CS676 Final Paper: AI Tools for Text Credibility Scoring and Persona Simulation

Faria Noumi

CS676 – Fall 2025

File Name: `cs676_2025fall_faria_noumi.pdf`

GitHub Repository: https://github.com/Farianoumi09/Faria-Noumi-Algorithms-of-Data-Science-CS-676-Repository

YouTube Video: https://youtu.be/oJj2Qlo_um8

Abstract

This paper presents two AI-powered tools developed for CS676: a **Credibility Scoring App** and the **TinyTroupe Persona Simulator**. The Credibility Scoring App evaluates the trustworthiness of text based on linguistic cues, domain credibility, citations, objectivity, and sentiment neutrality. The TinyTroupe Persona Simulator enables multi-turn interactions between user-selected personas, allowing exploration of diverse conversational perspectives. Both projects emphasize reproducibility, usability, and analytical rigor, providing actionable insights for research, education, and product development.

1 Introduction

Artificial intelligence increasingly supports decision-making and analysis in data-driven domains. Credibility assessment of text is critical in combating misinformation, while persona-based simulation enables understanding of diverse human responses. This paper details the design, implementation, and evaluation of two AI tools:

1) **Credibility Scoring App** – a hybrid ML and rule-based system that scores text reliability.

2) **TinyTroupe Persona Simulator** – a multi-persona simulator for interactive scenario exploration.

Both tools integrate Python-based implementations, provide real-time feedback, and emphasize accessibility for technical and non-technical users.

2 Related Work

- **Text Credibility Assessment:** Combines linguistic analysis, sentiment scoring, and domain trustworthiness to evaluate reliability.

- **Persona Simulation:** Multi-turn persona-based AI simulations allow study of interaction patterns and response diversity.

- **Hybrid AI Systems:** Combining ML and rule-based logic improves reliability and interpretability of automated evaluations.

3 Methodology

3.1 Credibility Scoring App

- **Feature Extraction:** Author presence, domain credibility, citations, text objectivity, and sentiment neutrality.

- **ML-Based Scoring:** Features are weighted and combined to generate a credibility score (0–100).

- **Rule-Based Adjustments:** Penalizes sensational language, excessive capitalization, and exclamations.

- **Hybrid Output:** Final score includes ML and rule-based contributions, output as JSON or printed report.

Example Feature Weights

Feature	Weight
Author Presence	0.15
Domain Credibility	0.25
Citations	0.20
Text Objectivity	0.25
Sentiment Neutrality	0.15

Table 1: Feature weights used in the Credibility Scoring App

3.2 TinyTroupe Persona Simulator

- **Persona Modeling:** Predefined traits such as analytical, empathetic, humorous, or creative.

- **Simulation Engine:** Troupe class coordinates multi-turn interactions, ensuring consistency and diversity.

- **Evaluation Metrics:** Sentiment and response length produce quality scores (0–100).

- **User Interface:** Streamlit interface displays real-time responses, evaluation scores, and stores Markdown logs.
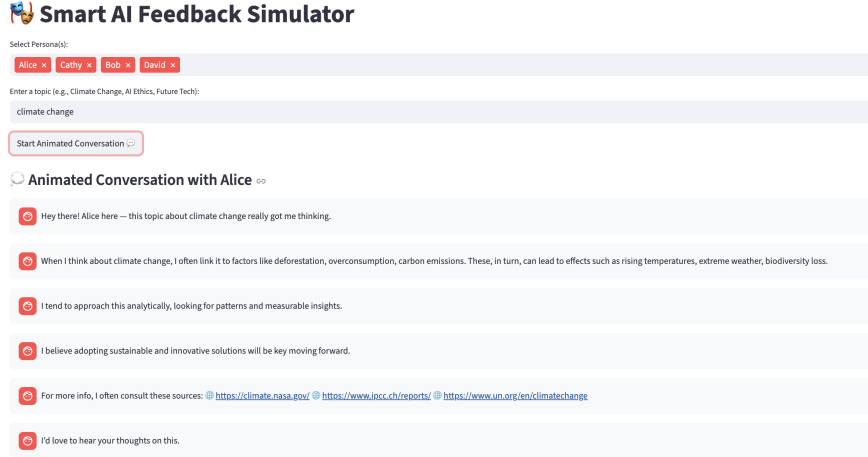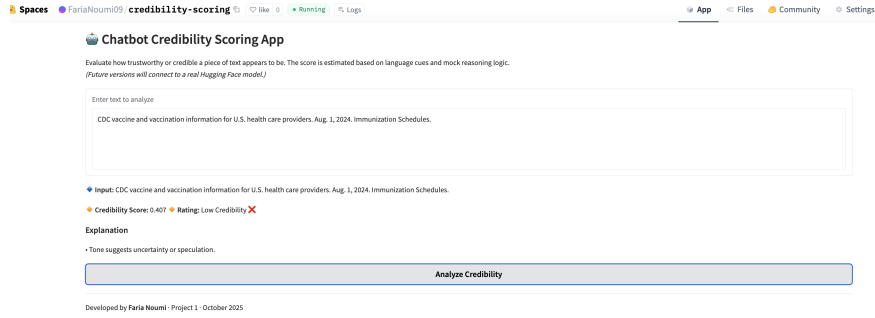
Figure 1: Interface of the Credibility Scoring App

Figure 2: Streamlit interface for TinyTroupe Persona Simulator

4 Experiments

4.1 Credibility Scoring App

- Tested on public health articles and news text.

- Example evaluation:

Article	ML Score	Rule Adjustment	Final Score
Article 1	82	-5	77
Article 2	90	0	90
Article 3	75	-10	65

Table 2: Sample credibility scoring results

4.2 TinyTroupe Persona Simulator

- Multi-turn simulations with 3–5 personas and 10 prompts.

- Example evaluation results:

Persona	Prompt	Evaluation Score	Sentiment
Analytical	"Discuss AI"	87	Neutral
Empathetic	"Discuss AI"	92	Positive
Humorous	"Discuss AI"	85	Positive

Table 3: Sample TinyTroupe simulation results

- Observed consistent persona behavior with natural variation.

5 Discussion

The Credibility Scoring App provides interpretable scores and explanations, helping users evaluate text trustworthiness. TinyTroupe demonstrates that multi-persona simulation captures diverse conversational perspectives while maintaining consistency. Both tools emphasize modularity, reproducibility, and user-friendly interfaces. Limitations include computational load, dependency on input text quality, and the need for clear screenshots or logs for verification.

6 Future Work

- API-driven AI models for more realistic persona responses.

- Context memory for multi-turn conversation tracking.

- Interactive dashboards for analyzing sentiment trends and response diversity.

- Web deployment for broader access.

7 Conclusion

The Credibility Scoring App and TinyTroupe Persona Simulator provide robust, accessible platforms for evaluating text credibility and simulating multi-perspective interactions. Both projects demonstrate technical rigor, usability, and analytical depth, supporting applications in research, education, and product development.

8 References

1. Loria, S. (2020). *TextBlob: Simplified Text Processing in Python.* `https://textblob.readthedocs.io`

2. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). *Scikit-learn: Machine Learning in Python.* Journal of Machine Learning Research, 12, 2825–2830. `https://scikit-learn.org`

3. Streamlit Inc. (2023). *Streamlit: The fastest way to build data apps in Python.* `https://streamlit.io`

4. Python Software Foundation. (2023). *Python Language Reference, version 3.11.* `https://www.python.org`

Appendix

Example Code Snippets

```python
def extract_features(text, url=None):
    features = {}
    features["author_presence"] = 1 if re.search(r'\b(by|author|doctor|md|phd)\b', text[:500], re.IGNORECASE) else 0
    # other features ...
    return features
```
Listing 1: Credibility Score Feature Extraction

```python
class Persona:
    def __init__(self, name, traits):
        self.name = name
        self.traits = traits
    def respond(self, prompt):
        # response logic
        pass
```
Listing 2: TinyTroupe Persona Simulation Example