Predicting Performance Problems Through Emotional Analysis

Ricardo Martins

Centro Algoritmi / Departamento de Informática
Universidade do Minho, Campus de Gualtar, Braga, Portugal
ricardo.martins@algoritmi.uminho.pt
https://orcid.org/0000-0003-1993-5343

José João Almeida

Centro Algoritmi / Departamento de Informática
Universidade do Minho, Campus de Gualtar, Braga, Portugal
jj@di.uminho.pt
https://orcid.org/0000-0002-0722-2031

Pedro Rangel Henriques

Centro Algoritmi / Departamento de Informática
Universidade do Minho, Campus de Gualtar, Braga, Portugal
prh@di.uminho.pt
https://orcid.org/0000-0002-3208-0207

Paulo Novais

Centro Algoritmi / Departamento de Informática
Universidade do Minho, Campus de Gualtar, Braga, Portugal
pjon@di.uminho.pt
https://orcid.org/0000-0002-3549-0754

Abstract

In the cartoons, every time a character is nervous he/she begins to count to ten to keep calm. This is a technique, among hundreds, that helps to control the emotional state. However, what would be the impact if the emotions would not be controlled? Are the emotions important in terms of impairing the ability to perform tasks correctly?

Using a case study of typing text, this paper is about a process to predict the number of writing errors from a person based on the emotional state and some characteristics of the writing process. Using preprocessing techniques, lexicon-based approaches and machine learning, we achieved a percentage of 80% of correct values, when considering the emotional profile on the writing style.

2012 ACM Subject Classification Computing methodologies → Natural language processing

Keywords and phrases emotion analysis, machine learning, natural processing language

Digital Object Identifier 10.4230/OASIcs.SLATE.2018.19

Category Short Paper

Funding This work has been supported by COMPETE: POCI-01-0145-FEDER-0070 43 and FCT – Fundação para a Ciência e Tecnologia – within the Project Scope UID/CEC/ 00319/2013.

7th Symposium on Languages, Applications and Technologies (SLATE 2018).
Editors: Pedro Rangel Henriques, José Paulo Leal, António Leitão, and Xavier Gómez Guinovart
Article No. 19; pp. 19:1–19:9
OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

In any kind of relationship, a golden rule to avoid problems is not taking decisions under emotional pressure. There are several strategies to do this: from counting to ten before responding an unpolished message until taking a break to "refresh the mind" before the decision.

However, there are situations, such as tests, where it is impossible to avoid emotional pressure and its consequences. When in a stressful situation, or under strong emotional conditions, people tend to make mistakes more frequently. This situation happens in any profession, and so being able to predict these errors that are consequences of emotional states is an important approach to plan a strategy to decrease or avoid them. For example, how important would it be for transportation companies to know the drivers' emotional state before travelling, aiming to reduce the risk of accidents? How important would be for a hospital to predict the errors of a doctor, based on his emotions?

Errors differ from profession to profession and also the effect of emotions over the work is different. Different data sets must be collected to identify these errors, and to correlate them with the emotional state of the worker.

As writers usually express their emotions in the texts they produce through the bag of words they use in each situations, and the typing errors they do along an editing session ca be measured, we intend to model the relation between emotions and errors, using the computer as a case of study. The purpose of the study here reported is to analyse a big collection of texts annotated with editing data to demonstrate that the relation between errors and emotions can be identified, and quantified in order to predict undesired situations.

In this paper, we present an approach using Sentiment Analysis and Machine Learning to characterize the impact of emotions in the number of errors during a typing process. After training the model, it will be used to predict new cases in order to assess it.

It is not our intention to claim that this approach is an alternative for predicting errors in all situations, however, we think that the approach will lead further future investigation in this relevant topic.

The remainder of the paper is as follows: Section 2 introduces the concept of basic emotions, which is a well-known theory used in Sentiment Analysis. Section 3 discusses related work concerned with the detection of emotions from a text that inspired this work, while Section 4 describes the creation of a dataset for our tests. Section 5 presents our proposal, explains the steps used in the analysis, discusses the results obtained from a set of tests performed. The paper ends in Section 6 with the conclusions and future work.

2 Theory of Basic Emotions

Basic emotion theorists explain that every human emotion is composed of a set of discrete basic emotions [2, 4, 11].

Many researchers have identified some basic universal emotions. One of the first attempts is a study by Paul Ekman [3] which concluded that there are six basic emotions are *Dislike*, *Happiness*, *Sadness*, *Anger*, *Fear* and *Surprise*. His work is based on the theory that human faces can represent this basic emotion as universal pictures.

For Plutchik [11], every sentiment is composed of a set of 8 basic emotions: *Anger*, *Anticipation*, *Disgust*, *Fear*, *Joy*, *Sadness*, *Surprise* and *Trust*, represented as a "wheel of emotions". Furthermore, the combination of basic emotions results in *dyads*. Plutchik created rules for building the *dyads*, defining the primary dyads emotions as the sum of two adjacent basic emotions, as $Optimism = Anticipation + Joy$. Meanwhile, secondary dyads emotions are composed of emotions that are one step apart on the “emotion wheel”, as *Unbelief* = *Surprise* + *Disgust*. The tertiary emotions are generated from emotions that are two steps apart on the wheel, as *Outrage* = *Surprise* + *Anger*.

Other well-known emotions model is the Five Factor Model (as known as Big Five), introduced by McCrae [8] which suggests that the personality is composed of 5 independent factors:

Openness to experience: People with high scores like news and tend to be creative. At the other end of the scale are the conventional and orderly, those who like the routine and have a keen sense of right and wrong;

Conscientiousness: It measures the level of concentration. Those with high scores are highly motivated, disciplined, committed and trustworthy. Those with low results are undisciplined and easily distracted;

Extroversion: It measures the sense of well-being, the level of energy, and the ability in interpersonal relationships. High scores mean affability, sociability, and ability to impose oneself. Lows indicate introversion, reservation, and submission;

Agreeableness: It refers to how we relate to others. Many points indicate a compassionate, friendly and warm person. At the other end are the withdrawn, critical and egocentric;

Neuroticism: It measures emotional instability. People with high scores on this scale are anxious, inhibited, melancholic and have low self-esteem. Those that get low scores are easy to deal with, optimistic and well-liked with themselves.

In this paper, we adopt the Plutchik’s model to represent emotions because we consider more realistic, easy to use and this model allows to represent several different emotions through dyads emotion. Moreover, there are some libraries and lexicons used in this work which represent and process the emotions according to this model.

3 Related work

There are several works using sentiment analysis for diversified objectives, and some are more relevant for the present paper as they have been used as a source for the idea proposed. In this section, we survey these inspiring works. However, predicting typing errors from an emotional analysis is an unexplored field and we have not found specific previous works to reference.

The usage of emotional labels for predictions was inspired by the work of Martins et al. [7], which uses emotional labels to improve the authorship identification. This is made using Facebook posts from personalities known and a hybrid approach containing lexicon and machine learning approaches.

Moreover, Thewall et al. [12] have presented a work to extract sentiment strength from the informal English text, using new methods to exploit the de facto grammars and spelling styles of cyberspace, which contributed with the idea of extract sentiment polarities from text.

Finally, the work presented by Meier et al. [9] contributed to the idea of predicting writing performance using affective variables to relate to efficacy expectations.

4 Data creation

In order to analyse the impact of the emotions during the text writing process, it was necessary to analyse texts containing emotional load and meta-information about its creation. For this purpose, the dataset provided by Banerjee et al [1] containing keystroke logs for

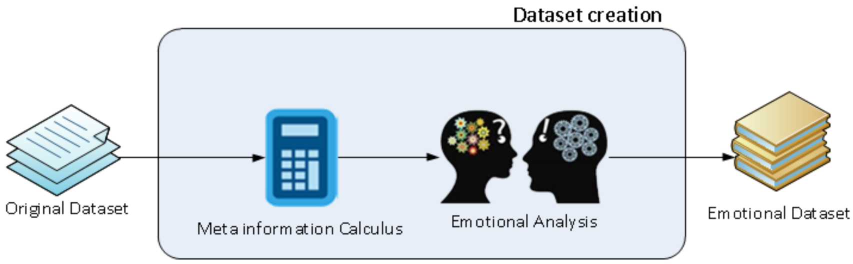

Figure 1 Dataset creation process.

opinion texts about gun control[1] was used as the basis for a new dataset creation. To perform the intended analysis we actually needed a text repository with emotional load and editing numerical data for each written piece. The option of a keystroke log is justified by the necessity of gather information about the text creation process. So, in our study all texts are considered written "from beginning to end", i.e., the first typing step without a posterior text revision phase. This is important for levelling possible errors and editing in a same identifiable pattern.

The process of the new datasets creation is performed in 2 steps: Meta-information Creation and Emotional Analysis, as presented in Figure 1.

4.1 Meta-information

The Meta-information Creation step analyses the keystroke log and calculates metrics about the text creation. For analysis purposes, we defined some metrics considered important in this study. These metrics are:

- **TimeText:** It represents the time spent during the text writing. It is the amount of time span in milliseconds between press and release for each character and white space key in the keyboard. Punctuations, numbers, and others are discarded;
- **AveragePerWord:** Is the average between the total words in the text and the time spent during typing process;
- **AmountErrors:** It is the number of errors during the typing process. It is important to emphasize that due to dataset limitations that do not store mouse movements or selections, it is impossible to detect all forms of removing characters (for example, single character or block removing). For convenience of this study, it is considered an *error* each *backspace occurrence*[2];
- **AverageErrors:** It is the average of the total words in the text and the number of errors during typing process;
- **TimeBetweenKeys:** It is the average time span between the keyboard press;
- **RepeatedCharacterFrequence:** It is the frequency of a character is repeated in the text. A repeated character is considered the same character those have been pressed immediately before and the time between them is at least 15% lower than TimeBetweenKeys. This is important to detect situations when a key is pressed for a long time, repeating the character.

[1] This is a hot, sensible, topic provoking emotive reactions on commenters.
[2] We are aware that counting backspaces is not the more adequate way to count errors, because other reasons can lead the writer to backspace and delete characters, and also many errors are made without being detected and corrected. Anyway we feel that there is a clear relation between both.

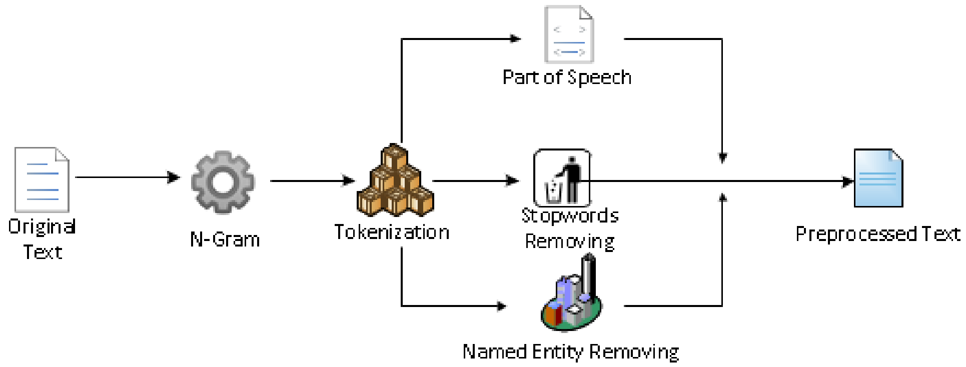

Figure 2 Preprocessing pipeline.

4.2 Emotional Analysis

The Emotional Analysis step is responsible for identification of each basic emotion according to Plutchik's model [11]. This model was chosen because there are many libraries to process information according to it and lexicons which contain the basic emotions for each word. To achieve this objective, all sentences are analysed using the EmoLex [10] lexicon.

For this analysis, all texts have been submitted to a preprocessing pipeline; at the end of this phase, only the relevant information remained.

This pipeline was composed of n-Gram identification, tokenization, stopwords removal, part of speech tagging and named entity removal, as presented in Figure 2.

Using the Stanford Core NLP toolkit [6] for these tasks, the preprocessing is divided into 3 parallel tasks. This is important because both Part of Speech Tagging and Named Entity Recognition need the text in the original format in order to identify the information.

The preprocessing begins with the N-Gram identification, where a predefined set of n-grams are identified in the text and labelled to be interpreted as a single word. Later, the tokenizer splits the text in a list of words (tokens) and these tokens are syntactically analysed in Part of Speech, where the nouns, verbs, adverbs, and adjectives are identified and stored for future purposes. In parallel, the tokens identified in a predefined stopwords list are removed and the tokens in named entity process are analysed in order to identify names (persons, locations or organizations) and discard them.

Later, the common tokens in these processes are stored and the emotions from each preprocessed text are identified through a process in R which queries the NRC lexicon [10] and identifies the basic emotions.

Finally, all information is stored in a new dataset containing the opinion and preprocessed text (from the original dataset), the 6 metrics created in Meta-Information Creation step, 8 basic emotions percentages and 2 polarities identified in the Emotional Analysis step.

5 Data analysis

The objective of this analysis is to find some evidence that emotions influence the writing process. In order to achieve this objective, some experiments were performed to relate emotions and writing patterns.

Table 1 Correlations between *AmountErrors* and emotions.

Emotion	r^2	Emotion	r^2	Emotion	r^2	Emotion	r^2
Anger	0.35	Optimism	0.30	Pessimism	0.42	Trust	0.36
Anticipation	0.31	Hope	0.39	Awe	0.38	Curiosity	0.37
Disgust	0.26	Anxiety	0.52	Despair	0.38	Pride	0.38
Fear	0.38	Love	0.35	Shame	0.38	Surprise	0.19
Joy	0.23	Guilt	0.41	Disapproval	0.30	Remorse	0.31
Sadness	0.30	Delight	0.25	Unbelief	0.27	Outrage	0.34

5.1 Emotional correlations

As initial step, the values for some Plutchik's basic emotions and defined dyad emotions[11] were calculated in order to provide more sources of information to analyse the data. To calculate these emotions, we used the package Syuzhet [5] in R, which analyses the text provided and returns the values of each basic emotion contained into the text, according to the NRC lexicon [10]. The dyad emotions were calculated according to the formula below:

- Optimism = Anticipation + Joy;
- Disapproval = Surprise + Sadness;
- Hope = Anticipation + Trust;
- Unbelief = Surprise + Disgust;
- Anxiety = Anticipation + Fear;
- Outrage = Surprise + Anger;
- Love = Joy + Trust;
- Remorse = Sadness + Disgust;
- Guilt= Joy + Fear;
- Delight = Joy + Surprise;
- Pessimism = Sadness + Anticipation;
- Curiosity = Trust + Surprise;
- Awe = Fear + Surprise;
- Despair = Fear + Sadness;
- Pride = Anger + Joy;
- Shame = Fear + Disgust;

Later, the Pearson correlation (r^2) was applied to both each basic emotion and dyad emotions, to obtaining the correlation between the number of backspaces in the writing process[3] (*AmountErrors* as presented in subsection 4.1) and the emotions, according to Table 1.

Despite no single emotion having a strong correlation, it is possible to identify that among all emotions analysed, anxiety is the most relevant for the number of errors during typing, having a moderate correlation.

5.2 Machine learning predictions

A machine learning analysis was applied to determine the influence of the meta-information and emotional labels on the number of errors prediction. For this purpose, 5 different scenarios were considered:

- Scenario A: Only text and opinion – no meta-information neither emotional labels;
- Scenario B: Only meta-information and emotional labels;
- Scenario C: Only meta-information;
- Scenario D: Only emotional labels;
- Scenario E: All dataset information – text, opinion, meta-information and emotional labels.

[3] Remember that we are using this measure to compute the errors number.

Table 2 Algorithms correlations and their Root Mean Squared Error (RMSE).

Scenario	Linear Regression	RMSE	SVM	RMSE	Random Forest	RMSE	Decision Table	RMSE
Scenario A	0.171	391.28	0.347	208.62	0.429	145.36	0.321	104.78
Scenario B	0.774	99.62	0.769	103.32	0.791	96.67	0.739	106.37
Scenario C	0.777	99.14	0.770	103.08	0.780	98.66	0.739	106.37
Scenario D	0.525	134.10	0.528	137.73	0.492	141.70	0.426	143.54
Scenario E	0.320	437.12	0.696	131.79	0.586	127.88	0.591	129.18

Table 3 Ranges of *AmountErrors* per emotions.

	Emotions							
AmountErrors	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
0.0-204.5	2.5-3.5	0.5-1.5	0.5-1.5	0.0-4.5	0.0-0.5	0.0-1.5	0.0-0.5	2.5-4.5
204.5-256.5	3.5-5.5	1.5-2.5	0.0-0.5	5.5-7.5	1.5-2.5	2.5-3.5	0.5-1.5	4.5-∞
256.5-340	3.5-5.5	0.5-1.5	1.5-2.5	5.5-7.5	0.5-1.5	3.5-∞	0.0-1	2.5-4.5
340-∞	5.5-∞	2.5-∞	2.5-∞	7.5-∞	2.5-∞	3.5-∞	2.5-∞	4.5-∞

The dataset used for both training and validation is the same created in subsection 4.2. In this analysis, the meta-information *AverageErrors* was discarded because it has strong correlation with *AmountErrors*, induced by $AmountErrors \approx AverageErrors * TimeText$.

For each scenario, the following machine learning algorithms were applied: Linear Regression, SVM, Random Forest and Decision Table.

All tests were performed using a 10-fold cross-validation in Weka; the correlation coefficients obtained with each algorithm between AmountErrors and the dimensions analysed in each scenario are shown in Table 2.

After the tests, the best correlation coefficient for predicting the number of errors was obtained with the Random Forest algorithm for Scenario B (only meta-information and emotional labels).

Once identified that the meta-information and emotional labels are an adequate to predict the number of errors, the next step was to identify the patterns for these predictions. For this purpose, all values were discretized into 4 groups and a K-Means algorithm was used to cluster the data (meta-information and emotional labels) into 4 different clusters representing respectively 39%, 17%, 23% and 21% of the information available.

In a preliminary analysis, all meta-information was identified with the same value range, and for this reason, it was considered irrelevant for this objective and removed from the visualization. Also, as the purpose of this analysis is to measure the emotional influence in the *AmountErrors* values, the dimensions *Positive* and *Negative* were removed too. Then the relevant information remaining was grouped by *AmountErrors* and is presented in Table 3, where the range in *AmountErros* refers to the number of errors identified, while the range for each emotion refers to the number of words containing the emotion in the text.

Having in mind the results in Table 3, it is possible to conclude that in general, as higher the emotions are, higher is the impact on the *AmountErrors* number.

6 Conclusion

A typist certainly will have fewer errors than a normal person when typing. However, even this typist will do more mistakes if he is, for instance, anxious or feeling guilty.

As a first step in a research direction we want to further explore – *the sentiment analysis in different tasks to understand the effect of the worker's emotional state on his performance, to reduce mistakes* – this paper presents a combination of lexicon-based and machine learning approaches to correlate the number of typing errors based on the emotional labels and metrics associated with the text creation (text first typing/editing). That model can be used to predict errors based on the information of the emotional state; in this way we will have a rigorous criterion to recommend people to stop doing some task under some personal states to avoid dangerous faults.

Everyone has particular characteristics of expressing himself and these personal characteristics can of course influence that prediction. Maybe on account of that, the study results so far obtained were a bit surprising, and the measured influence of emotions on users tendency to make mistakes is *moderate* (we were expecting bigger values). In our tests, the best approaches for predicting errors based on human behaviour were obtained using emotional information (emotions inferred from the text lexical analysis)and meta-information (metrics evaluated based on the text creation process) collected during text typing. Clustering the data revealed how the emotions can affect the number of errors. It is a promising result.

Despite the final outcomes, this work is the beginning of a research line, and some premises were assumed (like, calculate typing errors by counting the backspace-key strokes) in order to get objective data to produce numerical, trustable, results. However, when advancing this research line, these premises will be revisited and new detection/measuring approaches will be considered.

As future work, it is planned to increase the accuracy of the system to handle with character block removing (i.e. removing several characters at once with a mouse text selection and one backspace press). Moreover, it is planned to detect and handle typos as a different category of errors, as well as to determine on an individual scale, how the emotions affect people in their usual routine, increasing the risks of errors in their activities. Also, it is planned to analyse different texts from different authors. It is important to avoid bias in themes and identify "stressive patterns" which increase the number of errors in different authors and themes.

References

1 Ritwik Banerjee, Song Feng, Jun Seok Kang, and Yejin Choi. Keystroke patterns as prosody in digital writings: A case study with deceptive reviews and essays. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1469–1473, 2014.

2 Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

3 Paul Ekman and Richard J. Davidson. *The nature of emotion: Fundamental questions.* Oxford University Press, 1994.

4 Carroll E. Izard. *Human emotions.* Springer Science & Business Media, 2013.

5 Matthew L. Jockers. *Syuzhet: Extract Sentiment and Plot Arcs from Text*, 2015. URL: `https://github.com/mjockers/syuzhet`.

6 Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. The Stanford CoreNLP natural language processing toolkit. In *52nd Annual Meeting of the Association for Computational Linguistics*, pages 55–60, 2014.

7 Ricardo Martins, José João Dias de Almeida, Pedro Rangel Henriques, and Paulo Novais. Increasing authorship identification through emotional analysis. In *Trends and Advances in Information Systems and Technologies*, volume 745, pages 763–772. Springer International Publishing, 2018. `doi:10.1007/978-3-319-77703-0_76`.

8 Robert R. McCrae and Oliver P. John. An introduction to the five-factor model and its applications. *Journal of personality*, 60(2):175–215, 1992.

9 Scott Meier, Patricia R. McCarthy, and Ronald R. Schmeck. Validity of self-efficacy as a predictor of writing performance. *Cognitive therapy and research*, 8(2):107–120, 1984.

10 Saif Mohammad and Peter D. Turney. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465, 2013. `doi:10.1111/j.1467-8640.2012.00460.x`.

11 Robert Plutchik. *Emotion: A psychoevolutionary synthesis*. Harpercollins College Division, 1980.

12 Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment strength detection in short informal text. *Journal of the Association for Information Science and Technology*, 61(12):2544–2558, 2010.