Development and Evaluation of an Automated e-Counselling System for Emotion and Sentiment Analysis

Emmanuel Awuni Kolog, Calkin Suero Montero and Markku Tukiainen
School of Computing, University of Eastern Finland, Joensuu, Finland
emmanuk@uef.fi
calkins@uef.fi
markku.tukiainen@uef.fi

Abstract: Given the challenges associated with the analysis of emotions in text by counsellors, we present an intelligent e-counselling system for automatic detection of emotions and sentiments in text. The system- EmoTect- was developed using a supervised support vector machine learning classifier. Therefore, students' life stories were collected and developed into a corpus for training and evaluating of the classifier. EmoTect allows users to label instances of the training data based on their own perception of emotions, and then gradually learns to classify emotions according to the user's perceptions. The EmoTect interface provides a visualization of the emotional changes from automatically analysed students' submissions over a selectable period. In this paper, the EmoTect classifier is evaluated with a gold standard corpus obtained from students but annotated by counsellors. In addition to the classifier evaluation, the EmoTect prototype was evaluated with counsellors in their settings. From the experimental results, the EmoTect classifier for the sentiment classification achieved comparable accuracy to that achieved with a gold standard when presented with unknown data. The contextual evaluation of the system indicates counsellors' satisfaction and sense of enthusiasm for using EmoTect for counseling delivery.

Keywords: Counselling, Design science research, Emotion classification, Evaluation, Sentiment analysis, Support vector machine

1. Introduction

Emotion analysis forms an integral part of students' emotional and personal-social development (Barner & Brott, 2011). Personal-social domain of counselling partly encapsulates the strategies of helping students to unearth and manage their latent emotional and personal-social challenges (Alexandria, 2014). With variation in students' emotions and personal-social needs, school counsellors are mandated to facilitate the emotional learning opportunities and restorative approaches of students. Therefore, being able to deduce the emotional intensity of students, counsellors can predict the behaviour of their students towards academic engagements and performance (Ji, 2015; Pekrun *et al.,* 2002). In effect, counsellors are expected, as part of their profession, to devise strategies for understanding the emotions of their students, thereby motivating a successful academic achievement (Kolog, 2014, Paulus & Yu, 2012). Meanwhile, counsellors often counsel students through their own intuitional effort, especially in the context of counselling in Ghana (Kolog & Montero, 2017). For this reason, complementing the intuitional efforts of counsellors with Information and Communication Technologies (ICTs) is a way of ensuring efficiency and consistency in counselling delivery (Kolog & Montero, 2017).

While analysis of emotions in text has widely been studied in the domains of psychology and behavioural sciences, there is an increasing interest from the computer science researchers. This interest is mainly centred on the use of the computational approaches for detection and analysis of emotions in text. Computational tracking of emotions in text is useful in counselling delivery as it can be used to complement the consistencies in decision making processes (Kolog *et al.*, 2017). This is especially relevant when large volumes of data are involved. Baumeister and Bushman (2007: p. 61) define emotions as 'a subjective state, often accompanied by a bodily reaction and an evaluative response, to some event'. Thus, emotions can be manifested physiologically (Tobin *et al.,* 2016; Chopade, 2015) and linguistically (Hancock, 2007). In either case, emotions could be expressed through blood pressure, speech, facial expression (Damasio, 2000) and text (Schalke, 2014). However, the focus of this work is to detect emotions of students in text.

Textual submissions by students towards seeking counselling are most likely contain emotional expressions (Kolog *et al.,* 2016). With this, emotions in text can be spotted by analysing the linguistic information in them, usually through Natural language processing techniques. Linguistic information in text document is valuable for identifying emotion words, thereby aiding decision making processes. Paradoxically, manual analysis of

Reference this paper: Kolog, E. A., Suero Montero, C. and Tukiainen, M., 2018. Development and Evaluation of an Automated e-Counselling System for Emotion and Sentiment Analysis. *The Electronic Journal Information Systems Evaluation,* 21(1), pp. 1-19, available online at www.ejise.com

linguistic information in text is not efficient when it concerns large volumes of text documents. The application of automatic emotion analysis is diverse and the field is steadily being applied in different sectors, such as education, business and health sectors. For instance, business organisations use sentiment analysis (Coarse-grain emotion analysis) applications to study the perception and opinions of their customers regarding products and services (Balabantaray *et al*., 2012). Educational text mining is one of the domains in Computer science that has attracted a considerable interest for the understanding of learners' behaviour, usually through students learning diaries and course feedback (see for example: Munezero *et al.,* 2013).

Vinluan (2011) believes that counselling has benefited from ICT for many years. While we agree with Vinluan (2011), we also believe that more work in ICT-mediated counselling is still needed, especially for exploring how indigenous knowledge could be integrated into ICT (contextualisation). ICT-mediated counselling simplifies the work of counsellors and provides opportunities for students with diverse backgrounds to vary their counselling accessibility (Kolog, 2014). Since school counsellors are humans their emotional states, at a point, could directly or indirectly influence their judgement when analysing emotions of their clients in text (Kolog *et al.,* 2016; Lerner, 2003; Paulus & Yu, 2011). This is consistent with the findings of Han *et al*. (2007) who pointed out that different emotions expressed by individuals, could affect their perception of emotions, and perhaps clown their judgement when making decision. Therefore, to minimise the influence that counsellors' own emotions have on the perception of emotions on others while analysing text documents, we have developed a system for automatic emotion and sentiment classification, which is based on the Plutchik's basic emotions. The system, called *EmoTect*[1], is multi-functional and comprises two components: *contact counsellor* and *emotion detection*. The *contact counsellor* form 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. The primary purpose of EmoTect is to complement the work of counsellors by aiding them in their decision-making process concerning students. The development of EmoTect followed a design science research (DSR) process with end-users (students and counsellors) participation. The prototype version of EmoTect and its classifier is evaluated and presented in this paper. Hence, this paper forms part of the initial stages of an ongoing DSR project.

2. Background

Automatic detection of emotions in text has widely been applied in diverse domains. Its relevance is eminent given that a lot of people are accepting to use ICT in their daily life activities, especially through social media and other personalised platforms. This, in turn, brings about a considerable increase in online data, which has fuelled the need for big data analytics with the aid of computational algorithms (Shina & Choib, 2014). In this section, we delve into discussing machine learning, some related works on emotion detection and a Design science research with participatory design approaches.

2.1 Machine learning

Machine learning (ML) is an application of Artificial intelligence that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning has widely been used in diverse areas of Artificial Intelligence. For instance, domains such as NLP, affective computing, human computer interaction uses ML algorithms for extracting knowledge from data or text. Machine learning in text classification comprises two main learning approaches: *supervised* and *unsupervised*. On the one hand, the supervised machine learning involves the use of labelled data to train a classifier, thereby creating a model for classification. After the classifier is trained, a model is obtained and is made to predict unclassified or unknown data input. Mathematically, let *T* denote text and *s* denote a word, sentence or paragraph that contains emotion features, where $s \in T$. In addition, let *n* be the number of emotion class $E = \{x_i, \ldots\ldots x_{n,}\}$ where *i* = 1,2,3…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.

On the other hand, the unsupervised learning does not require labelled data for training a classifier. Rather, patterns from unclassified data are detected using cluster algorithms. When labelled and unlabelled data are combined to train a classifier, the resulting method is termed *semi-supervised learning*. The supervised ML has

[1] *EmoTect is available at: http://nlp4counselling.com/login.jsp*

been found to perform better than that of the unsupervised learning (Sigdel *et al.*, 2015). Most efficient, accurate and commonly used classifiers for supervised ML are support vector machine (SVM), neural network, decision tree and Naïve Bayes (Munezero *et al.*, 2013; Joachims, 1998; Hassan, Rafi and Shaikh, 2011; Matwin & Sazonova, 2012). With this in mind, we employed a multi-class SVM classifier in building the EmoTect system.

Support vector machine (SVM) is a ML classifier for supervised classification. The algorithm is predominantly used for text classification. SVM is a discriminative algorithm in a sense that it is defined by constructing a hyperplane or a set of hyperplane in a high dimensional space (Hashem & Mabrouk, 2014). The hyperplane in the higher-dimensional space is defined as the set of points whose dot product with a vector in that space is constant. When training data is presented to SVM, a model is built which consists of data points chosen from input data space and their class labels. SVM outputs 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 (the so-called functional margin) since in general the larger the margin the lower the generalization error of the classifier (Hashem & Mabrouk, 2014). SVMs find solutions of classification problems that have "generalization in mind" and they are able to find non-linear solutions efficiently.

2.2 Related works

A similar platform to EmoTect is the one developed by Munezero *et al.* (2013). Munezero *et al*. (2013) employed a supervised machine learning technique to build an emotion and sentiment classification system. According to the researchers, their system allows teachers to automatically track the emotional changes of students through their learning diaries. Munezero *et al.* (2013) used Plutchik's (1980) eight basic emotions as the predefined emotion category while the sentiment part used negative, positive and neutral. However, some of the Plutchik's emotions were combined to give secondary emotions of *anxiety* and *frustration* which were considered as part of the emotion categories. Ascertaining the efficacy of their system, Munezero *et al.* (2013) evaluated their system with a collection of students' learning diaries which they collected from Newman's *et al.* (2008) corpus. As reported by the researchers, their system was found to perform well to their expectation.

Furthermore, Lu (2006) developed a system for the retrieval of data from the web through keyword queries. The system further extracts emotion from the retrieved data in real time and output the results in visual form. As reported by the researchers, the purpose of system was to extract and visualise emotions from the web. With this approach, two natural language processing techniques were explored: *web mining engine* and *semantic labelling* techniques. While the web mining engine allow search by specific keywords and provide answers to lexical questions, the semantic labelling tool implements semantic role labelling for semantic parsing and eventual extraction of the emotions. Lu (2006) built the system based on a manual procedural approach for creating an emotional model. Hence, their model considered seven basic emotions: *happiness, Sadness, Anger, fear, disgust, surprise* and *neutral.*

Rahman and Ahmed (2014) developed a web-based emotion detection system for social media content processing and event monitoring. Their system, termed *MediaTagger*, is an open-source software that was built on Naïve-Bayes classifier. The *MediaTagger* has two different components which work together as one. The components are the *web data* and *emotion extractor*. While the web media extractor retrieves data from the social media sources such as Facebook and Twitter, the emotion extractor takes the data and further extracts emotion words. Rahman and Ahmed (2014) did not use any predefined emotion categories*,* rather their system returns the emotion words from the data. Contrary to our system in this work, we used predefined emotion categories from Plutchik on which unseen input text is predicted by a classifier model (see Figure 3). However, our system also returns emotional words from unseen text document. By ascertaining the efficiency of their web-data and classification algorithms, Rahman and Ahmed (2014) evaluated their system with sample text document, and with end-users. Rahman and Ahmed (2014) reported higher working efficiency of their algorithm with end-users appreciating the functionalities of *MediaTagger.*

2.3 Design science research

There are various ways by which design science research (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). An 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.

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).

Before developing EmoTect, preliminary studies were conducted to gather the requirements for the EmoTect development (see Section 3.1). Students and counsellors are the users of the EmoTect system, hence they were actively involved in the EmoTect development. For instance, we investigated the behavioural intentions of students towards accepting and using e-counselling in Ghana (Kolog *et al.,* 2015). The study was an exploratory research which was empirically conducted by involving students. The findings from the study motivated the need for DSR, thereby involving counsellors and students in the developmental process. Based on Peffer's *et al.* (2006) DSR model (redrawn in Figure 1), we have simplified the various steps into three main components interlinked iteratively and sequentially. From Figure 1, the three components: *analysis and requirement, implementation* and *evaluation* are deduced and explained elaborately in the subsequent subsections.

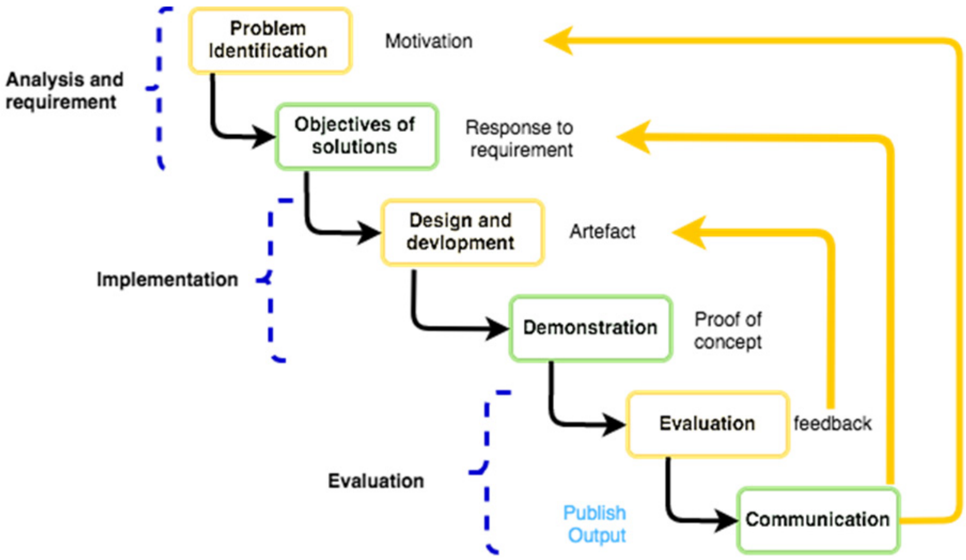

Figure 1: DSR model redrawn from Peffers' et al. (2006)

In the development of EmoTect, counsellors and students were consulted to give their input, especially during the requirement and the evaluation stages. This process is termed a participatory design (PD). 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 Roberston 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 counseling experts and experts in ICT were involved in the implementation process.

3. EmoTect implementation

In this section, we present the development of EmoTect in line with the DSR process shown in Figure 1. This section is divided into three sub-sections: *Problem* and *requirement identification, EmoTect development* and *the user guide.* To pre-empt is that EmoTect does not crawl data online for analysis, it is rather a tool that tracks emotions and sentiments from the textual submissions of students during counselling.

3.1 Problem and requirement identification

We have shown from our previous study that students have inherent life challenges that hinder their academic development (Kolog *et al.,* 2014*).* However, students rarely or do not discuss their life challenges with their school counsellors concerning their academic and life challenges (Kolog *et al.,* 2014). Awinsong *et al.* (2015) have reported similar findings in a study conducted in Ghana. The researchers found high rate of students' being reluctant to seek counselling face-to-face. This, Awinsong *et al.* (2015) attributed the reasons to lack of trust students have in their counsellors. A related study from other parts of the world have also attributed students' reluctance on seeking counselling to lack of trust in their counsellors (see for example: Le Surf & Lynch, 1999; Jenkins & Palmer, 2011; Mushaandja *et al*., 2013). Students who are willing to be counselled, in most cases, advocate for counselling through anonymity and this partly motivated this line of research. Currently, most of the counsellors in Ghana receive, analyse and adjudicate on students' emotions based on their intuitions. This process is not efficient, and as well a costly process, especially that students' population keep increasing over time. Most schools in Ghana do not have professional counsellors at the helm to give counselling to students (Essuman, 2014). For this reason, some schools have turned teachers into counsellors. This has become a problem since such counsellors lack the professional and technical expertise to give counselling to students. In this light, we believe that our e-counselling system could complement the work of counsellors regarding their decision making of students, especially the novice counsellors.

Given the fast growing of social media platforms and easy accessibilities of the internet, we found text-based medium of information exchange as the most commonly means by which most students exchange information. In addition, Witten (2005) pointed out that the most common vehicle for formal information exchange is through text. This motivated the idea of developing a text-based NLP system for automatic emotion and sentiment analysis. To the best of our knowledge, there is no existing NLP tool for emotion analysis to complementing the work of counsellors in Ghana. This partly informed the decision of taking up this research line by employing NLP for counselling, particularly for assisting counsellors to efficiently help students to manage their emotional and personal-social life challenges.

Coupled with our preliminary findings, we organised a semi-structured interactive discussion with selected counsellors and students from three senior high schools in Ghana before the commencement of EmoTect's development, forming part of the participatory design. 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 30 students contributed to the discussion. The essence of the meeting was to outline to the participants the preliminary requirements for the commencement of EmoTect's development based on previous encounters. By so doing, we 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 1.

Table 1: Proposed requirements from our preliminary studies and investigation

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

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 emotion categories (results) for future reference. Therefore, by 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, we still explained to them the basic idea of training a classifier. We 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 commence.

3.2 EmoTect Development

Peffers *et al.* (2006) have suggested that a critical attention should be given to a desired functionality and architecture of an intended artefact while eliciting or gathering requirements for its development. For such idea, March and Smith (1995) also agreed that design science research artefact should satisfy the needs of end-users. In this work, the implementation of EmoTect began after gathering the needed requirements for its development (see Section 3.1). In addition to the implementation phase, Peffers *et al.* (2006) suggest that "resources required moving from 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**,** we deduced a three tier process development for the EmoTect: *interface (presentation layer), domain logic* and *database design*. The presentation tier was developed from HTML5 and Java server pages. The presentation tier is the point of interaction between the users and the system. 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[2] and Weka's SVM classifier[3]. Part-of-speech (POS) tagger from the Natural language toolkit (NLTK) was used for syntactical parsing the input text. In addition, lemmatisation package form NLTK was used in Lemmatising of the input text. The last and third tier is the database. We used MySQL as the database server and Apache Tomcat for the webserver. After pre-processing and extracting of 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 being displayed in HTML5 tables and visualisation charts. Corpus building, training of the classifier and the classification phases are elaborated in the subsections below.

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. By so doing, emotional antecedents in a form of stories regarding students' academic development were collected from selection of schools in Ghana.

[2] *http://www.nltk.org/*
[3] *http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SMO.html*

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. Students who participated in the study were made to sign an informed consent form 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

After that, the students' life stories were pre-processed into a suitable granularity for annotation. The stories were then given out to the three counsellors (hereafter: C_1, C_2 and C_3) to annotate emotions in them (stories) using Plutchik's eight basic emotions. The use of the Plutchik's basic emotions was motivated by our previous study (Kolog *et al.,* 2016) where a focused group discussion was organised with counsellors to extract from them the common contextualised basic emotions that students expressed in counselling. Based on the outcome we concluded on using Plutchik basic emotions. Negative and positive polarities were used for the sentiment analysis in the EmoTect system.

Two rounds of annotation exercise were carried out by the counsellors on the same corpus in two different times and days. Hence, the exercise took a lot of time to complete since the data was large (2, 200 instances of the students' life stories). After the annotation exercise, we had meeting with the participants to consider some of the disagreements in the emotion and sentiment annotation. It resulted in an improved kappa scores. Intra- and inter-counsellors' annotation agreement of emotions was computed. While obtaining *Inter-annotators*[4] agreement kappa score of 70.3% and 80.5% for the emotion and sentiment respectively, the *intra-annotation*[5] agreements from all the counsellors yielded almost perfect average kappa greater than 85% in both the emotions and sentiments.

Training of the classifier did not only follow the traditional approach of using *inter-annotators'* agreement *gold standard* data, but counsellors' perception of emotions was considered as well. By default, the EmoTect system was trained on annotated life stories after obtaining a good inter-annotation agreement kappa score. However, we have made provision for the users in EmoTect to make changes to the emotion categories based on their personalised perception of emotions. With this, counsellors are expected to annotate the stories based on their own perception of emotions before using the EmoTect system, else the default setting trained from all-in-one inter-annotated corpus is maintained. For instance, different counsellors may tag different emotion categories to the same instance of a story. Counsellors are expected to do so at the front-end of EmoTect before starting to use the system. At any point in time, counsellors are able to make changes to the emotions they had labelled should the need arises. Figure 2 is the interface for users to effect changes to the default training data where necessary.

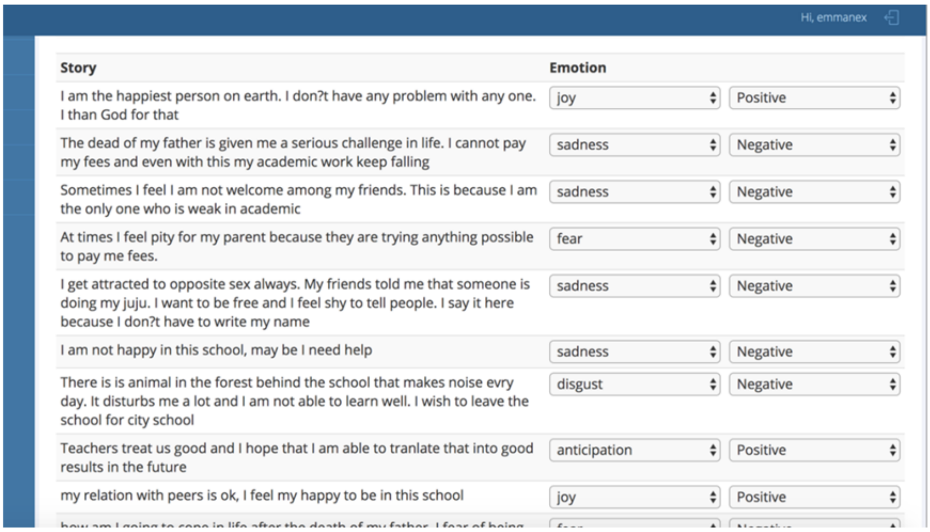

Figure 2: System training page

[4] *Inter-annotation agreement of emotions refers to the comparisons between counsellors´ annotated emotions in text corpora.*
[5] *Intra-annotation is the annotation agreement of emotions for each counsellor based on the various rounds of annotation*

Classification phase

From the EmoTect architecture in Figure 3, 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 support vector machine classifier learns from the training text and then predicts the 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 the pre-defined emotion and sentiment categories.

From the EmoTect architecture in Figure 3, the system works by first tokenise the training data (life stories). After that, the tokenised words are tagged by their parts of speech, which is accomplished by a Part-of-speech (POS) tagger from Natural language toolkit (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, 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.

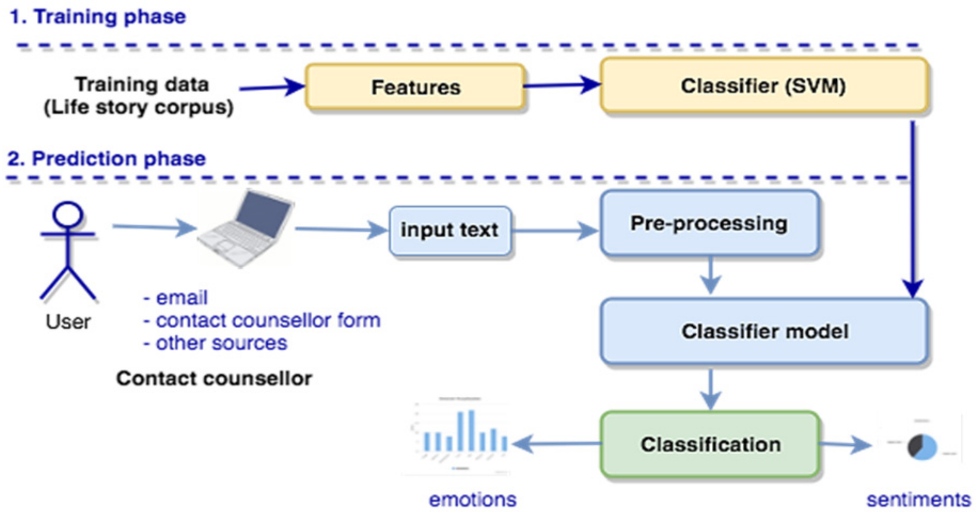

Figure 3: EmoTect Architecture

3.3 EmoTect description and user guide

In this section, we present the *Use case* model of the EmoTect accompanied with the step-by-step process of using the system to facilitate counselling delivery. Use case (UC) is a unified model language (UML) which identifies, clarifies and organises system requirements. Use case details the sequence of interactions between systems and users in a particular environment. Additionally, UC explicitly give general overview to the various stages of using an information system artefact. The term '*actors'* in UC model are referred to users of an EmoTect. The actors of EmoTect system are school counsellors, students and administrators of the system who is responsible for technically managing the system. Figure 4 represents the use case model of EmoTect showing the various actors and their interaction.

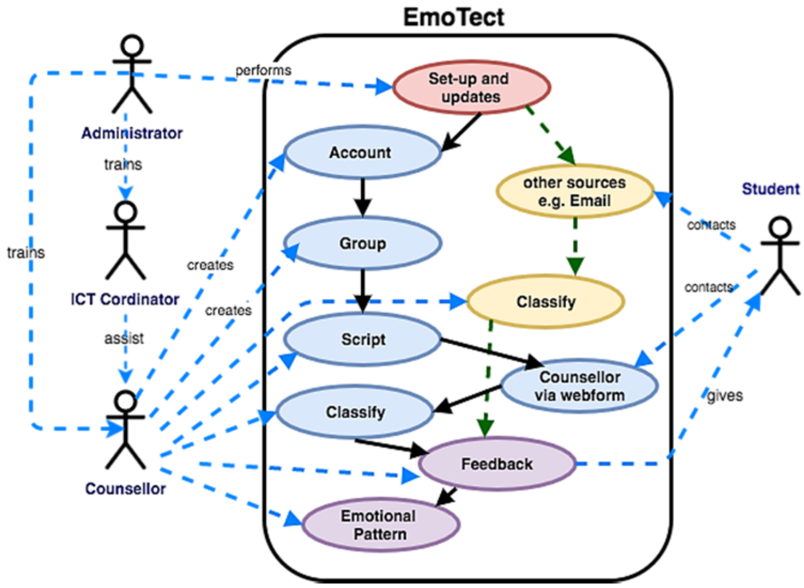

Figure 4: Use case model of the EmoTect

The various stages required to use the EmoTect system are: *Create group, view groups, emotion detection, sentiment analysis, system training* and *feedback report*. Figure 5 shows the stages as it appears in the EmoTect interface, and these are expanded in the subsequent paragraphs.

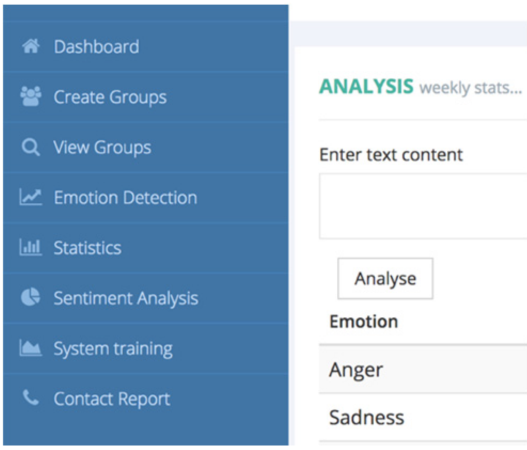

Figure 5: Navigation pane of EmoTect showing the various components

Creating access account

To use *EmoTect*, users are expected to create a user account for their respective schools by registering from the project's page (*http://nlp4counselling.com/dashboard.jsp*). Once the account is created, login credentials are obtained and can be used to access the system at any time provided there is internet connection.

Creating group

Once logged in, users are expected to create a group database for their schools. This can be done by clicking on the *"create group"* label at the navigation pane of the EmoTect main page (Figure 5). We recommend school administrators or counsellors to use the name of their schools as group names. At this point, EmoTect can be used to analyse students' textual submissions from other external sources such as email. This can be achieved by copying the content and pasting it into EmoTect for both emotion and sentiment classification.

Generating scripts

At the navigation pane of the EmoTect interface (Figure 5) is '*view group"*. The "*view group"* shows users their account database. On the same page, users are expected to generate and copy JavaScript codes to the HTML header of their school websites. As a result, a '*contact counsellor*' form appears on their respective school web pages. The widget form provides a frontend view for students to contact counsellors. As seen in Figure 6 all the entry fields are required except for the students' name field. This is in line with our previous findings where students have advocated for anonymity in counselling. The essence of the *email* field is for the counsellors to contact their students' after deciding on their textual submissions. Should the need arise, based on counsellors' discretion, counsellors can ask for face-to-face session with the students through their emails. Figure 6 shows sample text that was collected through the "*contact counsellor*" form. The text and the emails shown in Figure 7 are fictitious and do not represent any real data from the students. The following is a sample script that generates the "contact counsellor" widget form on the users' page.

```
<html>
    <head>
        <scriptsrc="http://nlp4counselling.com/assets/js/widget-
          contact.js"type="text/javascript"></script>        <script>document.createWidget
          ('2');</script>
    </head>
    <body>
     ------webpage content here -----------
    </body>
</html>
```

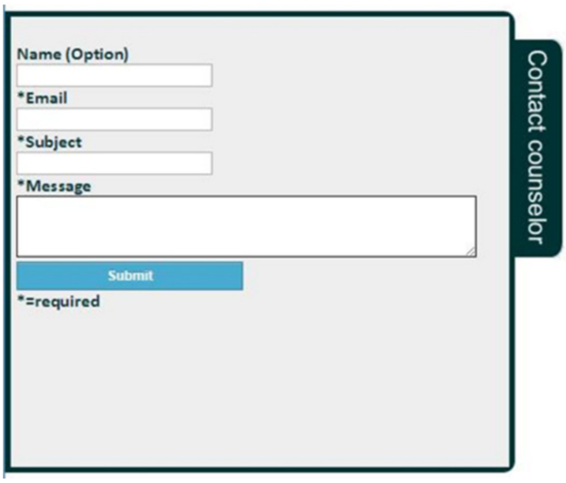

Figure 6: Contact counsellor widget form

	Subject	Message	Group	Action		
	Help	Dear Counsellor, I am the shy type hence I find it difficult to approach the teachers in my school. Is there any better way to help me out from this kind of a problem.	OKESS	2017-06-15 11:35:35.0	Extract sentiment	Extract emotion
com	Help	Dear Counsellor, I have lot of things helping me to go through life. Only that I have bad habit like stealing but my mother beat me over it. Please help me	OKESS	2017-06-15 10:59:10.0	Extract sentiment	Extract emotion
	I am shy	Dear counsellor, I am a shy type so I am not able to contact friends even if i need help for my academic work. Is it possible you could assist me to overcome my shyness. I have been yearning to master courage in that so that i can learn a lot from others. Thank you	OKESS	2017-06-08 21:18:09.0	Extract sentiment	Extract emotion
	Help	I write to inform you that I need help. I have been having challenges with my academic work and this is because of my addiction to smoking. I smoke weed everyday and I know the school would sack me if they find out. That is why I am contacting you from here, because i need help.	OKESS	2017-06-08 21:15:29.0	Extract sentiment	Extract emotion

Figure 7: Submissions from the "contact counsellor" widget form

Emotion and keywords extraction

Students' textual submissions through the '*contact counsellor*' form is analysed automatically by the EmoTect system. The results for both the emotion and sentiment are made available to the user at the feedback report section shown in Figure 5. The extracted keywords from the text are also outputted to the systems interface as well. Figure 8 shows the emotion detection and visualisation interface of EmoTect.

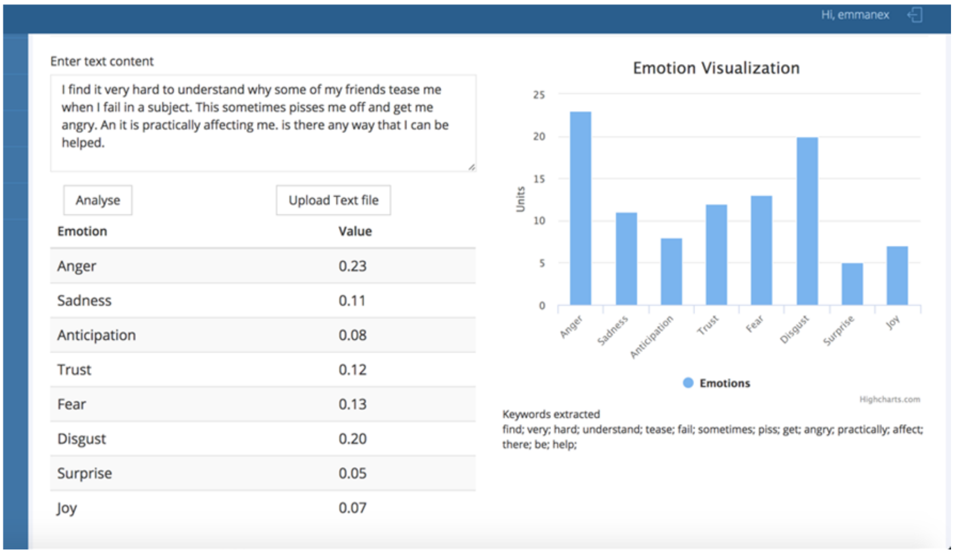

Figure 8: Front-end visualisation view of the emotion extraction part of EmoTect

The main essence of the EmoTect system is to help counsellors monitor the emotional changes of their students over a period. The automatically extracted emotions from students' submissions through the "*contact counsellor"* form can be visualised graphically for a selectable period. Figure 9 shows 4-month emotional pattern of test sample.

Figure 9: Sample view of the emotional changes in students' submission in text

Sentiment classification

The EmoTect system tracks, detects and outputs the emotional polarities (sentiment) from text. Negative and positive emotional polarities are the output. However, if there are no emotional words from input text, the system outputs the extracted keywords from the content and indicates *00%* for the polarities. Figure 10 shows the emotional polarities (negative and positive) from predicted sample students' text.

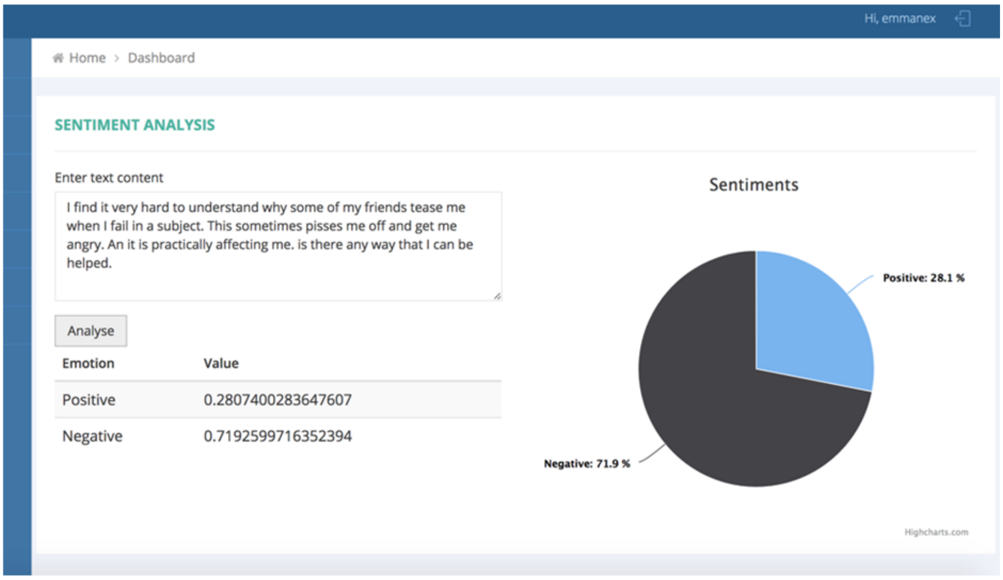

Figure 10: Visualising sentiments from students' textual submissions

There is a growing interest for e-counselling platforms, especially in the domains of health and education (King *et al.,* 2014). However, Ghana and perhaps some developing countries are still lagging behind in terms of ICT-mediated counselling (Kolog *et al.,* 2014). This may be due to the challenges associated with the high levels of illiteracy, poor internet connectivity and lack of ICT-mediated tools for counselling delivery. Regardless of these challenges, counsellors who are regarded as impeccable professionals by students, are expected to devise a strategy in helping students to fortify their motivation towards academic achievement. Consequently, advance in technology keeps widening the scope and methods on which counselling is delivered, which provides opportunities for diversification of counselling delivery.

4. Evaluation and discussion

The EmoTect classifier and the prototype evaluation are presented in this section. This is accompanied with the discussion of the results.

4.1 Classifier evaluation

Two evaluation processes were conducted in this study. First, we evaluated the SVM classifier to ascertain its efficacy by comparing with humans (i.e. counsellors) in terms of the sentiments and emotions. Secondly, the prototype version of the EmoTect was demonstrated in the study's environment with the intended end-users of the system in line with Peffer's *et al.* (2006) DSR framework.

Three hundred sixty instances of the students' life stories, representing about 16% of the total data (2,200 instances), 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, we 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. Table 3 displays test data and the output from the EmoTect (counsellors) and human prediction.

Table 3: Sample instance with prediction from counsellors and EmoTect

Instance	Emotion	Sentiment	Keywords extract
"I am the shy type hence I find it difficult to approach the teachers in my school but life in school is cool."	C_1: *Fear* C_2: *Disgust* C_3: *Disgust* EmoTect: *Fear*	C_1: *Negative* C_2: *Negative* C_3: *Negative* EmoTect: *Positive*	*Shy; difficult; approach; cool*

Instance	Emotion	Sentiment	Keywords extract
"I have lot of things helping me to go through life. Only that I have bad habit like stealing but my mother beat me over it."	C_1: *Disgust* C_2: *Sadness* C_3: *Sadness* EmoTect: *Sadness*	C_1: *Negative* C_2: *Negative* C_3: *Negative* EmoTect: *Negative*	*Have; help; only; have; bad; steal; beat*

Tables 4 and 5 are the proportion of the human annotated instances of the test data and the output from EmoTect (for the sentiment and emotion categories). Since the counsellors performed two rounds of annotation the tables are a single representation of the agreements of both rounds of annotation for each counsellor. The annotated training data obtained from the counsellors were, at different points, used to train the EmoTect classifier. The unlabelled test data was fed into EmoTect system for both emotion and sentiment classification (prediction). From our observation, whatever the difference in the training data, the same results were obtained after running the test data repeatedly with EmoTect. This may be so because there was no significant difference in the intra-annotation agreement kappa score for all the counsellors. In the end, the output from the EmoTect classifier was then compared with the *gold standards* corpus from each of the counsellors. We employed three standard evaluation matrices: *Recall, Precision and F-measure*. 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).

Table 4: Results from EmoTect's (sentiment) classifier with the gold standards

Sentiment	**EmoTect**	**C_1**	**Agreement (C_1 vs EmoTect)**	**C_2**	**Agreement (C_2 vs EmoTect)**	**C_3**	**Agreement (C_3 vs EmoTect)**
1. Positive	**130**	125	111	105	96	106	98
2. Negative	**201**	205	170	200	181	196	181

Table 5: Results from the EmoTect (emotion) classifier with the gold standards

Emotions	**EmoTect**	**C_1**	**Agreement (C_1 vs EmoTect)**	**C_2**	**Agreement (C_2 vs EmoTect)**	**C_3**	**Agreement (C_3 vs EmoTect)**
1. Anger	**7**	14	5	5	3	12	4
2. Happiness	**28**	51	26	35	20	62	14
3. Disgust	**26**	29	8	16	12	33	18
4. Sadness	**110**	116	68	170	76	97	66
5. Trust	**54**	65	26	40	22	61	38
6. Fear	**40**	38	20	22	16	23	17
7. Surprise	**9**	17	7	5	3	4	2
8. Anticipation	**35**	14	4	19	15	30	19

Classifier performance

Results from evaluating the EmoTect classifier and the prototype are reported and discussed in this section. We examined the performance of the EmoTect algorithm by comparing with the gold standards obtained from each of the three counsellors. We report the recall, precision and f-measure per each category of the emotions. In addition, the overall score of the recall, precision and f-measure for the emotion are reported. Whereas the precision takes into account the proportion of the retrieved data that was correctly labelled, we took a critical interest in the recall scores, which measured the instances of the baseline (gold standard) that was identified and extracted by the EmoTect from the extracted data. The test data was repeatedly ran for five times with the EmoTect algorithm. Upon repeated exercise, the same results were obtained. Our results from the entire evaluation of the classification algorithm is shown in Tables 6.

Table 6: Evaluation results of EmoTect classifier (Emotions)

	C_1			C_2			C_3		
Emotions	Recall (%)	Precision (%)	F-measure	Recall (%)	Precision (%)	F-measure	Recall (%)	Precision (%)	F-measure (%)
Anger	35.7	71.4	47.6	60.0	42.9	50.0	33.3	57.1	42.1
Happiness	92.9	51.0	65.8	75.0	46.2	57.1	22.6	50.0	31.1
Disgust	30.8	27.6	29.0	57.1	71.4	63.5	54.5	69.2	61.0
Sadness	61.8	58.6	60.2	44.7	69.1	68.6	68.0	60.0	61.0
Trust	48.1	40.1	43.7	55.0	40.7	46.8	62.3	70.4	66.6
Fear	50.0	52.6	51.3	72.7	40.0	51.6	73.9	42.5	54.0
Surprise	41.2	77.8	53.8	60.0	33.3	42.9	50.0	22.2	30.8
Anticipation	11.5	25.0	15.7	78.9	42.9	55.6	63.3	54.2	58.5
Overall	53.1	47.4	50.1	53.5	54.0	53.8	55.3	57.6	56.4

As indicated in Table 6, *happiness* (92.9%) and *sadness* (61.8%) yielded the highest recall with C_1. This implies the proportion of the labelled instances of the gold standard from C_1 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 C_1 gold standard, was found in *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 $C_{2,}$ shown in Table 6, 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 C_2 gold standard. *Surprise* had the highest recall (60%) at the cost of poor precision (33.3%). With the classifier performance with the C_2 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 C_3, 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 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 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 C_1 to C_3 for precision, recall and the F-measure, where: *recall:* C_3 *(55.3%)* > C_2 *(53.5%)* > C_1 *(53.1%), precision:* C_3 *(57.6%)* > C_2 *(54.0%)* > C_1 *(47.4%)* and *F-measure:* C_3 *(56.4%)* > C_2 *(53.8%)* > C_1 *(50.1%)*. With this observation, based on the methodology used in this dissertation, no particular reason could be attributed to the increasing trend from C_1 to C_3 other than the variations in the scores for each counsellor.

Though these findings came as no surprise, our 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, we believe that 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.

The performance of the EmoTect classifier was examined with regard 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. Table 7 depicts 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 standards, even when presented with unknown data. That is, EmoTect performed well when presented with unclassified data for prediction.

In the nutshell, the EmoTect algorithm for the sentiment detection achieved comparable accuracy to that achieved with a gold standard when presented with unknown data. However, the performance of classification algorithm for the emotion detection component was not as efficient as we expected. With this revelation, we deduced that more training data is required to train the algorithm thereby improving the level of accuracy in the prediction of emotions. That notwithstanding, we believe that EmoTect algorithm is capable for complementing the work of counsellors in the educational arena.

Table 7: Evaluation results of EmoTect classifier (Sentiments)

	C_1			C_2			C_3		
Sentiment	**Recall (%)**	**Precision (%)**	**F-measure**	**Recall (%)**	**Precision (%)**	**F-measure**	**Recall (%)**	**Precision (%)**	**F-measure**
Positive	82.9	84.6	87.1	90.5	90.0	81.7	92.3	90.0	83.1
Negative	88.8	85.4	83.7	91.4	73.8	90.3	92.5	75.4	91.2
Overall	**85.0**	**84.9**	**85.0**	**90.8**	**83.6**	**87.1**	**92.4**	**84.0**	**88.2**

The evaluation techniques used in this study has demonstrated that EmoTect can predict sentiments in text as accurate as human while the prediction of the emotions still requires more data for higher predictable accuracy. With this revelation, we deduced that more training data will be required to train the classifier, thereby improving the level of prediction against humans. The method of evaluation used in this study is a widely used technique, where its success largely depends on the quality of the training data and the annotation strategies. The implication is that if the data is poorly annotated there is a higher tendency for poor prediction. For this reason, it is imperative to give annotators a lot of time and training to carry out the annotation task. With this evaluation technique, EmoTect would have predicted poorly if the training data was poorly annotated with the emotions and sentiments. This is the reason why three different counsellors were trained to annotate the data for training the classifier. The challenge remains that since emotions are a subjective phenomenon it is difficult to actually ascertain a definite accuracy in terms of the human annotation. While in this study, we believe that the annotation strategies worked, we intend to compare the results from this study to that of the existing and similar tools such as WEKA. This is another level of evaluation to further ascertain the efficacy of the algorithm. From our experience, this evaluation technique should be conducted in ample time periods depending on the size of the data for annotation. Annotators should also be given enough training on the annotation exercise.

4.2 Prototype evaluation

Kies *et al.* (1998) proposed three iterative stages of evaluating information system artefact. The stages are *initial design, prototype,* and *final design evaluations.* At this stage of the implementation, we demonstrated the prototype version of the EmoTect to the end-users in their settings (schools). Students and counsellors were selected from three senior high schools to participate in the exercise. The various functionalities of the EmoTect prototype were explained to the participants. Altogether, six counsellors and thirty students participated in the evaluation exercise. The purpose, as part of the DSR framework, was to find from the participants if the EmoTect prototype met the requirements suggested in Table 2. The method of data collection was through observation and interview. This was conducted while the demonstration was ongoing. During the data collection, we were interested in the participants' perception of the usability and functional components of the EmoTect. All in all, this stage of evaluation is preliminary, since we intend to deploy the final version of the system in the environment and evaluate the system further in the future. For this reason, the data from the interview and the observations are qualitatively reported in this study.

Findings from Prototype evaluation

Given the capacity of humans to introspect, 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 that influence people's decision of adopting an Information Technology tool. We showed and demonstrated the EmoTect prototype to the participants as Hevner *et al.* (2007) advocated for rigorous DSR artefact evaluation. The demonstration was done by comparing the system with the prior requirements of EmoTect development (see Section 3.1).

A majority (90%) of the participants agreed that the EmoTect system had captured all the requirements they had proposed in the early stages of the work. To this end, the system was generally seen by the participants as a tool suitable for complementing their work in delivery of counselling to students. During the demonstration, we made some observations that is worth considering in the next phase of EmoTect development. One of the counsellors said and I quote:

> *"The software is good, because many students will be offered opportunity to contact counsellors. However, my concern is the unavailability of internet connection in my office. Usually if I have to do anything I move from my office to the ICT laboratory to use the internet. In this case, it will be difficult for us to use the software. Apart from that I think the software is good and it almost predicted exactly what I had predicted already."*

We found that all the participants were quite enthused about the capabilities of the EmoTect system. This was to be expected as there are no existing NLP tools that have been developed contextually in Ghana, with the capability of analysing emotions and sentiments in text. This could also be attributed to the doubt that people have over computational tracking of emotions in text. One of the counsellors said and I quote:

> *"This is amazing! Though I don't want to understand the technicalities of the process but I doubted from the onset how ICT can be used to determine peoples' emotions. I was particularly expecting to see how the intended platform would do. I really believe that ICT can do a lot of things that goes beyond the comprehension of our imagination."*

We demonstrated to the participants the novelty of our work regarding the training of the classifier. The counsellors saw the relevance in our approach but we could not evaluate their perception regarding the system's classifier training technique. This was because the counsellors did not really have any idea about computational analysis of emotions in text. Regarding the ease of use of the system, the participants did not raise a critical concern. The only aspect that got few of the participants taking was the part where *JavaScript* is expected to be copied from the EmoTect page to their webpages. In effect, the participants raised some concerns about the technicalities of copying the scripts from the EmoTect page to their respective pages in order to display the personalised "*contact counsellor*" widget form. Nonetheless, while majority (60%) of the participants had no difficulties configuring the EmoTect system before using it for the first time, some of them (40%) encountered configuration difficulties. Those who had configuration challenges are those who complained of the copying of the JavaScript code for embedding on their page. A counsellor expressed this views on the subject during the interview and we quote:

> *"I don't have problem with the system being used to help us, but my concern is the part that require us to do some programming in order to get the message box for our students to contact us. Since you said this is just a demonstration, I don't have much to say. However, at some point the systems output was not what I expected. Also take a good look at why the response at some point where slow."*

During the discussion, most of the participants did not find any difficulties with the interface, and as well as the aesthetic features of the EmoTect. All the participants found it easy to traverse the various pages of EmoTect. By explaining to the participants, we recommend for them to use the services of their ICT coordinators to help in the initial configuration settings. While demonstrating the system to the participants, minor technical glitches were found. This was to be expected since it is a prototype version. However, the glitches did not affect the output in terms of the text processing. All in all, the majority of the participants expressed satisfaction with the visualisation output.

In the nutshell, the responses from the participants were valuable as we intend to use their feed for the next phase of the EmoTect development. In fact, the technical glitches that were discovered during the demonstration have been considered. As part of the developmental process, we shall return to the participants to test the final version of EmoTect by allowing them to use for a while. This is, particularly, important in information system artefact development. EmoTect, in the future, shall fully be deployed in the environment for counsellors to use.

5. Conclusion

In this work, we have presented a contextualised intelligent e-counselling system for automatic emotion and sentiment detection in text. The system is implemented from a natural language processing using supervised machine learning technique. Therefore, annotated *life story corpus* was developed from students' stories and used as training data. The training of the classifier was based on the personalised emotion perception of counsellors where counsellors are given the opportunity to tag the training data with Plutchik basic emotions. The development of the system duly followed Peffer's design science research (DSR) framework. The system allows students to contact a counsellor through a personalised web form. The aim of the system is to complement the work of counsellors, thereby helping to automatically analyse emotions and sentiments of students' textual submissions to inform decision making. Changes in students' emotions as predicted by the EmoTect can be visualised over a period of time. As per the requirements of DSR and participatory design process, we evaluated the platform with gold standard data from counsellors. Results show that EmoTect algorithm for the sentiment detection achieved comparable accuracy to that achieved with a gold standard when presented with unknown data. However, the performance of classification algorithm for the emotion detection component was not as efficient as we expected. In addition, counsellors and students were enthused about the EmoTect system and expressed their satisfaction about its usability and functionalities in line with the intended purpose.

References

Alexandria, V., 2014. Mindsets and Behaviors for Student Success: K-12 College- and Career-Readiness Standards for Every Student. *American School Counselor Association* .

Alm, C., Roth, D. and Sproat, R., 2005. Emotions from text: machine learning for text-based emotion prediction. *In procedings of the conference on Human Language Technology and empirical methods in Natural Language Processing* (pp. 579-586). Association of Computational Linguistics.

Aman, S. and Szpakowicz, S., 2007. Identifying expressions of emotion in text. *In procedings of International Conference on Text, Speech and Dialogue*, pp. 196-205.

Archer, L. (1984). *Systematic method for designers.* (N. Cross, Ed.) London: John Wiley: Developments in design methodology.

Au, Y., 2001. Design Science: The Role of Design Science in Electronic Commerce Research. *Communications of the Association for Information Systems, 7*.

Awinsong, M., Dawson, O. and Gidiglo, B., 2015. Students' Perception of the role of Counsellors in the Choice of a Careeer: Astudy of the Mfatsipim Municipality in Ghana. *International journal of Learning and education research, 13*(3), 79-99.

Awinsong, M., Dawson, O. and Gidiglo, B. E., 2015. Students' Perception of the role of Counsellors in the Choice of a Careeer: Astudy of the Mfatsipim Municipality in Ghana. *International journal of Learning and education research, 13*(3), pp.79-99.

Barna, J. and Brott, P. E., 2011. How Important Is Personal/ Social Development to Academic Achievement? The Elementary School Counselor's Perspective. *Professional School Counseling journal, 14*(3), 242+.

Baumeister, R. F., DeWall, C. N., Vohs, K. D., and Alquist, J. L., 2010. Does emotion cause behavior (apart from making people do stupid, destructive things). *Then a miracle occurs: Focusing on behavior in social psychological theory and research*, pp. 12-27.

Bhowmwick , P.K, Mitra, P. and Basu, A., 2008. An agreement measure for determining Inter-annotator Reliability of Human Judges on Affective Text. *In proceedings of the workshop on human Judgements in Computational Linguistics.*, (pp. 58-65). Manchester.

Buitinck, L., Amerongen, J., E., T. and De Rijke, M., 2015. Multi-emotion detection in user-generated reviews. *In European Conference on Information Retrieval.* Springer International Publishing.

Chopade, C. (2015). Text based emotion recognition: A survey. *International journal of science and research , 4*(6), pp. 409-414.

Crowston, K., Liu, X., and Allen, E. E., 2010. Machine learning and rule-based automated coding of qualitative data. *In proceedings of the American Society for Information Science and Technology*, *47*, pp. 1-2.

Duveskog, M., Kemppainen, K., Bednarik, R., and Sutinen, E., 2009. Designing a story-based platform for HIV and AIDS counseling with Tanzanian children. *In Proceedings of the 8th International Conference on Interaction Design and Children* (pp. 27-35). ACM.

Eekels, J. and Roozenburg, N., 1991. A methodological comparison of the structures of scientific research and engineering design: their similarities and differences. . *Design Studies, 12*(4), pp. 197-203.

Fulcher, A.and Hills, P., 1996. Towards a strategic framework for design research. *Journal of Engineering Design, 7*(1), pp. 183-193.

Hancock, J., Landrigan, C. and Silver, C., 2007. Expressing emotion in text-based communication. *In procedingsof the SIGCHI conference on human factors in computing system systems* (pp. 929 - 932). ACM.

Hassan, S., Rafi, M. and Shaikh, M., 2011. Comparing SVM and naive bayes classifiers for text categorization with Wikitology as knowledge enrichment. *In Proceedings of Multitopic Conference (INMIC). 14*, pp. 31-34. IEEE.

Hevner, A., 2007. A three-cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(22), 87–92.

Hevner, A., March, S., Park, J. and Ram, S., 2004. Design Science in Information Systems Research. . *MIS Quarterly, 28*, 75-105.

Jenkins, P. and Palmer, P., 2012. At risk of harm‟? An exploratory survey of school counsellors in the UK, their perceptions of confidentiality, information sharing and risk management. *British Journal of Guidance & Counselling, 40*(5), pp. 545-559.

Joachims, T., 1998. Text categorization with support vector machines: Learning with many relevant features. *In European conference on machine learning* (pp. 137-142). Berlin Heidelberg: Springer.

Kolog , E. A., Montero, C. S. and Sutinen, E., 2016. Annotation Agreement of Emotions in Text: The Influence of Counselors' Emotional State on their Emotion Perception. *Proceedings of 16th IEEE International conference on Advanced Learning Technologies* (pp. 357-359.). Austin: IEEE.

Kolog, E. A., 2017. Contextuallising the application of human language technologies for counselling. *PhD thesis*, *University of Eastern Finland,*

Kolog, E. A., 2014. E-counselling implementation: Contextualized approach. Masters Thesis, publication of the *University of Eastern Finland, Finland*.

Kolog, E. A. & Montero, C. S., 2017. Towards automated e-counselling system based on counsellors emotion perception. *Education and Information Technologies*, pp. 1-23, Springer.

Kolog, E. A., Sutinen, E. and Vanhalakka-Ruoho, M., 2014. E-counselling implementation: Students' Life stories and counselling technologies in perspective. *International Journal of Education and Development using Information and Communication Technology, 10*(3), pp. 32-48.

Kolog, E. A., Sutinen, E. and Vanhalakka-Ruoho, M., 2015. Using Unified Theory of Acceptance and Use of Technology Model to Predict Students' Behavioral Intention to Adopt and Use E - Counseling in Ghana. *International Journal of Modern Education and Computer Science, 7*(11), 1-11.

Le Surf, L. & Leech, A., 1999. Exploring young people‟s perceptions relevant to counselling: a qualitative study . *British Journal of Guidance and Counselling, 27*, 152.

Lu, C.-Y., Hong, J.S. & Cruz-Lara, S., 2006. Emotion Detection in Textual Information by Semantic Role Labeling and Web Mining Techniques. *Third Taiwanese-French Conference on Information Technology – TFIT.* Nancy-France.

March , S. and Smith, G., 1995. Design and natural science research on information technology. *Decision support systems, 15*(4), 251-266.

Matwin, S. and Sazonova, V., 2012. Direct comparison between support vector machine and multinomial naive Bayes algorithms for medical abstract classification. *Journal of the American Medical Informatics Association, 19*(5), 917-917.

McKay, J. and Marshall, P., 2005. A review of design science in Information systems. In Proceedings of the 16th Australasian Conference on Information Systems. Sydney, Australia.

Mitchell, D. (2011). The nexus between decision-making and emotion regulation: a review of convergent neurocognitive substrates. *Behaviour brain research, 217*, pp. 215–231.

Mramba, N., Apiola, M., Kolog, E. A. and Sutinen, E., 2016. Technology for street traders in Tanzania: A design science research approach. *African Journal of Science, Technology, Innovation and Development, 8*(1), pp. 121-133.

Munezero , M., Montero, C., Sutinen, E. and Pajunen, J., 2014. Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text. *5*(2), pp. 111.

Mushaandja, J., Haihambo, C., Vergnani, T. and Frank, E., 2013. Major Challenges Facing Teacher Counselor s in Schools in Namibia. *Education journal*, *2*(3), 77-84.

Mushaandja, J., Haihambo, C., Vergnani, T. and Frank , E. 2013. Major Challenges Facing Teacher Counselors in Schools in Namibia. *Education Journal, 2*(3), 77-84.

Newman, M. L., Groom , C. J. and Handelman, L. D., 2008. Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, *45*, pp. 211-236.

Nyirenda-Jere , T. and Biru., T., 2015. Internet development and Internet governance in Africa. *Internet Society*, 1-43.

Paulus, M. and Yu. , A., 2012. Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in cognitive sciences , 16*(9), 476-483.

Peffers, K., Tuunanen, T., Gengler, C. E., Rossi, M., Hui, W., Virtanen, V. and Bragge, J., 2006. The design science research process: a model for producing and presenting information systems research. *In Proceedings of the first international conference on design science research in information systems and technology*, (pp. 83-106).

Pekrun, R., Goetz, T., Titz, W. and Perry, R. P., 2002. Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. *Educational Psychologist, 37*(2), 91 – 106.

Plutchik, R.,1980. Emotion: Theory, research, and experience: Theories of emotion. *New York: Academic, 1*, 399.

Porter, M.F., 1980. An algorithm for suffix stripping. *Program, 14*(3), 130-137.

Preston, M. and Mehandjiev, N. (2004). A Framework for Classifying Intelligent Design Theories. *The 2004 ACM workshop on Interdisciplinary software engineering research, Newport Beach, California:* (pp. pp. 49-54.). ACM.

Rahman, M. A. and Ahmed, A. M., 2014. Emotion-Based System for Social Media Content Processing and Event Monitoring. *Advances in Computing, 14*(2), pp. 31-40.

Robertson, T. and I., W., 2012. Ethics: Engagement, representation and politics-in-action. *In Routledge International Handbook of Participatory Design*, pp. 64–81.

Shelke, M., 2014. Approaches of emotion detection from text. *International Journal of Computer Science and Information Technology, 2*, pp. 123 - 128.

Shina, D.-H. and Choib, M., 2014. Ecological views of big data: Perspective and issues. *Telematics and informatics*, 311-320.

Shivhare, S. N. and Saritha, K., 2012. Emotion detection from text. arXiv preprint arXiv:1205.4944.

Simon, H., 1996. *The Sciences of Artificial* (Vol. 3rd edition). , Cambridge, MA.: MIT Press.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M., 2011. Lexicon-based methods for sentiment analysis. *Computational linguistics, 37*(2), 267-307.

Takeda, H., Veerkamp, P., Tomiyama, T. and Yoshikawam, H.,1990. Modeling Design Processes. *AI Magazine*, pp. 37-48.
Tobin, K., King, D., HEnderson, S., Bellocchi, A. and Ritchie, S., 2016. Expressing of emotions and physiological changes during teaching. *Cultural studies of science education, 11*, pp. 669-692.
Vinluan, L. V., 2010. The use of ICT in school guidance: Attitudes and Practices of Guidance in Metro Manila, the Philipines. *International journal of Advanced counselling, 33*, 22-36.
Witten, I., 2005. Text mining. *Practical handbook of Internet computing*, pp. 14-21.
Zmud, R., 1997. Editor's comments. *MIS Quarterly, 21*(2), xxi–xxii.