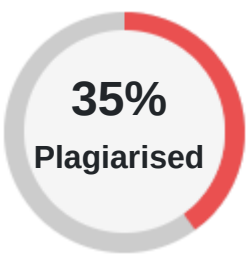
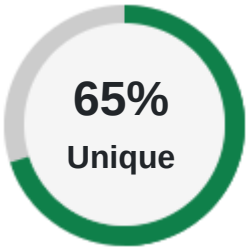
Plagiarism Detector.net

Date: April, 05 2021

PLAGIARISM SCAN REPORT





Excluded Url : None

Content Checked For Plagiarism

ABSTRACT: With the help of Internet, the online news can be instantly spread around the world. Most of the people now have the habit of reading and sharing news online, for instance, using social media like Twitter and Facebook. Typically, the news popularity can be indicated by the number of reads, likes or shares. For the online news stake holders, it's very valuable if the popularity of the news articles can be accurately predicted prior to the publication. Thus, it is interesting and meaningful to use the machine learning techniques to predict the popularity of online news articles. In this project, based on the dataset including 39,643 news articles from website Mashable, we try to find the best classification learning algorithm to accurately predict if a news article will become popular or not prior to publication. Our research reflects that predicting news popularity with the help of tweets in social media platforms based on their shares and hashtags used during weekdays and weekends being done using different classification algorithms i.e linear regression, random forest, adaboost gives different result and the most accurate result obtained by our research is from Random forest classification algorithm giving an accuracy rate of 0.6743 which is highest amongst all the other classification algorithms we've used. 2) INTRODUCTION: Working with machine learning algorithms in the large dataset is very common and especially with the expansion of online news, it became very useful. Random Forest, Linear Regression and Adaboost are most common machine learning algorithms used for classification. In this research, we aimed to find the best model and set of features to predict the popularity of online news, using machine-learning techniques and implement various machine learning algorithms on the selected features.The data source was Mashable, a well-known online news website. Precision, AUC (Area Under the Curve) and F-measure were used to evaluate the results and their results were compared to find the most accurate amongst all. Random Forest turns out to be the best model for prediction, and can achieve an accuracy of around 67%. Our work can help online news companies to predict news popularity before publication. Various works have been done in prediction of online news popularity. In [1], the popularity of online articles is analyzed based on the user's comments. [2] defines the popularity in terms of a competition where the popular articles are those which were the most appealing on that particular day. Ranking Support Vector Machine (SVM) is used to classify the appealing/non appealing of online news article. In [3], the number of retweets is predicted using both the features of the retweet content (length, words, number of hashtag, etc.) and the features of author (number of followers, friends, etc.). [4] collects a dataset with almost 40,000 articles from the Mashable website, compares five different methods on classifying popular/unpopular news articles and concludes that the Random Forest (RF) can achieve the best performance. 2.1) Data Exploration - The dataset consists of 39,643 news articles from an online news website called Mashable collected over 2 years from the time period of Jan. 2013 to Jan. 2015. It is downloaded from UCI Machine Learning Repository as https://archive.ics.uci.edu/ml/datasets/Online+News+ Popularity# and this dataset is generously denoted by the author of [4]. For each instance of the dataset, it has 61 attributes which includes 1 target attribute (number of shares), 2 non-predictive features (URL of the article and days between the article publication and the dataset acquisition) and 58 predictive features as shown in Fig. 1. For example, the categorical features like the published day of the week and article category have been transformed by one-hot encoding scheme, and the skewed feature like number of words in the article has been log transformed.

5% Plagiarized

With the help of Internet, the online news can be instantly spread around the world. Most of peoples now have the habit of reading and sharing news onli
instance, using social media like ...

https://www.researchgate.net/publication/339909346_A_Comprehensive_Review_on_Online_News_Popularity_Prediction_using_Machine_Learning_Ap

5% Plagiarized

Typically, the news popularity can be indicated by the number of reads, likes or shares. For the online news stake holders such as content providers or advertisers, it's very valuable if the ...

https://www.researchgate.net/publication/224187404_An_Approach_to_Model_and_Predict_the_Popularity_of_Online_Contents_with_Explanatory_Fact

10% Plagiarized

Various works have been done in prediction of online news popularity. Popularity of news depends upon various features like sharing of online news on social media, comments of visitors for news ...

https://www.researchgate.net/publication/220751760_Predicting_the_popularity_of_online_articles_based_on_user_comments

15% Plagiarized

by F Namous · 2018 · Cited by 1 — In this research, we aimed to find the best model and set of features to predict the popularity of online news, using machine-learning techniques and implement various data mining algorithms on the selected features. The data source was Mashable, a well-known online news website.

https://ieeexplore.ieee.org/document/8649529