EMMANUEL AWUNI KOLOG

This thesis contextualises the application of human language technologies for counselling. To this end, an e-counselling system – EmoTect – has been developed for the automatic detection and analysis of students’ emotions. A life story corpus has been built to train and test the EmoTect classifier. The corpus is made freely available for research purposes. Additionally, this work has demonstrated how an e-counselling system, built with machine learning capabilities, is trained and used based on individual user’s perception of emotions.
CONTEXTUALISING THE APPLICATION OF HUMAN LANGUAGE TECHNOLOGIES FOR COUNSELLING
CONTEXTUALISING THE APPLICATION OF HUMAN LANGUAGE TECHNOLOGIES FOR COUNSELLING
Author’s address: Emmanuel Awuni Kolog  
University of Eastern Finland  
School of Computing  
P.O. Box 111,  
FI-80101 JOENSUU, FINLAND  
email: emmanuk@uef.fi

Supervisors: Senior Researcher Calkin Suero Montero, PhD.  
University of Eastern Finland  
School of Computing  
P.O. Box 111  
FI-80101 JOENSUU, FINLAND  
email: calkin.montero@uef.fi

Professor Markku Tukiainen, PhD.  
University of Eastern Finland  
School of Computing  
P.O. Box 111,  
FI-80101 JOENSUU, FINLAND  
email: markku.tukiainen@uef.fi

Professor Erkki Sutinen, PhD.  
University of Turku  
Department of Future Technologies  
FI-20014 TURKU, FINLAND  
email: erkki.sutinen@utu.fi

Professor emerita Marjatta Vanhalakka-Ruoho, PhD  
University of Eastern Finland  
School of Educational Sciences and Psychology  
P.O. Box 111,  
FI- 80101 JOENSUU, FINLAND  
email: marjatta.vanhalakka-ruoho@uef.fi

Reviewers: Professor Venter Isabella, PhD  
University of Western Cape  
Department of Computer Science  
P.O. Box 7535,  
BELLVILLE, SOUTH AFRICA  
email: iventer@uwc.ac.za
Associate Professor Hugo Jair Escalante
Balderas, PhD
National Institute of Astrophysics, Optics and Electronics, INAOE
Department of Computer science
72840, PUEBLA, MEXICO
email: hugojair@inaoep.mx

Opponent:  Professor Alexander Gelbukh, PhD
Instituto Politécnico Nacional
Center for Computer Research
Av. Juan Dios Batiz,
Nueva Industrial Vallejo,
07738, MEXICO CITY, MEXICO
email: gelbukh@cic.ipn.mx
Kolog, Awuni Emmanuel
Contextualising the Application of Human Language Technologies for Counselling
Joensuu: University of Eastern Finland, 2017
Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences 2017; No. 281
ISBN: 978-952-61-2591-6 (print)
ISSNL: 1798-5668
ISSN: 1798-5668
ISBN: 978-952-61-2592-3 (PDF)
ISSN: 1798-5676 (PDF)

ABSTRACT

While efforts are being made to effectively integrate information and communication technologies into personal-social counselling, this thesis contextualises the application of human language technologies for counselling delivery. With this in mind, a web-based e-counselling system has been developed for aiding counsellors in their decision making of students. The system, called EmoTect, is multi-functional and comprises two components: contact counsellor and emotion detection. The ‘contact counsellor’ allows students to contact counsellors anonymously through text, and the textual submissions are then passed on to the ‘emotion detection’ phase for the automatic classification of emotions and sentiments.

Design science research was employed for developing the EmoTect system. Therefore, preliminary studies were first conducted to gather the needed requirements from the end-users for the implementation of EmoTect. The idea was to understand the needs of counsellors regarding emotion and personal-social decisions of students. Having gathered the needed requirements, the researcher was able to establish and further design an architecture for the EmoTect system. The technical core of EmoTect was developed using multi-class supervised support vector machine learning classifier. In that regard, an annotated life story corpus of students was built and used as training and testing of the classifier. Unlike the traditional approach of using all-in-one inter-annotation agreement gold standard training data for classifier training, the training of the SVM classifier, in this work, is based on each individual user’ perception of emotions. Therefore, EmoTect allows users to tag emotion and sentiment categories to unlabelled training data, based on their own perception of emotions, through the EmoTect interface before starting to use the system.

The EmoTect classification algorithm was evaluated with sample stories from the life story corpus to ascertain its efficacy. In addition, the final version of EmoTect was demonstrated with counsellors, teachers and students from three senior high schools in Ghana. Data regarding the functionalities, ease of use and impact of EmoTect in
counselling were collected from the participants. Results show that students and counsellors are willing to adopt e-counselling to support counselling delivery despite the challenges associated with its implementation in Ghana. In addition, the EmoTect classification algorithm achieved comparable accuracy to that achieved with a gold standard even when presented with unknown data. Moreover, counsellors and students’ curiosity was piqued about the capabilities of EmoTect, which made them express their desire to adopt it for counselling delivery. The users found the various functionalities of the system suitable, but expressed concern about the poor internet connection in Ghana, which is a potential challenge to the use of EmoTect. Toward this end and based on the outcome of this work, the researcher provides recommendations and guidelines for the implementation of e-counselling in Ghana.

**Universal Decimal Classification:** 004.85, 004.89, 004.912, 004.93, 159.942, 316.613.4, 364.624

**Library of Congress Subject Headings:** Information technology; Computational linguistics; Natural language processing (Computer science); Text processing (Computer science); Automatic classification; Supervised learning (Machine learning); Support vector machines; Corpora (Linguistics); Social service; Counseling; Counselor and client; Decision making; Decision support systems; Emotions; Ghanaian students; Africa

**Yleinen suomalainen asiasanasto:** tietotekniikka; kieliteknologia; tekstinlouhinta; luokitus; koneoppiminen; korpukset; sosiaalityö; neuvoa; päätöksenteko; päätöksentukijärjestelmät; tunteet; opiskelijat; Ghana; Afrikka
ACKNOWLEDGEMENTS

My profound gratitude goes to Prof. Erkki Sutinen who poached me during my master’s degree program. He took it upon himself, as a full professor, to supervise my master and PhD theses. Aside his busy schedules, he managed to devote a considerable amount of time for me throughout my studies. His comments, ideas and criticisms throughout my PhD studies has been a key to my success. Prof Sutinen introduced me into Natural Language Processing and flooded me with lots of ideas in the field. Prof Marjatta Vanhalakka-Ruoho was contacted by Prof Sutinen to help in the supervision of my master’s thesis, of which she gladly accepted the challenge and guided me with her experience in counselling. My master’s degree thesis was very successful, where I managed to obtain an excellent grade. As my PhD thesis is geared towards using NLP for counselling, her guide in the aspect of counselling was very instrumental. I say thank you for your contribution.

I am indebted to Dr. Calkin Suero Montero who accepted to supervise my work when I approached her. Ever since, she has been very helpful in our last three publications. Before a paper is submitted she makes sure that thorough effort is put into it. This is because she believes in targeting high quality forums so, she does not accept shoddy work. Her criticism, comments and ideas throughout this dissertation is highly appreciated.

Prof. Markku Tukiainen is a co-supervisor and acting head of the educational technology research group. He acts swiftly on any administrative request that I make. He was very instrumental in our last paper and provided insightful criticism and ideas in the final dissertation. I really appreciate his effort and relish to work with him in the future should the opportunity presents itself. I am equally grateful to Dr. Jarkko Suohon, who happens to be the coordinator for the Doctoral program, assisted me with every information I requested from him. He gave swift response and advice to me when I needed it. He contributed to my second paper and I feel deeply appreciated.

My indebtedness goes to the University of Eastern Finlad Foundation for awarding me a 2-year full research grant. The grant actually helped me to cope with the high cost of living in Finland and being able to attend conferences to present my research works. The effort of the preliminary examiners of this dissertation, Prof. Venter Isabella and Associate Prof. Hugo J. Escalante, is deeply appreciated.

I dedicate this work to my lovely late mother: Tiyempoka Awuni who worked hard to secure a better future for me. Her greatest desire was to witness my Doctoral graduation but she left unexpectedly to be with her God while I was compiling this dissertation. My father’s dedication to my course has been splendid and encouraging. He supported me financially and emotionally throughout my education. My wife -Betty Sumboh, and my two kids: Philippa Awuni and Emmanuel Awuni Jnr are gifts from God. Their love, understanding and dedication
while I was always busy studying for my PhD and MBA simultaneously, is deeply appreciated. I am also indebted to my siblings for the support.

Finally, my profound gratitude goes to Linus Atarah, a fellow Ghanaian, and who turned out to be a long lost relative, for being instrumental in proof-reading and patching the language potholes in all my publications and this dissertation as well. I am grateful and deeply appreciate your effort. Personalities such as Derrick Nii Odartey Lamptey, Samuel Adjei Konadu, Michael Osei Barima, counsellors and students from the various senior high schools contributed, in one way or the other, to this work. And I say thank you all for your efforts. I am glad to selflessly share the glory with all of you.

Joensuu, 1st September 2017
Emmanuel Awuni Kolog
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AASCB</td>
<td>American Association of State Counseling Boards</td>
</tr>
<tr>
<td>ACA</td>
<td>American Counseling Association</td>
</tr>
<tr>
<td>AC</td>
<td>Affective Computing</td>
</tr>
<tr>
<td>BDT</td>
<td>Behavioural Decision Theory</td>
</tr>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>BAC</td>
<td>British Association for Counselling</td>
</tr>
<tr>
<td>C</td>
<td>Counsellor</td>
</tr>
<tr>
<td>CL</td>
<td>Computational Linguistics</td>
</tr>
<tr>
<td>CORE</td>
<td>Council of Rehabilitation Education</td>
</tr>
<tr>
<td>CRCC</td>
<td>Commission of Rehabilitation Counselor Certification</td>
</tr>
<tr>
<td>CS</td>
<td>Computer Science</td>
</tr>
<tr>
<td>DSR</td>
<td>Design Science Research</td>
</tr>
<tr>
<td>EIC</td>
<td>Emotion-imbued Choice</td>
</tr>
<tr>
<td>e-NLP</td>
<td>Educational Natural Language Processing</td>
</tr>
<tr>
<td>GES</td>
<td>Ghana Education Service</td>
</tr>
<tr>
<td>HLT</td>
<td>Human Language Technology</td>
</tr>
<tr>
<td>HTML</td>
<td>Hypertext Mark-up Language</td>
</tr>
<tr>
<td>IAA</td>
<td>Inter-Annotation Agreement</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>ISEAR</td>
<td>International Survey on Emotion Antecedent and Reaction</td>
</tr>
<tr>
<td>JSON</td>
<td>JavaScript Object Notation</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>MOOC</td>
<td>Massive Open Online Courses</td>
</tr>
<tr>
<td>SEAL</td>
<td>Social and Emotional Aspect of Learning</td>
</tr>
<tr>
<td>MySQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>NBCC</td>
<td>National Board for Certified Counselors</td>
</tr>
<tr>
<td>NERIC</td>
<td>National Education Reform Implementation Committee.</td>
</tr>
<tr>
<td>NLG</td>
<td>Natural Language Generation</td>
</tr>
<tr>
<td>NLU</td>
<td>Natural Language Understanding</td>
</tr>
<tr>
<td>NLP</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>NLTK</td>
<td>Natural Language Tool Kit</td>
</tr>
<tr>
<td>NRC</td>
<td>National Research Council Ghana</td>
</tr>
<tr>
<td>PAD</td>
<td>Pleasure, Arousal and Dominance</td>
</tr>
<tr>
<td>PD</td>
<td>Participatory Design</td>
</tr>
<tr>
<td>POS</td>
<td>Part-Of-Speech</td>
</tr>
<tr>
<td>RQ</td>
<td>Research Question</td>
</tr>
<tr>
<td>SA</td>
<td>Sentiment Analysis</td>
</tr>
<tr>
<td>SHS</td>
<td>Senior High School</td>
</tr>
<tr>
<td>SMO</td>
<td>Sequential Minimum Optimisation</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>TAM</td>
<td>Technology Acceptance Model</td>
</tr>
<tr>
<td>TF-IDF</td>
<td>Term frequency-Inverse Document Frequency</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modelling Language</td>
</tr>
<tr>
<td>UTAUT</td>
<td>Unified Theory of Acceptance and Use of Technology</td>
</tr>
<tr>
<td>UCM</td>
<td>Use Case Model</td>
</tr>
<tr>
<td>WEKA</td>
<td>Waikato Environment for Knowledge Analysis</td>
</tr>
</tbody>
</table>
LIST OF ORIGINAL PUBLICATIONS

This thesis is based on data presented in the following articles, referred to by the Roman Numerals I-V.


OTHER RELATED PUBLICATIONS

The author, in addition to the original publications used in this dissertation, has also contributed to the following related publications:


AUTHOR’S CONTRIBUTION

The general contribution of this dissertation is the use of NLP techniques in developing automated emotion and sentiment analysis system to support counselling of students. However, preliminary research was conducted in the study’s context (Ghana) to understand and elicit requirements for the system’s development. And these partly led to [PI], [PII], [PIII] and [PIV]. Therefore, the major part of the contribution of this dissertation is based on the original publications [PI - PV].

All the original publications selected for this dissertation were led by the author under the supervision of Prof. Markku Tukiainen, Prof. Erkki Sutinen, Dr. Calkin Suero Montero and Prof. Marjatta Vanhalakka-Ruoho. Prof. Erkki Sutinen and Prof. Marjatta Vanhalakka-Ruoho were instrumental in the supervision of the publications of [PI] and [PII]. Dr. Jarkko Suhonen partly assisted in the supervision of [PII] concerned with/in relation to students motivational intent to use e-counselling and for that matter, educational NLP applications. The expertise of Dr. Calkin Suero Montero was indispensable in the work of [PIII], [IV] and [PV] that had to do with emotions and sentiments analysis. Dr. Calkin Suero Montero contributed to the ideation stage of EmoTect development and subsequently supervised the drafting of those papers ([PIII], [IV] and [PV]).

Apart from the author, the EmoTect development received a considerable ideas from Dr. Calkin Suero Montero who worked hand in hand with the author to implement the requirements. Derrick Nii Odartey Lamptey, who was introduced into NLP by the author, also implemented some part of the paltform as his master degree thesis. The experiments conducted to ascertain the efficacy of the EmoTect classifier were conducted by the author. Lastly, the contextual evaluation of EmoTect, in Ghana, was conducted by the author and assisted by Samuel Adjei Konadu.
# CONTENTS

**ABSTRACT** ................................................................. 7

**ACKNOWLEDGEMENTS** .................................................. 9

**1 INTRODUCTION** .......................................................... 19

1.1 Background and motivation .............................................. 20
1.2 Research questions ......................................................... 23
1.3 Structure of the thesis ....................................................... 25

**2 REVIEW OF LITERATURE** .............................................. 27

2.1 Counselling and e-counselling ......................................... 27
  2.1.1 Emotion and personal-social counselling ...................... 28
  2.1.2 e-Counselling .......................................................... 30
  2.1.3 Challenges of e-counselling implementation ................. 31
2.2 Emotion and decision making ............................................ 33
  2.2.1 Theories of emotions ................................................. 34
  2.2.2 Modelling emotions .................................................. 38
  2.2.3 Influence of emotions on decision making .................... 41
2.3 Human Language Technology ........................................... 44
  2.3.1 Text mining ............................................................. 44
  2.3.2 Natural Language Processing ........................................ 48
  2.3.3 Text-based emotion detection approaches .................... 50
  2.3.4 Related works in emotion classification ....................... 55
  2.3.5 Related works in sentiment analysis ............................. 57
  2.3.6 Evaluation of text classifier ........................................ 59
  2.3.7 The role of NLP in education ....................................... 62

**3 RESEARCH DESIGN** ....................................................... 65

3.1 Design science research ................................................... 65
3.2 Context and method ........................................................ 67
  3.2.1 Counsellors’ profiles ............................................... 68
  3.2.2 Requirement elicitation .............................................. 68
3.3 Ethical considerations ...................................................... 70

**4 IMPLEMENTATION** ....................................................... 71

4.1 EmoTect development .................................................... 72
  4.1.1 Corpus building and annotation ................................... 73
  4.1.2 Classifier training phase .......................................... 75
  4.1.3 EmoTect’s classification phase .................................... 76
4.2 EmoTect user guide ......................................................... 77

**5 EVALUATION AND RESULTS** ........................................ 83

5.1 EmoTect’s classification .................................................. 83
  5.1.1 Classifier performance evaluation ............................... 83
  5.1.2 Results from the classifier evaluation ........................... 86
5.2 EmoTect in contextual use ............................................... 90
1 INTRODUCTION

While in recent times information and communication technology (ICT) is being utilised efficiently to improve the life situation of humankind, little can be said about research in ICT-mediated counselling in Ghana, especially in the area of human language technologies. The application of human language technology (HLT), mainly for speech and text processing, is useful in counselling, as it supports more effective ways of exchanging information. With HLT, counsellors have the opportunity to communicate efficiently with their students.

This dissertation contextualises the application of HLT in Ghanaian school counselling services, aiming to enhance the decision making of students regarding their emotion and personal-social development. HLT is a developing interdisciplinary field that encompasses most sub-disciplines of linguistics, as well as computational linguistics, natural language processing (NLP), computer science, artificial intelligence, psychology, philosophy, mathematics and statistics. NLP has, in recent times, attracted much attention from related research communities, given its numerous practical applications, such as emotion and sentiment detection. These practical applications are useful in areas such as counselling and customer touchpoint analysis in business organisations.

Emotion constitutes the most basic form of communication and interaction among students, and between students and counsellors. Hence, counsellors can understand and communicate effectively with students once their (students’) emotional behaviours are construed. The subjective and subtle nature of emotion in text makes Shivhare and Khethawat (2012) believe the text-based expression of emotions is difficult to detect by computational approaches. Conversely, Miner (2012) argues that, despite the challenges associated with the morphological and linguistic analysis of text, detecting emotions in text is possible and has been a success in recent years though there are still some challenges to contend with (see Section 2.3.2).

In this dissertation, an NLP system is developed for the automatic detection of emotion and sentiment in text. The system is called EmoTect1 and is intended to complement the work of counsellors by enhancing their decision-making regarding students’ emotion and personal-social development. The EmoTect system has two functional components: a contact counsellor webform and emotion detection. The ‘contact counsellor’ webform allows students to contact their counsellors anonymously through text. The textual content of students’ submission is then passed on to the emotion detection part for the automatic classification of emotions and sentiments. The system also extracts emotion keywords from students’ textual submissions and further outputs the outcome to the system interface for counsellors. With the use of EmoTect, the emotional and sentimental changes of students over a selectable period

1 The system is available online at: http://nlp4counselling.com/.
could be monitored through a visualisation chart. The system uses a supervised support vector (SVM) machine learning classifier. Accordingly, a life story corpus collected from students was used to train and test the multi-class SVM classifier which forms part of the system evaluation.

In this research work, a design science research (DSR) is employed. Since the study involves an artefact creation (EmoTect) to solve a contextual problem, end-users (students and counsellors), at some stages of the implementation were consulted (thus: a participatory design process). Particularly, the end-users were made to participate in the requirement elicitation and evaluation phases of the EmoTect implementation. For instance, prior to EmoTect’s development, several empirical studies were conducted with the end users to gather requirements for EmoTect’s development. Hence, the first four of the original publications partly explore the requirements needed for EmoTect’s development.

1.1 BACKGROUND AND MOTIVATION

Research exploring ICT integration in education is considerably on the ascendancy. Most of the available research in educational technology is focused on improving teaching-related activities (Huang et al., 2016), school administrative activities (Wang & Yang, 2009) and workload management (Webb, 2006). However, less study has been conducted to integrate ICT into non-regular academic activities, such as counselling. Since education concerns the holistic development of the individual, then ‘the goal of education is not solely a cognitive knowledge of the facts, but also includes development of social and emotional maturity’ (Abroad & Kolb, 2011: p. 300). The relevance of counselling in education cannot be overlooked since school counsellors are meant to rely partly on their expertise with interpersonal relationships to help students to understand themselves and be able to adapt to the environment in which they find themselves.

The academic success of students, which usually produces tremendous delight in their parents, is hugely dependent on their emotion and personal-social stability. The stability of one’s emotional state is foundational to students’ academic and career achievement (Arbona, 2000; Daly et al., 2002; Elias et al., 2003). Therefore, the stability of one’s emotions acts as a catalyst in promoting the academic achievement of students (Jaeger & Eagan, 2007). Considerable research has been conducted on the influence of emotions on academic achievement (see, for examples: Valiente et al., 2012; Kannan & Miller, 2009). A typical example of such research is Liew’s (2012), who presented a succinct review of how emotions influence students’ academic achievement. Liew (2012) found emotion to be a key influence on students’ academic achievement and pointed out the need for counsellors to carefully understand students’ emotional changes and behaviours.
Even though counsellors are directly involved in counselling work, Low and Nelson (2005) believe that teachers also play an equally important role in guiding students towards academic success. In Ghana, most counselling committees established to regulate counselling activities in schools include teachers because they have more contact hours with students. Counsellors and teachers hold the responsibility of helping students to be aware of and manage their emotions and thus improve their inter- and intra-personal relationships. Mohzan (2013) investigated the influence of emotional intelligence on academic and occupational achievement. The researchers concluded that emotional intelligence influences academic achievement and improves students’ relationships with their teachers.

Until recently, not much attention has been paid to counselling students in Ghana, which has been a rather worrying situation because the sector is sensitive and a majority of students are adolescent. Adolescent students are prone to turbulent emotions that interfere with their academic work (Langelier, 2012). Counsellors are therefore mandated, as part of their profession, to respond to or help students manage their emotional challenges. In as much as this thesis is not advocating for a complete paradigm shift towards automatic technology-oriented counselling, complementing the work of counsellors with an automatic emotion and sentiment analysis system would make their work more efficient. Therefore, given the nature of counselling in schools, the relevance of automatically tracking the emotions of students in textual submissions should not be overlooked.

While many students in Ghana are reluctant to seek face-to-face counselling (Awinsong et al., 2015; Kolog et al., 2014), others do not even recognise the importance of counselling in their academic development (Awinsong et al., 2015). The reason could be attributed to the fact that students do not trust their counsellors (Kolog et al., 2015; Inman et al., 2009). Students fear divulging or sharing their sensitive and personal information with counsellors (Kolog et al., 2014), which is why many students favour anonymous counselling (Glasheen et al., 2013; Kolog et al., 2015). Although anonymous counselling may be the relatively less preferred choice of counselling by some counsellors concerning their engagement with students, students prefer it because of the comforting knowledge that they have nothing to lose in case their private data gets leaked; i.e., it cannot be traced to them.

Counselling in the Ghanaian schools has existed since 1975, but the sole method used by counsellors is face-to-face. Only on a few occasions is email used to contact students and their parents (Kolog et al., 2015). Quite recently, counselling has taken place through the social media platform called WhatsApp (Kolog et al., 2015). Research has revealed the counselling of students enhances their academic achievements (see Payton et al., 2008; Zins et al., 2004). However, counselling in schools is bedevilled with numerous problems concerning the challenges associated with its implementation (see Section 2.1.3). Fox and Butler (2007) believe that students have not been adequately informed about the availability and relevance of counselling in schools. NWOKOKOLO ET AL. (2010) INVESTIGATED COUNSELLING IN NIGERIA AND
discovered that most schools in Nigeria do not have counselling centres. Nwokolo et al. (2010) attributed this to a lack of awareness creation among students. In Tanzania, Kano (2012) found that counselling exists in schools but lacks the needed resources for the sector to flourish as it should.

Awinsong et al. (2015), in Ghana, have suggested that counselling in schools should place an emphasis on helping students to stabilise their emotional and personal-social challenges. However, given the nebulous (implicit) nature of emotions and personal-social counselling, school counsellors are faced with the herculean task of helping students manage their emotions. To introduce flexibilities and efficient ways of understanding emotions in text, several computational approaches have been proposed (see Section 2.3.3). With this, academic disciplines with differing interests have developed computational algorithms for tracking emotions and sentiments in text over many years. Some researchers are focusing their attention on how to explore and devise means of optimising the existing computational algorithms for classifying emotions in text, while others are also leveraging the existing approaches to natural language processing (NLP) in non-computer science (CS) disciplines such as counselling. Most notable of the CS areas concerned with emotion and sentiment analysis are affective computing (AC), computational linguistics, human computer interaction (HCI) and NLP.

In Ghana, when designing a system to recognise emotions in text, consideration should be given to its application and contextual aspects. On the one hand, counsellors can use such systems to monitor the changing emotional and academic trends of their students over a period of time. These systems can be used further to facilitate the decision-making process of students. For example, Munezero et al. (2013) developed a system that tracks sentiments and emotions from students’ learning diaries. The essence of their system is to assist teachers to evaluate the attitudes of their students towards their teaching methodologies and the retention rate of students. This helps teachers in decision making by, perhaps, prompting them to modify their teaching methodology. On the other hand, as in the contextual perspective, the text-based medium of information exchange is the most common and suitable means of information exchange in Ghana. Therefore, text-based emotion and sentiment analysis, for now, is an appropriate and perhaps efficient approach for sentiment/emotion extraction. This is also because of the poor state of internet connections in Ghana, which for instance do not allow fluent voice communication between counsellors and students. This informs the reason for the dominance of text-based information exchanges. Counsellors would appreciate the use of such systems if they adequately participated in the design and development of the application, as observed in this dissertation.
1.2 RESEARCH QUESTIONS

Since this study is set to explore a flexible way of helping counsellors to provide counselling service to students regarding emotion and personal-social development, Table 1.1 shows the framing research questions outlined to help realise the goals of this dissertation.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1 What are the emotional life challenges that threaten students’ academic pursuit?</td>
<td>PI</td>
</tr>
<tr>
<td>RQ2 What counselling technologies are being used in Ghana’s senior high schools?</td>
<td>PI</td>
</tr>
<tr>
<td>RQ3 What are the factors that motivate students to adopt and use e-counselling in Ghana?</td>
<td>PII</td>
</tr>
<tr>
<td>RQ4 Does the emotional state of counsellors influence their emotional perception while annotating emotions in text?</td>
<td>PIII, PIV</td>
</tr>
<tr>
<td>RQ5 How can a text-based emotion and sentiment classification system be constructed for counselling delivery?</td>
<td>PV</td>
</tr>
<tr>
<td>RQ6 How can a text-based emotion and sentiment classification system for counselling be evaluated?</td>
<td>PV</td>
</tr>
</tbody>
</table>

In Paper I [PI], RQ1 and RQ2 are investigated. The intended purpose of RQ1 was to explore the emotional challenges that students in Ghana encounter during their studies. This was examined from a collection of students’ life stories. Students were tasked to write about their life stories subjectively. The key essence of the research questions was to partly understand whether the stories of students could trigger sufficient emotions and attitudinal content to warrant using the stories to develop a corpus to train the supervised classifier in this study (i.e. a multi-class SVM). In addition, these researchers investigated the existing counselling technologies that are being used in the senior high schools (SHSs) for counselling service delivery. A counsellor and selected students were tasked to complete questionnaires, which were thematically analysed in line with qualitative research methods. Although the study was limited to one school, further studies were conducted afterwards in different SHSs in Ghana to ascertain the existing ICT tools that counsellors use to provide counselling services. This paper was geared towards developing an e-counselling system with an interest in game design. However, as the study progressed, the scope of the thesis was changed to lean more towards developing an automated e-counselling system for emotion and sentiment classification without the game component. The game component shall be considered in the future.
Paper II [PII] aimed at investigating the behavioural intention of students to adopt the use of e-counselling in Ghana, and that formed the basis for formulating RQ3. The unified theory of acceptance and use of technology (UTAUT) model with four constructs was adopted as independent variables while ‘behavioural intention’ was the dependent variable. The independent variables (UTAUT constructs) are performance expectancy, effort expectancy, facilitation condition and social influence (Venkatesh et al., 2003). The hypothesis was formulated while investigating which among the UTAUT constructs was/were the most important factor(s) for students to use and accept e-counselling in Ghana. The results influence the development of EmoTect as presented also in [PV].

Paper III [PIII] provided an answer to RQ4, which sought to investigate the influence of counsellors’ own emotions on their emotion perception while analysing the emotions of their students in text. This was established through the computation of intra- and inter-counsellors’ annotation agreement of emotions in text documents collected from students and a sample ISEAR (International Survey on Emotional Antecedents and Reaction) corpus. The findings strengthened the need to allow users of EmoTect to label and train the EmoTect classifier based on their own perception of emotions. In addition, in [PIII], the researchers used Plutchik’s (1980) eight basic emotions as baseline emotions. Plutchik’s emotions were confirmed as the baseline after a focus group discussion with selected counsellors from the SHS.

Paper IV [IV] answers RQ4. The paper is an expanded version of [PIII]. The two papers are closely related. However, further empirical studies were conducted in [PIV]. Apart from the results presented in [PIII], the level of disagreements in the emotion perception among the selected counsellors was explored in [PIV], and this led to the emotion category, which counsellors found difficult to annotate. Additionally, based on the findings, this researcher discussed the role of emotion and sentiment analysis in counselling, thereby retrospectively introducing the EmoTect system.

Paper V [PV] was compiled to provide answers to RQ5 and partly for RQ6. The EmoTect development process and its architecture are elaborated upon in [PV]. The paper provided an answer to RQ5. The part of [PV] that answered RQ6 is the evaluation of the EmoTect prototype and the classifier evaluation. However, the contextual evaluation of EmoTect is presented in this dissertation which also forms part of the answers to RQ6. The prototype evaluation was to ascertain from the users if EmoTect met the requirements gathered prior to its development. This formed part of the formative evaluation of the work. As evaluation of a classifier is an important aspect of NLP, and is strongly recommended in Design science research (DSR) project, the classification algorithm was evaluated to confirm how well the algorithm works to achieve its intended purpose. The findings from [PI]–[PIV] influenced the development of EmoTect, and this led to [PV].


2
1.3 STRUCTURE OF THE THESIS

This dissertation is organised into seven chapters. The first chapter introduces the subject matter of this dissertation, objectives of the research and grounds for discussing the background and motivation to employ NLP for counselling. In addition, research questions, which encapsulate all the published papers and the manuscript [PV], are outlined in the first chapter.

Chapter two is a literature review. Given the interdisciplinary nature of this dissertation, related literature from counselling, e-counselling and HLT are discussed. This enabled the author to define text mining and NLP, which are a subset of HLT. Emotion and its influence on decision making are elaborated upon in this chapter. Various processes, methods and techniques of text classification are discussed in the chapter as well. The web-based nature of EmoTect prompted the need to explore and discuss the challenges of e-counselling implementation in Ghana. The literature review considered the original publications used in this dissertation as well.

Chapter three is the research design. The research context and data collection procedure are discussed in this chapter. This takes due cognisance of the original publications used in this dissertation. The holistic approach of employing design science research is discussed in Chapter three. This chapter is drafted in line with the original publications of this dissertation. The initial requirements for EmoTect and the selected counsellors' profiles are outlined in this section.

Chapter four covers the implementation of EmoTect. This considers the various processes/procedures followed in designing and developing the EmoTect system. A case study of the use and architecture of EmoTect is presented in this section as well. The assembly of the corpus used in training and testing the EmoTect classifier is outlined in Chapter four in line with the various research methods. Much of the work in this chapter was adapted from the [PV] as it forms part of this dissertation.

The EmoTect system’s evaluation and results is presented in Chapter five. The evaluation encompasses the EmoTect classification algorithm and the contextual evaluation of EmoTect in the environment with end-users. The results arising from both the evaluation strategies are also presented in this chapter.

Chapter six is the discussion of the results and their implication. The discussion is presented in line with the literature. Therefore, the general discussion, implications and contributions of this dissertation are presented as well. Lessons learnt, constraints and limitations of the three-year study are presented in this chapter. The chapter also outlines guidelines and recommendations for stakeholders of education about the effective implementation of e-counselling to meet the needs of students.

The last and final chapter is the conclusion as it is captured in Chapter seven. The chapter concludes the entire work by providing answers to the research questions formulated in Chapter one. In addition, the chapter briefly outlines and discusses the major contribution of this dissertation.
2 REVIEW OF LITERATURE

As this study is interdisciplinary, this chapter is divided into three main sections. In the first section, literature from counselling and e-counselling is analysed. The second section includes a discussion of emotion and decision making. Since the study is particularly focused on emotion and sentiment classification, some existing works on emotions and their influence on decision-making are presented. The third and last section relates to HLT, which discusses text mining, natural language processing (NLP) and emotion classification in text.

2.1 COUNSELLING AND E-COUNSELLING

The term counselling is broad and has prompted various definitions from different scholars. There are also prominent counselling associations who have defined the term in accordance with their context of how the word should be understood. The British Association for Counselling and Psychotherapy (BAC)\(^3\) was the first professional counselling association to produce a definition of counselling in 1986. The BAC in 1986 defined counselling as a ‘skilled and principled use of relationship to facilitate self-knowledge, emotional acceptance and growth and the optimal development of personal resources’ (Gladding, 2004). Since then, several definitions have popped up, and Glenn (2015) observed that the myriad definitions of counselling stem from the diversity of counselling approaches or methods. Glenn (2015) proceeded to claim that counselling is rooted in the theoretical perspective, depending on the problems or challenges that one encounters or reports to a counsellor. Viewed from a different perspective, the American Counseling Association (ACA) in 1997 defined counselling as ‘the application of mental health, psychological or human development principles, through cognitive, affective, behavioural or systemic interventions, strategies that address wellness, personal growth, or career development, as well as pathology’. Other prominent counselling associations that have produced definitions of counselling are the American Association of State Counseling Boards (AASCB)\(^4\), the National Board for Certified Counselors (NBCC)\(^5\), the Council of Rehabilitation Education (CORE)\(^6\) and the Commission of Rehabilitation Counselor Certification (CRCC)\(^7\).

\(^3\)http://www.bacp.co.uk/
\(^4\)http://www.aascb.org/aws/AASCB/pt/sp/home_page
\(^5\)http://www.nbcc.org/
\(^6\)http://www.core-rehab.org/
\(^7\)https://www.crccertification.com/
Counselling in Ghanaian senior high schools (SHSs) was established in 1975, though work on it had begun already in 1960 (Essuman, 2001), with the ultimate aim of helping students in their academic development. Counselling in the school system is viewed as a tool for eliminating ignorance in young people to encourage holistic life development (Essuman, 2001). Thus, counselling aims to provide opportunities for people to develop holistically and satisfactorily (Gladding, 2004). A holistic model of counselling is an integrative counselling approach that combines a wide range of counselling techniques that focus on the well-being of individuals (Kolog et al., 2014). Although many students do not recognise the role of counselling in schools, it is an indispensable resource for helping students to achieve academic success (Carey & Harrington, 2010; Ramakrishnan & Jalajakumari, 2013).

In Ghana, SHS systems are divided into two streams: boarding and day school. The boarding school system is residential, and students and teachers reside on one campus; teachers take the role of parents during that period. Rules and regulations are designed to manage the behaviour of students. The students in the boarding system remain in the school for the entire duration of their education with intermittent holidays provided according to the division of the academic year. It is widely acknowledged that the advantage of the boarding school system is that it provides students with a peaceful learning environment without distractions by other sources of pleasure or harmful pursuits. However, one major shortcoming of the boarding system is that it is invariably more expensive, not only because of higher school fees but also because of room and board. In that regard, the boarding system is usually beyond the reach of students from economically disadvantaged backgrounds. The day school system, on the other hand, is non-residential, and students commute to school daily on school days. To some parents, the non-residential system is preferable because it allows them to closely monitor and influence the study patterns of their children. The major challenge with the day system is the cost of transportation to commute students daily. In both school systems, counselling is needed to help students to develop holistically.

2.1.1 Emotion and personal-social counselling

Emotion and personal-social counselling are interrelated and form part of school counselling activities. The goal of the emotion and personal-social wing of counselling is mostly geared towards orienting and helping students to develop their emotional intelligence (Jaeger & Eagan, 2007). In other words, emotion and personal-social counselling deals with assisting students to understand themselves and devise strategies to manage their emotions. Additionally, emotion and personal-social counselling fosters a sense of association and coexistence among students (Ofsted, 2007).

The school is a social unit which can be comprised of students from different cultural, religious and social backgrounds (Littlechild, 2012; Atria et al., 2007).
Therefore, counselling in the school system helps to promote the coexistence of students irrespective of their background. For instance, a religious-based conflict could easily erupt from a power struggle to make one religion dominate all others. An example can be illustrated from Nigeria, where Ushe (2015) has pointed out that there has been a spread of violent conflicts between Christians and Muslims in some schools in Nigeria.

Understanding the self and others is necessary for students' development in the short or long term. Given this necessity, students who decide to enter the work force after school or those who would like to continue their academic career need to understand themselves and learn to respect other values in coexistence. For instance, Quinn (2015) recognises emotional intelligence as one of the core competences that managers seek from employees. Aylward (2003: p. 33) believes that ‘companies recognise the importance of having a workforce of people with key transferable skills and attributes such as initiative and enterprise’. In educational settings, Ofsted (2007) viewed personal-social counselling as nebulous, attributed to its subjective nature, and that there are no curricula outlining how to deal with the personal-social development of students. With this challenge, counsellors use their intuition based on their psychological acumen to understand the emotional changes of their students.

Five elements of emotion and personal-social development in school counselling have been suggested by researchers who developed SEAL (Social and Emotional Aspect of Learning) in 2007: self-awareness, self-management, social awareness, relationship skills and responsible decision making (Devaney et al., 2006; Ofsted, 2007). Self-awareness encompasses awareness of our own emotions, moods and, perhaps, situations that define our behaviour. Self-management deals with how teachers or counsellors help students to be aware of and manage their feelings or emotions. Social awareness is concerned with managing feelings when interacting with others and demonstrating conflict resolution skills in social settings. While relationship skills are concerned with the management of one’s coexistence with others, responsible decision making is the ability to select choices from alternatives in line with one’s feelings. The aforementioned SEAL elements are geared towards helping students get to know themselves and be able to find effective solutions to their problems by making decisions based on a consideration of ethical standards, safety concerns, appropriate social norms and respect for others. This, in effect, encourages emotional stability and perhaps academic success. The five elements of emotional and personal-social counselling are presented in Figure 2.1.
2.1.2 e-Counselling

ICT-mediated counselling is also known as e-counselling (Tate et al., 2013) and online counselling (Shiller, 2009). In Ghana, e-counselling is viewed as ‘a digital form of receiving supportive counselling either through an exchange of emails or live webcam session over the internet’ (Kolog et al., 2015b: p. 3). Numerous attempts have been made to improve the delivery of counselling using ICT. Most of the existing counselling technologies are adopted and are even developed for global use, such as Skype and email. There is the need for counselling tools that are developed from indigenous ideas to allow users to appreciate their usefulness in solving indigenous challenges. Although the awareness creation on the use of ICT-mediated counselling has lately been intensified, Bambling et al. (2008) and Ralls (2011) believe that eliminating face-to-face counselling is dangerous and that both e-counselling and the traditional face-to-face counselling methods need to coexist to ensure efficiency in counselling work.

From this work’s preliminary study in Ghana, it was found that students are aware of the existence of counselling in schools, yet they have expressed reluctance to seek counselling face to face. This finding is contrary to that of Fox and Butler (2007), who undertook a similar study in the United Kingdom, where they found students were eager for counselling, especially with e-counselling tools. Many students prefer anonymous counselling to avoid the fear of stigmatisation if their private information leaked into the public domain. In addition, geographically isolated students and students with physical disabilities are not able to access face-to-face counselling in most cases. Furthermore, students who are on vacation and needing help do not have to be physically present to receive counselling. Likewise, counsellors can deliver counselling from any location, provided there is an internet connection. With e-counselling, remote students have the opportunity to be in contact with counsellors synchronously or asynchronously. Synchronous is real-time communication, such as chat and video conferencing, while asynchronous comes with a time lag in communication, such as email.

Counsellors may be living in the same academic environment with students, but when it comes to students sharing their secrets, a careful approach is required. It is nearly impossible for students to share their secret with a person they do not trust.
If trust becomes an issue for students to meet counsellors face to face, then students either keep their personal difficulties to themselves or try to find solace in anonymous counselling. For this reason, many students find safety within the anonymity of technology and feel comfortable sharing their privacy.

Some school counsellors take to online public platforms, where many students have access, to discuss and share ideas to help students to manage their life challenges. Counsellors post resources and information on these platforms that encourage students to feel that it is normal to seek help. Usually, this is done through personalised platforms and social media where the security levels are high. NLP technologies have made it possible to trawl for relevant data from these platforms for further processing to help with decision making. An example is a web-based system developed by Rahman and Ahmed (2014) for automatic emotion detection in social media content or event monitoring. Their system, termed MediaTagger, is open-source software that crawls for web content based on a keyword query and further extracts emotion keywords from the content.

2.1.3 Challenges of e-counselling implementation

In Ghana, implementing e-counselling in schools is bedevilled with several challenges. In this section, the researcher presents and discusses some existing challenges in e-counselling’s implementation in Ghana. Discussing the challenges of e-counselling implementation in this thesis is crucial, given that EmoTect is a web-based platform.

*Poor or no internet connection:* Poor or no internet connections in Ghana are a serious challenge not only in technology-based counselling delivery but also in other sectors that use the internet for their operations like e-learning and school administrative activities. Some schools in Ghana are located in rural communities where internet connections are very poor. While some schools are faced with a poor internet connection, schools in the rural areas, in particular, cannot access the internet at all. The situation of poor internet connections or the non-availability of the internet does not only pertain to rural communities but also some parts of the urban areas. With this challenge, streaming or uploading videos online becomes slow, which highlights the dominance of text-based information exchanges. Since EmoTect is a web-based system, it requires that users are connected to the internet. Therefore, a poor or no internet connection makes it difficult to implement in some schools. Although Ghana encourages private sector development, the government has a major role in the regulation of internet service providers in the country (De Heer-Menlah, 2002). A major setback is the government’s abuse of power. De Heer-Menlah (2002) raises concerns about the abuse of government influence in the acquisition of an operating licence. De Heer-Menlah (2002) has revealed that it can take anywhere from a month to years to acquire a license to operate the internet providing business in Ghana.
High cost of internet: The cost of an internet connection is another issue confronting most schools in Ghana. Given the economic condition of Ghana (a developing country), it is not surprising to note that the cost of an internet connection is more expensive than in Finland (a developed country) and most of the developed countries, if not all. While in 2017 a monthly connection for unlimited data access in Finland cost approximately 17 euros (85 cedis), the same quantity of data cost approximately 35 euros (175 cedis) in Ghana.

Counsellors’ lack of technical knowledge: A core challenge facing the integration of ICT into counselling delivery is that counsellors in Ghana are often not proficient in ICT and therefore unable to use it to provide student counselling. In that regard, counsellors may not necessarily adopt ICT-mediated counselling even if the infrastructure happens to exist. The problem could be traced to school counsellors not having acquired the needed training at the university level. Given this challenge, the IT coordinators of the various schools are recommended to help the counsellors with the initial configuration of EmoTect before use.

High cost of software and hardware: Software and hardware infrastructure is expensive to acquire. Often, schools, from internally generated funds, prefer to invest in acquiring ICT infrastructure for aiding teaching and learning instead of setting up counselling centres with a state-of-the-art ICT infrastructure. Tutu-Danquah (2016) pointed out that school administrators and students are aware of the relevance of ICT-mediated counselling, yet they would rather prioritise an ICT-mediated teaching and learning infrastructure. Many schools have attributed the bias to a lack of government support in the counselling sectors. Head teachers often complain of inadequate financial support from the government to finance the sector: therefore, the difficulty in resourcing schools with the needed infrastructure. Buttressing this claim is a counselling coordinator from Ghana’s education service, Bridgette Nzima-Mensah, who in 2016 called on the government to respond to their needs in terms of the infrastructure deficit in the counselling sector8.

Lack of poor electricity supply: The use of ICT in schools is largely determined by the availability of electricity which ultimately affects the use of ICT-based counselling, but Ghana is perennially plagued by a power crisis that affects every sector of the country. In most instances, schools in the rural communities of Ghana do not have electricity or are not connected to the national electricity grid. Naturally, e-counselling needs electricity for operationalisation. For instance, computers dedicated to keep students’ records and documents can only be powered if there is an electricity

connection. It is also a challenge for remote students to contact counsellors synchronously or asynchronously if there is no steady flow of electricity. Ghana’s electricity is not steady and has attracted a serious political reaction from the populace.

*Poor maintenance culture:* Even in instances where ICT-based counselling has been successfully established, sustaining the infrastructure through regular maintenance almost always emerges as a major problem. This issue is accentuated by a poor maintenance culture in Ghana where the public infrastructure, once procured, is often neglected and no longer receives financial attention to sustain it. The other cause is a lack of expertise in schools to maintain hardware and upgrade software, as well as an underlying factor concerning the lack of human resources in the ICT sector as a whole.

*Ghana Education Service directive:* Students in the various boarding houses are often not allowed to use mobile devices in SHSs, an adverse policy since many students may need the devices to search for information online and, most importantly, to communicate with counsellors. Some of the reasons advanced for banning the use of mobile devices in schools are that students use those devices to chat with their peers at odd times (such as lecture times) on social media and to watch adult recreation videos online, as well as a sense of students becoming addicted to mobile devices. While these reasons may be based on sound judgement, there is the need to balance these with the mental welfare that counselling services provide to students. The banning of the use of mobile devices in Ghana diverges from the 2004 policy framework of the government on ICT in education, which states:

> “As part of the mission to transform the educational system to provide the requisite educational, and training services and an environment capable of providing the right types of skills and human resource required for the developing and driving Ghana’s information and knowledge-based economy and society, the Government is committed to a comprehensive programme of rapid development, utilization and exploration of ICT within the educational system from primary school upward.” (Ghana National Report on basic Education-Ghana, 2004: p.18).

### 2.2 EMOTION AND DECISION MAKING

Given the scope of this dissertation, decision making is a process that requires one to make a choice from available options. It is essential, therefore, to weigh the negatives and positives of all the available options that would produce a logical outcome. Selecting a choice from equally important alternatives highlights the difficulties in decision making. To make an informed, good decision, one should have the capacity to make a forecast of the eventual consequences of all the available options, compare
them and select a choice based on rewards and liabilities. Emotion, on the other hand, is a feeling characterised by the state of mind (Shelke, 2004). Emotion in decision making constitutes what Lerner et al. (2014) has described as harmful and sometimes beneficial to decision making. This is because emotions sometimes take precedence over humans’ actions and influence their decision-making process (Baumeister et al., 2010). Conversely, Mitchell (2011) argues that conceptualising emotions and decision making together has not been well established despite the profound influences that emotion has on decision making. However, emotions and decision making are a dynamic, iterative process which aims to help individuals adapt to their environment. Paulus and Yu (2011) recognise that the success of helping students adapt to their environment is rooted in the internal states of mind of individual students, their personal characteristics and determinants of the valuation process, as well as the characteristics of the environment.

Emotion and decision making has widely been studied in psychology. It has emerged from research that many psychology scholars consider emotions to be a dominant driver for most meaningful decisions in life (Keltner et al., 2014; Lazarus, 1991; Loewenstein et al., 2001; Ekman, 2007). Lerner et al. (2014: p. 4) pointed out that ‘decisions serve as the conduit through which emotions guide everyday attempts at avoiding negative feelings (e.g., guilt, fear, regret) and increasing positive feelings (e.g., pride, happiness, love), even when we lack awareness of these processes’. In the subsequent section, the researcher presents some discussions on the influence of emotions in decision making.

2.2.1 Theories of emotions

Theories of emotions have long been formulated by several researchers depending on the schools of thought and the discipline. Many scholars have tried to correlate emotions with cognition since the era of Plato through the 19th century, when theories of emotions and cognition attracted much attention (Lazarus, 1966). There are over 150 existing theories of emotion that have connection with human behaviour (Strongman, 1978). The most prominent and applied ones in AI which are discussed in this section are: appraisal, James-Lange, Cannon-Bard theory and Schachter-Singer theory. These theories have been found to impact significantly the study and understanding of emotions and cognition.

Appraisal theory

Appraisal theory is focused on exploring the evaluation that individuals make about an event, situation or stimulus that causes emotional reactions (Petta & Gratch, 2009; Aronson et al., 2005). An emotional reaction or elicitation at a particular time can change one’s perception of or choice to make a decision. For instance, if a student’s perception of a teacher is positive, then he or she might feel emotions such as joy, trust and anticipation concerning the teacher because he or she had appraised the
teacher positively. Appraisal theory is strongly grounded in history, when Plato and Aristotle took interest in it. Arnold (1950) and Lazarus (1991) also spearheaded research in appraisal theories of emotions after its formulation in 1940. Now, the theory is predominantly used in psychological and computational fields due to its advance of demonstrating the connection between emotion and cognition (Marsalla et al., 2015). Some appraisal theorists emphasise discrete emotional categories – such as joy and anger – rather than continuous emotions (Petta & Gratch, 2009).

Figure 2.2 represents a component model view of the computational appraisal process adapted from Marsella et al. (2015). Marsella et al. (2015) have asserted that appraisal theory is the most used computational model of emotions and that the model has widely been used by AI researchers in the construction of AI artefacts. The figure presents a computational appraisal architecture which consists of different components that are interlinked and operate together. The information flow in the figure is cyclic, which has received criticism from appraisal theorists like Lazarus (1991), Scherer (2001) and Parkinson (2009). However, there are other appraisal theorists who have agreed to the architecture in Figure 2.2. Marsella et al. (2015: p. 61) summarise the model as ‘some representation of the person-environment relationship is appraised, this leads to affective response of some intensity; the response triggers behavioural and cognitive consequences; these consequences alter the person-environment; this change is appraised and so on’. The various components in the appraisal component view, shown in Figure 2.2, are described below:

**Person-environment relationship**: this component, as the name implies, is an imaginary representation of one’s (agent) relationship with his or her environment (Lazarus, 1991). Marsella et al. (2015: p. 62) has explained that ‘this representation should allow an agent in principle, to derive the relationship between external events – real or hypothetical – and the beliefs, desire, intentions of the agent or other significant entities in the (real or hypothetical) social environment’.

**Appraisal-derivation model**: transforms the person-environment relationship into an appraisal variable (Smith & Kirby, 2009). For example, Marsella et al. (2015: p. 62) postulate that ‘if an agent’s goal is potentially thwarted by some external action, an appraisal-derivation model should be able to automatically derive appraisals that this circumstance is undesirable, assess its likelihood, and calculate the agent’s ability to cope, i.e., by identifying alternative ways to achieve this goal’.

**Appraisal variables**: The appraisal variables are a set of judgements that an agent uses to produce different emotional responses, and this is generated from the appraisal derivation model.

**Affect-derivation model**: it lies between the appraisal variable and the affective state. After determining appraisal pattern, it functions as specifying how an individual or an agent will react.

**Affect-intensity model**: is closely related to the affect-derivation model. Hence its function is to specify the strength of emotional response from the specific appraisal.
Emotion/affect: the current emotional state of the individual or that of the agent after the aforementioned process.

Affect-consequent model: ‘maps affect (or its antecedents) into some behavioural or cognitive change’ (Marsella et al., 2015: p. 62).

James-Lange theory

In the early 19th century, James Williams and Carl Lange theorised that emotions are triggered based on physiological reactions to situations (Scherer, 1999). And this led to James-Lange theory. James-Lange theory was inspired by Appraisal theory which rather focused on the triggers of emotions based on the effects of reactions. James argued that ‘Instead of feeling an emotion and subsequent physiological (bodily) response, the theory proposes that the physiological change is primary, and emotion is then experienced when the brain reacts to the information received via the body’s nervous system’ (Cannon, 1972). For example, suppose you are walking in the woods, and you see a grizzly Lion. There is the likelihood that you begin to tremble, your heart begins to pound and perhaps tries to run from the Lion. The James-Lange theory proposes that you will interpret your physical reactions and conclude that you are frightened. In effect, the reaction triggers fear as demonstrated in the Figure 2.3. Since the inception of James-Lange theory, it has received a lot of criticisms where Cannon was a well-known critic of James-Lange theory (Cannon, 1972).
Cannon-Bard theory

Walter Cannon and Philip Bard, both psychologists, were critical of James-Lange theory. Together, the two proposed another theory of emotion in the 1920s, referred to as Cannon-Bard theory. Since Cannon-Bard theory was propounded based on the criticism of James-Lange theory, Cannon-Bard’s theory, instead, suggests that physiological reactions, such as crying and trembling, are caused by emotions. In essence, the theory is backed by neuro-biological science. Hence, the main concepts of the Cannon–Bard theory are that emotional expression results from the function of hypothalamic structures, and emotional feeling results from stimulations of the dorsal thalamus. More specifically, it is suggested that emotions occur when the thalamus sends a message to the brain in response to a stimulus, resulting in a physiological reaction. Therefore, contrary to James-Lange theory, arousal and emotions occur simultaneously. For instance, in Cannon–Bard theory, one reacts to a stimulus and experiences the associated emotion at the same time. For example, imagine that you are walking to your car through a darkened parking garage. You hear the sounds of footsteps trailing behind you and spot a shadowy figure slowly following you as you make your way to your car. According to the Cannon-Bard theory of emotion, you will experience feelings of fear and a physical reaction at the same time. You will begin to feel fearful, and your heart will begin to race.
**Schachter-Singer theory**

*Schachter-Singer* theory is also referred to as the ‘two-way factor theory of emotions’. The reason for the other name is because emotion is thought to occur based on two factors: physiological arousal and its cognitive label. The theory was propounded by Stanley Schachter and Jerome E. Singer, hence the name *Schachter-Singer theory*. As seen in Figure 2.5, the theory states that when an emotion is felt, a physiological arousal occurs and a person uses the immediate environment to search for emotional cues to label the physiological arousal (Schachter & Singer, 1962; Cotton, 1981). For example, when one encounters a deadly animal like a lion in the forest, the person will shake and tremble, sweat or even run away. Based on the emotional arousal that will occur, the person may shout. The key setback to the Schachter-Singer theory is that one may misinterpret the emotions based on the body’s physiological state. “When the brain does not know why it feels an emotion it relies on external stimulation for cues on how to label the emotion (Dutton & Aron, 1974).

![Schachter-Singer Two-factor Theory](https://www.thinglink.com/scene/513367126631776257)

**Figure 2.5. Demonstration of Schachter-Singer Theory (Redrawn from: Psychological science, https://www.thinglink.com/scene/513367126631776257)**

### 2.2.2 Modelling emotions

Several computational models of emotions have long been formulated to explain human emotions. However, the widely used models of emotion in AI are the *categorical* and *dimensional* emotion models. Arguments which has loomed for several years about emotion representation is whether categorical emotions can represent all other existing emotions or whether there are dimensions of emotions that can represent all other emotions. Many researchers in psychology have taken a stand as to which of the emotional divergent is worth considering. For instance, Russell & Barrett (1999) have strongly opposed basic emotions and taken a strong stand in dimensional representation of emotions. But on the other side of the debate, Plutchik (1980) and Ekman (1999) have stood by categorical emotions and further proposed some basic emotions which represent all emotions irrespective of the context.
The categorical model of emotions is a well-known model used in text, speech and computer vision. However, this work focuses on text classification. The categorical model marks text under discrete categories, such as basic emotions. The categorical emotion model is often associated with supervised machine learning where documents are labelled automatically (Gupta et al., 2010). Examples of categorical emotional models are the widely proposed basic emotions that have been used in the labelling of text documents. The study of basic emotions dates back to the early 1960s, when Darwin tried to reveal the original forms of human emotional behaviour (Griffiths, 2002). Basic emotion theorists claim that basic emotions exist, and Silvan and Tomkins (1962) first concluded eight basic emotions: surprise, interest, joy, rage, fear, disgust, shame and anguish. However, there are other basic emotions that have been proposed by scholars in the field. Table 2.1 shows other basic emotions and their proponents. In text mining, Ekman (1999) and Plutchik’s (1980) basic emotions have widely been used in emotion detection experiments. On the one hand, Ekman (1999) proposed six basic emotions: anger, disgust, fear, happiness, sadness and surprise. On the other hand, Plutchik’s basic emotions subsume Ekman’s basic emotions. The inclusion of trust and anticipation adds up to form Plutchik’s basic emotions. The wheel assisted in the understanding and interpretation of each of the basic emotions while annotating the students’ life stories with Plutchik’s basic emotions.

<table>
<thead>
<tr>
<th>Proponents</th>
<th>Basic emotions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plutchik’s (1980)</td>
<td>Anger, disgust, fear, happiness, sadness and surprise, trust, anticipation</td>
</tr>
<tr>
<td>Levenson et al. (1992)</td>
<td>Happiness, Sadness, Anger, Fear, Disgust</td>
</tr>
<tr>
<td>Ekman (1999)</td>
<td>Anger, disgust, fear, happiness, sadness and surprise</td>
</tr>
<tr>
<td>Frijda (1989)</td>
<td>Joy, Sadness, Anger, Fear, Surprise Regret, Relief, Hope</td>
</tr>
<tr>
<td>Izard (1994)</td>
<td>Joy, Sadness, Anger, Fear, Surprise, Disgust Shame, Interest</td>
</tr>
<tr>
<td>Damasio et al. (2000)</td>
<td>Happiness, Sadness, Anger, Fear</td>
</tr>
<tr>
<td>Vytal &amp; Hamann (2010)</td>
<td>Happiness, Sadness, Anger, Fear, Disgust</td>
</tr>
<tr>
<td>Phan et al. (2002)</td>
<td>Happiness, Sadness, (Anger), Fear, (Disgust)</td>
</tr>
</tbody>
</table>
Figure 2.6. Plutchik’s wheel of basic emotions (Plutchik, 1980)

**Dimensional model**

The dimensional emotion model started to be studied in the 19th century where Wilhelm Max Wundt, described as modern Psychologist, proposed that emotions can be described in three dimensions: "pleasurable versus unpleasurable", "arousing or subduing" and "strain or relaxation (Wundt, 1987). Since then there has been several proposition of dimensional emotions in psychology.

The dimensional emotion model is geared towards affect in a dimensional space. Kim (2011) asserted that emotions occupy a location in dimensional space. While the categorical model of emotion places emphasis on discrete emotion categories, the dimensional model emphasises continuous emotions such as mood, affect and core affect (Russell, 2003; Barrett, 2006). Dimensional emotions are viewed in 2- or 3-dimensional space. Mehrabian and Russell (1974) postulate that many computational models are built on a three-dimensional ‘PAD’ model, where ‘PAD’ represents pleasure, arousal and dominance. However, Russell’s (1980) 2-dimensional circumflex model (valence–arousal) has been explored widely in text classification. Figure 2.7 is Russell’s (1980) two-dimensional circumflex model of which the valence dimension shows the strength of pleasant and unpleasant emotions. The arousal dimension indicates the activation and deactivation of emotions in a pronged dimension. There are other emotion dimensional propositions with a similar structure. Watson et al. (1988) have developed a similar structure and focused on positive and negative affect, which is basically a 45-degree rotation of pleasure and arousal, and they describe the
same space. As opposed to the basic emotion theorists, dimensional classification uses emotional states, for instance, a valence dimension or positive versus negative emotions.

![Figure 2.7. Russell's circumflex model of affect (Russell, 1980)](image)

**2.2.3 Influence of emotions on decision making**

In recent times, decision-making researchers have noted the role that emotions play in decision-making (Lerner et al., 2015; Gutnik, et al., 2006). Until then, researchers on the subject had overlooked the role that emotion plays in decision-making (Lerner et al., 2003). Decision making was viewed mainly as a cognitive process which could be traced back to *behavioural decision-making theory (BDT)* (Loewenstein & Lerner, 2003). BDT emanated from the late 1960s and gave credence to explaining human behaviour mainly from a cognitive perspective, which merely considered the role of emotions. BDT explains how people make choices under the condition of uncertainty (Haward & Janvier, 2015). However, over a decade ago many strides were made by decision-making researchers regarding decision-making theories. Researchers have recognised the role that emotions play in decision making. Emotion determines human behaviour that triggers options for selection during decision making (Lerner et al., 2003; Loewenstein & Lerner, 2003). With this, researchers have further stressed the importance of environmental, social and emotional influences on decision-making (Medin & Bazerman, 1999; Loewenstein, 1996; Gold, 2000). This needs to be considered while analysing the influences of emotions on decision-making.
Indeed, research has shown that emotion influences decision making (Lerner et al., 2003; Loewenstein & Lerner, 2003). According to Loewenstein and Lerner (2003), emotion has two influences on decisions: expected and immediate emotional influence. The expected emotion influence, also known as anticipated emotions, is viewed as forecasted emotions that are not experienced immediately at the time of making a decision. Loewenstein and Lerner (2003) describe expected emotion as a cognitive process because it involves predicting the future. Expected emotions tend to maximise the effectiveness of decision makers since they are able to predict the emotional consequences of decisions. For example, students who engage in counselling activities might anticipate the positive influences of counselling on their life situation. Furthermore, Lerner et al. (2003) link expected emotional influence to the expected utility model, where people analysing emotions attempt to predict the emotional consequences irrespective of the facts or alternatives available. In line with Lerner et al.’s (2003) observations, counsellors might choose the option that would maximise their expected decision-making outcome at the expense of the alternatives available, so they would be considered responsible counsellors. Usually, humans’ expected or perceived emotions have a better influence which, in effect, transcends the actual reality on the ground. When a counsellor is carried away by such an influence, he or she is likely to make poor decisions, which leads to negative consequences for the client’s life.

Regarding the computational study of emotions, a text corpus for training a supervised learning classifier is often annotated by humans. For this reason, Kolog et al. (2016) investigated the influence of counsellors’ emotions on their judgement while annotating emotions in text [PIII]. The researchers did not delve deeply into the psychological or neurological process but found that counsellors’ emotions do influence their decisions while analysing the emotions of others in text. The study was empirically carried out by computing for the inter- and intra-annotation agreement of emotions by counsellors.

Unlike the expected emotions, immediate emotions are actual emotions experienced at the time of decision-making. An immediate emotion influences decision-making either by carrying information that people use as input for the decision they face by overwhelming deliberative decision making in high intensity or by changing the nature and/or depth of processing (Tiedens & Linton, 2001).

Several models have been proposed to demonstrate how emotions influence decision making. Lerner et al. (2014) have proposed the emotion-imbued choice model (EIC). The model, shown in Figure 2.8, descriptively summarises how emotions influence decisions or choices. According to Lerner et al. (2014), EIC was developed based on the ‘risk-as-feeling model’ proposed by Loewenstein et al. (2001), which is
a model for the determinant and consequence of emotions. Lerner et al. (2014) assert the ‘EIC model assumes that the decision maker faces a one-time choice between given options, without the possibility of seeking additional information or options’. The EIC model attempts to explain the influences of emotions on choices or decisions. In the EIC model shown in Figure 2.8, variables A to I are paths that explain the relationship between emotional variables (see Lerner et al., 2014: p. 815 for more details). As opposed to the expected emotions influence, Lerner et al. (2014) have stated that the EIC model ends at the moment of decision. This implies that the actual outcome or consequences, as a result of the decision, are not included. Lerner et al. (2014) assert the EIC model does not include the reflexive behaviour of humans, such as someone dying upon learning about something unexpected, either through text or any other means.

In the decision-making process, theorists found the need for evaluating options by assessing the utility of each outcome for each option. In Figure 2.8, utility outcomes are combined with characteristics of options (line C) such as probabilities and time delays, and characteristics of the decision maker (line B) like risk aversion and a discount rate. At that moment, the aforementioned factors are combined to form the overall evaluation of each option, and the best option is chosen (shown by line D in the figure). Emotions felt at the decision making in the figure represent current emotions. Green dotted lines in the model show potential sources of current emotions that include incidental influence, characteristics of the decision maker, characteristics of options, the expected outcome and the conscious and non-conscious evaluation. Based on triggers from various sources, current emotions directly influence the evaluation of the outcome (line G) by affecting which dimensions the decision maker focuses on.

Figure 2.8. EIC model for affective influences on decision making (Adapted: Lerner et al., 2014)
2.3 HUMAN LANGUAGE TECHNOLOGY

As this dissertation is geared towards contextualising the application of human language technology (HLT) into developing an e-counselling system, this section delves into discussing text mining and natural language processing in the context of HLT. Techniques and approaches of text classification are discussed as well.

HLT, also referred to language technology (LT), is an interdisciplinary field that comprises diverse academic domains which seek to understand and process human language. The academic domains, as earlier stated in the introductory section, encompass computational linguistics, natural language processing, computer science, artificial intelligence, psychology, philosophy, mathematics and statistics. Usually, in the study of HLT, two or more of the domains are combined to achieve a set goal. For instance, this study takes into account the domain of NLP, psychology (counselling) and linguistics. A German research centre for artificial intelligence\(^9\) view HLT as the composition of computational methods, computer programs and electronic devices that are specialised for analysing, producing or modifying texts and speech. Although there are different ways humans communicate, HLT mainly focuses on text and speech technologies.

This study is geared towards text analytics; hence, the common form of written communication is through printed documents such as newspapers, magazines and books, and in handwritten artefacts found in notebooks and personal letters. Given the relevance of written language in human transactions, its automatic recognition has practical significance to motivate the efficient and effective understanding of human communication. Although existing HLT systems are far from achieving human ability, they have numerous possible applications. With HLT, people have the opportunity to communicate efficiently and in a more natural way. In addition, HLT supports more effective ways of exchanging information and controlling its growing mass.

As HLT helps with the easy interpretation of human language, often, the goal of HLT is to create software products that have some knowledge of human language. They are urgently needed for improving the human-machine interaction since the main obstacle in the interaction between human and computer is the communication problem. Uszkoreit (2016: p. 1) believes that ‘even if the language the machine understands and its domain of discourse are very restricted, the use of human language can increase the acceptance of software and the productivity of its users’.

2.3.1 Text mining

The phrase ‘text mining’, also referred to as text analytics, is broad and consists of a wide range of technologies for analysing text. Text mining is the process of finding patterns in an unstructured or semi-structured natural language text for the purpose

\(^9\)https://www.dfki.de/lt/lt-general.php
of discovering knowledge. The unstructured texts are manipulated into structured text for the prediction or extraction of knowledge. The world’s data are mostly unstructured which highlights the relevance of text mining approaches, especially in the field of education, which is the focus of this dissertation. According to Witten (2005: p. 1), ‘the field of text mining usually deals with texts whose function is the communication of factual information or opinions, and the motivation for trying to extract information from such text automatically is compelling – even if success is only partial’.

Witten (2005) outlines some benefits of text mining in this era of data overload, pointing out that the most common vehicle for formal information exchange is through text (Witten, 2005). Sebastiani (2002) describes text mining as a ‘term used to denote any system that analyses large quantities of natural language text and detects lexical or linguistic usage patterns in an attempt to extract probable useful information’. Text mining became prominent two decades ago, after which the field has attracted overwhelming interest from researchers and practitioners (Miner, 2012). Given the rapid advances of text mining research, Miner (2012) predicted that text mining would top the lists of future trends in analytics. Since then, much research has been conducted from diverse domains to develop efficient and effective text mining approaches.

According to Miner (2012), text mining has seven practice areas. Each of the practice areas are unique, and different computational approaches are used in their analytics. The practice areas are: document clustering, document classification, information retrieval, web mining, information extraction, natural language processing and concept extraction. Each of these areas are unique but strongly interrelated in terms of their applicability. Usually, projects involving text mining require a combination of some or all the practice areas to achieve a desired outcome. For instance, while our e-counselling system uses NLP approaches, the emotion detection part is a classification technique based on the information extraction approach. Outputs from text mining are comprehensively summarised, which appears to form a different area of study called text summarisation. Irrespective of the aforementioned practice areas, text mining appears to place more focus on automatic NLP. Therefore, this thesis is delving deeper into NLP.

As a matter of fact, text mining and data mining are often used interchangeably. Nevertheless, the terms are not the same, as Witten (2005) and Patel and Soni (2012) have pointed out. For this reason, this author discusses briefly the significant difference between them in line with the literature. Both text and data mining is well grounded historically in terms of its evolution. While text mining looks for patterns in text, data mining looks for patterns in data. In more specific terms, Witten and Eibe (2000) have explained that ‘Data mining can be more fully characterised as the extraction of implicit, previously unknown, and potentially useful information from data’. Implicit information encompasses information that is unknown or hidden, or perhaps an automatic technique cannot be applied for extraction. However, in text
mining approaches, the information to be extracted is clearly and explicitly stated in text (Witten, 2005). Jadhav and Gadekar (2004) recognise that text and data mining can also be differentiated from the sources of data used for processing. Jadhav and Gadekar (2004) further pointed out that text mining sources of data are mostly unstructured, while data mining often uses structured sources of data. Figure 2.9 represents the relation between the various practice areas of text mining adapted from Miner (2012).

![Figure 2.9. Text mining practice areas (Adapted from Miner, 2012)](image)

Figure 2.10 represents a general context diagram of text mining. The figure shows three broad phases of text mining: establishing a corpus, pre-processing data and extracting knowledge (Miner, 2012).

**Establishing a corpus** involves the collection of text documents relevant for extracting meaning and knowledge. Text mining data sources may include but not be limited to text documents such as e-books, JSON files, HTML files, blog posts, feedback and reviews, web posts, emails, social media posts (e.g. Facebook, Twitter) and short notes. Miner (2012) emphasise that the quantity and quality of text documents must be important and contribute to text mining projects. The work of some researchers is to gather or collect text data and annotate them for research purposes. There are various types of annotated corpora that are freely available for research purposes. An example of such corpora is an international survey of emotional antecedents and
reactions (ISEAR). ISEAR is an annotated corpus collected from students in 37 different countries which was annotated for research purposes. For instance, sample data from the ISEAR corpus was used in [PIII] and [PIV].

The second phase, which is the pre-processing, is where a structured representation of the corpus is created for further analysis (i.e. knowledge discovery). The text pre-processing phase is the most time consuming one in knowledge discovery, and it undergoes a strict systematic process to introduce a structure into a corpus. However, the complexities of data pre-processing depend on the sources of data. For instance, a huge amount of data relatively contains more data outliers which, in effect, have weak predictive value compared to a small amount of data (Munkova et al., 2013). In text mining approaches, text pre-processing may generally take into account the various stages: data selection, cleaning, creating, integrating and formatting data (Chapman, 2001). In a narrower sense, Miner (2012) suggest that text pre-processing should take into account the following: 1) the scope of the source of text such as a word, sentence or paragraph, 2) tokenisation of the text into discrete words, 3) stemming of the tokenised words after the removal of stopping words and 4) the normalisation of spelling (correct misspelling) and the conversion of words into lower or upper case.

The last phase, which is knowledge extraction, is where novel patterns in text are extracted based on the type of text and the problem or the interest at hand. This phase is accomplished by using structured data from the pre-processing phase. The main categories of knowledge extraction in text mining involve prediction, clustering, association and trend analysis. For instance, in EmoTect, this researcher’s interest is to predict emotions and sentiments in text, thereby forming the basis of understanding human behaviour to inform good decision making in counselling.

---

2.3.2 Natural Language Processing

Natural language processing (NLP) is an area that focuses on the processing of natural languages such as English (Jones, 1994). NLP evolved in 1950 with the original idea of processing human language (Nadkarni et al., 2011). Reshamwala and Mishra (2013) provided a succinct historical review of NLP, as they posited that NLP is grounded in computer science, artificial intelligence and linguistics. Since its evolution, it has flourished into a high-level technology capable of finding, digesting, packaging and representing text-driven information. NLP has a wide array of applications spanning the fields of health, education and business, among others. Outputs from NLP are valuable for decision making, particularly for counselling students, the focus of this thesis.

NLP involves two components: natural language understanding (NLU) and natural language generation (NLG). On the one hand, NLU deals with machine reading comprehension which is basically focused on interpreting an input text fragment. NLU mimics the meaning of written text and produces data or text which encapsulates the meaning (Coupel, 2014). NLG, on the other hand, involves processing and generating useful sentences or phrases in natural language from some internal representation. The various stages associated with NLG are text planning, sentence planning and
text realisation. NLG ‘is revolutionary because it allows us automate tasks in the service industry and make knowledge workers more efficient. Knowledge workers perform at a higher level because they are assisted in their communication with clients’ (Coupel, 2014).

Extracted knowledge from unstructured text, in NLP, can be very useful for decision-making. However, NLP has several shortcomings, and the challenges associated with NLP are also reflected in the implementation of EmoTect. One key challenge is ambiguities in the text language, which may include homonymy (same word with different meanings in context) and synonymy (different words but similar or same meaning). Ambiguities in language make NLP difficult, so the context of language needs to be clarified. Another challenge is misspellings and abbreviations of words. Although several algorithms have been developed to normalise words before extracting meaning from them, some words are still difficult for researchers to write algorithms to correct. Generally, NLP in text undergoes the following stages, also shown chronologically in Figure 2.12.

**Lexical analysis**: is concerned with dividing a chunk of a textual document into a suitable granular such as word, sentence or paragraph for further processing. Identifying and analysing the structure of words in a sentence is also the goal of lexical analysis. A program that performs the lexical analysis may be termed as tokenizer, lexer or scanner.

**Syntactic analysis**: is concerned with the arrangement and ordering of words in a meaningful way. The results from the lexical analysis are exploited to build a structural description of the sentence. This process is referred to as parsing. For instance, ‘Emmanuel ate the food’ is syntactically more accurate than ‘ate food the Emmanuel’. Figure 2.8 contains synthetic tree representations of (Emmanuel (ate (the food))) and (The student (failed (the exams))). With EmoTect syntactic parsing, the POS (Part of speech) tagger parses the input data before stopping words are removed.

![Syntactic parser trees](https://example.com/syntactic_trees.png)

Figure 2.11. Syntactic parser trees for a) Emmanuel ate the food b) The student failed the exams
**Semantic analysis:** is concerned with the meaning of a word in a sentence or paragraph. For instance, *bank* in ‘I was at the river bank’ and ‘I went to the bank to withdraw money’ is not the same. While *bank* in the former sentence refers to a riverside or the seaside, *bank* in the context of the latter sentence refers to a financial institution where deposits and withdrawals of money are made.

**Discourse integration:** is concerned with the meaning of any sentence depending on the sentence before it. In addition, it brings about the meaning of the immediately succeeding sentence.

**Pragmatic analysis:** This stage is the last one that is concerned with deriving those aspects of language which require real-world knowledge.

---

![Diagram](image.png)

**Figure 2.12.** General NLP process for text classification

### 2.3.3 Text-based emotion detection approaches

Emotions in text are not only expressed in a sentence or paragraph form but also in words. Emotional words in text documents may be direct or a combination of words that gives meaning to a particular emotion category such as joy and sadness. Categorical basic emotions are predominantly used in emotion detection projects rather than dimensional emotion categories, which are rarely used. Two main categories of emotion classification in text are involved: *coarse-grained* and *fine-grained*. Coarse-grained classification involves the classification of text into either positive or negative. Sometimes, neutral is included as part of the categories. In fine-grained classification, more specific emotions are required, such as Ekman’s basic emotions. Unlike coarse-grained emotion classification, syntactic and semantic analyses of the text document are required during processing. Detecting emotions in a linguistic unit of text is mostly a multi-class classification problem, while sentiment analysis is, in most studies, a binary classification problem using positive and negative emotional valence. Generally, four main computational methods for emotion classification are identified: *keyword spotting, lexical affinity, learning-based* and *hybrid methods*. Although
this thesis employs a multi-class SVM learning approach in the development of the EmoTect system, the researcher takes interest in holistically presenting the various computational methods of emotion classification in text.

**Keyword spotting method:** This is the easiest and the most intuitive way of extracting emotions from text (Liu et al., 2009). The method uses a keyword pattern matching approach with which occurrences of keywords are identified. In this technique, categorical basic emotions, such as Ekman’s basic emotions, are used. Hence, keywords identified in a text document are classified according to emotion categories. Figure 2.13 represents the steps in a keyword spotting technique. From the figure, an input document in the form of text is first tokenised into individual words for further processing. Emotion words in the tokenised words are identified and detected. At this step, the intensity of the emotion words is analysed. Here, there are existing algorithms developed to determine the intensity of the emotion words, such as term frequency-inverse document frequency (tf-idf). Negation checks are performed and considered in the final output. The last step is to classify the emotions according to a predefined emotion class.

**Lexical Affinity method:** This uses lexical resources, such as bag-of-words, for emotion and sentiment classification in text. Bag-of-words is a commonly used model in document classification in which the frequency or occurrences of each word in a document are used as a feature. Unlike the “keyword spotting method”, the lexical affinity method further assigns a probabilistic affinity for a particular emotion to arbitrary words rather than identifying and picking emotion keywords. The probabilities assigned to the emotion keywords, in this method, are often part of linguistic cor-

Figure 2.13. Keyword-spotting method in NLP
The method, according to Shivhare and Khethawat (2012), has several disadvantages that affect accuracy. Shivhare and Khethawat (2012) explain that the assigned probabilities are biased towards a specific genre of text. Also, since the method operates solely on the word-level, the detection can easily be tricked by sentences such as “I avoided an accident” (negation) and “I saved my father from dearth by an accident” (positive) (Cambria & White, 2014).

**Learning-based:** This method is regarded as the most effective of all the methods. The learning-based method involves the use of machine learning approaches. Machine learning is a technique that teaches computers to learn human natural language through experience. Machine learning algorithms use computational methods to ‘learn’ information directly from data without relying on a predetermined equation as a model. Two main approaches are used in machine learning: supervised and unsupervised.

Supervised learning involves training a classifier with labelled data so the classifier learns and creates a model for prediction when presented with unseen data. By using the supervised machine learning method, let $T$ denote text and $s$ denote a word, sentence and paragraph that contains emotion features, where $s \in T$. In addition, let $n$ be the number of emotion class $E = \{x_1, \ldots, x_n\}$ where $i = 1,2,3,\ldots,n$. In some cases, researchers include ‘non-emotion’ or ‘neutral’ sentiments in the categories $E$. The aim of supervised machine learning in emotion classification is to tag $x_i$ to $s$ as accurately as possible, as in a mapping function $f(s) = x_i$, such that an ordered labelled pair of $(s,x_i)$ is obtained. The mapping is obtained based on a model after a classifier is trained.

Unlike supervised machine learning, unsupervised machine learning does not use training data. Rather, algorithms are used to draw inferences from datasets consisting of input data without labelled responses. Its usage is highly applicable in instances where one is not sure of what information is contained in a text corpus. One example of unsupervised learning is cluster analysis.

Machine learning began in 1959 (Arthur, 1959), but Miner (2012) asserts that machine learning started to be applied in text classification in the 1990s. Since then, several researchers in the field of NLP have used the technology for text analytics. There are various classification algorithms for text classification. The most efficient, accurate and commonly used ones are the support vector machine (SVM), neural network, decision tree and Naïve Bayes (see the works of: Munezero et al., 2013; Hassan et al., 2011; Matwin & Sazonova, 2012).

Text classification undergoes some sort of a process in classifying emotions or sentiments in text. Figure 2.14 represents the supervised machine learning classification process. In summary, a machine learning classifier is made to learn from training data, thereby creating a model. The model then classifies unseen or input data according to the various emotional or sentimental categories.
**Support vector machine**

SVM is a supervised machine learning classification algorithm that is used for both linear and non-linear data. The algorithm is discriminative in the sense that it is defined by constructing a hyperplane or a set of hyperplanes in a high-dimensional space (Hashem & Mabrouk, 2014). A hyperplane in higher-dimensional space is defined as a set of points whose dot products are constant with a vector in that space. The function that performs such a transformation is called the kernel function. There are many possible kernel functions; the most common are: linear, polynomial, sigmoid and radial basis function (RBF). When training data are presented to a SVM, a model is built which consists of data points chosen from input data space and their class labels. SVM outputs an optimal hyperplane which classifies unseen or unclassified data after a model is built. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general, the larger the margin, the lower the generalisation error of the classifier (Hashem & Mabrouk, 2014).

In this work, we used sequential minimal optimisation (SMO) (Platt, 1998) of SVM in the building of the EmoTect system. SMO has widely been used for training SVM in text classification problems. SMO solves a quadratic programming problem that arises during the training of a SVM. The algorithm is iterative for solving an optimisation problem and breaks the problem into a series of the smallest possible sub-problems. The sub-problems are then solved analytically by the algorithm. In each iteration, it solves a quadratic problem of size two. The main problem with the SMO is choosing a good pair of variables to optimise at each iteration (Rifkin, 2012). Figure 2.15 represents a graphical representation of SVM indicating the maximum margin and separation of the hyperplane.
Hybrid-based approach: As the name implies, this is an approach that combines any of the aforementioned approaches in text classification, such as the ‘keyword spotting’ method combined with the ‘learning-based’ method. Different text classification approaches are often combined with the intent of improving the level of accuracy (Villena et al., 2011). A typical example of a hybrid technique for emotion classification was proposed by Yang et al. (2012). Yang et al. (2012) combined lexicon-based keyword spotting, conditional random field emotion cue identification and machine learning-based emotion classification. Regarding the machine learning part, SVM, Naïve Bayes and maximum entropy classifiers were used. The researchers’ technique was integrated in a vote-based system where they evaluated the system with suicide notes as datasets. From their results after evaluation, Yang et al. (2012) achieved a 61% F-score alongside 58% precision and 64% recall.

Figure 2.16 represents a hybrid technique employed by Villena et al. (2011) for text categorisation. Initially, Villena et al. (2011) used a supervised machine learning method to create a model and to classify a text document. After that, they employed a rule-based expert system to post-process and improve the results provided by the machine learning classifier. Villena et al. (2011) suggested that, with this approach, if the recall and precision after the first stage are noisy, the rule-based approach can fine tune the output by adding specific rules to achieve a better result.
2.3.4 Related works in emotion classification

There are various systems that have been developed to detect emotions in text, spanning both structured and unstructured sources of text documents. The available systems are often developed from different computational approaches (see Section 2.3.1). However, in this section, some related works in emotion extraction are reported. Jain and Sandhu (2015) and Kao et al. (2009) provide systematic literature reviews on text-based emotion detection. Jain and Sandhu (2015) focused on reviewing the existing literature in text-based emotion classification which has used machine learning approaches. Jain and Sandhu (2015) presented a more coherent overview of both the supervised and unsupervised learning approaches, including some examples with different algorithms which are widely used. Jain and Sandhu (2015) also presented some works that have used machine learning approaches in emotion detection. In a similar vein are Kao et al. (2009), who provided an overview of emotion detection in text. Kao et al. (2009) investigated the various methods (keyword spotting, learning-based and hybrid) and models (categorical and dimensional) for emotion detection in text.

Burget et al. (2011) have carried out a study to detect emotions in Czech newspaper headlines. The study employed six basic emotions: anger, disgust, fear, joy, surprise and sadness. They collected several newspapers in the Czech language, pre-processed and fed them into a classification algorithm. Their approach placed much emphasis on data pre-processing, which was done at the sentence and word level. The data pre-processing went through the stages of POS tagging, lemmatisation and the removal of stopping words. After that, Burget et al. (2011) applied the term-frequency–inverse document frequency (tf-idf) technique to calculate the weight of each term.
and each emotion category. SVM was used as the classifier with 10-fold cross-validation. After evaluating the classifier, Burget et al. (2011) achieved a comparable overall average value of 80.28% accuracy.

Lu et al. (2006) developed a real-time information retrieval system. The system retrieves data from the web with a simple keyword query. Subsequently, emotions are extracted from the retrieved data and the output visualised. The researchers employed two NLP techniques: a web mining engine and semantic labelling. The web mining engine allows one to search from the web using a crawler, while semantic role labelling is responsible for semantic parsing and the eventual extraction of emotions. The categorical basic emotions of happiness, sadness, anger, fear, disgust, surprise and neutral were used. The researchers achieved good accuracy (70%) with their system as they evaluated with test data.

A related work is an emotion detection system developed by Hannan (2015). Hannan (2015) did not actually use ML or any existing affect lexicons such as WordNet Affect11 in developing the system. However, emotions were extracted from unseen text based on keywords in the content: hence, the keyword spotting technique. Unlike this researcher’s approach, Ekman’s six basic emotions were used as the emotion category. According to Hannan (2015), the system also extracts affect-word phrases and exclamatory keywords from input data. After experimenting with the system and gold standard data obtained from human annotation, Hannan (2015) reported a 77% overall level of accuracy, though the researcher was not clear about how the level of accuracy was computed.

Bentinck et al. (2015) pointed out that emotions expressed in user-generated content like blogs, reviews or social media have had less research in detecting emotions. With this assertion, Bentinck et al. (2015) explain that emotionally charged text from user-generated content may express multiple emotions at the same time. Hence, their research focused on classifying multiple emotions expressed in movie reviews. They collected movie reviews and annotated them with eight basic emotion categories: anger, disgust, fear, interest, joy, love, sadness and surprise. Bentinck et al. (2015) validated the dataset with two algorithms formulated from SVM: one-vs.-rest (OVR) and random k-labelsets (REKEL). After evaluation, Bentinck et al. (2015) found that both REKEL and OVR algorithms were not superior on all the labels, though the overall F-score was found for both algorithms to be above 80%.

Alm et al. (2005) focused on exploring the automatic classification of emotions in children’s fairy tales. Machine learning with a Sparse Network of Winnows (SNoW) learning architecture was used. A Naïve-Bayes classifier was employed in the classification. Unlike our approach in this study where we used Plutchik’s basic emotions, Alm et al. (2005) classified the fairy tales according to Ekman’s basic emotion categories. Alm et al. (2005) collected the fairy tales manually and gave them out for manual

11 http://wndomains.fbk.eu/wnaffect.html
annotation. They also compared the results from the positive, negative and neutral categories. After experimenting with 22 fairy tales, the results with the Naïve-Bayes classifier were found to be satisfactory, as reported by the researchers.

2.3.5 Related works in sentiment analysis

Sentiment analysis (SA), sometimes referred to as opinion mining, has been considerably researched in recent times, given its relevance to the understanding of the perceptions or opinions of people. A sentiment is an individual’s attitude, thought or judgement that gives rise to feelings (Fang & Zhan, 2015). Sentiments in text can be analysed on three different levels: document, sentence and entity. The document level involves the analysis of a whole document’s sentimental lexicons, such as negative and positive sentiments. This approach can be likened to how EmoTect tracks sentiments in text. Unlike the document level, the sentence level deals with each sentence within a document whether there is/are sentiments. The entity targets the exact word that is used to express a particular sentiment, such as like or dislike.

Textual information comes in two different forms: factual and opinionated (Liu, 2012). On the one hand, factual information in text content expresses objective information about events, situations or entities. On the other hand, opinions are subjective expressions about events, situations or entities (Liu, 2012). Clearly, there are more studies in business-oriented SA than education. Nevertheless, quite a few researchers have taken interest in educational data analytics. Sources of text researchers have used predominantly for SA are novels (John et al., 2006; Mohammad & Yang, 2011), customer reviews and newspaper headlines (Bellegarda, 2010), blog posts (Neviarouskaya et al., 2011; Genereux & Evans, 2006; Mihalcea & Liu, 2006), emails (Liu et al., 2003; Mohammad & Yang, 2011) and more recently, tweets (Mohammad, 2012). In a situation where large datasets are involved, it is economical to go by the semi-supervised method, given that only a small part of data is required to annotate, when training a classifier. In this case, the annotated data are combined with some unannotated data for training a classifier. Several linguistics and classification models are proposed to augment classification performance (see, for example: Wilson et al., 2005; Pang & Lee, 2008).

Emotion and sentiment analysis are often used interchangeably. Following that, Munezero et al. (2014) provide a succinct review of the two, where they considered sentiments as ‘partly social constructs of emotions that develop over time and are enduring’. Work in SA is typically focused on recognising valence. Usually, negative and positive valence is used, though some researchers consider neutral or no sentiment as well (see the works of: Boiy et al., 2007; Kiritchenko et al., 2004). SA has a wide range of applications, and several systems have been developed for sentiment classification. For instance, business organisations use the SA technique to determine
the opinions and the intent of customers regarding products and services. SA approaches are used in education to mine the opinions and perception of students towards teaching and learning.

Hancock et al. (2007) undertook a study on the classification of sentiments in text where only negative and positive emotional valence was used. The researchers employed content analysis inquiry and word count (LIWC) in the sentiments classification. In the end, Hancock et al. (2007) revealed that most positive sentiments are expressed with exclamation marks, while affective words are used, in most cases, to express negative emotions.

A closely related work to our sentiment classification component is the study by Pang et al. (2002). Pang et al. (2002) classified text documents according to positive and negative sentiments. With this, Pang et al. (2002) performed the classification on the content of documents rather than a more traditional approach of using document topics. Using movie reviews as their dataset, Pang et al. (2002) aimed to determine whether a review was positive or negative. The content of the movie review was manually annotated by humans, a method similar to this dissertation’s. The researchers compared the human annotation of the movie review to that of machine learning classifiers. Three machine learning classifiers were used: Naive-Bayes, SVM and maximum entropy. After the experiment, all the three classifiers did not perform well on sentiment classification, unlike traditional topic-based categorisation. Based on the results, they investigated further to find the factors that led to the underperformance of the machine learning approaches. Key among the factors, according to Pang et al. (2002), is that the position of a word in text documents can influence the outcome. Pang et al. (2002: p. 6) wrote, ‘movie reviews, in particular, might begin with an overall sentiment statement, proceed with a plot discussion, and conclude by summarising the author’s views’. Another factor that Pang et al. (2002) pointed out is the challenges associated with the identification of a negation sentence in the text document.

An important application of SA in education is a study conducted by Moretti et al. (2015). Moretti et al. (2015) aimed to understand how the media discuss and portray the policies of education in the United States. The researchers crawled for online data from The New York Times and Time magazine. Retrieval of the data was done with some twitter API with the keyword searching phrase teacher evaluation. Moretti et al. (2015) used a Natural Language Tool kit12 to harvest the text content. Supervised machine learning with a Naive-Bayes classifier was used to extract positive and negative sentiments in the data. Topic modelling was performed as well. Before then, the researchers trained their classifier with twitter sentiments as well as movie reviews. Moretti et al. (2015) concluded that SA and topic modelling are promising methods for analysing Internet data in ways that can inform policy decision making.

---

12 http://www.nltk.org/
2.3.6 Evaluation of text classifier

The efficacy of a classification algorithm needs to be ascertained to inform the validity and variability of experimental results when using machine learning techniques. Predominantly, precision, recall and an F-measure are used as measures for evaluating text classification algorithms. An annotation agreement is another tool for evaluating the efficacy of a classification algorithm. Although, there are other text classification matrices, we discuss the aforementioned evaluation measures for text classification in the subsections below.

Annotation agreement

The consistency and accuracy in text output from a supervised machine learning classifier is hugely dependent on the quality of the training data (Volkova et al., 2010; Canales & Martinez-Barco, 2014). In supervised machine learning, it is required that training data need to be annotated for a high level of accuracy (kappa >= 70%) before using them to train a classifier. Bhowmwick et al. (2008) indicated the ‘reliability of annotation is a key requirement for the usability of annotated corpus’. Poor annotation or inconsistencies in annotated corpora may lead to the poor performance of a supervised learning algorithm. Therefore, annotating emotions in text requires that proper training be given to selected annotators. Although the text-based annotation of emotions can be approached in diverse ways, humans are often used to carry out the annotation. The ultimate aim of annotation is to obtain a gold standard suitable for creating a reliable classifier model and determining the performance of a classifier. Three main approaches of annotation are used: manual, automatic and semi-automatic. In [PIV], this researcher described the various approaches of text annotation.

Regarding manual annotation, people with a good knowledge of specific text documents (such as students’ life stories, movie reviews etc.) should be selected to undertake annotation. For instance, in building this researcher’s life story corpus for training the EmoTect classifier meant for counselling delivery, three selected school counsellors were tasked to carry out the annotation (Section 4.1.1). Given the subjective and subtle nature of emotions, humans are not always perfect when analysing emotions in text (Kolog et al., 2016). This is because the cognitive and behavioural state of humans, as reported in Section 2.2, could influence their decision while annotating emotions in text. Therefore, it is highly recommended that two or more people be made to carry out annotations. Conversely, an annotated corpus from a single annotator is prone to errors which are likely to lead to the poor performance of a supervised learning classifier. Aman and Szpakowicz (2007) undertook a study on emotion annotation where they used blog posts they collected from the web as the main source of data. Aman and Szpakowicz (2007) used Ekman’s six basic emotions as their emotion categories, of which the blog post was annotated accordingly.
The reliability of annotated corpora from annotators is ascertained through a measure of agreement coefficient (Artstein & Artesio, 2008). Kappa statistics for computing inter-annotation agreement were proposed in 1960 (Cohen, 1960). Cohen’s kappa is two inter-annotators’ agreement measured from a chance-corrected agreement using nominal scale (Conger, 1980). Since its inception, it has widely been used in various academic disciplines due to its simplicity and robustness. For instance, fields like health informatics, electronics and geographical informatics have explored kappa statistics. Given that Cohen’s kappa is limited to only two annotators, researchers explored further and extended kappa computation to more than two annotators. Most notable are Light (1971), Fleiss (1975) and Krippendorff (1967). For the sake of reliability and validation, researchers often employ two or more of the kappa measures in their studies. In [PIII], we employed both Fleiss and Krippendorff’s measure in computing for inter-annotation agreement, though only Fleiss’ kappa was used in the analysis (refer to PIV for more details).

Cohen’s measure of two-rater agreement has been extended to compute multiple rater agreement. For instance, Fleiss (1975) proposed two independent formulations of the multiple rater agreement measure. Since there were three annotators for this study, we decided to use Fleiss’ measure for the computation of the kappa coefficient. Fleiss provided a multiple rater agreement measure based on the agreements among raters for classifying each person or object (Conger, 1980). Fleiss’ measure exclusively extended Cohen’s measure of agreement. However, Fleiss’ kappa allows a fixed number of raters, but different items can be rated by various raters (Fleiss, 1971). Fleiss’ kappa can only be used for nominal or binary scale ratings (Conger, 1980; Fleiss, 1971). The general definition of Fleiss’ kappa is shown in eq. 2.1, where $P_o$ and $P_e$ are observed and chance agreements respectively. As shown in eq. 2.1, the $1 - P_e$ is the level of agreement that occurs by chance, while the degree of agreement attainable above chance is $P_o - P_e$. If the annotators agree completely, then kappa is 1. If there is no agreement among annotators, then kappa is less than or equal to 0. With this, Landis and Kouch (1977) categorise the kappa according to a value < 0 as having no agreement, 0–0.20 has weak agreement, 0.21–0.40 has fair agreement, 0.41–0.60 has moderate agreement, 0.61–0.81 has a substantial agreement, and 0.81–1 has almost perfect agreement.

$$k_f = \frac{P_o - P_e}{1 - P_e}$$

Let the number of subjects (instances) be $n$, the number of evaluation categories be $k$ (thus, the emotion and sentiment class) and $m$ be the number of annotators for the subjects (thus, counsellors). The observed agreement ($P_o$) is computed by eq. 2.2 (Conger, 1980), and for each subject $i = 1, 2, 3, ..., n$ and the evaluation categories $j = 1, 2, 3, ..., k$. Let $h_{ij}$ be the number of annotators who assign category $j$ to subject $i$. In the same vein is the chance agreement ($P_e$), which is computed by eq. 2.3, where $P_j$ is the average annotator for category $j$. For example, in this study, 3 annotators were
tasked to assign 8 emotion categories to 360 instances of students’ life stories. Hence, \( n = 360 \), \( m = 3 \), and \( k = 8 \). In this study, each of the annotators performed equal tasks, but there were instances in which some of the annotators did not assign any category because they did not find any emotionally charged phrases or words. This makes Fleiss’ kappa robust because of its flexibilities in allowing such cases.

\[
P_o = \frac{1}{mn(m-1)} \left[ \sum_{i=1}^{n} \sum_{j=1}^{k} h_{ij}^2 - mn \right] \tag{2.2}
\]

\[
P_e = \sum_{j=1}^{k} P_j^2 \tag{2.3}
\]

Crowston et al. (2010) employed rule-based and machine learning techniques to code unlabelled data, a copy of which had already been coded qualitatively by humans. The rationale of their study was to determine which coding technique was more efficient and effective. They used data from emails collected from computer-supported groups and discussion forums. The codebook (themes) extracted from the data was used to categorise the test data. Two persons were made to manually code the data, of which a good inter-annotator agreement score of 80% was found. However, 75% of the data was used to train a machine learning linear classifier, while the remaining 25% was for testing. Crowston et al. (2010) developed rules using an NLP program. By comparing the machine learning output with the human-coded, they concluded that machine learning is an effective way to assist researchers in coding qualitative data, compared to the rule-based approach.

Aman and Szpakowicz (2007) undertook a study on emotion annotation in text in which they used blog posts as the main source of data. Aman and Szpakowicz (2007) employed Ekman’s six basic emotions (Ekman, 1992) with an inclusion of mixed emotions and no emotions categories, resulting in eight categories of emotions. Aman and Szpakowicz (2007) employed four judges to annotate the emotions in the blog posts based on the aforementioned emotion categories and found the average inter-annotators’ kappa values of the emotion categories ranging from 0.60 to 0.79. To this end, Aman and Szpakowicz (2007) reported an overall classification kappa of 73.89%, a value above the acceptable baseline of 70%. The overall kappa value obtained as a result of the human annotation was then used as a baseline to test the efficacy of their classification algorithms.

**Precision, recall and F-measure**

Precision, recall and F-measure are important evaluation measures for text classifiers. Some researchers only rely on the score from the F-measure to draw their conclusion (see, for example: Bentinck et al., 2015; Lu, et al., 2010), while others rely on precision, recall and/or F-measure for their conclusion (see, for example: Crowston, 2010;
Munezero et al., 2013). In this work, this researcher used the scores from precision, recall and the F-measure to draw conclusions regarding the efficacy of the EmoTect classifier. According to these evaluation approaches, gold standard test data are required. An unseen test datum is made to be classified by a classifier and the results compared with gold standard data (classified data). Therefore, the proportion of the labelled instances of the gold standard that are identified and extracted by a classifier is referred to as the recall. The fraction of the automatically extracted data that was labelled correctly as the gold standard by the classifier is precision. The F-measure, also called the F-score, is the harmonic mean of recall and precision. Formulas 2.4 and 2.5 represent recall and precision while Formula 2.6 is the F-measure.

\[
\text{Recall} = \frac{\text{number of relevant items extracted}}{\text{number of relevant items in collection}}
\]  

\[
\text{Precision} = \frac{\text{number of relevant items extracted}}{\text{total number of items extracted}}
\]

\[
F - \text{measure} = 2 \times \frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}}
\]  

Computing for precision, recall and F-measure for a binary classification is relatively straightforward compared to a multi-class binary classification. A multi-class classification problem can be quite confusing, given that several classes are used. In multi-class classification, the first thing to do is to generate a confusion matrix. Many existing machine learning packages already generate confusion matrices, but without that luxury, it is actually very easy to implement by keeping counters for the true positives (TP), false positives (FP) and total number of instances for each label.

2.3.7 The role of NLP in education

Before 1950, natural languages in educational settings were processed manually by relying solely on human intuition (Arnold, 1950). Although the practice was not efficient, it was the only option, since few studies had been conducted on assisting natural language processing with computational techniques. As technology advanced, however, after 1950, attention was directed towards improving the educational sector by developing NLP techniques to help with automatic natural language processing. Quite recently, there has been much research into developing educational NLP (e-NLP) tools to help the sector. The purpose is to create efficiency in understanding the natural language of learners (Alhawiti, 2014), thereby improving teaching and learning. NLP enhances academic performance and provides flexibilities in teaching and learning. Therefore, the role of e-NLP cannot be ignored. The role of e-NLP is discussed in this section, with some existing and emerging technologies in e-NLP.
NLP has multiple roles in education. This is because it has several applications in different sectors of education, such as counselling, teaching and learning, and school administrative activities. While some e-NLP researchers graciously focused on developing tools to help in the sector, others leveraged NLP resources to study and understand the behaviour of students and teachers. In a later approach, data from educational settings are required to extract knowledge. For example, Munezero et al. (2013) collected students' learning diaries and used an emotion detection system to track their emotions. The goal of the study was to understand and make recommendations regarding the behaviour of students for teachers. In this work, we also collected stories of students in texts with the aim of using the content as training and test data sets.

One of the areas in education that has attracted much attention from e-NLP researchers is e-learning. Since e-learning allows students to learn remotely, the application of e-NLP has become widespread and utilised for numerous occasions. Usually, some e-learning platforms incorporate NLP applications that help students to check and correct their grammar. Responses from students after classroom sessions are analysed in real time for decision making. Outputs from analysing student responses can be used by teachers to modify their teaching methods. The automatic classification of e-learning materials into desired categories could be integrated into e-learning platforms. The scoring of students' texts as well as developing text-based dialogue tutoring systems are NLP applications that have existed since 1960 (Littman, 2016). The automated and semi-automated generation of test items is another area of NLP in education. These tools help assessment developers identify appropriate sources of material for question item creation.

Given the unfavourable teacher-student ratio in many countries, dealing with students on MOOCs (massive open online courses) is a difficult task for teachers or instructors. Imagine a virtual or e-learning course running for over 100 students under the tutorship of one person. It is extremely difficult and challenging to identify and read all posts. An NLP application has the capability to alert a teacher if a post has not been read and graded. A related software is LanguageMuse℠, which was developed as a web-based system for instructional authoring. LanguageMuse℠ offers linguistic feedback that highlights vocabulary, sentence structures and discourse relations found in classroom texts. The linguistic feedback supports teachers in creating linguistically informed lesson plans, texts, activities and assessments with appropriate scaffolding techniques (Burstein et al., 2012).

There are standard non-NLP proofreading tools that have existed for some time now (Milkowski, 2010; De-Smedt, 2009). Most of them do not critically focus on natural language errors. Some available NLP tools, rather, focus on grammatical error

---

13 https://languagemuse.org/
detection (Leacock et al., 2010). Another application, in this regard, is an NLP semantic analyser, which has been developed to assess the meaning of both essay-length and short-answer student responses (Dzikovska et al. 2013). Summative language assessment tools also exist in education. They help with the evaluation of students’ reading, writing and speaking skills (Shermis & Burstein 2013).

With the help of NLP search platforms, students can search for refined and relevant information online from a simple query. In fact, there are various e-counselling systems that are developed to help students find relevant information online rather than the traditional web search platforms like Google and Yahoo. Often, these systems are developed to crawl the web for relevant and related information based on keyword queries. One example is a web-based e-counselling system by Khandelwal et al. (2013). The system was developed from the NLP technique, which returns relevant information to a query. The researchers justified that the web contains unorganised information that returns a large amount of irrelevant information; hence, their system filters out the irrelevant content and presents the most relevant information to users. Their system captures a ‘query from students in Natural Language, understands the logic of the query and returns a reply just as a human counsellor would, which helps the student to find proper career option with less efforts and time’ (Khandelwal et al., 2013: p. 1). This and many other related systems are available to play a similar role in students’ academic achievements.

In recent times, game developers from the educational technology research domain have made big strides to augment counselling delivery. Games involve one or more players who have goals, constraints, play-offs and consequences. Students who play counselling games get to learn while entertaining themselves. One example of such games is Façade14. Façade is an AI-based interactive story game developed specifically with NLP technologies (Mott, 2010). The game functions by interacting with users through automatically generated dialogues. Incorporating elements of both interactivity and drama, Façade takes advantage of voice acting and a 3-D environment, as well as natural language processing and other advanced artificial intelligence routines, to provide a robust interactive fictional experience. The player can take an active role in the conversation, pushing the topic one way or another, as in an interactive stage play. Stage plays are stored as script text files that can be read after the player has finished.

With games built with NLP capabilities, teachers could teach students about stress and anger management, decision making, communication skills, study or organisational skills, making friends and conflict resolution. Hence, games promote interpersonal and intra-personal interactions of students, thereby helping students to learn good socialisation and communication skills (Mahmoud, 2014). Examples of counselling games are anger and stress control games.

3 RESEARCH DESIGN

The methodological process on which the study was conducted for three years is elaborated upon in this section. This considers the design science research (DSR) framework adopted in the development of the EmoTect system. The context and method of the study is described here as well. This encompasses the building of the text corpus, counsellors' profiles and the requirement elicitation for EmoTect's implementation.

3.1 DESIGN SCIENCE RESEARCH

There are various ways DSR is characterised in terms of the guidelines associated with its application in artefact development. The guidelines are dependent on the discipline and the type of artefact. This is because several academic disciplines have gradually accepted design science as a research programme (Weber, 2010). Example of such disciplines include engineering (Archer, 1984; Eekels & Roozenburg, 1991; Fulcher & Hills, 1996), computer science (Preston & Mehandjiev, 2004; Takeda et al., 1990) and information systems (Hevner et al., 2004). As noted by Simon (1996: p. 9), design science ‘supports a pragmatic research paradigm that calls for the creation of innovative artefacts to solve real-world problems’. With this, DSR is thought to focus on artefact development with particular interest in the relevance of its applicability in real-life settings (McKay & Marshall, 2005). Thus, much value is placed on the impacts that artefacts developed from DSR have on individuals, organisations and teams.

In DSR, March and Smith (1995) acknowledged that two distinct knowledge programmes are required by researchers: natural and behavioural sciences. On the one hand, through Hevner et al. (2004: p. 75), March and Smith (1995) explained that ‘the design science paradigm seeks to create innovations that define the ideas, practices, technical capabilities, and products through which the analysis, design, implementation, and use of information systems can be effectively and efficiently accomplished’. On the other hand, ‘the behavioural science paradigm has its roots in natural science research methods. It seeks to develop and justify theories (i.e., principles and laws) that explain or predict organisational and human phenomena surrounding the analysis, design, implementation, and use of information systems.

DSR has widely been used in information systems and computer science artefact development. The fundamental base of DSR projects usually starts by understanding the context and identifying the problem on which the intended artefact is meant to solve (Au, 2001; Mramba et al., 2016). March and Smith (1995) consider DSR a pragmatic approach of building an artefact by involving end users in the developmental process. March and Smith (1995) further concluded that the relevant products of DSR
are either constructs, models, methods, instantiations or a combination thereof. With this, the end product adopting DSR, in this work, is the creation of an artefact (EmoTect).

Hevner (2007) pointed out that ‘the utility of the information system and characteristics of the organisation, its work systems, its people, and its development and implementation methodologies together determine the extent to which that purpose is achieved’. Therefore, it is necessary for information system researchers to have in-depth knowledge about the users and their requirements (Zmud, 1997; Johannesson & Perjons, 2012). In this dissertation, DSR was used in the development of the EmoTect system. Peffers et al.’s (2006) DSR framework was duly followed. Based on Peffers et al.’s (2006) DSR model (redrawn in Figure 3.1), this researcher has simplified the various steps into three whole components interlinked iteratively and sequentially. Figure 3.1 displays the three components: Analysis and requirement, implementation and evaluation, summarised from Peffer’s framework. The various components are also explained in [PV] which took into account how the requirements were formulated.

Subsequent sections will delve into the various processes of the DSR framework. The research problems and challenges that motivated this work are partly discussed in an earlier section of this dissertation (see Section 1.1). Moreover, some of the problems in the counselling sectors of Ghana, which motivated this study, have been elaborated upon in Section 2 of [PI]. These problems motivated the formulation of the research objectives of this work. A summary of the preliminary requirement of EmoTect is outlined in Section 3.2.2. The implementation and evaluation phases of
EmoTect are elaborated on in subsequent chapters of this work, though some parts are captured in [PV].

### 3.2 CONTEXT AND METHOD

Four different senior high schools (SHSs) from Ghana participated in the study. Three of the schools were intermittently involved in the development of the EmoTect system. The schools are Mfatsipim SHS, Osei Kyeretwie SHS, Bompata SHS and Agona Seventh Day Adventist SHS. The participants from the schools were involved in the design of the EmoTech system as part of participatory design (PD) approach taken. PD is a design approach that attempts to actively involve stakeholders (thus, students and counsellors) in the design process, thereby ensuring that the results based on the design meets stakeholder’s expectation. According to Robertson and Simonson (2012: p.6), “during a Participatory Design process all participants increase their knowledge and understandings: about technology for the users, about users and their practice for designers, and all participants learn more about technology design”. Duveskog et al., (2009) undertook a similar project in Tanzania regarding a digital tool for counselling people living with HIV/AIDS. Duveskog et al. (2009) adopted a PD consisting of teams of secondary school children, university counsellors, HIV counselling experts and experts in ICT were involved in the implementation process.

Counsellors and randomly selected students from the various schools participated in the project intermittently. Apart from the main counsellors, auxiliary counsellors from the various schools were also involved at some stages. A number of selected teachers participated in the project as well, especially during the contextual evaluation of the project. In all the varying stages of the project, where necessary, the participants were contacted in their various schools for their input. Questionnaires, interviews and observations were the main research techniques used in the collection of data throughout the studies. Preliminary data collections from the participants were mostly published in [PI] to [PIV]. The outcome influenced the development of EmoTect which ultimately led to [PV].

As indicated in Table 3.1, **RQ1** *(What are the emotional challenges that threatens students’ academic pursuit?)*, **RQ2** *(What counselling technologies are being used in the senior high schools of Ghana?)*, **RQ3** *(What are the factors that motivate students to adopt and use e-counselling in Ghana?)* and **RQ4** *(Does the emotional state of counsellors influence their emotion perception while annotating emotions in text?)* form part of the questions that led to understanding of the research environment. **RQ1** and **RQ2** were qualitatively investigated and the technique for data collection was through the use of questionnaires and interviews. In **RQ3**, we adopted a quantitative research method where statistical research approach was used. The study was investigated in [PII] where
only questionnaires were used to collect data from the selected students. Mixed research method was used in [PIV] where RQ4 (Does the emotional state of counsellors influence their emotion perception while annotating emotions in text?) was investigated. With this, only questionnaires were used in data collection. RQ5 (How can a text-based emotion and sentiment classification system be constructed for counselling delivery?) and RQ6 (How can a text-based emotion and sentiment classification system for counselling be evaluated?) were both investigated in [PV]. While RQ5 was in line with constructivism approach of requirement elicitation and implementation of the requirements in EmoTect system, RQ6 was the evaluation of the EmoTect system. The technique for gathering data in the RQ5 and RQ6, as seen on the table, are through interview, questionnaires and observation.

<table>
<thead>
<tr>
<th>Research Method</th>
<th>DSR process</th>
<th>Research technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1 Qualitative</td>
<td>Questionnaire</td>
<td>Questionnaire and interview</td>
</tr>
<tr>
<td>RQ2 Qualitative</td>
<td>Questionnaire</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>RQ3 Quantitative</td>
<td>Questionnaire</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>RQ4 Mixed research</td>
<td>Implementation</td>
<td>Interview, observation</td>
</tr>
<tr>
<td>RQ5 Constructivism</td>
<td>Evaluation</td>
<td>Observation, questionnaire</td>
</tr>
<tr>
<td>RQ6 Mixed research</td>
<td>Evaluation</td>
<td>Observation, questionnaire</td>
</tr>
</tbody>
</table>

### Table 3.1. Research methods and techniques for each research question

#### 3.2.1 Counsellors’ profiles

Three of the counsellors from the three active schools who participated in the study have ages ranging from 30 to 50 years. Two of them have bachelor’s degrees in educational counselling, while the other one has a bachelor’s degree in history. The one with the history background was once a teacher and subsequently was converted into a school counsellor because the sector lacks professional counsellors to help students. All the main counsellors reside on their respective school campuses where they counsel students and organise other symposia for students and teachers. The counsellors have worked in their respective schools for over 4 years as counsellors. Two of the counsellors are male, while the other one is female.

#### 3.2.2 Requirement elicitation

Coupled with our preliminary findings, we organised a semi-structured interactive discussion session with the selected counsellors and students before the commencement of EmoTect’s development, forming part of the DSR requirement. One counsellor and students, each from the three selected schools in Ghana, participated in the discussion. The session was organised in the respective schools of the participants. Altogether, three counsellors and thirty students contributed to the discussion. The essence of the meeting was to outline to the participants the preliminary require-
ments for the commencement of EmoTect’s development based on previous encounters. By so doing, this researcher expected the participants to contribute to the discussion before taking the next step to implement the initial (pre-defined) requirements for EmoTect. The initial preliminary requirements and brief description of the various components of EmoTect, as discussed with the participants, are shown in Table 3.2.

Table 3.2. Requirements from researcher’s preliminary studies

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact counsellor’s widget form</td>
<td>Presentation layer where students can contact counsellors on the grounds of seeking counselling (anonymous).</td>
</tr>
<tr>
<td>Emotion extraction and visualisation</td>
<td>The domain logic that extracts and visualises emotions of students’ submissions automatically.</td>
</tr>
<tr>
<td>Sentiment extraction and visualisation</td>
<td>Where emotional valence (sentiments) from students’ submissions are to be extracted. Only positive and negative valence is considered.</td>
</tr>
</tbody>
</table>

While the aforementioned proposed requirements were welcomed, the participants suggested further ideas that were worth considering in the development of EmoTect. Counsellors, for their part, requested that the intended system be able to assist them in monitoring the emotional changes of their students over time. This generated the idea of creating a database to store the extracted emotions’ categories (results) for future reference. Therefore, in designing the database, this author considered two different functionalities of EmoTect.

On the one hand, the database is meant to house the annotated corpus (life stories of students) meant to be used for classifier training. On the other hand, the database is to store the extracted emotions for future reference, which is made available in visual form at the EmoTect interface, as proposed by the participants. Additionally, counsellors were concerned about emotion keywords from students’ textual submissions, which according to the participants, would prompt further review of students’ submissions should the need arise. This resulted in the idea of extracting and outputting emotional keywords for counsellors at the EmoTect interface for users. The reason connected to the output of the emotional keywords is that counsellors will be able to consider students’ textual submission holistically, should there be any suspicious keywords. For instance, emotion keywords like kill, suicide, worry and die found in students’ submissions will give a clue to a counsellor about what a student may be up to. With the aforementioned emotion keywords, one may interpret that the student is either threatening suicide or he or she is only talking about suicide and death. Students expressed their concerns about the poor internet connection in the country, which partly forms the challenges associated with the implementation of e-
counselling in Ghana. This concern was raised given that EmoTect is a web-based platform.

Though the counsellors are not technically inclined, this researcher still explained to them the basic idea of training a classifier. This researcher has explained the concept as part of the requirement that users will be made to tag the training data based on their understanding of each instance of the training corpus. As expected, there were no opposing questions or ideas to that effect. The final requirements for the implementation were itemised for the development to commerce. In Chapter 4, we delve into the implementation of EmoTect based on the elicited requirements from the counsellors.

3.3 ETHICAL CONSIDERATIONS

Before undertaking this study, the selected schools were presented with an official letter asking for their co-operation to conduct this study. The schools were made aware that their participation in the study was voluntary. As a matter of ethical consideration, the researcher enquired from the schools if the essence of our research would not contradict their school ethics. In response, appendix 2 contains three letters inviting the researcher to conduct the study in the respective selected schools.

Students who participated in the study signed an informed consent form (Appendix 1) after they learned details about the essence of the study. It was, therefore, agreed that any details that might identify the students should not be included in their written response to the questionnaires. Students who were uncomfortable with sharing their life stories did not actually respond to the questionnaires. Participation in the study was voluntary and the students were made aware that they could stop their participation at anytime during the study. The questionnaires were also filled on voluntary basis and students that did not wish to share their personal stories were allowed to do so.

The life story corpus, which contains students' stories that border on their academic development, shall be made available for research purposes. This shall be done without linking the stories to the individual students. The text displayed in figures 4.7 and 4.8 are not real stories from the students; they are fictional texts to demonstrate how EmoTect works.
4 IMPLEMENTATION

As defined in Section 1.1, human language technology (HLT) was employed to contextually develop a web-based e-counselling system for aiding with counselling delivery. The system, called EmoTect, is multi-functional, comprising two components: the contact counsellor and emotion detection. Implementing an artefact is part of the core essence of DSR (March & Smith, 1995; Peffers et al., 2006). Therefore, this chapter presents the implementation process of EmoTect in line with the DSR framework adopted from Peffers et al. (2006). This work considers the development and architecture of both emotion and sentiment classification. The building of the life story corpus, which forms an integral part of EmoTect’s development, is outlined in this section as well.

Figure 4.1 shows the context view of EmoTect in terms of its use. The figure shows how students’ textual submissions to counsellors are analysed by EmoTect and the results presented to counsellors. The output from EmoTect is saved in a database for future reference. Ghana has over 50 languages, yet the national medium of instruction is British English because the country was colonised by the British. Since then, the educational system has been modelled on the British educational system. As EmoTect is contextualised in the Ghanaian context, it only processes the English language.

![Figure 4.1. Context view / process of EmoTect](image-url)
4.1 EMOTECT DEVELOPMENT

Peffers et al. (2006) suggest that critical attention should be given to the desired functionality and architecture of an intended artefact while eliciting or gathering requirements for its development. For such an approach, March and Smith (1995) agreed that a DSR artefact should satisfy the needs of end users. In this light, the implementation of EmoTect began after gathering the needed requirements outlined in Section 3.2.2 and [PV], which is based on [PI] to [PIV]. In addition to the implementation phase, Peffers et al. (2006: p. 13) suggest that the ‘resources required from moving objectives to design and development include knowledge of theory that can be brought to bear in a solution’.

As pointed out in the requirement identification phase, this researcher deduced that EmoTect’s process development needed three tiers: the interface (presentation layer), domain logic and database design. The presentation tier is the point of interaction between the users and the system. The presentation tier was developed from HTML5 and Java server pages, which allow user interaction with the system. The tier is comprised of user controls and webforms.

The second tier is the domain logic, which is responsible for the pre-processing and extraction of emotions and sentiments from input text. By developing the domain logic, the following main packages were used: NLTK\(^{15}\) and Weka’s multi-class SVM classifier\(^{16}\) (i.e. sequential minimum optimisation-SMO). The POS tagger from the NLTK was used for the syntactical parsing of the input text. In addition, the lemmatisation package from the NLTK was used in lemmatising the input text. The process of the text classification is elaborated in Section 4.13.

![Figure 4.2. Process diagram of EmoTect development](image)

The third and last tier is the database. This researcher used MySQL as the database server and Apache Tomcat for the webserver. After pre-processing and extracting emotion/sentiments, the output is provided in JSON (JavaScript Object Notation) format which is then sent to the front end for further processing, and it is displayed in HTML5 tables and visualisation charts. Figure 4.2 shows the various tiers with the

\(^{15}\) [http://www.nltk.org/](http://www.nltk.org/)

\(^{16}\) [http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SMO.html](http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SMO.html)
tools used in their respective implementation. Corpus building, training of the classifier and the classification phases are elaborated in the subsections below.

4.1.1 Corpus building and annotation

This researcher developed his own corpus, called a life story corpus, for training and testing the EmoTect system. As explained in [PV] and in line with [P1], emotional antecedents in the form of stories regarding students’ academic development were collected from a selection of students in their various schools (the selected schools stated earlier in Section 3.2). A questionnaire, a sample of which is shown in Appendix 1, was used in collecting the stories. Before then, the students were orientated to write stories that aligned with their emotional antecedents. As pointed out by Lugmayr et al. (2016), students can express their feelings and emotions when given the opportunity to write about their life stories in text, which was part of the research objectives in [P1]. In fact, the question that asked the students to write about their stories was clearly associated with the students’ academic development. The stories were pre-processed into a more readable format that was easy to annotate (see PV). The processed instances of the stories contained at least one sentence and at most five sentences. In the end, we managed a total of 2,200 instances of the stories in all as our life story corpus.

The stories were then given out to the three counsellors (hereafter: C1, C2 and C3) to annotate with eight basic emotions adopted from Plutchik. In [PIII], this author conducted a focused group discussion with selected counsellors to extract from them the emotions they often identify from students during regular counselling sessions. Based on the outcome, this researcher decided to use Plutchik’s basic emotions as baseline emotions. With regards to the sentiment analysis, negative and positive emotional valence were used, as counsellors were tasked to annotate sentiments in the stories alongside the emotional categories.

Like [PIII], this author employed two annotation strategies: inter17- and intra18-annotation agreement. Before the annotation exercise, counsellors were trained in how to do the annotation exercise. Figure 4.3 shows the flow process indicating the roles of the researcher and the counsellors during the annotation exercise. The counsellors who conducted the annotation exercise did so in two annotation rounds (two times). This researcher allowed a two-month interval between the first and second rounds of annotation. Therefore, the task took a long time to complete since the data were numerous for manual annotation. Unlike [PIII], a meeting was arranged afterwards

17 Inter-annotation agreement is the agreement of emotions annotated by annotators. An example is the emotion agreement between C1, C2 and C3.
18 Intra-annotation agreement is the agreement of emotions of each counsellor between his or her annotation from different exercises. An example is the annotation agreement for the first and second rounds of emotion annotation by C1.
with the annotators to discuss the most notable disagreements in the annotated dataset. In the end, most disagreements were agreed upon. This action was taken because a good kappa score was needed to train the classifier. In the end, this researcher computed both the intra- and inter-counsellors’ annotation agreement of emotions/sentiments from the annotated corpus using Fleiss’ kappa. The computation of kappa scores was assisted by the annotation agreement software developed by Geertzen (2012)\(^\text{19}\).

![Figure 4.3. Experimental set-up diagram for the annotation phase](https://nlp-ml.io/jg/software/ira/#demo)

While obtaining a weighted average of the *inter-annotation* agreement kappa score of 70.3% and 80.5% for the emotion and sentiment respectively, the *intra-annotation* agreements for all the counsellors yielded almost perfect average kappas greater than 85% in both the emotions and sentiments. With this, Landis and Kouch (1977) categorise a kappa < 0 as having *no agreement*, 0–0.20 has *weak agreement*, 0.21–0.40 has *fair agreement*, 0.41–0.60 has *moderate agreement*, 0.61–0.81 has *substantial agreement*, and 0.81–1 has *almost perfect agreement*. Table 4.1 represents the various intra-annotation agreement kappa scores for each of the counsellors.

\(^{19}\)https://nlp-ml.io/jg/software/ira/#demo
Table 4.1. Intra-annotation agreement kappa for each counsellor

<table>
<thead>
<tr>
<th></th>
<th>C1 (R1 ∩ R2)</th>
<th>C2 (R1 ∩ R3)</th>
<th>C3 (R1 ∩ R3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emotion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleiss kappa</td>
<td>87.6%</td>
<td>85.5%</td>
<td>87.3%</td>
</tr>
<tr>
<td><strong>Sentiment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleiss kappa</td>
<td>97.3%</td>
<td>96.1%</td>
<td>98.4%</td>
</tr>
</tbody>
</table>

4.1.2 Classifier training phase

The classifier was trained by considering the contextualisation strategies adopted throughout this study. With this in mind, the classifier training did not only follow the traditional approach of using all-in-one inter-annotators’ agreement gold standard data, but an individual’s perception of emotions/sentiment was also considered. This motive was driven by the fact that the *intra-annotation* agreement of emotions by individual counsellors was found to be almost perfect for all three counsellors. This fact was reported in this author’s empirical studies in [PIII], where he sought the influence of counsellors’ emotions on their emotion perception while analysing emotions of students’ textual submissions. The study, based on how individuals perceive emotions, justified the need for individual counsellors to be given an opportunity to label their own instances of training data when using supervised machine learning approaches.

By default, the EmoTect system was trained on annotated life stories after obtaining a good inter-annotation agreement kappa score indicated in Section 4.1.1. However, in the EmoTect interface, provisions were made for users to make changes to the default emotion categories based on their own perception of emotions in text (see Figure 4.4). With this, counsellors are expected to annotate the stories based on their own perception of emotions before using the EmoTect system. Otherwise, the default settings trained on the all-in-one inter-annotated training data are maintained. For instance, different counsellors may tag divergent emotion categories to the same instance of a story. At any time, counsellors can make changes to the emotions/sentiments they have labelled should they find undesirable outputs. Figure 4.4 shows a snapshot of the EmoTect training phase where counsellors annotate stories with emotions and sentiments based on their respective emotion perceptions.
4.1.3 EmoTect’s classification phase

As seen in the EmoTect architecture in Figure 4.5, the system comprises two classification phases: training and prediction. On the one hand, the annotated life stories are first made to train the EmoTect classifier for a model to be created. The implication is that the multi-class SVM (SMO) classifier learns from the training data to predict unlabelled or unseen text. The prediction phase, on the other hand, is where the classifier model extracts and classifies the emotions and sentiments from the input text according to defined emotion and sentiment categories.

From the architecture in Figure 4.5 and the training phase, the system works by first tokenising the training data (life stories) into words. After that, the tokenised words are tagged by their parts of speech, which is accomplished by a POS tagger from the NLTK package. The POS tagging helps to determine the ‘stopping words’; they are removed afterwards. To this end, the emotion features are extracted from the text after the removal of the stopping words. At the training phase, the feature words at this point are lemmatised before feeding them into the classifier. Lemmatisation refers to doing things properly with the use of a vocabulary and morphological analysis of words, normally aiming to remove inflectional endings only and to return the base or dictionary form of a word, which is known as the lemma. At this stage, the feature words are then fed into the classifier (SVM) as a training feature set. After the training, a classifier model is created, which then predicts unseen input text once it is fed into the classifier model. At the prediction phase, just like the training phase, the unseen input text goes through similar pre-processing stages where the unseen text is converted into feature sets. The feature sets are then fed into the classifier.
model, which generates the predicted labels (thus, emotions and sentiment). Emotional feature words are also spotted and output to the system interface. Figure 4.5 shows a pictorial view of EmoTect’s classification architecture.

1. Training phase

2. Prediction phase

Figure 4.5. EmoTect’s classification process

4.2 EMOTECT USER GUIDE

In this section, this researcher demonstrates how EmoTect can be used in counselling work. To launch and use EmoTect, one is expected to have access to the internet, given that EmoTect is a web-based system. In addition, users are told to launch the platform using the latest version of browsers, such as Google Chrome, Mozilla Firefox and Internet Explorer. Sources of data touted to be processed by the EmoTect system are text from the personalised ‘contact counsellor’ form elaborated upon in [PV] and external sources such as emails and social media content, among others.

Once EmoTect is launched, potential users of EmoTect (counsellors) are first expected to register on the project’s webpage. Counsellors, at any time, could access the page provided there are no technical glitches with their registration details. Once logged in, counsellors are expected to create their own database within the system where only their activities will be available to them.

After creating a database in the system, counsellors are expected to generate JavaScript codes and paste them in the header of their HTML webpage. The following script is a sample that generates the ‘contact counsellor’ widget form on a user’s page:

---

20 EmoTect is available here: http://nlp4counselling.com/.
In Figure 4.6, the ‘contact counsellor’ widget form is expected to be used by students to contact their counsellors. The ‘contact counsellor’ widget form is linked directly to the EmoTect system, where data are passed onto the emotion/sentiment extraction phases. Textual submissions through the ‘contact counsellor’ form are analysed automatically by EmoTect, and the results are presented to the counsellors in visual form (see Figure 4.8). To note is that the data captured in figures 4.6 and 4.7 are fictitious and do not represent real data or emails of the students, as the study seeks to protect their confidentiality.

For the sake of anonymity, the ‘name’ field in the ‘contact counsellor’ form is not compulsory, but students are required to enter their emails. Counsellors are expected to reach out to students through their emails in case they need to give feedback to students. Besides the personalised webform, other external sources of students’ textual submissions such as email and social media posts could be copied and analysed with EmoTect. However, it is strongly recommended for counsellors to allow their students to contact them through the personalised ‘contact counsellor’ form. In this case, accumulated automatic text analysis could be visualised in a time period as shown in the visualisation output in Figure 4.12.

![Contact counsellor widget form as it appears on users’ page](image-url)
Figure 4.7 depicts sample textual submissions from the ‘contact counsellor’ widget form. Here, emotions and sentiments are automatically extracted and saved in a database. The output is sent to the visualisation graph indicated in Figure 4.12. Counsellors, on the other hand, could visualise the output by clicking on ‘extract sentiment’ and ‘extract emotions’ as shown in Figure 4.7 and demonstrated in Figure 4.8.

As the researcher has recommended in [PV], ICT coordinators of the various schools are expected to assist with the initial configuration of EmoTect such as generating the JavaScript for embedding in their webpages. The administrator’s role is to update the webpage with new releases and training data. Figure 4.9 represents a UML (unified model language) diagram showing the various actors and their roles in EmoTect.
The extracted emotions from students’ textual submissions are stored for future reference. The intent of implementing this component is to give counsellors, and perhaps school administrators, the opportunity to monitor the emotional changes of their students over a selectable period as presented in Figure 4.12. Counsellors can use the emotional records of students to match with the performance of their students, thereby making decisions regarding any academic changes or flaws. The essence of the emotion keywords is to give counsellors a reason to be critical in their decision-making process regarding students’ emotional development. For instance, keywords like *kill*, *suicide*, *worry* and *die* are likely to trigger a suspicion that makes it worthwhile to take a second look at students’ submissions. Figures 4.10 and 4.11 are the EmoTect interfaces of the emotion and sentiment classification respectively, while Figure 4.12 is a snapshot of the graphical representation of students’ emotional changes over a selectable period.
Figure 4.10. Visualising emotion extracted from input text

Figure 4.11. Visualising sentiment from sample text
Figure 4.12. Visualisation output of emotional changes over a period
5 EVALUATION AND RESULTS

Results from evaluating the EmoTect system are presented in this section. In this light, during EmoTect’s developmental process, two forms of evaluation were considered: formative and summative (Scriven, 1967). On the one hand, a formative evaluation is a ‘method for judging the worth of a program while the program activities are forming or in progress’ (McGriff, 2000). This part of the evaluation focuses on the process of development. For instance, prototyping is used in formative evaluations to test a particular design aspect by using one or more iterations. During the formative evaluation of EmoTect, the prototype version was demonstrated to the counsellors in their respective schools [PV]. The rationale for the formative evaluation was in line with DSR framework to ascertain whether user requirements have been met after the first phase of development. Inputs from the prototype evaluation were considered a deliverable for the final phase of development.

On the other hand, a summative evaluation (sometimes referred to as external) ‘is a method of judging the worth of a program at the end of the program activities (summation)’ (McGriff, 2000). The main purpose of a summative evaluation is to match the outcome to goals. The summative evaluation of this work is the EmoTect classifier evaluation and the evaluation of EmoTect in the end users’ context. The classifier evaluation ascertained its efficacy, and this was done by comparing the performance of the EmoTect classifier with gold standards obtained from the counsellors. This is consistent with Peffers et al. (2006), who point out the DSR artefact must be demonstrated and evaluated with end users in their environment. This section, rather, reports the summative evaluation of the EmoTect system.

5.1 EMOTECT’S CLASSIFICATION

This section presents the evaluation of the EmoTect classifier. The results from the classifier evaluation are presented in this section as well.

5.1.1 Classifier performance evaluation

Three hundred sixty instances of the students’ life stories, representing about 16% of the total data, were used as test data, whereas the remaining 84% was used as the training data. The sampling of the test data was done randomly. Just like the training data, this researcher computed the kappa (k) scores for each of the counsellors (i.e. intra-annotation agreement). In the end, the intra-annotation agreement of emotions in the test data was found for each counsellor to be almost perfect [$C_1 (k = 0.94)$, $C_2 (k = 0.87)$ and $C_3 (k = 0.91)$], which was deemed suitable as gold standards for comparing
with the outputs from EmoTect. In terms of the sentiments, the intra-annotation agreement kappa scores from the annotated stories were equally found to be almost perfect for all the counsellors (all yielding beyond 90%). These represent the kappa scores for the test data. Figure 5.1 represents the entire process on which the EmoTect classification algorithm was evaluated.

Figure 5.1. Summative evaluative process of the EmoTect classifier

Tables 5.1 and 5.2 are the proportion of the counsellors’ annotated instances in the test data and the output from EmoTect for all the sentiment and emotion categories respectively. Since the counsellors performed two rounds of annotation, the tables are a single representation of the agreements of both rounds for each counsellor. The annotated training data obtained from each of the counsellors (see Section 4.1.1) were, at different points, used to train the EmoTect classifier. The test data were fed into the EmoTect system for both emotion and sentiment classification (prediction). From observation, whatever the counsellors’ variations in the training data, the same results were obtained after running the test data repeatedly with EmoTect. This may be because there was no significant difference in the intra-annotation agreement kappa score for all counsellors. In the end, the output from the EmoTect classifier was then compared with the gold standard corpus from each of the counsellors.

Table 5.1. Counsellors sentiment agreement and the EmoTect output

<table>
<thead>
<tr>
<th>Sentiment</th>
<th>EmoTect</th>
<th>C1 Agreement (C1 vs EmoTect)</th>
<th>C2 Agreement (C2 vs EmoTect)</th>
<th>C3 Agreement (C3 vs EmoTect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Positive</td>
<td>130</td>
<td>125</td>
<td>111</td>
<td>96</td>
</tr>
<tr>
<td>2. Negative</td>
<td>201</td>
<td>205</td>
<td>170</td>
<td>181</td>
</tr>
</tbody>
</table>
Table 5.2. Counsellors emotion agreement and the EmoTect output.

<table>
<thead>
<tr>
<th>Emotions</th>
<th>EmoTect</th>
<th>C1</th>
<th>Agreement (C1 vs EmoTect)</th>
<th>C2</th>
<th>Agreement (C2 vs EmoTect)</th>
<th>C3</th>
<th>Agreement (C3 vs EmoTect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Anger</td>
<td>7</td>
<td>14</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>2. Happiness</td>
<td>28</td>
<td>51</td>
<td>26</td>
<td>35</td>
<td>20</td>
<td>62</td>
<td>14</td>
</tr>
<tr>
<td>3. Disgust</td>
<td>26</td>
<td>29</td>
<td>8</td>
<td>16</td>
<td>12</td>
<td>33</td>
<td>18</td>
</tr>
<tr>
<td>4. Sadness</td>
<td>110</td>
<td>116</td>
<td>68</td>
<td>170</td>
<td>76</td>
<td>97</td>
<td>66</td>
</tr>
<tr>
<td>5. Trust</td>
<td>54</td>
<td>65</td>
<td>26</td>
<td>40</td>
<td>22</td>
<td>61</td>
<td>38</td>
</tr>
<tr>
<td>6. Fear</td>
<td>40</td>
<td>38</td>
<td>20</td>
<td>22</td>
<td>16</td>
<td>23</td>
<td>17</td>
</tr>
<tr>
<td>7. Surprise</td>
<td>9</td>
<td>17</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>8. Anticipation</td>
<td>35</td>
<td>14</td>
<td>4</td>
<td>19</td>
<td>15</td>
<td>30</td>
<td>19</td>
</tr>
</tbody>
</table>

Three standard evaluation matrices were employed: recall, precision and F-measure, as discussed in Section 2.3.6. Recall, also known as sensitivity, measures the fraction of labelled instances of the gold standard that were identified and extracted by the system (i.e., the coverage). Precision measures the fraction of the automatically extracted data that was labelled correctly in the gold standard (i.e., the accuracy). With reference to tables 5.1 and 5.2, Figure 5.2 is a graphical representation of the distributions of the gold standard annotations from counsellors and outputs from EmoTect for (a) emotions and (b) sentiments. This indicates the variations in the annotated instances of the test data by the counsellors and the EmoTect system.

Figure 5.2. Distribution of the gold standards and outputs from EmoTect

Table 5.3 contains sample test instances annotated by the counsellors, and classification by EmoTect according to emotion and sentiment categories. Also, the respective emotion keywords extracted from the samples are presented in a tabular form as well. EmoTect is able to detect multiple emotions in a text. Hence, the dominant emotion category is considered as demonstrated in the visualisation chart in Figure 4.10.
Table 5.3. Sample predictions from counsellors and the EmoTect

<table>
<thead>
<tr>
<th>Instance</th>
<th>Emotion</th>
<th>Sentiment</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I am the shy type hence I find it difficult to approach the teachers in my school but life in school is cool.”</td>
<td>C₁: Fear</td>
<td>C: Negative</td>
<td>Shy; difficult; approach; cool</td>
</tr>
<tr>
<td></td>
<td>C₂: Disgust</td>
<td>C: Negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C₃: Disgust</td>
<td>C: Negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EmoTect: Fear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“When I was in class six, my maths teacher conducted a test, a very cheap one. Since it was cheap, he took some strategy to make us score less which was if you solve it well without the units (cm, cm²) he marks you down, so when he brought our result, I had 1 of 10. It was as if I was mourning. I cried and cried until the next morning and even discourage me for the rest of the term.”</td>
<td>C₁: Fear</td>
<td>C: Negative</td>
<td>teacher; conduct; very; cheap; take; make; solve; well; mark; mourn; cry; next; even; discourage;</td>
</tr>
<tr>
<td></td>
<td>C₂: Sadness</td>
<td>C: Negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C₃: Sadness</td>
<td>C: Negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EmoTect: Sadness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“I am 18 years old from Bompata. Due to peer pressure on how to ‘propose’ to girls or chase girls, I do not get much time to learn. My parents do not have time to check my going and coming, so I choose to do whatever I do in the house. I thought staying in the house all alone was the situation destructing my studies the same thing is going on in boarding house. I follow what others do, I do not learn, I don’t often go to class, even if I go I do not go on time. It has really ruined my life, but I guess is too late nothing can be done because I am in my final year. I will be writing my WASSCE in no time.”</td>
<td>C₁: Disgust</td>
<td>C: Negative</td>
<td>old; peer; propose; chase; learn; check; choose; think; stay; alone; destruct; same; follow; learn; often; even; really; ruin; guess; late; final; write; WASSCE.</td>
</tr>
<tr>
<td></td>
<td>C₂: Sadness</td>
<td>C: Negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C₃: Sadness</td>
<td>C: Negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EmoTect: Sadness</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.1.2 Results from the classifier evaluation

*Emotion classifier performance:* The performance of the EmoTect classifier was compared with the gold standards (test data) obtained from each of the three counsellors as reported in Section 5.1. This author reports recall, precision and F-measure per each category of emotions and further reports the overall scores for the emotion categories using the same evaluation matrices mentioned above. The results from the evaluation of the EmoTect classifier are presented in tables 5.4 to 5.6. This is particularly for the emotion detection part of the EmoTect system.

Whereas precision takes into account the proportion of the extracted text data that was correctly labelled, the researcher took a particular interest in the recall, which measured the instances of the baseline (gold standard) that was identified and extracted by EmoTect. In addition, F-measure, which is the harmonic mean of the precision and recall, was computed as well.
Table 5.4. Results of the EmoTect (emotion) classifier with the C₁ gold standard

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Emotions</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life stories</td>
<td>Anger</td>
<td>35.7</td>
<td>71.4</td>
<td>47.6</td>
</tr>
<tr>
<td></td>
<td>Happiness</td>
<td>92.9</td>
<td>51.0</td>
<td>65.8</td>
</tr>
<tr>
<td></td>
<td>Disgust</td>
<td>30.8</td>
<td>27.6</td>
<td>29.0</td>
</tr>
<tr>
<td></td>
<td>Sadness</td>
<td>61.8</td>
<td>58.6</td>
<td>60.2</td>
</tr>
<tr>
<td></td>
<td>Trust</td>
<td>48.1</td>
<td>40.1</td>
<td>43.7</td>
</tr>
<tr>
<td></td>
<td>Fear</td>
<td>50.0</td>
<td>52.6</td>
<td>51.3</td>
</tr>
<tr>
<td></td>
<td>Surprise</td>
<td>41.2</td>
<td>77.8</td>
<td>53.8</td>
</tr>
<tr>
<td></td>
<td>Anticipation</td>
<td>11.5</td>
<td>25.0</td>
<td>15.7</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>53.1</td>
<td>47.4</td>
<td>50.1</td>
</tr>
</tbody>
</table>

As indicated in Table 5.4, happiness (92.9%) and sadness (61.8%) yielded the highest recall with C₁. This implies the proportion of the labelled instances of the gold standard from C₁ that were identified and extracted by the classifier. The corresponding precision scores for happiness and sadness are 51% and 58.6%, respectively. The F-measures for happiness and sadness are 65.8% and 60.2%. This implies that the classifier performed well for identifying the happiness and sadness categories of emotions. However, the lowest recall, as compared with the C₁ gold standard, was found in

Table 5.5. Results from the EmoTect (emotion) classifier with the C₂ gold standard

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Emotions</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life stories</td>
<td>Anger</td>
<td>60.0</td>
<td>42.9</td>
<td>50.0</td>
</tr>
<tr>
<td></td>
<td>Happiness</td>
<td>75.0</td>
<td>46.2</td>
<td>57.1</td>
</tr>
<tr>
<td></td>
<td>Disgust</td>
<td>57.1</td>
<td>71.4</td>
<td>63.5</td>
</tr>
<tr>
<td></td>
<td>Sadness</td>
<td>44.7</td>
<td>69.1</td>
<td>68.6</td>
</tr>
<tr>
<td></td>
<td>Trust</td>
<td>55.0</td>
<td>40.7</td>
<td>46.8</td>
</tr>
<tr>
<td></td>
<td>Fear</td>
<td>72.7</td>
<td>40.0</td>
<td>51.6</td>
</tr>
<tr>
<td></td>
<td>Surprise</td>
<td>60.0</td>
<td>33.3</td>
<td>42.9</td>
</tr>
<tr>
<td></td>
<td>Anticipation</td>
<td>78.9</td>
<td>42.9</td>
<td>55.6</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>53.5</td>
<td>54.0</td>
<td>53.8</td>
</tr>
</tbody>
</table>

Table 5.6. Results from the EmoTect (emotion) classifier with the C₃ gold standard

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Emotions</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life stories</td>
<td>Anger</td>
<td>33.3</td>
<td>57.1</td>
<td>42.1</td>
</tr>
<tr>
<td></td>
<td>Happiness</td>
<td>22.6</td>
<td>50.0</td>
<td>31.1</td>
</tr>
<tr>
<td></td>
<td>Disgust</td>
<td>54.5</td>
<td>69.2</td>
<td>61.0</td>
</tr>
<tr>
<td></td>
<td>Sadness</td>
<td>68.0</td>
<td>60.0</td>
<td>61.0</td>
</tr>
<tr>
<td></td>
<td>Trust</td>
<td>62.3</td>
<td>70.4</td>
<td>66.6</td>
</tr>
<tr>
<td></td>
<td>Fear</td>
<td>73.9</td>
<td>42.5</td>
<td>54.0</td>
</tr>
<tr>
<td></td>
<td>Surprise</td>
<td>50.0</td>
<td>22.2</td>
<td>30.8</td>
</tr>
<tr>
<td></td>
<td>Anticipation</td>
<td>63.3</td>
<td>54.2</td>
<td>58.5</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>55.3</td>
<td>57.6</td>
<td>56.4</td>
</tr>
</tbody>
</table>
anticipation (11.4%) and disgust (30.8%). In the same vein, the recall values for anticipation and disgust were respectively low at 25% and 27.6%. The corresponding F-measure for anticipation (15.7%) and disgust (29%) was found to be poor.

With regards to the comparison with the gold standard data from C2, shown in Table 5.5, the emotion categories with the highest recall are happiness (75%), fear (72.7%) and anticipation (78.9%). Meanwhile, sadness yielded the highest (69.1%) precision at the cost of low recall (44.7%) with the C2 gold standard. Surprise had the highest recall (60%) at the cost of poor precision (33.3%). With the classifier performance with the C2 gold standard, predicting sadness was found the highest for the F-measure (68.6%) score while predicting surprise was the lowest in the F-measure (30.8%).

By examining the performance of the EmoTect classifier with the gold standard obtained from C3, the emotion categories that yield the highest recall are fear (73.9%) and sadness (68%). However, surprise (22.2%) had the lowest precision, while the highest precision was found in trust (70.4%). In the corresponding score of the F-measure in Table 5.6, trust is the highest (66%), implying a somewhat good performance of the classifier.

As already anticipated, the perception of emotions by the counsellors varied (Kolog et al., 2016). Therefore, the overall performance of the system concerning each of the counsellors also varied, but slightly. The variation in the emotional perception by the counsellors could be attributed to the subjective and subtle nature of emotions (Zadra, 2011). In addition, the variation in the emotion perception by the counsellors shows how different people interpret and perceive emotions in text at a particular time. From tables 5.4 to 5.6, it can be deduced that the overall recall, precision and F-measure scores for all the counsellors looks promising for its purpose, though the performance of the EmoTect classifier with some of the emotion categories was found to be poor. All in all, an increasing trend was observed in the overall recall, precision and F-measure. That is, the performance of the EmoTect classifier against the counsellors increases from C1 to C3 for precision, recall and the F-measure, where: recall: C3 (55.3%) > C2 (53.5%) > C1 (53.1%), precision: C3 (57.6%) > C2 (54.0%) > C1 (47.4%) and F-measure: C3 (56.4%) > C2 (53.8%) > C1 (50.1%). With this observation, based on the methodology used in this dissertation, no particular reason could be attributed to the increasing trend from C1 to C3 other than the variations in the scores for each counsellor.

Though these findings came as no surprise, this researcher’s interest in the evaluation was to ascertain the performance of the EmoTect classifier with the counsellors’ gold standards. While the overall recall, precision and F-measure scores were somewhat good for each of the counsellors, this researcher believes the algorithm can still be improved when more of the annotated, emotionally charged students’ stories are used to train the algorithm further. With this, more attention will be given to the emotion categories that yielded the lowest recall and precision, such as happiness.
and surprise. To this end, it can be deduced that the EmoTect algorithm achieved comparable accuracy to the gold standard, even when presented with unknown data.

**Sentiment classifier performance:** The performance of the EmoTect classifier was examined with regards to the detection of sentiments in the test data. Just like the emotion detection part, the annotated sentiments by the counsellors were compared with the EmoTect algorithm. Tables 5.7 to 5.9 depict both the negative and positive sentiments yielding *almost perfect* for recall and precision with the gold standards’ corpora from the three counsellors. The implication is that EmoTect extracted a higher proportion (recall) of sentiments from the gold standards and predicted most of them correctly (precision). The same can be said about the overall recall and precision of the classifier with the gold standards from the counsellors. The F-measure for all the counsellors in terms of the classifier performance in the sentiments performed very well, and this can be said the same about the overall score for the F-measure. This indicates the EmoTect algorithm – the sentiment part – achieved accuracy comparable to the gold standard’s, even when presented with unknown data. That is, EmoTect performed well when presented with unclassified data for prediction. In a nutshell, while researchers believe the emotion detection part did not perform so well, the sentiment analysis part performed with higher accuracy. With this finding, it can be concluded that EmoTect performs better with the sentiment analysis than emotion extraction.

Table 5.7. Results from EmoTect’s (sentiment) classifier with the C₁ gold standard

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Sentiment</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life stories</td>
<td>Positive</td>
<td>82.9</td>
<td>84.6</td>
<td>87.1</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>88.8</td>
<td>85.4</td>
<td>83.7</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>85.0</td>
<td>84.9</td>
<td>85.0</td>
</tr>
</tbody>
</table>

Table 5.8. Results from EmoTect’s (sentiment) classifier with the C₂ gold standard

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Sentiment</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life stories</td>
<td>Positive</td>
<td>90.5</td>
<td>90.0</td>
<td>81.7</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>91.4</td>
<td>73.8</td>
<td>90.3</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>90.8</td>
<td>83.6</td>
<td>87.1</td>
</tr>
</tbody>
</table>

Table 5.9. Results from EmoTect’s (sentiment) classifier with the C₃ gold standard

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Sentiment</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life stories</td>
<td>Positive</td>
<td>92.3</td>
<td>90.0</td>
<td>83.1</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>92.5</td>
<td>75.4</td>
<td>91.2</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>92.4</td>
<td>84.0</td>
<td>88.2</td>
</tr>
</tbody>
</table>
5.2 EMOTECT IN CONTEXTUAL USE

Evaluation and results from the contextual use of EmoTect are presented in this section. The process is part of the summative evaluation of the EmoTect system.

5.2.1 Contextual evaluation of EmoTect

The final version of EmoTect was summatively evaluated in the environment with selected counsellors and teachers. As McGriff (2000) pointed out, a summative evaluation is used to assess whether the results of the artefact being evaluated (thus, EmoTect) met the stated goals (thus, aiding the decision making of students in counselling). By demonstrating EmoTect to the participants, this author could ascertain the users’ perception of the use of EmoTect for the intended purpose. This is consistent with Hevner et al. (2004), who believed that the quality and efficacy of a system can be ascertained by demonstrating it with end users during an evaluation.

From the three actively participating senior high schools, two counsellors each were selected. In addition, for the final evaluation, four teachers were picked who were considered to be actively involved in counselling work. Except for one of the schools where two of the teachers were selected, the remaining two teachers were respectively selected from the other two schools (refer to the schools in Section 3). In the end, 10 participants were involved in the contextual evaluation of EmoTect. In the respective schools of the participants, this researcher demonstrated the EmoTect system to the participants. Figure 5.3 shows the author demonstrating EmoTect to participants.

![Figure 5.3. Demonstrating EmoTect to participants in one of the selected schools](image)

While evaluating the system, the author observed the participants used the platform, and thereafter, they were allowed to continue to use it for a certain period. The focus during the observation was centred on, among other things, the counsellors’ ease of use of the system. A month later, a questionnaire consisting of both open- and closed-ended questions were administered to the participants. The authors used a 5-point Likert scale from 1 to 5 in the questionnaire, where 1–strongly disagree (NO), 2–disagree (NO), 3–neutral (no idea), 4–agree (YES) and 5–strongly agree (YES).
Based on the Likert scale, 1 and 2 represent ‘NO’, 3 represents ‘NO idea’ and 4 and 5 represent ‘YES’. However, some of the questions in the questionnaire also considered a simple ‘YES’ or ‘NO’ in the data analysis. This was the closed-ended part of the questions. Basically, the subject part of the questions was a follow up to the multiple-choice questions. This was aimed at gaining insight into the closed-ended response. The questionnaire is annexed to this dissertation in Appendix 1.

Given the small number of the participants (10), the data analysis was interpreted using percentages instead of descriptive and inferential statistics. The findings arising from analysing the content of the data are presented in Section 5.2.2. Altogether, 10 participants responded to the questionnaires. The essence of this stage of evaluation was to seek the perception of the participants regarding the functionalities, ease of use (usability) and impact of EmoTect in counselling delivery.

5.2.2 Results from the contextual evaluation

The perceived impact of EmoTect in counselling, part of the research objectives, was ascertained during an evaluation session with counsellors in their respective schools. During the demonstration, participants were observed while using the system. Therefore, the results presented in this section are the combined findings from both the questionnaires and researcher’s observations.

Results obtained from the questionnaires are presented in Table 5.10. While Table 5.10 represents responses to the objective questions, subjective follow-up questions were maintained and used as part of the questionnaires. Table 5.10 demonstrates that a majority (90%) of the participants were convinced that EmoTect is an effective tool to facilitate counselling delivery. Upon close examination, however, only one participant (10%) remained unconvinced of the counselling delivery capabilities of EmoTect, pointing out that most of the schools in Ghana do not have an internet connection, and since EmoTect is web-based, it would be difficult for it to serve its purpose. He said:

‘The application works well, but I think this country is not ready for this sort of platform, unless it is meant to be used by selected schools who have the resources for it. This is because, since the internet is a requirement to use the platform, an internet connection is a problem. Many schools do not have the internet in Ghana. Besides, others also find it difficult to foot high-speed internet bills. This is why I think it is not good for counselling work in Ghana.’

Notwithstanding that, those who became convinced that EmoTect was capable of supporting counselling delivery highlighted its efficiency in terms of the high speed with which it processes large volumes of text content. This is because EmoTect is able to aid counsellors to make decisions about their students’ emotional behaviour. It is also able to help counsellors and school administrators determine the mood of their
students at any particular time. Based on its relevance in counselling, the researcher argue that EmoTect is superior in terms of functionality against the existing email and WhatsApp platforms found to be the main tools counsellors use in counselling work. With the use of EmoTect, counsellors do not need to rely on manual processes to analyse submissions of students’ text content.

While demonstrating EmoTect, it was observed that participants were positively surprised about the capabilities of EmoTect, as already discussed in Section 1.2. This is consistent with [PV], where counsellors appeared perplexed about the output of EmoTect’s prototype during the formative evaluation. This was to be expected since there are no such existing systems available to complement counselling delivery contextually in terms of emotion and the personal-social development of students. One of the counsellors wrote:

‘Generally, I think the platform is a good one. It is able to perform the task of detecting emotions and sentiments in the text I tested with. Some of them, based on my perception, are difficult to tell which emotion it belongs to, but the platform made it easy for me. It is amazing to witness how computers can go to that extent of identifying if written or typed text contains emotions and the type of emotions it expresses. My only concern is that many schools may not be able to use it because of the poor state of internet connections in Ghana’.

As the EmoTect algorithm was evaluated in Section 5.2.1, the output was comparably favourable, though the researcher believes that more emotionally charged life stories of students are required to achieve more accuracy. Consistent with the findings from the contextual evaluation, a majority of the participants (70%) were satisfied with EmoTect’s output. An equal number, 70%, also considered EmoTect an efficient system for tracking emotions and sentiments in text. While 10% had no opinion about its accuracy level, 20% of the participants were not convinced that emotions and sentiments could be tracked accurately by EmoTect. The reason for their scepticism is that some of the output did not convince those participants. For instance, the participants found the system to have problems dealing with some sentences that contain negation phrases such as ‘I am not happy’.

Consistent with this researcher’s observation, the counsellors did not have much to complain about concerning the disagreements in EmoTect’s output (sentiments) against their own predictions. Nonetheless, there were a few instances where counsellors merely disagreed with the output from the emotion classification part. The minority of the participants who disagreed (30%) with the output from the emotion classification component were of the view that some of the outputs were contrary to their expectations. This was to be expected since emotion is a subjective phenomenon. This is why EmoTect allows users to modify the default training data to avoid these sorts of pitfalls. The keyword output met the expectations of the participants. None of the participants complained of any rejection of their proposed ideas when
gathering data for development. This made one of the participants express his opinion thus:

‘Since the platform is new to me, I cannot think of anything else. I only think that for now, everything we want is included. I will think of some new ideas in the future should I be contacted again’.

To verify the responses from the participants after the evaluation, they were asked to talk about the type of data they used in testing EmoTect. From their answers, this author deduced that random text from the web, students’ textual submissions from external sources and submissions from the ‘contact counsellors’ widget form were the main sources of data. However, participants reported that some of the text data did not produce any results. This was to be expected because the content of such data may not have contained emotionally charged words to generate predictions.

‘I used some random text coupled with some responses from the contact counsellor to test the platform. However, some of them gave no results in both the emotion and the sentiment parts’.

| Table 5.10. Frequency of the Impact of the EmoTect in counselling delivery |
|-----------------------------|----------------|----------------|----------------|
| **Questions**               | **Scale**      | **YES**        | **No Idea**    | **NO**         |
| 1 I believe that EmoTect can facilitate counselling delivery? | 1, 2 | 9 | 0 | 1 |
| 2 Are the outputs from EmoTect desirable for your expectation? | 3 | 7 | 0 | 3 |
| 3 EmoTect served the purpose on which it was developed? | 4, 5 | 10 | 0 | 0 |
| 4 EmoTect performed well to my expectation | 1 | 8 | 0 | 2 |
| 5 I agree that EmoTect is capable of tracking emotions and sentiment in text? | 2 | 7 | 1 | 2 |
| 6 Were you expecting something that was not included in the EmoTect system? | 3 | 0 | 0 | 10 |
| 7 I will use EmoTect for my counselling works? | 4 | 6 | 0 | 4 |
| 8 Have you ever used any system that works the same way as EmoTect before (for emotion and sentiment analysis)? | 10 | 0 | 0 | 10 |
| 9 Did you find any difficulties with EmoTect functionalities? | 10 | 0 | 0 | 10 |

Generally, Table 5.11 shows that all the participants (100%) voiced the opinion that EmoTect serves the purpose for which it was developed. According to the participants (90%), every part of EmoTect works, since it outputs emotions, sentiments and emotion keywords from text documents even though some participants did not highly regard some output. None of the participants (100%) found any difficulties with the operation of EmoTect. According to this group, every component of
EmoTect worked according to their expectations, without any functional setbacks. The participants highly anticipate the adoption of such a system in their counselling work (60%). They, however, pointed out the challenges of e-counselling implementation in most senior high schools, which according to the counsellors, is still in its infancy and lacks the needed resources for a widespread adoption of that kind of system. A typical challenge is a lack of or poor internet and electricity in most of the schools. This was the reason some of the participants (40%) were reluctant that they could use EmoTect in their counselling work.

Given the human capacity for introspection, people have a tendency to opt for an alternative that meets their requirements. The simplicity of an artefact is undeniably one of the key elements which influence people’s decision to adopt an IT tool (Prat et al., 2014). Table 5.12 represents the perception of participants about the simplicity of EmoTect. As seen in the table, most of the participants (90%) did not find any difficulties with the interface and admired the aesthetic view of EmoTect. All the participants (100%) found it easy to traverse EmoTect's pages, which include the various components of the platform: contact counsellor form, sentiment, emotion and keywords.

A majority (90%) of the participants expressed satisfaction with the visualisation output. While 60% of the participants had no difficulties configuring EmoTect before using it for the first time, 40% did encounter configuration difficulties. Participants raised concerns about the technicalities in copying scripts from the EmoTect page to their respective pages to bring up the personalised ‘contact counsellor’ widget form. The author recommends the school counsellors seek the help of IT experts in the configuration process. A participant wrote:

‘The initial stage where we have to copy something from your website into our website is confusing. This is because I have no idea about computing, and we did not develop the website. Although I understand that it needs to be done once, but I believe that it is the only problem. And since we have to use the service of the ICT people, then that is fine’.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Scale</th>
<th>1, 2</th>
<th>3</th>
<th>4, 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you find it difficult to use EmoTect for counselling works in terms of the interface?</td>
<td></td>
<td>9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Was it easy to traverse all the pages of the EmoTect page?</td>
<td></td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>The visualisation outputs from EmoTect are good and desirable?</td>
<td></td>
<td>9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Did you find the aesthetic view of EmoTect appropriate?</td>
<td></td>
<td>9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>The Configuration of EmoTect before its use was difficult?</td>
<td></td>
<td>4</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>
6 DISCUSSION

The discussion here will take a holistic view of the study, covering its initial stage of requirement elicitation to the development of EmoTect and including all of the original publications. The limitations, constraints and lessons learnt from the three-year study are also discussed.

6.1 GENERAL DISCUSSION

The challenges students face in the course of their lives often generate negative emotions that may diminish their capabilities to fulfil their academic potentials, as this study has just revealed. It is clear from [PI] that some students do have inner life stories that are closely guarded and never shared with anyone, including counsellors. Others may have discussed their challenges with peers but lack strategic capabilities to help them deal with the problems. Despite the challenges of providing counselling services in schools, it is an indispensable strategy for guiding students towards academic success. This concept largely hinges on the availability of the required resources for counselling in schools and the level of professionalism of counsellors. Understanding student emotions and fostering coexistence among them in schools creates a conducive environment for them to progress in their academic careers. Ghana is an emerging economy, yet it is full of ambition to improve its education sector. However, insufficient attention has been paid to provide the needed resources to create a robust counselling sector.

Given the advances of technology, counselling is no longer limited to face-to-face communication, where students have to meet counsellors in person (Watts, 2001). Existing ICT tools have shifted the paradigm; students can now receive counselling online (Rall, 2011; Shiller, 2009). Diverse technologies are available to assist in counselling delivery. For instance, artificial intelligence has made it possible to provide counselling to students without human intervention. Often, students who are geographically isolated and urgently needing counselling can now turn to online media platforms for such services.

As discussed in this dissertation, counsellors are likely to be overwhelmed with large amounts of textual submissions from students who may need remote counselling as students’ population keeps increasing. In Ghana, e-counselling is still at an infant stage, but when the demand for e-counselling increases – as it is likely to in the near future – the workload of counsellors will also inevitably grow. In effect, continuing to rely on manual processes of tracking emotions buried in large volumes of text will no longer be efficient in informing decision making. It is also going to be very costly. Therefore, the need to pivot to computational methods for recognising emotions in text will certainly take precedence over others, as discussed earlier.
Given the interdisciplinary nature of this dissertation, a literature review is provided as background as well as related research in the field of counselling, emotions and natural language processing (NLP). The author has explained both concepts and justified the need for integrating NLP into counselling as it falls under human language technology. NLP techniques for text mining, especially automatic emotion and sentiment classification, have been addressed in the review of the literature. In the same vein, existing contributions of NLP in education were delved into with the relevant literature. The challenges of e-counselling implementation in Ghana are discussed as well. The purpose is to create awareness of the various challenges in connection with the implementation of e-counselling in Ghana. This is not to dampen the motivation to further develop the field but to stimulate awareness in stakeholders about the need to address those challenges. Based on the e-counselling challenges, this researcher reckoned that web-based NLP systems are perhaps not the ideal platform for counselling in Ghana, but they still stand out to be the most preferred choice in terms of providing remote counselling to students. Since majority of Ghanaian youth are in the SHSs, policy makers need to step up their effort to augment the counselling centres with state-of-the-art ICT infrastructure.

Most students in Ghana shun the existing face-to-face delivery of counselling primarily because students are sceptical of exposing their privacy to unknown counsellors and rather prefer anonymous counselling, as this researcher’s finding in [PI] shows. This researcher’s study in [PI] has also uncovered trust issues in connection with the traditional face-to-face method which have created more preference for e-counselling; students think it provides solace. Kuhn (2004) concluded in his study that trust is a fundamental ingredient to prod an individual to open up or divulge vital information, yet many counsellors do not have the requisite skills to create an atmosphere of trust during face-to-face counselling. Although e-counselling is gradually being accepted in many parts of the world, Glasheen et al. (2013) revealed that many counsellors are reluctant to provide online counselling to students. Fletcher-Tomenius and Vossler (2009) identified trust as a particularly ‘important aspect of online interactions, especially in regard to the fact that cues and signals such as facial expression, tone of voice and gesture are not available online. However, there are only few studies that have investigated this factor in the context of an online therapeutic relationship’.

Counselling in the education sector aims to provide equal opportunities for students, irrespective of their background and location. Even though face-to-face counselling is still relevant, ICT-mediated counselling has introduced an option for students to choose their most preferred mode of counselling delivery. The need to verify further the factors that influence students in their selection of e-counselling led to [PII]. In [PII], social influence and performance expectancy were the perceived factors that would motivate students to turn to e-counselling. In other words, students have increased expectations in e-counselling to advance in their academic careers when colleagues such as peers, counsellors and teachers encourage its use. With
these findings, this author recommends that symposia regarding the use of ICT in counselling be encouraged in schools.

The emotional and personal-social development of students is important to their academic achievement and for the development of a school at large (Valiente et al., 2012; Reyes et al., 2012). Therefore, developing a computational system to extract emotions and sentiments based on students’ textual submissions as a mode of providing counselling is one way to complement the work of counsellors, especially in the decision-making process. As DSR was employed in this study, EmoTect’s development was broadly categorised into three processes, mainly focusing on requirements’ elicitation, implementation and evaluation. The various stages of development are illustrated in the resulting papers attached to this dissertation in Appendix 3. Understanding the environment includes the requirement identification for the development of EmoTect. As explained earlier, preliminary studies were conducted with selected counsellors and students in Ghana. [PI], [PII], [PIII] and PIV reported part of the understandings of the study’s context. Participants, especially counsellors, proposed ideas and their expectations of EmoTect before its development. The rationale for understanding the study’s environment is consistent with Simon (1996), who pointed out that DSR artefacts must be developed with a clear understanding of the environment, i.e., the intended users of the artefact. The counsellors and students who participated in the entire study expressed some scepticism of the possibility to extract emotions from text. This was to be expected, as this author’s unpublished preliminary research had found.

Psychologists have discovered that humans exhibit a high level of consistency in recognising emotions in text, but there is a great deal of variability in an individual’s ability to recognise emotions in text (Yoshihiro & Kato, 2011). This is consistent with the major findings of [PIII]. In [PIII], we justified that EmoTect and its related applications complement the work of counsellors and reduce the variability effect in recognising emotions in text. In effect, students’ emotional changes could be tracked by EmoTect for a period of time. Another finding in [PIV] was that counsellors found it easier to annotate emotions in one or two sentences rather than in a paragraph that contains five or more sentences. The challenge could be attributed to the multiple number of emotions that could be in one paragraph.

Apart from the initial evaluation during the development of EmoTect, a final evaluation of the EmoTect classification algorithm was carried out. The results from evaluating the EmoTect classification algorithm are promising and appear set to be adopted for counselling though more data are required to improve the level of accuracy. This was confirmed in the contextual evaluation, where most of the counsellors agreed that the output was satisfactory. Prior to the development of EmoTect, [PIII] and [PIV] confirmed the variations in counsellors’ annotation agreement of emotions in text. These papers are consistent with the findings of Mulcrone (2012), who believed that the subjective and subtle nature of emotions makes it difficult to achieve consistency and high levels of accuracy in tracking the emotions of others in text. The
user interface and the simplicity of using EmoTect met the desire of the participants without necessary calling for further modifications.

In a nutshell, this work led to the development of a supervised machine learning system (EmoTect) for tracking emotions and sentiments based on an individual’s perception of emotions. From the analysis in [PIII], this author found that counsellors, when analysing their students’ emotions in text, are likely to be influenced by external factors. For instance, a counsellor’s domestic difficulties could have a ripple effect on his or her judgement regarding a decision about students. It is in light of this that EmoTect was developed to allow users to label the training data based on the perception of emotions. This, in effect, alleviates the possible dissatisfaction of counsellors in supervised machine learning outputs, since it will depend much on their emotion perceptions. Working with EmoTect, external emotional influences are reduced, and that in turn, boosts the level of consistency and efficiency. EmoTect has the capability to extract emotions and sentiments in textual content, particularly from students’ textual submissions. Counsellors can use EmoTect to monitor the emotional trends and changes of their students over a selectable period.

6.2 LIMITATIONS, CONSTRAINTS AND LESSONS LEARNT

The entire three-year study was not without tremendous challenges. In this section, the author delves into discussing some of the difficulties and lessons learnt from the entire research.

Given that the DSR framework was employed in this study, there was the need to meet up with the participants periodically in their schools. There were many difficulties in meeting with the participants, especially the counsellors. These difficulties were in connection with organising a joint meeting of the students and counsellors from their respective schools to meet at the same venue. The inability to do this at a convenient time for all parties brought the work to a halt at one point; counsellors were either occupied with their official work, or students had lessons. This problem was not factored in before the research began. The most challenging element was the annotation exercise that counsellors were tasked to do manually; it took them several months to respond. Clearly, the slow pace created deep frustrations for the author. Besides, the annotation was done in two rounds on two different days. This in no way is meant to place any blame on the counsellors, merely to report the constraints of the study. In addition, the counsellors, students and assistants to the author helped voluntarily.

This research is limited to only text-based emotion detection for providing counselling, but emotions can be expressed in other forms, as elaborated earlier in the dissertation, which brings up one of the limitations of this work. For instance, emotions can also be expressed in bodily cues and speech, among others. Some personality traits of individuals do not show up in their written texts, and that constitutes
one of the limitations of this work. The author, therefore, concedes that these constitute a substantial limitation on the use of EmoTect in counselling, especially in the SHSs. Another limitation of this work is the detection of emotions in emoticons or emoji content. This was not considered in the development of EmoTect, on the assumption that Ghanaian culture does not recognise the formal submission of text content using emoticons or emoji, especially within counsellor-student interactions. Though the system does not recognise them, some students may still use them, and this is considered a limitation to the use of EmoTect. It is in light of this that EmoTect is proposed as a complementary tool for counsellors, rather than replacing their work.

One of the challenges faced during the study was building the life story corpus. Some of the students who were provided with questionnaires to write about their emotionally charged stories or emotional antecedents failed to return their questionnaires because they were allowed to do so at their own pace. Others did not even respond to the questions that touched on their life story. It took more effort to get students to write about their stories. As the author resided in Finland and had to conduct the study in far-off West Africa (Ghana), frequent travelling was needed between the two far-flung locations, which was also a substantial challenge to surmount. Though the author was ‘firing on all cylinders’ to get things done in a timely fashion, participants and assistants to the researcher were held down by their different agendas that slowed down the work. While the counsellors were convinced that EmoTect is a good platform for counselling in Ghana, there was still scepticism towards its widespread adoption in Ghana due to the reasons already explained about implementation discussed in Section 2.1.3.

The challenges notwithstanding, the author learnt many useful lessons from the study. For instance, the scope of this study included a game-based NLP platform at the beginning but was narrowed down to an automated e-counselling system for emotion and sentiment when the initial scope turned out to be too wide. The challenges and difficulties mentioned have armed the author with useful experience that could be used to minimise similar challenges in the future.

6.3 RECOMMENDATIONS FOR STAKEHOLDERS

The findings of this dissertation have shown some evidence of deviations in the planning and implementation of e-counselling in Ghana. In this section, recommendations and guidelines for the implementation of e-counselling in Ghana are proposed. These recommendations are primarily proposed for counselling practitioners and other stakeholders of education in Ghana (government, parents, teachers and non-government organisations, among others). The following are the proposed recommendations:
School counsellors should intermittently carry out a survey of their students regarding their counselling needs, especially about the mode in which students prefer to contact counsellors and share their life experiences. This will enable school counsellors to adjust their activities to meet the changing needs of students. For instance, there may be students who only wish to share vital information that threatens their academic development to counsellors only through anonymous means, as this study has found [PI].

As reported in this thesis (Section 2.1.3), counselling in schools is not well resourced. Hence, it is recommended that the Government of Ghana develop a policy framework for e-counselling, thereby consolidating it with the existing school curriculum. In this case, there should be a budget allocation for the running of the counselling in schools. This could be accomplished by getting input (expert advice) from the various stakeholders of education, including counsellors, teachers and policy makers.

Given technological advances, universities in Ghana should develop courses that will seek to produce more capable counselling personnel to use ICT infrastructures. This recommendation is based on the various challenges of e-counselling implementation found in this study, such as counsellors’ lack of technical knowhow regarding the use of ICT in counselling. The researchers did not delve deeper into the kind of courses the universities teach to counsellors. Nevertheless, based on the challenges of e-counselling in Ghana, the university curriculum should consider more ICT courses for counsellors.

The existing teacher professional development, which is organised intermittently, should include counsellors, thereby orienting them to the use of ICT in counselling. Counsellors could be introduced into the emerging counselling technologies, such as educational NLP-based systems, to support their counselling work. This could be accomplished through the organisation of symposia and workshops for counsellors.

N.b. that the government and other stakeholders of education should capitalise on the state-of-the-art ICT technologies, such as artificial intelligence, for counselling work. Artificial intelligence could be developed into a system capable of mimicking counsellors and delivering counselling to students. This can be likened to EmoTect, where the system can predict the emotional state of students in their textual submissions.

Given the sensitive nature of students’ academic challenges, e-counselling or online counselling platforms should be developed by due cognisance of the security requirements, thereby avoiding what could be disastrous in the case of an information leak. This should be critically examined by ICT experts by making sure that confidentialities and data protection are assured while developing any ICT tools for counselling.

To make use of e-counselling in the various schools, the author recommends that the government and other stakeholders of education take a critical look
at the ban on the use of mobile devices in the SHSs. That is, they should analyse the strength and weakness of the use of mobile devices in schools – through SWOT analysis, for instance.

➢ The electricity supply and the cost of the internet should be critically considered since it represents a great deal of the challenges confronting e-counselling implementation and adoption in Ghana. This would not only improve counselling delivery but other academic activities also, such as e-learning and improved administrative activities, among others.

➢ The researcher understands that the government of Ghana is implementing free SHS education from this year, 201721. As the aim is to gear up for a quality, affordable education for all, the Government should consider setting up committees to investigate further the requirements for the implementation of e-counselling.

➢ As found during this work, students are motivated to use e-counselling tools for counselling for their academic achievement through social influence [PII]. The author recommends that education stakeholders, such as parents, among others, should play their part in encouraging students to access e-counselling, especially students who desire anonymous counselling.

7 CONCLUSION

In this dissertation, requirements for the implementation of e-counselling for automatic emotion and sentiment classification were investigated. Human language technology (HLT) was employed to contextually develop a web-based e-counselling system for decision making. The system was developed from a supervised support vector machine (SVM) learning classifier. Because of this, students’ life story corpus was developed and used to train the classification algorithm (SVM). Training the EmoTect classifier was based on the user’s perception of emotions. Therefore, counsellors are given the opportunity, at the EmoTect interface, to personally label the training data based on their perception of emotions. By using EmoTect, counsellors would be able to monitor the emotional changes of their students over selectable periods. The DSR framework was employed in the development of the EmoTect system. The system was evaluated to determine the efficacy of the algorithm and its functional performance.

7.1 ANSWERS TO THE RESEARCH QUESTIONS

This researcher concludes this dissertation by reflecting on the various answers to the research questions formulated in Section 1.2. In this regard, the section also concludes the original publications and the manuscript selected for this dissertation (i.e. PI–PV) with regard to the objectives of this study.

RQ1: What are the emotional life challenges that threaten students in their academic pursuits?

Pursuant to this research question which was investigated in [PI], students were selected using a stratified sampling technique from a senior high school in Ghana. The counsellor of the school was also selected to be part of the participants. A questionnaire was designed and administered to the selected students for their response. A key question in the questionnaire touched on students’ emotional life stories. The life story of students was clearly defined to the students as stories that threaten their academic development. A codebook (themes) was developed from the content of the life stories: academic, career and psycho-social challenges. The content of the collected stories was classified manually based on the developed codebook. The data were thematically analysed.

In the end, the study revealed several latent life challenges that students encounter during their studies. It was found that students were reluctant to seek face-to-face counselling because of the lack of trust in their counsellors. In addition, students ex-
pressed their willingness to seek counselling anonymously. It was found that students’ life stories contained emotional and attitudinal content that raised the need to build a life story corpus from the students’ life stories. This resulted in the ‘life story corpus’. The life story corpus, which contains emotional elements, was used as a training and testing dataset for EmoTect.

**RQ2: What counselling technologies are being used in senior high schools of Ghana?**

This research question was investigated partly in [PI], and as well in the final evaluation of EmoTect in this dissertation. To understand the state of counselling technologies in the SHSs, selected counsellors were asked to respond to questions that tasked them to tell about the kind of technologies they use in conducting counselling with students in their various schools. The aim of this question is to understand the extent to which counsellors use e-counselling. Questionnaires and an interview were the main technique for the data collection. Just as in RQ1, the content of the data was analysed. After the analysis, it was found that most counsellors do not use any ICT tools for counselling delivery, except for a few of them who use email to contact students’ parents. However, email is rarely used in counselling. In further studies, it was found that WhatsApp is another platform that counsellors have adopted of late to contact students and their parents. As expected, no contextual NLP tools for emotion and sentiment analysis were found in use by counsellors to complement their work.

**RQ3: What are the factors that motivate students to adopt and use e-counselling in Ghana?**

This research question was examined in [PII] as part of the preliminary studies to gather requirements for EmoTect’s development. UTAUT, a model to investigate people’s intention to use technology, was adopted (Venkatesh et al., 2003). The four UTAUT constructs adopted in this dissertation which have been explained in [PII] are: performance expectancy, effort expectancy, social influence and facilitating condition. Randomly selected students were tasked to respond to close-ended questions using a seven-point Likert scale. The questions were developed based on the aforementioned UTAUT constructs. The collected data were analysed first using Cronbach’s alpha for reliability (Cronbach, 1951). The study revealed a good internal consistency of the data. Subsequently, MLR analysis performed on the data found performance expectancy and social influence as the influencing factors, while effort expectancy and facilitation condition emerged as insignificant to students’ behavioural intentions regarding the acceptance of e-counselling in Ghana. In effect, the contribution of this paper is to create interest in researchers concerning e-counselling in Ghana and perhaps other parts of the world about the factors that influence behaviour acceptance and the use of e-counselling. The role that parents, counsellors, peers and teachers play in the adoption of counselling technologies cannot be overlooked. It was for this
primary reason that counsellors and students were made to be part of the development of EmoTect.

**RQ4: Does the emotional state of counsellors influence their emotional perception while annotating emotions in text?**

Supervised machine learning systems for the automatic detection of emotion and sentiment require that training data be annotated for high-level accuracy (Maynard et al., 2012). This research question was investigated in [PIII] and [PIV]. In line with the context of the study, students’ life stories were collected and complemented with sample ISEAR data. The aim was to explore the extent to which counsellors’ inner state of emotions influence their perception of emotions when analysing others’ emotions in text. The text corpora (life story and ISEAR) were each pre-processed to a form suitable for manual annotation by counsellors. Counsellors were then tasked to annotate with Plutchik’s basic emotions in two rounds and on two different days. With this, the author computed the annotation agreements of emotions among the counsellors (inter-annotation agreement) and the agreement of emotions for each counsellor between the first and second rounds of the annotation exercise (intra-annotation agreement). Consistent with Lerner et al. (2003) and Loewenstein et al. (2001), it was discovered that counsellors’ inner emotions affect their perception of others’ emotions while annotating emotions in students’ textual submissions. Moreover, a perfect intra-annotation agreement from each of the counsellors was found when given the opportunity to annotate emotions in text in several sections. This called for the idea of designing EmoTect to allow counsellors to train the EmoTect classifier based on their own perception of emotions. Further investigation revealed that counsellors had difficulties in annotating paragraph-length text that contained more sentences than sentence-length text documents.

**RQ5: How can a text-based automatic emotion and sentiment classification system be constructed for counselling delivery?**

Given that EmoTect is meant to complement counsellors’ work, counsellors and students were intermittently involved in the design and development of the system. In essence, a DSR approach was employed by adopting Peffer’s DSR framework. The framework is systematic, which allowed the understanding of the environment by eliciting requirements for the implementation of IT artefacts. Therefore, the requirements for the development of EmoTect were partly elicited from the counsellors. In so doing, counsellors were intermittently tasked to evaluate the system until the final version was produced. In developing the system, the NLP technique was explored by employing a machine learning multi-class SVM in the development of the system. Other packages used in the development of EmoTect include NLTK and Weka’s multi-class SVM classifier. The POS tagger from NLTK was used for syntactically parsing the input data. In addition, the lemmatisation package from NLTK was used
in lemmatising the input data into their lemma. By training the classifier, the author used his developed life story corpus which is a collection of students’ life stories. Based on the findings from RQ4, the training of the classifier is optionally possible for counsellors to train the classifier based on their perception of emotions.

**RQ6: How can a text-based emotion and sentiment classification system be evaluated?**

Summatively, two phases of evaluating EmoTect were carried out – classifier evaluation and contextual evaluation. Before, the EmoTect prototype was formatively evaluated [PV]. In the final evaluation of the classifier, sample test data from the life story corpus were manually annotated by three different counsellors in two rounds and on two different days. As pointed out earlier, the annotation agreement of emotions was computed for each counsellor based on the various rounds of annotation. In the end, a gold standard annotation corpus from each counsellor was obtained. The test data were fed into EmoTect to classify according to Plutchik’s basic emotions and positive and negative emotional valence. Precision, recall and the F-measure were used as the evaluation matrices of the system. This was done by comparing the output from EmoTect to the gold standards obtained from the counsellors. In the contextual evaluation, the EmoTect classification algorithm achieved accuracy comparable to that achieved with the gold standard, even when presented with unknown data.

The second part of the evaluation was in the environment with selected counsellors and teachers. The EmoTect system was demonstrated to the participants, and data collected for analysis. From the analysis of the data, the interest of counsellors and teachers was aroused concerning the capabilities of EmoTect. They have expressed a desire to adopt it for counselling delivery. The users found the various functionalities of the system suitable but expressed concern over the poor internet connectivity in Ghana, which is a potential challenge to how the EmoTect system would be widely adopted. Nevertheless, the participants found it easy to utilise EmoTect during the testing phase. In addition, counsellors advocated for further refinements of the system in terms of the output, though a majority of the participants were satisfied with the output.

### 7.2 RESEARCH CONTRIBUTION

A major part of the contribution of this dissertation stems from the selected publications. The general contribution of this paper is where human language technologies were explored for counselling, thereby facilitating the efficient delivery of counselling to students in terms of their emotional and personal-social development. The key idea for the EmoTect system is rooted in the fact that many students prefer to be
counselled anonymously [PI]. The system allows students to contact counsellors and analyse automatically extracted emotions/sentiments in students’ submissions. The intent of the platform is to help counsellors in their decision making concerning students. The study used DSR, and this contributes to the growing pool of the use of DSR in artefact development.

One of the core contributions is the development of a life story corpus. To the best of the researcher’s knowledge, there is no existing corpus that has been developed which specifically targets students’ life stories bordering on their academic development. The collected students’ stories were given to trained counsellors to annotate with Plutchik’s basic emotions and for negative and positive valence. The researchers are intending to make the annotated life story corpus available online for research purposes. Hence, the study has contributed to the pool of annotated corpora on the grounds of research, specifically in the e-NLP of text.

As has been the traditional approach of training a supervised system with all-in-one inter-annotated training data, this study allows counsellors to label training data based on users’ own perception of emotions. This is because of the subjectivity of emotions and how individuals perceive emotions in text. The training data are made available in the system interface where users have the opportunity to label the training data based on their own perception of emotions, which does not require any technical knowledge to label or tag.

7.3 FUTURE STUDIES

Although the study considered Ghana as the context, EmoTect can be used globally. In the future, other classification (including deep learning approaches) algorithms such as Naïve-Bayes, convolutional neural network (CNN) and Long-short term memory (LSTM) network shall be considered and compared with the SVM used in this dissertation. In the future, this researcher is interested in broadening the scope of this study to embrace other sectors of education, such as e-learning. This would be geared towards learning analytics. The life stories of students containing emotive elements were collected with a questionnaire that took much time to annotate. Accordingly, in the future, the researcher intends to develop an NLP system that would enable the online collection of emotionally charged stories from students to alleviate the challenges encountered with the manual process.

As game-based NLP was one of the initial objectives of this dissertation, the researcher intends to upgrade EmoTect to incorporate a game-based part that can help counsellors in delivering efficient counselling to students. The EmoTect platform shall be upgraded to alleviate the challenges it poses in text classification, as found
during the evaluation phase. The researcher, in the future, is also interested in incorporating a major local language from Ghana, called *Twi*\(^{22}\), into the system. The motive for this is because many students decline to contact counsellors because of the difficulties in communicating with the English language. This will give opportunities to students who found themselves in that position to contact counsellors. In this light, EmoTect will be able to track emotions and sentiments in the local language. The task shall lead to teaming up with linguists in the area to develop a language corpus for the project.

---

8 BIBLIOGRAPHY


110


Maynard, D., Bontcheva, K., & Rout, D. (2012). Challenges in developing opinion mining tools for social media. In *proceedings of the NLP can u tag user generated content.*, (pp. 15-22.).


This thesis contextualises the application of human language technologies for counselling. To this end, an e-counselling system – EmoTect – has been developed for the automatic detection and analysis of students’ emotions. A life story corpus has been built to train and test the EmoTect classifier. The corpus is made freely available for research purposes. Additionally, this work has demonstrated how an e-counselling system, built with machine learning capabilities, is trained and used based on individual user’s perception of emotions.