



Plagiarism Detection Using Transformer Models

An Advanced Approach to Detecting Plagiarized and Paraphrased Text



Team Members

- VAISHNAVI KUKKALA
- HARIKRISHNA PARA
- SAI CHARAN CHANDU PATLA



CONTENT



01

Introduction

02

Statement of Value

03

Review of the State of Art

04

Approach

05

Evaluation & Results

06

Conclusion

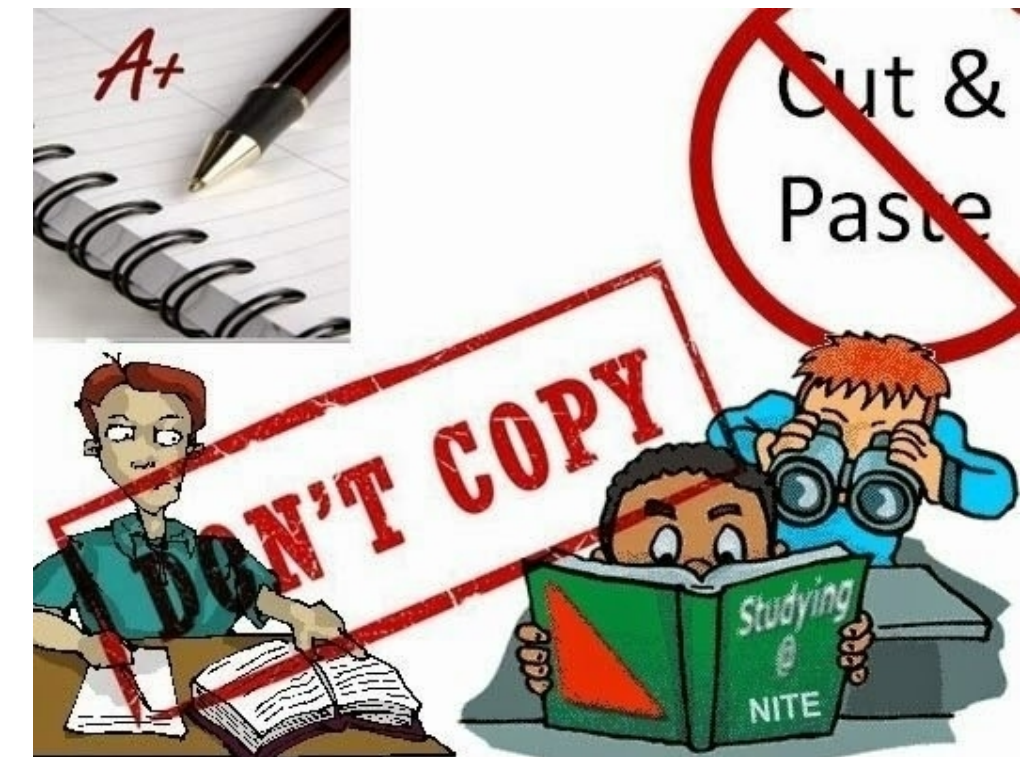
07

Future Work & Impact

Introduction

Our project focuses on developing a plagiarism detection system using natural language processing (NLP) techniques and transformer-based models to accurately find the text similarities and potential plagiarism, promoting academic integrity and protecting intellectual property. Our key objectives include:

- Build an advanced plagiarism detection system using **state-of-the-art NLP models** like BERT+LSTM, BERT, RoBERTa, and T5.
- Identify **plagiarism** and **paraphrased content** in text by leveraging deep learning techniques.
- **Importance:**
- Maintains **academic integrity**, protects **intellectual property**, and ensures **content originality**.



Statement Of Value



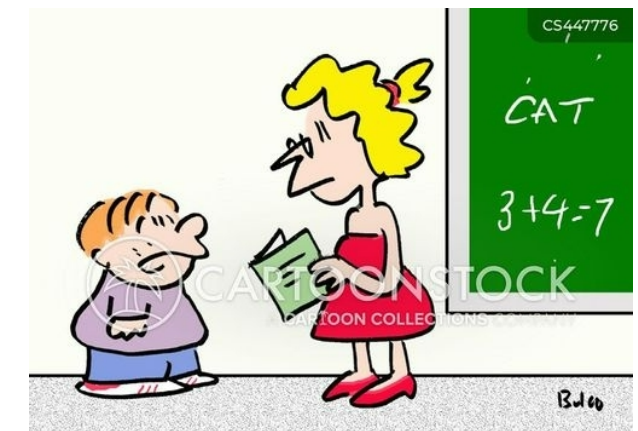
Promoting Academic Integrity: Our plagiarism detection system is designed to maintain the integrity of academic institutions by verifying the originality of submitted work and ensuring proper citation. This system fosters a culture of honesty and accountability among students, educators, and researchers.



Protecting Intellectual Property: By identifying instances of unauthorized content reuse, our project safeguards the intellectual property of creators, researchers, and organizations. This proactive detection helps prevent disputes and legal challenges related to content duplication.



Ensuring Quality and Credibility: A robust plagiarism detection system enhances the quality of academic and professional work by promoting originality and adherence to citation standards. This contributes to maintaining high standards in research, writing, and publishing, ultimately bolstering the credibility and reputation of educational institutions and organizations.



"A book report? — Wouldn't that be a copyright violation?"

State of Art



Plagiarism detection is essential for maintaining academic integrity, originality, and ethical standards in writing and research.



Existing plagiarism detection tools struggle to identify paraphrased or semantically similar content, limiting their effectiveness.



This project leverages state-of-the-art transformer models (e.g., BERT, RoBERTa, T5) to understand the context and meaning of text, improving accuracy in detecting reworded or subtle plagiarism.



It provides a scalable and modular framework that can be adapted to other NLP tasks, such as text similarity scoring, paraphrase detection, and fake news identification.



Automated plagiarism detection reduces manual effort for educators, publishers, and reviewers, saving time and improving efficiency.



Transformer-based models offer a significant improvement over traditional similarity metrics, addressing gaps in current plagiarism detection systems.



Approach

Algorithms

- Transformer-based Models:** These models utilize the transformer architecture, which is the basis for advanced NLP algorithms. They are designed to process sequential data efficiently by using self-attention mechanisms.
- Fine-Tuning:** Fine-tuning pre-trained models on a specific task (plagiarism detection, paraphrase detection) through supervised learning.
- Classification Algorithms:** These algorithms are used to classify text pairs into categories (e.g., plagiarized or not plagiarized) after feature extraction from transformer models.

Datasets

- SNLI (Stanford Natural Language Inference):** total of 570,152 pairs of sentences
- MRPC (Microsoft Research Paraphrase Corpus):** dataset with 5801 pairs of labelled sentence

Text	Judgments	Hypothesis
A man inspects the uniform of a figure in some East Asian country.	contradiction C C C C C	The man is sleeping
An older and younger man smiling.	neutral N N E N N	Two men are smiling and laughing at the cats playing on the floor.
A black race car starts up in front of a crowd of people.	contradiction C C C C C	A man is driving down a lonely road.
A soccer game with multiple males playing.	entailment E E E E E	Some men are playing a sport.
A smiling costumed woman is holding an umbrella.	neutral N N E C N	A happy woman in a fairy costume holds an umbrella.

S. No.	Sentence 1	Sentence 2	Gold annotation	Prediction	Remark
1	Ricky Clemons’ brief, troubled Missouri basketball career is over.	Missouri kicked Ricky Clemons off its team, ending his troubled career there.	paraphrase	paraphrase	correct
2	But 13 people have been killed since 1900 and hundreds injured.	Runners are often injured by bulls and 13 have been killed since 1900.	non-paraphrase	non-paraphrase	correct
3	I would rather be talking about positive numbers than negative.	But I would rather be talking about high standards rather than low standards.	paraphrase	paraphrase	correct
4	The tech-heavy Nasdaq composite index shot up 5.7 percent for the week.	The Nasdaq composite index advanced 20.59, or 1.3 percent, to 1,616.50, after gaining 5.7 percent last week.	non-paraphrase	paraphrase	incorrect
5	The respected medical journal Lancet has called for a complete ban on tobacco in the United Kingdom.	A leading U.K. medical journal called Friday for a complete ban on tobacco prompting outrage from smokers groups.	non-paraphrase	paraphrase	incorrect
6	Mrs. Clinton said she was incredulous that he would endanger their marriage and family.	She hadn’t believed he would jeopardize their marriage and family.	paraphrase	non-paraphrase	incorrect

Approach

- **BERT (Bidirectional Encoder Representations from Transformers):**

- A pre-trained model that uses bidirectional transformers for generating contextual embeddings of text. BERT is fine-tuned for text classification tasks, making it useful for identifying relationships between sentence pairs, such as plagiarism or paraphrasing.

- **BERT+BiLSTM:**

- Combines BERT's token embeddings with a **Bidirectional Long Short-Term Memory (BiLSTM)** network. BiLSTM helps capture sequential dependencies in text, improving performance in understanding complex paraphrasing.

- **RoBERTa (Robustly Optimized BERT Pretraining Approach):**

- An improved version of BERT that removes the Next Sentence Prediction (NSP) task and trains on more data. RoBERTa is typically more robust and performs better on sentence-pair tasks, like paraphrase detection.

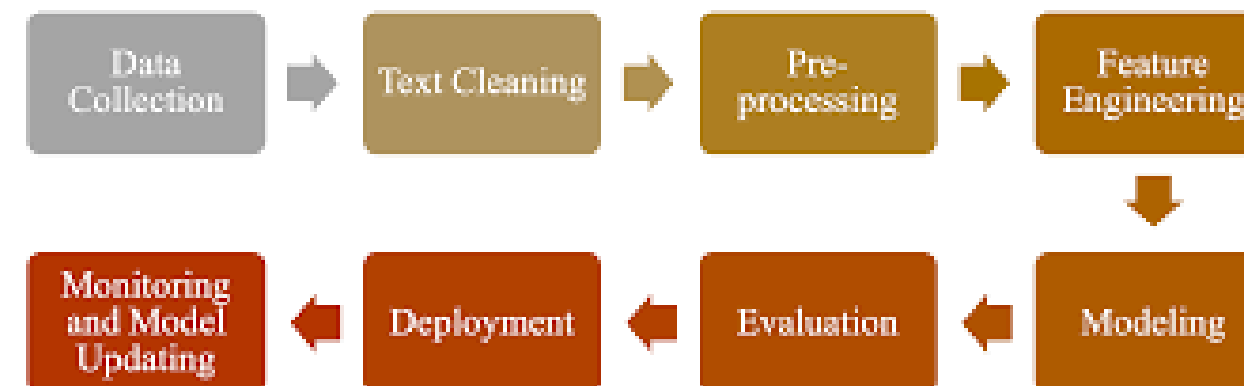
- **T5 (Text-to-Text Transfer Transformer):**

- A versatile model that reframes all NLP tasks as text-to-text problems. T5 can generate outputs like "Plagiarized" or "Not Plagiarized" from input sentence pairs, making it suitable for paraphrase and plagiarism detection.

Tools: Hugging Face Transformers, PyTorch, Tokenizers, Scikit-learn.

Techniques: Text tokenization and preprocessing. Fine-tuning transformer models for classification. Evaluation using metrics like accuracy, precision, recall, and F1 score.

NLP Pipeline



Evaluation & Results

Evaluation Metrics

- Accuracy:** Measures the overall correctness of the model.
- Precision:** The proportion of true positive results among the predicted positives.
- Recall:** The proportion of true positive results among all actual positives.
- F1 Score:** The harmonic mean of precision and recall, balancing both metrics.

✓ Comparison of model performances

1. D1 → SNLI Dataset

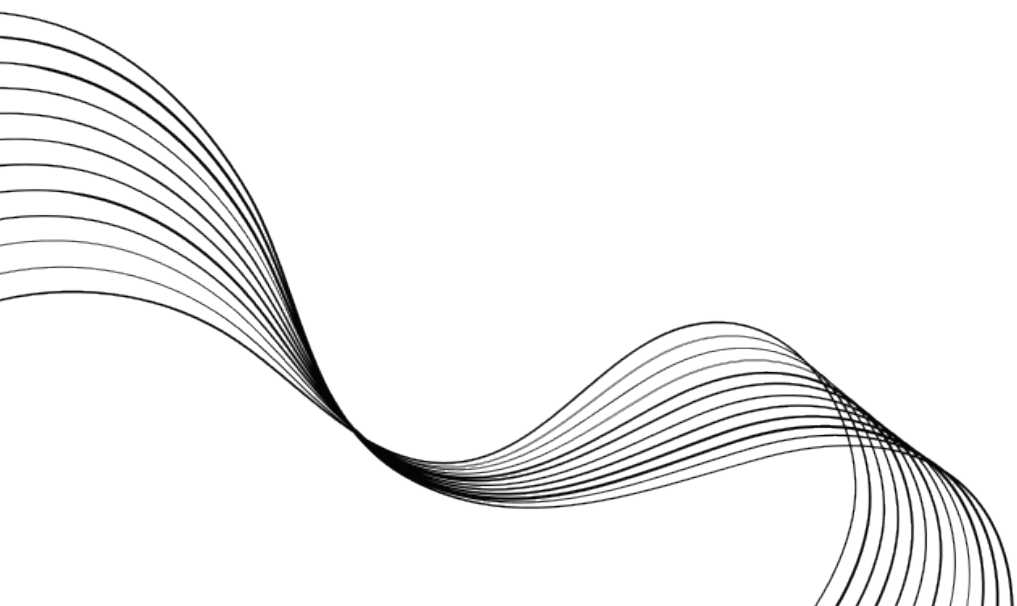
2. D2 → MSRP Dataset

Model	Accuracy (D1)	Recall (D1)	Precision (D1)	F1 Score (D1)	Accuracy (D2)	Recall (D2)	Precision (D2)	F1 Score (D2)
T5	0.7784	0.7779	0.7800	0.7776	0.6649	0.5000	0.3325	0.3994
RoBERTa	0.8370	0.8415	0.8400	0.8410	0.8203	0.8203	0.8170	0.8167
BERT	0.8587	0.8587	0.8594	0.8590	0.8371	0.7951	0.8295	0.8080
BERT + LSTM	0.8520	0.8520	0.8520	0.8284	0.90	0.7869	0.8174	0.7985

Conclusion

•Summary:

- Plagiarism detection using transformers provides an advanced, accurate solution.
- The project demonstrates the effectiveness of models like BERT, RoBERTa, BERT+LSTM and T5 for identifying both direct plagiarism and paraphrased content.
- **BERT** emerges as the best-performing model for plagiarism detection, demonstrating high accuracy and reliability on both SNLI and MSRP datasets, followed closely by **RoBERTa**.
- The **BERT + LSTM** hybrid model shows strong performance, particularly on the MSRP dataset, highlighting the potential of combining transformer embeddings with recurrent layers for specific tasks.
- **T5** underperforms in comparison, indicating that not all transformer architectures are equally suited for plagiarism detection.
- The project successfully validates transformer models as effective and scalable solutions for identifying both direct and paraphrased plagiarism, paving the way for future enhancements like multilingual and real-time detection capabilities.



Future Work & Impact

Future Work

- **Model Optimization:**

- Further optimize transformer models for faster inference.

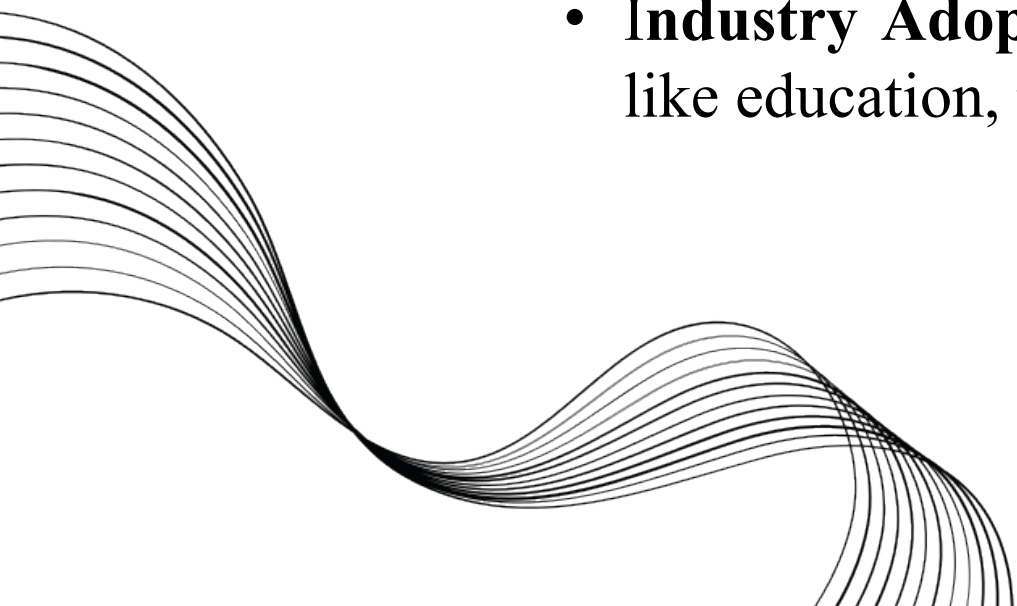
- **Scalability:**

- Scale the system to handle larger datasets and real-world use cases.

- **Deployment:**

- Deploy the solution as an API or integrate it into educational platforms and content platforms.

Impact

- Enhances **academic integrity**, ensures **content originality**.
 - **Legal Applications:** Detects content duplication in legal documents and contracts to protect intellectual property.
 - **Improved Ethics:** Promotes ethical writing practices and protects original content creators.
 - **Industry Adoption:** Provides a scalable, robust plagiarism detection tool for integration into various industries like education, media, and law.
- 
- A decorative graphic consisting of multiple thin, black, wavy lines that originate from the left side of the slide and curve downwards and to the right, creating a sense of motion and flow.

Citations

R. Patil, V. Kadam, R. Nakate, S. Kadam, S. Pattade and M. Mitkari, "A Novel Natural Language Processing Based Model for Plagiarism Detection," *2024 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2024, pp. 1-5, doi: 10.1109/ESCI59607.2024.10497386.

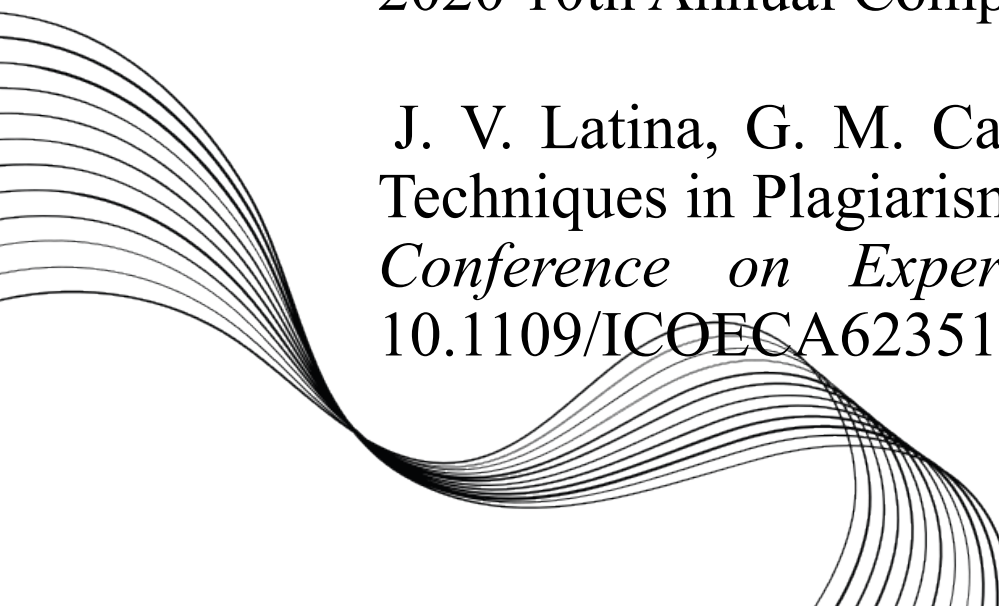
S. K. Pal, O. J. Raffik, R. Roy, V. B. Lalman, S. Srivastava and B. Sharma, "Automatic Plagiarism Detection Using Natural Language Processing," *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2023, pp. 218-222.

An, R., Yang, Y., Yang, F., & Wang, S. (2023). Use prompt to differentiate text generated by ChatGPT and humans. *Machine Learning with Applications*, 14 , 100497.

B. Ko and H. -J. Choi, "Paraphrase Bidirectional Transformer with Multi-task Learning," *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Busan, Korea (South), 2020, pp. 217-220, doi: 10.1109/BigComp48618.2020.00-72.

Barman, U., Das, S., & Das, A. (2020). An empirical analysis of plagiarism detection across programming languages. In *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0670-0675). IEEE.

J. V. Latina, G. M. Cabalsi, J. R. Sanchez, E. D. M. Vallejo, C. J. Centeno and E. A. Garcia, "Utilization of NLP Techniques in Plagiarism Detection System through Semantic Analysis using Word2Vec and BERT," *2024 International Conference on Expert Clouds and Applications (ICOECA)*, Bengaluru, India, 2024, pp. 347-352, doi: 10.1109/ICOECA62351.2024.00068.



**THANK
YOU!**

