iNeuron

LOW-LEVEL DOCUMENT

News Article Sorting

Revision number – 1.2

Last date of revision - 11/9/2023

Authored by: Krishna Chandra Yadav

1. Document Version Control

Date	Version	Description	Author
02/9/2023	1.0	Abstract	Krishna Chandra Yadav
08/9/2023	1.1	Design Flow	Krishna Chandra Yadav
11/9/2023	1.2	Performance Evaluation Conclusion	Krishna Chandra Yadav

Contents

Document Version Control 2
Abstract 4
Web Interface
1. Home Page 5
2. User Input 6
3. Output 7

Abstract:

This project focuses on improving the classification of news articles by leveraging the power of TF-IDF (Term Frequency-Inverse Document Frequency) for text vectorization and employing a range of machine learning algorithms for performance evaluation. In an era of information overload, efficient news article classification plays a pivotal role in helping user's access relevant content quickly and aiding advertisers in targeting their audience effectively.

The project employs TF-IDF, a widely used technique in natural language processing, to convert the textual content of news articles into numerical vectors. These vectors capture the importance of terms within articles, enabling the algorithms to understand the content better.

To assess the performance of the classification task, four powerful machine learning algorithms—Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Classifier, and AdaBoost are employed. Each algorithm brings its unique strengths, from SVM's ability to handle high-dimensional data to Random Forest's ensemble-based robustness.

The project evaluates these algorithms based on accuracy to determine their effectiveness in classifying news articles accurately. Through extensive experimentation and analysis, we aim to identify the algorithm that offers the best trade-off between precision and computational efficiency for news article classification.

Keywords: Natural Language Processing, machine learning, TF-IDF, content recommendation, news article classification, Support Vector Machine, Random Forest, Gradient Boosting Classifier, Ada-Boost, classification system.

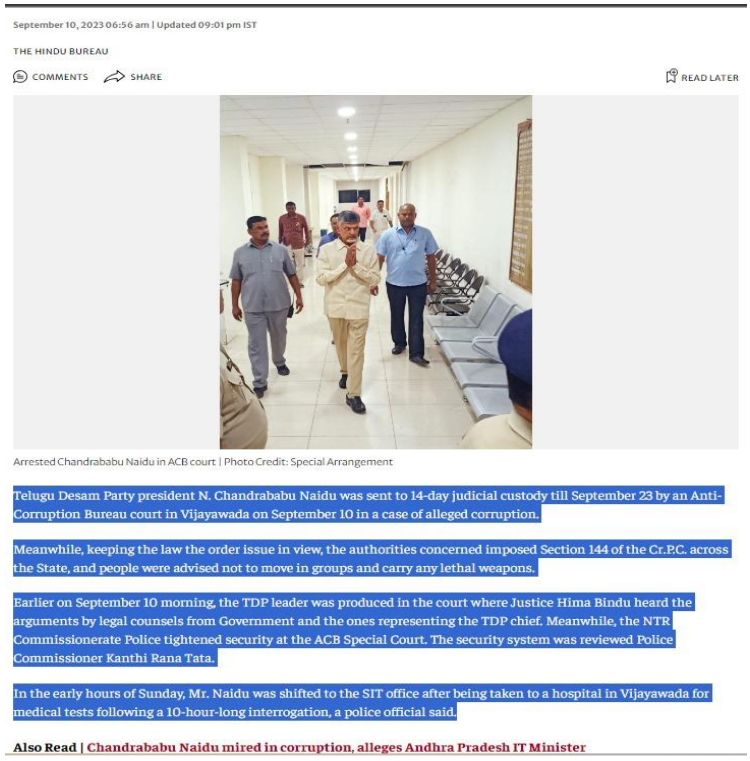
Web Interface

1. Home Page

This is the landing page of our application.

For the User to provide the Article paragraph, there are text boxe.

2. User Input



September 10, 2023 06:56 am | Updated 09:01 pm IST

THE HINDU BUREAU

COMMENTS SHARE READ LATER

Arrested Chandrababu Naidu in ACB court | Photo Credit: Special Arrangement

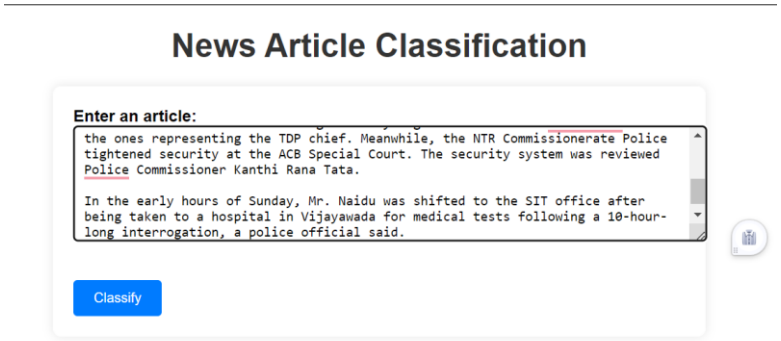
Telugu Desam Party president N. Chandrababu Naidu was sent to 14-day judicial custody till September 23 by an Anti-Corruption Bureau court in Vijayawada on September 10 in a case of alleged corruption.

Meanwhile, keeping the law the order issue in view, the authorities concerned imposed Section 144 of the Cr.P.C. across the State, and people were advised not to move in groups and carry any lethal weapons.

Earlier on September 10 morning, the TDP leader was produced in the court where Justice Hima Bindu heard the arguments by legal counsels from Government and the ones representing the TDP chief. Meanwhile, the NTR Commissionerate Police tightened security at the ACB Special Court. The security system was reviewed Police Commissioner Kanthi Rana Tata.

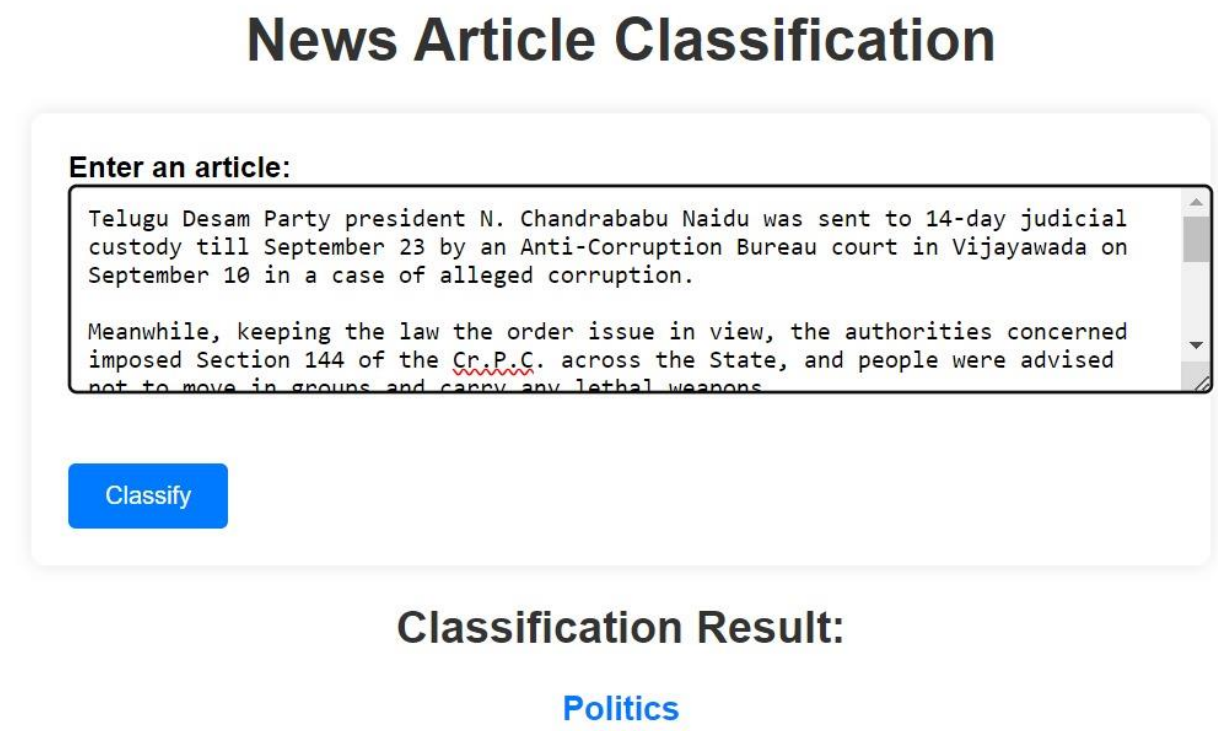
In the early hours of Sunday, Mr. Naidu was shifted to the SIT office after being taken to a hospital in Vijayawada for medical tests following a 10-hour-long interrogation, a police official said.

Also Read | Chandrababu Naidu mired in corruption, alleges Andhra Pradesh IT Minister



Here, we are taking the latest Article from political news. This input data will also follow the same preprocessing steps that we did before training the model. The model will classify the customer using the SVM algorithm.

3. Output



As we can see clearly our model classified the article correctly.