International Journal of Synergy in Engineering and Technology Vol. 3 No. 1 (2022) 79-87

International Journal of Synergy in Engineering and Technology IJSET

Text Classification on Social Media using Bidirectional Encoder Representations from Transformers (BERT) for Zakat Sentiment Analysis

Azhar Hamid*[1], Raja Mohd Tariqi Raja Lope Ahmad[1], Wan Azlan Wan Hassan[1], Suziyanti Majudi[2], Syahrul Fahmy[3]

[1] Faculty of Communication, Visual Art and Computing, Universiti Selangor, MALAYSIA
[2] Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn, MALAYSIA
[3] Big Data Institute, University College TATI, MALAYSIA

*Corresponding author: azhar08@unisel.edu.my

KEYWORDS

Social Media
Sentiment Analysis
Text Classification
BERT
Zakat
Transformer
Big Data
Deep Learning
Machine Learning

ABSTRACT

Social media has impacted social communities in many aspects. They provide a platform to share information, express opinions and discuss common interests. They have become important part of everyday life. Companies used them for marketing and engage with customers for their services and products. The usage of social media has great influence on company image reputation. A viral positive message is good for a company but negative viral message may affect image reputation of the company. Zakat institution manage zakat payment and distribution of zakat to the zakat receivers. There are zakat payers and receivers that did not satisfy with the current mechanism of zakat and they use social media to communicate not only on personal matters but also their satisfaction with services or products that they received. Therefore, sentiment analysis on zakat is important to resolve the customer dissatisfaction or problems. BERT model is used to analyze the sentiment analysis since it is a powerful Transformer-based machine learning model for Natural Language Processing (NLP). Data is obtained from social media based on keywords of zakat. Then a text pre-processing and word-embedding features of BERT are used to build a text classification model. This classification model is used to analyze the sentiment analysis on the zakat institution. The result from this model can be used by zakat institutions to resolve zakat payer and receiver matters in a more personalized customer service systematically and efficiently.

Received 20 April 2022; received in revised form 20 May 2022; accepted 25 May 2022

1.0 Introduction

Businesses can improve their performance in a variety of ways by embracing social media, including increasing brand loyalty (Hamid et al., 2019) (Laroche, Habibi, & Richard, 2013), improving the sales process (Andzulis, Panagopoulos, & Rapp, 2012) and consumer spending (Wang, Yu, & Wei, 2012), developing digital brands and earning more income (Geiser, 2017).

Text analytics from social media can be used for a variety of reasons. Sentiment analysis is one of the most well-known. The emotion of customers towards a brand is identified by classifying texts from blogs and other media (Mostafa, 2013). These findings reveal customers input that is beneficial to a company. Text analytics can also be used to help in crisis management. Customers complaining about product quality on a Facebook fanpage, for example, can have a detrimental impact on sales. As a result, marketing departments should recognise the impact of social media and begin using tools to uncover metrics and keywords, which will help them enhance marketing campaigns and find new methods to communicate with customers (Chumwatana & Wongkolkitsilp, 2019).

Zakat is the universal welfare system that has played an important role in helping the poor and needy. There was a time in history that Muslim community rarely needed charity due to effective zakat management (Khan, 2015). Therefore, zakat clearly can be an effective way to overcome poverty. For a competent Muslim it is obligatory to pay zakat, which is one of the five pillars of Islam that a Muslim must fulfil. The main activity of zakat is the collection and distribution of zakat to the poor and needy people. However, in Malaysia, issues of inefficiency in collecting and distributing zakat have become a major issue in poverty alleviation (Nazri, Rahman, & Omar, 2012)(Ab Rahman, Alias, & Omar, 2012) (Hamid et al., 2016). The Covid-19 pandemic has caused many businesses to close their operations, lay off staff and decrease of revenues. It causes many people lost their jobs, reduce the number of zakat payers and increase the number of zakat receivers. This people express their needs, satisfactions and dissatisfactions on zakat via social media. The negative viral give very bad image to the zakat institution which need people trust and confidence to participate in zakat programmes. Therefore, it is very important for zakat institution to manage the sentiment matters and this research suggest towards the contribution to resolve the matters by Text Classification on social media using Bidirectional Encoder Representations from Transformers (BERT) for the Zakat Sentiment Analysis.

This study used text analytics which are part of big data technology to analyse sentiment on zakat to resolve the problems of collection and distribution. This method involves huge data that we choose to use deep learning method instead of machine learning. Deep learning method specifically text classification using BERT approach is more suitable in this scenario due to its ability to do processes faster and more effectively.

2.0 Literature Review

BERT (Bidirectional Encoder Representations from Transformers) is a natural language processing (NLP) model that was created to pretrain deep bidirectional representations from unlabelled text and then fine-tune them using tagged text for various NLP tasks (Devlin, Chang, Lee, & Toutanova, 2019). As a result, we may develop state-of-the-art models for a variety of NLP tasks using the BERT model. The results of BERT in several NLP tasks may be found in (Devlin et al., 2019). BERT has recently obtained best-in-class scores in a variety of NLP tasks (Devlin et al., 2019). To put it another way, BERT can be thought of as a more dynamic generalisation of the traditional NLP pipeline (González-Carvajal & Garrido-Merchán, 2020).

2.1 Bidirectional Encoder Representations from Transformers (BERT)

There are two steps in BERT model framework which are pre-training and fine-tuning. In pre-training, the BERT model is trained on unlabelled large corpus and for fine-tuning, the model is initialized with pre-trained parameters and they are fine-tuned using labelled data for specific jobs (González-Carvajal & Garrido-Merchán, 2020).

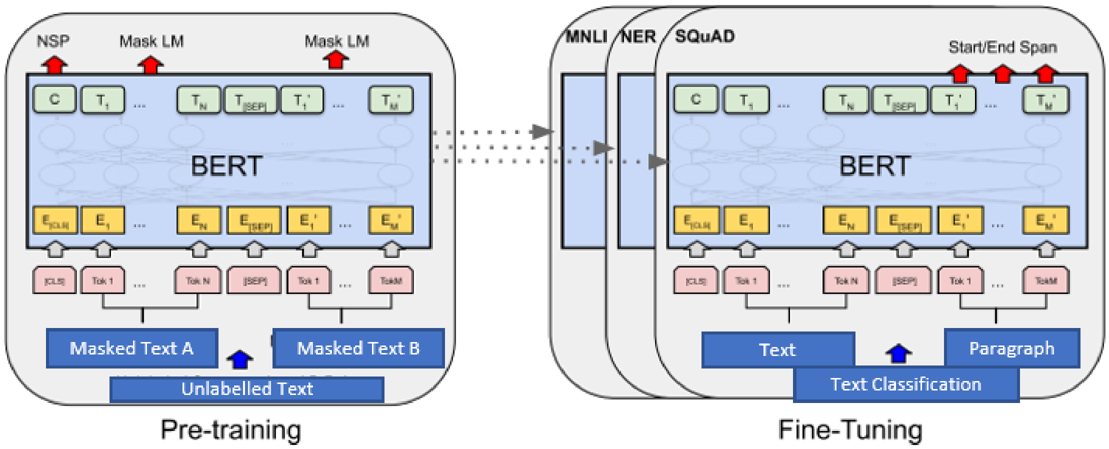


Figure 1: Pre-training and Fine-tuning procedures for BERT

Overall pre-training and fine-tuning procedures for BERT. Apart from output layers, the same architectures are used in both pre-training and fine-tuning. The same pre-trained model parameters are used to initialize models for different down-stream tasks. During fine-tuning, all parameters are fine-tuned. [CLS] is a special symbol added in front of every input example, and [SEP] is a special separator token (e.g. separating questions/answers).

BERT's model architecture is a multi-layer bidirectional Transformer encoder based on original implementation described by (Vaswani et al., 2017), which is available in the tensor2tensor library (Devlin et al., 2019).

2.2 Sentiment Analysis

Sentiment analysis is a broad phrase that refers to the automatic extraction of valence, emotions, and other affective states from text or speech using computer algorithms. The most typical application of sentiment analysis it the process of automatically identifying the valence of a piece of text, whether it is positive, negative, or neutral, the star rating of a product or movie review, or a real-valued score in a 0 to 1 range that reflects the degree of positivity of a piece of text. However, it can also apply to determining one's attitude toward a specific target or topic in a broader sense. An evaluative assessment, such as positive or negative, or an emotional or affective attitude, such as frustration, joy, rage, grief, enthusiasm, and so on, are examples of attitude.

Sentiment analysis can also refer to the process of assessing an individual's emotional state based on their words (irrespective of whether the text is expressing an attitude towards an outside entity). The term sentiment analysis comes from early research that focused significantly on customer reviews' assessments. Some now refer to the field as emotion analysis because it has expanded to include emotions and feelings (Mohammad, 2016).

2.3 Zakat Issues

Managing zakat is extremely difficult because Zakat institution is responsible to ensure that the zakat fund contributed by the zakat payers is properly managed. There is a growing concern over how the zakat fund was dispersed. This is when social media reports on the abandonment of the poor and needy, which may tarnish the reputation of the zakat institutions (Rusydiana & Al Farisi, 2016).

In Selangor, there were reports of mismanagement of zakat distributions, prompting political parties to intervene. On social media, a member of the public wrote an open letter to the zakat's high authority, challenging the effectiveness of zakat in addressing societal (Sakeenah Anuar, Mohd Alwi, & Ariffin, 2019). Most zakat payers have doubts about the zakat institutions' ability to manage the zakat fund (Ismail & Masturah, 2014).

Furthermore, numerous previous research, such as those by (Sawandi, Abdul Aziz, & Saad, 2017), found that the distribution was inefficient. In contrast, our study found that zakat institutions actually had good zakat allocations to the eight categories of asnaf. The different reports finding show that more research is needed to understand the reasons for public criticism and unhappiness with zakat institution (Sakeenah Anuar et al., 2019).

Therefore, it is very important to resolve the zakat problems because of the usage of social media is now part of everyday life and the viral negative can give huge impact to zakat institution. This research analyse sentiment of public on zakat matters with text classification using BERT model.

3.0 Methodology

The overview of methodology in Figure 2 shows data is collected from social media using crawler with keywords related to zakat. Keywords are identified by using Sketchengine which is an application used by many scholars for their research. Sketchengine is an efficient searching in learner corpora with annotated errors that is built into the corpus querying tool Sketch Engine (Kilgarriff et. al., 2004). The collected data is cleaned into dataset which is trained with BERT model. The training model is used to analyse sentiment on zakat and evaluated with accuracy score.

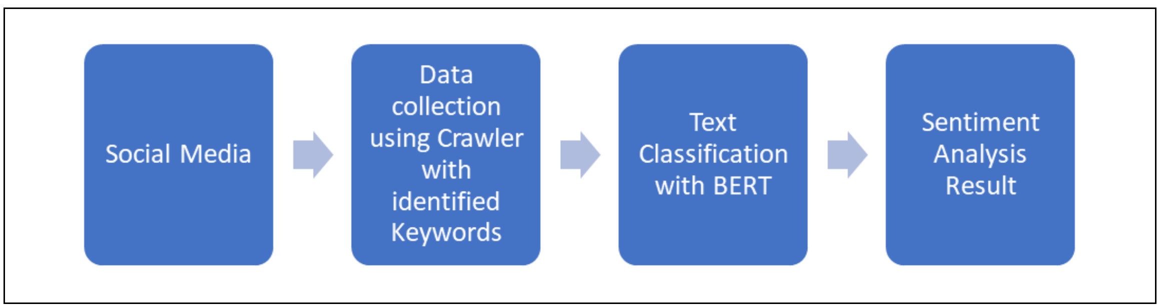


Figure 2: Research Methodology

4.0 Result and Discussion

Data is obtained from the social media using a social media crawler as shown in figure 3. First the data from social media is scraped based on the keywords of zakat. The data that is collected in this stage is in raw format and it requires parsing of the data. It needs to be cleaned from all unwanted symbols and duplicated entries. The cleaned data is structured into database based on the requirements of zakat.

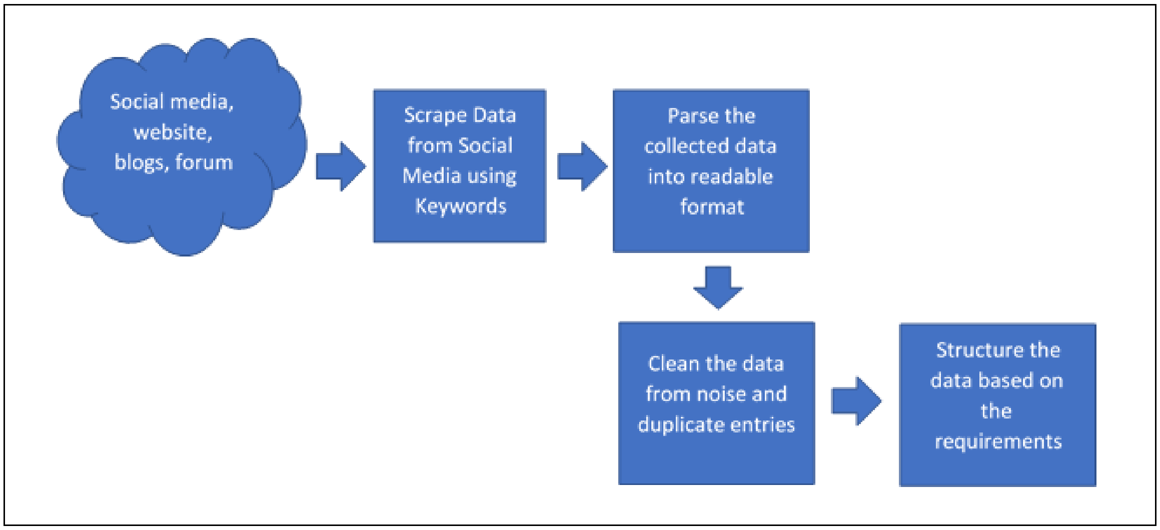


Figure 3: Crawling Process

The data from the crawler process are stored into a database to be trained into zakat model based on text classification. Data from the crawler are classified into two labels as show in Table 1 as follows:

Table 1: Dataset for BERT modelling.

Statement	Sentiment
Zakat sgt cekap membantu kami diwaktu musim banjir dulu. Terima kasih pihak zakat kerana Berjaya melakukan kerja dgn cepat dan cekap.	1
Saya perlukan duit tp susah betul nak daptkan bantuan zakat. Tp nasib baik staf zakat sgt baik dan banyak bantu sy.	0
Tak reti kira zakat jadi malas plak nak bayar zakat.	2

The sentiment for the text is classified into 3 categories which are number 1 is for positive, 0 is neutral and 2 is for negative sentiment.

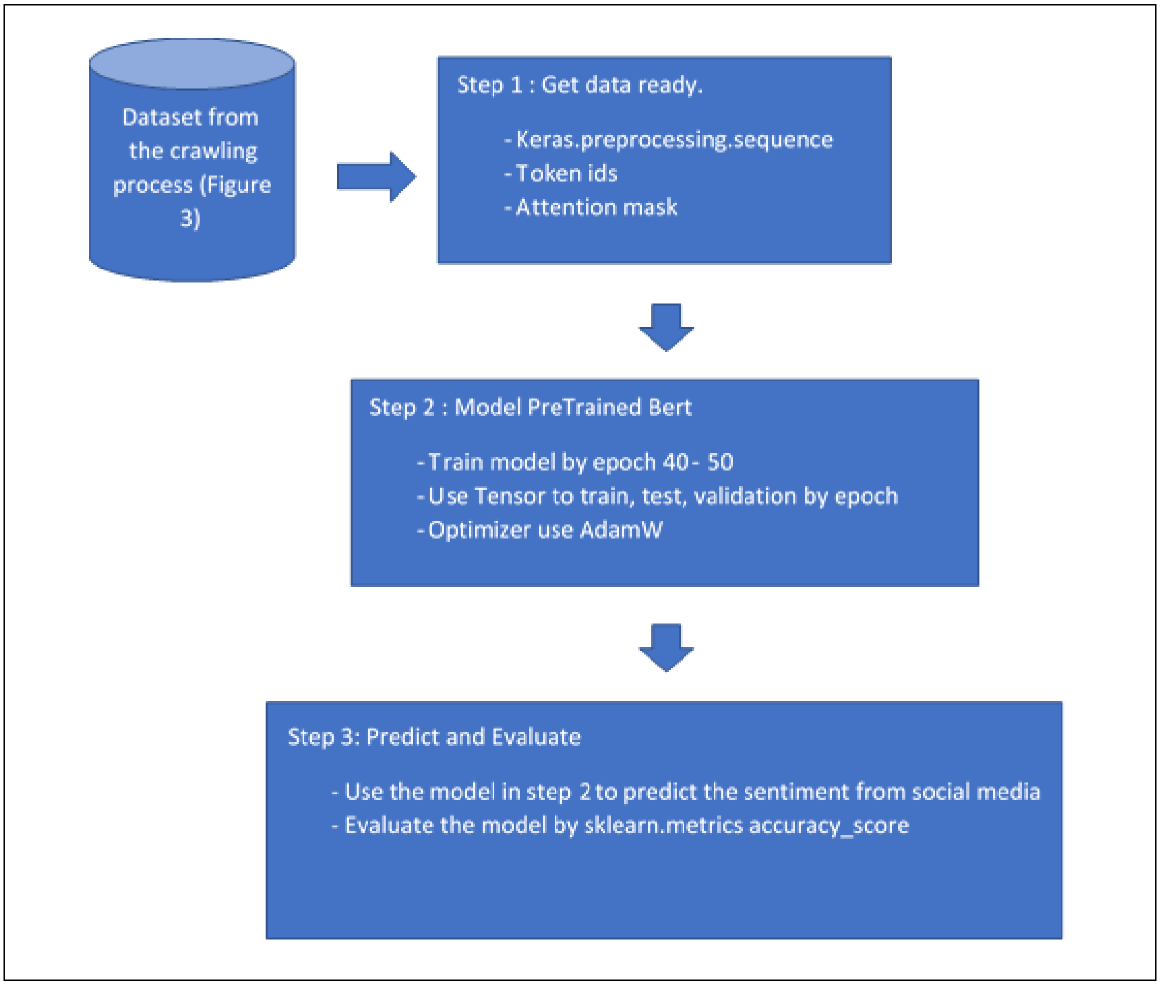
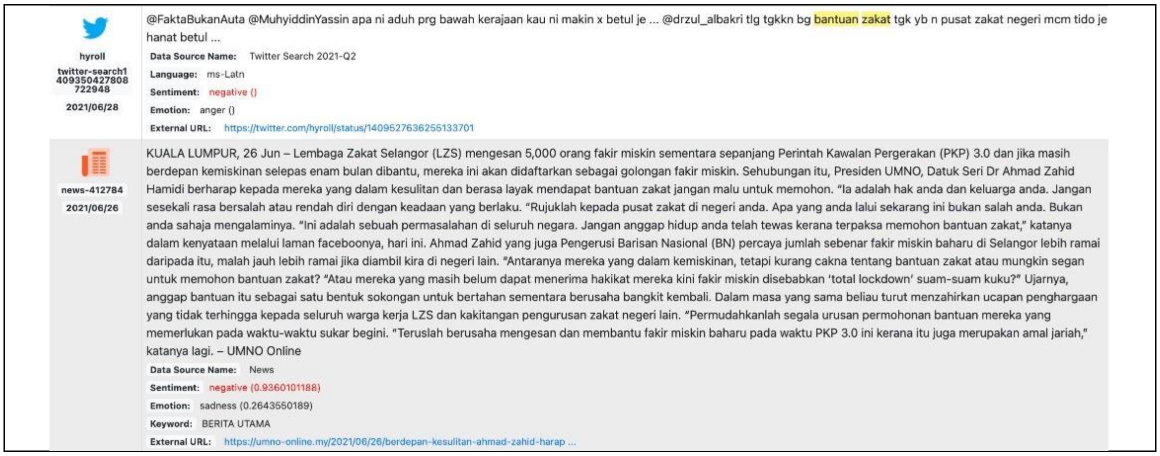


Figure 4: Text classification using BERT

The dataset is prepared for modelling by using keras.preprocessing sequence where statement data is tokenized and converted to token index. Attention mask is used to show the transformer which tokens are padding, placing 0s in the positions of padding tokens and 1s in the position of actual tokens. This is done to perform inference on many samples with batching which will be much faster in processing the tensor. Then data is split into training data and testing data. The training data model is trained by epoch set between 40 to 50. The model is fine-tuned by using optimizer AdamW. The next step is to perform validation to the training model to evaluate the accuracy of the model. The accuracy of the trained model is 96.5%. This prove that the model has high accuracy to predict sentiment on zakat efficiently. Finally, the testing data is tested using the model to predict the sentiment of zakat. The result of the prediction is evaluated by using sklearn metrics to get the accuracy score. The accuracy score obtained from this method is 97% which is a very good and trusted result of sentiment on zakat.

The following sample of sentiment analysis results show the statement from the social media, sentiment and emotion:



hyroll
twitter-search1409350427808722948
2021/06/28

@FaktaBukanAuta @MuhyiddinYassin apa ni aduh prg bawah kerajaan kau ni makin x betul je ... @drzul_albakri tlg tgkkn bg bantuan zakat tgk yb n pusat zakat negeri mcm tido je hanat betul ...

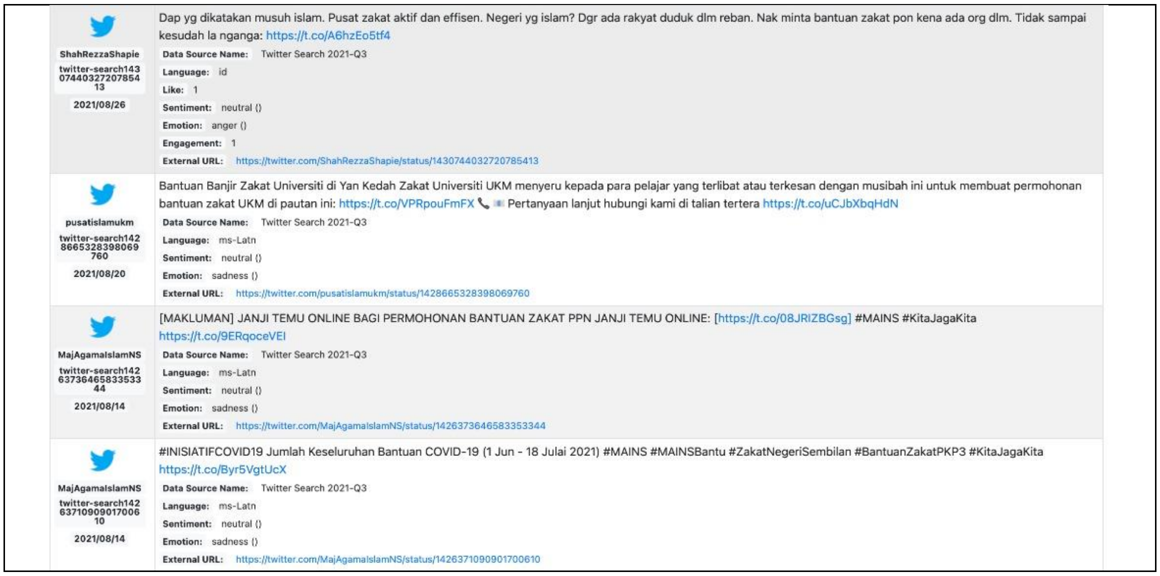
Data Source Name: Twitter Search 2021-Q2
Language: ms-Latn
Sentiment: negative ()
Emotion: anger ()
External URL: https://twitter.com/hyroll/status/1409527636255133701

news-412784
2021/06/26

KUALA LUMPUR, 26 Jun – Lembaga Zakat Selangor (LZS) mengesan 5,000 orang fakir miskin sementara sepanjang Perintah Kawalan Pergerakan (PKP) 3.0 dan jika masih berdepan kemiskinan selepas enam bulan dibantu, mereka ini akan didaftarkan sebagai golongan fakir miskin. Sehubungan itu, Presiden UMNO, Datuk Seri Dr Ahmad Zahid Hamidi berharap kepada mereka yang dalam kesulitan dan berasa layak mendapat bantuan zakat jangan malu untuk memohon. "Ia adalah hak anda dan keluarga anda. Jangan sesekali rasa bersalah atau rendah diri dengan keadaan yang berlaku. "Rujuklah kepada pusat zakat di negeri anda. Apa yang anda lalui sekarang ini bukan salah anda. Bukan anda sahaja mengalaminya. "Ini adalah sebuah permasalahan di seluruh negara. Jangan anggap hidup anda telah tewas kerana terpaksa memohon bantuan zakat," katanya dalam kenyataan melalui laman faceboonya, hari ini. Ahmad Zahid yang juga Pengerusi Barisan Nasional (BN) percaya jumlah sebenar fakir miskin baharu di Selangor lebih ramai daripada itu, malah jauh lebih ramai jika diambil kira di negeri lain. "Antaranya mereka yang dalam kemiskinan, tetapi kurang cakna tentang bantuan zakat atau mungkin segan untuk memohon bantuan zakat? "Atau mereka yang masih belum dapat menerima hakikat mereka kini fakir miskin disebabkan 'total lockdown' suam-suam kuku?" Ujarnya, anggap bantuan itu sebagai satu bentuk sokongan untuk bertahan sementara berusaha bangkit kembali. Dalam masa yang sama beliau turut menzahirkan ucapan penghargaan yang tidak terhingga kepada seluruh warga kerja LZS dan kakitangan pengurusan zakat negeri lain. "Permudahkanlah segala urusan permohonan bantuan mereka yang memerlukan pada waktu-waktu sukar begini. "Teruslah berusaha mengesan dan membantu fakir miskin baharu pada waktu PKP 3.0 ini kerana itu juga merupakan amal jariah," katanya lagi. – UMNO Online

Data Source Name: News
Sentiment: negative (0.9360101188)
Emotion: sadness (0.2643550189)
Keyword: BERITA UTAMA
External URL: https://umno-online.my/2021/06/26/berdepan-kesulitan-ahmad-zahid-harap ...

Figure 5: Negative Sentiment



ShahRezzaShapie
twitter-search1430744032720785413
2021/08/26

Dap yg dikatakan musuh islam. Pusat zakat aktif dan effisen. Negeri yg islam? Dgr ada rakyat duduk dlm reban. Nak minta bantuan zakat pon kena ada org dlm. Tidak sampai kesudah la nganga: https://t.co/A6hzEo5tf4

Data Source Name: Twitter Search 2021-Q3
Language: id
Like: 1
Sentiment: neutral ()
Emotion: anger ()
Engagement: 1
External URL: https://twitter.com/ShahRezzaShapie/status/1430744032720785413

pusatislamukm
twitter-search1428665328398069760
2021/08/20

Bantuan Banjir Zakat Universiti di Yan Kedah Zakat Universiti UKM menyeru kepada para pelajar yang terlibat atau terkesan dengan musibah ini untuk membuat permohonan bantuan zakat UKM di pautan ini: https://t.co/VPRpouFmFX 📞 Pertanyaan lanjut hubungi kami di talian tertera https://t.co/uCJbXbqHdN

Data Source Name: Twitter Search 2021-Q3
Language: ms-Latn
Sentiment: neutral ()
Emotion: sadness ()
External URL: https://twitter.com/pusatislamukm/status/1428665328398069760

MajAgamaIslamNS
twitter-search1426373646583353344
2021/08/14

[MAKLUMAN] JANJI TEMU ONLINE BAGI PERMOHONAN BANTUAN ZAKAT PPN JANJI TEMU ONLINE: [https://t.co/08JRIZBGsg] #MAINS #KitaJagaKita https://t.co/9ERqoceVEI

Data Source Name: Twitter Search 2021-Q3
Language: ms-Latn
Sentiment: neutral ()
Emotion: sadness ()
External URL: https://twitter.com/MajAgamaIslamNS/status/1426373646583353344

MajAgamaIslamNS
twitter-search1426371090901700610
2021/08/14

#INISIATIFCOVID19 Jumlah Keseluruhan Bantuan COVID-19 (1 Jun - 18 Julai 2021) #MAINS #MAINSBantu #ZakatNegeriSembilan #BantuanZakatPKP3 #KitaJagaKita https://t.co/Byr5VgtUcX

Data Source Name: Twitter Search 2021-Q3
Language: ms-Latn
Sentiment: neutral ()
Emotion: sadness ()
External URL: https://twitter.com/MajAgamaIslamNS/status/1426371090901700610

Figure 6: Neutral Sentiment

drzul_albakri
twitter1410416546235275776
2021/07/01

@UncleKiddd Ini pembaharuan inisiatif khusus melibatkan bantuan zakat di WP https://t.co/G2rBWoI77k

Data Source Name: Twitter
Language: in
Sentiment: positive (0.8422526259)
Emotion: happy (0.2926881879)
External URL: https://twitter.com/drzul_albakri/status/1407851109499031558

news-399974
2021/06/24

MAJLIS Agama Islam Wilayah Persekutuan (MAIWP) terus ke depan membantu golongan asnaf yang memerlukan melalui agihan zakat secara langsung dan tidak langsung. Bayangkan lebih RM1 bilion diagihkan sejak Mac 2020 sehingga Mei 2021. Bagi tempoh Mac hingga Disember tahun lalu sebanyak RM732.55 juta zakat diagihkan MAIWP manakala sehingga Mei tahun ini pula RM303 juta zakat. Menteri di Jabatan Perdana Menteri (Hal Ehwal Agama), Datuk Dr Zulkifli Mohamad Al-Bakri, berkata agihan zakat itu membabitkan 30 skim bantuan, inisiatif bantuan perniagaan (Musa'adah Tijariyyah) dan bantuan zakat kecemasan termasuk belanja pengurusan, prasarana dan kemudahan asnaf," katanya mengumumkan pencapaian MAIWP bagi tempoh Mac 2020 hingga Mei lalu dalam hantaran di Facebook, pada Khamis. Zulkifli berkata, MAIWP turut mengadakan beberapa pembaharuan sejak Mac tahun lalu seperti memperkenalkan aplikasi Mobile Jejak Asnaf (MoJA) bagi membantu orang ramai melaporkan kewujudan golongan asnaf. Katanya, aplikasi itu turut dilengkapi teknologi sistem pengesanan lokasi (GPS) bagi memudahkan Skuad Jejak Asnaf Baitulmal MAIWP mengesan golongan asnaf yang memerlukan bantuan di Wilayah Persekutuan. "Perkhidmatan bantuan lain bakal diperkenalkan adalah Projek Foodtruck MAIWP yang akan meyediakan makanan secara percuma kepada 750 golongan asnaf dan yang memerlukan setiap hari untuk tempoh lima hari. "Selain itu, projek Kedai Makan Usahawan MAIWP membabitkan tiga kedai milik usahawan asnaf akan memberi manfaat makanan percuma kepada 300 golongan asnaf dan yang memerlukan selama lima hari seminggu. "Projek Foodrider MAIWP membabitkan 10 kakitangan MAIWP yang ditugaskan menghantar Kit Bantuan Makanan kepada asnaf akan dilaksanakan," katanya. -WILAYAHKU

Data Source Name: News
Sentiment: positive (0.5699940181)
Emotion: happy (0.2435763404)
Keyword: BERITA, UTAMA
External URL: https://www.wilayahku.com.my/maiwp-agih-lebih-rm1-bilion-wang-zakat-se ...

Figure 5: Positive Sentiment

5.0 Conclusion and Future Work

This study has acceptable performance in terms of precision and accuracy with the accuracy more than 90%. The result of zakat sentiment analysis has high accuracy when the model is trained with epoch from 40 to 50. This shows that the higher epoch loop result in higher accuracy score. With this sentiment analysis result, zakat institution can take necessary action to resolve the customer problems on zakat distribution and collection effectively. Future work is to develop implementation framework using BERT model for zakat institution.

References

Ab Rahman, A., Alias, M. H., & Omar, S. M. N. S. (2012). Zakat institution in Malaysia: Problems and issues. Global Journal Al-Thaqafah, 2(1), 35–41. https://doi.org/10.7187/GJAT122012.02.01

Andzulis, J., Panagopoulos, N. G., & Rapp, A. (2012). A review of social media and implications for the sales process. Journal of Personal Selling and Sales Management, 32(3), 305–316. https://doi.org/10.2753/PSS0885-3134320302

Chumwatana, T., & Wongkolkitsilp, K. (2019). Using classification technique for customer relationship management based on Thai social media data. ACM International Conference Proceeding Series, 7–11. https://doi.org/10.1145/3313991.3314010

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1(Mlm), 4171–4186.

Geiser, F. (2017). Social media as a communication channel – Is it possible to build a digital brand and generate revenue streams simultaneously by applying influencer marketing ? (June), 1–47. Retrieved from https://kth.diva-portal.org/smash/get/diva2:1189730/FULLTEXT01.pdf

González-Carvajal, S., & Garrido-Merchán, E. C. (2020). Comparing BERT against traditional machine learning text classification. (Ml). Retrieved from http://arxiv.org/abs/2005.13012

Hamid, A., Raja Lope Ahmad, R.M.T., Azlan, W., Hashim, M., Syakiran, M., Fahmi, M., & Marjudi, S. (2016). Customer Relationship Management (CRM) Implementation Perception and Barriers for Zakat Institution: Case Study Approach. International Journal of Computer Applications, 148(14), 17–21. https://doi.org/10.5120/ijca2016911271

Hamid, A., Raja Lope Ahmad, R. M. T., Majudi, S., Wan Hassan, W. A., Wan Ismail, W. H. B., Setik, R., & Hashim, M. (2019). Social CRM Analytics Model on Malaysian Institutes of Higher Learning (IHLs). TIMES-ICON 2018 - 3rd Technology Innovation Management and Engineering Science International Conference. https://doi.org/10.1109/TIMES-iCON.2018.8621834

Ismail, A., & Masturah, M. (2014). The Efficiency of Zakat Collection and Distribution: Evidence from Two Stage Analysis. Journal of Economic Cooperation and Development, 35(4), 133–170.

Khan, T. (2015). Access to finance and human development — Essays on zakah, awqaf and microfinance: An introduction to the issues and papers. In H A El-Karanshawy et al. (Eds.), Access to finance and human development — Essays on Zakah, Awqaf and microfinance (Vol. 1).

Laroche, M., Habibi, M. R., & Richard, M. O. (2013). To be or not to be in social media: How brand loyalty is affected by social media? International Journal of Information Management, 33(1), 76–82. https://doi.org/10.1016/j.ijinfomgt.2012.07.003

Mohammad, S. M. (2016). Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. Emotion Measurement, 201–237. https://doi.org/10.1016/B978-0-08-100508-8.00009-6

Mostafa, M. M. (2013). More than words: Social networks' text mining for consumer brand sentiments. Expert Systems with Applications, 40(10), 4241–4251. https://doi.org/10.1016/j.eswa.2013.01.019

Nazri, F. A. A., Rahman, R. A., & Omar, N. (2012). Zakat and Poverty Alleviation : Roles of Zakat Institutions in Malaysia. International Journal of Arts and Commerce, 1(7), 61–72.

Rusydiana, A. S., & Al Farisi, S. (2016). The Efficiency of Zakah Institutions Using Data Envelopment Analysis. Al-Iqtishad: Journal of Islamic Economics, 8(2). https://doi.org/10.15408/aiq.v8i2.2876

Sakeenah Anuar, F., Mohd Alwi, N., & Ariffin, N. M. (2019). Financial Management Practices and Performance of Zakat Institutions in Malaysia. Financial Management Practices and Performance of Zakat Institutions in Malaysia IPN Journal of Research and Practice in Public Sector Accounting and Management, 9(1), 1–26.

Sawandi, N., Abdul Aziz, N. M., & Saad, R. A. J. (2017). Zakat accountability practices: Case study of zakat institution AA. International Journal of Economic Research, 14(16), 223–232.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 2017-Decem(Nips), 5999–6009.

Wang, X., Yu, C., & Wei, Y. (2012). Social Media Peer Communication and Impacts on Purchase Intentions: A Consumer Socialization Framework. Journal of Interactive Marketing, 26(4), 198–208. https://doi.org/10.1016/j.intmar.2011.11.004