चित्रण: An Automated Festive Poster Generator with Wishes in Nepali Language

Team members:
Kristina Bhandari (THA077BCT022)
Kristina Ghimire (THA077BCT023)
Pradeepti Dongol (THA077BCT033)
Namita Bhatta (THA077BCT048)

Supervised By:
Er. Saroj Shakya

Department of Electronics and Computer Engineering
Institute of Engineering, Thapathali Campus

August, 2024

Outline

- Introduction

- Motivation

- Objectives

- Scope of Project

- Project Applications

- Methodology

- Results

- Analysis of Results

- Remaining Tasks

- Tentative Timeline

- Estimated Project Budget

- References

Introduction

- Automates the creation of visually appealing festival posters.

- Creates beautiful posters with Nepali texts for Nepali festivals (Dashain, Tihar, Chhath, New Year, and Holi).

- Takes prompts in English and creates posters that include festive wishes in Nepali text, based on the prompts given.

Motivation

- Repetitive task of poster creation for more than 50 Nepali festivals.

- Costly, tedious and inefficient traditional process

- Divert manpower to less prioritized tasks

- Automation improves organization's performance

Objectives

- To analyze input prompt to extract festival themes, then generate concise Nepali short title.

- To generate a Nepali festival-themed image and integrate it with a styled title in Nepali to create a digital poster.

Scope of Project

Capabilities

• Nepali Text Generation
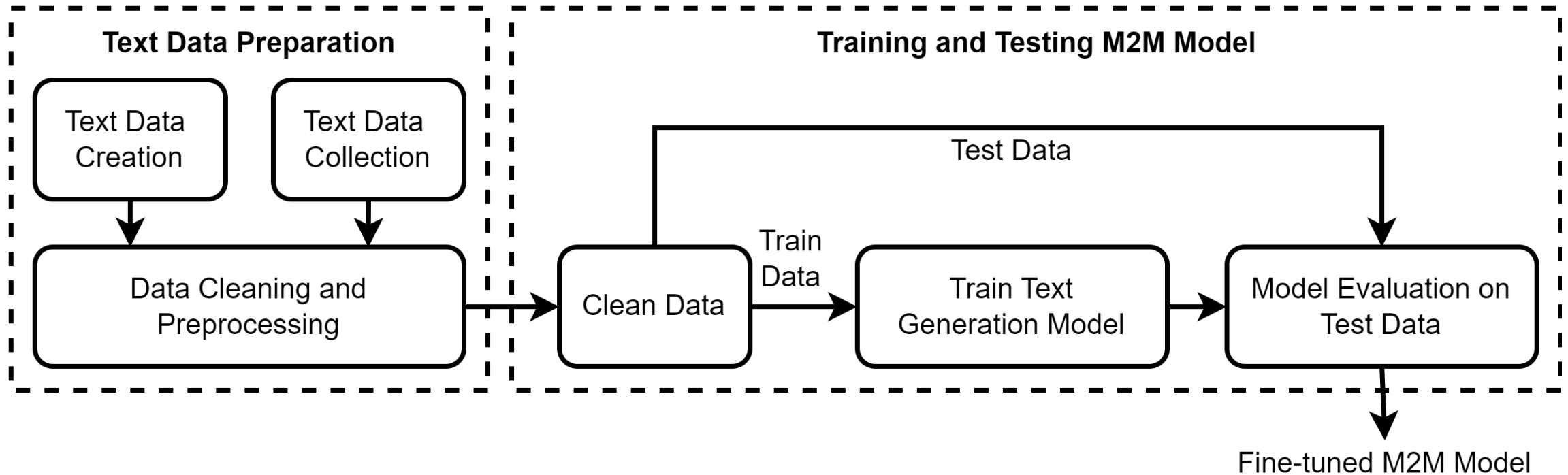
• Text to Image Synthesis

• Automated Poster Design

Limitations

• Only Five Festivals Poster generation

• Less diversity in text styling

Project Applications

- Personal Use

- Business and Corporate Use
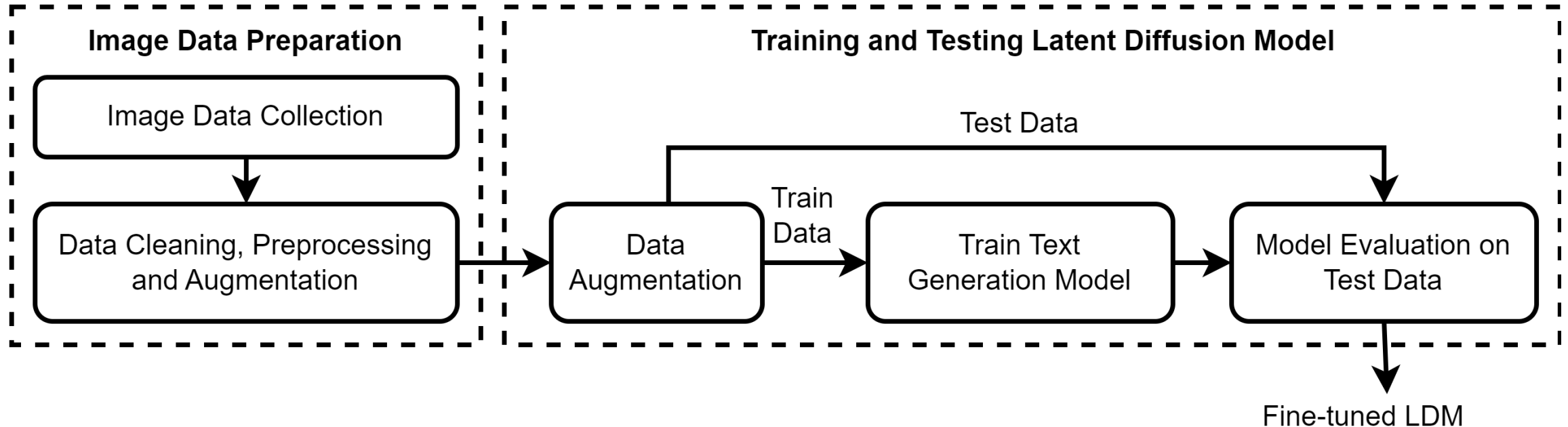
- Educational Institutions
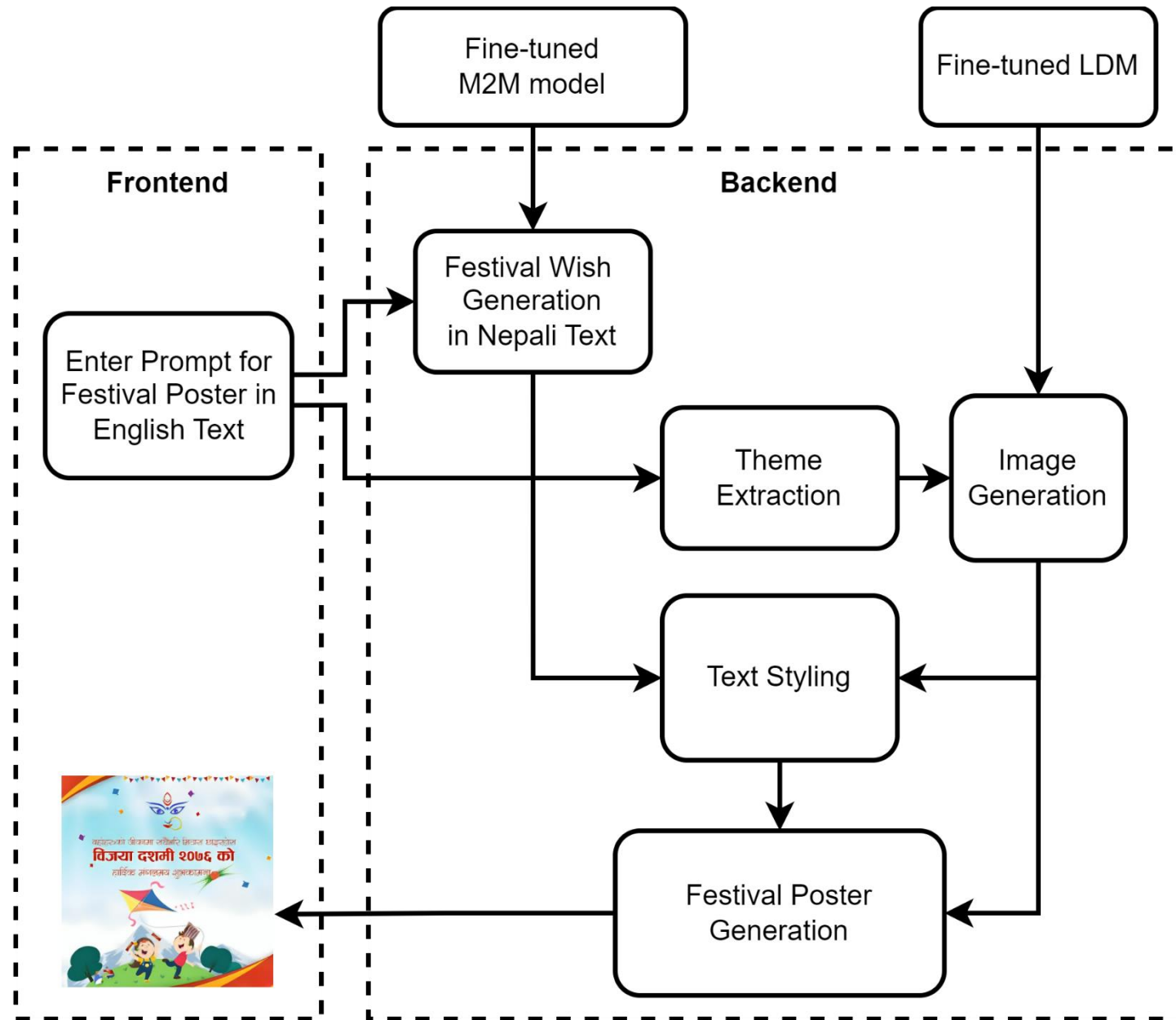
Methodology
(System Block Diagram)

Methodology – [contd] (System Block Diagram)
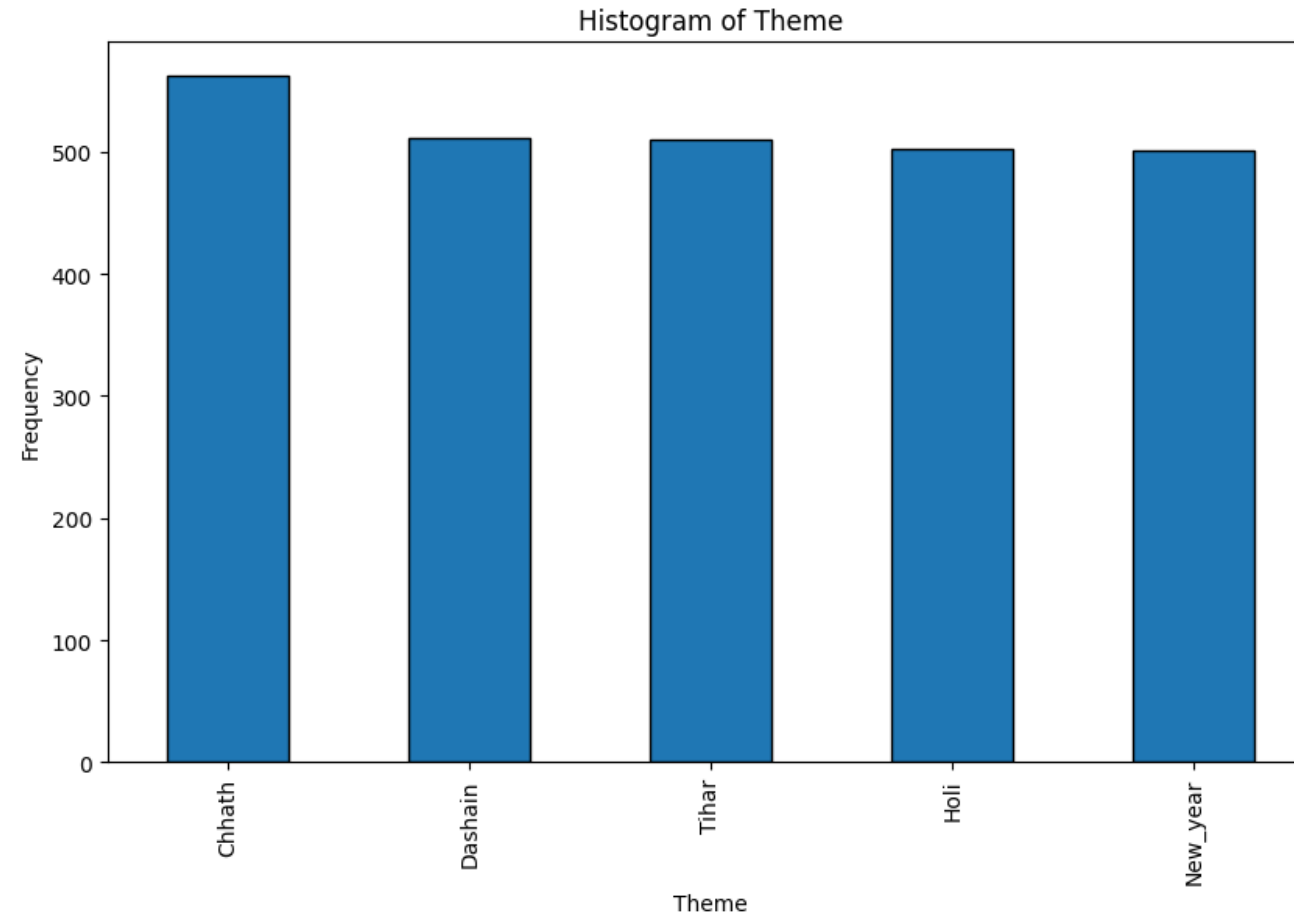
Methodology – [contd] (System Block Diagram)

Methodology – [contd]
(Text Dataset Preparation)

- Collected manually from different sources such as social media posters, greeting cards, and so on.

- Dataset contains text prompt in English and its corresponding wishes in Nepali font.

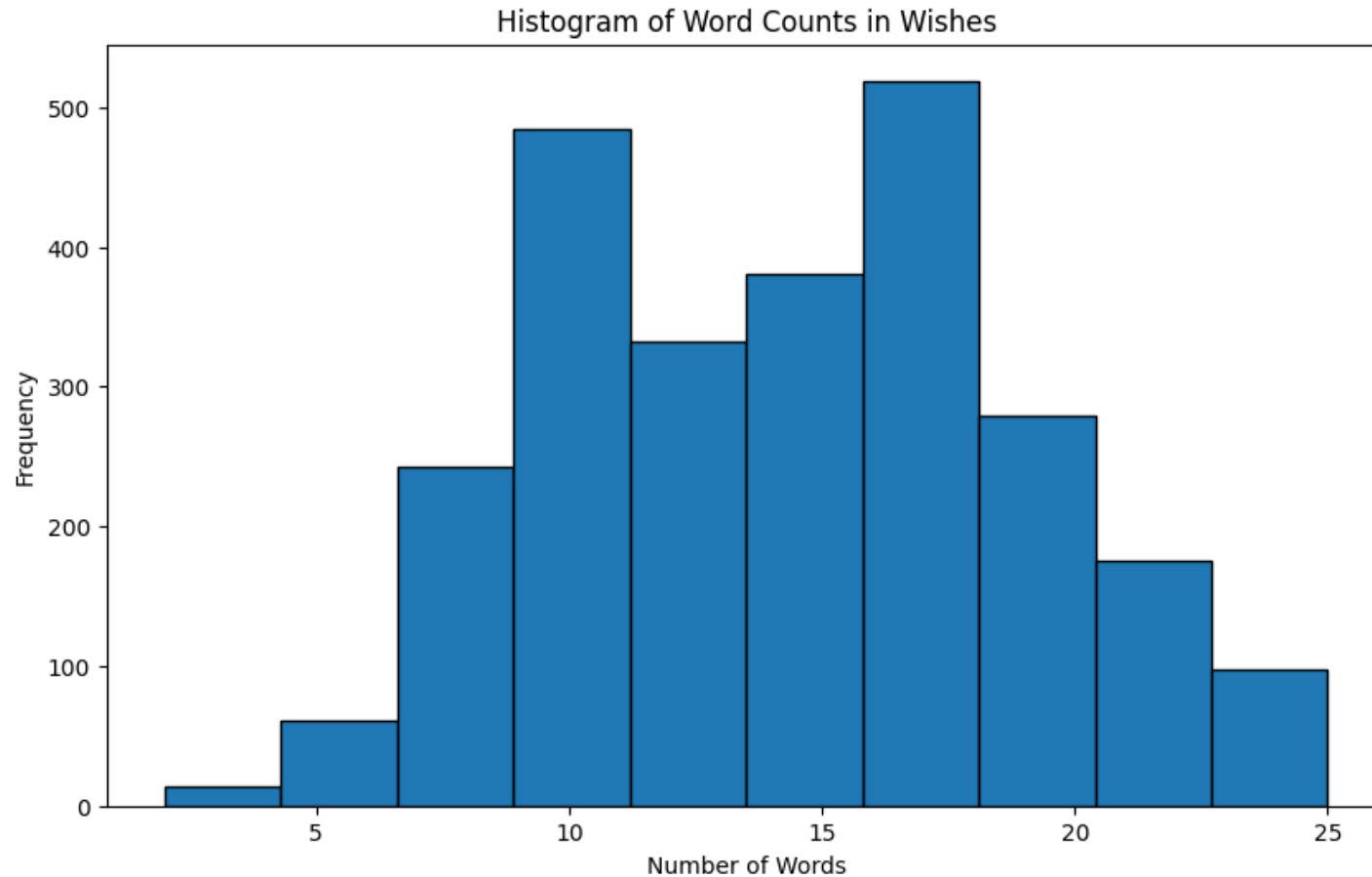- Collected 2586 pairs of prompts and wishes so far with token limit 25

Methodology – [contd]
(Text Dataset Preparation)

Methodology – [contd]
(Text Dataset Preparation)

Histogram of Word Counts in Wishes

Methodology – [contd] (Text Dataset Preprocessing)

- Spelling correction

- Grammar correction

- Year correction

- Dataset diversity

  - Varying length (in short, in detail)

  - Addressing general and specific audience

  - Including festival symbolic wishes and so on

Methodology – [contd]
(Image Dataset Preparation)

- Focused on five Nepali festivals: Dashain, Tihar, Chhath, New Year, and Holi.

- Festive images are collected from various websites like Pinterest, Facebook, and Dribble.

- Any text, logo and other information is removed.

- The image quality is enhanced.

Methodology – [contd]
(M2M-100 Model)

- Transformer-based model that can directly translate between 100 languages

- Uses 12 encoder layers and 12 decoder layers with 1024 hidden dimensions and 16 attention heads

- Pre-trained on a large-scale multilingual corpus

- Outperforms bilingual models and English-centric models

Methodology – [contd]
(M2M-100 Model)
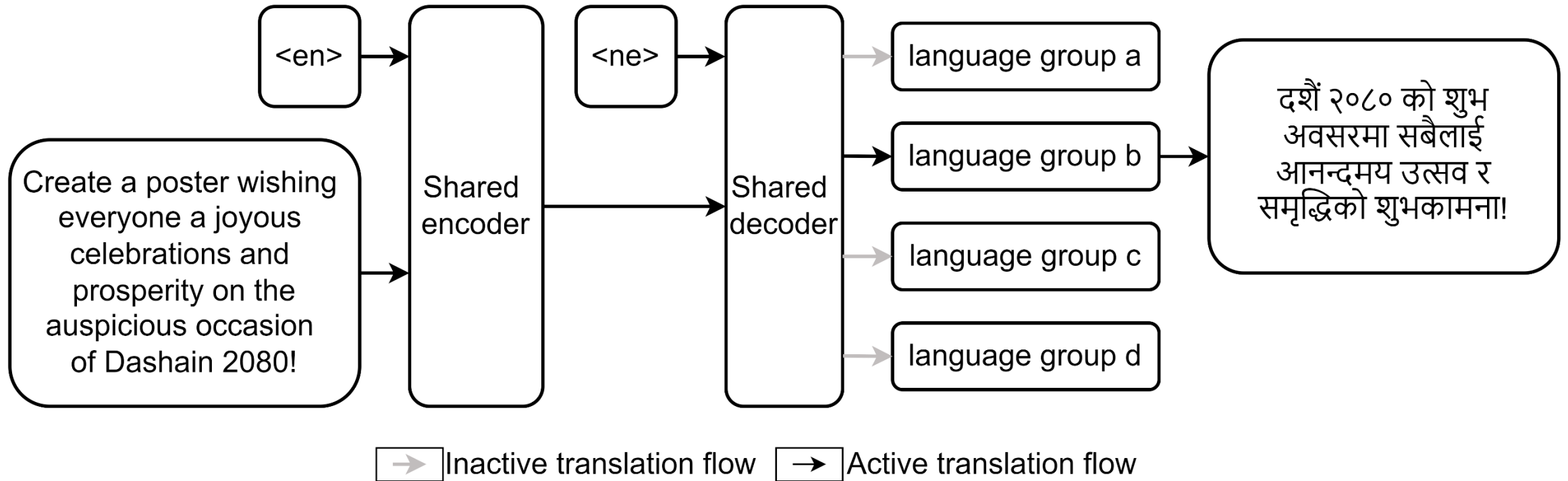
Methodology – [contd]
(M2M-100 Model)

- Encoder converts sequence of tokens $W = (w_1 \ldots, w_s)$ and

  source language l_s into a high-dimensional representation H.
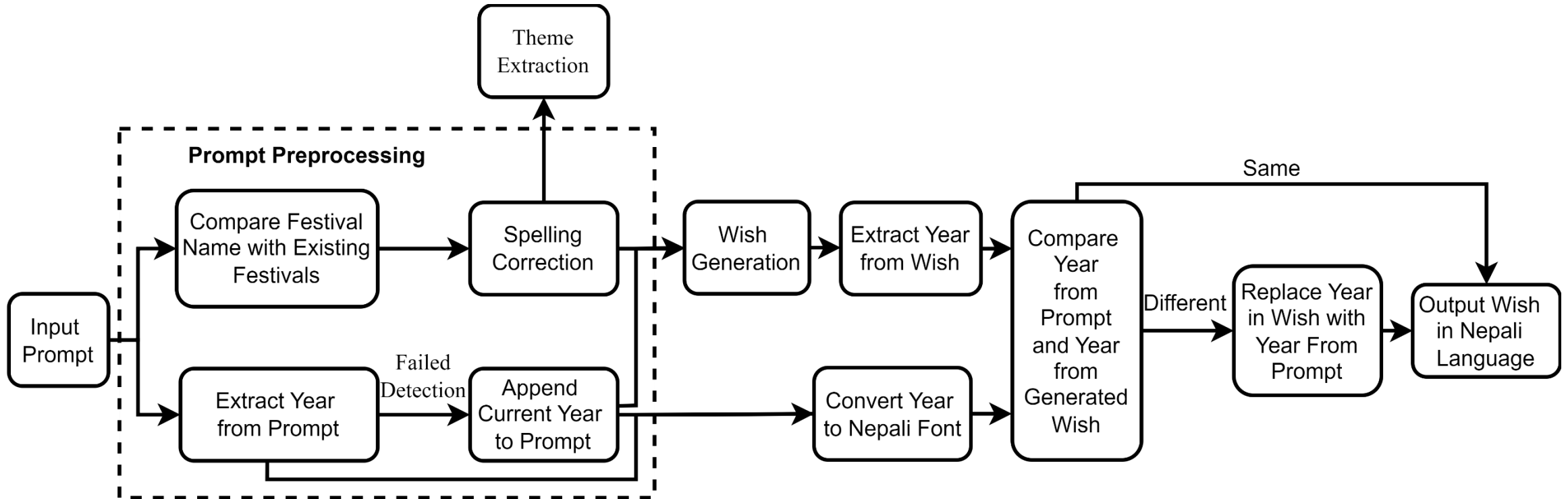
$$H = encoder(W, l_s)$$

- Decoder generates next token using the target language l_t,

  sequence of target tokens $(v_1 \ldots v_i)$ and encoder's

  output H.

$$v_{i+1} = decoder\ (H, l_t, v_1 \ldots v_i)$$

Methodology – [contd]
(Title Generation)

Methodology – [contd]
(Title Generation)

- For title generation, pretrained M2M-100 model is fine tuned on our dataset.

- BLEU score, ROUGE score, F1-score, METEOR score, TER score, ChrF score is used for evaluation.

- M2M-100 model takes English prompt and generates festival wish in Nepali language.

Methodology – [contd] (BLEU Score)

- Compare generated and reference text by measuring n-gram overlap

- Ranges from 0 too 100

$$BLEU = BP \times exp\left(\sum pn\right)$$

- where, BP (Brevity Penalty) adjust score for translation shorter than reference and pn is precision of n-grams

Methodology – [contd] (METEOR Score)

- Incorporates exact, stem, synonym and paraphrase matches

- Ranges from 0 to 1

$$M = F_{mean}(1 - Penalty)$$

where,

$$F_{mean} = \frac{P.R}{\propto. p + (1 - \propto). R}$$

$$Penalty = \gamma \left(\frac{C}{m} \right)^{\beta}$$

Methodology – [contd] (ROUGE Score)

- Rouge-N measures similarity using n-gram overlap

- Ranges from 0 to 1

$$ROUGE - N = \frac{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count(gram_n)}$$

- Rouge-L measures longest common subsequence between generated and reference

$$ROUGE - L = \frac{LCS(generated.reference)}{Length(reference)}$$

Methodology – [contd] (Self-BLEU Score)

- Measure diversity of generated texts

- Ranges from 0 to 1

$$self - BLEU = \frac{1}{N} \sum_{i=1}^{N} BLEU(g_i, G\backslash\{g_i\})$$

- where, $G = \{g_1, g_2, ..., g_N\}$ is the set of generated texts and

  BLEU $(g_i, G\backslash\{g_i\})$ is the BLEU score of g_i against the rest

Methodology – [contd]
(ChrF Score)

- Character n-gram F-score

- Evaluates translation quality based on precision and recall

$$ChrF = (1 + \beta) \frac{ChrP \cdot ChrR}{\beta^2 \cdot ChrP + ChrR}$$

- where, β parameter balances precision and recall

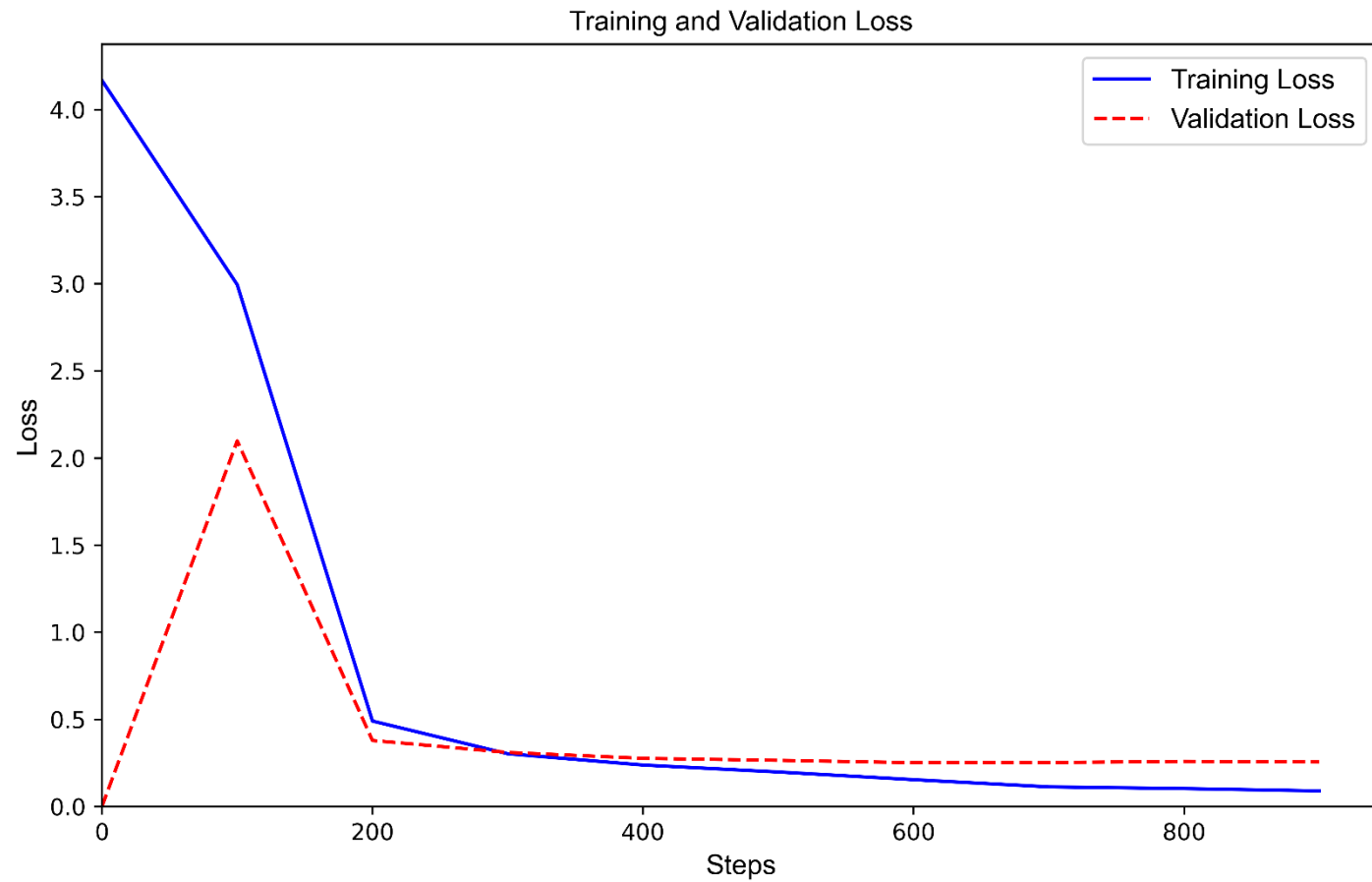# Methodology – [contd]
(F1 Score)

- Ranges from 0 to 1

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
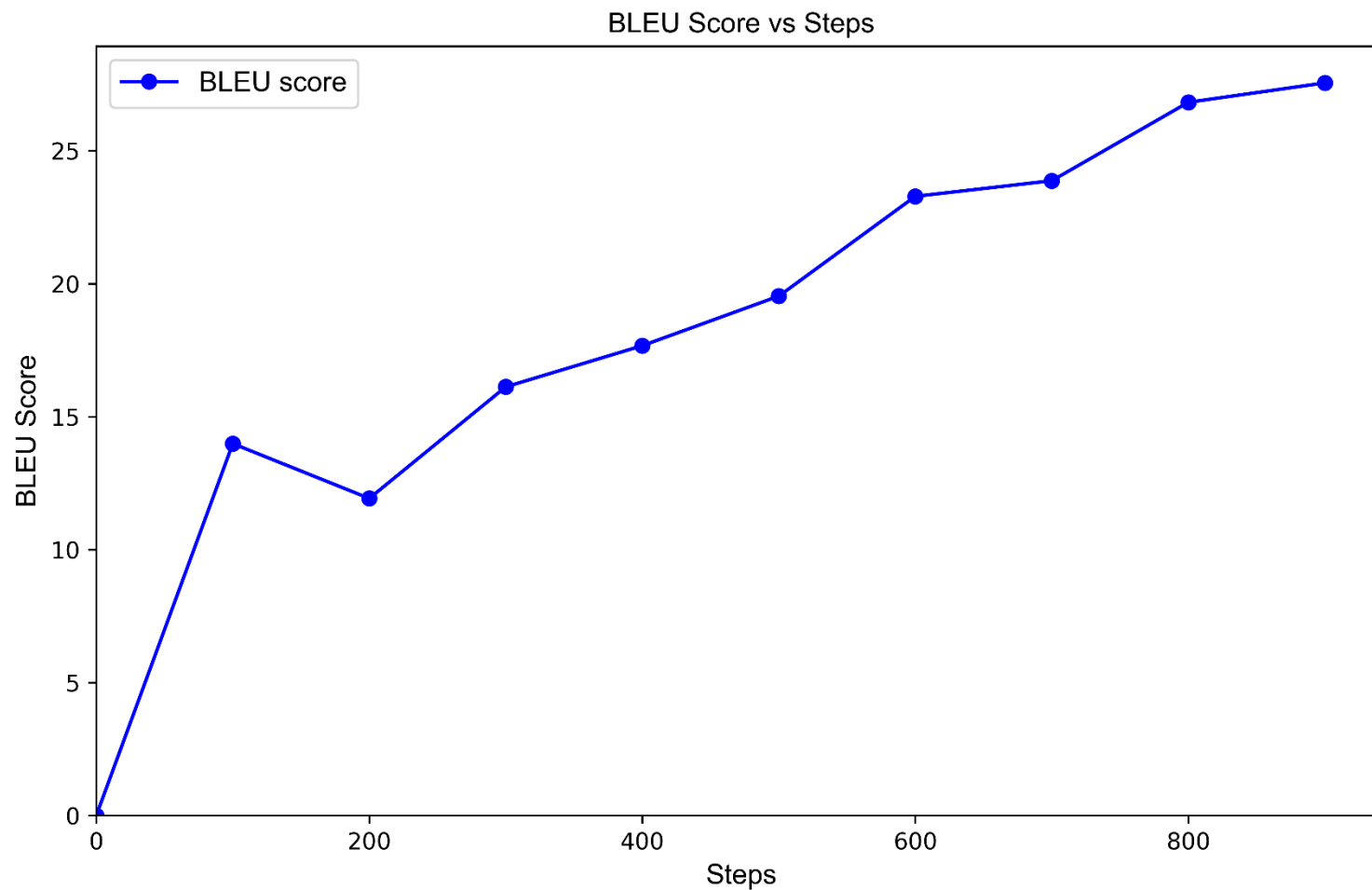
Methodology - [contd]
(Software tools and library)

- PIL/ Pillow

- ColorThief

- OpenCV

- FastAPI

- NumPy

- Pandas

- ReactJS

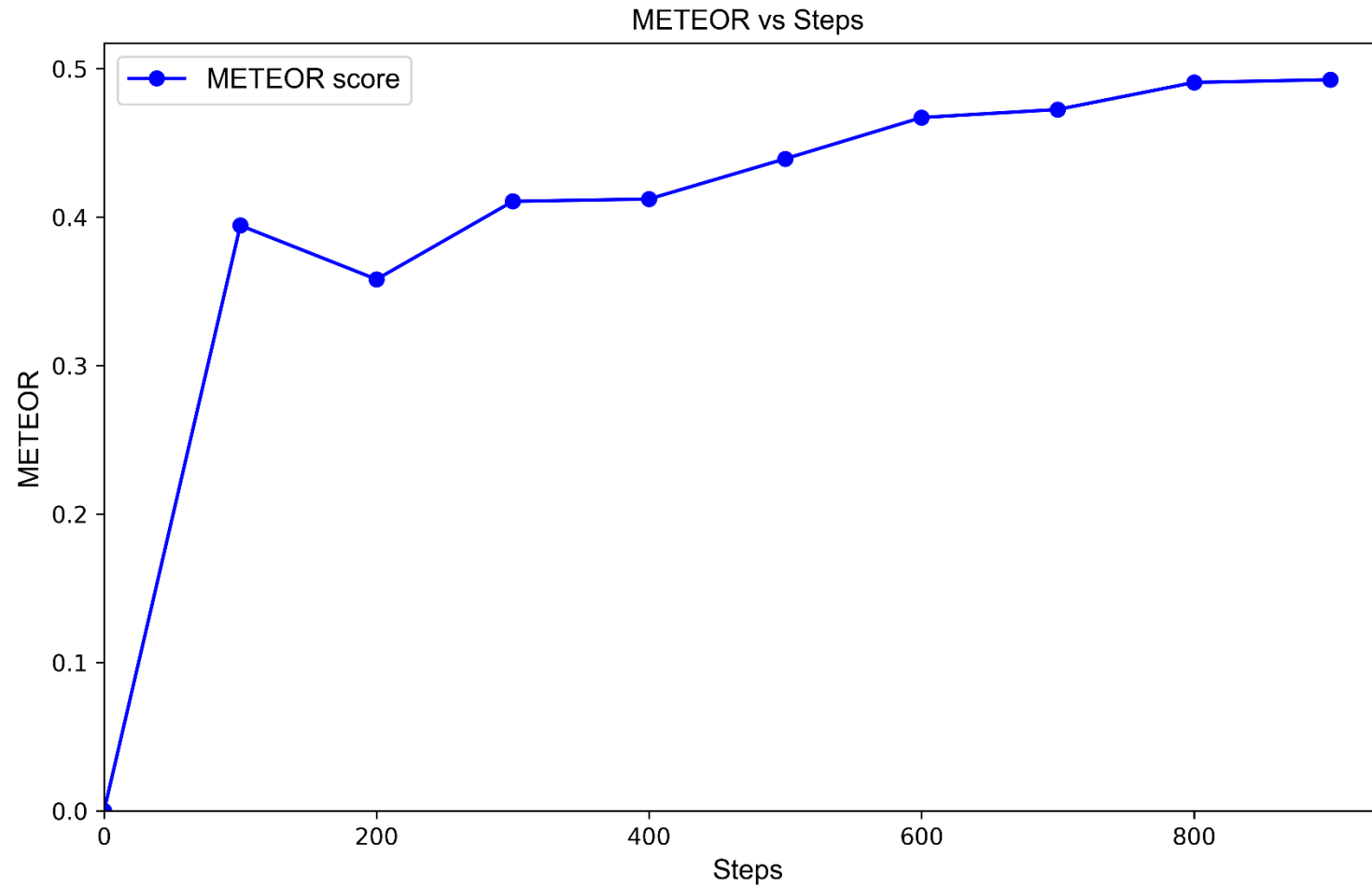- HTML/ CSS

Results
(Loss Plot)

Results - [contd]
(BLEU Score Plot)
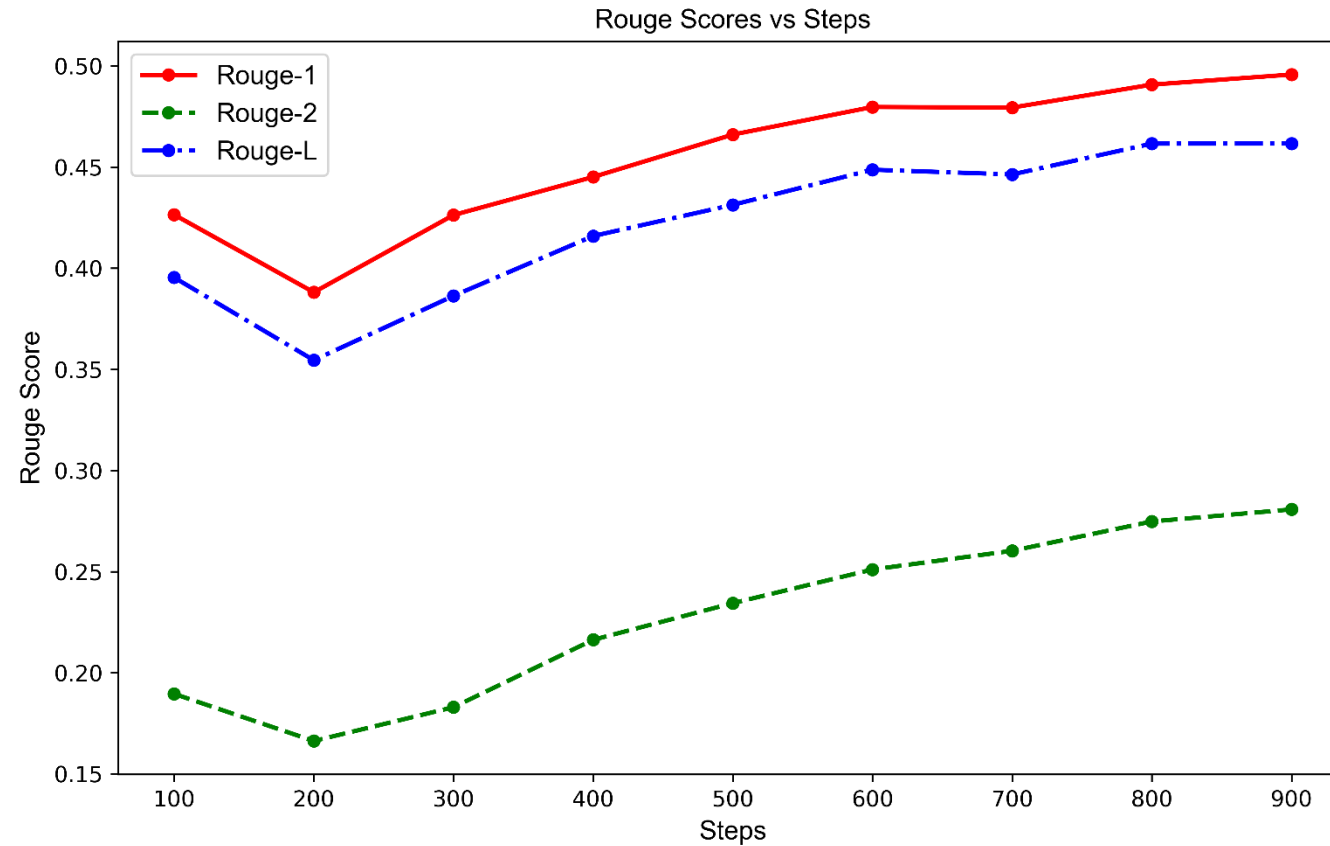
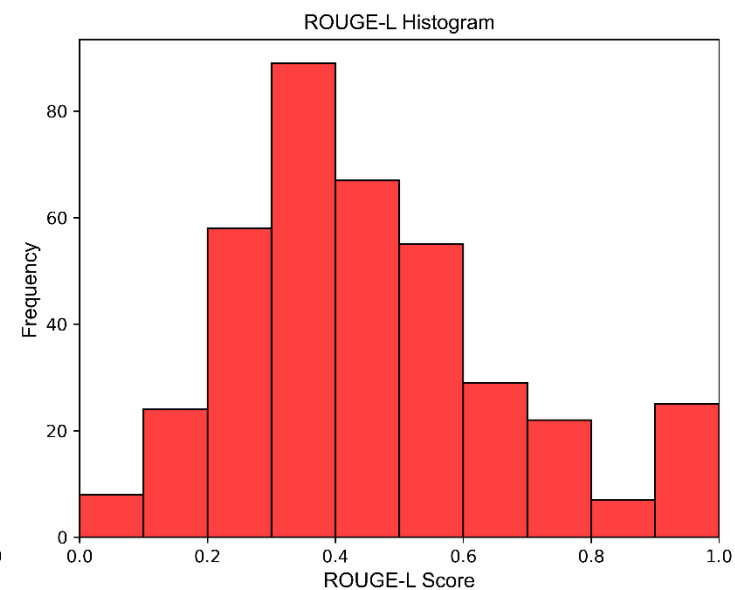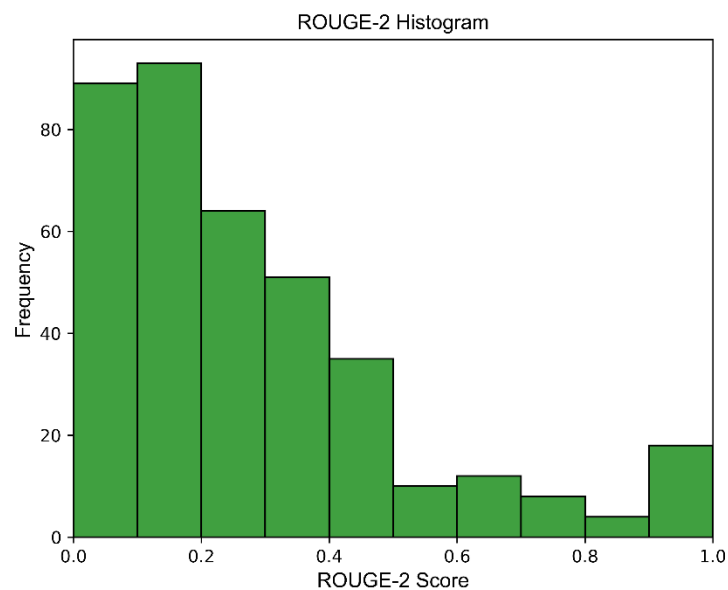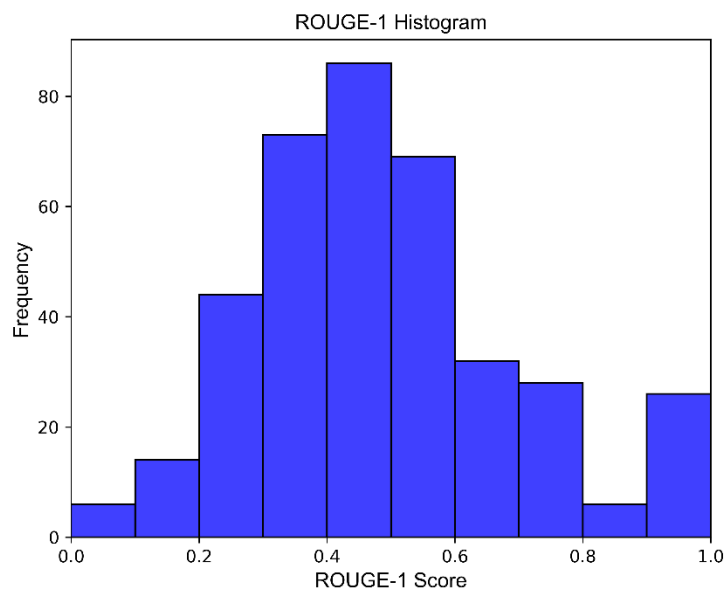# Results - [contd]
(Meteor Score Plot)

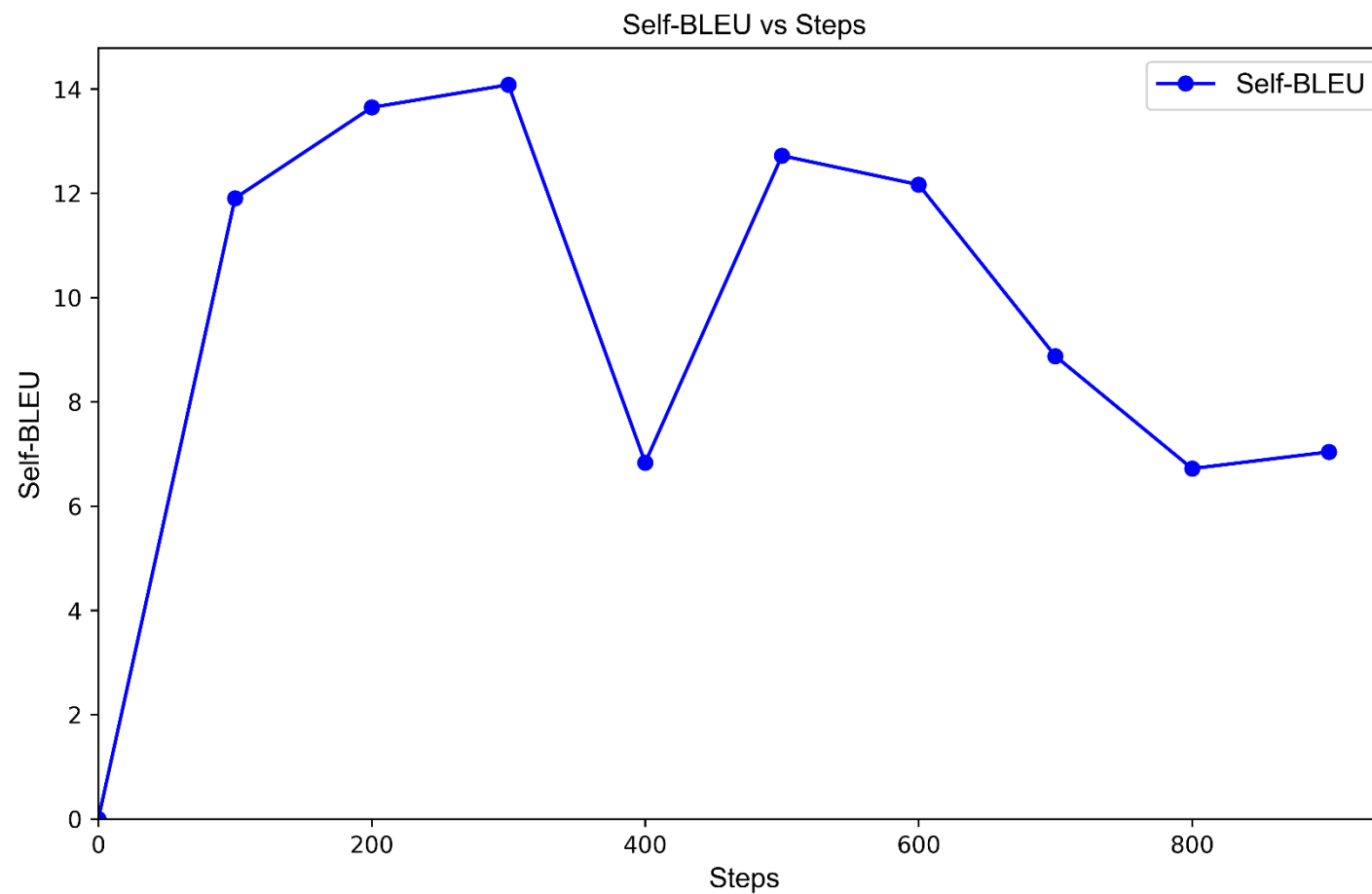# Results - [contd] (ROUGE Score Plot)
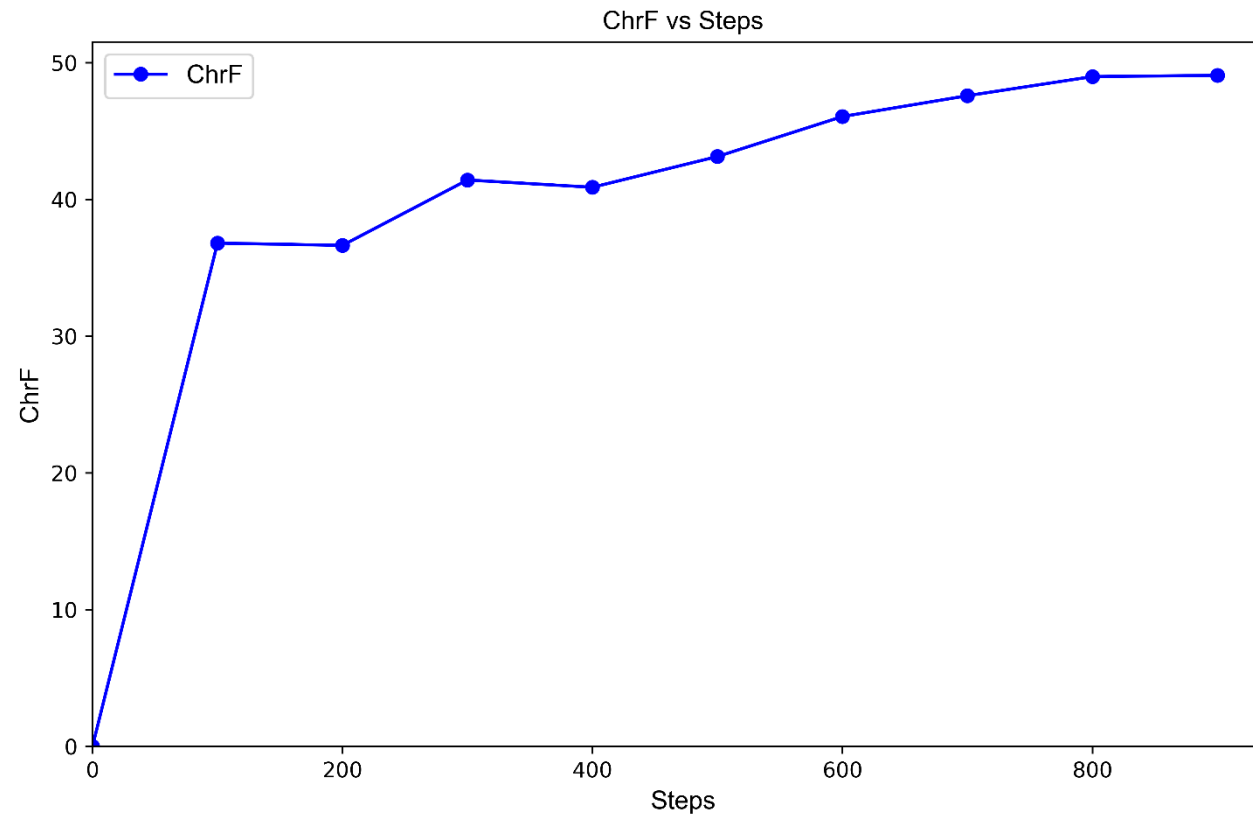
Results - [contd]
(ROUGE Score Plot)

Results - [contd]
(Self-BLEU Score Plot)

Results - [contd]
(ChrF Score Plot)

Results - [contd]
(F1 Score Plot)
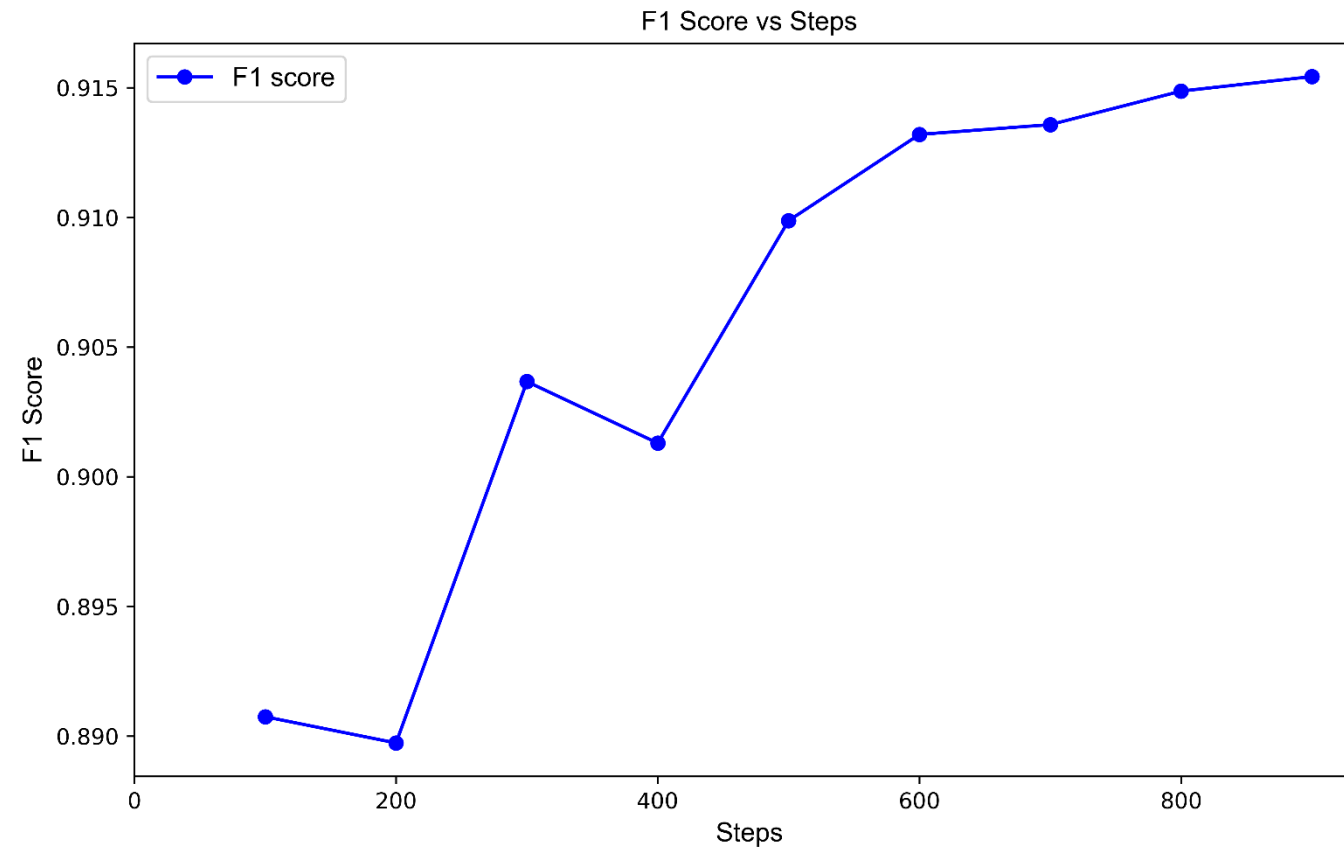
Results - [contd]
(Text Generation)

- Captures context of prompt

```
Input Text: Give me Chhath puja 2084 poster talking about the holly food eaten in Chhath
Generated Text: छठ पूजा २०८४ को शुभकामना! छठ पूजामा ठेकुवा, पुरी, र खीरको मिठास सँगै जीवनमा खुसीयाली साट्नुहोस्।
Theme: Chhath
Extracted Year from Prompt: 2084
Extracted Year from Wish: २०८४
Updated Prompt: Give me Chhath puja 2084 poster talking about the holly food eaten in Chhath
Corrected Year: छठ पूजा २०८४ को शुभकामना! छठ पूजामा ठेकुवा, पुरी, र खीरको मिठास सँगै जीवनमा खुसीयाली साट्नुहोस्।
```

- Addresses specific audience

```
Input Text: Generate a poster for Holi 2078 wishing my friends a festival filled with fun and laughter
Generated Text: सबै मित्रहरूलाई होली २०७८ को हार्दिक शुभकामना! रंगहरूको यो पर्वले तपाईंको जीवनमा खुशी र हर्ष ल्याओस्।
Theme: Holi
Extracted Year from Prompt: 2078
Extracted Year from Wish: २०७८
Updated Prompt: Generate a poster for Holi 2078 wishing my friends a festival filled with fun and laughter
Corrected Year: सबै मित्रहरूलाई होली २०७८ को हार्दिक शुभकामना! रंगहरूको यो पर्वले तपाईंको जीवनमा खुशी र हर्ष ल्याओस्।
```

Results - [contd]
(Text Generation)

- Different length of wishes based on prompt


```
Input Text: Develop a festive poster to wish my neighbors a happy Tihar 2075 in short.
Generated Text: तिहार २०७५ मा छिमेकीहरूलाई सुख, शान्ति र समृद्धिको शुभकामना!
Theme: Tihar
Extracted Year from Prompt: 2075
Extracted Year from Wish: २०७५
Updated Prompt: Develop a festive poster to wish my neighbors a happy Tihar 2075 in short.
Corrected Year: तिहार २०७५ मा छिमेकीहरूलाई सुख, शान्ति र समृद्धिको शुभकामना!
```



```
Input Text: Develop a festive poster to wish my family a happy Dashain 2076 in detail.
Generated Text: मेरो परिवारलाई २०७६ सालको बडा दशैंको हार्दिक शुभकामना! यो दशैंले सबैको जीवनमा आनन्द, समृद्धि र स्वास्थ्यको आशिर्वाद ल्याओस्।
Theme: Dashain
Extracted Year from Prompt: 2076
Extracted Year from Wish: २०७६
Updated Prompt: Develop a festive poster to wish my family a happy Dashain 2076 in detail.
Corrected Year: मेरो परिवारलाई २०७६ सालको बडा दशैंको हार्दिक शुभकामना! यो दशैंले सबैको जीवनमा आनन्द, समृद्धि र स्वास्थ्यको आशिर्वाद ल्याओस्।
```

Results - [contd]
(Text Generation)

- Deals with absence of year in prompt


```
Input Text: Develop a poster to wish everyone a Happy Tihar with Goddess Laxmi's blessing
Generated Text: सबैलाई तिहार २०८१ को शुभकामना! लक्ष्मी देवीको आशिर्वादले तपाईंको जीवनमा खुशी र समृद्धि ल्याओस्।
Theme: Tihar
Extracted Year from Prompt: 2081
Extracted Year from Wish: २०८१
Updated Prompt: Develop a poster to wish everyone a Happy tihar 2081 with Goddess Laxmi's blessing
Corrected Year: सबैलाई तिहार २०८१ को शुभकामना! लक्ष्मी देवीको आशिर्वादले तपाईंको जीवनमा खुशी र समृद्धि ल्याओस्।
```

- Year Correction


```
Generated Text: बडा दशैं २०७५ को हार्दिक शुभकामना! यस दशैंले तपाईंको जीवनमा खुशी, समृद्धि, र उत्साह ल्याओस्।
Theme: Dashain
Extracted Year from Prompt: 2081
Extracted Year from Wish: २०७५
Updated Prompt: In detail, create a visually appealing poster for dashain 2081 that captures the spirit and excitement of the celebration.
Corrected Year: बडा दशैं २०८१ को हार्दिक शुभकामना! यस दशैंले तपाईंको जीवनमा खुशी, समृद्धि, र उत्साह ल्याओस्।
```

Results Analysis

- Training and evaluation loss decreases overtime

- On training 2586 text dataset using M2M model, we achieved maximum of 27.55 BLEU score.

- Self- BLEU score of 7.04 that represents diversity.

- Obtained an maximum of 49% of Meteor score.

- Obtained an maximum of 49.06 CHRF.

Results Analysis[contd]

- F1 score was observed to be 91.54%
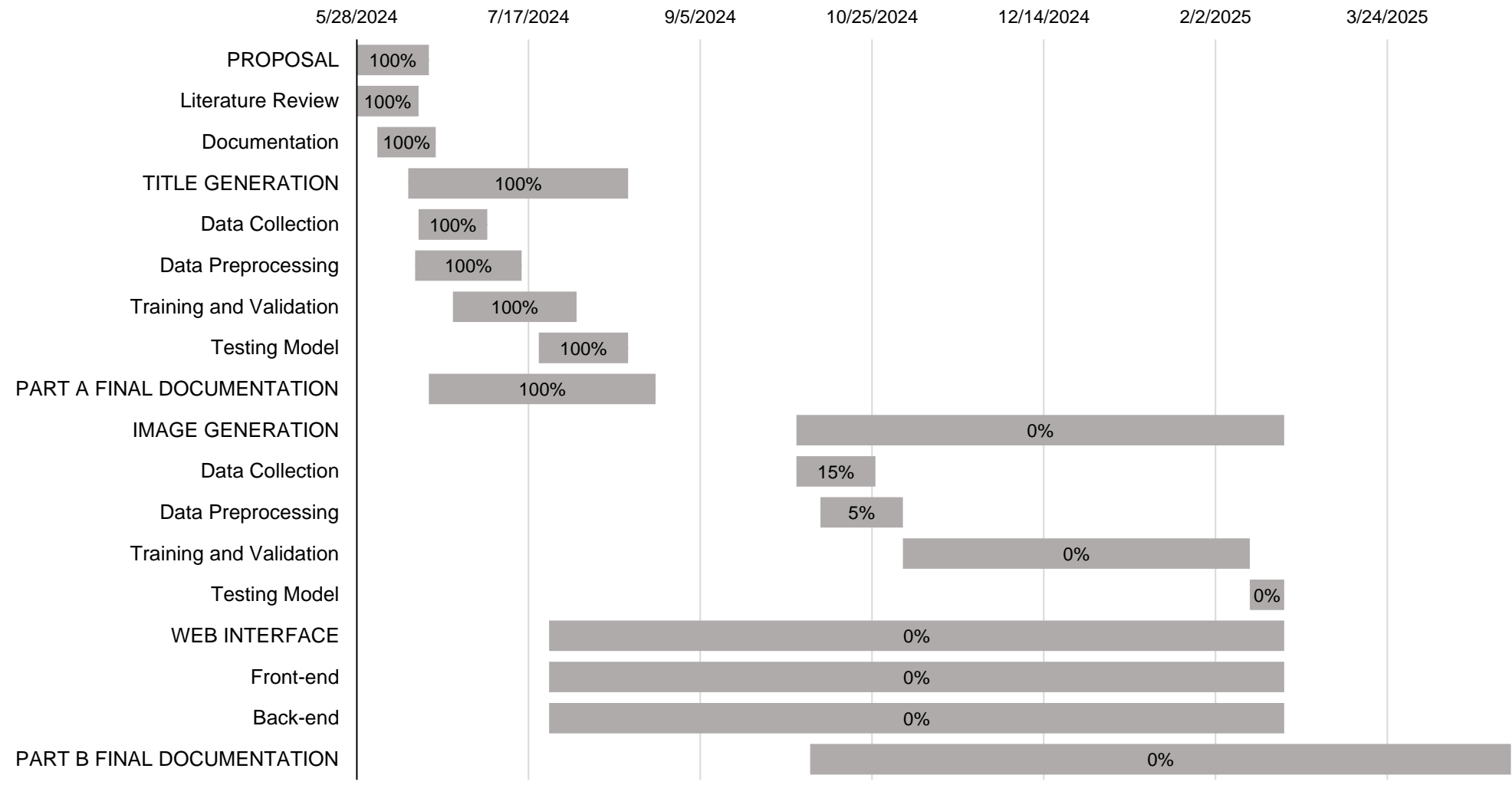
- ROUGE 1 score of 0.49 representing effective individual word matches

- ROUGE 2 score of 0.28 representing struggle preserve word pair sequences

- ROUGE L score of 0.46 indicating balanced performance for longer sequences

Remaining Tasks

- More image data collection and preprocessing

- Finetuning LDM

- Text styling and integration

- Creating a user-friendly interface

Tentative Timeline

Task	Completion
PROPOSAL	100%
Literature Review	100%
Documentation	100%
TITLE GENERATION	100%
Data Collection	100%
Data Preprocessing	100%
Training and Validation	100%
Testing Model	100%
PART A FINAL DOCUMENTATION	100%
IMAGE GENERATION	0%
Data Collection	15%
Data Preprocessing	5%
Training and Validation	0%
Testing Model	0%
WEB INTERFACE	0%
Front-end	0%
Back-end	0%
PART B FINAL DOCUMENTATION	0%

Timeline dates: 5/28/2024, 7/17/2024, 9/5/2024, 10/25/2024, 12/14/2024, 2/2/2025, 3/24/2025

References

[1] A. Fan et al., "Beyond english-centric multilingual machine translation," Journal of Machine Learning Research, vol. 22, no. 107, pp. 1–48, 2021.

[2] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 10684–10695.

References

[1] A. Fan et al., "Beyond english-centric multilingual machine translation," Journal of Machine Learning Research, vol. 22, no. 107, pp. 1–48, 2021.

[2] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 10684–10695.