Classification of Quranic Topics Using Ensemble Learning

Arkok B., Zeki A.M.

Abstract

the real datasets in the world usually are imbalanced; the number of samples for their classes is not equal. Classifying these datasets makes the classifiers pay attention to the class with more samples
than the classes with fewer samples. The Qur’anic dataset can be considered an imbalanced dataset because verses of the Qur’anic topics are not equal. Many studies have been performed to classify Qur’anic text using different classifiers. However, few studies classified the Qur’anic verses based on Imbalanced Learning (IL). So, this work aims to classify the Qur’anic text using Ensemble methods, Boosting and Bagging. The base classifiers of these methods were LibSVM, Naive Bayes, KNN, and J48.

Three techniques are conducted in this paper based on the standard classifiers. The three techniques are: implementing the base classifiers alone, implementing these classifiers with the Boosting method, and implementing the classifiers with the Bagging method. The results showed that the Qur’anic classification performance was improved when the ensemble methods were applied for the imbalanced Qur’anic verses in the standard classifiers. © 2021 IEEE.

Author keywords
classification; ensemble methods; imbalanced; Quran; topics

Indexed keywords

SciVal Topics

Metrics

Funding details

References (39)

View in search results format

1. Khalilia, M., Chakraborty, S., Popescu, M.
Predicting disease risks from highly imbalanced data using random forest (Open Access)
http://www.biomedcentral.com/bmcmedinformdecismak/
View at Publisher

2. Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., Bing, G.
Learning from class-imbalanced data: Review of methods and applications
View at Publisher

3. Krawczyk, B., Schaefer, G.
An improved ensemble approach for imbalanced classification problems
ISBN: 978-1-4673-6400-3
doi: 10.1109/SACI.2013.6609011
View at Publisher
4 López, V., Fernández, A., García, S., Palade, V., Herrera, F.
An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics
doi: 10.1016/j.ins.2013.07.007
View at Publisher

5 Siddiqui, M.A., Faraz, S.M., Sattar, S.A.
Discovering the Thematic Structure of the Quran using Probabilistic Topic Model
ISBN: 978-147992823-1
doi: 10.1109/NOORIC.2013.55
View at Publisher

6 Al-Kabi, M.N.
Statistical classifier of the holy Quran verses (Fatiha and Yaseen chapters

7 Nassourou, M.
Using Machine Learning Algorithms for Categorizing Quranic Chapters by Major Phases of Prophet Mohammad’s Messengership

8 Al-Kabi, M.N.
A topical classification of Quranic Arabic text. In Proceedings of the 2013

9 Abdelnasser, H., Mohamed, R., Ragab, M., Mohamed, A., Farouk, B., El-Makky, N., Torki, M.
Al-Bayan: An Arabic Question Answering System for the Holy Quran
https://aclanthology.org/W14-36
ISBN: 978-193728496-1
View at Publisher

10 Akour, M., Alasmadi, I., Alazzam, I.
MQVC: Measuring quranic verses similarity and sura classification using N-gram
11 Sabbah, T., Selamat, A.
Support vector machine based approach for quranic words detection in online textual content
ISBN: 978-147995439-1
doi: 10.1109/MySec.2014.6986038
View at Publisher

12 Hassan, G.S., Mohammad, S.K., Alwan, F.M.

13 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)
doi: 10.3844/jcsssp.2016.169.177
View at Publisher

14 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)
doi: 10.18517/ojaset.7.4.2198
View at Publisher

15 Jamil, N.S., Ku-Mahamud, K.R., Din, A.M.

16 Jamil, N.S., Ku-Mahamud, K.R., Din, A.M., Ahmad, F., ChePa, N., Ishak, W.H.W., Din, R., (..), Ahmad, F.K.
A subject identification method based on term frequency technique (Open Access)
doi: 10.1910/IJACR.2017.730020
View at Publisher
17 Adeleke, A., Samsudin, N.
A Hybrid Feature Selection Technique for Classification of Group-based Holy Quran Verses

18 Adeleke, A., Samsudin, N., Mustapha, A., Ahmad Khalid, S.
Automating quranic verses labeling using machine learning approach (Open Access)

View at Publisher

19 Adeleke, A.O.

20 Izzaty, A.M.K., Mubarok, M.S., Huda, N.S., Adiwijaya
A multi-label classification on topics of quranic verses in English translation using Tree Augmented Naive Bayes
ISBN: 978-153864571-0
doi: 10.1109/ICoICT.2018.8528802

View at Publisher

21 Khaira Batubara, D., Arif, M., Adiwijaya, B.

22 Pane, R.A., Mubarok, M.S., Huda, N.S., Adiwijaya
A multi-label classification on topics of Quranic verses in English translation using multinomial naive bayes
ISBN: 978-153864571-0
doi: 10.1109/ICoICT.2018.8528777

View at Publisher
<table>
<thead>
<tr>
<th>ID</th>
<th>Author(s)</th>
<th>Title</th>
<th>Journal</th>
<th>Volume</th>
<th>Issue</th>
<th>Pages</th>
<th>Year</th>
<th>DOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>Hidayat, R., Minati, S.</td>
<td>Comparative Analysis of Text Mining Classification Algorithms for English and Indonesian Qur’an Translation. IJD</td>
<td>International Journal on Informatics for Development</td>
<td>8</td>
<td>1</td>
<td>47-51</td>
<td>2019</td>
<td></td>
</tr>
</tbody>
</table>
29 Ulumudin, G., Adiwijaya, A., Mubarak, M. 

30 Elmitwally, N.S., Alsayat, A. 
The multi-class classification for the first six surats of the Holy Quran (Open Access) 

http://theses.org/Publications/Archives?code=IJACSA 
View at Publisher

31 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) 

www.journals.elsevier.com/journal-of-king-saud-university-computer-and-information-sciences/ 
doi: 10.1016/j.jksuci.2019.03.007 
View at Publisher

32 Huda, N.S., Mubarak, M.S., Adiwijaya 
A multi-label classification on topics of qur'anic verses (english translation) using backpropagation neural network with stochastic gradient descent and adam optimizer 

http://ieeexplore.ieee.org/xp/mostRecentIssue.jsp?punumber=8825728 
ISBN: 978-153868052-0 
doi: 10.1109/ICoICT.2019.8835362 
View at Publisher

33 Abu Al-Khair, A., Kabbani, M. 
Koran teacher intonation 

34 Khedher, M.Z. 
Multword corpus of the Holy Quran 
35. Khorsheed, M.S., Al-Thubaity, A.O.
Comparative evaluation of text classification techniques using a large diverse Arabic dataset
View at Publisher

36. Alabbas, W., Al-Khateeb, H.M., Mansour, A.
Arabic text classification methods: Systematic literature review of primary studies (Open Access)
ISBN: 978-150900751-6
doi: 10.1109/CIST.2016.7805072
View at Publisher

37. Tang, Y., Zhang, Y.-Q., Chawla, N.V.
SVMs modeling for highly imbalanced classification (Open Access)
doi: 10.1109/TSMCB.2008.2002909
View at Publisher

38. Al-Azani, S., El-Afy, E.-S.M.
Using Word Embedding and Ensemble Learning for Highly Imbalanced Data Sentiment Analysis in Short Arabic Text (Open Access)
http://www.sciencedirect.com/science/journal/18770509
doi: 10.1016/j.procs.2017.05.365
View at Publisher

39. Patel, H., Thakur, G.S.
Classification of imbalanced data using a modified fuzzy-neighbor weighted approach (Open Access)
http://www.ijies.org/share/2017022807.pdf
doi: 10.22266/ijies2017.0228.07
View at Publisher

© Copyright 2021 Elsevier B.V, All rights reserved.