

Prediction of ISO 9001:2015 Audit Reports According to its Major Clauses using Recurrent Neural Networks

Ken Jon M. Tarnate, Madhavi Devaraj

Check for updates

Abstract: The Quality Assurance Department of the educational sectors is rapidly generating digital documents. The continuous increase of digital documents may become a risk and challenge in the future. Interpreting and analyzing those digital data in a short period of time is very critical and crucial for the top management to support their decisions. By this purpose, this paper explored the possibility of machine learning and data mining process to improve the Quality Assurance Management System process, specifically in the Quality Audit procedures and generation of management reports. The researchers developed a machine learning model that predicts an audit report according to the major clauses of the ISO 9001:2015 Quality Management System (QMS) Requirements. The proposed data mining process helps the top management to identify which principles of the ISO 9001:2015 QMS Requirements they are lacking. The authors used four different Recurrent Neural Networks (RNNs) as a classifier; (1) Long Short Term-Memory (LSTM), (2) Bidirectional-LSTM, (3) Deep-LSTM and a (4) Deep-Bidirectional-LSTM Recurrent Neural Networks with a combine word representation models (word encoding plus an embedding dimension layer). The Deep-Bidirectional-LSTM outperformed the other three RNN models. Where it achieved an average classification accuracy of 91.10%

Index Terms: Quality Assurance, Quality Management System (QMS), Recurrent Neural Networks, Text Classification

I. INTRODUCTION

The quality management system (QMS) is considered as an organizational structured that consists of quality procedures, processes and resources which is setup systematically to achieve quality objectives. Today, State Universities and Colleges decided to implement the principles and methodologies of the quality management system for two main reasons. First, is to improve their operations and increase customer's satisfactions. Second is their wish for formal international recognition to drive corporate reputation. [1] In this paper, the researchers identified the current challenges of the quality assurance department of the academic sectors. First, their quality assurance staff is doing manual classification of digital documents such as audit findings and reports. This normally results to the delay of data analysis and late submission of management reports to the top

Revised Manuscript Received on 30 July 2019.

* Correspondence Author

Ken Jon M. Tarnate, Mapua University, Manila, Philippines Madhavi Devaraj, Mapua University, Manila, Philippines

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

management. In addition, it creates a bottleneck for a large data to be analyzed in a short period of time and this becomes a challenge for the administrative sectors. Their main tasks are to create reports punctually and provide adequate and latest data for the top management to make critical decisions. Second, most auditors during quality audits and management reviews had difficulty on identifying and categorizing the audit findings according to its major clauses of the ISO 9001:2015 QMS Requirements. And this is where automatic text categorization fits to solve the problem of the quality assurance auditing team of the academic institutions. However, the text categorization process is also facing a challenge, in terms of picking the right "word representation model" for the traditional machine learning algorithms. [15] For this purposed, the researchers developed a deep neural networks models with a combine word representation models (word encoding plus an embedding dimension layer) to classify an audit reports according to its major clauses of the ISO 9001:2015 QMS Requirement. The authors built a two layered LSTM and Bi-LSTM neural networks with a total number of 225 hidden units and compare the performance of those models to the traditional LSTM and Bi-LSTM models. The main contributions of this paper are the development of machine learning model to help the internal auditors to identify and classify an audit reports according to the major clause of the ISO 9001:2015 QMS Requirements. Second, the researchers investigated the learning ability of the four recurrent neural network models using a combined word representation models (word encoding with an embedding dimension layer). And most importantly, the authors able to speed up the process of the extractions and generations of management reports based from the results of the internal and external quality audits conducted in the university.

II. LITERATURE REVIEW

A. ISO 9001:2015 Quality Management System (QMS) Requirements

The quality assurance department of the higher education was assigned to adopt and implement the principles and methodologies of the ISO 9001 Series of quality management systems (QMS) to the institution. [2], [3]. This international standard has basically ten (10) major clauses; (1) Scope, (2) Normative References, (3) Terms and Definition, (4) Context of the organization, (5) Leadership, (6) Planning, (7) Support,



(8) Operation, (9) Performance Evaluation and (10) Improvement. During an actual quality audit, the internal and external auditors will use this international standard as their guide for auditing a management process. Eventually, the auditors will start their audit from clause four (4) to clause ten (10). [4]

B. Application of Data Mining and Deep Learning Models to Support the Quality Management Systems

The possibility of data analytics and data mining in educational management institutions was discussed by Agasisti and Bowers in 2017, their paper entitled "Data Analytics and Decision-Making in Education." indicated that the use of data mining process and analytics can improve the decision-making of the top management of the educational institutions. [5] Guans, Sun, Wang & Zheng (2016) built a machine learning models to support the management activities of their quality assurance department. Their task includes event and short text classification they also extracted valuable information on the quality assurance databases. They analyzed it and gain meaningful insights to improve the nuclear quality assurance management system process. [6] Pandey et al. (2017) used machine learning for automated classification of software issues reports. They worked makes a great impact on the improvement of software quality assurance field. They used traditional machine learning models to classify the reports such as Naïve Bayes, K-NN, Support Vector Machines (SVM) and Random Forest. [7] In this paper, the authors used the deep learning models because of its recent advancement to solve classification problems such as text categorization.

C. Neural Networks and Deep Learning for Text Categorization

Text categorization is a process of classifying text documents under a predefined set of classes. Basically, there are two types of text classification; first is the "single label classification" it is considered as a single label if there is only one class assigned in the document. The best example of this single label is the "binary classification" which predicts if a certain document belongs to a particular class or not. And second type of text classification, is called the "multi-label classification" it is considered as a multi-label if there are two or more classes assigned in a document. [8] And today, neural networks and deep learning models create major advances in the natural language processing (NLP) field such as text categorization. HD Wehle (2017) defined deep learning as a form of machine learning that can be either utilized by a supervised or unsupervised learning or both. [8] Recently, the success of deep learning models in the image classification have attracted considerable attentions to used it in the text classification problem. [9] In 2014 Yoon Kim, used the Convolutional Neural Network (CNN) to classify sentences. He used the pre-trained word2vec, a type of word embedding text representation model developed by Mikolov et al. in 2013. Yoon Kim conducted a series of experiments with the CNN model, and he established concrete evidence that unsupervised pre-training of word vectors shows an important ingredient in the application of deep learning algorithms for the natural language processing problems. [10] In 2016, Liu,

Qui & Huang used the Recurrent Neural Networks (RNN) model for text classification with multi-task learning. They used a word embeddings word representation to convert text data into a set of vectors then fed it to their LSTM model. In their work they have investigated the effect of word representation models on the neural networks. They concluded that the performance of the neural networks model can be improved by exploring and choosing common text features.[11] Nowak, Taspinar & Scherer (2017) used the traditional LSTM and Bi-LSTM for short text and sentiment classification of spam base, farm advertisement and amazons' book reviews. In their experiments the LSTM and Bi-LSTM models were almost par with the Gated Recurrent Unit and AdaBoosting model. This only proves that the Recurrent Neural Network models shows significant used for the text categorization problem even though it was mostly used for time series forecasting problems. [12] Shih, Yan, Liu & Chen (2017) also used the LSTM and Bi-LSTM to classify IMDB and 20Newsgroup datasets. They even created their own word representation model based from LSTM networks they called it "Siamese LSTM" their state-of-the-art model outperforms the other traditional machine learning models and almost par with Bi-LSTM and even in the Convolutional Neural Network Model. [13] Finally, last 2018, Radhika, Bindu & Parameswaran used the recurrent neural networks and CNN model to classify essays according to its author. Their dataset consists of more than 2,000 essays with different authors. According to their experiment, RNN model outperformed the CNN model. They evaluated their models using maximum cross entropy loss and classification accuracy. [14]

D. Word Representation

In the machine learning process, the process of converting unstructured data to structured data is called "text or word representation" this is mandatory for all machine learning process, in order for the computer to process the data in a computable format. [15], [23], [24]. It was very important to have an effective word representation model to build an efficient machine learning classification system. [16], [25]. Currently, a lot of word representation models existed, the most commonly known is the Bag-Of-Words, this model learns from the vocabulary of the corpus or the whole documents. It counts the number of times the words appear in the documents. It is the simplest text representation where the frequency of the occurrence of each word is used to train the classifier. [17] However, this word representation model has some limitations it does not consider the order of the sentences and semantic relationship of the words in the documents. This is the main reason why researchers are continuously developing and exploring new techniques to represents unstructured data to be mathematically computable. And recently, deep neural networks have been used to improve and develop new approaches to convert unstructured data to a suitable data format for machine learning. [15], [16], [18]

eering Exploring Innovation



E. Long Short Term-Memory (LSTM) and Bidirectional-LSTM Recurrent Neural Networks

The standard Recurrent Neural Networks (RNN) has a problem to learn long-term dependencies; this is called "the vanishing gradient decent" during training, the gradient becomes less, and the weights becomes harder to updates, that results to a longer training. [19], [22], [24]. To address this issue, LSTM and Bi-LSTM architecture has been introduced they are able to store and propagate historical information through chains of neural networks which preserved the information and updates the weights easily. It is best suitable for sequential data such as video frames, images and sequence of text and words [20], [21], [24]

III. METHODOLOGY

In the Exploratory Data Analysis, the researchers used the RStudio IDE version 1.1463. Using the "R" programming language and libraries such as "ggplot2", "dplyr", "quanteda", "caret" and "doSnow" using these methods the authors able to analyze and graph the dataset. in a sophisticated way using the visualization techniques in data mining. Meanwhile, the researchers also used the MATLAB Software version 2018b in doing the entire standard procedures of the text categorization process of audit reports according to the ISO 9001:2015 QMS Requirements.

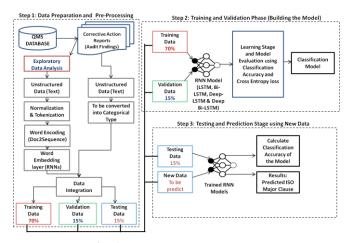


Fig. 1: Research Process Model

This figure shows the visual representation of the categorization process of audit reports; Data gathering and extraction, Exploratory Data Analysis, Data Pre-processing, Training, Validation and Testing the RNN(s) Model and Evaluation of RNN Models.

A. Data gathering and Extraction

The datasets were extracted from the Quality Assurance Databases on one of the State University in the Philippines which their quality management system is ISO certified. The entire datasets were composed of 2583 audit reports. Each audit reports have a labeled according to its major clauses and specific clauses, status and its types of nonconformities; major nonconformance (NC), minor (MiN), and opportunity for improvement (OFI).

B. Exploratory Data Analysis on the Datasets

Using the RStudio IDE version 1.1.463. The authors explored the data of the quality assurance department of the higher education institutions. Using the data visualization techniques, it helped the top management to foresee, on which principles and methodologies of the ISO 9001:2015 QMS Requirements they are lacking with. Through this exploratory data analysis, the top management of the institution had been able to identify which areas of the management are needed to be prioritized.

Table I: Breakdown of Audit Reports Per Major Clause of ISO 9001:2015 QMS Requirements

Major Clauses	Actual Count	Percentage
Content of the Organization	294	11%
Leadership	144	6%
Planning	255	10%
Support	918	35%
Operations	480	19%
Performance Evaluation	444	17%
Improvement	48	2%

Table I. Shows the breakdown of the entire datasets according to its major clauses. Basically, during quality audits, the auditors are looking for this major clause of the ISO 9001:2015 QMS Requirements. It shows that the area of "Support" has the greatest number of findings. This means that the current Quality Management System of the institution is lacking on the "Support Systems", this also means, that their current weakness is in the area of support groups and programs. However, the lowest number of findings can be found in the area of "Improvement" this can be means that the current QMS of the institution, is stable enough to continue on the day to day operations.

C. Data Preparations: Text Processing and Word Representation

In this stage, the researchers cleaned and normalized the text data and converted it into a computable format so that the deep learning model can process and train by it. The researchers tokenized first the documents, then removed all the stopwords, punctuations and lowercase all the characters. Then converted the unstructured text data into a structured data; this can be done by using a word representation model. The authors used the "doc2sequence" n-gram model to convert the text data into vectors. Then reduced the dimensionality of the data by assigning a fixed number or length of attributes through word embedding.

D. Modeling

After the data preparations, the researchers divided the datasets into to three states, 70% for the training of the model, 15% for the validation of the model and another 15% for the testing of RNN(s) the model. Then the authors fed the data in

the recurrent neural networks model to train and learn from it.



Categorization of ISO 9001:2015 Audit Reports Using Recurrent Neural Networks

In the experiment a fully dense connected network was used and each RNN models have 225 hidden units' layers and have an embedding dimension of 100. The authors used the "Adam" optimization solver to iteratively update the network weights during training then used the "SoftMax" activation functions to normalize the input values to outputs vectors and assigned probability distributions on the output values. And to avoid over fitting on the training data, the researchers used the "dropout" regularization technique for the Deep-LSTM and Deep-Bidirectional-LSTM.

IV. RESULTS AND DISCUSSIONS

Fig. 2: Sample R Code for Automated Process of Generating a Management Reports based from the Internal and External Quality Audits.

This figure shows the actual "R" code to generate a certain report needed by the quality assurance department. This code able to summarize the results of the internal and external quality audits findings. It also able to graphs and calculates the average nonconformities and status of the findings.

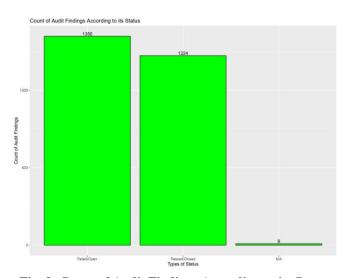


Fig. 3: Count of Audit Findings According to its Status.

In this figure, based from automated extraction process of audit reports, there are 1,350 or 52% of audit findings are still open or not resolve, however there are 1224 or 47% of audit findings are already resolved. Meanwhile there nine (9) or 1% of audit findings is undetermined.

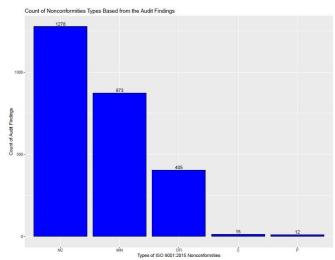


Fig. 4: Count of Nonconformities Types Based from the Quality Audit Findings.

In this figure, based from the data mining process, the system identified the count of nonconformities of the entire university on the ISO 9001 Standard. According to these results, there are 1278 processes that are major nonconformance on the ISO 9001 QMS standard and 873 minor findings. However, there are 405 findings which can be a good opportunity to improve for the entire QMS process of the university. Lastly, some auditors submitted 15 compliance and 12 positive audit findings. It indicates that there are existing processes that needs to be maintained and retained because it causes positive impact on the university.

A. Evaluation Metrics

The authors evaluated the model by getting its classification accuracy with respect to its cross-entropy loss.

 $Accuracy = \frac{Number of correct predictions}{Total number of predictions}$

Fig. 5. Formula of Classification Accuracy

Table II. Performance of the RNN(s) Models

Model	Classification Accuracy During Validation Stage	Classification Accuracy During Testing Stage	Average Classification Accuracy
LSTM	96.90%	83.5%	90.2%
Deep-LSTM	92.51%	81.9%	87.20%
Bidirectional - LSTM	94.32%	84.2%	89.26%
Deep- Bidirectional - LSTM	92.51%	89.7%	91.10%

Table II. Shows the classification accuracy of each RNN(s) Model. Based from these results, the Deep-Bidirectional LSTM outperformed the other three RNN(s) model. However, the plain LSTM model outperformed the two deep neural networks. Based from these results, the authors conclude that the adding of another hidden layers is not

necessarily increases the accuracy of the neural networks model.





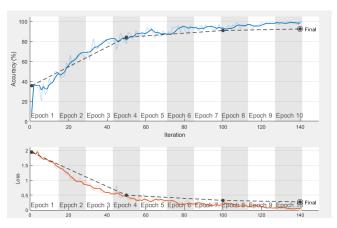


Fig. 7: Accuracy and Cross-Entropy Loss of the Best RNN Model based from the experiments (Deep-Bidirectional-LSTM)

		Confusion Matrix							
	Context of the Organization	37 9.6%	0 0.0%	2 0.5%	2 0.5%	2 0.5%	3 0.8%	4 1.0%	74.0% 26.0%
	Improvement	0 0.0%	5 1.3%	0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
co.	Leadership	0 0.0%	2 0.5%	13 3.4%	0 0.0%	0 0.0%	2 0.5%	2 0.5%	68.4% 31.6%
Output Class	Operation	0 0.0%	0 0.0%	3 0.8%	69 17.8%	2 0.5%	1 0.3%	0 0.0%	92.0% 8.0%
Jutpul	Performance Evaluation	2 0.5%	1 0.3%	1 0.3%	0.0%	61 15.8%	0 0.0%	1 0.3%	92.4% 7.6%
Ü	Planning	0 0.0%	0.0%	0.0%	0.0%	1 0.3%	32 8.3%	0 0.0%	97.0% 3.0%
	Support	5 1.3%	0 0.0%	2 0.5%	2 0.5%	0 0.0%	0 0.0%	130 33.6%	93.5% 6.5%
			37.5%	61.9% 38.1%	94.5% 5.5%	92.4% 7.6%	84.2% 15.8%		89.7% 10.3%
Corner of the Constitution and a state of the state of th									
Congraduti. Safething									
	Target Class								

Fig. 8: Confusion Matrix of the Best RNN Model based from the experiments. (Deep-Bidirectional-LSTM)

Figure 7 and Figure 8. Shows the Classification Accuracy (CA) and Cross Entropy (CE) Loss of the Deep-Bidirectional LSTM Model, based from the experiments, the Deep-Bidirectional LSTM model outperformed the other RNN Models. Where it achieved an average Classification Accuracy of of 91.10 % and has a Cross-Entropy Loss of 1.59 %

Table III. Sample Predictions and Results

Table 111. Sample I redictions and Kesuits						
Audit Findings	Predicted	Actual Major Clause				
	Clause	Clause				
No CAR	Improvement	Improvement				
No QMS-indexed Forms	Support	Support				
No CSF Result*	Performance	Leadership				
110 CSI Result	Evaluation*	Leadership				
No QMS internal audit results for Q3 FY 2018	Performance Evaluation	Performance Evaluation				
Audit Findings	Predicted	Actual Major				
ō	Clause	Clause				
No IPR	Context of the	Context of the				
Nonk	Organization	Organization				
No Action plan to achieve						
quality objectives when	Planning	Planning				
they are not met						

No definite date of				
issuance of Transcript of				
Records after graduation		Leadership		
of students without	Leadership			
bearing Board				
Resolutions numbers.				
This violates RA 11032				
Ease of doing Business				
No 5s technical audit				
results for Q3 FY 2018	Omenation	Operation		
due to unavailability of	Operation			
the process owner.				
No accomplished NAP	C	Support		
form 1.	Support			
The storage room of the				
P.E, department is not	Support	Support		
properly maintained	_ ~	_ <u></u>		

Table III. Shows the actual predictions of the Deep-Bidirectional LSTM. The model predicted nine (9) correct major clauses out of 10 trials.

V. CONCLUSIONS AND FUTURE WORKS

In general, the proposed deep learning model helps the internal auditors of the university. They used the machine learning model as a cross reference and as a validation tool for labelling an audit findings or reports. In addition, the proposed data mining and machine learning framework helps the administrative staffs of the quality assurance department to generate reports in a short period of time. And the top management had been able to receive an updated and adequate data where their can based and support their decisions from the results of the data mining process. Lastly, the Deep-Bidirectional LSTM outperformed the other three RNN(s) model having an average classification accuracy of 91.10% However, during the validation stage; the traditional LSTM has the highest validation accuracy and outperformed the two layered Deep-LSTM Model having an average classification accuracy of 90.2%. This signifies that the adding a neural network layer has no significant factor in the increase of the accuracy of the deep neural networks. This is also due to the vanishing gradient descent problem of the deep neural networks. For further studies, researchers should also consider and look on this problem when using a deep learning model for a classification task. By this means, as advice for the future researchers, they should carefully examined the hyper parameters, dropouts and hidden units before adding a neural network layer when building a deep neural networks model.

ACKNOWLEDGMENT

This research was accomplished through the grace and help of our Lord Jesus Christ. All praises belongs to God. The authors would like express their gratitude's to the Quality Assurance Department of the Technological University of the Philippines (TUP) for their approval to use their ISO 9001:2015 audit reports as a dataset for this research and special thanks to Dr. Ralph Sherwin A. Corpuz, for giving us access to the TUP Online

Quality Management Systems.



Categorization of ISO 9001:2015 Audit Reports Using Recurrent Neural Networks

REFERENCES

- Knowledge2Innovation. (2010). Application of quality management systems in research organizations technology centres and universities. https://sgitt-otri.ua.es/es/proyectos-internacionales/documentos/k2i/training-set/qms-guide.pdf pages 6-8
- Janette Ruiz Rodriguez, Madonna Valenzuela and Nunilon Ayuyao, "TQM paradigm for higher education in the Philippines", Quality Assurance in Education, https://doi.org/10.1108/QAE-12-2015-0048
- Ruiz, A. J., & Junio-Sabio, C. (2012): Quality Assurance in Higher Education in the Philippines. The Asian Society of Open and Distance Education. ISSN 1347-9008 Asian J D E 2012 vol 10, no 2, pp 63 – 70
- ISO. (2015). ISO 9001:2015 Quality Management Systems Requirements. International Organization for Standardization
- Agasisti, Tommaso & J. Bowers, Alex. (2018). Data Analytics and Decision-Making in Education: Towards the Educational Data Scientist as a Key Actor in Schools and Higher Education Institutions.
- Guan, Y., Sun, Y., Wang, Z., & Zheng, Q. (2017). Natural Language Process: A New Kind of Nuclear Quality Assurance Management Tool. Energy Procedia, 127, 201-219
- Pandey, N., Sandal, D. K., Hudait, A., & Sen, A. (2017). Automated classification of software issue reports using machine learning techniques: an empirical study. Innovations in Systems and Software Engineering, 13(4), 279–297. doi:10.1007/s11334-017-0294-1
- Aggarwal, A., Singh, J., & Kapil Gupta, D. (2018). A Review of Different Text Categorization Techniques. International Journal of Engineering & Technology, 7(3.8), 11-15. DOI: http://dx.doi.org/10.14419/ijet.v7i3.8.15210
- HD Wehle. (2017). Machine Learning, Deep Learning and AI: What's the Difference?
- Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. EMNLP
- Pengfei Liu, Xipeng Qiu, and Xuanjing Huang. 2016. Recurrent neural network for text classification with multi-task learning. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16), Gerhard Brewka (Ed.). AAAI Press 2873-2879
- Nowak, J., Taspinar, A., & Scherer, R. (2017). LSTM Recurrent Neural Networks for Short Text and Sentiment Classification. Lecture Notes in Computer Science, 553–562. doi:10.1007/978-3-319-59060-8_50
- Shih, C.-H., Yan, B.-C., Liu, S.-H., & Chen, B. (2017). Investigating Siamese LSTM networks for text categorization. 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). doi:10.1109/apsipa.2017.8282104
- K. Radhika, Bindu, K. R., and Parameswaran, L., "A text classification model using convolution neural network and recurrent neural network", International Journal of Pure and Applied Mathematics, vol. 119, pp. 1549-1554, 2018.
- 15. Jindal, R., Malhotra, R., & Jain, A. (2015). Techniques for text classification: Literature review and current trends. Webology, 12.
- D. S. Guru, B. S. Harish, and S. Manjunath. 2010. Symbolic representation of text documents. In Proceedings of the Third Annual ACM Bangalore Conference (COMPUTE '10). ACM, New York, NY, USA, Article 18, 4 pages. DOI: https://doi.org/10.1145/1754288.1754306
- S. Deepu, Pethuru Raj, S. Rajaraajeswari, "A Framework for Text Analytics using the Bag of Words (BoW) Model for Prediction", 1st International Conference on Innovations in Computing & Networking (ICICN16) CSE RRCE ISSN: 0975–0282.
- 18. Haider, S. (2018). Urdu Word Embeddings. LREC.
- Hu, Y., Huber, A.E., Anumula, J., & Liu, S. (2018). Overcoming the vanishing gradient problem in plain recurrent networks. CoRR, abs/1801.06105.
- Zhou, C., Sun, C., Liu, Z., & Lau, F.C. (2015). A C-LSTM Neural Network for Text Classification. CoRR, abs/1511.08630.
- Zhou, P., Qi, Z., Zheng, S., Xu, J., Bao, H., & Xu, B. (2016). Text Classification Improved by Integrating Bidirectional LSTM with Two-dimensional Max Pooling. COLING.
- Fernandez, A. C. T., Tarnate, K. J. M., & Devaraj, M. (2018).
 Deep Rapping: Character Level Neural Models for Automated Rap Lyrics Composition. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-2S
- Hassan, A., & Mahmood, A. (2017, December). Efficient deep learning model for text classification based on recurrent and convolutional layers. In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1108-1113). IEEE
- De Guia, J.M. & Devaraj, M. (2019). Methods and Trends in Natural Language Processing Applications in Big Data. International Journal of

- Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7 Issue-6S5.
- J. N. Rao and M. Ramesh (2019). A Review on Data Mining & Big Data, Machine Learning Techniques. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7-Issue-6S2

AUTHORS PROFILE



Ken Jon M. Tarnate, He is currently working in Technological University of the Philippines as a college instructor in College of Science, Mathematics Department. He is also a part of the Quality Assurance Auditing Team of the said University. He works as an Internal Auditor and Technical Assistant in the quality

assurance department. He obtained his bachelor's degree in information and communication technology education from Philippine Normal University – Manila, Philippines, in 2014. He is currently working toward a M.S. degree in computer science at Mapua University, Philippines. He is also a full-time scholar of Engineering Research and Development for Technology (ERDT) under Department of Science and Technology of the Philippines.



Dr. Madhavi Devaraj, She graduated with a PhD in Computer Science from Dr. A.P.J. Abdul Kalam Technical University (formerly Uttar Pradesh Technical University) in Lucknow, Uttar Pradesh, India in 2016. She took up her master's in philosophy in Computer Science from the Madurai Kamaraj

University and Master of Computer Applications from V.V. Vannaiaperumal College for Women both in India in 2004 and 2000 respectively. She finished her Bachelor of Science in Mathematics from the Bharathidasan University - Government Arts College for Women in 1997.

