Indonesian Journal of Electrical Engineering and Computer Science • Volume 22, Issue 2, Pages 70 - 79 • May 2020

Document type
Article

Source type
Journal

ISSN
25024752

DOI
10.11591/ijeecs.v22.i2.pp70-79

Publisher
Institute of Advanced Engineering and Science

Original language
English

Classification of Quranic topics based on imbalanced classification

Arkok B. , Zeki A.M.

Save all to author list

[a] Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, Selangor, Malaysia

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 Quranic classification due to the different in the number of verses. Many studies classified Quranic verses, which depended on the traditional classification . However, no study classified Quranic topics based on the techniques of imbalanced classification . Therefore, this paper aims to apply the methods of imbalanced classification as synthetic minority oversampling technique (SMOTE), random over sample (ROS), and random under sample (RUS) methods to classify the Quranic 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 Quranic classification performance improved when imbalanced classification techniques were applied.

Author keywords

Imbalanced classification Quran Re-sampling methods Text classification Topics

SciVal Topics

Metrics

References (28) View in search results format >

All

Export Print E-mail Save to PDF Create bibliography

1 Siddiqui, M.A., Faraz, S.M., Sattar, S.A.

Discovering the Thematic Structure of the Quran using Probabilistic Topic Model

(2013) *Proceedings - 2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences, NOORIC 2013*, art. no. 7277252, pp. 234-239. Cited 10 times.
ISBN: 978-147992823-1
doi: 10.1109/NOORIC.2013.55

View at Publisher

2 Salloum, S.A., AlHamad, A.Q., Al-Emran, M., Shaalan, K.

A survey of Arabic text mining

(2018) *Studies in Computational Intelligence*, 740, pp. 417-431. Cited 38 times.
http://www.springer.com/series/7092
doi: 10.1007/978-3-319-67056-0_20

View at Publisher

3 Ali, H.
A review on data pre-processing methods for class imbalance problem
(2019) *International Journal of Engineering Technology*, 8 (3), pp. 390-397. Cited 6 times.

4 Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., Bing, G.

Learning from class-imbalanced data: Review of methods and applications

(2017) *Expert Systems with Applications*, 73, pp. 220-239. Cited 844 times.
doi: 10.1016/j.eswa.2016.12.035

View at Publisher

5 Nayal, A., Jomaa, H., Awad, M.

KerMinSVM for imbalanced datasets with a case study on arabic comics classification

(2017) *Engineering Applications of Artificial Intelligence*, 59, pp. 159-169. Cited 7 times.
doi: 10.1016/j.engappai.2017.01.001

View at Publisher

6 Khalilia, M., Chakraborty, S., Popescu, M.

Predicting disease risks from highly imbalanced data using random forest (Open Access)

(2011) *BMC Medical Informatics and Decision Making*, 11 (1), art. no. 51. Cited 313 times.
http://www.biomedcentral.com/bmcmedinformdecismak/
doi: 10.1186/1472-6947-11-51

View at Publisher

7 Alibeigi, M., Hashemi, S., Hamzeh, A.

DBFS: An effective Density Based Feature Selection scheme for small sample size and high dimensional imbalanced data sets

(2012) *Data and Knowledge Engineering*, 81-82, pp. 67-103. Cited 55 times.
doi: 10.1016/j.datak.2012.08.001

View at Publisher

8 Al-Azani, S., El-Alfy, E.-S.M.

Using Word Embedding and Ensemble Learning for Highly Imbalanced Data Sentiment Analysis in Short Arabic Text (Open Access)

(2017) *Procedia Computer Science*, 109, pp. 359-366. Cited 72 times.
http://www.sciencedirect.com/science/journal/18770509
doi: 10.1016/j.procs.2017.05.365

View at Publisher

9 Mountassir, A., Benbrahim, H., Berrada, I.

Some methods to address the problem of unbalanced sentiment classification in an arabic context

(2012) *CiSt 2012 - Proceedings: 2012 Colloquium in Information Science and Technology*, art. no. 6388061, pp. 43-48. Cited 11 times.
ISBN: 978-146732725-1
doi: 10.1109/CIST.2012.6388061

View at Publisher

10 Al-Kabi, M. N
Statistical classifier of the holy Quran verses (Fatiha and Yaseen chapters)
(2005) *Journal of Applied Sciences*, 5 (3), pp. 580-583. Cited 24 times.

11 Al-Kabi, M. N.
A topical classification of Quranic Arabic text
(2013) *Proceedings of the 2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences*. Cited 15 times.

12 Nassourou, M
Using Machine Learning Algorithms for Categorizing Quranic Chaptersby Major Phases of Prophet Mohammad's Messengership, "
(2012) *International Journal of Information and Communication Technology Research: Citeseer*, 2. Cited 5 times.

13 Akour, M., Alsmadi, I., Alazzam, I.
MQVC: Measuring quranic verses similarity and sura classification using N-gram
(2014) *WSEAS Transactions on Computers*, 13, pp. 485-491. Cited 9 times.
http://www.wseas.org/multimedia/journals/computers/2014/a105705-605.pdf

14 Hassan, G. S., Mohammad, S. K., Alwan, F. M.
Categorization of Holy Quran-Tafseerusing K-nearest neighbor algorithm
(2015) *Int. J. Comput. Appl*, 129 (12), pp. 1-6. Cited 8 times.

15 Hamed, S.K., Aziz, M.J.A.
A question answering system on Holy Quran translation based on question expansion technique and Neural Network classification (Open Access)
(2016) *Journal of Computer Science*, 12 (3), pp. 169-177. Cited 32 times.
http://thescipub.com/PDF/jcssp.2016.169.177.pdf
doi: 10.3844/jcssp.2016.169.177
View at Publisher

16 Adeleke, A.O., Samsudin, N.A., Mustapha, A., Nawi, N.M.
Comparative analysis of text classification algorithms for automated labelling of Quranic verses (Open Access)
(2017) *International Journal on Advanced Science, Engineering and Information Technology*, 7 (4), pp. 1419-1427. Cited 17 times.
http://www.insightsociety.org/ojaseit/index.php/ijaseit/article/download/2198/1170
doi: 10.18517/ijaseit.7.4.2198
View at Publisher

17 Rostam, N.A.P., Malim, N.H.A.H.

Text categorisation in Quran and Hadith: Overcoming the interrelation challenges using machine learning and term weighting (Open Access)

(2019) *Journal of King Saud University - Computer and Information Sciences*. Cited 11 times.
www.journals.elsevier.com/journal-of-king-saud-university-computer-and-information-sciences/
doi: 10.1016/j.jksuci.2019.03.007

View at Publisher

18 Ulumudin, G.I., Adiwijaya, A., Mubarok, M.S.

A multilabel classification on topics of qur'anic verses in English translation using K-Nearest Neighbor method with Weighted TF-IDF (Open Access)

(2019) *Journal of Physics: Conference Series*, 1192 (1), art. no. 012026. Cited 4 times.
http://iopscience.iop.org/journal/1742-6596
doi: 10.1088/1742-6596/1192/1/012026

View at Publisher

19 Adeleke, A., Samsudin, N.A., Othman, Z.A., Ahmad Khalid, S.K.

A two-step feature selection method for quranic text classification (Open Access)

(2019) *Indonesian Journal of Electrical Engineering and Computer Science*, 16 (2), pp. 730-736. Cited 10 times.
http://ijeecs.iaescore.com/index.php/IJEECS/article/view/19977/13101
doi: 10.11591/ijeecs.v16.i2.pp730-736

View at Publisher

20 Elmitwally, N.S., Alsayat, A.

The multi-class classification for the first six surats of the Holy Quran (Open Access)

(2020) *International Journal of Advanced Computer Science and Applications*, 11 (1), pp. 327-332. Cited 4 times.
http://thesai.org/Publications/Archives?code=IJACSA
doi: 10.14569/ijacsa.2020.0110141

View at Publisher

21 Al-Khair, A. A, Kabbani, M. A.
"Quran teacher intonation
(2003)
The Tunisian Company for Distribution, 2003

22 Khedher, M. Z.
Multiword corpus of the Holy Quran
(2017) *International Journal on Islamic Applications in Computer Science And Technology*, 5 (1), pp. 1-11. Cited 3 times.

23 Khorsheed, M.S., Al-Thubaity, A.O.
Comparative evaluation of text classification techniques using a large diverse Arabic dataset

(2013) *Language Resources and Evaluation*, 47 (2), pp. 513-538. Cited 73 times.
doi: 10.1007/s10579-013-9221-8

View at Publisher

24 Alabbas, W., Al-Khateeb, H.M., Mansour, A.
Arabic text classification methods: Systematic literature review of primary studies (Open Access)

(2016) *Colloquium in Information Science and Technology, CIST*, 0, art. no. 7805072, pp. 361-367. Cited 18 times.
ISBN: 978-150900751-6
doi: 10.1109/CIST.2016.7805072

View at Publisher

25 Liu, X.-Y., Wu, J., Zhou, Z.-H.
Exploratory undersampling for class-imbalance learning

(2009) *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 39 (2), pp. 539-550. Cited 1227 times.
doi: 10.1109/TSMCB.2008.2007853

View at Publisher

26 Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.
SMOTE: Synthetic minority over-sampling technique (Open Access)

(2002) *Journal of Artificial Intelligence Research*, 16, pp. 321-357. Cited 10938 times.
http://www.cs.cmu.edu/afs/cs/project/jair/pub/volume16/chawla02a.pdf
doi: 10.1613/jair.953

View at Publisher

27 Tang, Y., Zhang, Y.-Q., Chawla, N.V.
SVMs modeling for highly imbalanced classification (Open Access)

(2009) *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 39 (1), pp. 281-288. Cited 589 times.
doi: 10.1109/TSMCB.2008.2002909

View at Publisher

28 Patel, H., Thakur, G.S.
Classification of imbalanced data using a modified fuzzy-neighbor weighted approach (Open Access)

(2017) *International Journal of Intelligent Engineering and Systems*, 10 (1), pp. 56-64. Cited 23 times.
http://www.inass.org/share/2017022807.pdf
doi: 10.22266/ijies2017.0228.07

View at Publisher

Arkok, B.; Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, Selangor, Malaysia; email:bassam.arkok@gmail.com

About Scopus

What is Scopus

Content coverage

Scopus blog

Scopus API

Privacy matters

Language

日本語に切り替える

切换到简体中文

切換到繁體中文

Русский язык

Customer Service

Help

Tutorials

Contact us