

THE
SPECIALIZED SVM FAKE NEWS
DETECTION MODEL

Cavite State University

July 1, 2025

* SPECIALIZED SVM FAKE NEWS DETECTION MODEL *

USER MANUAL

“Detecting Deception. One Category at a Time”



Authors:

Alian Gem Axel N. Crooc | John Joshua F. Garcia | Dinothelo P. Quiroga

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



BRIEF INTRODUCTION



This software application is designed to detect fake news using advanced machine learning and natural language processing techniques. Aimed at researchers, students, developers, media professionals, and anyone interested in understanding the nature of misinformation, the tool serves as a practical platform for evaluating the credibility of digital news content. Whether used in academic settings, media organizations, or personal investigations, it provides a structured and interactive way to analyze and classify news articles.

The software allows users to either manually input a headline and article content or supply a URL, from which the system automatically extracts relevant information. It then categorizes the article into a specific topic such as Politics, Health, Entertainment, or Others, and applies a trained Support Vector Machine (SVM) model enhanced with sentiment analysis and TF-IDF similarity scoring.

Users can expect clear outputs indicating the likely credibility of the article, a breakdown of the topic classification, and insights into the emotional tone and lexical similarity of the content. The interface is designed to be user-friendly, even for those with limited technical backgrounds. However, a basic understanding of how fake news operates, or familiarity with common online media practices, will enhance the user experience.

Important notes

The system is built using Python, integrates various NLP libraries, and is most effective when analyzing English-language articles. It relies on pre-trained models and a curated dataset of real and fake news, so its effectiveness depends on the quality and relevance of the input. While the tool does not guarantee a perfect classification, it provides a meaningful and data-driven starting point for identifying potentially misleading content online.

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



TRY OUR ONLINE DEMO WEB APP



You can explore or demo the app directly online via Hugging Face Spaces:

Visit:



<https://huggingface.co/spaces/DinowRwr/fake-news-detector>

This hosted version allows users to test the app on any device with an internet browser. No installation needed. However, for full functionality and speed, running the app locally is still recommended.

THE

SPECIALIZED SVM FAKE NEWS

DETECTION MODEL

Cavite State University

July 1, 2025



TABLE OF CONTENTS



1. System Requirements	1
2. Installation Guide	2
3. UI, Features, and Functionality	3
4. User Interface of the Web App	5
5. Troubleshooting	6
6. FAQs	7
7. User Feedbacks	10
8. Conclusion	11

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



SYSTEM REQUIREMENTS



1. **Operating System:** Windows 11 (64-bit)
2. **Processor:** Intel Core i5 or equivalent
3. **RAM:** 8GB minimum
4. **GPU:** 4GB VRAM
5. **Storage:** At least 1 GB

Software Requirements:

1. Python 3.13
2. VS Code

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



INSTALLATION GUIDE



Follow these steps to install and run the Fake News Detection Web App on your local machine:

Step 1: Download the ZIP file

Scan the QR code provided (or visit the link, if available) to download the ZIP file containing the source code and assets.

https://drive.google.com/file/d/1wN7B20YFENpDcBqgYod_BUH0OkP2WM7M/view?usp=sharing



Step 2: Unzip the file

Extract the ZIP file to a folder on your computer.

Step 3: Open the folder in VS Code

1. Launch Visual Studio Code.
2. Click on File > Open Folder.
3. Navigate to the folder where you unzipped the project and open it.

Step 4: Install the required Python libraries

1. Open a new terminal in VS Code.
2. Copy and paste the following command into the terminal and press Enter:

```
pip install gradio joblib nltk numpy pandas newspaper3k spacy
scikit-learn vaderSentiment scipy lxml_html_clean
```

Step 5: Run the main app

In the terminal, type the following commands one by one:

```
cd python
python main_app5.py
```

Step 6: Open the web application

Once the app is running, look for the following line in the terminal output:

Running on local URL: http://127.0.0.1:0000

Step 7: Hold Ctrl+ Left Click the URL to open the web application in your browser.

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025

* USER INTERFACES, FEATURES, AND FUNCTIONALITY *

Below is a detailed breakdown of each user interface component and its corresponding functionality:

1. URL Input Field

- A text input box where users can paste the URL of a news article.
- This is optional, the system also supports manual headline and content input.

2. Scrape Article Button

- Upon clicking, the system will scrape:
 - Headline
 - Content
- This automates data collection from online news sources.

3. Headline and Content Output

- Editable textboxes where scraped or manually entered headline and content are displayed.
- Users can freely edit these fields before analysis.

4. Scraping Status

- A live status message or alert that informs the user whether the scraping was:
 - Successful
 - Failed
- Helps users understand what's happening behind the scenes.

5. Analyze Headline and Content Button

- Executes the complete analysis pipeline:
 - Runs the news categorizer model
 - Sends the input to the appropriate SVM fake news detector
 - Displays analysis results

To be continued next page

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025

*** USER INTERFACES, FEATURES, AND FUNCTIONALITY ***

6. Results Section

Displays results of the analysis in four levels:

a. Predicted Category

- Displays the detected news topic category:
 - Politics, Health, Entertainment, or Others

b. Category Confidence Scores

- Shows the confidence for each of the four categories.

c. Predicted Credibility

- Indicates whether the news article is Credible or Not Credible

d. Credibility Confidence Scores

- Displays the model's confidence for both "Credible" and "Not Credible" predictions.

7. Warnings Section

- Alerts users to potential issues such as:
 - Headline and content are different
 - URL is not provided, sign of fabricated content
- Enhances user awareness and ensures input validity.

8. Calculation Results

- Provides insight into how the model reached its decision:

Feature	Description
Sentiment	Shows the VADER Compound Score , indicating whether the tone is generally positive or negative.
LDA Distribution	Reflects topic-based separation. Negative value implies credibility, Positive implies deception.
Word Count	Displays the number of words in the news content useful to detect very short articles.
SVM Decision Boundary Distance	Distance from the separating line: Negative = Credible, Positive = Not Credible
TF-IDF Similarity to Real News	Score representing similarity to known real news articles. A value > 0.3 implies higher credibility.

9. Clear All Button

- Resets all input fields, clears outputs, and removes previous analysis results.
- Allows users to start a new session without refreshing the page.

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



USER INTERFACE OF THE WEB APP



Fake News Detection Tool

Enter news URL
Paste a news article URL here

Scrape Article

Headline

Content

Scraping Status

Analyze Headline & Content

Predicted Category

Category Confidence Scores

Predicted Credibility

Credibility Confidence Scores

Warnings

Sentiment (VADER Compound Score)

LDA Distribution (Negative = Credible, Positive = Not Credible)

Word Count

SVM Decision Boundary Distances (Negative = Credible, Positive = Not Credible)

TF-IDF Similarity to Real News (Greater than 0.3 = More Likely Credible)

Clear All

Thesis Proposal by Alian Gem Axel N. Croce
Papers Organized by John Joshua F. Garcia
WebApp Creation by Dinothelo P. Quiroga
Use via API · Built with Gradio · Settings

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



TROUBLESHOOTING



The app doesn't start after running python main_app5.py

- Possible Causes:
 - Required Python libraries are not installed
 - You are not in the correct directory
- Solution:
 - Make sure you are inside the "python" folder
 - In the terminal, type the following commands:
 - cd python
 - python main_app5.py
 - If an error about missing libraries appears, install them using:
 - pip install gradio joblib nltk numpy pandas newspaper3k spacy scikit-learn vaderSentiment scipy lxml_html_clean

Nothing happens after I click "Scrape Article"

- Possible Causes:
 - The URL is invalid or unsupported
 - The news website blocks scraping
 - There is no internet connection
- Solution:
 - Make sure the URL starts with http:// or https://
 - Try a different news article from a different website
 - If scraping fails, manually copy and paste the headline and content

"ModuleNotFoundError" or "ImportError" appears in the terminal

- Cause:
 - A required library is not installed
- Solution:
 - Open the terminal and install the required libraries by typing:
 - pip install gradio joblib nltk numpy pandas newspaper3k spacy scikit-learn vaderSentiment scipy lxml_html_clean

The web app opens but no results are displayed after clicking analyze

- Possible Causes:
 - Headline and content fields are empty
 - Scraping failed and input was not manually filled
- Solution:
 - Check if the headline and content fields are filled
 - If not, enter the text manually before clicking analyze

"Running on local URL: http://127.0.0.1:0000" appears, but I can't click it

- Solution:
 - Press and hold Ctrl, then left-click the URL in the terminal
 - If that doesn't work, copy the URL and paste it into your browser manually

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



FREQUENTLY ASKED QUESTIONS



Do I need to enter a URL?

- Answer:
 - The URL field is optional. You can type or paste the headline and content manually if you prefer.

Can I edit the headline and content after scraping?

- Answer:
 - Yes. The fields are editable. You may correct or change the text before analyzing it.

What if the detected category is wrong?

- Answer:
 - The news categorizer may sometimes misclassify content. However, the app will still proceed with analysis using the closest match. You may retry with a clearer input.

What browsers are compatible with the web app?

- Answer:
 - The app works on most modern browsers including:
 - Google Chrome
 - Mozilla Firefox
 - Microsoft Edge
 - Opera GX

Can I use this app offline?

- Answer:
 - No. Internet is required to scrape articles, download models, and access dependencies like news APIs and sentiment models.

Why does the app say "TF-IDF similarity is low"?

- Answer:
 - A low TF-IDF similarity score means the article content is not very similar to the patterns found in real news training data. This might suggest it's less credible or highly unusual.

What does "Credibility Confidence Score" mean?

- Answer:
 - It shows how confident the model is about whether the article is credible or not. A higher score means the model is more certain.

To be continued next page

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



FREQUENTLY ASKED QUESTIONS



What does "Predicted Category" mean?

- Answer:
 - It refers to the news topic classification detected by the system. The categories include: Politics, Health, Entertainment, and Others. This helps the app route the input to the most appropriate fake news detection model.

What do the "Category Confidence Scores" represent?

- Answer:
 - They show how confident the system is in assigning the input to each news category. Higher scores mean the model is more certain that the article fits that specific topic.

What does "Predicted Credibility" mean?

- Answer:
 - This is the final judgment of the fake news detection model. The result will either be "Credible" or "Not Credible" based on multiple features such as text patterns, sentiment, similarity, and model distance.

Why does the credibility have two confidence scores?

- Answer:
 - The model outputs the probability for both classes: "Credible" and "Not Credible". These scores help you understand how confident the system is in its final decision.

What does "VADER Compound Score" mean?

- Answer:
 - It's a sentiment analysis score ranging from -1 (very negative) to +1 (very positive). It reflects the overall emotional tone of the content. Neutral or mildly negative tones are more typical of credible news.

What is "LDA Distribution" used for?

- Answer:
 - It shows the topic projection value after applying Linear Discriminant Analysis. Negative values usually indicate alignment with credible news, while positive values may indicate non-credible content.

To be continued next page

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



FREQUENTLY ASKED QUESTIONS



What does "SVM Decision Boundary Distance" indicate?

- Answer:
 - It represents how far the input lies from the model's decision boundary. A negative distance means it's likely credible; a positive distance suggests it's more likely fake or misleading.

How is "TF-IDF Similarity to Real News" calculated?

- Answer:
 - This compares the input to real news articles in the dataset. A score higher than 0.3 usually means the article shares language patterns common to verified news, suggesting credibility.

Why does the app include a word count?

- Answer:
 - Word count helps identify incomplete or shallow articles. Very short content may not provide enough information for a reliable credibility assessment and might raise a warning.

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025



USER FEEDBACK



Please rate your experience using the Fake News Detection Web App and explain any issues, suggestions, or feedback you have. Your input is highly appreciated and will help improve the system.

Scan the QR code or visit the feedback link below to submit your responses.

QR Code:



Link:

<https://forms.gle/TYiowHe9YJxNGoeY7>

Privacy Note

No personal information is required unless the user chooses to provide it. All feedback will be used solely for research and system refinement purposes.

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025

CONCLUSION

The Fake News Detection Web Application provides a streamlined and intelligent solution for identifying the credibility of online news articles. By combining natural language processing (NLP), sentiment analysis, topic categorization, and Support Vector Machine (SVM) classification, the system delivers accurate, category-specific assessments of news content.

The application supports both automatic article scraping through URLs and manual input, giving users flexibility and control over the analysis process. The integration of multiple analysis layers such as sentiment scoring, LDA projection, TF-IDF similarity, and SVM boundary distance, offers transparency in how the model arrives at its predictions.

Throughout the development process, the use of categorized datasets, sentiment tools like VADER, and dimensionality reduction through LDA proved effective in enhancing classification performance. However, the findings also highlighted that generalized models may sometimes outperform specialized ones in overall accuracy, suggesting that fake news detection requires both linguistic versatility and focused evaluation strategies.

Ultimately, this project contributes to the ongoing fight against misinformation by offering a user-friendly, explainable, and practical tool for verifying the authenticity of news. Future improvements may include expanding to more categories, incorporating multilingual datasets, and refining the news categorizer model for better routing accuracy.

THE SPECIALIZED SVM FAKE NEWS DETECTION MODEL

Cavite State University

July 1, 2025

CONCLUSION

The Fake News Detection Web Application provides a streamlined and intelligent solution for identifying the credibility of online news articles. By combining natural language processing (NLP), sentiment analysis, topic categorization, and Support Vector Machine (SVM) classification, the system delivers accurate, category-specific assessments of news content.

The application supports both automatic article scraping through URLs and manual input, giving users flexibility and control over the analysis process. The integration of multiple analysis layers such as sentiment scoring, LDA projection, TF-IDF similarity, and SVM boundary distance, offers transparency in how the model arrives at its predictions.

Throughout the development process, the use of categorized datasets, sentiment tools like VADER, and dimensionality reduction through LDA proved effective in enhancing classification performance. However, the findings also highlighted that generalized models may sometimes outperform specialized ones in overall accuracy, suggesting that fake news detection requires both linguistic versatility and focused evaluation strategies.

Ultimately, this project contributes to the ongoing fight against misinformation by offering a user-friendly, explainable, and practical tool for verifying the authenticity of news. Future improvements may include expanding to more categories, incorporating multilingual datasets, and refining the news categorizer model for better routing accuracy.

**SPECIALIZED SVM FAKE NEWS
DETECTION MODEL**
