

# Classification of Qur'anic topics based on imbalanced classification

**Bassam Arkok, Akram M. Zeki**

Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, Malaysia

---

## Article Info

### Article history:

Received Jun 30, 2020

Revised Apr 7, 2021

Accepted Apr 13, 2021

---

### Keywords:

Imbalanced classification

Quran

Re-sampling methods

Text classification

Topics

---

## ABSTRACT

Imbalanced classification techniques have been applied widely in the field of data mining. It is used to classify the imbalanced classes that are not equal in the number of samples. The problem of imbalanced classes is that the classification performance tends to the class with more samples while the class with few samples will obtain poor performance. This problem can be occurred in the Qur'anic classification due to the different in the number of verses. Many studies classified Qur'anic verses, which depended on the traditional classification. However, no study classified Qur'anic topics based on the techniques of imbalanced classification. Therefore, this paper aims to apply the methods of imbalanced classification as synthetic minority over-sampling technique (SMOTE), random over sample (ROS), and random under sample (RUS) methods to classify the Qur'anic topics that are imbalanced. Many metrics were used in this research to evaluate the experimental results. These metrics are sensitivity/recall, specificity, overall accuracy, f-measure, g-mean, and matthews correlation coefficient (MCC). The results showed that the Qur'anic classification performance improved when imbalanced classification techniques were applied.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



---

## Corresponding Author:

Bassam Arkok

Department of Computer Science

Kulliyyah of Information and Communication Technology

International Islamic University Malaysia, Jalan Gombak, 53100, Selangor, Malaysia

Email: bassam.arkok@gmail.com

---

## 1. INTRODUCTION

The Holy Quran is one of the holy books of Allah SWT. It is considered one of the primary references for an estimated 1.6 billion Muslims around the world. Many artificial intelligence (AI) researches have been conducted on the holy Quran, such as automatic text classification (ATC). The topical classification of Qur'anic text is assigning one or more topics automatically for the Qur'anic text based on their content. The Qur'anic verses classification can be considered one of the essential researches in Machine learning. The researchers have been interested in the classification of the Qur'anic verses because the Quran is the first source of religious texts for Muslims. The Qur'anic text is presented as either instructive or narrative, and it has an underlying deep connection that joins the entire text into one whole related concept [1]. The essential ideas and meaning of the Qur'an overlap from chapter to chapter and verse to verse. Therefore, finding out the implied connections would need time and more in-depth study to discover the hidden topics in the Qur'an as there are many topics found [2]. Besides, the Qur'anic verses should be classified to extract the revealed knowledge in the Quran to help Muslims understand the Qur'anic text well.

Imbalanced text classification assigns one or more classes to a document according to their content for the text, which is highly imbalanced. Imbalanced classification is a popular technique for imbalanced

datasets in different languages. The data set is called an imbalanced dataset when one class consists of more samples significantly than the others [3]-[5]. The class with more samples is called the majority class, while the class with a few samples is called the minority class. The classes are imbalanced if the imbalanced classes' rate range is between 0.01-29.1% [6]. Researchers have paid attention to the imbalanced datasets classification due to their importance in many real-world practices such as text classification, biological data analysis, fraud detection, and image classification [7].

The traditional classification fails when applied to classify the imbalanced data because the high accuracy tends for the majority class while the low accuracy tends for the minority class. These are several methods to solve the problem of imbalanced classification. One solution is re-sampling methods, which are among the most common techniques discovered in solving this problem. They are considered a pre-processing phase before the classification task that relies on generating new samples for the minority class or eliminating samples from the majority class to re-balance the imbalanced datasets. Hundreds of algorithms have been proposed in different languages in the past decade [4] to classify the imbalanced datasets. However, few studies were conducted to classify imbalanced Arabic text [5], [8]-[9].

Many Qur'anic topics can be considered as imbalanced datasets due to the difference in the number of verses. As it is noticed in the used index for this research, most Qur'anic topics will be imbalanced when classified altogether. For example, the Tawheed topic has 940 verses, while the Shirk topic has 94 verses only. Then, tawheed and shirk are considered as imbalanced classes if they are classified together. Therefore, the classification performance will be tended to the samples of Tawheed topics, while Shirk topic samples will be obtained poor performance. Nevertheless, many studies classified Qur'anic verses in Arabic and non-Arabic language corpus based on different techniques, but these studies did not pay attention to the imbalanced issue. So, this research aims to classify Qur'anic topics using the methods of imbalanced classification.

The purpose of this research is twofold: first, the classification of Qur'anic topics using some techniques of imbalanced classification and comparison with the traditional classification. The second is to determine the best method among these techniques for Qur'anic classification later. The rest of this paper is organized as follows: the next section is a literature review that will mention some studies based on classified Qur'anic datasets by different methods. The methodology of the research will be offered to clarify the steps and phases of the classification process. Finally, the results and the experiments will be revealed to conclude some of the findings and show the performance of the classification for the Holy Quran with or without using the imbalanced classification.

## 2. LITERATURE REVIEW

There are many studies classified Qur'anic verses for Arabic and non-Arabic language corpus to determine its corresponding topic. The first attempt of classifying Qur'anic Arabic topics was by [10]. The authors designed a statistical model to categorize the verses automatically for the chapters of Fatiha and Yaseen for predefined classes. A classifier classifies the verses in each chapter based on computing a score for every verse against each category as a first stage, and then the verses were assigned to classes with the highest score. This study is considered limited and primitive in the topical classification of the Qur'anic verses [11].

Nassourou [12] discussed the classification of Qur'anic chapters according to the major phases of the Prophet Mohammad's messengership using machine learning algorithms. First, the chapters were classified according to places of the revelation using Naïve Bayes and Support vector machine classifiers. Moreover, the chapters were classified into Meccan and Medinan. After that, the Meccan and Medinan chapters were clustered into groups depended on cities of revelation. This study concluded that the Medinan chapters got a higher performance from the Meccan chapters, whether for SVM or Naïve Bayes.

In [11], the authors classified verses of the Quran according to their topics. They applied well-known classification algorithms: k-nearest neighbours (K-NN), decision trees, naïve bayes (NB), and support vector machines (SVM) for three selected topics. The pre-processing phase was performed on the Qur'anic dataset, and six performance metrics were used for the evaluation. However, only 1,227 verses were used out of the whole 6,236 verses of the Holy Quran for training and evaluating the selected topics. The topics were (Ignorant of religion, "*Aljahilun bialdiyn*"), (Oneness of God, "*Tawhid Allah*"), and (Penalty of Apostates, "*Jaza' almurtadin*"). These topics were considered based on the verses of the topics being unbalanced. There are 1,218 verses for Oneness of God, and three verses only for Ignorant of religion, while the Penalty of Apostates has six verses. Therefore, the authors classified these imbalanced topics by using traditional classification.

Akour *et al.* [13] aimed to evaluate similarity in the documents and the text in the Arabic language based on the Holy Quran. The Qur'anic verses were used as queries to search and evaluate similarity. The

measuring Qur'anic verses similarity and sura classification (MQVC) approach were employed to retrieve the most similar verses compared to a user input verse as a query. The dataset consists of over 2000 verses from the Quran, and they implemented the MQVC approach randomly on 70 of the 114 chapters of the Holy Quran. The experiment was performed using the N-gram technique besides the machine learning algorithm (LibSVM classifier) to classify the selected Quran chapters into chapters of Makki and Madani.

While Hassan *et al.* [14] applied KNN algorithm to classify the Tafseer verses of the Holy Quran into predefined categories and to conduct the classification, a database of 1000 Ayat of the Quran was divided into two document sets with the training set consists of 800 verses and the test set consist of 200 verses. For this classification, seven of the Tafseer texts were chosen: Inheritance, Marriage, Respecting Parents, Pray, Zakat, Jihad, and Halal after translating them from the Arabic to the Malay language. This dataset was experimented based on Recall and Precision metrics. 'Inheritance' has the lowest recall value of 0.74, while the 'marriage' category has the highest recall value of 0.9.

In another research [15], it aimed to develop the efficiency of information retrieval from the Holy Quran based on question answering system (QAS) through classifying the verses by using the neural network (NN) technique. This research had used the most popular English translation of the Quran of Abdullah Yusuf Ali as the data set. The QAS tackled some problems by expanding the question, using WordNet, and benefitting from collecting Islamic terms to avoid differences in the questions and translations. This QAS classified the Al-Baqarah surah into two classes, fasting and pilgrimage. It is based on the NN classifier, to reduce the retrieval of irrelevant verses since a user's questions are asked about Fasting and Pilgrimage. Hence, the QAS retrieves the relevant verses to the question based on the N-gram technique, then ranking the retrieved verses based on the highest similarity score to satisfy the user's desire.

In [16], the authors presented a feature selection approach to label Qur'anic verses automatically to relate the verses with the five pillars of Islam; the 'Shahada' (profession of faith), the 'Solat' (daily prayers), the 'Zakat' (almsgiving), the 'Saum' (fasting during Ramadan), and the 'Hajj' (pilgrimage to Mecca). The proposed feature selection approach used a group-based term count to represent the features extracted from three different Qur'anic translations. The approach adopted two common feature selection algorithms to reduce the feature space for 200 randomly selected verses based on the manual index, limited to Islam's first two pillars. Finally, the k-NN, SVM, and NB classifiers were implemented independently on the feature selection algorithms to determine class membership for each verse and measure the results in terms of area under the receiver operating characteristics curve (AUC).

Rostam and Malim [17] proposed a method to classify selected classes by determining the interrelation between the resources of the Qur'an and Hadith using the techniques of text classification. Several experiments were implemented by using a combination of different Quran and Hadith datasets. Three categories were selected, Hajj, Prayer, and Zakat, which were compared using three classifiers, SVM, Naïve Bayes, and KNN via term frequency-inverse document frequency (TF-IDF) as well as term weighting. Furthermore, ensemble methods were experimented with to find out whether they are fit to classify the interrelated datasets as the resources of the Qur'an and Hadith.

Ulumudin *et al.* [18] built a model to classify a multi-label Qur'anic corpus translated to English using KNN classifier. In this study, KNN classifier was combined with TF-IDF and Weighted TF-IDF methods to determine the best approach. According to the results obtained in this research, Weighted TF-IDF method achieved better results compared to TF-IDF.

In [19], the authors combined two feature selection techniques to improve classification performance and at the same time to reduce the computational runtime. The combined algorithms were chi-square and wrapper correlation-based (CFS). Firstly, CHI was implemented to reduce the features space dimensionality, and CFS was implemented later to select the essential features from the reduced features. This proposed technique was applied to classify Qur'anic verses using the standard classifiers, SVM, NB, and J48. The Quranic classification was achieved better results than the basic selection feature methods when the proposed approach was applied.

Elmitwally and Alsayat [20] proposed a model that can classify the Qur'anic verses automatically. It can also conclude the essential features through the classification of the first six chapters in the Qur'an. This study used SVM, Naïve Bayes, J48, and KNN classifiers to classify the Qur'anic text. In the end, a multi-class classification model was built for the chapter names of these Qur'anic verses using GaussianNB and support vector classifier (SVC).

### 3. RESEARCH METHOD

This section explains the research methodology in detail, which comprises many steps before the classification task. In Figure 1, it is shown all these steps using the following flowchart.

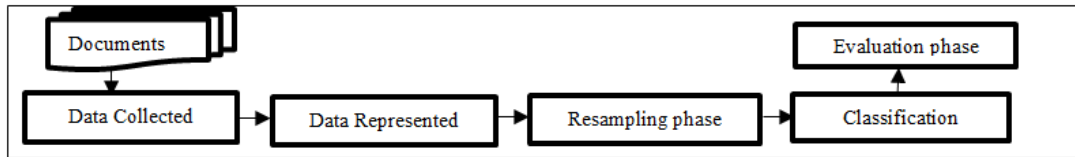


Figure 1. Flowchart of this research

**3.1. Dataset collected**

This work classifies verses of 8 Qur’anic topics declared by the lexicon for the Quran’s topics by Islamic scholars [21]. These topics are imbalanced dataset which cannot be classified by the traditional methods correctly. Table 1 presents the collected data with the number of samples for each binary dataset, imbalanced ratios (IR) among these topics, and the features number.

Table 1. Presents the collected data with the number of samples for each binary dataset, IR and the number of features

		# of majority class	# of minority class	IR	# of features
1	Tawheed_Shirk	940	94	10	1119
2	Prayer_Zakat	186	37	5.03	645
3	Labor_Science	537	76	7.07	947
4	Prophecy_Religion	322	86	3.74	743
5	Q_stories_Political_R	301	32	9.41	726
6	Jihad_Fasting	153	29	5.28	624
7	Islam_Faith	1291	993	1.30	1477
8	Organization of financial Relations_Call_to_Allah	211	140	1.51	760

**3.2. Data representation**

The collected topics’ Quranic roots were extracted by [22] that has been published in IJASAT journal. So, the pre-processing step to get the Qur’anic roots is cancelled and ignored in this work. The collected topics were represented using the term frequency-inverse document frequency (TF-IDF) technique in the classification matrix. Firstly, all Qur’anic roots were formatted in a vector space. Then, TF-IDF for all Qur’anic words for each verse was computed, and its value was put under its feature in the space vector. After that, the matrix was represented that collected all vectors for these verses. Finally, these vectors are labelled in the matrix based on the lexicon used.

Moreover, all common verses for both classes are removed from the matrix. For example, 24 verses exist in Tawheed and Shirk topics at the same time, so these verses are deleted from Tawheed and Shirk classes. Deleting the common verses from the datasets helps to get pure data to prevent confusion in the classification phase and avoid overlapping between the training and testing datasets.

**3.3. Features selection**

In this stage, the formatted data features in the matrix are reduced and chosen by the Chi-square algorithm. Chi-square algorithm has been used to select the important features which help in improving the performance of the classification. The features are picked out by the Chi-square algorithm, which was used widely for the Arabic text classification [23]-[24]. After the features are selected, the duplicated samples are deleted again from the represented data. The repeated samples are removed to get pure data and also to evaluate the different methods more precisely. In other words, when the same sample exists in both training and testing datasets, this sample has 100% accuracy. Therefore, this issue is taken into account to get results more accurate and strictly evaluate the different methods.

**3.4. Division of datasets**

The datasets are divided into 70% as training data and 30% as testing data to build the training models and test the classifiers. This division considers the imbalanced ratios of the Qur’anic dataset, which means 70% of every class is chosen for the training data and the same policy for the testing data. This division is based on the study done by [23], which concluded that classification accuracy is improved when the testing dataset is smaller than the training dataset. Table 2 shows the final datasets after applying the previous steps with the number of samples for each topic in the training phase, the imbalanced ratios (IR), and the number of each dataset’s features.

Table 2. Shows the final datasets after applying the previous steps

		# of majority class	# of minority class	IR	# of features
1	Tawheed_Shirk	185	57	3.25	53
2	Prayer_Zakat	32	16	2	18
3	Labor_Science	228	51	4.47	98
4	Prophecy_Religion	95	51	1.86	36
5	Q_stories_Political_R	121	22	5.5	86
6	Jihad_Fasting	27	20	1.35	34
7	Islam_Faith	778	529	1.47	133
8	Organization of financial Relations_Call_to_Allah	90	67	1.34	60

### 3.5. Re-sampling phase

The most common techniques in the re-sampling techniques are applied in this research. These methods are random under sampling (RUS), random over sampling (ROS), and synthetic minority over-sampling technique (SMOTE). RUS method eliminates the majority class samples randomly for the re-balancing of samples in both classes [25]. While in ROS method, the samples of minority class are duplicated randomly also for the re-balancing. RUS and ROS methods rely on re-balancing the samples randomly, so they are implemented 1000 times in this research. The average of all these experiments is computed to obtain the experimental results for them. But, the SMOTE method oversamples the minority class based on the neighbouring samples [26]. New samples of synthetic minority class are added by interpolating pairs of closest neighbours randomly in the minority class.

### 3.6. Classification phase

This work uses waikato environment for knowledge analysis (WEKA 3.6) tool to upload the Qur'anic datasets collected and implement the classifiers required. 4 different base classifiers were used for the classification task: Naive Bayes, J48, Voted perceptron, and LibSVM classifiers. But, before the classification process, the training datasets should be re-sampled by the re-sampling methods. Moreover, many classifiers were combined with these re-sampling methods in this study to show the performance of the best combination for the Quranic topics.

### 3.7. Evaluation phase

This step comprises a phase of the evaluation to test the experimented methods. Testing of the experiments is evaluated using trained models to reveal the performance of the classifiers. As long as the dataset is imbalanced, then the accuracy metric is not enough to test the imbalanced classification performance because the prediction is tended for the majority class [8], [27]-[28]. So, for this phase, different metrics are used here to evaluate the performance of imbalanced classification. These metrics are sensitivity/recall, specificity, overall accuracy, f-measure, g-mean, and matthews correlation coefficient (MCC) used in the imbalanced classification [4]. The formulas for these metrics are:

$$\text{Sensitivity/Recall (the accuracy of majority class)} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity (the accuracy of minority class)} = \frac{TN}{TN + FP} \quad (2)$$

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$F - \text{Measure} = 2 \times \left( \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (4)$$

$$G - \text{mean} = \sqrt{\text{Sensitivity} \times \text{Specificity}} \quad (5)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6)$$

**4. EXPERIMENTAL RESULTS AND ANALYSIS**

In this section, the experimental results obtained by this research are presented in 4 formats. Firstly, the classifiers' average results without using any re-sample method. These are the results of traditional classification to prove that the traditional classification is poor when used to classify Qur'anic topics that are imbalanced. Second, the results average obtained using SMOTE method for the required. Third, the results average achieved using RUS method in all classifiers. Finally, the results average of the classifiers using ROS method in all Qur'anic datasets. Table 3 shows the average results for all Qur'anic datasets for all techniques used. All results above will be categorized based on the used classifiers: LibSVM, Naïve Bayes, J48, and Voted perceptron.

**4.1. Experimental results obtained by LibSVM classifier**

Table 3 and Figure 2 reveal the experimental results obtained using LibSVM classifier in all used techniques. As is noticed in Table 3 and Figure 2, the traditional classification, which was implemented by LibSVM classifier, it obtained higher results in Sensitivity/Recall metrics, but it was the worst among the others in the Specificity metric. It means the traditional classification tends to achieve higher accuracy for the majority class, while the samples of minority class get poor performance. This weakness in traditional classification performance was not observed because it earned a higher value in the Overall accuracy metric. So, it was decided in the field of imbalanced classification that the Overall accuracy metric is not enough to evaluate the imbalanced classification performance. However, the imbalanced classification techniques applied in this research obtained similar results in the Overall accuracy metric. It is also shown that the traditional classification was the weakest in the performance in G-mean and MCC metrics. The traditional classification outperformed the re-sampling methods in F-measure slightly in 1% or 2% percentages only. Finally, in Table 3, it is seen that the SMOTE method improved the results of traditional classification significantly when it was combined with LibSVM classifier, especially in Specificity and G-mean metrics.

Table 3. Reveals the experimental results obtained using libsvm classifier

Technique	Sensitivity/Recall	Specificity	Overall Accuracy	F-Measure	G-Mean	MCC
LibSVM	0.93	0.63	0.84	0.89	0.76	0.62
RUS_LibSVM	0.87	0.77	0.84	0.87	0.81	0.65
ROS_LibSVM	0.88	0.76	0.84	0.88	0.82	0.65
SMOTE_LibSVM	0.87	0.80	0.85	0.88	0.83	0.66

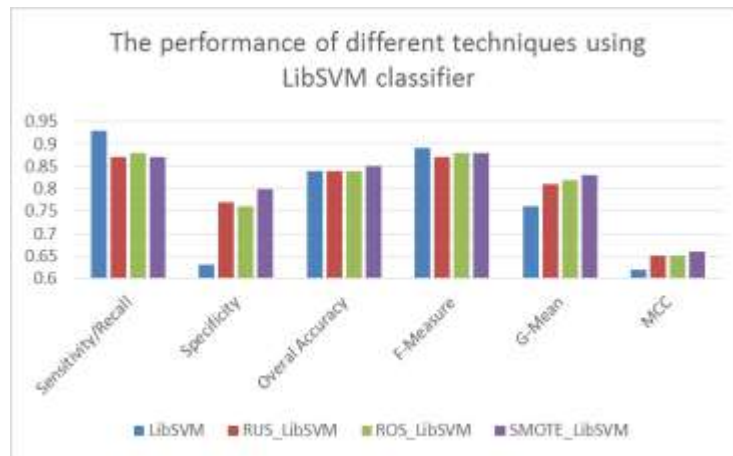


Figure 2. Shows the performance of all techniques using LibSVM classifier

**4.2. Experimental results obtained by KNN classifier**

Table 4 and Figure 3 show the experimental results were revealed using KNN classifier for the collected Qur'anic in all required techniques in this research. It is noticed in the table and figure above that the majority class's accuracy was significantly higher than the accuracy of the minority class in KNN classifier. But these accuracies were adjusted somewhat in SMOTE and ROS methods. Moreover, the performance was similar in all techniques in Overall accuracy and F-measure metrics. Also, it is seen that SMOTE method outperformed slightly in G-mean metric and significantly in MCC on the others. Also, in

Table 4, it is seen that SMOTE method improved the results of traditional classification significantly when it was combined with KNN classifier in most metrics.

Table 4. Shows the experimental results revealed using KNN classifier

Technique	Sensitivity/Recall	Specificity	Overall Accuracy	F-Measure	G-Mean	MCC
KNN	0.86	0.55	0.77	0.83	0.68	0.45
RUS_KNN	0.80	0.67	0.76	0.81	0.73	0.47
ROS_KNN	0.85	0.56	0.78	0.83	0.69	0.46
SMOTE_KNN	0.84	0.74	0.81	0.85	0.79	0.58

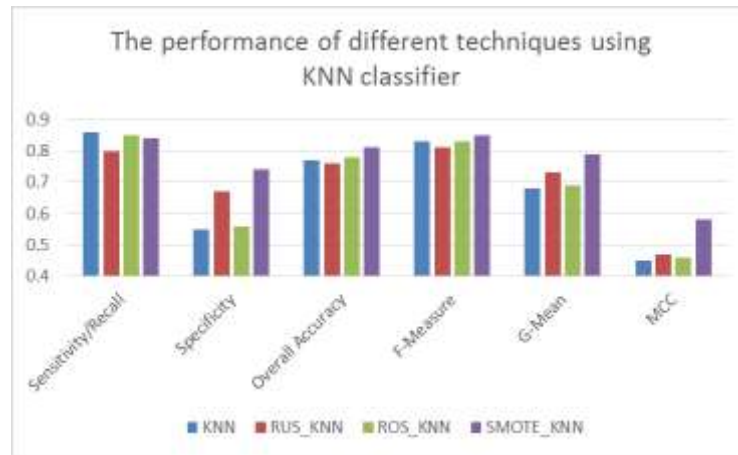


Figure 3. Shows the performance of all techniques using KNN classifier

### 4.3. Experimental results obtained by J48 classifier

Table 5 and Figure 4 show the experimental results obtained using the J48 classifier for required techniques. In the Table and Figure above, it is noticed that the performance of traditional classification, using J48 classifier, also demonstrated the problem of imbalanced classification when was used to classify imbalanced Qur’anic topics. Therefore, the classification performance tended to the majority class and paid less attention to the minority class. Moreover, the re-sampling methods outperformed the traditional classification in the other metrics besides the Overall accuracy metric except for F-measure. In F-measure, the traditional classification achieved similar performance to the performance of re-sampling methods. Furthermore, as in Table 5, the SMOTE method outperformed the others in many metrics.

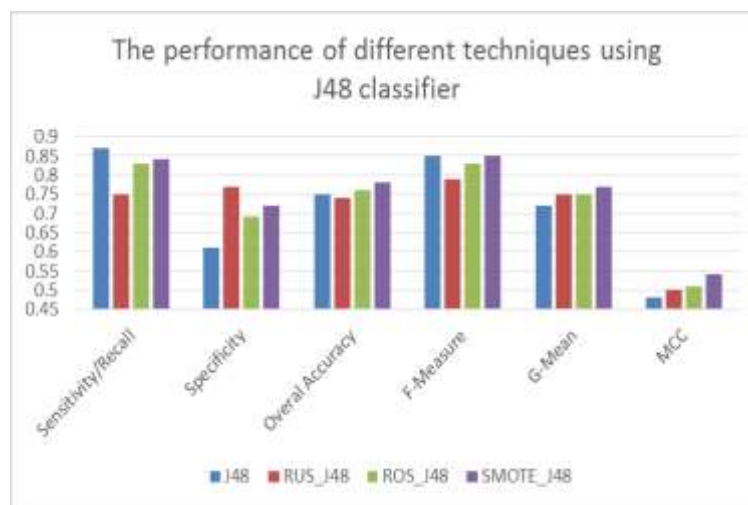


Figure 4. Shows the performance of all techniques using J48 classifier



Table 5. Shows the experimental results obtained using J48 classifier

Technique	Sensitivity/Recall	Specificity	Overall Accuracy	F-Measure	G-Mean	MCC
J48	0.87	0.61	0.75	0.85	0.72	0.48
RUS_J48	0.75	0.77	0.74	0.79	0.75	0.50
ROS_J48	0.83	0.69	0.76	0.83	0.75	0.51
SMOTE_J48	0.84	0.72	0.78	0.85	0.77	0.54

**4.4. Experimental results obtained by voted perceptron classifier**

The experimental results obtained using Voted perceptron classifier are explored in Table 6 and Figure 5 for the applied techniques. As in the table and figure above, the majority class samples achieved higher performance, while the samples of the minority class obtained poor performance in traditional classification performance in all metrics. But, the performance of re-sampling methods was better in all metrics. Moreover, the MCC metric is improved significantly than the traditional classification in SMOTE and ROS methods. It means that the classification of Qur’anic topics, which are imbalanced, will be enhanced when classified by imbalanced classification techniques. Finally, as shown in Table 6, the SMOTE method improved traditional classification results significantly when combined with a Voted perceptron classifier. This improvement was proven in all metrics except in Sensitivity/Recall metric that got a similar performance in this metric.

Table 6. Presents experimental results obtained using voted perceptron classifier

Technique	Sensitivity/Recall	Specificity	Overall Accuracy	F-Measure	G-Mean	MCC
VP	0.85	0.64	0.79	0.84	0.73	0.50
RUS_VP	0.74	0.79	0.75	0.80	0.76	0.49
ROS_VP	0.82	0.79	0.81	0.85	0.81	0.59
SMOTE_VP	0.84	0.87	0.85	0.88	0.85	0.68

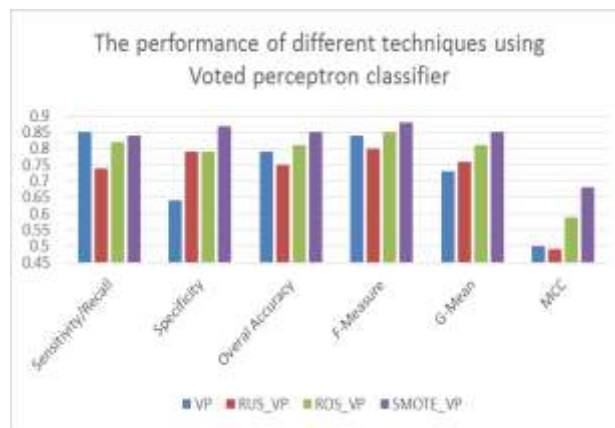


Figure 5. Shows the performance of all techniques using voted perceptron classifier

**5. RESULTS ANALYSIS IN GENERAL**

The performance of re-sampling methods is compared in all required classifiers in this section. This comparison is conducted to pick out the best one for the Qur’anic classification. In Table 7, it is shown the performance of all re-sampling techniques in all used metrics. As noticed in the Table above, all re-sampling methods were the best methods in Sensitivity/Recall metrics when combined with LibSVM classifier, while SMOTE with Voted perceptron classifier was the best one in Specificity metric. Moreover, all these combinations were the best techniques in Overall Accuracy and F-Measure metrics. In contrast, the combination of SMOTE with Voted perceptron achieved higher performance practically in G-mean and MCC metrics. As a result of the above, the candidate as the best performer in this research is SMOTE with Voted perceptron classifier due to his superiority over the others in 5 out of 6 metrics. According to the results obtained in Section 4, SMOTE was the best re-sampling method compared to the other techniques in this research. Therefore, SMOTE can be considered the best approach to classify imbalanced Qur’anic topics.



Table 7. Reveals the performance of all re-sampling methods

Technique	Sensitivity/Recall	Specificity	Overall Accuracy	F-Measure	G-Mean	MCC
RUS_J48	0.75	0.77	0.74	0.79	0.75	0.50
ROS_J48	0.83	0.69	0.76	0.83	0.75	0.51
SMOTE_J48	0.84	0.72	0.78	0.85	0.77	0.54
RUS_LibSVM	0.87	0.77	0.84	0.87	0.81	0.65
ROS_LibSVM	0.88	0.76	0.84	0.88	0.82	0.65
SMOTE_LibSVM	0.87	0.80	0.85	0.88	0.83	0.66
RUS_VP	0.74	0.79	0.75	0.80	0.76	0.49
ROS_VP	0.82	0.79	0.81	0.85	0.81	0.59
SMOTE_VP	0.84	0.87	0.85	0.88	0.85	0.68
RUS_KNN	0.80	0.67	0.76	0.81	0.73	0.47
ROS_KNN	0.85	0.56	0.78	0.83	0.69	0.46
SMOTE_KNN	0.84	0.74	0.81	0.85	0.79	0.58

## 6. CONCLUSION

Text classification is one of the most important aspects used widely nowadays. Several studies classified the Qur'anic text using different methods. However, there is no study for the classification of Qur'anic text based on imbalanced classification. This work aimed to classify imbalanced Qur'anic topics using Re-sampling methods to re-balance the Qur'anic dataset before the classification phase. Also, in this research, classifiers' performance is compared before and after using these re-sampling methods. The results showed that the Qur'anic classification performance improved significantly by applying the imbalanced classification almost for many of the metrics with the stability of the accuracy metric. The future work of this research is applying the imbalanced classification techniques for the Sunnah books to improve the Hadiths classification.

## ACKNOWLEDGEMENTS

This paper was supported by International Islamic University of Malaysia under (FRGS19-083-0691) research project.

## REFERENCES

- [1] Siddiqui, M. A., S. M. Faraz, and S. A. Sattar. "Discovering the thematic structure of the Quran using probabilistic topic model," in *2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences*. 2013. IEEE, doi: 10.1109/nooric.2013.55.
- [2] Salloum, S. A, et al., "A survey of Arabic text mining," in *Intelligent Natural Language Processing: Trends and Applications*, Springer. pp. 417-431, 2018, doi: 10.1007/978-3-319-67056-0\_20.
- [3] Ali, H., et al., "A review on data pre-processing methods for class imbalance problem," *International Journal of Engineering & Technology*, vol. 8, no. 3, pp. 390-397, 2019, doi: 10.14419/ijet.v8i3.29508.
- [4] Haixiang, G., et al., "Learning from class-imbalanced data: Review of methods and applications," *Expert Systems with Applications*, vol. 73, pp. 220-239, 2017, doi: 10.1016/j.eswa.2016.12.035.
- [5] Nayal, A., H. Jomaa, and M. Awad, "KerMinSVM for imbalanced datasets with a case study on arabic comics classification," *Engineering Applications of Artificial Intelligence*, vol. 59, pp. 159-169, 2017, doi: 10.1016/j.engappai.2017.01.001.
- [6] Khalilia, M, S. Chakraborty, and M. Popescu, "Predicting disease risks from highly imbalanced data using random forest," *BMC medical informatics and decision making*. vol. 11, no. 1, pp. 51, 2011, doi: 10.1186/1472-6947-11-51.
- [7] Alibeigi, M, S. Hashemi, and A. Hamzeh, "DBFS: An effective Density Based Feature Selection scheme for small sample size and high dimensional imbalanced data sets," *Data & Knowledge Engineering*. vol. 81, pp. 67-103, 2012, doi: 10.1016/j.datak.2012.08.001.
- [8] Al-Azani, S and E.-S. M. El-Alfy. "Using Word Embedding and Ensemble Learning for Highly Imbalanced Data Sentiment Analysis in Short Arabic Text," in *ANT/SEIT*, 2017, doi: 10.1016/j.procs.2017.05.365.
- [9] Mountassir, A, H. Benbrahim, and I. Berrada, "Some methods to address the problem of unbalanced sentiment classification in an arabic context," in *2012 Colloquium in Information Science and Technology*, 2012, IEEE, doi: 10.1109/CIST.2012.6388061.
- [10] Al-Kabi, M. N, et al., "Statistical classifier of the holy Quran verses (Fatiha and Yaseen chapters)," *Journal of Applied Sciences*. vol. 5, no. 3 pp. 580-583, 2005, doi: 10.3923/jas.2005.580.583.
- [11] Al-Kabi, M. N., et al., "A topical classification of Quranic Arabic text," in *Proceedings of the 2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences*. 2013.
- [12] Nassourou, M, "Using Machine Learning Algorithms for Categorizing Quranic Chapters by Major Phases of Prophet Mohammad's Messengership," *International Journal of Information and Communication Technology Research: Citeseer*, vol. 2, 2012.

- [13] Akour, M. I. M. Alsmadi, and I. Alazzam, "MQVC: Measuring quranic verses similarity and sura classification using N-gram," *WSEAS Transactions on Computers*, vol. 13, pp. 485-491, 2014.
- [14] Hassan, G. S., S. K. Mohammad, and F. M. Alwan, "Categorization of 'Holy Quran-Tafseer' using K-nearest neighbor algorithm," *Int. J. Comput. Appl.*, vol. 129, no. 12, pp. 1-6, 2015.
- [15] Hamed, S. K and M. J. Ab Aziz, "A Question Answering System on Holy Quran Translation Based on Question Expansion Technique and Neural Network Classification," *J. Comput. Sci.*, vol. 12, no. 3, pp. 169-177, 2016, doi: 10.3844/jcssp.2016.169.177.
- [16] Adeleke, A. O., *et al.*, "Comparative analysis of text classification algorithms for automated labelling of Quranic verses," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 7, no. 4, p. 1419, 2017, doi: 10.18517/ijaseit.7.4.2198.
- [17] Rostam, N. A. P and N. H. A. H. Malim, "Text categorisation in Quran and Hadith: Overcoming the interrelation challenges using machine learning and term weighting," *Journal of King Saud University-Computer and Information Sciences*, 2019, doi: 10.1016/j.jksuci.2019.03.007.
- [18] Ulumudin, G, A. Adiwijaya, and M. Mubarak. "A multilabel classification on topics of qur'anic verses in English translation using K-Nearest Neighbor method with Weighted TF-IDF," in *Journal of Physics: Conference Series*. 2019. IOP Publishing, doi: 10.1088/1742-6596/1192/1/012026.
- [19] Adeleke, A., *et al.*, "A two-step feature selection method for quranic text classification," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 2, pp. 730-736, 2019, doi: 10.11591/ijeecs.v16.i2.pp730-736.
- [20] Elmitwally, N. S and A. Alsayat, "The Multi-Class Classification for the First Six Surats of the Holy Quran," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 1, pp. 327-332, 2020, doi: 10.14569/IJACSA.2020.0110141.
- [21] Al-Khair, A. A and M. A. Kabbani, "Quran teacher intonation," The Tunisian Company for Distribution, 2003. 2003. علي أبو الخير، محمد عربي القباني، مصحف معلم التجويد، الشركة التونسية للتوزيع، 2003.
- [22] Khedher, M. Z., "Multiword corpus of the Holy Quran," *International Journal on Islamic Applications in Computer Science And Technology*. vol. 5, no. 1, pp. 1-11, 2017.
- [23] Khorsheed, M. S. and A. O. Al-Thubaity, "Comparative evaluation of text classification techniques using a large diverse Arabic dataset," *Language resources and evaluation*, vol. 47, no. 2, pp. 513-538, 2013, doi: 10.1007/s10579-013-9221-8.
- [24] Alabbas, W., H. M. Al-Khateeb, and A. Mansour. "Arabic text classification methods: Systematic literature review of primary studies," in *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)*, IEEE. 2016, doi: 10.1109/CIST.2016.7805072.
- [25] Liu, X.-Y, J. Wu, and Z.-H. Zhou, "Exploratory undersampling for class-imbalance learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 2, pp. 539-550, 2009, doi: 10.1109/TSMCB.2008.2007853.
- [26] Chawla, N. V, *et al.*, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*. vol. 16, pp. 321-357, 2002, doi: 10.1613/jair.953.
- [27] Tang, Y, *et al.*, "SVMs modeling for highly imbalanced classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 1, pp. 281-288, 2009, doi: 10.1109/TSMCB.2008.2002909.
- [28] Patel, H and G. S. Thakur, "Classification of imbalanced data using a modified fuzzy-neighbor weighted approach," *International Journal of Intelligent Engineering and Systems*, vol. 10, no. 1, pp. 56-64, 2017, doi: 10.22266/ijies2017.0228.07.