

## Mining social media opinion on online distance learning issues during and after movement control order (MCO) in Malaysia using topic modeling approach

Noor Afni Deraman<sup>1\*</sup>, Alya Geogiana Buja<sup>2</sup>, Siti Daleela Mohd Wahid<sup>3</sup>, and Mohd Ali Mohd Isa<sup>2</sup>

Lecturer, Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka, Malaysia<sup>1</sup>

Senior Lecturer, Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka, Malaysia<sup>2</sup>

Senior Lecturer, Faculty of Business Management, Universiti Teknologi MARA Cawangan Melaka, Malaysia<sup>3</sup>

Received: 30-October-2020; Revised: 14-February-2021; Accepted: 18-February-2021

©2021 Noor Afni Deraman et al. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

*The implementation of the Movement Control Order (MCO), which resulted in the closing of non-essential operations, led to the implementation of online learning at the university. This sudden announcement places university stakeholders in a state of unpreparedness to face the challenge of Open and Distance Learning (ODL). As this occurs unexpectedly and affects various people from all backgrounds, social media's views and debates need to be checked. This is necessary such that support can be given and their concerns heard, and action can be taken. This study is done by scrapping data from Facebook and Twitter with specific keywords from 17th March 2020 to 10th October 2020. A total of 2000 data were collected, but only 1283 were used after the pre-processing of the document. The results of the study show that the issues often addressed include "fees," "tired," "ODL," "information," and "zakat."*

### Keywords

*Movement control order (MC), Open and distance learning (ODL), Opinion mining, Social media analytic, Topic modeling.*

## 1.Introduction

The Covid-19 pandemic surprised the whole world and transformed the way of life to a new standard. The first phase of covid-19 in Malaysia began on 25th January 2020 with the confirmation of the first three cases. The whole situation is under control until the second wave of the covid-19 hit Malaysia in March 2020. COVID-19 pandemic leads to the implementation of Movement Control Orders (MCO) and the immediate suspension of civic and commercial operations involving all parties and sectors. The effect is also being felt by universities that are the driving force in Malaysia. In response to the MCO implementation, Malaysia's Ministry of Higher Education (MoHE) has requested all universities to adopt online learning or Open & Distance Learning (ODL) to ensure that students can continue learning throughout the running semester.

Online learning is not new in Malaysia, as it has long been recognized as e-learning, blended learning, and a Massive Open Online Course (MOOCs). Before Covid-19, online learning implementation was flexible with at least 50% face-to-face learning and 50% online learning. Online learning also uses application-specific e-learning tools developed by the university itself, such as the Learning Management System (LMS) and other supportive systems by third-party applications such as Google Classroom, Kahoot, Padlet, and Moodle [1]. In line with technological advances, communication between lecturers and students has also changed where instant messaging platforms such as WhatsApp and Telegram have become a quicker and more efficient area of information delivery. Supporting this additional approach and face-to-face meeting with lecturers would allow students to better understand the topic with additional resources to improve their understanding. However, online learning challenges are things that lecturers and students need to face

\*Author for correspondence

because they are unexpectedly implemented without having time to make careful plans [2]. There are various challenges for students who from different backgrounds and household income groups, such as internet access, devices, and data. University management must take this need into account for the continuity of student learning [3, 4].

Further research needs to be carried out in order to understand better the challenges and difficulties that university stakeholders face in implementing ODL. The objective of this paper is therefore to answer the following research questions:

RQ1: Is there any correlation between MCO's implementation and MoHE's public announcement to the social media user?

RQ2: What is the subject of debate when open and distance learning is carried out in all universities in Malaysia?

Approach taken to deal with these are by analyzing the content on social media. Students are regarded as avid social media users. Social media is a platform for them to share information, express their opinions, and connect with other people globally. This study uses Twitter and Facebook as a source of data and RapidMiner as a medium for implementing the opinion mining process.

The rest of the paper is structured as follows. Section 2 outlines the relevant issues regarding MCO, social media analytics, and topic modeling. Section 3 detail out the process of opinion mining and operator used in RapidMiner. Section 4 addresses the overall summarization result of the opinion mining process and answering the research questions stated. Section 5 draws some conclusions and possible work.

## 2.Literature review

The deadly Wuhan virus, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), also known as coronavirus disease 2019 (COVID-19) affected 213 countries with 1,524,162 confirmed cases and 92,941 deaths. The first case reported in Malaysia is on 25th January 2020 from a man that returns from Wuhan. Since then, several new cases have been reported with a controllable number [4]. However, the death toll increases drastically, and with this unexpected global epidemic, the World Health Organization (WHO) declared on 11th March 2020, Covid-19 has spread as a pandemic. In response to this update, the Malaysian Government takes initiative steps by implementing MCO on 18th March 2020 to break the chain. At that time, Malaysia was hit with a second wave, where a new

cluster originated from religious gatherings attended by more than 10,000 participants. As a result, all public facilities such as universities, schools are closed, and non-essential activities are canceled.

University press released have been issued by the MoHE, and the students continue to study online during the MCO. The Government of Malaysia has implemented 7 MCO phases, as shown in *Table 1*. MoHE will issue a press release each time Majlis Keselamatan Negara (MKN) extends the MCO [5]. The enforcement of MCO in Malaysia has greatly affected all citizens, and in this case, the university itself. All students are asked to leave the university before 18th March 2020, which brings outcry.

Malaysians use social media networks to voice their fears about the new developments in the Covid-19 pandemic. Rated fourth in the world in the field of mobile social penetration, Malaysians have publicly addressed MCO issues and the closure of non-essential activities from the date of introduction of the MCO [6]. University students are also inclined to make social media a forum for dialogue and communication of their feelings.

Social media such as Facebook, Twitter, and Instagram are described as an application or website that enables users to share their ideas or views and engage in social networking. It has become a critical resource for coordinating different real-life activities by offering a platform for a massive audience immediately [7]. The material shared in social media is up-to-date and sometimes only available for a brief period when posts that suit the latest trend and virality draw the audience. Social media's participatory nature provides an atmosphere that enables people to exchange knowledge, express their opinions, and participate in discussions. It is not unprecedented to find fresh and specific information with intimate details for an event on social media sites compared to conventional media [8]. The gathering of the most valuable data from the social media site is known as the social media analysis, where the processed data can provide some insight and lead to actionable conclusions for the responsible parties, such as government agencies, private and non-profit organizations (NGOs) and companies. In Malaysia, Facebook, Twitter, and Instagram are the most commonly used social media sites, as seen in *Tables 2 and 25* to 34-year-old users are the primary demographic users of social media [9]. The expanding number of Twitter and Facebook users in Malaysia has garnered many opportunities to gain

better insight into issues, trends, and correlation. Improved decision-making and policymaking can be

taken based on demographic data and targeted consumers.

**Table 1** Chronology of MCO stages and MoHE press statement

Stages	Date	MoHE press statement
Phase 1: MCO	18th March 2020- 31st March 2020	Operation of Academic Programs in Public Universities (IPTA) and Private Universities (IPTS) (28th March 2020)
Phase 2: MCO	1st April 2020 - 14th April 2020	Message to IPT Students Following the Announcement of Phase 3 Movement Control Order (12th April 2020)
Phase 3: MCO	15th April 2020 - 28th April 2020	IPT Student Return Movement to Kampung Halaman Starting from 27th April 2020 (26th April 2020)
Phase 4: MCO	29th April 2020 - 12th May 2020	
Phase 5: Conditional MCO	13th May 2020 - 9th June 2020	Conduct of Academic Activities at IPTA & IPTS During and Post Movement Control Orders (27th May 2020)
Phase 6: Conditional MCO	10th June 2020 - 31st August 2020	
Phase 7: Recovery MCO	1st September 2020 - 31st December 2020	Postponement of Physical Registration/ Student IPT Admission October 2020 (2nd October 2020)

**Table 2** Social media statistic in Malaysia [9]

Social media	Total audience
Facebook	24 Million
Instagram	12 Million
Twitter	2.63 Million
Snapchat	1.65 Million

Mining opinion on social media has become a large subject of study. One of the methods used in addition to sentiment analysis is topic modeling. Topic modeling is one of the text analytics tools whose primary purpose is to derive information from a large volume of text data. It will reflect the information in the set using an unsupervised machine learning method [10,11]. Topic modeling has two main objectives: to discover hidden themes in text data by clustering similar words groups accordingly and classifying the document into the discovered theme. The effect of topic modeling is useful in improving classification by grouping related terms together in topics and identifying patterns in social media [12]. It is often used in the recommender system with the use of similarity measures.

There are two basic topic models Probabilistic Latent Semantic Indexing (PLSI) and Latent Dirichlet Allocation (LDA). PLSI is a technique that learns the latent topic by performing matrix decomposition. Meanwhile, LDA is a generative probabilistic corpus model that uses Dirichlet over the latent topic. The basic concept is that documents are represented as random mixtures over latent topics, where each topic is defined by a distribution over words [13]. Although the PLSI is easier to train than the LDA, it has low accuracy. Another advantage of LDA over PLSI, LDA is usually better suited for documents

with a scale of at least 1000 words. As the data source for this study comes from Twitter and Facebook, the LDA approach is most fitting.

A lot of research on the topic of modeling using the LDA technique has been done. A study by [14] used LDA to examine the topic discussed during Brexit on Twitter and find the connection between Brexit sentiment and the British pound sterling exchange rate. A similar idea is applied by [15], where Twitter is mined using the topic modeling approach to understand consumer attitudes towards vaccination. LDA technique is used to interpret public perception of the non-pharmaceutical intervention (NPI) method during Covid19 was explored [16]. This study's findings defined issues relevant to keywords across six countries and assisted the government in the implementation of NPI decision-making and strategies.

Apart from that, LDA-based topic modeling has been adopted by [17] to identify product opportunities among Twitter users and quantify the importance of the topic identified. Similar work has been done by [18] where post related to climate change is clustered according to countries using geotagged tweet data.

### 3.Methods

A classical opinion mining process comprises of 4 stages: opinion retrieval, pre-processing data, topic modeling, and summarization, as shown in *Figure 1*.

#### Stage 1: Opinion retrieval

Opinion retrieval is the mechanism by which social media views and feedback on an issue are made. Postings related to ODL and MCO on Facebook and Twitter were selected for analysis. A collection of keywords has been established to scrap the related post. Some of the keywords used are "odl", "online learning", "online class", and the hashtag of all university's name in Malaysia. This project uses the Facepuger tool to scrape data from Facebook and Python script to scrape Twitter data. Data collected from the Facebook post are the post itself, comments received, and the readers' number of liked and shared. Meanwhile, attribute collected in Tweeter are the tweet itself, retweet count, and the number of likes received. The data collected is from 17th March 2020 to 10th October 2020. All collected data is stored in a CSV file to be processed.

#### Stage 2: Pre-processing

At this stage, raw data collected from Facebook and Twitter are pre-processed using RapidMiner. All raw data will go through 2 processes: text preparation and text pre-processing.

Text preparation is the process of readying the data for the machine learning process through various operators in Rapid Miner. *Figure 2* shows the detailed process done using RapidMiner. First, the raw data in excel format will be added to the RapidMiner repository. Then process Extract, Transform, and Load (ETL) is done on the data using RapidMiner operator. Standard operators were used at this process: "Select Attribute", "Nominal to Text", and "Process Document to Data". The "Select Attribute" operator is used to select the number of attributes available from the data collection. For this project, only a text message is selected as an attribute parameter, so the filtering type attribute is set to single. "Nominal to Text" operator is used purposely to convert all nominal data type attributes to readable text before mapping to the next operator, "Process Document to Data" where text input is needed. The "Process Document to Data" operator will process

selected attributes and create a word vector by applying pre-processing text.

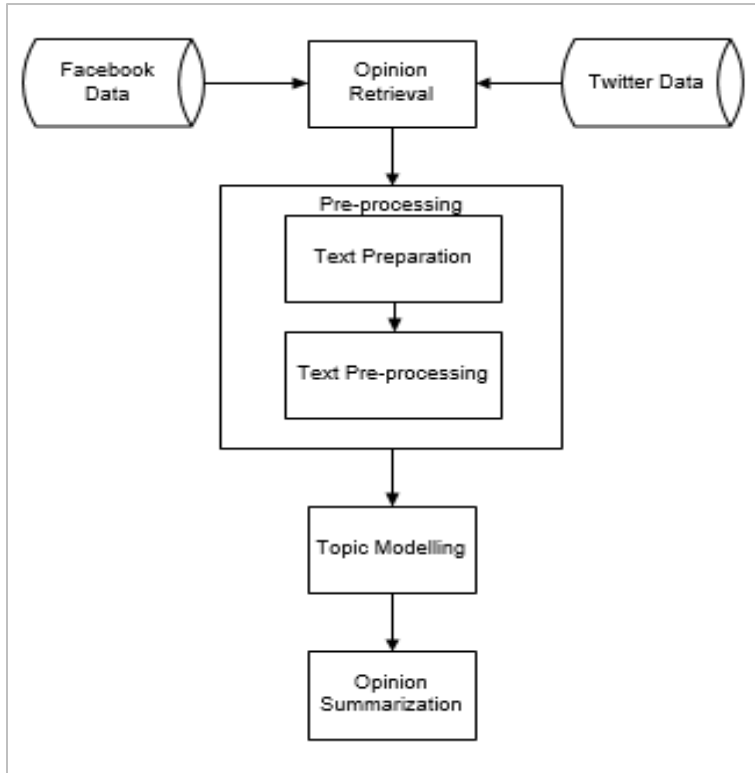
After all the data has been prepared, it will be converted into documents for further processing. Four main text pre-processing components consist of tokenization, transform case, remove stop words, and stemming process, as shown in *Figure 3*. Tokenization is the process of extracting (also called tokenizing) words from the body of the text. Then all text will be converted to lowercase using the "Transform Case" operator of RapidMiner. After that, all terms with a high recurrence rate and little analysis value in the document are omitted using the text "Filter Stopwords" operator. The remaining words will be stemmed using the "Stem" operator to remove suffixes. Similarly, stemming turns noun plurals into singular forms. The result of this process will be saved in the excel file using the "Write Excel" operator before being used for the topic modelling stage.

#### Stage 3: Topic Modeling

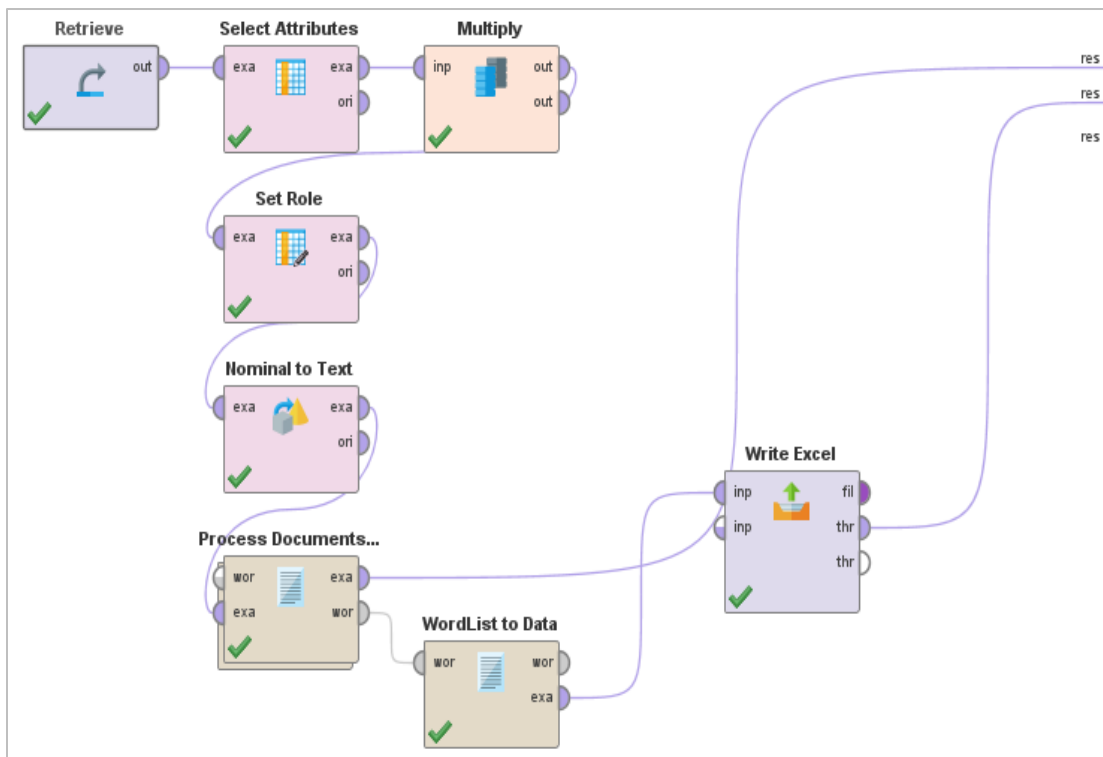
In RapidMiner, topic modeling is done using Operator Toolbox extension where the LDA method to find hidden theme in processed data. Referring to *Table 1* information, topic modeling is applied for all 7 phases of MCO, where the data collected are divided into seven sets of dates based on MCO phases. Since topic modeling is unsupervised learning, the data is unlabelled, and the process is set to iterate five times. The process of extracting the topic from the data collection requires a dictionary of the term document frequency and the corpus of the terms collected from the previous stage. *Figure 4* illustrates the process of extracting topic from a document using LDA. The operator "Extract Topics from Documents (LDA)" is used to provides topic diagnostics in the model object. After that, the process of topic annotation is done to group the topic into a meaningful theme.

#### Stage 4: Summarization

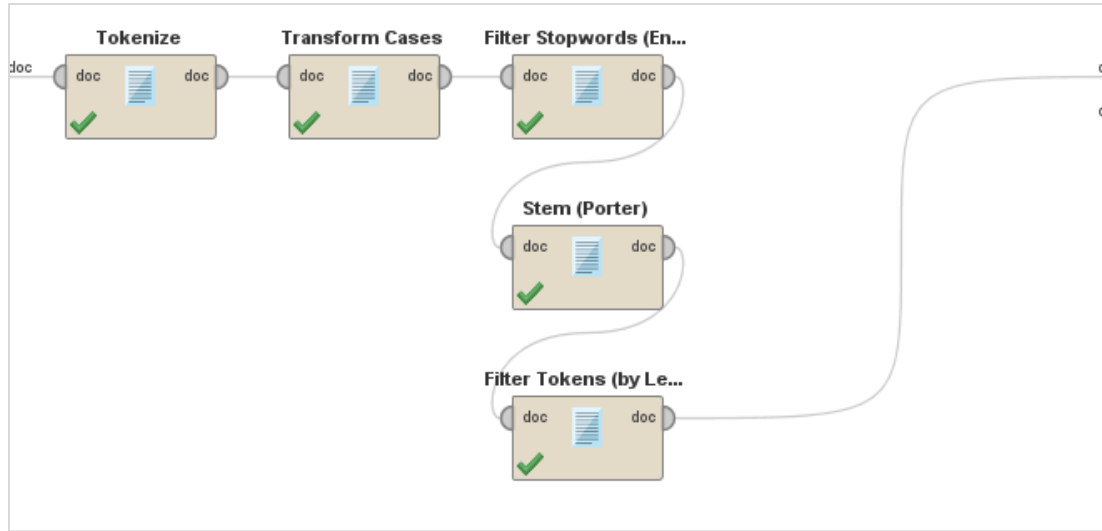
Finally, the modeling result is visualized in graph and word cloud format to comprehend better and answer RQ1. The Pearson correlation and the p-value are calculated in order to address RQ2.



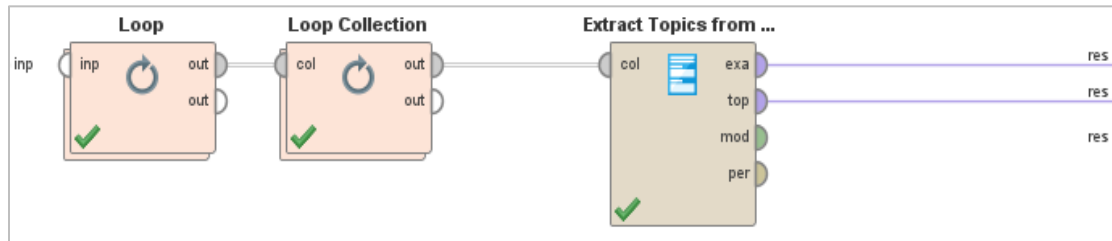
**Figure 1** Architecture of opinion mining



**Figure 2** Data preparation process



**Figure 3** Text pre-processing activities



**Figure 4** Topic modeling process

#### 4.Result and discussion

A total of 2000 data was scrapped on Facebook and Twitter, with 1,283 data from Facebook and 717 from Twitter. After the data cleaning process has been completed, 3371 keywords are obtained for analysis from 855 posts. The reduction is attributable to the removal of URL link, repetition, comments in the form of gif file and emoji, and an unimportant statement by the reader that is not relevant to the subject being scrapped. Non-English words that are meaningless and irrelevant to the issues are omitted manually. Furthermore, words with a document frequency (DF) of less than ten are removed, and tweets with less than ten tokenized words are also omitted. This section will answer two research questions of this study.

RQ1: What is the subject of debate when open and distance learning is carried out in all universities in Malaysia?

In answering RQ1, the detailed process of LDA is analyzed. Topic modeling using LDA required three

inputs: the corpus of processed data, the word dictionary that was acquired at the pre-processing stage, and the number of topics, which is RapidMiner set to 10 by default.

Table 3 shows part of the LDA analysis result where the document is grouped into a relevant topic. The topic extracted from the document is calculated using the LDA analysis. For example, Document 0 has a high confidence value for Topic 6: 0.3794 and Topic 7: 0.401, while the rest are less than 0.01. The topic represented is manually defined by referring to Document 0 and the high-frequency keyword. Document 0 has been linked to the online class problem where Topic 6 is ODL word, and Topic 7 is learning. Upon completion of the text annotation process, the theme is listed, as shown in Table 4. The keywords that define the subject are described based on the value of confidence received. Figure 5 indicates ten keywords commonly found in the corpus and is considered topics frequently discussed on social media.

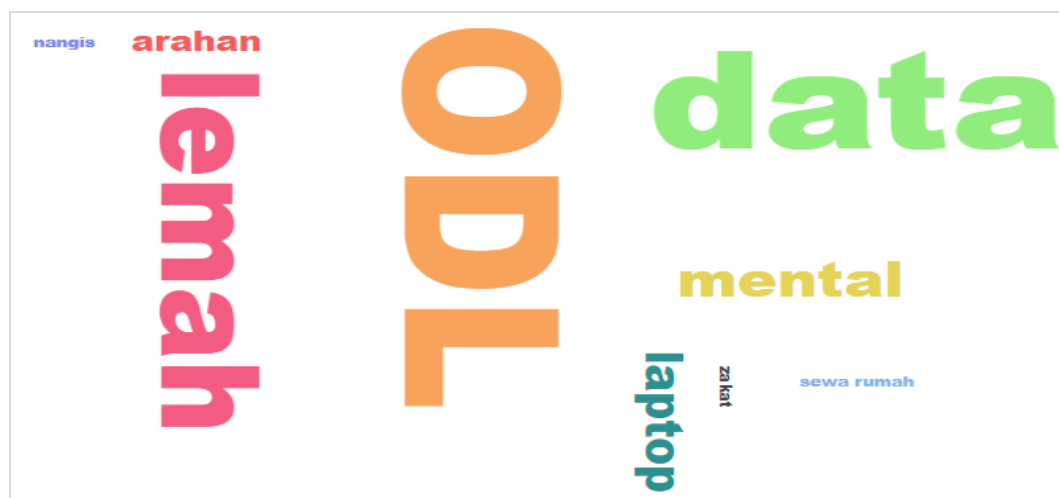


**Table 3** Part of LDA analysis

Document	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
0	0.0019	0.0251	0.0318	0.022	0.0238	0.3795	0.4016	0.0263	0.0721	0.0159
1	0.0250	0.3771	0.0272	0.0232	0.0428	0.0148	0.3974	0.0292	0.0312	0.0321
2	0.3785	0.3817	0.0289	0.0233	0.0243	0.0271	0.0232	0.0297	0.0279	0.0554
3	0.2581	0.1285	0.0303	0.0151	0.0232	0.0267	0.4135	0.0295	0.0313	0.0438
4	0.0279	0.0284	0.0287	0.0219	0.0142	0.4651	0.0437	0.0263	0.0227	0.3211

**Table 4** Topic discovered using LDA

Topic	Major keywords and representative data
Finance	<p>“yuran”, “sewa rumah”, “zakat”, working, “susah”, “kewangan”</p> <ol style="list-style-type: none"> <li>1. “Ada tak bantuan untuk pelajar yang kurang berkemampuan yang sedang belajar di *** sebab saya ada kawan yang berniat nak berhenti belajar atas sebab masalah kewangan”</li> <li>2. “Saya nk tanya bagaimana zakat untuk sem ni ye dapat ke tidak ye”</li> <li>3. “Mohon yuran dikurangkan bagi ODL”</li> </ol>
Emotion	<p>Stress, “penat”, mental, “nangis”, broken, miss, family</p> <ol style="list-style-type: none"> <li>1. “Odl is the death of me Smiling face cepatlah buka *** semula Pensive face”</li> <li>2. “Haritu ada student uni mentally exhausted disebabkan oleh odl”</li> <li>3. “Missing my uni, taknak belajar odl ha Frowning face Frowning face”</li> </ol>
Internet	<p>Data, connection, poor, Sabah</p> <ol style="list-style-type: none"> <li>1. “house electricity bill goes up nanti, money spent sbb buy data &amp; living expense”</li> <li>2. “Kalau rasa takleh odl kt rumah then balik lah ye. and internet pulak boleh apply yes 4G tu”</li> <li>3. “Mohon agar pihak *** dapat memberikan insentif untuk student membeli plan internet bagi ODL ni sebab tak semua student mampu beli plan internet yang mahal sepanjang ODL ni.”</li> </ol>
Device	<p>Laptop, handphone</p> <ol style="list-style-type: none"> <li>1. “*** nak bayar ke laptop aku yang dah semakin rosak sbb on 24/7 sepanjang odl ni? Siang malam siapkan assignment. Saya bukan anak orang kaya. Take note.”</li> <li>2. “Just realized assignment byk + no laptop thanks *** for odl”</li> <li>3. “ODL ni yes efficient, sebab assessment buat online, everything is just in the laptop”</li> </ol>
Class	<p>ODL, lecturer, assignment, project, test</p> <ol style="list-style-type: none"> <li>1. “Online classes won't be as effective as the physical class”</li> <li>2. “Since *** wants to implement online class/lectures until sept.... can we get our hostel's fee back?”</li> <li>3. “*** tolong laa final sem bagi student balik semua kolej aku dah tak sanggup nak odl, siapkan assignment dekat rumah. Kurang facilities dan banyak gangguan hmm”</li> </ol>

**Figure 5** Issues that are regularly addressed in social media on the basis of frequency value

RQ2: Is there any correlation between MCO's implementation and MoHE's public announcement to the social media user?

To address this question, the data were divided into 7 phases of the MCO. The number of tweets or posts collected during that period, as shown in *Figure 6*. The outcomes of the findings, MCO 1, MCO 2, and MCO 3 had a significant effect on university stakeholders. Details number of posts and tweets collected, as shown in *Figure 7*.

To assess the strength of the association between MCO's implementation and MoHE's public announcement to social media users, the Pearson correlation ( $r$ ) between the number of tweets and the number of posts counted for MKN and MoHE is measured. The results obtained are  $r = 0.8511$ , which means that there is a positive connection between the MCO implementation and the MoHE announcement. A hypothesis test is conducted to determine the significance of the correlation result.  $H_0$  and  $H_1$  are defined as followed:

$H_0$ : There is no correlation between MCO implementation and the MoHE announcement.

$H_1$ : There is a correlation between MCO implementation and the MoHE announcement.

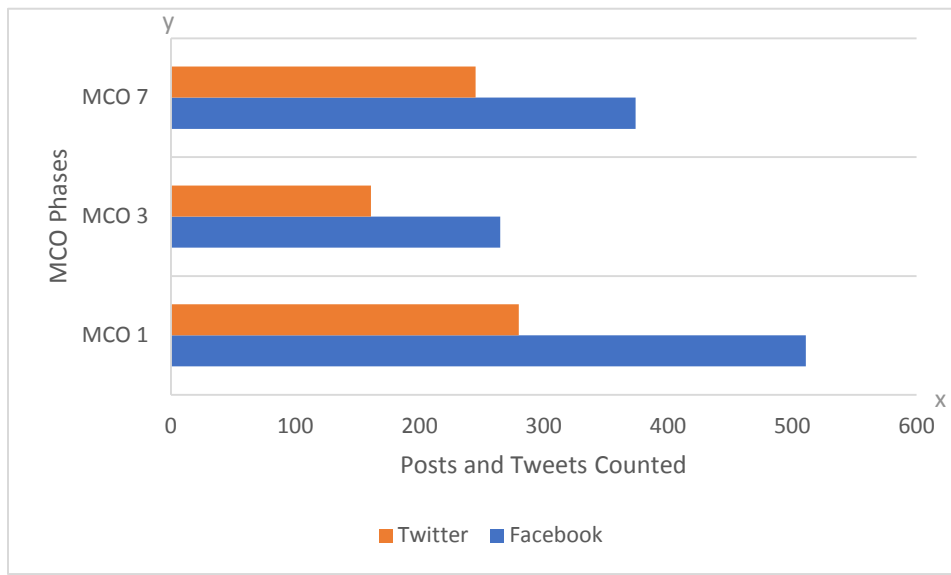
The t-test outcome provides a  $p$ -value of 0.01514 in which it is less than 0.05. This indicates that there is clear evidence against  $H_0$ , and it is rejected while  $H_1$  is supported.

As a result of this development, this research suggests a significant correlation between MCO and MoHE announcements to student posts and tweets. The findings of MCO 1 suggest a condition of concern due to the unreadiness of students and families. The majority of MCO 3 issues related to students are tired of studying online, like the anonymous comment in *Table 5*. MCO 7 is concerned about students' frustration as they are ready to start a face-to-face class, but the notification of cancelation is late. This is due to an unforeseen situation as Malaysia is facing the third wave of Covid-19.



**Figure 6** Total data collected During MCO1 to MCO 7





**Figure 7** Posts and tweets counted during MCO 1, MCO 3 and MCO 7

**Table 5** Posting and tweets by social media user

No.	Tweets and posting
1.	“Penatnya. Bngun pagi lps mndi buat assignment. Stop tngahari smyg zohor pstu smbung balik. Stop ptg smyg asar n smbung buat assignment smpai sekarang. This is student life during ODL.”
2.	“*** ws so lucky to have odl in the end of their semester, while *** started the odl on our week 4. How suppose I'm gonna get my cgpa constantly 3.5 above for this semester???? Odl ws really not for me suffered stressed and depressed for *** odl.”
3.	“***...tlg la sem depan tangguh sem...aq x larat la nk odl ni Loudly crying face”
4.	“It's week 12, typically for *** students to be so busy and start to breaking down, stressful hectic. But hell yeah this ODL make it worst. I really start to lost interest in study, i don't feel like I wanna care anymore about it.”
5.	“Cepat la odl ni habis , phone aku makin rosak siapa nak bayar *** ke”

#### 4.1 Limitation

There are two shortcomings identified in this study. The lack of Bahasa Melayu word dictionary and word list stemming will certainly have a significant impact during the pre-processing phase. The resulting keywords become very broad and require researchers to remove irrelevant keywords manually. This also occurs when the result of the topic modelling is analyzed. The topic of modelling developed use general labels: Topic 1, Topic 2, etc. To learn more about the subject under discussion, researchers need to examine the document in the corpus one by one and correlate it with the document's word frequency occurrence. Both of these factors have an effect on the overall outcomes of this study, as the resulting evaluation is based on the interpretation of the researcher, which is subjective, in order to classify the keywords and topics of the relevant group.

#### 5. Conclusion and future work

The main aim of this paper is to study the opinion of the Malaysian citizen on the introduction of the ODL with the implementation of the MCO through two research questions. This paper analyses as many as 2000 cleaned data from Facebook and Twitter on ODL issues. The topics of discussion on social media during the MCO and ODL periods were analyzed using topic modelling techniques. Economic, emotional and Internet connectivity problems have been described as the key themes of the debate. In addition, the findings of the study have showed that there is a relationship between the MCO announcement and the MoHE press release with the increased number of tweets and Facebook posts during the actual event. With these results, government officials and NGOs will provide a better interpretation of public views of these issues. The effect of this information and public voice allows the government and NGOs to carry out their duties

properly and to implement policies that support all walks of life. Hashtags used by social media users are also voice knowledge that needs to be taken into account, such as #SiswaJagaSiswa and #NormaBaharu. In spite of the contributions made in this study, further work still needs to be done. A comparison of techniques is required to analyze secret subjects in the corpus. A functional comparison with the LDA solution can be made in terms of performance, as the LDA cannot handle implicit aspects.

### Acknowledgment

Sincere appreciation goes to Universiti Teknologi MARA Cawangan Melaka for the support given to this research endeavor, TEJA: Internal Grant (GDT2020-17).

### Conflicts of interest

The authors have no conflicts of interest to declare.

### References

- [1] <https://www.mohe.gov.my/en/mediakpt/newspaper-cutting/1322-cabaran-pendidikan-tinggi-berdepan-impak-covid-19?showall=1>. Accessed 12 October 2020.
- [2] Arulogun OT, Akande ON, Akindele AT, Badmus TA. Survey dataset on open and distance learning students' intention to use social media and emerging technologies for online facilitation. *Data in Brief*. 2020; 31:1-8.
- [3] Dwidienawati D, Abdinagoro SB, Tjahjana D, Gandasari D. Forced shifting to E-learning during the COVID-19 outbreak: information quality, system quality, service quality, and goal orientation influence to E-learning satisfaction and perceived performance. *International Journal of Advanced Trends in Computer Science and Engineering*. 2020; 9(2): 1518-5.
- [4] Shah AU, Safri SN, Thevadas R, Noordin NK, Abd Rahman A, Sekawi Z, et al. COVID-19 outbreak in Malaysia: actions taken by the Malaysian government. *International Journal of Infectious Diseases*. 2020; 97:108-16.
- [5] <https://www.mohe.gov.my/en/media-mohe/press-statement/1372-penangguhan-kemasukan-pelajar-baharu-dan-lama-dari-kawasan-zon-merah-covid-19-ke-ipt>. Accessed 12 October 2020.
- [6] <https://www.bharian.com.my/bisnes/lain-lain/2019/01/526175/kadar-penembusan-media-sosial-malaysia-tertinggi-di-asia-tenggara>. Accessed 12 October 2020.
- [7] Kawash J, editor. *Online social media analysis and visualization*. Springer International Publishing; 2014:229-33.
- [8] Reese SD, Rutigliano L, Hyun K, Jeong J. Mapping the blogosphere: professional and citizen-based media in the global news arena. *Journalism*. 2007; 8(3):235-61.
- [9] <https://marketingsignallab.com/social-media-marketing-statistic-in-malaysia/>. Accessed 12 October 2020.
- [10] Tribhuvan PP, Bhirud SG, Deshmukh RR. Product features extraction for feature based opinion mining using latent Dirichlet allocation. *International Journal of Computer Sciences and Engineering*. 2017; 5(10):128-31.
- [11] Lamba M, Madhusudhan M. Application of topic mining and prediction modeling tools for library and information science journals. *Library Practices in Digital Era*. Eds. MR Murali Prasad et al. Hyderabad: BS Publications. 2018: 395-401.
- [12] Stokes DC, Andy A, Guntuku SC, Ungar LH, Merchant RM. Public priorities and concerns regarding COVID-19 in an online discussion forum: longitudinal topic modeling. *Journal of General Internal Medicine*. 2020; 35(7):2244-7.
- [13] Ankarali E, Külcü Ö. RapidMiner ile Twitter Verilerinin Konu Modellemesi. *Bilgi Yönetimi*. 2020; 3(1):1-10.
- [14] Ilyas SH, Soomro ZT, Anwar A, Shahzad H, Yaqub U. Analyzing Brexit's impact using sentiment analysis and topic modeling on Twitter discussion. In the international conference on digital government research 2020 (pp. 1-6).
- [15] Zhang H, Wheldon C, Dunn AG, Tao C, Huo J, Zhang R, et al. Mining twitter to assess the determinants of health behavior toward human papillomavirus vaccination in the United States. *Journal of the American Medical Informatics Association*. 2020; 27(2):225-35.
- [16] Doogan C, Buntine W, Linger H, Brunt S. Public perceptions and attitudes toward COVID-19 nonpharmaceutical interventions across six countries: a topic modeling analysis of Twitter data. *Journal of Medical Internet Research*. 2020; 22(9):1-11.
- [17] Jeong B, Yoon J, Lee JM. Social media mining for product planning: a product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management*. 2019; 48:280-90.
- [18] Dahal B, Kumar SA, Li Z. Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*. 2019; 9(1):1-20.



**Noor Afni Deraman** is currently a lecturer at Faculty of Computer and Mathematical Sciences, UiTM Cawangan Melaka Kampus Jasin. She obtains her bachelor's degree in Computer Science majoring in Software Engineering from Universiti Sains Malaysia. She completed her master's degree in Science (Computer Science) in 2006. Her research interest is in Data Science and Artificial Intelligence.  
Email: noora465@uitm.edu.my



**Dr Alya Geogiana Buja** is a Senior Lecturer at the Faculty of Computer and Mathematical Sciences in Universiti Teknologi MARA (UiTM) Cawangan Melaka. She is a PhD holder in the field of Information Security and graduated from Universiti Teknikal Malaysia Melaka, MSc in

Computer Science and BSc in Netcentric Computing from Universiti Teknologi MARA. Her research interests are Networking and Information Security, Cryptanalysis and Cyber Security.

Email: geogiana@uitm.edu.my



**Siti Daleela Mohd Wahid** is a senior lecturer at Faculty of Business Management, Universiti Teknologi MARA, Melaka. She completed her BBA and MBA at Universiti Teknologi MARA, Malaysia. Currently, pursuing PhD under supervision of Dr Wan Mohd Hirwani Wan Hussain in social

entrepreneurship at The National University of Malaysia. Her areas of interest are Societal Innovation, Social Entrepreneurship, Community Entrepreneurship, Entrepreneurship, Marketing and Service Quality.

Email: sitid365a@uitm.edu.my



**Mohd Ali Mohd Isa** is a senior lecturer at MARA University of Technology (UiTM). Previously he was the Head of Department for System Management at the Center of e-Learning UiTM from 2005 until 2015. He was responsible for the setting up of the first Learning Management System for UiTM and

oversee the working of the system for 10 years. Currently he is teaching the subjects of Web Technology and Information Security, which are 2 of his research areas.

Email: mohdali@uitm.edu.my