Fake News Predictor

By
Manan Arya
01996303116

Abstract of Project:

This project is based on Natural Language Processing on textual data. The project classifies news data into 'possibly fake' and 'authentic' by creating a sparse matrix and creating a 'bag-of-words' model

Company Profile

OSAHUB Technologies is India's leading training provider. Focus on education solutions, covering all the latest technologies - be it in the field of Computer Science, Electronics, Mechanical Engineering, Robotics, Animation, etc.Provide Workshops, Trainings, Faculty Development Programs, Certifications Courses, Seminars, Career Counseling, etc. to engineering colleges and schools all across India. We have an extremely skilled network of subject matter experts and highly experienced trainers who have several years of experience in the industry. We are pioneers in the education industry and are proud of our excellent team.

Week wise breakdown of Summer Training

Week 1
* Introduction
* Python loops and Data Structures

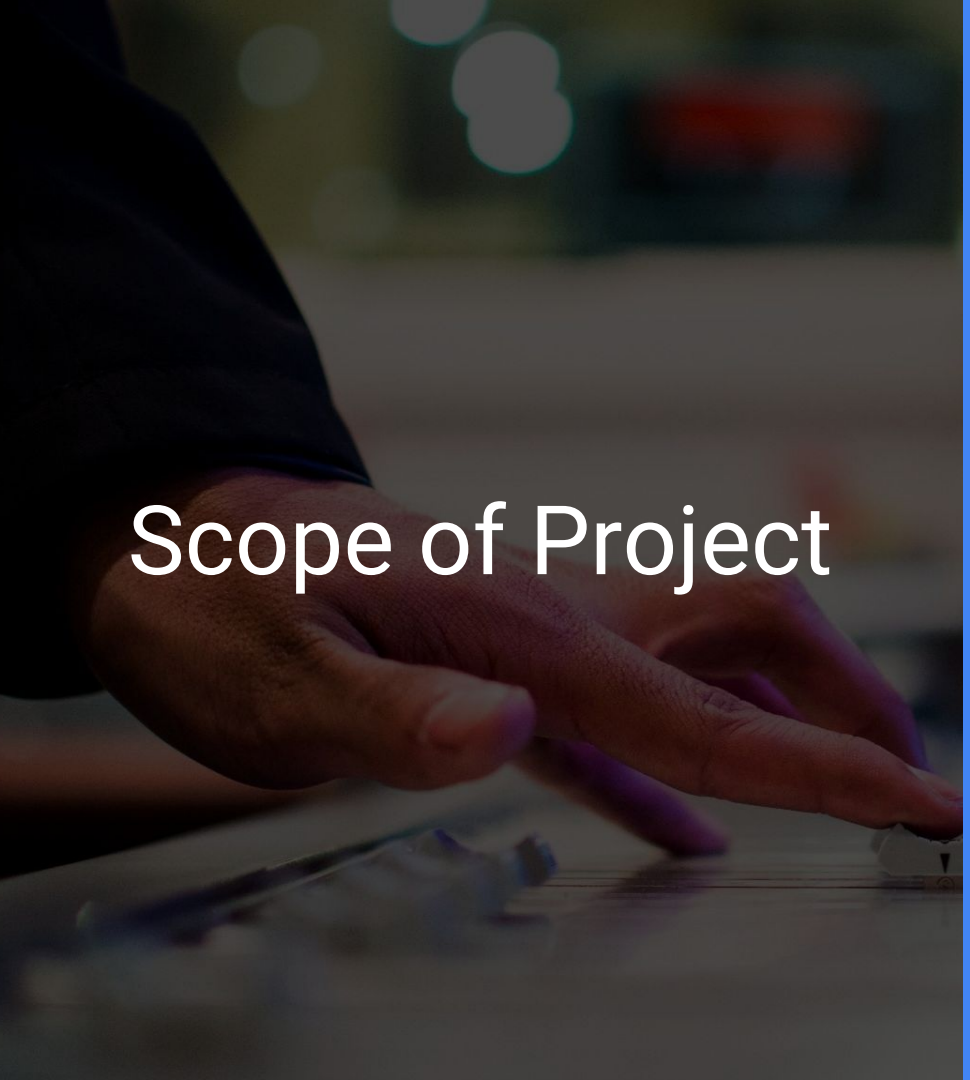
Week 2
* Pandas
* Numpy
* Bike Share Project
* Introduction to Machine Learning

Week 3
* Supervised Algorithms
* KNN
* Linear Regression
* Optimization using gradient Descent
* sk-learn
* Predicting house prices

Week 4
* Decision Trees
* XG Boosting
* Identifying Titanic Survivorship
* Identifying Iris Flowers
* Identifying Iris Flowers ,Unsupervised Learning
* K Means
* Clustering stocks using k means

Week 5
* Non-negative Matrix Factorization Building recommender Systems
* Fake News Classifier
* Support Vector Machines
* Histogram of Oriented Gradients
* DLIB
* Face Detection using DLIB

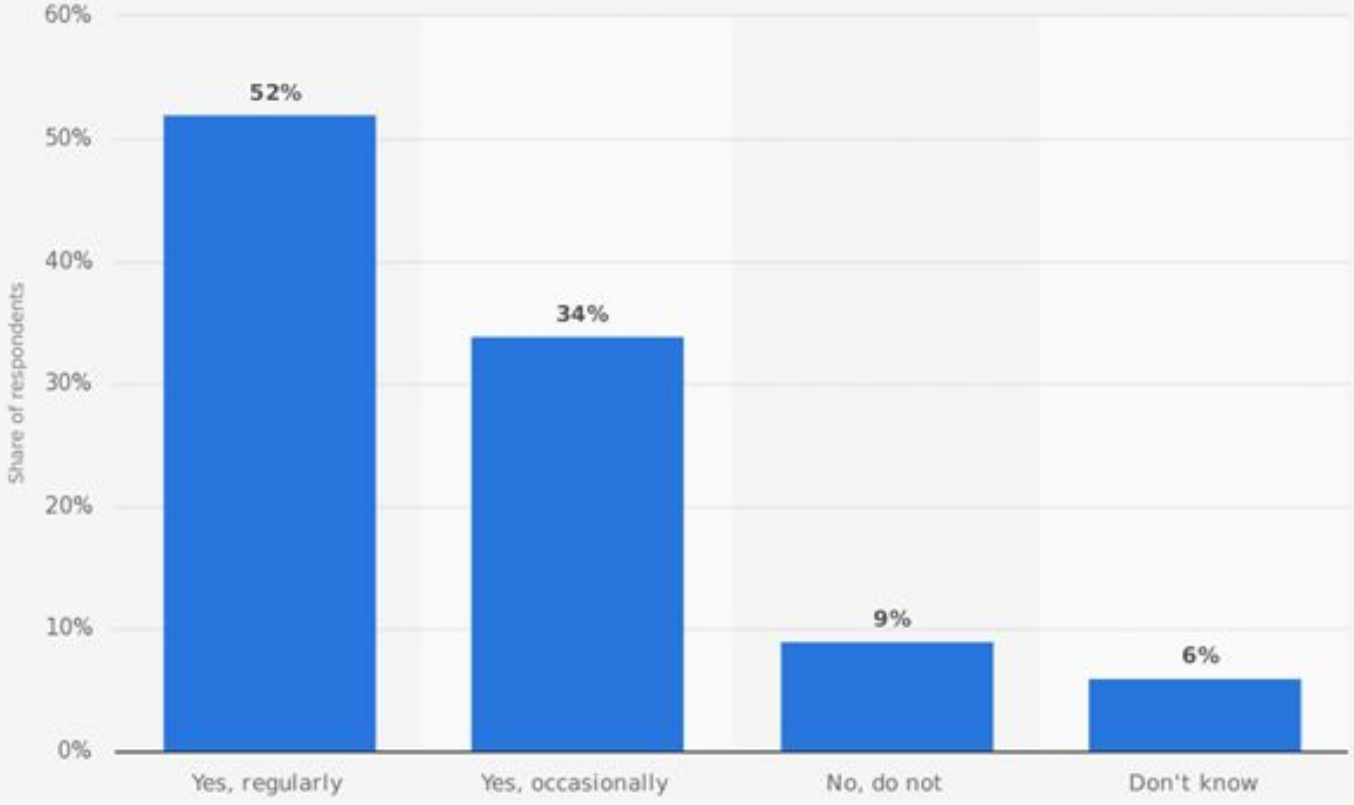
Week 6
* Neural Networks
* Multi level perceptron
* Back Propagation
* Convolution Neural Networks
* MNIST classification
* CIFAR Classification
* Dog Breed Classification.



Scope of Project

This Project tries to determine the authenticity of a news. This is a step forward to prohibit the generation of fake news these days. It can further upgraded to increase the accuracy upto 100%

Perceived frequency of online news websites reporting fake news stories in the United States as of March 2018

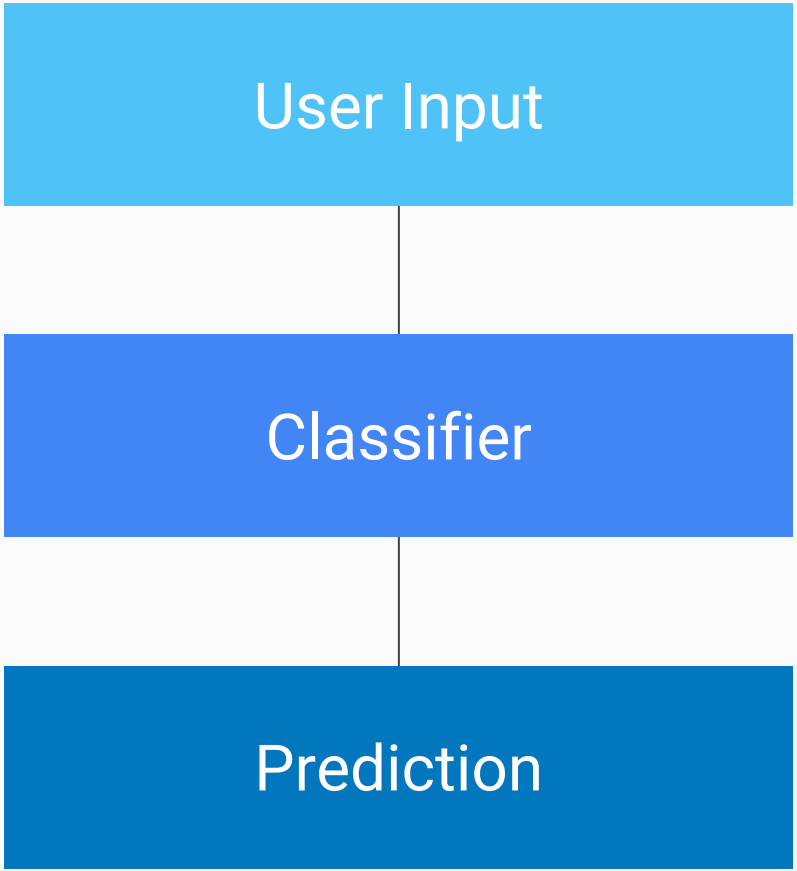


Source
Monmouth University

Additional Information:
United States; March 2 to 5, 2018; 803 respondents; 18 years and older

Project Model

This Project is basically classified in three steps.

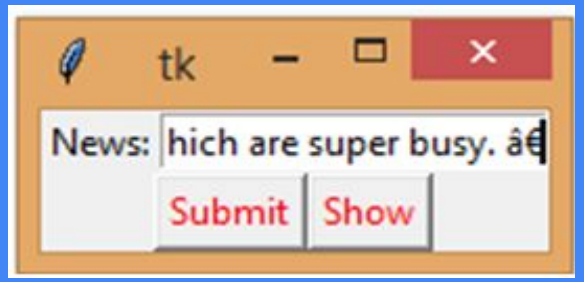


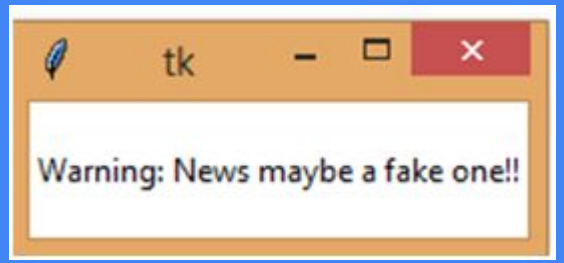
Bag of words

- sparse matrix
- numbers can be:
  - binary - 0/1
  - simple term frequency
  - weight - e.g. TF-IDF

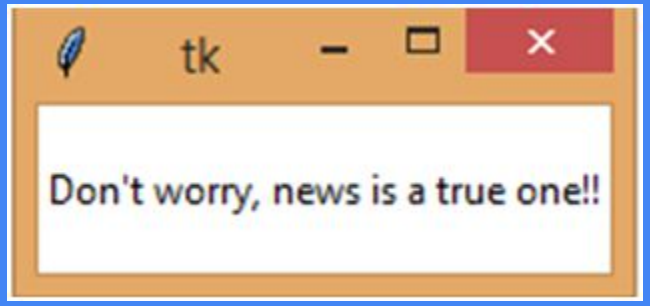
Terms / Documents	Investment Risk	Project Management	Software Engineering	Development	SAP	...
Document 1	1			1		
Document 2		1				
Document 3			3		1	
Document 4		1				
Document 5			2	1		
Document 6	1			1		
...						

Sample Outputs






tk
News:
warships travel to Syria
Submit
Show


tk
Don't worry, news is a true one!!

Conclusion

The Fake News Detector predicts the authenticity of any news on the basis of mapping the input news text with the 'bag of words' created by training a machine learning model, making a matrix of more than 16000 columns,on a dataset having labels specified to each corresponding news as 'fake' or 'authentic'. The detector can be used to predict authenticity of any news preferably originating from the United States as the machine learning model is treated primarily on the U.S dataset.

The model is reliable with an accuracy of 96%.


Thankyou