1-2021

Analyzing tweets on new norm: Work from home during COVID-19 outbreak

Swapna GOTTIPATI
Singapore Management University, SWAPNAG@smu.edu.sg

Kyong Jin SHIM
Singapore Management University, kjshim@smu.edu.sg

Hui Hian TEO
Singapore Management University, huihian.teo.2019@mitb.smu.edu.sg

Karthik NITYANAND
Singapore Management University, karthikn.2019@mitb.smu.edu.sg

Shreyansh SHIVAM
Singapore Management University, shreyanshs.2019@mitb.smu.edu.sg

Citation
GOTTIPATI, Swapna; SHIM, Kyong Jin; TEO, Hui Hian; NITYANAND, Karthik; and SHIVAM, Shreyansh. Analyzing tweets on new norm: Work from home during COVID-19 outbreak. (2021). *2021 11th IEEE Annual Computing and Communication Workshop and Conference: January 27-30, Virtual: Proceedings*. 500-507.
Available at: https://ink.library.smu.edu.sg/sis_research/6027

Analyzing Tweets on New Norm: Work from Home during COVID-19 Outbreak

Swapna Gottipati, Kyong Jin Shim, Hui Hian Teo, Karthik Nityanand, Shreyansh Shivam
School of Information Systems, Singapore Management University, Singapore
Swapnag,kjshim,huihian.teo.2019,karthikn.2019,shreyanshs.2019@smu.edu.sg

***Abstract*—The COVID-19 pandemic triggered a large-scale work-from-home trend globally in recent months. In this paper, we study the phenomenon of "work-from-home" (WFH) by performing social listening. We propose an analytics pipeline designed to crawl social media data and perform text mining analyzes on textual data from tweets scrapped based on hashtags related to WFH in COVID-19 situation. We apply text mining and NLP techniques to analyze the tweets for extracting the WFH themes and sentiments (positive and negative). Our Twitter theme analysis adds further value by summarizing the common key topics, allowing employers to gain more insights on areas of employee concerns due to pandemic.**

Keywords—data analytics, COVID-19, social media, work-from-home, NLP

I. Introduction

When the COVID-19 outbreak first occurred in the city of Wuhan, the Chinese government quickly put the city into lockdown. Since then, with the outbreak spreading globally and difficult to manage, more than 100 countries worldwide had instituted a partial or full lockdown by the end of March 2020, affecting billions of people and many have started to work from home [1].

However, not all employees may be comfortable or prepared to WFH. Also, not all countries are equipped with the necessary infrastructure to facilitate their workforce in working from home. With mandatory or increased working from home arrangements becoming a 'new normal', it would be of interest for employers and policymakers to understand employees' concerns and sentiments. While calls for feedback and surveys may be done, some may be uncomfortable sharing their opinions without anonymity [2]. Hence, online social platforms are alternative sources for analysis.

In this project, we collected around 78K tweets for one month related to WFH during the lockdown time using a predefined list of twelve hashtags to find the relevant tweets. The word cloud in Fig 1. for the hashtags on the data, (excluding the search hashtags), provides us with a better understanding of the key topics being discussed. We observe that the prominent phrases such as lockdown, home_office, remote_working, new_normal, etc., represent general thoughts of the employees.

Besides the overall insights, Fig. 1 also provides some granular insights e.g. "Zoom", "ergonomic" which show the digital tools and physical support that employees require as part of the working from home arrangement. However, these general results are insufficient to analyze the pulse of the employees. Therefore, we study in-depth issues, topics of concerns, sentiments, and overall statistics on the voice of employees for making better decisions on the policy matters by both companies and Government.

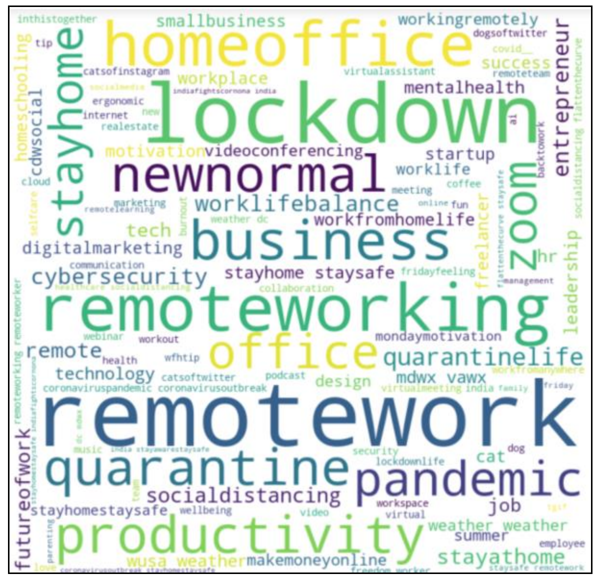

Fig. 1. Word cloud of hashtags from our tweets dataset

Through this project, using NLP and text mining techniques, we aim to discover the topics of interest and sentiments raised by people through an online social platform, Twitter, regarding the new norm of working from home. The rest of the paper is organized as follows. Section II presents the methodology in terms of data and project tasks. In section III, we describe our solution design along with the justification of choices of the techniques at each stage. Section IV describes the experiments at each stage, and section V we present analysis of results. In section VI, we describe the insights and recommendations with examples. We present limitations in Section VII, and we conclude in Section VIII.

II. Methodology

A. Dataset

Tweepy Package [3] in Python is used for data crawling by leveraging on the search API, for four weeks (24 May 2020 to 20 June 2020). English tweets comprising popular hashtags related to the WFH theme are focused in the data collection process. As our focus is only on unique user-generated content, retweets which may result in duplicated content, are excluded from datasets. An average of 2,800 tweets are collected daily, and by 20 June 2020, 78,606 tweets are obtained.

B. *Analytics Tasks*

To discover insights on employee's voices on WFH, three main text analytics tasks are performed; advertisement classification, theme analysis, and sentiment analysis. In our preliminary analysis of a sample of tweets, we observed that the tweets dataset contains both advertisements and non-advertisements content. As advertisements are irrelevant to our project's objective, we filter them first before analyzing the content using the classification approach [4].

Topic modelling task aims to identify topics of interest in tweets [5, 27]. In our project, we need to study the changes in topic distribution over time from the output of topic modelling [6]. We apply topic modelling on the non-advertisements tweets to glean the underlying issues faced or felt by the employees across the four-week period during WFH situation.

Complementing topic modelling is sentiment analysis that has been widely used for instance to understand the public's sentiments on political issues [7, 27]. In our preliminary analysis, we observed that the employees not only present the positive and negative feelings in the tweet but also state factual information. Hence, it is important to identify the objectivity and subjectivity in the tweets. In other words, we study positive, negative, and neutral tweets. Combined with topics, sentiment polarity enables us to study the subjectivity across topics over four weeks.

III. SOLUTION DESIGN

The overview of the solution approach to analyze the tweets is depicted in Fig. 2. We first process the tweets to remove textual noise. The tweets with advertisements are removed in the second stage. In the final steps, theme analysis and sentiment modelling analysis are applied on non-advertisement tweets to identify the underlying themes together with the sentiment polarity.

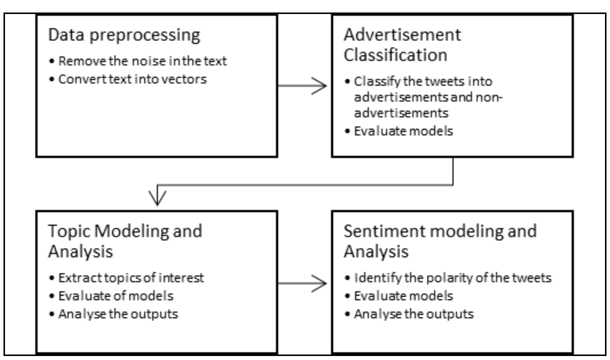

Fig. 2. Solution design overview

A. *Data Pre-Processing*

First, tweet-specific attributes (e.g. @user mentions, URLs, #hashtags), digits, and punctuations are removed using the Preprocessor package and Regex. The text is tokenized and converted to lower case. At this point, the pre-processed text is used for sentiment analysis as removal of stop words (e.g. not in "not happy") could affect the sentiment analysis. For classification and topic modelling, we proceed to remove stop words that are listed in NLTK package [9] and user-defined from the manual analysis. This is followed by lemmatization to generate high quality topics from LDA models. Lastly, the text is converted into vectors which will be the inputs for topic models and classification.

B. *Classification of Advertisements and Non-Advertisements*

Recall that the tweets contain WFH advertisements-related content such as advertisements for jobs, courses, etc. Fig. 3 describes the classification steps to achieve the goal of filtering the noisy tweets.

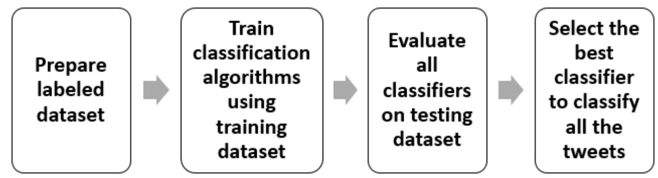

Fig. 3. Advertisement classification steps

For classifier evaluations, 750 tweets are manually labelled by human judges. Finally, we apply the best model on the remaining 77,856 tweets for finding tweets that are non-advertisements. Six algorithms as shown in Table 1 are evaluated on the training dataset (46% are non-adverts and 54% are adverts). [10, 11].

Table 1. Advertisement classification algorithms

Naïve Bayes Models	1. SKLearn Multinomial Naïve Bayes Algorithm 2. SKLearn Bernoulli Naïve Bayes Algorithm
Linear Models	1. SKLearn Linear Logistic Regression Algorithm 2. SKLearn Linear Stochastic Gradient Descent
Support Vector Machines	1. SKLearn SVC Algorithm 2. SKLearn Linear SVC Algorithm

The tweets with "non-adverts" label are subsequently filtered from the datasets before we apply topic modelling and sentiment analysis algorithms.

C. *Theme Analysis*

The objective of topic modelling is to discover the abstract topics that occur in a collection of documents using algorithms such as Latent Dirichlet Allocation (LDA) in an unsupervised manner [12, 13, 26]. Recall that, topic analysis aims to identify the issues or topics raised by individuals regarding WFM. As there could be shifts in the topics over the four-week period as countries start to relax their lockdown measures, the topic modelling is conducted monthly and weekly basis to observe differences across the weeks as well overall statistics of topics in a month. The overall approach is summarized in Fig. 4.

We train two topic modelling algorithms (LDA Gensim [24], and LDA Mallet [13]) on the non-advert dataset (23,267 tweets) and compared their performance using the coherence scores for a varying number of topics (k).

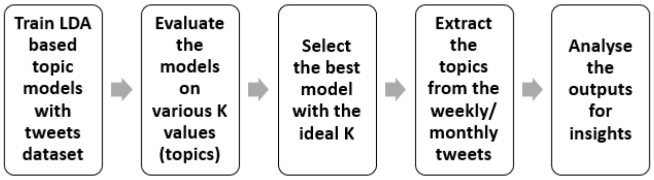

Fig. 4. Topic modelling steps

The ideal number of topics in the month and weeks is decided separately using both quantitative (e.g. using coherence score) and qualitative assessment (e.g. topic-word distribution). After the best model is selected, the weekly topic analysis is generated by adjusting the number of topics K using the coherence scores.

D. Sentiment Analysis

Sentiment analysis is a process where information is analyzed using natural language processing (NLP) and the goal is to discover the polarity of a document; negative, positive, or neutral sentiment [14, 25]. Using sentiment analysis, we try to understand the overall monthly sentiments across topics(i.e. the entire 4 weeks) and how the sentiments could have changed each week, weekly topic-sentiment analysis. Fig. 5 provides an overview of the sentiment analysis steps.

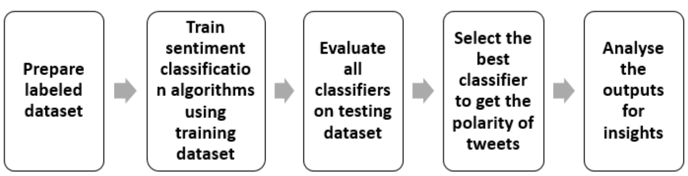

Fig. 5. Sentiment analysis steps

In our study, we use lexicon-based sentiment classification, also known as rule-based sentiment classification, and refers to the study conducted by the language experts. A classifier is a set of rules (also known as sentiment lexicon) where the words are classified as either positive or negative along with their corresponding intensity measure. Since the domain of employment is unique and we believe that generic lexicon methods tend to provide better performance unlike domains such as movies or products.

The two most popular lexicon-based sentiment classifiers are VADER [15] and TextBlob [16]. TextBlob is an NLP library, based on NLTK that comes prepackaged with a sentiment analysis functionality and returns the polarity score of a tweet. VADER (Valence Aware Dictionary for Sentiment Reasoning) takes into consideration punctuation, capitalization, degree modifiers, conjunctions, tri-gram preceding when assigning sentiment values: negative, positive, neutral. These features make VADER sentiment analyzer achieve remarkable results when classifying social media texts like tweets and hence, is a suitable tool to conduct our analysis. VADER's output is a dictionary of scores in four categories: negative, neutral, positive, compound where the latter is computed by normalizing the scores of the positive, negative, and neutral scores. The range of scores is from -1.0 to 1.0.

IV. Experiments & Evaluations

A. Classification of Advertisements & Non-Advertisements

As the training dataset is relatively balanced with 46% of data labelled non-adverts and accuracy is used as the performance comparison metric across models. The accuracy of the six algorithms based on the testing dataset is shown in Table 2.

Table 2. Advertisement classifiers & performance results

Classifier	Accuracy (%)
SKLearn Multinomial Naïve Bayes Algorithm	68.00%
SKLearn Bernoulli Naïve Bayes Algorithm	**74.00%**
SKLearn Linear Logistic Regression Algorithm	65.30%
SKLearn Linear Stochastic Gradient Descent Algorithm	61.30%
SKLearn SVC Algorithm	71.30%
SKLearn Linear SVC Algorithm	60.00%

The Bernoulli Naïve Bayes algorithm where each term frequency is given a binary-value had the highest accuracy at 74% and is chosen as the final model. Besides accuracy, the weighted average of precision and recall are 0.69 and 0.68 respectively. These results are also higher than the next two best-performing models by accuracy - Support Vector Classification and Multinomial Naïve Bayes.

Misclassification Analysis

The best model has an accuracy of 74% and examples of the misclassified tweets are shown in Table 3.

Table 3. Examples of misclassified tweets

Tweet	Actual label	Predicted label
#IIFLAAA is changing lives of independent financial advisors in India forever. India's only #workfromhome tablet for broking and financial advisory industry is here. #IIFL	Advert	Non-Advert
Launch Postponed to Saturday, May 30! Looking forward to making history, @NASA & @SpaceX...ðŸš€	Advert	Non-Advert
My boss really needs to get off my back. Im staying logged in and he still gives me this look. Lol!	Non-advert	Advert
Putting on a sweater a minute before standup starts	Non-advert	Advert

The reasons for the misclassification could be due to the nature of the text. Tweets by nature are very short text and after pre-processing (e.g. removal of stop words), the number of remaining terms became even smaller. This issue is further exacerbated with a relatively small training dataset of 600 that limited the number of terms in the dictionary which in turn could have affected the models' performance. In addition, certain terms are acronyms (e.g. dhs) or short forms (e.g. lol, pod) which create noise in the features for classification.

Analysis of Non-Advertisements

The best model is used to classify the tweet dataset (77,856 tweets) where about 29% of the data is classified as non-adverts and 71% are classified as adverts. All the non-advertisement data is combined and analyzed. 30% of the overall data (23,267 out of 78,606 tweets) are non-advertisements and its distribution across the four weeks is shown in Fig. 6. 23,267 non-advertisement tweets will be used for subsequent topic modelling and sentiment analysis tasks.

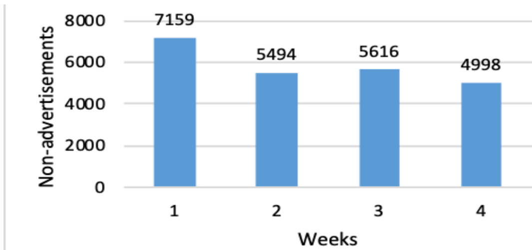

Fig. 6. Non-advertisement tweets over four weeks

B. *Themes Extraction Evaluations*

Recall that the outputs from both topic modelling algorithms on the entire month's data is compared to help us choose the best model. The results and observations are as described in this sub-section.

<u>Evaluations of LDA Models (Coherence Scores)</u>

Both LDA Gensim and LDA Mallet support the calculation of coherence and perplexity scores. Since the focus of our analysis is to obtain human interpretable topics and results, Coherence Score which is more intuitive to business users is chosen as the comparison metric and the results are depicted in Fig. 7. Comparing the coherence scores for both models over a range of K (# of topics), we observe that LDA Gensim consistently provides better scores when the number of topics is greater than 6. As we assume that our topic results should have more than 6 topics, we select LDA Gensim for topic analysis on the non-advertisement tweets (weekly and monthly).

```
2020-07-19 14:26:44.487995
#Topics: 4 Score: 0.1687123713936342
#Topics: 6 Score: 0.2082729273196795
#Topics: 8 Score: 0.24508254642659077
#Topics: 10 Score: 0.24448769574752954
#Topics: 12 Score: 0.26955560541848467
#Topics: 14 Score: 0.27501992159013067
#Topics: 16 Score: 0.3184841791903466
#Topics: 18 Score: 0.2854836559964184
#Topics: 20 Score: 0.2934668243227591
2020-07-19 14:32:29.731238
```

(a) LDA Gensim.

```
2020-07-19 14:34:17.470964
#Topics: 4 Score: 0.26270329560246936
#Topics: 6 Score: 0.23794709735645236
#Topics: 8 Score: 0.232800306981381
#Topics: 10 Score: 0.23903871556326528
#Topics: 12 Score: 0.23598422797344876
#Topics: 14 Score: 0.23984809609987753
#Topics: 16 Score: 0.23198925164788367
#Topics: 18 Score: 0.23180309376788413
#Topics: 20 Score: 0.23193959621766436
2020-07-19 14:39:47.797628
```

(b) LDA Mallet

Fig. 7. Coherence scores of LDA Gensim & LDA Mallet for K topics

C. *Sentiment Analysis Evaluations*

The evaluations on sentiment classifiers show that VADER has a higher accuracy of 67% compared to TextBlob's with 59% and hence, VADER is chosen for our project. The output of the VADER sentiment analysis is then visualized using Tableau to gain insights on the overall sentiment over the entire month period and how the sentiments might have varied across the weeks.

V. FINDINGS & ANALYSIS

A. *Theme Analysis*

<u>Monthly Topic Analysis</u>

From the topic-word distribution and sample tweets in each topic, we manually assigned a topic label to each topic to facilitate interpretability. LDA Gensim topic model on monthly data and the breakdown of the topics and percentage of tweets are shown in Fig. 8. From Fig. 8, we observe that WFH morale, Companies & Culture, and News/Others are top topics.

Topics	% of Tweets
WFH Morale	14.39%
Companies & Culture	13.73%
News/Others	12.51%
WFH Changes	9.54%
WFH Environment	9.06%
WFH Recommendations	7.97%
Health	6.08%
WFH questions	5.59%
Family	5.33%
Balance life	5.10%
WFH Preparations	4.81%
New normal	3.15%
Virtual working	2.73%

Fig. 8. Percentage distribution of tweets across topics sorted ascendingly

<u>Weekly Topic Analysis</u>

From Fig. 8, it is difficult to gauge if the topics are consistent throughout or have evolved across the 4-weeks period. We are also unsure if new topics related to returning to work might appear in the later weeks. Hence, we proceed to perform topic modelling analysis for the 4 individual weeks to obtain more granular insights. The breakdown of the topics and percentage tweets for each week is shown in Fig. 9.

From Fig. 9, we observe that there are consistent topics across the weeks (e.g. Companies & culture, News/Others, WFH morale) which indicated that these are time-independent during the analysis period and are perennial issues for employees. However, there are also slightly different topics of focus in each week.

At the start of lockdown and WFH is imposed in many countries, during Week 1, a larger proportion of tweets (20.74% of the week's total tweets) are on 'WFH Changes' and 'Family'. However, these topics are not very prominent in the later weeks.

Topics	% of Twe..
WFH Changes	20.74%
Family	16.91%
Companaies & Culture	13.78%
WFH Morale	13.42%
WFH Recommendatio..	13.33%
WFH Environment	11.02%
News/Others	10.79%

Week 1: 7 topics

Topics	% of Twe..
WFH Recommendatio..	16.36%
Companaies & Culture	15.94%
WFH questions	14.95%
WFH Morale	14.74%
WFH Environment	14.51%
Virtual working	11.85%
News/Others	11.65%

Week 2: 7 Topics

Topics	% of Twe..
Companaies & Culture	16.60%
News/Others	14.86%
Balance life	13.57%
New normal	13.07%
WFH Morale	12.65%
WFH Preparations	10.25%
Health	10.13%
WFH questions	8.87%

Week 3: 8 Topics

Topics	% of Twe..
WFH Morale	17.41%
Health	17.05%
WFH Changes	14.12%
News/Others	13.32%
WFH Preparations	10.98%
WFH Environment	10.54%
Balance life	8.57%
Companaies & Culture	8.00%

Week 4: 8 Topics

Fig. 9. Percentage distribution of tweets across topic labels (Week 1 to 4 in chronological order and human-labeled topics) *% is by dominant topic

In week 4 when the lockdowns in some countries started to ease, tweets revolving around 'Morale' (17.41%) and 'Health' (17.05%) started to gain traction. Whether there is positivity or negativity on these topics is not answered by topic analysis.

For those in countries that are still under strict lockdown measures, we observe several employees grew restless and stressed out from working from home and we observed an increasing proportion of tweets (9.28% for Week 3 and 16.81% for Week 4 of the respective weeks' total tweets) regarding 'Health' and 'Morale'. Such detailed observations and trend analysis on a weekly-basis help in understanding the human behaviors and response to the WFM situations.

B. Sentiment Analysis

Fig. 10 shows the monthly sentiment distribution. The following insights are observed from the results in Fig. 10. The bar chart reveals that 61% of the tweets are positive and only 17.5% of the tweets are negative. This shows that people are speaking more positively about working from home arrangements.

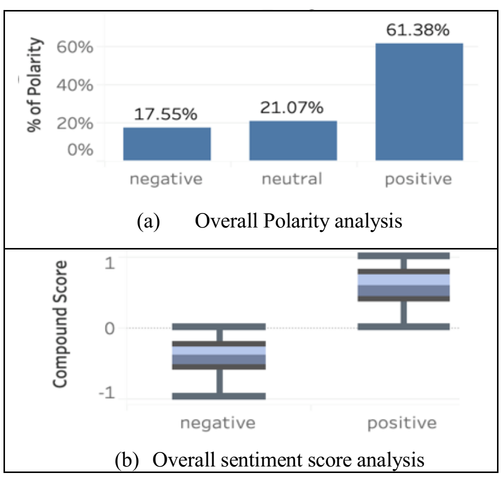

Fig. 10. Sentiment distribution of tweets - overall

From the sentiment spread in the box plot, we observe that the spread is wider for the positive tweets compared to the negative tweets with a median compound score of 0.598. We infer that the positive tweets would likely contain more positive comments leading to a higher compound score and correspondingly, a high median score.

The negative tweets have a median score of -0.38 indicating that these comments are mildly negative. Given that the median score is relatively high, we infer that the negativity may not have been too critical and perhaps entailed only mild dissatisfaction.

We visually analyze how the sentiments varied across the weeks as depicted in Fig. 11. We observe that there is a gradual decrease in the positive sentiment from 62.9% in Week 1 to 60.9% in Week 4, revealing that the positivity is gradually wearing off as the weeks passed. This could indicate restlessness, anxiety, and compounded negative feelings due to the various challenges noted in the topic modelling results such as the negative impact on work-life balance and productivity due to a lack of proper working environment.

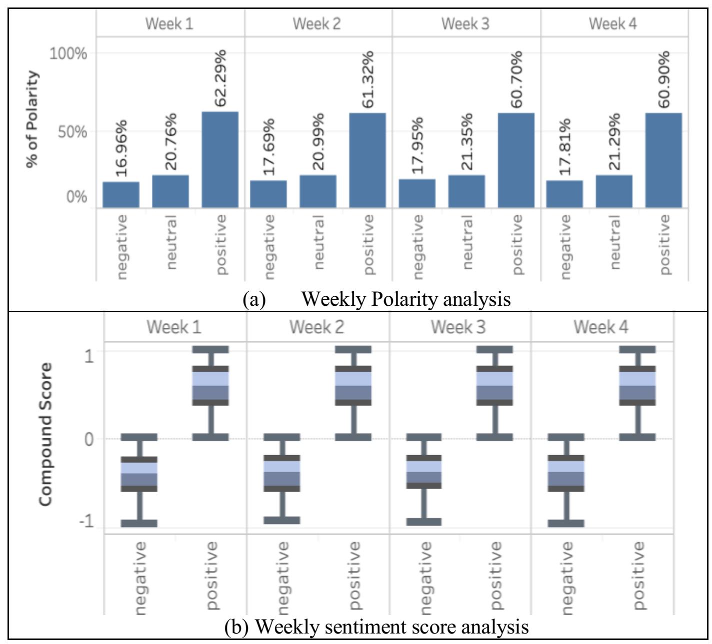

Fig. 11. Sentiment distribution of tweets - weekly

C. Analysis by Themes and Sentiments

To gain insights on the positivity and negativity for each topic discovered through topic modelling, we visualize the polarity of the tweets for each topic as shown in Fig. 12. We also analyzed the insights by weeks and the graph is not depicted due to space constraints.

Overall, positive sentiments are high for all the topics. For "New normal", the positive sentiment is the highest compared to other topics. The negative sentiment is highest for "Virtual working". "Virtual working", "Companies & Culture" (50%) and "Family" (50%) have the lowest proportion of positive tweets and high negative sentiments. "Morale" has a fluctuation but by the end of week 4, we observe more positivity (19%). In terms of "WFH Changes", "Environment", "Balance life" and "Questions" we observe that the positivity decreases.

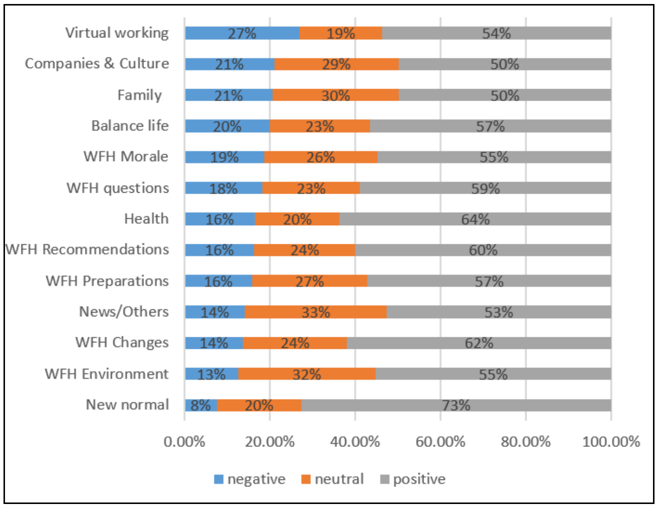

Fig. 12. Analysis by themes and sentiments

At the same time, concerns on "Health" and accepting "New normal" started late. In terms of "Health" the positivity and negativity increased over time. This shows that people have started embracing new norm while tweets on "Changes" decreased. Health has both polarity impacts indicating that some employees have taken measures independently while others could not balance the work. Similarly, some employees managed to have more "Balance life" by the end of week 4 while others could not achieve the same.

VI. Insights & Recommendations

Insights discovered by studying the topics discovered via topic modelling, the hashtags word cloud, and the sentiment analysis helped us to brainstorm some suggestions.

Topics such as 'Family', 'WFH recommendations', and 'WFH environment' (Fig. 13) reveal both material and immaterial needs of people working from home.

Fig. 13. Tweets from topics 'Family', 'WFH recommendations' and 'WFH environment'

These topics also suggest that people are experiencing problems adjusting to WFH and signs of overwork. These findings are augmented by words discovered in the word cloud such as 'wellbeing', 'worklifebalance', and 'mentalhealth'. For the insights stated here, policymakers and employers may need to take care of the whole spectrum of people's needs and look into the regulation of work hours during WFH. A deeper analysis of such tweets by location, gender, age, etc., gives inputs to the decisions.

Topics on 'Companies & culture', and 'WFH morale' are augmented by words such as 'team', 'remoteteam', 'inthistogether' and leadership' in the word cloud. We observe that employees are unhappy, anxious, and expect that employers should rethink certain working culture related to online collaboration and meetings (Fig. 14).

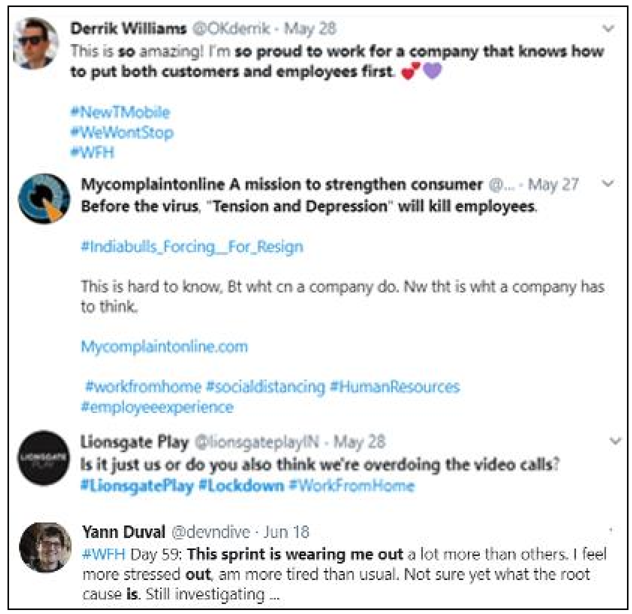

Fig. 14. Tweets from topics 'Companies & culture' and 'WFH Morale'

From the sentiment analysis results, the relatively low positivity and high negativity on topics such as 'Companies & culture' and 'Balance life' further unveil that these topics are time-critical for employers and policy makers to act on. In our tweet analysis by topics, we observe that the tweets about some companies being ahead of the curve while others lagged in cultivating a conducive WFH culture for their employees. Policymakers and employers should consider investing in psychological support for burnt-out employees and improve WFH culture in order to retain their talents and make WFH more efficient.

Other than the shortcomings of WFH, we identify topics such as 'WFH recommendations' and 'Virtual working' (see Fig. 15). These topics are in-line with words discovered in the word cloud such as 'tip' 'webinar' and 'virtualassistant'. These topics revealed encouragements, tips, digital tools and equipment for people to manage both material and immaterial needs. Policymakers and employers should consider tapping on social network platforms or refer to the recommendations to improve support for mental wellness and other practical support.

Fig. 15. Tweets from topics ‘WFH recommendations’ and ‘Virtual working’

In the later weeks (Weeks 3), the topic ‘New normal’ enabled us to discover more tweets (see Fig. 16) on people’s feelings about returning to their work premises. We infer that policymakers and employers should look into aiding employees to transit to and from WFH, especially if another lockdown is needed. The results from sentiment analysis on the overall positivity related to the topic ‘Embracing the New Normal’ also reveals that people are still optimistic about the new normal of working from home. This presents employers and policy makers an opportunity to ensure continued and rising satisfaction from the general workforce by addressing critical issues like work-life balance on time.

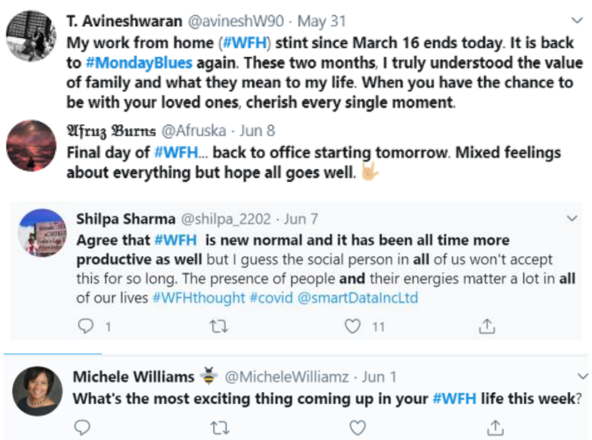

Fig. 16. Tweets from topic ‘New normal’

Finally, other concerns (see Fig. 17) include the lack of adequate network infrastructure and connectivity, improper work environment, and drops in physical fitness.

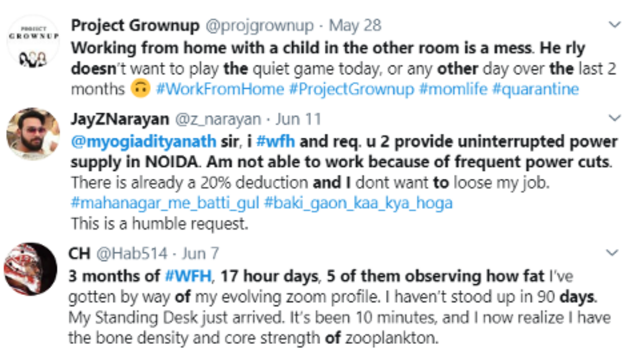

Fig. 17. Tweets about Other Concerns

VII. Limitations

Tweets with no or little textual content: We observed tweets with no or little textual content among our scrapped dataset. Instead, such tweets had images, URLs, GIFs, or videos attached as an expression. Such expression could not be analyzed with text mining and requires image processing models. Advertisements in tweets irrelevant to project: For the 78,606 tweets crawled using the list of hashtags mentioned in Section 1, we discovered almost 70% of tweets are of advertisements. This resulted in a lower data sample for analysis and may require more data crawling.

Classification of tweets into advertisements and non-advertisements. We had to manually label the training dataset into ‘adverts’ and ‘non-adverts’. This task of labeling the training dataset is laborious as we had to look into each tweet to check if it originated from a person or organization/company and discern via the text and URLs if the tweet is of an advertisement or non-advertisement nature. Our classification with 70% accuracy affected the topics and there is a need for better classification models for advertisements.

Impact of data on topic modelling and sentiment analysis: As a result of misclassification, tweets of advertisement nature appeared under the non-advertisement dataset, which is subsequently used for topic modelling and sentiment analysis. When manually giving labels to the topics output, we observed that some top-10 words in some topics are commonly found in the tweets of advertisement nature, implying that the misclassification had some impact on topic modelling and some are related to the news. Also, we noticed that tweets of advertisement nature generally used positive words. Hence, the inclusion of such tweets in sentiment analysis could cause a slight leaning towards positivity. Improving the performance of the classifier would mitigate the impact of misclassification on topic modelling and sentiment analysis.

Impact of existing library ‘tweet-preprocessor’ on topic modelling and sentiment analysis: The use of such package removed URLs and other content from the tweets rendering some tweets comprising entirely of hashtags. Such text has a performance impact on topic modelling and sentiment analysis. Related work on MA-LDA (Multi-Attribute Latent Dirichlet Allocation) by Wang et al [17] suggests that important topics and sentiments could be yielded from hashtags.

Manual topic labeling: Much effort is put into ensuring suitable and consistent labels across weeks of tweets. We had to look into the tweets containing the top-10 words for each topic by week to gain insights on the suitable topic label to assign to the topics as well the tweets falling under the topics. Since the human judges are working professionals, the task is managed with some inter-agreements.

Other than the limitations and improvements suggested previously to mitigate the impact of gaps we discussed, the scope of the project can be extended to other features to have deeper insights on each topic. Extraction of text from images: In our data analysis, we observed some images containing textual content. Such textual content is relevant and useful for topic modelling and sentiment analysis. In fact, there is ongoing work in this area of study [18] [19] and an existing

Python library (Tesseract) [20] that can be used for image analysis for text extraction.

Analyzing images for expressions and emotions: Expressions of emotions representing different sentiments are observed among tweets with no or little textual content but with images. The future work is to detect emotions and perform sentiment analysis on images with humans and faces [21] [22]. Analyzing accompanying videos/gifs for sentiments: With the increasing usage of videos and GIFs on social platforms to express sentiments, another future work is to perform sentiment analysis on videos or GIFs accompanying the tweets. Ongoing work in this area includes work done by Kumar et al [23].

Topic modelling and sentiment analysis based on geolocation: Scrapped tweets include the geolocation of Twitter users. Geolocation data is useful to analyze the topics and sentiments of tweets from various countries or regions. We believe that topic modelling and sentiment analysis based on geolocation provide relevant insights to policymakers and employers seeking to understand the local situation or comparing across regions or countries (for the case of multinational companies).

VIII. CONCLUSION

In this work, we apply NLP and text mining techniques to pre-process and analyze the textual data from tweets scrapped based on hashtags related to work from home in COVID-19 situation. Insights derived from the topics and sentiments across weeks, together with manual investigations into the tweets provided recommendations for employees and policy makers to make decisions related to WFH situation. With the given one month data and the useful insights, which may not usually surface in employee feedback or surveys, policymakers and employers may prepare necessary action items during the COVID-19 pandemic. Future work suggested in section VII offers deeper insights based on locale and across a longer duration to implement policies and measures for people/employees working from home.

ACKNOWLEDGEMENT

We thank Hui Ling Tan and Danny Kai Wing Yam for their assistance in this research study.

REFERENCES

[1] Dunford, D., Dale, B., Stylianou, N., Lowther, E., Ahmed, M., de la Torres Arenas, I. (2020, Apr 7) Coronavirus: The world in lockdown in maps and charts. BBC News. https://www.bbc.com/news/world-52103747

[2] Alon, T. M., Doepke, M., Olmstead-Rumsey, J. and Tertilt, M. (2020), ‘The impact of covid-19 on gender equality’, National Bureau of Economic Research Working Paper 26947.

[3] Tweepy API. 2020 http://docs.tweepy.org/en/latest/api.html

[4] Kateb, F., Kalita, J., (2015). Classifying Short Text in Social Media: Twitter as Case Study. International Journal of Computer Applications, 111(9), 1-12.

[5] Yang, S., Zhang H., (2018) Text Mining of Twitter Data using a Latent Dirichlet Allocation Topic Model and Sentiment Analysis. International Journal of Computer and Information Engineering 12(7) 525-529.

[6] Westerlund, M., Leminen, S., Rajahonka M., (2018) A Topic Modelling Analysis of Living Labs Research Technology Innovation Management Review, 8(7), 40-51.

[7] Li, M., Ch’ng, E., Chong, A., See, S., (2016) Twitter Sentiment Analysis of the 2016 US Presidential Election using an Emoji Training Heuristic. Applied Informatics and Technology Innovation Conference (AITIC).

[8] Angiani, G., Ferrari, L., Fontanini, T., Fornacciari, P., Iotti, E., Magliani, F., & Manicardi, S. (2016). A Comparison between Preprocessing Techniques for Sentiment Analysis in Twitter. KDWeb.

[9] Bird, S., Klein, E. and Loper, E. (2009). Natural Language Processing with Python. O’Reilly Media Inc. http://nltk.org/book

[10] Hartmann, J., Huppertz, J., Schamp, C., Heirmann, M. (2019). Comparing automated text classification methods. International Journal of Research in Marketing, 36, 20-38.

[11] Pedregosa, et al., 2011. Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830, 2011.

[12] Rehurek, R., & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora.

[13] McCallum, A. K.. "MALLET: A Machine Learning for Language Toolkit." http://mallet.cs.umass.edu. 2002.

[14] Liu, B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139084789

[15] Hutto, C. J. & Gilbert, E. E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

[16] Textbolb.com, “Textblob Tutorial, Quickstart“, [online] Available at: https://textblob.readthedocs.io/en/latest/quickstart.html#quickstart

[17] Wang, J,, Li, L., Tan, F., Zhu, Y., Feng, W. (2015, October) Detecting Hotspot Information Using Multi-Attribute Based Topic Model. PLoS ONE 10(10).

[18] Sarshogh, R., Hines, K. (2019, Jan 29) Learning to Read: Computer Vision Methods for Extracting Text from Images https://medium.com/capital-one-tech/learning-to-read-computer-vision-methods-for-extracting-text-from-images-2ffcdae11594

[19] Chen, T., Lu, D., Kan, M-Y., Cui, P. (2013) Understanding and Classifying Image Tweets. MM ’13: Proceedings of the 21st ACM International Conference on Multimedia https://doi.org/10.1145/2502081.2502203

[20] Andreou, C. (2020, Apr 10) How to Extract Text from Images with Python https://towardsdatascience.com/how-to-extract-text-from-images-with-python-db9b87fe432b

[21] You, Q., Jin, H., Luo J. (2017) Visual Sentiment Analysis by Attending on Local Image Regions. AAAI’17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. 231-237.

[22] Gajarla, V., Gupta, A. Emotion Detection and Sentiment Analysis of Images https://www.cc.gatech.edu/~hays/7476/projects/Aditi_Vasavi.pdf

[23] Kumar, A., Garg, G. (2019) Sentiment analysis of multimodal twitter data. Multimedia Tools and Applications. 78, 24103-24119.

[24] Rehurek, R, Sojka, P (2011) Gensim-statistical semantics in python

[25] Gottipati, S., Shankararaman, V. & Lin, J.R. Text analytics approach to extract course improvement suggestions from students’ feedback. RPTEL 13, 6 (2018). https://doi.org/10.1186/s41039-018-0073-0

[26] Gottipati, S., Shankararaman, V., & Lin, J. Latent Dirichlet Allocation for textual student feedback analysis. (2018). Proceedings of the 26th International Conference on Computers in Education ICCE 2018: Manila, Philippines, November 28-30. 220-227.

[27] S. Gottipati, K. J. Shim, A. W. Jiang and A. J. S. W. Lee, "Exploring the Impact of COVID-19 on Aviation Industry: A Text Mining Approach," 2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2020, pp. 0208-0215, doi: 10.1109/IEMCON51383.2020.9284858.