University of Sheffield

Identifying complaints in social media using deep learning with transformers

Nitin Sunny Mathew

Supervisor: Nikolaos Aletras

A report submitted in fulfilment of the requirements
for the degree of MSc in Data Analytics

in the

Department of Computer Science

September 13, 2023

Declaration

All sentences or passages quoted in this report from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure in this project and the degree examination as a whole.

Name: Nitin Sunny Mathew

Signature: Nitin Sunny Mathew

Date: 13-Sep-2023

Abstract

A complaint is a statement made by a person or an entity with the intent to indicate something is unacceptable or unsatisfactory. This is commonly used in various aspects of day-to-day life including when conducting business operations. With the proliferation of social media across our lives and the active enablement of such platforms by organisations for user engagement, it has become a common medium for users to raise complaints. With such complaints being publicly visible, it is imperative for organisations to identify, prioritise and respond to these complaints swiftly. Automatically identifying complaints in social media is an active area of research. In the past few years, the focus has been on using NLP approaches driven by developments in transfer learning and transformer-based models.

In this paper, the use of these approaches are extended by assessing 'lightweight' transformer based models such as DistillBERT, MobileBERT, BERT tiny/small which are meant to reduce the time required for fine-tuning as well as inference. The performance of these 'lightweight' models is compared with the traditional transformer models including BERT, ROBERTA, BERTweet for this particular task. The dataset used consists of anonymised and annotated(complaint or not) Twitter data utilized in previous research and currently available in the public domain. In addition, the act of complaining and the nature of complaints are analysed from a linguistic perspective along with discussions on state-of-the-art approaches for such NLP tasks.

Update with high level results

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

In the act of complaining, dissatisfaction or annoyance is expressed by a person or entity in response to a previous or ongoing event that has negatively impacted them [14]. There is a breach of expectation and the act of complaining provides an avenue to direct dissatisfaction to the appropriate organisation or individual with the hope of rectification or redressal. It could also be used as a means to issue a Face Threatening Act [6], to the detriment of the receipient's reputation of the complaint. The event or action could be concerning a product or service procured by the concerned person or entity. The need to recognise, acknowledge and act on complaints is of significant importance to businesses and organisations to retain their customers while maintaining their reputations.

Until the advent of online platforms and specifically social media, the impact of negative word-of-mouth was confined to a relatively limited audience. However, since then complaints posted online have the potential to rapidly go viral, reaching millions of individuals and significantly damaging a company's brand reputation and goodwill in a short period [21]. Customers are able to express their complaints directly, conveniently, and with enhanced effectiveness to organisations through multiple social media channels and platforms [2].

By examining instances of complaints on social media and specifically Twitter, we find them in alignment with the previously described act of complaining. These examples as shown in Table 1.1 are of individuals who have encountered breaches of their expectations. Regarding the intentions underlying these complaints, we find the objective being rectification in the first and second examples. In the first tweet, there is a request for a specific software version to resolve an issue, while the second tweet seeks clarification on a policy due to the perceived violation arising from a wrongly advertised product. Additionally, one could argue that the second tweet encompasses an element of a Face Threatening Act, given that implicating false advertising in the context of a company potentially could harm the brand's reputation. In contrast, the third and fourth tweets are instances of issuing a Face Threatening Act. They

No.	Example complaints from Twitter
1	hi please i cant find a driver for video card (nvidia geforce 8500 gt) for mac please send me a link when i can download a driver
2	what is your policy on false advertising regarding sale items ? i was refused a sale in westfield due to a company error on pricing
3	thanks to <user> ' s incompetence i now can't work till october 4th , when the ati card arrives .
4	you jave the worst customer service #pissed #useless #worstbrand

Table 1.1: Sample complaints extracted from Twitter, exhibiting diverse degrees of complaint expression and severity. These complaints are sourced from data that has undergone the preprocessing steps outlined in Chapter 3.

are written with the intention to harm the brand's value considering the use of terms such as *incompetence*, *worst customer service* and hashtags like *#pissed*, *#useless* and *#worstbrand*.

In addition to the timely addressing of customer complaints, automated detection of complaints in natural language has several other purposes. Linguists could gain a more detailed understanding of the context, intent, and various types of complaints on a larger scale while psychologists could utilise this information to identify the underlying human traits that drive the behaviour and expression of complaints. Developing downstream natural language processing (NLP) applications, such as dialogue systems is another use case of this task [15].

Attempting to identify complaints manually through the multitude of posts and streams coming through the various social media channels is neither practical nor scalable. Various approaches to automate this task have been explored. The traditional vector-space method utilizing dictionaries has been applied in other text classification tasks [11]. Latent Semantic Indexing based on Singular Value Decomposition along with linguistic style features has been utilised to classify emails as complaints or not [7]. In recent years, we have seen the use of various Machine learning and Natural Language Processing (NLP) based approaches for similar classification problems. The performance of logistic regression over various types of feature spaces against neural-network based models like Multi-layer Perceptron (MLP) and Long Short Term Memory (LSTM) has been analysed by [15] on Twitter feeds. The use of more advanced approaches using transformer networks has shown to have better results as explored by [9]. As part of this paper, the use of the BERT and its many variants, including that of recently created lightweight versions will be assessed further on a publicly available Twitter dataset.

1.2 Aims and Objectives

TO UPDATE

Chapter 2

Literature Survey

2.1 The act of complaining

As per [14], the speech act of complaining in the traditional sense can be understood from the perspective of the speaker stating their displeasure or dissatisfaction to a target entity or individual. This is done as a reaction to an unfavourable event that is currently taking place or has already occurred. The authors believe a few preconditions have to be satisfied to result in a complaint being made. This includes the speaker's belief the entity or individual is responsible for the unfavourable outcome and that the speaker in question suffers from the consequences. The result is a verbally expressed complaint.

This expression of complaint could be carried out in various ways. The speaker might choose to directly communicate their complaints or concerns to the individual or entity, either immediately after the incident or at a later time. Or they might voice their grievances to others through word-of-mouth or they could even opt to escalate the issue by involving a third party, such as a consumer advocacy office [19].

The authors of [14] further delve into the intentions of the speaker in making the complaint. They argue this is carried out with either the hope of repair of the situation or as a 'Face Threatening Act' [6], with the purpose being to damage the face of the individual or entity against whom the complaint is made. In this scenario, a face-threatening act refers to an action that challenges the reputation of the recipient by going against what the recipient desires. These acts can manifest in a verbal form including with variation in tone or inflection or using non-verbal methods.

While such complaints could be considered direct complaints as per [5], the authors additionally highlight the use of indirect complaining in speech. In the case of indirect complaints, the speaker does not attribute responsibility for the cause of the complaint to the individual or entity being addressed. The authors theorise, an indirect complaint is used to bring about 'solidarity' between speakers, which is contrary to the use of direct complaints. It can serve as a means to initiate conversations and establish temporary connections with others. The scope of the data for this project (described in the subsequent chapter) is primarily focused on direct complaints as they are selected based on tweets being addressed to a brand's customer service handle. However it is possible, tweets which fall into the category of indirect complaints are also included in the dataset.

Analysing deeper into which types of customers complain more, [18] have looked at how personality traits like impulsivity and self-monitoring impact customer complaining behaviour. *Impulsivity*, as defined by [16], refers to a consistent inclination of customers to act spontaneously and immediately, without much reflection or careful consideration of available options or potential consequences. This trait remains relatively stable over time for such customers. [3] defines *self-monitoring* as the propensity to adjust one's behaviour based on the actions or behaviour of others. High self-monitoring individuals are sensitive to others' expressions and behaviour, relying on social cues for their actions, while low self-monitoring individuals may be influenced by personal traits. From their experiments, [18] concluded that individuals with high impulsiveness tend to complain more than those with low impulsiveness, whereas individuals with high self-monitoring tend to complain less than those with low self-monitoring. However, these effects are more pronounced in situations where the level of dissatisfaction is high.

2.2 Complaining online

The act of complaining exists online in various forms and with varying degrees of intensity and this prevalence lead to the emergence of third-party organisations that provide online channels for customers' ease and convenience [21]. Notably, there are complaint websites like complaintsboard.com, review websites like trustpilot.com as well as consumer organisations' sites such as consumeraffairs.com, where customers can share their negative experiences and exchange information with others. The impact of negative word of mouth is quite high due to the ease with which negative reports can rapidly reach millions of people, potentially causing significant harm to a company's brand. Various user-generated content platforms such as YouTube, Twitter, and Facebook serve as spaces for expressing complaints. Brands use these platforms for user engagement and this provides the users with the required visibility to potentially raise or escalate an issue. With numerous such options available online, companies can experience significant repercussions arising from actions taken by dissatisfied customers [21].

Of the 431 online complaints assessed by [21], 96% followed what they call a double deviation. This occurs when customers experience both a product or service failure followed by multiple unsuccessful attempts to resolve the issue, resulting in them feeling they have been violated twice. Such customers then resort to online complaining. Their urge to complain online is driven by how they felt betrayed rather than simply being dissatisfied or with any form of malicious intentions to hinder business operations.

Complaining online is also associated with electronic word-of-mouth or EWOM, which involves sharing information online with a wider group, and it remains accessible over an extended period while often being anonymous [8]. This type of communication can take place on various platforms, ranging from official company-sponsored sites to unaffiliated blogs. The Internet offers consumers a convenient and anonymous platform to express negative word-of-mouth by sharing their viewpoints and complaints with others. Among the different forms of EWOM, consumer reviews are particularly noteworthy, as they provide valuable insights about products, whether positive or negative [19]. Such Negative electronic WOM (EWOM), can significantly damage a brand's reputation and influence potential customers to seek alternative products or services.

Technology provides an accessible channel that allows consumers to complain with significant ease, making it available to anyone with internet access, even those who may be hesitant to complain directly to the company [19]. The reviews and comments posted by consumers online can hold considerable influence over decisions made by other fellow consumers. From an organisation's perspective, the use of online complaining by consumers has some indirect negative consequences as well. The potential experience and knowledge frontline personnel could gain from addressing the complaints directly are lost and this has long-term implications for the organisation [19].

2.3 Complaining in social media

TO UPDATE

2.4 Self-expression on Twitter

TO UPDATE

2.5 Transformers

TO UPDATE

2.6 Ongoing research

TO UPDATE

Chapter 3

Methodology

3.1 Task

For a short text segment, $T = \{t_1, t_2, ..., t_n\}$ where t_i is defined as a token, classify if the sequence of tokens T is a complaint or not.

3.2 Data and pre-processing

The data used for the experiments is from Twitter. Twitter provides a good representation of social media text due to the direct connection consumers have with organisations and brands as well as the ability to express oneself [15]. **Add content on why Twitter**

The data set created by [15] and further used by [9] is utilised for this project. The original process for collection and annotation employed by them is breifly described below. The particular version [1] used for the experiments is the one enhanced by [10] with the addition of labels for the severity of complaints. These additional labels are not used for the experiments in this project.

3.2.1 Domains and organisations

A cross-industry representative collection of 93 customer service handles of organisations on Twitter were identified manually. These handles were then categorised into 9 domains based on their industry type. Since an organisation could have business activities across domains, the assigned domain was based on the products or services receiving the most number of complaints. All the domains used in the experiments are listed in Table 3.1.

[1] The data can be found here - https://archive.org/details/complaint_severity_data

Domains	Complaints	Non-Complaints	Total Tweets
Food & Beverage	95 (73%)	35 (27%)	130 (7%)
Apparel	141 (55%)	117 (45%)	258 (13%)
Retail	124 (62%)	75 (38%)	199 (10%)
Cars	67 (73%)	25 (27%)	92 (4%)
Services	207 (61%)	130 (39%)	337 (17%)
Software & Online Services	189 (65%)	103 (35%)	292 (15%)
Transport	139 (56%)	109 (44%)	248 (12%)
Electronics	174 (61%)	112 (39%)	286 (15%)
Other	96 (79%)	33 (21%)	129 (7%)
Total	1232 (63%)	739 (37%)	1971

Table 3.1: The nine domains and the distribution of tweets that are complaints and those that are not. The percentages indicate how the splits are distributed.

3.2.2 Data Extraction

The data was extracted from Twitter via the Twitter API [2]. The latest 3,200 tweets at the time of the collection exercise were extracted and the original tweets to which the customer service handles responded were identified. Then, random sampling equally for each handle, 1,971 tweets were identified where there was a response from the support's handle. To ensure a more balanced and diverse dataset, 1,478 randomly sampled tweets were added to the dataset. 739 tweets were replies to other handles (outside the 93 identified) and the remaining 739 tweets were not addressed to any Twitter handle. Table 3.2 shows the breakdown of the total population of the tweets dataset. Tweets were filtered for English using langid.py [12]. Retweets were excluded and all usernames and URLs were anonymised and replaced with placeholder tokens.

Extraction Criteria	Complaints	Non-Complaints	Total Tweets
Addressed to and replied by the identified 93 customer service handles	1239 (63%)	739 (37%)	1971 (58%)
Addressed to other customer service handles	0	739 (100%)	739 (21%)
Not addressed to any Twitter handle	0	739 (100%)	739 (21%)
Total	1232 (36%)	2217 (64%)	3449

Table 3.2: Selection of tweets based on random sampling and where they have received replies when addressed to the 93 customer service handles combined with random sampled tweets that are addressed to other handles and tweets that are not addressed to any handle.

[2]https://developer.twitter.com/en

3.2.3 Annotation

The classification of the 1,971 tweets as complaints or not was carried out using a binary annotation task (complaint or not). Since tweets are concise and typically express a single idea, an entire tweet was classified as a complaint if it contained at least one speech act of complaining. To guide the annotation process, a complaint definition from [14], stating that a complaint portrays a situation that contradicts the writer's positive expectation was used. Two of the authors with extensive annotation experience in linguistics independently labelled the 1,971 tweets. They had substantial agreement [1] with Cohen's Kappa of $\kappa = 0.731$. In the end, 1,232 tweets (63%) and 739 tweets (37%) were identified as complaints and non-complaints. Table 3.1 gives the breakdown of the complaint and non-complaint tweets for each domain.

3.3 Environment and models

3.3.1 Hardware

The key details of the environment used for the experiments are listed below.

- **CPU Count:** 8
- **Memory:** 45 GB
- **GPU Count:** 1
- **GPU Model:** NVIDIA RTX A4000
- **GPU Memory:** 16 GB

3.3.2 Software

For the experiments, the BERT transformer large language models along with a number of its variants are used to classify the tweets and compare the performance. The models are based on the *transformers* library implementation from Hugging Face[3] along with the *datasets* and *evaluate* libraries. From scikit-learn[4] the *sklearn* library is used to generate the stratified splits for the nested cross-fold validation. The versions for each library are shown in the table 3.3.

3.3.3 Models

The transformer models used are listed in table 3.4 along with the no. of parameters for each of them. The no. of parameters is based on the embedding and output layers along with the attention heads. The models chosen are such that there is a wide range of parameter counts for the models. This allows for a comparison of the model performance both in terms of the

[3]https://huggingface.co/
[4]https://scikit-learn.org/stable/

Provider	Library Name	Version
Hugging Face	transformers	4.21.3
	datasets	2.4.0
	evaluate	0.4.0
Scikit-Learn	sklearn	1.1.2
Numpy	numpy	1.23.4
Pandas	pandas	1.5.0

Table 3.3: Software and library versions used for this project. Other more

predictions as well as the inference time in relation to the number of parameters used in the models. **Add content on the impact of layers and parameters on model performance**

Model	Parameter Count	Tokenizer Type	Vocab. Size
ROBERTA base	125M	Byte-level BPE	50,265
BERT base (uncased)	110M	WordPiece	30,522
BERTweet base	110M	Byte-Pair Encoding (BPE)	64,000
DistilBERT base (uncased)	66M	WordPiece	30,522
MobileBERT (uncased)	25.3M	WordPiece	30,522
ALBERT base	11M	SentencePiece	30,000
BERT Tiny	4.4M	WordPiece	30,522

Table 3.4: The transformer models used for the experiments along with type of tokenization, and vocabulary size and sorted by the number of parameters for each of them. The parameter counts are from [4] for Roberta, Bert, Albert and Bert Tiny. For Bertweet it is from [13], MobileBert from [20] and Distilbert from [17].

3.4 Data tokenisation

The tokenization process is required to be applied to input data for it to be prepared appropriately for use by BERT and its variants. The tokenization process involves dividing the input text into tokens based on a predefined set of rules. These tokens are subsequently transformed into numerical representations and tensors, along with any extra inputs needed by the model. Tokens in general could be words, subwords, phrases or even characters. The models in scope of the experiments use one of the Byte-Pair Encoding (BPE), Byte-level BPE, WordPiece or SentencePiece tokenization processes and are shown in Table 3.4.

Byte-Pair Encoding or BPE works by iteratively combining the most frequently occurring pairs of characters or subwords within a corpus until a predefined vocabulary size is reached or after reaching a maximum number of iterations. The vocabulary will consist of a set of subwords, which can include characters, character sequences, or partial words. Byte-level BPE works similarly to BPE but operates at byte-level, treating each byte of a text as a token and merging the most frequent pairs of bytes in a text corpus.

The tokenizer provided by the *transformers* library is used for tokenizing the input tweets. The library provides various model-specific tokenizers such as, *BertTokenizer* [5] or *RobertaTokenizer* [6]. Certain models like BERTweet do not use dedicated tokenizers and leverage existing ones. For the experiments, the *AutoTokenizer* [7] has been used which conviniently selects the appropriate tokenizer relevant for the model in use.

** Talk about the default settings from the preprocess page**

Tweet from input dataset

love it when i almost die rear ended by a semi cause my jeep turns off again . one day they will fix it #jeepsucks #chrysler

Post tokenisation

[101, 2293, 2009, 2043, 1045, 2471, 3280, 4373, 3092, 2011, 1037, 4100, 3426, 2026, 14007, 4332, 2125, 2153, 1012, 2028, 2154, 2027, 2097, 8081, 2009, 1001, 14007, 6342, 10603, 1001, 17714, 102, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

Decoding the tokenized input

[CLS] love it when i almost die rear ended by a semi cause my jeep turns off again. one day they will fix it # jeepsucks # chrysler [SEP] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD] [PAD]

3.4.1 Before and after tokenization

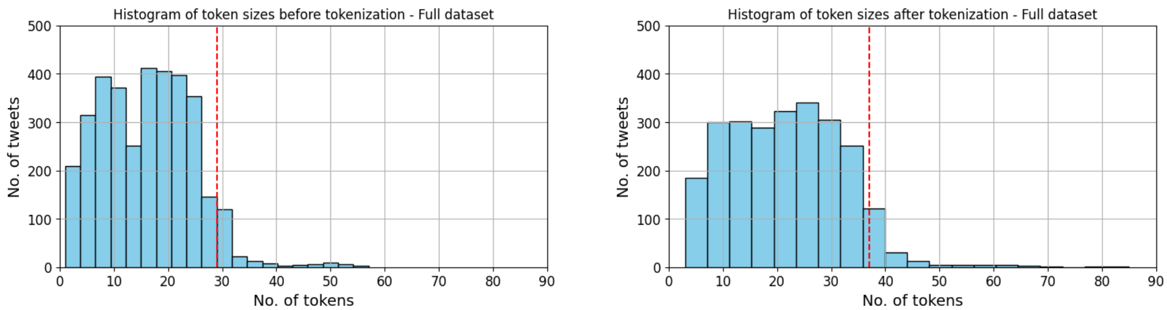
The distribution of the number of tokens in the tweets from the pre-processed data from [10] before applying the model-specific tokenization is shown in Figure 3.1a. We have over 95% of the tweets having 29 tokens or less. Using the BertTokenizer as an example, as shown in Figure 3.1b, we find we need about 37 tokens to completely cover 95% of the tweets. This analysis assists in the decision on the appropriate value for the *max_ value* for the tokenizer. This when used in conjunction with *truncation=True*, it sets the maximum number of tokens

[5] `https://huggingface.co/docs/transformers/v4.21.3/en/model_doc/bert#transformers.BertTokenizer`

[6] `https://huggingface.co/docs/transformers/v4.21.3/en/model_doc/roberta#transformers.RobertaTokenizer`

[7] `https://huggingface.co/docs/transformers/v4.21.3/en/model_doc/auto#transformers.AutoTokenizer`

for each input tweet. Anything that follows is truncated and not used for the training or inference. Based on this, a *max_value=50* is selected for the experiments.



(a) Before tokenization

(b) After tokenization

Figure 3.1: The token count distribution for the full dataset of 3,449 tweets before and after tokenization with the red dashed line indicating 95% coverage of tweets. BertTokenizer is used here.

3.5 Nested cross validation

** To UPDATE **

3.6 Domain splits with nested cross-validation

** To UPDATE **

3.7 Ethical, Professional and Legal Issues

** TO UPDATE **

Chapter 4

Results and discussion

4.1 Risk Analysis

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem. Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue. Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus. Maecenas tempus, tellus eget condimentum rhoncus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum. Nam quam nunc, blandit vel, luctus pulvinar, hendrerit id, lorem. Maecenas nec odio et ante tincidunt tempus. Donec vitae sapien ut libero venenatis faucibus. Nullam quis ante. Etiam sit amet orci eget eros faucibus tincidunt. Duis leo. Sed fringilla mauris sit amet nibh. Donec sodales sagittis magna. Sed consequat, leo eget bibendum sodales, augue velit cursus nunc.

4.2 Project Plan

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem. Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue.

Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus. Maecenas tempus, tellus eget condimentum rhoncus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum. Nam quam nunc, blandit vel, luctus pulvinar, hendrerit id, lorem. Maecenas nec odio et ante tincidunt tempus. Donec vitae sapien ut libero venenatis faucibus. Nullam quis ante. Etiam sit amet orci eget eros faucibus tincidunt. Duis leo. Sed fringilla mauris sit amet nibh. Donec sodales sagittis magna. Sed consequat, leo eget bibendum sodales, augue velit cursus nunc.

4.3 Another Section if You Need It

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem. Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue. Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus. Maecenas tempus, tellus eget condimentum rhoncus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum. Nam quam nunc, blandit vel, luctus pulvinar, hendrerit id, lorem. Maecenas nec odio et ante tincidunt tempus. Donec vitae sapien ut libero venenatis faucibus. Nullam quis ante. Etiam sit amet orci eget eros faucibus tincidunt. Duis leo. Sed fringilla mauris sit amet nibh. Donec sodales sagittis magna. Sed consequat, leo eget bibendum sodales, augue velit cursus nunc.

Chapter 5

Conclusions

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem. Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue. Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus. Maecenas tempus, tellus eget condimentum rhoncus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum. Nam quam nunc, blandit vel, luctus pulvinar, hendrerit id, lorem. Maecenas nec odio et ante tincidunt tempus. Donec vitae sapien ut libero venenatis faucibus. Nullam quis ante. Etiam sit amet orci eget eros faucibus tincidunt. Duis leo. Sed fringilla mauris sit amet nibh. Donec sodales sagittis magna. Sed consequat, leo eget bibendum sodales, augue velit cursus nunc.

Bibliography

[1] Artstein, R., and Poesio, M. Inter-Coder Agreement for Computational Linguistics. *Computational linguistics - Association for Computational Linguistics 34*, 4 (2008), 555–596. Place: One Rogers Street, Cambridge, MA 02142-1209, USA Publisher: MIT Press.

[2] Balaji, M. S., Jha, S., and Royne, M. B. Customer e-complaining behaviours using social media. *The Service industries journal 35*, 11-12 (2015), 633–654. Place: London Publisher: Routledge.

[3] Becherer, R. C., and Richard, L. M. Self-Monitoring as a Moderating Variable in Consumer Behavior. *The Journal of consumer research 5*, 3 (1978), 159–162. Place: CHICAGO Publisher: Journal of Consumer Research.

[4] Bhargava, P., Drozd, A., and Rogers, A. Generalization in NLI: Ways (Not) To Go Beyond Simple Heuristics, Oct. 2021. arXiv:2110.01518 [cs].

[5] Boxer, D. Social distance and speech behavior: The case of indirect complaints. *Journal of pragmatics 19*, 2 (1993), 103–125. Place: AMSTERDAM Publisher: Elsevier B.V.

[6] Brown, P., and Levinson, Stephen C. *Politeness : some universals in language usage.* Studies in interactional sociolinguistics ; 4. Cambridge University Press, Cambridge [Cambridgeshire] ; New York, 1987.

[7] Coussement, K., and Van den Poel, D. Improving customer complaint management by automatic email classification using linguistic style features as predictors. *DECISION SUPPORT SYSTEMS 44*, 4 (2008), 870–882. Place: AMSTERDAM Publisher: Elsevier B.V.

[8] Hennig-Thurau, T., Gwinner, K. P., Walsh, G., and Gremler, D. D. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of interactive marketing 18*, 1 (2004), 38–52. Place: Hoboken Publisher: Elsevier Inc.

[9] Jin, M., and Aletras, N. Complaint Identification in Social Media with Transformer Networks.

[10] Jin, M., and Aletras, N. Modeling the Severity of Complaints in Social Media.

[11] Liang, C.-Y., Guo, L., Xia, Z.-J., Nie, F.-G., Li, X.-X., Su, L., and Yang, Z.-Y. Dictionary-based text categorization of chemical web pages. *Information Processing & Management 42*, 4 (2006), 1017–1029.

[12] Lui, M., and Baldwin, T. langid.py: An Off-the-shelf Language Identification Tool. In *Proceedings of the ACL 2012 System Demonstrations* (Jeju Island, Korea, July 2012), Association for Computational Linguistics, pp. 25–30.

[13] Nguyen, D. Q., Vu, T., and Nguyen, A. T. BERTweet: A pre-trained language model for English Tweets, Oct. 2020. arXiv:2005.10200 [cs].

[14] Olshtain, E., and Weinbach, L. Complaints: A study of speech act behavior among native and non-native speakers of Hebrew. 195–208.

[15] Preotiuc-Pietro, D., Gaman, M., and Aletras, N. Automatically Identifying Complaints in Social Media.

[16] Rook, D. W., and Fisher, R. J. Normative Influences on Impulsive Buying Behavior. *The Journal of consumer research 22*, 3 (1995), 305–313. Place: CARY Publisher: University of Chicago Press.

[17] Sanh, V., Debut, L., Chaumond, J., and Wolf, T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, Feb. 2020. arXiv:1910.01108 [cs].

[18] Sharma, P., Marshall, R., Alan Reday, P., and Na, W. Complainers versus non-complainers: a multi-national investigation of individual and situational influences on customer complaint behaviour. *Journal of marketing management 26*, 1-2 (2010), 163–180. Place: Helensburg Publisher: Taylor & Francis.

[19] Sparks, B. A., and Browning, V. Complaining in Cyberspace: The Motives and Forms of Hotel Guests' Complaints Online. *Journal of hospitality marketing & management 19*, 7 (2010), 797–818. Publisher: Taylor & Francis Group.

[20] Sun, Z., Yu, H., Song, X., Liu, R., Yang, Y., and Zhou, D. MobileBERT: a Compact Task-Agnostic BERT for Resource-Limited Devices, Apr. 2020. arXiv:2004.02984 [cs].

[21] Tripp, T. M., and Gregoire, Y. When unhappy customers strike back on the Internet. *MIT Sloan management review 52*, 3 (2011), 37. Place: Cambridge Publisher: Sloan Management Review.

Appendices

Appendix A

References

A.1 Model References

Model	Model Documentation
ALBERT base	`https://huggingface.co/albert-base-v2`
BERT base (uncased)	`https://huggingface.co/bert-base-uncased`
BERT Tiny	`https://huggingface.co/prajjwal1/bert-tiny`
BERTweet base	`https://huggingface.co/vinai/bertweet-base`
DistilBERT base (uncased)	`https://huggingface.co/distilbert-base-uncased`
MobileBERT (uncased)	`https://huggingface.co/google/mobilebert-uncased`
ROBERTA base	`https://huggingface.co/roberta-base`

Table A.1: The transformer models used for the experiments and links to their documentation.

Appendix B

Another Appendix

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem. Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue. Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus. Maecenas tempus, tellus eget condimentum rhoncus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum. Nam quam nunc, blandit vel, luctus pulvinar, hendrerit id, lorem. Maecenas nec odio et ante tincidunt tempus. Donec vitae sapien ut libero venenatis faucibus. Nullam quis ante. Etiam sit amet orci eget eros faucibus tincidunt. Duis leo. Sed fringilla mauris sit amet nibh. Donec sodales sagittis magna. Sed consequat, leo eget bibendum sodales, augue velit cursus nunc.