Teacher Issues and Concerns in Dutch Secondary Education: A Text Mining Approach

Mikola Haagen

2820455

University of Twente

Faculty of Behavioral Management and Social Sciences

MSc Psychology

Positive Clinical Psychology and Technology

Cito Foundation

CitoLab

Master Thesis

First Supervisor: Dr Peter ten Klooster

Second supervisor: Eva de Schipper, MSc

September 4, 2022

Author Note

Special thanks to Eva de Schipper MSc, researcher at the Cito Foundation in Arnhem, and Dr Peter M. ten Klooster from the department of Psychology, Health and Technology, at the University of Twente in Enschede, for guiding me through the process of writing my Master Thesis, reviewing it, and giving feedback on it. Correspondence concerning this thesis should be addressed to Mikola Haagen, 6, rue de la montée, L-7423 Dondelange, Grand-Duché de Luxembourg. E-mail: m.haagen@student.utwente.nl

Abstract

This study explores the potential of two text mining techniques, topic modeling and sentiment analysis, in the educational and psychological domains. It applied an exploratory approach to learn what issues teachers are facing at work and how they feel about these, through analyzing online blogs written by Dutch secondary education teachers on two blogging platforms. Topic modelling and lexicon-based sentiment analysis were combined to explore the topics the teachers discuss in their blogs and the sentiments they hold about these topics. Nine topics were found and interpreted, such as “Workshops”, “Online education (digital tools)” and “Pupil-teacher interactions”. All of which had a positive sentiment score. Generally, the teachers in this study sample used more positive than negative words when writing blogs. The teachers appeared to value topics which involve social interactions the most and they generally had a positive attitude about topics that involve teaching structure. In contrast, the teachers valued topics involving online education the least. Although, due to the lack of negatively valanced topics, this study was not able to provide an overview of issues teachers are facing at work, it nonetheless provides important insights into aspects that may influence teachers’ morale and job satisfaction. Conclusively, the present study suggest that text-mining can be used to offer a holistic overview of what teachers do at work, what seems to be important to them, what they experience and how they feel about it. Additionally, the present study proposes an efficient alternative to traditional (and often time-consuming) qualitative research methods, which can, when cautiously applied, produce a general overview of relevant topics and sentiments from large corpora of pre-existing textual data.

(Word count: 271)

Keywords: Natural language processing, NLP, text mining, topic modelling, sentiment analysis, SA, education, educational, blogposts, blog posts, latent Dirichlet allocation, LDA.

Teacher Concerns in Dutch Secondary Education: A Text Mining Approach

Introduction

Text is one of the most important types of data that humans use to store and pass on information (Zhang & Han, 2019). More than 80% of human knowledge is estimated to be concealed in the form of unstructured text and it is rapidly expanding in volume, accessibility, and relevance across academic fields (Gandomi & Haider, 2015; Antons et al., 2020). A myriad of textual contents is created, shared and analyzed every day, ranging over a wide spectrum of domains such as news articles, books, research articles, technical papers and dictionaries, medical reports, product reviews, e-mail messages, social media posts, fora, web pages and blog posts (Zhang & Han, 2019; Han et al., 2012; Stedman, 2020). The type of textual data analyzed in the present study comes from blog posts written by teachers, i.e., unstructured textual data. The present study explored the potential of an automated technique called *Text Mining*, suited to derive information from unstructured textual data, in analyzing blog posts written by Dutch secondary education teachers to obtain an overview of issues teachers are facing at work and how they are feeling about these issues.

One way to derive meaning and information from unstructured text for research purposes is by conducting analyses such as qualitative content analysis (Kuckartz, 2014). However, these techniques require researchers to manually label and code entries in the data set, meaning that these qualitative analysis approaches suit best for comparably small data sets. When analyzing big data, i.e., data sets that are too complex and large for manual coding, or that take a disproportionate amount of time or costs to analyze by traditional methodologies, computer-based approaches are the preferred methodological choice (Mohanty, 2015). In order to bridge the gap between unstructured natural language written for people to read and structured data that can be processed and *understood* by computers, a methodology that involves an understanding of natural language and symbols present in texts needs to be used.

Text mining, also known as text data mining or text analytics, is an artificial intelligence technology that uses natural language processing (NLP) to convert unstructured text into structured data that is suited for automated analysis (Milward, 2019). It is an interdisciplinary field that is based on a number of advanced statistical, machine learning, and linguistic techniques (Talabis et al., 2015; Han et al., 2012). Text mining refers to the procedure of analyzing and extracting meaningful information, such as themes discussed or sentiments expressed, from large amounts of unstructured text into a structured format in order to identify new insights, hidden relationships and meaningful patterns within written text bodies (Kao, 2020; Stedman, 2019; IBM Cloud Education, 2020a). This procedure usually entails the use of programming languages, such as Python, to structure the raw natural text, identify meaningful patterns and trends, and to eventually interpret the results (Kao, 2020; Talabis et al., 2015). In this study, text mining will be used in order to gain a better understanding of issues that teachers in the Netherlands face during their work.

The birth of text mining can be dated back to the 1990s, when researchers started using unstructured text as data. Early text mining was primarily concerned with different forms of *information retrieval* and *summarization*, i.e., creation of abstracts and indexes as well as the grouping of documents (Miner et al., 2012). Later, the focus shifted to *information extraction*, where contents and underlying relationships hidden in text corpora are unveiled. Nowadays there are many applications for text mining, such as text clustering, topic modelling, sentiment analysis and concept extraction (Talabis et al., 2015).

Present-day applications of text mining techniques involve mostly commercial settings (Kotu & Deshpande, 2019; Han et al., 2012). For example, text mining techniques are used in e-mail spam filtering (Verma & Nasib Singh, 2020; Alurkar et al., 2017), patent mining

(Tseng et al., 2007; Kim et al., 2019), consumer sentiment analysis (Humphreys & Wang, 2017; Berger et al., 2019), finances (Gupta et al., 2020; Ahmad et al., 2007), biomedical analyses (Krallinger et al., 2011; James et al., 2017), risk management (Shah et al., 2021; Chu et al., 2019), public attitude (Huang et al., 2022; Merson & Mary, 2017) and online media analyses (Agichtein et al., 2008; Yue et al., 2018). The common denominator in the listed applications is that they all analyze large amounts of textual data. Text mining techniques are not domain specific, and therefore can be used in any field of research as long as the studied material is comprised of large amounts of textual data.

In this study, the potential of text mining techniques is explored in the educational domain by examining whether it can offer a useful overview of issues teachers are facing at work and how they are feeling about these issues. Such research can be done through extensive interviews or questionnaires which would be time consuming, costly to administer and evaluate, and would only be able to reach a limited amount of people. Conversely, the automated nature of text mining tools allows for time and cost-efficient large-scale data analysis (Tahmasebi & Hengchen, 2019; McDonald & Kelly, 2012; Antons et al., 2020). Furthermore, by using pre-existing data, no further investments of time and effort are required from the target group. Moreover, the naturalistic nature of the information extracted from blog posts offers an unbiased real-world representation of teacher perspectives (Given, 2008).

Case Study: Teacher Issues and Concerns

Teachers have an important role in the educational system. They provide pupils with resources, motivation and assistance required to academically succeed. Teachers also require motivation and assistance in order to maintain their morale, i.e., the enthusiasm and professional interest an individual displays towards (group) achievements in relation to their job (Rempel & Bentley, 1970). High teacher morale and job satisfaction is indicated by topics centering around enthusiasm about and interest in teaching, job related goals and opportunities, inclusion in policy making and decision making, administrative support and good communication (Rempel & Bentley, 1970). Meanwhile, topics that center around job-related stress, frustration, feelings of being undervalued or low respect, low pay, low levels of recognition and low achievement are indicative for low teacher morale and job dissatisfaction (Rempel & Bentley, 1970).

High teacher morale and job satisfaction is important for both pupils and teachers themselves. First, low levels of teacher morale and job satisfaction can lead to reduced productivity in teachers and in some cases to burnout and depression (Hooftman et al., 2014). Furthermore, it can lead to low feelings of belongingness or feelings of detachment, greater use of sick leave, negative or cynical perceptions of pupils and decreased teaching quality (Rempel & Bentley, 1970). Second, high teaching quality is positively associated with pupil achievement and productivity, and greater teacher job satisfaction and morale can lead to better test outcomes in pupils (Goldhaber, 2004; Mackenzie, 2007; Miller, 1981). Thus, teacher morale and job satisfaction have a great impact on both pupil academic achievement and teacher health.

Already in 1985, an American study noted that it was difficult to recruit new teachers and retain those who are currently in the field because the benefits of teaching do not outweigh the work-related frustrations, putting the remaining teachers under further strain (Rosenholtz, 1985). This trend can also be observed at present in the Netherlands, as in the academic year of 2021/22 a shortage of 10,022 teachers was noted by the Dutch Ministry of Education (De Wit et al., 2006). Since decades, the workload for teachers in the Netherlands has increased due to factors such as the aforementioned lack of qualified teachers, a full and growing curriculum and an increased number of children with special educational needs (Ministerie van Onderwijs, 2012).

To promote teacher morale and job satisfaction and therewith pupils' academic achievements, it is important to first gain an overview of relevant issues teachers are facing at work and how they are feeling about these issues. Traditionally, such an overview would be gained through the use of time-intensive methods such as interviews or questionnaires, which are valuable and important but also demand time and effort from the participating teachers. This study therefore proposes an exploratory approach that may inform intervention and innovation without adding any burden to the target group's workload, by analyzing pre-existing blogs written by teachers in secondary education in the Netherlands using topic modelling and sentiment analysis (SA).

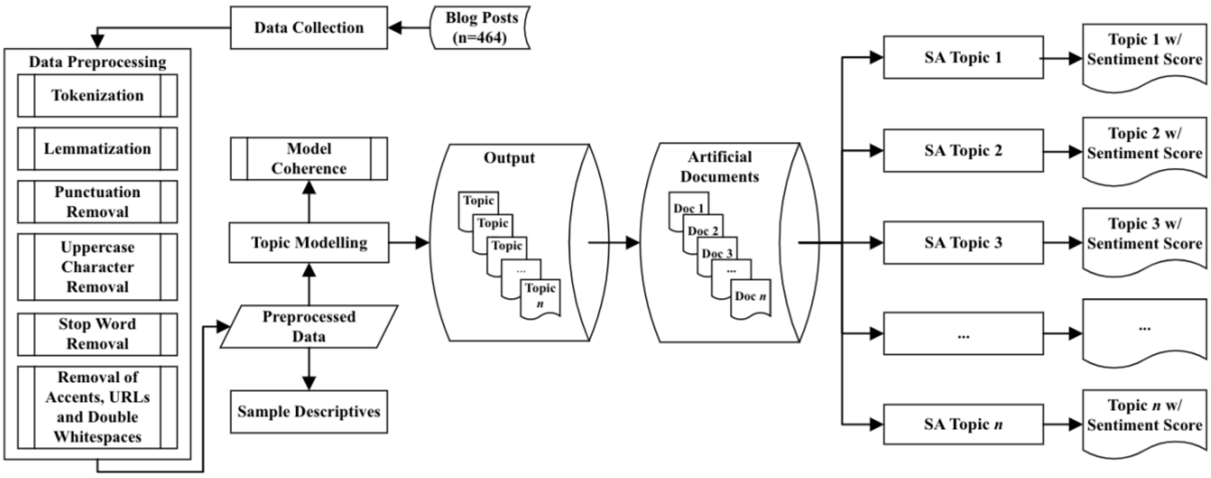
The aim of the present study is to examine and illustrate the usefulness of text mining techniques to gain insight into relevant issues teachers are facing at work and how they feel about these issues, by combining two types of text mining techniques to answer the following research questions: (1) What are the topics Dutch secondary education teachers are discussing in online blogs? (2) How do Dutch secondary education teachers feel about the topics they discuss in online blogs?

Methods

The following sections describe the data collection, data handling and analyses performed in this study. All steps of the present study's workflow are outlined in Figure 1.

Figure 1

Flowchart



Data Collection

The data for the present study was *scraped*, from a list of Dutch educational bloggers provided by a website called trendmatcher.nl using the Selenium library in Python and Chromedriver (Karssenberg, n.d.; Van Rossum & Drake, 1995; Huggins, 2004). The scraped data was consequently imported into the Python-based open-source data visualization, machine learning and data mining toolkit Orange (Demšar et al., 2013). At the time of writing, Trendmatcher.nl listed 298[1] teachers from different educational levels in Dutch education (PO[2], VO[3], MBO[4], HO[5], and SO[6]), who write blogs on 16 different blogging platforms. Given that the current study's target population are Dutch *secondary education* teachers, only bloggers who exclusively teach in VO were retained, resulting in a total of 28

[1] Figure may not be accurate as Trendmatcher.nl regularly updates their list.
[2] PO (Primair Onderwijs): Primary education.
[3] VO (Voortgezet Onderwijs): Secondary education.
[4] MBO (Middelbaar beroepsonderwijs): Vocational education.
[5] HO (Hoger onderwijs): Higher education.
[6] SO (Speciaal onderwijs): Special-needs education.

blog domains potentially suitable for the analyses. Since scraping entails a tedious process (writing and running the platform-specific scripts), blog domains from the two largest platforms mentioned on Trendmatcher.nl were chosen to be scraped, resulting in a total of 22 blog domains. It was expected that this decision would not introduce a bias in the present study's sample as there was no reason to suspect systematic differences in blog content across platforms. The scripts did not manage to scrape data from seven domains due to technical difficulties related to the format of the webpages. One blog domain was excluded from further analyses as its content was exclusively composed of mathematical formula explanations and a last blog domain was excluded as all of its documents were composed of English text, which cannot be analyzed using a Dutch language model. Conclusively, 464 blog entries composed by 13 teachers from secondary education in the Netherlands were collected from two blogging platforms (*Blogspot* and *Wordpress*) that were posted between 24.04.2012 and 08.03.2022 and contained on average 685.14 words ($SD = 542.172$).

Table 1 lists the blog domains of the 13 teachers, the corresponding blogging platform, the number of blogs collected per blogger, and the mean length of blogs in number of words. 334 out of 464 scraped blogs were composed by the six teachers on the blogging platform *Blogspot* ($M_{n.words} = 700.91$, $SD_{n.words} = 512.559$) and seven teachers composed 130 blogs on the blogging platform *Wordpress* ($M_{n.words} = 644.64$, $SD_{n.words} = 612.013$). The following sections describe the methods that were used to preprocess and analyze the data.

Table 1

Sample Descriptives

Blog Domain	**Blogging Platform**	**Number of Blogs**	**Length of Blogs in Number of Words**	
			M	***SD***
1. adriennedekock	Blogspot	22	780.73	311.565
2. astridwillemse	Blogspot	64	337.77	270.393
3. kingboko	Blogspot	31	581.87	266.417
4. martijnsytsma	Blogspot	57	705.86	408.015
5. nieknijenkamp	Blogspot	11	549.00	385.958
6. media-ictenmedia	Blogspot	149	879.19	603.414
7. kaizenonderwijs	Wordpress	20	582.50	172.887
8. bernardblogt	Wordpress	20	416.85	351.401
9. jasperrijpmablog	Wordpress	19	1276.94	1005.536
10. ipadindeklas	Wordpress	20	286.80	228.130
11. jvremoortere	Wordpress	20	1171.45	547.402
12. robalberts	Wordpress	20	336.45	183.368
13. suvanomics	Wordpress	11	332.73	95.141
Total Blogspot	-	334	700.91	512.559
Total Wordpress	-	130	644.64	612.013
Total	-	464	685.14	542.172

Data Preprocessing

The text data that was analyzed in the present study was written and meant to be understood by humans and thus it includes elements such as punctuation, accents, function words and capitalization. These grammatical elements of natural language help convey a message in a readable, meaningful and understandable way from one person to another. Contrary to humans, who extract the important information from text while reading, machine learning algorithms struggle differentiating meaningful from meaningless textual elements. Therefore, thorough preprocessing of raw text data is required.

Text preprocessing refers to the *cleaning* of text data for the use of text mining algorithms (Wesslen, 2018). The steps taken during data preprocessing aim at standardizing the text data and to remove text, characters and symbols that are not relevant for the analysis (Danubianu, 2015). Hence, the cleaning performed during preprocessing organizes the data into a suitable form for text mining algorithms and increases the performance and quality for the models obtained by text mining procedures (Danubianu et al., 2012; Vaishali & Rashmi, 2018). The following sections present an overview of the preprocessing procedures that were applied in the current study.

Tokenization

Tokenization refers to a process of splitting sequences of text strings into *tokens*, which are smaller units of text, mostly containing a single word (Mayo, 2017; Vaishali & Rashmi, 2018). Tokenization is a mandatory step that allows to work with text data since machine learning algorithms process the data at the level of tokens (Aravind, 2020). Thus, in order to get an algorithm to *understand* text data, sequences of text strings need to be converted into tokens. The current study applied the pretrained *UDPipe Regex tokenizer*[7] for Dutch language.

Normalization

Normalization refers to procedures that are applied during data preprocessing that aim at converting text into a standard form and removing unnecessary information. Normalization is an important part of data preprocessing, as it reduces the number of unique tokens present in a document (Ganesan, 2019). There are mainly two ways to normalize textual data: *Stemming* and *lemmatization*.

Stemming refers to a rule-based process for reducing the number of variations of a token to one single token by removing all of a word's suffixes such that only the word's stem remains (Saxena, 2021). For example, the words "smile", "smiles", "smiling" and "smiled" would all become "smil" after the stemming process. Stemming is a simple but suboptimal choice for normalization, as it occasionally creates meaningless terms that are not included in dictionaries such as the term "smil" from the previous example (Saxena, 2021).

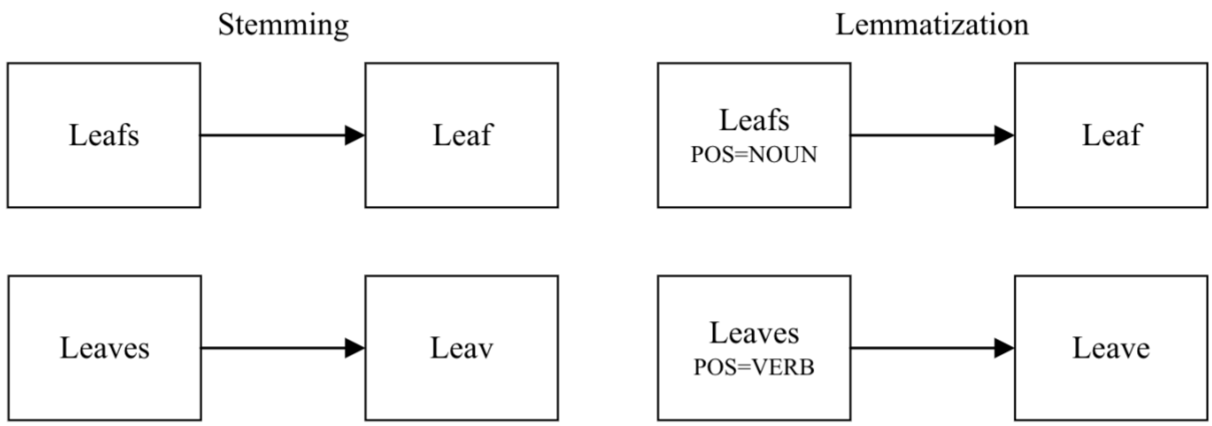
Lemmatization refers to a systematic process that removes inflectional forms (i.e. *lexemes*) of words and transforms them into *lemmas*. In linguistics, a lemma stands for the base word of its lexemes, i.e. the word in its dictionary form. The main difference between lemmatization and stemming, and the reason why lemmatization is generally preferred over stemming is that lemmatization considers the context in which the lexeme is used and normalizes it into its meaningful lemma (Prabhakaran, 2018). Since a word can have multiple lemmas, the context in which the terms are lemmatized needs to be specified through the identification of *part-of-speech (POS) tags* (Prabhakaran, 2018). POS tagging refers to a supervised machine learning method that defines the syntactical function of a word (e.g. noun

[7] Refer to Straka and Strakova's article on "Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe" for a detailed explanation of the UDPipe Regex tokenizer *(2017).*

or verb) depending on its surrounding words, whether its first letter is capitalized, etc. (Kargın, 2021). The POS tags thus help lemmatization to find the correct lemma for a given lexeme. Figure 2 depicts the differences between stemming and lemmatizing and demonstrates the benefits of the latter. The current study applied the *UDPipe Lemetizer*[8] for Dutch language.

Figure 2

The Difference between Stemming and Lemmatizing



Transformation and Filtering

Punctuation Mark Removal. While punctuation marks aid at facilitating the readability of text for humans by for example indicating the end of a sentence, it does not help an algorithm to better *understand* text. Contrarily, since punctuation marks are converted into standalone tokens, when unremoved, they make up a large portion of tokens within a sample and thus may hinder the algorithm to identify meaningful patterns (Etaiwi & Naymat, 2017).

Uppercase Character Removal. In most languages (including Dutch), the uppercase character is primarily used for special purposes, such as the first letter of a proper noun or the first letter in the beginning of a sentence. Sometimes, uppercase letters are also used to highlight the (sentimental) importance of a word or to express anger. Consequently, the use of

[8] Refer to Straka and Strakova's article on "Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe" for a detailed explanation of the UDPipe Lemetizer *(2017).*

uppercase letters in text helps to convey extra information about the word it is applied to. At first, it may sound counterintuitive that uppercase characters are removed from the text data. However, since the algorithm does not truly *read* the words within a document, it tokenizes uppercase and lowercase variants of the same word into two distinct tokens. In order to avoid obtaining two tokens of the same word, and thus weakening the algorithm's abilities to find patterns in the text data, uppercase characters have been removed, even though this means that some information about the data might be lost.

Stop Word Removal. *Stop words* refer to the terms in a *stop list* (i.e., negative dictionary) which do not convey any significant or desired meaning and are therefore removed from the text data (i.e. stopped) before conducting analyses (Rajaraman & Ullman, 2011). Generally, a stop word list is composed of functional words (e.g., and, on, off, at, to, so, etc.)[9] which only convey syntactic information. Stop words are frequently used in natural languages and would make up a large portion of tokens within a sample if they were included and thus may impede the algorithm's usability to identify meaningful patterns. In the present study, an open-source negative dictionary was used to remove stop words from the data. Furthermore, a list of further words and symbols that were frequently used in the present study's text data and which had not yet been removed through the pre-existing stop word list (such as 1, 2, 3, x, etc.) was added after manual inspection of the preprocessing results.

Further Transformations. Accents, URLs and double whitespaces were also removed from the text data as these elements do not add any meaningful information to the present study's text data. Furthermore, the removal of these elements promotes the uniformity of the text data, which in turn aids the text mining algorithms to perform better analyses.

[9] translated from Dutch "en, aan, af, bij, te, zo, etc.".

Figure 3 shows a visual representation of the effects of data preprocessing on the present study’s data set. Table 2 shows an example of the effects of preprocessing on a phrase segment from the present study’s data.

Figure 3

Word clouds for before and after preprocessing of the Blog posts

a. b.

Note. a. visualization of the most frequent terms in the unprocessed data. b. visualization of most frequent terms in the preprocessed data.

Table 2

Data Preprocessing Example

Unprocessed text	**Preprocessed text**
En toen de laatste auto door de carwash wegreed, gebeurde het. Mijn onderwijsmoment van 2020. Wat er gebeurde? Daarvoor ga ik eerst even terug […]	auto carwash weg rijd gebeur onderwijsmoment gebeur daarvoor eerst even terug […]

Note. Refer to Table A in the Supplementary Materials for the English translation.

Topic Modelling

With the rise in popularity of social media platforms and the digitalization of information, extracting relevant insights from these vast data sources (i.e. big data) becomes more and more interesting to researchers across academic fields. Big data can be analyzed though *supervised* or *unsupervised machine learning methods*. Supervised machine learning methods are a subtype of machine learning techniques in which an algorithm is trained using *labelled* training data to classify unseen data into pre-specified categories (Jaiswal, 2021). Labelling refers to manual tagging of the input data with the correct output prior to the training phase (Jaiswal, 2021). Unsupervised machine learning methods are used to automatically analyze and cluster unlabeled datasets using machine learning algorithms (IBM Cloud Education, 2020b). These algorithms uncover hidden patterns or data groupings without the need for human interaction (IBM Cloud Education, 2020). Topic modelling is an example of an unsupervised machine learning method that allows for extraction of hidden semantic structures (i.e *topics*) from a collection of text documents (Mujahid et al., 2021, Jelodar et al., 2018; Blei et al., 2003; Lossio-Ventura et al., 2021; Dahal et al., 2019; Blei, 2012). Its algorithms classify text bodies into coherent themes that describe the collection of documents without human intervention (Mutanga & Abayomi, 2020; Kanan et al. 2015). The resulting themes (i.e., topics) can range from very specific to very general depending on the predefined number of topics to be extracted and on the content of the input data (Blair, et al., 2020). Simply put, topic modelling scans text corpora, analyses the frequency and similarity of words used, and classifies the text bodies according to automatically defined topics.

Latent Dirichlet Allocation (LDA) is a popular and well-established topic modelling technique (Jelodar et al., 2018). LDA is a three-layered hierarchical Bayesian model which relies on the bag-of-words model, where a text (such as a document or sentence) is represented as a bag (multiset) of its words, which ignores grammar, syntax and word order while maintaining multiplicity (Harris, 1954). In this model, each document is made up of different words, and each topic has its own set of words, i.e., each topic is represented as a probability distribution over the words in a dictionary (Blei et al., 2003; Reed, 2012). LDA looks at the word co-occurrences within documents and assumes that words appearing in the same document are more likely to be related to the same topic compared to words that are not, and that documents that share common words are more likely to contain the same topics than documents that do not. In the context of the current study, a document refers to a blog post. Thus, LDA's goal is to determine which topics a document belongs to, based on the words it contains. In this study, LDA was used in order to extract topics that teachers discuss in online blog posts.

In order to determine the interpretability of the topic modelling's output, it was looked at the *coherence scores* for topic models at different numbers of topics. In topic modelling, the coherence score measures how similar the analyzed words within one topic are to each other (Zvornicanin, 2021). For instance, the words {cat, dog, mouse, cow} are more closely connected than {cat, talk, radio, plant} and therefore they will have a better coherence score than the latter.

The coherence score for any given topic model is calculated as follows:

$$C(t; V^{(t)}) = \Sigma_{m=2}^{M} \Sigma_{l=1}^{m-1} log \frac{D\left(v_m^{(t)}, v_l^{(t)}\right) + 1}{D(v_l^{(t)})}$$

where

$D(v)$ = number of documents including at least one token of type *v*

$D\left(v_m^{(t)}\right)$ = the conditional probability of each word given each of the higher-ranked words in the topic

$V^{(t)} = (v^{(t)}, \dots , v_M^{(t)})$ is a list of the M most probable words in topic *t*

Note. Since the result of log 0 would be undefined a *smoothing count* of 1 is added to the equation.

Given that the coherence score formula includes a logarithm of a fraction, the output of this formula always results in a value lower than one, such that a low coherence score indicates high topic coherence, and a coherence score closer to zero means low topic coherence for a given model (range -100 to 1). There is no guideline that determines what value signifies a good or a bad coherence score, because the score depends on the context of the input data. Generally, a greater number of topics results in a smaller (i.e. better) coherence score, and this decrease becomes progressively smaller (Zvornicanin, 2021). Therefore, the point at which the decrease of the coherence score was so small that it was no longer worth the additional increase in the number of topics was chosen. This technique is referred to as the elbow technique.

Research has shown that models with a relatively small number of topics are the easiest to interpret by human judgment (Remmits, 2017). Human judgments are however prone to interpretation bias, and therefore for this study, two researchers independently rated which topics did not convey enough meaning to be retained for further analyses. The topics on which both researchers agreed that they were not meaningful were excluded. The remaining topics were labeled independently by the two researchers and were consequently compared and the final labels were determined by consensus.

Sentiment Analysis

Sentiment analysis (SA), also known as opinion mining, is another text mining tool that uses NLP to automatically extract meaningful information from text data (Turney, 2022; Pang & Lee, 2008). It is used to reveal opinions within a text without the need of extensive manual analysis. Cambria, 2016; Liu, 2015). SA refers to the technique used in determining whether a segment of text has a negative, neutral, or positive valence, i.e., the computational assessment of feelings, opinions, and subjectivity in textual data (Martínez-Cámara et al., 2012). SA is mostly used by companies to analyze customer feedback and product reviews (Medhat et al., 2014). However, SA is also increasingly used in reputation management (Colleoni et al., 2011; Olaleye et al., 2018), social media monitoring (Neri et al., 2012; Arunachalam & Sarkar, 2013) and public opinions (Singh et al., 2019; Sabatovych, 2019). In short, SA is used to automatically identify whether a text expresses positive, negative, or neutral feelings about an idea, product, or a particular topic (Kaur & Bhatia, 2016).

There are two major approaches for SA, viz. supervised machine learning SA and lexicon-based SA (Eisenstein, 2016). In the present study, lexicon-based, also known as dictionary-based, SA was applied. Lexicon-based SA compares the words present in a given document to an already existing lists of positive and negative terms (i.e. *sentiment lexicon*) and counts the number of positive and negative words used in the document (Chen & Skiena, 2014). Lexicon-based SA then calculates sentiment scores based on the frequency of positive and negative words present in the document (Chen & Skiena, 2014). The sentiment lexicon used in the present study is the Data Science Lab open-source multilingual sentiment lexicon for Dutch language which includes a total of 3994 words (Chen & Skiena, 2014). This lexicon was developed as part of a larger project in which lexica for 136 languages were generated (Chen & Skiena, 2014).

The sentiment score for a segment of text is calculated as follows:

$$Sentiment\ Score = 100 * \frac{n_{pos} - n_{neg}}{n_{pos} + n_{neu} + n_{neg}}$$

where

Sentiment score = percentage of sentiment difference in the document

n_{pos} = number of positive words in a document

n_{neu} = number of neutral words in a document

n_{neg} = number of negative words in a document

Range = -100 to 100

Unlike the results from topic modelling which are obtained through analyzing the whole set of documents, a sentiment score computed through SA normally refers to a single document within the whole set of documents. Moreover, a document has a probability distribution over topics, which implies that it is impossible to collect the exact text sequences that *belong* to a topic. Hence, getting a sentiment score for a topic requires some extra steps. There are two approaches commonly applied to obtain sentiment scores for topics, namely Sen*timent-Based Topic Modelling* and *Aspect-Based Sentiment Analysis.*

Sentiment-Based Topic Modelling (SBTM) refers to a commonly applied combination of SA and topic modelling, where first a SA on all of the documents present in the set is conducted. Then each of the documents is classified according to their sentiment score in either of the three categories “positive”, “neutral” and “negative”, and finally one topic model per category is conducted to obtain a list of topics per sentiment category. Unfortunately, this approach does not suit the structure of the present study’s documents. The analyzed data set includes documents of an average length of ca. 685 words, which raises the likelihood that multiple topics are discussed within a document. If a document would include for example one positive, one negative and one neutral topic, the SA would fail to catch this and would assign the entirety of the document to one of the three sentiment categories. In contrast, when analyzing smaller documents, such as Twitter posts which are limited to 40-70 words, it can be expected that it is less likely that multiple topics are discussed within one document and thus this approach may be more reasonable to apply compared to the present study’s approach.

Aspect-Based Sentiment Analysis (ABSA) refers to a text analysis technique that first categorizes data by aspect, and then identifies the sentiment related to each aspect. ABSA simultaneously identifies aspects, and sentiment expressions in each phrase of a document, such that each phrase includes at least one syntactically related aspect-opinion pair (see Figure 4). The obtained pairs are then manually classified into coherent topics and the sentiment expressions are evaluated in the same way as it is done in SA (see Table 3).

An important feature of this approach is that the number of aspect-opinion pairs per document depends on the length of the document, meaning that large documents produce many highly specific aspect-opinion pairs which all need to be classified into overarching topics. While this approach performs very well in a predefined scope such as restaurant or product reviews where only a limited number of topics is expected (eg. food, service, location, price, quality, etc.), it does not perform well on broad research questions as it would produce too many aspect-opinion pairs which are challenging to classify into coherent topics that would truly represent a general overview of what a diverse population discusses in online blogs. Furthermore, applying an ABSA requires a lot of manual labor, which in itself defeats the purpose of conducting *automatic* text analyses as opposed to qualitative research methods.

Figure 4

Example of Aspect-Based Sentiment Analysis (ABSA) conducted on a Restaurant Review

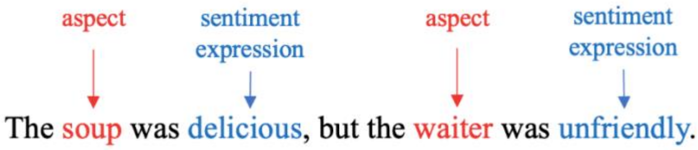


Table 3

Example of Aspect-Based Sentiment Analysis Output (continued)

Topic	Aspect	Sentiment Expression	Sentiment
Food	Soup	delicious	positive
Service	Waiter	unfriendly	negative

Note. The words “soup” and “delicious” form one aspect-opinion pair for the positive topic “food”.

Both alternatives, Sentiment-Based Topic Modelling and Aspect-Based Sentiment Analysis are valuable text mining tools, but they do not fit the nature of the present study’s data and would result in either too specific or too vague findings which would be hard to interpret and generalize. Therefore, the present study proposed a new approach to analyze sentiments related to topics, through the creation of artificial documents that include estimates of the probabilistic distribution of terms for each topic. The process of obtaining a sentiment score for a topic goes as follows. First, topic modelling is conducted on all of the documents within the sample. Second, an artificial document per topic is created which contains the terms in the corpus with a frequency proportionate to the probability of the term belonging to the topic. Finally, the SA is conducted on the artificial documents to estimate sentiment scores for each topic to a precision of five positions behind the decimal point.

The present study’s sentiment analysis, while being a deterministic model, bases its computations on the probabilistic output of the topic model through the creation of artificial documents that include estimates of the probabilistic distribution of terms for each topic. These artificial documents each include approximately 100,000 terms, such that each term’s probability to appear in a topic is multiplied by 100,000 and thus allows for estimates which approximate their probabilistic counterparts to the precision of five positions behind decimal point. By way of illustration, a term that has a probability of 0.03 to appear in a given topic would display 3000 times in its corresponding artificial document and would make up 3% of the 100,000 words within this topic. Sentiment scores were computed for these artificial documents.

Results

The aim of the present study was to examine the usefulness of text mining techniques for gaining insight into relevant issues teachers are facing at work and how they feel about these issues. A combination of topic modelling and SA was conducted to reveal what topics Dutch secondary education teachers are discussing in online blogs and how they feel about these topics. The following section reports the descriptive statistics, followed by the results obtained from topic modelling and finally the results obtained from SA.

After preprocessing, the 10 most common words used by the 13 bloggers across 464 blogs were “pupil”, “good”, “school”, “teaching/learning”, “education”, “class/lesson”, “year”, “child/kid”, “read” and “new”. Throughout the 464 blog entries, the teachers used a total of 82,594 words of which 18,789 (22.75%) were unique words (see Table 4 and Figure 5).

Table 4

The ten Most Frequently Used Terms in Blog Posts composed by Dutch Secondary Education Teachers

Term	Frequency (Percentage)
leerling	2275 (2.75%)
goed	1121 (1.36%)
school	940 (1.14%)
leren	929 (1.12%)
onderwijs	905 (1.10%)

les	607 (0.73%)
jaar	544 (0.66%)
kind	503 (0.61%)
lezen	483 (0.58%)
nieuw	470 (0.57%)
Total words	82,594 (10.62%*)
Total unique words	18,789 (22.75%)

Note. The word frequency was assessed after preprocessing. * The ten most frequently used words make up 10.62% of all words used throughout the blog posts. Refer to Table B in the Supplementary Materials for the English translation.

Figure 5

Word Cloud of the 50 Most Frequently Used Words in Online Blogs composed by Dutch Secondary Education Teachers

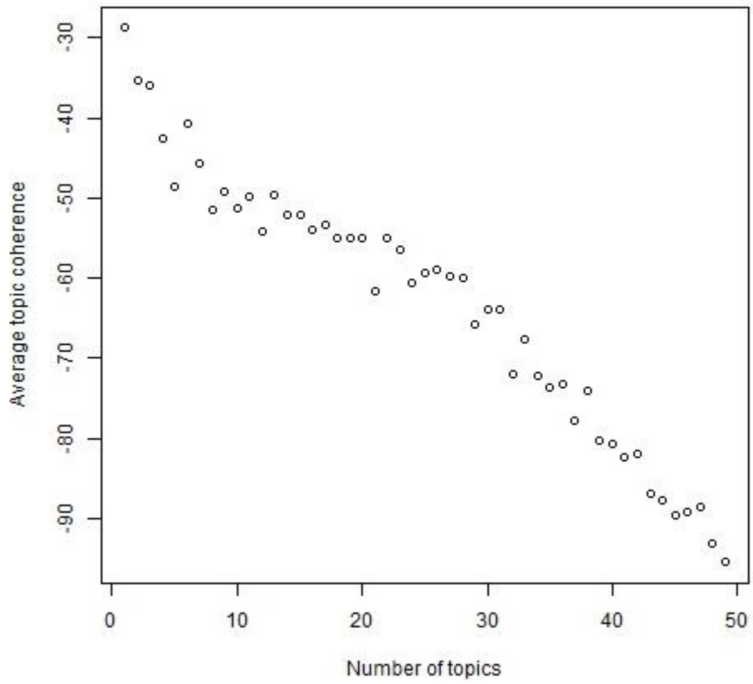
Topic Modelling

The first step in topic modelling is to define the most meaningful number of topics which was done by analyzing the coherence scores of potential models with different numbers of topics as well as a preliminary interpretation of a sample of the most promising models. As aforementioned, the coherence score of a given model refers to the connectedness of words within each topic, and thus informs about the interpretability of the model. Since an LDA topic model is probabilistic in nature, slightly different outputs are obtained each time the script is run. Accordingly, it does not make sense to choose the model with the *best* coherence score, but it does however make sense to analyze the visual representation of the average coherence scores for a range of topic numbers (k), i.e. the model performance, and identify the point at which the decrease of the coherence score is so small that it is no longer worth the additional increase in k (i.e the *elbow* of the plot).

The model performance was assessed by plotting the coherence scores for topic numbers (k) ranging from 2 to 50 and the elbow in the plot appeared to lie between k = 10 and k = 20 with coherence scores ranging from *C* = -51,5609 to *C* = -54,8982 (see Figure 6). A sample of three models at k = 10, k = 15 and k = 20 (*C* = -51,5609, *C* = -52,1548, *C* = -54,8982) from the elbow in the plot was drawn and a preliminary interpretation of the resulting topics for the three models was made by two researchers and it was concluded that the most interpretable model included 10 topics. The models at k = 15 and k = 20 each split some of the resulting topics from the model at k = 10 into subtopics which did not lead to an increase in interpretability.

Figure 6

Model Performance for Different Numbers of Topics



The 10 topics of the chosen model were first independently interpreted and labeled by two researchers, then mutually discussed to reach consensus and the agreed-upon labels were retained. One of the resulting 10 topics was excluded from further analyses[10] as it was composed of English words and thus did not convey any meaning in a Dutch text mining analysis. The final model suggests that Dutch secondary education teachers primarily discussed themes that center around the following 9 topics: “pupil-teacher interactions”, “education at school”, “online education (digital sources)”, “online education (digital tools), “Learning through play/development”, “literature about the Netherlands/Dutch literature”, “Literature (books and stories)”, “assessment” and “workshops” (see Table 5).

[10] Excluded topic: “Engelse woorden: {Ipad, student, school, Twitter, leuk, like, laden, Facebook, share, gerelateren, education, year, you, -, learning, from, this, vinden, will}”

Table 5

Nine Topics Dutch Secondary Education Teachers Discuss in Online Blogs

Topic	Term (Probability)*
Workshops	workshop (0.0249), gaan (0.0248), school (0.0177), worden (0.0158), een (0.0146), dag (0.0135), maken (0.0114), mee (0.0113), mooi (0.0111), jaar (0.0107), tijd (0.0106), elkaar (0.0103), veel (0.0094), samen (0.0087), kommen (0.0079), twee (0.0079), gesprek (0.0077), doen (0.0073), echt (0.0070), allemaal (0.0066).
Lesmateriaal over Nederland/ Nederlands lesmateriaal	vinden (0.0259), veel (0.0209), weten (0.0194), nederland (0.0151), hier (0.0137), zien (0.0101), project (0.0096), Nederlands (0.0088), deze (0.0081), wereld (0.0078), lesmateriaal (0.0074), materiaal (0.0069), bieden (0.0066), water (0.0063), staan (0.0061), informatie (0.0059), beeld (0.0055), onderwerp (0.0054), groot (0.0053), land (0.0050).
Onderwijs op school	worden (0.0320), onderwijs (0.0317), school (0.0204), leraar (0.0196), veel (0.0172), gaan (0.0159), deze (0.0141), goed (0.0127), denken (0.0096), leerling (0.0093), hier (0.0078), eigen (0.0076), willen (0.0073), vak (0.0065), manier (0.0062), laten (0.0061), liggen (0.0061), zien (0.0059), vaardigheid (0.0057), ict (0.0055).
Literatuur (Boeken en verhalen)	lezen (0.0345), boek (0.0333), gaan (0.0261), verhaal (0.0182), schrijven (0.0157), veel (0.0144), jaar (0.0142), deze (0.0123), laten (0.0122), worden (0.0118), tekst (0.0113), komen (0.0107), jong (0.0095), vinden (0.0088), nederlands

	(0.0085), nemen (0.0083), vertellen (0.0075), literatuur (0.0075), willen (0.0073), vaak (0.0067).
Formatief handelen	leerling (0.0638), leren (0.0280), veel (0.0256), geven (0.0219), les (0.0211), maken (0.0178), gaan (0.0174), zich (0.0139), doen (0.0130), een (0.0115), komen (0.0110), werken (0.0109), deze (0.0100), goed (0.0082), opdracht (0.0082), docent (0.0081), volgen (0.0078), formatief (0.0078), laten (0.0076), kijken (0.0075).
Spelenderwijs leren/ ontwikkeling	kind (0.0323), leren (0.0266), worden (0.0173), maken (0.0166), veel (0.0137), belangrijk (0.0132), spelen (0.0127), deze (0.0110), stellen (0.0110), doen (0.0097), ontwikkeling (0.0092), weten (0.0090), mens (0.0085), goed (0.0084), mogelijkheid (0.0076), probleem (0.0068), willen (0.0066), denken (0.0064), spel (0.0064), ontwikkelen (0.0060).
Online onderwijs (Digitale bronnen)	digitaal (0.0245), veel (0.0167), onderwijs (0.0154), media (0.0157), worden (0.0153), docent (0.0137), online (0.0137), onderzoek (0.0116), bron (0.0112), school (0.0102), informatie (0.0095), internet (0.0090), zich (0.0085), maken (0.0084), open (0.0082), vinden (0.0079), gebruik (0.0075), vaak (0.0073), sociaal (0.0073), eens (0.0072).
Online onderwijs (Digitale hulpmiddelen)	maken (0.0357), worden (0.0342), nieuw (0.0223), gebruik (0.0159), leerling (0.0143), zien (0.0125), hier (0.0125), openen (0.0109), klik (0.0106), delen (0.0104), vraag (0.0099), werken (0.0097), zelf (0.0086), film (0.0084),

	venster (0.0083), video (0.0077), willen (0.0071), tool (0.0067), mogelijkheid (0.0095), foto (0.0055).
Leerling-docent interactie	leerling (0.0383), vraag (0.0257), goed (0.0213), doen (0.0206), geven (0.0158), willen (0.0149), komen (0.0136), gaan (0.0126), docent (0.0123), worden (0.0119), vinden (0.0115), iets (0.0105), groep (0.0099), deze (0.0097), zitten (0.0089), zeggen (0.0088), les (0.0086), kijken (0.0083), twee (0.0083), weten (0.0080).

Note. Refer to Table C in the Supplementary Materials for the English translation.

Blue color: positive term. Red color: negative term. All other terms are neutral terms.

*Probability of a term belonging to a topic represented as percentage of total terms within a topic.

Sentiment Analysis

The results obtained through the SA suggest that Dutch secondary education teachers compose texts using mostly neutral words and generally slightly more positive words than negative ones. The black terms in the table above refer to the neutral words used by Dutch secondary teachers, which make up 80% of all words used when teachers discussed the nine topics (Refer to Table 4). The blue terms refer to the positive words, which make up 11 % of all the words used, and finally the red terms refer to the negative words and make up the remaining 9% of words used to describe the nine topics. Table 6 lists the positive, neutral and negative terms which teachers used across nine topics.

Table 6

Positive, Neutral and Negative Terms and their Frequencies across Nine Topics Dutch Secondary Education Teachers Discuss in Online Blogs

Sentiment	Term (Frequency)
Positive	Belangrijk (311), bieden (288), bron (281), echt (296), geven (165), goed (1121), groot (312), mooi (250), nieuw (470), open (318), spelen (108), vaardigheid (187), werken (289), willen (470), zelf (536)
Neutral	Allemaal (152), beeld (147), boek (406), dag (243), delen (425), denken (197), deze (1152), digitaal (468), docent (729), eens (307), eigen (445), elkaar (306) , film (171), formatief (135), foto (120), gaan (1137), gebruiken (242), gesprek (108), groep (190), hier (541), ict(234), iets (347), informatie (264), internet (186), jaar (544), jong (205), klik (186), komen (289), land (102), leerling (2275), leraar (238), leren (929), les (607), lesmateriaal (117), literatuur (136), maken (833), manier (337), materiaal (122), media (236), mee (366), mens (284), mogelijkheid (224), nederland (193), nederlands (235), nemen (413), onderwerp (199), onderwijs (905), onderzoek (359), online (376), ontwikkelen (187), ontwikkeling (196), opdracht (310), openen (318), project (189), samen (216), school (940), schrijven (261), sociaal (181), spel (184), staan (203), stellen (251), tekst (198), tool (102), twee (296), vaak (349), vak (203), veel (711), venster (111), verhaal (164), vertellen (97), video (175), vinden (292), volgen (413), vraag

	(339), water (102), wereld (203), weten (196), worden (957), workshop (112), zeggen (211), zich (434), zien (401), zitten (80).
Negative	Doen (465), gebruik (318), kijken (279), kind (503), laten (355), lezen (317), liggen (30), probleem (151), tijd (315)

Note. Refer to Table D in the Supplementary Materials for the English translation.

The following section describes each topic's sentiment score, as well as its composition of positive, neutral and negative words. As aforementioned, the present study's approach to computing sentiment scores per topic takes both the most relevant words per topic and their probabilities to appear in a given topic into account to produce a topic's corresponding sentiment score. Refer to Table 5 (above) to see each topic's positive terms in blue, neutral words in black and negative words in red. Furthermore, Table 7 (below) lists each topic's sentiment score as well as the overall sentiment score of the entire data set.

No single topic had a negative sentiment score, resulting in a general percentage of sentiment difference across topics of 1.88%, which signifies an overall slightly positive attitude towards all topics discuss by Dutch secondary education teachers in the analyzed online blogs.

The topic with the highest sentiment score and herewith the topic teachers discuss using the most positive and least negative words is "Workshops" (3.09%). The 20 most relevant words used by teachers who discussed this topic were composed of 16 neutral words, 2 positive words and 2 negative words. The topic with the second highest sentiment score is "Literature about the Netherlands/ Dutch literature" (2.81%). The 20 most relevant words used by teachers who discussed this topic were composed of 18 neutral words, 2 positive words and no negative words. The topic with the third highest sentiment score is "Education at school" (2.40%). The 20 most relevant words used by teachers who discussed this topic were composed of 15 neutral words, 3 positive words and 2 negative words. The next topic is "Literature (books and stories)" with a sentiment score of 1.81%. The 20 most relevant words were composed of 15 neutral words, 1 positive word and 2 negative words. The following topic "Assessment" has a sentiment score of 1.77% and is composed of 14 neutral words, 3 positive words and 3 negative words. The sentiment score of the topic "Learning through play/ development" is 1.76%. Its 20 most relevant terms are composed of 13 neutral words, 4 positive words and 3 negative words. The third lowest sentiment score was found in the "Online education (digital sources)" topic (1.55%). The corresponding 20 most relevant terms are composed of 17 neutral words, 2 positive words and one negative word. The second lowest sentiment score was found in the "Online education (digital tools)" topic (0.99%). Its 20 most relevant terms are composed of 15 neutral, 4 positive words and 1 negative word. The topic with the lowest sentiment score and herewith the topic teachers discuss using the least positive and most negative words is "Pupil-teacher interactions" (0.76%). The 20 most relevant words used by teachers who discussed this topic were composed of 15 neutral words, 3 positive words and 2 negative.

Table 7

Sentiment Scores of Nine Topics Dutch Secondary Education Teachers Discuss in Online Blogs

Topic	Sentiment Score
Workshops	3.09
Lesmateriaal over Nederland/ Nederlands lesmateriaal	2.81
Onderwijs op school	2.40
Literatuur (Boeken en verhalen)	1.81
Formatief handelen	1.77
Spelenderwijs leren/ ontwikkeling	1.76
Online onderwijs (Digitale bronnen)	1.55
Online onderwijs (Digitale hulpmiddelen)	0.99
Leerling-leraar interactie	0.76
Overall Sentiment	1.88

Note. Refer to Table C in the Supplementary Materials for the English translation.

Discussion

By combining two types of text mining techniques, topic modelling and SA, the present study aimed at examining the usefulness of text mining techniques to gain insight into relevant issues teachers are facing at work and how they feel about these issues. More broadly, it was aimed at exploring whether text mining techniques are also valuable tools that can be used outside of their traditional domains of application. Hence, the present study answered the following exploratory research questions using a combination of topic modelling and SA: (1) What are the topics Dutch secondary education teachers are discussing in online blogs? (2) How do Dutch secondary education teachers feel about the topics they discuss in online blogs?

Performing topic modelling on blogs written by Dutch teachers in secondary education led to 10 discovered topics, of which one was excluded from further analyses as it did not convey any meaningful information. Sentiment analysis on the remaining 9 topics revealed a mostly positive valency to these topics.

Interestingly, the topic "workshop" had the highest sentiment score, meaning that Dutch secondary education teachers expressed themselves the most positively about this topic. Furthermore, the topic with the third highest sentiment score was "education at school". A common factor that seems to offer an explanation for the high sentiment scores in these two topics might be the social interactions that play an important role in both topics. Social interactions, i.e. the process through which people exert reciprocal influence on one another during interpersonal contacts, can have a positive effect on both mental and physical health and can promote, among other things, happiness (Little, 2016; Tough et al., 2017; Quoidbach et al., 2019). Consequently, it can be expected that topics which involve social interactions lead to amplified use of positively valanced words. This was however not the case for each topic that involves social interactions. For instance, the topic "pupil-teacher interactions" surprisingly scored the lowest – although still positive – of all topics.

The nature of a social interaction largely influences its effect on individuals, meaning that on the one hand pleasant social interactions indeed lead to positive feelings such as happiness, feelings of belonging and meaning, while on the other hand unpleasant social interactions are predictive for negative feelings and behaviors, such as decreased happiness and avoidant behavior (Quoidbach et al., 2019). This may give an explanation to the fact that the topic "pupil-teacher interactions" scores the lowest, while all other topics that involve social interactions score among the highest in terms of sentiment. It cannot be concluded from the present study that social exchanges between teachers and pupils indeed involve less pleasant social interactions, but it is however interesting to further explore this topic in order to find substantial information that help understanding the reasons for which this topic does not score among the highest comparable to the other topics that involve social interactions.

Another common theme can be found in the topics “Literature about the Netherlands/ Dutch literature” and “Literature (books and stories)”, namely the focus on the school curriculum and learning material. Both topics rank high with respect to sentiment (second and fourth place), which may indicate teacher satisfaction with the curriculum and learning material. Moreover, this may also indicate that the teachers use their blogs to recommend material to other teachers. This interpretation could potentially extend to further incorporate *teaching strategies* when including the topics “Learning through play/ development” and “Assessment” and form an overarching theme of *teaching structure*. A good teaching structure, i.e. a high quality school curriculum combined with high quality student activities that promote the acquisition of knowledge, high quality assessment and high teacher cooperation, facilitates the ability for teachers to work together efficiently to increase student academic achievement, as well as enhances teacher job satisfaction (Wilard & Sithulisiwe, 2021; Cheon et al., 2020; Bates, 2019). The present findings do not necessarily suggest a certain presence of good teaching structure, they do however indicate the presence of high teacher satisfaction with the school curriculum and learning material, and in extension with teaching strategies, ergo a potentially satisfying teaching structure. Therefore, it is interesting for further research to inquire whether the teaching structure Dutch secondary education teachers are subjected to is truly as positive as the current study suggests and beyond to analyze what exactly brings about the teachers’ positive experiences regarding teaching structure in pursuance of promoting these elements such that both student academic achievement as well as teacher morale and satisfaction can further benefit.

Unsurprisingly, the SARS-CoV-2 pandemic, which reached the Netherlands in early 2020, also may have influenced the present study's findings. The two topics "Online education (digital sources) and "Online education (digital tools)" are indicative for the timing in which posts have been composed. In the past two years, the digital transformation of education was massively stimulated and digital tools as well as digital sources have become integral parts of modern-day education, which resulted in completely online education during the peaks of the pandemic and blended education during times when in-school education was (partially) possible (Jain & Lamba, 2021). This digital transformation is however not welcomed equally among teachers; some value the novel greater versatility of education, while others are concerned about the quality of education and prefer traditional means of teaching (Razkane et al., 2022). Keeping this wide range of attitudes towards online and blended education in mind, it can be reasonably assumed that the comparatively low sentiment scores among these two topics are due to the divisive nature of the issue. Even though the current paper's findings suggest that Dutch secondary education teachers in the current sample have a generally slightly positive opinion about online education, it is important to further inquire about issues and concerns that arise from the digital transformation for the sake of guaranteeing delivery of high-quality education as well as to promote teacher job satisfaction and morale in a changing work environment.

Topic modelling is a powerful method that allows for the detection of hidden semantic structures within large corpora of documents without human intervention, but it needs to be applied with caution. The famous expression "Garbage in, garbage out" from the field of computer science perfectly illustrates how important the quality of data as well as preprocessing procedures are. This expression refers to the concept that input data that is flawed or nonsense, i.e. garbage, always produces nonsense output. (Babbage, 1864; Mellin, 1957). Thus, the performance and output of any algorithm is always at most as good as the input data is and moreover, largely depends on the quality of data preprocessing steps taken prior to analysis.

During preprocessing, a lot of choices need to be made which can tremendously impact the interpretability of the results. Any kind of preprocessing choice – even the choice to not preprocess the data at all – leads to a result, which means that preprocessing fallacies and mistakes can go undetected. Therefore, it is of utmost importance to critically choose preprocessing steps that make sense with regards to the nature of the analyzed input as well as to the desired output. Even though the idea behind text mining techniques such as topic modelling is that human intervention is not needed, human supervision during the preprocessing phase is very much needed. One of such supervision tasks a researcher has is to check the fruitfulness of preprocessing steps through visual inspection.

A last important note is that the topic modelling algorithm does not understand the content of the topics it produces, therefore human reasoning is always needed in the interpretation of the topics. Unfortunately however, humans differ in their perceptions and interpretations, leading to potentially poor uniformity in topic interpretation and labeling. A guide that instructs researchers how to interpret topics can facilitate consistency and may thus promote inter-rater agreement.

The expression “garbage in, garbage out” is just as applicable to sentiment analysis as it is to topic modelling. A high-quality dictionary for positive and negative words is necessary to obtain meaningful sentiment scores. In an effort to illustrate the importance the validity of a well-grounded dictionary, it is worth considering the following hypothetical example: a sentiment analysis was conducted on a product review with a dictionary that is designed in a way that words are randomly assigned to the positive and negative lists. The algorithm worked and a result is received. The sentiment score suggests that the author has a positive stand on the product they reviewed, as the document includes many words from the positive list and only a few from the negative. Even though the result obtained in this example seems valid, it is nonetheless produced randomly and therefore meaningless.

While correct classification of words in a dictionary is imperative, accurately identifying the sentiment attached to a word can sometimes be complicated. Natural language is highly complex and can include polysemes, homonyms, synonyms, neologisms, idiomatic expressions, sarcasm and so forth, which makes it hard for algorithms to *understand* human language and to meaningfully identify the sentiment attached to a word. One such difficulty can be observed in the Dutch polyseme “echt” which is included in the positive list of the Dutch sentiment lexicon. “Echt” can however be translated into “really” and “genuine” depending on the context in which it is used. In the English counterpart of the Dutch Data Science Lab open-source multilingual sentiment lexicon, the word “genuine” appears in the positive list as well, while the word “really” does not appear in any of the lists, meaning that it is a neutral word (Chen & Skiena, 2014). Hence, some words can vary in their actual sentiment depending on the context in which they are used, and not all present-day algorithms are able to adjust sentiment scores to the context in which words are used.

As aforementioned, the present study used the Dutch Data Science Lab open-source multilingual sentiment lexicon. This lexicon was used as no other Dutch sentiment lexicon was available for Orange. During inspection of the SA’s output, some irregularities were observed. The sentiment lexica for both positive and negative terms included words that seemed neutral instead of indicating positive or negative sentiment. This discovery can be explained by the technique that was used to develop the lexicon. The algorithm that generated the Dutch sentiment dictionary was trained on Dutch Wikipedia articles which were priorly labeled as positive or negative in sentiment. Researchers identified a number of Wikipedia articles as conveying an either positive or negative sentiment through the use of sentiment analysis in English. It was assumed that an article that conveys a given sentiment in English language would also convey the same sentiment across other languages including Dutch. The same Wikipedia articles were then accessed in Dutch and analyzed by an algorithm that was provided with the information of whether an article was of positive or negative sentiment. Consequently, the algorithm generated lists of words that were associated with positive or negative sentiment and all remaining words were considered to be of neutral sentiment (Chen & Skiena, 2014).

One of Wikipedia's fundamental principles states that "*All encyclopedic content on Wikipedia must be written from a neutral point of view (NPOV), which means representing fairly, proportionately, and, as far as possible, without editorial bias, all the significant views that have been published by reliable sources on a topic*" and it is further clarified that nonjudgmental language should be applied as well as that stating opinions should be avoided (Sanger, 2001). In brief, Wikipedia does not promote the expression of opinions and endorses a neutral tone. Given that the algorithm was trained on a set of documents that is sparse in sentimental expressions, it should thus be considered that the algorithm may have overestimated the sentimental load of some of the terms and it may have inappropriately included some neutral terms into either of the sentiment lists.

Another limitation of the current study's approach to sentiment analysis is that the algorithm does not differentiate between the sentimental valency of different terms. While in natural language the two terms "good" and "fantastic" clearly express different levels of intensity of positive sentiment, the sentiment analysis applied in the current study does not take sentiment intensity into account, such that both terms are evaluated equally positive. The lack of a sentiment intensity indicator in the present study's sentiment analysis may have caused less variability across the sentiment scores obtained. Hence, results may be less interpretable. For further research it is therefore advisable to incorporate a sentiment analysis approach that takes the intensity of a sentiment into account for the purpose of obtaining more fine-grained results. Including such a sentiment analysis model would require a more elaborate lexicon which would not only list terms according to whether they are positive or negative but would rank the terms and give them numerical values which would indicate their sentimental intensity.

The present study explored the suitability of two already existing approaches, namely Sentiment-Based Topic Modelling and Aspect-Based Sentiment Analysis, that aim at combining topic modelling and SA in a meaningful way and proposed a new approach to analyze sentiments related to topics. A strength of the present study's approach is that it does not require manual labelling such as it is the case in ABSA, so that analyses of large text data sets are achievable. Another noteworthy aspect is that the proposed approach can analyze the sentiment scores for multiple topics within a single document, which is not the case in SBTM, where the sentiment score for a document is determined prior to topic modelling. A weakness of the proposed approach is that the sentiment score related to a given topic is estimated from the term probabilities of that topic, meaning that the SA's output largely depends on how well topic modelling has been performed. Hence, the proposed approach may be useful for other studies in which researchers choose to estimate sentiment scores for topics from data sets which are structurally similar to the present study's data set and for which the research question is similarly exploratory.

There are several software options available for researchers, all with their strengths and limitations. These software options range from subscription-only to open-source programming languages and they vary in their level of programming proficiency required and range from point and click interfaces, such as Orange, to programming languages which require the researcher to write their own scripts, such as in R. Interfaces such as Orange allow researchers with little programming experience to conduct various text mining analyses, including topic modelling and sentiment analysis. Orange offers the researcher a convenient interface with widgets on which they can simply click to conduct an analysis or obtain figures. Furthermore, there are open-source program extensions available that can be downloaded into the interface, and Orange offers tutorials on their website. A downside to Orange is that the researcher cannot see the scripts that are actually being run and therefore the researcher's advanced understanding of what the algorithm is exactly doing might be impeded. Furthermore, scripts cannot be exported from Orange and thus proofreading of the script or continuing the analyses in another interface is complicated. Open-source programming languages such as R do not have this limitation as the researcher can write, read and edit, as well as export scripts. Moreover, the use of programming languages such as R enables the researcher to make use of script libraries which can simplify the process of writing and understanding scripts. However, the learning curve required may be too steep for researchers with little programming experience. Both options have their strengths and limitations, and programming languages should be chosen according to the researcher's programming skills.

It was expected that teachers would discuss both topics that indicate high work morale and job satisfaction, potentially leading to improved mental health, and topics indicating low work morale and job satisfaction, potentially leading to diminished mental health. Contrary to expectations, this study did not find a single negatively valanced topic and thus no topic indicating the presence of teacher concerns. The lack of negative topics discussed by Dutch secondary education teachers may root in the stigmatization of mental health concerns and the desire to be socially accepted by others which may eventually result in concealment of negatively valanced topics. The present study's findings may thus be biased in favor of positively valanced topics through concealment of negative self-disclosures motivated by the desire to be socially accepted by peers which can be referred to as *social desirability bias*. Social desirability bias is a validity issue that is not unique to text mining techniques, or other qualitative and quantitative research methods, but it is however especially relevant for the present study’s text mining techniques (topic modelling and SA) as existing data from authors who published their thoughts on social media was used (Massara et al., 2012). While other qualitative research methods often guide the participants with an open question or statement, topic modelling and SA, which use pre-existing data from blogs, allow authors to freely choose what they write about and do not provide any guidance to participants. Each individual differs in the extent to which they are prone to socially desirable responding, therefore it can be suggested to use a higher number of different bloggers in future analyses in hopes that this may cancel out some of the socially desirable posts (Fordyce, 1956). This is however not an ideal solution, as it is impossible to know whether this measure is fruitful. Another solution would be to contact the authors and ask them to complete a social desirability scale, such as the Marlowe-Crowne Social Desirability Scale, and using the resulting insights as help to exclude data from authors who are concerned with social approval (Crowne & Marlowe, 1960).

Conclusion

The present study aimed at examining the usefulness of text mining techniques in the educational and psychological domain to gain insight into relevant issues teachers are facing at work and how they feel about these issues. Furthermore, through these insights it was aimed at highlighting important aspects that affect teachers’ everyday work in furtherance of promoting teacher morale and job satisfaction and therewith pupils’ academic achievements. Two types of text mining techniques, namely topic modelling and sentiment analysis, were combined to explore the topics Dutch secondary education teachers discuss in online blogs and the sentiments they hold about these topics.

Unexpectedly, the present study did not find any negatively valanced topics which limits the ability to offer an overview of *issues* Dutch secondary education teachers face at work. Nonetheless, the current research was able to display nine topics Dutch teachers from secondary education discussed in blog posts. Sentiment analysis inidcated that topics which involve social interactions were valued the most. In contrast, it was observed that teachers valued topics that involved online education less. Another noteworthy finding is that teachers generally have a positive attitude about topics that involve the teaching structure. Even though the present study was not able to provide a list of (negative) issues teachers are facing at work it was able to provide important insights into aspects that influence teachers' morale and job satisfaction.

Conclusively, the present study offers a holistic overview of what teachers do at work, what seems to be important to them, what they experience and how they feel about it. In addition, the present study illustrates the use of text mining as an efficient alternative to traditional methods such as questionnaires and interviews which can, when cautiously applied, produce a general overview of relevant topics from large corpora of pre-existing textual data.

References

Agichtein, E., Castillo, C., Donato, D., Gionis, A., & Mishne, G. (2008). Finding high-quality content in social media. *Proceedings of the International Conference on Web Search and Web Data Mining - WSDM '08*. https://doi.org/10.1145/1341531.1341557

Ahmad, K., Cheng, D., & Almas, Y. (2007). Multi-lingual sentiment analysis of financial news streams. *Proceedings of 1st International Workshop on Grid Technology for Financial Modeling and Simulation PoS(GRID2006)*. https://doi.org/10.22323/1.026.0001

Alurkar, A. A., Ranade, S. B., Joshi, S. V., Ranade, S. S., Sonewar, P. A., Mahalle, P. N., & Deshpande, A. V. (2017, November 1). *A proposed data science approach for email spam classification using machine learning techniques*. IEEE Xplore. https://doi.org/10.1109/CTTE.2017.8260935

Antons, D., Grünwald, E., Cichy, P., & Salge, T. O. (2020). The application of text mining methods in innovation research: current state, evolution patterns, and development priorities. *R&D Management*. https://doi.org/10.1111/radm.12408

Aravind, P. (2020). Tokenization In NLP. *Analytics Vidhya.* https://www.analyticsvidhya.com/blog/2020/05/what-is-tokenization-nlp/

Arunachalam, R., & Sarkar, S. (2013). The New Eye of Government: Citizen Sentiment Analysis in Social Media. *SocialNLP@IJCNLP*.

Babbage, C. (1864). Passages from the Life of a Philosopher. *Longman, Green, Longman, Roberts, and Green, London.*

Bates, A. W. T. (2019). Teaching in a Digital Age. *BCcampus.*

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2019). Uniting the Tribes: Using Text for Marketing Insight. *Journal of Marketing*, *84*(1), 1–25. https://doi.org/10.1177/0022242919873106

Blair, S. J., Bi, Y., & Mulvenna, M. D. (2019). Aggregated topic models for increasing social media topic coherence. *Applied Intelligence*. https://doi.org/10.1007/s10489-019-01438-z

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77. https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3, 993--1022. doi: http://dx.doi.org/10.1162/jmlr.2003.3.4-5.993

Cambria, E. (2016). Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems*, *31*(2), 102–107. https://doi.org/10.1109/mis.2016.31

Chen, Y., & Skiena, S. (2014). Building Sentiment Lexicons for All Major Languages. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, *2*, 383–389. https://doi.org/10.3115/v1/P14-2063

Cheon, S. H., Reeve, J., & Vansteenkiste, M. (2020). When teachers learn how to provide classroom structure in an autonomy-supportive way: Benefits to teachers and their students. *Teaching and Teacher Education*, *90*, 103004. https://doi.org/10.1016/j.tate.2019.103004

Chu, C.-Y., Park, K., & Kremer, G. E. (2019). Applying Text-mining Techniques to Global Supply Chain Region Selection: Considering Regional Differences. *Procedia Manufacturing*, *39*, 1691–1698. https://doi.org/10.1016/j.promfg.2020.01.271

Colleoni, E., Arvidsson, A., Hansen, L. K., & Marchesini, A. (2011). Measuring Corporate Reputation using Sentiment Analysis. *The 15th International Conference on Corporate Reputation*, New Orleans, United States.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349–354. https://doi.org/10.1037/h0047358

Dahal, B., Kumar, S. A. P., & Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*, *9*(1). https://doi.org/10.1007/s13278-019-0568-8

Danubianu, M. (2015). Step by Step Data Preprocessing for Data Mining: A Case Study. *International Conference on Information Technologies*, Bulgaria.

Danubianu, M., Popa, V., & Tobolcea, I. (2012). Unsupervised information-based feature selection for speech therapy optimization by data mining techniques. In *Proceedings of the Seventh International Multi-Conference on Computing in the Global Information Technology ICCGI* (pp. 29-39).

Demšar, J., Erjavec, A., Gorup, C., Hočevar, T., Možina, M., Toplak, M., Umek, L., & Zupan, B. (2013). Orange: Data Mining Toolbox in Python Tomaž Curk Mitar Milutinovič Matija Polajnar Anže Starič. *Journal of Machine Learning Research*, *14*, 2349–2353. http://jmlr.org/papers/volume14/demsar13a/demsar13a.pdf

De Wit, A., Rutgers, E. P., & Steijn, A. J. (2006). Het lerarentekort: een problem of een uitdaging. *Scriptie ter afronding van de studie Bestuurskunde Afstudeerrichting Arbeid, Organisatie, en Management.*

Eisenstein, J. (2016). Unsupervised Learning for Lexicon-Based Classification. https://arxiv.org/pdf/1611.06933.pdf

Etaiwi, W., & Naymat, G. (2017). The Impact of applying Different Preprocessing Steps on Review Spam Detection. *Procedia Computer Science*, *113*, 273–279. https://doi.org/10.1016/j.procs.2017.08.368

Fordyce, W. E. (1956). Social desirability in the MMPI. *Journal of Consulting Psychology*, *20*(3), 171–175. https://doi.org/10.1037/h0048547

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Ganesan, K. (2019). All you need to know about text preprocessing for NLP and Machine Learning. Kdnuggets.com. https://www.kdnuggets.com/2019/04/text-preprocessing-nlp-machine-learning.html

Given, L. M. (2008). Naturalistic Data. *The SAGE Encyclopedia of Qualitative Research Methods*. https://doi.org/10.4135/9781412963909.n279

Gupta, A., Dengre, V., Kheruwala, H. A., & Shah, M. (2020). Comprehensive review of text-mining applications in finance. *Financial Innovation*, *6*(1). https://doi.org/10.1186/s40854-020-00205-1

Han, J., Kamber, M., & Pei, J. (2012). Data Mining Trends and Research Frontiers. *Data Mining*, 585–631. https://doi.org/10.1016/b978-0-12-381479-1.00013-7

Harris, Z. S. (1954). Distributional Structure. WORD, *10*(2-3), 146–162. https://doi.org/10.1080/00437956.1954.11659520

Hooftman, W. E., Mars, G. M. J., Janssen, B., Vroome, E. M. M. de, & Bossche, S. N. J. van den. (2014). *Nationale Enquête Arbeidsomstandigheden 2014. Methodologie en globale resultaten*. TNO.

Huang, H., Long, R., Chen, H., Sun, K., & Li, Q. (2022). Exploring public attention about green consumption on Sina Weibo: Using text mining and deep learning. *Sustainable Production and Consumption*, *30*, 674–685. https://doi.org/10.1016/j.spc.2021.12.017

Huggins, J. R. (2004). The Selenium Browser Automation Project. *Selenium.* https://www.selenium.dev/documentation/

Humphreys, A., & Wang, R. J.-H. (2017). Automated Text Analysis for Consumer Research. *Journal of Consumer Research*, *44*(6), 1274–1306. https://doi.org/10.1093/jcr/ucx104

IBM Cloud Education. (2020a). What is Text Mining? *Www.ibm.com.* https://www.ibm.com/cloud/learn/text-mining

IBM Cloud Education. (2020b). What is Unsupervised Learning? *Www.ibm.com.* https://www.ibm.com/cloud/learn/unsupervised-learning

Jain, E., & Lamba, J. (2021). Management and Digitalization Strategy for Transforming Education Sector. *Emerging Challenges, Solutions, and Best Practices for Digital Enterprise Transformation*, 69–83. https://doi.org/10.4018/978-1-7998-8587-0.ch004

Jaiswal, S. (2021). Supervised Machine learning. *Javatpoint. Www.javatpoint.com.* https://www.javatpoint.com/supervised-machine-learning

James, T. L., Villacis Calderon, E. D., & Cook, D. F. (2017). Exploring patient perceptions of healthcare service quality through analysis of unstructured feedback. *Expert Systems with Applications*, *71*, 479–492. https://doi.org/10.1016/j.eswa.2016.11.004

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2018). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, *78*(11), 15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Kanan, T., Ayoub, S., Saif, E., Kanaan, G.G., Chandrasekarar, P., & Fox, E.A. (2015). Extracting Named Entities Using Named Entity Recognizer and Generating Topics Using Latent Dirichlet Allocation Algorithm for Arabic News Articles.

Kao, R. (2020). What is Text Mining? In *www.heavy.ai*. https://www.heavy.ai/technical-glossary/text-mining

Kargın, K. (2021). NLP: Tokenization, Stemming, Lemmatization and Part of Speech Tagging. *Medium.* https://medium.com/mlearning-ai/nlp-tokenization-stemming-lemmatization-and-part-of-speech-tagging-9088ac068768

Karssenberg, W. (n.d.). Edubloggers. *Trendmatcher Tussen ICT En Onderwijs*. Retrieved September 4, 2022, from https://www.trendmatcher.nl/p/edubloggers.html

Kaur, F., & Bhatia, R. (2016). Sentiment Analyzing by Dictionary based Approach. *International Journal of Computer Applications*, *152*(5), 32–34. https://doi.org/10.5120/ijca2016911814

Kim, K. H., Han, Y. J., Lee, S., Cho, S. W., & Lee, C. (2019). Text Mining for Patent Analysis to Forecast Emerging Technologies in Wireless Power Transfer. *Sustainability*, *11*(22), 6240. https://doi.org/10.3390/su11226240

Kotu, V., & Deshpande, B. (2019). Text Mining. *Data Science*, 281–305. https://doi.org/10.1016/b978-0-12-814761-0.00009-5

Krallinger, M., Vazquez, M., Leitner, F., Salgado, D., Chatr-aryamontri, A., Winter, A., Perfetto, L., Briganti, L., Licata, L., Iannuccelli, M., Castagnoli, L., Cesareni, G., Tyers, M., Schneider, G., Rinaldi, F., Leaman, R., Gonzalez, G., Matos, S., Kim, S., & Wilbur, W. J. (2011). The Protein-Protein Interaction tasks of BioCreative III: classification/ranking of articles and linking bio-ontology concepts to full text. *BMC Bioinformatics*, *12*(S8). https://doi.org/10.1186/1471-2105-12-s8-s3

Little, W. (2016). Introduction to Sociology (2nd ed.). *BCcampus.*

Lossio-Ventura, J. A., Gonzales, S., Morzan, J., Alatrista-Salas, H., Hernandez-Boussard, T., & Bian, J. (2021). Evaluation of clustering and topic modeling methods over health-related tweets and emails. *Artificial Intelligence in Medicine*, *117*, 102096. https://doi.org/10.1016/j.artmed.2021.102096

Mackenzie, N. (2007). Teaching Excellence Awards: An Apple for the Teacher? *Australian Journal of Education*, *51*(2), 190–204. https://doi.org/10.1177/000494410705100207

Martínez-Cámara, E., Martín-Valdivia, M. T., Ureña-López, L. A., & Montejo-Ráez, A. R. (2012). Sentiment analysis in Twitter. *Natural Language Engineering*, *20*(1), 1–28. https://doi.org/10.1017/s1351324912000332

Massara, F., Ancarani, F., Costabile, M., & Ricotta, F. (2012). Social Desirability in Virtual Communities. *International Journal of Business Administration*, 3(6). https://doi.org/10.5430/ijba.v3n6p93

Mayo, M. (2017). A General Approach to Preprocessing Text Data. *KDnuggets*. https://www.kdnuggets.com/2017/12/general-approach-preprocessing-text-data.html

McDonald, D., & Kelly, U. (2012). Value and benefits of text mining. In *Jisc*. https://www.jisc.ac.uk/reports/value-and-benefits-of-text-mining

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Mellin, W. (1957). Work With New Electronic 'Brains' Opens Field for Army Math Experts. *The Hammond Times.*

Merson, F., & Mary, R. (2017). A Text Mining Approach to Identify and Analyse Prominent issues from Public Complaints. *IJARCCE*, *6*(3), 54–59. https://doi.org/10.17148/ijarcce.2017.6313

Miller, W.C. (1981). Staff Morale, School Climate, and Educational Productivity. *Educational Leadership, 38*.

Milward, D. (2019, August 28). What is Text Mining, *Text Analytics and Natural Language Processing?* Linguamatics. https://www.linguamatics.com/what-text-mining-text-analytics-and-natural-language-processing

Miner, G., Elder, J., Hill, T., Delen, D., Fast, A., & Nisbet, R. A. (2012). The History of Text Mining. *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*, 3–27. https://doi.org/10.1016/b978-0-12-386979-1.00001-3

Ministerie van Onderwijs. (2012). Aanpak lerarentekort - Werken in het onderwijs. *Rijksoverheid.nl. Www.rijksoverheid.nl.* https://www.rijksoverheid.nl/onderwerpen/werken-in-het-onderwijs/aanpak-tekort-aan-leraren

Mohanty, H. (2015). Big Data: An Introduction. *Studies in Big Data*, 1–28. https://doi.org/10.1007/978-81-322-2494-5_1

Mujahid, M., Lee, E., Rustam, F., Washington, P. B., Ullah, S., Reshi, A. A., & Ashraf, I. (2021). Sentiment Analysis and Topic Modeling on Tweets about Online Education during COVID-19. *Applied Sciences*, *11*(18), 8438. https://doi.org/10.3390/app11188438

Mutanga, M. B., & Abayomi, A. (2020). Tweeting on COVID-19 pandemic in South Africa: LDA-based topic modelling approach. *African Journal of Science, Technology, Innovation and Development*, 1–10. https://doi.org/10.1080/20421338.2020.1817262

Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., & By, T. (2012). Sentiment Analysis on Social Media. *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. https://doi.org/10.1109/asonam.2012.164

Olaleye, S. A., Sanusi, I. T., & Salo, J. (2018). Sentiment analysis of social commerce: a harbinger of online reputation management. *International Journal of Electronic Business*, *14*(2), 85. https://doi.org/10.1504/ijeb.2018.094864

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135. https://doi.org/10.1561/1500000011

Prabhakaran, S. (2018). Lemmatization Approaches with Examples in Python. *Machine Learning Plus*. https://www.machinelearningplus.com/nlp/lemmatization-examples-python/#:~:text=Lemmatization%20is%20the%20process%20of

Quoidbach, J., Taquet, M., Desseilles, M., de Montjoye, Y.-A., & Gross, J. J. (2019). Happiness and Social Behavior. *Psychological Science*, *30*(8), 1111–1122. https://doi.org/10.1177/0956797619849666

Rajaraman, A., & Ullman, J. D. (2011). Data Mining. *Mining of Massive Datasets*, 1–17. https://doi.org/10.1017/cbo9781139058452.002

Razkane, H., Sayeh, A. Y., & Yeou, M. (2022). University Teachers' Attitudes Towards Distance Learning During COVID-19 Pandemic: Hurdles, Challenges, and Take-away Lessons. *European Journal of Interactive Multimedia and Education*, *3*(1). https://doi.org/https://doi.org/10.30935/ejimed/11436

Reed, C. (2012). Latent Dirichlet Allocation: Towards a Deeper Understanding. *University of Iowa*.

Remmits, Y. (2017). Finding the Topics of Case Law: Latent Dirichlet Allocation on Supreme Court Decisions.

Rempel, A. M., & Bentley, R. R. (1970). Teacher Morale: Relationship with Selected Factors. *Journal of Teacher Education*, *21*(4), 534–539. https://doi.org/10.1177/002248717002100414

Rosenholtz, S. J. (1985). Effective Schools: Interpreting the Evidence. *American Journal of Education*, *93*(3), 352–388. https://doi.org/10.1086/443805

Sabatovych, I. (2019). Use of Sentiment Analysis for Predicting Public Opinion on Referendum: A Feasibility Study. *The Reference Librarian*, *60*(3), 202–211. https://doi.org/https://doi.org/10.1080/02763877.2019.1595260

Sanger, L. (2001). Wikipedia: Neutral point of view. *Wikipedia.* https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view#:~:text=A%20neutral%20point%20of%20view%20neither%20sympathizes%20with%20nor%20disparages

Saxena, S. (2021). Tokenization and Text Normalization. *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2021/03/tokenization-and-text-normalization/#:~:text=Normalization%20is%20helpful%20in%20reducing

Shah, S. M., Lütjen, M., & Freitag, M. (2021). Text Mining for Supply Chain Risk Management in the Apparel Industry. *Applied Sciences*, *11*(5), 2323. https://doi.org/10.3390/app11052323

Singh, P., Dwivedi, Y. K., Kahlon, K. S., Sawhney, R. S., Alalwan, A. A., & Rana, N. P. (2019). Smart Monitoring and Controlling of Government Policies Using Social Media and Cloud Computing. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-019-09916-y

Stedman, C. (2019). What is text mining (text analytics)? *Search Business Analytics.* https://searchbusinessanalytics.techtarget.com/definition/text-mining

Straka, M., & Straková, J. (2017). Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. *Semantic Scholar.* https://doi.org/10.18653/v1/K17-3009

Tahmasebi, N., & Hengchen, S. (2019). The Strengths and Pitfalls of Large-Scale Text Mining for Literary Studies. *Samlaren: Tidskrift För Forskning Om Svensk Och Annan Nordisk Litteratur*, *140*, 198–227. http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-406938

Talabis, M. R. M., McPherson, R., Miyamoto, I., Martin, J. L., & Kaye, D. (2015). Analytics Defined. *Information Security Analytics*, 1–12. https://doi.org/10.1016/b978-0-12-800207-0.00001-0

Tough, H., Siegrist, J., & Fekete, C. (2017). Social relationships, mental health and wellbeing in physical disability: a systematic review. *BMC Public Health*, *17*(1). https://doi.org/10.1186/s12889-017-4308-6

Turney, P. (2022). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. https://aclanthology.org/P02-1053.pdf

Tseng, Y.-H., Lin, C.-J., & Lin, Y.-I. (2007). Text mining techniques for patent analysis. *Information Processing & Management*, *43*(5), 1216–1247. https://doi.org/10.1016/j.ipm.2006.11.011

Kuckartz, U. (2014). Qualitative text analysis: a guide to methods, practice et using software. *Sage.*

Vaishali, K., & Rashmi, A. (2018). Importance of Text Data Preprocessing & Implementation in RapidMiner. *Proceedings of the First International Conference on Information, Technology and Knowledge Management*, *14*, 71–75. https://doi.org/http://dx.doi.org/10.15439/2017KM46

Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Verma, T., & Nasib Singh, G. (2020). Email Spams via Text Mining using Machine Learning Techniques. *International Journal of Innovative Technology and Exploring Engineering*, *9*(4), 2535–2539. https://doi.org/10.35940/ijitee.d1915.029420

Wesslen, R. (2018). Computer-Assisted Text Analysis for Social Science: Topic Models and Beyond. *ArXiv*, *2*. https://doi.org/https://doi.org/10.48550/arXiv.1803.11045

Wilard, N., & Sithulisiwe, B. (2021). Strategies Used by Teachers for High Academic Performance in High Schools: Views from Head Teachers and Teachers. *International*

Journal of Humanities, Social Sciences and Education, *8*(4). https://doi.org/10.20431/2349-0381.0804003

Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2018). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, *60*(2), 617–663. https://doi.org/10.1007/s10115-018-1236-4

Zhang, C., & Han, J. (2019). Multidimensional Mining of Massive Text Data. In *Google Books*. Morgan & Claypool Publishers.

Zvornicanin, E. (2021). When Coherence Score is Good or Bad in Topic Modeling? *Baeldung on Computer Science*. Www.baeldung.com. https://www.baeldung.com/cs/topic-modeling-coherence-score#:~:text=We%20can%20use%20the%20coherence

Supplementary Materials

Table A

Data Preprocessing Example (English Version)

Unprocessed text	Preprocessed text
And when the last car drove away through the carwash, it happened. My teaching moment of 2020. What happened? I'll go back for that first […]	auto carwash drive away happen teaching moment happen go back first […]

Table B

The ten Most Frequently Used Terms in Blog Posts composed by Dutch Secondary Education Teachers (English Version)

Term	Frequency (Percentage)
pupil	2275 (2.75%)
good	1121 (1.36%)
school	940 (1.14%)
teaching/learning	929 (1.12%)
education	905 (1.10%)
class/lesson	607 (0.73%)
year	544 (0.66%)
Child/kid	503 (0.61%)
reading	483 (0.58%)
new	470 (0.57%)
Total words	82,594 (10.62%*)
Total unique words	18,789 (22.75%)

Note. The word frequency was assessed after preprocessing. * The ten most frequently used words make up 10.62% of all words used throughout the blog posts.

Table C

Nine Topics Dutch Secondary Education Teachers Discuss in Online Blogs (English Version)

Topic	Term (Probability)*	Sentiment Score
Workshops	workshop (0.0249), go (0.0248), school (0.0177), become (0.0158), one (0.0146), day (0.0135), make (0.0114), with (0.0113), nice (0.0111), year (0.0107), time (0.0106), each other (0.0103), many (0.0094), together (0.0087), come (0.0079), two (0.0079), conversation (0.0077), do (0.0073), really/genuine (0.0070), everybody (0.0066).	3.09
Literature about the Netherlands/ Dutch literature	find (0.0259), many (0.0209), know (0.0194), netherlands (0.0151), here (0.0137), see (0.0101), project (0.0096), dutch (0.0088), this (0.0081), world (0.0078), literature (0.0074), material (0.0069), offer (0.0066), water (0.0063), stand (0.0061), information (0.0059), image (0.0055) subject (0.0054), large/great (0.0053), land (0.0050).	2.81
Education at school	become (0.0320), education (0.0317), school (0.0204), teacher (0.0196), many (0.0172), go	2.40

	(0.0159), this (0.0141), good (0.0127), think (0.0096), pupil (0.0093), here (0.0078), own (0.0076), want (0.0073), course (0.0065), method (0.0062), let (0.0061), lay (0.0061), see (0.0059), skill (0.0057), ict (0.0055).	
Literature (books and stories)	read (0.0345), book (0.0333), go (0.0261), story (0.0182), write (0.0157), many (0.0144), year (0.0142), this (0.0123), let (0.0122), become (0.0118), text (0.0113), come (0.0107), young (0.0095), find (0.0088), dutch (0.0085), take (0.0083), tell (0.0075), literature (0.0075), want (0.0073), often (0.0067).	1.81
Assessment	pupil (0.0638), learn (0.0280), many (0.0256), give (0.0219), class (0.0211), make (0.0178), go (0.0174), oneself (0.0139), do (0.0130), one (0.0115), come (0.0110), work (0.0109), this (0.0100), good (0.0082), assignment (0.0082), teacher (0.0081), follow (0.0078), formative (0.0078), let (0.0076), look (0.0075).	1.77
Learning through play/ development	kid (0.0323), learn (0.0266), become (0.0173), make (0.0166), many (0.0137), important (0.0132), play (0.0127), this (0.0110), ask (0.0110), do (0.0097), development (0.0092), know (0.0090), human (0.0085), good (0.0084), possibility (0.0076), problem (0.0068), want	1.76

	(0.0066), think (0.0064), game (0.0064), develop (0.0060).	
Online education (digital sources)	digital (0.0245), many (0.0167), education (0.0154), media (0.0157), become (0.0153), teacher (0.0137), online (0.0137), research (0.0116), source (0.0112), school (0.0102), information (0.0095), internet (0.0090), oneself (0.0085), make (0.0084), open (0.0082), find (0.0079), use (0.0075), often (0.0073), social (0.0073), once (0.0072).	1.55
Online education (digital tools)	make (0.0357), become (0.0342), new (0.0223), use (0.0159), pupil (0.0153), see (0.0125), here (0.0125), open (0.0109), click (0.0106), share (0.0104), question (0.0099), work (0.0097), oneself (0.0086), film (0.0084), window (0.0083), video (0.0077), want (0.0071), tool (0.0067), possibility (0.0095), photo (0.0055).	0.99
Pupil-teacher interactions	pupil (0.0383), question (0.0257), good (0.0213), do (0.0206), give (0.0158), want (0.0149), come (0.0136), go (0.0126), teacher (0.0123), become (0.0119), find (0.0115), something (0.0105), group (0.0099), this (0.0097), sit (0.0089), say (0.0088), class (0.0086), look (0.0083), two (0.0083), know (0.0080).	0.76

Note. Blue color: positive term. Red color: negative term. All other terms are neutral terms.

*Probability of a term belonging to a topic represented as percentage of total terms within a topic.

Table D

Positive, Neutral and Negative Words and their Frequencies across Nine Topics Dutch Secondary Education Teachers Discuss in Online Blogs (English Version)

Sentiment	Term (Frequency)
Positive	give (165), good (1121), important (311), large/great (312), new (470), nice (250), offer (288), oneself (536), open (318), play (108), really/genuine (296), skill (187), source (281), want (470), work (289).
Neutral	Book (406), assignment (310), become (957), class (607), click (186), come (289), conversation (108), course (203), day (243), develop (187), development (196), digital (468), dutch (235), docent (729), each other (306), education (905), everybody (152), film (171), find (292), follow (413), formative (135), game (184), go (1137), group (190), here (541), human (284), ict (234), image (147), information (264), internet (186), know (196), land (102), learn (929), learning material (117), literature (136), make (833), many (711), material (122), media (236), method (337), Netherlands (193), often (349), once (307), one (8051), oneself (434), online (376), open (318), own (445), photo (120), possibility (244), project (189), pupil (2275), question (339), research (359), say (211), school (940), see (401), share (425), sit (80), social

	(181), something (347), stand (251), story (164), subject (199), take (413), teacher (238), tell (97), text (198), think (197), this (1152), together (216), tool (102), two (296), using (242), video (175), water (109), window (111), with (366), workshop (98), world (203), write (261), year (544), young (205).
Negative	Do (465), kid (503), lay (317), let (355), look (279), problem (151), read (317), time (315), use (318).