Fiction

Empirical works have established the relation between reading and emotions or emotional response in areas of culture, media, and arts.

Miall and Kuiken (1994) observed how foregrounding, or stylistic variation in the text, influences the affect of the reader. *Foregrounding* refers to a range of stylistic variations that occur in literature at a phonetic level (alliteration, rhyme), grammatical level (inversion, ellipsis), or the semantic level (metaphor, irony). They noted that readers are more likely to report the phrases that "struck them" or "caught their eye" when presented with the original text as opposed to one written in neutral terms. Hence, this means that foregrounding strikes interest.

Their experiment consists of three literary stories, each short enough for an hour-long, single reading session. Each story contains a variety of foregrounded features and they were divided roughly into equal segments using phrase and sentence divisions while still retaining meaningful units as far as possible (approximately 77-86 segments per story). Their test subjects include both those with high and low literary competencies. For each participant, there were two readings. In the first reading, story segments are presented one at a time, with reference to the previous segment, and the current segment is highlighted. By the second reading, story segments are presented again one at a time, with reference to the two previous and succeeding segments. Then the participants are asked to rate the current segment based on strikingness (1 to 5 scale) and affect (no feeling to strong feeling).

Their findings show that despite the test subjects having different interests and literary competencies, the affect of the reader to the story is independent. It may imply that only those with higher literary competencies draw more meaning on the text. Another finding they discovered is that the more foregrounded the segment, the more emotions it may elicit from the segment. In relation to the research project, this set-up seems like an appropriate methodology to adopt in the data gathering phase.

Cupchik et al. (1998) conducted a similar experiment but instead of focusing on what aspects of the texts elicited the emotional response, it is concerned with what type of emotional response the reader evoked. Their selected short stories were divided into four equal segments. Two of the stories have unifying emotional themes, while the other two were descriptively dense. The participants were instructed to either be a spectator and feel sympathy for the main character or to imagine oneself as the main character. After reading each segment, they were to rate the experience on 11-point scale measuring pleasure, intensity, and tension. Aside from rating, they were asked if they experienced *fresh emotions* (emotions of sympathy, identification, or empathy) or *emotional memories* (relived or remembered emotions).

Chapter 3

Theoretical Framework

This chapter discusses the theories and concepts needed to build the reader affect model.

3.1 Emotions

3.1.1 Importance of Emotions

Kleinginna Jr. and Kleinginna (1981) compiled a categorized list of emotion definitions. This list is comprised of 92 definitions and 9 skeptical statements from various emotion theorists and sources of emotion literature. Despite the plethora of definitions, the relevance of emotions in a person's day-to-day activities has been proven by various studies.

In the field of neurology, Bechara et al. (2000) posited that the decision-making process is consciously or unconsciously influenced by marker signals arising from bioregulatory processes. These bioregulatory processes include those that are expressed in emotions. Decision-making is subserved in the orbitofrontal cortex of the brain as well as other cortical and subcortical areas. They have shown that lesions in the ventromedial (VM) prefontal cortex, a region of the brain involved in emotional response, seriously impairs the efficiency of decision-making. The same holds true for substance abusers and people with psychiatric disorders.

In social psychology, Schwarz (2000) pointed out the interplay of emotions, cognition, and decision-making. He identified scenarios and studies showing the role of emotions in various stages of decision-making. Before actually making a decision, a person's emotional state serves as a bias towards one of the options. Similarly, after a decision is made, there is an emotional reaction (e.g. regret, disappointment). Aside from the emotions elicited before and after making a decision, anticipated affect and memories of past affect also come into play. Anticipated affect simply means that a person is motivated to choose a certain option because he wants to avoid, for example, disappointment (anticipated affect). The same goes for memories of past affect, wherein instead of anticipating a certain emotion, the person already knows that a particular emotion is a likely outcome due to having already experienced it before.

A body of intelligence terminologies concerning affect was coined by Salovey and Mayer (1990). They defined *emotional intelligence* (EI) as the subset of social intelligence that involves the ability to monitor one's own and others' feelings and emotions, to discriminate among them, and to use this information to guide one's thinking and actions. Ciarrochi et al. (2000) presented a critical evaluation of the importance of EI versus the intelligence quotient (IQ). They concluded that while a number of people claim that EI is more important than IQ, there is little scientific evidence to support said claim. However, using the Multi-factor Emotional Intelligence Scale (MEIS), they were able to justify some of the claims of EI like being an important determinant of success in relationships and careers.

3.1.2 Emotion Models

A challenge for affect recognition is determining the appropriate emotion model to be used. The reason for this problem can be traced back to the unclear definition of emotions. With reference to the numerous research done in emotion theory, it has resulted to two general kinds of emotion models: the categorical models and the dimensional models.

Categorical models are those that define a number of discrete basic emotions. Ekman (1972) is an example of this model. He defined six basic emotions: *happiness*, *sadness*, *fear*, *anger*, *disgust*, and *surprise*. In contrast, *dimensional models* describe the components of emotions and are often represented as a two- or three-dimensional space where the emotions are presented as points in the coordinate space of these dimensions. The arousal-valence scale by Russell (1980) is an example of this model, as shown in Figure 3.1. The dimension *valence* provides information about the degree of pleasantness of the content and ranges from pleasant (positive) to unpleasant (negative), while the dimension of *arousal* represents the inner activation and ranges from energized to calm.

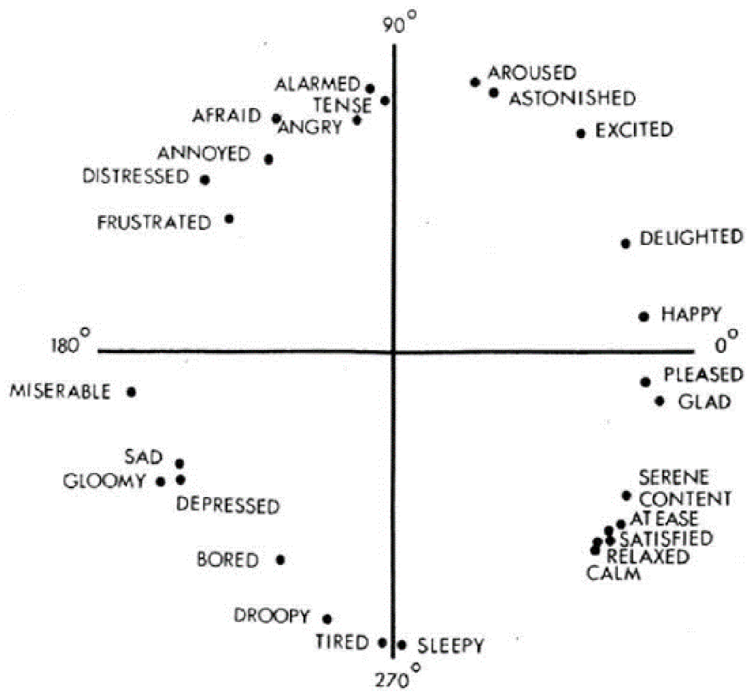


Figure 3.1: Dimensional emotion model by Russell (1980).

Given the difference in these two kinds of emotion models, Cambria et al. (2012) proposed a novel biologically inspired and psychologically motivated emotion categorization model they named as the *Hourglass of Emotion*. They described the model as one that is able to represent affective states both through its labels (categorical) and its four independent but co-occurring affective dimensions (dimensional). In this way, their model can potentially describe a full range of emotional experiences. For this reason, this emotion model was chosen for this research.

The Hourglass of Emotions (HoE) model, as shown in Figure 3.2, reinterprets Plutchik (2001) by organizing the primary emotions around four independent but related dimensions. These dimensions measure the user's level of amusement by interaction modalities (*Pleasantness*), of interest in interaction contents (*Attention*), of comfort with interaction dynamics (*Sensitivity*), and of confidence in interaction benefits (*Aptitude*). Each affective dimension is described by six *sentic levels* which is a measure of the strength of an emotion.

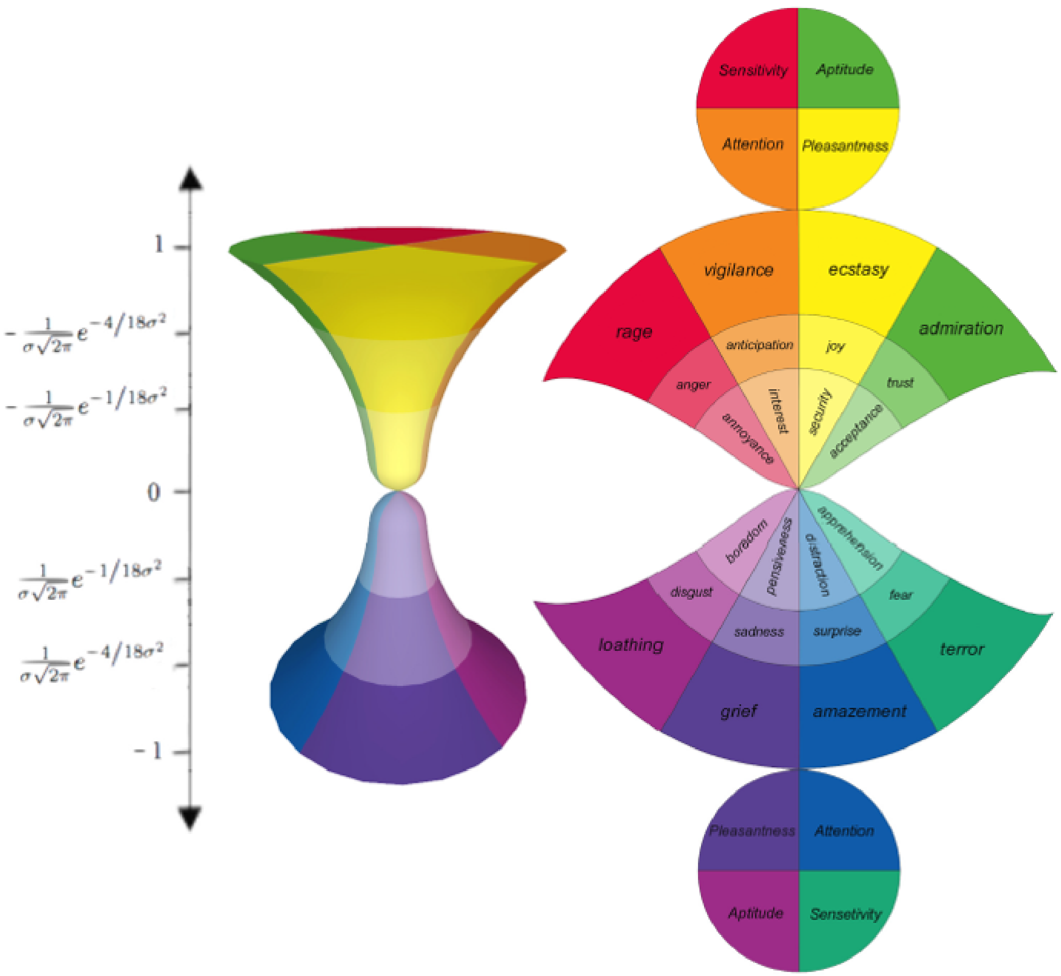


Figure 3.2: The Hourglass of Emotions model by Cambria et al. (2012).

3.2 Electroencephalography

Electroencephalography (EEG) is the recorded electrical activity generated by the brain (Rossetti & Laureys, 2015; NeuroSky, 2009). In the brain, millions of neurons generate small electric voltage fields. The aggregation of these electrical activities is what the EEG electrodes are able to detect and record. EEG is an effective means of neuro-imaging because it is a noninvasive and safe procedure which can record in milliseconds. *Event-related potentials* (ERP) are the EEG signals time-locked to a stimulus that the participant reacts to. The *baseline* are the EEG signals during the time before the stimulus is presented. In the case of this research, the brainwaves during the time of reading the short stories serve as the ERP.

EEG is generally described in terms of its frequency bands. Each band has a certain frequency range and relates to various brain states as shown in Table 3.1.

Table 3.1: EEG frequency band distribution and it corresponding mental state.

Brainwave Type	**Frequency Range**	**Mental States and Conditions**
Delta δ	0.1Hz to 3Hz	Deep, dreamless sleep, non-REM sleep, unconscious
Theta θ	4Hz to 7Hz	Intuitive, creative, recall, fantasy, imaginary, dream
Alpha α	8Hz to 12Hz	Relaxed, but not drowsy, tranquil, conscious
Low Beta β_1	12Hz to 15Hz	Formerly SMR, relaxed yet focused, integrated
Midrange Beta β_2	16Hz to 20Hz	Thinking, aware of self & surroundings
High Beta β_3	21Hz to 30Hz	Alertness, agitation
Gamma γ	30Hz to 100Hz	Motor functions, higher mental activity

3.3 Emotiv Insight EEG Headset

The Emotiv Insight[5] is a 5-channel (AF3, AF4, T7, T8, Pz), wireless EEG headset that records brainwaves. It is a commercial product marketed worldwide and is designed for everyday use. It uses a polymer sensor that is safe to use and offers great electrical conductivity with the convenience of a dry sensor. These sensors read the brainwave signals and then transmits these signals to a computer via Bluetooth. Figure 3.3a and 3.3b show what the device looks like, whereas Figure 3.3c shows a screen capture of the device's control panel when it is connected to a computer.

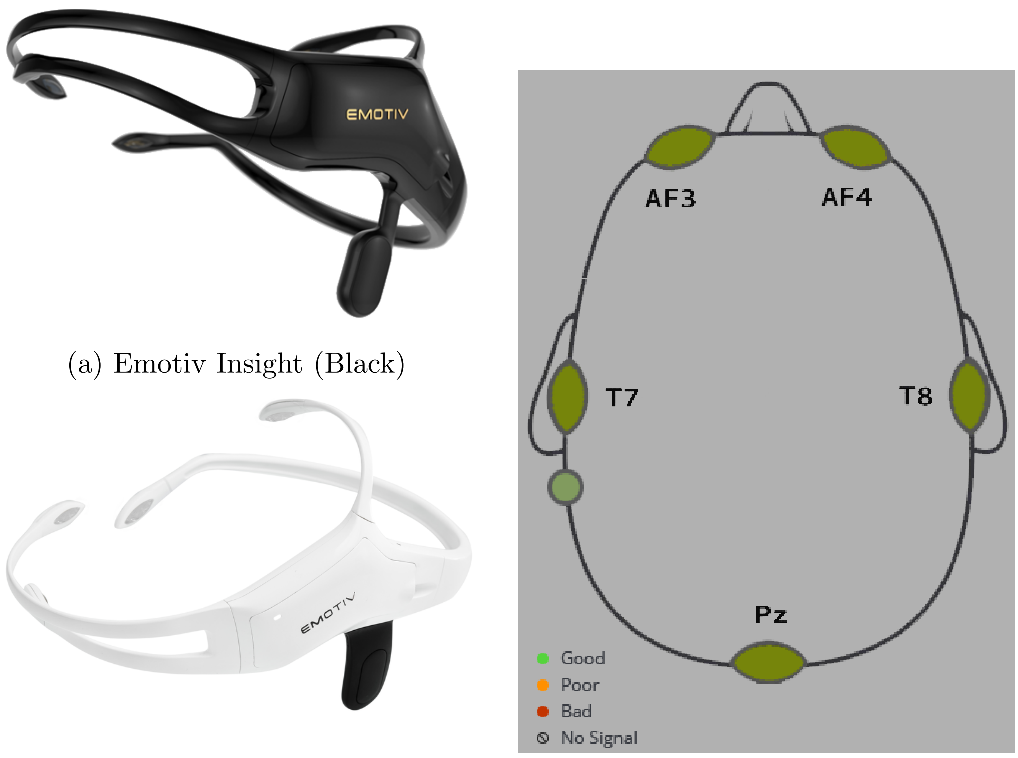
This device has global recognition for its personal usage such as assessing athletic performance, cognitive training, or health and well-being. The technology is also backed and trusted by the scientific, academic, engineering, and media communities and has been validated by many independent research papers[6].

3.4 Reading Fiction

The act of reading literary fiction is part of a broader aspect of human growth and development based on understanding one's own experiences and the social world (Freire & Slover, 1983). It is an experience that is never the same from one reading to the next (Tompkins, 1980). Mar et al. (2009) and Kidd and Castano (2013) proved that reading improves a person's empathy and *theory of mind*, defined as the ability to impute mental states (i.e. beliefs, intents, desires, pretending, knowledge) to the self and others and to understand that others have their own beliefs, desires, intentions, and perspectives that are different from one's own (Doherty, 2008). Vezzali et al. (2015) proved this when they showed how reading Harry Potter improves attitudes (reduces prejudices) towards out-group individuals (e.g. immigrants, homosexuals, refugees).

[5]Emotiv Insight, `http://emotiv.com/insight/`

[6]Independent researches that used Emotiv, `http://emotiv.com/the-science/`



(a) Emotiv Insight (Black)

(b) Emotiv Insight (White)

(c) Emotiv Xavier Control Panel

Figure 3.3: Emotiv Insight device and control panel.

3.4.1 Reader-Response Criticism

Reader-response criticism is a school of literary theory that focuses on the *reader*, the *reading process*, and *response*, rather than the literary text itself. It is said to have started with I.A. Richard's discussion of emotional response (Tompkins, 1980). Emotions are central to the experience of reading literary narrative fiction. A person's affect and mood influence, and are being influenced before, during, and after the actual reading. Mar et al. (2011) cited current empirical studies on emotions and narrative fiction at each stage of reading. As for this present research, it focuses on the emotions evoked *during* reading.

3.4.2 Emotions of Literary Response

Oatley (1995) presented a taxonomy of emotions in literary response (ELR). Emotions evoked when looking at a piece of fiction from a distance, evaluating its craft, style, and the like are called *aesthetic emotions*. This research focuses on the *narrative emotions*, which are emotions concerned with entering the narrative world. Narrative emotions are further subdivided into three kinds. *Sympathy* is when the reader is a witness to the scene but unable to affect the action in any way. *Emotion memories* is when the text makes the reader remember a past feeling, thus, emotion memories are not only recalled but also relived. *Emotions of identification* is when the reader takes on the characteristics and goals of the character. This taxonomy is further extended by Mar et al. (2011). It included *emotions of empathy*, which is closely related to emotions of sympathy and identification. Here, the reader does not identify with the character but rather empathizes with it. Emotion memories are split into relived emotions and remembered emotions. *Relived emotions* come from a recollection of past personal experiences whereas *remembered emotions* is a recalled emotion that does not fall under past personal experiences. This latest taxonomy is shown in Figure 3.4.

Miall and Kuiken (2002) described another classification of ELR. Basic emotions, such as anger or fear, cannot fully encapsulate the feelings evoked during reading. Hence, they posited that these feelings can be roughly sorted into four domains. There is no specific distinction among these four because the reader may experience them simultaneously. However, they argued that each feeling domain can be differentiated by certain structures and processes. *Evaluative feelings* are those concerned with the overall enjoyment, pleasure, or satisfaction of reading the text. They can be emotions in response to the other domains. *Narrative feelings* are

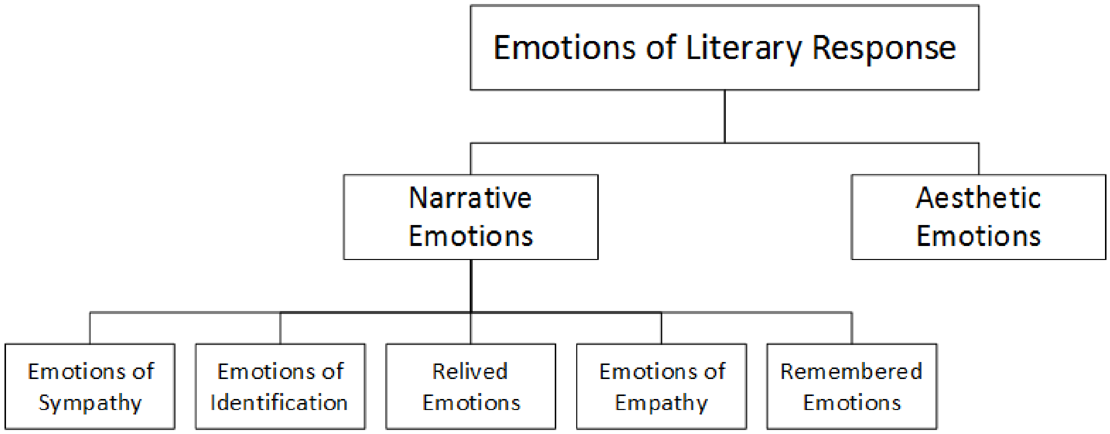


Figure 3.4: Taxonomy of Emotions of Literary Response.

evoked by events or characters in the fictional world of the text. Emotions of sympathy or empathy are in this domain. *Aesthetic feelings* are prompted by the formal (generic, narrative, or stylistic) components of a text. These reflect the heightened interest of the reader when they encounter a passage so striking that it holds their attention. Lastly, *self-modifying feelings* restructure the reader's understanding of the textual narrative, and simultaneously, the reader's sense of self. In relation to this research, it is only concerned with the first three. Notice that this classification coincides with the third specific objective of this research. Hence, this will be the chosen classification of emotions for literary response along with the notion that it also covers the other classification.

3.4.3 Age-Related Changes in Emotion Understanding and Reading Comprehension

Hannon and Daneman (2009) conducted an experiment to show the age-related changes in reading comprehension between young adults (people of ages 18-25) and older adults (people of ages 64-87) based on four components: text memory (tested memory for information explicitly presented in the paragraph), text inferencing (tested inferences about information that was implied in the paragraph), knowledge integration (required participants to access prior knowledge and integrate that knowledge with text information), and knowledge access (tested access to prior knowledge). Their results show that those four components are likely to decline with aging. However, they also concluded that overall reading comprehension ability remained the same regardless of the age of their participants. With regard to understanding and decoding emotions from written passages, Phillips et al. (2002) revealed that there is no difference between young adults (people of ages 20-40) and older adults (people of ages 60-80). In summary, younger and older adults more or less have similar reading comprehension abilities as well as understanding and decoding emotions from written passages.

3.5 Feature Modeling

3.5.1 Features

The signals from the AF3, AF4, T7, T8, and Pz channels are the raw EEG recordings obtained from the Emotiv device. These signals are initially in the time-domain. They are transformed into the frequency-domain using the Fourier transform, after which, the theta θ, alpha α, beta β, and gamma γ frequency bands are extracted. Four feature types were then computed for each frequency band. For each feature type, the mean, maximum value, and minimum value serve as the features. The following are the feature types extracted from the EEG:

Magnitude, computed using the formula in Equation 3.1, refers to the absolute value of the signal.

$$x_{mag}(n) = |x(n)| \tag{3.1}$$

Power Spectral Density (PSD) shows the strength of the signal as a function of frequency. In other words, it shows at which frequencies the signals are strong or weak. This was computed as the square of the signal in the frequency-domain as shown in Equation 3.2.

$$\begin{aligned} x_{psd}(n) &= |x(n)|^2 \\ &= x_{mag}(n)^2 \end{aligned} \tag{3.2}$$

Spectral power of the asymmetric electrode pairs (i.e. AF3-AF4 and T7-T8) are also extracted. *Differential Asymmetry* (DASM) was calculated using power subtraction, as shown in Equation 3.3. *Rational Asymmetry* (RASM), on the other hand, was calculated using power division, as shown in Equation 3.4. Note that C refers to the asymmetric electrode pairs.

$$C_{dasm}(n) = C_1(n)^2 - C_2(n)^2 \tag{3.3}$$

$$C_{rasm}(n) = C_1(n)^2 / C_2(n)^2 \tag{3.4}$$

3.5.2 Feature Selection

Principal Component Analysis (PCA) is a technique used to emphasize variation and bring out strong patterns in a dataset. It is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated features into a set of values of uncorrelated features called *principal components*. The number of principal components is less than or equal to the number of original attributes. This transformation is defined in such a way that the first principal component's variance is as high as possible (accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it should be orthogonal to (uncorrelated with) the preceding components.

3.6 Classification Methods

The goal of machine learning is generalizing properties from observed data onto new and unobserved data. The models being built in machine learning tasks impose structure that is not necessarily present in the data. Supervised machine learning deals with predicting class labels from given attributes or features whereas unsupervised learning is organizing data and finding structure in them (Welling, 2011). The research explored the use of MLP and SVM as they are the two classification methods used in the reviewed studies, as well as Decision Trees.

3.6.1 Decision Tree

Decision Trees (DT) are inverted tree-like graphs that represent a prediction model wherein each branch node corresponds to an input attribute or feature; and the leaf nodes are the predicted values. Examples of decision tree algorithms are ID3 and C4.5. These algorithms employ an information gain measure to designate the root and subsequent nodes.

Advantages of decision trees include, but are not limited to, (a) being easy and simple to interpret as it can be displayed graphically; (b) requiring little data preparation; (c) large datasets can be analyzed within a reasonable time as opposed to neural networks, for example, that typically require numerous iterations; and (d) can describe the model by the relevance of features. The important features are easily identified because they appear as the root node or the nodes near the root of the tree. These features can be used in feature selection experiments. The disadvantage of this method is that the results do not tend to be as accurate as the other classification models (James, Witten, Hastie, & Tibshirani, 2013).

3.6.2 Multilayer Perceptron

Multilayer Perceptrons (MLP) are a type of feed-forward artificial neural networks that are composed of several layers of smaller processing units known as neurons (Bishop, 1995). Neural networks are universal approximators. They can approximate any mathematical function, which makes them very flexible classifiers that can adapt to a variety of machine learning tasks.

The MLP architecture consists of one input layer, one or more hidden layers, and one output layer. Each neuron's input is connected to the output of the previous layer's neurons and the neurons of the output layer determine the class of the input feature vector. AutoMLP is an implementation of MLP that combines genetic algorithm and stochastic optimization. It trains a series of ensemble networks in parallel with different rates and different numbers of hidden layers. After a small, fixed number of epochs, the error rate is determined on a validation set, in which the worst performers are replaced with copies of the best networks, modified to have different numbers of hidden units and learning rates. Hidden unit numbers and learning rates are drawn according to probability distributions derived from successful rates and sizes (Breuel & Shafait, 2010).

3.6.3 Support Vector Machine

Support Vector Machines (SVM) essentially work by finding the optimal hyperplane between classes in a dataset through minimizing the empirical error and maximizing the margin (Shashua, 2009). In a p-dimensional space, a hyperplane is defined as a flat affine subspace of dimension $p - 1$. This means that the hyperplane divides the p-dimensional space in half. The idea of maximizing the margin pertains to the hyperplane that is the farthest minimum distance to the training observations. Although this is often successful, it may lead to overfitting when p is large (James et al., 2013). All in all, SVMs have been shown to perform well in different machine learning tasks, most especially in EEG-based studies (Lotte, Congedo, Lcuyer, Lamarche, & Arnaldi, 2007).

Chapter 4

Design, Implementation, and Challenges

This chapter presents the design of the reader affect model and its implementation. It covers the research framework, how some phases of the research methodology were realized, as well as the challenges encountered during implementation.

4.1 Research Framework

The framework for the research experiment undertaken is presented in Figure 4.1. It begins with the collection of the EEG data from a subject while he is reading the short story. After data processing and preparation, different machine learning techniques are applied to build the reader affect model. A detailed discussion is presented in the subsequent sections.

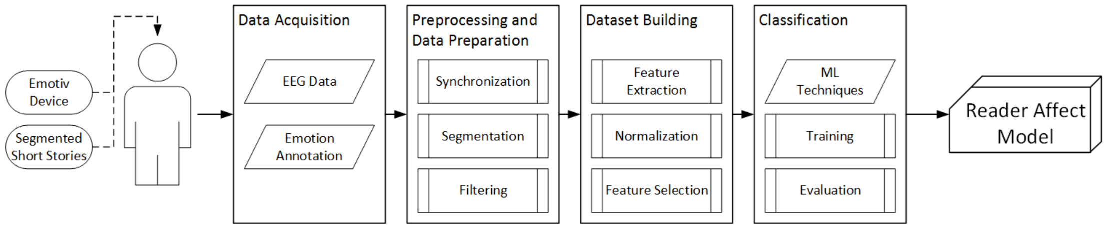


Figure 4.1: Research framework for building the reader affect model.

4.2 Selection of Short Stories

The *short story* is considered the most viable form of narrative for this research since it will entail only a brief reading time (more or less than an hour). This delimitation is set in response to the perceived short attention span of teenage students. Nonetheless, the proposed selections are deemed some of the best in the world and have appeared in must-read lists for high school students. Another criterion is that these selections are deemed to contain prominent levels of emotional resonance (e.g. surprise, confusion, horror, pathos, shock, amusement) due to the clever use of narrative devices by the authors (e.g. suspense, atmosphere, mood, intrusion of the fantastic, characterization). These selections are recommended by the resource person from the creative writing department of the university. The proposed selections are:

- ***The Lovely House* by Shirley Jackson** tells of a young girl Margaret who spends one summer vacation with her friend in the latter's ancestral house. It is a ghostly tale concocted in the Gothic tradition. Shirley Jackson (American author, 1916-1965) is known for the *subversive impact* of her work, for her disquieting portrayal of history and society.

- ***The Veldt* by Ray Bradbury** is a story of a family who lives in an automated house with machines and gadgets that do all the work for them. This house has a nursery room that is able to create virtual realities in response to the telepathic wishes of the children. The finale of this science fiction masterpiece is quite disturbing. Ray Bradbury (American author, 1920-2012) is one of the foremost science fiction writers of the 20th century, known for his compelling visions of a dystopic world.

- ***Gangrene* by F. Sionil Jose** is a moving story about a brief reunion between a father and his son in an army hospital in Sierra Madre. The father is a brain surgeon of the Philippine Army, while the son is with the insurgents. The boy lies dying in the hospital room after he was captured and tortured. Francisco Sionil Jose (Filipino author, 1924-present) is National Artist for Literature. His stories, vividly realistic and socially relevant, are canonical representations of Philippine traditions.

- ***Man from the South* by Roald Dahl** is set in a Jamaican hotel, where a chance encounter takes place between a South American man and a boy. It is a story about a strange bet. Roald Dahl (British author, 1916-1990) wrote such children's classics as *Charlie and the Chocolate Factory* and *James and* *the Giant Peach*. He is famous for his macabre short fiction, tinged with dark humor, irony, and surprise ending.

- ***The Fisherman and the Jinni*** **from** ***One Thousand and One Nights*** is a tale-within-a-tale, filled with magic and sorcery. It begins with an unfortunate meeting between a fisherman and a jinni, and ends with a marvelous tale of a prince and his enchanted kingdom. *One Thousand and One Nights* (ca. 850 AD) is deemed the first novel in the annals of world literature. Composed of several tales of wonder and framed within other narratives, this significant work, of Indian, Persian and Arabic origins, is a classic of all times. Its juicy tales are satires of human follies and foibles.

From these proposed selections, each story was read while being timed to get an idea of how long reading each story will take. *Gangrene* and *The Lovely House* took relatively longer to read (approximately 1 hour and 15 minutes versus approximately 40 minutes for the others). Thus, the remaining three (*The Veldt*, *Man from the South*, and *The Fisherman and the Jinni*) were considered as they fit the duration constraint for each participant's data acquisition session.

4.3 Data Collector Tool

Figure 4.2 presents the architectural design of the data collector tool. It consists of three main modules, which are the connector, parser, and logger, and were developed in C#. The connector module is responsible for integrating external devices, such as the Emotiv device, to the data collector tool. The parser converts the story file into a data structure for the tool. Finally, the logger handles the creation of the output files. Figure 4.3 shows a screenshot of the tool when in use. The emojis used in Figure 4.3b were sourced from EmojiOne[7].

4.3.1 Connector Module

This module has two classes, `EmotivConnector` and `CameraConnector`, that handles the connection of the Emotiv Insight and the machine's native webcam, respectively. `EmotivConnector` uses the Emotiv SDK[8] whereas `CameraConnector`

[7]EmojiOne: The Open Emoji Standard, `http://emojione.com/`
[8]Emotiv SDK, `https://github.com/Emotiv`

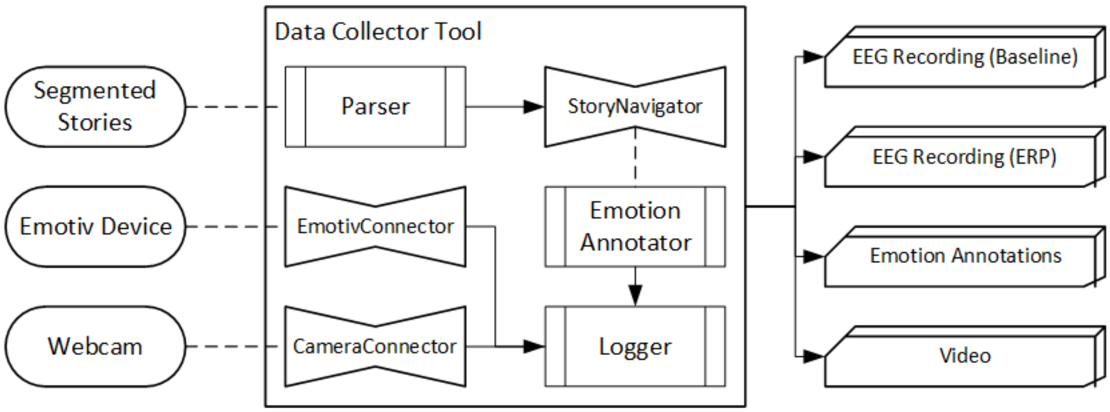


Figure 4.2: Architecture of the data collector tool.

makes use of AForge.NET[9]. Both connectors essentially provide methods for initializing the connection to the devices, thread for data recording, and means of stopping the data capture.

Connector Initialization

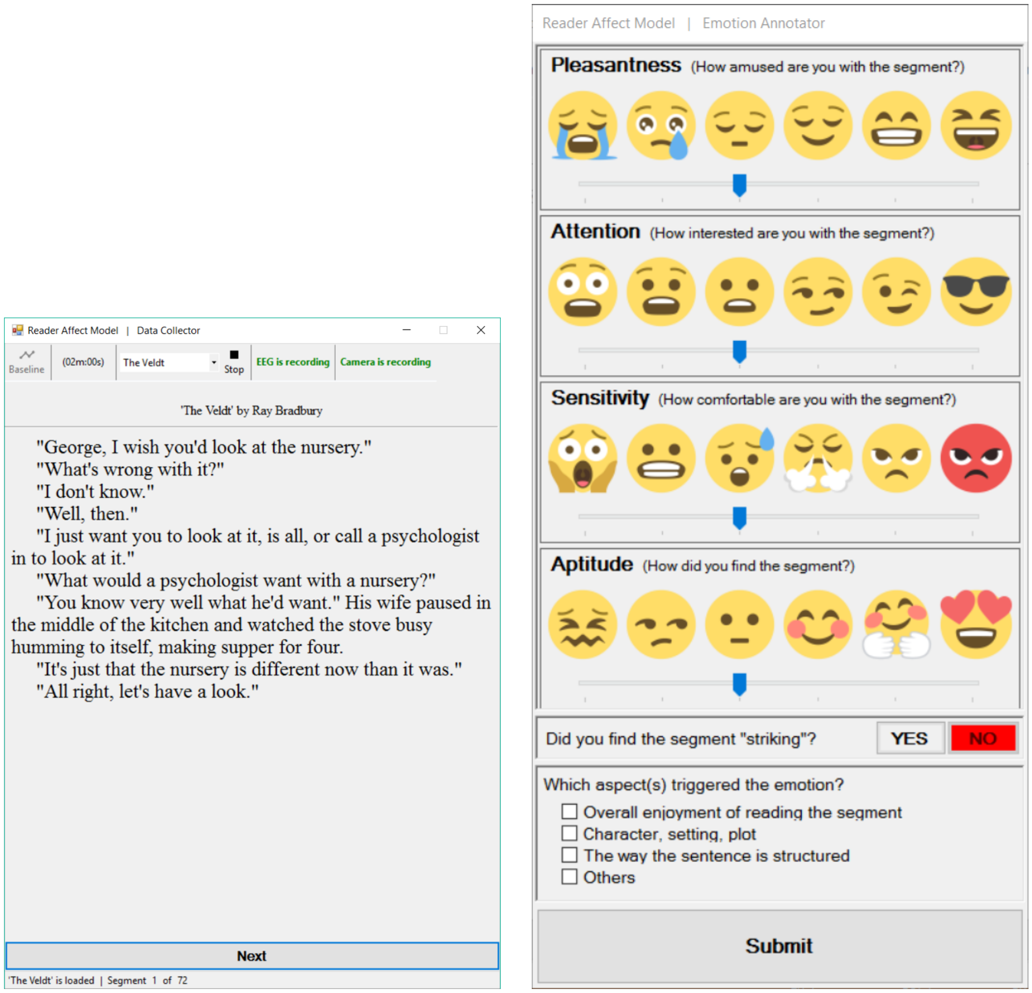
For both cases, it starts by referencing the accompanying .NET dll files. Access to the Emotiv device is through the `EmoEngine` object while access to the webcam is through the `VideoCaptureDevice` object. Listing 4.1 and 4.2 show the code snippet for initialization of the `EmotivConnector` and `CameraConnector`, respectively.

```
EmoEngine engine;
int userID = -1;
  . . .
// Constructor
public EmotivConnector() {
  // Create an instance of the EmoEngine
  engine = EmoEngine.Instance;

  // Add user event handler
  engine.UserAdded += new EmoEngine.UserAddedEventHandler(engine_UserAdded_Event);
}
```

Listing 4.1: Code snippet for initializing the `EmotivConnector`.

[9]AForge.NET framework, `http://www.aforgenet.com/`



(a) Presentation of the segments. (b) Emotion annotation

Figure 4.3: Screenshot of the data collector tool.

```
VideoCaptureDevice videoSource;
  . . .
// Constructor
public CameraConnector() {
  // List all available video sources. (Video sources can be webcams as well as tv cards, etc)
  FilterInfoCollection videosources = new FilterInfoCollection(FilterCategory.VideoInputDevice);

  // Check if at least one video source is available
  if(videosources != null) {
    // Set the videoSource
    videoSource = new VideoCaptureDevice(videosources[0].MonikerString);

    // Create NewFrame event handler. Triggers every time a new frame/image is captured.
    videoSource.NewFrame += new AForge.Video.NewFrameEventHandler(videoSource_NewFrame);
  }
}
```

Listing 4.2: Code snippet for initializing the `CameraConnector`.

Start Data Capture

After initializing the access objects, methods for handling connection events are added. When an Emotiv device connects to the tool, the `engine_UserAdded_Event` method is invoked, as shown in Listing 4.3. Note that this only prepares the tool for data capture, not record the EEG data yet. The method for capturing the EEG data is located in another thread. This is to allow simultaneous execution with the other methods, such as going to the next story segments, recording the emotions annotations, or recording the video. Listing 4.4 shows the code snippet on how the threading is done while Listing 4.5 shows the code for recording the EEG data.

```
private void engine_UserAdded_Event(object sender, EmoEngineEventArgs e) {
  // Record the user
  userID = (int)e.userId;

  // Enable data aquisition for this user.
  engine.DataAcquisitionEnable((uint)userID, true);

  // Ask for up to 1 second of buffered data
  engine.EE_DataSetBufferSizeInSec(1);
}
```

Listing 4.3: Code snippet for `engine_UserAdded_Event` method.

Likewise, the `videoSource_NewFrame` method is triggered whenever a new frame is captured for video recording, as shown in Listing 4.6. To begin video recording, simply call the `StartRecording` method. There is no need to explicitly create a thread for this part of the module. The code snippet is shown in Listing 4.7.

```
// Main
public partial class MainFrame : Form {
  private static EmotivConnector emoConnector;
  private Thread thdEmotivConnector;
    . . .
  emoConnector.Connect();
  StartEegComponent();
    . . .
  // Starts the EEG recording
  private void StartEegComponent() {
    // Create the thread object. This does not start the thread.
    thdEmotivConnector = new Thread(emoConnector.StartRecording);

    // Start the worker thread.
    thdEmotivConnector.Start();
  }
}

public class EmotivConnector {
  // Flag for telling the thread to begin terminating.
  private volatile bool _shouldStop;
    . . .
  // Connects the tool to the headset.
  public void Connect() {
    _shouldStop = false;
    userID = -1;

   // Connect to EmoEngine.
   engine.Connect();
  }

  // Prompts the tool to start capturing data from the device.
  public void StartRecording() {
    while(!_shouldStop) {
      Record();
      Thread.Sleep(delay);
    }
    engine.Disconnect();
  }
}
```

Listing 4.4: Code snippet for Emotiv data capture threading.

```
// Logs the values captured from the device to the output CSV file.
public void Record() {
  // Handle any waiting events
  engine.ProcessEvents();

  // If the user has not yet connected, do not proceed
  if((int)userID == -1)
    return;

  Dictionary<EdkDll.EE_DataChannel_t, double[]> data = engine.GetData((uint)userID);

  if(data == null)
    return;

  int _bufferSize = data[EdkDll.EE_DataChannel_t.ES_TIMESTAMP].Length;

  // Write the data to a file
  for(int i = 0; i < _bufferSize; i++)
    log.Log(Utilities.GetCsvTimestamp(), data[EdkDll.EE_DataChannel_t.AF3][i],
        data[EdkDll.EE_DataChannel_t.T7][i], data[EdkDll.EE_DataChannel_t.O1][i],
        data[EdkDll.EE_DataChannel_t.T8][i], data[EdkDll.EE_DataChannel_t.AF4][i]);
}
```

Listing 4.5: Code snippet for capturing data from the Emotiv device.

```
private void videoSource_NewFrame(object sender, AForge.Video.NewFrameEventArgs eventArgs) {
    . . .
  // Cast the frame as Bitmap object
  videoWriter.WriteVideoFrame((Bitmap)eventArgs.Frame.Clone(), elapse);
}
```

Listing 4.6: Code snippet for `videoSource_NewFrame` method.

```
public partial class MainFrame : Form {
  private static CameraConnector camConnector;
    . . .
  camConnector.StartRecording();
}

public class CameraConnector {
    . . .
  // Starts the camera recording.
  public void StartRecording() {
      . . .
    videoSource.Start();
  }
}
```

Listing 4.7: Code snippet for video recording.

Stop Data Capture

Let's say the main execution is handled by `Thread A`. When the data capture is prompted, `Thread B` is spawned to handle EEG recording. The main execution of going through the story segments and logging the data annotations still resides in `Thread A`. Hence, it will be `Thread A` that will send the signal to `Thread B` to stop recording the EEG through `volatile`[10] `bool _shouldStop`, as shown in Listing 4.8.

```
public partial class MainFrame : Form {
    . . .
  StopEegComponent();
    . . .
  // Stops the EEG recording.
  private void StopEegComponent() {
    // Request that the worker thread stop itself:
    emoConnector.StopRecording();

    // Use the Join method to block the current thread until the object's thread terminates.
    thdEmotivConnector.Join();
  }
}

public class EmotivConnector {
    . . .
  // Prompts the tool stop capturing data from the device.
  public void StopRecording() {
    _shouldStop = true;
  }
}
```

Listing 4.8: Code snippet for stopping EEG recording.

In contrast, simply call the `StopRecording` method to stop video recording. However, there are still additional measures to add to the code for the program not to crash. `StopRecording` stops the current video recording and closes the video file stream. This does not mean that the connection to the camera has closed. Hence, this allows the data collector tool to have many video recordings in a single run. To properly terminate the program, `CloseConnector` must also be called. This method is invoked via the `MainFrame_FormClosed` when the user exits the tool. `CloseConnector` stops any current video recording and then sets the `videoSource` to `null`. These are seen in Listing 4.9.

[10] `volatile` C# Reference, https://msdn.microsoft.com/en-us/library/x13ttww7.aspx

```
public partial class MainFrame : Form {
    . . .
  private void MainFrame_FormClosed(object sender, FormClosedEventArgs e) {
    camConnector.CloseConnector();
  }
    . . .
  camConnector.StopRecording();
}

public class CameraConnector {
    . . .
  // Stops the camrea recording and closes the video file stream.
  public void StopRecording() {
    //Stop and free the webcam object.
    videoSource.SignalToStop();

    // Close the video file stream.
    videoWriter.Close();
  }

  // Closes the CameraConnector.
  public void CloseConnector() {
    if(videoSource.IsRunning)
      StopRecording();

    videoSource = null;
  }
}
```

Listing 4.9: Code snippet for stopping video recording.

4.3.2 Parser Module

This module handles the conversion of the short story segments into a data structure for the tool. These segmented stories are represented in XML documents, as shown in Listing 4.10. The `<story>` tag serves as the root element of the XML document, with `title` and `author` as attributes. Each segment is denoted by the `<segment>` tag. The `<part>` tag represents the paragraphs in each segment.

```
&lt;?xml version="1.0" encoding="utf-8" ?&gt;
&lt;story title="Man from the South" author="Roald Dahl"&gt;
  &lt;segment&gt;
     &lt;part&gt;It was getting on towards six o'clock so I thought I'd buy myself a beer and go out and sit in a deckchair by the swimming pool and have a little evening sun.&lt;/part&gt;
     &lt;part&gt;I went to the bar and got the beer and carried it outside and wandered down the garden towards the pool.&lt;/part&gt;
  &lt;/segment&gt;
  &lt;segment&gt;
      &lt;part&gt;It was a fine garden with lawns and beds of azaleas and tall coconut palms, and the wind was blowing strongly through the tops of the palm trees, making the leaves hiss and crackle as though they were on fire. I could see the clusters of big brown nuts hanging down underneath the leaves.&lt;/part&gt;
  &lt;/segment&gt;
    . . .
  &lt;segment&gt;
      &lt;part&gt;This is the last segment.&lt;/part&gt;
  &lt;/segment&gt;
&lt;/story&gt;
```

Listing 4.10: XML representation of the segmented *Man From the South* by Roald Dahl

The presentation of the segments is controlled by the `StoryNavigator`. This class is a counter that knows the current and previous segment numbers. When the user proceeds to the next segment, both counters increment by one.

4.3.3 Logger Module

This module contains two classes, `EegLogger` and `EmotionLogger`, that implement the interface `ILogger`. `ILogger` forces the two classes to have an `Initialize` and `Log` method. The `Log` method is implemented in such a way that should there be an unexpected interruption (i.e. the program crashed), the output files are not corrupted. Each time `Log` is called, it opens the file stream, appends the contents, and closes the stream.

Initial testing of the tool shows that it appears to be functioning as intended as it was able to capture the EEG data, emotion annotations, and video recording without the program abruptly terminating. However, a logic error was encountered when the collected data were put through the preprocessing phase.

The error was located at the line of code for logging the timestamp. This error is peculiar because it works in the `EegLogger` while it did not in the `EmotionLogger` despite the fact they use the same function. In `EmotionLogger`, the program takes note of the time the participant annotates the emotions because this time period will be removed from the ERP EEG recording. Listing 4.11 shows the logical error where this occurred. The output for those logged in the `EegLogger` is something like `1468300340536.31`. Those in the `EmotionLogger` looks like `1.4683E+12`.

```
// Method for getting the timestamp.
public static double GetCsvTimestamp();

// Usage in EegLogger
eeglog.Log(GetCsvTimestamp(), <otherParameters>);

// Usage in EmotionLogger
startTime = GetCsvTimestamp();
  . . .
endTime = GetCsvTimestamp();
emoLog.Log(startTime, endTime, <otherParameters>);
```

Listing 4.11: Logic error in code for logging timestamp.

This minor error was resolved by simply converting the `double` value timestamp to `string` before storing it in the variable as shown in Listing 4.12.

```
startTime = GetCsvTimestamp().ToString();
  . . .
endTime = GetCsvTimestamp().ToString();
emoLog.Log(startTime, endTime, <otherParameters>);
```

Listing 4.12: Resolution to timestamp logging logic error.

4.4 Data Acquisition

Figure 4.4 shows the procedural flow of the data acquisition the participants underwent.

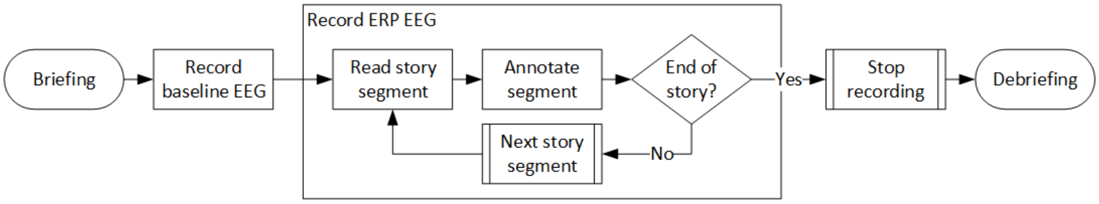


Figure 4.4: Data acquisition flowchart.

1. **Briefing.** Each participant is informed about the experiment. They are briefed about the process and the definition of the labels they will annotate in each segment. At this point, they also signify their consent for participation in the experiment by signing the informed consent form (see Appendix B).

2. **Record baseline EEG.** The participants are then asked to close their eyes and relax for 2 minutes while wearing the Emotiv headset. This recording serves as the baseline.

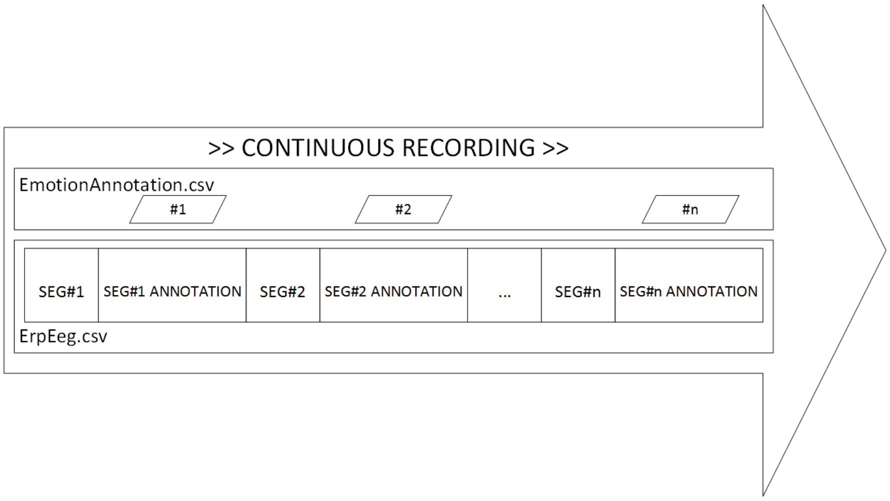


Figure 4.5: Representation of the ERP recording step and its output files.

3. **Record ERP EEG.** A continuous recording of EEG occurs at this stage. The participants are to read the current story segment presented in the screen (see Figure 4.3a). After reading, they must annotate the segment (see Figure 4.3b) based on what they felt when they read it (based on the HoE model by Cambria et al. (2012)) and what they think triggered the emotion (based on the classification of ELR by Miall and Kuiken (2002)). This process is repeated until the last segment is reached, only then will the EEG recording stop. There is no time limit in reading the segments. The participants are encouraged to read at their own pace. Figure 4.5 shows a graphical representation of this step. Each block in `EmotionAnnotaion.csv` represents a row of data in the file. Each block in `ErpEeg.csv` represents the time blocks between reading and annotating the segment.

4. **Debriefing.** At this last part of the experiment, each participant is interviewed regarding what they thought of the whole process. Feedback, comments, and suggestions regarding the data collector tool and the overall experiment itself are solicited. They are also asked about the emotion annotations they tagged in the segments, on why they feel a segment is *striking* or on what triggered that emotion. They also answered a profile questionnaire about their reading preferences.

The data were collected from undergraduate and graduate students of De La Salle University, aged 18 and above. A total number of 32 students participated in this experiment. The short story that they read was *The Veldt* by Ray Bradbury, which was manually divided into 72 segments. The participants have not yet read *The Veldt* nor are they familiar with its plot. Hence, the data gathered from them are their first impressions towards the story.

Note that despite the Emotiv Insight having the convenience of a dry sensor, it still suffered the technical factors Picard discussed. The headset is not adjustable as it only comes in one size. The target is to have all the indicators on the Emotiv Xavier control panel light up as green before starting the experiment, as shown in Figure 3.3c. To achieve this, the sensors on the headset must have contact with the skin. It was observed that those participants with short hair, mostly the males, were the easiest to set-up the headset to. With the females, their hair needed to be parted on the location where the sensors are. This was so that the sensors will have contact with the scalp. If all the green lights could not be achieved, then at least a combination of green and orange lights was sufficient.

4.5 EEG Processing

4.5.1 Preprocessing

For each participant, two sets of EEG recordings were collected: the baseline and the ERP recording. The ERP recording is first processed by synchronizing and merging it with its accompanying annotations as shown in Figure 4.6. Note that the time block spent in annotating the segment was removed so that the data being processed is only the EEG of the person while they are reading.

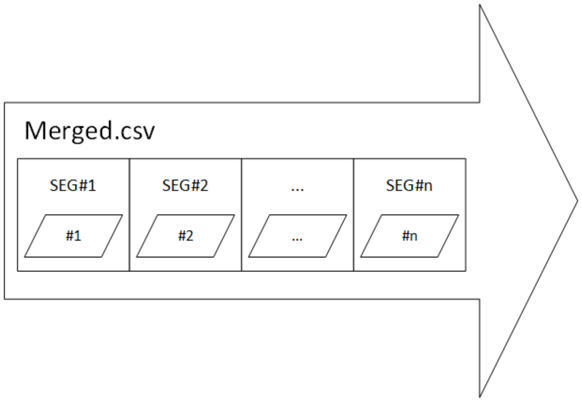
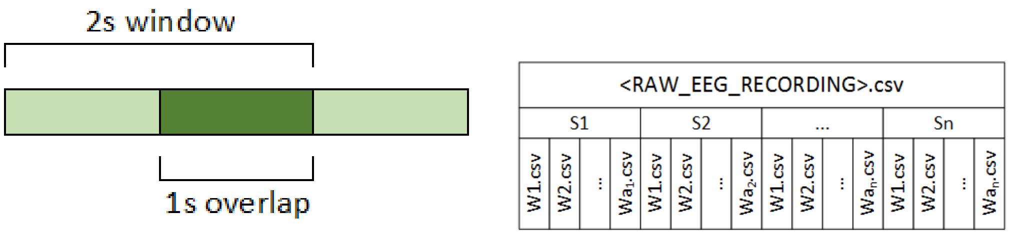


Figure 4.6: `Merged.csv` representation.

The resulting merged file is then divided into the corresponding story segments and then further divided into 2-second windows with 1-second overlap as shown in Figure 4.7. Note that $\texttt{Wa}_\texttt{n}$ represents the number of windows for Segment #`n` and it is also the number of ERP instances for that particular participant.



(a) 2-second windows with 1-second overlap

(b) Resulting files after windowing

Figure 4.7: Preprocessing for EEG data during reading session.

4.5.2 Feature Extraction

While not using the interactive interface of EEGLAB[11], this MATLAB toolbox provides the specifications of useful functions for EEG processing. EEGLAB is used as a reference and then the default MATLAB program is used to recreate the EEGLAB functions as needed.

According to EEGLAB, `eegfilt()` is specified as:

```
eegfilt() - (high|low|band)-pass filter data using two-way least-squares FIR
    filtering. Optionally uses the window method instead of least-squares.
    Multiple data channels and epochs supported. Requires the MATLAB Signal
    Processing Toolbox.

Inputs:
   data         =  (channels,frames*epochs) data to filter
   srate        =  data sampling rate (Hz)
   locutoff     =  low-edge frequency in pass band (Hz)  {0 -> lowpass}
   hicutoff     =  high-edge frequency in pass band (Hz) {0 -> highpass}
   epochframes =  frames per epoch (filter each epoch separately
                  {def/0: data is 1 epoch}
   filtorder    =  length of the filter in points
                  {default 3*fix(srate/locutoff)}
   revfilt      =  [0|1] reverse filter (i.e. bandpass filter to notch
                  filter). {default 0}
   firtype      =  'firls'|'fir1' {'firls'}
   causal       =  [0|1] use causal filter if set to 1 (default 0)

Outputs:
   filtwts = filter coefficients [smoothdata <- filtfilt(filtwts,1,data)]
```

Listing 4.13: `eegfilt()` specification

Using the specification in Listing 4.13, the creation of the theta θ, alpha α, beta β, and gamma γ band pass filters is done with MATLAB's Filter Design and Analysis Tool[12]. This tool is accessed by calling `fdatool` in MATLAB. The creation of the filters is done by tweaking the parameters indicated in the tool to match that of `eegfilt()`. The `Response Type`, `Design Method`, and `Magnitude Specifications` are the same among the different frequency band filters. The computation to get the `Filter Order` is defined in `filtorder` of Listing 4.13. The content of the `Frequency Specifications` depends on which frequency band is being created. `Fpass1` and `Fpass2` refer to the lower and upper bound frequency of the chosen frequency band (refer to Table 3.1). `Fstop1` and `Fstop2` are the ± 0.5 to the `Fpass` values. Figure 4.8 shows the creation of the alpha α band pass filter in the Filter Design and Analysis Tool.

[11] EEGLAB, `https://sccn.ucsd.edu/eeglab/`
[12] MATLAB `fdatool`, `https://www.mathworks.com/help/signal/ref/fdatool.html`

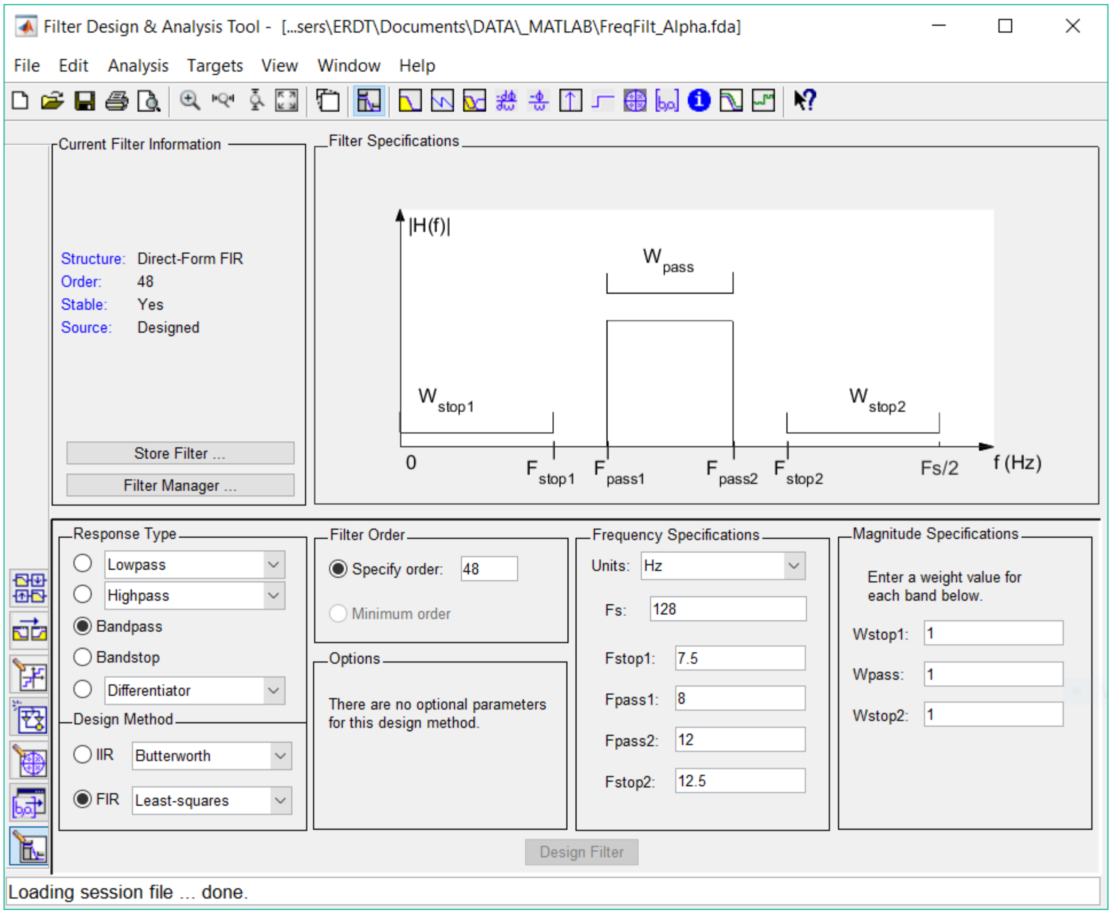


Figure 4.8: MATLAB `fdatool` session for alpha α band pass filter

Before extracting the frequency bands, smoothing via a moving average filter was applied to the EEG signals in the baseline and `Wa`$_n$`.csv`. Smoothing removes the noise and sees the trends in the data by showing the gradual change in the values. In extracting the frequency bands, the output of the created filters is used as one of the parameters of `filter()`. For each band, the signals were transformed to the frequency domain using the Fast Fourier Transform `fft()` and then the magnitude, PSD, DASM, and RASM features were computed (refer to Section 3.5.1 for description of each feature type). There was a total of 252 extracted features. Figure 4.9 shows the summary of these features.

<EEG_CHANNEL>																																			
theta						alpha						beta_lo						beta_mid						beta_hi						gamma					
MAG			PSD			MAG			PSD			MAG			PSD			MAG			PSD			MAG			PSD			MAG			PSD		
MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN

<ELECTRODE_PAIRS>																																			
theta						alpha						beta_lo						beta_mid						beta_hi						gamma					
DASM			RASM			DASM			RASM			DASM			RASM			DASM			RASM			DASM			RASM			DASM			RASM		
MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN

Figure 4.9: Summary of features extracted from the data.

4.6 Self-Reported Emotions and Assignment of Instance Labels

As mentioned previously, each `Wa`$_n$ is an instance for each participant. That particular instance is labeled according to the self-reported emotions. With regard to the HoE, the participants were to rate each of the 4 dimensions from -3 to +3. If the value is negative, the assigned label is *low*. If the value is positive, then the assigned label is *high*. With regard to the ELR, whichever the participant chose is the assigned label. It is possible for the instance to have multiple assigned labels for ELR.

4.7 Dataset Building

This research is concerned with the brainwaves of participant p during reading time. Those brainwaves (ERP) are isolated by employing baseline correction, which is simply subtracting the ith feature value of the baseline $b_{feat}(p,i)$ from all the ERP instances $Wa_{n_{feat}}(p,i)$ as shown in Equation 4.1 (Woodman, 2010).

$$Wa_{n_d}(p,i) = b_{feat}(p,i) - Wa_{n_{feat}}(p,i) \tag{4.1}$$

The results of these, as denoted by $Wa_{n_d}(p,i)$, are then standardized by computing for their z-score values $Wa_{n_z}(p,i)$ as shown in Equation 4.2, where $\mu(i)$ and $\sigma(i)$ are the mean and standard deviation, respectively, of the ith feature for all participants in the particular dataset being built. Extreme z-score values were clipped to at least -3 and at most +3.

$$Wa_{n_z}(p,i) = \frac{Wa_{n_d}(p,i) - \mu(i)}{\sigma(i)} \tag{4.2}$$

The research also identified certain *reader profiles* as a means of grouping the participants and building the datasets around those groups. These profiles were based on what the participants reported on the profile questionnaire. Table 4.1 shows the number of participants per group per profile.

Table 4.1: Number of participants per dataset per reader profile.

Profile	Profile Groups	# of Participants
P1: Sex	Female	23
	Male	9
P2: Reading Preference (Traditional books vs. eBooks)	RP1: Prefers reading traditional books only	8
	RP2: Prefers traditional books over eBooks	12
	RP3: Fine with both traditional books and eBooks	9
	RP4: Prefers eBooks over traditional books	3
	RP5: Prefers reading eBooks only	0
P3: Reading Frequency (How often do you read for fun?)	RF1: Almost all the time (approx. 1-2 books a week)	2
	RF2: Now and then (approx. 1-2 books a month)	14
	RF3: Not very often (approx. 1-2 books in 6 months)	14
	RF4: Never	2

Chapter 5

Results and Analysis

This chapter presents the experiments conducted as well as the observations and findings.

5.1 Experiments

The experiments conducted in this research are binary classifications of the HoE and ELR models on different datasets and classification methods. The different variables are listed below:

- **Dataset:** Female, Male, RP1 (prefers reading traditional books only), RP2 (prefers reading traditional books over eBooks), RP3 (fine with reading on both traditional books and eBooks), RP4 (prefers reading eBooks over traditional books), RF1 (reads books for fun almost all the time), RF2 (reads books for fun every now and then), RF3 (does not read for fun very often), RF4 (never reads for fun), Sex-merged, RP-merged, RF-merged
- **Classification Method:** Decision Tree, Support Vector Machine, Multi-layer Perceptron
- **Hourglass of Emotions class labels (high/low)**: Aptitude (`AP`), Attention (`AT`), Pleasantness (`PL`), Sensitivity (`SE`)

- **Emotions of Literary Response class labels (true/false)**: Aesthetic feelings (AE), Evaluative feelings (EV), Narrative feelings (NA), Others (OT)

Because of the uneven distribution of participants among the reading profiles, the datasets were filtered such that the number of participants in all the datasets in a particular reader profile is equal to the dataset in that profile with the least number of participants. For example, the number of participants of the Female dataset will then be 9 and these 9 participants were chosen at random. Tables 5.1 and 5.2 show the number of all the instances per class label per dataset. Note that those in blue are the balanced or relatively balanced (±10%) datasets.

Table 5.1: Number of all the instances per HoE class labels per dataset.

	FEMALE (n = 23120)		**MALE (n = 21843)**		**RP1 (n = 8555)**		**RP2 (n = 8414)**		**RP3 (n = 9003)**	
	LOW	**HIGH**	**LOW**	**HIGH**	**LOW**	**HIGH**	**LOW**	**HIGH**	**LOW**	**HIGH**
AP	12205 (52.8%)	10915 (47.2%)	13428 (61.5%)	8415 (38.5%)	6529 (76.3%)	2026 (23.7%)	3164 (37.6%)	5250 (62.4%)	4126 (45.8%)	4877 (54.2%)
AT	15842 (68.5%)	7278 (31.5%)	10614 (48.6%)	11229 (51.4%)	6836 (79.9%)	1719 (20.1%)	3658 (43.5%)	4756 (56.5%)	6436 (71.5%)	2567 (28.5%)
PL	10889 (47.1%)	12231 (52.9%)	9601 (44.0%)	12242 (56.0%)	6444 (75.3%)	2111 (24.7%)	2249 (26.7%)	6165 (73.3%)	3601 (40.0%)	5402 (60.0%)
SE	17966 (77.7%)	5154 (22.3%)	15269 (69.9%)	6574 (30.1%)	7141 (83.5%)	1414 (16.5%)	4521 (53.7%)	3893 (46.3%)	7374 (81.9%)	1629 (18.1%)

	RP4 (n = 7561)		**RF1 (n = 3487)**		**RF2 (n = 3517)**		**RF3 (n = 6019)**		**RF4 (n = 5853)**	
	LOW	**HIGH**	**LOW**	**HIGH**	**LOW**	**HIGH**	**LOW**	**HIGH**	**LOW**	**HIGH**
AP	5183 (68.5%)	2378 (31.5%)	1125 (32.3%)	2362 (67.7%)	1052 (29.9%)	2465 (70.1%)	5501 (91.4%)	518 (8.6%)	2766 (47.3%)	3087 (52.7%)
AT	3728 (49.3%)	3833 (50.7%)	1566 (44.9%)	1921 (55.1%)	1816 (51.6%)	1701 (48.4%)	5505 (91.5%)	514 (8.5%)	3693 (63.1%)	2160 (36.9%)
PL	4148 (54.9%)	3413 (45.1%)	967 (27.7%)	2520 (72.3%)	801 (22.8%)	2716 (77.2%)	5341 (88.7%)	678 (11.3%)	2022 (34.5%)	3831 (65.5%)
SE	6627 (87.6%)	934 (12.4%)	2374 (68.1%)	1113 (31.9%)	998 (28.4%)	2519 (71.6%)	5990 (99.5%)	29 (0.5%)	4479 (76.5%)	1374 (23.5%)

	SEX-merged (n = 44963)		**RP-merged (n = 33533)**		**RF-merged (n = 18876)**	
	LOW	**HIGH**	**LOW**	**HIGH**	**LOW**	**HIGH**
AP	25633 (57.0%)	19330 (43.0%)	19002 (56.7%)	14531 (43.3%)	10444 (55.3%)	8432 (44.7%)
AT	26456 (58.8%)	18507 (41.2%)	20658 (61.6%)	12875 (38.4%)	12580 (66.6%)	6296 (33.4%)
PL	20490 (45.6%)	24473 (54.4%)	16442 (49.0%)	17091 (51.0%)	9131 (48.4%)	9745 (51.6%)
SE	33235 (73.9%)	11728 (26.1%)	25663 (76.5%)	7870 (23.5%)	13841 (73.3%)	5035 (26.7%)

Table 5.2: Number of all the instances per ELR class labels per dataset.

	FEMALE (n = 23120)		MALE (n = 21843)		RP1 (n = 8555)		RP2 (n = 8414)		RP3 (n = 9003)	
	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE
AE	20139 (87.1%)	2981 (12.9%)	17536 (80.3%)	4307 (19.7%)	7780 (90.9%)	775 (9.1%)	7305 (86.8%)	1109 (13.2%)	7491 (88.2%)	1512 (16.8%)
EV	12916 (55.9%)	10204 (44.1%)	14795 (67.7%)	7048 (32.3%)	4966 (58.0%)	3589 (42.0%)	5396 (64.1%)	3018 (35.9%)	3489 (38.8%)	5514 (61.2%)
NA	12653 (54.7%)	10467 (45.3%)	8303 (38.0%)	13540 (62.0%)	4322 (50.5%)	4233 (49.5%)	3530 (42.0%)	4884 (58.0%)	4105 (45.6%)	4898 (54.4%)
OT	22925 (99.2%)	195 (0.8%)	21083 (96.5%)	760 (3.5%)	7555 (88.3%)	1000 (11.7%)	8377 (99.6%)	37 (0.4%)	8938 (99.3%)	65 (0.7%)

	RP4 (n = 7561)		RF1 (n = 3487)		RF2 (n = 3517)		RF3 (n = 6019)		RF4 (n = 5853)	
	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE
AE	6123 (81.0%)	1438 (19.0%)	3062 (87.8%)	425 (12.2%)	2738 (77.9%)	779 (22.1%)	5577 (92.7%)	442 (7.3%)	4988 (85.2%)	865 (14.8%)
EV	4030 (53.3%)	3531 (46.7%)	1431 (41.0%)	2056 (59.0%)	1160 (33.0%)	2357 (67.0%)	1866 (31.0%)	4153 (69.0%)	2474 (42.3%)	3379 (57.7%)
NA	5307 (70.2%)	2254 (29.8%)	964 (27.6%)	2523 (72.4%)	2432 (69.1%)	1085 (30.9%)	5230 (86.9%)	789 (13.1%)	3420 (58.4%)	2433 (41.6%)
OT	6905 (91.3%)	656 (8.7%)	3471 (99.5%)	16 (0.5%)	3432 (97.6%)	85 (2.4%)	5980 (99.4%)	39 (0.6%)	5788 (98.9%)	65 (1.1%)

	SEX-merged (n = 44963)		RP-merged (n = 33533)		RF-merged (n = 18876)	
	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE
AE	37675 (83.8%)	7288 (16.2%)	28699 (85.6%)	4834 (14.4%)	16365 (86.7%)	2511 (13.3%)
EV	27711 (61.6%)	17252 (38.4%)	17881 (53.3%)	15652 (46.7%)	6931 (36.7%)	11945 (63.3%)
NA	20956 (46.6%)	24007 (53.4%)	17264 (51.5%)	16269 (48.5%)	12046 (63.8%)	6830 (36.2%)
OT	44008 (97.9%)	955 (2.1%)	31775 (94.8%)	1758 (5.2%)	18671 (98.9%)	205 (1.1%)

Across all classification experiments, the classifiers were trained with a *participant-fold cross-validation* (can also be known as *leave-one-participant-out cross-validation*), in which one participant from the dataset serves as the test set and the rest of the participants are the training sets. The validation is repeated until all of the participants have experienced being a test set. This type of validation scheme was employed to maintain mutually exclusive training and test sets.

To have a picture of the emotions the participants felt while reading *The Veldt*, Figures 5.1 and 5.2 map the averaged self-reports of the participants per class label per segment. The idea behind the Hourglass of Emotions model is to determine what the reader is feeling. The values for HoE range from -3 to +3. The idea behind the Emotions of Literary Response model is to determine what aspect of the story triggered the emotions. The values for ELR range from 0 to 1, with 1 denoting that the particular ELR class label is the cause of the emotions. Note that the 5-act dramatic structure of *The Veldt* is as follows: the *exposition* consists of segments 1 to 17, the *rising action* consists of segments 18 to 46, the *climax* consists of segments 47 to 58, the *falling action* consists of segments 59 to 70, and the *resolution* consists of segments 71 to 72.

Referring to Figure 5.1, generally, the participants feel a mild love or hate towards the events in the segments as their Aptitude ranges from only -1 to +1; their response towards the events is more of fear, whereas their perception towards details is more of surprise because almost all of the averaged Sensitivity and Attention self-reports are in the negative range; it is observed that there are generally happy feelings at the beginning of the story but by the end of the rising action, the Pleasantness fluctuates from the positive range to the negative range. Referring to Figure 5.2, it is observed that the participants report that the cause of their emotions are more on the narrative and evaluative aspects than aesthetics. Note that these are their first impressions towards the story, which makes the aesthetic aspect unnoticable at the moment.

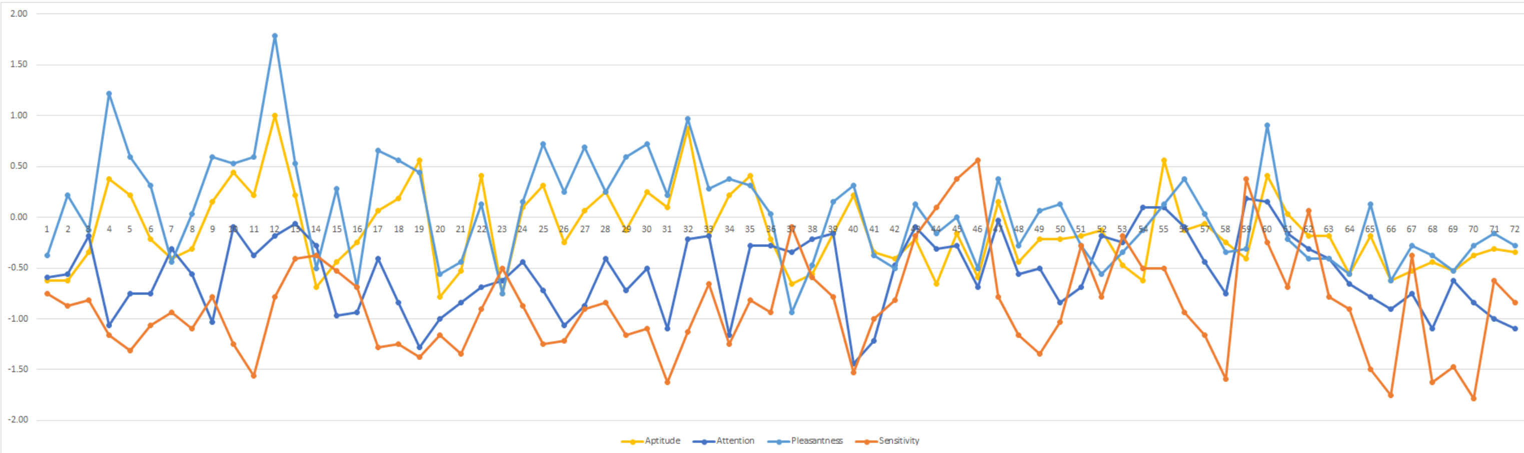


Figure 5.1: Mapping of the averaged HoE class label self-reports to *The Veldt* segments.

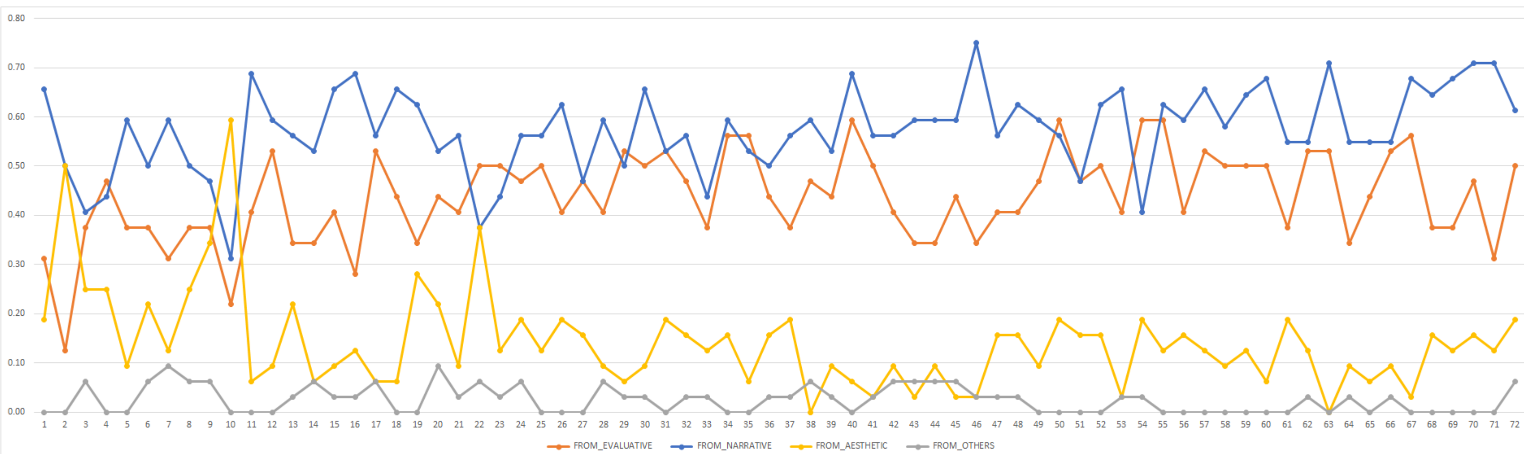


Figure 5.2: Mapping of the averaged ELR class label self-reports to *The Veldt* segments.

5.2 Evaluation Metrics

In evaluating the performance of the classifiers, precision, recall, accuracy, f-measure, and Cohen's kappa are well established metrics in machine learning tasks.

Table 5.3: Confusion matrix for the evaluation metrics computation.

	True N	True P
Pred. N	True Negative (TN)	False Negative (FN)
Pred. P	False Positive (FP)	True Positive (TP)

Using the confusion matrix defined in Table 5.3 as basis, the evaluation metrics are computed as follows:

Accuracy, as shown in Equation 5.1, is the number of correct predictions divided by the number of predictions made (total population). This metric alone is not enough because of the accuracy paradox, wherein the model yields a high accuracy despite having a significant amount of False Positives and False Negatives.

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{5.1}$$

Precision, as shown in Equation 5.2, is the measure of the exactness of a classifier. A low precision can also indicate a large number of False Positives. Precision is also known as positive predictive value (PPV).

$$precision = \frac{TP}{TP + FP} \tag{5.2}$$

Recall, as shown in Equation 5.3, is the measure of the completeness of a classifier. A low recall indicates many False Negatives. If the classifier, for example, predicts all the instances as positive, then it has a perfect recall value of 1.

$$recall = \frac{TP}{TP + FN} \tag{5.3}$$

F-measure, as shown in Equation 5.4, is the harmonic mean of precision and recall.

$$F = 2 \times \frac{precision \times recall}{precision + recall} \tag{5.4}$$

Cohen's kappa, as shown in Equation 5.5, is a metric that compares an observed accuracy p_o with an expected accuracy p_e. It is generally thought to be a more robust measure because it considers the possibility of the agreement occurring by chance. p_o is the number of instances classified correctly throughout the entire confusion matrix. p_e is the accuracy that any random classifier would be expected to achieve based on the confusion matrix. $\kappa < 0$ means no agreement, whereas when $\kappa = 1$ means perfect agreement.

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{5.5}$$

where: $p_o = accuracy$

$$p_e = \frac{\frac{(TN+FP)\times(TN+FN)}{TP+FP+TN+FN} + \frac{(FN+TP)\times(FP+TP)}{TP+FP+TN+FN}}{TP+FP+TN+FN}$$

5.3 Results

This section presents and discusses the results of each experiment. Refer to Appendix D for the confusion matrices of these experiments.

5.3.1 Baseline Classification Performance Based on Decision Trees

In this experiment, the goal was to set the baseline performance using Decision Trees and to see whether there is an improvement in performance among the profile-specific datasets. The classifiers were trained using all 252 EEG features. Note that DTs describe the classification model by its relevant features, which are interestingly from the alpha α and theta θ bands only (refer to Appendix E). The theta θ band is concerned with intuition, creativity, recall, and imagination, whereas the alpha α bands are concerned with relaxed, but not drowsy, and conscious mental conditions.

As shown in Figures 5.3, 5.4, 5.5, 5.6, 5.7, and 5.8, it is observed that on an average basis, there is no significant improvement in the performance between the general dataset and the profile-specific datasets. However, on a label-specific basis, there are significant improvements on certain profiles, as discussed below. Refer to Appendix F to see the summary of the performance results of this experiment.

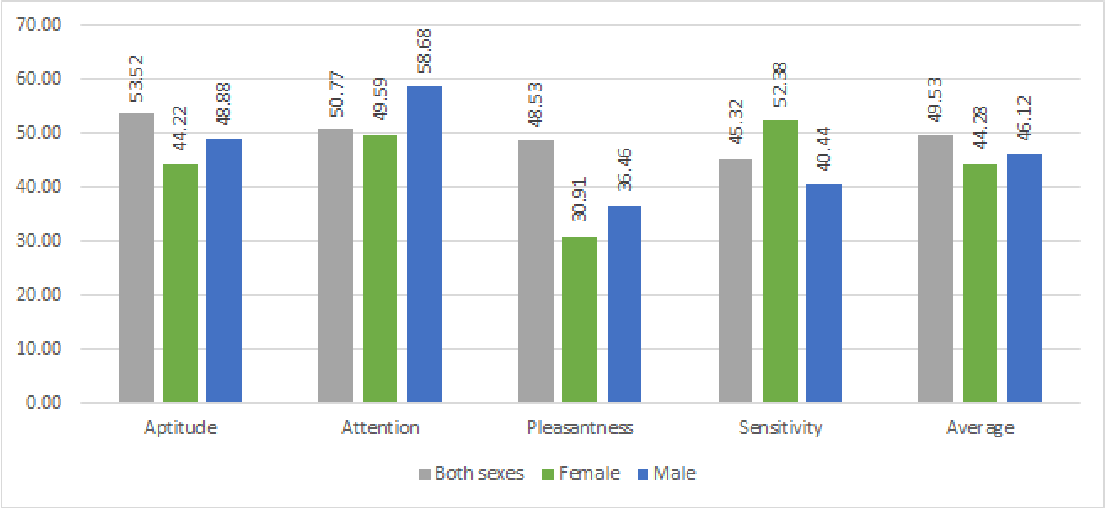


Figure 5.3: Comparison of DT f-measure values per HoE class label of the Sex-merged, Female, and Male datasets.

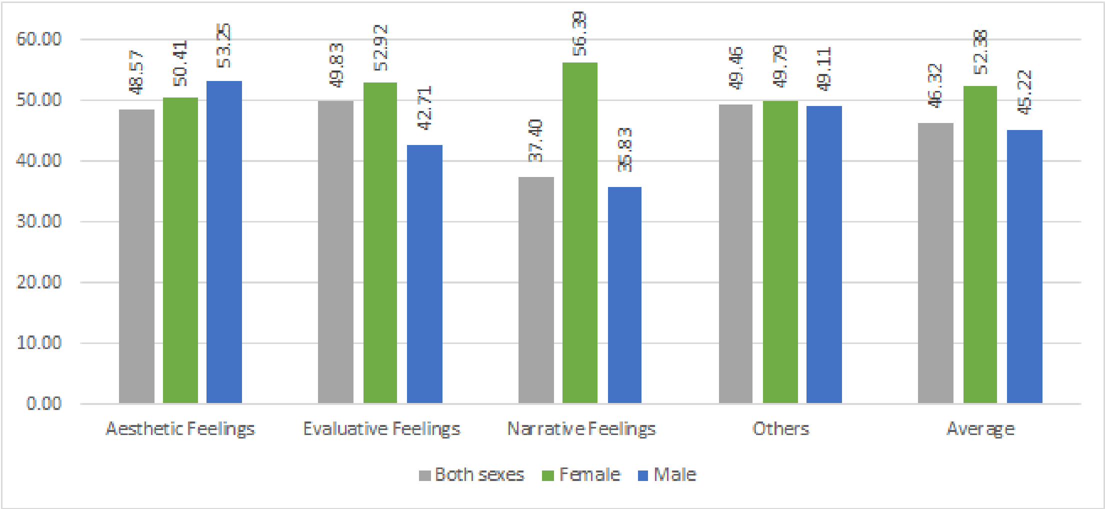


Figure 5.4: Comparison of DT f-measure values per ELR class label of the Sex-merged, Female, and Male datasets.

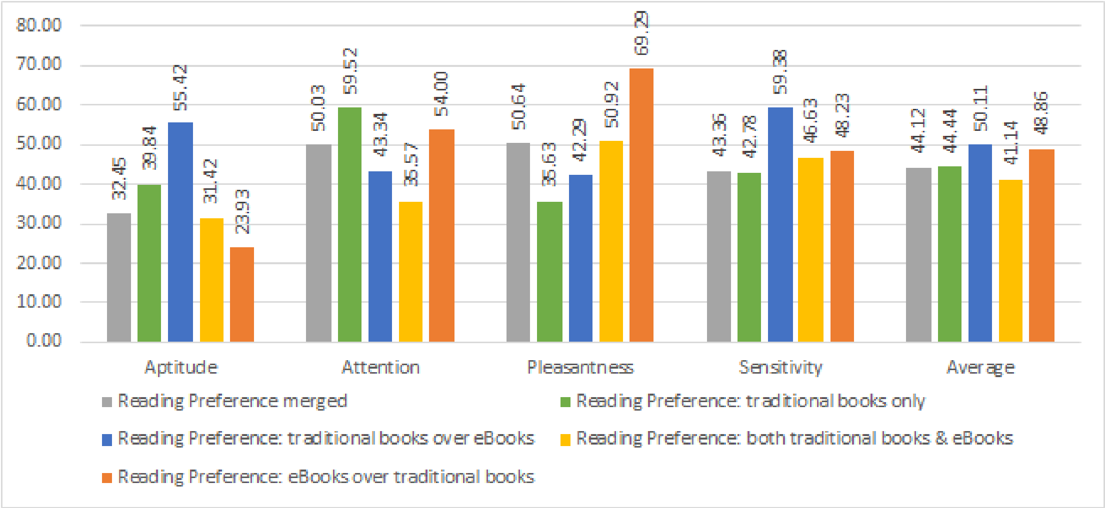


Figure 5.5: Comparison of DT f-measure values per HoE class label of the RP-merged, RP1, RP2, RP3 and RP4 datasets.

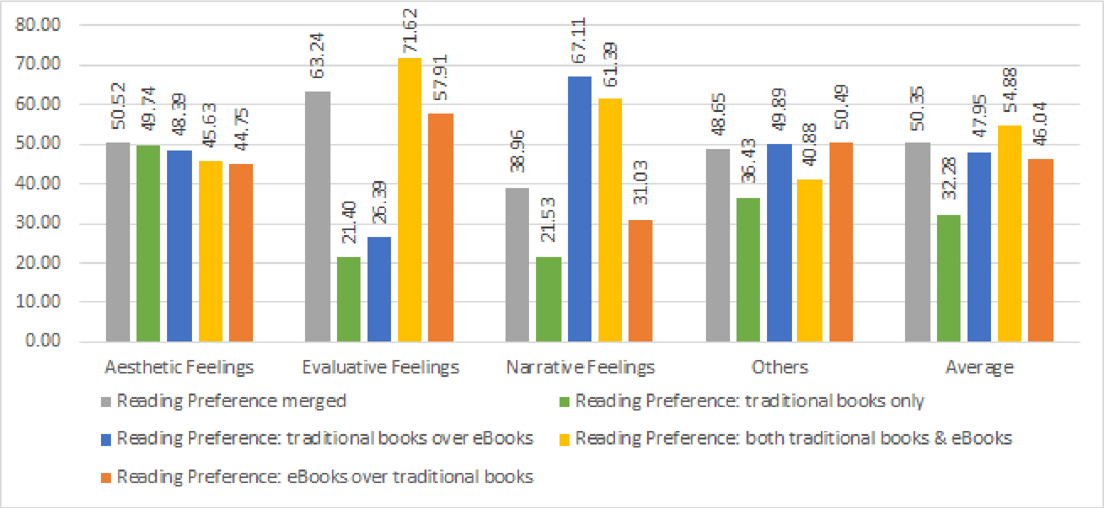


Figure 5.6: Comparison of DT f-measure values per ELR class label of the RP-merged, RP1, RP2, RP3 and RP4 datasets.

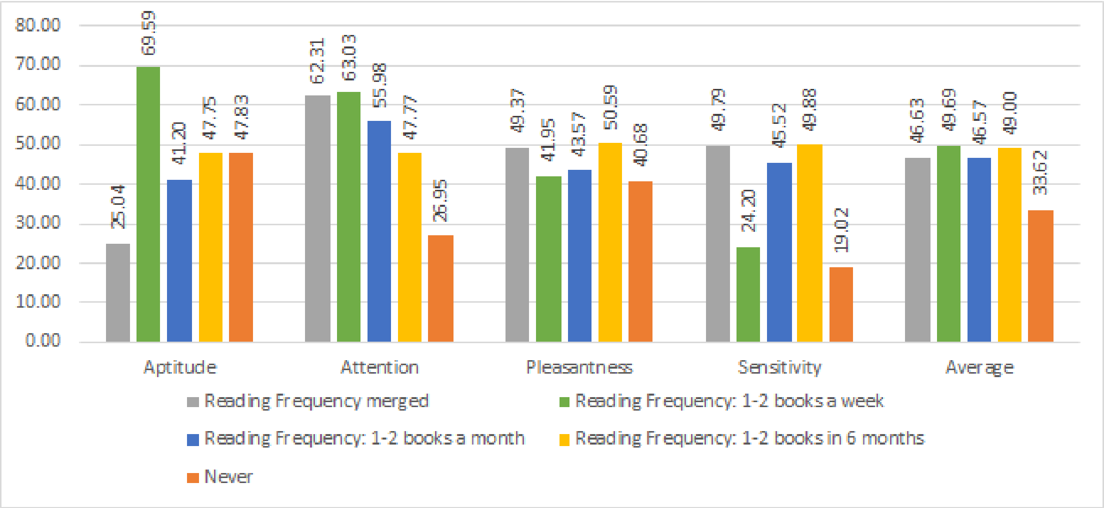


Figure 5.7: Comparison of DT f-measure values per HoE class label of the RF-merged, RF1, RF2, RF3 and RF4 datasets.

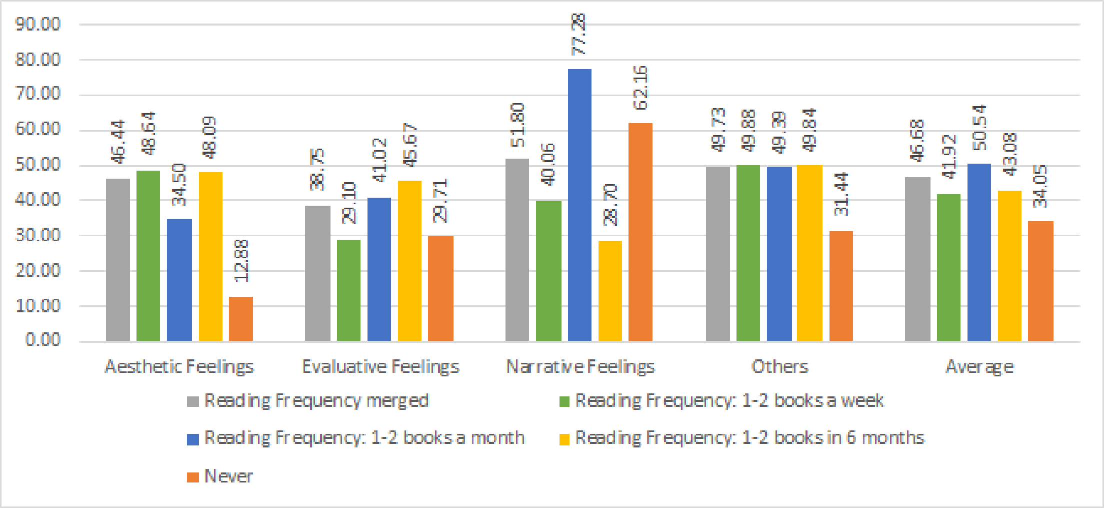


Figure 5.8: Comparison of DT f-measure values per ELR class label of the RF-merged, RF1, RF2, RF3 and RF4 datasets.

Hourglass of Emotions Class Labels

Aptitude pertains to whether the reader hated (*low*) or loved (*high*) what happened in the segment. The RF2 and RF3 datasets are cases of the accuracy paradox, wherein RF2 predicted all instances as *high* and RF3 predicted all instances as *low*. This may be due to RF2 having more *high* instances than *low* and the inverse for RF3. It is interesting to note that males and females have no distinguishing difference in loving or hating a story, as opposed to people who prefer reading traditional books (RP1, RP2), people who always read for fun (RF1), and those who rarely read for fun (RF4).

Attention pertains to the reader's level of interest to the events in the segment. It means that the reader is either amazed at the event that happened (*low*) or already expected it to happen (*high*). RP4 predicted most of its instances as *high*, that is, only 6 instances as *low*. In terms of percentage, the Male dataset has more *high* instances than the Female and combined sex datasets. This is consistent with what the male participants have said during the debriefing, wherein they reported that they already have an inkling on what was going to happen. This is in line with Kret and De Gelder (2012), wherein they state the men respond better to threatening cues. Another observation worth noting is in the RF profiles. The scenario is that the RF1 and RF2 datasets in this class label are relatively balanced, while the RF3 predicted all instances as *high* and the RF4 predicted all instances as *low*. However, it is intriguing to see the decline in the classification performances (63.03 for RF1, 55.98 for RF2, 47.77 for RF3, and 26.95 for RF4). People who always or frequently read for fun are accustomed to keeping track of the events in the story, and this accustomedness was reflected in their brainwaves.

Pleasantness pertains to the reader's amusement towards the events in the segment, on whether they were sad (*low*) or happy (*high*) about it. The results show that the classifier for people who prefer reading eBooks over traditional books (RP4) yielded the best performance among all the PL DT classifications. Since the story segments were presented via a software program on a laptop (i.e. not a traditional book), it can be presumed that those who prefer reading traditional books have a certain bias clouding their enjoyment of the story. The said presumption is consistent because of the increase in performance of the f-measure values (35.63 for RP1, 42.29 for RP2, 50.92 for RP3, and 69.29 for RP4).

Sensitivity pertains to whether the reader feared (*low*) or was angered (*high*) with the events in the segment. The RP2 dataset classifier may have yielded the best performance but it may be because it was the only relatively balanced dataset among the RP profiles. With regard to the females, it is consistent with their debriefing reports wherein they expressed more anger or fear on certain elements of the story, such as the characters or the plot events, than the males.

Emotions of Literary Response Class Labels

Aesthetic feelings are caused by the formal (generic, narrative, or stylistic) components of the text in the segment. This is where the foregrounding, discussed by Miall and Kuiken (1994), can be found. Note that in all cases of the datasets, there is a significant percentage that marks the *false* presence of this class label. Only an average of 15% of the instances are marked as *true*. It is observed that there is a gradual increase in the Sex datasets (50.41 for Female and 53.25 for Male) and a gradual decrease in RP datasets (49.74 for RP1, 48.39 for RP2, 45.63 for RP3, and 44.75 for RP4). With regard to the RF datasets, RF4 yielded the worst performance because its classifier predicted all of its instances as *true* (which is approximately 15% of the total instances). Following what Miall and Kuiken (1994) presented, wherein it can be considered that their reading profile of having a high level of literary competence or not, it can be assumed that the response or emotion evoked towards formal components of the text is the same on other reading profiles (i.e. sex, reading preference, reading frequency). Note that further testing could be done to prove or disprove the assumption due to the imbalanced datasets and that these are the first impressions of the participants towards the story. Most of them reported the other two class ELR class labels as the cause of their emotions during their debriefings.

Evaluative feelings are concerned with the overall enjoyment, pleasure, or satisfaction in reading the segment. The Female dataset yielded the best performance in predicting this class label than the Male and combined sex datasets. This is consistent with the notion that women are more expressive than men (Kret & De Gelder, 2012).With regard to the RP datasets, it is interesting to note that those who prefer traditional books (RP1, RP2) have classifiers with a lower performance as opposed to those who prefer eBooks (RP3, RP4). It can be surmised that this is somewhat similar to the case of the RP datasets in predicting Pleasantness, wherein the displaying the story segments via a laptop screen presents an unconscious bias in the reader.

Narrative feelings are evoked by the events or characters in the segment. It is observed that the classifiers for the RP3, RP4, RF2, and RF4 datasets yielded significantly better performances than their respective general datasets. Further analysis and research into this case are needed as no inference or assumption comes to mind. On the other hand, the Female dataset gave the best performance than the Male and combined sex dataset. In the debriefing reports, both sexes have expressed that the plot or the characters are the main causes of their emotions. However, it is worth noting that females expounded more on why the plot or characters evoked such emotions. Hence, it is inferred that women's sympathy, empathy, or identification with the character and responses towards the plot events are registered more on their brainwaves rather than the men's, a notion consistent with the study of Kret and De Gelder (2012).

Tables 5.4 and 5.5 show a summary of the case-by-case analysis on the HoE and ELR DT classification results.

Table 5.4: Summary of the analysis on the HoE DT classification results.

<table>
<tr><th></th><th>Aptitude</th><th>Attention</th><th>Pleasantness</th><th>Sensitivity</th></tr>
<tr><td>Female</td><td rowspan="2">Men and women have no distinguishing difference in loving or hating a story.</td><td></td><td></td><td>It is consistent with their debriefing reports wherein they expressed more anger or fear on certain elements of the story, such as the characters or the plot events, than the males.</td></tr>
<tr><td>Male</td><td>Consistent with Kret and De Gelder (2012), in which men respond better to threatening cues.</td><td></td><td></td></tr>
<tr><td>RP1</td><td rowspan="2">Yielded better performance results than the combined RP dataset but need further research as to why.</td><td></td><td rowspan="4">People who prefer reading eBooks over traditional books (RP4) yielded the best performance among all the PL DT classifications. Since the story segments were presented via a software program on a laptop (i.e. not a traditional book), it can be presumed that those who prefer reading traditional books have a certain bias clouding their enjoyment of the story. The said presumption is consistent because of the increase in performance of the f-measure values.</td><td></td></tr>
<tr><td>RP2</td><td></td><td>The RP2 dataset classifier may have yielded the best performance, but it may be because it was the only relatively balanced dataset among the RP profiles.</td></tr>
<tr><td>RP3</td><td></td><td></td><td></td></tr>
<tr><td>RP4</td><td></td><td></td><td></td></tr>
<tr><td>RF1</td><td>Yielded better performance results than the combined RF dataset but need further research as to why.</td><td rowspan="4">Despite the results of RP3 and RP4 being accuracy paradoxes, it is interesting to note the decline in the classification performances. People who always or frequently read for fun are accustomed to keeping track of the events in the story, and this accustomedness was reflected in their brainwaves.</td><td></td><td></td></tr>
<tr><td>RF2</td><td></td><td></td><td></td></tr>
<tr><td>RP3</td><td></td><td></td><td></td></tr>
<tr><td>RF4</td><td>Yielded better performance results than the combined RF dataset but need further research as to why.</td><td></td><td></td></tr>
</table>

Table 5.5: Summary of the analysis on the ELR DT classification results.

<table>
<tr><th></th><th>Aesthetic Feelings</th><th>Evaluative Feelings</th><th>Narrative Feelings</th></tr>
<tr><td>Female</td><td rowspan="10">It can be assumed that the response or emotion evoked towards formal components of the text is the same on all the reading profiles. Further testing could be done to prove or disprove this assumption due to the imbalanced datasets and that these are the first impressions of the participants towards the story. Noted that most of the participants reported the other two ELR class labels as the cause of their emotions during their debriefings.</td><td>The Female dataset yielded the best performance in predicting this class label than the Male and combined sex datasets. This is consistent with the notion that women are more expressive than men. (Kret & De Gelder, 2012)</td><td>The Female dataset gave the best performance than the Male and combined sex dataset. It is inferred that women's sympathy, empathy, or identification with the character and responses towards the plot events are registered more on their brainwaves rather than men's, a notion consistent with the study of Kret and De Gelder (2012).</td></tr>
<tr><td>Male</td><td></td><td></td></tr>
<tr><td>RP1</td><td rowspan="4">Those who prefer traditional books (RP1, RP2) have classifiers with a lower performance as opposed to those who prefer eBooks (RP3, RP4). It can be surmised that displaying the story segments via a laptop screen presents an unconscious bias in the reader.</td><td></td></tr>
<tr><td>RP2</td><td></td></tr>
<tr><td>RP3</td><td rowspan="2">Yielded better performance results than the combined RP dataset but need further research as to why.</td></tr>
<tr><td>RP4</td></tr>
<tr><td>RF1</td><td></td><td></td></tr>
<tr><td>RF2</td><td></td><td>Yielded better performance results than the combined RF dataset but need further research as to why.</td></tr>
<tr><td>RP3</td><td></td><td></td></tr>
<tr><td>RF4</td><td></td><td>Yielded better performance results than the combined RF dataset but need further research as to why.</td></tr>
</table>

5.3.2 Improving Classification Performance with Support Vector Machines or Multilayer Perceptrons

The goal of this experiment is to attempt to improve the classification performance with SVM and MLP classifiers. The same conditions in the DT experiment were subjected to the SVM and MLP training, that is using all 252 EEG features with a participant-fold cross-validation. As stated previously, there is generally no significant improvement in profile-specific classification with the exception of certain cases. Also, the profile-specific classifications tend to fall into the accuracy paradox due to small number of instances. Thus, the combined datasets are the ones discussed below as the distribution of their instances is approximately ±30% from being a balanced dataset.

Figures 5.9, 5.10, 5.11, 5.12, 5.13, and 5.14 show the comparisons of the DT, SVM, and MLP f-measure values per class label on the Sex-merged, RP-merged, and RF-merged. It is observed that on average, SVM or MLP yields a slightly better performance in predicting class labels than DT, or at least on par with DT. Tables 5.6 and 5.7 show the summary of this comparison, wherein the ones with checkmarks yielded the best performance. Not considering the performance of the classifiers with accuracy paradoxes (distinguished by having a recall of 50% and kappa of 0), DT outperformed the other two 3 times, SVM outperformed the other two 4 times, and MLP outperformed the other two 5 times for HoE, as shown in Table 5.6; whereas DT outperformed the other two 1 time, SVM outperformed the other two 2 times, and MLP outperformed the other two 9 times for ELR, as shown in Table 5.7. Note that the SVM and MLP classifications were done with the default parameters RapidMiner has provided. Further tests by tweaking their respective parameters may yield higher results. Refer to Appendix G to see the summary of the performance results of this experiment.

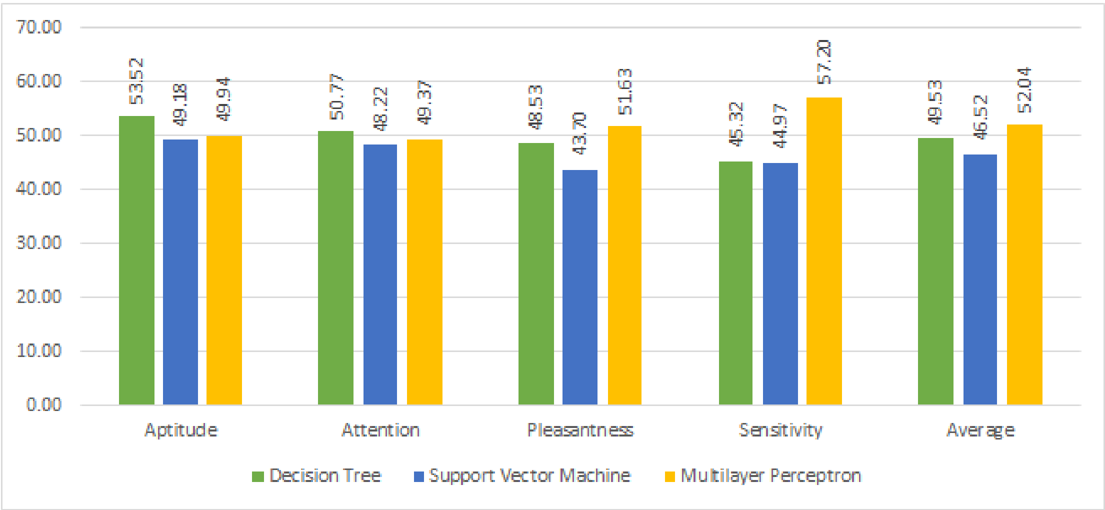


Figure 5.9: Comparison of DT, SVM, and MLP f-measure values per HoE class label of the Sex-merged dataset.

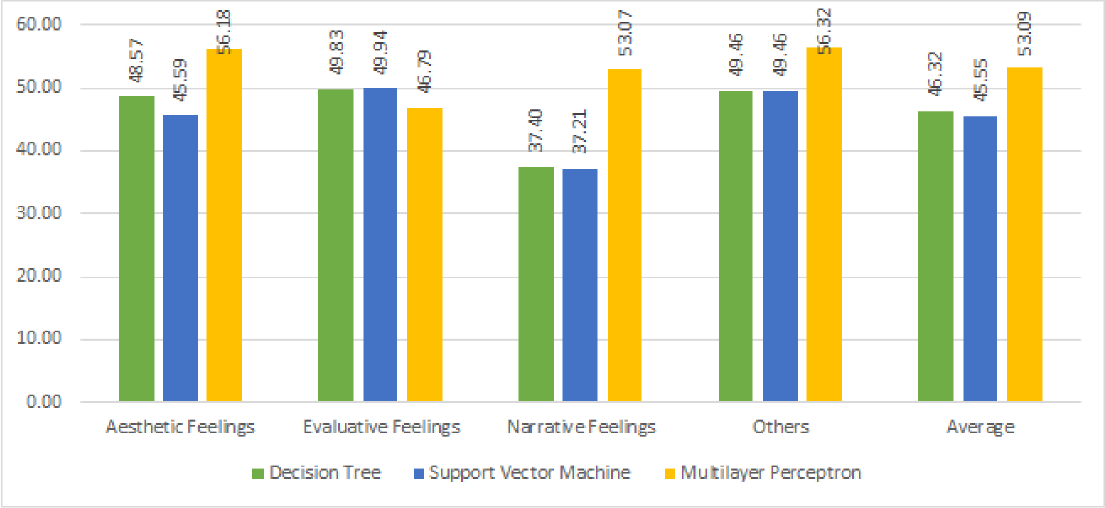


Figure 5.10: Comparison of DT, SVM, and MLP f-measure values per ELR class label of the Sex-merged dataset.

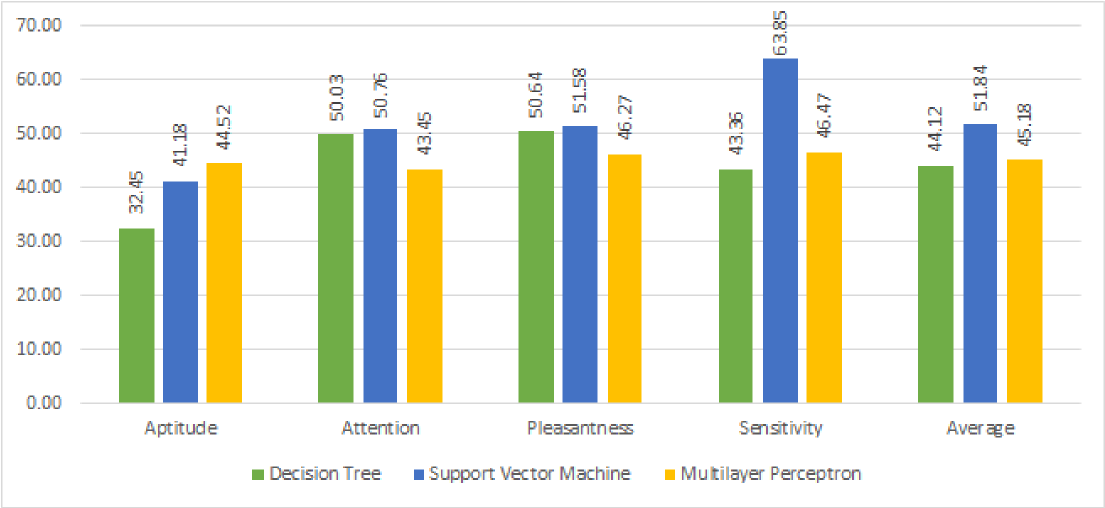


Figure 5.11: Comparison of DT, SVM, and MLP f-measure values per HoE class label of the RP-merged dataset.

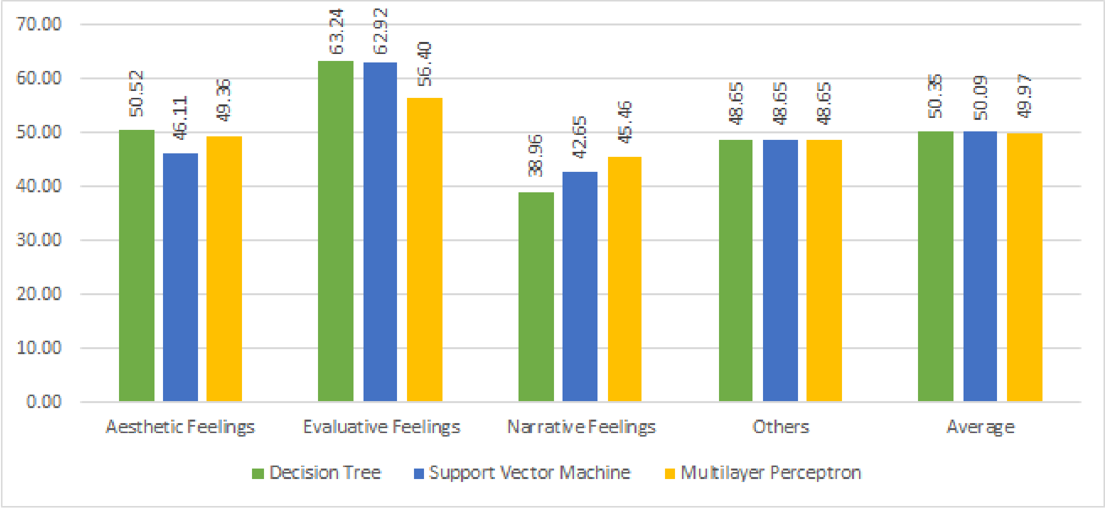


Figure 5.12: Comparison of DT, SVM, and MLP f-measure values per ELR class label of the RP-merged dataset.

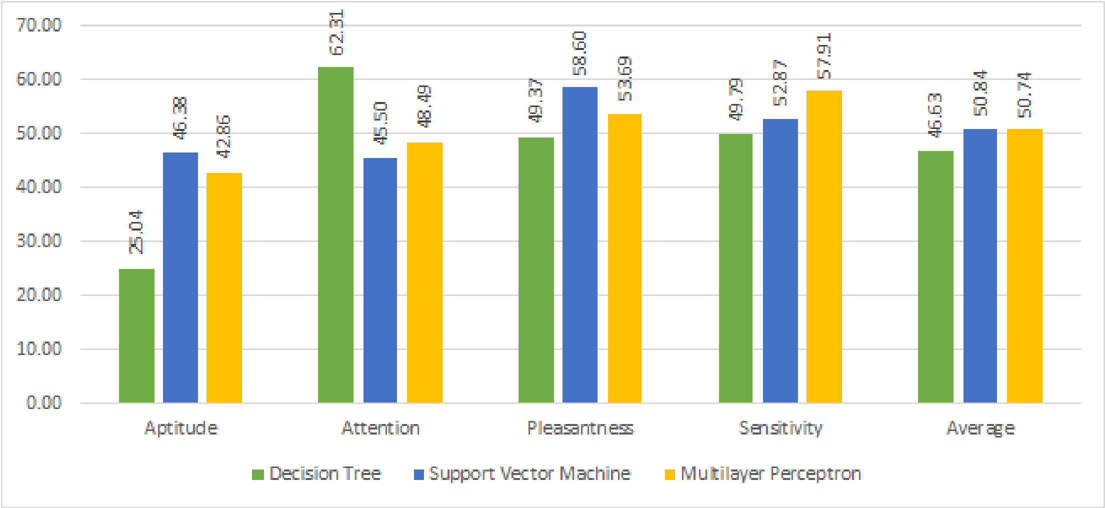


Figure 5.13: Comparison of DT, SVM, and MLP f-measure values per HoE class label of the RF-merged dataset.

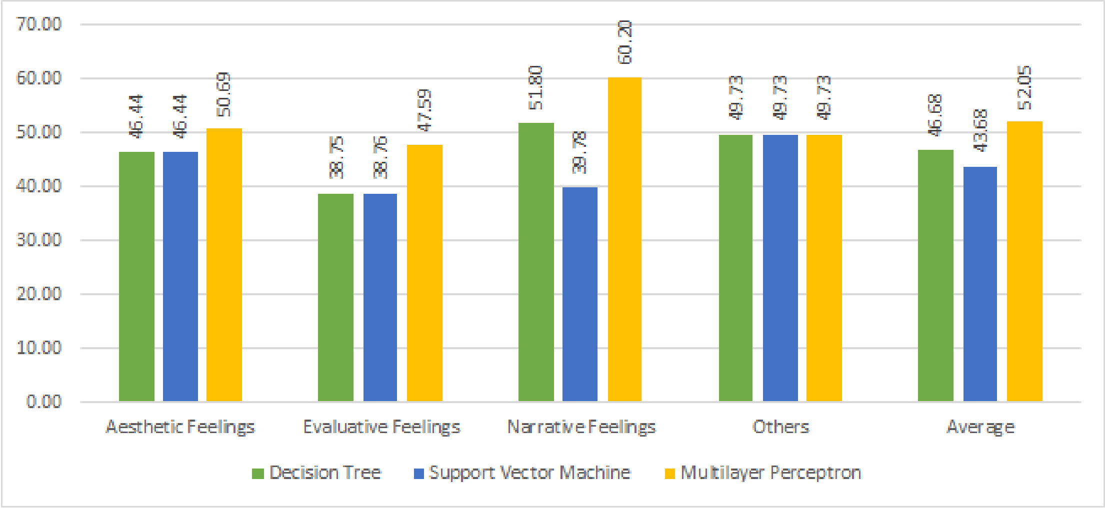


Figure 5.14: Comparison of DT, SVM, and MLP f-measure values per ELR class label of the RF-merged dataset.

Table 5.6: Summary of the DT, SVM, and MLP f-measure values comparison for the HoE class labels.

		DT	SVM	MLP
SEX	AP	✓		
	AT	✓		
	PL			✓
	SE			✓
RP	AP			✓
	AT		✓	
	PL		✓	
	SE			✓
RF	AP		✓	
	AT	✓		
	PL		✓	
	SE			✓
		3	**4**	**5**

Table 5.7: Summary of the DT, SVM, and MLP f-measure values comparison for the ELR class labels.

		DT	SVM	MLP
SEX	AE			✓
	EV		✓	
	NA			✓
	OT			✓
RP	AE	✓		
	EV		✓	
	NA			✓
	OT			✓
RF	AE			✓
	EV			✓
	NA			✓
	OT			✓
		1	**2**	**9**

5.3.3 Feature Selection with Principal Component Analysis

In this experiment, the goal is to see whether reducing the number of features could yield a result that is higher than or at least at par with the base feature set. The SVM and MLP classifiers are, again, trained with a participant-fold cross-validation. Like in the previous subsection, the combined datasets are the ones discussed below.

Figures 5.15, 5.16, 5.17, 5.18, 5.19, and 5.20 show the comparisons of the SVM (Base) to SVM (PCA) and the MLP (Base) to MLP (PCA) f-measure values per class label on the Sex-merged, RP-merged, and RF-merged. On an average basis, it is observed that the performance of the classifiers with PCA feature sets yields subpar results to that of its counterpart with the base set of 252 EEG features. The average difference in the f-measure value of the Base classifiers and PCA classifiers for both SVM and MLP is ±5.

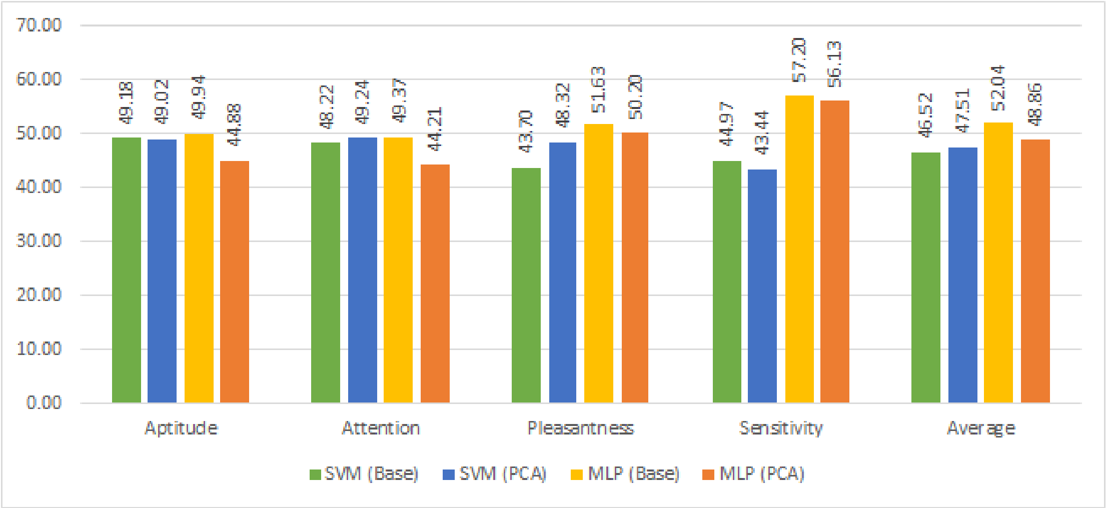


Figure 5.15: Comparison of SVM (Base) to SVM (PCA) and MLP (Base) to MLP (PCA) f-measure values per HoE class label of the Sex-merged dataset.

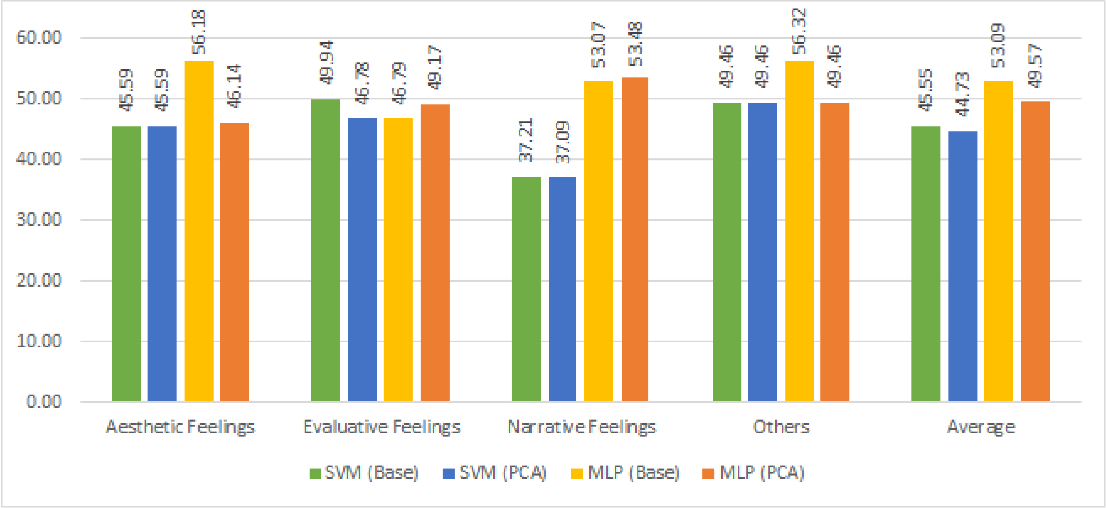


Figure 5.16: Comparison of SVM (Base) to SVM (PCA) and MLP (Base) to MLP (PCA) f-measure values per ELR class label of the Sex-merged dataset.

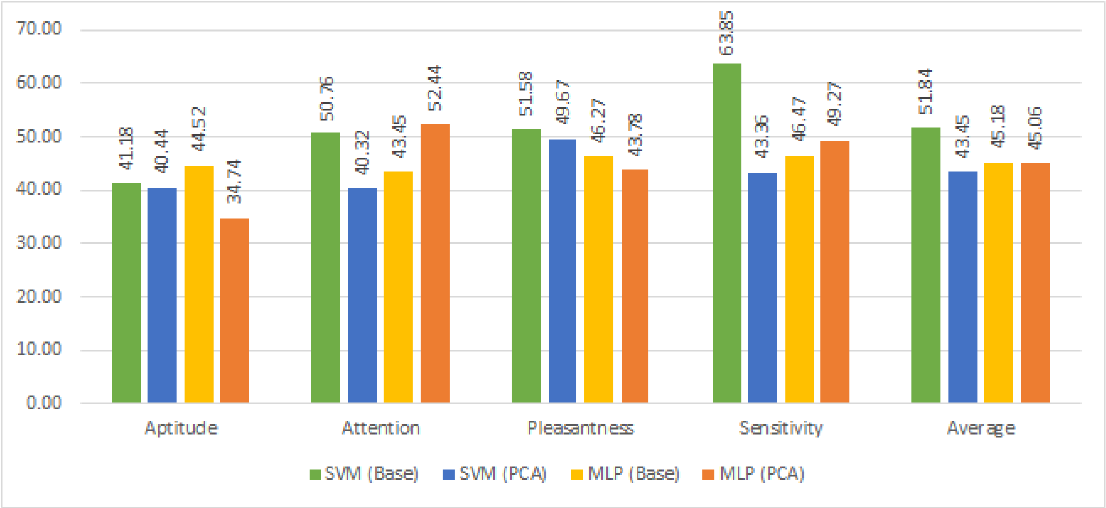


Figure 5.17: Comparison of SVM (Base) to SVM (PCA) and MLP (Base) to MLP (PCA) f-measure values per HoE class label of the RP-merged dataset.

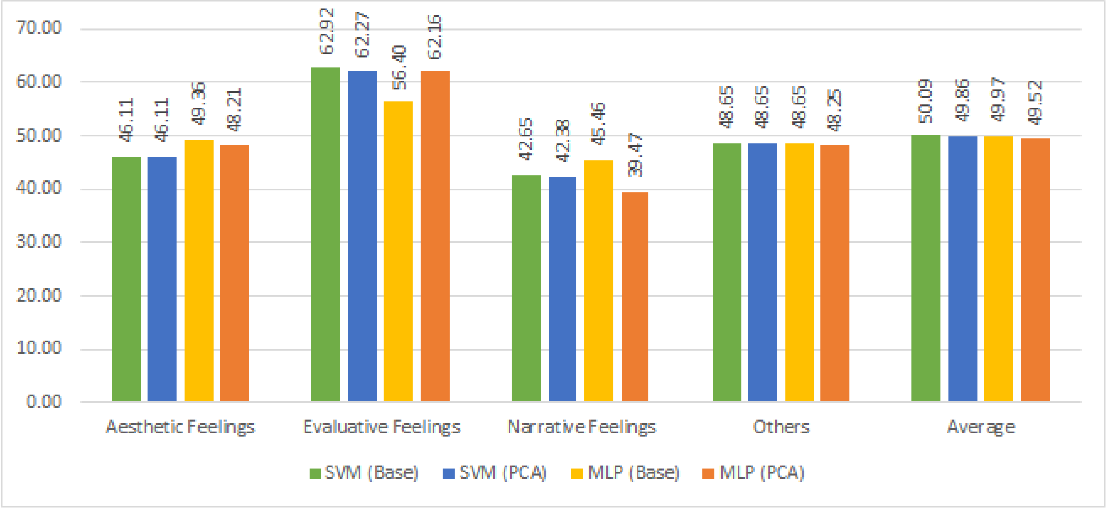


Figure 5.18: Comparison of SVM (Base) to SVM (PCA) and MLP (Base) to MLP (PCA) f-measure values per ELR class label of the RP-merged dataset.

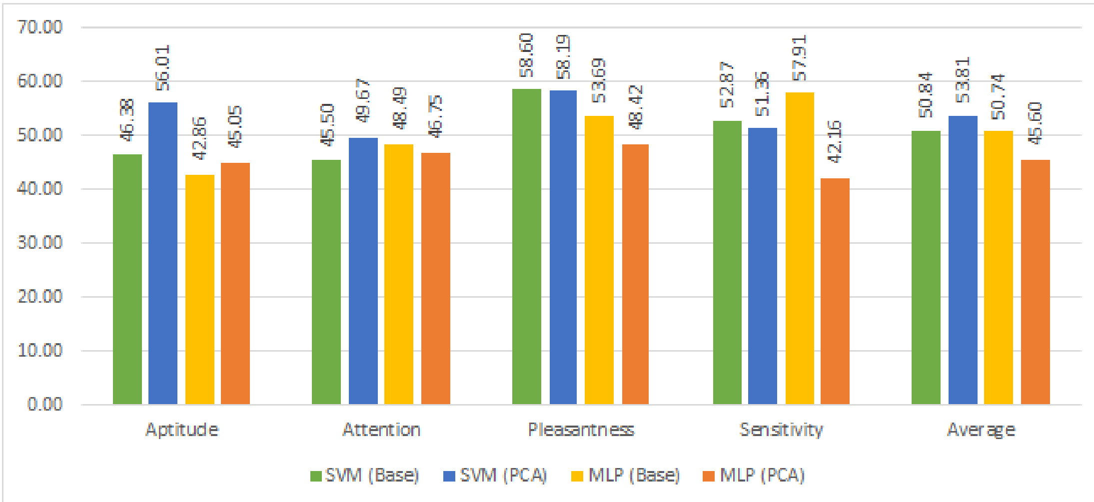


Figure 5.19: Comparison of SVM (Base) to SVM (PCA) and MLP (Base) to MLP (PCA) f-measure values per HoE class label of the RF-merged dataset.

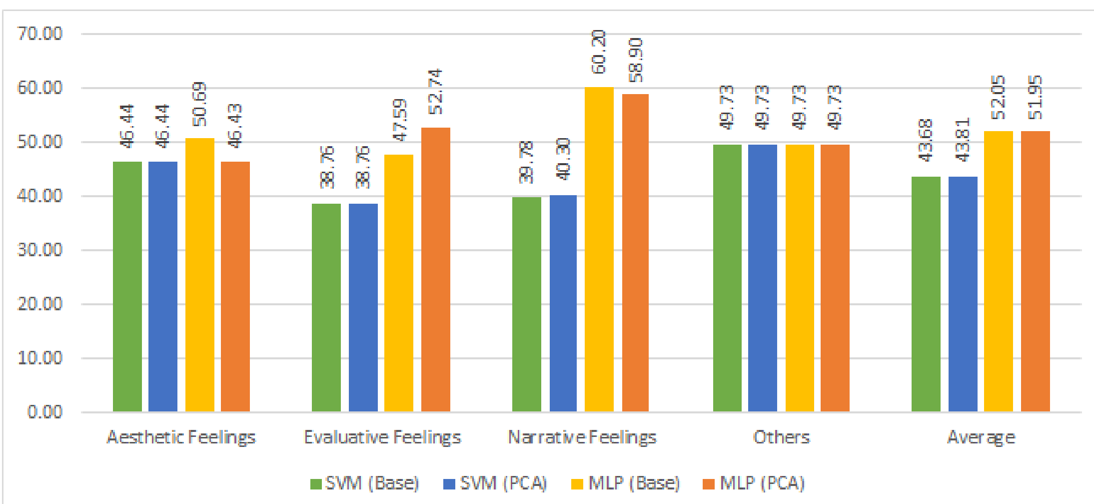


Figure 5.20: Comparison of SVM (Base) to SVM (PCA) and MLP (Base) to MLP (PCA) f-measure values per ELR class label of the RF-merged dataset.

Note that the processing of the base feature set on either MLP or SVM takes a lot of computer resources and processing time. The Sex-merged dataset, having approximately 45,000 instances, took a little more than 3 days to finish processing the MLP (Base) classifiers. Meanwhile, the MLP (PCA) classifier for that same dataset took approximately 10 hours to finish processing. Hence, if it is acceptable to have a ±5 margin of error, then using the PCA feature set would suffice as compensation for faster processing time. Like in the previous subsection, the SVM and MLP PCA classifications were done with the default parameters RapidMiner has provided. Higher performance results may be obtained by tweaking these parameters. Refer to Appendix H to see the summary of the performance results of this experiment.

5.4 Participant Feedback

This section discusses the consolidated responses from the participants during debriefing regarding their experiences with the experiment. The responses cover their thoughts on their reading experience, their emotional responses, and on software usage.

5.4.1 Reading Experience and Emotional Response

A limitation of the study is that it does not take into account the emotions evoked due to external factors. An example of this is the environment. Though most have stated that they were comfortable with the environment, there are others that said that: (a) because of the silence, they are either more focused or not used to it; (b) the temperature is fine for some, while too cold for others; (c) sitting is not their preferred position while reading; (d) they are not used to reading in front of a screen; and (e) the fact that they know that they are being recorded made them feel pressured. One participant even noted that the dysmenorrhea she was experiencing during her session might have skewed her data. Around half of the participants said that they are comfortable when they are reading while lying down.

A common complaint among all the participants is the slight discomfort in wearing the headset. They noted that the longer they wore the headset, the more they felt its two front nodes pressing on their foreheads. For some, they became used to the discomfort, while for others, it became a slight or severe distraction.

5.4.2 Software Usage

Their response on the usage of the tool was generally positive, that it was fairly easy and straightforward to use although there were some suggestions on user interface improvement. Some noted that the segmentation of the story seemed geared towards creating suspense. Nevertheless, each segment was coherent and sufficient by itself. The annotation of the emotions became tedious later on. They said that tagging emotions was not a problem for the minor or transitional segments but then became a bit troublesome on the high intensity parts such as the climax. This was because they wanted to read the next part but they had to annotate first before proceeding to the next segment. However, there are those that have expressed that the read-annotate cycle process is preferable than reading the whole story and then annotate everything at the end.

Another limitation of this research is the accuracy of the ground truth. As this research used the participants' self-report as the ground truth, this meant that if any instruction on the evaluation of their emotions was unclear to them, then it adds more inaccuracies to obtaining the ground truth. It is interesting that one participant noted that it felt as if they were evaluating what they should feel rather than what they actually felt.

Chapter 6

Conclusions and Recommendations

6.1 Conclusions

The current work presents a study on brainwaves or EEG signals and their association to emotions while a person is reading literary fiction. This is done by building an affect model that maps the brainwaves collected from readers as they are reading short stories to specific emotions. This general objective was achieved as discussed subsequently.

In addition to related literature of existing EEG-based affect detection studies, literature on emotions, on EEG, and on reading fiction were also reviewed. This is to determine and design the methodologies and experiments conducted in this research. For example, the data acquisition methodology used in this work is an amalgamation of methodologies from affect recognition and reader-response studies.

During the theoretical literature review, two emotion models were encountered that fit the objectives of this research. The *Hourglass of Emotions* model by Cambria et al. (2012) is a model that combines the categorical and dimensional emotion models. In this way, it can potentially describe a full range of emotions through different levels of four independent but related affective dimensions. These dimensions are *aptitude* (hate/love), *attention* (surprise/anticipation), *pleasantness* (sad/happy), and *sensitivity* (fear/anger).

If the HoE model answers what the participant *feels*, the other model answers what *causes* the evoked emotions, that is, which elements of the story induced the emotions. The use of the *Emotions of Literary Response* model by Miall and Kuiken (2002) allows the classifier to determine if the participant's emotions are caused by the formal components of a text (*aesthetic feelings*), their overall enjoyment, pleasure, or satisfaction in reading the text (*evaluative feelings*), or evoked by the events or characters in the text (*narrative feelings*). Referring to Figures 5.1 and 5.2, the former shows what the participants are feeling per segment, whereas the latter shows what caused the emotion per segment.

EEG data from 32 participants were collected while they were reading *The Veldt*, a short story by Ray Bradbury. These EEG signals were collected with the use of an Emotiv Insight EEG headset, attached to the head of the participant while reading the story segments presented via the developed data collector tool. After which, features were extracted and different datasets were built according to sex, reading preference, and reading frequency profiles.

Based on the reviewed literature, *Support Vector Machines* and *Multilayer Perceptrons* were the machine learning algorithms employed. In addition to SVM and MLP, *Decision Trees* were also used to establish baseline performance results due to its relatively fast processing time. *Principal Component Analysis* was an approach used for feature selection. The experiments conducted in this research were binary classifications of the HoE and ELR models on different datasets and machine learning algorithms. All these experiments were trained with a participant-fold cross-validation scheme.

Lastly, these different classifiers were evaluated with accuracy, precision, recall, f-measure, and Cohen's kappa. The main metric used for comparison is the f-measure. This is because of several cases of the accuracy paradox wherein there is a high accuracy despite having a significant amount of false positive or false negatives. An accuracy paradox case can also be distinguished with a recall of 50% and kappa value of 0.

From the experiments conducted, the following are the analyses, inferences, and assumptions deduced:

- According to the DT models, all the relevant features are from the alpha α and theta θ bands only. This is quite interesting to note because the mental activity concerned with reading is in line with the accompanying descriptions of these two frequency bands.

- Generally, there is no significant improvement in the performance between the combined datasets and its corresponding profile-specific datasets. However, in reviewing the results on a case-by-case basis, it is observed that they are consistent and in line with the debriefing interviews and other reviewed studies.

- Generally, SVM or MLP yielded better performance results than DT by a small degree. Note that these classifiers were implemented with the default parameters RapidMiner has already provided. So, there is a possibility of increasing the performance results by adjusting the parameters.

- Classification experiments with feature selection via PCA yielded comparable performance results with an average ±5 margin of error. If such a margin is acceptable, then using the PCA feature set would suffice as compensation for faster processing time. Like the previous item, there is a possibility of increasing the performance results since only the default parameters were used.

6.2 Recommendations

The collected data are one of the most relevant components in researches such as this. This is the input for the experiments and tests, from which results and new findings are generated. Hence, here are recommendations for further work on data collection, processing, dataset building, and methodology:

- Improvements in the data collection process is advisable because the current acquisition process introduces added distractions that may have affected the gathered EEG data. It can be something as simple as tweaking the user interface to make it more intuitive and easier to use. A data acquisition process may be designed in such a way that what the participant would be doing is as near as to what is naturally possible. Refer to Section 5.4.2 to see the feedback with regard to the current experiment design and then future work can make adjustments accordingly.

- Other types of profiles could be considered such as personality types or degrees of literary competence. In the current work, specific profiles did not outperform the general classifiers but interesting inferences could still be obtained on a case-by-case basis.

- The usage of more feature types and/or more EEG channels could be considered. The current work only uses AF3, AF4, T7, T8, and Pz channels. The other channels may contain relevant features in these kinds of classification tasks. From the results, Decision Trees deem that features from alpha α and theta θ frequency bands are relevant. The next step could be to determine from which of, for example, the standard 32 EEG channels are relevant. This is to minimize sensors being placed as well as already optimizing the feature set.

- The current classification models suffered from the accuracy paradox. It is recommended to have more participants and apply techniques to balance the datasets such as oversampling the label with the least amount of instances.

- Applying the current methodology on other stories or different stimuli may yield different and notable results. Certain stories can evoke only certain kinds of emotions. For example, a suspense or thriller story is geared towards the extreme levels of attention, whereas a horror story is geared towards low sensitivity. This research could work in tandem with the 6 core emotional arcs of stories that Reagan et al. (2016) proposed. The same methodology may be repeated except that the stories presented subscribe to one of the 6 core emotional arcs. Given this, the participant would read, at least, 6 different stories. The experiment would then become an *intra-subject* classification, instead of the *inter-subject* classification that the current work does. This intra-subject classification works by creating datasets from the data for each story and a general dataset of the combined data from all the stories from one participant. From these datasets, the same classification experiments may be conducted.

- On the other hand, applying the current methodology with stimuli of another genre could also be done. Poems or news articles can be read in one sitting, wherein the latter could be considered as a story told in a factual narrative. These two present different writing styles, which are aligned with aesthetic feelings.

- The current work makes use of the first impressions of the participants towards the story. Following what Tompkins (1980) said that reading is an experience that is never the same from one reading to the next, this could be tested by having the participant re-do the data acquisition process for a number of times. In this way, the fourth domain in the Emotions of Literary Response model, *self-modifying feelings* which involves the restructuring of the reader's understanding of the textual narrative, could potentially be mapped. The trajectory in the change of emotions for the same stimuli could be observed.

It was mentioned in the previous chapter that there are still possibilities for improving the performance of the classification. Below are future work that can be done on classification and feature selection:

- Default parameters provided by RapidMiner were utilized in the classification experiments. Further testing with adjusting the parameters is recommended.
- Other feature sets that can be used are the relevant features output by the DT models. Likewise, feature selection by k-Means clustering could also be employed. The current work made use of all 252 features. These 252 are composed of MSP, PSD, RASM, and DASM. Selecting only a specific feature type (similar to the experiment conducted by Lin et al. (2010)) or a combination of these can be done.
- Only DT, SVM, and MLP were the machine learning algorithms used in the experiments. Tests exploring other supervised and unsupervised machine learning algorithms are recommended.

Lastly, analyzing and interpreting why the classification experiments yielded particular results are just as important. In addition, recommendations for the use of the findings in this research are listed as well:

- Generally, this is an interdisciplinary research. Despite the researcher having read various literature on the subject matter, expertise on certain aspects of the research cannot be claimed. Hence, consultation with experts on reader-response and EEG is recommended.
- Means of visualizing and showing the trajectory may lead to the discovery of notable patterns that other affect-aware systems can utilize. Another visualization that can be done is extending Figures 5.1 and 5.2 and make them on an individual basis or profile-specific basis, and then draw comparisons from them. This visualization can also work with the emotional arcs discussed by Reagan et al. (2016). For example, a story with a particular emotional arc looks like this for females and like that for males.
- Embodied conversational agents (ECA) can use the results of the models as basis for the topic of its conversation with the reader. For example, the

model has identified the cause of the emotion (e.g. Emotions of Literary Response by Miall and Kuiken (2002)). Given this, the ECA can converse with the reader in-depth regarding the evoked emotions. The ELR by Oatley (1995) can come into play here, wherein the ECA can specifically ask if the emotion is due to sympathy or empathy to the character or if it is due to emotion memories.

Appendix A

Research Ethics Forms and Checklists

RESEARCH ETHICS CLEARANCE FORM[1] **For Thesis Proposals**
Names of Student Researcher(s): Kristine Ma. Dominique F. KALAW
College: College of Computer Studies
Department: Software Technology
Course: Master of Science in Computer Science
Expected Duration of the Project: from: January 2016 to: April 2017
Ethical considerations *(The Ethics Checklists may be used as guides in determining areas for ethical concern/consideration)*
To the best of my knowledge, the ethical issues listed above have been addressed in the research. Ethel Chua Joy ONG ______________________________ **Name and Signature of Adviser/Mentor:** **Date:**
Noted by: ______________________________ **Name and Signature of the Department Chairperson:** Dr. Rafael A. CABREDO **Date:**

[1] The same form can be used for the reports of completed projects. The appropriate heading need only be used.

DE LA SALLE UNIVERSITY
General Research Ethics Checklist

This checklist is to ensure that the research conducted by the faculty members and students of De La Salle University is carried out according to the guiding principles outlined in the Code of Research Ethics of the University. The investigator is advised to refer to the De La Salle University Code of Research Ethics and Guide to Responsible Conduct of Research before completing this checklist. Statements pertinent to ethical issues in research should be addressed below. The checklist will help the researcher/s and advisers/readers/evaluators determine whether procedures should be undertaken during the course of the research to maintain ethical standards. The University's Guide to the Responsible Conduct of Research provides details on these appropriate procedures.

Faculty/ASF Researcher Details	
Principal Investigator	Kristine Ma. Dominique F. KALAW
Department	College of Computer Studies Software Technology Department
Proposed Title of the Research	Recognizing Reader's Affect Using EEG Data
Term(s) and academic year in which research is to be conducted	AY2015-2016, T2-T3; AY2016-2017, T1-T3
Other researchers involved in project including their positions (e.g., student, faculty)	

Student Researcher Details (for students who are co-proponents)	
Course Title	N/A
Department	N/A
Thesis Adviser	Ethel Chua Joy ONG
Email Address	ethel.ong@delasalle.ph

This checklist must be completed AFTER the **De La Salle University Code of Ethics** ***has been read and BEFORE gathering data.***

Questions	Yes	No
1. Does your research involve human participants (this includes new data gathered or using pre-existing data)? If your answer is **yes**, please answer **Checklist A (Human Participants).**	✓	
2. Does your research involve animals (non-human subjects)? If your answer is **yes**, please answer **Checklist B (Animal Subjects).**		✓
3. Does your research involve Wildlife? If your answer is **yes**, please answer **Checklist C (Wildlife)**.		✓
4. Does your research involve microorganisms that are infectious, disease causing or harmful to health? If your answer is **yes**, please answer **Checklist D (Infectious Agents).**		✓
5. Does your research involve toxic/chemicals/ substances/materials? If your answer is **yes**, please answer **Checklist E (Toxic Agents).**		✓

Research with Ethical Issues to address:

If you have a YES answer to any of the above categories, you will be required to complete a detailed checklist for that particular category. A YES answer does not mean the disapproval of your research proposal. By providing you with a more detailed checklist, we ensure that the ethical concerns are identified so these can be addressed in adherence to the University Code of Ethics.

Declaration of Conflict of Interest

[✓] 1. I do not have a conflict of interest in any form (personal, financial, proprietary, or professional) with the sponsor/grant-giving organization, the study, the co-investigators/personnel, or the site.

[ ] 2. I do have a conflict of interest, specifically:

[ ] A. I have a personal/family or professional interest in the results of the study (family members who are co-proponents or personnel in the study, membership in relevant professional associations/organizations).

Please describe the personal/family or professional interest:

____

____

____

[ ] B. I have propriety interest vested in this proposal (with the intent to apply for a patent, trademark, copyright, or license)

Please describe propriety interest:

____

____

____

[ ] C. I have significant financial interest vested in this proposal (remuneration that exceeds P250,000.00 each year or equity interest in the form of stock, stock options or other ownership interests).

Please describe financial interest:

____

____

____

Declaration

I certify that I have read and understood the De La Salle University Code for the Responsible Conduct of Research and will abide by the ethical principles in this document. To the best of my knowledge that my research proposal does not involve any of the above-mentioned categories. I will submit a final report of the proposed study to the DLSU-Research Ethics Office. I will not commence with data collection until I receive an ethics review approval from the College Research Ethics Committee.	
Kristine Ma. Dominique F. KALAW ______________________________ **Name and Signature of Principal Investigator**	____________________ **Date**

FOR GRADUATE and UNDERGRADUATE DLSU STUDENTS ONLY		
I confirm that the student(s) is/are capable of undertaking this research in a safe and ethical manner.		
Ethel Chua Joy ONG		
Adviser's Name	Signature	Date

DE LA SALLE UNIVERSITY

Checklist A
Research Ethics Checklist for Investigations involving Human Participants

This checklist must be completed AFTER the* De La Salle University Code of Research Ethics and Guide to Responsible Conduct of Research *has been read and BEFORE gathering data. The University Code of Research Ethics is available at http://www.dlsu.edu.ph/offices/urco/forms/URCO-Code-of-Research-Ethics_August2011.pdf

NOTE: This checklist is completed after the research proponent fills out the General Checklist Form.

Only answer this Checklist if you answered YES on question 1 of the General Checklist.

Researcher Details	
Lead Researcher's Signature	
Lead Researcher's Name (Please Print)	Kristine Ma. Dominique F. KALAW
Email Address(es)	kristine_ma_kalaw@dlsu.edu.ph
Department/College	College of Computer Studies Software Technology Department
Proposed Title of the Research	Recognizing Reader's Affect Using EEG Data
Term(s) and academic year in which research project is to be undertaken	AY2015-2016, T2-T3; AY2016-2017, T1-T3
Other faculty members involved in project and their department affiliation(s)	Ethel Chua Joy ONG Thesis Adviser AdRIC Director

Provide a brief description of the data collection procedure to be undertaken in the research:
The participants will wear an EEG headset while they read a short story that is divided into segments. Before proceeding to the next segment, they will have to annotate the emotions they felt for the segment. An interview is also conducted after the reading session.

The following should be attached to the checklist:
• A copy of the informed consent form to be used in the study. • A copy of the instrument/tool that will be administered to the participants. • If applicable, a copy of the letter seeking permission to collect data from participants who are under the supervision of an agency, institution, department, or office. • If applicable, a copy of the parental consent form for participants below 18 years old.

The following items refer to important ethical considerations in the conduct of research with human participants. Provide a check for the appropriate answer to each question.

Source of data *Please check all that apply:*		
	1. New data will be collected from human participants If you checked this item, how will the new data be gathered? Please check all that apply. **After answering this question, please proceed to page 3**	
	✓	**Experimental Procedures/Intervention/ Treatments**
		Focus Group
	✓	**Personal Interviews**
	✓	**Self-administered Questionnaire**
		Researcher-administered Questionnaire
		Internet survey
	✓	**Observation**
		Telephone survey
	✓	**Others, please specify:** Video recording, audio recording
	2. Pre-existing data from human participants, i.e., from a dataset **If you checked this item, please proceed to page 7**	

If both options are checked (both new data and pre-existing data), **answer all of the questions in this document.**

Only answer if new data will be collected (item 1 above)	
Sampling Details	
Number of Participants/Subjects	at least 30 participants
Location where the participants will be recruited/ where subjects will be obtained?	De La Salle University
How long will the data collection take place?	Jan 2017 - Feb 2017
Who will perform the data collection?	Kristine Ma. Dominique F. KALAW
Location(s) where data collection will take place	De La Salle University
What procedures will be employed to ensure voluntary consent from participants?	Meals will be provided.
Data Retention	
How long will data with participant identifiers be kept after the publication of the first paper from the project?	Participants are given the option to include or have their name be anonymous as indicated in the attached informed consent form.
How long will anonymized data be kept after the publication of the first paper from the project?	Participants are given the option to have their data only be used for this specific research or also be used for future studies as indicated in the attached informed consent form.
Procedure for Informed Consent	
How will informed consent be recorded? (check all that applies) Reminder: please attach informed consent that will be used in the study	[✓] Written Consent [ ] Audio-recorded Consent [ ] Online/Email recorded Consent [ ] Others, please specify:

If you will not obtain a recorded informed consent, answer the questions that follow:

Why does the waiver of informed consent not pose a threat to the welfare and rights of the participants?

Why is recording an informed consent not practical for the proposed study?

	Yes	No	Not Applicable
1. Will the research involve students who will be receiving course credits for their participation? **If YES, please attach a copy of the consent form and a summary of the debriefing process that will help participants understand how their participation in the research has provided a relevant learning experience to the crediting course.**	☐	☑	☐
2. Does the study involve participants below 18 years old or those who are unable to give their informed consent? **If YES, please attach a copy of the parental consent form.**	☐	☑	☐
3. Is there a possibility that the research can induce physical and/or psychological harm to the participants? Will they experience pain or some discomfort as a result from their participation in the research? **If YES, please attach an acceptable argument that outlines the benefits of doing the research and how they outweigh the cost of harming the participants.**	☐	☑	☐
4. Will the participants be deliberately falsely informed or made unaware that they are being observed? Will they be misled in a way that they will possibly object to or show unease when told of the real purpose of the study? **If YES, please attach an acceptable argument that outlines the benefits of doing the research and how they outweigh the cost of harming the participants.**	☐	☑	☐
5. Will the research involve the discussion of, or questions on, sensitive topics (e.g. sexual activity, substance abuse, or mental health)? **If YES, please make sure that the informed consent form explicitly states that sensitive questions will be posed and that you will safeguard the anonymity of the participants and ensure confidentiality. Please attach a copy of your informed consent form and your instrument.**	☐	☑	☐

	Yes	No	Not Applicable
6. Will the research involve the administration of drugs, or other substances to the participants? **If YES, please attach an acceptable argument that outlines the benefits of doing the research and how they outweigh the cost of harming the participants.** **Please also attach a description of the procedure that will ensure that the participants will be brought back to their physical and psychological states prior to their participation in the research.**	☐	☑	☐
7. Will biological samples (e.g. blood, saliva, urine) be obtained from the participants? **If YES, will this involve invasive procedures? Please attach a description of these procedures.**	☐	☑	☐
8. Will genetic materials be obtained from the biological samples? **If YES, please attach a description of the procedures that will ensure confidentiality. Please attach the informed consent form.**	☐	☑	☐
9. Will financial inducements (other than reasonable expenses, like transportation or meal allowances) be offered to the participants for their participation in their research? **If YES, the researcher(s) should be mindful of how the inducements can influence the participants' responses or behaviors during the research. Indicate the financial inducements offered to the participants:** ____________________________________	☐	☑	☐
10. Is there a possibility for groups or communities to be harmed by the dissemination of the research findings? **If YES, please attach a description of procedures to ensure the anonymity and confidentiality of the research findings.**	☐	☑	☐

Answering <u>YES</u> to most of the above items will signal an ethical issue that needs to be addressed. Some actions that will allow adherence to research ethical principles are provided with each item. The researcher is advised to refer to the University's Guide to the Responsible Conduct of Research for the appropriate procedures to ensure adherence to ethical principles in the conduct of research.

<u>Declaration</u>

I certify that I have read and understood the De La Salle University Code for the Responsible Conduct of Research and will abide by the ethical principles in this document. I will submit a final report of the proposed study to the DLSU-Research Ethics Office. I will not commence with data collection until I receive an ethics review approval from the College Research Ethics Committee.

Kristine Ma. Dominique F. KALAW **Name and Signature of Principal Investigator**	**Date**

FOR GRADUATE and UNDERGRADUATE DLSU STUDENTS ONLY		
I confirm that the student(s) is/are capable of undertaking this research in a safe and ethical manner.		
Ethel Chua Joy ONG Adviser's Name	Signature	Date

FOR PROPONENTS WHO WILL GATHER NEW DATA ONLY, PLEASE STOP ANSWERING.

Use of Pre-existing Data collected from Human Participants		
Indicate the dataset from which the data for the study will be sourced		
Is the data publicly available, i.e., the access to which does not necessitate an approval process?		Yes Please indicate where the dataset is available:
		No Please indicate/attach the approval authority for access:
Was the original dataset originally collected for the present study's purpose?		Yes **Please attach the Consent Form used in the original study.**
		No **Please attach the Information Collection Statement (i.e., the statement given to informants providing them with the rationale for the collection of specific information).**
Does the original data set contain sensitive data, that is information that an individual would not likely want to be disclosed publicly, e.g., data on sexual activities, substance use?		Yes **Please describe the type of sensitive data to be used in the present research:**
		No
Does the original dataset have personal identifiers?		No *(This means that neither the researcher nor the participant provided any personal identifiers)*
		Yes, specifically: ______ Direct *(i.e., the participant provided personal details like name and address)* _____ Indirect (*i.e., the participant was given a respondent code to make the participant identifiable)*
Will new data be collected and analyzed along with data from the existing dataset?		Yes **Please answer questions on page 3-5.**
		No

Declaration

I certify that I have read and understood the De La Salle University Code for the Responsible Conduct of Research and will abide by the ethical principles in this document. I will submit a final report of the proposed study to the DLSU-Research Ethics Office. I will not commence with data collection until I receive an ethics review approval from the College Research Ethics Committee.

_______________________________ **Name and Signature of Principal Investigator**	_____________ **Date**

FOR GRADUATE and UNDERGRADUATE DLSU STUDENTS ONLY
I confirm that the student(s) is/are capable of undertaking this research in a safe and ethical manner.
Adviser's Name   Signature   Date

Appendix B

Informed Consent Form

De La Salle University
College of Computer Studies

RESEARCH INFORMATION SHEET (VERSION-DECEMBER 9, 2016)

You are being asked to participate in a research entitled,

RECOGNIZING READER'S AFFECT USING EEG DATA

You must be 18 years or older to participate in this study. Your participation is voluntary. Please carefully read the information below and do not hesitate to ask any questions regarding the experiment that may not be clear to you.

This study is conducted by Kristine Ma. Dominique F. Kalaw, supervised by Prof. Ethel Chua Joy Ong, as part of her work towards a Master's degree in Computer Science at De La Salle University.

A. INTRODUCTION/PURPOSE

This study aims to associate brainwave patterns to specific emotions while reading literary fiction.

B. PROCEDURE

You are asked to read at least two (2) short stories. One session involves reading one story in a single sitting. The second session may not necessarily be immediately after the first. Each session will last approximately 1 hour and 30 minutes. If you agree to participate in this research, this would imply the following:

1. During the session, you will wear an Emotiv Insight headset to measure and capture your brain activity. While wearing the headset, hair ornaments or accessories should be removed.
2. A camera will be used to record the whole session.
3. Prior to reading the story, a baseline of your brain activity will be recorded. You are asked to close your eyes and be as relaxed and as comfortable as possible for a period of two (2) minutes.
4. You will be asked to read the pre-selected short story, which is presented in segments, using the software developed by the researcher. Please note that you may not go back to previous segments. Only the immediate previous segment of the current segment is displayed for your reference.
5. For each current segment, you will identify the degree of *pleasantness*, *attention*, *sensitivity*, and *aptitude* it causes you. You will also indicate if whether the segment has struck you or caught your attention. Lastly, you will indicate whether the emotion you are experiencing is an evaluative feeling, narrative feeling, aesthetic feeling, or others. All this is done using the software developed by the researcher.
6. Steps 4 and 5 are repeated until you have finished the story.
7. After, you will answer a participant profile questionnaire as well as have a short interview about the experience. Both of these are used as part of the data collected from you. Please note that the short interview will be voice-recorded.

C. DATA ANNOTATION

I. What did you feel towards the segment?

- **Pleasantness** refers to your amusement to the segment. It ranges from grief to ecstasy.
- **Attention** refers to your interest in the segment. It ranges from amazement (negative surprise) to vigilance (positive surprise).
- **Sensitivity** refers to your comfortability towards the segment. It ranges from terror to rage.
- **Aptitude** refers to your confidence (trust) in the segment. It ranges from loathing to admiration.

II. What triggered your feelings?

- **Evaluative feelings** toward the text, such as the overall enjoyment, pleasure, or satisfaction of reading a short story.
- **Narrative feelings** toward specific aspects of the fictional event sequence, such as empathy with a character or resonance with the mood of a setting.
- **Aesthetic feelings** in response to the formal (generic, narrative, or stylistic) components of a text, such as being struck by a metaphor.

D. POTENTIAL RISKS AND DISCOMFORT

The Emotiv Insight is a 5-channel, wireless EEG headset that records your brainwaves. It is a commercial product marketed worldwide and is designed for everyday use. It uses a polymer sensor that is safe to use and offers great electrical conductivity with the convenience of a dry sensor. These sensors read the brainwave signals and then transmits these signals to a computer via Bluetooth.

This device has global recognition for its personal usage such as assessing athletic performance, cognitive training, or health and well-being (*Source: http://emotiv.com/insight/*). The technology is also backed and trusted by the scientific, academic, engineering and media communities and has been validated by many independent research papers (*Source: http://emotiv.com/the-science/*).

EEG or brainwave recording procedures are quite safe and has been in use for over 30 years; it is used routinely in hospitals to test brain function and to diagnose illness such as temporal epilepsy. There are no known major risks associated with this procedure other than a mild discomfort for some people who have sensitive skin when wearing the headset. This is not permanent and is of no serious consequence (*Source: https://emotiv.zendesk.com/hc/en-us/articles/204701495-EEG-Basic-Participant-Information-and-Safety*). However, if you find the headset uncomfortable to use or if you decide to stop the session for any other reason, the experiment will be halted immediately.

E. POTENTIAL BENEFIT TO SUBJECTS AND/OR TO SOCIETY

Apart from some personal insights with regards to your awareness of your emotions during the reading process, you will not directly benefit from your participation in this research study.

Since the study will attempt to associate brainwave patterns to specific emotions while reading literary fiction, the results of this study would provide a useful baseline data for studies involved in reader-response or emotion theory, or may serve as reference for future EEG-based affect recognition studies.

F. CONFIDENTIALITY

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. The information collected about you will be coded using a fake name (pseudonym) or initials and numbers. The information which has your identifiable information will be kept separately from the rest of your data. All the information collected will be stored in a private archive.

The EEG recordings of your sessions will be used solely for experimental purposes, and after the data collection is over, they will be stored in a private archive. Portions of these EEG recordings may be published and/or presented in scientific journals and/or scientific conference proceedings, but will never be published in a non-scientific venue. Further, no information, such as name, address, or other private information, will be included in these publications. Likewise, for this research, your video and voice recordings will be used solely for analysis and gathering insights regarding the experiment. It will also be stored in a private archive.

Apart from this possible usage, such data will only be viewed/used for experimental purposes. At any time during or after the experiment, you may request to review or edit the tapes and/or request that your files be destroyed.

G. PARTICIPATION AND WITHDRAWAL

Your participation in this study is completely voluntary. If at any point in time you wish to withdraw during or after the experiment, you may do so without penalty or consequence of any kind. Any data collected, should you withdraw, will be disposed properly.

H. IDENTIFICATION OF RESEARCHER

If you have any questions or concerns about the research, please feel free to contact:

KRISTINE MA. DOMINIQUE F. KALAW
Master's Student
De La Salle University
Contact No.: (+63) 927 854 4201
Email: kristine_ma_kalaw@dlsu.edu.ph

PROF. ETHEL CHUA JOY ONG
Adviser
De La Salle University
Software Technology Department
Email: ethel.ong@delasalle.ph

CONSENT TO PARTICIPATE IN "RECOGNIZING READER'S AFFECT USING EEG DATA"

I, the undersigned, confirm that (please tick the appropriate boxes):

	TAKING PART IN THIS PROJECT	
1.	I have read and understood the information about the research as provided in the Information Sheet dated December 9, 2016.	☐
2.	I have been given the opportunity to ask questions about the study and my participation.	☐
3.	I voluntarily agree to participate in the study and am aware that taking part in it includes being interviewed and voice-recorded.	☐
4.	I understand that my taking part is voluntary; I can withdraw from the study at any time and I do not have to give any reason for why I no longer want to take part.	☐
	USE OF THE INFORMATION I PROVIDE FOR THIS PROJECT ONLY	
5.	The procedures regarding confidentiality have been clearly explained to me.	☐
6.	Select only **one** of the following:	
	a. I would like my name used and understand that what I have said or written as part of this study will be used in reports, publications, and other research outputs so that anything I have contributed to this research can be recognized	☐
	b. I would like to remain anonymous.	☐
	USE OF THE INFORMATION I PROVIDE BEYOND THIS PROJECT	
7.	I agree for the data I provide be archived by the researchers.	☐
8.	Select only **one** of the following:	
	a. I am allowing other researchers to have access to this data if they agree to preserve the confidentiality of the data and if they agree to the terms I have specified in this form.	☐
	b. I am **not** allowing other researchers to have access to this data and consent only of its use to this project.	☐
9.	If you are **allowing** other researchers to have access to this data (**8a**), select the type/s of data you consent for the other researchers to use. Skip this item if you are **not allowing** other researchers to have access to this data (**8b**). ☐ EEG recording ☐ Video recording ☐ Voice recording	
	SO THAT THIS STUDY CAN USE THE INFORMATION I PROVIDE LEGALLY	
10.	I agree to assign the copyright I hold in any materials related to this project to the Researcher.	☐

PARTICIPANT:

Name of Participant	Signature	Date

RESEARCHERS:

Name of Researcher	Signature	Date
KRISTINE MA. DOMINIQUE F. KALAW		01/10/17
ETHEL CHUA JOY ONG		01/10/17

Appendix C

Sample Segments of Selected Short Stories

The Veldt by Ray Bradbury

Segment #1

"George, I wish you'd look at the nursery."

"What's wrong with it?"

"I don't know."

"Well, then."

"I just want you to look at it, is all, or call a psychologist in to look at it."

"What would a psychologist want with a nursery?"

"You know very well what he'd want." His wife paused in the middle of the kitchen and watched the stove busy humming to itself, making supper for four.

"It's just that the nursery is different now than it was."

"All right, let's have a look."

Segment #2

They walked down the hall of their soundproofed, Happy-life Home, which had cost them thirty thousand dollars installed, this house which clothed and fed and rocked them to sleep and played and sang and was good to them. Their approach sensitized a switch somewhere and the nursery light flicked on when they came within ten feet of it. Similarly, behind them, in the halls, lights went on and off as they left them behind, with a soft automaticity.

Segment #3

"Well," said George Hadley.

They stood on the thatched floor of the nursery. It was forty feet across by forty feet long and thirty feet high; it had cost half again as much as the rest of the house. "But nothing's too good for our children," George had said.

Segment #4

The nursery was silent. It was empty as a jungle glade at hot high noon. The walls were blank and two dimensional. Now, as George and Lydia Hadley stood in the center of the room, the walls began to purr and recede into crystalline distance, it seemed, and presently an African veldt appeared, in three dimensions, on all sides, in colors reproduced to the final pebble and bit of straw. The ceiling above them became a deep sky with a hot yellow sun.

Segment #5

George Hadley felt the perspiration start on his brow.

"Let's get out of this sun," he said. "This is a little too real. But I don't see anything wrong."

"Wait a moment, you'll see," said his wife.

Man from the South by Roald Dahl

Segment #1

It was getting on towards six o'clock so I thought I'd buy myself a beer and go out and sit in a deckchair by the swimming pool and have a little evening sun.

I went to the bar and got the beer and carried it outside and wandered down the garden towards the pool.

Segment #2

It was a fine garden with lawns and beds of azaleas and tall coconut palms, and the wind was blowing strongly through the tops of the palm trees, making the leaves hiss and crackle as though they were on fire. I could see the clusters of big brown nuts hanging down underneath the leaves.

Segment #3

There were plenty of deck-chairs around the swimming pool and there were white tables and huge brightly coloured umbrellas and sunburned men and women sitting around in bathing suits. In the pool itself there were three or four girls and about a dozen boys, all splashing about and making a lot of noise and throwing a large rubber ball at one another.

Segment #4

I stood watching them. The girls were English girls from the hotel. The boys I didn't know about, but they sounded American, and I thought they were probably naval cadets who'd come ashore from the U.S. naval training vessel which had arrived in harbour that morning.

Segment #5

I went over and sat down under a yellow umbrella where there were four empty seats, and I poured my beer and settled back comfortably with a cigarette.

It was very pleasant sitting there in the sunshine with beer and cigarette. It was pleasant to sit and watch the bathers splashing about in the green water.

The Fisherman and the Jinni from *One Thousand and One Nights*

Segment #1

It is said, oh wise and happy King, that a very poor fisherman who swore by Almighty God that he would only cast his net three times each day, went down to the sea late one afternoon as usual, waited until he saw the moon shining above him, and then threw his net very carefully into the water.

Segment #2

He sat there for a time, and then, when he pulled on his net and felt that it had grown heavy, he sang to himself:

"Glide over to me, my magnificent fish

And slither into my waiting net

So that someone asleep on his soft silken bed

Will awaken and buy you with his silver bread."

Segment #3

He opened his net and there, to his horror, found a dead donkey. "A donkey?" he cried out. "My wretched luck. You send me a donkey when you know that my family and I are starved out of our brains?" He managed to free it from his net with one hand while pinching his nose with the other to block out the horrible smell.

Segment #4

He cast his net carefully into the sea again, waited for it to sink, tugged on it and to his amazement felt that the net was even heavier than the first time. It was so heavy he had to climb back on to the shore, drive a stake into the ground, and tie the rope of the net to the stake. Then he hauled with all his might until he managed to pull the net up out of the sea.

Segment #5

But instead of an abundance of fish jumping and playing in the net he found a broken, rusty wooden chest filled with sand. He shouted in a loud voice, "A chest? Is this how you compensate my work? My labour? Or are you telling me that the key to my good fortune lies inside this coffin?"

He kicked the chest as hard as he could, but then managed to recover his patience, and washed out his net once again.

Appendix D

Confusion Matrices

Hourglass of Emotions Confusion Matrices (Base Features)

HoE confusion matrices for the Sex-merged dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	15589	10399
(P) high	10044	8931

DT-AT

	(T) low	(T) high
(P) low	18324	12550
(P) high	8132	5957

DT-PL

	(T) low	(T) high
(P) low	7220	9319
(P) high	13270	15154

DT-SE

	(T) low	(T) high
(P) low	29195	11016
(P) high	4040	712

SVM-AP

	(T) low	(T) high
(P) low	17768	13691
(P) high	7865	5639

SVM-AT

	(T) low	(T) high
(P) low	18811	13753
(P) high	7645	4754

SVM-PL

	(T) low	(T) high
(P) low	9177	14055
(P) high	11313	10418

SVM-SE

	(T) low	(T) high
(P) low	31876	11572
(P) high	1359	156

MLP-AP

	(T) low	(T) high
(P) low	14542	10989
(P) high	11091	8341

MLP-AT

	(T) low	(T) high
(P) low	17273	12307
(P) high	9183	6200

MLP-PL

	(T) low	(T) high
(P) low	9804	10909
(P) high	10686	13564

MLP-SE

	(T) low	(T) high
(P) low	29638	9165
(P) high	3597	2563

HoE confusion matrices for the Female dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	5168	5883
(P) high	7037	5032

DT-AT

	(T) low	(T) high
(P) low	10659	4956
(P) high	5183	2322

DT-PL

	(T) low	(T) high
(P) low	1304	5739
(P) high	9585	6492

DT-SE

	(T) low	(T) high
(P) low	17758	5059
(P) high	208	95

SVM-AP

	(T) low	(T) high
(P) low	5448	5856
(P) high	6757	5059

SVM-AT

	(T) low	(T) high
(P) low	12530	6693
(P) high	3312	585

SVM-PL

	(T) low	(T) high
(P) low	4864	6955
(P) high	6025	5276

SVM-SE

	(T) low	(T) high
(P) low	17961	5154
(P) high	5	0

MLP-AP

	(T) low	(T) high
(P) low	5033	4290
(P) high	7172	6625

MLP-AT

	(T) low	(T) high
(P) low	9966	4785
(P) high	5876	2493

MLP-PL

	(T) low	(T) high
(P) low	3467	5585
(P) high	7422	6646

MLP-SE

	(T) low	(T) high
(P) low	15424	3981
(P) high	2542	1173

HoE confusion matrices for the Male dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	10879	6963
(P) high	2549	1452

DT-AT

	(T) low	(T) high
(P) low	7056	5538
(P) high	3558	5691

DT-PL

	(T) low	(T) high
(P) low	1227	4457
(P) high	8374	7785

DT-SE

	(T) low	(T) high
(P) low	12744	6334
(P) high	2525	240

SVM-AP

	(T) low	(T) high
(P) low	9874	7178
(P) high	3554	1237

SVM-AT

	(T) low	(T) high
(P) low	8052	6030
(P) high	2562	5199

SVM-PL

	(T) low	(T) high
(P) low	2068	3191
(P) high	7533	9051

SVM-SE

	(T) low	(T) high
(P) low	10570	5781
(P) high	4699	793

MLP-AP

	(T) low	(T) high
(P) low	7469	6222
(P) high	5959	2193

MLP-AT

	(T) low	(T) high
(P) low	8828	5588
(P) high	1786	5641

MLP-PL

	(T) low	(T) high
(P) low	3994	3080
(P) high	5607	9162

MLP-SE

	(T) low	(T) high
(P) low	13270	5058
(P) high	1999	1516

HoE confusion matrices for the RP-merged dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	2865	6818
(P) high	16137	7713

DT-AT

	(T) low	(T) high
(P) low	16033	9986
(P) high	4625	2889

DT-PL

	(T) low	(T) high
(P) low	11051	11279
(P) high	5391	5812

DT-SE

	(T) low	(T) high
(P) low	25663	7870
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	8613	9166
(P) high	10389	5365

SVM-AT

	(T) low	(T) high
(P) low	13341	8121
(P) high	7317	4754

SVM-PL

	(T) low	(T) high
(P) low	10461	10348
(P) high	5981	6743

SVM-SE

	(T) low	(T) high
(P) low	25663	7869
(P) high	0	1

MLP-AP

	(T) low	(T) high
(P) low	9894	9157
(P) high	9108	5374

MLP-AT

	(T) low	(T) high
(P) low	15031	10806
(P) high	5627	2069

MLP-PL

	(T) low	(T) high
(P) low	3522	4746
(P) high	12920	12345

MLP-SE

	(T) low	(T) high
(P) low	23256	7461
(P) high	2407	409

HoE confusion matrices for the RP1 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	4166	1738
(P) high	2363	288

DT-AT

	(T) low	(T) high
(P) low	2376	224
(P) high	4460	1495

DT-PL

	(T) low	(T) high
(P) low	1479	1130
(P) high	4965	981

DT-SE

	(T) low	(T) high
(P) low	4714	1180
(P) high	2427	234

SVM-AP

	(T) low	(T) high
(P) low	1228	1412
(P) high	5301	614

SVM-AT

	(T) low	(T) high
(P) low	6836	1719
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	6444	2111
(P) high	0	0

SVM-SE

	(T) low	(T) high
(P) low	7141	1414
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	3613	1653
(P) high	2916	373

MLP-AT

	(T) low	(T) high
(P) low	6836	1719
(P) high	0	0

MLP-PL

	(T) low	(T) high
(P) low	1480	1141
(P) high	4964	970

MLP-SE

	(T) low	(T) high
(P) low	7139	1414
(P) high	2	0

HoE confusion matrices for the RP2 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	3094	4953
(P) high	70	297

DT-AT

	(T) low	(T) high
(P) low	1	4
(P) high	3657	4752

DT-PL

	(T) low	(T) high
(P) low	0	0
(P) high	2249	6165

DT-SE

	(T) low	(T) high
(P) low	1572	706
(P) high	2949	3187

SVM-AP

	(T) low	(T) high
(P) low	1432	4510
(P) high	1732	740

SVM-AT

	(T) low	(T) high
(P) low	2160	3782
(P) high	1498	974

SVM-PL

	(T) low	(T) high
(P) low	686	6256
(P) high	1563	909

SVM-SE

	(T) low	(T) high
(P) low	3085	3285
(P) high	1436	608

MLP-AP

	(T) low	(T) high
(P) low	1001	2663
(P) high	2163	2587

MLP-AT

	(T) low	(T) high
(P) low	1356	2316
(P) high	2302	2440

MLP-PL

	(T) low	(T) high
(P) low	550	4691
(P) high	1699	1474

MLP-SE

	(T) low	(T) high
(P) low	2943	3185
(P) high	1578	708

HoE confusion matrices for the RP3 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	4126	4877
(P) high	0	0

DT-AT

	(T) low	(T) high
(P) low	1983	1589
(P) high	4453	978

DT-PL

	(T) low	(T) high
(P) low	2213	3220
(P) high	1388	2182

DT-SE

	(T) low	(T) high
(P) low	5039	1242
(P) high	2335	387

SVM-AP

	(T) low	(T) high
(P) low	0	0
(P) high	4126	4877

SVM-AT

	(T) low	(T) high
(P) low	6436	2567
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	0	0
(P) high	3601	5402

SVM-SE

	(T) low	(T) high
(P) low	7374	1629
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	1377	2204
(P) high	2749	2673

MLP-AT

	(T) low	(T) high
(P) low	6436	2567
(P) high	0	0

MLP-PL

	(T) low	(T) high
(P) low	1631	2324
(P) high	1970	3078

MLP-SE

	(T) low	(T) high
(P) low	7374	1629
(P) high	0	0

HoE confusion matrices for the RP4 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	0	0
(P) high	5183	2378

DT-AT

	(T) low	(T) high
(P) low	4	2
(P) high	3724	3831

DT-PL

	(T) low	(T) high
(P) low	2974	1129
(P) high	1174	2284

DT-SE

	(T) low	(T) high
(P) low	4470	674
(P) high	2157	260

SVM-AP

	(T) low	(T) high
(P) low	5167	2370
(P) high	16	8

SVM-AT

	(T) low	(T) high
(P) low	1746	3408
(P) high	1982	425

SVM-PL

	(T) low	(T) high
(P) low	1096	2244
(P) high	3052	1169

SVM-SE

	(T) low	(T) high
(P) low	6627	934
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	5183	2378
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	2149	3525
(P) high	1579	308

MLP-PL

	(T) low	(T) high
(P) low	4148	3413
(P) high	0	0

MLP-SE

	(T) low	(T) high
(P) low	6627	934
(P) high	0	0

HoE confusion matrices for the RF-merged dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	4943	7912
(P) high	5501	520

DT-AT

	(T) low	(T) high
(P) low	9627	3290
(P) high	2953	3006

DT-PL

	(T) low	(T) high
(P) low	1931	2160
(P) high	7200	7585

DT-SE

	(T) low	(T) high
(P) low	13295	4843
(P) high	546	192

SVM-AP

	(T) low	(T) high
(P) low	4546	4282
(P) high	5898	4150

SVM-AT

	(T) low	(T) high
(P) low	7026	4107
(P) high	5554	2189

SVM-PL

	(T) low	(T) high
(P) low	3400	2109
(P) high	5731	7636

SVM-SE

	(T) low	(T) high
(P) low	13762	4986
(P) high	79	49

MLP-AP

	(T) low	(T) high
(P) low	3260	3803
(P) high	7184	4629

MLP-AT

	(T) low	(T) high
(P) low	6622	3514
(P) high	5958	2782

MLP-PL

	(T) low	(T) high
(P) low	4013	3578
(P) high	5118	6167

MLP-SE

	(T) low	(T) high
(P) low	10267	2908
(P) high	3574	2127

HoE confusion matrices for the RF1 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	694	515
(P) high	431	1847

DT-AT

	(T) low	(T) high
(P) low	758	451
(P) high	808	1470

DT-PL

	(T) low	(T) high
(P) low	0	0
(P) high	967	2520

DT-SE

	(T) low	(T) high
(P) low	0	0
(P) high	2374	1113

SVM-AP

	(T) low	(T) high
(P) low	1	4
(P) high	1124	2358

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1566	1921

SVM-PL

	(T) low	(T) high
(P) low	363	1895
(P) high	604	625

SVM-SE

	(T) low	(T) high
(P) low	2374	1113
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	431	1847
(P) high	694	515

MLP-AT

	(T) low	(T) high
(P) low	1566	1921
(P) high	0	0

MLP-PL

	(T) low	(T) high
(P) low	0	0
(P) high	967	2520

MLP-SE

	(T) low	(T) high
(P) low	2374	1113
(P) high	0	0

HoE confusion matrices for the RF2 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	0	0
(P) high	1052	2465

DT-AT

	(T) low	(T) high
(P) low	812	562
(P) high	1004	1139

DT-PL

	(T) low	(T) high
(P) low	0	0
(P) high	801	2716

DT-SE

	(T) low	(T) high
(P) low	396	1248
(P) high	602	1271

SVM-AP

	(T) low	(T) high
(P) low	222	1894
(P) high	830	571

SVM-AT

	(T) low	(T) high
(P) low	1816	1701
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	57	2086
(P) high	744	630

SVM-SE

	(T) low	(T) high
(P) low	0	0
(P) high	998	2519

MLP-AP

	(T) low	(T) high
(P) low	1052	2465
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	1816	1701
(P) high	0	0

MLP-PL

	(T) low	(T) high
(P) low	0	6
(P) high	801	2710

MLP-SE

	(T) low	(T) high
(P) low	0	0
(P) high	998	2519

HoE confusion matrices for the RF3 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	5501	518
(P) high	0	0

DT-AT

	(T) low	(T) high
(P) low	5505	514
(P) high	0	0

DT-PL

	(T) low	(T) high
(P) low	5229	661
(P) high	112	17

DT-SE

	(T) low	(T) high
(P) low	5590	29
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	5501	518
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	5505	514
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	5341	678
(P) high	0	0

SVM-SE

	(T) low	(T) high
(P) low	5990	29
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	5501	518
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	5505	514
(P) high	0	0

MLP-PL

	(T) low	(T) high
(P) low	5341	678
(P) high	0	0

MLP-SE

	(T) low	(T) high
(P) low	5990	29
(P) high	0	0

HoE confusion matrices for the RF4 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	1436	1736
(P) high	1330	1351

DT-AT

	(T) low	(T) high
(P) low	0	0
(P) high	3693	2160

DT-PL

	(T) low	(T) high
(P) low	836	2331
(P) high	1186	1500

DT-SE

	(T) low	(T) high
(P) low	0	0
(P) high	4479	1374

SVM-AP

	(T) low	(T) high
(P) low	532	544
(P) high	2234	2543

SVM-AT

	(T) low	(T) high
(P) low	1736	1652
(P) high	1957	508

SVM-PL

	(T) low	(T) high
(P) low	0	1
(P) high	2022	3830

SVM-SE

	(T) low	(T) high
(P) low	4479	1374
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	0	0
(P) high	2766	3087

MLP-AT

	(T) low	(T) high
(P) low	1551	1600
(P) high	2142	560

MLP-PL

	(T) low	(T) high
(P) low	0	0
(P) high	2022	3831

MLP-SE

	(T) low	(T) high
(P) low	4479	1374
(P) high	0	0

Emotions of Literary Response Confusion Matrices (Base Features)

ELR confusion matrices for the Sex-merged dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	37559	7274
(P) high	116	14

DT-AT

	(T) low	(T) high
(P) low	23374	14593
(P) high	4337	2659

DT-PL

	(T) low	(T) high
(P) low	571	3387
(P) high	20385	20620

DT-SE

	(T) low	(T) high
(P) low	44005	955
(P) high	3	0

SVM-AP

	(T) low	(T) high
(P) low	37675	7288
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	23547	14675
(P) high	4164	2577

SVM-PL

	(T) low	(T) high
(P) low	4249	10644
(P) high	16707	13363

SVM-SE

	(T) low	(T) high
(P) low	44008	955
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	12237	6843
(P) high	13396	12487

MLP-AT

	(T) low	(T) high
(P) low	20617	15365
(P) high	5839	3142

MLP-PL

	(T) low	(T) high
(P) low	8684	8901
(P) high	11806	15572

MLP-SE

	(T) low	(T) high
(P) low	28162	8646
(P) high	5073	3082

ELR confusion matrices for the Female dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	17416	2553
(P) high	2723	428

DT-AT

	(T) low	(T) high
(P) low	3143	1992
(P) high	9773	8212

DT-PL

	(T) low	(T) high
(P) low	10146	7241
(P) high	2507	3226

DT-SE

	(T) low	(T) high
(P) low	22925	195
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	20139	2981
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	7026	5996
(P) high	5890	4208

SVM-PL

	(T) low	(T) high
(P) low	9486	3604
(P) high	3167	6863

SVM-SE

	(T) low	(T) high
(P) low	22925	195
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	20117	2978
(P) high	22	3

MLP-AT

	(T) low	(T) high
(P) low	8545	6063
(P) high	4371	4141

MLP-PL

	(T) low	(T) high
(P) low	3033	1033
(P) high	9620	9434

MLP-SE

	(T) low	(T) high
(P) low	22925	195
(P) high	0	0

ELR confusion matrices for the Male dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	13876	3116
(P) high	3660	1191

DT-AT

	(T) low	(T) high
(P) low	5138	3522
(P) high	9657	3526

DT-PL

	(T) low	(T) high
(P) low	2393	7844
(P) high	5910	5696

DT-SE

	(T) low	(T) high
(P) low	21080	760
(P) high	3	0

SVM-AP

	(T) low	(T) high
(P) low	17536	4307
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	13431	5604
(P) high	1364	1444

SVM-PL

	(T) low	(T) high
(P) low	1721	4725
(P) high	6582	8815

SVM-SE

	(T) low	(T) high
(P) low	21083	760
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	15072	3232
(P) high	2464	1075

MLP-AT

	(T) low	(T) high
(P) low	10135	4962
(P) high	4660	2086

MLP-PL

	(T) low	(T) high
(P) low	654	2121
(P) high	7649	11419

MLP-SE

	(T) low	(T) high
(P) low	21083	760
(P) high	0	0

ELR confusion matrices for the RP-merged dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	26597	4444
(P) high	2102	390

DT-AT

	(T) low	(T) high
(P) low	17548	13367
(P) high	333	2285

DT-PL

	(T) low	(T) high
(P) low	6813	10014
(P) high	10451	6255

DT-SE

	(T) low	(T) high
(P) low	31775	1758
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	28699	4834
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	12751	7181
(P) high	5130	8471

SVM-PL

	(T) low	(T) high
(P) low	9852	11579
(P) high	7412	4690

SVM-SE

	(T) low	(T) high
(P) low	31775	1758
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	25168	4296
(P) high	3531	538

MLP-AT

	(T) low	(T) high
(P) low	12952	9448
(P) high	4929	6204

MLP-PL

	(T) low	(T) high
(P) low	8894	9850
(P) high	8370	6419

MLP-SE

	(T) low	(T) high
(P) low	31764	1758
(P) high	11	0

ELR confusion matrices for the RP1 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	5363	540
(P) high	2417	235

DT-AT

	(T) low	(T) high
(P) low	2327	3588
(P) high	2639	1

DT-PL

	(T) low	(T) high
(P) low	216	2409
(P) high	4106	1824

DT-SE

	(T) low	(T) high
(P) low	4901	1000
(P) high	2654	0

SVM-AP

	(T) low	(T) high
(P) low	7780	775
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	342	2298
(P) high	4624	1291

SVM-PL

	(T) low	(T) high
(P) low	4317	4233
(P) high	5	0

SVM-SE

	(T) low	(T) high
(P) low	7555	1000
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	7749	775
(P) high	31	0

MLP-AT

	(T) low	(T) high
(P) low	2974	2284
(P) high	1992	1305

MLP-PL

	(T) low	(T) high
(P) low	4253	4224
(P) high	69	9

MLP-SE

	(T) low	(T) high
(P) low	7555	1000
(P) high	0	0

ELR confusion matrices for the RP2 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	5234	840
(P) high	2071	269

DT-AT

	(T) low	(T) high
(P) low	0	0
(P) high	5396	3018

DT-PL

	(T) low	(T) high
(P) low	1699	804
(P) high	1831	4080

DT-SE

	(T) low	(T) high
(P) low	8377	37
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	7305	1109
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	1135	1337
(P) high	4261	1681

SVM-PL

	(T) low	(T) high
(P) low	1840	4102
(P) high	1690	782

SVM-SE

	(T) low	(T) high
(P) low	8377	37
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	4113	658
(P) high	3192	451

MLP-AT

	(T) low	(T) high
(P) low	1626	1675
(P) high	3770	1343

MLP-PL

	(T) low	(T) high
(P) low	3	2
(P) high	3527	4882

MLP-SE

	(T) low	(T) high
(P) low	8377	37
(P) high	0	0

ELR confusion matrices for the RP3 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	2721	714
(P) high	4770	798

DT-AT

	(T) low	(T) high
(P) low	3267	2888
(P) high	222	2626

DT-PL

	(T) low	(T) high
(P) low	3232	2804
(P) high	873	2094

DT-SE

	(T) low	(T) high
(P) low	6227	65
(P) high	2711	0

SVM-AP

	(T) low	(T) high
(P) low	7491	1512
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	1290	4134
(P) high	2199	1380

SVM-PL

	(T) low	(T) high
(P) low	1956	3468
(P) high	2149	1430

SVM-SE

	(T) low	(T) high
(P) low	8938	65
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	5226	1075
(P) high	2265	437

MLP-AT

	(T) low	(T) high
(P) low	1116	1586
(P) high	2373	3928

MLP-PL

	(T) low	(T) high
(P) low	1956	3468
(P) high	2149	1430

MLP-SE

	(T) low	(T) high
(P) low	8938	65
(P) high	0	0

ELR confusion matrices for the RP4 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	6123	1438
(P) high	0	0

DT-AT

	(T) low	(T) high
(P) low	1233	604
(P) high	2797	2927

DT-PL

	(T) low	(T) high
(P) low	2432	1950
(P) high	2875	304

DT-SE

	(T) low	(T) high
(P) low	6871	652
(P) high	34	4

SVM-AP

	(T) low	(T) high
(P) low	3352	1114
(P) high	2771	324

SVM-AT

	(T) low	(T) high
(P) low	2745	2911
(P) high	1285	620

SVM-PL

	(T) low	(T) high
(P) low	5303	2254
(P) high	4	0

SVM-SE

	(T) low	(T) high
(P) low	6905	656
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	3147	1090
(P) high	2976	348

MLP-AT

	(T) low	(T) high
(P) low	4027	3528
(P) high	3	3

MLP-PL

	(T) low	(T) high
(P) low	3742	1981
(P) high	1565	273

MLP-SE

	(T) low	(T) high
(P) low	6905	656
(P) high	0	0

ELR confusion matrices for the RF-merged dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	16365	2511
(P) high	0	0

DT-AT

	(T) low	(T) high
(P) low	0	2
(P) high	6931	11943

DT-PL

	(T) low	(T) high
(P) low	8358	4500
(P) high	36880	2330

DT-SE

	(T) low	(T) high
(P) low	18671	205
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	16365	2511
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	6931	11945

SVM-PL

	(T) low	(T) high
(P) low	3390	3312
(P) high	8656	3518

SVM-SE

	(T) low	(T) high
(P) low	18671	205
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	15716	2394
(P) high	649	117

MLP-AT

	(T) low	(T) high
(P) low	794	1751
(P) high	6137	10194

MLP-PL

	(T) low	(T) high
(P) low	9805	4304
(P) high	2241	2526

MLP-SE

	(T) low	(T) high
(P) low	18671	205
(P) high	0	0

ELR confusion matrices for the RF1 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	1012	156
(P) high	2050	269

DT-AT

	(T) low	(T) high
(P) low	1431	2056
(P) high	0	0

DT-PL

	(T) low	(T) high
(P) low	187	1022
(P) high	777	1501

DT-SE

	(T) low	(T) high
(P) low	3471	16
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	3062	425
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	262	947
(P) high	1169	1109

SVM-PL

	(T) low	(T) high
(P) low	0	0
(P) high	964	2523

SVM-SE

	(T) low	(T) high
(P) low	3471	16
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	3062	425
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1431	2056

MLP-PL

	(T) low	(T) high
(P) low	0	0
(P) high	964	2523

MLP-SE

	(T) low	(T) high
(P) low	3471	16
(P) high	0	0

ELR confusion matrices for the RF2 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	854	520
(P) high	1884	259

DT-AT

	(T) low	(T) high
(P) low	562	1581
(P) high	598	776

DT-PL

	(T) low	(T) high
(P) low	1914	229
(P) high	518	856

DT-SE

	(T) low	(T) high
(P) low	3432	85
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	2738	779
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1160	2357

SVM-PL

	(T) low	(T) high
(P) low	518	856
(P) high	1914	229

SVM-SE

	(T) low	(T) high
(P) low	3432	85
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	855	520
(P) high	1883	259

MLP-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1160	2357

MLP-PL

	(T) low	(T) high
(P) low	1	0
(P) high	2431	1085

MLP-SE

	(T) low	(T) high
(P) low	3432	85
(P) high	0	0

ELR confusion matrices for the RF3 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	5576	442
(P) high	1	0

DT-AT

	(T) low	(T) high
(P) low	752	2060
(P) high	1114	2093

DT-PL

	(T) low	(T) high
(P) low	2423	789
(P) high	2807	0

DT-SE

	(T) low	(T) high
(P) low	5980	39
(P) high	0	0

SVM-AP

	(T) low	(T) high
(P) low	5577	442
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1866	4153

SVM-PL

	(T) low	(T) high
(P) low	5230	789
(P) high	0	0

SVM-SE

	(T) low	(T) high
(P) low	5980	39
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	2692	115
(P) high	2885	327

MLP-AT

	(T) low	(T) high
(P) low	750	2057
(P) high	1116	2096

MLP-PL

	(T) low	(T) high
(P) low	5230	789
(P) high	0	0

MLP-SE

	(T) low	(T) high
(P) low	5980	39
(P) high	0	0

ELR confusion matrices for the RF4 dataset (Base Features).

DT-AP

	(T) low	(T) high
(P) low	0	0
(P) high	4988	865

DT-AT

	(T) low	(T) high
(P) low	2474	3379
(P) high	0	0

DT-PL

	(T) low	(T) high
(P) low	2278	1025
(P) high	1142	1408

DT-SE

	(T) low	(T) high
(P) low	2684	65
(P) high	3104	0

SVM-AP

	(T) low	(T) high
(P) low	4988	865
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	2474	3379

SVM-PL

	(T) low	(T) high
(P) low	1198	1504
(P) high	2222	929

SVM-SE

	(T) low	(T) high
(P) low	5788	65
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	0	0
(P) high	4988	865

MLP-AT

	(T) low	(T) high
(P) low	2474	3379
(P) high	0	0

MLP-PL

	(T) low	(T) high
(P) low	2222	929
(P) high	1198	1504

MLP-SE

	(T) low	(T) high
(P) low	5788	65
(P) high	0	0

Hourglass of Emotions Confusion Matrices (PCA Features)

HoE confusion matrices for the Sex-merged dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	17678	13679
(P) high	7955	5651

SVM-AT

	(T) low	(T) high
(P) low	20676	14695
(P) high	5780	3812

SVM-PL

	(T) low	(T) high
(P) low	11463	14504
(P) high	9027	9969

SVM-SE

	(T) low	(T) high
(P) low	33058	11722
(P) high	177	6

MLP-AP

	(T) low	(T) high
(P) low	14605	12958
(P) high	11028	6372

MLP-AT

	(T) low	(T) high
(P) low	13332	11493
(P) high	13124	7014

MLP-PL

	(T) low	(T) high
(P) low	7747	9157
(P) high	12743	15316

MLP-SE

	(T) low	(T) high
(P) low	29976	9545
(P) high	3259	2183

HoE confusion matrices for the Female dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	6856	7774
(P) high	5349	3141

SVM-AT

	(T) low	(T) high
(P) low	14140	6822
(P) high	1702	456

SVM-PL

	(T) low	(T) high
(P) low	5813	10374
(P) high	5076	1857

SVM-SE

	(T) low	(T) high
(P) low	17960	5154
(P) high	6	0

MLP-AP

	(T) low	(T) high
(P) low	5525	5490
(P) high	6680	6425

MLP-AT

	(T) low	(T) high
(P) low	12585	6437
(P) high	3257	841

MLP-PL

	(T) low	(T) high
(P) low	5019	6386
(P) high	5870	5845

MLP-SE

	(T) low	(T) high
(P) low	15862	4792
(P) high	2104	362

HoE confusion matrices for the Male dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	9665	7007
(P) high	3763	1408

SVM-AT

	(T) low	(T) high
(P) low	6366	5464
(P) high	4248	5765

SVM-PL

	(T) low	(T) high
(P) low	1982	3465
(P) high	7619	8777

SVM-SE

	(T) low	(T) high
(P) low	10884	5891
(P) high	4385	683

MLP-AP

	(T) low	(T) high
(P) low	11789	7626
(P) high	1639	789

MLP-AT

	(T) low	(T) high
(P) low	5945	3420
(P) high	4669	7809

MLP-PL

	(T) low	(T) high
(P) low	5676	5495
(P) high	3925	6747

MLP-SE

	(T) low	(T) high
(P) low	13263	5230
(P) high	2006	1344

HoE confusion matrices for the RP-merged dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	7949	8881
(P) high	11053	5650

SVM-AT

	(T) low	(T) high
(P) low	16170	11875
(P) high	4488	1000

SVM-PL

	(T) low	(T) high
(P) low	9333	9813
(P) high	7109	7278

SVM-SE

	(T) low	(T) high
(P) low	25663	7870
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	9422	11535
(P) high	9580	2996

MLP-AT

	(T) low	(T) high
(P) low	15138	8854
(P) high	5520	4021

MLP-PL

	(T) low	(T) high
(P) low	10710	13054
(P) high	5732	4037

MLP-SE

	(T) low	(T) high
(P) low	24986	7690
(P) high	677	180

HoE confusion matrices for the RP1 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	1228	1412
(P) high	5301	614

SVM-AT

	(T) low	(T) high
(P) low	6836	1719
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	6444	2111
(P) high	0	0

SVM-SE

	(T) low	(T) high
(P) low	7141	1414
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	1245	1412
(P) high	5284	614

MLP-AT

	(T) low	(T) high
(P) low	6826	1718
(P) high	10	1

MLP-PL

	(T) low	(T) high
(P) low	6396	2093
(P) high	48	18

MLP-SE

	(T) low	(T) high
(P) low	7118	1412
(P) high	23	2

HoE confusion matrices for the RP2 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	1432	4510
(P) high	1732	740

SVM-AT

	(T) low	(T) high
(P) low	2160	3782
(P) high	1498	974

SVM-PL

	(T) low	(T) high
(P) low	686	5256
(P) high	1563	909

SVM-SE

	(T) low	(T) high
(P) low	2757	3185
(P) high	1764	708

MLP-AP

	(T) low	(T) high
(P) low	7	45
(P) high	3157	5205

MLP-AT

	(T) low	(T) high
(P) low	1352	2312
(P) high	2306	2444

MLP-PL

	(T) low	(T) high
(P) low	6	44
(P) high	2243	6121

MLP-SE

	(T) low	(T) high
(P) low	1215	2486
(P) high	3306	1407

HoE confusion matrices for the RP3 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	0	0
(P) high	4126	4877

SVM-AT

	(T) low	(T) high
(P) low	6436	2567
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	0	0
(P) high	3601	5402

SVM-SE

	(T) low	(T) high
(P) low	7374	1629
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	2781	3520
(P) high	1345	1357

MLP-AT

	(T) low	(T) high
(P) low	4089	2147
(P) high	2347	420

MLP-PL

	(T) low	(T) high
(P) low	0	0
(P) high	3601	5402

MLP-SE

	(T) low	(T) high
(P) low	7374	1629
(P) high	0	0

HoE confusion matrices for the RP4 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	5183	2378
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	373	2965
(P) high	3355	868

SVM-PL

	(T) low	(T) high
(P) low	1093	2242
(P) high	3055	1171

SVM-SE

	(T) low	(T) high
(P) low	6627	934
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	4908	2094
(P) high	275	284

MLP-AT

	(T) low	(T) high
(P) low	3294	3655
(P) high	434	178

MLP-PL

	(T) low	(T) high
(P) low	2162	2147
(P) high	1986	1266

MLP-SE

	(T) low	(T) high
(P) low	6594	927
(P) high	33	7

HoE confusion matrices for the RF-merged dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	4344	2527
(P) high	6100	5905

SVM-AT

	(T) low	(T) high
(P) low	6227	3161
(P) high	6353	3135

SVM-PL

	(T) low	(T) high
(P) low	3498	2265
(P) high	5633	7480

SVM-SE

	(T) low	(T) high
(P) low	13539	4896
(P) high	302	139

MLP-AP

	(T) low	(T) high
(P) low	5538	5301
(P) high	4906	3131

MLP-AT

	(T) low	(T) high
(P) low	7011	3936
(P) high	5569	2360

MLP-PL

	(T) low	(T) high
(P) low	4014	4589
(P) high	5117	5156

MLP-SE

	(T) low	(T) high
(P) low	8932	4074
(P) high	4909	961

HoE confusion matrices for the RF1 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	0	0
(P) high	1125	2362

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1566	1921

SVM-PL

	(T) low	(T) high
(P) low	349	1832
(P) high	618	688

SVM-SE

	(T) low	(T) high
(P) low	2374	1113
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	0	0
(P) high	1125	2362

MLP-AT

	(T) low	(T) high
(P) low	808	1470
(P) high	758	451

MLP-PL

	(T) low	(T) high
(P) low	0	0
(P) high	967	2520

MLP-SE

	(T) low	(T) high
(P) low	2374	1113
(P) high	0	0

HoE confusion matrices for the RF2 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	225	1907
(P) high	827	558

SVM-AT

	(T) low	(T) high
(P) low	1816	1701
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	57	2086
(P) high	744	630

SVM-SE

	(T) low	(T) high
(P) low	0	0
(P) high	998	2519

MLP-AP

	(T) low	(T) high
(P) low	217	1853
(P) high	835	612

MLP-AT

	(T) low	(T) high
(P) low	1815	1701
(P) high	1	0

MLP-PL

	(T) low	(T) high
(P) low	0	0
(P) high	801	2716

MLP-SE

	(T) low	(T) high
(P) low	312	1037
(P) high	686	1482

HoE confusion matrices for the RF3 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	5501	518
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	5505	514
(P) high	0	0

SVM-PL

	(T) low	(T) high
(P) low	5341	678
(P) high	0	0

SVM-SE

	(T) low	(T) high
(P) low	5990	29
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	5501	518
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	5505	514
(P) high	0	0

MLP-PL

	(T) low	(T) high
(P) low	5321	678
(P) high	20	0

MLP-SE

	(T) low	(T) high
(P) low	5990	29
(P) high	0	0

HoE confusion matrices for the RF4 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	449	465
(P) high	2317	2622

SVM-AT

	(T) low	(T) high
(P) low	1589	1604
(P) high	2104	556

SVM-PL

	(T) low	(T) high
(P) low	0	0
(P) high	2022	3831

SVM-SE

	(T) low	(T) high
(P) low	4479	1374
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	1345	1357
(P) high	1421	1730

MLP-AT

	(T) low	(T) high
(P) low	1608	1609
(P) high	2085	551

MLP-PL

	(T) low	(T) high
(P) low	58	106
(P) high	1964	3725

MLP-SE

	(T) low	(T) high
(P) low	4362	1348
(P) high	117	26

Emotions of Literary Response Confusion Matrices (PCA Features)

ELR confusion matrices for the Sex-merged dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	37675	7288
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	22503	14835
(P) high	5208	2417

SVM-PL

	(T) low	(T) high
(P) low	5250	12057
(P) high	15706	11950

SVM-SE

	(T) low	(T) high
(P) low	44008	955
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	32847	6840
(P) high	4828	448

MLP-AT

	(T) low	(T) high
(P) low	19260	12261
(P) high	8451	4991

MLP-PL

	(T) low	(T) high
(P) low	8958	8625
(P) high	11998	15382

MLP-SE

	(T) low	(T) high
(P) low	44006	955
(P) high	2	0

ELR confusion matrices for the Female dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	20139	2981
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	6946	5104
(P) high	5970	5100

SVM-PL

	(T) low	(T) high
(P) low	7688	2764
(P) high	4965	7703

SVM-SE

	(T) low	(T) high
(P) low	22925	195
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	20136	2981
(P) high	3	0

MLP-AT

	(T) low	(T) high
(P) low	7366	6294
(P) high	5550	3910

MLP-PL

	(T) low	(T) high
(P) low	3640	2685
(P) high	9013	7782

MLP-SE

	(T) low	(T) high
(P) low	22925	195
(P) high	0	0

ELR confusion matrices for the Male dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	17536	4307
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	13506	5638
(P) high	1289	1410

SVM-PL

	(T) low	(T) high
(P) low	1744	4750
(P) high	6559	8790

SVM-SE

	(T) low	(T) high
(P) low	21083	760
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	15543	4133
(P) high	1993	174

MLP-AT

	(T) low	(T) high
(P) low	9251	4392
(P) high	5544	2656

MLP-PL

	(T) low	(T) high
(P) low	1181	2891
(P) high	7122	10649

MLP-SE

	(T) low	(T) high
(P) low	21083	760
(P) high	0	0

ELR confusion matrices for the RP-merged dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	28699	4834
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	13702	8341
(P) high	4179	7311

SVM-PL

	(T) low	(T) high
(P) low	9717	11546
(P) high	7547	4723

SVM-SE

	(T) low	(T) high
(P) low	31775	1758
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	26294	4558
(P) high	2405	276

MLP-AT

	(T) low	(T) high
(P) low	15237	10079
(P) high	2644	5573

MLP-PL

	(T) low	(T) high
(P) low	4047	7029
(P) high	13217	9240

MLP-SE

	(T) low	(T) high
(P) low	31258	1758
(P) high	517	0

ELR confusion matrices for the RP1 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	7780	775
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	342	2298
(P) high	4624	1291

SVM-PL

	(T) low	(T) high
(P) low	4267	4227
(P) high	55	6

SVM-SE

	(T) low	(T) high
(P) low	7555	1000
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	7780	775
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	362	2298
(P) high	4604	1291

MLP-PL

	(T) low	(T) high
(P) low	2180	3938
(P) high	2142	295

MLP-SE

	(T) low	(T) high
(P) low	7555	1000
(P) high	0	0

ELR confusion matrices for the RP2 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	7305	1109
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	1135	1337
(P) high	4261	1681

SVM-PL

	(T) low	(T) high
(P) low	1840	4102
(P) high	1690	782

SVM-SE

	(T) low	(T) high
(P) low	8377	37
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	7305	1109
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	5361	3003
(P) high	35	15

MLP-PL

	(T) low	(T) high
(P) low	1063	2601
(P) high	2467	2283

MLP-SE

	(T) low	(T) high
(P) low	8377	37
(P) high	0	0

ELR confusion matrices for the RP3 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	7491	1512
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	1290	4134
(P) high	2199	1380

SVM-PL

	(T) low	(T) high
(P) low	1956	3468
(P) high	2149	1430

SVM-SE

	(T) low	(T) high
(P) low	8938	65
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	7491	1512
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	1011	3669
(P) high	2478	1845

MLP-PL

	(T) low	(T) high
(P) low	2025	3501
(P) high	2080	1397

MLP-SE

	(T) low	(T) high
(P) low	8938	65
(P) high	0	0

ELR confusion matrices for the RP4 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	3566	1149
(P) high	2557	289

SVM-AT

	(T) low	(T) high
(P) low	2567	2786
(P) high	1463	745

SVM-PL

	(T) low	(T) high
(P) low	5302	2254
(P) high	5	0

SVM-SE

	(T) low	(T) high
(P) low	6905	656
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	879	1009
(P) high	5244	429

MLP-AT

	(T) low	(T) high
(P) low	3904	3468
(P) high	126	63

MLP-PL

	(T) low	(T) high
(P) low	5167	2205
(P) high	140	49

MLP-SE

	(T) low	(T) high
(P) low	6905	656
(P) high	0	0

ELR confusion matrices for the RF-merged dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	16365	2511
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	6931	11945

SVM-PL

	(T) low	(T) high
(P) low	3532	3329
(P) high	8514	3501

SVM-SE

	(T) low	(T) high
(P) low	18671	205
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	16361	2511
(P) high	4	0

MLP-AT

	(T) low	(T) high
(P) low	666	831
(P) high	6265	11114

MLP-PL

	(T) low	(T) high
(P) low	5361	1811
(P) high	6685	5019

MLP-SE

	(T) low	(T) high
(P) low	18671	205
(P) high	0	0

ELR confusion matrices for the RF1 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	3062	425
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	262	947
(P) high	1169	1109

SVM-PL

	(T) low	(T) high
(P) low	0	0
(P) high	964	2523

SVM-SE

	(T) low	(T) high
(P) low	3471	16
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	3062	425
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	262	947
(P) high	1169	1109

MLP-PL

	(T) low	(T) high
(P) low	777	1501
(P) high	187	1022

MLP-SE

	(T) low	(T) high
(P) low	3471	16
(P) high	0	0

ELR confusion matrices for the RF2 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	2738	779
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1160	2357

SVM-PL

	(T) low	(T) high
(P) low	518	856
(P) high	1914	229

SVM-SE

	(T) low	(T) high
(P) low	3432	85
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	2738	779
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1160	2357

MLP-PL

	(T) low	(T) high
(P) low	513	849
(P) high	1919	236

MLP-SE

	(T) low	(T) high
(P) low	3432	85
(P) high	0	0

ELR confusion matrices for the RF3 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	5577	442
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	1866	4153

SVM-PL

	(T) low	(T) high
(P) low	5230	789
(P) high	0	0

SVM-SE

	(T) low	(T) high
(P) low	5980	39
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	5577	442
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	5	10
(P) high	1861	4143

MLP-PL

	(T) low	(T) high
(P) low	5230	789
(P) high	0	0

MLP-SE

	(T) low	(T) high
(P) low	5980	39
(P) high	0	0

ELR confusion matrices for the RF4 dataset (PCA Features).

SVM-AP

	(T) low	(T) high
(P) low	4988	865
(P) high	0	0

SVM-AT

	(T) low	(T) high
(P) low	0	0
(P) high	2474	3379

SVM-PL

	(T) low	(T) high
(P) low	1198	1504
(P) high	2222	929

SVM-SE

	(T) low	(T) high
(P) low	5788	65
(P) high	0	0

MLP-AP

	(T) low	(T) high
(P) low	4988	865
(P) high	0	0

MLP-AT

	(T) low	(T) high
(P) low	11	11
(P) high	2463	3368

MLP-PL

	(T) low	(T) high
(P) low	3181	2340
(P) high	239	93

MLP-SE

	(T) low	(T) high
(P) low	5788	65
(P) high	0	0

Appendix E

Relevant Features Based on Decision Trees

Relevant features based on Decision Trees for HoE classification.

Feature	#
alpha_T_minDASM	18
theta_T_avgDASM	15
theta_T_minDASM	12
alpha_T_maxRASM	11
alpha_AF_maxRASM	10
alpha_AF_minDASM	9
theta_AF_maxRASM	9
theta_AF_minDASM	8
theta_T7_avgPSD	8
alpha_AF_maxDASM	7
alpha_Pz_avgPSD	7
alpha_T_avgDASM	7
theta_AF_avgRASM	7
theta_AF3_maxMag	7
theta_T_maxDASM	7
theta_AF_avgDASM	6
theta_AF4_maxMag	6
theta_Pz_maxMag	6
theta_T_avgRASM	6
theta_T8_maxMag	6
alpha_AF_minRASM	5
alpha_AF3_avgPSD	5
alpha_AF3_maxMag	5
alpha_AF3_minPSD	5
alpha_T7_maxMag	5
theta_T8_avgPSD	5
alpha_AF_avgDASM	4
alpha_AF4_maxMag	4
alpha_Pz_minPSD	4
alpha_T_avgRASM	4
alpha_T_maxDASM	4
alpha_T_minRASM	4
theta_AF_maxDASM	4
theta_AF3_minMag	4
theta_Pz_maxPSD	4

Feature	#
theta_T_minRASM	4
theta_T7_maxMag	4
alpha_AF3_avgMag	3
alpha_Pz_avgMag	3
alpha_T7_avgPSD	3
theta_AF3_avgPSD	3
theta_T7_avgMag	3
theta_T8_maxPSD	3
alpha_AF_avgRASM	2
alpha_Pz_maxMag	2
alpha_T7_avgMag	2
alpha_T8_avgMag	2
theta_AF_minRASM	2
theta_AF3_avgMag	2
theta_AF4_avgPSD	2
theta_AF4_maxPSD	2
theta_Pz_minMag	2
theta_T7_minMag	2
theta_T8_avgMag	2
theta_T8_minMag	2
alpha_AF3_minMag	1
alpha_AF4_maxPSD	1
alpha_AF4_minMag	1
alpha_Pz_maxPSD	1
alpha_Pz_minMag	1
alpha_T7_minMag	1
alpha_T8_minMag	1
theta_AF3_minPSD	1
theta_AF4_avgMag	1
theta_Pz_avgMag	1
theta_Pz_avgPSD	1
theta_Pz_minPSD	1
theta_T_maxRASM	1
theta_T7_maxPSD	1

Relevant features based on Decision Trees for ELR classification.

Feature	#	Feature	#
alpha_T_maxRASM	11	alpha_AF3_maxMag	2
alpha_T_avgDASM	10	alpha_AF4_minMag	2
alpha_T_minDASM	9	alpha_Pz_maxPSD	2
alpha_T_avgRASM	8	alpha_T7_avgPSD	2
alpha_T_maxDASM	8	alpha_T8_avgPSD	2
alpha_AF_minDASM	7	alpha_T8_minPSD	2
alpha_AF_minRASM	7	theta_AF3_avgMag	2
theta_T_minRASM	7	theta_AF3_avgPSD	2
theta_T7_avgPSD	7	theta_AF3_minMag	2
theta_AF_avgRASM	6	theta_AF4_avgMag	2
theta_AF_maxRASM	6	theta_T_maxDASM	2
theta_AF_minDASM	6	theta_T_maxRASM	2
theta_T_avgDASM	6	theta_T7_maxMag	2
alpha_AF_maxDASM	5	theta_T8_avgPSD	2
alpha_Pz_maxMag	5	theta_T8_maxMag	2
theta_AF_maxDASM	5	theta_T8_minMag	2
theta_AF_minRASM	5	alpha_AF_avgDASM	1
theta_Pz_maxMag	5	alpha_AF4_avgMag	1
alpha_AF_avgRASM	4	alpha_AF4_maxMag	1
alpha_AF3_avgMag	4	alpha_AF4_maxPSD	1
alpha_AF3_minPSD	4	alpha_AF4_minPSD	1
alpha_T_minRASM	4	alpha_Pz_avgMag	1
alpha_T7_avgMag	4	alpha_Pz_minMag	1
alpha_T7_maxMag	4	alpha_T7_maxPSD	1
theta_AF3_maxMag	4	alpha_T7_minMag	1
theta_Pz_avgPSD	4	alpha_T7_minPSD	1
theta_T_avgRASM	4	alpha_T8_avgMag	1
theta_T_minDASM	4	alpha_T8_maxMag	1
theta_T7_avgMag	4	alpha_T8_minMag	1
alpha_AF3_avgPSD	3	theta_AF3_minPSD	1
theta_AF_avgDASM	3	theta_AF4_avgPSD	1
theta_AF3_maxPSD	3	theta_AF4_maxPSD	1
theta_AF4_minMag	3	theta_Pz_minMag	1
theta_Pz_maxPSD	3	theta_T7_maxPSD	1
theta_T7_minMag	3	theta_T7_minPSD	1

Appendix F

Summary of Decision Tree Classification Results

Hourglass of Emotions Decision Tree Classification Results

Summary of Sex-merged dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	54.53	53.53	53.51	53.52	0.070
AT	54.00	50.82	50.72	50.77	0.015
PL	49.76	48.48	48.58	48.53	-0.029
SE	66.51	43.79	46.96	45.32	-0.075

Summary of Female dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	44.12	44.23	44.22	44.22	-0.115
AT	56.15	49.60	49.59	49.59	-0.008
PL	33.72	29.45	32.53	30.91	-0.356
SE	77.22	54.59	50.34	52.38	0.010

Summary of Male dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	56.45	48.63	49.14	48.88	-0.019
AT	58.36	58.78	58.58	58.68	0.171
PL	58.74	34.88	38.19	36.46	-0.247
SE	59.44	37.74	43.56	40.44	-0.154

Summary of RP-merged dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	31.55	30.96	34.08	32.45	-0.296
AT	56.43	50.03	50.03	50.03	0.001
PL	50.29	50.68	50.61	50.64	0.012
SE	76.53	38.27	50.00	43.36	0.000

Summary of RP1 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	52.06	40.71	39.01	39.84	-0.199
AT	45.25	58.24	60.86	59.52	0.113
PL	28.76	36.59	34.71	35.63	-0.190
SE	57.84	44.39	41.28	42.78	-0.129

Summary of RP2 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	40.30	59.69	51.72	55.42	0.026
AT	56.49	38.26	49.97	43.34	-0.001
PL	73.27	36.64	50.00	42.29	0.000
SE	56.56	60.47	58.32	59.38	0.160

Summary of RP3 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	45.83	22.91	50.00	31.42	0.000
AT	32.89	36.76	34.46	35.57	-0.233
PL	48.82	50.93	50.92	50.92	0.017
SE	60.27	47.22	46.05	46.63	-0.063

Summary of RP4 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	31.45	15.73	50.00	23.93	0.000
AT	50.72	58.69	50.00	54.00	0.001
PL	69.54	69.27	69.31	69.29	0.386
SE	62.56	48.83	47.64	48.23	-0.028

Summary of RF-merged dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	28.94	23.54	26.75	25.04	-0.478
AT	66.93	62.49	62.14	62.31	0.246
PL	50.41	49.25	49.49	49.37	-0.010
SE	71.45	49.66	49.93	49.79	-0.002

Summary of RF1 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	72.87	69.24	69.94	69.59	0.391
AT	63.89	63.61	62.46	63.03	0.255
PL	72.27	36.13	50.00	41.95	0.000
SE	31.92	15.96	50.00	24.20	0.000

Summary of RF2 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	70.09	35.04	50.00	41.20	0.000
AT	55.47	56.12	55.84	55.98	0.116
PL	77.22	38.61	50.00	43.57	0.000
SE	47.40	45.97	45.07	45.52	-0.083

Summary of RF3 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	91.39	45.70	50.00	47.75	0.000
AT	91.46	45.73	50.00	47.77	0.000
PL	87.16	50.98	50.21	50.59	0.006
SE	99.52	49.76	50.00	49.88	0.000

Summary of RF4 dataset HoE classification results (Base Features).

	A	P	R	F	K
AP	47.62	47.83	47.84	47.83	-0.043
AT	36.90	18.45	50.00	26.95	0.000
PL	39.91	41.12	40.25	40.68	-0.172
SE	23.48	11.74	50.00	19.02	0.000

Emotions of Literary Response Decision Tree Classification Results

Summary of Sex-merged dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	83.56	47.27	49.94	48.57	-0.002
EV	57.90	49.79	49.88	49.83	-0.003
NA	47.13	32.36	44.31	37.40	-0.120
OT	97.87	48.94	50.00	49.46	0.000

Summary of Female dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	77.18	50.40	50.42	50.41	0.008
EV	49.11	53.43	52.41	52.92	0.045
NA	57.84	57.31	55.50	56.39	0.115
OT	99.16	49.58	50.00	49.79	0.000

Summary of Male dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	68.98	53.11	53.39	53.25	0.065
EV	39.66	43.04	42.38	42.71	-0.124
NA	37.03	36.23	35.44	35.83	-0.279
OT	96.51	48.26	49.99	49.11	0.000

Summary of RP-merged dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	80.48	50.67	50.37	50.52	0.009
EV	59.14	72.02	56.37	63.24	0.134
NA	38.97	38.97	38.96	38.96	-0.221
OT	94.76	47.38	50.00	48.65	0.000

Summary of RP1 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	65.44	49.86	49.63	49.74	-0.004
EV	27.21	19.69	23.44	21.40	-0.551
NA	23.85	19.49	24.04	21.53	-0.517
OT	57.29	41.53	32.44	36.43	-0.205

Summary of RP2 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	65.40	48.83	47.95	48.39	-0.028
EV	35.87	17.93	50.00	26.39	0.000
NA	68.68	68.45	65.83	67.11	0.330
OT	99.56	49.78	50.00	49.89	0.000

Summary of RP3 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	39.09	46.77	44.55	45.63	-0.053
EV	65.46	72.64	70.63	71.62	0.362
NA	59.16	62.06	60.74	61.39	0.207
OT	69.17	49.48	34.83	40.88	-0.014

Summary of RP4 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	80.98	40.49	50.00	44.75	0.000
EV	55.02	59.13	56.74	57.91	0.130
NA	36.19	32.53	29.66	31.03	-0.364
OT	90.93	50.93	50.06	50.49	0.002

Summary of RF-merged dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	86.70	43.35	50.00	46.44	0.000
EV	63.27	31.64	49.99	38.75	0.000
NA	56.62	51.86	51.75	51.80	0.036
OT	98.91	49.46	50.00	49.73	0.000

Summary of RF1 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	36.74	49.12	48.17	48.64	-0.013
EV	41.04	20.52	50.00	29.10	0.000
NA	48.41	40.68	39.45	40.06	-0.196
OT	99.54	49.77	50.00	49.88	0.000

Summary of RF2 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	31.65	37.12	32.22	34.50	-0.219
EV	38.04	41.35	40.69	41.02	-0.153
NA	78.76	75.81	78.80	77.28	0.536
OT	97.58	48.79	50.00	49.39	0.000

Summary of RF3 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	92.64	46.33	49.99	48.09	0.000
EV	47.27	46.00	45.35	45.67	-0.082
NA	40.26	37.72	23.16	28.70	-0.257
OT	99.35	49.68	50.00	49.84	0.000

Summary of RF4 dataset ELR classification results (Base Features).

	A	P	R	F	K
AE	14.78	7.39	50.00	12.88	0.000
EV	42.27	21.13	50.00	29.71	0.000
NA	62.98	62.09	62.24	62.16	0.243
OT	45.86	48.82	23.19	31.44	-0.022

Appendix G

Summary of Decision Tree, Support Vection Machine, and Multilayer Perceptron Classification Results for the General Profile Datasets (Base Features)

Hourglass of Emotions Classification Results

Summary of Sex-merged dataset HoE classification results (Base Features).

	DT					SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AP	54.53	53.53	53.51	53.52	0.070	52.06	49.12	49.24	49.18	-0.016	50.89	49.94	49.94	49.94	-0.001
AT	54.00	50.82	50.72	50.77	0.015	52.41	48.05	48.40	48.22	-0.034	52.21	49.35	49.40	49.37	-0.012
PL	49.76	48.48	48.58	48.53	-0.029	43.58	43.72	43.68	43.70	-0.125	51.97	51.63	51.64	51.63	0.033
SE	66.51	43.79	46.96	45.32	-0.075	71.24	41.83	48.62	44.97	-0.038	71.62	58.99	55.52	57.20	0.130

Summary of RP-merged dataset HoE classification results (Base Features).

	DT					SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AP	31.55	30.96	34.08	32.45	-0.296	41.68	41.25	41.12	41.18	-0.176	45.53	44.52	44.53	44.52	-0.110
AT	56.43	50.03	50.03	50.03	0.001	53.96	50.77	50.75	50.76	0.015	50.99	42.53	44.42	43.45	-0.121
PL	50.29	50.68	50.61	50.64	0.012	51.30	51.63	51.54	51.58	0.031	47.32	45.73	46.83	46.27	-0.064
SE	76.53	38.27	50.00	43.36	0.000	76.53	88.27	50.01	63.85	0.000	70.57	45.12	47.91	46.47	-0.054

Summary of RF-merged dataset HoE classification results (Base Features).

	DT					SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AP	28.94	23.54	26.75	25.04	-0.478	46.07	46.40	46.37	46.38	-0.071	41.79	42.67	43.06	42.86	-0.134
AT	66.93	62.49	62.14	62.31	0.246	48.82	45.69	45.31	45.50	-0.089	49.82	48.58	48.41	48.49	-0.029
PL	50.41	49.25	49.49	49.37	-0.010	58.47	59.42	57.80	58.60	0.158	53.93	53.76	53.62	53.69	0.073
SE	71.45	49.66	49.93	49.79	-0.002	73.17	55.84	50.20	52.87	0.006	65.66	57.62	58.21	57.91	0.158

Emotions of Literary Response Classification Results

Summary of Sex-merged dataset ELR classification results (Base Features).

	DT					SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AE	83.56	47.27	49.94	48.57	-0.002	83.79	41.90	50.00	45.59	0.000	54.99	56.19	56.17	56.18	0.118
EV	57.90	49.79	49.88	49.83	-0.003	58.10	49.92	49.96	49.94	-0.001	52.84	46.14	47.45	46.79	-0.055
NA	47.13	32.36	44.31	37.40	-0.120	39.17	36.48	37.97	37.21	-0.245	53.95	53.13	53.01	53.07	0.061
OT	97.87	48.94	50.00	49.46	0.000	97.87	48.94	50.00	49.46	0.000	69.49	57.15	55.51	56.32	0.122

Summary of RP-merged dataset ELR classification results (Base Features).

	DT					SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AE	80.48	50.67	50.37	50.52	0.009	85.58	42.79	50.00	46.11	0.000	76.66	49.32	49.41	49.36	-0.013
EV	59.14	72.02	56.37	63.24	0.134	63.29	63.13	62.72	62.92	0.256	57.13	56.77	56.04	56.40	0.123
NA	38.97	38.97	38.96	38.96	-0.221	43.37	42.36	42.95	42.65	-0.142	45.67	45.43	45.49	45.46	-0.091
OT	94.76	47.38	50.00	48.65	0.000	94.76	47.38	50.00	48.65	0.000	94.72	47.38	49.98	48.65	-0.001

Summary of RF-merged dataset ELR classification results (Base Features).

	DT					SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AE	86.70	43.35	50.00	46.44	0.000	86.70	43.35	50.00	46.44	0.000	83.88	51.03	50.35	50.69	0.010
EV	63.27	31.64	49.99	38.75	0.000	63.28	31.64	50.00	38.76	0.000	58.21	46.81	48.40	47.59	-0.037
NA	56.62	51.86	51.75	51.80	0.036	36.60	39.74	39.83	39.78	-0.174	65.33	61.24	59.19	60.20	0.197
OT	98.91	49.46	50.00	49.73	0.000	98.91	49.46	50.00	49.73	0.000	98.91	49.46	50.00	49.73	0.000

Appendix H

Summary of Support Vection Machine and Multilayer Perceptron Classification Results for the General Profile Datasets (PCA Features)

Hourglass of Emotions Classification Results

Summary of Sex-merged dataset HoE classification results (PCA Features).

	SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AP	51.88	48.95	49.10	49.02	-0.019	46.65	44.80	44.97	44.88	-0.102
AT	54.46	49.10	49.38	49.24	-0.013	45.25	44.27	44.15	44.21	-0.116
PL	47.67	48.31	48.34	48.32	-0.033	51.29	50.21	50.20	50.20	0.004
SE	73.54	38.55	49.76	43.44	-0.007	71.52	57.98	54.40	56.13	0.107

Summary of RP-merged dataset HoE classification results (PCA Features).

	SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AP	40.55	40.53	40.36	40.44	-0.190	37.03	34.39	35.10	34.74	-0.303
AT	51.20	37.94	43.02	40.32	-0.156	57.13	52.62	52.26	52.44	0.047
PL	49.54	49.67	49.67	49.67	-0.007	43.98	43.20	44.38	43.78	-0.111
SE	76.53	38.27	50.00	43.36	0.000	75.05	48.73	49.82	49.27	-0.005

Summary of RF-merged dataset HoE classification results (PCA Features).

	SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AP	54.30	56.21	55.81	56.01	0.112	45.93	45.03	45.08	45.05	-0.099
AT	49.60	49.69	49.65	49.67	-0.006	49.65	46.90	46.61	46.75	-0.064
PL	58.16	58.87	57.53	58.19	0.152	48.58	48.42	48.43	48.42	-0.031
SE	72.46	52.48	50.29	51.36	0.008	52.41	42.52	41.81	42.16	-0.156

Emotions of Literary Response Classification Results

Summary of Sex-merged dataset ELR classification results (PCA Features).

	SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AE	83.79	41.90	50.00	45.59	0.000	74.05	45.63	46.67	46.14	-0.075
EV	55.42	45.98	47.61	46.78	-0.053	53.94	49.12	49.22	49.17	-0.016
NA	38.25	36.77	37.41	37.09	-0.255	54.13	53.56	53.41	53.48	0.069
OT	97.88	48.94	50.00	49.46	0.000	97.87	48.94	50.00	49.46	0.000

Summary of RP-merged dataset ELR classification results (PCA Features).

	SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AE	85.58	42.79	50.00	46.11	0.000	79.24	47.76	48.66	48.21	-0.033
EV	62.66	62.89	61.67	62.27	0.237	62.06	64.01	60.41	62.16	0.215
NA	43.06	42.10	42.66	42.38	-0.148	39.62	38.84	40.12	39.47	-0.195
OT	94.76	47.38	50.00	48.65	0.000	93.22	47.34	49.19	48.25	-0.024

Summary of RF-merged dataset ELR classification results (PCA Features).

	SVM					MLP				
	A	**P**	**R**	**F**	**K**	**A**	**P**	**R**	**F**	**K**
AE	86.70	43.35	50.00	46.44	0.000	86.68	43.35	49.99	46.43	0.000
EV	63.28	31.64	50.00	38.76	0.000	62.41	54.22	51.33	52.74	0.032
NA	37.26	40.31	40.29	40.30	-0.167	54.99	58.82	58.99	58.90	0.156
OT	98.91	49.46	50.00	49.73	0.000	98.91	49.46	50.00	49.73	0.000

Appendix I

Resource Persons

Ms. Ethel Chua Joy Ong
Thesis Adviser
Software Technology Department
College of Computer Studies
De La Salle University - Manila
ethel.ong@delasalle.ph

Dr. Judith J. Azcarraga
Assistant Professor
Computer Technology Department
College of Computer Studies
De La Salle University - Manila
jay.azcarraga@delasalle.ph

Dr. Shirley O. Lua
Director
Bienvenido N. Santos Creative Writing Center
De La Salle University - Manila
shirley.lua@dlsu.edu.ph

Mr. Roy Francis Navea
Assistant Professor
Electronics and Communications Engineering Department
Gokongwei College of Engineering
De La Salle University - Manila
roy.navea@dlsu.edu.ph

References

Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. *Science(Washington)*, *228*(4698), 456–462.

Azcarraga, A., Talavera, A., & Azcarraga, J. (2016). Gender-Specific Classifiers in Phoneme Recognition and Academic Emotion Detection. In *International Conference on Neural Information Processing* (pp. 497–504). Springer.

Azcarraga, J., & Suarez, M. T. (2012, September). Predicting academic emotions based on brainwaves, mouse behaviour and personality profile. In P. Anthony, M. Ishizuka, & D. Lukose (Eds.), *PRICAI 2012: Trends in Artificial Intelligence* (pp. 728–733). Springer Berlin Heidelberg. (DOI: 10.1007/978-3-642-32695-0_64)

Bechara, A., Damasio, H., & Damasio, A. R. (2000, March). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307. doi: 10.1093/cercor/10.3.295

Bishop, C. M. (1995). *Neural networks for pattern recognition.* Oxford university press.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59.

Breuel, T., & Shafait, F. (2010). Automlp: Simple, effective, fully automated learning rate and size adjustment. In *The Learning Workshop.* Utah.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, *45*(1), 16.

Cambria, E., Livingstone, A., & Hussain, A. (2012). The Hourglass of Emotions. *Cognitive Behavioural Systems*, 144–157. doi: 10.1007/978-3-642-34584-5_11

Cassell, J. (2000). *Embodied conversational agents.* MIT press.

Ciarrochi, J. V., Chan, A. Y. C., & Caputi, P. (2000, March). A critical evaluation of the emotional intelligence construct. *Personality and Individual Differences*, *28*(3), 539–561. doi: 10.1016/S0191-8869(99)00119-1

Clynes, M., & Menuhin, Y. (1977). *Sentics: The touch of emotions.* Anchor Press

Garden City, NY.

Cowie, R., Douglas-Cowie, E., Savvidou, S., McMahon, E., Sawey, M., & Schrder, M. (2000). 'FEELTRACE': An instrument for recording perceived emotion in real time. In *ISCA tutorial and research workshop (ITRW) on speech and emotion.*

Cupchik, G. C., Oatley, K., & Vorderer, P. (1998). Emotional effects of reading excerpts from short stories by James Joyce. *Poetics*, *25*(6), 363–377.

Doherty, M. J. (2008). *Theory of mind: How children understand others' thoughts and feelings.* New York, NY, USA: Psychology Press.

Ekman, P. (1972). Universal and cultural differences in facial expression of emotion. In *Nebraska symposium on motivation* (Vol. 19, pp. 207–284). University of Nebraska Press Lincoln.

Freire, P., & Slover, L. (1983). The importance of the act of reading. *Journal of Education*, 5–11.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10–18.

Hannon, B., & Daneman, M. (2009, September). Age-related changes in reading comprehension: An individual-differences perspective. *Experimental Aging Research*, *35*(4), 432–456. doi: 10.1080/03610730903175808

Hofmann, M., & Klinkenberg, R. (2013). *RapidMiner: Data mining use cases and business analytics applications.* Chapman and Hall/CRC.

Iser, W. (1972). The reading process: A phenomenological approach. *New Literary History*, *3*(2), 279–299.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 6). Springer.

Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, *342*(6156), 377–380.

Kleinginna Jr., P. R., & Kleinginna, A. M. (1981, December). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, *5*(4), 345–379. doi: 10.1007/BF00992553

Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, *50*(7), 1211–1221.

Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T.-L., Jeng, S. K., Duann, J. R., & Chen, J. H. (2010). EEG-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, *57*(7), 1798–1806.

Lotte, F., Congedo, M., Lcuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based braincomputer interfaces. *Journal of neural engineering*, *4*(2), R1.

Mar, R. A., Oatley, K., Djikic, M., & Mullin, J. (2011). Emotion and narrative fiction: Interactive influences before, during, and after reading. *Cognition & Emotion*, *25*(5), 818–833.

Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications*, *34*(4), 407–428.

Miall, D. S., & Kuiken, D. (1994, August). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics*, *22*(5), 389–407. doi: 10.1016/0304-422X(94)00011-5

Miall, D. S., & Kuiken, D. (2002). A feeling for fiction: Becoming what we behold. *Poetics*, *30*(4), 221–241.

Morris, J. D. (1995). Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. *Journal of Advertising Research*, *35*(6), 63–68.

NeuroSky, I. (2009). *Brain Wave Signal (EEG) of NeuroSky, Inc.* Author

Nie, D., Wang, X. W., Shi, L. C., & Lu, B. L. (2011). EEG-based emotion recognition during watching movies. In *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on* (pp. 667–670). IEEE.

Oatley, K. (1995, January). A taxonomy of the emotions of literary response and a theory of identification in fictional narrative. *Poetics*, *23*(12), 53–74. doi: 10.1016/0304-422X(94)P4296-S

Phillips, L. H., MacLean, R. D., & Allen, R. (2002). Age and the understanding of emotions neuropsychological and sociocognitive perspectives. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*(6), P526–P530.

Picard, R. W. (1997). *Affective computing*. Cambridge, Massachusetts, United States: Mit Press.

Picard, R. W. (2000). Toward computers that recognize and respond to user emotion. *IBM Systems Journal*, *39*(3.4), 705–719.

Picard, R. W. (2003). Affective computing: challenges. *International Journal of Human-Computer Studies*, *59*(1), 55–64.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *23*(10), 1175–1191.

Plutchik, R. (2001). The Nature of Emotions. *American Scientist*, *89*(4), 344–350.

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, *5*(1), 31.

Rossetti, A. O., & Laureys, S. (2015). *Clinical neurophysiology in disorders of consciousness: Brain function monitoring in the ICU and beyond*. Springer.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*(3), 185–211.

Schwarz, N. (2000, July). Emotion, cognition, and decision making. *Cognition & Emotion*, *14*(4), 433–440. doi: 10.1080/026999300402745

Shashua, A. (2009). Introduction to machine learning.

Tompkins, J. P. (1980). The reader in history: The changing shape of literary response. *Reader-Response Criticism: From Formalism to Post-Structuralism*, *201*.

Vezzali, L., Stathi, S., Giovannini, D., Capozza, D., & Trifiletti, E. (2015). The greatest magic of Harry Potter: Reducing prejudice. *Journal of Applied Social Psychology*, *45*(2), 105–121.

Welling, M. (2011). A first encounter with Machine Learning. *Irvine, CA.: University of California*, 1–93. doi: 10.1.1.441.6238

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*(8), 2031–2046.

Yazdani, A., Lee, J.-S., Vesin, J.-M., & Ebrahimi, T. (2012, March). Affect Recognition Based on Physiological Changes During the Watching of Music Videos. *ACM Transactions on Interacive Intelligent Systems*, *2*(1), 7:1–7:26. doi: 10.1145/2133366.2133373

Zeng, Z., Hu, Y., Fu, Y., Huang, T. S., Roisman, G. I., & Wen, Z. (2006). Audio-visual emotion recognition in adult attachment interview. In *Proceedings of the 8th International Conference on Multimodal Interfaces* (pp. 139–145). New York, NY, USA: ACM. doi: 10.1145/1180995.1181028

Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2007). A survey of affect recognition methods: Audio, visual and spontaneous expressions. In *Proceedings of the 9th International Conference on Multimodal Interfaces* (pp. 126–133). New York, NY, USA: ACM. doi: 10.1145/1322192.1322216