CLICKBAIT DETECTOR USING MACHINE LEARNING

Students' information
Jatin Meena
meenajatin145@gmail.com 9981751813
Manya Nayak
manyanayak07@gmail.com 9301678982
Mohit Chouhan
mc3984925@gmail.com 9630570135

Supervisor
Mr. Lalit Gehlod

25/8/25

Department of Computer Engineering

Keywords: Clickbait, Machine Learning, Text Classification, Natural Language Processing (NLP), Web Application

1. INTRODUCTION

In today's digital era, online platforms and social media have become the primary sources of news and information. The growing competition for user attention has given rise to the practice of **clickbait**, where headlines are deliberately exaggerated or misleading to encourage users to click. While this strategy may boost engagement metrics, it often results in disappointment for readers, reduces the credibility of genuine content, and contributes to the spread of misinformation.

Detecting clickbait automatically has emerged as an important challenge in the field of text analytics. Traditional rule-based methods have shown limited effectiveness due to the dynamic nature of language and headline styles. On the other hand, **Machine Learning (ML)** combined with **Natural Language Processing (NLP)** provides an efficient way to analyze linguistic patterns and classify text. Techniques such as tokenization, stopword removal, and lemmatization enable computers to interpret text more effectively, while feature extraction methods like **TF-IDF** help capture the importance of words in distinguishing between clickbait and non-clickbait content.

This project proposes a **Clickbait Detection System** that uses **Logistic Regression** as the baseline machine learning model. The system will be developed in a modular manner so that it can be extended with advanced models like Support Vector Machines or Random Forests. For deployment, the trained model will be integrated into a **FastAPI backend**, with a simple **web application interface** for user interaction.

The motivation for this project lies in building a lightweight and practical solution that improves the trustworthiness of online content while also strengthening our understanding of ML and NLP applications in real-world problems.

2.Literature Survey

Clickbait detection has attracted significant research interest due to its role in misleading readers and reducing the credibility of digital media. Several approaches have been proposed, ranging from traditional machine learning to advanced deep learning techniques.

Potthast et al. (2016) introduced one of the earliest benchmark datasets for clickbait detection, demonstrating that linguistic cues such as sensational wording and exaggerated expressions could be effectively used for classification. While valuable for academic research, this dataset is limited in diversity and domain coverage.

Bhatt and Sharma (2018) explored deep learning models including LSTMs and CNNs for headline classification, achieving better accuracy than traditional approaches. However, such models demand large training datasets and significant computational resources, making them unsuitable for lightweight, real-time systems.

Community-driven platforms like Kaggle have hosted competitions on clickbait detection, where models such as logistic regression, random forest, and support vector machines were widely applied. These models are simple, interpretable, and efficient but may fall short in capturing the complex patterns of modern clickbait headlines.

Practical tools such as the "BS Detector" browser extension have also attempted to flag misleading news and clickbait content. While user-friendly, these extensions rely heavily on predefined rules, which limits adaptability to evolving patterns of clickbait.

From this survey, it is evident that existing approaches are either computationally heavy or lack deployability in real-world scenarios. To bridge this gap, our project proposes a **lightweight, modular, machine learning-based clickbait detection system**, deployed through **FastAPI** and accessible via a **web application** for real-time usage

3. PROBLEM DOMAIN

With the rapid growth of online media, users are constantly exposed to headlines that are designed more to attract clicks than to convey genuine information. Such **clickbait headlines** mislead readers, degrade the quality of journalism, and contribute to the spread of misinformation. Since clickbait is often crafted in subtle ways, manual detection is neither scalable nor reliable. Hence, there is a pressing need for an automated system that can accurately distinguish between clickbait and non-clickbait headlines.

The objective of this project is to design and implement a **machine learning-based clickbait detection system** that classifies headlines into two categories: clickbait and non-clickbait. The project aims to:

- Apply **Natural Language Processing (NLP)** techniques for effective text preprocessing.
- Use **TF-IDF** for feature extraction and **Logistic Regression** as a baseline classifier.
- Develop a **modular system** that can be extended with advanced models.
- Deploy the solution through a **FastAPI backend** integrated with a **web application** for ease of use.

This problem domain highlights the need for a practical, efficient, and user-friendly system to improve the reliability of digital content.

4. SOLUTION DOMAIN

The proposed project aims to develop a **Clickbait Detection System** that leverages Machine Learning and Natural Language Processing to classify news headlines as clickbait or non-clickbait. The solution is designed to be lightweight, modular, and practical for real-world deployment.

The workflow begins with **text preprocessing**, where headlines undergo transformations such as lowercasing, removal of stopwords, tokenization, and lemmatization. These steps ensure that the text is standardized and free from noise, thereby improving the quality of features extracted for model training. The **Natural Language Toolkit (NLTK)** library will be used to implement these operations effectively.

For feature extraction, the **Term Frequency–Inverse Document Frequency (TF-IDF)** technique will be employed. TF-IDF converts the textual data into numerical vectors by measuring the importance of words relative to the dataset. This representation helps capture discriminative patterns between clickbait and non-clickbait text.

The **Logistic Regression algorithm** will be used as the baseline classifier due to its simplicity, interpretability, and strong performance in binary classification tasks. The system will be implemented in a modular fashion, enabling the integration of more advanced models such as Support Vector Machines (SVM) or Random Forests if required in future enhancements.

To make the model accessible, it will be deployed through a **FastAPI backend**, which serves as an interface between the machine learning model and external applications. FastAPI is chosen for its speed, scalability, and ease of integration with Python-based ML workflows. On the client side, a **web application** built using HTML, CSS, and JavaScript will allow users to input headlines and instantly view predictions.

This solution ensures a seamless combination of **data preprocessing, machine learning, and web deployment**, resulting in a system that is efficient, extensible, and user-friendly.

5 SYSTEM DOMAIN

The implementation of the proposed clickbait detection system requires a combination of programming languages, libraries, and deployment tools. **Python** is chosen as the primary programming language due to its extensive support for machine learning and natural language processing. Libraries such as **NLTK** will be used for text preprocessing, while **scikit-learn** will provide efficient implementations of TF-IDF and Logistic Regression. For model persistence, **joblib** will be used to serialize and load trained models.

To deploy the solution, **FastAPI** has been selected for its lightweight, high-performance capabilities and seamless integration with Python-based ML models. The frontend will be developed using **HTML, CSS, and JavaScript** to create a simple yet interactive web interface for user interaction.

The project is platform-independent and can run on any system with at least **4 GB RAM and a dual-core processor**, making it suitable for standard personal computers. This combination of tools and technologies ensures that the system remains **efficient, scalable, and easy to extend** with future improvements.

6 APPLICATION DOMAIN

The proposed clickbait detection system has wide applications in the field of digital media and online information dissemination. It can be integrated into **news aggregators, blogging platforms, and social media sites** to automatically filter or flag misleading headlines, thereby improving the overall credibility of content. Browser extensions or plugins can also be developed using this system to help end-users avoid wasting time on deceptive links. Beyond clickbait, the framework can be extended to related domains such as **fake news detection, spam filtering, and sentiment analysis**. Thus, the project has both academic significance and practical real-world impact.

7. EXPECTED OUTCOME

The project is expected to deliver the following outcomes:

1. A **trained machine learning model** capable of classifying news headlines into clickbait and non-clickbait categories.
2. A **modular NLP pipeline** with preprocessing (tokenization, stopword removal, lemmatization) and TF-IDF feature extraction.
3. A **FastAPI-based backend** for serving model predictions in real time.
4. A **web application interface** where users can input headlines and view instant results.
5. A **scalable framework** that can be extended with advanced ML models for improved accuracy.
6. A practical system contributing to **enhanced trustworthiness of online content**.

References :

[1] M. Potthast, S. Köpsel, B. Stein, and M. Hagen, "Clickbait detection," in *Proceedings of the 38th European Conference on Information Retrieval (ECIR)*, Padua, Italy, 2016, pp. 810–817.

[2] P. Bhatt and P. Sharma, "Automatic detection of clickbait headlines using deep learning," in *Proceedings of the International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, Greater Noida, India, 2018, pp. 908–912.

[3] A. Anand, M. Chakraborty, and S. Park, "We used neural networks to detect clickbaits: You won't believe what happened next!," in *Proceedings of the European Conference on Information Retrieval (ECIR)*, 2017, pp. 541–547.

[4] A. Agrawal, "Clickbait detection using machine learning," *International Journal of Computer Applications*, vol. 176, no. 38, pp. 22–27, July 2020.

[5] Kaggle, "Clickbait title classification dataset," Available: https://www.kaggle.com/datasets/amananandrai/clickbait-dataset, Accessed: Aug. 2025.

[6] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, O'Reilly Media, 2nd ed., 2009.

[7] scikit-learn, "Machine learning in Python," Available: https://scikit-learn.org, Accessed: Aug. 2025.

[8] NLTK, "Natural Language Toolkit Documentation," Available: https://www.nltk.org, Accessed: Aug. 2025.

[9] FastAPI, "FastAPI documentation," Available: https://fastapi.tiangolo.com, Accessed: Aug. 2025.

[10] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed., Pearson, 2020.