UNDERSTANDING THE HUMAN PSYCHOLOGY TOWARDS RELATIONSHIPS USING THE REDDIT DATASET

A Thesis presented to

the Faculty of the Graduate School

at the University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

PRANJAL DUBEY

Dr. KSM Tozammel Hossain, Thesis Supervisor

DECEMBER 2022

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled:

UNDERSTANDING THE HUMAN PSYCHOLOGY TOWARDS RELATIONSHIPS USING THE REDDIT DATASET

presented by Pranjal Dubey, a candidate for the degree of Master of Science and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. KSM Tozammel Hossain

Dr. Grant Scott

Dr. Ed Mirielli

DEDICATION

The thesis is wholeheartedly dedicated to my parents, who has been my source of inspiration and supporter throughout my research, they encouraged me attentively with their fullest and truest attention to accomplish my work with truthful self-confidence.

To my mentors, friends and classmates who shared their works of advice and encouragement to complete the thesis.

I would like to give special thanks to my life partner for his cheerful motivation throughout the accomplishment of my work.

ACKNOWLEDGMENTS

I am highly indebted to Dr. KSM Tozammel Hossain, and obliged for giving me the autonomy of functioning and experimenting with ideas. I would like to take this opportunity to express my profound gratitude to him not only for his academic guidance but also for his personal interest in my report and constant support coupled with confidence boosting and motivating sessions which proved very fruitful and were instrumental in infusing self-assurance and trust within me. The nurturing and blossoming of the present work is mainly due to his valuable guidance, suggestions, astute judgment, constructive criticism and an eye for perfection. My mentor always answered myriad of my doubts with smiling graciousness and prodigious patience, never letting me feel that I am a novice by always lending an ear to my views, appreciating and improving them and by giving me a free hand in my report. It's only because of his overwhelming interest and helpful attitude, the present work has attained the stage it has. I would like to express my gratitude towards Dr. Grant Scott and Dr. Ed Mirielli for their interest to be a part of my thesis committee.

Finally, I am very grateful for the consistent support and love I got from my beloved parents, family and friends which helped me leave no stone unturned to productively complete this thesis.

Pranjal Dubey

Contents

List of Figures

ABSTRACT

There is an increasing interest in exploiting social media data to address problems in various domains, including social science and psychology. Reddit is a well-known social media platform that promotes user interactions on various topics and accumulates massive amounts of data daily. Reddit users participate in discussions by posting reddits, which could be text posts, links, images, and videos, or by reacting to reddits via up- or down-votes. These reddits can be analyzed to investigate mass opinion on particular issues. In this thesis, we analyze users' behavioral trends and morals of expiry dating using reddits. A relationship can start from casual dating and may not come with future bonding; such a relationship starts with an expiration date, which is known at the initial stage. People tend to seek advice on this type of relationship on social media, including Reddit. In this study, we extract reddits on relationships and apply machine-learning-based methods to infer trends on expiry dating. We adapt anchor-based topic models to infer topics from reddit corpus and estimate the trends over these topics. In addition, we infer the moral intentions—care/harm, fairness/cheating, purity/degradation, authority/subversion, loyalty/betrayed—of individuals while getting into expiry relationships based on the updated moral foundation theory framework. We collected $\sim$665K reddits under various subreddits from 2019 to 2022. Our analysis shows an increasing trend in the topic 'dating with end date' where the moral is 'fairness' over these years.

Chapter 1

Introduction

Language analysis relies heavily on the human ability to operationalize psychological constructs as they manifest in language. Either by drawing on one's own experience of psychological states and using language, or imagining the verbal tendencies of hypothetical individuals, extracting psychological inferences from text data requires placing one's self within a language user's perspective. It is a distinctly human ability, much as a conversation, reading a book, or listening to a speech requires the universal human ability to communicate. This, however, presents a pragmatic challenge: how can the subjective, human experience of language be applied to a framework for objective, evidence-based science? To generate psychological knowledge from language data, a dedicated, systematic approach must be developed. A quantitative language science must be built on principles of objectivity, reproducibility, and transparency. This is not so different from other quantitative sciences, but the complexity of language and the inherent subjectivity in assigning psychological meaning to language make the challenge unique. Psychological inferences about language use are rooted in subjectivity, dependent on who is producing the inferences (i.e., the researcher), and difficult to dissect given that they come from inscrutable psychological processes, such as reflecting on the likely intention of the speaker when coding a piece of text. Computation offers a hybridized approach, wherein human subjectivity is treated systematically and inferences are expressed within a statistical framework [1]. Computerization requires that we make everything in our analysis formal and explicit. Everything — data representation, data processing, algorithms and decisions — is deterministic and programmed. Instructions for a computer in a computational language analysis include text preprocessing (i.e., including and excluding certain characters and words based on predefined rules), pattern matching (finding words of interest in text), and statistical modeling. By scripting the language coding process, we can trace model behavior and gain insight into the measures produced. For example, if we could produce a measure of a text's suicidal ideation by instructing a computer program to follow a sequence of commands, we can justify (and critique) the measure based on the steps that produced it. This is fundamentally different from a measure produced by a human coder, given that human interpretations are not guided by pure formalism but by their mental processes and insights. However, a computer's objective advantage over a human is counterbalanced by its disadvantage in the skills and intuitions that humans have for understanding and interpreting language. Humans have evolved as social animals, using language to transmit their internal states to others and navigate social life. Not only can we efficiently use language to convey complex social information about what we are thinking and feeling, but we can also use our rich store of life experiences to turn other people's words into meaning. That is why humans are excellent at reading texts and inferring what the author means, what psychological state the author is in, and so on. Computers, on the other hand, have no intuition about the texts they process: they see sentences as character strings and know little about what a text means or how it would be interpreted by a human reader. Thus, computation affords us a variety of advantages, at the cost of losing intrinsic human socio-cognitive abilities. Below, we discuss a complementary framework for combining the best computational methods with human interpretation.

LeFebvre et al. (2018) have argued that research and scholarship have not fully captured the current lay lexicon use (e.g., ghosting, benching, catfishing, cuffing season) to describe dating and relationships. In addition, there is limited research and insight into short-term committed relationships with no expectations of long-term commitment, otherwise known as relationships, that lack the long-term orientation component of commitment. Further investigation into this concept would add clarity to our existing knowledge of romantic and sexual relationships.

1.1 Computational Psychology

The application of computational principles to understanding human behavior covers a broad range of topics and approaches. A prime example of this is simulation modeling, i.e., the development of computer programs that simulate human behavior. It uses modeling tools to analyze large data sets, reflecting, for example, the instantaneous behavior of millions of Reddit users, the moment-to-moment ontogenetic history of a single chick raised in a sensor-rich environment, or the analysis of more standard forms of multivariate problems in human behaviorMorteza Dehghani al. [2].

1.2 Sentimental Analysis

Analyzing human language is often a complicated process as it requires dealing with grammatical nuances and linguistic variations of great magnitude. Lack of contextual understanding makes the process of extracting sentiment problematic, but the advances in computational techniques can facilitate the process of regulating the emotion and tone behind a series of texts and derive the attitude, emotion, and opinion of the speaker. Sentiment analysis or opinion mining observes conversations and calculates language and voice inflections to quantify the opinions or emotions of the given database Padmanabhan B [3]. It is a technique used for determining the expression and tone underlined in the data. Sentiment analysis can be automated wholly or centered on the analysis of humans or a combination of these two aspects. It provides a visual representation of sentiment density and text polarity and enhances the possibilities of interpretation. Analyzing personality traits based on computational methods in select novels of Cormac McCarthy. An analysis of the characters portrayed in literary texts is made by referring to psychological measurements, and employing text analysis tools. The focus is on the efficiency of bringing together advanced text analytics tools and theories of human personality for a better understanding of human personality.

1.3 Moral foundation

Moral foundations theory was created by a group of social and cultural psychologists to understand why morality varies so much across cultures yet still shows so many similarities and recurrent themes [4]. In brief, the theory proposes that several innate and universally available psychological systems are the foundations of 'intuitive ethics.' Each culture then constructs virtues, narratives, and institutions on top of these foundations, thereby creating the unique moralities we see around the world, and conflicting within nations too. The five foundations for which we think the evidence is best are:

1) Care/harm: This foundation is related to our long evolution as mammals with attachment systems and an ability to feel (and dislike) the pain of others. It underlies the virtues of kindness, gentleness, and nurturance.

2) Fairness/cheating: This foundation is related to the evolutionary process of reciprocal altruism. It generates ideas of justice, rights, and autonomy.

3) Loyalty/betrayal: This foundation is related to our long history as tribal creatures able to form shifting coalitions. It underlies the virtues of patriotism and self-sacrifice for the group. It is active anytime people feel that it's 'one for all and all for one.'

4) Authority/subversion: This foundation was shaped by our long primate history of hierarchical social interactions. It underlies the virtues of leadership and

followership, including deference to legitimate authority and respect for traditions.

5) Sanctity/degradation: This foundation was shaped by the psychology of disgust and contamination. It underlies religious notions of striving to live in an elevated, less carnal, more noble way. It underlies the widespread idea that the body is a temple that can be desecrated by immoral activities and contaminants.

1.4 Existing approaches

In this section, we describe the pertinent research.

LeFebvre et al. [5] has argued that research and scholarship has not fully captured the current lay lexicon use (e.g., ghosting, benching, cat fishing, cuffing season) to describe dating and relationships. In addition, there is limited research and insight into short term committed relationships with no expectations of long-term commitment, otherwise known as relationships, that lack the long-term orientation component of commitment. Further investigation into this concept would add clarity to our existing knowledge of romantic and sexual relationships.

LeFebvre et al. [5] here the study compares machine learning methods for sentiment analysis with the n-gram feature in the case of financial news and corporate headlines. The number of neutral sentiments is 2879, the number of positive sentiments is 1363, and the number of negative sentiments is 604. The data is preprocessed by deleting things that are not needed. Preprocessing results are given weight by the n-gram method. The n-gram features are analyzed by sentiment analysis using several machine learning methods.Sentiment analysis performance is calculated using precision, recall, and f1-measure.

Trager et al. [6] focused our corpus compilation effort on Reddit for a number of reasons. First, in comparison to Twitter, Reddit shares many of the same research friendly features (e.g., responsiveness to current events, public posts, and available APIs) [7] but is organized into what are called subreddits. Different subreddits have distinct topics and consistent communities with varying cultures and norm. In relation to morality, these distinct communities and norms have led researchers to use Reddit to investigate moral conflicts across groups a phenomena that is harder to investigate on Twitter (which does not have organized groups) or Facebook (where many groups are private). Second, Reddit provides more anonymity than many other social media platforms, potentially enabling users to more freely speak their minds and express their opinions Third, in addition to general differences in language usage [8], the lack of restriction on the length of posts on Reddit may be particularly beneficial for training models. Fourth, we believe that Reddit has played a distinct role in contemporary politics. For example, the r/The Donald and r/incels subreddits have been linked to political extremism [9] and mass shootings [10] In order to understand the validity and relative performance of different text classification methods for identifying the moral concerns manifested in the MFRC, we report baseline results for a number of NLP approaches: Distributed Dictionary Representations Bidirectional Encoder Representations from Transformers. BERT models. These baselines can both be used as evidence about the relative performance of various NLP models on the task of moral concern detection and also inform future methodological work including to help calibrate new models of moral sentiment classification. a set of baseline classification results using a range of methods, and a cross-corpus comparison with MFTC.

Therefore, it is concluded that there is limited research and insight into short term committed relationships with no expectations of long-term commitment, otherwise known as relationships, that lack the long-term orientation component of commitment or relationship with an end-date.

1.5 Objectives

This research aims to analyze and estimate the trends and characterize the moral foundation of expiration dating using social media data.

1.5.1 Relationship Trend Analysis

Does dating comes with an end date? What are the trends of dating with an end date over years? This justifies the psychological behaviour of the user by their ways of discussion of the social media. What is the moral of the individual who get in relationship with an end-date? This justifies that weather the user get in relationship with end date with some positive moral towards the other person on with negative moral. This is observed through vice and virtue of the moral foundation refer Figure 1.1.

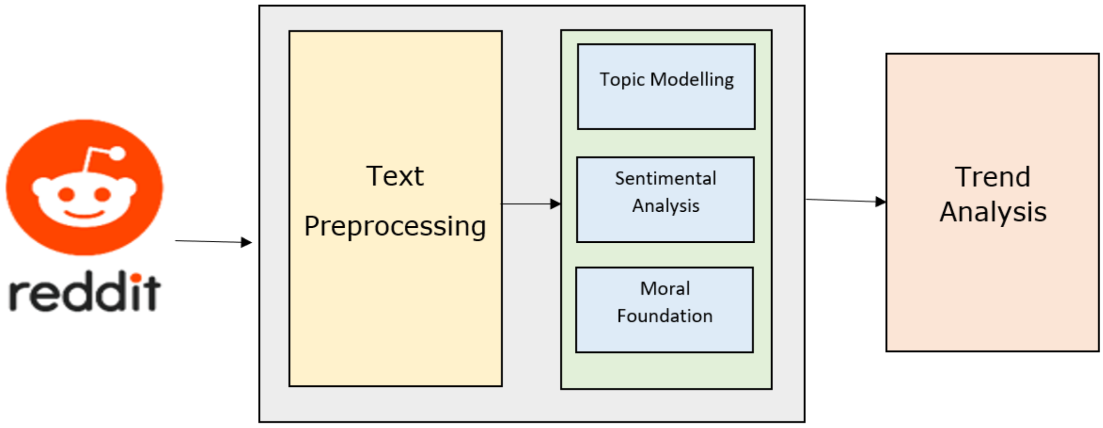

Figure 1.1: Overview of the analysis. The key idea in this framework is to extract data from 3rd Party API- Reddit and analysis the trends using semi-supervised machine learning methods.

1.5.2 Moral Foundation Analysis

Moral intuitions are a central motivator in human behavior. Recent work highlights the importance of moral intuitions for understanding a wide range of issues ranging from online radicalization to vaccine hesitancy. Extracting and analyzing moral content in messages, narratives, and other forms of public discourse is a critical step toward understanding how the psychological influence of moral judgments unfolds at a global scale. Extant approaches for extracting moral content are limited in their ability to capture the intuitive nature of moral sensibilities, constraining their usefulness for understanding and predicting human moral behavior. Here we introduce the computational Moral Foundations, manually annotated data set performs the supervision on the large corpus. The CMFD, unlike previous methods, is followed by supervised computational models. It demonstrate that the CMFD outperforms existing approaches in a variety of domains. I anticipate that the cMFD will contribute to advance the study of moral intuitions and their influence on social, psychological, and communicative processes.

1.6 Proposed Solution

In this thesis, we study trends over relationship with end-date and moral foundation of the users", which is an upgrade on the existing moral foundation methods for time-series data processing with increased emphasis on the time information present in the data. This analysis contributes the computational analysis of the user behavior towards relationship with end-date to the society gives the morals of the user which inclined towards relationship with end-date. Specifically, our solution approach provides benefits such as e.g., utilizing time-related information, unlimited events, and a variable number of subreddit over time. Realizing that manual analysis over related issue is time consuming and using computational linguistic - Semi-supervised Learning Corex model for finding the insights from the corpus and the semi-supervised method for label propagation are the proposed modelling techniques.

We extracted our dataset from the third-party Reddit API using psaw library between 2019-2021. In this work, we first evaluate the efficiency of the model using metrics like Accuracy, Precision, Recall, etc. We then compare the performance of the models and suggest the best suitable model with the highest accuracy score in Machine Learning.

1.7 Research Workflow

According to the research objectives, the report will describe the work flow as below (refer to Figure 1.2):

Step 1: Data Extraction: The process involves extraction of data through Reddit API. The data is extracted separately for both Post and Comments.

Step 2: Data Cleaning and Preprocessing: The process involves formatting the unprocessed initial data and standardizing the data. This work removes duplicate data and removes stopwords and punctuation's and vectorize the clean data according to the requirements which will be explained in the next chapters.

Step 3: Data Modeling and Evaluation: Semi-Supervised- Corex Topic modeling is performed to achieve the required results; Sentimental analysis is performed over each obtained topic; Data annotation and label propagation is performed for moral foundation as a semi-supervised machine learning modeling. After training, we evaluate our model with the following metrics Accuracy, Precision, Recall and F1-Score. The best metric model is taken and moral foundation are analysed for the un-annotated data, As a part of moral foundation the vice and virtue of the data set is estimated using the positive and negative sentiments.

Step 4: Trends and Outcomes: After topic model the trends are analysed over time three years, at an interval of six month and at an interval of 3 months. The sentiments over each topic is analysed over years and moral foundation results are analysed over years and topics.

1.8 Thesis Organisation

The thesis has been divided into 5 chapters for readability and easy access to information.

- The first chapter introduces us to the problem described in the thesis. It gives the context of the problem and the objectives to be achieve in order

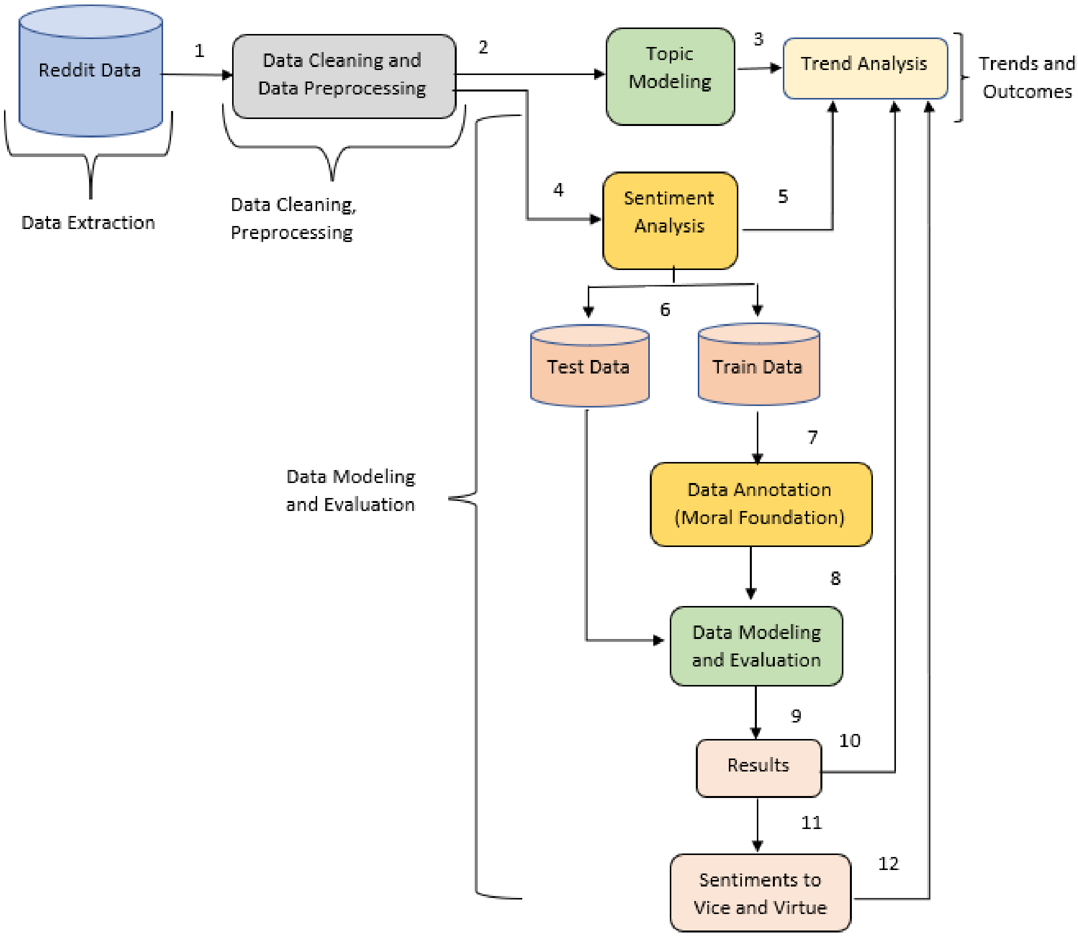

Figure 1.2: Research Workflow: There are 4 broad step: Data extraction; Data cleaning,Preprocessing; Data Modeling and Evaluation and Trends and Outcomes

to perform the trend analysis.

- The second chapter includes the individual tools and techniques we will be using to solve the problem. It contains the key research work in this domain.
- The third and fourth chapter guides us the process of working and creation of our model. This involves preprocessing the data, standardizing it, and other mechanisms. They will be described clearly in chapter 3 and 4.
- The fifth chapter deals with the experiments performed on the model and various results involving it.The performance is checked using several trends analysing graphs which we will see in detail in this chapter.
- The sixth chapter gives us the future scope of the project and conclusions that are made from studying this topic.

Chapter 2

Related Works

Nicholas Proferes et al. [11] Researchers have used social media data for a wide range of purposes—from predicting postpartum depression from Facebook posts De Choudhury et al. [12] to try to predict movements in the stock market based on the sentiment of Tweets Mittal et al. [13]. However, not all uses of social media data have been welcomed by users or seen as acceptable in the research community. For example, personally identifiable information from more than 87million Facebook users was collected in an academic study, but then the data were used by Cambridge Analytica to micro-target political advertisements Isaak et al. [14] Transgender YouTubers have had their images collected without consent to train facial recognition software and as a part of automatic-gender recognition research and development [15]. And a group of Danish researchers was criticized after they publicly released a data set of nearly 70,000 users of the online dating site OkCupid, which included "usernames, age, gender, location, what kind of relationship (or sex) they are interested in, personality traits, and answers to thousands of personal profiling questions used by the site" Zimmer et al. [16] Within research communities and among Institutional Review Boards (IRBs), there is disagreement about the ethical practices that should follow from the use of public data for research purposes and if, or when, using social media constitutes human subject research Metcalf et al. Vitak et al. [17, 18]. In the United States, institutions that house research with human subjects and who receive federal funds are required to have an IRB. However, the use of publicly available data from social media platforms often do not meet the threshold criteria of "research involving human subjects" according to many IRBs. Thus, some IRBs may exempt these kinds of studies from ongoing compliance review and informed consent practices, though others may not Vitak et al. [19].

Ongoing questions about using public social media data have led researchers to question how users' feel about their data being used for research. Fiesler and Proferes [20] conducted a survey of Twitter users to assess how they felt about their data being used. Their findings showed that users were largely unaware that their data were used for research purposes and that perceptions of data use varied by contextual factors, such as who the researchers are and the topic of study, a finding echoing that of Beninger [21].

This prior work reveals inconsistencies in the way "human subjects research" is defined and applied by ethical bodies and researchers, as well as potential discomfort among many social media users in being research subjects. Increasingly, researchers are using Reddit data as a source; however, there are no systematic reviews of the contexts on Reddit that researchers are studying, nor the ethical practices they are engaging in relation to their work.

2.1 Reddit Dataset

Discussions on Reddit are primarily public in that anyone, with or without a Reddit account, can view the content (with the exception of private subreddits). Both original shared content and discussion comments are "voted" on by users, which determine their visibility. To become a Reddit user, all users need is to select a unique username and a password—email verification is not required. The terms of service dictate users must be at least 13years of age to sign up. Site-wide norms discourage participation with one's real name as a privacy-protecting measure.

Participation history on the site is also public, meaning that anyone can see all of a user's public comments and posts by clicking on their username. The ease with which users can create accounts means that it is possible, and not uncommon, for one person to have multiple accounts. "Throwaway" accounts, or single-purpose accounts created for limited time use, are commonly used when users do not want a post or comment associated with their main or primary account, such as sharing sensitive or personal information Ammari et al. Leavitt et al. [22],[23]. Because participation on Reddit is pseudonymous, demographic information is somewhat difficult to obtain. According to Reddit's site administrators [24] a majority (58 percent) of users are between 18 and 34years old and male (57 percent).

Subreddits are both user-created and user-moderated. While there are a few overarching Reddit rules about content, subreddits vary considerably regarding what they allow, and in their specific cultures and norms Chandrasekharan et al. Fiesler et al. [25] [20]. As part of their subreddit specific-rules, some subreddits carry warnings to researchers about data collection in the communities. For example, r/depression and r/ SuicideWatch state all research-related posts and surveys must be approved by the moderator team, and r/IndianCountry prohibits unauthorized research and requests that anyone interested in using the subreddit for research purposes must complete a form for review by moderators.

In addition to individual subreddit rules, Reddit also has a site-wide user agreement. Reddit.com [24] user agreement includes the following prohibition related to collecting data:

"Access, search, or collect data from the Services by any means [automated or otherwise] except as permitted in these Terms or in a separate agreement with Reddit. We conditionally grant permission to crawl the Services in accordance with the parameters outlined in our robots.txt file but scraping the Services without Reddit's prior consent is prohibited."

These terms are fairly standard in their ambiguity Fiesler et al. [20] but do suggest that data collected outside the confines of specific allowances—for example, using their API—may be a violation of this user agreement. However, Reddit's API is freely available and can be used to access content on the site.

Reddit posts, comments, and metadata can be accessed via the site itself, or via its APIs. Reddit's official API is free and publicly available and provides an array of functions. For these reasons, Reddit has an ecosystem of bots created by its user base to help in several ways, such as content moderation Jhaver et al. [26] adding functionality through summarizing information and linking to other websites, or providing humor through parody bot accounts Long et al. [27]. There are additional ways of accessing Reddit data outside of means provided directly by the platform. One of the largest is known as Pushshift, a social media data collection, analysis, and archiving platform founded in 2015 by Jason Baumgartner. Pushshift ingests data from Reddit's official API and collates the data into public data dumps and a live stream of new comment and post data that can be accessed by Pushshift's own unique API. The Pushshift dataset contains submissions and comments posted on Reddit since June 2005 and has been popular for researchers due to its ease of use and larger querying limits Baumgartner et al. [28]. However, PushShift is not an exact mirror of data from Reddit. After posts, comments, and metadata from Reddit's API are ingested by PushShift they are functionally distinct. So, for example, once a person deletes their user history on Reddit those public comments and posts may still exist on Pushshift.

This section deals with the review of literature that has been done in the area of the topic model and moral foundation. This review has been divided into 2 parts:

- Semi-Supervised Topic Modeling- Corex Model
- Moral Foundation using sentiment analysis

2.2 Semi-Supervised Topic Modeling

Based on a survey of the literature, numerous researchers discovered techniques to use machine learning on topic modeling. Latent Dirichlet Allocation (LDA), almost synonymous with topic modeling, is the dominant method in this field. Developed in 2003 by [29] it postulates a generative model for how a topic gives rise to the words in a document. We won't go into the details here because it is covered extensively in many other sources, but essentially it assumes two things:

- That each topic is a distribution over words and
- That each document is a distribution of topics, i.e. every word in a document is generated by first sampling a topic and then sampling from the selected topic's word distribution.

Given this probabilistic model of the probabilistic model p(data—topics), the actual LDA algorithm then tries to determine the posterior p(topics—data); the authors show this can be solved with tools from variational inference.

While LDA makes assumptions about the generative process and the form of the probability distributions (i.e. Dirichlet priors), it has proven itself to be a very successful method for extracting topics. It has also served as a springboard for many, many derived topic models which aim to extend LDA to related situations. Notable examples include "dynamic topic modeling", which models the evolution of topics through time, or "correlated topic modeling", which acknowledges that some topics may be correlated with one another.

However, one key drawback of LDA is that it is unsupervised and this can result in unsatisfactory topics. For example, emerging areas of research in Medline might not be identified as a topic by vanilla LDA, simply because there aren't many articles on the subject. We might think that re-running with an increased number of topics (a hyperparameter) might force topics to break apart into narrower themes and eventually produce results that indicate the new research area has a topic of its own. In practice, however, it has been observed that, with

LDA, underrepresented topics tend to be washed out by ones that have a stronger presence in the corpus.

To get around this difficulty, semi-supervised topic modeling allows the user to inject prior knowledge into the topic model. In particular, there are versions where the user can supply the model with topic "seed" words, and the model algorithm then encourages topics to be built around these seed words. This not only resolves the issue illustrated above (the absence of known topics), but more broadly gives us the flexibility to steer topics towards relevant themes by simply adding keywords, whilst also leaving room to uncover "unknown" topics.

The topic modeling approach is called anchored CorEx (short for "correlation explanation") by [30], which implements a completely novel approach to topic modeling compared to LDA. Unlike LDA, CorEx makes no assumption about the data generation process, instead approaching topic modeling in an information-theoretic manner. CorEx treats each word as a random variable and seeks to find collections of words (i.e. topics) that best "explain" the data. Explaining translates into decorrelating the observed word count data, which proceeds by optimizing a mutual information-based objective function. The authors observe that this objective function realizes an information bottleneck and this can be exploited to modify the model to accept seed, or "anchor", words as input. The flexibility of anchor word selection is a great feature of this model. Anchor words can be grouped in any way, they can be given different weights (a model hyperparameter) and interestingly the same anchor words applied to multiple topics can be used to uncover topic "aspects". The original paper provides an example where a CorEx model with 55 topics in total is fitted, but five of these topics are equally anchored with the words "protest" and "protests". The model then uncovers topics that tease apart five different aspects of protests.

To summarise, when faced with a situation where we wish to uncover topics in text data but have knowledge about what words we want to base some of these topics, then semi-supervised topic modeling can help. This setup can help to uncover underrepresented topics or improve the relevance of topics over unsupervised cases.

2.3 Moral Foundation Using Sentiment Analysis

Moral rhetoric and framing play a role in increasing polarization and divisions in our societies Marietta, Dehghani, Brady et al. [31],[32],[33] , but also in a wide range of pro-social behaviors that can potentially bring people togetherVoelkel et al,Wolsko,Kidwell et al. [34],[35],[36] . In order to understand the relationship between hate, division, compassion, and unity in the digital age, we need to understand the dynamics of moral language online. In particular, capturing and investigating the moral sentiment of text can allow for the study of how individuals and groups expressed moral sentiment related to various downstream online and offline behaviors.

Moral sentiment assessment and classification are subjective tasks, and when done automatically using Natural Language Processing (NLP) techniques, this subjectivity results in the need for large and diverse, both in terms of topics and coders, sets of annotations. The Moral Foundations Twitter Corpus by Hoover et al [37] a collection of 35,108 tweets that have been curated from seven distinct domains of discourse and hand-annotated for 10 categories of moral sentiment (care, harm, fairness, inequality, loyalty, betrayal, authority, subversion, purity, and degradation) based on the Moral Foundations Theory by Haidt et al , Graham et al [38],[39]. This corpus has been used to design novel methods for moral sentiment classification [40] used in models to investigate the impacts of moral framing in other domains (e.g., misinformation and polarization Mutlu et al , Ruch et al [41],[42] and has been applied to train models that produce morally salient text (e.g., arguments and jokes, Alshomary et al. Yamane et al. [43],[44]) However, as useful as MFTC is, its training dataset is limited to the social media platform Twitter.

Different online social media platforms have different linguistic and social structural environments that may result in variations in moral language and behavior Curiskis et al. [45]. Beyond differences in social structure, different platforms have varying character limits (e.g. 280 characters on Twitter compared to 10k-40k on Reddit) which alters the language usage of users and therefore may contribute to differences in the use and effectiveness of moral rhetoric Candia et al. [46]. Additionally, different platforms have different policies concerning the levels of user anonymity and sensitive content moderation which may additionally influence the domains of morality discussed given the potential judgments from others. Research has argued that higher levels of anonymity reduce the feeling of responsibility and alter moral behavior online Lastly, modern NLP methods are known to require massive training data for producing sufficiently accurate, generalizable, and robust models. It has empirically been shown that diverse sets of training data, from different platforms and on different topics, can help improve the classification results by allowing the models to obtain generalized domain Kennedy et al. [47] , rather than surface knowledge restricted to a particular platform and a small set of topics.

As mentioned previously, the MFTC relied on on the Moral Foundation Theory's framework which is a pluralistic perspective of moral cognition and identifies multiple dimensions of moral values that have evolved to facilitate individual well-being, coalitional unity, and cooperation with strangers Haidt et al. Graham et al. [48, 49] The original version of the theory identified five separate but interrelated categories of moral concerns: Care/Harm, Fairness/Cheating, Loyalty/Betrayal, Authority/Subversion, and Purity/Degradation. Recently, a revision to the Moral Foundations Theory Atari et al. [50] split Fairness into distinct and new foundations of Equality and Proportionality, while retaining the other four existing foundations of Care, Loyalty, Authority, and Purity. This split aims to capture the distinct moral concerns of fairness in the procedure (proportionality) and equality of outcome (equality). In order to better understand the different nuances that will result from this theoretical change, we need to have an updated annotated corpus that complies with the latest theoretical revisions.

Together, the above reasons call for a corpus from a different platform, focused on a diverse set of topics and annotated by a diverse group of annotators. Here we address this need by introducing the Moral Foundations Reddit Corpus (MFRC), a collection of 16,123 Reddit.com comments annotated for 8 categories of moral sentiment and curated from 12 morally-relevant subreddits. Reddit is a public social media platform with approximately 52 million daily users who post content in over 138,000 active "subreddits" which are user-created and user-moderated communities about different subjects. Shared content and comments in subreddits are "voted" on by users which are used to decide the visibility of the post. Activity on Reddit has been the center of several prominent cultural moments including coordinated attempts to challenge short-sellers of GameStop stock or attempts to identify the Boston-city bombing terrorists.

The author focused on the corpus compilation effort on Reddit for several reasons. First, in comparison to Twitter, Reddit shares many of the same research-friendly features (e.g., responsiveness to current events, public posts, and available APIs) Proferes et al. [51] but is organized into what are called subreddits. Different subreddits have distinct topics and consistent communities. In relation to morality, these distinct communities and norms have led researchers to use Reddit to investigate moral conflicts across groups a phenomenon that is harder to investigate on Twitter (which does not have organized groups) or Facebook (where many groups are private). Second, Reddit provides more anonymity than many other social media platforms, potentially enabling users to speak their minds and express their opinions. Third, in addition to general differences in language usage, the lack of restrictions on the length of posts on Reddit may be particularly beneficial for training models. Fourth, we believe that Reddit has played a distinct role in contemporary politics. For example, the r/TheDonald and r/incels subreddits have been linked to political extremism Gaudette et al. [52] and mass shootings

Nicholas Helm et al. [53].

To understand the validity and relative performance of different text classification methods for identifying the moral concerns manifested in the MFRC, we report baseline results for several NLP approaches: Distributed Dictionary Representations, Bidirectional Encoder Representations from Transformers BERT models. These baselines can both be used as evidence about the relative performance of various NLP models on the task of moral concern detection and also inform future methodological work including helping calibrate new models of moral sentiment classification.

Finally, to facilitate research into annotator response patterns and bias, as recommended Prabhakaran et al. [54], the author provided psychological and demographic metadata of our annotators. The background and biases of human annotators have been shown to impact their annotations with particularly damaging effects that amplify pre-existing biases. Annotators' biases may be particularly relevant in domains characterized by high subjectivity, such as moral values. While, for example, an annotator's political ideology might not have a substantial influence on how they annotate "positive" and "negative" sentiment in a corpus of hotel reviews, it seems likely that their ideology could substantially influence how they annotate expressions of justice and respect in a politically relevant corpus.

The author provided a detailed description of the corpus, our annotation procedures, a set of baseline classification results using a range of methods, and a cross-corpus comparison with MFTC [4].

We have found from the theories of moral foundation that no research has been made on the subreddit /relationship, /dating, etc. This has given us the idea to perform the same analysis in order to analyze and find trends on the 'topic: dating with an end date'.

Chapter 3

Dataset and Preprocessing

To develop a mechanism for continuous forecasting of the Reddit dataset. The process involves the extraction of data from API, data cleaning, and preprocessing of the data. In this section, we will also discuss the data collection and preparation process.

3.1 Data Sources

Reddit is an American social news aggregation, web content rating, and discussion website. It is an excellent source of data for a social scientist, with over 8,000 reddits sent per second. Registered members submit content to the site such as links, text posts, and images, which are then voted up or down by other members. Posts are organized by subject into user-created boards called “subreddits”, which cover a variety of topics including news, science, movies, video games, music, books, fitness, food, and image-sharing. This can be as basic as looking for keywords and phrases like ‘dating’ or ‘dating with end date’ or can be more advanced, aiming to discover general topics contained in a dataset. Submissions with more up-votes appear towards the top of their subreddit and, if they receive enough votes, ultimately on the site’s front page. The reddits that millions of users send can be downloaded and analyzed to try and investigate mass opinion on particular issues.

3.1.1 Data Extraction

The data is extracted for both posts and submissions(comments) in two different csv files as we have collected the data, the csv files are created for each day. These csv are concatenated. There are separate csv for posts and comments for each year as the duration for which the data is collected is 2019, 2020, and 2021. These files are then merged over the post id column for each year (refer to Figure 3.1). The subreddits for which the data is collected are: 'dating', 'relationship', 'love', and 'long distance'. Each subreddit has three file for each year.

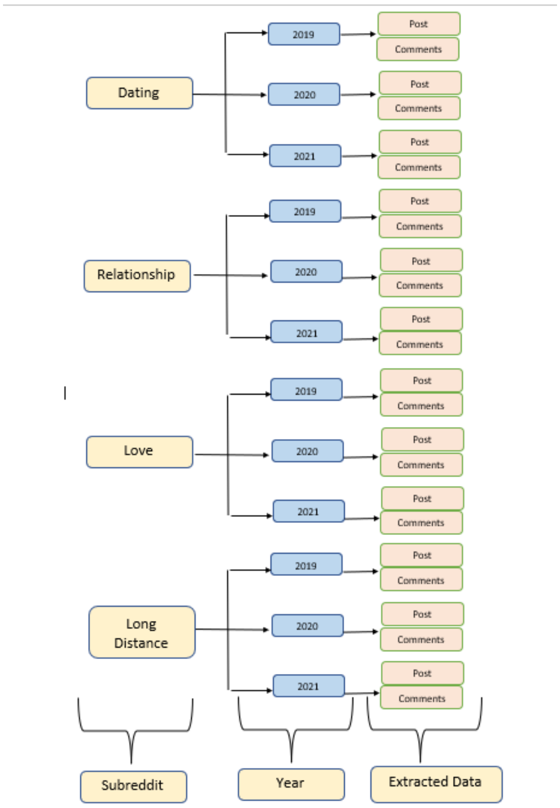

Figure 3.1: The data extraction structure using 3rd party Reddit API: Data is extracted separately for the subreddits and the year 2019, 2020 and 2021

3.1.2 Data cleaning

Data cleaning refers to the series of steps that ensures that the underlying data used for high-end analysis or modeling is complete, correct, accurate, and relevant. These csv files are loaded in python pandas are data cleaning is performed on the column 'post' where the punctuation's, blank space, empty rows, duplicate values, nltk stopwords and digits are removed and stored in a new data-frame column.

3.1.3 Data Preprocessing

The cleaned posts are then preprocessed or vectorized using tfidf vectorizer. TF-IDF is an abbreviation for Term Frequency Inverse Document Frequency. This is very common algorithm to transform text into a meaningful representation of numbers which is used to fit machine algorithm for prediction. TF-IDF Vectorizer is a measure of originality of a word by comparing the number of times a word appears in the document with the number of documents the word appears in.

We choose TF-IDF over count vectorizers and bag-of-words because it not only focuses on the frequency of words present in the corpus but also provides the importance of the words. We can then remove the words that are less important for analysis, hence making the model building less complex by reducing the input dimensions. And, Bag of Words vectors are easy to interpret. However, TF-IDF usually performs better in machine-learning models.

The usage of the TF-IDF vectorizer is: First, we will create a vectorizer object using 'TfidfVectorizer()' and fit and transform the text data into vectors. After that, we will use vectorizers to extract the names of the words. We will now use TF-IDF tokens and vectors to create a pandas data frame.

3.1.4 N-gram Analysis

An N-gram model is built by counting how often word sequences occur in corpus text and then estimating the probabilities. Since a simple N-gram model has limitations, improvements are often made via smoothing, interpolation and back off.

An N-gram model is one type of a Language Model (LM), which is about finding the probability distribution over word sequences. A model that simply relies on how often a word occurs without looking at previous words is called unigram. If a model considers only the previous word to predict the current word, then it's called bigram. If two previous words are considered, then it's a trigram model. For example, consider a sentence: "I think I've been approaching the dating scene all wrong" (refer to Figure 3.2)

unigram	I	think	I've	been	approaching	the	dating	scene	all	wrong
bigram	I think	I've been	Approaching the	dating scene	all wrong					
trigram	I think I've	been approaching the	Dating scene all	wrong						

An n-gram model for the above example would calculate the following probability:

P('I think I've been approaching the dating scene all wrong') =

P('I', 'think', 'I've', 'been', 'approaching', 'the', 'dating', 'scene', 'all', 'wrong') =

P('I')P('think '|'I')P('I've been '|'approaching the')P('dating' | ' scene all wrong")

Figure 3.2: N-gram analysis process for unigram, bigram and trigram analysis

N-gram models including part-of-speech tagging, natural language generation, word similarity, sentiment extraction and predictive text input

The ngrams for the subreddits are: 'expiration', 'dating', 'dating expiration', 'end date', 'dating with end date', 'short term committed relationship', 'dating with expiration date', 'relationship with no future', 'dating end', 'dating end date', 'short term', 'not a committed relationship', 'relationship no future', 'expiration date' , 'no future'.

The figure 3.3 defines the number of occurrence of words related to dating followed by short-term are involved

- Highest number of word occurrence by the user are involved in dating followed by 'short term'.
- In 2021 have used the term highest number of times followed by 2019 and 2020.

Phrases/ Year	2019	2020	2021
expiration'	3215	3047	5884
dating'	2830955	2787424	5241447
'dating expiration'	0	0	0
end date'	6129	4662	8163
'dating with end date'	0	0	0
'short term committed relationship'	0	0	0
'dating with expiration date'	0	1	5
'relationship with no future'	24	30	255
'dating end'	520	168	540
'dating end date'	0	0	0
'short term'	20921	14527	27200
'not a committed relationship'	46	57	54
' relationship no future'	0	1	0
'expiration date'	2608	2460	4470
'no future'	6671	8747	11997

Figure 3.3: Statistics of the data extracted using n-grams analysis: A sanity check process which shows the number of occurrence of words related to dating followed by short-term

3.2 Conclusion

In this section, we specified the data extraction using the 3rd party API. We have also created a pipeline for data cleaning and preprocessing and used in further topic modeling. After setting up a pipeline for data preprocessing, according to the criteria defined in the previous sections, we are now ready to formulate the semi-supervised topic modeling, sentimental analysis, and moral foundation and use it on the processed data.

Chapter 4

Methodology

In this section, we formulate the problem. We discuss all the methods we used to analyze the relationship trends that come with an end date and moral foundation of the user.

4.1 Problem Formulation

Based on the research questions, our model has to be formulated such that the results can be achieved efficiently. The modeling methodology is divided into two parts refer Figure 4.1:

- Part 1: A semi-supervised learning Corex topic modeling. It assigns a topic to each post followed by sentimental analysis which assigns positive, negative, and neutral sentiments to each post.
- Part 2: Semi-supervised label propagation method is applied to the post where the best machine learning model is selected which is created over the annotated data of size 500. The sentiments are then used as the vice and virtue of the moral foundation.

The problem is formulated to analyze the trends of relationships with end dates

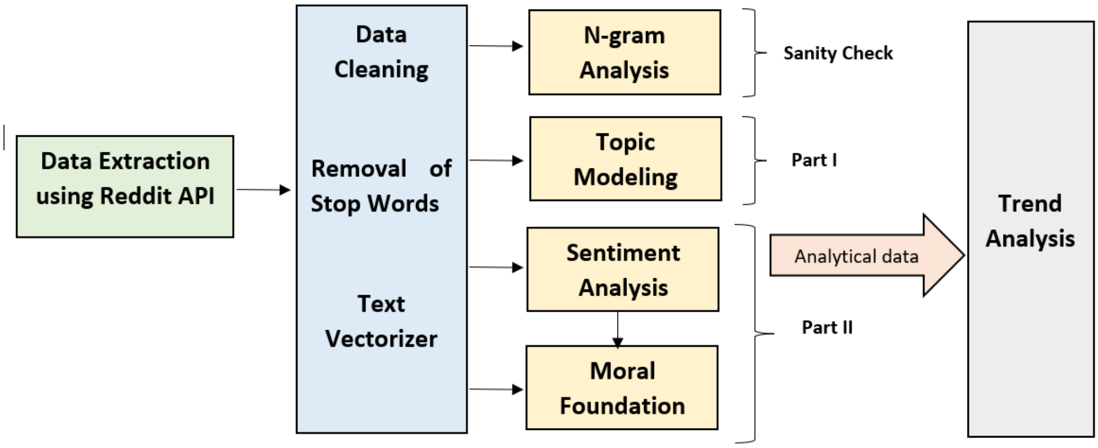

Figure 4.1: Modeling methodology: Is divided into two parts: Topic modeling and sentiment analysis and moral foundation

over the years and the moral foundation behind such posts by the user. As the problem involves forecasting tends over the Reddit post.

4.2 Corex Topic Modeling

Supervised/ Semi-supervised topic modeling allows the user to inject prior knowledge into the topic model. In particular, there are versions where the user can supply the model with topic "seed" words, and the model algorithm then encourages topics to be built around these seed words. This not only resolves the issue but more broadly gives us the flexibility to steer topics toward relevant themes by simply adding keywords, whilst also leaving room to uncover "unknown" topics. Here we, have used the anchored CorEx modeling technique. which implements a completely novel approach to topic modeling compared to LDA. Unlike LDA, CorEx 'correlation explanation' does not assume the data generation process, instead approaching topic modeling in an information-theoretic manner. CorEx treats each word as a random variable and seeks to find collections of words (i.e. topics) that best "explain" the data. Explaining translates into decorrelating the observed word count data, which proceeds by optimizing a mutual information-based objective function. In figure 4.2 The authors observe that this objective function realizes an information bottleneck and this can be exploited to modify the model to accept seed, or “anchor”, words as input [30].

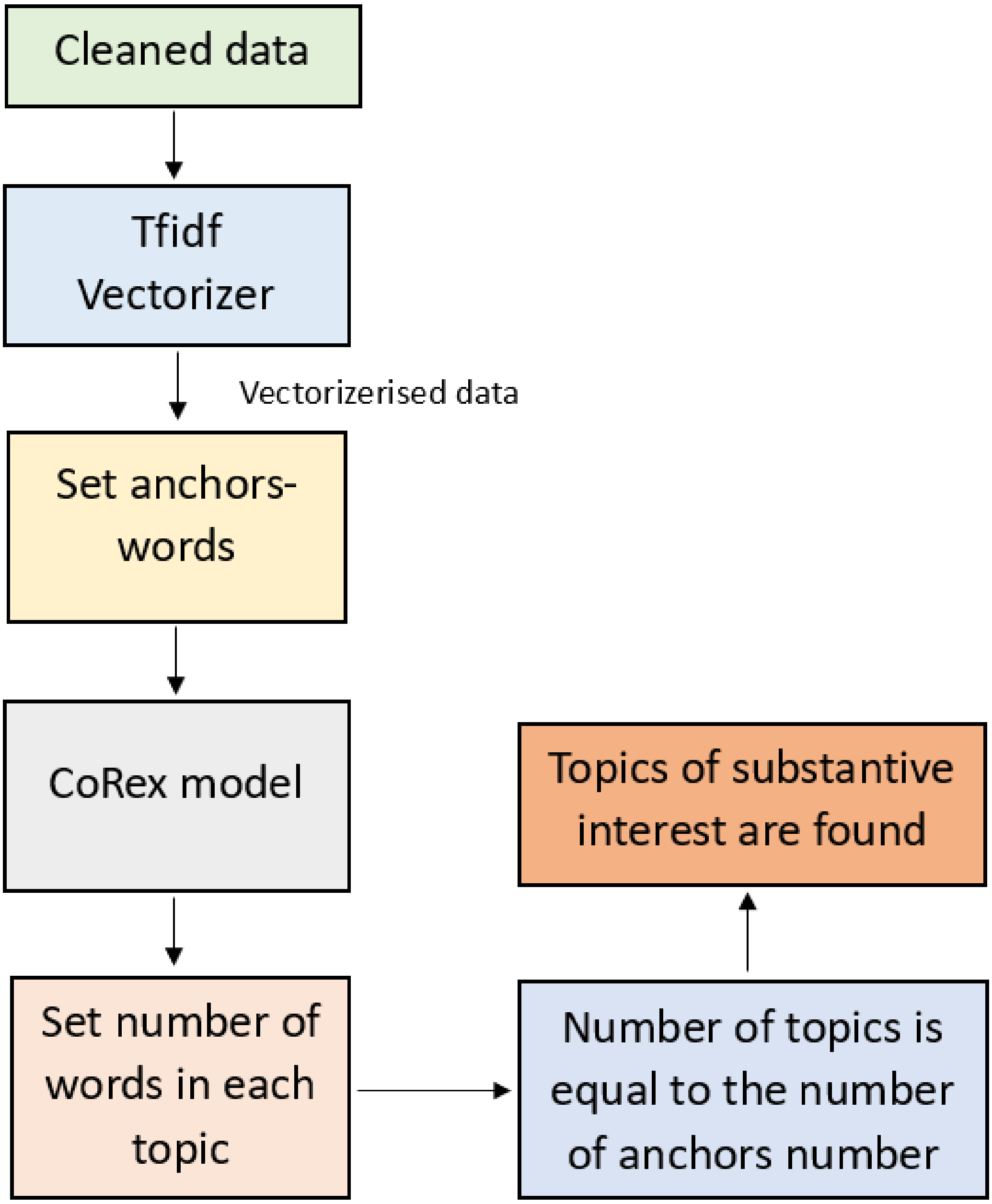

Figure 4.2: CoRex Topic Modeling Architecture: Semi-supervised modeling technique that accepts anchors to yield rich topics that are maximally informative and of substantive interest in the corpus

We employed the CorEx topic model using the anchors:

- ‘dating’, ‘expire’, ‘end’, ‘date’
- ‘short’, ‘term’
- ‘break’, ‘relationship’

- ‘toxic’, ‘relationship’
- ‘never’, ‘committed’,‘cheating’, ‘jerk’, ‘not’, ‘serious’, ‘relationship’

Topic	Topic Name	Topic Aspects
1	Dating with end date	dating, date, end, first, guy, someone, met, seeing, first time, else
2	Short term	like, feel, long, distance, feel like, long distance, term, make, time, love
3	Break up	break, up, years, know, wants, don, months, want, together, back
4	Toxic relationship	relationship, toxic, distance relationship, year relationship, relationship boyfriend, toxic relationship, new relationship, relationship advice, term relationship, abusive relationship
5	Never committed	relationship, never, cheating, serious, not, committed, jerk, ve never, never met, boyfriend never

Figure 4.3: Topics formed that are equally anchored with the words/anchors provided

4.3 Sentiment Analysis

The objective of Sentiment Analysis [55] is related to contextual mining of text which identifies and extracts subjective information in the source material, and helps a business to understand the social sentiment of their brand, product, or service while monitoring online conversations. However, analysis of social media streams is usually restricted to just basic sentiment analysis and count-based metrics. A key aspect of sentiment analysis is polarity classification. Polarity refers to the overall sentiment conveyed by a particular text, phrase, or word. This polarity can be expressed as a numerical rating known as a “sentiment score”. For example, this score can be a number between -100 and 100 with 0 representing neutral sentiment. This score could be calculated for an entire text or just for an individual phrase. Here we have used the VADER method. VADER (Valence Aware Dictionary for Sentiment Reasoning) is a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion refer figure 4.4. It is available in the NLTK package and can be applied directly to unlabeled text data.

VADER sentimental analysis relies on a dictionary that maps lexical features to emotion intensities known as sentiment scores. The sentiment score of a text can be obtained by summing up the intensity of each word in the text.

For example- Words like 'love', 'enjoy', 'happy', and 'like' all convey a positive sentiment. Also, VADER is intelligent enough to understand the basic context of these words, such as "did not love" as a negative statement.

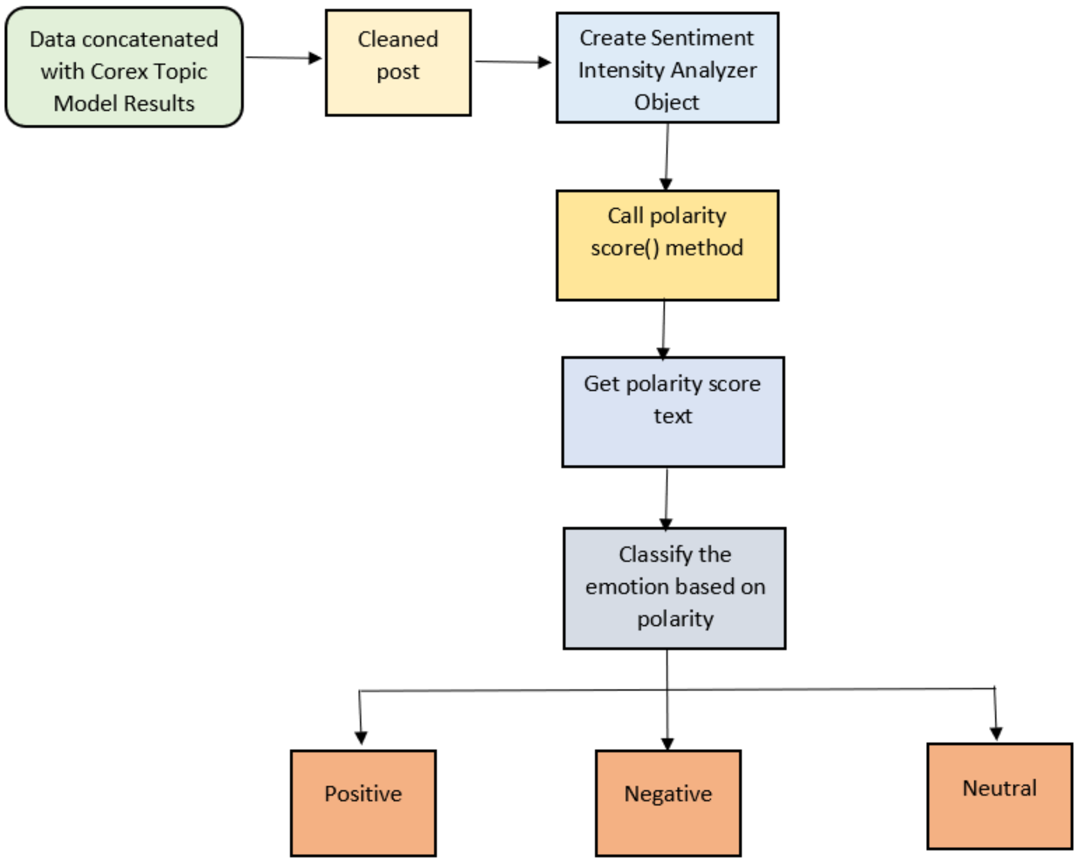

Figure 4.4: Sentiment analysis model flow: Using the nltk sentiment Vader library sentiment intensity analyzer used, followed by polarity score method to classify the emotion threshold value of 0.05 is set

The descriptive statistics of sentiment analysis are (refer figure 4.5):

- Positive sentiments: 316028
- Neutral sentiments: 291364
- Negative sentiments: 270368

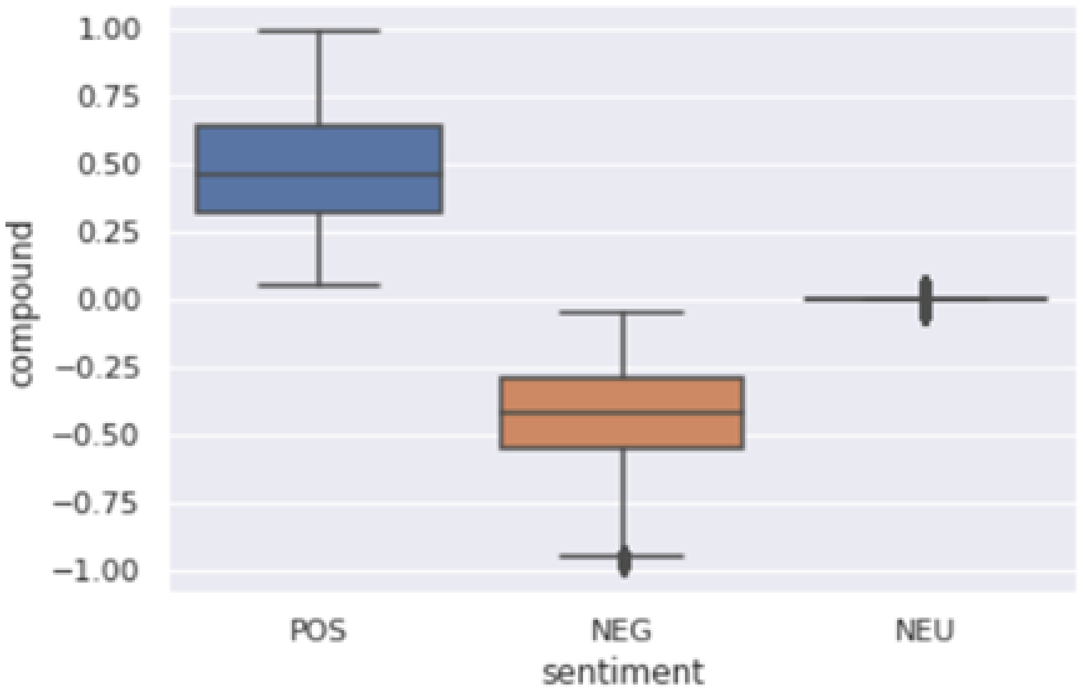

Figure 4.5: Descriptive statistics of sentiment analysis: The graph justifies that positive sentiment scores are above the threshold, negative sentiment scores are below the threshold and neutral are aligned with the threshold

4.4 Moral Foundation Analysis

Moral Foundations Theory was created by a group of social and cultural psychologists to understand why morality varies so much across cultures yet still shows so many similarities and recurrent themes. In brief, the theory proposes that several innate and universally available psychological systems are the foundations of "intuitive ethics." Each culture then constructs virtues, narratives, and institutions on top of these foundations, thereby creating the unique moralities we see around the world, and conflicting within nations too. The five foundations for which we think the evidence is best are:

1) Care/harm: This foundation is related to our long evolution as mammals with attachment systems and an ability to feel (and dislike) the pain of others. It underlies the virtues of kindness, gentleness, and nurturance.

2) Fairness/cheating: This foundation is related to the evolutionary process of reciprocal altruism. It generates ideas of justice, rights, and autonomy. [Note: In

our original conception, Fairness included concerns about equality, which are more strongly endorsed by political liberals. However, as we reformulated the theory in 2011 based on new data, we emphasize proportionality, which is endorsed by everyone, but is more strongly endorsed by conservatives]

3) Loyalty/betrayal: This foundation is related to our long history as tribal creatures able to form shifting coalitions. It underlies the virtues of patriotism and self-sacrifice for the group. It is active anytime people feel that it's "one for all and all for one."

4) Authority/subversion: This foundation was shaped by our long primate history of hierarchical social interactions. It underlies virtues of leadership and followership, including deference to legitimate authority and respect for traditions.

5) Purity/degradation: This foundation was shaped by the psychology of disgust and contamination. It underlies religious notions of striving to live in an elevated, less carnal, more noble way. It underlies the widespread idea that the body is a temple that can be desecrated by immoral activities and contaminants (an idea not unique to religious traditions).

These five foundations comprise the building blocks of morality, regardless of the culture. In other words, while every society constructs its morality, it is the varying weights that each society allots to these five universal foundations that create the variety. Haidt likens moral foundation theory to an "audio equalizer," with each culture adjusting the sliders differently.

Colin Prince [56] suggests that the researchers, however, were not content to simply categorize moral foundations—they have tied the foundations to political leanings. And it is here that moral foundation theory becomes a truly practical tool for the lawyer.

Brendan Kennedy et.al [57] Moral framing and sentiment can affect a variety of online and offline behaviors, including donation, pro-environmental action, political engagement, and even participation in violent protests. Various computational methods in Natural Language Processing (NLP) have been used to detect moral sentiment from textual data, but to achieve better performances in such subjective tasks, large sets of hand-annotated training data are needed. Previous corpora annotated for moral sentiment have proven valuable, and have generated new insights both within NLP and across the social sciences, but have been limited to Twitter.

To facilitate improving our understanding of the role of moral rhetoric, we present the Moral Foundations Reddit Corpus, a collection of English Reddit comments that have been manually annotated 500 data count from 4 distinct subreddits, by an annotator for 5 categories of moral sentiment (i.e., Care, Fairness, Purity, Authority, Loyalty) based on the updated Moral Foundations Theory (MFT) framework, which is further extended to Vice and Virtue. We have performed the moral foundation using the computational machine learning process. Therefore, it is called Computational Moral Foundation(CMF).

Refer figure 4.6, as we computed multiple model using the trained data for semi-supervised learning- Label Propagation: Dummy Classifier, Decision Tree, Random Forest, Multinomial Naive Bayes, and Gradient Boosting. Each model is developed using Gridsearch Cross Validation. Among which Gradient Boosting is the highest on evaluation metrics. Later, each label is further classified into vice and virtue based on the sentiments.

4.5 Gridsearch Cross Validation

GridSearchCV is the process of performing hyperparameter tuning in order to determine the optimal values for a given model.The performance of a model significantly depends on the value of hyperparameters.There is no way to know in advance the best values for hyperparameters so ideally, it tries all possible values to know the optimal values. Doing this manually could take a considerable amount of time and resources and thus we use GridSearchCV to automate the tuning of hyperparameters. GridSearchCV is a function that comes in Scikit-learn's (or

SK-learn) model selection package. GridSearchCV tries all the combinations of the values passed in the dictionary and evaluates the model for each combination using the Cross-Validation method. Hence after using this function we get accuracy/loss for every combination of hyperparameters and we can choose the one with the best performance.

4.6 Gradient Boosting Algorithm

Gradient boosting is one of the most powerful techniques for building predictive models. This algorithm builds an additive model in a forward stage-wise fashion; it allows for the optimization of arbitrary differentiable loss functions. In each stage n-classes- regression trees are fit on the negative gradient of the loss function, e.g. binary or multiclass log loss. Binary classification is a special case where only a single regression tree is induced. There is a technique called the Gradient Boosted Trees whose base learner is CART (Classification and Regression Trees) refer figure 4.7.

There is an important parameter used in this technique known as Shrinkage.

Shrinkage refers to the fact that the prediction of each tree in the ensemble is shrunk after it is multiplied by the learning rate (eta) which ranges between 0 to 1. There is a trade-off between eta and number of estimators, decreasing learning rate needs to be compensated with increasing estimators in order to reach certain model performance. Since all trees are trained now, predictions can be made. Each tree predicts a label and final prediction is given by the formula:

$$y(pred) = y1 + (eta * r1) + (eta * r2) + ... + (eta * rN) \tag{4.1}$$

The class of the gradient boosting regression in scikit-learn is GradientBoostingRegressor. A similar algorithm is used for classification known as GradientBoostingClassifier.

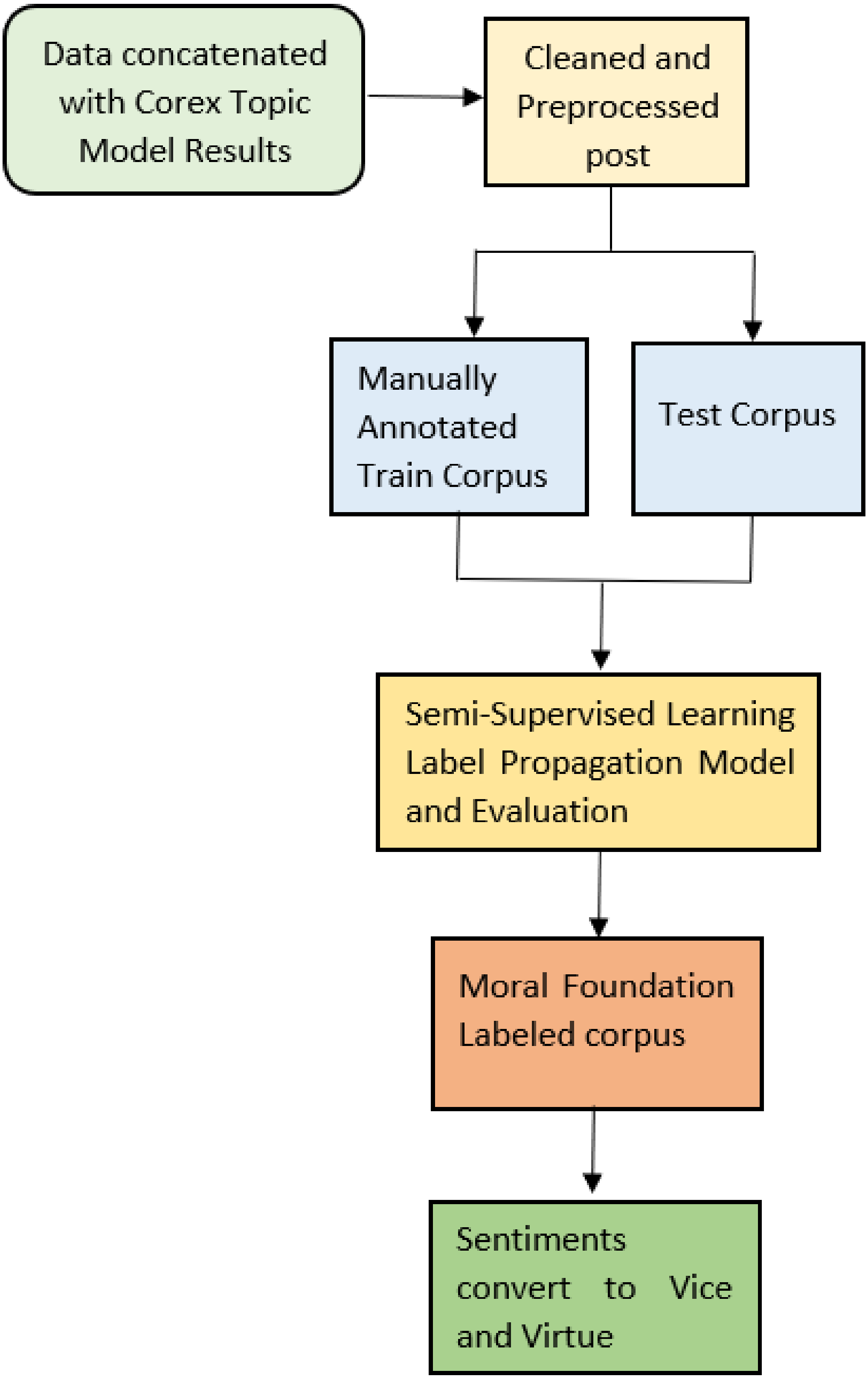

Figure 4.6: Architecture of Moral Foundation: Moral foundation is manually curated in a subset of data which is used as the trained data for semi-supervised learning- Label Propagation; and sentiment analysis is replaced to vice and virtue of the morals

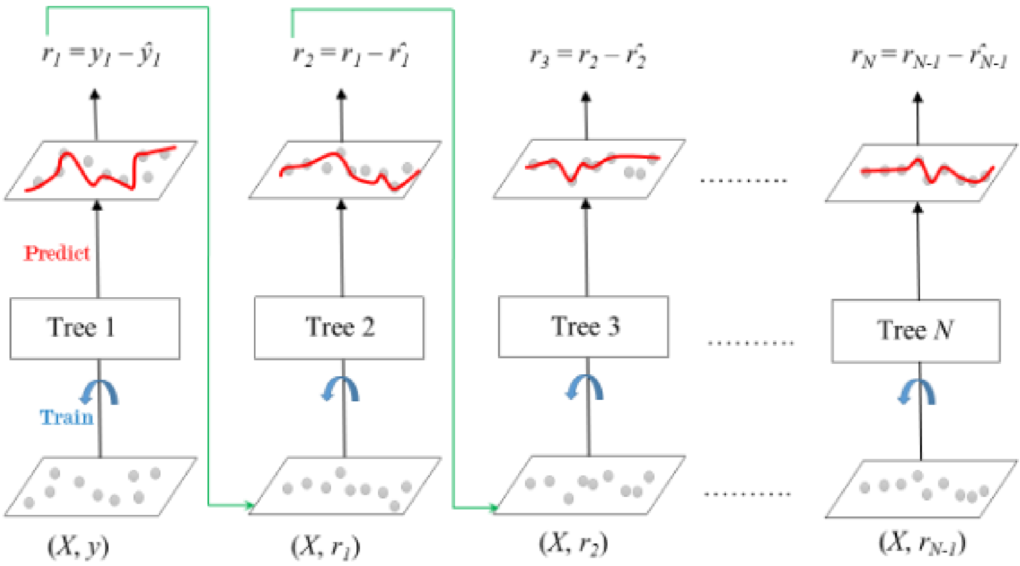

Figure 4.7: Gradient Boosted Trees for Regression: The ensemble consists of N trees. Tree1 is trained using the feature matrix X and the labels y. The predictions labelled y1(hat) are used to determine the training set residual errors r1. Tree2 is then trained using the feature matrix X and the residual errors r1 of Tree1 as labels. The predicted results r1(hat) are then used to determine the residual r2. The process is repeated until all the N trees forming the ensemble are trained

4.7 Conclusion

With the above steps, we have completed our preprocessing and we build a model using the extracted data. We will compare different models and perform the yearly analysis.

Chapter 5

Experiments

This chapter discusses the various experiments pertaining to the proposed hypothesis and their findings. The experiments are performed based on objectives that are discussed in chapter 1. Every objective is addressed in the experiment section and is achieved in the course of implementation. The Results that have obtained after following the steps from the previous chapter are discussed in this chapter. Trend analysis over years is also done in the results section.

The following experiments are done and the results are reported systematically.

- Does dating comes with an end date? What are the trends of dating with an end date over years, six months, or three months? (Experiment 1)

- What is the trend of dating with an end date over sentiments? - This may be a motivation behind participating in a relationship with an expiration date. (Experiment 2)

- What is the trend of the moral foundation of the users while getting into expiry dating? (Experiment 3)

5.1 Experimental Setup

After extracting the data through Reddit. It was observed the subreddit 'relationship' has the highest number in the year 2019 followed by subreddit 'dating' then 'long-distance' and then 'love' (Refer figure 5.1).

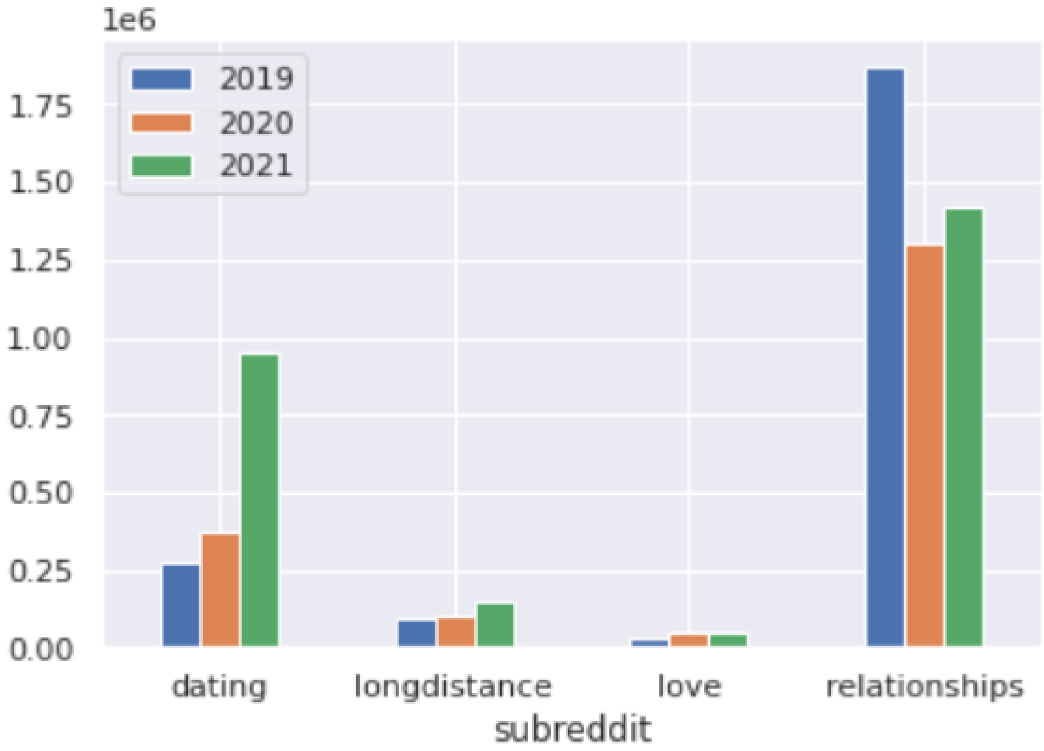

Figure 5.1: Extracted data counts per subreddit

After performing the data cleaning, preprocessing, n-gram analysis, and Corex-Semi-supervised Topic Modeling we have applied the topics to the anchors.

5.1.1 Experiment 1: Trends of expiry dating

Here, we have performed the analysis over a year, six-month, and three-month duration. This will give a clear understanding of the topic trends over the entire duration and subset of duration.

While performing the yearly trend analysis, It was observed in figure 5.2 that 'Topic 1 : Dating with end date' has the highest trends among other topics and a higher number of counts in the year 2021. 'Topic 2: Short term with the highest number of count in 2021. 'Topic 3: Break up' is with the highest number of counts in the year 2019. ‘Topic 4: Toxic relationships’ is with the highest number of counts in the year 2021. ‘Topic 5: Never committed’ is with the highest number of counts in the year 2021.

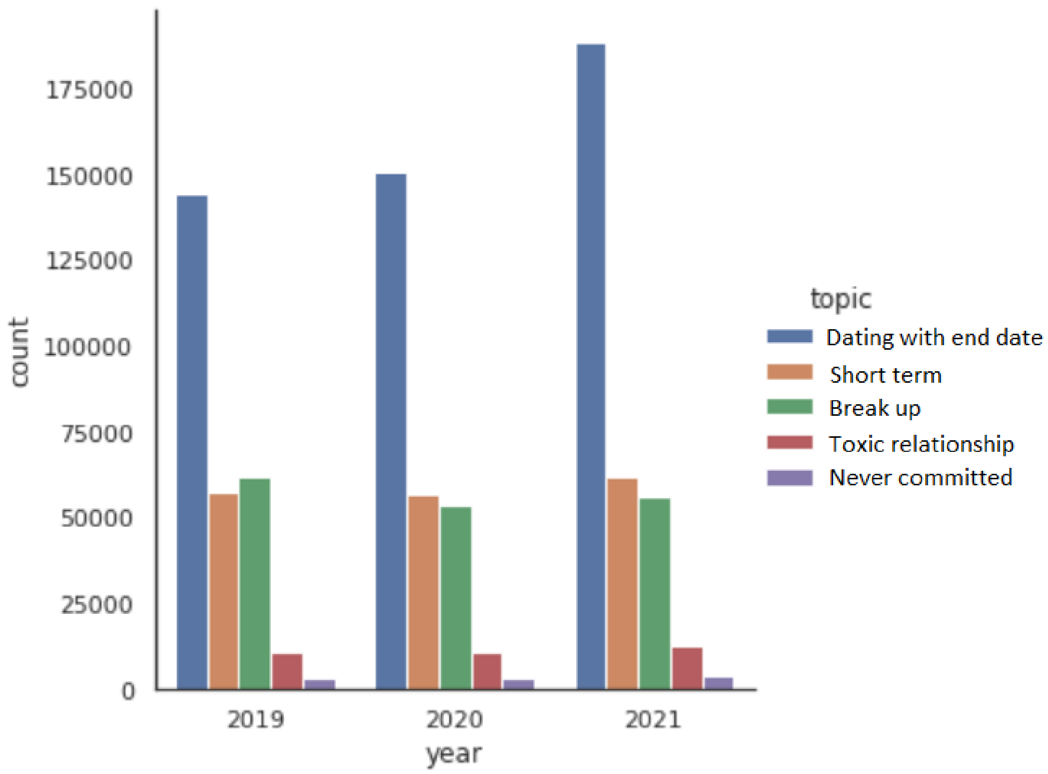

Figure 5.2: Topic trends per year: ‘Topic 1 : Dating with end date’ has the highest trends among other topics and a higher number of count is in the year 2021

While performing the six-month trend analysis, It was observed in figure 5.3 that ‘Topic 3 : Break up’ has the highest trends followed by ‘Topic 2 : Short term’. The higher number of counts is in the duration of ‘2019-01-31’ to ‘2019-07-31’. ‘Topic 1 : Dating with end date’ has the highest number of trends in the duration of ‘2021-07-31’ to ‘2022-01-31’. ‘Topic 5: Never committed’ has the highest number of trends in the duration of ‘2019-01-31’ to ‘2019-07-31’.

While performing the three-month trend analysis, It was observed that in figure 5.4 ‘Topic 3 : Break up’ has the highest trend count in the duration of ‘2019-01-31’ to ‘2019-04-30’. ‘Topic 2 : Short term with the highest trend count is in the duration of ‘2020-10-31’ to ‘2021-01-31’ then ‘Topic 1 : Dating with end date’ has the highest trend in the duration of ‘2021-07-31’ to ‘2021-10-31’.‘Topic 5: Never committed’ the highest trend is seen in the duration of ‘2020-10-31’ to ‘2021-01-

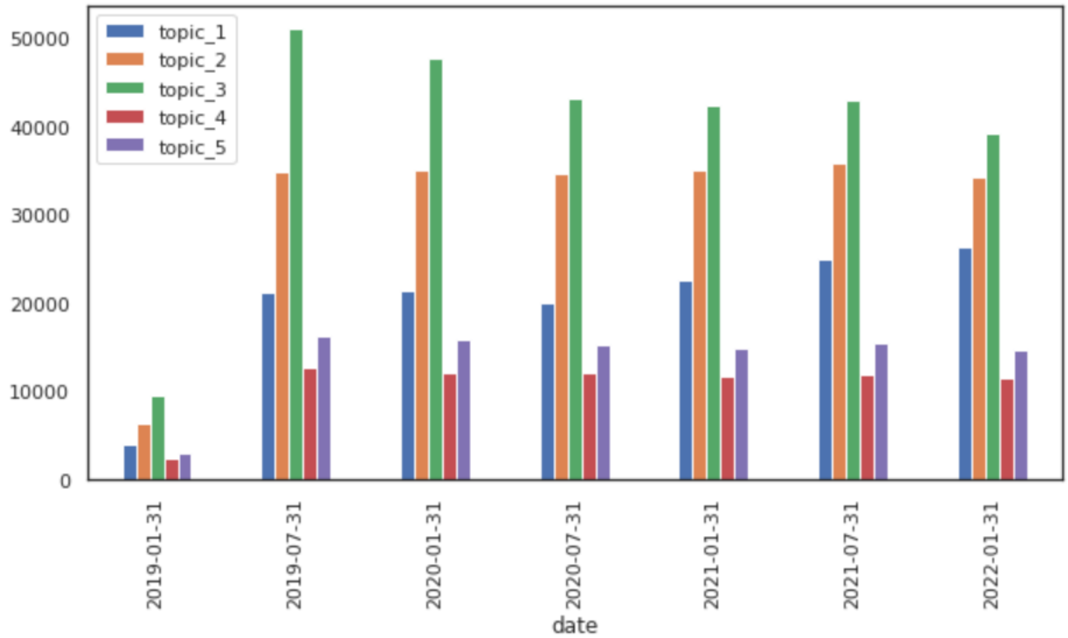

Figure 5.3: Topic trends in Six-month duration: 'Topic 3 : Break up' has the highest trends in the duration of '2019-01-31' to '2019-07-31'

31'. There is not much fluctuation in 'Topic 4 : Toxic relationship' but the highest trend is seen in the duration of '2020-10-31' to '2021-01-31'.

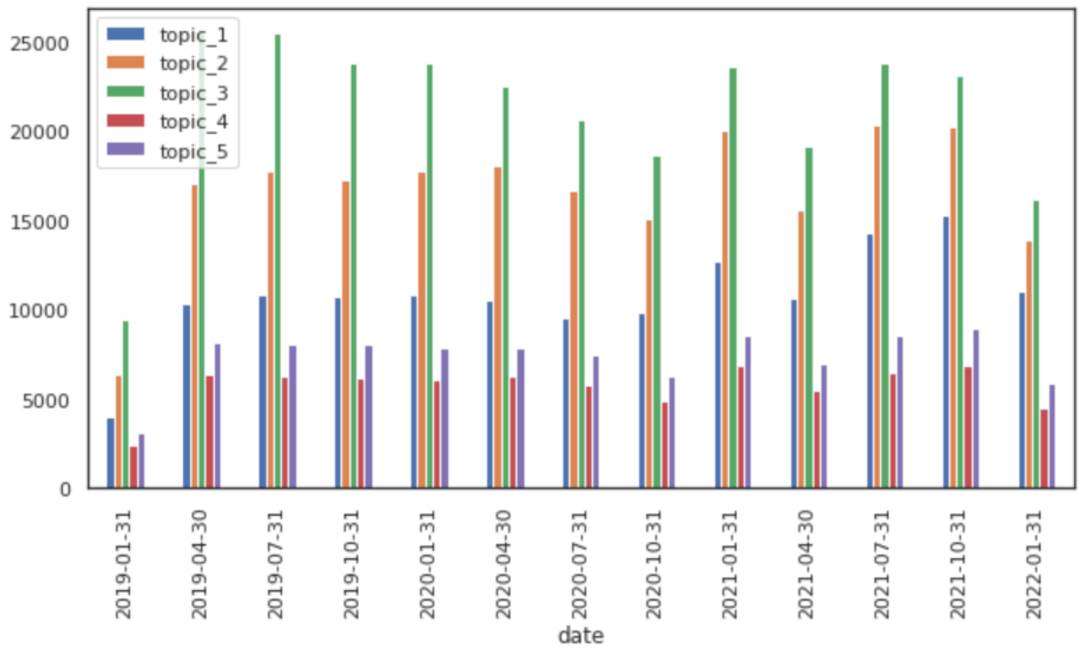

Figure 5.4: Topic trends in Three-month duration: 'Topic 3 : Break up' has the highest trends in the duration of '2019-01-31' to '2019-04-30'

5.1.2 Experiment 2: Dating/ Relationship trends over sentiment analysis and subreddits

This analysis is the result of the sentiment analysis performed. It gives a clear understanding of the user's sentiment about getting in relationships or dating. Also, it gives a variety of sentiments over the subreddit and the topics.

It was observed in the figure 5.5 'Topic 2 : Short term' has the highest number of positive count in the subreddit 'love', followed by 'Topic 1 : Dating with end date' has the highest number of positive count in the subreddit 'love' which shows users are more interested in the short term love and not interested in the affectionate long term love with each other. 'Topic 5: Never committed' has the highest negative count in the subreddit 'relationships', followed by 'Topic 5: Never committed' in the subreddit 'long distance' which shows users do not like to get committed in a relationship, especially in long distance relationships.

5.1.3 Experiment 3: Trends of Moral Foundation

Let us recall, here we have spilt the dataset in train and test. For training, we manually annotated 500 data items and built multiple models using gridsearch cross validation pipeline. Gradient Boosting is considered as it has the highest accuracy. The test data is then labeled using gradient boosting label propagation (refer figure 5.6)

After the label propagation, it was observed (refer figure 5.7) moral foundation 'Fairness' has the highest number of counts followed by 'Care', 'Loyalty', ' Purity', and 'Authority'.

Next, we have performed the comparative analysis of the sentiment score and the morals (refer figure 5.8). 'Care' has the highest positive sentiments which show users get into the relationship due to care, followed by 'Loyalty', 'Purity', 'Authority', and 'Fairness'. 'Loyalty' has the highest negative sentiments which shows users are less loyal towards the relationship, followed by 'Fairness', 'Care', 'Authority', and 'Purity'.

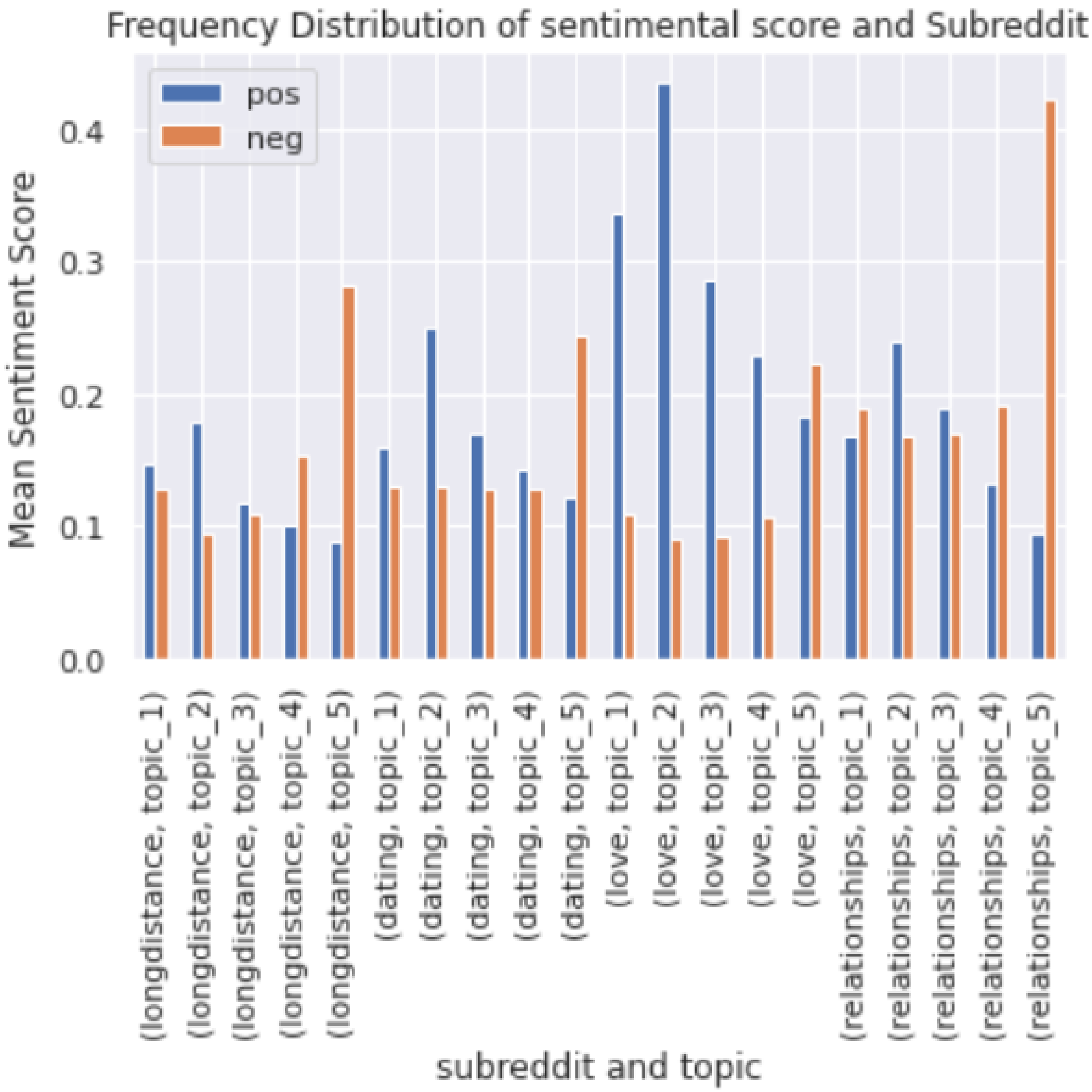

Figure 5.5: Topic trends over sentiments: 'Topic 2 : Short term' has the highest number of positive count in the subreddit 'love' and 'Topic 5: Never committed' has the highest negative count in the subreddit 'relationships'

The sentiment analysis is again performed for positive and negative thresholds. This is done to remove the neutral sentiments. The sentiments negative sentiments are then replaced by vice and positive sentiments are then replaced by virtue. 'Fairness- Vice' has the highest frequency among the other moral foundations, followed by 'Care- Vice', 'Loyalty- Vice', 'Purity- Vice', and 'Authority- Vice'. The same sequence is followed by 'Virtue' (refer figure 5.9).

It was observed in the figure 5.10 'Fairness- Vice' has the highest count followed by 'Fairness- Virtue'. This shows users are more getting relationship with the 'Fairness- Vice' morals.

Model/ measures	Accuracy	F1 average	Precision average	Recall average
Dummy classifier	0.65	0.51	0.42	0.65
Decision tree	0.63	0.62	0.62	0.63
Naive Bayes	0.65	0.51	0.67	0.65
Gradient Boosting	0.68	0.66	0.67	0.68
Random Forest	0.65	0.53	0.69	0.65

Figure 5.6: Model Development Metrics: Gradient Boosting cross validation is considered as it has the highest accuracy for training the semi-supervised learning label propagation model

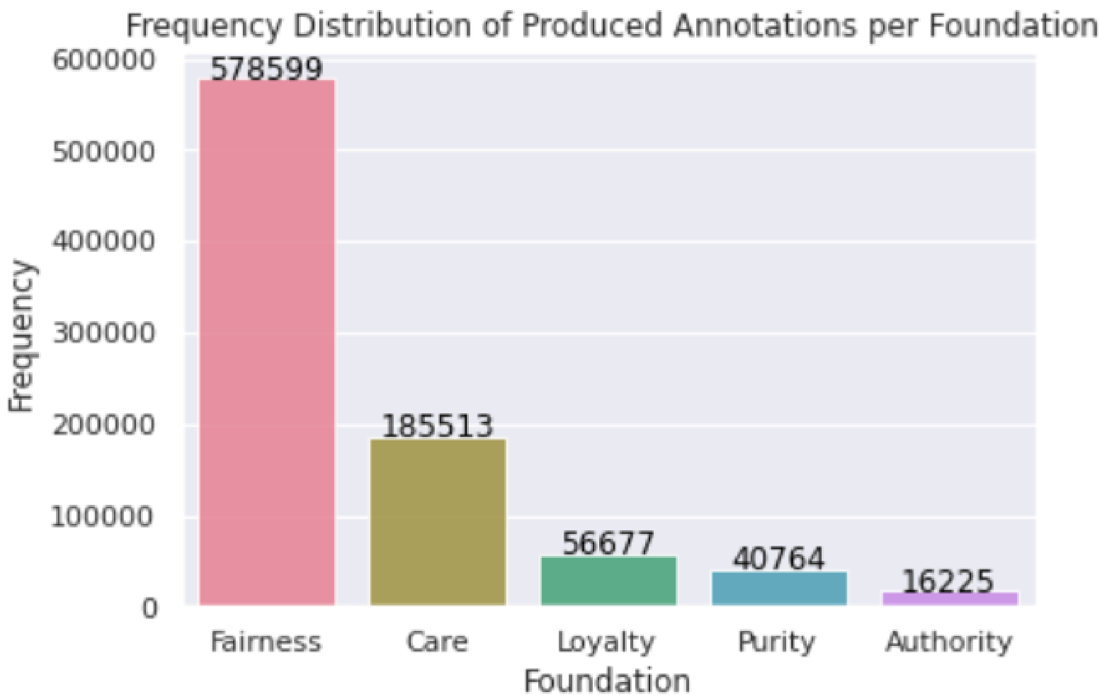

Figure 5.7: Label Propagation counts: 'Fairness' has the highest number of count

In the trend analysis of moral foundation over vice and virtue in figure 5.11 'Fairness- Vice' and 'Fairness- Virtue' has the highest frequency distribution in the year 2021 followed by the year 2020 and 2019.

Let us observe the variation of morals over topics in figure 5.12 'Topic 1

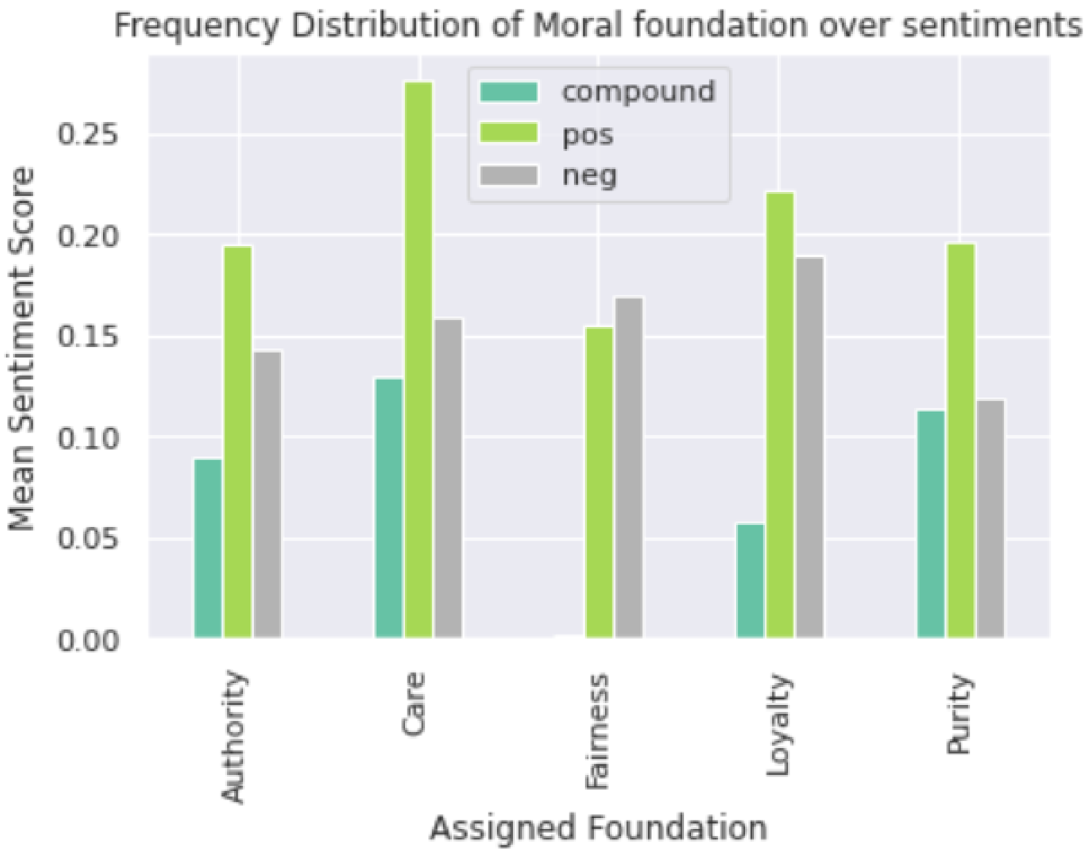

Figure 5.8: Frequency distribution of Moral foundation over sentiments: 'Care' have the highest positive sentiments and 'Loyalty' has the highest negative sentiments.

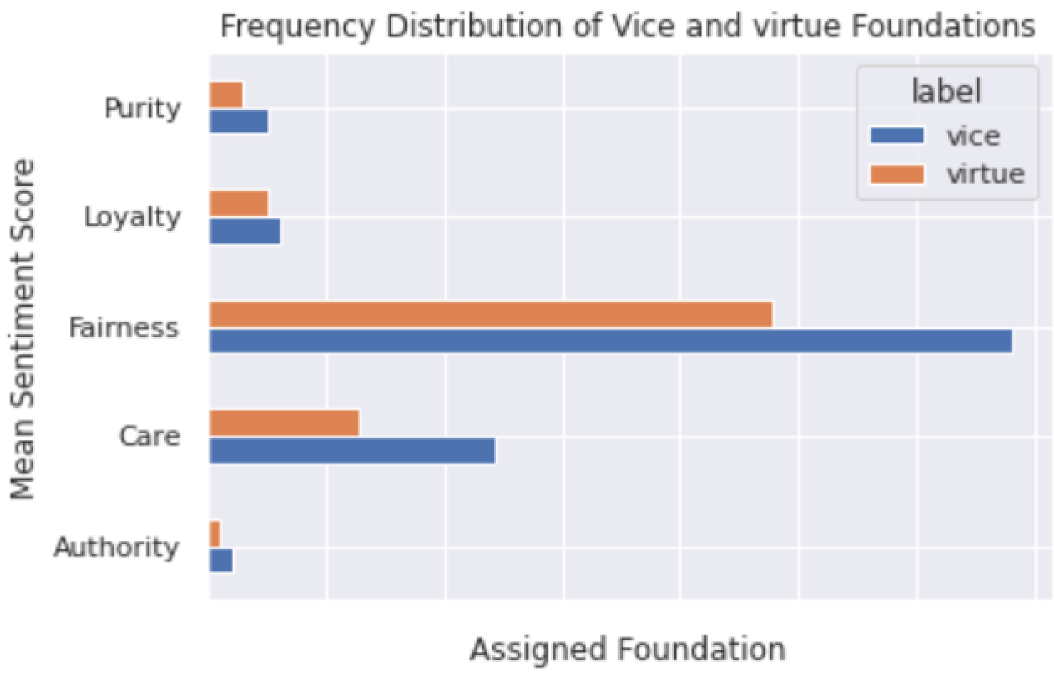

Figure 5.9: Frequency distribution of Moral foundation over Vice and Virtue: 'Fairness- Vice' and 'Fairness- Virtue' has the highest frequency.

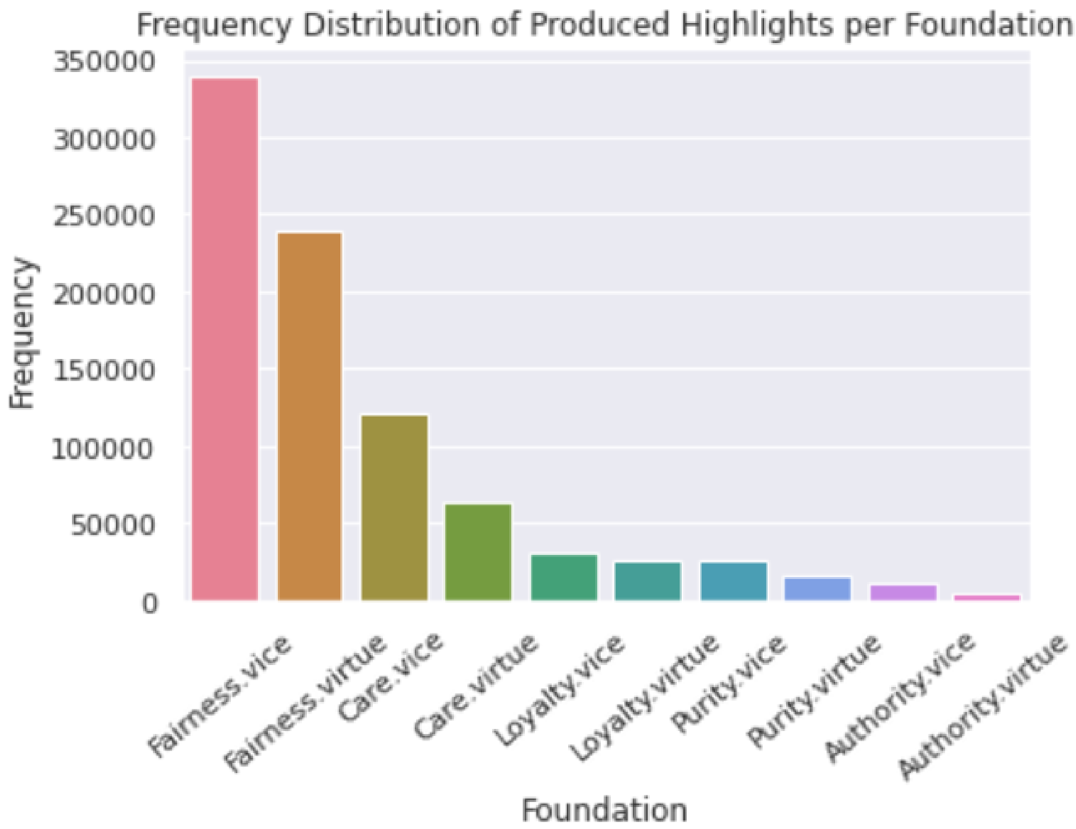

Figure 5.10: Frequency distribution of Produced Highlights per Foundation foundation over Vice and Virtue: 'Fairness- Vice' has the highest count followed by 'Fairness- Virtue'

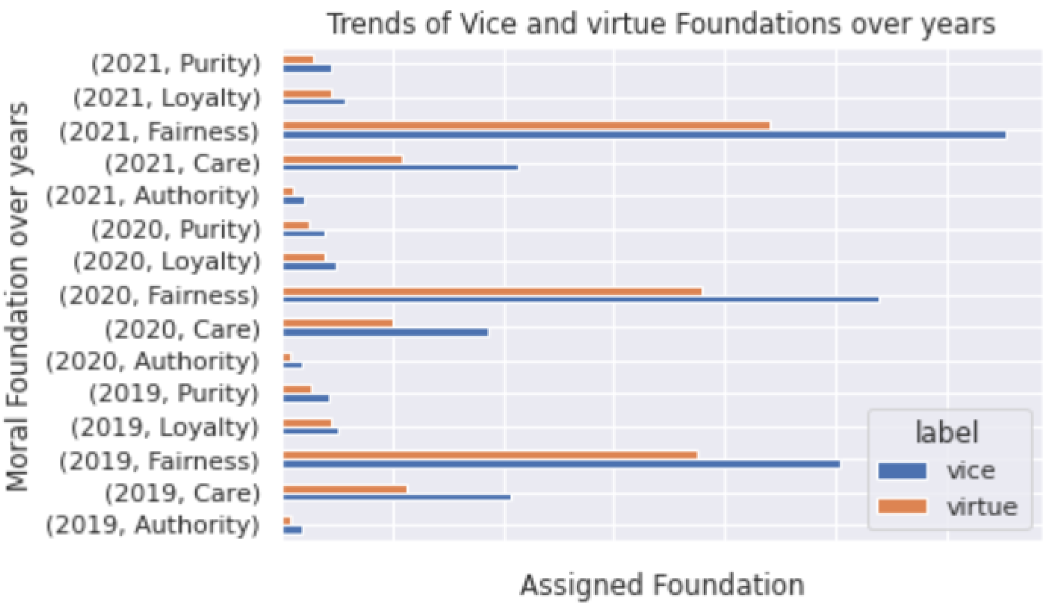

Figure 5.11: Trends of Moral foundation over Vice and Virtue: 'Fairness- Vice' and 'Fairness-Virtue' has the highest frequency distribution in the year 2021

: Dating with end date' has the highest frequency over moral 'Fairness Vice' followed by 'Fairness Virtue'. Followed by 'Topic 2: Short term' with the moral

'Fairness Vice' and 'Topic 1: Dating with end date' with the moral 'Care Vice'. This shows users are more involved in a relationship with 'Vice' morals.

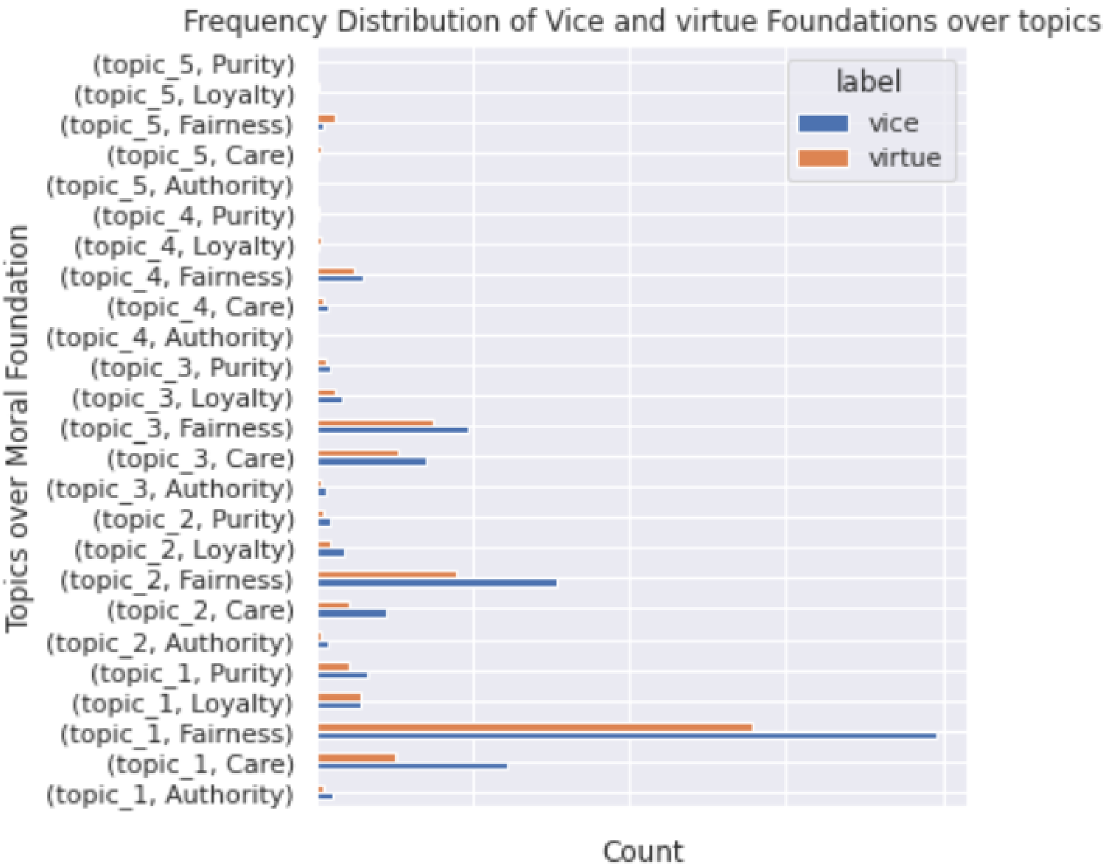

Figure 5.12: Variation of Moral foundation over topics: 'Topic 1 : Dating with end date' has the highest frequency over moral 'Fairness Vice'

In the yearly trend analysis of morals over topics in figure 5.13 'Topic 1 : Dating with end date' has the highest frequency over moral 'Fairness Vice' in the year 2021 followed by the year 2020 and 2019. Followed by 'Topic 2 : Short term' over moral 'Fairness Vice'.

5.2 Analytical Workflow by adding subreddits

To use this model for other subreddits, we must follow some existing preprocessing techniques for the data that have been discussed in the earlier sections. The model is developed for the analysis of a large corpus for the related Reddit. If any other Reddit is used data annotation and model training needs to be performed.

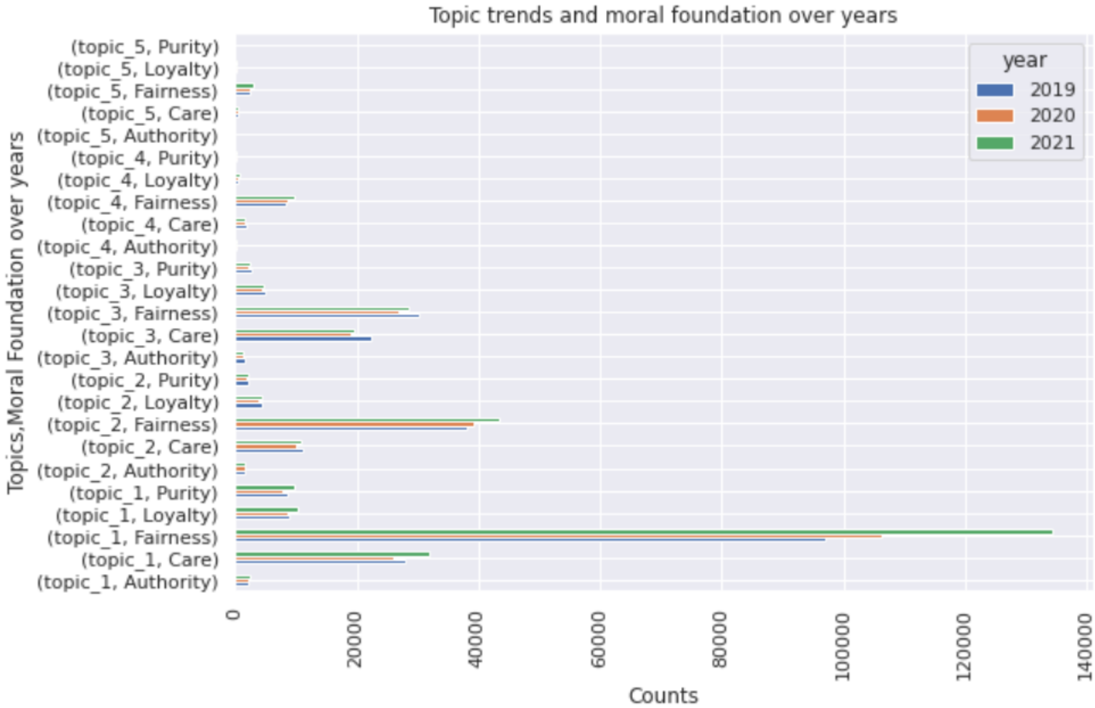

Figure 5.13: Yearly trend of Moral foundation over topics: 'Topic 1 : Dating with end date' has the highest frequency over moral 'Fairness Vice' in the year 2021

5.3 Conclusion

In this chapter, we have done several analyses pertaining to the objectives this work was trying to achieve. The results obtained from these experiments reveal that the model can train and predict trends. The computational trends found are sufficient to justify the objectives. The next chapter concludes the work and gives an idea of future works that can be performed to enhance the analysis.

Chapter 6

Discussion

To evaluate the trends of 'dating with end date', and morals behind of the user behind 'dating with end date'. We have extracted, cleaned and preprocessed 665523 Reddit posts in our dataset.

Our key contributions are as follows:

- We performed the Semi-supervised Corex topic modeling for finding topics using user-specified anchors. This gave us the relevant topics for each post.

- We demonstrated sentiment analysis on our Reddit dataset. This gave the sentiments of each post.

- We demonstrate the concept of moral foundation on our Reddit dataset and performed a comparative trend analysis of each topic, morals over years. These results suggest that the proposed framework produces expected insights from the Reddit posts and trends over years.

Our proposed trend analysis model flow offers a viable framework for analyzing the trends over social media analysis related to user-specified topics. The method is quite robust in terms of handling large dataset.

6.1 Future Work

This work can be progressed in several directions. The dataset can be formed by using any set of subreddits related to the area of research.

In this work, we are using five topics. We plan to add more anchors and form more topics which will give us deep insights from the Reddit dataset.

In this work, we have used a dataset of count 500 to train our model for the moral foundation. We plan to introduce more robustness into the model, to do that we will annotate more data count of 500 per subreddit.

The data extract we used contains the data for the last three years. As we are confirmed that 'dating comes with an end date'. In the analysis 'Topic: dating with end date' is highly discussed in the year 2021. The user has fairness morals while getting into a relationship/dating but vice moral. This shows users are not getting relationships or being in a relationship they want to be fair with their partners but vice manner. Therefore, they choose dating with an end date.

Bibliography

[1] B. Kennedy, A. Ashok kumar, R. Boyd, and M. Dehghani, "Text analysis for psychology: Methods, principles, and practices," in *The Handbook of Language Analysis in Psychology*, M. Dehghani and R. Boyd, Eds. Guilford Press, 2022. [Online]. Available: https://doi.org/10.31234/osf.io/h2b8t

[2] T. M. S. J. R. B. T. J. W. J. Z. Morteza Dehghani, Jack McArdle, *https://dornsife.usc.edu/psyc/computational-psychology.*

[3] P. B, "Computational personality recognition and sentiment analysis of select novels of cormac mccarthy," *The IUP Journal of English Studies, Vol. XV, No. 3,Available at SSRN: https://ssrn.com/abstract=3812925*, 2020.

[4] *Moral foundation: https://moralfoundations.org.*

[5] LeFebvre, "Swiping me off my feet: Explicating relationship initiation on tinder," *Journal of Social and Personal Relationships*, vol. 35(9), 2018.

[6] J. Trager, A. S. Ziabari, A. M. Davani, P. Golazazian, F. Karimi-Malekabadi, A. Omrani, Z. Li, B. Kennedy, N. K. Reimer, M. Reyes, K. Cheng, M. Wei, C. Merrifield, A. Khosravi, E. L. Álvarez, and M. Dehghani, "The moral foundations reddit corpus," *ArXiv*, vol. abs/2208.05545, 2022.

[7] S. G. C. F. Nicholas Proferes, Naiyan Jones and M. Zimmer, "Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics," *Social Media Society*, vol. 7(2):20563051211019004, 2021.

[8] K. D. Arnout B Boot, Erik Tjong Kim Sang and R. A. Zwaan, "How character limit affects language usage in tweets," *Palgrave Communications*, vol. 5(1):1–13, 2019.

[9] G. D. Tiana Gaudette, Ryan Scrivens and R. Frank, "Upvoting extremism: Collective identity formation and the extreme right on reddit," *New Media Society*, vol. 23(12):3491–3508, 2021.

[10] T. J. H. S. C. Brenna Helm, Ryan Scrivens and R. Frank, "Examining incel subculture on reddit," *Journal of Crime and Justice*, p. pages 1–19, 2022.

[11] S. G. C. F. Nicholas Proferes, Naiyan Jones and M. Zimmer., "Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics," *Social Media Society*, 2021.

[12] C. S. H. E. J. . H. A. De Choudhury, M., "Characterizing and predicting postpartum depression from shared facebook data." *In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work Social Computing (pp. 626–638). http://www.munmund.net/pubs/cscw*$_{1}4_{1}$*.pdf*, 2014.

[13] . G. Mittal, A., "Stock prediction using twitter sentiment analysis." *http://cs229.stanford.edu/proj2011/GoelMittalStockMarketPredictionUsingTwitterSentimen* 2010.

[14] . H. M. J. Isaak, J., "User data privacy: Facebook, cambridge analytica, and privacy protection." *Computer, 51(8), 56–59.*, 2018.

[15] J. Vincent, "Transgender youtubers had their videos grabbed to train facial recognition software. the verge." *https://www.theverge.com/2017/8/22/16180080/transgenderyoutubers-ai-facial-recognition-dataset*, 2017, August 22.

[16] M. Zimmer, "Addressing conceptual gaps in big data research ethics: An application of contextual integrity." *Social Media + Society, 4(2), 2056305118768300.*, 2018.

[17] . C. K. Metcalf, J., "Where are human subjects in big data research? the emerging ethics divide." *Big Data Society, 3(1), 2053951716650211*, 2016.

[18] S. K. . A. Z. Vitak, J., "Beyond the belmont principles: Ethical challenges, practices, and beliefs in the online data research community." *In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work Social Computing (pp. 941–953). https://dl.acm.org/doi/abs/10.1145/2818048.2820078*, 2016.

[19] P. N. S. K. . A. Z. Vitak, J., "Ethics regulation in social computing research: Examining the role of institutional review boards." *Journal of Empirical Research on Human Research Ethics, 12(5), 372–382.*, 2017.

[20] . P. N. Fiesler, C., ""participant" perceptions of twitter research ethics." *Social Media + Society, 4(1), 2056305118763366.*, 2018.

[21] K. Beninger, "Social media users' views on the ethics of social media research." *In L. Sloan A. Quan-Haase (Eds.), The Sage handbook of social media research methods (pp. 57–73). SAGE.*, 2017.

[22] S. S. . R. D. Ammari, T., "Self-declared throwaway accounts on reddit: How platform affordances and shared norms enable parenting disclosure and support." *Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1–30*, 2019.

[23] A. Leavitt, ""this is a throwaway account": Temporary technical identities and perceptions of anonymity in a massive online community." *In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work Social Computing (pp. 317–327). https://dl.acm.org/doi/abs/10.1145/2675133.2675175*, 2015.

[24] Reddit.com., "Reddit.com." *User agreement—October 15, 2020—Reddit. Reddit User Agreement. https://www.redditinc.com/policies/ user-agreement-october-15-2020*, 2020b.

[25] S. M. J. S. C. H. B. A. L. C. E. J. . G. E. Chandrasekharan, E., "The internet's hidden rules: An empirical study of reddit norm violations at micro, meso, and macro scales." *Proceedings of the ACM on Human-Computer Interaction, 2, 32*, 2018.

[26] B. I. G. E. . B. A. Jhaver, S., "Human machine collaboration for content regulation: The case of reddit automoderator." *ACM Transactions on Computer Human Interaction, 26(5), 31.*, 2019.

[27] V. J. S. S. B. P. F. T. K. B. B. J. . L. S. Long, K., ""could you define that in bot terms"? requesting, creating and using bots on reddit." *In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 3488–3500). Association for Computing Machinery.*, 2017.

[28] Z. S. K. B. S. M. . B. Baumgartner, J., "The pushshift reddit dataset." *Proceedings of the International AAAI Conference on Web and Social Media, 14, 830–839.*, 2020.

[29] J. D. L. David M. Blei, "Dynamic topic models," *https://mimno.infosci.cornell.edu*, 2003.

[30] D. K. Ryan J. Gallagher, Kyle Reing1 and G. V. Steeg, "Anchored correlation explanation: Topic modeling with minimal domain knowledge," *cs.CL*, 2018.

[31] M. Marietta, "From my cold, dead hands: Democratic consequences of sacred rhetoric." *The Journal of Politics, 70(3):767–779.*, 2008.

[32] J. H. E. S. J. G. N. J. P. S. V. R. I. Morteza Dehghani, Kate Johnson and J. Graham, "Purity homophily in social networks." *Journal of Experimental Psychology: General, 145(3):366.*, 2016.

[33] M. J. C. William J Brady and J. J. V. Bavel, "The mad model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online," *The role of motivation, attention, and design in*

the spread of moralized content online. Perspectives on Psychological Science, 15(4):978–1010., 2016.

[34] C. R. Jan G Voelkel, Mashail Malik and R. Willer, "Changing americans' attitudes about immigration: Using moral framing to bolster factual arguments." *The ANNALS of the American Academy of Political and Social Science, 700(1):73– 85.*, 2022.

[35] C. Wolsko., "Expanding the range of environmental values: Political orientation, moral foundations, and the common ingroup." *Journal of Environmental Psychology, 51:284–294.*, 2017.

[36] A. F. Blair Kidwell and D. M. Hardesty, "Getting liberals and conservatives to go green: Political ideology and congruent appeals." *Journal of Consumer Research, 40(2):350–367.*, 2017.

[37] L. Y. S. H. A. M. D. Y. L. B. K. M. A. Z. K. M. M. Joe Hoover, Gwenyth Portillo-Wightman, "Moral foundations twitter corpus: A collection of 35k tweets annotated for moral sentiment," *Social Psychological and Personality Science*, 2020.

[38] S. K. M. M. R. I. S. P. W. Jesse Graham, Jonathan Haidt and P. H. Ditto, "Moral foundations theory: The pragmatic validity of moral pluralism," *In Advances in experimental social psychology*, 2013.

[39] J. H. R. I. S. K. Jesse Graham, Brian A Nosek and P. H. Ditto., "Mapping the moral domain." *Journal of personality and social psychology, 101(2):366.*, 2011.

[40] S. D. G. A. G. L. M. Luigi Asprino, Luana Bulla and M. Mongiov, "Uncovering values: Detecting latent moral content from natural language with explainable and non-trained methods," *Deep Learning Inside Out (DeeLIO 2022): The 3rd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*, 2022.

[41] E. T. u. u. u. J. J. Ece C¸ igdem Mutlu, Toktam Oghaz and I. Garibay., "Quantifying latent moral foundations in twitter narratives: The case of the syrian white helmets misinformation." *arXiv preprint arXiv:2004.13142.*, 2020.

[42] A. D.-F. Alexander Ruch and R. Batra, "Millions of co-purchases and reviews reveal the spread of polarization and lifestyle politics across online markets." *arXiv preprint arXiv:2201.06556*, 2022.

[43] T. G. Milad Alshomary, Roxanne El Baff and H. Wachsmuth, "The moral debater: A study on the computational generation of morally framed arguments." *arXiv preprint arXiv:2203.14563*, 2022.

[44] Y. M. Hiroaki Yamane and T. Harada, "Humor meets morality: Joke generation based on moral judgement." *Information Processing Management*, 2021.

[45] T. R. O. Stephan A Curiskis, Barry Drake and P. J. Kennedy, "An evaluation of document clustering and topic modelling in two online social networks: Twitter and reddit." *Information Processing Management, 57(2):102034.*, 2020.

[46] N. K. Cristian Candia, Mohammad Atari and B. Uzzi, "Overuse of moral language dampens content engagement on social media." *Under review*, 2022.

[47] R. L. B. Brendan Kennedy, Ashwini Ashokkumar and M. Dehghani, "Text analysis for psychology: Methods, principles, and practices." *Handbook of Language Analysis in Psychology*, 2022.

[48] J. Haidt, "The righteous mind: Why good people are divided by politics and religion," *Vintage.*, 2012.

[49] S. K. M. M. R. I. S. P. W. Jesse Graham, Jonathan Haidt and P. H. Ditto, "Moral foundations theory: The pragmatic validity of moral pluralism." *In*

Advances in experimental social psychology, volume 47, pages 55–130. Elsevier, 2013.

[50] J. G. S. K. S. T. S. Mohammad Atari, Jonathan Haidt and M. Dehghani, “Morality beyond the weird: How the nomological network of morality varies across cultures.” *PsyArXiv*, 2022.

[51] S. G. C. F. Nicholas Proferes, Naiyan Jones and M. Zimmer, “Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics.” *Social Media+ Society, 7(2):20563051211019004*, 2021.

[52] G. D. Tiana Gaudette, Ryan Scrivens and R. Frank, “Upvoting extremism: Collective identity formation and the extreme right on reddit.” *New Media Society, 23(12):3491–3508*, 2021.

[53] T. J. H. S. C. Brenna Helm, Ryan Scrivens and R. Frank, “Examining incel subculture on reddit,” *Journal of Crime and Justice, pages 1–19*, 2022.

[54] A. M. D. Vinodkumar Prabhakaran and M. Diaz., “On releasing annotator-level labels and information in datasets,” *arXiv preprint arXiv:2110.05699*, 2021.

[55] Y. Mejova, “Sentiment analysis: An overview,” *Academia*, 2009.

[56] C. Prince, “Moral foundation theory and the law,” *https://digitalcommons.law.seattleu.edu*, 2010.

[57] B. K. Jackson Trager, “The moral foundations reddit corpus,” *Academia*, 2022.