Identifying Human Needs through Social Media:
A study on Indian cities during COVID-19

Sunny Rai[1], Rohan Joseph[1], Prakruti Singh Thakur[2] and Mohammed Abdul Khaliq[3]

[1]Mahindra University, Hyderabad, India.
[2]Arizona State University, Tempe, USA
[3]University of Stuttgart, Stuttgart, Germany
sunny.rai@mahindrauniversity.edu.in, rohan18545@mechyd.ac.in
psthakur@asu.edu, st181091@stud.uni-stuttgart.de

Abstract

In this paper, we present a minimally-supervised approach to identify human needs expressed in tweets. Taking inspiration from Frustration-Aggression theory, we trained RoBERTa model to classify tweets expressing *frustration* which serves as an indicator of unmet needs. Although the notion of frustration is highly subjective and complex, the findings support the use of pretrained language model in identifying tweets with unmet needs. Our study reveals the major causes behind feeling *frustrated* during the lockdown and the second wave of the COVID-19 pandemic in India. Our proposed approach can be useful in timely identification and prioritization of emerging human needs in the event of a crisis.

1 Introduction

India reported its first case of COVID-19 from Kerala in the month of Jan, 2020 (Andrews et al., 2020). Several control measures including restrictions on international travel, screening of flight passengers, and institutional quarantine were undertaken shortly after to combat the transmission. The Government of India (GoI) imposed a nation-wide lockdown[1] on Mar 25, 2020 as a preventive measure to curb the spread of COVID-19. *Lockdown* is an emergency protocol that restricts non-essential movement of people as well as goods. This lockdown was eventually extended till May 31, 2020, making it one of the longest lockdowns imposed during the pandemic. This resulted in a huge gap in demand and supply of goods (Mahajan and Tomar, 2021), increased stress (Rehman et al., 2021) and mass exodus of migrant workers from cities due to lack of earning opportunities in the informal economy (Das and Kumar, 2020).

Amidst the growing panic, Twitter emerged as the go to platform to express one's feelings and needs such as *travel, food*, *hospital beds, oxygen, cremation* and *funds*[2]. An overwhelming number of tweets seeking support, lack of timely response and inadequate after-care are a few motivating factors behind this study. We particularly study the tweets from metropolitan Indian cities posted during the COVID-19 pandemic. The main contributions are as follows:

- Using topic modeling and minimal supervision, we create a dataset of tweets labelled with needs as described in Maslow's Theory of Motivation (Maslow and Lewis, 1987). This dataset with tagged needs is available for public research[3].

- Taking inspiration from Frustration-Aggression theory (Dollard et al., 1939), we finetuned a state of the art neural language model, RoBERTa (Liu et al., 2019) to detect the unmet *needs*.

The rest of the paper is organized as follows. Section 2 describes the prior work pertinent to the research presented here. We present our approach to gather the needs from Twitter in Section 3. We introduce a RoBERTa based classifier to detect unmet needs in Section 4. We discuss the social impact of the proposed work in Section 5 and list down the limitations in Section 6. We conclude our work in Section 7.

2 Background

Understanding human needs is a widely researched domain by state agencies as well as commercial organizations (Costanza et al., 2007). Prior research has shown that fulfilled needs have a positive impact on a person's feelings of well-being

---

[1] 'Coronavirus in India: 21-day lockdown begins; key highlights of PM Modi's speech', Business Today (Mar 25, 2020). Available at Link

[2]Reuters: https://graphics.reuters.com/HEALTH-CORONAVIRUS/INDIA-TWITTER/oakpekqlrpr/
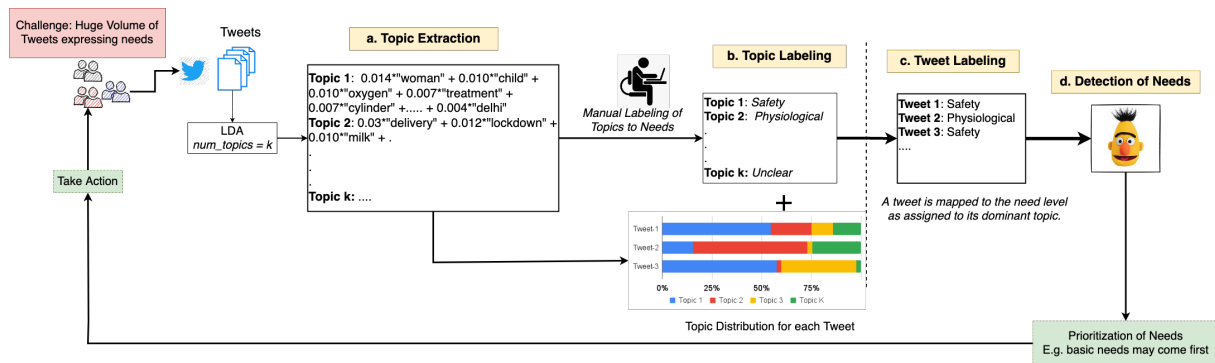[3]https://github.com/AxleBlaze3/Covid_19_Tweets_with_Tagged_Needs

Figure 1: Block diagram for the proposed approach

(Ryff and Keyes, 1995). From stockpiling basic household items during the initial phase of the pandemic to embracing digital technologies such as *zoom*, the market witnessed quite a shift in consumer needs since the outbreak of COVID-19 (Becdach et al., 2020; Mehta et al., 2020). Identifying one's true needs, however, is a challenging task. Yang and Li (2013) took inspiration from Maslow's theory of motivation to predict consumer's needs and purchasing behavior using social media. Ko et al. (2020) used Korean twitter and blogs to discover customer's unmet needs through Hierarchical Concept Search Space algorithm. Their approach aimed to facilitate idea generation for home appliances.

More recently, Yang et al. (2021) advocated the use of Weibo[4] to identify unmet non-COVID-19 healthcare needs. Suh et al. (2021) studied the transition in needs during COVID-19 through the *search queries* on Bing. The product type in search queries were manually marked with the needs as described in Maslow's theory of motivation to automate the task of need identification. Their results affirmed a human tendency to first satisfy basic needs such as *food* and *shelter* before exploring advanced needs such as *creativity* and *love*. Jolly et al. (2020) performed a psychometric analysis of tweets posted in response to official bulletins on COVID-19 by state agencies, revealing the causality between bulletins and the feeling of medical emergency on Twitter.

Prior studies (Saha et al., 2020; Guntuku et al., 2020; Mendoza et al., 2010) have consistently demonstrated the efficacy of social media platforms such as Twitter in capturing the feelings of society at scale. However, this unfurls the challenge of annotating the posts with their expressed needs. In

this paper, we propose an approach to automate labeling of tweets with their expressed needs. We also build a model to detect unmet needs from tweets. To the best of our knowledge, this is the first study on the needs expressed through tweets from Indian cities during the pandemic.

3 Identifying Human *needs* from Tweets

We illustrate the different components of the proposed model in Fig. 1. The block diagram represents the steps of the proposed approach that are, (a) *extracting key topics from Twitter discourse*, (b) *manual mapping of the topics to the expressed human needs*, (c) *mapping tweets to needs assigned to their dominant topics* and (d) detection of unmet needs from tweets. The components in green represent the use case of detecting unmet needs and categorization when given a live stream of tweets.

3.1 Tweets Collection

The GoI officially declared the nationwide lockdown on Mar 25, 2020. The second wave of COVID-19 peaked in the mid of May, 2021. Taking into account the baseline considered by Suh et al. (2021) and the number of COVID-19 cases[5] in India, we set the duration of study from Dec 1, 2019 to Jun 30, 2021, comprising a total of nineteen months. The first three months that is from Dec 1, 2019 to Feb 28, 2020 is the baseline period that serves as an indicator of pre-COVID-19 *needs pattern*. We here assume that a tweet does not need to be marked with hashtags related to COVID-19 to have a *need* affected or emerged due to ongoing COVID-19 pandemic.

Using *snscrape*[6], we extract Indian tweets posted between Dec, 2019 and June, 2021. We set the

---

[4]Weibo.com

[5]WHO, Coronavirus disease 2019 (COVID-19) Situation Report – 39. Feb 28, 2020. LINK

[6]https://pypi.org/project/snscrape/

Figure 2: Areas covered during Tweet Collection

parameter "geocode" of the form [*latitude, longitude, radius*] to the (latitude, longitude) of cities namely *Nagpur, Bangalore, Jaipur, Kolkata* and *Patna* with the radii as *500km, 400km, 350km, 50km* and *100km* respectively in an attempt to encompass representative metropolitan cities situated in different parts of India. The covered region is depicted in Fig. 2.

As a pre-processing step, we removed duplicate and non-English tweets. We also filtered out tweets having less than twelve words. The word limit threshold was decided empirically after analysing the content of tweets. These tweets were either related to marketing/greetings such as *good morning*, *happy birthday* or comments on the original tweets without substantial semantic content of its own. We have a total of $1.4M$ unique tweets for further study.

3.2 Mapping Tweets to Human Needs

In this research, we study only *expressed needs* in tweets. Bradshaw (1972) defined expressed need as "*the felt need turned into action*". To identify the different types of *expressed need*, we take inspiration from Maslow's Hierarchy of Needs (MHoN) (Maslow and Lewis, 1987) which categorizes the human needs into five distinct levels namely *physiological*[L_1], *Safety* [L_2], *Love and Belonging* [L_3], *Love and Belonging* [L_3], *Esteem* [L_4] and *Self-Actualization* [L_5]. *Physiological* and *safety* needs are considered basic needs that need to be satisfied first before one begins to explore the advanced needs related to *esteem* and *self-actualization*.

3.2.1 Topic Extraction

Manual annotation of over $1.4M$ tweets with their expressed need is time consuming as well as prohibitively expensive. We therefore employ topic modeling[7] (Blei et al., 2003) to identify the major topics of discourse for monthwise set of tweets which we then manually label to a level as described in Maslow's Theory of Motivation.The number of topic words is set to 20 and the rest of the parameters were set to default values. The number of topics is decided empirically after analysing the coherence score and execution time for a randomly picked sample of three months. We set the #*topics* to 30 after analysing different number of topics, #*topics* = $\{10, 15, 20, ...45, 50\}$. We thus obtain a set $\mathcal{T}$ having 570 topics (30 topics $*19$ months).

3.2.2 Manual Labeling of Topics

We asked a team of three human annotators to map the extracted topics $t \in \mathcal{T}$ to the levels $\{L_1, L_2, L_3, L_4, L_5\} \in$ MHoN. Each annotator is an undergraduate student, aged 19-21 years and highly proficient in the English language. Two were male and one was female. Given a topic t and few tweets elaborating its context of usage, the task is to map the topic t to either $L_i \in$ MHoN or as '*unclear*'. Annotators were encouraged to choose *unclear* if they find a topic ambiguous. A topic is assigned a *need level* from MHoN only if all annotators choose the same level.

Out of 570 topics, 59 topics were assigned to L_1, 150 topics mapped to L_2, 84 topics to L_3, 66 topics to L_4 and 95 topics to the level L_5. Rest were *unclear*.

The key categories emerged after mapping topics with needs are provided in Table 1. The *physiological* need majorly comprised *food staples and beverages*, *hygiene concerns*, *mobility*. Clearly, the meaning of safety has evolved and included topics such as *housing*, *infection*, *unemployment*, *domestic violence*, *market* and *financial liabilities*. Relationships and concern for loved one's are discussed under *love and belonging*. *Esteem* covers *online learning*, *ideologies*, postponed *examinations*, lack of internet and smart devices. Self-actualization comprises recreational tasks such as DIY, sports, entertainment and skill acquisition.

---

[7]Gensim LDA: https://radimrehurek.com/gensim_3.8.3/models/wrappers/ldamallet.html

Table 1: Mapped topics and Need Level in Maslow's Theory of Motivation

Need	#Tweets	#Topics	Key Topics
Physiological	165114	59	**staples** such as *food, beverages, apparel, household products*, **Hygiene** such as *toilet paper*, **basic daily services** such as *grocery delivery, milk, bread*, **rest**, **medicine**, **transport**
Safety	367774	150	**Housing** such as rental/mortgages, evictions, **COVID-19 Safety** such as *masks*, quarantine or *sanitizers*, **Domestic violence**, justice, **Financial liabilities** such as tax, loans, or bankruptcy, stock market, business **Job posts/application & unemployment**
Love & Belonging	192843	84	Expression of or resources for **mental health** or **emotional issues** such as *anxiety, depression, loneliness, isolation, suicide, nervousness, rejection, fear* or *sadness*; **Social media**, Search for **relationships** with significant others, dating, issues such as *divorce* or *breakup*
Esteem	151873	66	**Education** or learning materials, University/Schools; **Online classroom** learning, Examinations; Educational degrees or programs; Knowledge/Skill, zoom meetings, **Ideologies/religions**
Self-Actualization	258287	95	**Recreational task**s such as *self-care, home decor, music* etc., parenting, wedding, **Talent/Skill acquisition**, **Life goals**, Charity/Donation, **volunteering**, **Entertainment** such as *Netflix, Prime, TV shows, movie, sports (IPL), News*

3.2.3 Mapping Tweets to MHoN

At this step, we have a list of topics mapped with the need levels as described in MHoN. We also have the probability distribution of topics for each tweet. The dominant topic of a tweet is the topic with the highest probability. A tweet t is thus marked to a need level $L_i \in$ MHoN on the basis of the mapping assigned to its dominant topic. We illustrate this process in Fig. 1 where the topic distribution along with mapped topics are used to infer the expressed need in tweets. If the tweet has multiple topics with same probability, we only assign a need level if all dominant topics are marked to the same need level else it is marked as *unclear*.

3.3 Analysis

After excluding *unclear* tweets, we have over $1.1M$ tweets marked as expressing a need. Below are few examples[8] that were mapped to their relevant level and those that were *unclear*:

"*Nobody staying at hotels, So why not convert them into covid centers* " –Physiological

"*When I see some people attacking doctors, i get scared about the corona situation*" –Safety

"*Not everyone can work from home. Feeling kinda unsafe or its just fear of getting sucked up in situation and putting life of my family friends in danger.*" –Love and Belonging

"*As a teacher I thank pm for cancelling the Std. 12th board examination.*" –Esteem

"*xxx movie are a big hit, an average human would have to watch his movies multiple times to*

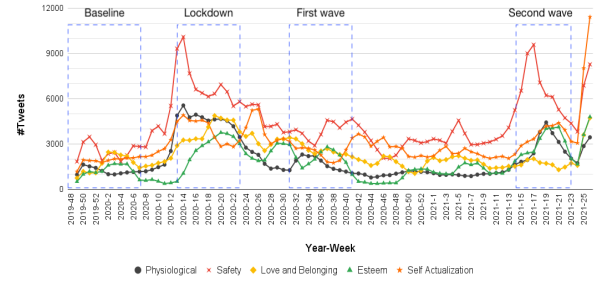[8]Tweets are rephrased to protect user's privacy however, the message remains the same.

Figure 3: Volume of tweets expressing needs from Dec, 2019 to Jun, 2021.

understand." –Self-Actualization

"*Met her while traveling she was selling fruits adding on to the income of her parents, with her impressive salesman skills* ." –Unclear

It may be noted that we have not considered tweets on political topics such as Citizen Amendment Act (Wikipedia contributors, 2021b), National Register of Citizens (Wikipedia contributors, 2021c) and Farm laws (Wikipedia contributors, 2021a) which were part of public discourse during the time of study.

We illustrate the time-wise distribution of tweets tagged with different need levels in Fig. 3. We have considered two weeks moving average to nullify noisy fluctuations in the data. For the baseline, we consider the tweets posted in the first twelve weeks that is, between week-48'2019 to week-7'2020, to understand the pre-COVID pattern of needs. *Lockdown* phase is the period between week-13'2020 to week-23'2020. The *first wave* ranges from week-31'2020 to week-41'2021.

The *second wave* started from week-14'2021 and ended in week-23'2021. The phases namely *baseline, lockdown*, *first wave* and *second wave* are annotated with boxes in Fig. 3. The first peak is placed in the lockdown period and the second peak occurred during the second wave of the pandemic. The volume of needs were slightly higher than pre-COVID levels during the first wave. There is also a huge surge in self-actualization and safety needs starting *week-24'2021*.

Indian Twitter users voiced the *safety* need most often followed by *physiological* need during the lockdown. Both needs peak at the same time. A total of 45% more tweets expressing basic needs were posted during lockdown compared to the second wave of the pandemic. Over twice the number of physiological tweets was expressed during the lockdown when compared to the second wave. The relatively advanced needs namely *love and Belonging* and *esteem* display a delay during the lockdown and peak almost 3-4 weeks after the basic needs. Soon after the lockdown was lifted, the needs started to return to pre-COVID pattern of needs.

During the second phase of the pandemic, *safety* turned out to be the foremost concern and *physiological* needs peaked only after a delay of two weeks. There is no clear precedence for physiological needs over advanced needs during the second wave. Moreover, *love and belonging* needs stayed at pre-COVID levels during the second wave unlike lockdown phase where concern for loved ones was expressed in large volumes.

Safety has indeed emerged as the dominant concern in the both phases of the pandemic. Lockdown was a special scenario where essential commodities were in shortage due to lack of production as well as black marketing [9]. It is thus not conclusive from our data if physiological needs always take precedence over the advanced needs in the event of a crisis in today's world.

The most advanced need, *self-actualization* surged and ebbed through out the months of our study without any clear correlation with the different phases of pandemic. The huge surge in self-actualization and safety needs starting *week-24'2021* is due to large volume of tweets discussing Indian Premier League 2021 (Wikipedia contributors, 2022) and mass gatherings.

---

[9]The Hindu "Coronavirus lockdown: Invoke Essential Commodities Act to curb black marketing, Home Secretary tells States" (Apr 8, 2020). Available at Link

4 Detecting unmet Needs

Unmet needs are widely characterized by *frustration* (Dollard et al., 1939; Killgore et al., 2021). Through Frustration-Aggression theory, Dollard et al. (1939) defined *frustration* as an impediment or blockage in achieving one's needs or goals. An impediment to a goal is considered frustration if and only if the person is actively striving to reach this goal. We thus hypothesize that an unmet need can be detected by identifying whether a tweet with expressed need has *frustration* or not.

4.1 Approach

Our task is to classify whether a given tweet tagged with need is expressing *frustration* or not. We fine tuned the RoBERTa pretrained model (Wolf et al., 2020) with a learning rate of $2e^{-5}$ and dropout of 0.3 for this classification task. For training, we collected tweets containing the hashtag *#frustrated*. For negative class that is, *Not frustrated*, we extracted tweets with hashtags that symbolise satisfaction (ex: *#satisfied*, *#FeelingContent*). This dataset has a total of 13970 tweets with equal number of instances for positive and negative class. We provide a representative tweet from each class below:

" *HOW fast does one have to be to book a slot on COWIN? I saw slots available at a hospital; I selected the time slot; entered the CAPTCHA in not more than 15 seconds... and still it didn't book the slot. And then when I refreshed, all the slots were gone*" - `Frustrated`

"😊😊😊🙏*I did it! ... I officially completed my undergraduate program and received my bachelors degree. may the glory be to God for blessing me with the gifts to achieve this great milestone*" - `Not Frustrated`

As a preprocessing step, we remove hashtags and mentions from the tweet text. We consider 80% of tweets for training and the rest 20% is equally divided for validation and test set. We achieved an accuracy of 93.4% on the validation set. We obtained an accuracy of 92.2% with a precision of 91% and recall of 93% on the test set.

4.2 Performance Evaluation

Out of $1.1M$ tweets, our model predicted a total of 792533 tweets as *frustrated*. 77.36% of physiological needs and 77.5% of safety needs expressed frustration. Under advanced needs, 54.13% of love and

Figure 4: Tweets predicted as *Frustrated*

Figure 5: Percentage of Frustrated Tweets

belonging, 70.43% of esteem needs and 62.91% of self-actualization needs were marked frustrated.

To evaluate the quality of predictions, we randomly sampled 100 tweets which were annotated as *frustrated* or *not frustrated* by three undergrad students proficient in English. The majority vote was considered as the final label. A total of 45 samples were labelled as frustrated out of 100. The inter-annotator agreement (fleiss kappa) obtained for this task was 0.638 indicating its subjective nature. The trained model achieved an accuracy of 76% with a weighted precision of 78% and weighted recall of 76% on this set of annotated tweets. Below are two example tweets which were classified as *frustrated*:

> "*Please complete the pending projects in Telangana State. Sir please do the needful. There is no direct train from Karimnagar to Hyderabad*"

> "*Need of hour Free Education Free/Affordable health care No freebies , let people work*".

We observe that the above tweets clearly express *frustration* as described in Dollard et al. (1939). Another point worth noting is the subjectivity when labelling frustration. Consider the below tweets predicted as *frustrated* but annotated as *not frustrated* by human annotators.

> "*She lost her life in line of duty. She had been performing her duty in adverse circumstances amid lockdown.She should be declared "Corona Warrior"and all benefits and compensation should be given to her family by the govt.*"

> "*Finally, I am buying an Iphone , twelfth edition but next year. As i also thought about Iphone last year.*"

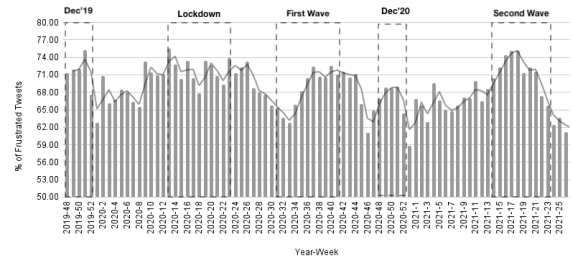Whether the above tweets express frustration or not, is quite debatable. Therefore, the performance metrics need to be interpreted accordingly.

4.2.1 Decoding *frustration* through RoBERTa

On a random sample of 30 tweets predicted as *frustrated*, we used integrated gradients method (Sundararajan et al., 2017) to identify the type of input features that attribute to the prediction to the class *frustrated*.

We provide few example tweets from this set in Fig. 4. Here, the shade of red signifies the importance of input features in prediction. The greater the significance, the deeper the hue of red. For instance, the words highlighted with deeper red such as *shortage*, *oxygen*, *where*, *loose*, *ridicule*, and *all* led to the classification into frustrated class for the first tweet in Fig. 4. Likewise for other tweets, the words namely *have*, *transport*, *electricity*, *delay*, *infected*, *ventilators*, *expensive*, *treatment* are input features that derived the prediction to the class *frustrated*. Since these terms intrinsically reflect constraints or impediments in leading a purposeful life, we may conclude that the model correctly learned to detect tweets expressing frustration.

4.3 Discussion

We illustrate the week wise percentage of tweets predicted as *frustrated* in Fig. 5. At first glance, Twitter appears to be a land of frustration with dissatisfaction rate of around 62% even before COVID-19. The jump in frustration rate in the fourth week of December'19 is due to the mulling over the passing year and eventually settled down in the next two months of Jan, 2020 and Feb, 2020.

The percentage of frustrated tweets hovered between 71 − 74% during the *lockdown*, the *first wave* and the *second wave* of COVID-19. Clearly, there is an increment of over 4% in frustration rate when compared to non-stressful phases of the pandemic.

We illustrate the week wise transition for the volume of frustrated tweets expressing basic and advanced needs in Fig. 6. There is a huge jump
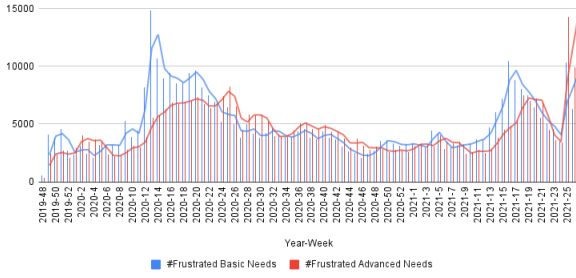
59

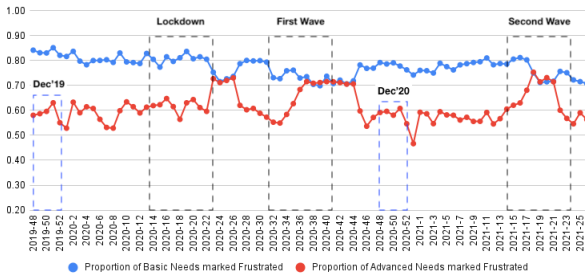Figure 6: #Frustrated Tweets for *Basic* & *Advanced* Needs

Figure 7: Proportion of Basic and Advanced Needs marked Frustrated

in the volume of both categories of tweets. More tweets expressing frustration due to basic needs were posted during the lockdown in comparison to the second wave. The volume of basic tweets during the first wave remained slightly above the pre-COVID level.

The proportion of frustrated tweets across *basic* and *advanced* level of needs is illustrated in Fig. 7. We observe that almost 80% of tweets expressing basic needs are unmet irrespective of the time of the year. Despite the fact that a large number of basic needs were posted throughout the lockdown, the dissatisfaction rate remained constant. It is thus safe to assume that users discuss basic needs only when these needs are unfulfilled. The general rate of frustration for advanced needs is 60%. We also note that as soon as the frustration due to basic needs reduces, the frustration due to advanced needs increased by over 10%. There are three such peaks in Fig. 7. This does support the belief that once the basic needs are secured, one quickly moves to the advanced needs. On analysis, education with key terms such as *board exams, national level entrance exams, graduation degree* was found to be the dominant concern across each peak. Another common concern was consumer-centric services with worries revolving around delayed *refunds*, cancelled travel plans, delayed delivery of
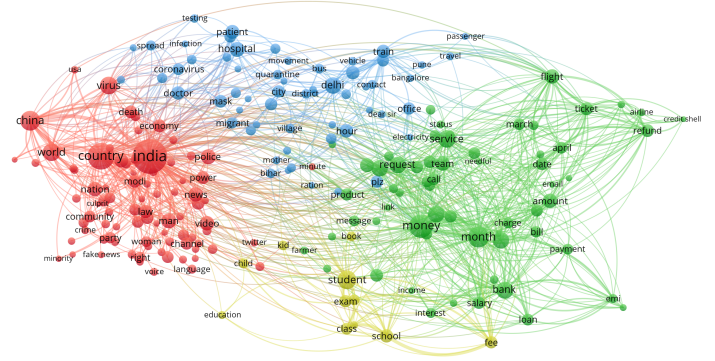

Figure 8: Themes for *frustrated* tweets: Lockdown

online orders etc.

Moreover, time specific events such as *call against products made in China*, Bollywood scandals, football, entertainment were found during the second peak (week-36'2020 - week-44'2020). In contrast, the third peak (week-18'2021 - week-22'2021 discussed the lack of availability of vaccines and further called for inclusivity and transparency in distribution.

4.3.1 Key Themes behind *frustration*

To discover the themes behind the increased volume of frustrated tweets during lockdown and the second wave of COVID-19, we used a computer program called VOSviewer (van Eck and Waltman, 2011) to create a term co-occurrence map for the tweets labelled as *frustrated*. Fig. 8 and Fig. 9 illustrate the oft-discussed terms in frustrated tweets posted during the lockdown and the second wave respectively.

Lockdown: Travel concerns due to the imposed nationwide lockdown are evident from the terms in cluster **blue** in Fig. 8. Major Indian cities namely *bengaluru, bihar, pune* coupled with transportation choices such as *bus, train, vehicle* can be seen. We also note terms such as *quarantine, doctor, patient, office* in the same cluster indicating the traveling problems faced during daily life activities. The nodes in **green** reveal the challenges faced by logistics and travel industry. Terms such as *refund, ticket, airline, flight, credit* reflect the chief complaints by customers along with *bill* and other *payments*.

The nodes in cluster **red** highlight the discussion on digital media and news channels. Growing concern due to increasing toll of infections in the *USA* and a sense of anger towards *China* were expressed through tweets. *Fake news*, *channel*, *minority* and *economy* were also a few topics of online discussion. The nodes in cluster **yellow** depict the
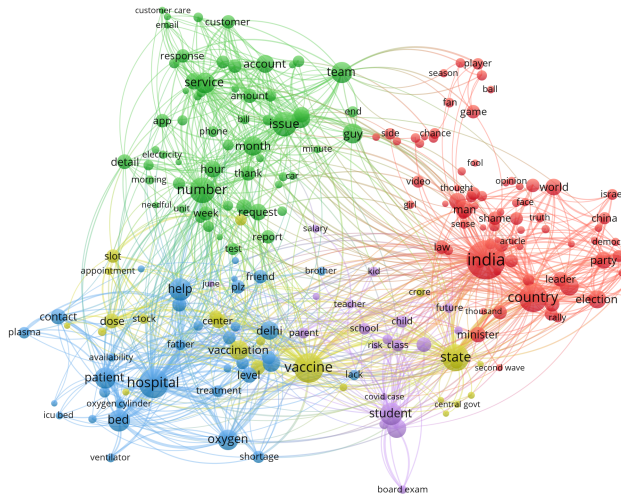
60

Figure 9: Themes for *frustrated* tweets: Second wave

concerns revolving around closed educational institutions, payment of *fees*, *online classes* and *exams*.

Second Wave: The usual customer care complaints are depicted in cluster **green**. The nodes in cluster **red** particularly reveal the frustration against political parties and elections. There are also terms such as *player, season, ball, game* due to upcoming IPL cricket matches. The anxiety due to shortage of *ventilator, patient, hospital, icu bed* and *oxygen cylinder* is captured through nodes in cluster **blue**. Words such as *refer, friend, help* reveal anxious attempts to locate healthcare through contacts on Twitter. Availability of *vaccine* and booking of *slots* were also a cause of frustration amongst Indians. Education remained a concern during the second wave as evident from nodes marked in **purple**.

5 Social Impact

Tsao et al. (2021) highlighted the paucity of action driven research on the COVID-19 data. Early detection of human needs will enable public agencies and independent organizations to provide prompt support including *food supplies*, *medical care*, *transport* and timely awareness about the crisis amongst masses. Our approach can facilitate timely identification and prioritization of emerging human needs in the event of a crisis. When coupled with geo-location tag, the proposed approach can be customized to retrieve closest support available. *Unmet needs scoping* can help in designing public policies to cater to emerging needs of a society. During the COVID-19 pandemic in India, people expressed distinct needs at different stages of each

wave. Public needs on social media can thus serve as an immediate feedback mechanism for public agencies to improvise their relief efforts and policies. Our model to detect unmet needs leverages a pre-trained neural language model that generalises well and is capable of transfer learning from previously labelled data at the start of a crisis. It is thus easy to extend our approach for other languages using publicly available pre-trained multilingual language models.

6 Limitations

Due to our focus on understanding the pattern of needs emerged in India during the COVID-19 pandemic, we performed rigorous filtering to retain only those tweets geo-tagged with locations within India. This significantly reduced the quantity of tweets gathered for our study. Human needs are innately complex and ever evolving concept. As we transition from basic to advanced needs, the needs become more obscure and implicit. To optimize the time and effort for human annotation, we assumed that the dominant topic of a tweet would reflect its need type as discussed in Section 3. This had an impact on the quality of tweet-need mapping and resulted in incorrect labeling in some cases.

7 Conclusion

In this paper, we examined the human needs expressed in Indian cities during the COVID-19 pandemic. We described a minimally supervised approach to annotate tweets with their need level as in Maslow's Hierarchy of Needs. This greatly reduced the time and human effort without much impact on the quality of annotation. We observed a recurring pattern in the needs, indicating predictability in the emerging needs in the event of a crisis. The results support the use of pretrained language model for the task of unmet needs detection. In future, we will extend the proposed model to detect needs in regional languages. We will further work upon incorporating theories better suited to capture advanced psychological needs.

References

MA Andrews, Binu Areekal, KR Rajesh, Jijith Krishnan, R Suryakala, Biju Krishnan, CP Muraly, and PV Santhosh. 2020. First confirmed case of COVID-19 infection in India: A case report. *The Indian Journal of Medical Research* 151, 5 (2020), 490.

Camilo Becdach, Brandon Brown, Ford Halbardier, Brian Henstorf, and Ryan Murphy. 2020. Rapidly forecasting demand and adapting commercial plans in a pandemic. *URL https://www. mckinsey. com/industries/consumer-packaged-goods/our-insights/rapidly-forecasting-demand-and-adapting-commercial-plans-in-a-pandemic* (2020).

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research* 3 (2003), 993–1022.

Jonathan Bradshaw. 1972. The concept of social need. *New society* 30 (1972).

Robert Costanza, Brendan Fisher, Saleem Ali, Caroline Beer, Lynne Bond, Roelof Boumans, Nicholas L Danigelis, Jennifer Dickinson, Carolyn Elliott, Joshua Farley, et al. 2007. Quality of life: An approach integrating opportunities, human needs, and subjective well-being. *Ecological economics* 61, 2-3 (2007), 267–276.

Ritanjan Das and Nilotpal Kumar. 2020. Chronic crisis: migrant workers and India's COVID-19 lockdown. *South Asia@ LSE* (2020).

John Dollard, Neal E Miller, Leonard W Doob, Orval Hobart Mowrer, and Robert R Sears. 1939. Frustration and aggression. (1939).

Sharath Chandra Guntuku, Garrick Sherman, Daniel C Stokes, Anish K Agarwal, Emily Seltzer, Raina M Merchant, and Lyle H Ungar. 2020. Tracking mental health and symptom mentions on Twitter during COVID-19. *Journal of general internal medicine* 35, 9 (2020), 2798–2800.

Baani Leen Kaur Jolly, Palash Aggrawal, Amogh Gulati, Amarjit Singh Sethi, Ponnurangam Kumaraguru, and Tavpritesh Sethi. 2020. Psychometric Analysis and Coupling of Emotions Between State Bulletins and Twitter in India during COVID-19 Infodemic. *arXiv preprint arXiv:2005.05513* (2020).

William DS Killgore, Sara A Cloonan, Emily C Taylor, Ian Anlap, and Natalie S Dailey. 2021. Increasing aggression during the COVID-19 lockdowns. *Journal of affective disorders reports* 5 (2021), 100163.

Taehoon Ko, Ilsun Rhiu, Myung Hwan Yun, and Sungzoon Cho. 2020. A Novel Framework for Identifying Customers' Unmet Needs on Online Social Media Using Context Tree. *Applied Sciences* 10, 23 (2020), 8473.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692* (2019).

Kanika Mahajan and Shekhar Tomar. 2021. COVID-19 and Supply Chain Disruption: Evidence from Food Markets in India. *American journal of agricultural economics* 103, 1 (2021), 35–52.

Abraham Maslow and KJ Lewis. 1987. Maslow's hierarchy of needs. *Salenger Incorporated* 14 (1987), 987.

Seema Mehta, Tanjul Saxena, and Neetu Purohit. 2020. The New Consumer Behaviour Paradigm amid COVID-19: Permanent or Transient? *Journal of Health Management* 22, 2 (2020), 291–301.

Marcelo Mendoza, Barbara Poblete, and Carlos Castillo. 2010. Twitter under crisis: Can we trust what we RT?. In *Proceedings of the first workshop on social media analytics*. 71–79.

Usama Rehman, Mohammad G Shahnawaz, Neda H Khan, Korsi D Kharshiing, Masrat Khursheed, Kaveri Gupta, Drishti Kashyap, and Ritika Uniyal. 2021. Depression, anxiety and stress among Indians in times of Covid-19 lockdown. *Community mental health journal* 57, 1 (2021), 42–48.

Carol D Ryff and Corey Lee M Keyes. 1995. The structure of psychological well-being revisited. *Journal of personality and social psychology* 69, 4 (1995), 719.

Koustuv Saha, John Torous, Eric D Caine, Munmun De Choudhury, et al. 2020. Psychosocial effects of the COVID-19 pandemic: large-scale quasi-experimental study on social media. *Journal of medical internet research* 22, 11 (2020), e22600.

Jina Suh, Eric Horvitz, Ryen W White, and Tim Althoff. 2021. Population-scale study of human needs during the covid-19 pandemic: Analysis and implications. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. 4–12.

Mukund Sundararajan, Ankur Taly, and Qiqi Yan. 2017. Axiomatic attribution for deep networks. In *International Conference on Machine Learning*. PMLR, 3319–3328.

Shu-Feng Tsao, Helen Chen, Therese Tisseverasinghe, Yang Yang, Lianghua Li, and Zahid A Butt. 2021. What social media told us in the time of COVID-19: a scoping review. *The Lancet Digital Health* 3, 3 (2021), e175–e194.

Nees Jan van Eck and Ludo Waltman. 2011. Text mining and visualization using VOSviewer. arXiv:1109.2058 [cs.DL]

Wikipedia contributors. 2021a. 2020 Indian agriculture acts — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=2020_Indian_agriculture_acts&oldid=1025817251. [Online; accessed 9-July-2021].

Wikipedia contributors. 2021b. Citizenship (Amendment) Act, 2019 — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Citizenship_(Amendment)_Act,_2019&oldid=1032618250 [Online; accessed 9-July-2021].

Wikipedia contributors. 2021c. National Register of Citizens — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=National_Register_of_Citizens&oldid=1028089793. [Online; accessed 9-July-2021].

Wikipedia contributors. 2022. 2021 Indian Premier League — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=2021_Indian_Premier_League&oldid=1079776521. [Online; accessed 1-April-2022].

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2020. HuggingFace's Transformers: State-of-the-art Natural Language Processing. arXiv:1910.03771 [cs.CL]

Huahai Yang and Yunyao Li. 2013. Identifying user needs from social media. *IBM Research Division, San Jose* (2013), 11.

Wei-Fa Yang, Danping Zheng, Reynold CK Cheng, Jingya Jane Pu, and Yu-Xiong Su. 2021. Identifying unmet non-COVID-19 health needs during the COVID-19 outbreak based on social media data: a proof-of-concept study in Wuhan city. *Annals of Translational Medicine* 9, 18 (2021).