

Problem Statement:

The main goal of this research is to explore how AI can be used to detect and handle fake news in practical, real-world environments, while focusing on challenges like language diversity, explainability, and user trust.

In today's digital age, people rely more and more on online platforms and social media to get their news. While this makes information more accessible, it also opens the door for fake news to spread quickly sometimes faster than the truth. This creates confusion, affects how people think, and can even lead to serious real-world consequences, especially in politics, health, or public safety.

To address this, researchers have turned to artificial intelligence (AI) as a way to automatically detect and stop the spread of fake news. Many recent models use transformer-based architectures like BERT, RoBERTa, and MARBERTv2, which represent major progress in the field. These models are designed to better understand language context, and they usually perform well in controlled settings. However, they still face clear limitations when applied to real-world scenarios especially in multilingual environment.

One major limitation is that they are often trained on specific datasets, so when they are applied to different platforms, languages, or types of content, their performance usually drops. For example, one study focused on Chinese-language fake news and others on Arabic content. While these models performed well in their specific contexts, they had difficulty generalizing across languages, platforms, or dialects. Also, these models often act like "black boxes" they give an answer, but don't explain why.

Another issue is that most models focus only on the text itself and ignore how users interact with the content. One of the studies I reviewed (Park & Chai, 2023) showed that tweets with emotional language, high user activity, or repetitive patterns were more likely to be fake. These behavioral features are very relevant, but they're rarely included in modern models.

Also, the nature of fake news is always changing. Some articles are now being generated by AI tools themselves, making them harder to detect. Models that were accurate a year ago may not be effective anymore. So, we need systems that can adapt and keep up with these changes.

The main goal of this research is to explore how AI can be used to detect and handle fake news in practical, real-world environments, while focusing on challenges like language diversity, explainability, and user trust.

Research Objectives:

1. To explore the main challenges AI models, face when detecting fake news in real-life environments.
 - Look at how data quality, model design, and platform differences affect performance.
2. To evaluate how transformer models like BERT or MARBERTv2 perform outside the lab.
 - Compare their behavior using different types of content and across multiple platforms.
3. To understand how using different languages, like Chinese and Arabic, affects model accuracy.
 - See how models trained on one language or dialect might struggle with others.
4. To find out how explainable models are and why that matters when it comes to trust.
 - Explore the need for transparency, especially when AI is used in sensitive places.
5. To check if user behavior can help improve detection.
 - Look at things like emotional tone or how content spreads, and if these signals can be added to AI systems.
6. To suggest ways, we can design fake news detection systems that actually work in the real world.
 - Give realistic suggestions based on what I find in the research.