TRIBHUVAN UNIVERSITY

Institute of Science and Technology

School of Mathematical Sciences

Kritipur, Kathmandu

A Dissertation Report on

NEPALI LEGAL DOCUMENT SUMMARIZATION USING MULTILINGUAL LARGE LANGUAGE MODELS

Submitted by:

Arpan Sapkota

(Registration No. 3-2-368-168-2016)

Under the Supervision of:

Rupak Raj Ghimire

Submitted to:
School of Mathematical Sciences
Kirtipur, Kathmandu, Nepal

Submitted in partial fulfillment of the requirements for the degree of Master in Data Science

July, 2025

Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES

Date: 2025 July 29

To Whom It May Concern

This is to certify that **Mr. Arpan Sapkota** has submitted **Master in Data Science** dissertation (MDS 699) entitled "NEPALI LEGAL DOCUMENT SUMMARIZATION USING MULTILINGUAL LARGE LANGUAGE MODELS" on **July 29, 2025** self-declaring that there is no difference in the contents of hard copy and soft copy. The provided copy has been checked as a whole by iThenticate plagiarism checker software under the Tribhuvan University anti-plagiarism procedure - 2080 and shows below 20% overall similarity index.

_____

Keshab Raj Phulara

Program Coordinator

TRIBHUVAN UNIVERSITY

Institute of Science and Technology

School of Mathematical Sciences

Kritipur, Kathmandu

STUDENT'S DECLARATION

I hereby declare that the contents written in this report are original and no sources other than mentioned here have been used in this work.

_____

Arpan Sapkota

Date: 2025 July 29

TRIBHUVAN UNIVERSITY

Institute of Science and Technology

School of Mathematical Sciences

Kritipur, Kathmandu

SUPERVISOR'S RECOMMENDATION

I hereby recommend that this Dissertation prepared under my supervision by **Arpan Sapkota** titled **"Nepali Legal Document Summarization using Multilingual Large Language Models"** in partial fulfillment for the degree of Master in Data Science be processed for evaluation.

Rupak Raj Ghimire,

School of Mathematical Sciences,

Institute of Science and Technology,

Tribhuvan University,

Kritipur, Kathmandu, Nepal

Date: 2025 July 29

TRIBHUVAN UNIVERSITY

Institute of Science and Technology

School of Mathematical Sciences

Kritipur, Kathmandu

Date: 2025 July 29

APPROVAL SHEET

This is to certify that we have read this Dissertation and, in our opinion, it is applicable for the scope and the quality of Dissertation in partial fulfillment of the requirement for the degree of Master in Data Science.

Evaluation Committee

..... **Assoc. Prof. Nawaraj Paudel** **(Director)** School of Mathematical Sciences IOST, Tribhuvan University, Nepal	..... **Rupak Raj Ghimire** **(Supervisor)** School of Mathematical Sciences IOST, Tribhuvan University, Nepal
..... **Prof. Dr. Shashidhar Ram Joshi** **(External Examiner)** Tribhuvan University, Nepal	..... **Sarbin Sayami** **(Internal Examiner)** Tribhuvan University, Nepal

Date: 2025 July 29

ACKNOWLEDGMENT

First and foremost, I would like to express my sincere gratitude to my supervisor, **Rupak Raj Ghimire**, for his continuous guidance, encouragement, and invaluable insights throughout this research journey. His expertise in Natural Language Processing and unwavering support made this dissertation possible. I am deeply thankful to **Prof. Dr. Yubaraj Sangroula**, Executive Director and Professor at Kathmandu School of Law, for his expert guidance and insightful perspectives on the structure and interpretation of Nepali legal case documents in the context of the Supreme Court of Nepal. My appreciation further goes to **Assoc. Prof. Nawaraj Paudel**, Director of the School of Mathematical Sciences, Tribhuvan University for fostering a supportive academic environment and for his encouragement during the completion of this dissertation. I am also grateful to the faculty and staff of the School of Mathematical Sciences, for providing a conducive academic environment and necessary resources for completing my Master's in Data Science. Special thanks to my classmates and friends for their constant motivation and constructive feedback during my research. I would also like to acknowledge the contributions of open source community, especially the Hugging Face, WandB and Kaggle platforms, for making the models and GPU easily accessible. My heartfelt appreciation goes to my family for their endless support, patience, and encouragement throughout my academic career.

Lastly, I am grateful to the Government of Nepal and the publishers of the *Nepal Kanun Patrika* for providing access to public legal documents, which formed the core of my research dataset.

Arpan Sapkota

2025 July 29

ABSTRACT

Legal documents in the Nepali language are often lengthy, complex, and written in domain specific language, ...

Keywords: LLM, Legal Text Summarization, NLP, Abstractive Summarization, mBART, mT5, Transformer Models, Low-Resource Language, ROUGE, BLEU, Hugging Face.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
API	Application Programming Interface
BART	Bidirectional and Auto-Regressive Transformers
BERT	Bidirectional Encoder Representations from Transformers
BLEU	Bilingual Evaluation Understudy
CUAD	Contract Understanding Atticus Dataset
GPT	Generative Pretrained Transformer
GPU	Graphics Processing Unit
LLM	Large Language Model
METEOR	Metric for Evaluation of Translation with Explicit Ordering
mLLM	Multilingual Large Language Model
MLSUM	Multilingual Summarization Dataset
mBART	Multilingual Bidirectional and Auto-Regressive Transformers
mBERT	Multilingual Bidirectional Encoder Representations from Transformers
mT2	Multilingual Text-to-Text Transfer Transformer
NER	Named Entity Recognition
NKP	Nepal Kanun Patrika
NLG	Natural Language Generation
NLP	Natural Language Processing
OCR	Optical Character Recognition
POS	Part-of-Speech
ROUGE	Recall-Oriented Understudy for Gisting Evaluation
T2	Text-to-Text Transfer Transformer
TF-IDF	Term Frequency-Inverse Document Frequency
XLM-R	Cross-Lingual Language Model-RoBERTa

CHAPTER 1: INTRODUCTION

1.1 Background

1.1.1 Text Summarization

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Types of S

Text summarization techniques can be broadly classified into the following categories:

1. T1 Summarization
2. T2 Summarization
3. H Summarization

1. Extractive Summarization

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

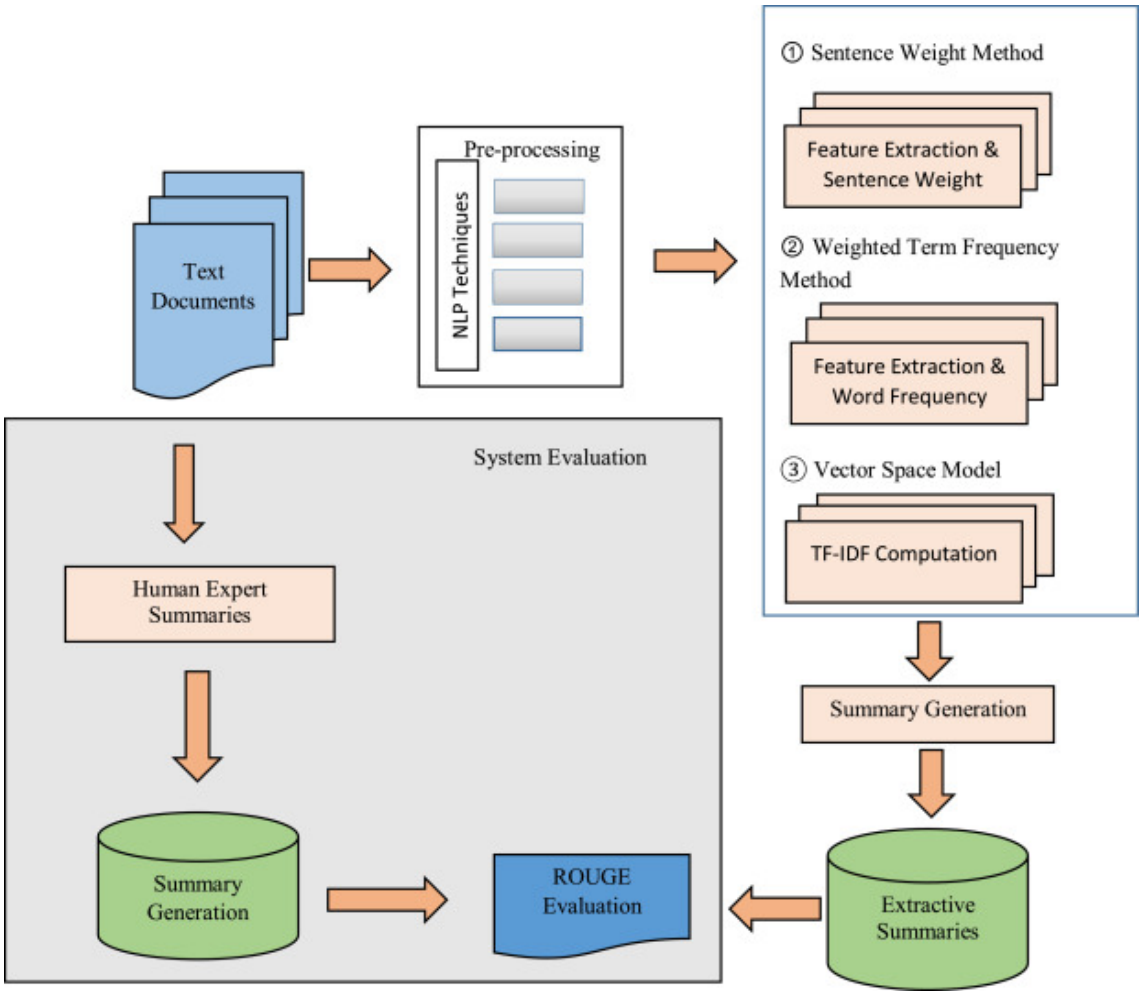

Figure 1.1: Extractive Text Summarization (Nawaz et al., 2020)

Example Techniques of Extractive Summarization:

- TextRank (Mihalcea & Tarau, 2004): A graph-based ranking model where sentences are nodes, and edges represent similarity between them .
- LexRank (Erkan & Radev, 2004): Similar to PageRank, it ranks sentences based on eigenvector centrality.

2. Abstractive Summarization

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat

non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

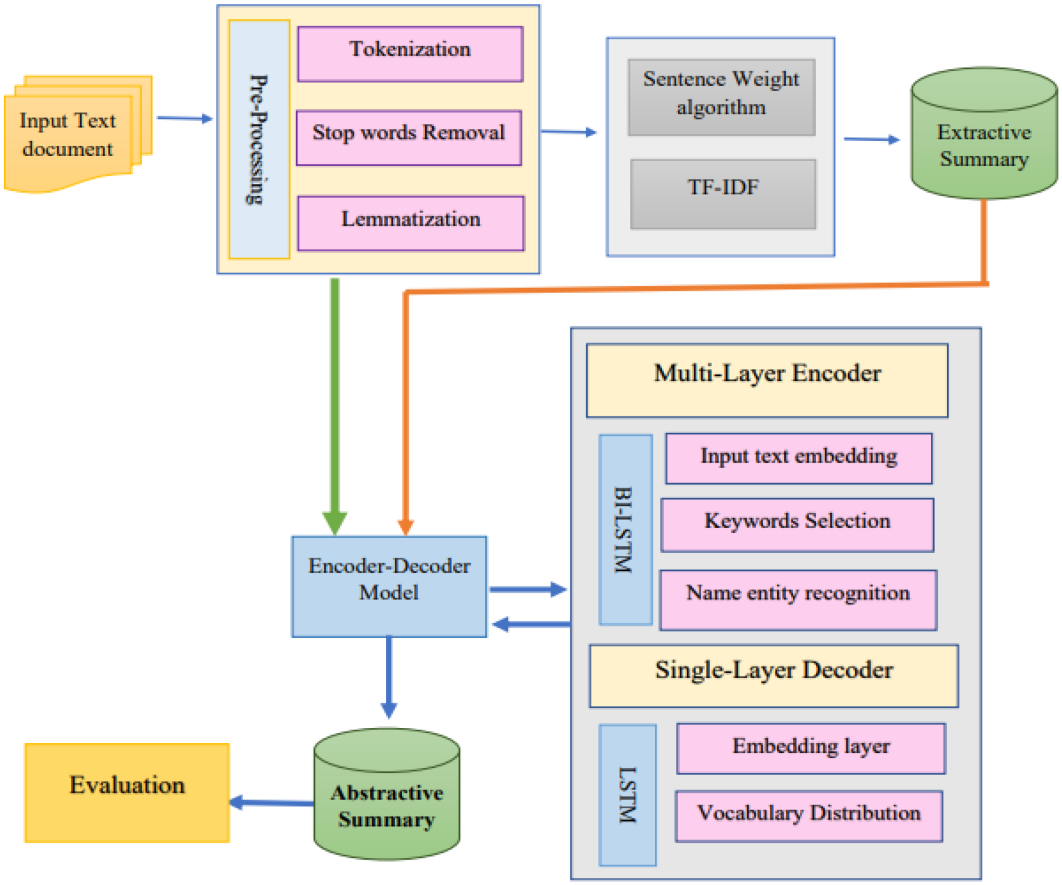

Figure 1.2: Abstractive Text Summarization (Shafiq et al., 2023)

Example Techniques of Abstractive Summarization:

- Pointer-Generator Networks (See, Liu, & Manning, 2017): Combines the benefits of copying words from the source text and generating new words.
- Text-to-Text Transfer Transformer (T5) : A transformer-based model pre-trained on multiple Natural Language Processing (NLP) tasks including summarization .
- Bidirectional and Auto-Regressive Transformers (BART) (Lewis et al., 2020): Denoising autoencoder for pretraining sequence-to-sequence models, especially effective for the abstractive summarization.

3. Hybrid Summarization

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.2 Problem Statement

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.3 Research Objectives

The objectives of this research are:

1.4 Research Questions

1.5 Significance of Study

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in

voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.6 Scope and Limitations

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.7 Report Organization

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

CHAPTER 2: LITERATURE REVIEW

2.1 Overview of Summarization

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

2.2 International Advances

2.3 Nepali and Devanagari Scripts Challenges

2.4 Legal Document Summarization

2.5 Research Gaps

2.6 Theoretical Framework

CHAPTER 3: METHODOLOGY

3.1 Research Method

3.2 Data Collection

3.3 Data Preprocessing

3.3.1 Text Cleaning

3.3.2 Data Normalization

3.3.3 Tokenization

3.4 LLM based Data Augmentation

3.5 Fine- Tuning the mBART Model

3.5.1 mBART: Architecture and Working Mechanism

3.5.2 Transformer Architecture in mBART

Encoder:

Decoder:

Language-Aware Training

3.5.3 Pretraining Objective

3.5.4 Fine-Tuning for Summarization

3.5.5 System Architecture of mBART for Summarization

Multilingual Denoising Pretraining: Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Transformer Encoder (Pretraining Stage): Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Transformer Decoder (Pretraining Stage): Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Fine-Tuning on Summarization Task: Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Transformer Encoder (Fine-tuning Stage): Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Transformer Decoder (Fine-tuning Stage): Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Output: Summary (Nepali): Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

3.6 Implementation Details

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

3.7 Model Deployment

To make the fine-tuned models publicly accessible and reusable, both versions of the Multilingual Bidirectional and Auto-Regressive Transformers (mBART) model were deployed to the Hugging Face Model Hub Wolf et al. (2020).

Deployment Details

Two separate models were deployed based on the datasets used during the fine-tuning process:

- **Raw Data Fine-Tuned Model:** This model was trained on the original Nepali legal summaries, without any paraphrasing or augmentation. It is hosted at [1]

- **Paraphrased Data Fine-Tuned Model:** This version was fine-tuned using the Large Language Model (LLM) augmented summaries, where paraphrasing was performed using the LLaMA3 8B 8192 model (AI, 2024) through the Groq Application Programming Interface (API). It is accessible at [2]

3.8 Model Evaluation

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

3.8.1 Evaluation Dataset

The dataset was split using the hold-out validation technique.

- **Training Set:** 80% of the data was used for training the model.

- **Validation Set:** 10% of the data was used to evaluate the model performance every after 200 steps.

- **Test Set:** 10% of the data was reserved for testing the model's generalization and summarization ability.

A fixed seed (42) was used to ensure consistent and reproducible splits during model fine-tuning.

[1] *https://huggingface.co/arpansapkota/Legal_mbart_large_50_finetuned_RawData*
[2] *https://huggingface.co/arpansapkota/Legal_mbart_large_50_finetuned_ParaData*

3.8.2 Evaluation Metrics

ROUGE (Recall-Oriented Understudy for Gisting Evaluation):

BLEU (Bilingual Evaluation Understudy):

CHAPTER 4: RESULTS AND DISCUSSION

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

4.1 Experimental Setup

4.2 Model Performance

4.2.1 Loss Curves Comparison

Figures 4.1 illustrate training and validation loss curves. The RawData model exhibits rapid convergence and stabilizes at a significantly lower loss (0.0902) compared to the ParaData model (0.4254). This suggests better adaptation and generalization capabilities when fine-tuned directly on the original legal dataset.

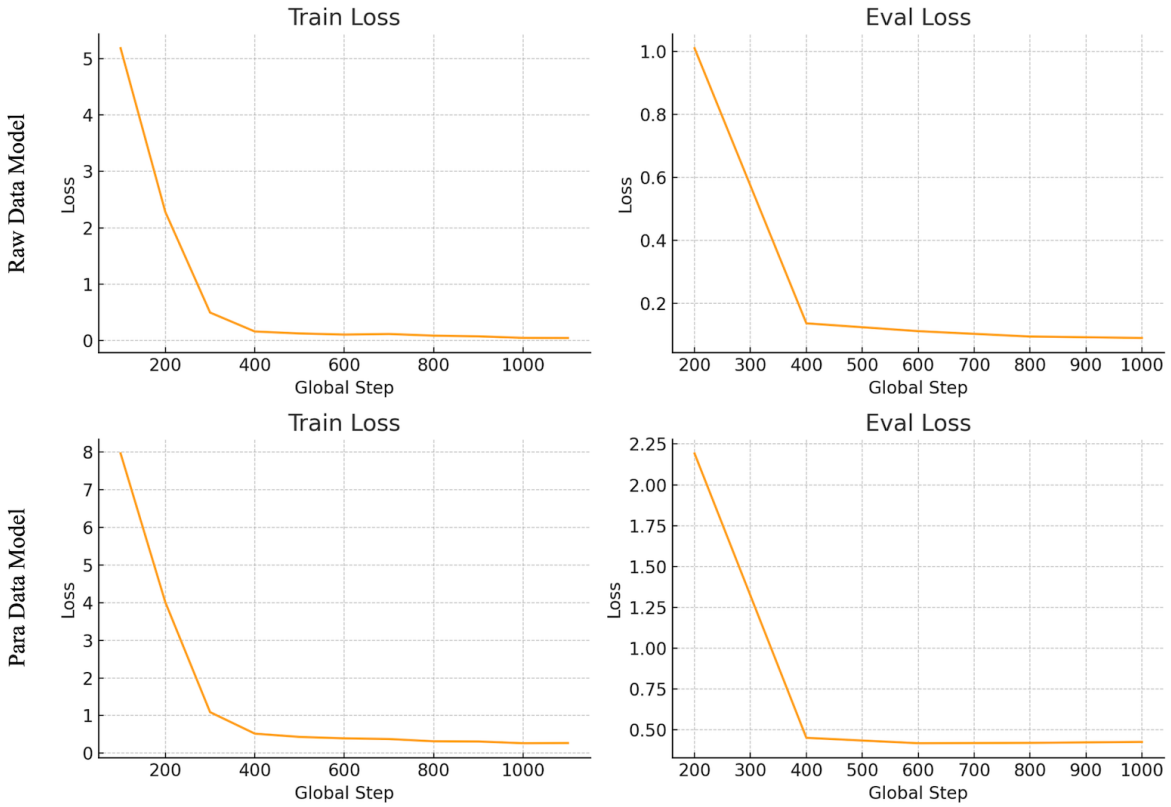

Figure 4.1: Training and validation loss observed during training of Raw and Paraphrased Data Model

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

4.2.2 Training Speed and Runtime Analysis

Analysis of runtime and evaluation speed metrics further provides insights into training dynamics. Curves such as `eval/runtime`, `eval/steps_per_second`, and `eval/samples_per_second` (illustrated in Figure 4.2) indicate consistent evaluation efficiency. Specifically, the RawData model maintains a stable runtime with higher evaluation throughput, suggesting slightly more efficient utilization of computational resources.

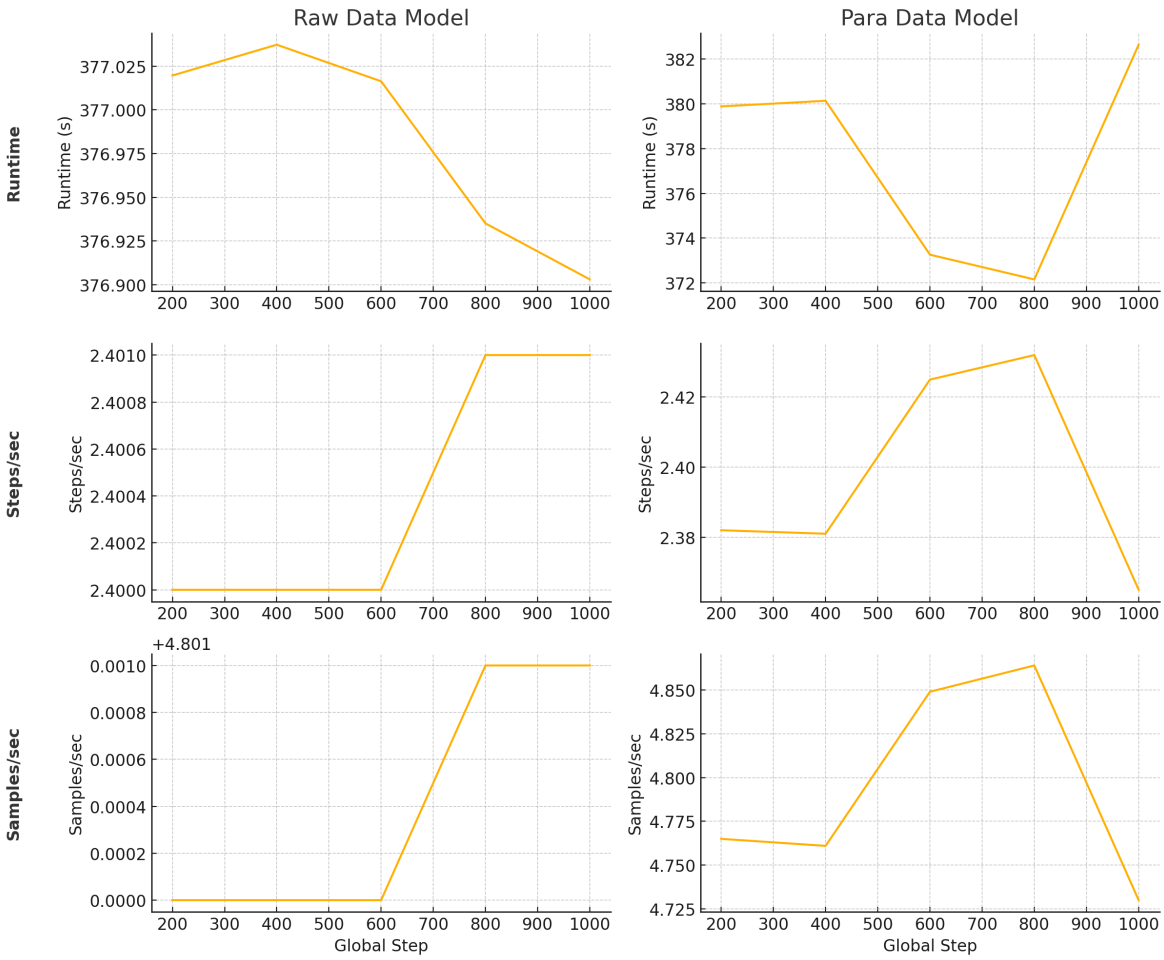

Figure 4.2: Evaluation runtime and throughput metrics for Raw and Paraphrased Data Model

4.2.3 Training Dynamics and Hyperparameter Trends

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

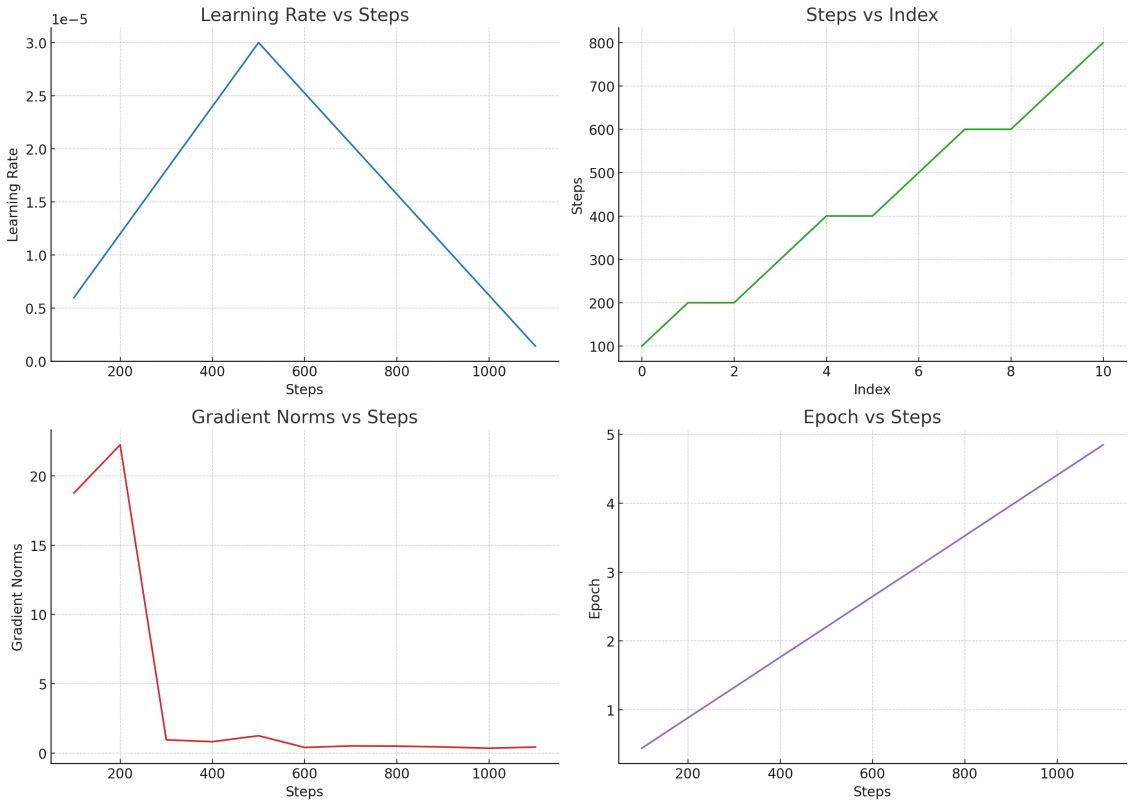

Figure 4.3: Training dynamics (learning rate, gradient norms, steps, and epochs) for Raw Data Model.

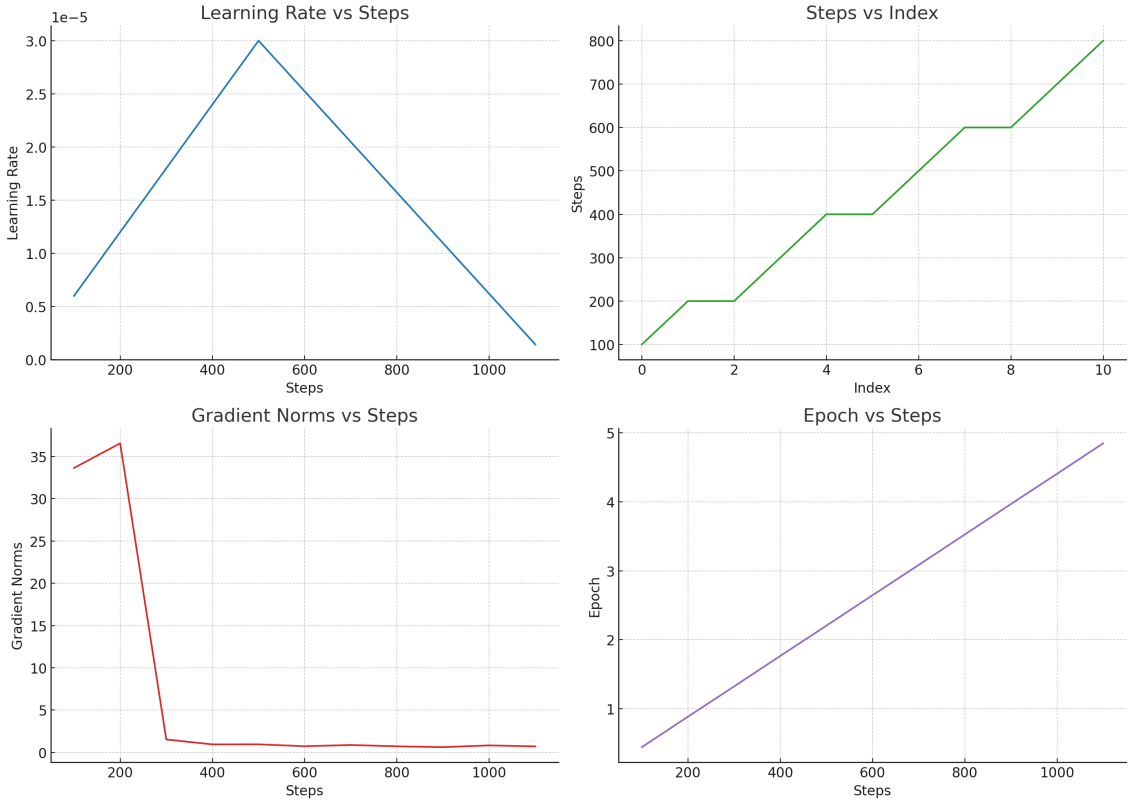

Figure 4.4: Training dynamics (learning rate, gradient norms, steps, and epochs) for Paraphrased Data Model.

4.3 Model Inference

4.4 Evaluation Results

4.5 Qualitative Observations

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

4.6 Discussion and Interpretation

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

4.7 Comparison with Existing Approaches

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

4.8 Error Analysis

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in

voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

CHAPTER 5: CONCLUSION AND RECOMMENDATION

5.1 Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

5.2 Findings

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

5.3 Limitations

5.4 Contributions to the Field

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

5.5 Future Work

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco

laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

REFERENCES

AI, M. (2024). *Llama 3 technical report.* https://llama.meta.com/llama3. (Accessed: 2025-01-15)

Erkan, G., & Radev, D. R. (2004, July). LexPageRank: Prestige in multi-document text summarization. In D. Lin & D. Wu (Eds.), *Proceedings of the 2004 conference on empirical methods in natural language processing* (pp. 365–371). Barcelona, Spain: Association for Computational Linguistics. Retrieved from https://aclanthology.org/W04-3247/

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... Zettlemoyer, L. (2020). Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of acl* (pp. 7871–7880). doi: 10.18653/v1/2020.acl-main.703

Mihalcea, R., & Tarau, P. (2004, July). TextRank: Bringing order into text. In D. Lin & D. Wu (Eds.), *Proceedings of the 2004 conference on empirical methods in natural language processing* (pp. 404–411). Barcelona, Spain: Association for Computational Linguistics. Retrieved from https://aclanthology.org/W04-3252/

Nawaz, A., Bakhtyar, M., Baber, J., Ullah, I., Noor, W., & Basit, A. (2020). Extractive text summarization models for urdu language. *Information Processing & Management*, *57*(6), 102383. Retrieved from https://www.sciencedirect.com/science/article/pii/S0306457320308785 doi: https://doi.org/10.1016/j.ipm.2020.102383

See, A., Liu, P. J., & Manning, C. D. (2017, July). Get to the point: Summarization with pointer-generator networks. In R. Barzilay & M.-Y. Kan (Eds.), *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 1073–1083). Vancouver, Canada: Association for Computational Linguistics. Retrieved from https://aclanthology.org/P17-1099/ doi: 10.18653/v1/P17-1099

Shafiq, N., Hamid, I., Asif, M., Nawaz, Q., Aljuaid, H., & Ali, H. (2023, 01). Abstractive text summarization of low-resourced languages using deep learning. *PeerJ. Computer science*, *9*, e1176-e1176. doi: 10.7717/peerj-cs.1176

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... Brew, J. (2020,

October). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: System demonstrations* (pp. 38–45). Association for Computational Linguistics. Retrieved from https://aclanthology.org/2020.emnlp-demos.6