

Fake News Detection Using Machine Learning Algorithms

Project overview

EVALUATION MATRICES

Evaluate the performance of algorithms for fake news detection problem; various evaluation metrics have been used. In this subsection, we review the most widely used metrics for fake news detection. Most existing approaches consider the fake news problem as a classification problem that predicts whether a news article is fake or not:

True Positive (TP): when predicted fake news pieces are actually classified as fake news;

True Negative (TN): when predicted true news pieces are actually classified as true news;

False Negative (FN): when predicted true news pieces are actually classified as fake news; False

Positive (FP): when predicted fake news pieces are actually classified as true news.

Confusion Matrix:

A confusion matrix is a table that is often used to describe the performance of a classification model (or “classifier”) on a set of test data for which the true values are known. It allows the visualization of the performance of an algorithm. A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix. The confusion matrix shows the ways in which your classification model is confused when it makes predictions. It gives us insight not only into the errors being made by a classifier but more importantly the types of errors that are being made [26].

Table 1: Confusion Matrix

| Total | Class 1 (Predicted) | Class 2 (Predicted) |
|------------------|---------------------|---------------------|
| Class 1 (Actual) | TP | FN |
| Class 2 (Actual) | FP | TN |

By formulating this as a classification problem, we can define following metrics-

$$1. \text{ Precision} = \frac{|T P|}{|T P| + |F P|}$$

$$2. \text{ Recall} = \frac{|T P|}{|T P| + |F N|}$$

$$3. \text{ F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$4. \text{ Accuracy} = \frac{|T P| + |T N|}{|T P| + |T N| + |F P| + |F N|}$$

These metrics are commonly used in the machine learning community and enable us to evaluate the performance of a classifier from different perspectives. Specifically, accuracy measures the similarity between predicted fake news and real fake news.

4.5 SNAPSHOTS OF SYSTEM WORKING

A. Static System-

```
UserWarning)
The given statement is True
The truth probability score is 0.6202405257600963

(base) C:\Users\HP\Desktop\fake news detetction\Fake_News_Detection>
```

Figure 3: Static output (True)

```

The given statement is False
The truth probability score is 0.3221557972557687

(base) C:\Users\HP\Desktop\fake news detetction\Fake_News_Detection>

```

Figure 4: Static Output (False) B.

Dynamic System-

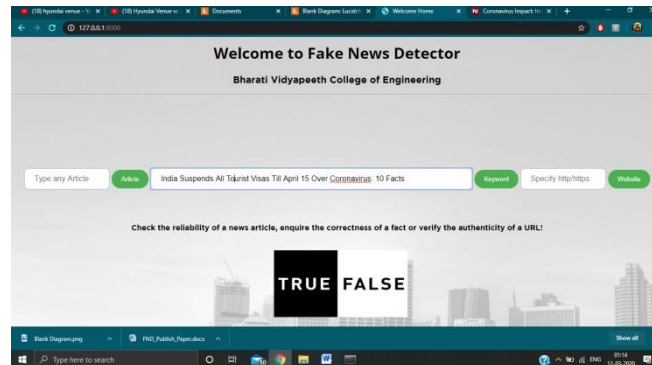


Figure 5: Fake News Detector (Home Screen)

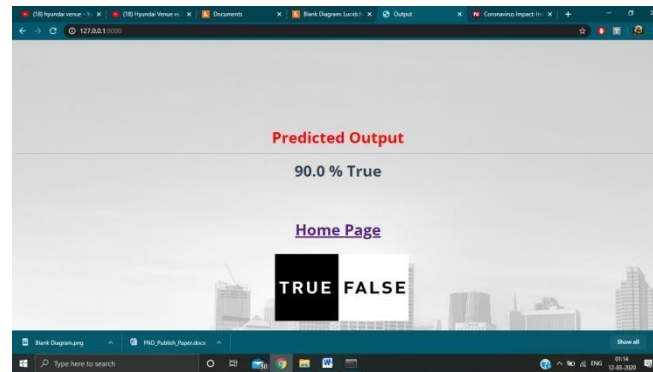


Figure 6: Fake News Detector (Output page)

V. RESULTS

Implementation was done using the above algorithms with Vector features- Count Vectors and Tf-Idf vectors at Word level and Ngramlevel. Accuracy was noted for all models. We used K-fold cross validation technique to improve the effectiveness of the models.

A. Dataset split using K-fold cross validation

This cross-validation technique was used for splitting the dataset randomly into k-folds. (k-1) folds were used for building the model while k^{th} fold was used to check the effectiveness of the model. This was repeated until each of the k-folds served as the test set. I used 3fold cross validation for this experiment where 67% of the data is used for training the model and remaining 33% for testing.

B. Confusion Matrices for Static System

After applying various extracted features (Bag-of-words, Tf-Idf, N-grams) on three different classifiers (Naïve bayes, Logistic Regression and Random Forest), their confusion matrix showing actual set and predicted sets are mentioned below:

Table 2: Confusion Matrix for Naïve Bayes Classifier using Tf-Idf features-

| Total= 10240 | Naïve Bayes Classifier | |
|---------------|-------------------------|-------------------------|
| | <i>Fake (Predicted)</i> | <i>True (Predicted)</i> |
| Fake (Actual) | 841 | 3647 |
| True (Actual) | 427 | 5325 |

Table 3: Confusion Matrix for Logistic Regression using Tf-Idf features-

| Total= 10240 | Logistic Regression |
|--------------|---------------------|
|--------------|---------------------|

| | <i>Fake (Predicted)</i> | <i>True (Predicted)</i> |
|---------------|-------------------------|-------------------------|
| Fake (Actual) | 1617 | 2871 |
| True (Actual) | 1097 | 4655 |

Table 4: Confusion Matrix for Random Forest Classifier using Tf-Idf features-

| Total= 10240 | Random Forest | |
|---------------|-------------------------|-------------------------|
| | <i>Fake (Predicted)</i> | <i>True (Predicted)</i> |
| Fake (Actual) | 1979 | 2509 |
| True (Actual) | 1630 | 4122 |

Table 5: Comparison of Precision, Recall, F1-scores and Accuracy for all three classifiers-

| Classifiers | Precision | Recall | F1-Score | Accuracy |
|---------------------|-----------|--------|----------|----------|
| Naïve Bayes | 0.59 | 0.92 | 0.72 | 0.60 |
| Random Forest | 0.62 | 0.71 | 0.67 | 0.59 |
| Logistic Regression | 0.69 | 0.83 | 0.75 | 0.65 |

As evident above our best model came out to be Logistic Regression with an accuracy of 65%. Hence we then used grid search parameter optimization to increase the performance of logistic regression which then gave us the accuracy of 80%.

Hence we can say that if a user feed a particular news article or its headline in our model, there are 80% chances that it will be classified to its true nature.

C. Confusion Matrix for Dynamic System

We used real_or_fake.csv with passive aggressive classifier and obtained the following confusion matrix-

Table 6: Confusion Matrix for passive aggressive Classifier-

| Total= 1267 | Passive Aggressive Classifier | |
|---------------|-------------------------------|-------------------------|
| | <i>Fake (Predicted)</i> | <i>True (Predicted)</i> |
| Fake (Actual) | 588 | 50 |
| True (Actual) | 42 | 587 |

Table 7: Performance measures-

| Classifier | Precision | Recall | F1-Score | Accuracy |
|------------|-----------|--------|----------|----------|
| PAC | 0.93 | 0.9216 | 0.9257 | 0.9273 |

VI. CONCLUSION

In the 21st century, the majority of the tasks are done online. Newspapers that were earlier preferred as hard-copies are now being substituted by applications like Facebook, Twitter, and news articles to be read online. Whatsapp's forwards are also a major source. The growing problem of fake news only makes things more complicated and tries to change or hamper the opinion and attitude of people towards use of digital technology. When a person is deceived by the real news two possible things happen- People start believing that their perceptions about a particular topic are true as assumed. Thus, in order to curb the phenomenon, we have developed our Fake news Detection system that takes input from the user and classify it to be true or fake. To implement this, various NLP and Machine Learning Techniques have to be used. The model is trained using an appropriate dataset and performance evaluation is also done using various performance measures.

The best model, i.e. the model with highest accuracy is used to classify the news headlines or articles. As evident above for static search, our best model came out to be Logistic Regression with an accuracy of 65%. Hence we then used grid search parameter optimization to increase the performance of logistic regression which then gave us the accuracy of 75%. Hence we can say that if a user feed a particular news article or its headline in our model, there are 75% chances that it will be classified to its true nature. The user can check the news article or keywords online; he can also check the authenticity of the website. The accuracy for dynamic system is 93% and it increases with every iteration.

We intend to build our own dataset which will be kept up to date according to the latest news. All the live news and latest data will be kept in a database using Web Crawler and online database.