UNIVERSITY OF GDAŃSK

PHILOLOGY DEPARTMENT

Celina Czyszczoń

album number: 264718

Field of Study: English Philology

Specialty: Natural Language Processing

Sentiment and Emotion Analysis of the Black Mirror Series

Master thesis written
at the Institute of English and American Studies
under the supervision of
dr Joanna Redzimska

Gdańsk 2020

UNIWERSYTET GDAŃSKI

WYDZIAŁ FILOLOGICZNY

Celina Czyszczoń

numer albumu: 264718

Kierunek studiów: Filologia angielska

Specjalność: Ptrzetwarzanie języka naturalnego

Analiza sentymentu i emocji w serialu Black Mirror

Praca magisterska wykonana

w Instytucie Anglistyki i Amerykanistyki

pod kierunkiem

dr Joanny Redzimskiej

Gdańsk 2020

Table of contents

Introduction

Sentiment analysis is a field in computational linguistics studies focused on automatic detecting of the polarity of a given text and returning a score of how positive or negative the analysed text sample is. It is commonly used to analyse product reviews to determine users' opinions on products and their particular features, as well as on social media content to investigate current attitudes of the society towards particular events, political parties or persons, for example, to predict the results of presidential elections.

This study is focused on applying sentiment analysis methods to research larger text samples from the domain of films and television series. The work offers a comprehensive overview of the processes and theories applied in sentiment analysis approached from a linguistic perspective. The first chapter discusses how corpora can be used for a linguistic research, providing examples of widely used corpora, explaining how new corpora can be compiled and describing possible applications of a corpus. In the subsequent chapter, axiology and sentiment analysis are discussed to explain why a positive-negative scale is used to describe semantic aspects of language, how axiological charge can be elicited for individual lexical items and how the sentiment of whole sentences and documents is calculated by the lexicon-based sentiment analysis engines. The research chapter includes an overview of recent projects in the field of sentiment analysis, methodology section with a detailed description of the tools used for the research – VADER Sentiment Analysis and Regressive Imagery Dictionary, as well as findings of the study on a corpus compiled of transcripts of episodes of the Black Mirror series. The study includes the examination of the sentiment expressed in the series as computed by a sentiment analysis engine, as well as calculated based on the frequencies of emotion words. The analysis of the results of the research is followed by suggestions on how the process of sentiment analysis on television series could be improved and ideas of possible applications of such analysis for commercial use in the film industry.

1. Using corpora for linguistic research

1.1 Introduction

This chapter focuses on using corpora in linguistic research. Firstly, it defines the notion of a corpus, explains how a corpus can be compiled and provides examples of general corpora that are available on-line. Then, it discusses the practical applications of corpora in computational linguistics. Finally, the chapter presents corpus linguistics as a methodology employed by researches of many different fields of linguistics and shortly discusses its advantages and limitations.

1.2 Defining a corpus

According to Lindquist (2009,) the word corpus (meaning 'body' in Latin) has been previously used to refer to the total amount of works written by a given author or a certain mass of texts e.g. "The Shakespeare corpus" and has been also used to refer to collections of linguistic examples written down on slips of paper, which Kennedy (as quoted in Lindquist 2009) refers to as 'pre-electronic corpora'. At present time, the word corpus is usually used synonymously with electronic corpus, i.e. "a collection of texts which is stored on some kind of digital medium and used by linguists to retrieve linguistic items for research or by lexicographers for dictionary-making" (Lindquist 2009: 3). This definition entails that corpora differ from dictionaries, which, according to Oxford Dictionary, are books or electronic resources which list the words of a language and provide their meaning, or corresponding words in another language, often also giving additional information such as pronunciation, origin or usage. A definition provided by the Expert Advisory Group on Language Engineering Standards (EAGLES), however, allows such a list to also be considered as a corpus, as it states that "a corpus can potentially contain any text type, including not only prose, newspapers, as well as poetry, drama, etc., but also word lists, dictionaries, etc." (https://www.cs.vassar.edu/CES/CES1-0.html)

For the purpose of this work, the term corpus will be used to denote a collection of texts or parts of texts upon which further research can be conducted. Word lists and dictionaries will be considered as data that may be a result of the corpus analysis or resources that can be used to compute the frequencies of their elements to provide statistical analysis of their occurrence within the corpus (e.g. a list of words expressing negative sentiment can be compared to texts within a corpus to find those negative words which occur in it and to compile a list of most common negative words for those particular texts).

1.3 Compiling a corpus

Planning the construction of a corpus and collecting the data should be guided by the ultimate use of the corpus – the research that is to be conducted on its basis. Multi-purpose corpora should be balanced in terms of the number of text samples from different genres, the proportion of spoken and written language, age, gender and social background of the speakers to cover for different varieties of a language (Meyer 2004). In order to conduct a thorough analysis of the language of a specific variety, researchers often choose to compile their own, specialized corpora, which is restricted to that particular language field.

To collect the data needed for the corpus, scholars can either scan written documents, use literature available on-line or harvest the web to find text samples that they would use to compile the corpora (Michael Wilkinson 2006). Michael Wilkinson (2006) highlights that while choosing the texts, linguists should make sure that the samples are written by native speakers (otherwise, they could include non-idiomatic expressions). Moreover, he claims that it is better to use full texts than just extracts to ensure that the corpus covers the language used in each section of the text. He also advises using recent texts to guarantee that the information retrieved from the corpus is up-to-date.

After collecting the data, texts can be uploaded to a software toolkit such as AntConc, WordSmith or Sketch Engine to allow concordancing, finding word patterns and searching for other linguistic information within the collection of texts. Such text analysis software usually requires plain .txt files, so the files need to be converted to meet these requirements.

1.4 Examples of available corpora

Results retrieved from a specialized corpus compiled by a linguist may be compared to the already existing, more general and freely accessible corpora. This section provides examples of contemporary corpora available on-line which can be used for that purpose.

1.4.1 The Brown Corpus

The Brown Corpus can function as an example of a “balanced” written corpus because “it is divided into 2,000-word samples representing different types (or genres) of written English, including press reportage, editorials, government documents, technical writing, and fiction” (Meyer 2004). Such structure of the corpus permits a general study of a language as a whole but also it allows for research on particular genre and as well as it facilitates making comparisons between different kinds of texts.

1.4.2 The British National Corpus (BNC)

According to Lindquist (2009), the British National Corpus (BNC) is the result of a big commercial and academic project run by Oxford University Press Longman, Chambers, the British Library, Oxford University and Lancaster University. Its main aim is to give a balanced representation of British English. It is divided into spoken (10%) and written part (90%) which consists of fiction (25%) and non-fiction (75%). The samples used in the corpus are relatively big, between 40,000 and 50,000 words. This is particularly important, as texts tend to have different vocabulary and structures used in the beginning and at the end of a work. Because of how big this corpus is, of how carefully it has been compiled and how easily accessible it is through free interfaces it is very popular among linguists.

The spoken component of the British National Corpus (BNC) is a good example of a spoken corpus. It was compiled in the early 1990s and consists of recordings of lectures, homilies, radio shows and other public events (5 million words) and tapes containing everyday conversations which people recorded using portable tape recorders (Lindquist 2009). The speakers were of different age, came from different regions of England and had various social backgrounds, so the samples are demographically evenly distributed (Lindquist 2009).

1.4.3 The Corpus of Contemporary American English (COCA)

The American counterpart for BNC is the Corpus of Contemporary American English (COCA), which has been created in 2008 by Mark Davies at Brigham Young University. It is widely approved by researchers, available for free (with a limited number of queries per day), has a user-friendly interface and therefore may be used as a tool for interesting investigations. (Lindquist 2009)

“The corpus contains more than 560 million words of text (20 million words each year 1990-2017) and it is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts.” (https://www.english-corpora.org/coca/) It is arguably the most commonly used corpus of English, and as it is related to other corpora compiled by Mark Davies (e.g. The TV Corpus, The Movie Corpus, Corpus of American Soap Operas, Corpus of US Supreme Court Opinions) it can be used for comparative research of variations in English (https://www.english-corpora.org/coca/).

1.4.4 The Oxford English Dictionary (OED) used as a corpus

Sometimes electronic versions of dictionaries contain excerpts from authentic texts as illustrations of language use and therefore can be used by researchers as corpora. An example of such dictionary containing contemporary texts is the Oxford English Dictionary (OED) (Lindquist 2009).

1.4.5 Parallel corpora and Polish corpora

The examples of corpora provided above consist only of texts in English; however, there are also corpora containing two or more languages. Typically, they include original texts and their translations and are useful for researchers interested in comparative linguistic and translation studies (Lindquist 2009). Those are called parallel corpora and *Paralela* is an example of such a corpus consisting of Polish-English and English-Polish translations. It includes text collections from large open source parallel corpora such as the *European Parliament Proceedings*, manually aligned scientific texts from *Academia* and Center for Eastern Studies, as well as literary classics in Polish-English and English-Polish translations (Pęzik 2016). The most popular corpus of Polish texts is the National Corpus of Polish (Narodowy Korpus Języka Polskiego – NKJP), which contains over fifteen hundred million words (nkjp.pl) and can be browsed using an advanced search engine *PELCRA NKJP*, available at http://www.nkjp.uni.lodz.pl/.

1.4.6 The web used as a corpus

Yet another source of research data is the Internet, which provides the researcher with an enormous amount of text samples, which are crucial for studies on morphological productivity or collocations as they require very large databases (Lindquist 2009). It is also the source containing most up-to-date language and often the only one to provide examples of newly-coined words and expressions. Searching the web for specified linguistic queries is possible through advanced settings in Google search engine.

1.5 Possible applications of corpora

1.5.1 Concordances, frequencies and KWIC

Thanks to electronic corpora it is possible to quickly access reliable information such as concordances and frequency figures. Conventionally, concordances are lists of all contexts in which a word occurs in a given text. In linguistics, the data are most often shown by means of keyword-in-context (KWIC) concordances which present the keyword in the middle of approximately one line of context (Lindquist 2009).

(LAUGHTER) OBAMA: Still , I guess Governor Romney is feeling	good	about things , because he took a few hours off the other
if you add 10,000 advisers , are you not just sending	good	after bad here ? GRAHAM# Well , the surge did work .
some days were bad . But actually , every day was	good	and bad because you 'd meet with all these doctors and sc
of course , reject arguments that they are an obstacle to	good	and economical government , or that restricting their pow
Rock growled . " It 'll just scour your old butt	good	and raw when it comes out the other end . " "
while longer . At last he said , " Illusory .	Good	as any , illusory . That will do . " I lay
Senior Citizens ' Home and all the rest , he was	good	as dead . There are some things that do n't bear thinking
of talk , Turner . Our budgeting system is just as	good	as the Soviets ' . # Meanwhile , chalk one up for
OFF BY IMPERFECT . # An ugly tomato tastes just as	good	as (sometimes better than) a pretty one . But avoid
miles from Stanford , and it 's Deathville . Either no	good	at all or cold , steep , spooky Deathville . It 's

Figure 1.1 Example of KWIC search results in COCA
Source: https://www.english-corpora.org/coca/

1.	To super, na dodatek jeszcze mnie wywalą z roboty. Dzwonię, dzień	**dobry**	, przepraszam, ale musiałem pilnie załatwić sprawę w urzędzie
2.	chyba uda mi się ją spławić. Może za dwa tygodnie? Tak, to chyba	**dobry**	pomysł. I kombinuję, że przecież mogę zmienić numer komórki. Tamtego
3.	albo gdzieś, zapewne po kolacji przy świecach, nocą zapowiada się	**dobry**	seks albo, jeszcze lepiej, kolacja od razu była połączona z seksem,
4.	napije, robi się strasznie upierdliwy. O, to będzie chyba ten. Halo,	**dobry**	wieczór, przepraszam, że tak późno, ale mam pilną sprawę do Arka, bo
5.	nich przez dziesięć minut nie można dodzwonić. No, wreszcie. Dzień	**dobry**	, dzwonię z expressu, z kroniki wypadków, kolega już dzwonił, no tak,
6.	- Butelka od firmy - powiedział. - Dzisiaj jest	**dobry**	dzień. Czasem się przez tydzień tyle nie zarobi.
7.	Chłopaki chyba uważali, że to	**dobry**	pomysł, bo się roześmiali.
8.	-	**Dobry**	wieczór pani. Można się przysiąść?
9.	to mógł kopa dostać a nie worek cementu. A tutaj nikt z rachunków	**dobry**	nie był. Stary też nie mógł podskoczyć, jak za rękę nie złapał.
10.	i się ukłonił a gadali, że on teraz takie panisko chociaż zawsze był	**dobry**	chłop. Jeszcze pamiętam jak żeśmy razem wiosną na pstrąga chodzili,

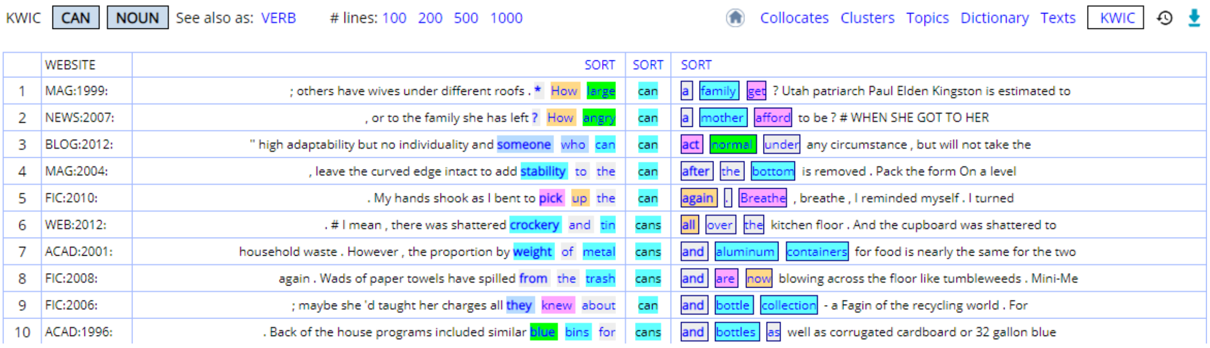
Figure 1.2 Example of KWIC search results in NKJP
Source: http://www.nkjp.uni.lodz.pl/

Figure 1 presents the results of KWIC search for the word *good* in the Contemporary Corpus of American English (COCA) while figure 1.2 presents the search for its Polish counterpart *dobry* in the National Corpus of Polish. It can be observed that the COCA search engine while displaying the keyword in context automatically colour-codes the words in its nearest proximity according to their part-of-speech tags; adjectives are marked with green, verbs with pink, nouns with blue etc.

Frequencies of given words or structures may be counted automatically and easily retrieved from large collections of texts thanks to electronic corpora. This data can be later used by linguists to compare and observe differences between geographical varieties genres of a language, spoken and written registers or text from different periods of time (Lindquist 2009).

1.5.2 Part-of-speech tagging

In order to obtain more reliable conclusions of a study, words in a corpus may be tagged in relation to which part of speech (word class) they function as in a given context. Having a corpus annotated in such manner allows a researcher to search for the word ‘can’ for example, only when it is used as a noun and ignoring the instances where the word appears as a verb (Lindquist 2009). Figure 1.3 presents an example of search with such a parameter in COCA. As can be observed, the currently analysed part of speech is displayed in the top left-hand corner with the option of swithich to other parts of speech in which the word occurs in the corpus.



	WEBSITE	SORT	SORT	SORT
1	MAG:1999:	; others have wives under different roofs . * How large	can	a family get ? Utah patriarch Paul Elden Kingston is estimated to
2	NEWS:2007:	, or to the family she has left ? How angry	can	a mother afford to be ? # WHEN SHE GOT TO HER
3	BLOG:2012:	" high adaptability but no individuality and someone who can	can	act normal under any circumstance , but will not take the
4	MAG:2004:	, leave the curved edge intact to add stability to the	can	after the bottom is removed . Pack the form On a level
5	FIC:2010:	. My hands shook as I bent to pick up the	can	again . Breathe , breathe , I reminded myself . I turned
6	WEB:2012:	. # I mean , there was shattered crockery and tin	cans	all over the kitchen floor . And the cupboard was shattered to
7	ACAD:2001:	household waste . However , the proportion by weight of metal	cans	and aluminum containers for food is nearly the same for the two
8	FIC:2008:	again . Wads of paper towels have spilled from the trash	cans	and are now blowing across the floor like tumbleweeds . Mini-Me
9	FIC:2006:	; maybe she 'd taught her charges all they knew about	can	and bottle collection - a Fagin of the recycling world . For
10	ACAD:1996:	. Back of the house programs included similar blue bins for	cans	and bottles as well as corrugated cardboard or 32 gallon blue

Figure 1.3 KWIC search results for the word *can* as a noun in COCA

Source: https://www.english-corpora.org/coca/

Small corpora can be tagged manually by an individual or a team of researchers trained to tag according to a certain standard. Larger corpora require automatic taggers – computer programs responsible for tagging, which rely on different methods such as grammatical rules and probabilities. The taggers also make use of lexicon of common words and their tags. The classification made by those programs is very precise as the tagsets they use are quite big. For example, the CLAWS (Constituent Likelihood Automatic Word-tagging System) which is used for the BNC uses as much as 160 tags (Lindquist 2009).

1.5.3 Parsing

Parsing is the process of "assigning a syntactic analysis to the corpus" (Lindquist 2009: 48). It is a more complex task than POS tagging as its application may vary across linguistic theories. Phrase structure analysis, where phrases and clauses are labelled according to their function such as adjective phrase, adverbial phrase, past participle clause, etc., is the most popular type of parsing. Corpora annotated in such a manner are called treebanks, as the structures can be represented as tree diagrams. According to Lindquist (2009), the treebank which is most widely used by linguists is the parsed version on ICE-GB and its interface ICECUP, which is shown in figure 1.4. Lindquist (2009) also claims that automatic parsing is particularly important in natural language processing where it can be applied in developing machine translation or speech to text conversion.

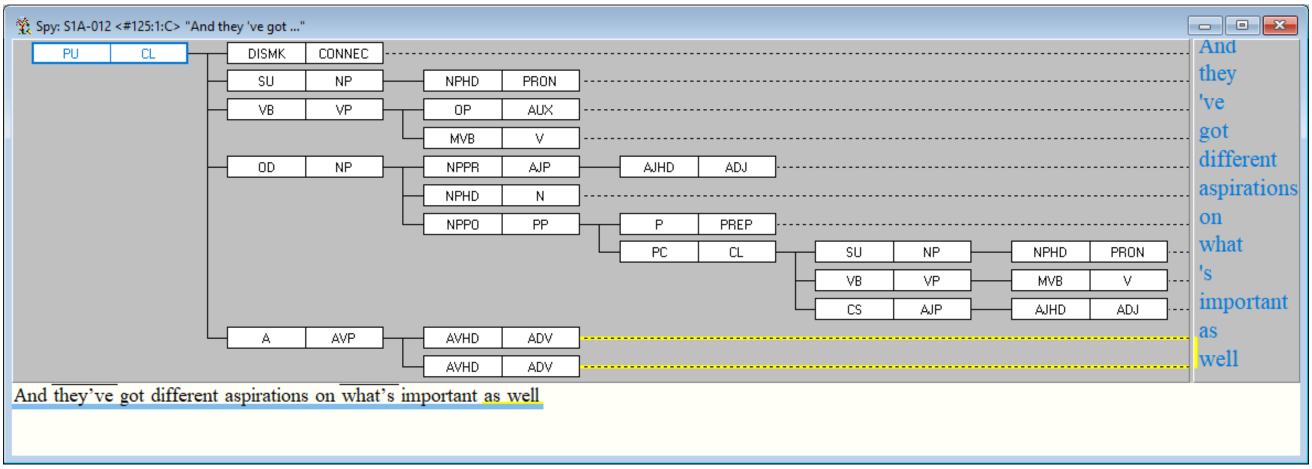


Figure 1.4 An example of a tree diagram from the ICE-GB corpus

1.5.4 Using corpora in discourse analysis

Even though 'standard' corpora such as BNC cannot be directly used for the discourse analysis, as it usually does not contain sufficient amount of data within the specific area of language that a linguist may want to examine, there is a connection between discourse analysis and corpus linguistics (McEnery & Wilson 2001). Firstly, the same computer-aided methodology as the one that is used in corpus linguistics is often applied in the discourse analysis. Thanks to that, linguists can analyse, for instance, word frequencies or concordances in texts they are studying (McEnery & Wilson 2001). Moreover, McEnery and Wilson (2001) claim that general corpora may function as control data in the discourse analysis. The findings from a study of a given text can be confronted with the data from a standard corpus to determine how they occur in a language as a whole.

1.5.5 Corpus-driven approach to sentiment analysis

Compiling and analysing a corpus of texts from a specific domain is crucial for developing a tool for sentiment analysis of similar data. For instance, in a study conducted by Stuart, Botella and Ferri (2016) the linguistic analysis of a corpus consisting of patients' reviews was used to develop and test software to conduct sentiment analysis on the said corpus. Such study of a corpus enables the researchers to come up with lists of positive, negative and intensifying words and expressions used in a specific domain (Stuart et. al., 2016). Then with applying more advanced linguistic procedures such as POS tagging and parsing it is possible to develop a lexicon and rule-based tool for sentiment analysis of texts within the investigated domain.

1.6 Corpus linguistics

Corpus linguistics differs from other linguistic studies, as it is not a particular linguistic paradigm (Meyer 2004) but rather a methodology including a variety of related methods which can be applied by scholars in various areas of language research (Lindquist 2009).

However, it is often associated with a particular outlook on language, which proposes that rules of a language are based on its use and that studying it is a good way of examining the workings of a given language (Meyer 2004). Therefore, we may differentiate between the methodological approaches taken within the same area of research and, for example, distinguish corpus-based syntax from non-corpus-based syntax. (McEnery & Wilson, 2001)

There is also a distinction between corpus-driven and corpus-based linguistics. The first one, proposed by Tognini-Bonelli in 2001 (as cited in Lindquist 2009), suggests that a researcher should not have any predetermined ideas when starting to work with a corpus, yet inductively arrive at an analysis. On the other hand, corpus-based linguistics uses corpora to test hypotheses based on already existing theories (Lindquist 2009).

1.6.1 Advantages of corpus linguistics

The advantages of corpus linguistics presented by Svartvik in 1992 (as cited in Lindquist 2009) mostly concern the quality of data that it uses for research. Firstly, the data is objective and can be verified by other researchers. Also, the same data can be used by various researchers with no need of compiling their own database each time, as computerized corpora can be accessed by scholars around the world. Moreover, data of this type is crucial for research on differences between dialects and styles as it provides the frequency of occurrence of linguistic items. It also provides vital information for applied areas of enquiry such as the machine translation or the speech synthesis.

1.6.2 Limitations of corpus linguistics

Svartvik (as cited in Lindquist 2009), however, also highlights the corpus needs to be of good quality in order for the findings to be relevant. He also deems careful manual analysis to be vital for the research, as he finds a bare presentation of the numbers highly insufficient.

Lindquist (2009) also points out that because of the infinite number of possible sentences in a language, corpora will never be big enough to cover everything within a speaker's knowledge of a language and therefore "the intuition of a native speaker will always be needed to identify what is grammatical and what is not" (Lindquist 2009). Apart from that intuition, Lindquist (2009) claims that a theory of language is always necessary for the researcher to know what to search for and how to describe their findings

Another criticism against corpus linguistics (especially raised by generative linguists) is that some of the discoveries based on corpora may seem to be trivial, as for example "that the fact that *I live in New York* is more frequent than *I live in Dayton, Ohio* does not have any relevance for linguistic theory or description" (Lindquist 2009: 9).

1.7 Conclusions

Corpora are useful for different kinds of linguistic research, including the discourse analysis and the sentiment analysis. A specified field of language can be investigated with the use of custom-made corpora consisting of texts relevant for that field. Such collection of texts can be compiled by means of exploiting the web resources and uploading retrieved data to one of the freely available software toolkits for concordancing and text analysis. There are many types of general corpora available on-line, which may function as control data that can be compared to the findings retrieved from the study of a specialized corpus. With the application of part-of-speech tagging and parsing, the corpus can be searched for even more complex linguistic queries and provide more accurate results.

2. Axiology and sentiment analysis

2.1 Introduction

This chapter focuses on the study of values and their reflection within natural languages. Firstly, it discusses the link between axiology and language and explains why a positive-negative opposition is useful for describing semantic aspects of language. Secondly, the concept of axiological charge is defined and examples of how it can be elicited for particular lexical items are provided. Then, the following section describes how sentiment analysis models based on lexicons and rules can automatically determine the polarity of a given text. The last section is devoted to possible applications of sentiment analysis.

The sections of this chapter which deal with axiology are primarily based on Tomasz Krzeszowski's book *Angels and Devils in Hell: Elements of Axiology in Semantics* published in 1997, which remains the most comprehensive work on the subject of axiology from the linguistic perspective. The work offers a profound insight into multiple aspects of the study of value represented in a language and therefore it was chosen as the main source for the research on the subject.

2.2 Axiology and language

Axiology, which is also referred to as the theory of value, is defined by Britannica (Britannica.com) as "the philosophical study of goodness, or value, in the widest sense of these terms". As far as we can tell from the evidence available, the systems of values which govern the world of the humankind, have always been expressed and reflected through language (Alba-Juez & Thompson 2014). Therefore, both philosophers and linguists examine natural languages in search of representations of value. Leibniz (as cited in Alba-Juez & Thompson 2014: 3) claim that language is "the best mirror of the human mind" and therefore suggested that studying language in depth can give a significant insight into the understanding of human reasoning. Ludwig Wittgenstein (as cited in Alba-Juez & Thompson 2014), developed a theory which supposed that language mirrored reality and presented an account of the relation between words, sentences and propositions to the real situation or event to which they referred.

The study of values is also popular in linguistic research, as valuations and categorizations "directly manifest themselves in language" (Krzeszowski 1997: 15). According to Puzynina (as cited in Krzeszowski 1997:15) linguistics and axiology meet in two domains: “the domain of valuative words and their meaning” and “the ways in which valuations are expressed in a language and in the structure of texts”. However, Puzynina (as cited in Krzeszowski 1997) also notices that while axiologists concern themselves with *what* concepts such as ‘value’, ‘beauty’, ‘goodness’ etc are, linguists are more concerned with *how* the meanings of the corresponding words are understood by people in a given language and at a given time.

2.3 Using a positive-negative scale for describing semantic aspects of language

While the models of language based on the truth conditional logic, such as transformational-generative grammar, use the ‘true-false’ distinction to describe semantic aspects of language, Krzeszowski (1997) argues, that it is the ‘positive-negative’ opposition which should be the most central in semantics and highlights that:

1) The largest part of general variance (33%) in langauge is connected with the evaluation on the ‘good-bad’ scale (Osgood, Suci & Tannenbaum as cited in Krzeszowski 1997)
2) The distinction between good and bad is the first categorization that infants learn to recognize and implement. It precedes the distinction between ‘ugly’ and ‘beautiful’ or ‘true’ and ‘false’.
3) While the ‘true-false’ distinction applies only on the sentence level, ‘positive-negative’ opposition is relevant for both individual words and to their combinations.
4) The ‘postive-negative’ scale is the most general one, as the concepts of good and bad can be used to replace more specific evaluative adjectives (e.g. beauty can be referred to with the expression *good* looks)

2.4 The domain of values

Krzeszowski (1997) describes the domain of values as two dimensional. The vertical (UP-DOWN) dimension of the domain corresponds to the hierarchy of values from the lowest ones to the highest ones and (in this model) it is directly derived from the Great Chain of Being. That supposes that “the world as experienced by human beings consists of things arranged in a certain hierarchical order” (Krzeszowski 1997: 64); these things fall into 5 categories, which are also regarded as consecutive levels of hierarchy (from the highest ones to the lowest ones): God, humans, animals, plants and inorganic things. The horizontal (LEFT-RIGHT) level corresponds to scales of values at each level of that hierarchy between the negative (-) and the positive (+) pole. The scale is a continuum between two extremes, with a negative pole situated on the left, a positive pole situated on the right and a neutral zone between them.

Each concept can be assigned a category in the vertical dimension and a value on the horizontal scale, which can be used as two coordinates describing a concept's position in the domain of values. In this work, the focus is more on the horizontal dimension, corresponding to the axiological charge of a concept.

2.5 Axiological charge

As explained in the previous subchapter, the axiological charge is the position of an entity on a positive-negative scale. Krzeszowski (1997) argues that all lexical items can be assessed on an axiological scale and distinguishes three types of techniques which enable eliciting the axiological charge of a lexical item:

1) those which elicit absolute values of particular lexical items
2) those which elicit values of lexical items in comparison with some other lexical items
3) those which elicit actual values of lexical items in specific contexts.

2.5.1 Eliciting absolute values of particular lexical items

Krzeszowski (1997) describes two methods which can be used to elicit the axiological charge of a lexical item. Those are direct rating and sentence formation.

Direct rating is a technique commonly used in psychological and sociological questionnaires in which respondents need to evaluate a concept by choosing one of the suggested options. These options consist of two polar terms and a certain number of intermediate options between them. The number of these intermediate options may vary depending on the needs of investigators. Some researchers use only three-grade scales (with one middle element), but five-, seven- and ten-grade scales can be used in that type of questionnaires if there is a need for more precise evaluation (Krzeszowski 1997). With a larger scale there is a need to use quantifiers such as 'extremely', 'very', 'quite', 'slightly', to describe the level of intensity of polar terms. For example, Krzeszowski (1997) conducted an experiment in which he asked a group of 100 respondents to evaluate two concepts: smile (uśmiech) and grimace (grymas) on a five-grade scale with polar terms from very pleasant (bardzo przyjemny) to very unpleasant (bardzo nieprzyjemny). On that scale ‘0’ was used to indicate lack of decision. The results obtained in that survey are presented in Figure 2.1

	very pleasant	pleasant	0	unpleasant	very unpleasant
Smile	83	17	-	-	-
Grimace	-	-	9	58	33

Figure 2.1 Results of direct rating survey

Results of such questionnaires allow us to calculate the mean absolute axiological charge of lexical items (regardless of context in which they might be used).

Another technique to elicit the absolute axiological charge of lexical items is sentence formation. In this method, participants are asked to form sentences with a given lexical item. While doing so, participants tend to reveal their attitudes by explicitly condemning or applauding the concept which the item describes, expressing their subjective like or dislike for the concept or identifying the item in a definitely positive or negative concept. The result of such survey is a ratio of the positive to negative sentences created with a given concept.

2.5.2 Eliciting values of lexical items in comparison to other lexical items

Another method to elicit the axiological charge is a survey in which participants are asked to order given lexical items in relation to the intensity of a given property. For example, Krzeszowski (1997) conducted a survey in which he asked participants to arrange six nouns from the most to the least beautiful one. Each position had a proportional number of points ascribed to it; from six points for the top position to one point for the last position. Then, a ranking of those nouns was compiled by calculating the number of points acquired by each item.

2.5.3 Eliciting actual values of lexical items in specific contexts

Axiological charge can also be estimated by calculating the instances of positive and negative context in which they appear. Arguably, the most reliable set of data to examine are the examples provided by a monolingual dictionary (Krzeszowski 1997). For instance, six out of seven examples provided by SJP (Słownik języka polskiego – the Dictionary of the Polish Langauge) (as cited in Krzeszowski 1997) to illustrate the word uśmiech (smile) present the word in a positive context and therefore the item can be evaluated as positive. However, as the examples may be ambiguous and therefore difficult to evaluate, Krzeszowski (1997) advises this method only to provide supplementary evidence to the findings retrieved from more independent assessments.

Theoretically, a balanced corpus can be used, for example, as a source of contexts within which a lexical item occurs. However, in this case, the big amount of data is a serious disadvantage, as it would be a tedious and time-consuming task to manually assess whether a given lexical item is used in a positive or negative context in each of the examples. For instance, National Corpus of Polish (Narodowy Korpus Języka Polskiego NKJP) contains 5531 examples of paragraphs in which the word smile (uśmiech) occurs (http://www.nkjp.uni.lodz.pl/). This set can be downsized by either randomly choosing a smaller number of examples or by narrowing to a specific category or time frame; such process, however, seems to negate the purpose of using corpora and is likely to result in unreliable findings. The process could be facilitated by choosing a corpus which is already annotated for sentiment, yet this cannot be considered a perfect solution either. Firstly, such corpora are rare and usually only contain texts from a certain category (for example *An Annotated Corpus for Sentiment Analysis in Political News* or *Awais Athar - Citation Sentiment Corpus*). Moreover, it does not seem reliable to assess the value of a particular linguistic item basing on polarity scores assigned to entire sentences or even paragraphs; for that purpose, one would need a corpus with sentiment values assigned specifically to phrases containing a given keyword. Unfortunately, such corpora are not available.

2.6 Sentiment Analysis

Sentiment analysis is a field in computational linguistics which concerns the automatic evaluation the axiological charge contained within a given text (Taboada 2016). Three main aspects of such analysis, according to Esuli and Sebastiani (as cited in Taboada 2016), are first determining the subjectivity of the text - deciding whether a text contains subjective opinions whose polarity can be measured and, if so, determining a text's polarity and lastly evaluating the strength of that polarity.

In order to successfully estimate the polarity of a new text, a sentiment analysis model needs to be driven by means of information on how to process and classify characteristics extracted from that text. Two main approaches in building a sentiment analysis model are a lexicon-based one and a machine learning one (Taboada 2016). This work focuses on the lexicon-based approach as it is more closely related to the linguistic information contained in the text.

2.6.1 Lexicon-based approach in sentiment analysis

In a lexicon-based (also called dictionary-based or rule-based) method, dictionaries which consist of evaluative words are applied to classify individual words in a text as positive and negative and then aggregate their values to determine the semantic orientation of the entire text.

Dictionaries used in such models often include information about how positive or negative a word is expressed on a scale. For example, the dictionary in the Semantic Orientation Calculator (SO-CAL) uses a 10-point scale, from -5 to +5 (Taboada 2016). The sizes of such dictionaries vary between models and range from 5,000 words include in SO-CAL (Taboada et al. 2011) to almost 76,000 words included in the Macquarie Semantic Orientation Lexicon (Mohammad et al. 2009). Arguably, large dictionaries tend to capture more noise which leads to less accurate results (Taboada et al. 2011).

A rule-based method also needs to account for intensifiers – “devices that change the intensity of an individual word, whether by bringing it up or down” (Taboada 2016: 13-14). Taboada et al. (2011: 275) propose a percentage scale for modifiers to calculate values affected by them:

- most +100%
- really +25%
- very +15%
- somewhat -30%
- arguably -20%

Therefore, the most advanced lexicon-based models also need to contain dictionaries of positive and negative modifiers with assigned values. Then if an opinion word is found to be modified by one of them its value is calculated accordingly. In the model described by Taboada (2011) the base value of an opinion word is multiplied by the percentage assigned to the modifying word which affects it and is added to the base value. For example, if the opinion word ‘good’ is assigned a value of +3 its value would be calculated as +3.75 when accompanied by the modifier ‘really’.

$$3 \times 0.25 + 3 = 3.75.$$

In the same way, its value would be reduced to +2.4 if it occurred in a phrase ‘arguably good’.

$$3 \times -0.2 + 3 = 2.4$$

Another factor that needs to be taken into consideration while processing a text is negation. An approach suggested by Taboada (2016: 18) is to use shift negation, “method where the effect of a negator is to shift the negated term in the scale by a certain amount, but without making it the polar opposite of the original term”. While other researchers such as Asghar et al. (2017) propose reversing the value of a negated opinion word by multiplying it by -1, this approach assumes that negations are not used to reverse the meaning of an evaluative phrase, but only to tone it down. For instance, if the adjective ‘excellent’ had an assigned strength of +5, in this model its value would be reduced to +1 when negated ‘not excellent’ Taboada (2011: 277). This approach is therefore more natural than assigning the opposite value of -5, as the sentiment of the phrase ‘not excellent’ is usually not perceived as a strongly negative one but it is rather close to neutral.

2.6.2 Applications of sentiment analysis

As sentiment analysis programmes can instantly provide a researcher with the information about a text’s polarity it is often used to analyse reviews of movies, books and consumer products to determine a reviewer’s attitude towards the product. Results of sentiment analysis conducted on the content on social media are often used by companies, marketers and political analysts (Taboada 2016). However, there are also projects which perform sentiment analysis on larger samples of texts to observe patterns within them; for example, “Analyzing Plot Structure of Novels Using Sentiment Analysis” by Zakir Hossain (2018).

2.7 Conclusions

The fact that valuations are manifested in a language makes it possible to evaluate lexical items on a positive-negative scale. In sentiment analysis, such evaluations can be performed automatically for larger samples of texts taking into account the context in which each of the items is used by examining factors such as modifiers and negation. In a lexicon-based approach, the models use lists of positive and negative words with their assigned polarity scores as well as lists of modifiers with designated percentage values which determine how strongly they influence the polarity value of the item they modify. While Krzeszowski (1997) claims that axiological charge can be elicited for every lexical item and that it is always semantically relevant, contemporary linguists and scientists developing sentiment analysis programmes assess most items as neutral and therefore having no impact on the total polarity score of an analysed sample. As programmes for sentiment analysis can quickly process large amounts of data and provide fairly precise estimations of sentiment within analysed text, they are useful not only for linguistic research but also for commercial and political purposes.

3. Sentiment analysis of the Black Mirror series

3.1 Introduction

This chapter aims to present the findings of the study on sentiment analysis of the Black Mirror series. It starts with a short overview of recent projects in the domain of sentiment analysis to provide the reader with some background knowledge about the current trends in that field. The following sections explain what the purpose of the research is and what methodology has been employed to conduct it. The methodology section includes a presentation of the Black Mirror Corpus, which was compiled for the study, as well as an explanation of how the tools – VADER Sentiment Analysis and Martindale's Dictionary – have been used to analyse it. Subsequently, the results of the research are discussed together with difficulties which occurred in the process of sentiment analysis and ideas for improvements. The interpretation of these findings is followed by a section suggesting their possible applications in the film industry.

3.2 State of the art in sentiment analysis

Sentiment analysis, being a means of automatic classification of texts as positive or negative, is commonly used in various types of studies and as a process in developing more complex tools for the text analysis. This section reviews three recent scientific articles describing the application of sentiment analysis in different projects, in which it helped to develop more advanced tools such as the venue recommender system, the news classification system based on the sentiment of the text and software for recognizing hate speech on social media.

In the article *Comparison of Sentiment Analysis and User Ratings in Venue Recommendation* Wang, Ounis and Macdonald (2019) analysed different types of sentiment analysis engines to determine which approach is most useful in classifying venue reviews on Location-Based Social Networks (LBSNs) such as Yelp to improve their recommender systems. Their study compares four engines for sentiment analysis: one lexicon-based classifier - SentiWordNet-Based Classifier (SWN) and three classifiers based on different types of neural networks – SVM-Based Classifier (support vector machine), CNN-Based Classifier (Convolutional Neural Network) and LSTM-Based Classifier (Long short-term memory). Based on the results of their study, the researchers argue that CNN-based classifiers and LSTM-based classifiers are the most accurate and result in the recommendation models with the highest effectiveness.

Taj, Shaikh and Meghji (2019) have recently presented a study in which lexicon-based sentiment analysis using SentiWordNet was paired with TF-IDF (Term Frequency-Inverse Document Frequency) statistical technique to develop a classifier tailored to the language used in the domain of news articles. The TF-IDF technique was used to find the most frequently used words in articles and assign weight to words based on their frequencies. Lists of words compiled in that process were later assigned polarity scores based on WordNet and used to calculate total polarity scores of news articles from the BBC news dataset. The study found that categories of sports and business had more positive articles, while entertainment and tech had a majority of negative articles. Researchers argue that such sentiment analysis methods can help to customize the news feed for particular users.

The article *Automatic Detection of Hate Speech on Facebook Using Sentiment and Emotion Analysis* (2019*)* written by Rodríguez, Argueta and Chen describes a project in which lexicon and rule-based sentiment analysis VADER was used together with JAMMIN – an emotion analysis method which relies on pre-computed emotion-loaded patterns to determine the emotions expressed in social posts. In further stages of the study, unsupervised clustering was implemented to determine the most discussed hate-promoting topics. The ultimate goal of the project was to identify the pages that promote hate speech by posting comments related to controversial topics on Facebook and the researchers claim that the approach they employed shows very encouraging results.

In conclusion, both lexicon-based and machine-learning approaches are used in recent projects implementing sentiment analysis. Also, the types of texts to which sentiment analysis is applied vary from short texts such as venue reviews and social media content to longer documents such as news articles. Supplemented by other statistical and analytical techniques such as TF-IDF or automatic emotion detection sentiment analysis can be used to develop advanced tools such as recommender systems, text classifiers of hate speech detecting software.

3.3 Purpose of the research

As it has been explained in the previous chapters, sentiment analysis is commonly and successfully used for automatic categorization of reviews, tweets and other types of short texts expressing opinions. This work aims at assessing the value of applying sentiment analysis to texts of moderated discourse such as transcripts of TV programmes. For this purpose, a commonly used open-source tool for sentiment analysis was applied to transcripts of a popular Netflix series The Black Mirror, which is a British dystopian science fiction anthology picturing the future society of our planet as new technologies emerge and result in unpredicted consequences. The choice of the series was dictated by the fact that the series is renowned for its negativity and dark, satirical tone.

The research has been conducted with an inductive approach, aiming to evaluate the possible application of the already well-researched process of sentiment analysis to a field in which it is still not commonly used. The focus was on examining that process, finding difficulties which can occur while analysing texts of such type and testing ideas for improving the analysis so that the conclusions could help to develop a tool designed specifically for the analysis of the language used in television programmes and films.

3.4 Methodology

3.4.1 Collecting and pre-processing data

The data for this research consists of transcripts of 22 episodes of *The Black Mirror* Netflix series. Hearing-impaired subtitles (including information about music and environmental sounds) were downloaded on the 17th of January, 2020 from https://www.opensubtitles.org/ in .srt format and later converted to .txt files using the tool available at https://subtitletools.com/convert-subtitles-to-plain-text-online. The text files were then used to compile a corpus using the PlaintextCorpusReader function of the NLTK (Natural Language Toolkit) Python package. All newline ('\n') and carriage return ('\r') characters were deleted and hash ('#') and musical note ('♪') symbols were replaced with periods to facilitate tokenizing excerpts from music lyrics which lacked any other punctuation marks. Consequently, the sents_tokenize function of the NLTK library was used to tokenize the text into sentences. Original punctuation and spelling (including the use of capital letters) was preserved in the texts as it is needed for calculating the sentiment score. As a result, a custom corpus of Black Mirror episodes' transcripts (further referred to as the Black Mirror Corpus) consisting of 154183 words and 22169 sentences was compiled to be further researched.

3.4.2 Sentiment Analysis with VADER

3.4.2.1 Tool description

Sentiment analysis was conducted using VADER (Valence Aware Dictionary and sEntiment Reasoner) library for Python, which was created by C.J. Hutto Eric Gilbert and is freely available at https://github.com/cjhutto/vaderSentiment. In the research the latest (3.3.1) version of the tool (as available on May 17, 2020) was used. The key components of the package are a lexicon containing data about intensity and polarity of concepts and a Python code for rule-based sentiment analysis engine. The lexicon currently consists of 7518 items (which include sentiment words, emojis, acronyms, initialisms and sentiment carrying slang expressions) with their corresponding sentiment value between -4 and +4. The value is an average of "sentiment ratings from 10 independent human raters (all pre-screened, trained, and quality checked for optimal inter-rater reliability)" (https://github.com/cjhutto/vaderSentiment). Figure 3.1 presents examples of words from VADER lexicon with their corresponding values.

Great	3.1
Good	1.9
Okay	0.9
Sad	-2.1
Horrible	-2.5
Hopelessness	-3.1

Figure 3.1 Examples of words from the VADER lexicon with their corresponding values

The analysis engine works basing on that lexicon but also by implementing grammatical and syntactical rules, which recalculate sentiment score on a level of a sentence. What is important, these heuristics incorporate word-order sensitive relationships between terms such as intensifiers impact the strength of a concept. The package takes into account also punctuation and phenomena such as capitalization of sentiment words (which strengthens the polarity value) and negation within a tri-gram preceding the sentiment word (which in this model flips the sentiment value of the whole sentence). Furthermore, the engine also analyses constructive conjunction with *but*, which according to Hutto and Gilbert (2014: 6) "signals a shift in sentiment polarity, with the sentiment of the text following the conjunction being dominant". They explain this function on the example of the phrase "The food here is great, but the service is horrible" whose sentiment is mixed, yet dominated by the negativity in the second clause.

As a result of all of these processes a user can obtain four values for any analysed texts: pos (positive), neu (neutral), neg (negative) and compound scores. Compound scores are computed by summing the intensity scores of words which can be found in the lexicon and later adjusted according to the rules. They are presented in a normalized score between -1 and +1, where values below -0.05 are usually considered as negative, values between -0.05 and 0.05 neutral and values above 0.05 are classified as positive. The pos, neg and neu scores represent proportions of text which fall into each category and therefore their sum should be equal 1 (or close to it with float operation). All these metrics together provide the user with multidimensional measures of sentiment for a given text.

```
The food is good
{'neg': 0.0, 'neu': 0.508, 'pos': 0.492, 'compound': 0.4404}
The food is very good
{'neg': 0.0, 'neu': 0.556, 'pos': 0.444, 'compound': 0.4927}
The food is VERY good
{'neg': 0.0, 'neu': 0.505, 'pos': 0.495, 'compound': 0.6028}
The food is VERY GOOD
{'neg': 0.0, 'neu': 0.462, 'pos': 0.538, 'compound': 0.6867}
The food is not good
{'neg': 0.376, 'neu': 0.624, 'pos': 0.0, 'compound': -0.3412}
The food is not good
{'neg': 0.376, 'neu': 0.624, 'pos': 0.0, 'compound': -0.3412}
The service wasn't great
{'neg': 0.523, 'neu': 0.477, 'pos': 0.0, 'compound': -0.5096}
The food was good, but the service wasn't great
{'neg': 0.332, 'neu': 0.523, 'pos': 0.146, 'compound': -0.5409}
The food is good, but the service was even better
{'neg': 0.0, 'neu': 0.58, 'pos': 0.42, 'compound': 0.7003}
```

Figure 3.2 Examples of pos, neu, neg and compound scores computed using VADER sentiment for sentences with examples of positive sentiment words, intensifiers, capitalization, negation and contrastive conjunction with *but*

Examples of sentences provided in figure 3.2 show how the sentiment score of a positive sentence ‘the food is good’ is increased when an intensifier is used. The use of capitalization of sentiment words increases the score even more, whereas negation causes the adjectives good and great to no longer be recognized as positive and therefore flips the sentiment value of the whole sentences. The programme detects also the conjunction with *but* and reassigns the sentiment value of the following clause. An occurrence of a more negative clause after the conjunction is interpreted as magnifying the strength of the second clause, while in other cases the clause after the conjunction is considered to be less relevant and therefore its score is divided by two.

3.4.2.2 Analysing the sentiment

The methods employed in this part of the research were predominantly quantitative and based on computer calculations using the VADER Python package described above. Firstly, the average positive, negative and compound score values of each episode were calculated to compare them with each other and find the most positive and most negative episode of the series. For this part of the research, only the non-neutral sentences (i.e. with polarity score below -0.05 or above +0.05) were taken into account in order to obtain average polarity score which would not be affected by sentences without any axiological charge, as they were deemed irrelevant at this stage of the study. Secondly, the polarity scores were compared to the rating of episodes to see if there occurred a correlation between these figures. The programme was also used for finding sentences with the strongest intensity (highest pos and neg scores). The frequencies of sentiment words (based on VADER lexicon) in the Black Mirror Corpus were calculated and the most common negative and positive words were analysed.

The quantitative study was supplemented by a qualitative analysis of the results. Therefore, the most positive and negative sentences were analysed to discover contexts in which the VADER sentiment analysis tool is not able to evaluate text properly. The contexts in which most common positive and negative words were analysed in a similar manner, to observe whether the lexicon or rules set can be improved to facilitate analysis of texts similar to the ones included in the Black Mirror Corpus.

3.5 Martindale's Dictionary

For the following, also quantitative part of the research, the Martindale's regressive dictionary was used. The dictionary was created as a tool for content analysis to measure primordial (associative, concrete, and taking little account of reality) and conceptual (abstract, logical, reality-oriented, and aimed at problem-solving) thinking. It consists of approximately 3200 words and roots divided into 29 categories of primary process cognition, 7 categories of secondary process cognition and 7 categories of emotion, which were derived from the theoretical and empirical literature on regressive thought (Martindale as cited by https://www.kovcomp.co.uk/wordstat/RID.html - accessed on May 17, 2020).

> *"The rationale behind the dictionary is that psychological processes will be reflected in the content of a text. Thus, for example, the more primordial the thought involved in producing a text, the less abstract and the more drive- and sensation-oriented words it should contain"*
> - *https://www.kovcomp.co.uk/wordstat/RID.html - accessed on May 17, 2020*

The texts from the Black Mirror series corpus were searched for words which according to this dictionary fall into the following categories: positive affect, anxiety, sadness, affection, aggression, expressive behaviour and glory. To enable that, the .CAT file of the dictionary was downloaded from https://www.kovcomp.co.uk/wordstat/RID.html on 10th of November, 2019 and words corresponding to each of emotion subcategories were saved to a python list object.

Frequencies of these words were then calculated to find the dominating emotions in each of the texts. Sentences containing these emotion words were also analysed together with their sentiment scores, to investigate whether their presence affects the intensity of sentiment and whether the VADER sentiment engine can recognize it correctly. Figure 3.3 presents examples of words in each category.

Positive Affect	Cheerful, enjoy, fun
Anxiety	Afraid, fear, phobic
Sadness	Depression, dissatisfied, lonely
Affection	Affectionate, marriage, sweetheart
Aggression	Angry, harsh, sarcasm
Expressive Behavior	Art, dance, sing
Glory	Admirable, hero, royal

Figure 3.3 Examples of words corresponding to different emotion subcategories as classified in Martindale's dictionary

3.6 Sentiment analysis of the Black Mirror Corpus

3.6.1 Average compound values and episodes rating

Average values of non-neutral sentences within each episode were computed to compare them across each other and with the ratings of corresponding episodes according to the most popular cinematography database filmweb.pl. However, despite eliminating neutral sentences from the calculations, the average polarity scores of all episodes were still close to zero, with episode 5 of the fourth season obtaining the lowest score of -0.13 and third episode of the first season being the most positive with the score of +0.19. The average neu scores calculated for these episodes vary between 0.79 and 0.86 even though only the non-neutral sentences (with a compound score lower than -0.05 or higher than +0.05) were taken into account at this stage of the study. These facts lead to the conclusion that the transcript of a TV series, even the one that is widely perceived as having a very strong emotional charge, is predominantly neutral in terms of its sentiment value as calculated by a sentiment analysis engine.

As far as the ratings are concerned, there is little correlation between the polarity scores of episodes and their corresponding sentiment value. Graph 3.4 displays the rating of each of the episodes (on the scale from 1 to 10) and the compound scores of the episodes multiplied by 10 for the ease of reading the graph.

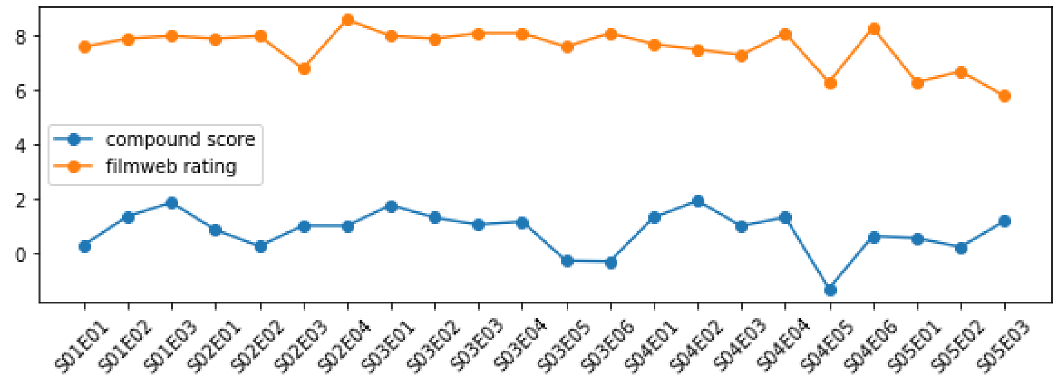


Figure 3.4. Compound scores and filmweb.pl ratings of all Black Mirror episodes

As it can be observed, the lowest polarity score of the fourth episode of the fourth season was paired with the second lowest rating (6.3) among all the episodes, yet it is impossible to find more dependencies between the two plots.

3.6.2 Sentences with highest pos and neg values

As was explained in the previous chapter, the pos neg and neu scores correspond to the proportions between words classified as positive, negative or neutral within a given sample. Therefore, it was not an unexpected result that the sentences which showed the highest pos and neg scores were these short clauses which contained only one axiologically charged word such as ‘yes’ or ‘no’. Figure 3.5 shows examples of ten sentences with the highest intensity scores found in episode 1 of the series.

```
Yes.--------------------------  1.0
Interesting.------------------  1.0
Please!-----------------------  1.0
Yeah.-------------------------  1.0
Yes!--------------------------  1.0
Confident?--------------------  1.0
Yeah.-------------------------  1.0
Yeah.-------------------------  1.0
Yeah.-------------------------  1.0
Yes.--------------------------  1.0

No.---------------------------  1.0
No.---------------------------  1.0
No.---------------------------  1.0
No.---------------------------  1.0
Fuck!-------------------------  1.0
Poor bastard!-----------------  1.0
You're lying!-----------------  0.787
Then you're a stupid bitch.---  0.783
Bad outcome.------------------  0.778
No sign of struggle.----------  0.692
```

Figure 3.5 Sentences with highest intensity scores found in episode 1 of the series.

3.6.3 Highest and lowest compound scores

While sentences with the highest intensity were very short, consisting of only one or two words, these were usually the more complex sentences which obtained the highest compound scores. In figure 3.6 the sentences with the highest compound score from each episode are presented. As it can be observed in these examples, the episodes subtitles included hearing impaired information about the manner of speaking or background noises such as *cheering giggling* and *laughter* and those words contributed to the sentiment scores of the analysed sentences, very often increasing their score. Such calculations are not necessarily always correct, as can be seen on the example of sentence from S01E02 where sounds such as *manic giggling* or *crazed laughter* are evaluated as highly positive. Unfortunately, even though VADER analysis engine recognizes intensifiers and negation it does not have a feature which would re-evaluate the polarity score of a word based on preceding adjective in phrases such as presented in that example. The most positive sentence from episode 2 of the second season also shows a problem, as the sentence in fact contains a threat followed by a cry for help. However, the sentiment analysis engine calculated it as positive basing on the fact that it contains an endearing word darling and the word *help* which in the VADER lexicon is listed as positive, with a score of +1.7.

Lyrics of songs playing in the background were also included in the transcripts and therefore these also contributed to the overall scores of the episodes. Frequently, sentences from the lyrics were evaluated as having a high positive value, as in episodes S05E02 and S05E3, where the sentences with highest compound scores were not part of the dialogue but lines from songs included in the show. Another tendency which can be observed upon analysing the sentences with the highest compound values is that the repetition of words with high polarity scores such as *yes* or *please* significantly increases the total compound score of a sentence.

Analogically, repetition of negative words such as *no* or *shit* contribute to a lower polarity score of a sentence and therefore sentences such as *[PHONE BEEPS] No, no, no, no, no, no.* or *Shit, shit.* were computed as the most negative sentences in their episodes. In general, sentences with the lowest compound scores frequently contained common swear words such as *fuck* or *shit* but also vocabulary connected with crime such as *murder, kill* or *stolen*. Figure 3.7 presents the most negative sentences from each of the files in the Black Mirror Corpus.

CHEERING 'I trust this will bring about the safe return of Princess Susannah.'	0.8816
(MALE HEAVY BREATHING) (ABBY MOANING) (ABBY SINGING FAINTLY) (TV) 'The hottest girls... ' MANIC LAUGHTER MANIC GIGGLING CRAZED LAUGHTER CONTINUES (BREATHES HEAVILY) PARTY MUSIC ECHOES INDISTINCTLY Could you stand aside, please?	0.9393
VOICES AT HIGH SPEED OK, great, this has been really great, Liam, we really hope to... look forward to seeing you again.	0.9363
It's perfectly natural, there's nothing wrong with it.	0.8502
What I'm going to do here darling, is just going to push this straight through your... Help me!	0.8367
So's YouTube and I don't know if you've seen it but the most popular video is a dog farting the theme tune to Happy Days.	0.8804
Just, like, pills, but they must've been pretty strong because I actually enjoyed myself.	0.9428
- [cheering] - But, of course, my most heartfelt thanks, my warmest thanks, are to someone who's my best friend, my lover and now, I'm honoured to say, my wife.	0.9911
Yeah, well, I'm afraid Britain's pretty tame as far as daredevil opportunities go.	0.8625
_ He was so funny in the play.	0.7304
Wish me love a wishing well to kiss and tell .	0.9136
Anderson, Hague, I want you to stay here and help them torch the food, keep them calm and reassured. 0.7906	
Well, that's great, save the planet, hallelujah!	0.9286
For he's a jolly good fellow For he's a jolly good fellow .	0.9081
Wow, OK. - [TABLET BEEPS] - [WOMAN] These are her vitals.	0.8331
["ANYONE WHO KNOWS WHAT LOVE IS" PLAYING] God, yeah, yeah, that, that was on.	0.8868
I felt safe, happy, comfortable.	0.872
SWEET CLANGS 993, 992... SWEET CLANGS - Yes.	0.8797
Please, please, please, please, please.	0.8591
Love is the greatest feeling .	0.872
And I thank God I'm alive .	0.7351
Oh, yes, yes, yes, yes, yes, yes.	0.9349

Figure 3.6 Sentences with highest compound score values in each episode

I considered it necessary... FUCKING HELL!	-0.8249
They may as well have cut to a war crime!	-0.8268
It's like I've had a bad tooth for years and I'm just finally getting my tongue in there and I'm digging out all the rotten shit.	-0.836
Shit, shit.	-0.802
'The couple were caught after harrowing footage of Jemima's 'torture and murder was discovered on a mobile phone 'in Ski llane's possession.	-0.8625
It's a death trap anyway, look at it!	-0.7574
You also failed to report a murder - fatal poisoning, in case you've forgotten.	-0.9531
So when he died I thought, fuck it.	-0.8107
- [screams] - Stop that, that could kill you!	-0.8016
- A dirty, sick, disgusting pervert?	-0.9531
But so is gas mileage... [woman] No, no, no.	-0.8126
We've got reports of a food store broken into, shit stolen.	-0.872
Picking along dead ends just to prove they're dead ends, that's most of the job.	-0.8625
[WALTON EXHALES] But you threw my son out of an airlock, so fuck you to death.	-0.9272
It's not like I use that shit.	-0.6917
- [PHONE BEEPS] No, no, no, no, no, no.	-0.8807
No, no, no, no, no.	-0.8402
ROBOT WHIRRS Shit, shit, shit.	-0.8957
38-year-old Clayton Leigh was today found guilty of the horri fic murder of WNL weather reporter Denise Stockley, whose mut ilated remains were discovered one year ago today.	-0.9169
About to kick your ass, motherfucker.	-0.8442
Seriously, my uncle killed himself and it fucked my whole fam ily up.	-0.891
Catherine is so fucking full of shit, you two have no clue ho w much of a bitch that woman has been.	-0.8709

Figure 3.7 Sentences with lowest compound scores in each episode

3.6.4 Most commonly occurring positive and negative words

Interestingly, there were far more words recognized as positive (6776 occurrences of positive words in total) in the Black Mirror corpus than there were negative ones (4140 occurrences). By far, the negative word that was used most often was simple *no*, followed by two popular swear words *fuck* and *shit*. Other negative words which appeared in the corpus fairly often were *stop*, *screams*, *seriously*, *hell*, *gun* and *wrong*, which suggests a lot of action and drama in the series. The sixth most popular negative word (with 66 occurrences) was 'mia', which has a -1.2 polarity score in the VADER lexicon, most likely as an acronym standing for 'missing in action'. In the series, however, it was used simply as a name of one of the characters in the third episode of the fourth season. Unfortunately, VADER sentiment analysis tool does not account for word sense disambiguation and classifies words as positive or negative based on their score in the VADER lexicon, regardless of the meaning that was used in a particular context.

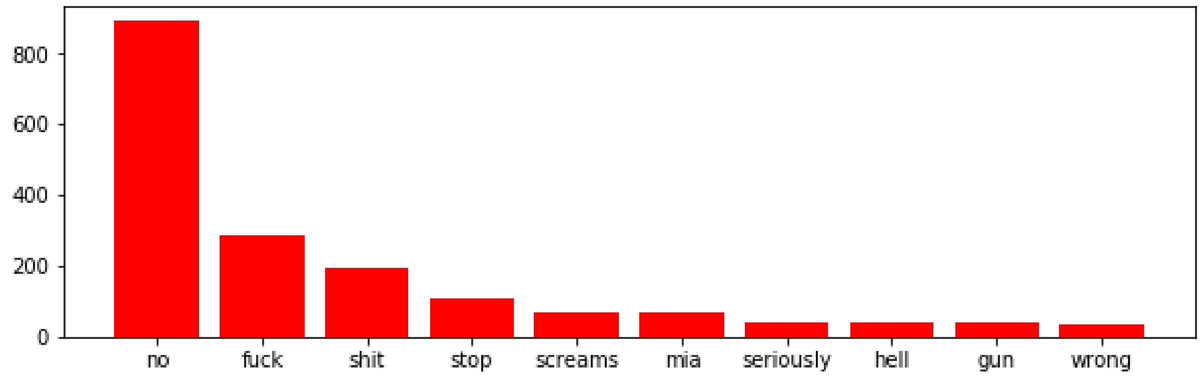


Figure 3.8 Most frequently occurring negative words

The word which occurred most often and was classified as positive was *yeah* which stands in the VADER lexicon with a score of +1.2, and therefore can be interpreted as moderately positive. Such valuation of these words is questionable in the context of the Black Mirror Corpus, where the word was used to introduce sentences of all types; from highly positive ones to strongly negative ones and also used simply as phatic expressions, often repeated two or three times within a sentence.

```
[ LOWRY ] Yeah , you ' re seriously gonna need to sit the fuck down .
Yeah , that feels quite nice .
Yeah , yeah , yeah , piss off , Jonas .
Yeah , yeah , right .
```

Figure 3.9 Examples of sentences with the word yeah from Black Mirror Corpus

Similar conclusions can be drawn about words such as *ok* (spelled also as *okay)* and *well* which also belong to the group of most commonly used positive words in the Black Mirror Corpus. Similarly, the word *god* was a frequently used positive word, yet in the majority of cases it was used in phrases such as *oh, my God* and in sentences *Oh ... Oh, God, no.* or *Oh, God, I don't want to die.* In those sentences, the use of expressions with the noun *God* intensifies the polarity of the whole sentence rather than adds a positive value. Nevertheless, the current version of VADER sentiment analysis engine does not take that phenomenon into account.

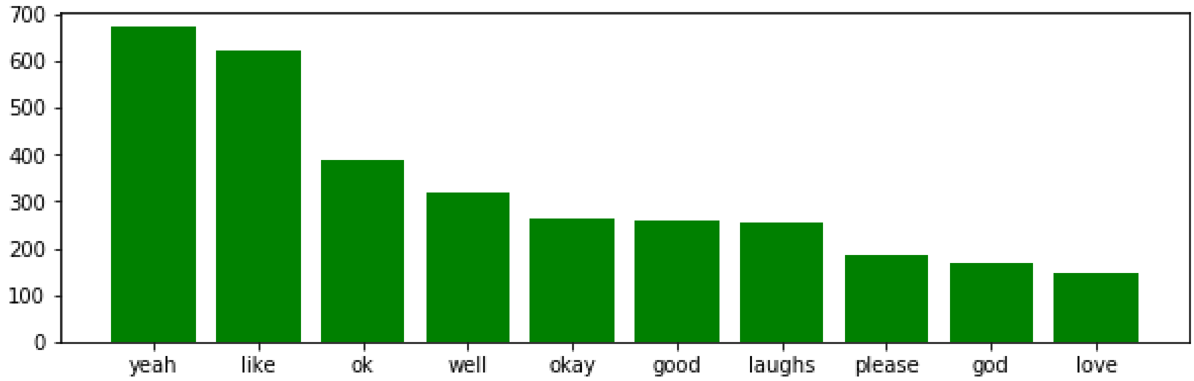


Figure 3.10 Most frequently occurring positive words

As can be observed in Figure 3.10, the word *like* was the second most commonly used positive word with 624 occurrences in the corpus. As was mentioned before, VADER sentiment tool does not account for word sense disambiguation or even part of speech recognition and therefore classifies the word *like* as a positive one not only when used as a verb but also in other cases, for example when used as a preposition. Therefore, the sentences *I like her.* and *He snores like a bison*. will be computed as having the same sentiment value.

```
I like her
{'neg': 0.0, 'neu': 0.286, 'pos': 0.714, 'compound': 0.3612}
He snores like a bison
{'neg': 0.0, 'neu': 0.545, 'pos': 0.455, 'compound': 0.3612}
```

Figure 3.11 Sentiment scores for sentences with *like*

A sentence with a misinterpreted use of like was also present among the negative sentences in figure 3.11. The negation of the word *like* in the sentence *It's not like I use that shit* resulted in a very negative compound score while the word like was used as a proposition and therefore should be considered neutral. The negation should then be applied to the verb *use*, as the sentence is similar in meaning to *I don’t use that shit*, which would then have a positive compound value.

Other words which were frequently used in the series transcripts were *laughs*, *please*, and *love*, which in the majority of cases were indeed used in a positive context. It is worth mentioning, however, that there were instances of sentences which would unmistakably be classified as negative by a human judge, yet the VADER sentiment tool incorrectly calculated their value as positive due to the use of those words.

```
Make love to a pig!
{'neg': 0.0, 'neu': 0.4, 'pos': 0.6, 'compound': 0.6696}
```

Figure 3.12 Example of an incorrectly recognized sentiment of a sentence with the word *love*

A good example of such a sentence is shown in figure 3.12, where sentence *Make love to a pig!* is clearly charged negatively, yet its compound score is calculated as approximately +0.7 and therefore highly positive.

3.6.5 The analysis of emotion words with Martindale's dictionary

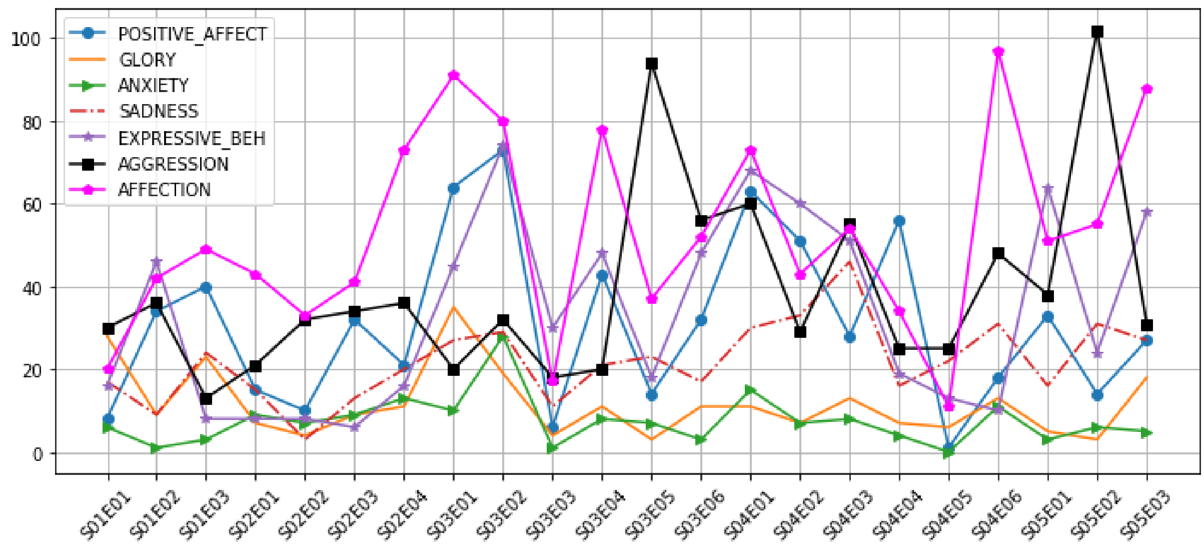


Figure 3.13 Distribution of emotion words across the Black Mirror episodes

The current part of the research is focused on determining the dominating emotions in the Black Mirror Corpus. Figure 3.13 presents the distribution of emotions accounted for by the Martindale dictionary across analysed episodes of the Black Mirror series. While the overall polarity scores of those episodes are fairly similar, the emotion words count allows us to compare the episodes by terms of their dominating emotions. As can be read from the graph, anxiety and glory are the least represented emotions within the corpus and remain on a low and stable level with averages of 12 and 7 words representing them in an episode respectively. Words expressing sadness are not frequently used in the episodes either, yet their numbers started growing at the beginning of the fourth season of the season to peak in its third episode, where the programme found 46 words corresponding to that emotion. The average value for sadness equals to 22. For a positive affect, aggression and affection the average values are higher (31, 39 and 52 words per episode) and the distribution of these words is more varied across episodes. It can be therefore observed that there are episodes such as S03E01 with very high numbers of affection words and episodes S03E03 and S04E05 with scores below 20. Similarly, episodes S03E05 and S05E02 have by far the highest scores of aggression. Interestingly, the numbers of all emotion words were much lower in the first two seasons of the series and only later can there be observed episodes that are dominated by a particular emotion. Figure 3.14 presents a diagram of the distribution of emotion words in the first episode of the series (on the left-hand side) contrasted with a diagram presenting the emotion words in the fifth episode of the third season (on the right hand-side) dominated by words corresponding to aggression.

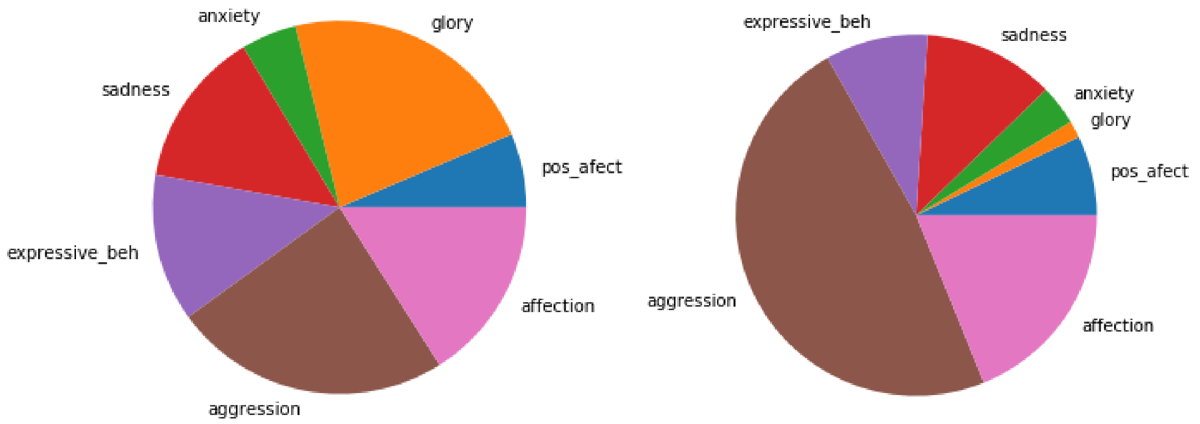


Figure 3.14 distribution of emotion words in episodes S01E01 and S03E05

The distributions of all emotion words have been also compared to the ratings and calculated polarity scores of the series, yet there has been no correlation observed.

3.6. Interpretation

While sentiment analysis has a potential use in the film industry, the study has shown that there are still a lot of problems with implementing a rule and lexicon-based sentiment analysis tool to such long and complex texts as the transcripts of TV series. Firstly, it has to be noted that unlike reviews or social media content text of this type is predominantly neutral and therefore it is difficult to classify whole text files as clearly positive or negative. It can, however, help to analyse individual sentences or scenes in terms of their sentiment value. A tool for the analysis of this type should be more advanced than the one used in the study (VADER sentiment analysis) and implement functions such as part of speech recognition and word sense disambiguation for more precise calculations of sentiment values. The list of fixed phrases should be updated to contain expressions such as *oh, God* so that they can be evaluated according to their meaning as a whole rather than calculated basing on the sentiment word they contain. Moreover, the programme would benefit greatly from adding a list of adjectives and adverbs modifying the meaning of words describing the manner of speaking and noises used in closed captions and hearing-impaired subtitles. Such list could be then implemented with a set of rules to flip the sentiment score of phrases such as *manic giggling* or *crazed laughter*. An idea worth considering would be also to introduce rules accounting for repetition of the same sentiment word within one sentence, as in this type of text it hardly ever increases the actual sentiment value of a sentence. The study of particular emotions within a text is a promising idea, which can be used to supplement sentiment analysis and which seems to be a more detailed yet simpler to implement measure of sentiment of larger text samples.

3.7 Possible applications in the film industry

While evaluating the polarity of films and episodes on a positive-negative scale based just on their transcripts may be problematic and inconclusive due to the predominant portion of neutral sentiment within them, the study of particular emotions is a promising concept which can be further researched and exploited in the film industry. Classifying films, shows and even individual episodes based on dominating emotions can help improving recommender systems used on streaming platforms to provide users with suggestions tailored to their taste with reference to their previous choices and reviews or even suggest positions to watch on the basis of the viewer’s current mood. With a classification based on the content of the movie, rather than just belonging to a certain genre, such suggestions can be more accurate. Moreover, statistical data about the language used in a movie or other television shows can be a very good indicator of whether the content is appropriate for children. Parents themselves or a parental control software could easily check whether the show features aggression or contains inappropriate language for a young viewer. The analysis of the content would offer more information than just the maturity ratings provided by the producers and therefore could allow personalisation of content filtering based on custom criteria.

3.8 Conclusions

The overview of recent projects in sentiment analysis has shown that both lexicon-based and machine-learning approaches are used in recent projects implementing sentiment analysis. Paired with other statistical and analytical techniques such as TF-IDF or automatic emotion detection sentiment analysis can be used on various text types and help to develop advanced tools such as recommender systems, text classifiers of hate speech detecting software.

The results of the research conducted on the Black Mirror Corpus with the use of VADER Sentiment Analysis and Regressive Imagery Dictionary has helped to identify problems occurring in the analysis of longer text samples, such as incorrect classification of words as positive or negative, insufficient list of modifying words and unprecise calculations of sentiment scores for sentences which contained word repetition. Those issues can be solved by expanding the list of modifiers, adding rules for calculating the sentiment value in the case of word repetition and applying POS tagging and word sense disambiguation. The results of the study show that sentiment analysis of films and television series has a lot of potential for use in the film industry. The automatic detection of dominating emotions is particularly promising as it is easy to apply and provides more detailed information about a specific text, which can be used to improve parental control and film recommender systems.

4. Overall conclusions

Sentiment analysis is a good example of how linguistic research can be enriched and facilitated by combining linguistic theories with computer science. The fact that valuations are manifested in a language and therefore lexical items can be evaluated on a positive-negative scale was observed by linguists such as Krzeszowski (1997) already in the previous century, yet their studies were only theoretical. With the use of today's technology that premise can be used to elicit the axiological charge of a large number of lexical items, create lexicons and on their base automatically compute polarity scores for whole sentences. As programmes for sentiment analysis can quickly process large amounts of data and provide fairly precise estimations of sentiment within the analysed text, they are useful not only for linguistic research but also for commercial purposes.

As sentiment is "at the intersection of linguistics and computer science" (Taboada 2016: 2) studies in this field require at least a basic knowledge of a modern programming language. The entire research conducted for this work was based on a custom-developed Python code, including data pre-processing, compiling a custom corpus, conducting sentiment analysis and a quantitative study on the emotions expressed in the examined text. While there are many freely available software toolkits for concordancing and text analysis (as described in the first chapter), using the NLTK Python library to compile the corpus not only did not require downloading any additional software but also allowed for further research on the text collection. The tool chosen for the sentiment analysis was another Python package, called VADER Sentiment Analysis, which includes a lexicon and rule-based engine for determining the polarity of a text sample. It is based on a lexicon consisting of 7518 items with their corresponding sentiment value between -4 and +4. Those sentiment values are synonymous with what Krzeszowski (1997) refers to as axiological charge (which was widely discussed in the second chapter) and were elicited by the developers of the library in the same fashion as the method Krzeszowski (1997) suggests for eliciting absolute values of particular lexical items. The subsequent part of the research, the analysis of emotions expressed in the corpus, was based on custom-written functions applied to the corpus to calculate the frequencies of emotion words classified in the Regressive Imagery Dictionary.

The study was undertaken with an inductive approach and aimed to assess the possible application of sentiment analysis to the language of films and television series. The focus was on examining the process, finding difficulties which can occur while analysing texts of such type and suggesting solutions. Problems which were identified included incorrect classification of words as positive or negative, insufficient list of modifying words and unprecise calculations of sentiment scores for sentences which contained word repetition. Those issues can be addressed on the level of altering the code of the sentiment analysis engine by expanding the list of modifiers it uses, adding rules for calculating the sentiment value in the case of word repetition and applying POS tagging and word sense disambiguation. The calculations of emotion words within the episodes allowed for analysis of the distribution of emotions across the episodes, finding dominating emotions for each text and also recognizing the episodes with highest aggression and sadness values. Being able to automatically obtain this type of information about television shows and films could be very valuable for the film industry, as it can be used to improve parental control and film recommender systems.

All things considered, even though lexicon and rule-based sentiment analysis engines such as VADER are successfully used for product reviews, social media content and other short text samples, they need substantial improvements to be effectively used for analysing larger texts such as transcripts of television series. Emotion analysis, however, is a less complex procedure which can be used both as a complement to sentiment analysis and as an independent process. Automatic detection of dominating emotions can potentially improve other programmes such as recommender systems or parental control software.

Bibliography

About the NKJP Project. Retrieved May 26, 2019 from http://nkjp.pl/index.php?page=0&lang=1

Alba-Juez, L., Thompson, G. (2014), *Evaluation in Context.* Philadelphia: John Benjamins Publishing Company

Asghar, M. Z., Khan, A., Ahmad, S., Qasim, M., Khan, I. A. (2017) “Lexicon-enhanced sentiment analysis framework using rule-based classification scheme”. *PLoS ONE* 12(2): e0171649. Retrieved January 4, 2020 from https://doi.org/10.1371/journal.pone.0171649

Axiology, Retrieved January 4, 2020 from https://www.britannica.com/topic/axiology

Corpus Encoding Standard. Retrieved June 8, 2019 from https://www.cs.vassar.edu/CES/CES1-0.html

Corpus of Contemporary American English. Retrieved May 26, 2019 from https://www.english-corpora.org/coca/

Dictionary definition. *oxforddictionaries.com.* Retrieved June 8, 2019 from https://en.oxforddictionaries.com/definition/dictionary

Hutto, C.J. & Gilbert, E.E. (2014). *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014. Retrieved May 26, 2020 from https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/download/8109/8122

Krzeszowski, T. (1997). *Angels and Devils in Hell: Elements of Axiology in Semantics*. Warszawa: Wydawnictwo Energeia

Lindquist, H. (2009). *Corpus Linguistics and the Description of English* Edinburgh University Press Ltd

McEnery, T., Wilson, A. (2001) *Corpus Linguistics: An Introduction, 2nd ed.* Edinburgh University Press Ltd

Meyer, C. F. (2004). *English Corpus Linguistics: An Introduction*. Cambridge University Press

Pęzik, P. (2016). "Exploring phraseological equivalence with Paralela". *In:* Gruszczyńska, Ewa; Leńko-Szymańska, Agnieszka. (eds.) Polskojęzyczne korpusy równoległe. Polish-language Parallel Corpora. Warszawa: Instytut Lingwistyki Stosowanej Retrieved June 20, 2020 from https://depot.ceon.pl/handle/123456789/13396

Regressive Imagery Dictionary Retrieved May 5, 2020 from https://www.kovcomp.co.uk/wordstat/RID.html

Rodríguez A, Argueta C. and Chen Y. (2019). "Automatic Detection of Hate Speech on Facebook Using Sentiment and Emotion Analysis," *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, Okinawa, Japan, 2019, pp. 169-174, doi: 10.1109/ICAIIC.2019.8669073.

Stuart, K., Botella, A., Ferri, I. (2016) "A Corpus-Driven Approach to Sentiment Analysis of Patient Narratives" *EPiC Series in Language and Linguistics* 1, 381-395. Retrieved May 26, 2019 from https://easychair.org/publications/download/r8SR

Taboada, M. (2016). "Sentiment Analysis: An Overview from Linguistics". *Annual Review of Linguistics*, 2. Retrieved January 4, 2020 from https://www.researchgate.net/publication/283954600_Sentiment_Analysis_An_Overview_from_Linguistics

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M. (2011). "Lexicon-Based Methods for Sentiment Analysis". *Computational Linguistics*, 37, 267-307. Retrieved January 4, 2020 from https://doi.org/10.1162/COLI_a_00049 .

Taj S., Shaikh B. B. and Fatemah Meghji A. (2019). Sentiment Analysis of News Articles: A Lexicon based Approach, *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, Sukkur, Pakistan, 2019, pp. 1-5, doi: 10.1109/ICOMET.2019.8673428.

VADER Sentiment Analysis Retrieved May 5, 2020 from https://github.com/cjhutto/vaderSentiment

Wang X., Ounis I., Macdonald C. (2019). Comparison of Sentiment Analysis and User Ratings in Venue Recommendation. In: Azzopardi L., Stein B., Fuhr N., Mayr P., Hauff C., Hiemstra D. (eds) Advances in Information Retrieval. ECIR 2019. Lecture Notes in Computer Science, vol 11437. Springer, Cham

Wilkinson, M. (2006) *Compiling Corpora for Use as Translation Resources.* Retrieved June 8, 2019 from https://translationjournal.net/journal/35corpus.htm

Abstract

Sentiment analysis is a field in computational linguistics which deals with the automatic recognition of the sentiment of a text and classifying it as positive or negative. Currently, this process is commonly applied to short texts such as product reviews and content published on social media. In this work, that process is used for longer text samples exemplified by the transcripts of episodes of the popular Black Mirror series. The study aimed to assess the extent to which computer analysis of sentiment can be useful in examining texts of such type, identifying problems posed by the specificity of the domain and proposing improvements and applications for sentiment analysis in the film industry.

The theoretical part of the work focuses on describing terms related to the corpora, axiology and sentiment analysis itself. It describes examples of commonly used corpora, such as Corpus of Contemporary American English (COCA) or Brown Corpus, procedures needed for compiling custom corpora and the processes which a corpus can undergo, in particular part-of-speech tagging and parsing. Axiology, generally understood as a study of the nature of values, is described from a linguistic perspective to explain how the positive-negative scale can be used to describe the semantic aspects of language.

The study presented in the work is based on the sentiment analysis of a custom made corpus consisting of transcripts of episodes of the Black Mirror series (the Black Mirror Corpus) using VADER Sentiment Analysis library and statistical analysis of the frequencies of words corresponding to particular emotions, classified basing on Martindale's Regressive Imagery Dictionary. The results of the research allowed the identification of problems occurring in the analysis of longer texts, which could be solved by expanding the list of modifiers, adding rules for calculating the sentiment value in the case of word repetition and applying POS tagging and word sense disambiguation. The results of the study show that sentiment analysis of films and television series, especially the automatic detection of dominating emotions, has potential use in the film industry. Additional information obtained as a result of such analysis can help to improve the parental control and film recommender systems.

Streszczenie po polsku

Analiza sentymentu jest dziedziną zajmująca się automatycznym rozpoznawaniem wydźwięku tekstu i klasyfikowaniu go jako pozytywny lub negatywny. Obecnie proces ten często jest aplikowany do krótkich tekstów, takich jak recenzje produktów i treści publikowane na mediach. W niniejszej pracy proces ten został zastosowany dla dłuższych próbek tekstu, którymi były transkrypcje odcinków serialu Black Mirror. Celem badania było ocenienie w jakim stopniu komputerowa analiza wydźwięku może być przydatna w badaniu tekstów tego typu. Dodatkowo, rozpoznano problemy jakie stwarza specyfika dzedziny i zaproponowano usprawnienia oraz zastosowania dla analizy sentymentu w przemyśle filmowym.

Część teoretyczna pracy skupia się na przybliżeniu zagadnień związanych z korpusami, aksjologią oraz samą analizą sentymentu. Opisuje przykłady powszechnie stosowanych korpusów, takich jak Korpus współczesnego amerykańskiego angielskiego (Coprus of Contemporary American English – COCA) czy Korpus Browna (the Brown Corpus), sposoby tworzenia własnych korpusów oraz procesy, jakim można poddać korpus, w szczególności anotację części mowy oraz parsowanie składniowe. Aksjologia, ogólnie pojmowana jako analiza natury wartości, jest opisana z perspektywy językoznawczej w celu przybliżenia w jaki sposób skala pozytywne-negatywne może być wykorzystwana do opisywania semantycznych aspektów języka.

Badanie opisane w pracy opierało się na analizie sentymentu korpusu złożonego z transkrypcji odcinków serialu Black Mirror za pomocą biblioteki VADER Sentiment Analysis oraz poprzez statystyczną analizę częstosci występowania słów odpowiadającym poszczególnym emocjom sklasyfikowanym na podstawie słownika Martindale'a. Wyniki badań pozwoliły na rozpoznanie problemów związanych z analizą dłuższych tekstów, które mogłyby zostać rozwiązane poprzez rozszerzenie listy słów modyfikujących sentyment, dodanie reguł dotyczących obliczania wartości sentymentu w przypadku powtarzających się słów oraz zastosowanie anotacji części mowy i rozróżniania znaczeń słów. Rezultaty badania wskazuja na to, że analiza sentymentu tekstów seriali i filmów, zwłaszcza automatyczne rozpoznawanie dominujących emocji danego tekstu, posiada potencjał do wykorzystania w przemyśle filmowym. Dodatkowe informacje otrzymane w wyniku takiej analizy mogą pozwolić na usprawnienie systemów rekomendujących filmy oraz systemów kontroli rodzicielskiej.

OŚWIADCZENIE

Oświadczam, że przedłożona praca dyplomowa została przygotowana przeze mnie samodzielnie, nie narusza praw autorskich, interesów prawnych i materialnych innych osób.

Wyrażam zgodę na udostępnienie osobom zainteresowanym mojej pracy dyplomowej dla celów naukowo-badawczych.

Zgoda na udostępnienie pracy dyplomowej nie oznacza wyrażenia zgody na kopiowanie pracy dyplomowej w całości lub w części.

Potwierdzam, że jest mi znana treść przepisu § 4 ust. 1 zarządzenia nr 35/R/20 Rektora Uniwersytetu Gdańskiego z dnia 31 marca 2020 r. w sprawie szczególnego trybu składania prac dyplomowych i przeprowadzania egzaminów dyplomowych w okresie zagrożenia zakażeniem koronawirusem SARS-Cov-2, zgodnie z którym przesłanie pracy dyplomowej z dołączonymi oświadczeniami w postaci elektronicznej jest równoznaczne ze złożeniem przez studenta pisemnych oświadczeń dołączanych do pracy.