AI-Facts-Verification

📌 Project Description

An open-source project supporting fact verification using AI.
Goal: To create a system for labeling content credibility in a **neutral and educational way**, helping people analyze information instead of imposing a single "truth."

🛠️ How It Works?

✅ AI analyzes content and provides context rather than making absolute judgments.
✅ The system assigns **credibility scores** on a scale of 1-100.
✅ **Does not block content**, but highlights potential inconsistencies and manipulations.
✅ Can be integrated as a **browser extension**, **API**, or **standalone platform**.

📖 Documentation

1. Overview of AI Content Verification

AI-Facts-Verification utilizes a multi-layered approach to assess content credibility, analyzing factors such as:

- **Source reliability** – cross-checking information against trusted databases.
- **Logical consistency** – evaluating if claims are internally coherent.
- **Manipulation detection** – identifying emotionally charged language, logical fallacies, and selective omissions.
- **Cross-referencing facts** – checking statements against known factual records.
- **Historical analysis** – determining whether similar claims have been previously debunked.

2. Credibility Scoring System

The system assigns a **credibility score from 1 to 100** based on:

- **High credibility (80-100):** Well-researched, multiple verified sources.
- **Moderate credibility (50-79):** Some supporting sources, minor inconsistencies.
- **Low credibility (1-49):** Few or no sources, high manipulation risk.
- **Unknown (0):** Not enough data to determine credibility.

3. AI Processing Pipeline

1️⃣ **Text Ingestion** – Content is extracted from the webpage or uploaded document.
2️⃣ **Metadata Analysis** – AI checks sources, timestamps, and author credibility.
3️⃣ **Semantic & Sentiment Analysis** – Detects emotional bias and manipulative phrasing.
4️⃣ **Fact-Checking & Cross-Validation** – Compares claims against established factual databases.
5️⃣ **Pattern Recognition** – Identifies propaganda techniques, deepfake content, and misinformation tactics.
6️⃣ **Final Credibility Score Calculation** – Assigns a rating and provides a summary.

4. AI Algorithms & Techniques

The system is powered by:

- **Natural Language Processing (NLP)** – Enables AI to understand and interpret textual content.
- **Machine Learning (ML) Classifiers** – Detect misinformation patterns using supervised learning.
- **Neural Networks for Context Understanding** – Identifies relationships between claims and sources.
- **Sentiment Analysis** – Determines emotional bias in writing.
- **Knowledge Graphs** – Maps connections between factual data sources.

5. Implementation Options

- 🔹 **Browser Extension** – Real-time evaluation of online content.
- 🔹 **API Integration** – Websites and platforms can query AI for credibility assessments.
- 🔹 **Standalone Web Platform** – Users can paste text to receive AI analysis.
- 🔹 **Mobile Application** – AI-powered verification tool for users on the go.

6. Data Sources & Verification Methods

✅ **Fact-checking databases** – AI pulls data from verified sources like PolitiFact, Snopes, and academic research. ✅ **Blockchain for Data Integrity** – Prevents tampering with verification results. ✅ **Community Reporting System** – Users can submit questionable claims for review. ✅ **Multi-language Support** – AI adapts credibility analysis for various languages and cultural contexts.

7. Future Roadmap

✅ **Phase 1:** Develop prototype for browser extensions.
✅ **Phase 2:** Expand AI dataset and improve accuracy of credibility assessments.
✅ **Phase 3:** API deployment for large-scale integration.
✅ **Phase 4:** Community-driven expansion and open-source collaboration.
✅ **Phase 5:** Advanced AI for deepfake detection and real-time misinformation tracking.

📌 License

🚀 **Join the project development!** If you have ideas, submit an "Issue" or contact us.

Contributors:

- **CoolTomGPT** (GitHub: @cooltomgpt)
- **ChatGPT AI System**