

Optimizing Sentiment Analysis using LLM

THESIS

Submitted in partial fulfillment of the requirement for the Degree of

MASTER OF TECHNOLOGY

(Information Technology)

(Session: 2024-25)



Submitted By

RANJAN KUMAR

Roll No.: 23037104

Enrollment No.: GGV/19/1237

Under the Supervision of

MR. DEEPAK KANT NETAM

Assistant Professor

And

DR. AGNIVESH PANDEY

Assistant Professor

Department of Information Technology
Guru Ghasidas Vishwavidyalaya, Bilaspur (C.G.)
(A Central University established by the Central Universities Act 2009 No. 25 of 2009)

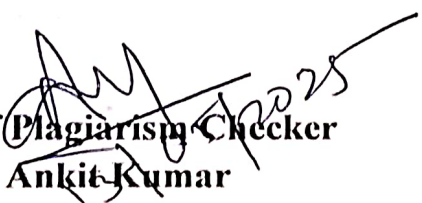
May 2025



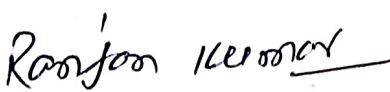
Department of Information Technology
Guru Ghasidas Vishwavidyalaya, Bilaspur (C.G.)
(A Central University established by the Central Universities Act 2009 No. 25 of 2009)

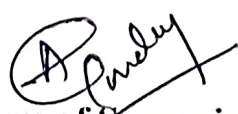
CERTIFICATE

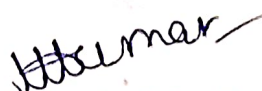
As per University Grant Commission (promotion of academic integrity and prevention of plagiarism in higher institute) regulation 2023 dated 8 May 2025 the thesis "Optimizing sentiment analysis using LLM" of Ranjan Kumar, student of M. TECH (Information Technology) final thesis has been checked by Turnitin software by Dr. Ankit Kumar (Assistant Professor), Guru Ghasidas Vishwavidyalaya (A Central University), Bilaspur, Chhattisgarh, INDIA. The amount of similarity in the thesis is 7% percent. Signed Turnitin similarity report is attached with certificate.


Signature of Plagiarism Checker
Dr. Ankit Kumar
(Assistant Professor)


Signature of Supervisor
Mr. Deepak Kant Netam
(Assistant Professor)


Signature of Student
Ranjan Kumar


Signature of Supervisor
Dr. Agnivesh Pandey
(Assistant Professor)


Signature of Head of the Department
Dr. Manoj Kumar
(Professor)

HEAD
Department of Information Technology
SoS, Engg. & Technology
Guru Ghasidas Vishwavidyalaya
(Central University) Bilaspur (C.G.)

ABSTRACT

Determining the emotional tone of textual input is a vital task in NLP and is frequently employed in applications like opinion mining and customer feedback analysis. BERT, a potent pre-trained transformer-based language model renowned for its profound contextual knowledge of text, is used in this dissertation to offer an improved sentiment analysis approach. The study is conducted on the Amazon Fine Food Reviews dataset and introduces two innovations to improve BERT's performance: TF-IDF-based term weighting and guided Gaussian noise injection. TF-IDF emphasizes significant terms in each review, while the guided Gaussian noise—proportional to these weights—is added to the BERT input embeddings during fine-tuning. By concentrating on relevant terms and avoiding overfitting, this hybrid strategy improves the model's capacity for generalization. The suggested model obtained an accuracy of 86.59% on training and an accuracy of 82.30% on validation, and demonstrated stable learning with training and validation losses of 0.3122 and 0.3607, respectively. By blending statistical insights with deep learning precision, this research contributes a robust and scalable method for sentiment analysis, offering practical benefits in real-world applications such as automated review summarization, product ranking, and user satisfaction analytics.

Keywords: Amazon Fine Food Reviews, BERT, Sentiment Analysis, Gaussian Noise, TF-IDF, NLP