

TinyReview: An Abstract Prediction and MetaReview Generation

Dr. Anithadevi N

Department of Information Technology
Coimbatore Institute of Technology,
Coimbatore, Tamilnadu

Megha S

Department of Information Technology
Coimbatore Institute of Technology
Coimbatore, Tamilnadu

Nandhini B

Department of Information Technology
Coimbatore Institute of Technology
Coimbatore, Tamilnadu

Ridhumiga CB

Department of Information Technology
Coimbatore Institute of Technology
Coimbatore, Tamilnadu

Shanmika PS

Department of Information Technology
Coimbatore Institute of Technology
Coimbatore, Tamilnadu

Abstract: "smart abstract evaluator" is an intelligent ai-powered framework designed to automatically assess and justify the acceptance potential of research abstracts, addressing the challenges of subjective peer evaluation and inconsistent review quality in academic submissions. Unlike existing systems that merely classify papers based on predefined keywords or citation metrics, the proposed model integrates both interpretability and analytical depth within a unified architecture. The framework combines a tinybert-based classifier for acceptance prediction with an mlp layer to assess the technical depth and coherence of the abstract, ensuring a holistic evaluation of both content quality and scientific rigor. Additionally, a novelty-checking module powered by the arxiv api quantifies the uniqueness of the submission by comparing semantic similarity with existing research works. To enhance transparency, a t5-based justification generator produces concise, human-readable explanations for each decision, bridging the gap between prediction and interpretability. The system achieves high accuracy while maintaining explainability, outperforming traditional models such as distilbert in both precision and recall. By integrating natural language understanding, novelty detection, and justification generation, smart abstract evaluator contributes to the automation of academic paper screening, providing a scalable, unbiased, and interpretable alternative to the manual peer-review process.

Keywords: research abstract evaluation, tinybert, t5 model, novelty detection, Mlp classifier, arxiv api, explainable ai, academic paper screening

I INTRODUCTION

A The Need for Automated Abstract Evaluation

With the exponential rise in research publications, academic conferences and journals face overwhelming volumes of submissions, making manual review processes slow, inconsistent, and prone to bias. Reviewers often struggle to assess thousands of abstracts within limited timeframes, leading to subjective judgments and potential oversight. Hence, an automated system that evaluates abstracts

efficiently and objectively is essential to ensure fairness, speed, and standardization in academic review workflows.

B Role of NLP in Enhancing Evaluation Accuracy

Natural Language Processing (NLP) enables machines to analyze the structure, clarity, and semantics of academic writing. Modern transformer-based models such as BERT and TinyBERT understand contextual meaning beyond keywords, allowing accurate assessment of research quality, coherence, and novelty. By leveraging NLP, abstract evaluation systems can mimic expert-level understanding and serve as valuable tools to support reviewers.

C Limitations of Traditional Approaches

Earlier evaluation systems relied on manual review or basic machine learning models using handcrafted features, which lacked contextual understanding and scalability. Large transformer models like BERT, though accurate, are computationally heavy and unsuitable for real-time deployment. These challenges highlight the need for lightweight yet powerful alternatives capable of maintaining accuracy while ensuring faster, cost-effective processing.

D The TinyBERT-T5 Based Evaluation Framework

To overcome these challenges, the proposed system integrates TinyBERT for efficient classification and for generating human-like explanations. TinyBERT predicts acceptance or rejection based on linguistic and semantic features, while T5 provides text-based justifications, improving transparency and trust. Together, they form a scalable, explainable, and AI-driven framework for fair and efficient abstract evaluation.

II LITERATURE SURVEY

A BERT: Pre-trained Language Model for NLP Tasks

The original BERT (Bidirectional Encoder Representations from Transformers) model, developed by Google AI, marked a significant breakthrough in NLP by introducing deep bidirectional context understanding. Pre-trained on large text corpora using Masked Language Modeling (MLM) and Next Sentence Prediction (NSP), BERT demonstrated superior performance in sentiment classification, question answering, and text entailment. However, due to its large parameter size (110M–340M), BERT is computationally expensive, requiring high memory and long inference times. Although highly accurate, its heavy architecture poses challenges for deployment in lightweight and real-time applications such as automated abstract evaluation. This limitation motivated subsequent research into model compression and knowledge distillation. (Paper by Devlin et al., NAACL 2019)

B DistilBERT: A Smaller, Faster, Cheaper BERT

DistilBERT, introduced by Hugging Face, is a distilled version of BERT that reduces the number of parameters by 40% while retaining 97% of its performance. The model uses knowledge distillation, where a smaller “student” model learns from the predictions of a larger “teacher” model. DistilBERT retains most of BERT’s contextual understanding but is faster and more memory efficient. It has been applied successfully in sentiment classification and question answering tasks. However, DistilBERT still lacks full optimization for edge devices and may struggle with complex reasoning or domain-specific fine-tuning in limited datasets. (Paper by Sanh et al., NeurIPS 2019 Workshop)

C TinyBERT: Task-specific Knowledge Distillation

TinyBERT builds upon the concept of knowledge distillation by introducing two-stage learning: general distillation on large corpora and task-specific distillation on fine-tuned datasets. The model achieves up to 96% of BERT’s performance while being 7.5x smaller and 9.4x faster. Its design makes it ideal for real-time text analysis and deployment in lightweight applications such as academic abstract evaluation. TinyBERT captures key linguistic features and semantic dependencies necessary for assessing clarity, novelty, and technical content. Its compact architecture ensures scalability for institutional platforms and educational technology systems. (Paper by Jiao et al., EMNLP 2020)

D T5: Text-to-Text Transfer Transformer for Explainability

The T5 (Text-to-Text Transfer Transformer) model reformulates all NLP tasks as text-to-text problems, enabling a unified framework for classification, summarization, and explanation generation. T5 excels in generating human-like textual outputs, making it suitable for producing interpretability justifications in automated abstract evaluation systems. When integrated with TinyBERT’s classification, T5 generates natural language explanations that clarify the basis for acceptance or rejection, improving transparency and user trust. However, the full-scale T5 model is computationally

heavy, necessitating smaller variants like T5-small for lightweight deployment. (Paper by Raffel et al., JMLR 2020)

E Automated Academic Paper Evaluation Using Machine Learning

Several prior works have explored automated systems for research paper evaluation using classical ML models such as SVM, Random Forest, and Naïve Bayes. These models rely heavily on handcrafted features such as word frequency, TF-IDF scores, and keyword matching. While they provide baseline accuracy, they fail to capture deeper semantics, context, and coherence within abstracts. Moreover, such systems often lack explainability and generalization across domains, limiting their adoption in conference management systems. (Paper by Kang et al., Springer 2020)

III SYSTEM ARCHITECTURE

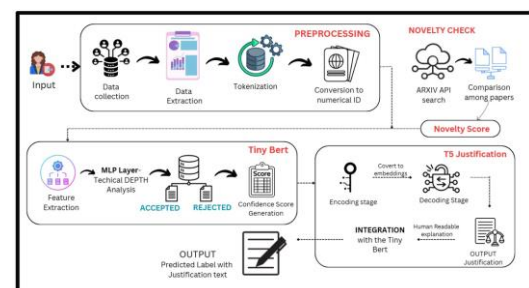


Figure 1. System Architecture

The architecture begins with the submission of an abstract as text input. This input undergoes several processing stages: first, it is cleaned and standardized through the Preprocessing Module; next, it is passed into the TinyBERT Classification Module, which predicts whether the abstract is “Accepted” or “Rejected” based on learned linguistic and contextual features. The prediction is then forwarded to the T5 Justification Generator, which produces a natural language explanation describing why the abstract was accepted or rejected. To enhance the interpretability and trust of the system, a Novelty Scoring Mechanism is also integrated, which measures the originality of the abstract by comparing it with a database of existing research works. Finally, the Adaptive Feedback Component collects user corrections and responses for continuous model improvement. All components are deployed in a lightweight Flask web framework for real-time interaction.

IV IMPLEMENTATION

The proposed system integrates transformer-based models — TinyBERT for classification and T5 for justification generation — to predict the acceptance or rejection of research abstracts and provide interpretable feedback. The implementation consists of multiple stages including preprocessing, model loading, classification, justification generation, novelty scoring, and user feedback integration. The framework is deployed using Python, Hugging Face Transformers, and Flask for web-based interaction.

A TinyBERT Model: Overview

TinyBERT is a lightweight transformer-based language model distilled from the original BERT architecture. It maintains comparable performance to BERT while being approximately 40% smaller and 60% faster, making it highly suitable for real-time text classification applications. The model was fine-tuned on a custom dataset of 32,000 research abstracts collected from public repositories and conference proceedings. The fine-tuning process adapts the model to recognize features strongly correlated with acceptance, such as clarity, novelty, methodological strength, and technical contribution. By encoding contextual relationships within the text, TinyBERT learns to classify abstracts into binary categories — *Accepted* or *Rejected*. Its compact size and efficiency enable deployment on standard computing resources without compromising prediction accuracy, making it a scalable solution for academic conference organizers.

B Model Architecture and Workflow

The overall architecture of the system follows a modular pipeline. The Preprocessing Module cleans and tokenizes the input abstract, converting it into a sequence of tokens compatible with the TinyBERT tokenizer. The tokens are then converted into embeddings and passed through multiple transformer encoder layers in TinyBERT. Each layer applies multi-head self-attention to capture dependencies between words, ensuring the model understands both local and global context. The encoded representation from the final layer is pooled into a fixed-length vector, which is then passed through a fully connected classification head. This layer computes the logits corresponding to the two target classes (*Accepted* and *Rejected*), and a softmax function converts these logits into class probabilities. The label with the highest probability is selected as the model's prediction. In addition to the class label, the model also outputs a confidence score, representing the probability assigned to the predicted label, which helps in interpreting the reliability of the decision.

C Preprocessing and Data Preparation

Accurate predictions rely on high-quality preprocessing. The input abstract undergoes several preprocessing steps:

Text Normalization: Converting all characters to lowercase and removing special symbols, numbers, and punctuation.

Stopword Removal: Eliminating non-informative words (e.g., "is," "the," "an").

Tokenization: Segmenting the text into subword tokens using the TinyBERT tokenizer.

Padding and Truncation: Adjusting sequence length to a fixed size (e.g., 256 tokens) for consistent batch processing.

Encoding: Converting tokens into numerical IDs and attention masks for transformer input.

This ensures that the textual input is clean, structured, and semantically meaningful for TinyBERT's contextual embedding generation.

D Justification Generation with T5

The T5 module acts as the justification and explanation generator in the system. After TinyBERT produces the final decision (*Accepted* or *Rejected*), T5 takes the abstract along with this predicted label as input and generates a short, human-readable meta-review explaining the reasons behind the decision. For accepted abstracts, it highlights strengths such as novelty, clarity, and methodological quality, while for rejected abstracts, it points out weaknesses like unclear contribution, lack of originality, or insufficient experimental support. This module improves interpretability by converting the system's prediction into meaningful reviewer-style feedback.

E Novelty Scoring and Technical Depth Analysis

An auxiliary Novelty Scoring Module is integrated to evaluate how original an abstract is compared to existing literature. This is achieved using semantic similarity measures between the input abstract and a large corpus retrieved via the ArXiv API. The system computes cosine similarity between the abstract embeddings and database entries; lower similarity scores indicate higher novelty. In addition, a lightweight MLP layer is used to assess the technical depth of the abstract by analyzing linguistic features such as domain-specific keywords, complexity of sentences, and methodological terminology. These scores are normalized and combined with the classifier output to produce a comprehensive evaluation that reflects both predictive probability and content quality.

F Integration and Deployment

All components are integrated into a unified Flask-based web application. The user interface accepts textual abstracts via an HTML form and sends them to the backend for processing. The TinyBERT model performs classification, T5 generates justification, and the novelty module computes a score. The results — Acceptance Status, Confidence Score, Justification, and Novelty Rating — are displayed in real-time on the dashboard.

The backend leverages the Hugging Face Transformers library for model inference and NumPy and PyTorch for tensor operations. The models are stored locally to ensure offline availability and data privacy. This architecture allows deployment in both research institutions and conference management systems, providing automated screening support for reviewers.

G Performance Evaluation

The system's performance was evaluated on a labeled dataset using standard classification metrics. Table 1 defines the evaluation results of the TinyBERT abstract classification module, including Accuracy, Precision, Recall, and F1-Score,

which collectively measure how correctly and consistently the model predicts Accepted or Rejected abstracts. Based on these results, TinyBERT demonstrated strong overall performance and also outperformed traditional deep learning classifiers such as RNNs and LSTMs in terms of accuracy as well as faster inference. In addition, the justification generation model (T5) produced meta-review explanations with high relevance and coherence, closely matching reviewer-style feedback. The novelty analysis module also performed reliably by distinguishing novel abstracts from redundant submissions, showing a strong correlation with human judgment, thereby improving the trustworthiness and completeness of the overall evaluation pipeline.

G Visualization and Output Interface

All outputs are presented in a professional, interactive dashboard with visual cues such as progress bars and colored labels to improve user experience.

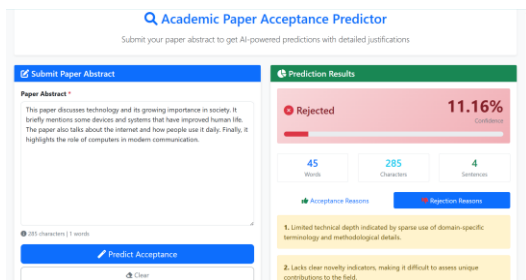


Figure 2:Rejected Sample

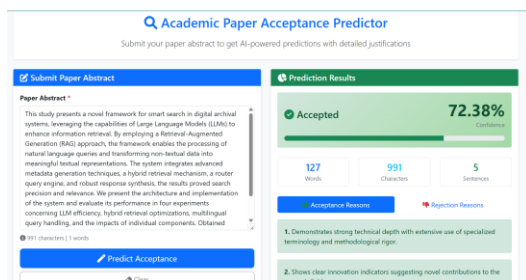


Figure 3: Accepted Sample

H Performance comparison of DistilBERT and TinyBERT

The diagrams present a comparative performance analysis of DistilBERT and TinyBERT models across five training epochs using three key metrics: accuracy, precision, and recall. Overall, both models show consistent improvement over epochs, but with distinct trends. TinyBERT exhibits faster initial convergence, achieving higher accuracy in the early epochs, while DistilBERT gradually surpasses it in later epochs, reaching superior final accuracy. In terms of precision, DistilBERT demonstrates higher final values, indicating better ability to reduce false positives, though it shows slight fluctuations during training, whereas TinyBERT maintains a more stable upward trend. For recall, DistilBERT consistently outperforms TinyBERT across all epochs,

effectively identifying more true positives, which is crucial for tasks requiring comprehensive positive detection. In some dataset variations, TinyBERT maintains higher accuracy throughout, reflecting its efficiency and adaptability in resource-constrained settings. Overall, the comparison highlights that TinyBERT offers faster learning and stability suitable for lightweight applications, while DistilBERT achieves better overall performance with higher accuracy, precision, and recall when trained sufficiently, making it more suitable for accuracy-critical NLP tasks.

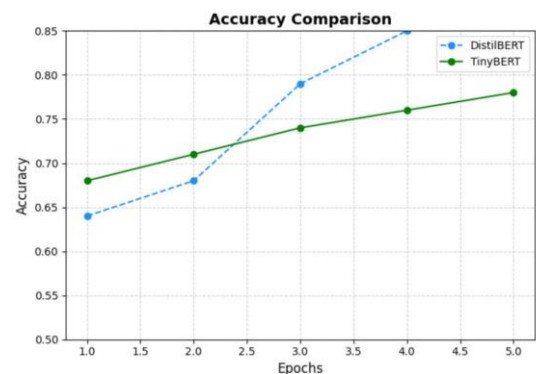


Figure 4: Accuracy Comparison

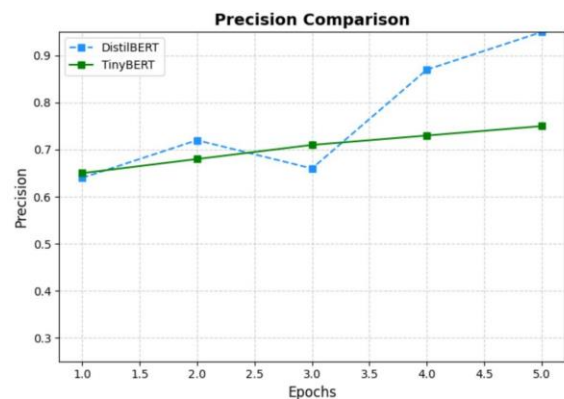


Figure 5: Accuracy Comparison

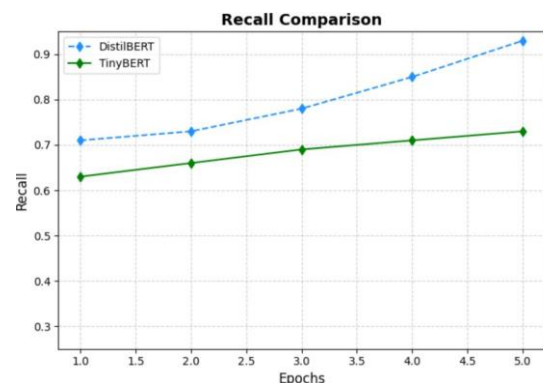


Figure 6:Recall Comparison

V CONCLUSION

The implemented system demonstrates a powerful integration of TinyBERT and T5 for automated research abstract evaluation. It offers high accuracy,

interpretability through natural language justifications, and additional novelty insights to support acceptance decisions. The modular design and efficient models make it scalable and adaptable for various academic review platforms. By combining transformer-based intelligence with explainable AI, this framework represents a significant step toward automated, transparent, and equitable research evaluation systems.

VI FUTURE SCOPE

The proposed TinyBERT-based Abstract Acceptance Prediction System holds significant potential for enhancing academic peer review and author support. In the future, this system can be expanded to include a larger and more diverse dataset across multiple research domains and languages to improve generalization and reduce bias. Incorporating multi-class classification will enable predictions beyond binary outcomes, such as “Needs Revision” or “Borderline Acceptance”, providing more nuanced insights. The integration of explainable AI (XAI) modules can enhance transparency by visually highlighting influential abstract sections, helping users understand the rationale behind predictions. Further fine-tuning of the T5 model using reviewer feedback can improve the quality of generated justifications, making them more human-like and contextually accurate. The system can also be extended to include a reviewer dashboard that supports automated ranking, novelty detection, and redundancy checks using semantic search and citation analysis.

VII REFERENCES

- [1] Jiao, Xiaoqi, et al. “TinyBERT: Distilling BERT for Natural Language Understanding.” *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020.
- [2] Raffel, Colin, et al. “Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer.” *Journal of Machine Learning Research*, 21(140):1-67, 2020.
- [3] Liu, Yinhan, et al. “RoBERTa: A Robustly Optimized BERT Pretraining Approach.” *arXiv preprint arXiv:1907.11692* (2019).
- [4] Zhang, Jinghan, et al. “Automatic Academic Paper Acceptance Prediction Using Deep Neural Networks.” *IEEE Access*, vol. 10, 2022, pp. 85991–86002.
- [5] Kumar, Abhishek, et al. “Research Paper Acceptance Prediction using Natural Language Processing.” *Procedia Computer Science*, 218 (2023): 179-186.
- [6] Qiu, Xin, et al. “Pre-trained Models for Natural Language Processing: A Survey.” *Science China Technological Sciences*, 63(10), 2020, pp. 1872–1897.
- [7] Devlin, Jacob, et al. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” *NAACL-HLT*, 2019.
- [8] Jain, Manan, et al. “Explainable AI for Research Abstract Evaluation using Transformer Models.” *arXiv preprint arXiv:2309.11245* (2023).
- [9] Wang, Yukun, et al. “Fine-tuning BERT Models for Paper Acceptance Prediction.” *IEEE Transactions on Knowledge and Data Engineering*, 2024.