“iSAY”: Blockchain-based Intelligent Polling System for Legislative Assistance

Deshan Wattegama¹, Pasindu S.Silva², Chamika R. Jayathilake³
Kalana Elapatha⁴, Kavinga Abeywardena⁵, N. Kuruwitaarachchi⁶
Faculty of Computing, Sri Lanka Institute of Information Technology, Sri Lanka¹,²,³,⁴,⁵
Faculty of Technology, University of Sri Jayewardenepura, Sri Lanka⁶

Abstract—“iSAY” is a Blockchain-based polling system created for legislative assistance. Sri Lanka is a democratic country. Country follows a representative democracy and voters in Sri Lanka vote for their preferred government based on their election mandate. However, governments implement legislative decisions that are not stated in the election mandate. People won’t get a chance to state their opinion on this legislative matter and the government also doesn’t know whether people like this or not. To solve this issue, in this paper the authors propose a blockchain-based intelligent polling application for legislative assistance. “iSay” is an application where blockchain technology gets together with machine learning to add value into the public opinion. The government can create a poll about a legislative decision and people can state their opinion which could be further discussed in the legislature. Adding a significant change to the blockchain based e-voting solutions this paper proposes a novel feature where users can add their idea to a relevant poll. Using machine learning algorithms all these user ideas will be classified and analyzed before presenting to the government. Through this research, it is expected to deploy scalable elections among the general public and get their vote and ideas about specific legislations to generate an overview of general public opinion about legislative decisions.

Keywords—Blockchain; machine learning; distributed systems; e-voting; legislative assistance; liquid democracy; natural language processing

I. INTRODUCTION

In Sri Lanka every five years people vote for their preferred political parties and the political party who wins most votes will govern the country. Government will make legislative decisions and most of the time these decisions will be unchallenged due to less opportunities for the general public to state their opinions on these matters. Social activists will protest against these decisions if they think these legislations are not suitable for the society. Then the government will change this legislation where they can meet activists in a middle ground. However, the government cannot know the general public opinion for sure since there are no methods or tools for that. In this COVID-19 pandemic situation the government took legislative decisions which had a clear impact on society. Because of the Social Distancing and New Normality protests were not allowed. Hence social activists didn’t have an opportunity to express their opinion while the government also had no clue about general public opinion. Since “iSAY” blockchain based intelligent legislative polling system solves all these issues this application will be a key application to protect democracy in a post COVID-19 era.

The role played by public opinion in a democracy, particularly as it affects the legislative process, has long been a subject for speculation by political scientists [1]. In Sri Lanka every five years people vote for their preferred political parties and the political party who wins most votes will govern the country. This is called representative democracy and most of the countries follow this [2]. People will look at the mandate of political parties and they will decide which party to vote for. Sometimes governments will implement some Legislative Decisions which were not stated in their election mandates. In the recent past, the Sri Lankan government took a legislative decision to sign an agreement with Millennium Challenging Cooperation of the United States of America [3]. The opposition party debated that voters in Sri Lanka do not want to sign that agreement, however they did not have a way to show this. On the other hand, the government said that this will bring big benefits to the country, nevertheless they also did not have a way to know what people think about this. This is a problem in representative democracy because after election voters do not participate in the legislative decision-making process. Simon Tormey stated “Representative Democracy” as a Contemporary Crisis [4] in Australia Parliament paper no 66. To solve this issue specialists introduced a new kind of democracy called Liquid Democracy. In Liquid Democracy voters actively take part in the decision-making process [5].

As found in the preliminary survey of this research, most of the public want a part in the legislative decision-making process. Therefore, authors Introduced “iSay”, a Blockchain-Based Polling System to fill this void created by representative democracy. There was a survey done in Michigan, on smoking in public places in the consideration of non-smoking legislation to measure the opinion of the public [6]. In the study it is reported that the poll results taken from the public helped to restrict smoking in public places.

This system is far different from current blockchain based e-voting solutions. Current blockchain based e-voting applications support only quantitative votes such as 1 (yes) and 0 (no). Since the main aspect of this application is to present an opportunity to the voters to state their ideas on legislative matters, authors added a novel feature which is adding amendments. Since most of these legislations are
specific to special groups such as Doctors and Engineers, therefore the authors have decided to add a feature called clustered elections.

Since this system targets the general public there will be thousands of users who are adding their ideas as amendments to these polls. Government admins are unable to read these ideas one, therefore the authors are introducing machine learning algorithms to cluster these amendments and analyse these amendments. Amendments and suggestions which are added by the users, will be filtered out and clustered into groups using Natural Language Processing and Machine Learning Algorithms. Moreover, this system facilitates cross platform support.

With all these novel features the government can add a legislative poll about some legislative matter targeting specific crowds and these voters can participate in this poll actively while sitting at home. They can just say they like the legislation or else can add their ideas on that matter.

This application is not a traditional Blockchain based e-voting application. This application is a combination of several areas. Blockchain technology is used to host a secure election where users’ identity and vote is protected. Several Networking technologies have been used by authors to create an election network which will allow users to vote from their own devices. To generalize user amendments authors used Machine Learning and Deep Learning techniques. After every successful poll decision maker will be presented with all the user ideas and poll results which are extracted in a secure environment where voters’ identity is completely protected.

II. LITERATURE SURVEY

Since the arrival of bitcoin [7] Blockchain technology has become a buzzword in the information technology field. Most countries are now trying to convert their manual election system to e-voting systems. Because of its inherently mutable and secure nature, blockchain quickly took over the e-voting applications all around the world.

Satoshi Nakamoto, the founder of bitcoin states, “New transactions are broadcast to all nodes” in his first ever research paper on bitcoin. Bitcoin is used as an electronic currency and when an asset transfer happens the transaction record will be added to a blockchain and it will be distributed among all the nodes [7]. This nature is good for an electronic voting system. Since every user node has their own copy of blockchain hackers cannot tamper with data. Due to that election will be more reliable and transparent. If all the election data stored in a centralized database fault tolerance will be very difficult. As “Blockchain Versus Database: A Critical Analysis” research paper discussed blockchain is using cryptographic measures and it has more advantages than centralized databases [8].

In 2014, Russia used a blockchain based polling system on diverse subjects for an active citizen program [9]. South Korea launched a blockchain based E-voting system for voting purposes in 2017 [10]. This system had no involvement from any kind of administration for the management. With the development of Estonia’s blockchain based E-voting systems since 2007, during the time of 2015 Estonia was able to cast 30% of the parliamentary votes using blockchains voting systems [11]. Estonia used the national ID cards of citizens to verify the voter identity. A Swiss blockchain startup known as Agora launched an E-voting system for Swiss government and institutions in 2018 [12]. An E-Voting system based on public chain was proposed by McCorry in 2017, which was applied on small scale voting scenarios [13]. This system was implemented on Ethereum and used zero knowledge proof. This has been able to support two candidate voting only. Another small scale voting system implemented on Ethereum was introduced by author Zhu where it used bind signatures and ring signatures [14]. This system supported multi-candidate voting. Author Yu proposed an E-voting system based on the alliance chain which was applied on large scale voting scenarios [15]. Based on the alliance chain, this system used both ring signature and homomorphic encryption. Author Jonathan proposed a system called Netvote in 2018 which is implemented using Ethereum [16]. This system proposed three decentralized applications such as, an Admin, Voter and Tally decentralized applications.

Considering the literature survey it is clear that there are a number of blockchain based E-voting systems. However there are no findings of a blockchain based polling/voting application which is specifically targeted for legislative assistance with the enhancement of artificial intelligence for decision making.

III. METHODOLOGY

“iSay” application has only software components. These components are connected to each other in a distributed manner. Fig. 1 above illustrates the main system diagram of the “iSay” application. Admin node and User node will be responsible for creating and holding elections while machine learning components inside the flask server will analyze the amendments given by users.
A. Liquid Democracy based Considerations

In this proposed system authors are eying to achieve Liquid Democracy while helping the government to get a public opinion as well. To achieve this there are few constraints which need to be considered;

1) Every voter should have an equal opportunity to vote.
2) Every vote should be completely validated.
3) Every vote should be protected.
4) This voting system should only allow eligible individuals to vote in an election.

In this system all these constraints are satisfied through using blockchain technology. Key to protecting every one of these constraints is protecting users’ identity. For every user a pair of private and public keys is generated by the system based on user data and credentials.

B. Create an Election using Smart Contracts

Smart contracts are a binary file or a method which takes the blockchain as input parameter. Since blockchain technology is still in its infancy, there is no exact way to use smart contracts. Consequently most proposed election frameworks authors have studied use smart contracts in various different ways. Most of the proposed systems are using “Elections as a Smart Contract” [19]. Since “iSay” e-voting system is based on a blockchain framework implemented by authors, it is required to use a unique smart contract language as well. Exonum, Quorum and Go-Ethereum blockchain frameworks use Rust and Solidity as their smart contract language. Authors deviated from these existing blockchain frameworks mainly because of storage and computing power related constraints. Unlike most of the other e-voting systems authors decided to run smart contracts in each user node when the user is voting. Therefore, it is proposed to make use of Javascript as the smart contract language due to its lightweight and less resource consuming nature.

C. Validating users in an Election

The ECDSA (Elliptic Curve Digital Signature Algorithm) on secp256k1 curve is used to generate the public key. Authors of this research chose ECDSA over RSA (Rivest-Shamir-Adleman) algorithm because ECDSA uses shorter encryption keys that use fewer memory in CPU resources [17]. The ECDSA equation \( E \) over a finite prime number field \( F_p \) can be written as follows:

\[
y^2 = x^3 + ax + b
\]  
(1)

Where

\[ a, b \in F_p, 4a 3 + 27b 2 \neq 0(modp), \] and \( p \) is a randomly selected large prime number.

The recommended parameters for the secp256k1 curve is as follows in Table I, which is extracted from the Recommended Elliptic Curve Domain Parameters [18].

D. Voting in an Election Blockchain Network

To achieve the transparency of the poll blockchain needs to be stored in a decentralized network. By doing so it is possible to achieve basic blockchain qualities as well. To achieve these authors, suggest a peer-to-peer distributed network.

As authors found in the literature review most of the proposed blockchain based election systems have static blockchain nodes. Voters have to go to a specific location to cast their votes. Since this system is specially focusing on legislative matters governments cannot ask their voters to go and cast their vote in that specific location several times a month since these legislations are frequently taken out in parliament.

Solve these issues and authors treat all the users as blockchain nodes. So that they can vote from home with proper authorization.

Left hand side of Fig. 2 illustrates a normal blockchain mostly consisting of static nodes. These nodes have computational power to mine the blocks and store them. However, in the proposed network blockchain cannot be stored in every node. To solve this issue through the proposed network authors have created three types of nodes.

- Signaling Node
- Half nodes
- Peer nodes

| TABLE I. RECOMMENDED ELLIPTIC CURVE DOMAIN PARAMETERS |
|---------------------------------|----------------|
| Parameter | Value |
| \( p \)      | FFFFFFF FFFFFFF FFFFFFF FFFFFFF FFFFFFF FFFFFFF FFFFFFF FFFFFFF FFFFFFF |
| \( a \)      | 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 |
| \( b \)      | 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 |
| \( G \)      | 0279BE667E F9DCCBBAC 55A06295 CE870B07 029BFCDB 2DCE28D9 59F2815B 16F81798 |
| \( n \)      | FFFFFFF FFFFFFF FFFFFFF FFFFFFF FFFFFFF |
| \( h \)      | 01 |

Fig. 2. Blockchain Network Comparison.
Peer nodes are the user nodes which will be voting. In these nodes the blockchain is stored as a hash, hence these can be retrieved in the validation process and recheck the chain. Signaling nodes are the full privilege admin nodes. There will be several signaling nodes with more computational power. These nodes will save the election blockchain. Every other node will be initially connecting to this node when they are restarting. And then these peer nodes will be redirected to another node using Leadership Selection Algorithm. As illustrated in the Fig. 2 every node will act as a client and at the same time as a server.

Admin node will be a privileged user node. It will save the whole blockchain as it is. Giving a novel experience to blockchain based e-voting applications authors propose Clustered Polling. Most of the blockchain based e-voting applications are proposed to handle general elections [19] [20]. Unlike general elections, certain legislative matters may only affect a specific group of participants.

If the legislation added by the government is about farmers, only farmers should be allowed to vote. Otherwise polling results and suggestions will not be accurate as expected. In order to resolve this issue a clustered polling mechanism is used. In the registration process all the voters are classified into clusters using their electorate and profession. Using these clusters government admin can add legislative polls to specific groups. To create this clustered poll authors are using half nodes which were introduced before. When a clustered poll is created they will be disconnected from the main network and they will be connected to a cluster. If this legislation is about farmers the whole cluster consists of farmer nodes.

E. Amendments Adding Feature

If users are satisfied or unsatisfied with the legislation they can just vote ‘yes’ or ‘no’. Or else they want some improvements they can add their opinion on them. However, not all of the opinions are accepted. These suggestions are listed to the other voters and they have to accept them.

F. Amendments Generalization

A main feature of this system is to allow voters to post their own amendment/suggestion to a legislative matter. Since these elections are open to the general public there will be thousands of suggestions and amendments. Anyone whose taking decisions or presenting election results cannot read all the amendments. To solve this issue authors introduced amendments generalization using machine learning algorithms.

A major issue faced by collecting large sets of text data is, it will result in visualizing and managing the voter amendments in an orderly manner. There are amendments which belong to the same groups with similar meaning. Therefore, authors propose a solution of generalizing these amendment text documents using Machine Learning and Deep Learning techniques. This generalization of amendments process will be responsible for partitioning the amendments by taking the inputs as text documents, then cluster into groups and later summarizing the results to give the best generalized output.

| ALGORITHM 1: The Proposed Algorithm to Generalize Amendments |
|------------------------------------------------------------|
| 1. **Input**: Raw amendment text documents                  |
| 2. **Output**: Clusters of summarized amendments              |
|                                                            |
| 3. First step                                               |
| 4. Preprocess raw amendments using NLP (Natural Language    |
|    Processing) pre-processing procedures and return a word  |
|    corpus.                                                   |
| 5. Convert the word corpus into a numerical matrix.         |
| 6: Second step                                              |
| 7. Calculate the maximum silhouette score to determine the  |
|    optimal number of clusters of clustering.                |
| 8. Apply k-means to selected features and create a model.   |
| 9. For each text document                                    |
| 10. Predict the cluster with the model                      |
| 11. Store cluster text documents.                           |
| 12: Third step                                              |
| 13. Calculate normalized word frequencies of text documents.|
| 14. Calculate the sentence score using normalized word      |
|    frequencies.                                              |
| 15. Select 30% of sentences with large sentence frequency.  |
| 16. Return final clustered and summarized text.             |

\[ JV= \sum_{i=1}^{c} \sum_{j=1}^{c} || x_i - v_j ||^2 \]  

Where,

\[ ||x_i - v_j|| \] is the Euclidean distance between \( x \) and \( v \)'s points, \( c \) is the number of data points in \( i \) cluster, and \( c' \) is the number of cluster centers. Since this is an automated process, the real amount of clusters cannot be identified. Therefore this system ran the k-means algorithm for multiple times and created multiple models. Out of all these models, the one with the best Silhouette score is chosen. This chosen k-means model is then stored in the server and the respective cluster for the text document will be predicted.

After the prediction of all these cluster documents, the cluster documents will be combined into single documents. In order to reduce the content of the documents a text document summarization is performed. This is an extractive summarization procedure and the python library Spacy is used. The steps taken for the summarization is mentioned in the above Algorithm 1.

IV. RESULTS AND DISCUSSION

Most of the components of this application are working individually as separate components. Since the rise of Blockchain Technology it has proven that Blockchains are more secure than centralized database systems. However, there were some issues in current blockchain frameworks. One of the main issues was all these frameworks are designed for general purposes. Hence most of them are resource consuming.
Therefore, a new framework is proposed called “iZigma” which is implemented using JavaScript for this system, with multiple chain support and giving freedom for the developer to handle the network according to the system requirement. The framework is created as a node module which allows to use the framework easily in a node project. A platform that gives full customization to the blockchain is needed due to the following two features. In the Table II is a comparison between the proposed framework and the other blockchain frameworks. Fig. 3 illustrates the blockchain framework performance test results.

1) The system is designed to have several elections at the same time and each election uses a separate blockchain.
2) Network in this system is divided into clusters.

TABLE II. COMPARISON BETWEEN PROPOSED FRAMEWORK AND OTHER BLOCKCHAIN FRAMEWORKS

|                     | Hyperledger Fabric | Go-Ethereum                               | iZigma                      |
|---------------------|--------------------|------------------------------------------|-----------------------------|
| Consensus           | PBFT               | PoW (Proof of Work), PoS (Proof of Stake), PoA (Proof of Authority) | PoW                         |
| Transactions p/s    | 3,500 transactions p/s | Depends                                  | Depends (Can be configured) |
| Private Support     | Yes                | Yes                                      | Yes                         |
| Smart Contract      | Go, Java, Kotlin   | Solidity                                 | JavaScript                  |
| Language            |                    |                                          |                             |
| Programming Language| Go, JavaScript, Python | Go, C, JavaScript                        | JavaScript                  |
| Multiple Chains     | No                 | No                                       | Yes                         |

Using this newly created blockchain framework authors were able to create and deploy multiple elections which can run concurrently. As illustrated in Fig. 3 these elections were less resource consuming and more scalable than most of the existing frameworks.

The file size of the blockchain file after doing this process was 2.45 KB. This new Blockchain network was able to do the blocks mining in an improved pace aiding the application to generate real time results.

Using Blockchain the authors were able to hold secure elections. If election results were corrupted by changing a block authors were created validation methods to identify these external changes instantly. If the blockchain was identified as invalid the main application will ask the half node copy of that blockchain and it will be replaced.

Networking part of this blockchain based voting system is important. Here authors faced challenges due to the large number of users connecting concurrently to the application since this application is opened for the general public. To solve this problem, authors propose a distributed network where nodes are loosely coupled.

User nodes can connect and disconnect to the main server anytime they want. Using this approach, blockchain concepts can be protected while allowing thousands of users to connect to the network and vote. Sri Lanka has a population of around 20 million. This approach helps to facilitate a large number of nodes as well. All the nodes will not connect to the main server. Therefore the load to the signalling server is reduced preventing single point of failure due to large amounts of requests. Only oldest nodes will be connected to the signalling server. User nodes are connected with each other and they will contact the signalling server through the oldest nodes in the network.

Fig. 4 represents an image taken from an actual election network. It is clear that all the nodes are creating real time socket connections. Peer node users can install this application and it will run as a local node server. Since authors cannot access the local process of a machine while in the outside world, authors use ngrok [21] to tunnel these servers with one another. Ngrok creates a random URL each time the node is restarting. This random URL generation will protect the identity of the clients.

![Image of blockchain performance results](image-url)
Voters can connect to this network with proper authentication and add their vote on specific elections. These elections can be clustered elections where the election targets a specific group of people.

When users add their amendments on legislations these amendments will be run through deep learning and machine learning algorithms. This machine learning component is a separate component where this can be used for general purpose as well. Using this amendment generalization component user amendments will be categorized and will be presented with the final election results.

In the amendment generalization as mentioned in the methodology section, authors are using two methods for the numerical representation of text documents. For testing the accuracy of the model authors used live responses from citizens as amendments and suggestions regarding a specific legislative matter. By using the TF-IDF a Silhouette score of 0.012434112109754505 was obtained. With the model created with Doc2Vec, authors were able to calculate a Silhouette score of 0.24956775. Fig. 5 illustrates the Silhouette coefficient against the cluster labels plot for 4 clusters.

![Fig. 5. Silhouette Coefficient Plot.](image)

Therefore, it’s clear that Doc2Vec can generate a more accurate k-means model from text documents since Doc2Vec more focuses on the semantic level of the documents [22]. In the proposed extractive text document summarization, only 30% of the sentences from the paragraph with the highest frequency score is taken. Since this is not an abstractive text summarization, it does not focus on the semantic level of the summary. Therefore authors suggest a text summarization implemented using a RNN (Recurrent Neural Network).

V. CONCLUSION
Due to the current COVID-19 pandemic situation “iSay” application now has a value more than ever. In the recent past the general public participated in protests against some legislative decisions. However, the government had no idea about general public opinion about that legislation before presenting that into the parliament. With the current situation there is no chance for protesting and with the “iSay” application both parties, the general public and government can gain benefits since both can express their ideas using this application.

This paper proposes a polling system which will allow the general public to be a part of the legislative decision making process. This polling application will help both government and opposition parties as well. This application will enact liquid democracy while solving the problems of representative democracy. Authors were able to create a stable election network, validate votes real time and give election results real time using the newly proposed blockchain framework and distributed network. As a value addition to the e-voting domain. Authors were able to classify suggestions given by users and run a sentiment analysis on these suggestions.

REFERENCES
[1] Cantwell, F., 1946. Public Opinion and the Legislative Process. American Political Science Review, 40.
[2] S. Gamage, “Democracy in Sri Lanka: past, present and future,” Asian Studies Review, vol. 17, no. 1, 1993.
[3] “Sri Lanka Compact,” Millennium Challenge Corporation. [Online]. Available: https://www.mcc.gov/where-we-work/program/sri-lanka-compact.
[4] “Papers on Parliament no. 66 - Parliament of Australia.” [Online]. Available: https://www.aph.gov.au/About_Parliament/Senate/Powers_practice_n_procedures/pwpapers_on_parliament_66/The_Contemporary_Crisis_of_Representative_Democracy.
[5] Blum, Christian & Zuber, Christina. (2015). Liquid Democracy: Potentials, Problems, and Perspectives. Journal of Political Philosophy. 24. 10.1111/jopp.12065.
[6] PerlHadt, H. and Holmes, R., 1987. The role of public opinion polling in health legislation. American Journal of Public Health, 77.
[7] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf.
[8] Chowdhury, Mohammad & Colman, Alan & Kabir, Asad & Han, Jun & Sarda, Paul. (2018). Blockchain Versus Database: A Critical Analysis. 1348-1353. 10.1109/TrustComBigDataSE.2018.00186.
[9] M. Hochstein, “Moscow’s Blockchain Voting Platform Adds Service for High-Rise Neighbors,” CoinDesk, 15 Mar. 2018; https://www.coindesk.com/moscow-blockchain-voting-platformadds-service-for-high-rise-neighbors, 2018.
[10] “South Korea Uses Blockchain Technology for Elections,” KryptoMoney, https://kryptomoney.com/south-korea-uses-blockchain-technology-for-elections, 2017.
[11] Andrew Barnes, Christopher Brake and Thomas Perry, “Digital Voting with the use of Blockchain Technology”, https://www.economist.com/sites/default/files/plymouth.pdf, 2016.
[12] Agora: Bringing our voting systems into the 21st century Available at: https://agora.vote/Agora_Whitepaper_v0.1.pdf, 2017.
[13] McCorry, P., Shahandashti, S.F., Hao, F.: A smart contract for boardroom voting with maximum voter privacy. In: International Conference on Financial Cryptography and Data Security. pp. 357–375. Springer (2017).
[14] Zhu, Y., Zeng, Z., Lv, C.: Anonymous voting scheme for boardroom with blockchain. International Journal of Perforrnability Engineering 14(10), 2414–2422 (2018).
[15] Yu, B., Liu, J.K., Sakzad, A., Nepal, S., Steinfeld, R., Rimba, P., Au, M.H.: Platform-independent secure blockchain-based voting system. In: International Conference on Information Security. pp. 369–386. Springer (2018).
[16] Jonathan Alexander, Steven Landers and Ben Howerton, “Netvote: A Decentralized Voting Network”, https://netvote.io/wpcontent/uploads/2018/02/Netvote-White-Paper-v7.pdf, 2018.
[17] Mahto, Dindayal & YADAV, DILIP. (2017). RSA and ECC: A comparative analysis, International Journal of Applied Engineering Research, 12. 9053-9061.
[18] Secg.org, 2020. [Online]. Available: http://secg.org/sec2-v2.pdf.
[19] F. P. Hjalmarsson, G. K. Hreioarsson, M. Hamdaqa, and G. Hjalmysson, “Blockchain-Based E-Voting System,” 2018 IEEE 11th International Conference on Cloud Computing (CLOUD), 2018.

[20] R. Bulut, A. Kantarci, S. Keskin, and S. Bahtiyar, “Blockchain-Based Electronic Voting System for Elections in Turkey,” 2019 4th International Conference on Computer Science and Engineering (UBMK), 2019.

[21] "Ngrok - secure introspectable tunnels to localhost", Ngrok.com, 2020. [Online]. Available: https://ngrok.com/.

[22] D. Kim, D. Seo, S. Cho and P. Kang, "Multi-co-training for document classification using various document representations: TF-IDF, LDA, and Doc2Vec"; Information Sciences, vol. 477, pp. 15-29, 2019. Available: 10.1016/j.ins.2018.10.006.