|  |  |  |  |
| --- | --- | --- | --- |
| Title | Multi-category Bangla News Classification using Machine Learning Classifiers and Multi-layer Dense Neural Network | | |
| Author(s) Name | Sharmin Yeasmin, Ratnadip Kuri, A R M Mahamudul Hasan Rana, Ashraf Uddin, A. Q. M. Sala Uddin Pathan, Hasnat Riaz | | |
| Contact Email(s) | dr.ashraf@aiub.edu | | |
| Published Journal Name | International Journal of Advanced Computer Science and Applications(IJACSA) | | |
| Type of Publication | Article | | |
| Volume | 12 | Issue | 5 |
| Publisher | The Science and Information (SAI) Organization | | |
| Publication Date | 31-5-2021 | | |
| ISSN | 2156-5570 | | |
| DOI | https://dx.doi.org/10.14569/issn.2156-5570 | | |
| URL | https://thesai.org/Publications/ViewPaper?Volume=12&Issue=5&Code=IJACSA&SerialNo=88 | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| Online and offline newspaper articles have become an integral phenomenon to our society. News articles have a significant impact on our personal and social activities but picking a piece of an appropriate news article is a challenging task for users from the ocean of sources. Recommending the appropriate news category helps find desired articles for the readers but categorizing news article manually is laborious, sluggish and expensive. Moreover, it gets more difficult when considering a resource-insufficient language like Bengali which is the fourth most spoken language of the world. However, very few approaches have been proposed for categorizing Bangla news articles where few machine learning algorithms were applied with limited resources. In this paper, we accentuate multiple machine learning approaches including a neural network to categorize Bangla news articles for two different datasets. News articles have been collected from the popular Bengali newspaper Prothom Alo to build Dataset I and dataset II has been gathered from the famous machine learning competition platform Kaggle. We develop a modified stop-word set and apply it in the preprocessing stage which leads to significant improvement in the performance. Our result shows that the Multi-layer Neural network, Naïve Bayes and support vector machine provide better performance. Accuracy of 94.99%, 94.60%, 95.50% has been achieved for SVM, Logistic regression and Multi-layer dense Neural network, respectively. | |