Modelling the effect of Seattle City Council Covid-19 policy on the emotionality and sentiment of citizen submitted work requests.

Peter C. Hardisty [1] **Sloan C. Stinnett** [1]

1. Introduction

The COVID-19 pandemic has profoundly impacted various aspects of urban life, including the functioning and accessibility of city services. This project aims to investigate how the pandemic influenced the types and frequency of service requests made by residents of Seattle, as well as how the city's responses to these requests may have changed.

Seattle offers a wide range of municipal services through agencies such as the police, sanitation, and public works departments. Residents can request these services using various methods, including the "Find-It/Fix-It" app, which provides a direct line of communication with the city. The app provides resources for citizens to utilize to address/report problems that they see, as seen in Figure 1.

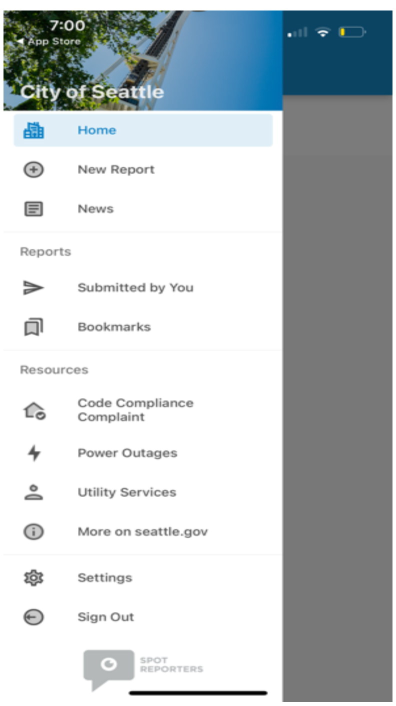

Figure 1. Homepage of "Find-It/Fix-It"

Through the app, citizens are able to describe the problem and give a location to the city to address the problem. And if the city needs additional information, they are able to contact the citizen. An example ticket submission can be seen in Figure 2.

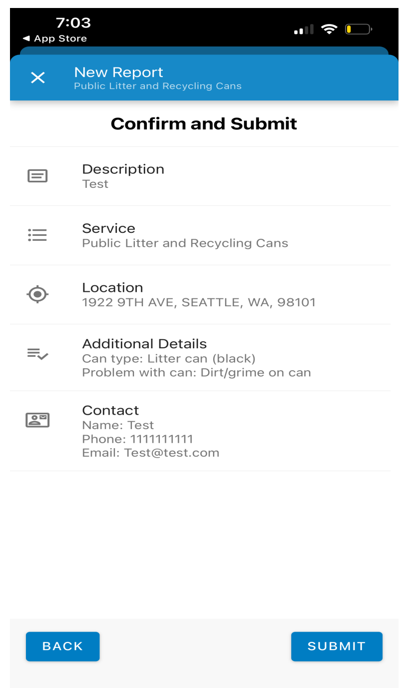

Figure 2. "Find-It/Fix-It" Example Submission

A complete list of the services provided are as follows:

1. Abandoned Vehicle/72 hr Parking Ordinance
2. Parking Enforcement
3. Scooter or Bike Share Issue
4. Found Pet
5. Graffiti
6. Lost a Pet
7. Nuisance dogs in a park
8. Unauthorized Encampment

9. Clogged Storm Drain
10. Overgrown Vegetation
11. Parks and Recreation Maintenance
12. Pothole
13. Street Sign Maintenance
14. Streetlight
15. Traffic Signal Maintenance
16. Dead Animal
17. Illegal Dumping/Needles
18. Public Litter and Recycling Cans

Given the significant disruptions caused by the COVID-19 pandemic, it is reasonable to assume that both the nature of these requests and the city's ability to respond were affected.

This analysis utilizes data obtained by Matthew Beattie of the University of Oklahoma through a Freedom of Information Act request, which was subsequently redacted and ingested into a PostgreSQL database for exploration. In this project, we leverage large language models (LLMs) to delve into the unstructured text fields of these service requests. Specifically, we used a fine-tuned version of the DistilRoBERTa model for emotional analysis and DistilBERT models for reclassification and sentiment analysis. Our goal is to uncover patterns and insights that reveal the pandemic, and its related policy changes, impact on city services, thereby informing future strategies for urban management and resilience.

2. Data

2.1. Find-It/Fix-It

The data for this project was obtained by Matthew Beattie through a Freedom of Information Act request to the city of Seattle for data relating to the "Find-It/Fix-It" app. Included in the data provided, were tables outlining the request tickets from citizens outlining the service request type, details of the request, creation of the ticket, and other information detailing the requests from citizens. Key variables used in the creation of the model were the date of creation of the ticket and details of the service request. There was approximately 777,500 requests in the database. This data was redacted of personal information such as SSNS, phone numbers, license plate numbers and driver's license numbers by city of Seattle Staff before being transferred to Matthew Beattie who ingested the records into a PostgreSQL database. For general exploration of the text fields and request occurrences the complete dataset was used. However, it was necessary to clean and sample the data for reclassification and sentiment analysis. the researchers handled rows that lacked information in the details section of the data by omitting these rows from sentiment and emotional analysis. After dropping these rows, there was 365,205 requests that had viable information for sentiment and emotion analysis.

2.2. King County Covid-19 Data

King county, the county that contains Seattle and it's suburbs, collected and released daily data on both hospitalizations due to COVID-19 and the number and variety of COVID vaccinations administered. Daily counts were released starting with February 8th 2020 and ending with April 30th 2024 (Kin). Data on vaccinations and county population were used to find the first day past which 75% of the population of Kings county had received their full regiment of vaccines December 11, 2021. This date was used to represent the end of the COVID pandemic in Kings county as vaccinations and past exposures would have past the keystone number of immunizations to have conferred herd immunity to the county's populace (Dowdy & D'Souza).

2.3. Data Preparation

2.3.1. Reclassification

Prior to sentiment and emotional classification, a process of data cleaning and preparation was required. Most central to this effort was ensuring consistency of classification scheme across the data set. A critical class of interest , Unauthorized Encampment, was not available across the entire span of the data set. This critical class for examining how the citizens of Seattle reacted to homelessness before, during, and after the pandemic was only available after June 27, 2022. However, several requests prior to this data had clearly fallen into the category and only been label otherwise due to a lack of the option in the application. In order to ameliorate this, it was decided that the use of a LLM for text classification would be the best option for allowing us to alter the classification scheme used on prior requests to match that of the most recent requests.

To accomplish this, the DistilBERT base model from Hugging Face was fined tuned on a random sample of 47,585 work requests that had detail text and came after the addition of the Unauthorized Encampment class in June of 2022. The remaining 11,896 records were reserved as a testing set. After fine-tuning for reclassification the DistilBERT model was able to achieve an 80% accuracy after 4 epochs of training (Beattie). Please note, Hugging Face is an open-source platform that facilitates the creation, training, and deployment of machine learning (ML) models, with a particular focus on natural language processing (NLP).

This reclassification had a noticeable effect on the distri-

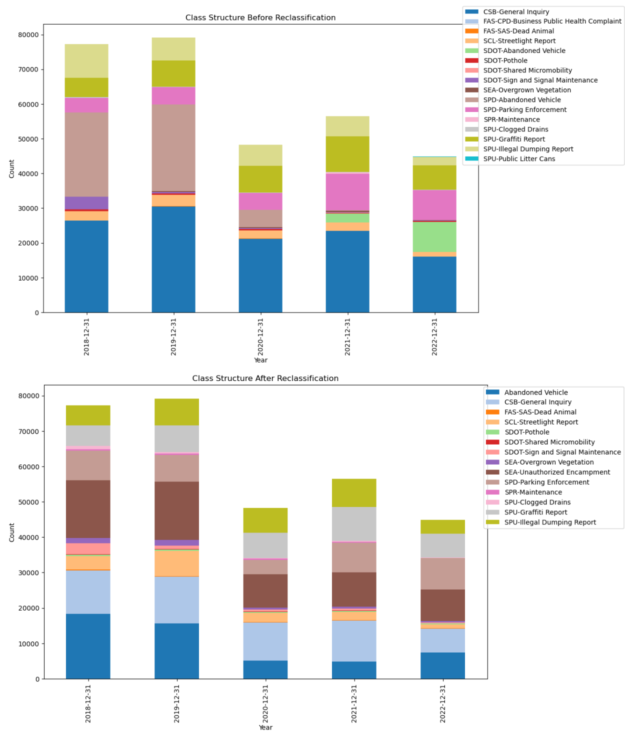

Figure 3. Distribution of Text Bearing Request Before and After Reclassification

bution of text bearing work requests. The most apparent change in the distribution is the disappearance of the class SPU-Public Litter Cans that was split nearly evenly by the model between the General Enquiry class and the Illegal Dumping class with many other classes receiving a small number of the remaining requests. The Unauthorized Encampment class received 60,787 work requests with the majority of those requests coming from the General Enquiry class (47,856) with other major contributors being Abandoned Vehicle (5,230), Illegal Dumping (3,789), and Parking Enforcement (2,232). For the majority of the remaining classes, save General Inquiry and Business Public Health Complaint, the model reclassified most of the request back into the same class.

2.3.2. Preliminary analysis and Grouping

Grouping

In order to simplify the analysis to a manageable size, the seventeen classes that existed in the original data were assigned into one of four groups: Parking, Homelessness, Clean-up, and Other. The parking group centered on classes that related to parking enforcement; namely parking enforcement and abandoned vehicles. Homelessness covered unauthorized encampments. Clean-up covered all requests

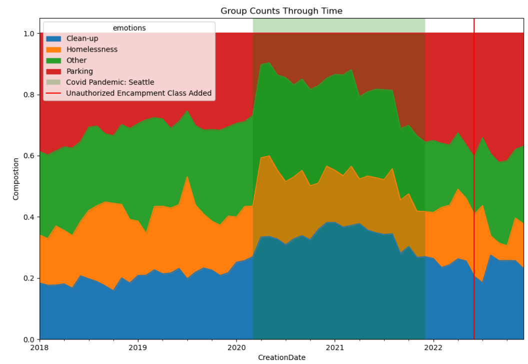

Figure 4. Distribution of Groups by Month

directed to SPU, SEA, and FAS. This includes requests such as clogged drain and illegal dumping. All other requests were put in the other group, including general inquiry and maintenance.

Looking to Figure 4 we see that there are few major swings in the distribution of of requests among the four groups outside a sharp decline in the parking group at the beginning of the pandemic. As we see later in the results section, this drop in parking related requests is most attributable to changes in Seattle parking policy due to the pandemic.

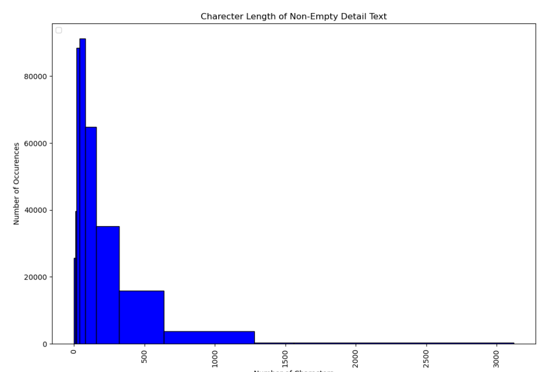

Figure 5. Distribution of Text Length for All Detail Text

Text Length

Of the roughly 777,500 service requests that were received 412,294 had no detail text what so ever. Of the remaining 365,206 requests, the average number of characters in the Detail Text was 95.2508. As we can see in Figure 5, the length of the Detail text has a heavy right tail with many Detail text extending into the hundreds of characters. Looking to Figure 6, we see text length broken out by group with clean-up and parking requests have relatively thin right tails and a high concentration of short messages while homelessness and the other category had more long messages. General inquiry and unauthorized encampment were both classes that had large messages, both due to the complexity of the information needed to specify the issues at hand and due to the fact that these classes were more often used to communicate more general frustration or political beliefs than the more utilitarian clean-up and parking related classes.

Key Words for Each Group

$$TFIDF_{i,j} = TF_{i,j} \cdot IDF_i$$

$$TF_{i,j} = \frac{\text{number of appearances of i in j}}{\text{number of words in j}}$$

$$IDF_i = log_2(\frac{\text{number of documents}}{\text{number of documents where i appears}})$$

Where i is a term and j is a document

To better grasp the overall themes of each group and ensure that they roughly match expectations, the top 5 most descriptive words for the group were selected using term frequency inverse document frequency (TFIDF). Looking to the preceding formulae we see that words with higher values are those that are frequent to a document and appear in few documents. For this calculation all detail text for every request in each category where concatenated and used as a single document. This gave us one document for each group. The top five words for each group can be seen in the following table with the terms with the highest TFIDF appearing higher up.

Parking	Clean-up	Homelessness	Other
Vehicle	Graffiti	Encampment	Street
PEO	Garbage	Homeless	Light
Parked	Trash	Camping	Sidewalk
Complaint	Sidewalk	Park	Ave
Area	Sign	Illegal	Customer

PEO stands for Parking Enforcement Officer

The top words for each group align with our general expectation of the groups. Parking has terms relating to parked vehicles, Clean-up to debris and possible locations that the trash is present, Homelessness has terms relating to camping. Other is insight full in that the other category contains both maintance and general inquiry and it seems that particular classes with other such as street light report lend its most distinctive language.

3. Model

3.1. BERT and RoBERTa for Sentiment and Emotion Analysis

Historically sentiment and emotion analysis has been a purely lexical task. Bag-of-words or Bag-of-ngram models were used to train classification models. These methods however have a natural disadvantage in that they treat text as simply a collection of words or small phrases rather than as a sequence with a natural order where far off words can have a large effect on the meaning of a certain word or phrase within context(El-Din, 2016).

The new state of the art in sentiment and emotion analysis is to leverage the sequence comprehension abilities of LLMS to improve context dependent classifications of documents. A popular choice for this task is stacked encoder models, especially BERT and RoBERTa. These models offer several advantages over traditional lexical methods. First among these advantages is the ability of these models to learn a context dependant understanding of words thanks to positional and segment encoding allowing for the model to properly understand word and sentence order and how the same token functions differently based on position. Combining this with the attention mechanism,which learns which words are most connected within and between sentences, allows for encoder based models to learn import order based relationships between tokens and integrate important contextual information both within and between sentences in a document (Devlin et al., 2019). These benefits are found in both the BERT and RoBERTa models.

The chief difference between the BERT and Roberta models is in the use of dynamic rather than static masking in the pre-training process. Dynamic masking varies which tokens are masked for each epoch of training and encourages the model to improve word and context level information rather than simply focus on the position of the masked tokens.Other changes in pre-training between the BERT and RoBERTa models include a change in the encoding scheme with RoBERTa using a byte level encoding scheme that allows for a fuller coverage of possible words than the token level encoding of the original BERT model and an increase in batch size and learning rate over BERT (Liu et al., 2019)

3.2. Sentiment Model

To assess the sentiment of service requests made by Seattle residents during the COVID-19 pandemic, we employed a robust NLP model from the Hugging Face platform.

3.2.1. Model Selection

After evaluating several models available on Hugging Face, we selected the DistilBERT model, for our sentiment analy-

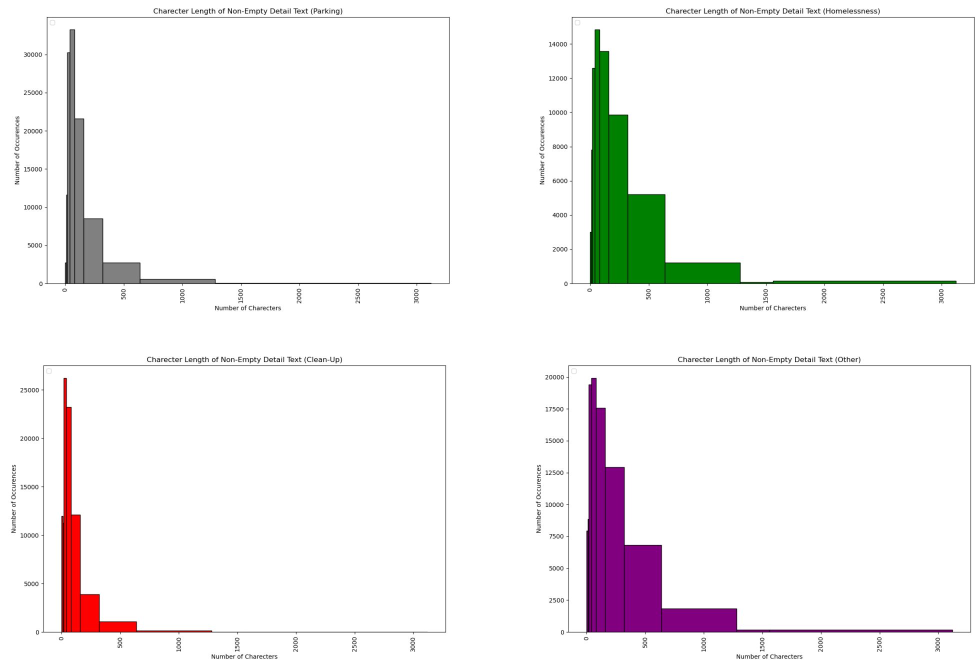

Figure 6. Distribution of Text Length by Group

sis. DistilBERT, short for "Distilled BERT," is an advanced model built upon the BERT (Bidirectional Encoder Representations from Transformers) architecture (NLPTown, 2022). A bidirectional encoder in models like BERT processes text by considering context from both directions (left-to-right and right-to-left), allowing for a more comprehensive understanding of each word's meaning. While a bidirectional decoder generates output text by using context from both previous and future tokens within the sequence. We chose the bert-base-multilingual-uncased-sentiment model by NLP Town (NLPTown, 2022) for sentiment analysis. The decision to use this DistilEBERT was based on its reliable performance in ranged sentiment analysis, making it well-suited for classifying text in a range from 1 to 5, with 1 being the most negative and 5 being the most positive.

3.2.2. Advantages of DistilBERT

DistilBERT is a "smaller, faster, cheaper, and lighter" version of the BERT model (Sanh et al., 2020). DistilBERT has the ability to perform the same NLP tasks, particularly sentiment analysis, as the BERT model with minimal loss in performance. The team was able to leverage powerful NLP capabilities without needing access to high-performance computing resources. DistilBERT was fine-tuned and evaluated on a wide range of datasets covering diverse languages and text domains. This extensive fine-tuning process enhanced the model's ability to generalize across different languages and text genres, surpassing models trained on single-domain datasets. For example, DistilBERT achieved superior performance across tasks such as sentiment analysis, question answering, and text classification when compared to models trained on specialized benchmarks like IMDb reviews or news articles. Its adaptability and efficiency make DistilBERT a robust choice for various NLP applications requiring rapid inference and broad domain coverage. NLP Town fine tuned their DistilBERT model's accuracy against 5,000 product reviews. The model had a 67 percent accuracy rating for exactly predicting a review's rating and a 95 percent accuracy rating when off by one or less. Given the high accuracy of the DistilBERT model by NLP Town, the team made the decision to rely on its pre-trained accuracy for our analysis.

3.2.3. Sentiment Analysis Process

1. Tokenization: Each text entry was tokenized using the DistilBERT tokenizer, which converts the text into a format that the model can process.

2. Model Inference: The tokenized text was then fed into the DistilEBERT model, which predicted the sentiment for each entry. The model output was numerical, indicating on a scale from 1 to 5 how positive or negative the text was.

3. Result Aggregation: The sentiment predictions were aggregated to analyze trends over time, particularly focusing on changes during the pandemic.

By employing the DistilEBERT model for sentiment analysis, we aimed to uncover patterns and insights from the unstructured text fields of the service requests. This analysis helps in understanding the impact of COVID-19 on the sentiments expressed by Seattle residents in their interactions with city services.

To analyze the data, we created a timeline from 2018 through 2022 and recorded the number of positive responses, comparing them to the number of negative responses during this period. We grouped the data by month to provide a detailed view of how sentiments evolved over time. This approach allowed us to observe the impact of COVID-19 on the sentiments expressed by Seattle residents in their service requests, providing valuable insights into the pandemic's effect on city services and resident satisfaction.

3.3. Emotion Model

3.3.1. Model Selection

Two major points of consideration were used while selecting an emotion model for this analysis.

1. Neutrality Class: Given that the data at hand consisted mainly of reports of fact combined with expressions of opinion and complaints, a key area of concern was the sometimes spurious emotional attachments that language models can create between idea such as color, size, social status when a neutral class is not present in the training data. For example, the Hartman model selected will classify ”red car” and ”silver car” both as neutral, however a similar emotional model by bhadresh savani fine tuned without a neutral class classifies ”red car” as anger and ”silver car” as joy. This large swing in emotional classification caused by an insignificant change in content is not ideal (Savani). For this reason, only models that had a neutral emotional class were considered.

2. Contextual Emotion Analysis: Given the relative brevity of the documents under consideration, short pieces of detail text, the ability of the model to integrate context effectively to maximize information integration was critical. For this reason, most traditional methods of lexical and bi-gram analysis were taken out of consideration due to their relative ineffectiveness in integrating contextual information into their classifications.

These two criteria pointed to the use of an encoder based model thanks to the encoders ability to easily integrate context into their classifications. Specifically models that were fine tuned, at least in part, on emotion classification data sets that contained a neutral class such as the Crowdflower, GoEmotion and MELD data sets.

These considerations pointed to the use of the emotion-english-distilroberta-base model fined tuned by the user J-Hartmann on Hugging Face (Hartmann, 2022). This model is based on the DistilRoBERTa which has shown very strong performance on text classification tasks thanks in part to RoBERTas use of dynamic masking during training time leading to better context awareness (Hartmann, 2022). It also has been fine tuned on a balanced sample of emotion texts that contained a neutrality class, which meant that the model passed both of the hurdles for use in this analysis.

3.3.2. Emotion Analysis Process

The focus of the emotion analysis was on the proportion of messages classified into a given emotion over the course of a week. To allow for a uniform analysis of emotionality through time, independent of the number of reports that were received in each time period. The entire corpus of non-null report details were tokenized and classified using the Hartmann model, the proportion of weekly messages classified into the given emotions was then calculated.

4. Results

4.1. Policy Change and Major Events

4.1.1. Halted Parking Enforcement and Transfer of Parking Enforcement to SDOT

Quite early in the pandemic, April 3rd 2020, the city of Seattle made the decision to cease the enforcement of seventy two hour parking rules within the city (Savransky, 2020). This was in reaction to stay at home orders and the subsequent decrease in mobility around the city. A secondary goal of this program was to allow for those living in their vehicles greater stability, and easier access to resources during the pandemic. This cessation of parking enforcement is directly reflected in the number of services requests relating to parking as service requests in that group dropped to their all time low the same week that enforcement stopped, dropping from a pre-pandemic median of 1,067 to a pandemic median of 407. Parking requests remained low through out the pandemic and only recovering late in 2021 as can be seen in Figure 7.

Another change in parking policy that occurred during the period of study was that the city council for Seattle decided to move administration of ticketing for parking violations from the police department to the Seattle Department of Transportation (SDOT) (Bowman, 2021). The change in control, combined with the majority of individuals in Seattle having received their full regiment of COVID vaccines led to an up-tick in reports for parking violations to a level that had not been seen since the cessation of parking enforcement, this tick up yet again, to the highest levels yet, once parking

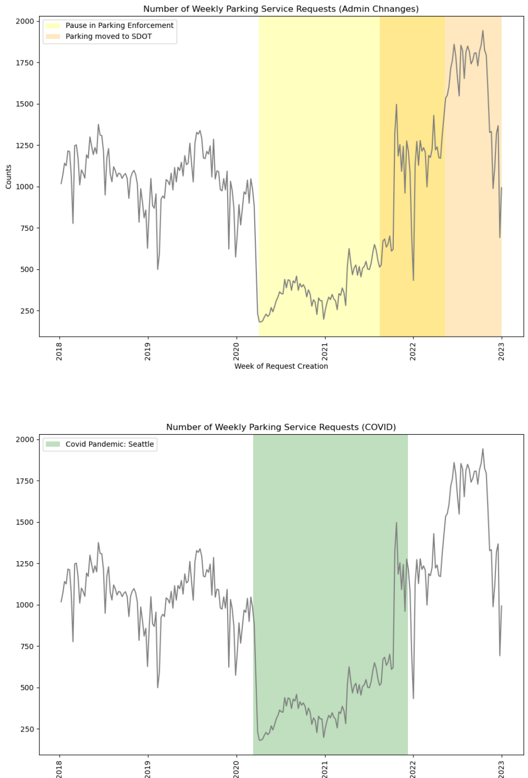

Figure 7. Number of Service Requests Related to Parking Through Time

enforcement rules where reinstated on May 13th 2022 (Cit).

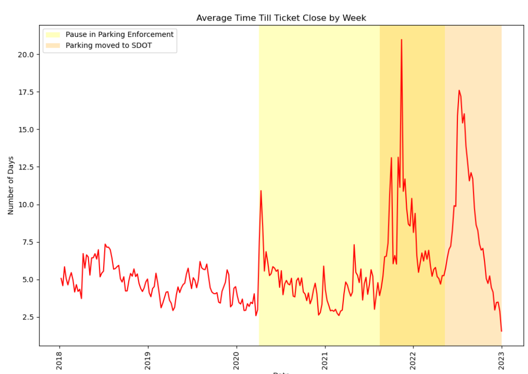

Figure 8. Time Till Ticket Close

In Figure 8, we see that the average number of days till ticket closure had a slow decrease up until the COVID pandemic and its resultant change in parking enforcement policy. This major change in policy likely caused confusion and lead to a sharp, but short lived, increase in completion times. This pattern would recur again after the transfer of parking enforcement to SDOT and again when the parking enforcement hiatus was lifted showing that large scale changes in bureaucratic procedure along side pandemic precautions had a markedly negative effect on the effectiveness of the cities personnel and their ability to fulfill work requests in a timely manner.

4.1.2. COVID Pandemic

As can be seen in Figures 7 and 9, different kinds of request reacted differently to the break out of the COVID pandemic. Parking requests dropped to historic lows most likely due to the change in enforcement policy. Other reports saw only a marginal decline that seemed to have been concentrated to the beginning of the pandemic with pre-pandemic median weekly requests of 679 falling only slightly to median weekly requests of 644 during the pandemic before rising after the pandemic to a all time high of 1,008. The most unusual feature of the Other groups report rates is the dramatic increase in pot hole reports that directly follows our assigned end to the pandemic. This sharp up-tick is mainly composed of work requests in the SDOT-Pothole class. This sudden spike could indicate that as people began to more freely move about the city that more potholes where noticed and subsequently reported, however the concentration of the reports to mainly a single week indicates that it might also have been due to a structural change in the reporting method. However, unlike Unauthorized Encampment, SDOT-Pothole appears throughout the timeline indicating that it had always been a valid option within the app.

Reports of homelessness, however seem to have had an interesting pattern with reporting rates generally decreasing on the lead up to the pandemic with a median of 275 median weekly reports, excepting a sharp spike past 1,200 reports that occurred after posters requesting that people report tents to the Find-It/Fix-It app appeared around Seattle and led to a short term reporting movement that included both valid and obviously satirical reports (Red, 2019). However, after the generally low and stable reporting rate of the COVID era, median weekly reports feel to 177, the number of reports quickly rises to all time highs reaching a median weekly reports of 426.

Clean up related requests, which include things like illegal dumping, drug waste, and clogged sewage drains, continued their pre-pandemic rising trend through out the pandemic, going from a median weekly requests of 849 before the pandemic to 1,089 median weekly requests during the pandemic, and continuing after the majority of the populace had completed their vaccinations and everyday activities had begun to return to normal with post pandemic median

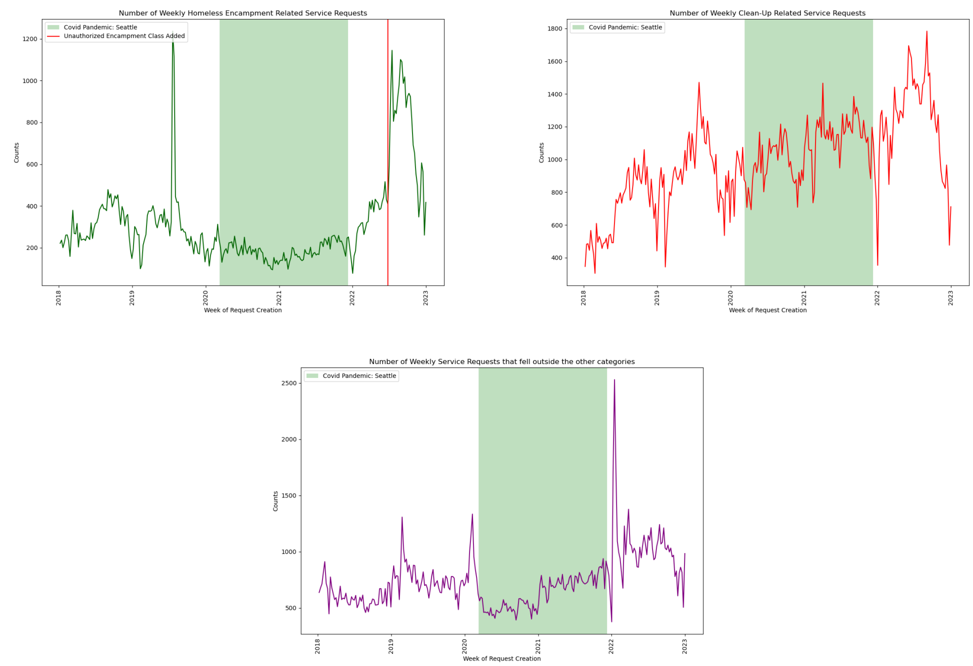

Figure 9. Service Requests relating to homeless encampments and city clean up efforts through time

weekly requests of 1,270.5. [1]

4.2. Sentiment Analysis

The following plot depicts the results of the DistilBERT model.

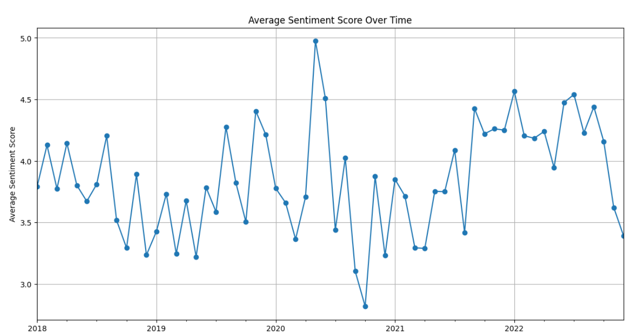

Figure 10. Sentiment Analysis Over Time

From this model, the impact of the COVID-19 pandemic had an effect on citizens' service requests. Right before the start of the pandemic (March 2020), there was a surge in the overall passivity of the service requests before taking a steep drop in negativity. We can speculate the reason for the drop in service request's sentimentality is due to citizens viewing the city's handling of the pandemic negatively, specifically in regards to the city's policies on parking enforcement and unauthorized encampments. Below you will see another sentimentality model. This model is specific to service request tickets that had the tag 'Unauthorized Encampment'.

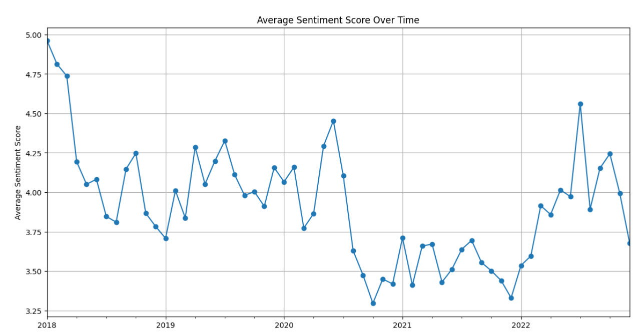

Figure 11. Sentiment Analysis for Unauthorized Encampments

As seen in the model in Figure 11, the sentiment of tickets labeled "Unauthorized Encampments" follows the general trend of the main sentiment model with a noticeable decline at the start of the pandemic but the variability is noticeably less. We are to ask ourselves what is the cause for the decrease in sentimentality? It is most likely citizens not

[1]*homeless encampment was not a category before June 27th 2022 the counts from before that date are from those that were reclassified into that category using a fine-tuned DistilBERT LLM.

looking kindly upon the city's policies on unauthorized encampments.

4.3. Emotion Analysis

4.3.1. Reduced Neutrality and Spike in Fear

Looking to Figure 12, we see that there is drop in neutral classified work requests across all four groups. Parking sees a drop in median proportion of requests with a neutral classification from 0.6736 before the pandemic to 0.4228 during the pandemic.Homelessness drops from 0.3702 to 0.2659, and Clean-up drops from 0.3086 to 0.2143. Other is the only category that escapes this pitfall with a drop from 0.4372 to 0.4151 which seems to have more been a continuation of a pre-pandemic trend rather than a direct impact of the pandemic and Seattle's policy response. This constriction of the neutral class can be interpreted in the light of disaster response. During times of uncertainty and stress people language naturally follows their general mood. The COVID pandemic saw a noticeable increases in rates of insomnia,depression and anxiety (May, 2024). All of which points to a generally more emotional atmosphere.

[2]

The second general trend that can be observed in all four groups, excepting possible clean-up as the not nearly as pronounce, is a uptick in the proportion of messages labelled as fear. Unlike many of the quite pronounce spikes in individual group emotional profiles this spike does not have a clear answer responses on several topics from downed street lights to graffiti all show a increase in fear. This increase in language that the model recognizes as fear indicates a structural change in how issues within the city are being discussed, however after examination what this change may be exceeds the scope of this examination.

4.3.2. Parking: Abandoned Cars, Sadness and Anger

The emotional rating of requests related to parking saw two major swings during the period under study as seen in figure 13. The first was after the parking moratorium was put in place. Immediately the group switch from being composed on mainly parking violations to mainly being composed of reports for abandoned cars. This drastic increase in texts mainly speaking about abandonment is reflected in the emotional composition as an increase in sadness.

However, as quickly as this spike came about it vanishes with an increase in anger taking its place coinciding with the hand off of parking enforcement from Seattle Police (SPD) to Seattle Department of Transportation (SDOT). This change does not seem to be in the quality of service rendered, however as the majority of the increase in requests labelled with anger came from requests with detail text that contained text of the form "(Area) PEO {agent name} is responding to this complaint". With the change in management came the addition of a service that automatically responded to requests with information on the responding agent and these requests were misclassified by the model as angry messages. Changing the classification of these messages to neutral removes the abnormality as can be seen in the bottom half of Figure 13.

4.3.3. Homelessness reporting and Policy Engagement

There are three interesting features in change in emotional composition of reports in the homelessness group. The first of these features emerges in June of 2018 and lasts roughly till the beginning of 2019. This coincides with an increase in the number of homeless reports and with the addition of messages that have the structure "Date Ticket is Closed/resolution/MIR-NAME". However, these messages do not seem be the prevailing cause of the increase like in other groups, instead it appears that this increase in neutral messages coincides with a real increase in the homeless population as seen in the 2018 King County Count Us In report, a digital copy of which could not be found, and later reported on in the Seattle Times where the majority of the homeless interviewed cited job loss as their reason for homelessness (Greenstone, 2019).

The second of these features is a sharp spike in overall and especially neutral reporting between July 20th and 29th 2019. This sharp increase came as a direct response to what appears to be a private posturing effort in Seattle that encouraged citizens to report tents to the general inquiry section of Find-It/Fix-It app, this was before unauthorized encampment was added as a self contained class in June 2022 (Red, 2019). This effort was clearly successfully in increasing reporting, however the authenticity of the reports questionable as the reactions to this seeming escalation in the policing of homelessness was received by some with hostility.

Messages included with the reports ranged from the informative and most likely genuine such as "There are garbage cans and a shopping cart spilling over from encampment near HOV lane. It is within inches of the lane. There are people often exiting their tents into the HOV lane. Very unsafe". To castigation," Shame on you.I have used this app for good. to see your new addition of pin pointing tents within the city is heartless. This is a gross misuse of information.". To the satirical "So many tents plz help rei is a gang"[3], "There is an entire band of Gypsies with tents set

[2]* In Figure 12 all graphs use values after the adjustments mentioned in subsequent passages have been applied

[3]REI is a popular chain store specializing in equipment for

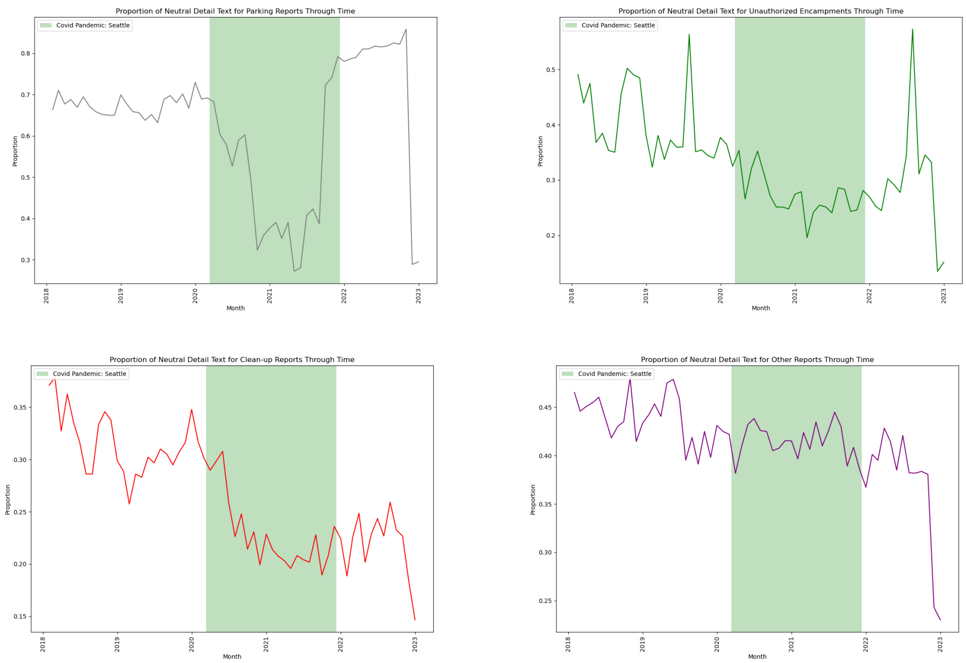

Figure 12. Proportion of Text Bearing Reports Classified as Neutral

up in the park. I think they have tambourines. I saw it in my crystal ball. Down with the tents!"

Included among these messages were at least 230 that were some variation of simply saying tent. Although the posters themselves do not appear to have been posted by the city to encourage the reporting of illegal encampments and instead appear to be part of a private effort. It is interesting that this exact strategy of integrating homelessness reporting into the Find-It Fix-It, when repeated in an official manner in 2022 would lead to a similar increase in reporting but would not receive the same kind of negative messages as the earlier private effort.

When we look at the final spike that comes in the July 2022 directly following the addition of the Unauthorized Encampment class to the Find-It/Fix-It app, we see a sharp increase in homelessness related reports from a pandemic average of 177 median weekly reports to 426 median weekly reports post pandemic. Unlike the original private effort, this change did not seem to illicit the same kind of negative response from the public. This could be for a few reasons.

1. The populace that commonly uses the Find-It/Fix-It app would notice the addition are separate from the populace that were exposed to the original posturing campaign. The addition of the class to the Find-It/Fix-It app was done with little announcement and so would not have the wide reaching engagement that a public posturing campaign could garner.
2. A change in taste or belief in the general populace from the 2019 posturing campaign to the 2022 official change could account for the lack of dissenting voices being raised in the requests as opinions on the effectiveness or ethics of the city encouraging citizens to share information on possible homeless communities may have become more favorable.

hiking,camping and other outdoor activities

4.3.4. Other Questions

The Other group is unique in that it has an apparent shift in emotional profile that precedes the pandemic and then falls away as the pandemic begins. However, like the increase in anger seen in the parking reports this increase also seems to be an artifact. The responses that generated this shift chiefly came from the SCL-Streetlight Report class and contained only the word gone, either upper or lower case and sometimes including the date. Though it is possible that this was a legitimate trend in how citizens were filling out the particular questions posed to them by the app it is also possible that this was a generic response that those city service workers completing the task were advised to input into the app from their end given that the rows that do have dates have data that follow the Creation date by a few days. Removing the rows that have start with gone shows that the Other group had a generally stable emotional profile through out the period of study. Reports were chiefly neutral with a

Figure 13. Emotion Analysis for Parking, Clean-up, and Homelessness reports through time

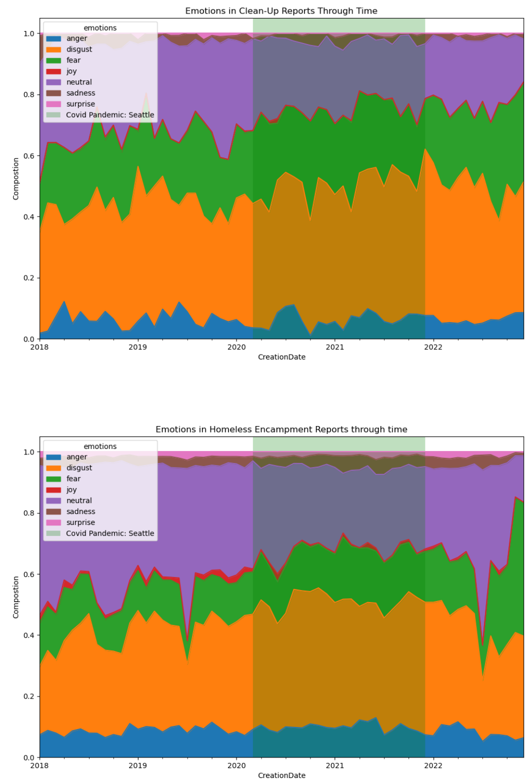

Figure 14. Emotion Analysis for Parking, Clean-up, and Homelessness reports through time

consistent mix of other emotions.

5. Goals for Further Research

5.1. Recurring Reports and Time to Resolution

In Figure 8, we looked at the average time till a ticket was closed as completed. However, the values shown are most likely skewed downward from what the average citizen would think of as completed by the fact that many requests are resubmitted multiple times either by multiple interested citizens that happen to come across and report the same problem or the same individual submitting reports for the same issue multiple times. This could be because the issue has recurred or the issue was not satisfactorily completed by the city before the first request was closed. The lack of identifying information on the individuals that submitted each report makes traditional longitudinal matching of reports and tracking of issues extremely difficult.

This leaves open the question of the time between first reporting and resolution. A possible route for further study would be to apply non-traditional methods such as text-classification or speaker identification to either track persistent reporters or to identify recurring requests so that a fuller picture of time to resolution could be developed.

5.2. Expanded reaction domain

Given the prevalence of social media as the modern replacement for the town square paring the data on emotional reactivity from this study with similar results on social media would give an interesting insight into how citizens speak about problems within their city when they are under public observation as opposed to when they are communicating directly with city services and are working under the assumption that their communications will be viewed by only a small number of people.

The brief increase in the number of homelessness reports that seemed to have been initiated by the posting of flyers, by a party other than the city, encouraging citizens to report homeless encampments to the Find-It Fix-It app does point to the fact that at least under certain conditions citizens did find the app to be an appropriate local for the airing grievances and displeasure with city administration.

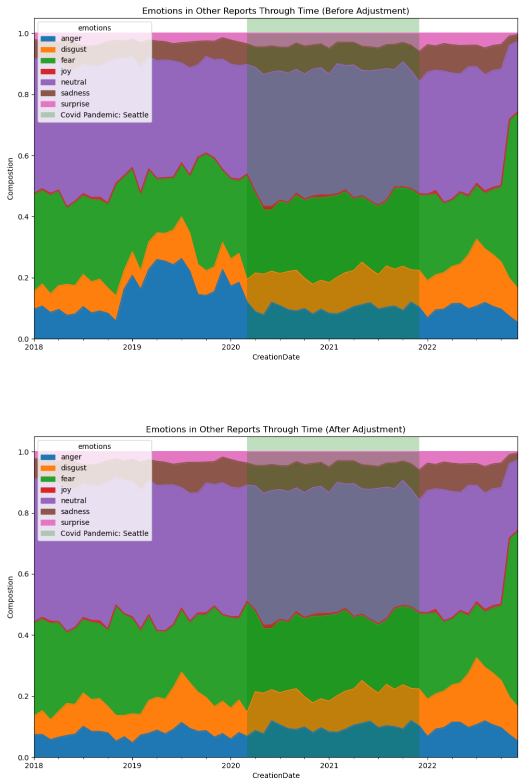

Figure 15. Emotion Analysis for Parking, Clean-up, and Homelessness reports through time

However, whether or not the general tone of these messages aligns with the tone taken on explicitly public forums is an open area for research.

6. Self Assessment

6.1. Sloan C. Stinnett

Going into this project my main goals were to:

1. Learn more about transformer based models and why they are good at sequence based tasks.
2. Improve my understanding of natural language processing and the different techniques and tools available.

I believe that I accomplished both of these goals. Through this project I both learned a lot about encoder based models, BERT and RoBERTa especially, and how mechanisms like multi-head attention and positional encoding when applied to sequenced based learning tasks like masked value prediction and sequence ordering allow for the model to gain deep insights into how complex sequences, such as sentences in English, are formed and how different parts of the sequence can relate to one another in a dynamic manner.

I also gained a much greater understanding of what state of the art natural language processing looks like in the age of transformers. I have worked on sentiment and emotional modeling in the past and found lexical methods too weak to be useful. However, now that I have experience with applying BERT I am much more confident that I would have powerful tools ready at hand for exploring sentiment and emotion based questions. It was also interesting learning about TFIDF and the part that it plays in document summarizing and it seems like a wonderful technique for quickly finding what makes a document in a corpus unique.

Having a good grasp on the fundamentals of working with databases, cleaning and manipulating data tables, and finding data sources were the skills that I found the most useful for this project. I enjoyed learning about how transformer models function and how to apply them using the hugging face package in python.

This project was supervised by Dr.Matthew Beattie

6.2. Peter C. Hardisty

Going into this project my main goals were to:

1. Test the skills I've gained in DSA and apply them in a more professional manner.
2. Learn about natural language processors and transformer models and practically apply them.

I believe that I achieved both of my goals to test the skills I've learned in DSA and to learn about and incorporate natural language processors and transformer models. Before this project, I had only heard about natural language processors and transformer models. It was challenging to implement these processes at first, but as I learned more of the methods and the practical parts for implementation, things became smoother. I was particularly interested in how the BERT and DistilBERT models operated. How they tokenized the text and then ran these tokens through different layers was fascinating to learn about. I hope to use transformer models more as a professional. As for using skills I've learned while being in the DSA program at the University of Oklahoma, I believe that I incorporated them efficiently and practically for this project. The different methods I used for data exploration was particularly helpful in investigating and understanding the data I was given. Also, the amount of coding done in the DSA program readied me for this project. It was helpful having already gotten over the learning curve of having good coding practices before beginning the project.

Finally, the DSA skills that I thought were most useful for this project would be the different methods needed for efficient data exploration and general database knowledge. It was helpful to have a few methods at the ready to explore the data so that I could focus on what the data was telling me. And being able to navigate a PostgreSQL database and write up queries quickly was very beneficial in getting the data in a format that was useful for analysis.

This practicum was for four credit hours and was a research project for Matthew Beattie.

This project was supervised by Matthew Beattie

7. Source Code

Source code will be made available at this github page upon request.

References

URL https://www.seattle.gov/transportation/projects-and-programs/programs/parking-program/parking-regulations/72-hour-on-street-parking-ordinance.

Jul 2019. URL https://www.reddit.com/r/SeattleWA/comments/cfrzeu/find_it_fix_it_app_posters_around_town/.

Apr 2024. URL https://www.mayoclinic.org/diseases-conditions/coronavirus/in-depth/mental-health-covid-19/art-20482731#:~:text=Worldwide%20surveys%20done%20in%202020,an%20issue%20for%20many%20people.

Beattie, M. Mjbeattie/fifi$_{c}lassificationhuggingface.URL.$

Bowman, N. Seattle council votes to move parking enforcement officers out of spd, Aug 2021. URL https://mynorthwest.com/3093949/seattle-council-votes-move-parking-enforcement-away-from-spd-into-sdot/.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. 2019.

Dowdy, D. and D'Souza, G. Early herd immunity against covid-19: A dangerous misconception. URL https://coronavirus.jhu.edu/from-our-experts/early-herd-immunity-against-covid-19-a-dangerous-misconception#:~:text=To%20reach%20herd%20immunity%20for,would%20need%20to%20be%20immune.

El-Din, D. M. Enhancement bag-of-words model for solving the challenges of sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 7 (1), 2016.

Greenstone, S. Why are people homeless in seattle?, May 2019. URL https://www.seattletimes.com/seattle-news/homeless/why-are-people-in-seattle-homeless/.

Hartmann, J. Emotion english distilroberta-base. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/, 2022.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. Roberta: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692, 2019. URL http://arxiv.org/abs/1907.11692.

NLPTown. Bert base multilingual uncased sentiment. https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment, 2022.

Sanh, V., Debut, L., Chaumond, J., and Wolf, T. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108, 2020.

Savani, B. Bhadresh-savani/distilbert-base-uncased-emotion · hugging face. URL https://huggingface.co/bhadresh-savani/distilbert-base-uncased-emotion?text=Silver%2Bcar.

Savransky, B. Seattle announces free parking in the city amid covid-19 outbreak, Apr 2020. URL https://www.seattlepi.com/coronavirus/article/seattle-announces-free-parking-amid-covid-19-15177
php.