

**A REPORT**

**ON**

**Swecha Fake News Detection**

**BY**

**Abhinav Verma**

**2020A7PS1093H**

**AT**

**Swecha Telangana**

**A Practice School-I Station of**

**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI**

**June 2022 - July 2022**

**A REPORT ON**

**Swecha Fake News Detection**

BY

Abhinav Verma

B.Tech Computer Science, BITS Hyd

Prepared in partial fulfillment of the Practice  
School-I Course Nos.

BITS C221/BITS C231/BITS C241 AT

Swecha Telangana

A Practice School-I Station of

**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI**

**June 2022 - July 2022**

## Acknowledgement

With immense pleasure, I , Abhinav Verma, 2020A7PS1093H, undergraduate student of BITS Pilani, Hyderabad Campus present end report of Practice School I (PS I ) held during June 2022 - July 2022 at PS station - **Swecha Telangana** .

I wish to thank my Swecha mentor - Mr. Harisai and my faculty incharge - Mr. Supratim Ghosh for the unending support and guidance they gave me during my entire 2 month long internship.

# **BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI (RAJASTHAN)**

**Practice School Division**

**Station:** Swecha Telangana

**Duration** June 2022 - July 2022 **Date of Start.** 1 June 2022

**Date of Submission** 22 July 2022

## **SWECHA FAKE NEWS DETECTION**

**ID/NAME/DISCIPLIN -** 2020A7PS1093H  
Abhinav Verma  
B.Tech CSE

**NAME AND DESIGNATION OF EXPERT -** Mr Harisai  
Senior Executive Swecha

**NAME OF THE PS FACULTY -** Mr Supratim Ghosh

**Key Words:** NLP, Python, Web Scraping, Solr Database, Keras, Pandas

**Project Areas:** Artificial Intelligence, Natural Language Processing, Web Scraping, Solr Database, Fake News

**Abstract:** A PYTHON MODULE TO DETECT FAKE NEWS USING NLP and SOLR DATABASE.



Signature(s) of Student(s)



Signature of PS Faculty

# Table Of Content

Introduction	1
Main Text	2
Conclusion / Recommendation	4
References	5

# INTRODUCTION

At Swecha Telangana, we students solved a complex yet simple looking problem of fake news detection using state of the art python libraries which includes Numpy, Pandas, Scikit Learn, NLP, BeautifulSoup and Keras.

The application uses pre-stored database of authentic news collected from reliable sources using web scraping.

As an intern and the member of Data Science team, I was incharge of the web scraping. I ended up creating a python script which extracted data from [www.inshorts.com](http://www.inshorts.com) and merged the data with already existing data in Solr database.

We were divided into four teams, each team looking over a particular segment of the project.

## MAIN TEXT

Fake news has existed since the advent of the internet. With over 560 million users, India is the second largest online market just after China, With almost no stringent rules for publishing news for publishers, use of flashy news has led to an increase in fake news on the internet. At Swecha this problem was targeted by the advent of Swecha Fake News Detection module.

In brief, a lot of data was collected from scraping websites, using an automatic web scraper which on periodic intervals used to merge freshly collected news with the already present news in the Solr database, a practice also called as online learning in machine learning world. This ensures easy access to all the data up to date.

The project used BeautifulSoup a popular python module for extracting data from web.

Further using NLP each news was mapped to the corresponding article from where it was picked. It is essential at the later half of the project.

When a particular news was fed to the system, the application tried to map the news

with the already present data from database, if matched it outputs the result as true else as fake.

In order to increase the efficiency a lot of common ML practices were conducted such as data cleaning, batch learning, negative false inputs.



## **Conclusion**

In conclusion, I learned a lot about how to work in team by being a part of Data Science team and also how to meet deadlines in a corporate organization.

Along with that Swecha turned out to be an excellent organization to work with, which taught me a lot of technical skills, which includes Python and Beautiful Soup Module.

I also learned how difficult a simple looking problem could be when we try to implement it.

## References

<https://www.python.org/>

<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

<https://karyalayam.in/s/2Ex67Ws32SSQd3T>

<https://www.nltk.org>

<https://keras.io>

<https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.html>