

**School Of Computer Science**

**UNIVERSITY OF PETROLEUM & ENERGY STUDIES,**

**DEHRADUN- 248007. Uttarakhand**

**PROJECT REPORT**

**on**

**Fake News Detection**

**Submitted By:**

|  |  |  |
| --- | --- | --- |
| **Name** | **SAP ID** | **Specialization/Batch** |
| Nihar | 500091867 | AI & ML (Hons) - B1 |

**Submitted To:**

Dr. Achala Shakya

**INDEX**

|  |  |  |
| --- | --- | --- |
| **Topic** | | **Page No** |
| Revision History | | 2 |
| 1 | Introduction | 3 - 5 |
|  | 1.1 Literature Review |  |
|  | 1.2 Problem Statement |  |
|  | 1.3 Motivation |  |
|  | 1.4 Object |  |
|  | 1.5 System Requirement |  |
|  |  |  |
| 2 | Project Description | 6 - 8 |
|  | 2.1 Flowchart |  |
|  | 2.2 Methodology |  |
|  | 2.3 Code Snippets |  |
|  | 2.4 Result |  |
|  | 2.5 Future Scope |  |
|  |  |  |
| 3 | Conclusion and References | 8 |
|  | 3.1 Conclusion |  |
|  | 3.2 References |  |

**Introduction: -**

Fake news has become a widespread issue in today’s digital age, with misinformation spreading rapidly across online platforms. To solve this problem, we developed a fake news detection model using machine learning. Our system utilizes sensitivity analysis, TF-IDF vectorization, and logistic regression to classify news stories as fake or true.

**Literature Review: -**

We have performed a comprehensive literature review comparing different research papers on Fake News Detection. So, let's start:

**FAKE NEWS DETECTION USING LOGISTIC REGRESSION ALGORITHM WITH MACHINE LEARNING [1]**

The paper addresses the issue of fake news on social media and the importance of authenticating news articles circulated on platforms like WhatsApp, Facebook, Twitter, and other social networking sites. The authors emphasize the societal harm caused by fake news and the need to focus on correct, authenticated news articles, especially in developing countries like India. The paper presents a model and methodology for detecting fake news using machine learning and natural language processing techniques. The proposed model utilizes logistic regression to determine the authenticity of news and claims to achieve an accuracy of 97.21%. The authors note that fake news can cause significant harm before it is debunked and that social media sites play a major role in spreading false news.

**A Comprehensive Review on Fake News Detection with Deep Learning [2]**

The paper provides a comprehensive review of fake news detection with deep learning, comparing existing surveys, discussing various deep learning approaches, and highlighting the consequences of fake news. It mentions the use of Graph Neural Networks (GNN) for fake news detection and the advantages of Gated Recurrent Unit (GRU) over Long Short-Term Memory (LSTM) in certain applications. Additionally, it suggests combining GRU and Convolutional Neural Networks (CNN) for improved results and encourages researchers to experiment with models like SeqGAN and Deep Belief Network (DBN). The model also notes the shift towards Transformers in natural language processing tasks and recommends using Generative Pre-trained Transformer (GPT) for fake news detection.

**Fake News Detection using Machine Learning [3]**

The paper discusses fake news detection using machine learning techniques such as Support Vector Machine, Naive Bayes, and Passive Aggressive Classifier. The model presented in the document achieved an accuracy of 95.05% using feature extraction techniques like TF-IDF. The rise of fake news on social media is highlighted as a major concern, and the importance of detecting and combating fake news is emphasized.

**Fake News Detection Using AdaBoost & Bagging Classifier [4]**

The paper discusses the detection of fake news using machine learning, focusing on the LIAR dataset collected from POLITIFACT.COM. It highlights the challenges posed by fake news and the importance of accurate detection. Various techniques, including ensemble methods like Bagging Classifier and AdaBoost, are used to improve accuracy. The conclusion emphasizes the significance of distinguishing between real and fake news, with models achieving accuracies comparable to human judgment.

**Automatic Detection of Fake News [5]**

The document discusses the automatic detection of fake news, presenting datasets and models for fake news detection. It highlights challenges with using satirical sources and fact-checking websites. The study includes experiments to build accurate fake news detectors and compares automatic and manual identification methods. The document also addresses challenges in collecting deceptive data and presents a dataset of fake news from web sources. Additionally, it mentions the use of Amazon Mechanical Turk (AMT) for data collection and the creation of a crowdsourced dataset called FakeNewsAMT.

**Problem Statement:-**

Develop an automated system that accurately distinguishes between real and fake news articles using machine learning and NLP techniques, such as sentiment analysis and TF-IDF, with the ultimate goal of enhancing information credibility and combating misinformation.

**Motivation:-**

In our interconnected digital landscape, misinformation spreads rapidly, leading to conflicts and undermining trust. By creating robust tools for fake news detection, our aim is to empower individuals with accurate information, foster informed decision-making, and ultimately contribute to a more reliable and trustworthy information ecosystem.

**Objective:-**

The main objective of our project is to develop an automated system for detecting fake news. By leveraging machine learning algorithms, such as logistic regression, sentiment analysis, and TF-IDF vectorization, we aim to classify news articles as either fake or real.

**System Required:-**

Python:

All operating systems require Python as it's the primary programming language for machine learning and data science.

Install Python from the official website or via package managers like Homebrew (macOS) or Chocolatey (Windows).

Integrated Development Environment (IDE):

Choose an IDE for coding convenience. Popular options include:

* Jupyter Notebook / JupyterLab
* Visual Studio Code
* Google Colab

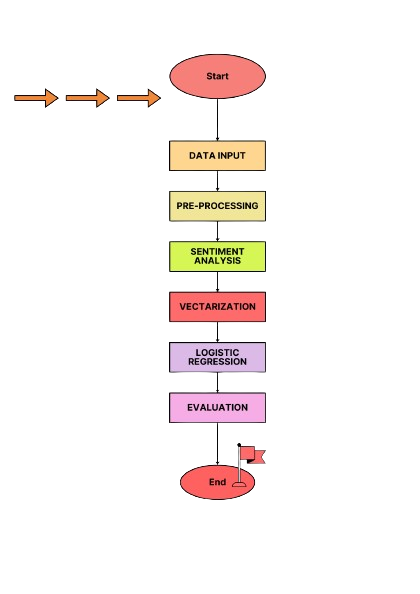
Ensure compatibility with your operating system.

Install the necessary Python libraries for machine learning, such as:

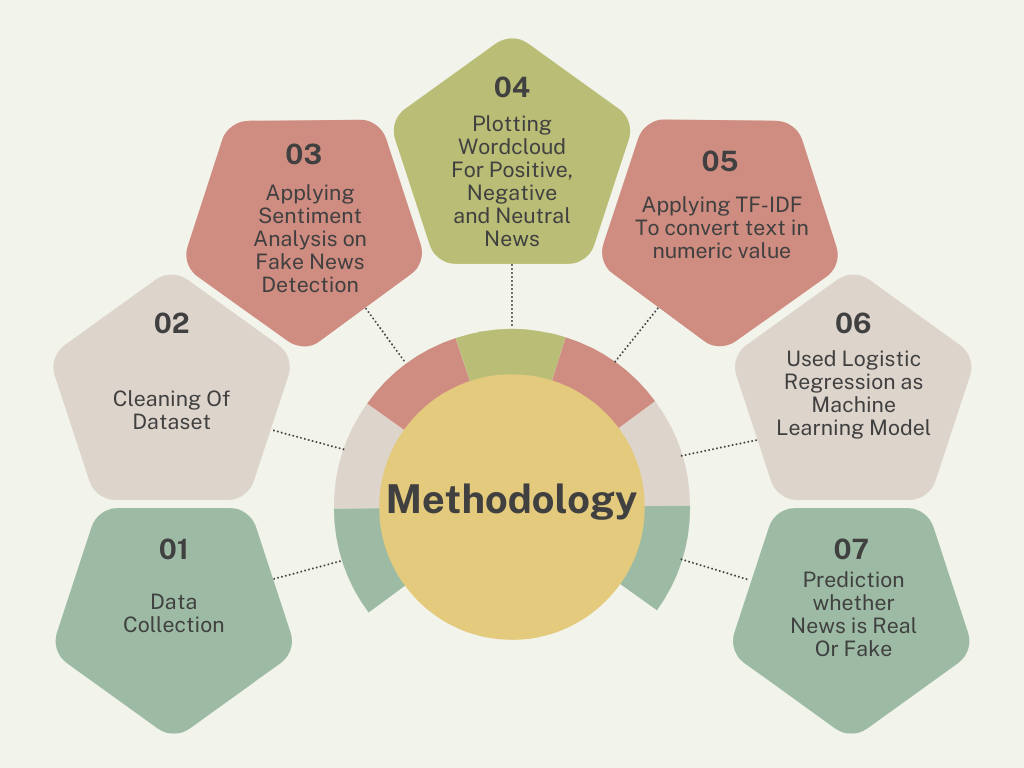
* NumPy
* Pandas
* TextBlob
* NLTK

**Project Description**

1. **Flow Chart Diagram:-**



1. **Methodology:-**



1. **Code Snippets:-**

**Dataset:-** [**https://upesstd-my.sharepoint.com/:x:/g/personal/500091867\_stu\_upes\_ac\_in/ERwvXEqqhB1MpxAdgxQ79KUB9i4wL7x\_84IASYG5CUrhlw?e=pHB62U**](https://upesstd-my.sharepoint.com/:x:/g/personal/500091867_stu_upes_ac_in/ERwvXEqqhB1MpxAdgxQ79KUB9i4wL7x_84IASYG5CUrhlw?e=pHB62U)

1. **Result:-**
2. **Future Scope:-**

In the future, our fake news detection model can be used to detect fake news in real-time, which can directly integrate with social media platforms, improve user interface, and explore advanced machine learning techniques for better accuracy.

**Conclusion and References**

* **Conclusion:-**

In conclusion, our project has successfully developed a fake news detection system using machine learning techniques. Through the implementation of logistic regression, sentiment analysis, and TF-IDF vectorization, we have created a tool capable of classifying news articles as fake or real.

* **References**

1. Dr.M. RAJESWARI, A.SRINIRANJANI, “FAKE NEWS DETECTION USING LOGISTIC REGRESSION ALGORITHM WITH MACHINE LEARNING” , JETIR June 2022, Volume 9, Issue 6, ISSN-2349-5162
2. M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar and M. S. Rahman, "A Comprehensive Review on Fake News Detection With Deep Learning," in IEEE Access, vol. 9, pp. 156151-156170, 2021, doi: 10.1109/ACCESS.2021.3129329
3. J. Shaikh and R. Patil, "Fake News Detection using Machine Learning," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur Odisha, India, 2020, pp. 1-5, doi: 10.1109/iSSSC50941.2020.9358890.
4. Lilapati Waikhom, Rajat Subhra Goswami “Fake News Detection Using AdaBoost & Bagging Classifier”(October 2, 2019). Proceedings of International Conference on Advancements in Computing & Management (ICACM) 2019, Available at SSRN: <https://ssrn.com/abstract=3462938> or [http://dx.doi.org/10.2139/ssrn.3462938](https://dx.doi.org/10.2139/ssrn.3462938)
5. [Verónica Pérez-Rosas](https://arxiv.org/search/cs?searchtype=author&query=P%C3%A9rez-Rosas,+V), [Bennett Kleinberg](https://arxiv.org/search/cs?searchtype=author&query=Kleinberg,+B), [Alexandra Lefevre](https://arxiv.org/search/cs?searchtype=author&query=Lefevre,+A), [Rada Mihalcea](https://arxiv.org/search/cs?searchtype=author&query=Mihalcea,+R) “Automatic Detection of Fake News” 2017 ARXIV. arXiv:1708.07104(or arXiv:1708.07104v)
6. Xichen Zhang, Ali A. Ghorbani” Information Processing & Man Information Processing & Management”. Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB), Fredericton, NB E3B 5A3, Canada