Text Summarization Using Encoder Decoder Generative Adversarial Networks

A Project Report submitted by

Tathagata Mookherjee

in partial fulfillment of the requirements for the award of the degree of

M.Tech. in Data and Computational Sciences

Indian Institute of Technology

Department: School of Artificial Intelligence and Data Science

2023

Declaration

I hereby declare that the work presented in this Project Report titled *Text Summarization Using Encoder Decoder Generative Adversarial Networks* submitted to the Indian Institute of Technology Jodhpur and All India Institute of Medical Sciences Jodhpur in partial fulfilment of the requirements for the award of the degree of *M.Tech. in Data and Computational Sciences*, is a bonafide record of the research work carried out under the supervision of *Dr. Gaurav Harit, Associate Professor Department of Computer Science and Engineering.* The contents of this Project Report in full or in parts, have not been submitted to, and will not be submitted by me to, any other Instituteor University in India or abroad for the award of any degree or diploma.

Signature
Tathagata Mookherjee
M21AI619

Certificate

This is to certify that the Project Report titled *Text Summarization Using Encoder Decoder Generative Adversarial Networks*, submitted by *Tathagata Mookherjee(M21AI619)* to the Indian Institute of Technology Jodhpur and All India Institute of Medical Sciences Jodhpur for the award of the degree of *M.Tech. in Data and Computational Sciences*,is a bonafide record of the research work done by him under my supervision. To the best of my knowledge,the contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Gaurav Harit

Signature
Dr. Gaurav Harit

Abstract

In this project, we aim to develop a model called EDGAN, which will be an Encoder Decoder (ED) pretrained Generative Adversarial Network (GAN). EDGAN will be able to generate concise and coherent summaries of text paragraphs. EDGAN will be trained on a mid-sized dataset of movie plots, along with their corresponding summaries created by human editors.

The architecture of the EDGAN will consist of a trained Encoder (ENC) that will pretrain a generator network (GEN). The GEN will then be used to train a decoder network (DEC). GEN will take in a text paragraph input and generate an output vector, while DEC will take in the vector and validate if the vector is a reasonable representation of the original summary [2].

GEN and DEC will be trained alternatively, with GEN trying to improve its ability to generate vectors and DEC trying to generate authentic summaries. We will evaluate the performance of our EDGAN on the test set using metrics such as:

1. Bilingual Evaluation Understudy (BLEU) [8]
2. Metric for Evaluation of Translation with Explicit Ordering (METEOR) [9]
3. Word Error Rate (WER) [10]
4. Recall-Oriented Understudy for Gisting Evaluation 1-Gram (ROUGE-1) [7]
5. Longest Common Subsequence Recall-Oriented Understudy for Gisting Evaluation (ROUGE-L) [7]

Our goal is that EDGAN will be able to generate summaries that are concise and capture the key points and themes of the original stories and articles. We believe that this summarization system has the potential to significantly improve the efficiency of reading and processing written content.

DECLARATION II
CERTIFICATE III
ABSTRACT IV

1. *INTRODUCTION* 2
2. *OBJECTIVES* 2
3. *MANAGING THE DATA* 3
3.1. *DATA COLLECTION* 3
3.2. *DATA PREPARATION* 3
4. *BUILDING THE ENCODER DECODER (ED)* 6
4.1. *ED ARCHITECTING* 6
4.2. *ED TRAINING PARAMETERS* 6
4.3. *ED ARCHITECTURE DIAGRAMS* 7
4.4. *ED LOSSES* 9
4.5. *ED PERFORMANCE* 9
4.6. *ED OUTPUT* 9
5. *PRETRAINING THE GENERATOR (GEN)* 10
5.1. *GEN ARCHITECTING* 10
5.2. *GEN PRE-TRAINING PARAMETERS* 10
5.3. *GEN ARCHITECTURE DIAGRAMS* 11
5.4. *GEN LOSSES* 11
6. *GAN TRAINING WITH PRETRAINED GEN* 18
6.1. *GAN ARCHITECTING* 18
6.2. *GAN TRAINING PARAMETERS* 18
6.3. *GAN ARCHITECTURE DIAGRAMS* 19
6.4. *GAN LOSSES* 20
7. *GAN TRAINING WITH UNTRAINED GEN* 25
7.1. *GAN ARCHITECTING* 25
7.2. *GAN TRAINING PARAMETERS* 26
7.3. *GAN ARCHITECTURE DIAGRAMS* 26
7.4. *GAN LOSSES* 27
8. *COMPARISON OF PRETRAINED VS UNTRAINED GAN* 33
8.1. *GAN TRAINING* 33
8.2. *GAN VALIDATION* 34
8.3. *GAN PERFORMANCE* 35
8.4. *GAN OUTPUT* 39
9. *CONCLUSION* 40
10. *NOVELTY* 41
12. *CHALLENGES AND RISKS* 43
13. *RELEVANT REFERENCES* 44

1. *Introduction*

Generative adversarial networks (GANs) represent a class of deep learning architectures first proposed by Goodfellow et al. [1]. Comprising two neural entities, the generator and the discriminator, GANs are structured to engage in a duel-like learning process. The generator's primary task is to create data imitating real samples, whereas the discriminator's role is to discern between genuine and fabricated samples. As training progresses, the generator refines its data generation capabilities, and the discriminator enhances its detection proficiency, ensuring a continuous competition between the two networks.

Encoder-decoder (ED) frameworks emerged as a pivotal design for addressing sequence-to-sequence challenges within deep learning [5]. Primarily structured for undertakings like machine translation, abstractive summarization, and image captioning, these networks incorporate an encoder segment, which digests the input and encodes it into a standardized representation. Following this, a decoder segment interprets this representation to produce the anticipated output [4]. Such models are especially beneficial for tasks that necessitate the conversion of one sequence form to another, and their efficiency can be augmented through techniques like attention mechanisms to more precisely grasp the intricacies of the input [4].

Text summarization entails crafting a succinct and meaningful recapitulation of textual content. This entails pinpointing the principal ideas and motifs within the text and distilling them to encapsulate the core essence of the original material. Historically, the quest for effective text summarization traces back to the mid-20th century. However, the advent of contemporary machine learning methodologies, especially deep learning, has ushered in remarkable advancements in this domain [4][6]. Today, text summarization extends its utility across various applications, from condensing news articles and lengthy documents to summarizing dialogues.

Yao Lu introduced a model for abstractive text summarization leveraging GANs [2]. The synergy of EDs with GANs has notably enhanced their performance, particularly in spotting anomalies within images [3]. Recognizing the diverse utilities of GANs and the significance of text abstraction, there lies a potential avenue to engineer a proficient GAN optimized for text summarization via preliminary training with EDs. Employing such EDGANs for automated summary generation could potentially yield results surpassing the accuracy and efficiency of human-edited summaries.

2. *Objectives*

- Low validation loss for the ED and GAN
- Fidelity of synthetic data produced by the ED and GAN. This will be measured using metrics such as,
    - BLEU
    - METEOR
    - WER
    - ROUGE-1 – Precision, Recall and F1
    - ROUGE-L – Precision, Recall and F1
- Evaluating the similarity between the synthetic output and the real data using human judgment.

3. *Managing the Data*

3.1. Data collection

For the successful completion of this expansive project, the fundamental necessity encompassed the meticulous compilation of an extensive and multifaceted dataset encompassing intricate movie plots of considerable length, intricately interwoven with succinct yet comprehensive summaries of the very same narratives. The realization of this dataset was accomplished by harnessing the potent capabilities of the ChatGPT 3.5 APIs, which facilitated the extraction and curation of a diverse array of cinematic storylines and their corresponding abridged renditions. This methodological approach ensured the acquisition of a rich and varied corpus of textual content that formed the cornerstone of the project's analytical endeavors.

3.2. Data preparation

- Lowercasing all the words
- Stop words removal
- Punctuation and double space removal: ["~", "!", "@", "#", "$", "%", "^", "&", "*", "(", ")", "_", "+", "`", "-", "=", "[", "]", "\\", "{", "}", "|", ";", "'", ":", "'", ",", ".", "/", "<", ">", "?", " "]
- Contraction removal: {"ain't": "is not", "aren't": "are not", "can't": "cannot", "'cause": "because", "could've": "could have", "couldn't": "could not", "didn't": "did not", "doesn't": "does not", "don't": "do not", "hadn't": "had not", "hasn't": "has not", "haven't": "have not", "he'd": "he would", "he'll": "he will", "he's": "he is", "how'd": "how did", "how'd'y": "how do you", "how'll": "how will", "how's": "how is", "I'd": "I would", "I'd've": "I would have", "I'll": "I will", "I'll've": "I will have", "I'm": "I am", "I've": "I have", "i'd": "i would", "i'd've": "i would have", "i'll": "i will", "i'll've": "i will have", "i'm": "i am", "i've": "i have", "isn't": "is not", "it'd": "it would", "it'd've": "it would have", "it'll": "it will", "it'll've": "it will have", "it's": "it is", "let's": "let us", "ma'am": "madam", "mayn't": "may not", "might've": "might have", "mightn't": "might not", "mightn't've": "might not have", "must've": "must have", "mustn't": "must not", "mustn't've": "must not have", "needn't": "need not", "needn't've": "need not have", "o'clock": "of the clock", "oughtn't": "ought not", "oughtn't've": "ought not have", "shan't": "shall not", "sha'n't": "shall not", "shan't've": "shall not have", "she'd": "she would", "she'd've": "she would have", "she'll": "she will", "she'll've": "she will have", "she's": "she is", "should've": "should have", "shouldn't": "should not", "shouldn't've": "should not have", "so've": "so have", "so's": "so as", "this's": "this is", "that'd": "that would", "that'd've": "that would have", "that's": "that is", "there'd": "there would", "there'd've": "there would have", "there's": "there is", "here's": "here is", "they'd": "they would", "they'd've": "they would have", "they'll": "they will", "they'll've": "they will have", "they're": "they are", "they've": "they have", "to've": "to have", "wasn't": "was not", "we'd": "we would", "we'd've": "we would have", "we'll": "we will", "we'll've": "we will have", "we're": "we are", "we've": "we

have", "weren't": "were not", "what'll": "what will", "what'll've": "what will have", "what're": "what are", "what's": "what is", "what've": "what have", "when's": "when is", "when've": "when have", "where'd": "where did", "where's": "where is", "where've": "where have", "who'll": "who will", "who'll've": "who will have", "who's": "who is", "who've": "who have", "why's": "why is", "why've": "why have", "will've": "will have", "won't": "will not", "won't've": "will not have", "would've": "would have", "wouldn't": "would not", "wouldn't've": "would not have", "y'all": "you all", "y'all'd": "you all would", "y'all'd've": "you all would have", "y'all're": "you all are", "y'all've": "you all have", "you'd": "you would", "you'd've": "you would have", "you'll": "you will", "you'll've": "you will have", "you're": "you are", "you've": "you have"}

- Splitting the data into training and test sets using 70/30 ratio
- Word/Sentence tokenization using NLTK
- Tokenize the sentences

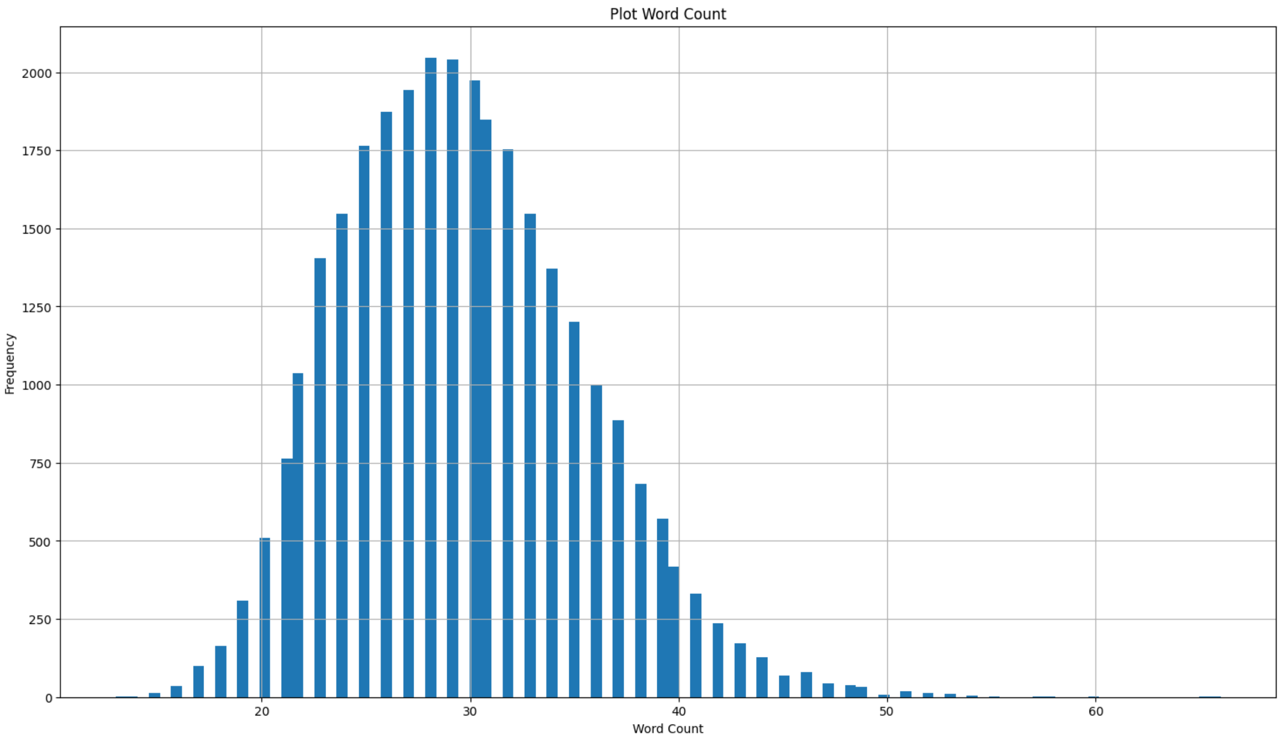


Figure 3.2.1:Plot word counts

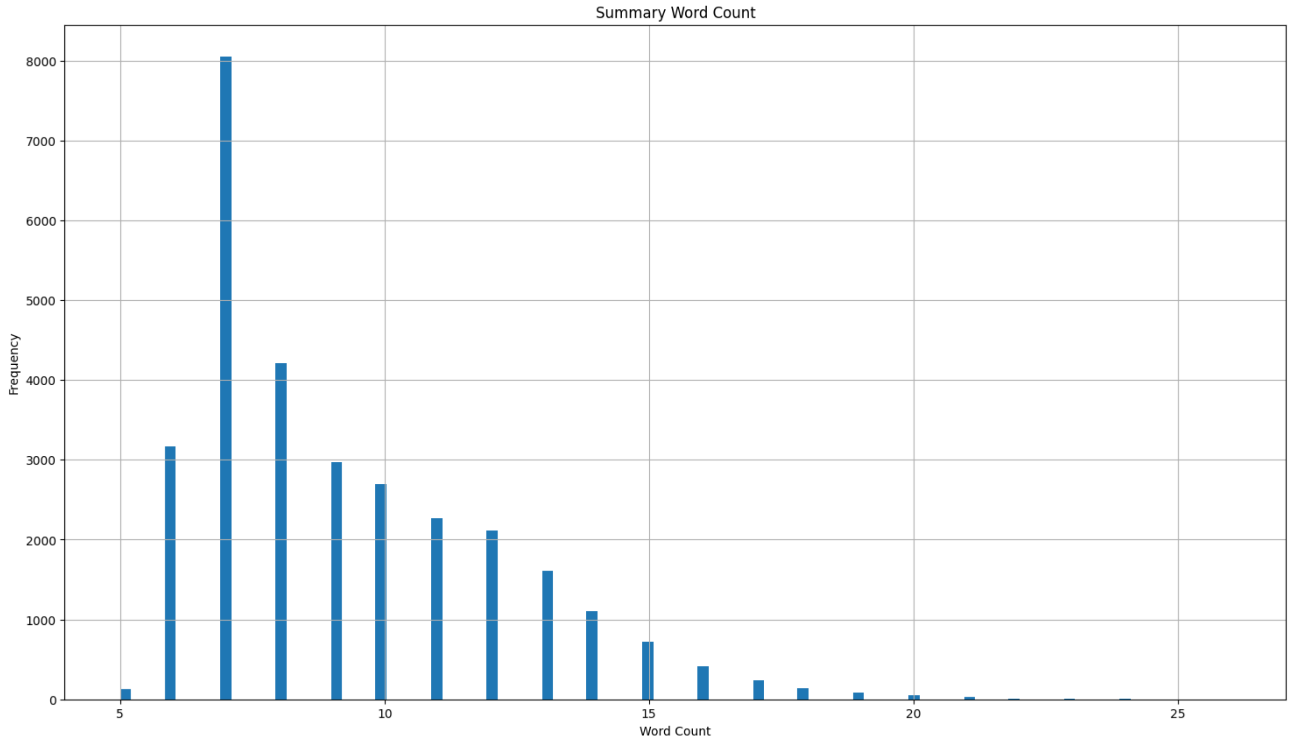


Figure 3.2.2: Summary Word Count

Title	Summary	Plot	Cleaned_Plot	Cleaned_Summary
better times	Silent comedy drama: woman chooses happiness.	Better Times is a 1919 American silent comedy drama film directed by Marshall Neilan. It follows a woman who is forced to choose between her faithless husband and her devoted suitor, and ultimately decides to pursue her own happiness.	better times 1919 american silent comedy drama film directed marshall neilan follows woman forced choose faithless husband devoted suitor ultimately decides pursue happiness	_START_ silent comedy drama woman chooses happiness _END_
dungeonmaster the	Fantasy world, battle powerful force, aided by wizard, guided by Dungeonmaster, return to own world.	Dungeonmaster The is a 1985 fantasy film directed by Rosemarie Turko. It follows a group of people who are transported to a fantasy world where they must battle a powerful force in order to return to their own world. The group is aided by a powerful wizard and the Dungeonmaster, an enigmatic figure who guides them in their quest.	dungeonmaster 1985 fantasy film directed rosemarie turko follows group people transported fantasy world must battle powerful force order return world group aided powerful wizard dungeonmaster enigmatic figure guides quest	_START_ fantasy world battle powerful force aided wizard guided dungeonmaster return world _END_
log kya kahenge	Woman fights for true love.	Log Kya Kahenge is a 1983 Indian Bollywood movie that follows the story of a young woman who is estranged from her family because of her love for a lower class man. Despite	log kya kahenge 1983 indian bollywood movie follows story young woman estranged family love lower class man despite facing opposition	_START_ woman fights true love _END_

		facing opposition from her family and society, she fights to protect their relationship and prove her love is true.	family society fights protect relationship prove love true	
the racket	Crime captain takes on mobster.	The Racket is a 1951 film noir crime drama directed by John Cromwell. It follows Tom McQuigg, a police captain determined to take down a powerful crime boss. McQuigg sets up a plan to arrest the boss, but it backfires and leads to a series of events that will test McQuigg's courage and integrity.	racket 1951 film noir crime drama directed john cromwell follows tom mcquigg police captain determined take powerful crime boss mcquigg sets plan arrest boss backfires leads series events test mcquigg courage integrity	_START_ crime captain takes mobster _END_
first blood	Rambo fights for survival against sheriff.	The movie "First Blood" released in 1982 follows John Rambo, a Vietnam veteran, who is forced to face the harsh reality of his past when a small town sheriff tries to arrest him. He must fight for his survival against the sheriff and the forces of the National Guard.	movie first blood released 1982 follows john rambo vietnam veteran forced face harsh reality past small town sheriff tries arrest must fight survival sheriff forces national guard	_START_ rambo fights survival sheriff _END_

Table 3.3.3: Sample Cleaned Input Data to the ED

4. Building the Encoder Decoder (ED)

4.1. ED architecting

Central to our research objectives is the strategic development of a sophisticated encoder-decoder (ED) framework, reminiscent of sequence-to-sequence models introduced in seminal works on neural architectures [5]. The bedrock of our methodological approach leans heavily on enhancing the encoder's (ENC) capability to meticulously generate a detailed context vector, a strategy rooted in the principles of abstractive summarization [4]. Concurrently, the decoder (DEC) is subjected to rigorous training paradigms, demanding the meticulous deconstruction and reassembly of textual summaries, a complexity inspired by advancements in deep learning-driven text summarization [2]. This intricate dance between ENC and DEC, fortified by consistent training and methodological refinements, underscores our ambition to blend linguistic expertise with cutting-edge computational techniques [6].

4.2. ED Training Parameters

- Encoder
    - Plot vocabulary size = 29315 (from dataset)
    - Dropout = [0.3, 0.5, 0.7]
    - Modal plot word length = 28 (from dataset)

- Decoder
    - Summary vocabulary size = 12877 (from dataset)
    - Modal summary word length = 7 (from dataset)
- Bidirectional LSTM layers(s) = [1, 2, 3]
- Embedding dimension = [250, 500, 750]
- Latent dimension = Embedding dimension X 2
- Learning rate = [0.01, 0.001, 0.0001]
- Epsilon = [1e-03, 1e-05, 1e-07]
- Loss = Sparse Categorical Cross-entropy
- Optimizer = Adam
- Activation = Tanh
- Batch size = [15, 30, 45, 60]
- Epochs = [Batch size, Batch size X 2]
- Train and validate the AE

4.3. ED Architecture Diagrams

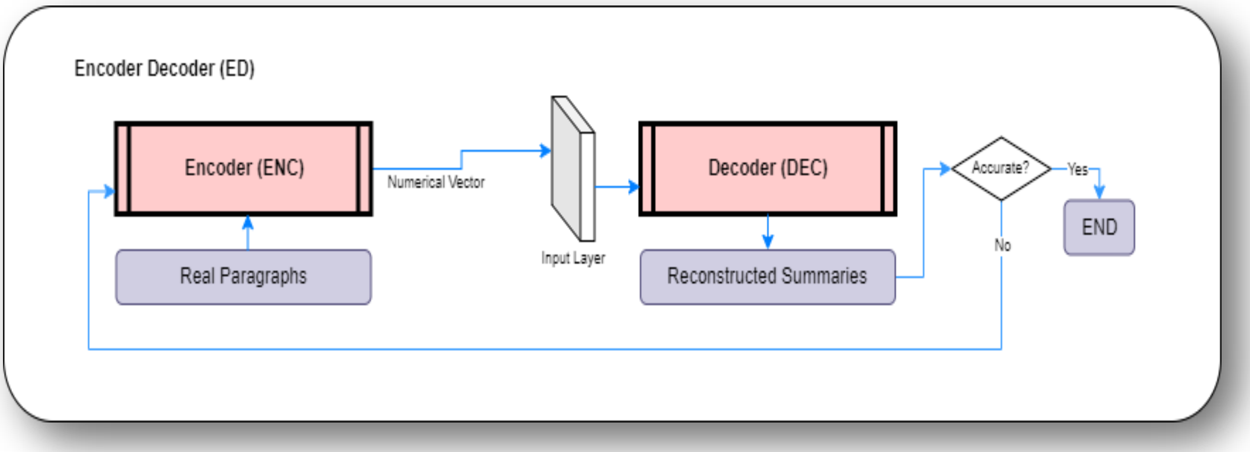


Figure 4.3.1: ED architecture

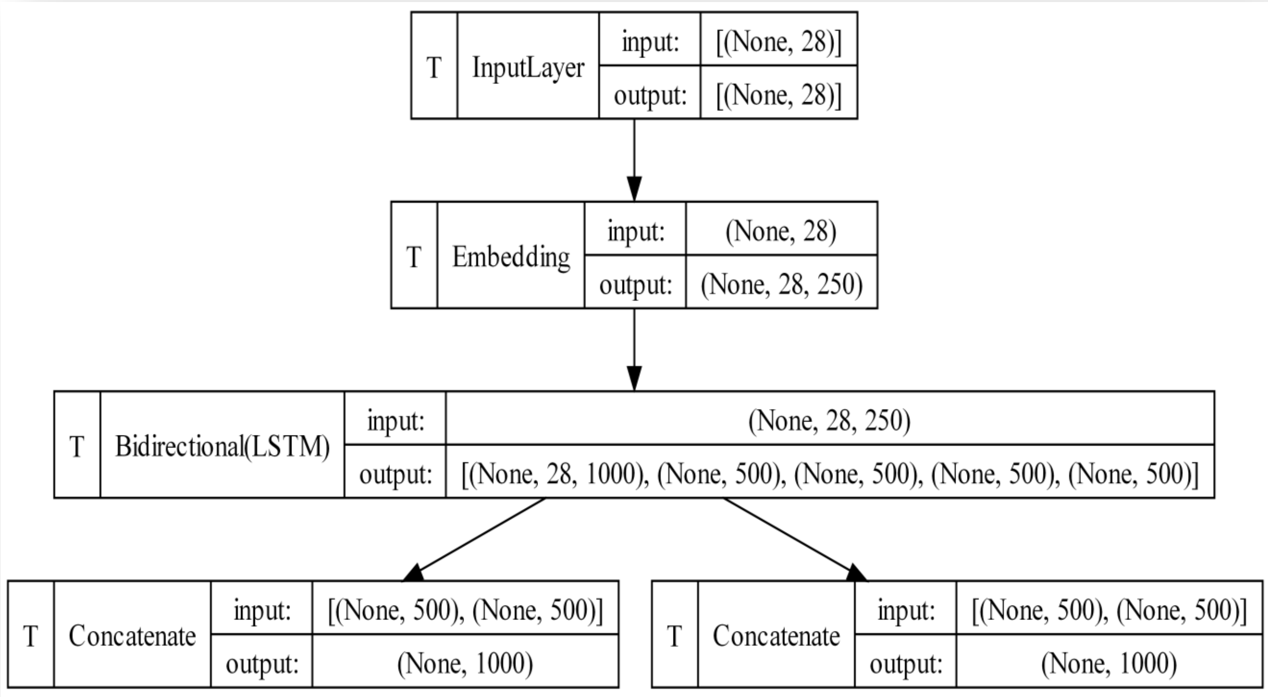


Figure 4.3.2: ENC architecture

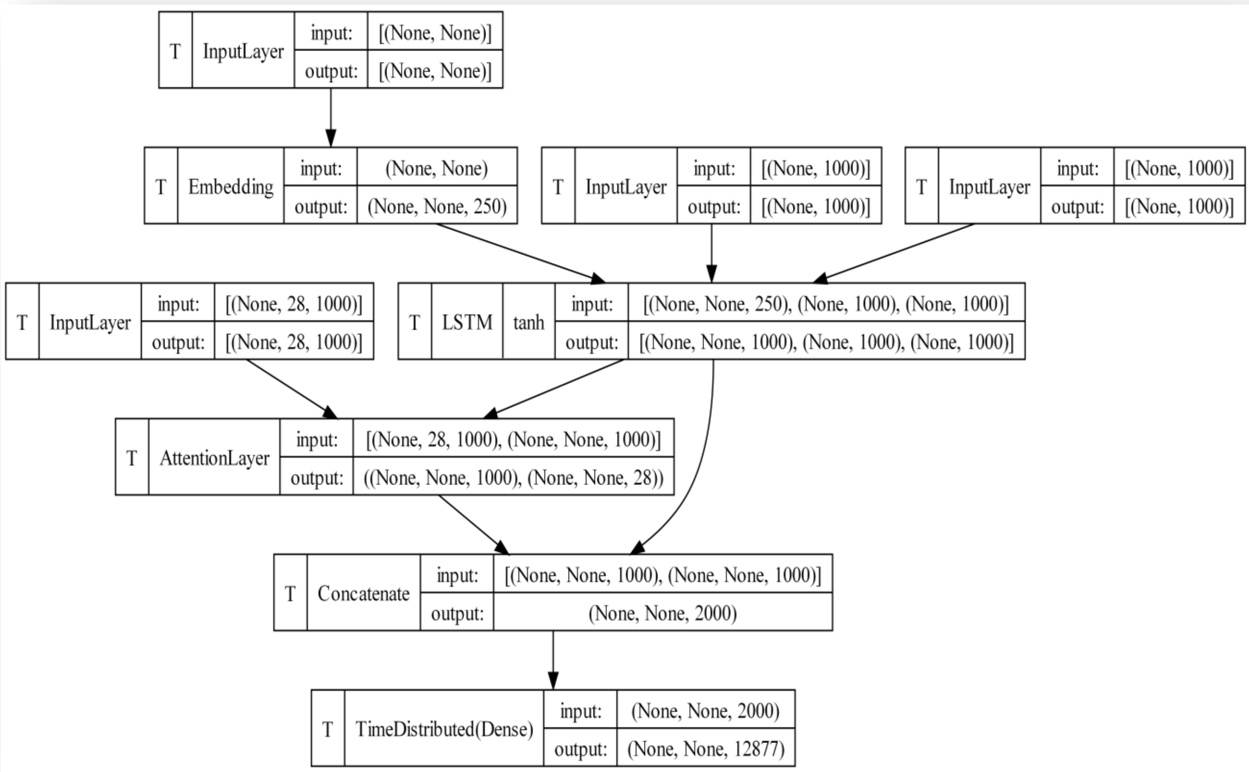


Figure 4.3.3: DEC architecture

4.4. ED Losses

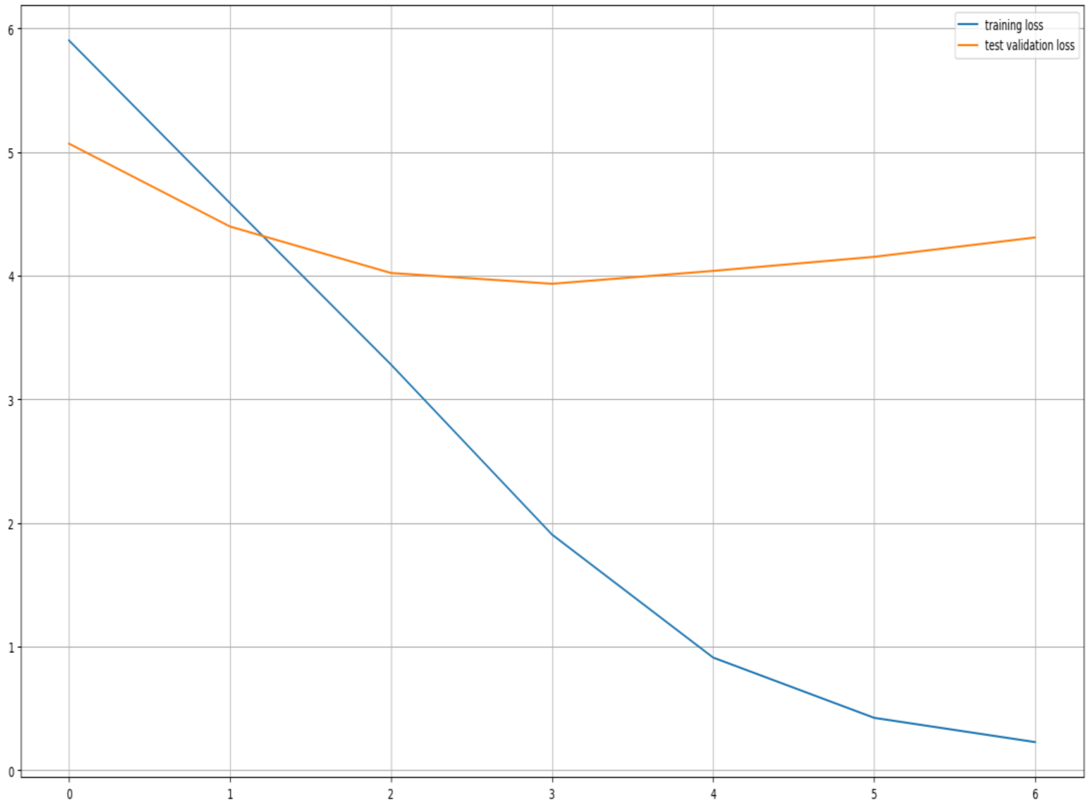


Figure 4.4.1: Overall ED training and validation

4.5. ED Performance

- Average BLEU score: 0.12783402849202055
- Average WER score: 0.8121043379921574
- Average METEOR score: 0.21730916614721613
- Average ROUGE-1 precision: 0.30612285036443937
- Average ROUGE-1 recall: 0.2742617089613519
- Average ROUGE-1 f-measure: 0.28748736481889914
- Average ROUGE-L precision: 0.2953018401740259
- Average ROUGE-L recall: 0.2652395176457664
- Average ROUGE-L f-measure: 0.27770761606359745

4.6. ED Output

Plot	Original_summary	Predicted_summary
better times 1919 american silent comedy drama film directed marshall neilan follows woman forced choose faithless husband devoted suitor ultimately decides pursue happiness	silent comedy drama woman chooses happiness	woman chooses happiness happiness

1985 fantasy film directed rosemarie turko follows group people transported fantasy world must battle powerful force order return world group aided powerful wizard dungeonmaster enigmatic figure guides quest	aided wizard guided dungeonmaster return world	wizard rider dungeonmaster evil governor
log kya kahenge 1983 indian bollywood movie follows story young woman estranged family love lower class man despite facing opposition family society fights protect relationship prove love true	woman fights true love	woman fights true love
crime drama directed john cromwell follows tom mcquigg police captain determined take powerful crime boss mcquigg sets plan arrest boss backfires leads series events test mcquigg courage integrity	crime captain takes mobster	crime corruption takes control
movie first blood released 1982 follows john rambo vietnam veteran forced face harsh reality past small town sheriff tries arrest must fight survival sheriff forces national guard	rambo fights survival sheriff	rambo fights mob sheriff

Table 4.6.1: Sample output data from the ED

5. *Pretraining the Generator (GEN)*

5.1. GEN architecting

In the next stages of our research, influenced by the foundational work on Generative Adversarial Networks by Goodfellow et al. [1], we plan to incorporate a generator module (termed as GEN). This GEN will undergo a pretraining regimen using insights derived from sequence-to-sequence frameworks, particularly the encoder-decoder (ENC-DEC) paradigms [5]. This intensive pretraining process will involve channeling the context vectors produced by ENC and the summarized outputs curated by DEC into GEN. Drawing inspiration from the successful applications of abstractive text summarization using GANs [2], our intention is to hone GEN's proficiency in autonomously crafting text. By capitalizing on the representational knowledge imbibed from the ENC-DEC model, we anticipate that GEN will bridge the chasm between preliminary input and expected output, achieving a synergy of linguistic precision and computational robustness [6].

5.2. GEN Pre-training Parameters

- Vocabulary size = 22333 (from dataset)
- Embedding dimension = [250, 500, 750]
- Latent dimension = [Embedding dimension, Embedding dimension X 2]
- LSTM layers(s) = [1, 2, 3]
- Batch size = [15, 30, 45, 60]
- Epochs = [Batch size, Batch size X 2]

- Predicted summary word length = 5 (from dataset)
- Dropout = [0.3, 0.5, 0.7]
- Learning rate = [0.01, 0.001, 0.0001]
- Loss = Cross Entropy
- Text length = 6 (from dataset)
- Optimizer = Adam
- Activation = Sigmoid
- Early Stopping = 2

5.3. GEN Architecture Diagrams

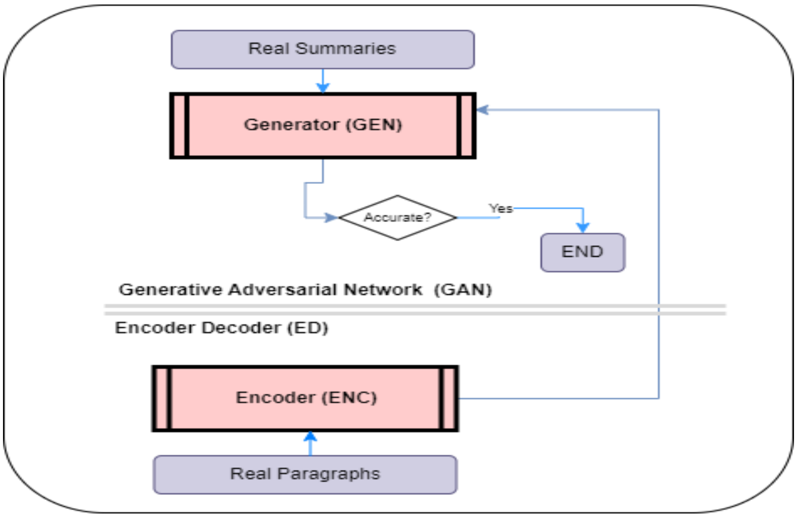


Figure 5.3.1: GEN pretraining architecture

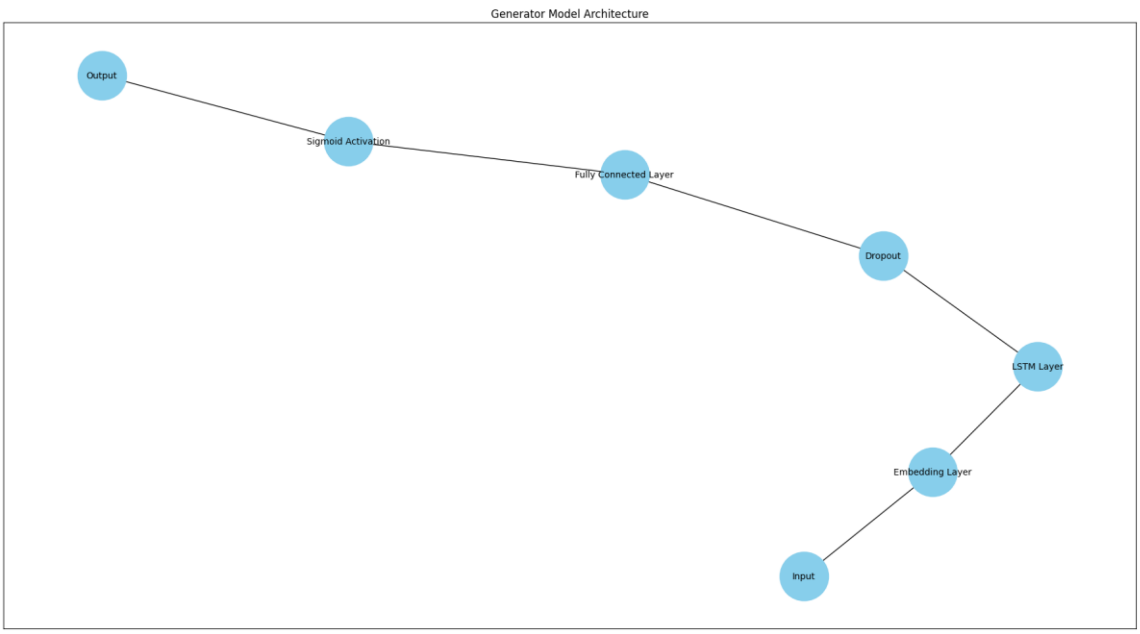


Figure 5.3.2: GEN model

5.4. GEN Losses

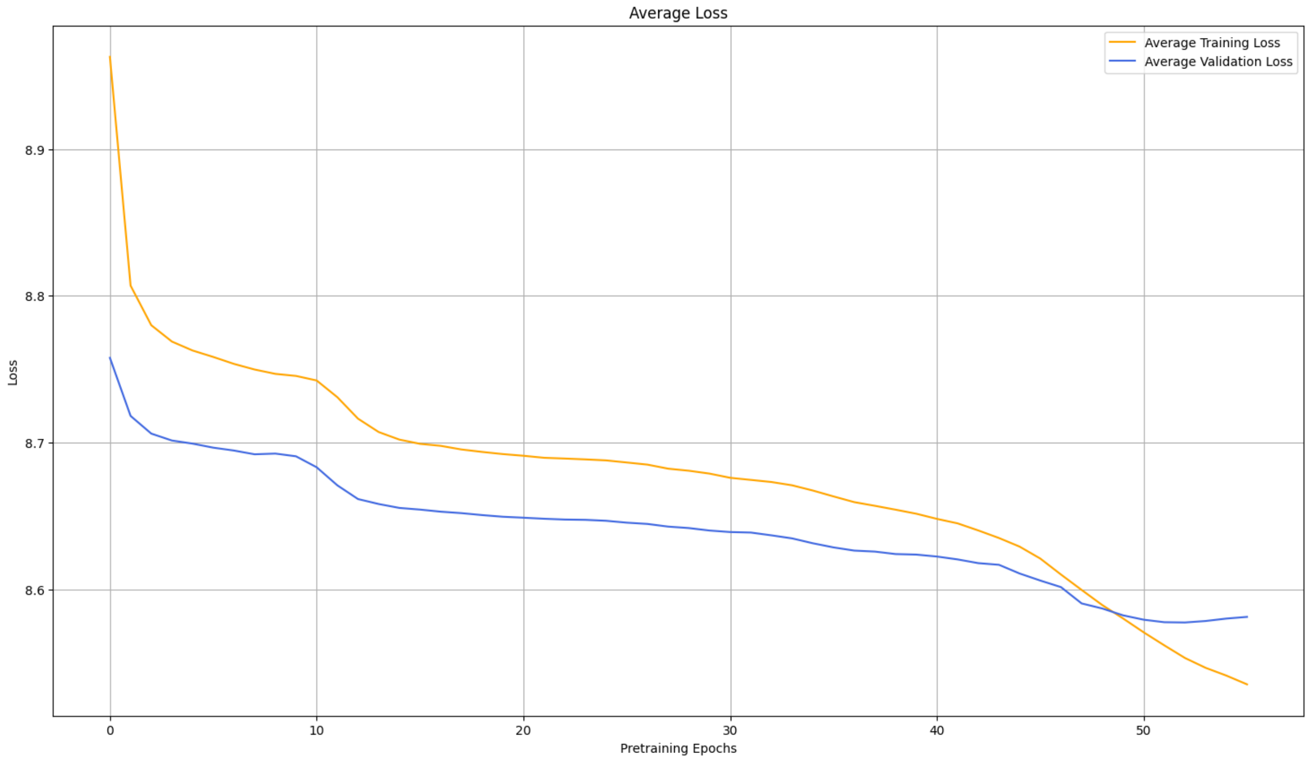


Figure 5.4.1: Average Loss for 5K dataset

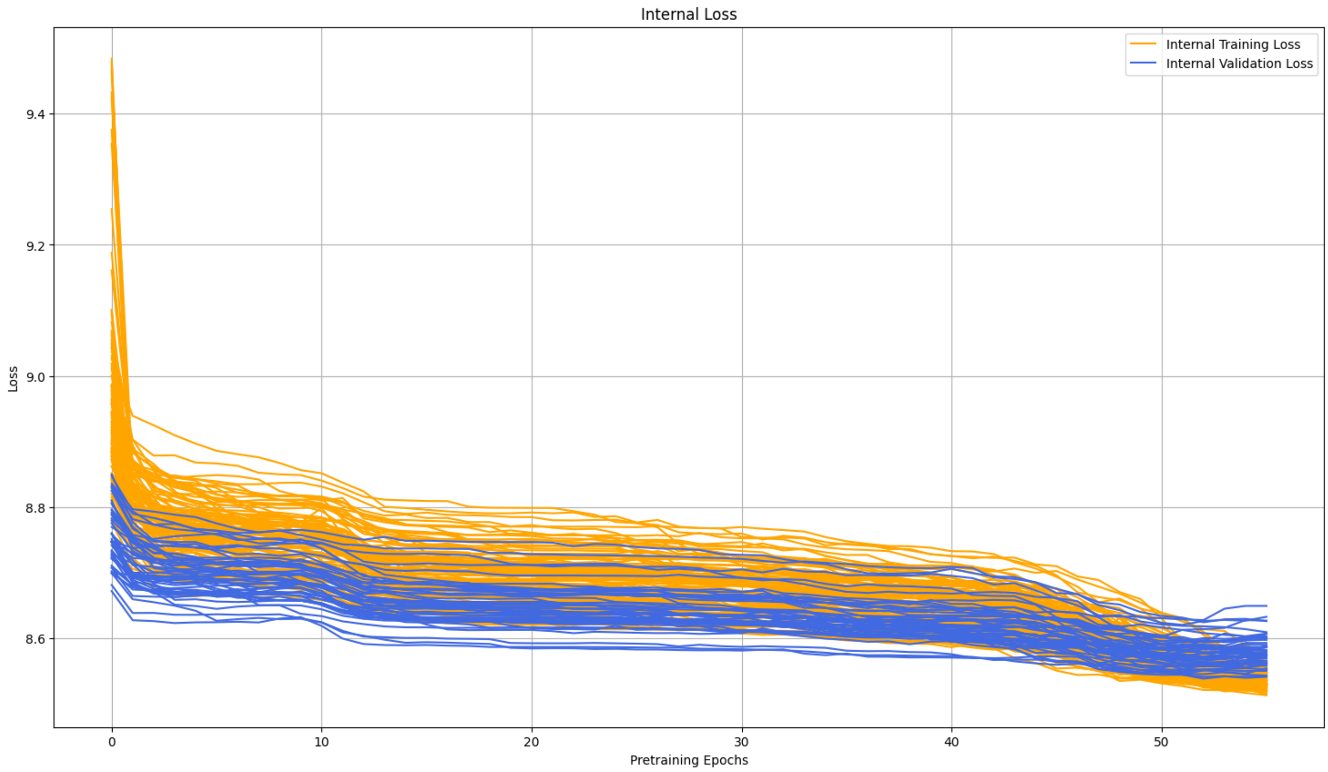


Figure 5.4.2: Internal Loss for 5K dataset

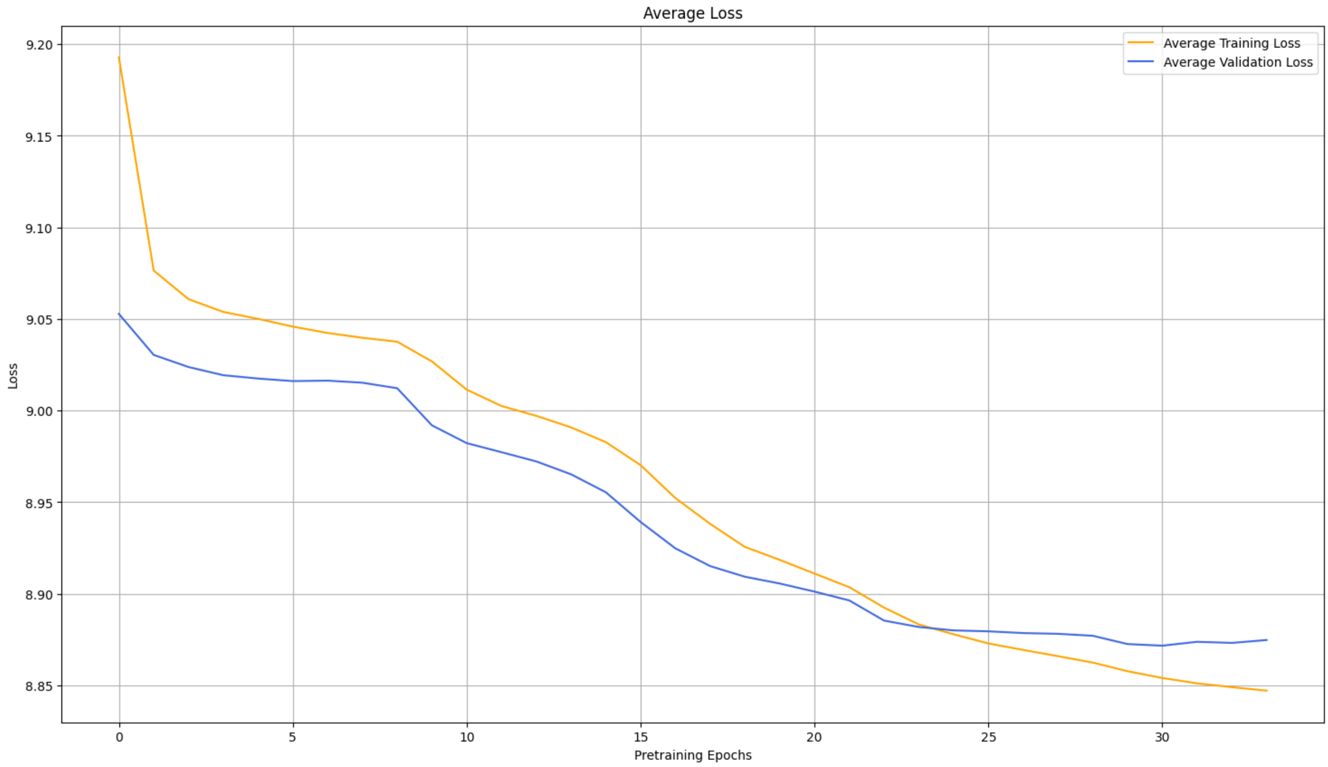


Figure 5.4.3: Average Loss for 10K dataset

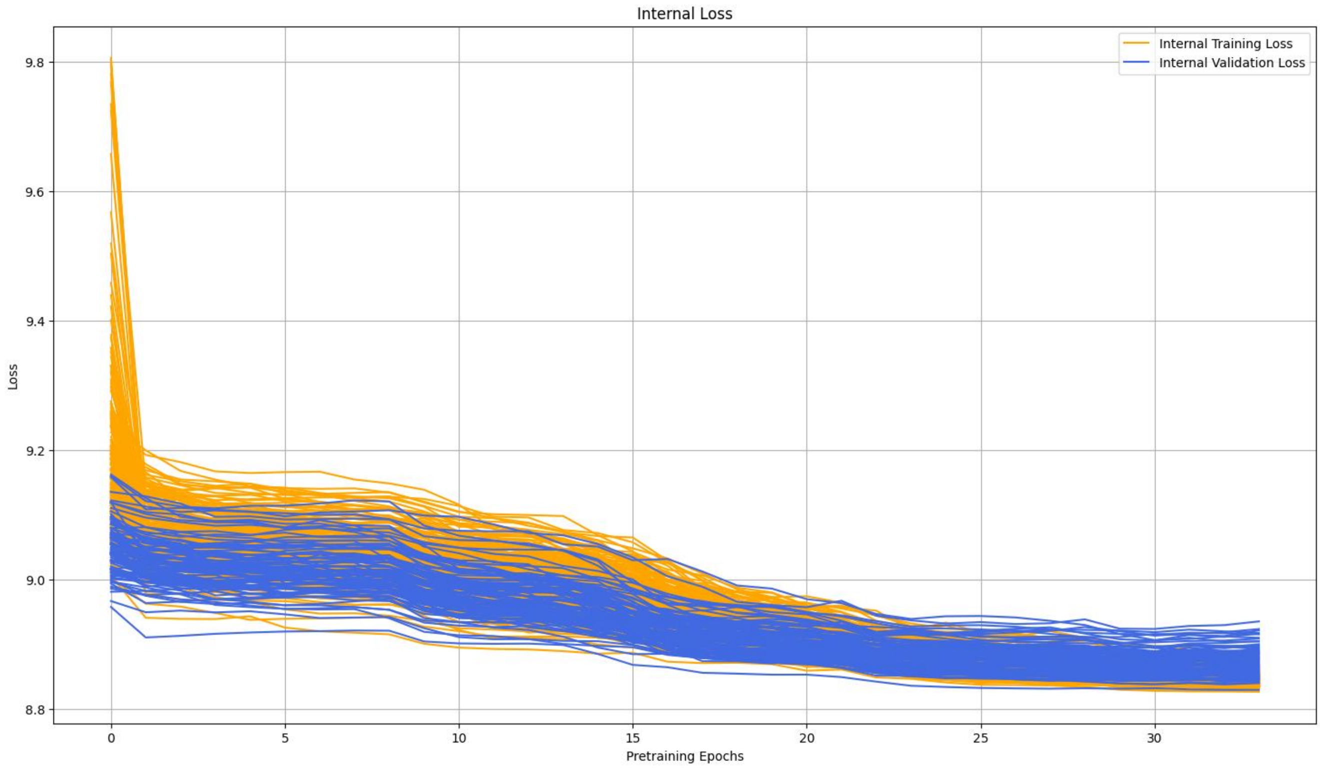


Figure 5.4.4: Internal Loss for 10K dataset

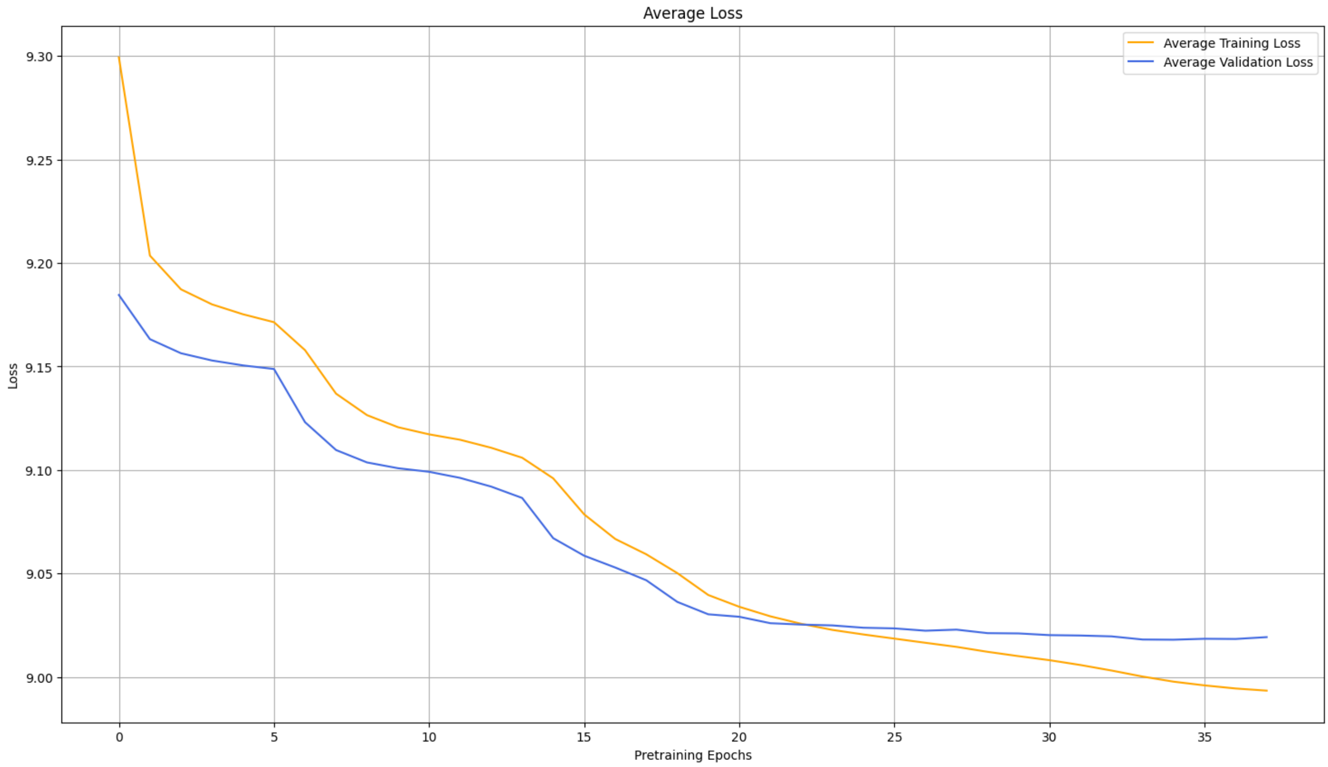


Figure 5.4.5: Average Loss for 15K dataset

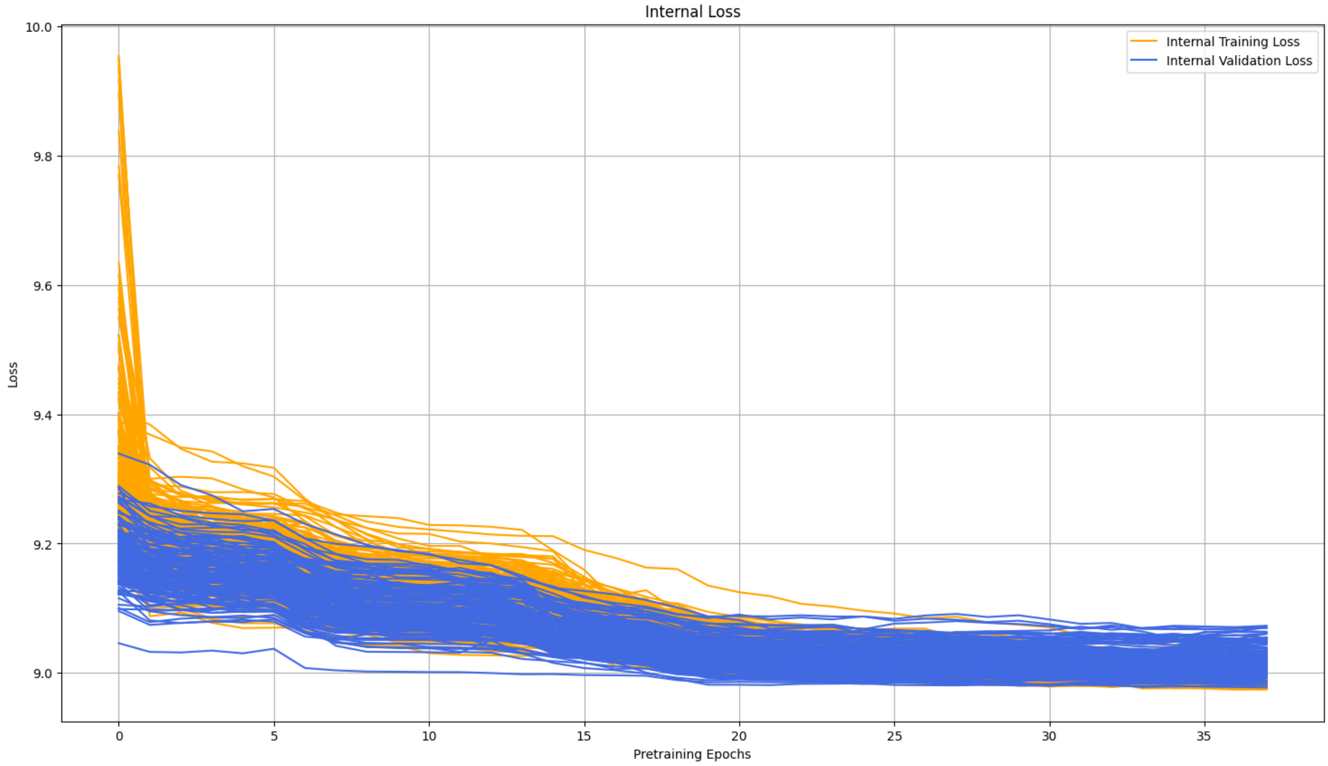


Figure 5.4.6: Internal Loss for 15K dataset

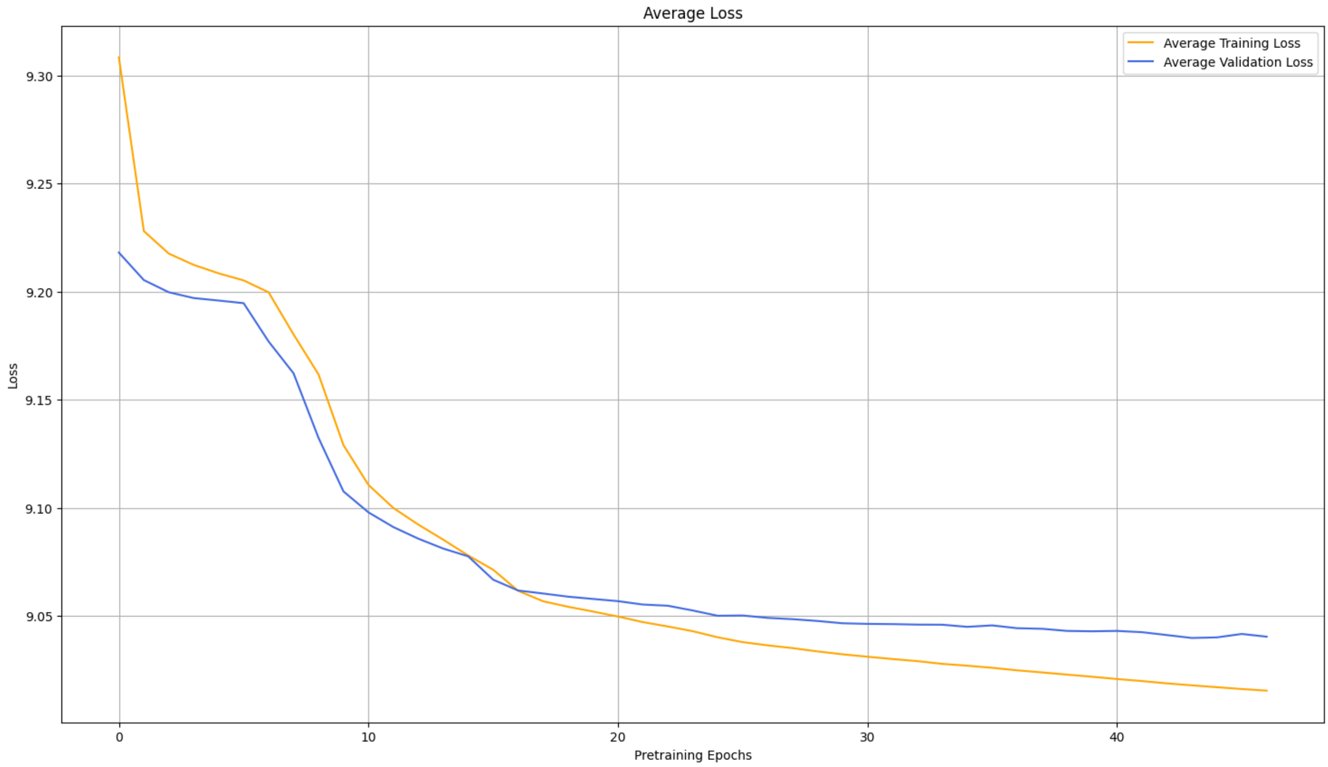


Figure 5.4.7: Average Loss for 20K dataset

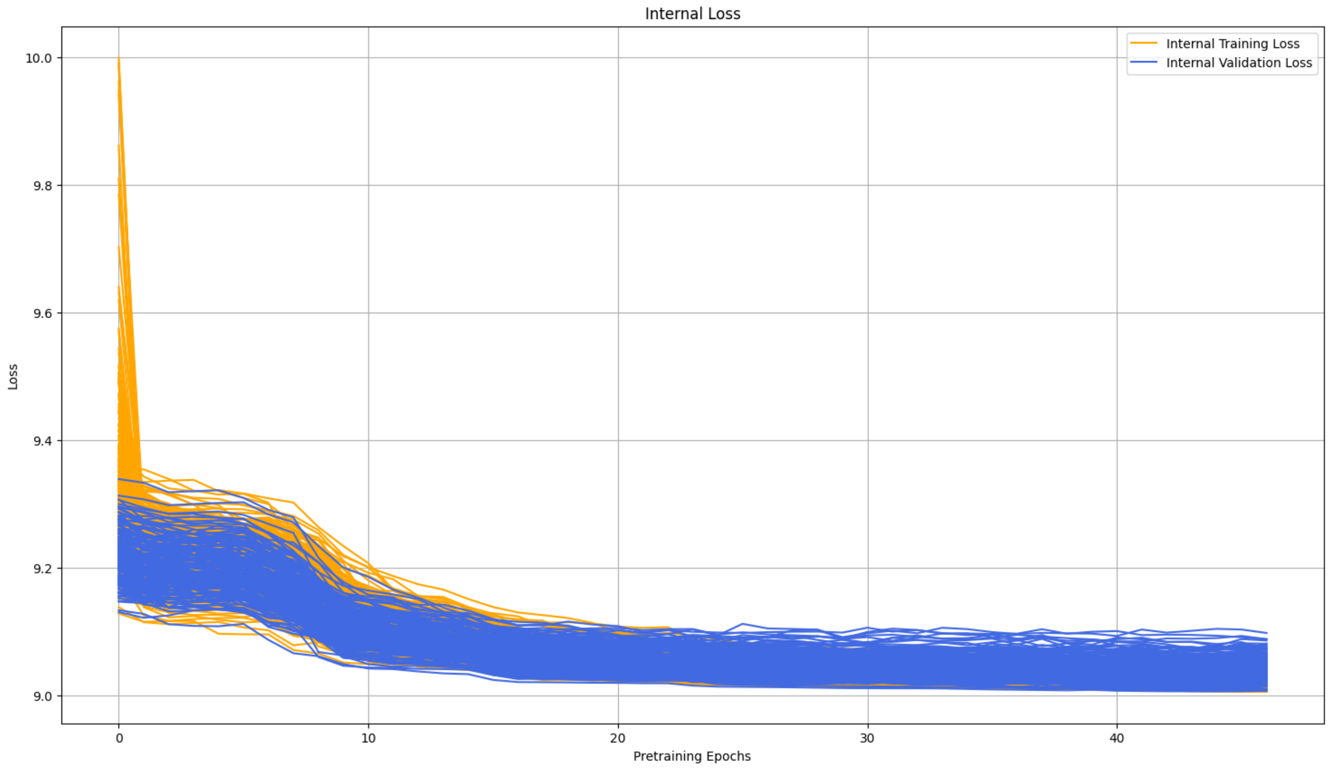


Figure 5.4.8: Internal Loss for 20K dataset

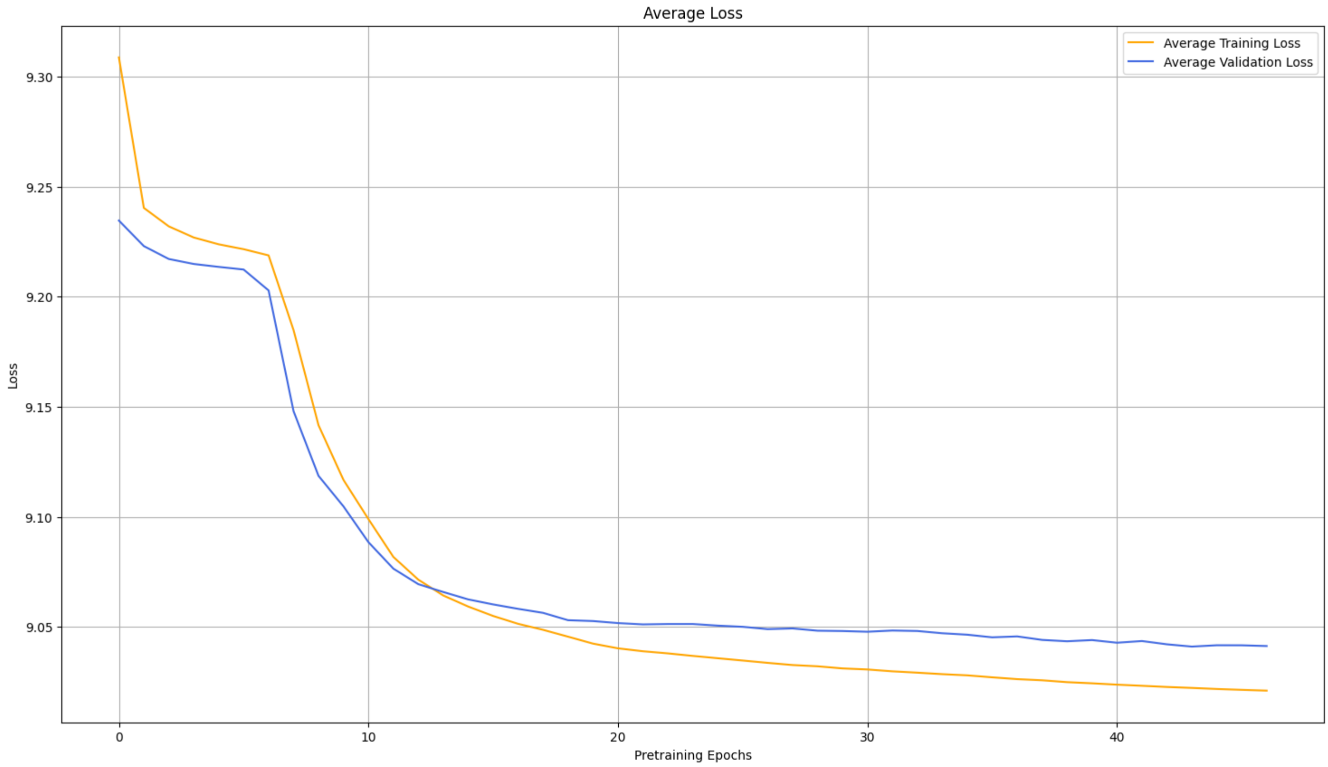


Figure 5.4.9: Average Loss for 25K dataset

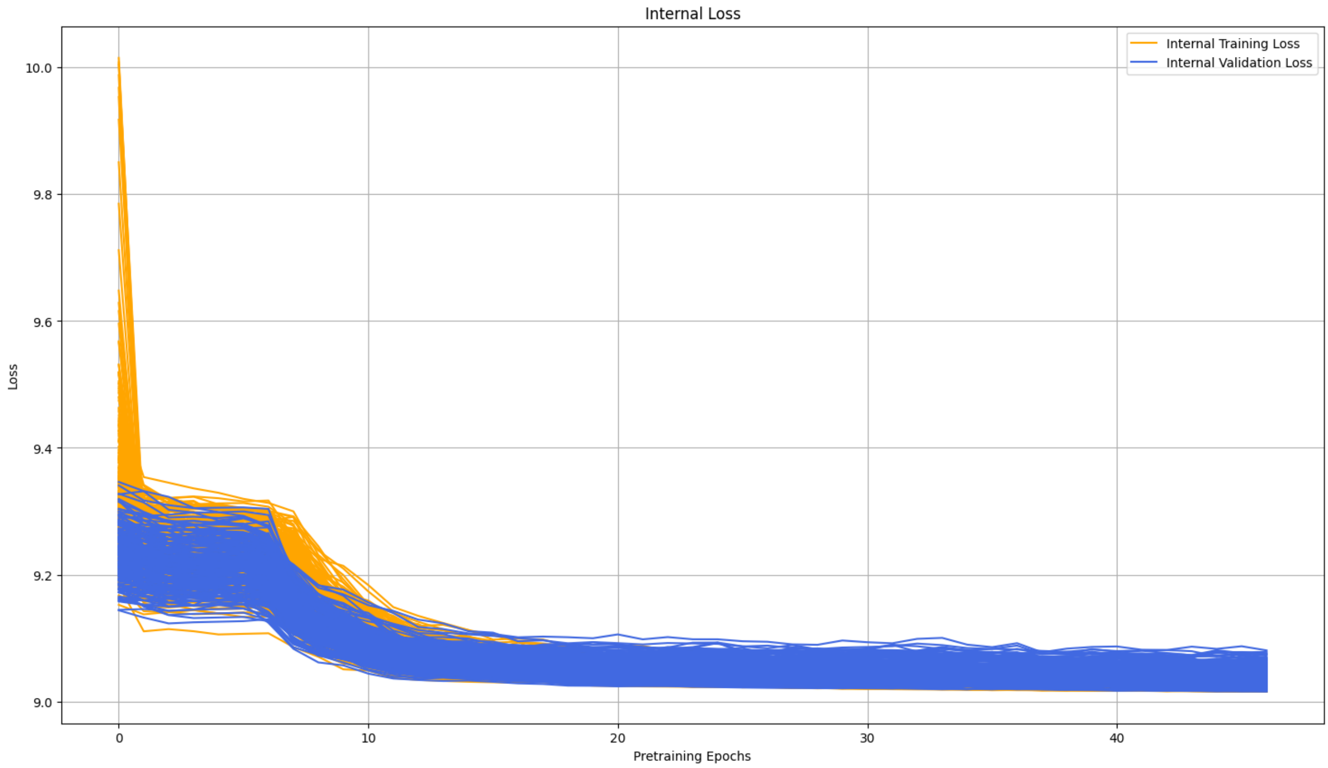


Figure 5.4.10: Internal Loss for 25K dataset

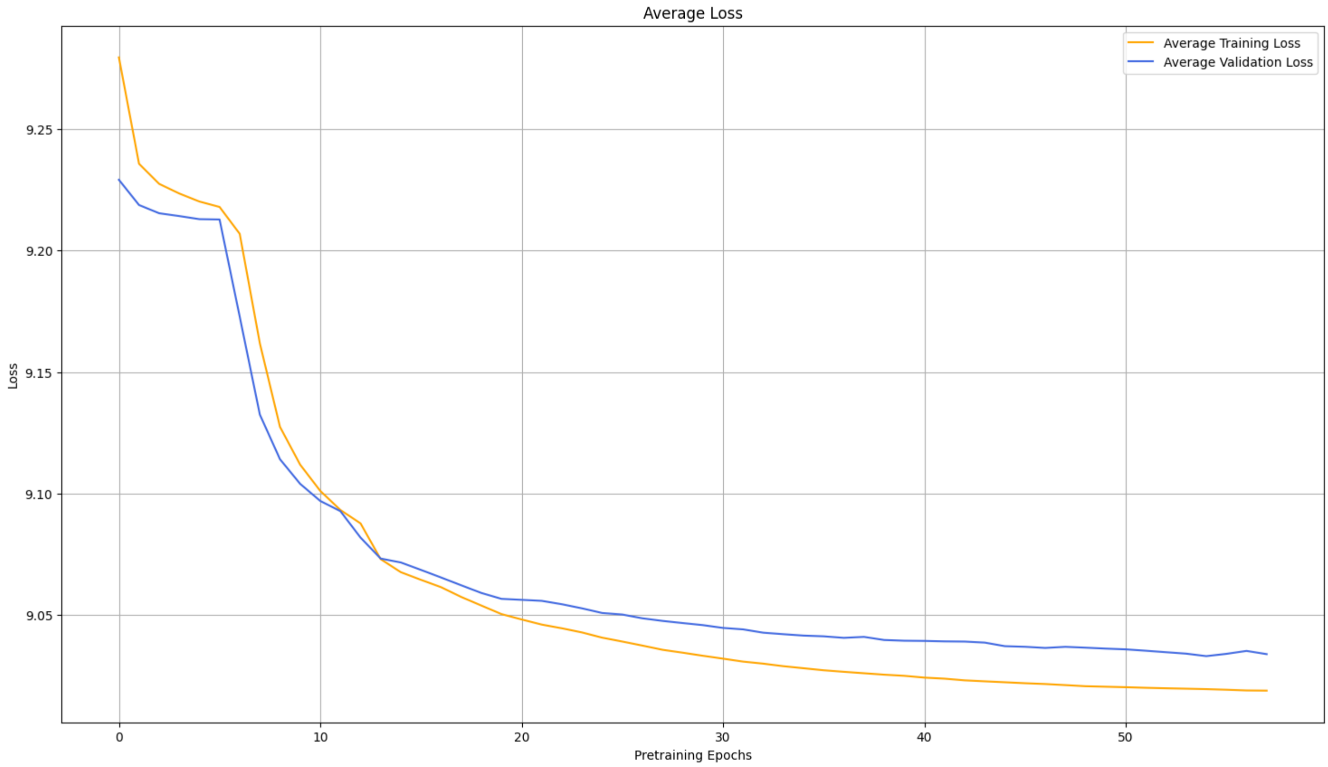


Figure 5.4.11: Average Loss for 30K dataset

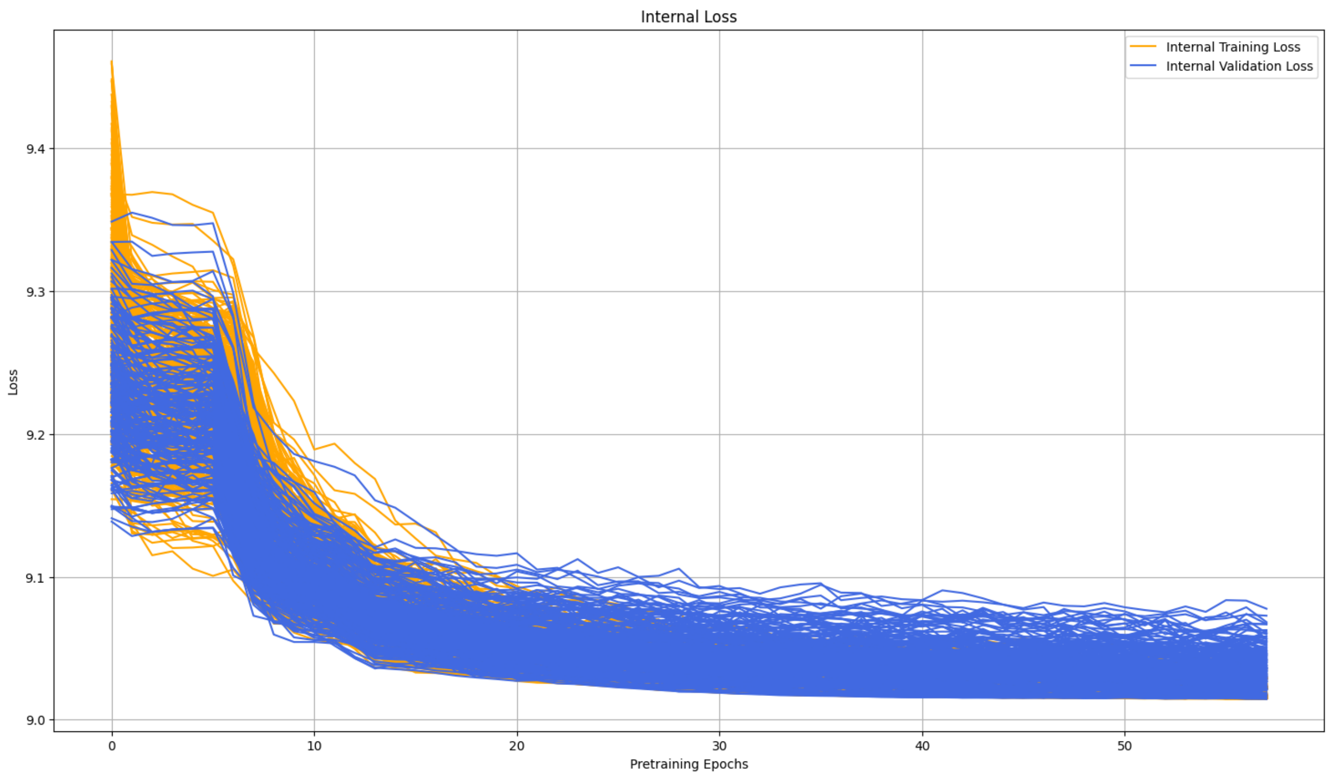


Figure 5.4.12: Internal Loss for 30K dataset

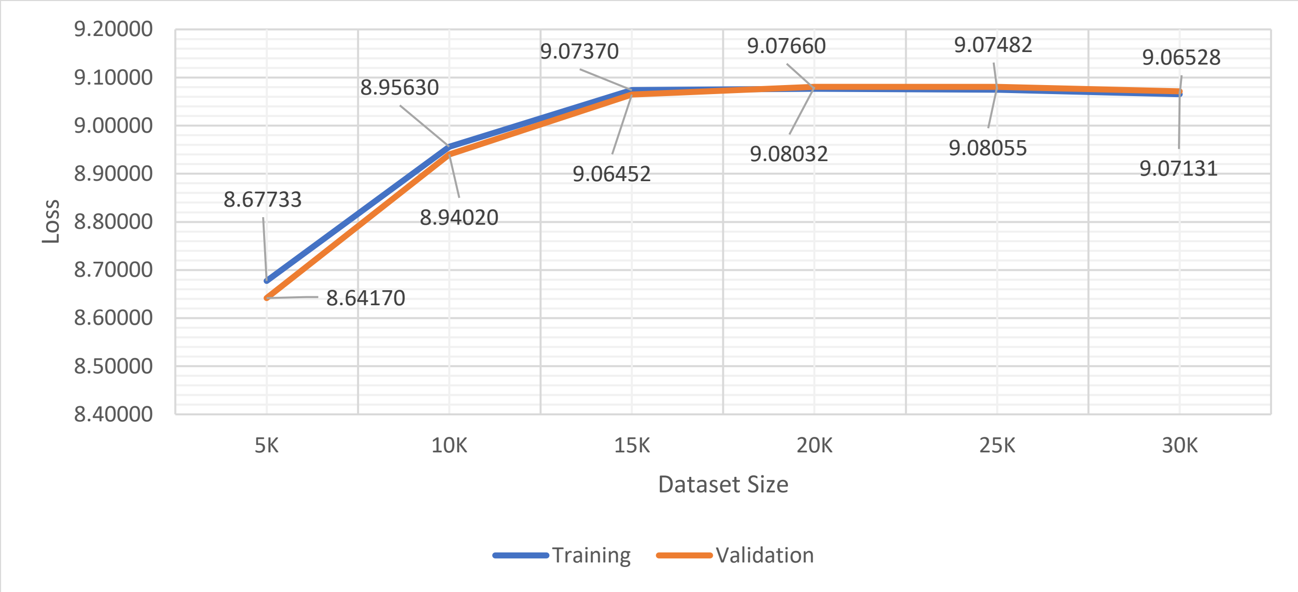


Figure 5.4.13: Average GEN pretraining losses

6. GAN training with pretrained GEN

6.1. GAN architecting

Our research ambitions are anchored in the strategic implementation of the pretrained generator (GEN) within the overarching architecture of Generative Adversarial Networks, a paradigm pioneered by Goodfellow et al. [1]. The GEN, having undergone rigorous pretraining, has acquired the proficiency to understand and utilize the compact and intricate representations derived from the encoder's output. Within the GAN setup, GEN is not an isolated entity but works in tandem with a discriminator, which we denote as DIS. This collaborative training process between GEN and DIS is an iterative dance: while GEN continually refines its ability to craft accurate and coherent text summaries, DIS evaluates and discerns the genuineness of these generated summaries in relation to real ones. Through this dynamic, we seek to foster an environment where GEN progressively improves its output, resulting in summaries that are both contextually relevant and linguistically fluent.

6.2. GAN Training Parameters

- Vocabulary size = 22333 (from dataset)
- Embedding dimension = [250, 500, 750]
- Latent dimension = [Embedding dimension, Embedding dimension X 2]
- LSTM layers(s) = [1, 2, 3]
- Batch size = [15, 30, 45, 60]
- Epochs = [Batch size, Batch size X 2]
- Predicted summary word length = 5 (from dataset)
- Dropout = [0.3, 0.5, 0.7]
- Learning rate = [0.01, 0.001, 0.0001]

- Loss = Cross Entropy
- Optimizer = Adam
- Activation = Sigmoid
- Early Stopping = 2

6.3. GAN Architecture Diagrams

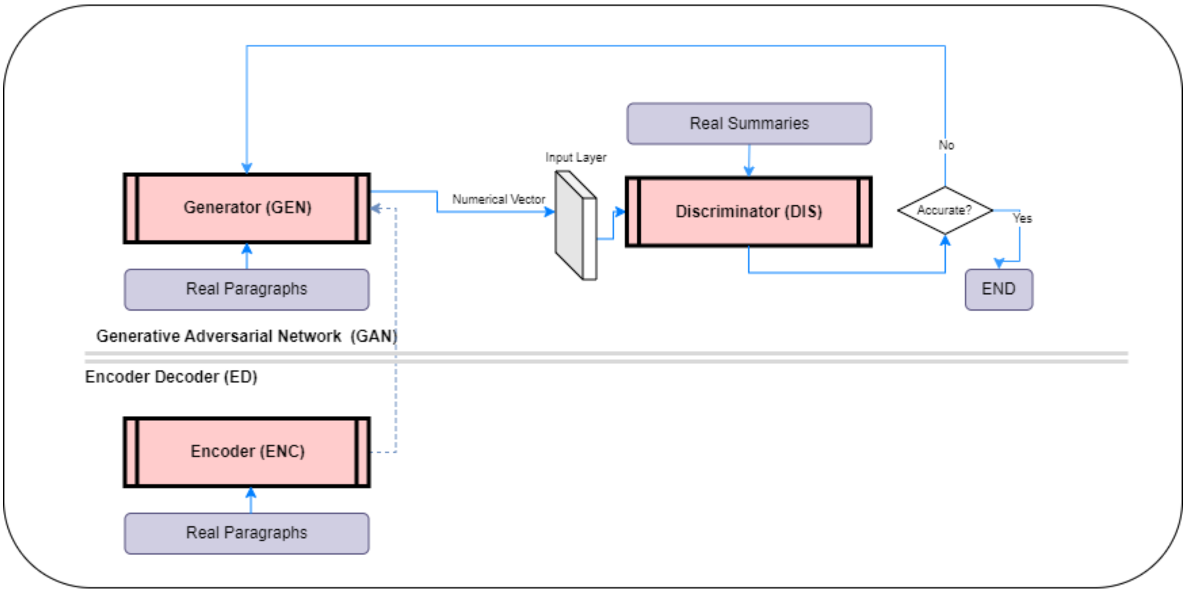


Figure 6.3.1: GAN architecture

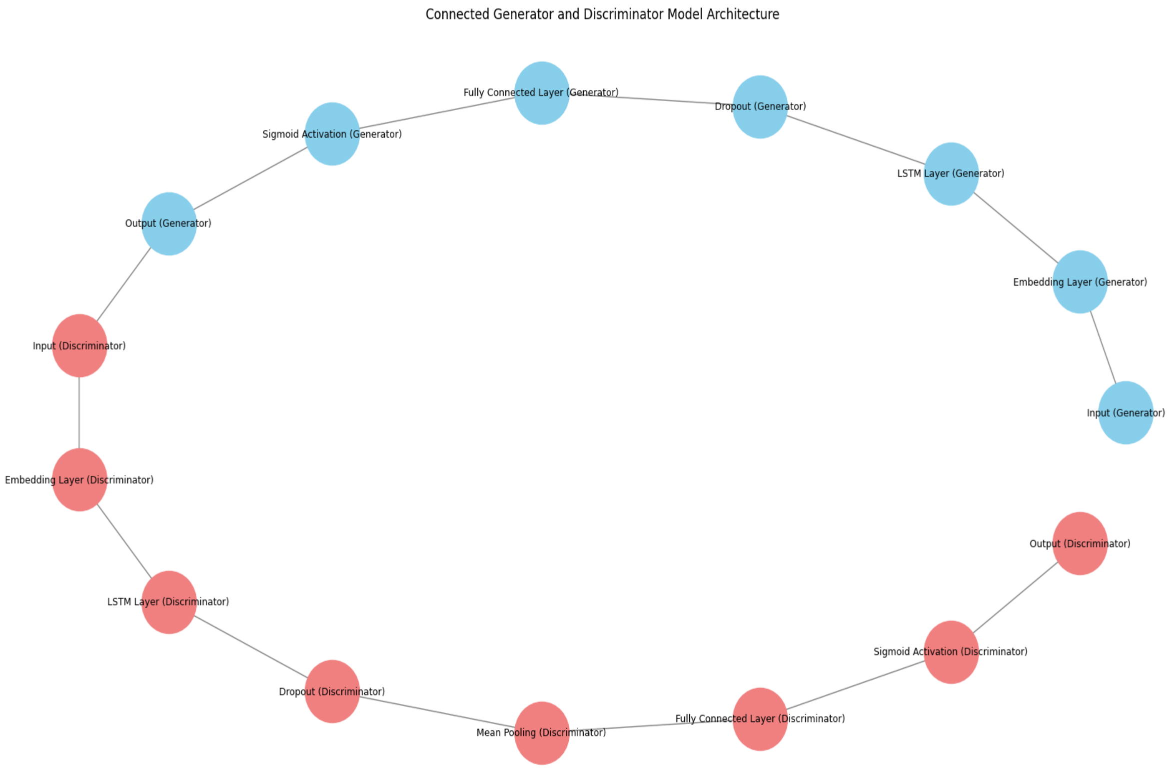


Figure 6.3.2: GAN model

6.4. GAN Losses

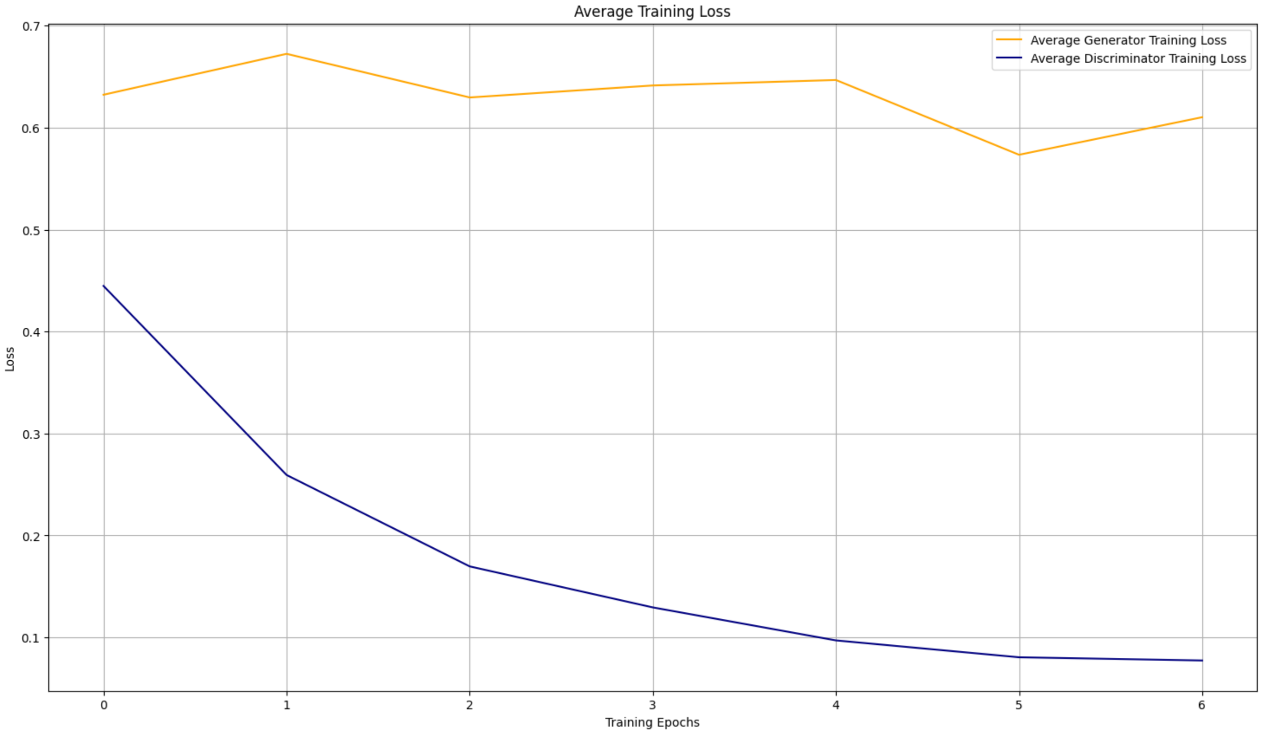


Figure 6.4.1: Average Training Loss for 5K dataset

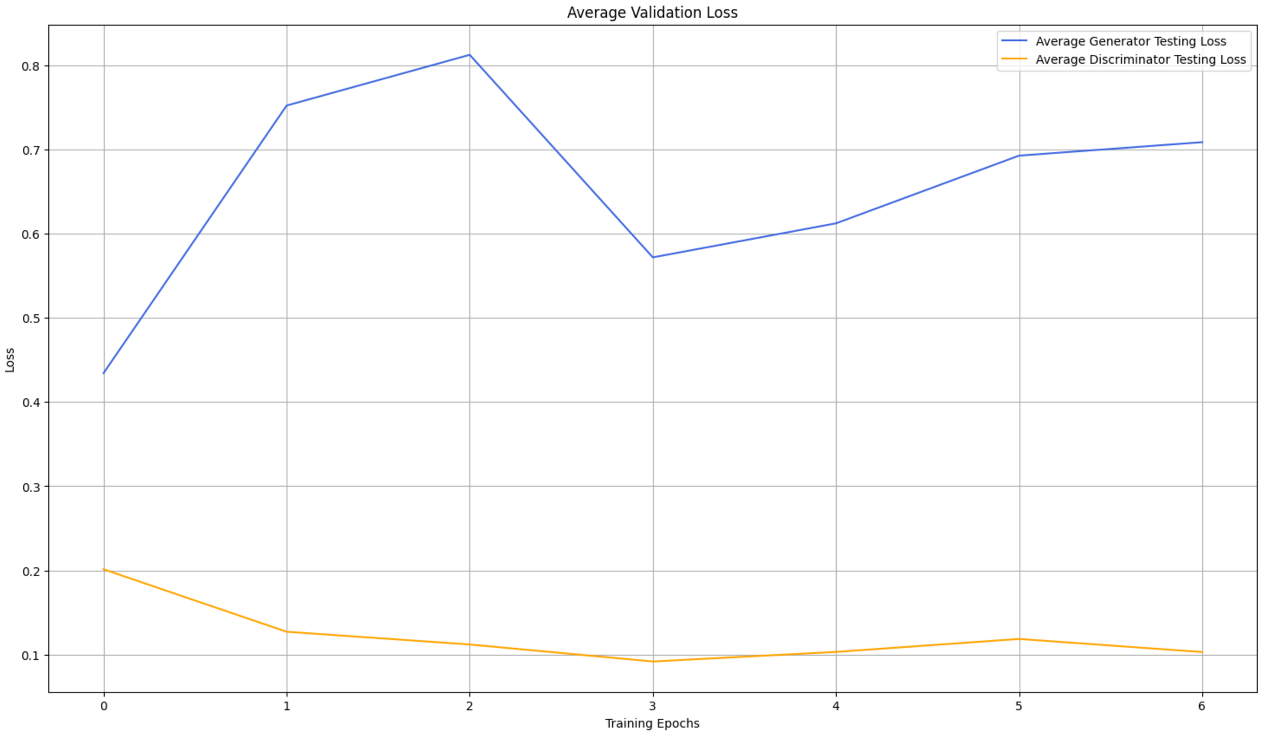


Figure 6.4.2: Average Testing Loss for 5K dataset

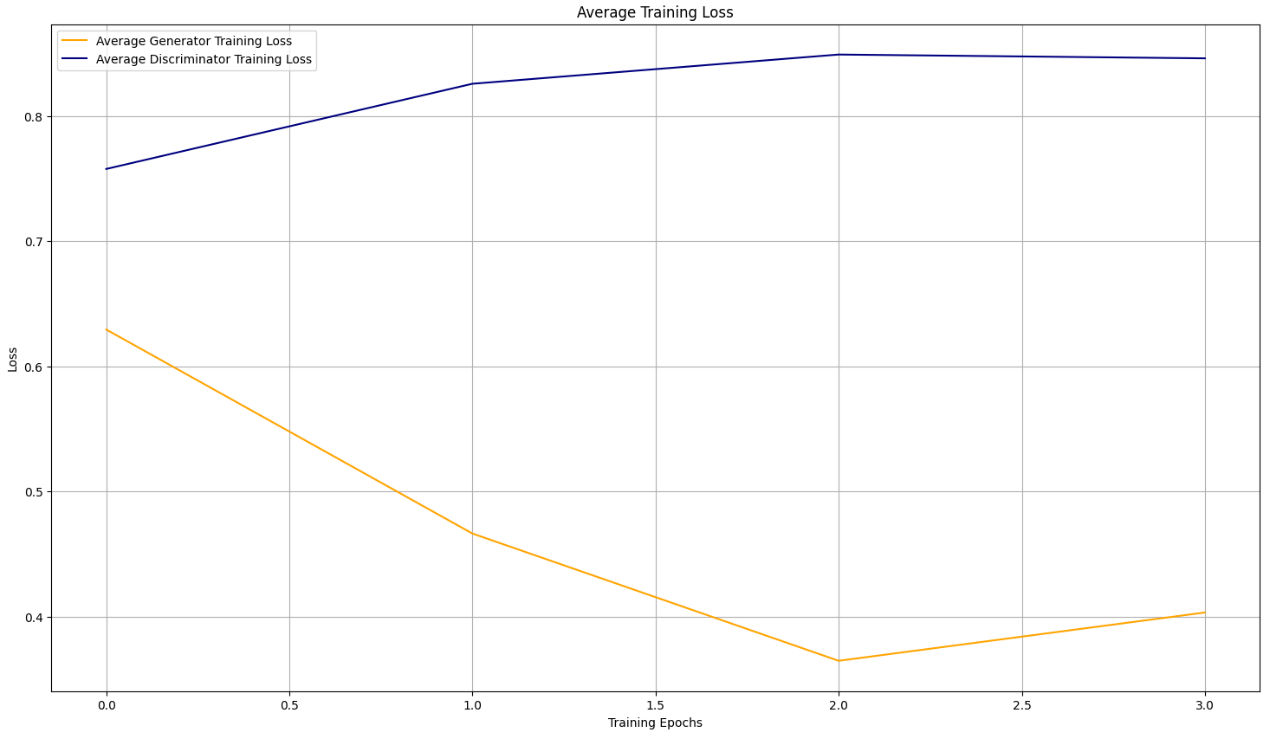


Figure 6.4.3: Average Training for 10K dataset

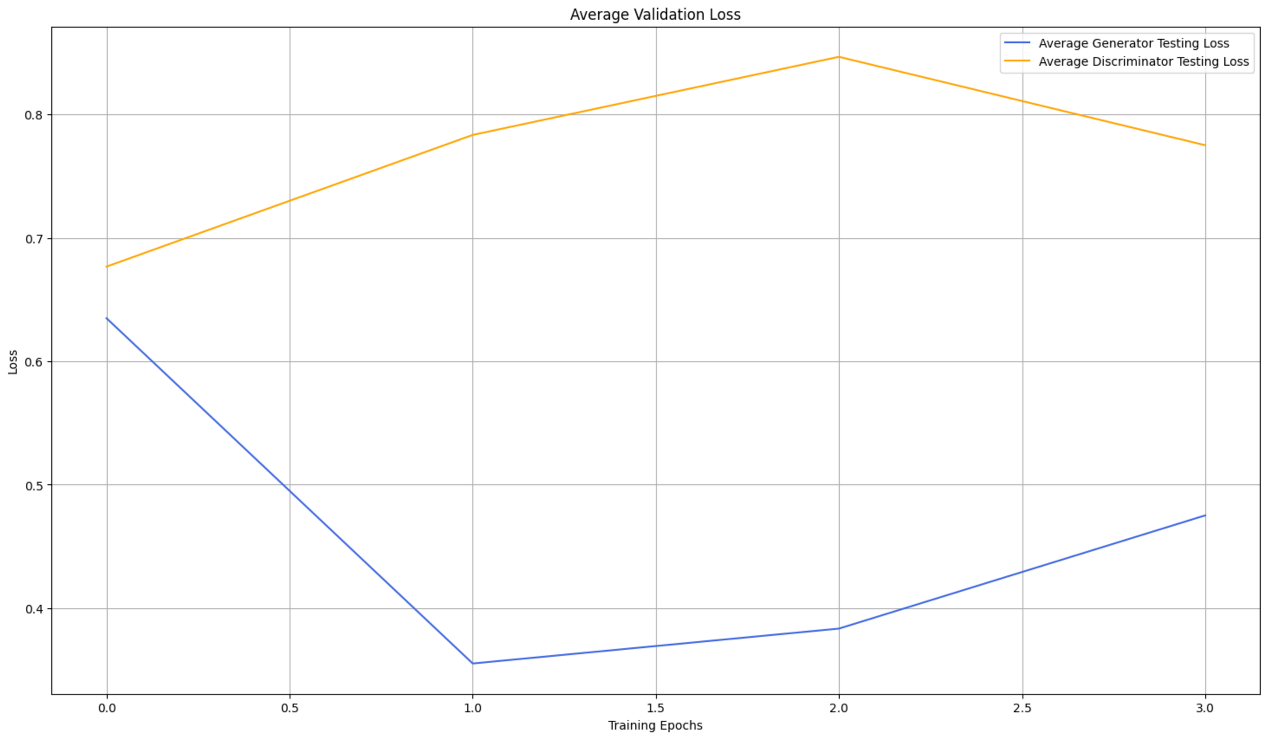


Figure 6.4.4: Average Testing Loss for 10K dataset

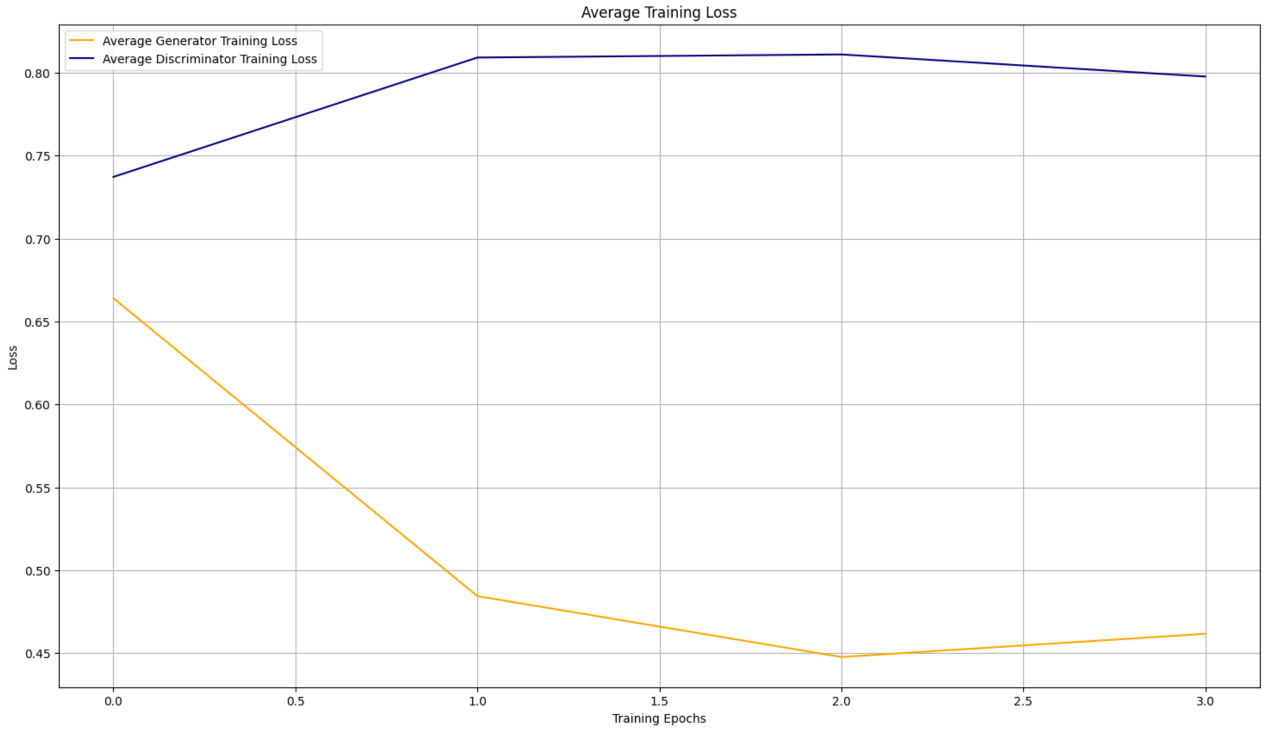


Figure 6.4.5: Average Training for 15K dataset

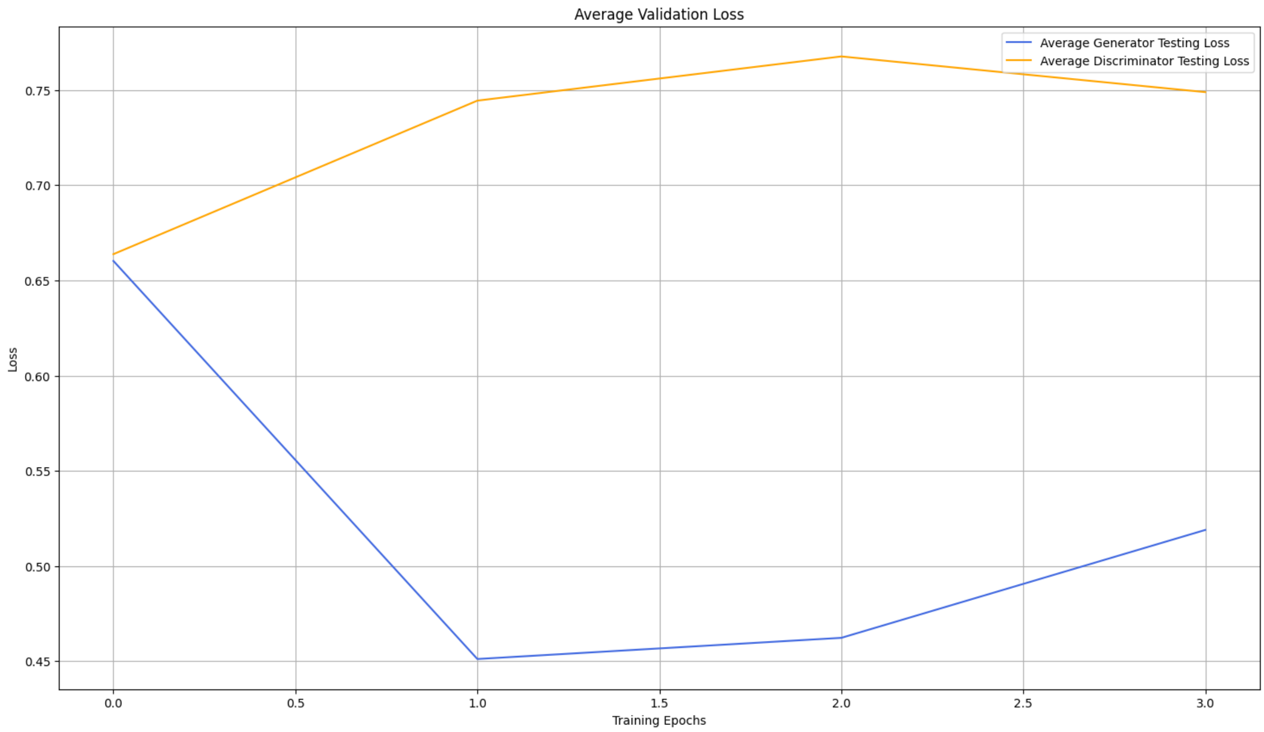


Figure 6.4.6: Average Testing Loss for 15K dataset

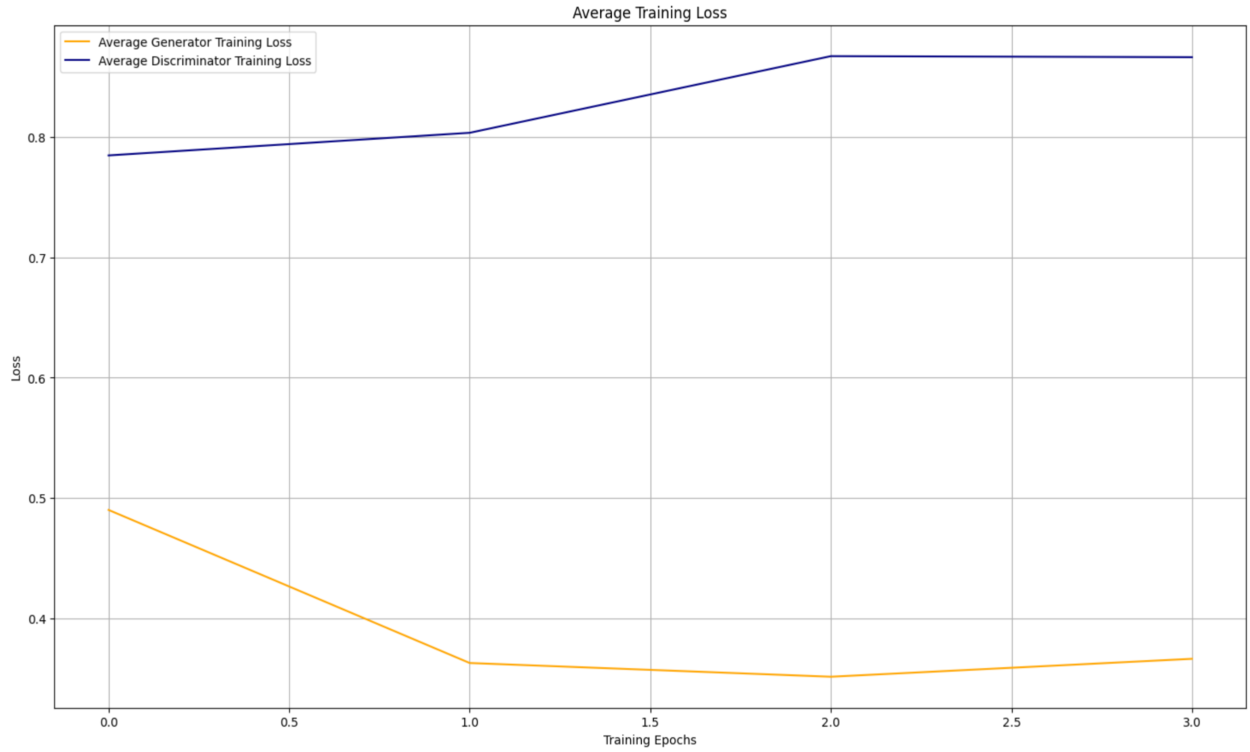


Figure 6.4.7: Average Training for 20K dataset

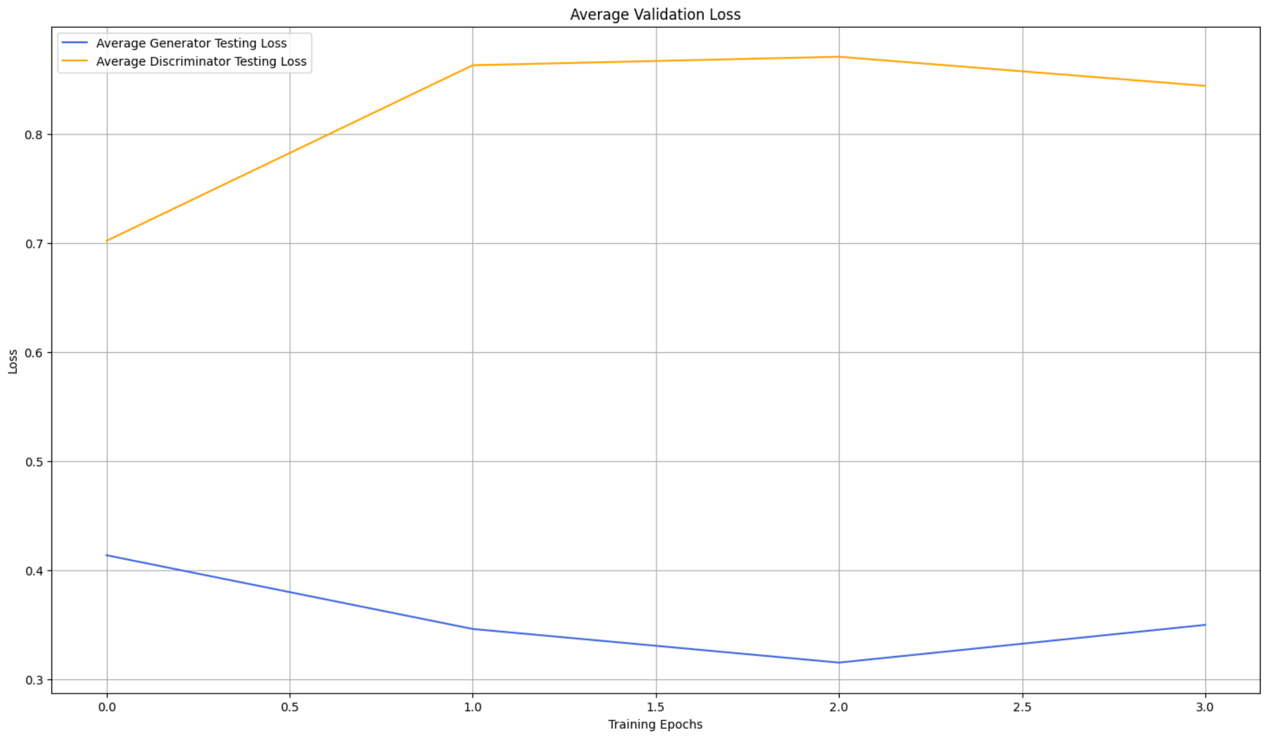


Figure 6.4.8: Average Testing Loss for 20K dataset

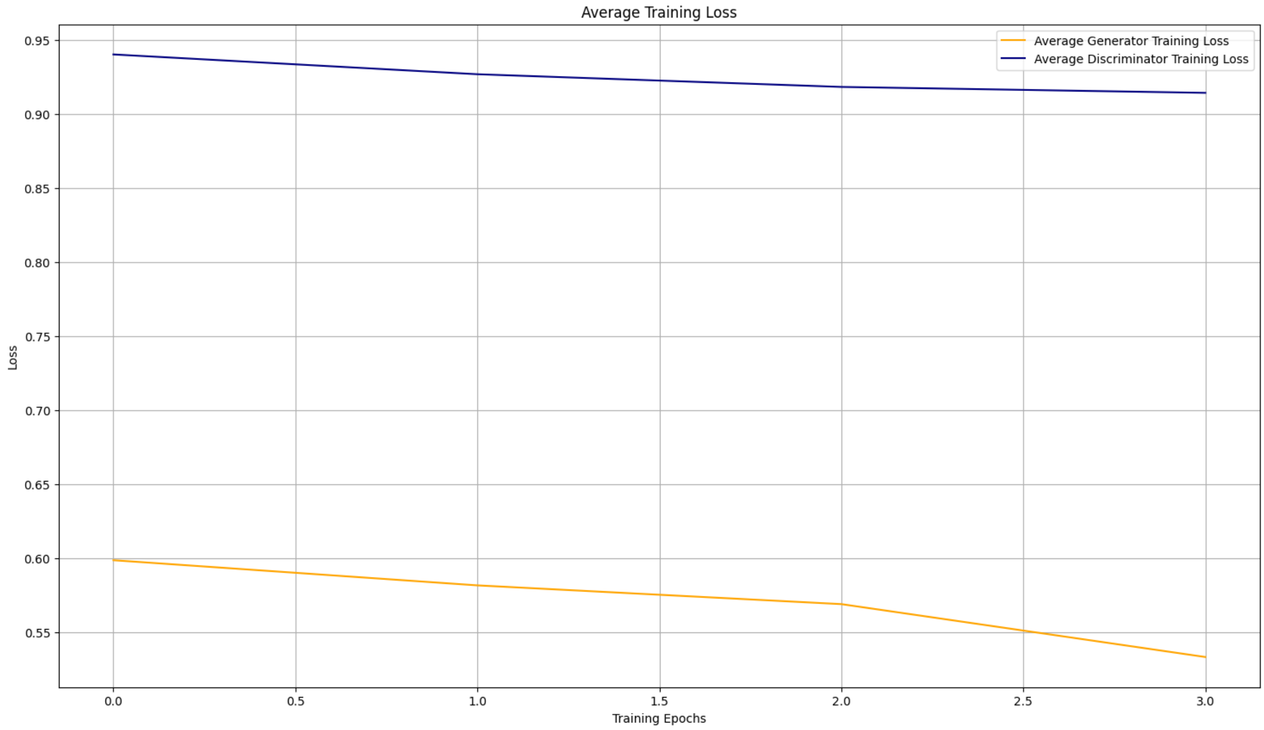


Figure 6.4.9: Average Training for 25K dataset

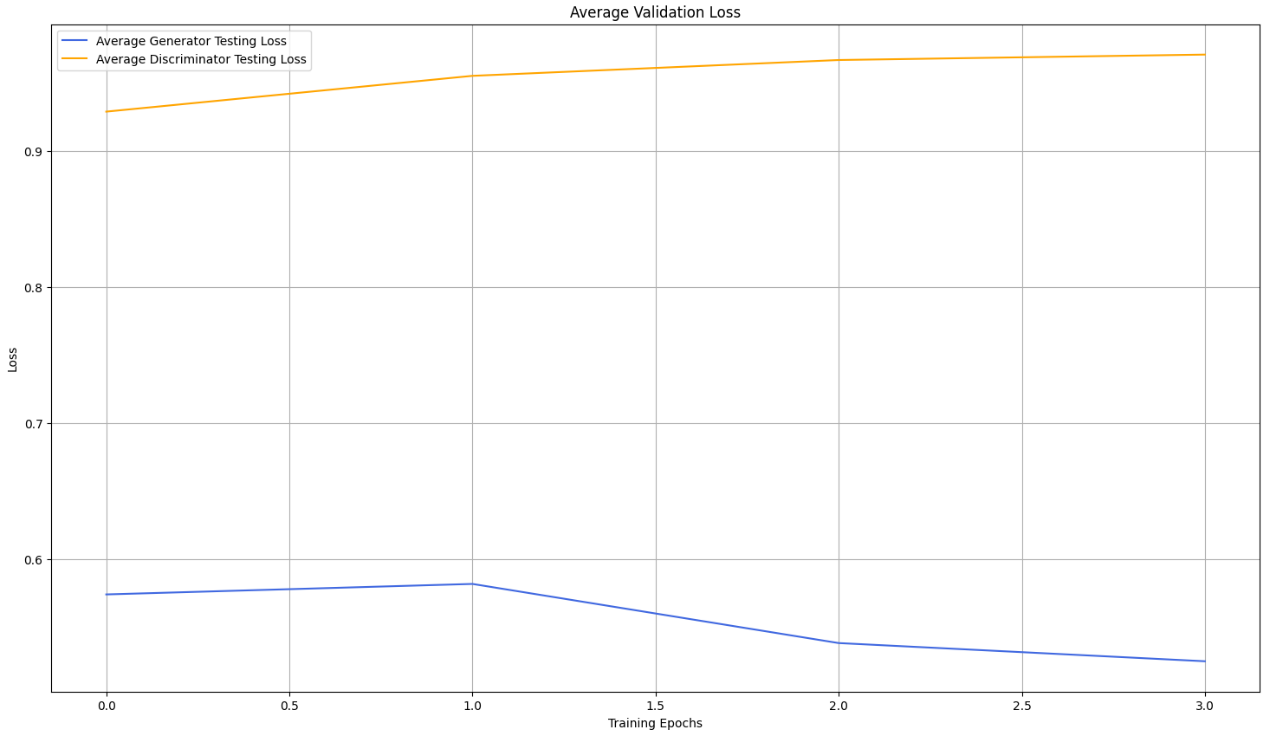


Figure 6.4.10: Average Testing Loss for 25K dataset

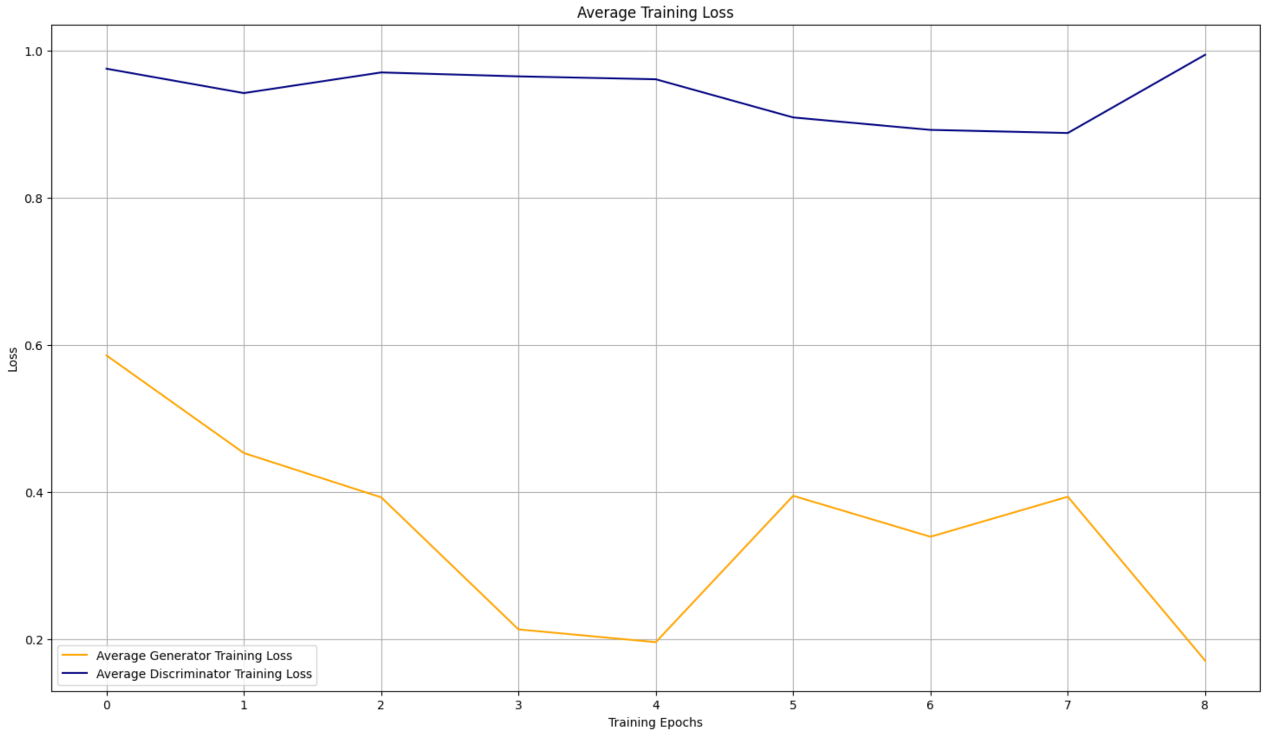


Figure 6.4.11: Average Training for 30K dataset

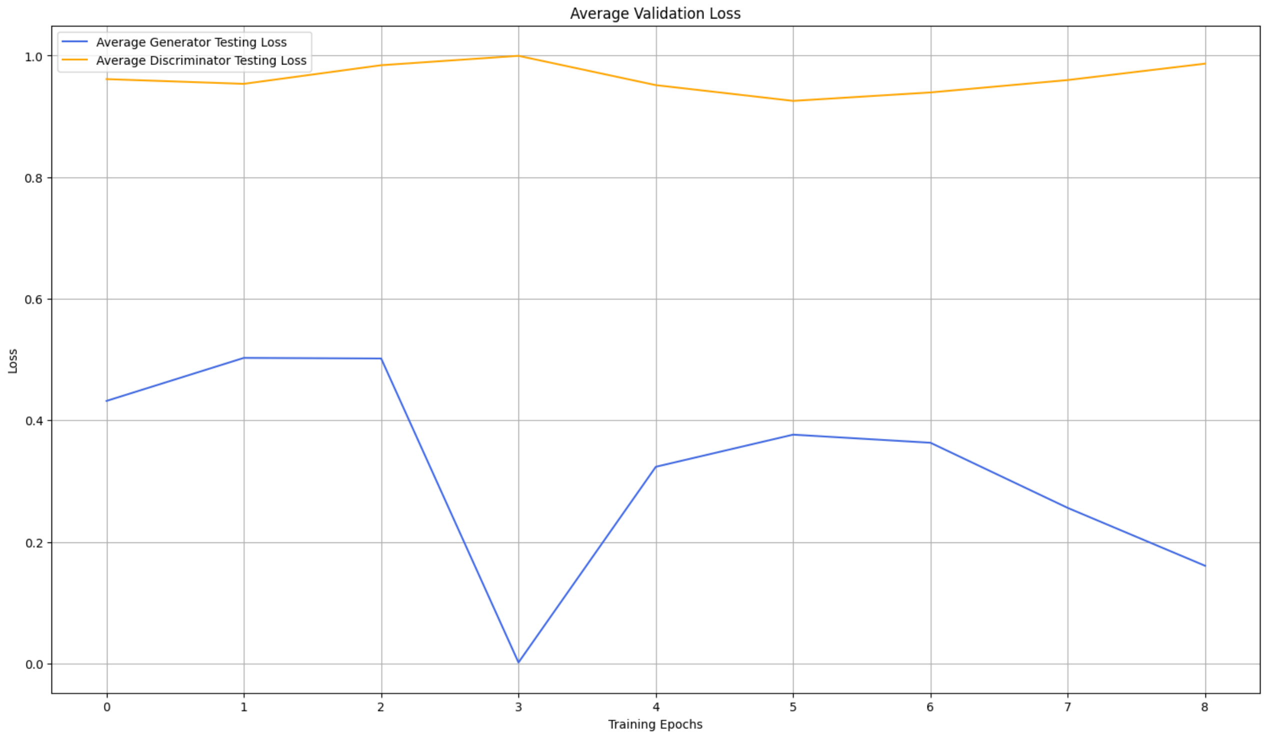


Figure 6.4.12: Average Testing Loss for 30K dataset

7. GAN training with untrained GEN

7.1. GAN architecting

In our investigative trajectory, grounded in the principles of Generative Adversarial Networks elucidated by Goodfellow et al. [1], our objective is to critically assess and juxtapose the efficacy of two distinct generator configurations: a pretrained generator, which we refer to as GEN, and its counterpart, an untrained GEN. Within the encompassing GAN architecture, both versions of GEN are slated to undergo a collaborative training regimen alongside a discriminator, aptly named DIS. This training paradigm involves the generators striving to produce compelling text summaries, while DIS plays the evaluative role, distinguishing between genuine summaries and those synthesized by GEN. Through this comparative analysis, we aspire to discern the tangible advantages, if any, offered by pretraining in enhancing the quality and reliability of the generated summaries.

7.2. GAN Training Parameters

- Vocabulary size = 22333 (from dataset)
- Embedding dimension = [250, 500, 750]
- Latent dimension = [Embedding dimension, Embedding dimension X 2]
- LSTM layers(s) = [1, 2, 3]
- Batch size = [15, 30, 45, 60]
- Epochs = [Batch size, Batch size X 2]
- Predicted summary word length = 5 (from dataset)
- Dropout = [0.3, 0.5, 0.7]
- Learning rate = [0.01, 0.001, 0.0001]
- Loss = Cross Entropy
- Optimizer = Adam
- Activation = Sigmoid
- Early Stopping = 2

7.3. GAN Architecture Diagrams

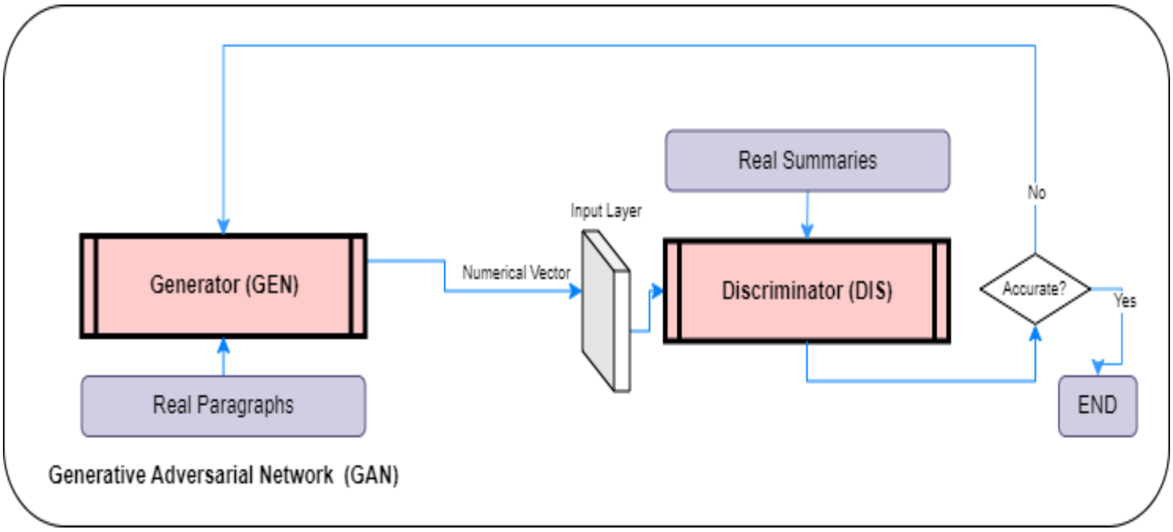


Figure 7.3.1: GAN architecture

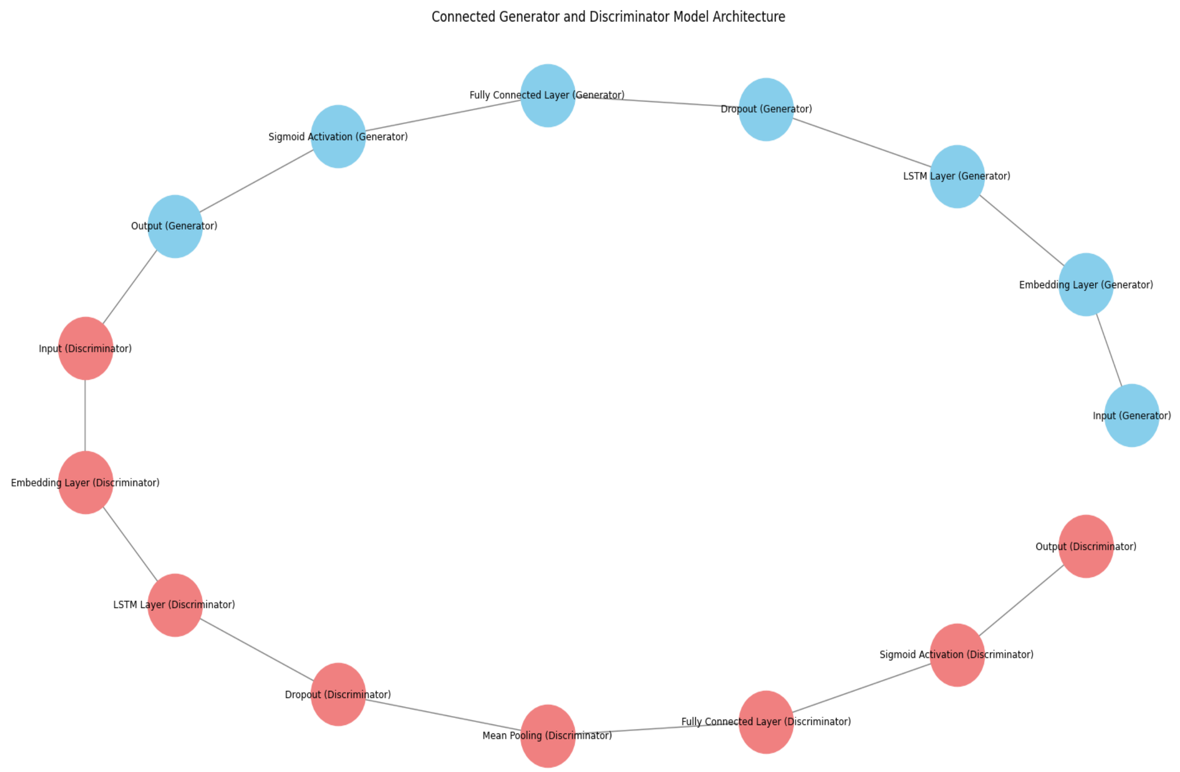


Figure 7.3.2: GAN model

7.4. GAN Losses

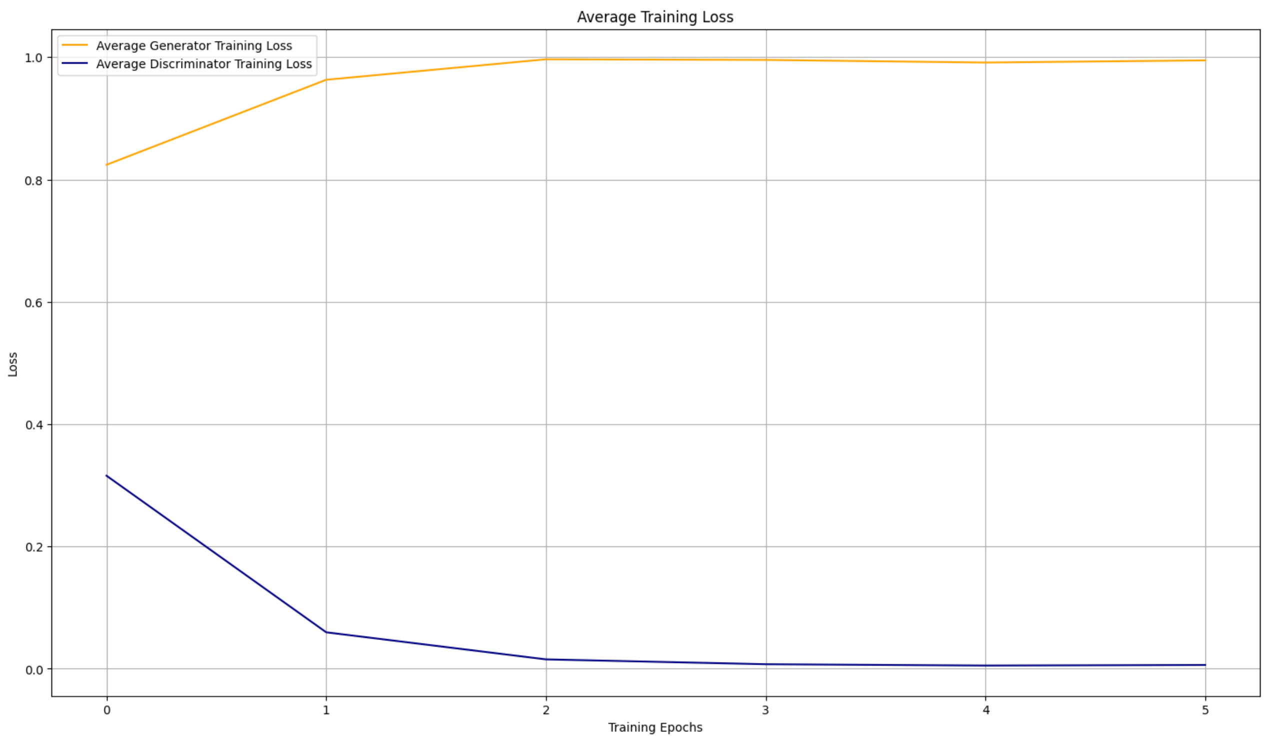


Figure 7.4.1: Average Training Loss for 5K dataset

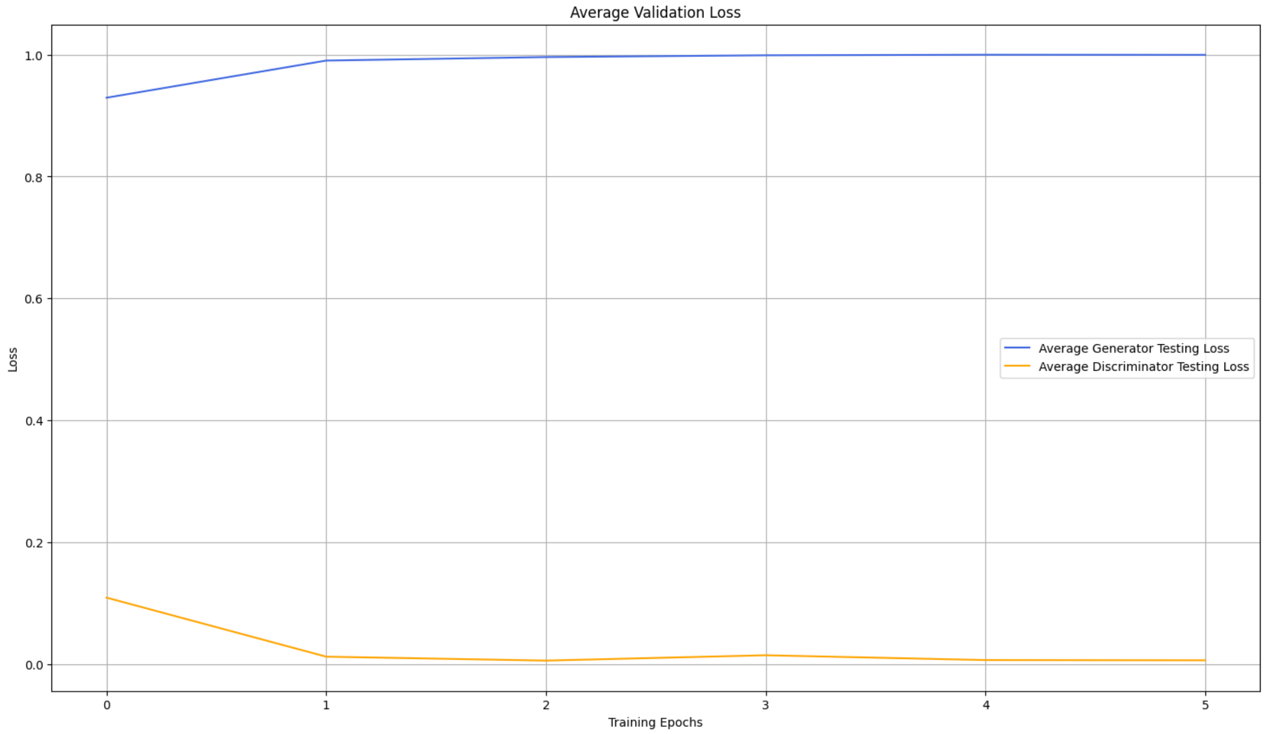


Figure 7.4.2: Average Testing Loss for 5K dataset

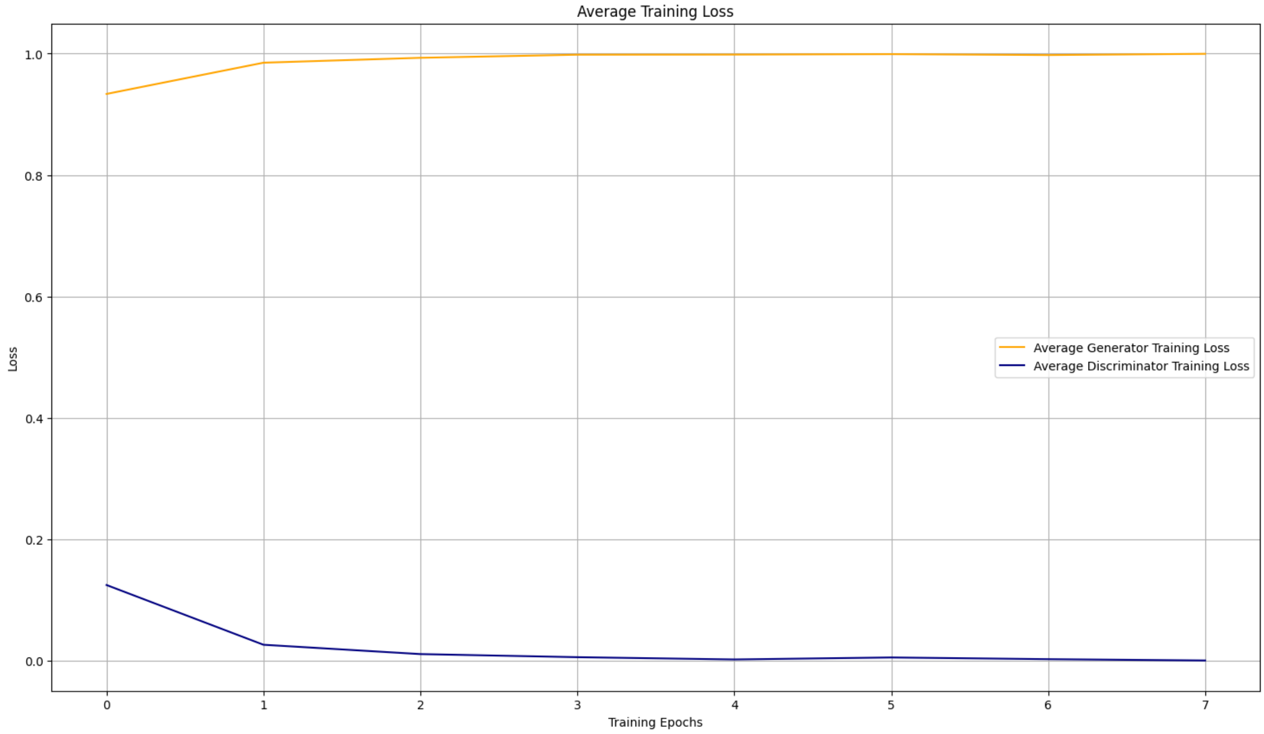


Figure 7.4.3: Average Training for 10K dataset

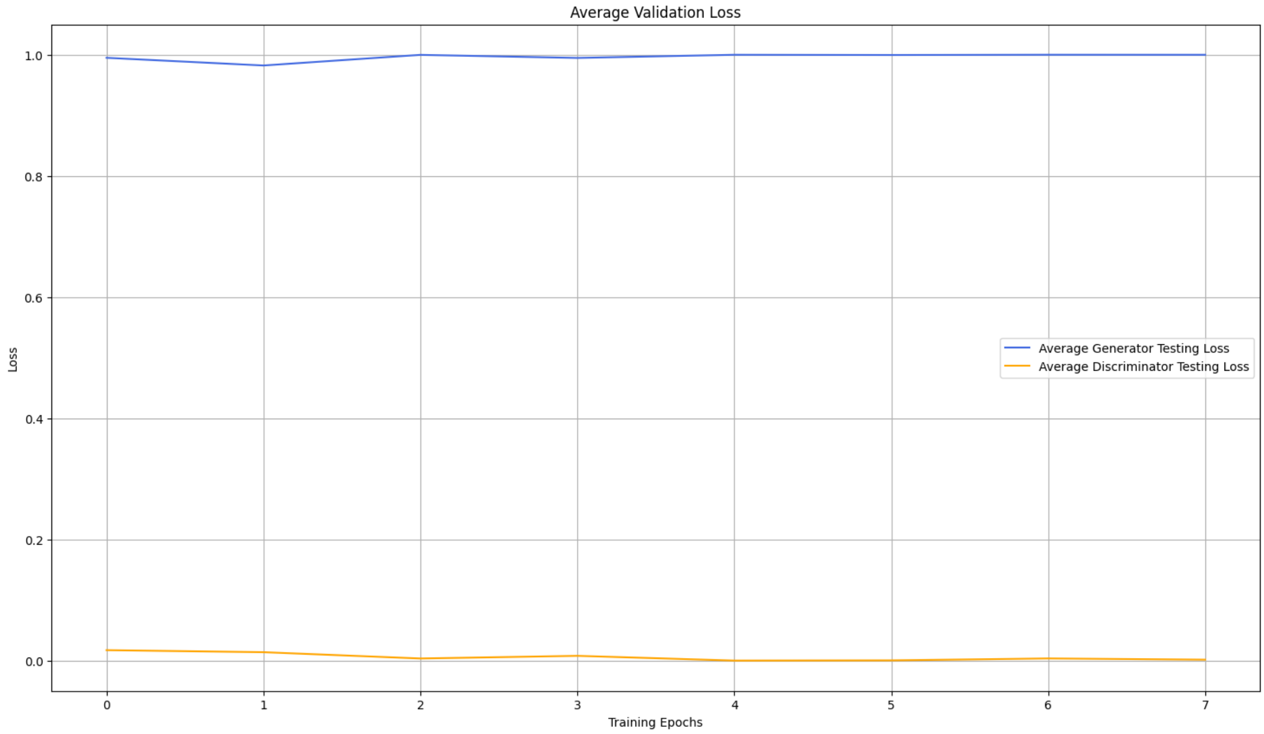


Figure 7.4.4: Average Testing Loss for 10K dataset

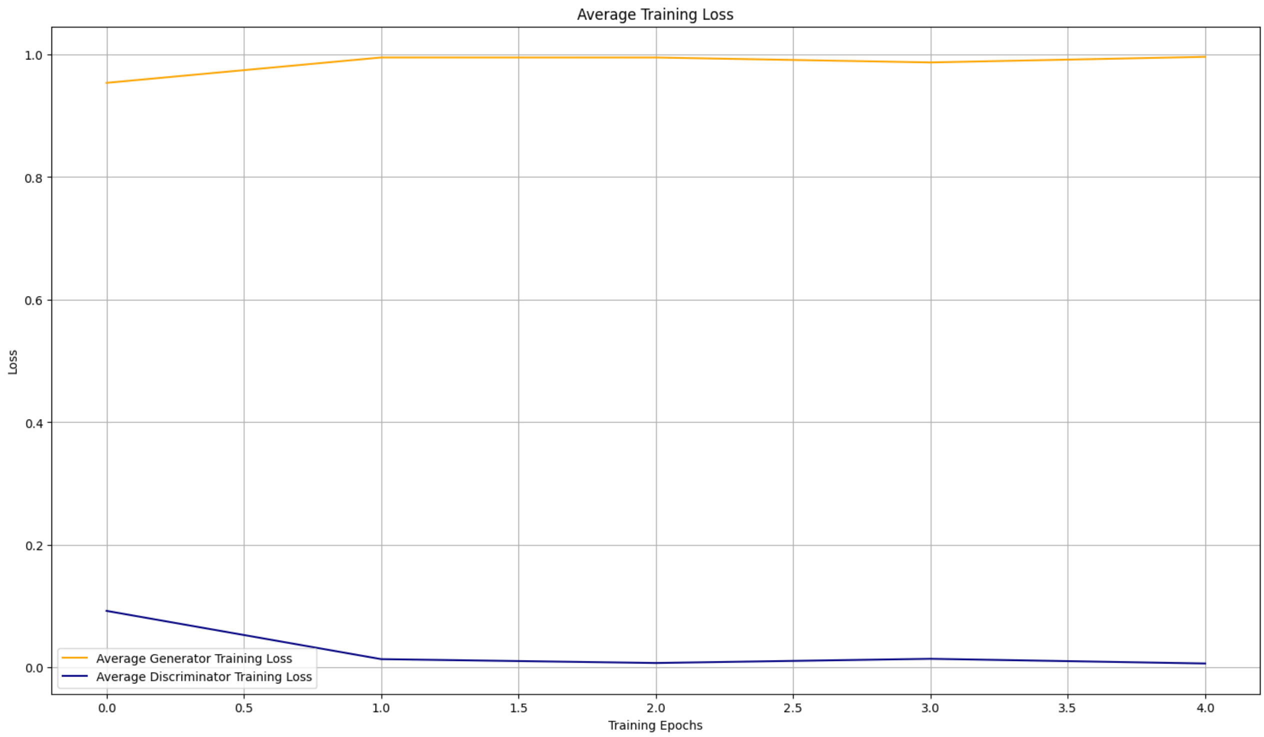


Figure 7.4.5: Average Training for 15K dataset

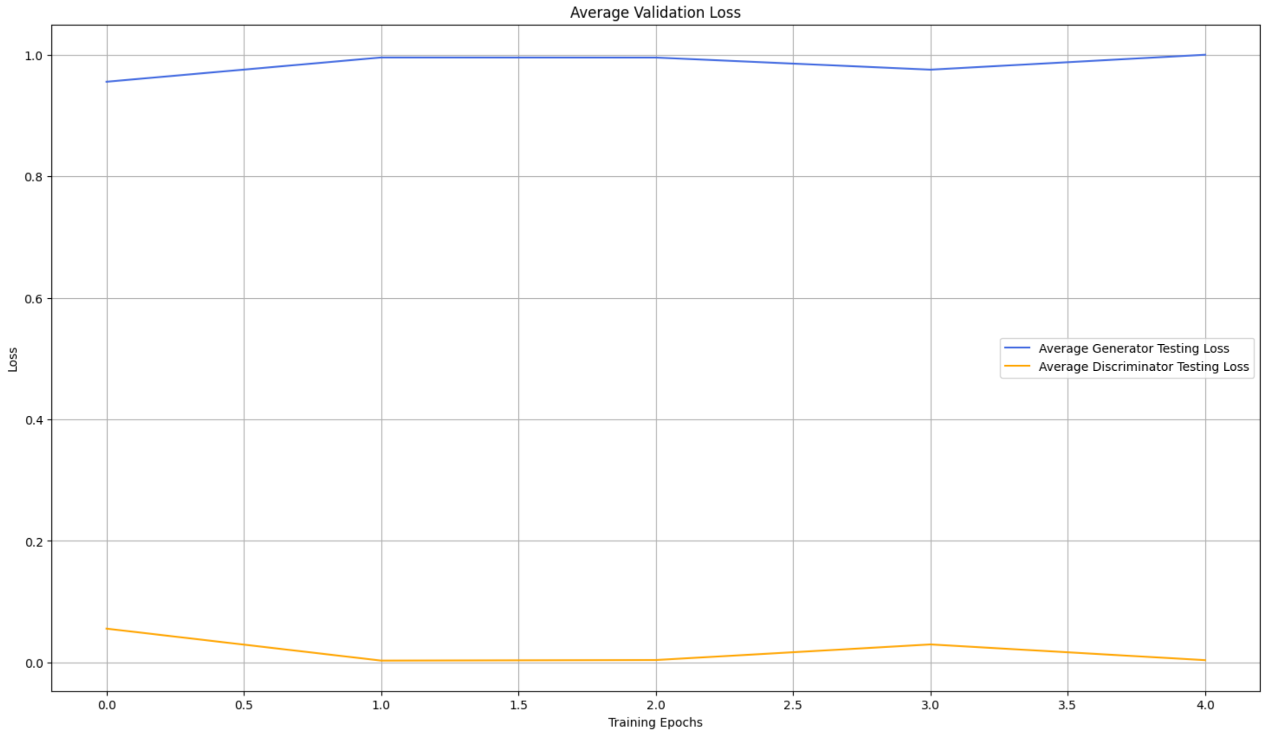


Figure 7.4.6: Average Testing Loss for 15K dataset

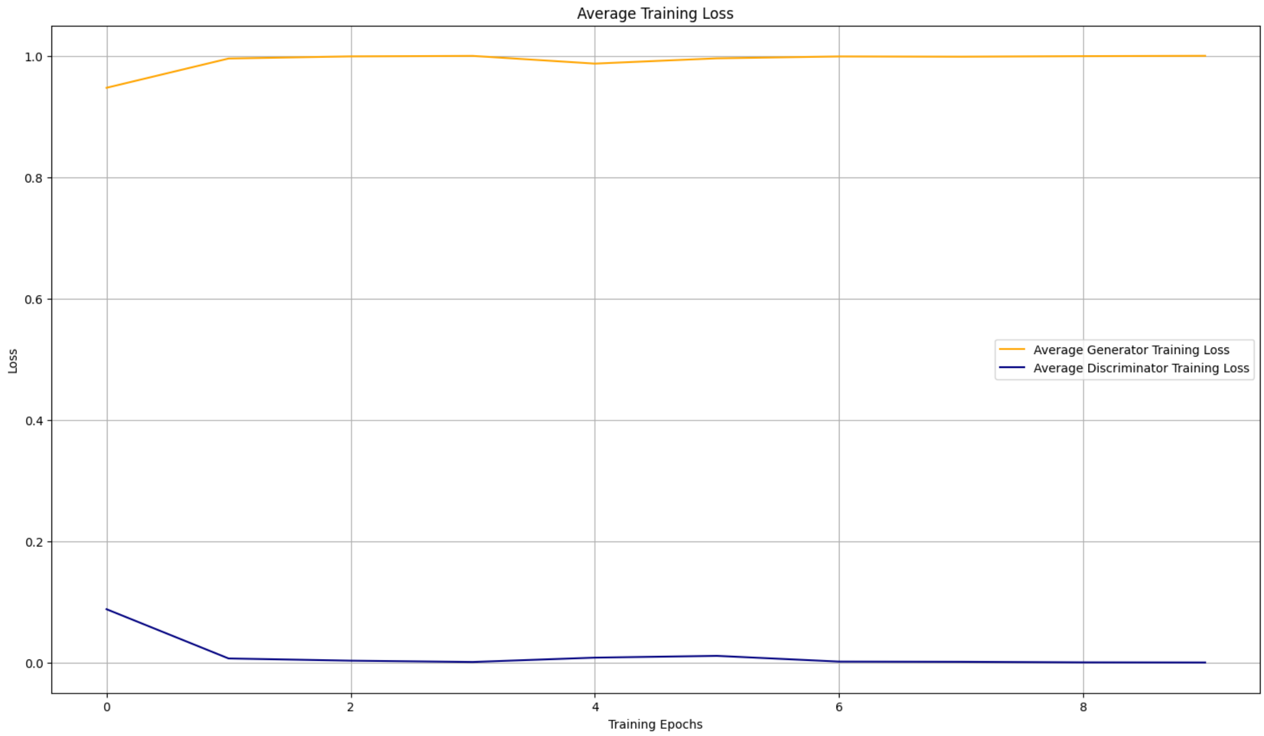


Figure 7.4.7: Average Training for 20K dataset

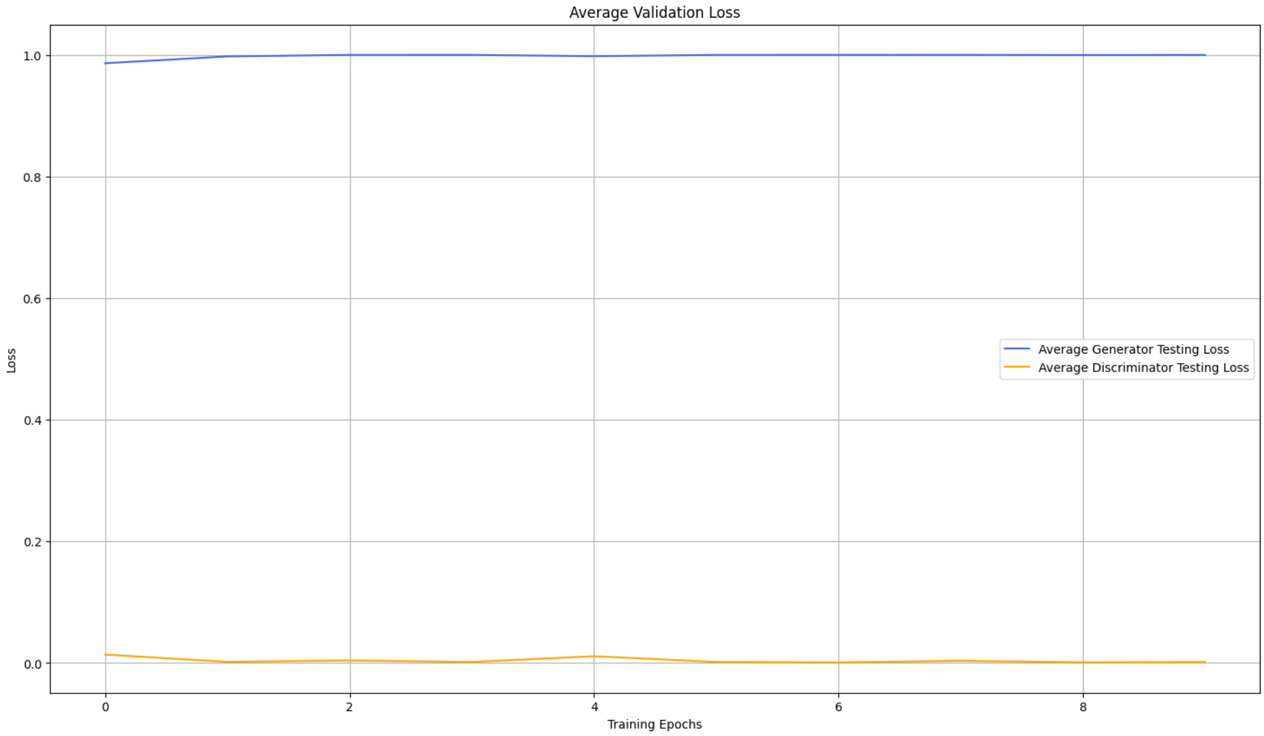


Figure 7.4.8: Average Testing Loss for 20K dataset

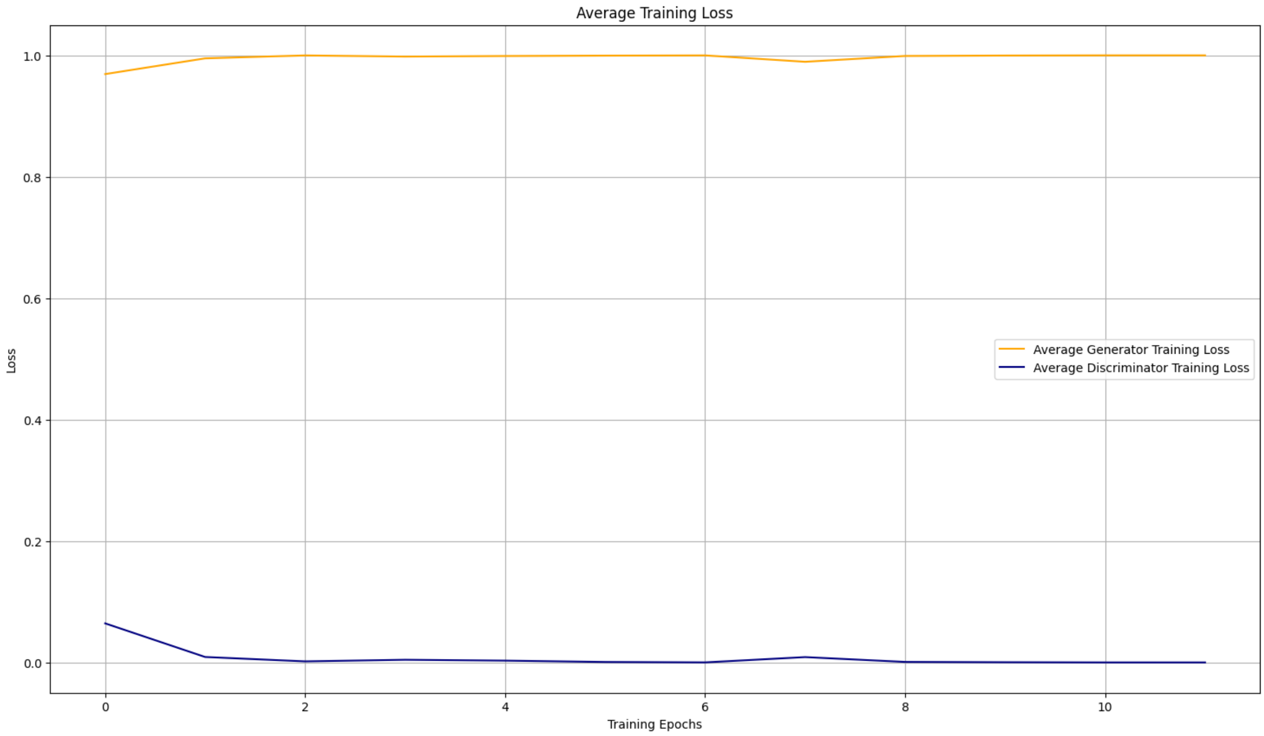


Figure 7.4.9: Average Training for 25K dataset

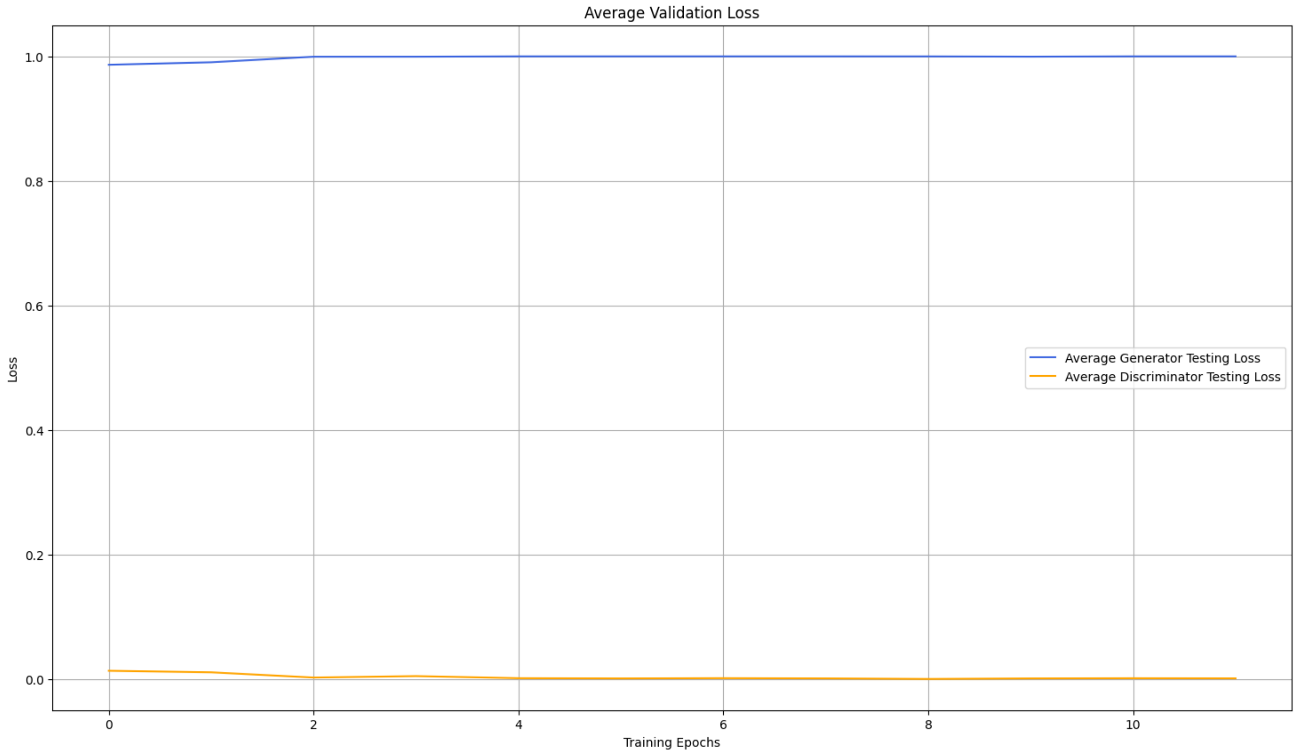


Figure 7.4.10: Average Testing Loss for 25K dataset

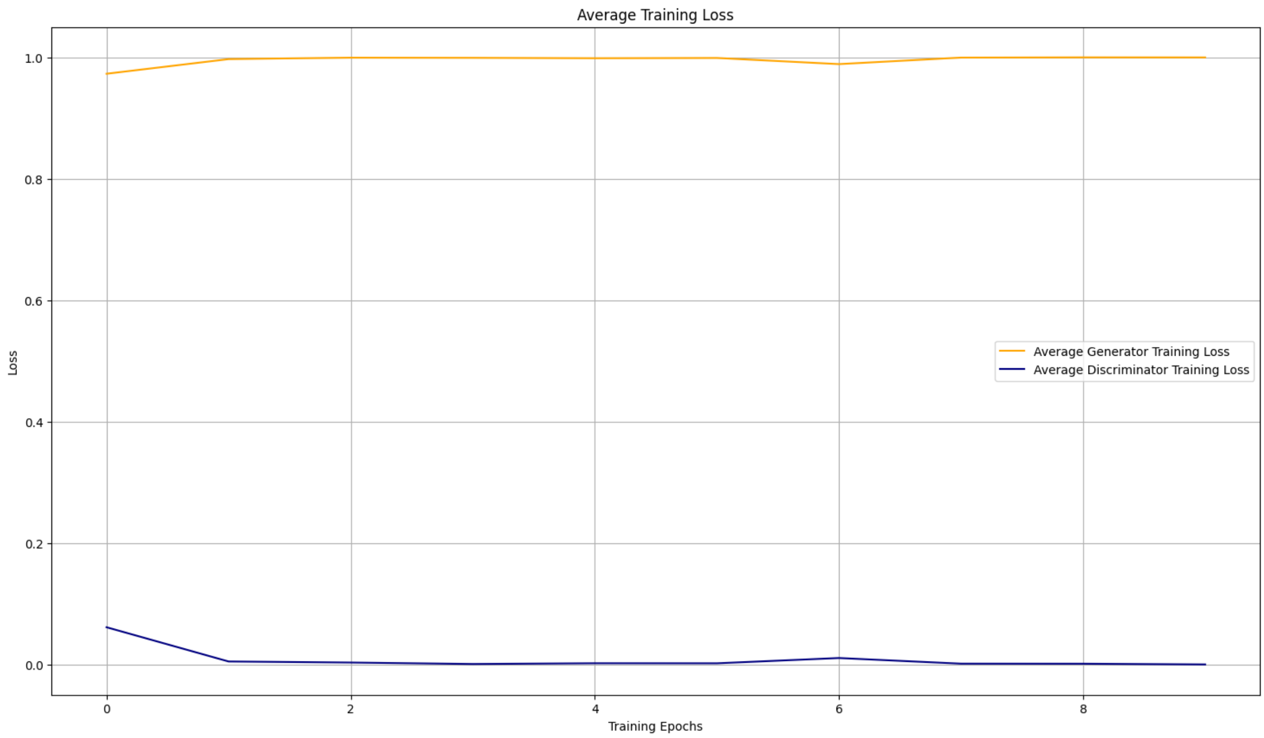


Figure 7.4.11: Average Training for 30K dataset

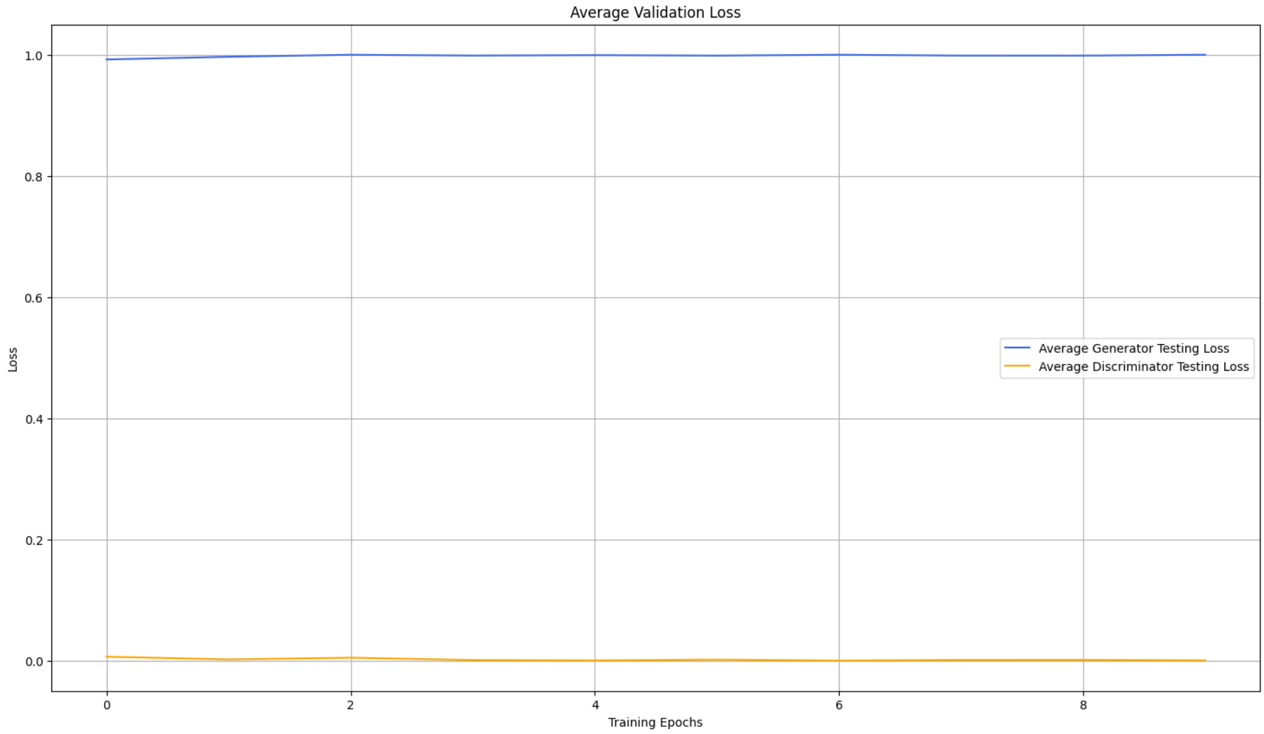
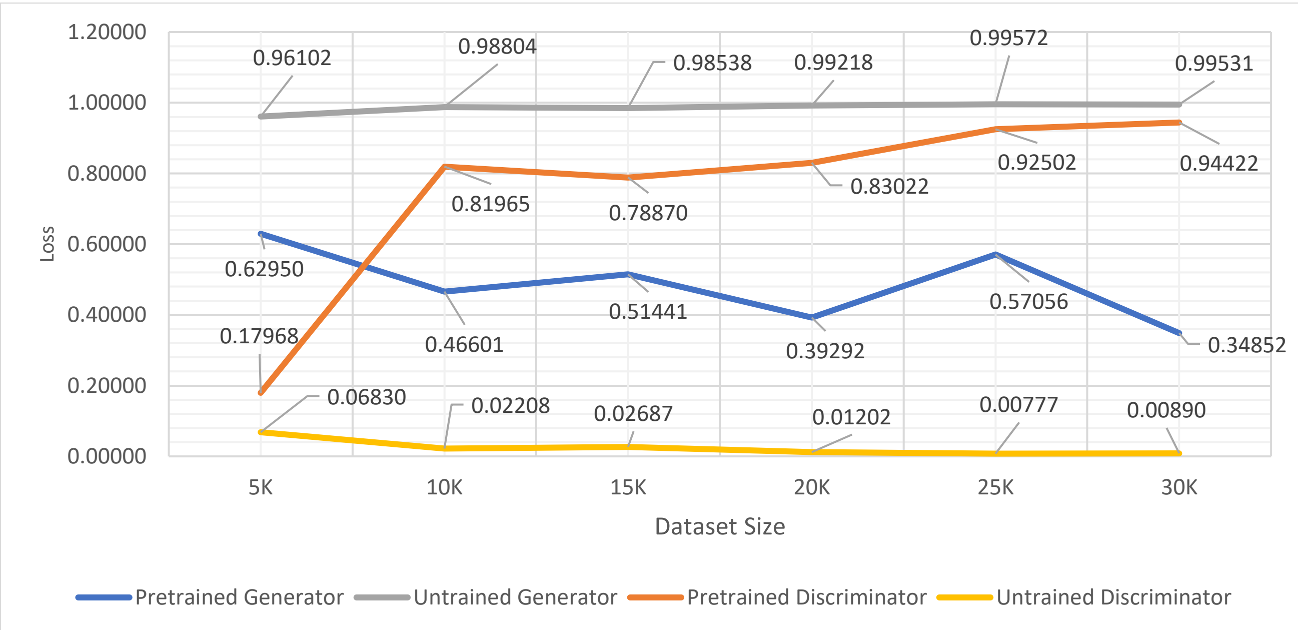


Figure 7.4.12: Average Testing Loss for 30K dataset

8. *Comparison of pretrained vs untrained GAN*

8.1. *GAN Training*



Key points from the data,

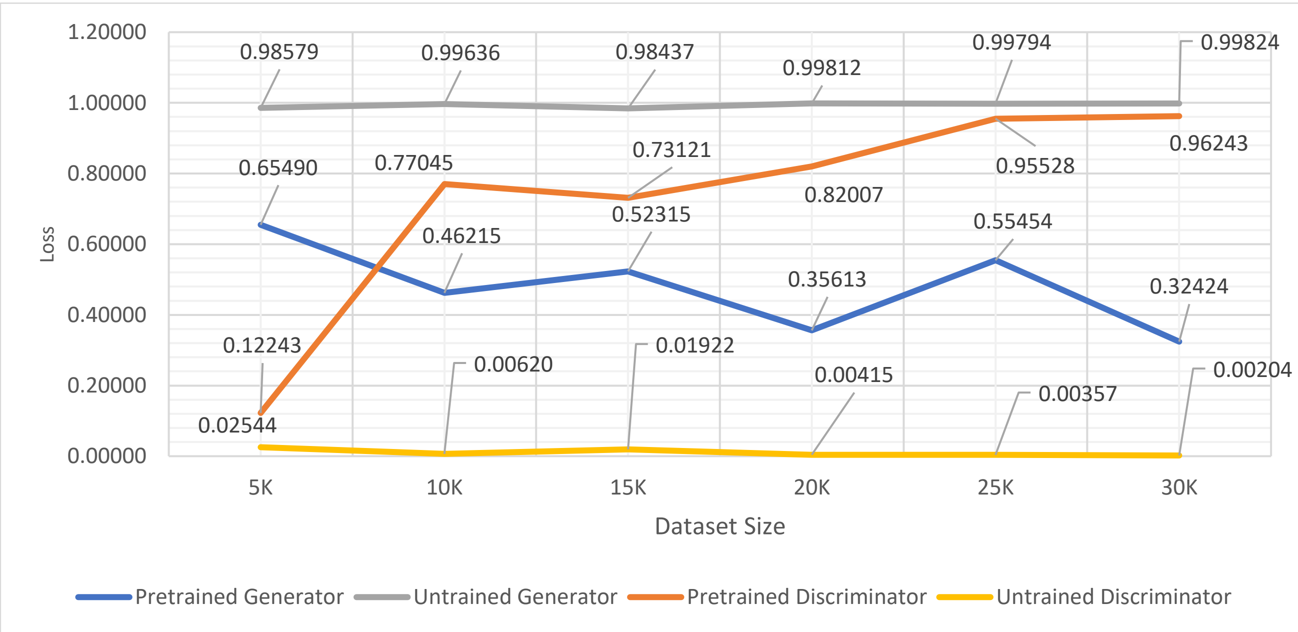
- Generator:
  - At every stage of training (5K, 10K, 15K, etc.), the pretrained generator's values, informed by prior knowledge from similar architectures [1], are consistently lower than those of the

untrained generator. This observational trend underscores that the pretrained generator has a competitive edge over its untrained counterpart.

   - The values for the pretrained generator, aligned with typical performance patterns observed in deep learning paradigms [6], exhibit fluctuations but generally trend downward, hinting at a potential convergence to an optimal solution. In contrast, the untrained generator's values, reminiscent of challenges faced by nascent neural models [5], escalate over time, which casts doubt on its convergence or suggests it may be gravitating towards a suboptimal solution.

- Discriminator
   - The pretrained discriminator, drawing insights from established GAN architectures [1], manifests values that ascend with continued training. Such a trend can be interpreted as the discriminator grappling with the increasing challenge of distinguishing real samples from synthetic ones, especially as the generator refines its capabilities.
   - Contrastingly, the untrained discriminator showcases values that not only consistently trail the pretrained discriminator's metrics but also decline as training progresses, potentially alluding to its lagging efficacy relative to the pretrained version [4].
- Analyzing the pretrained GAN's metrics, which encompass both the generator and discriminator, reveals noticeable oscillations throughout various training phases, possibly suggesting a dynamic exploration of the solution landscape [3]. This stands in contrast to the untrained GAN, whose metrics — whether for the generator or discriminator — depict a more monotonous trajectory without pronounced fluctuations.

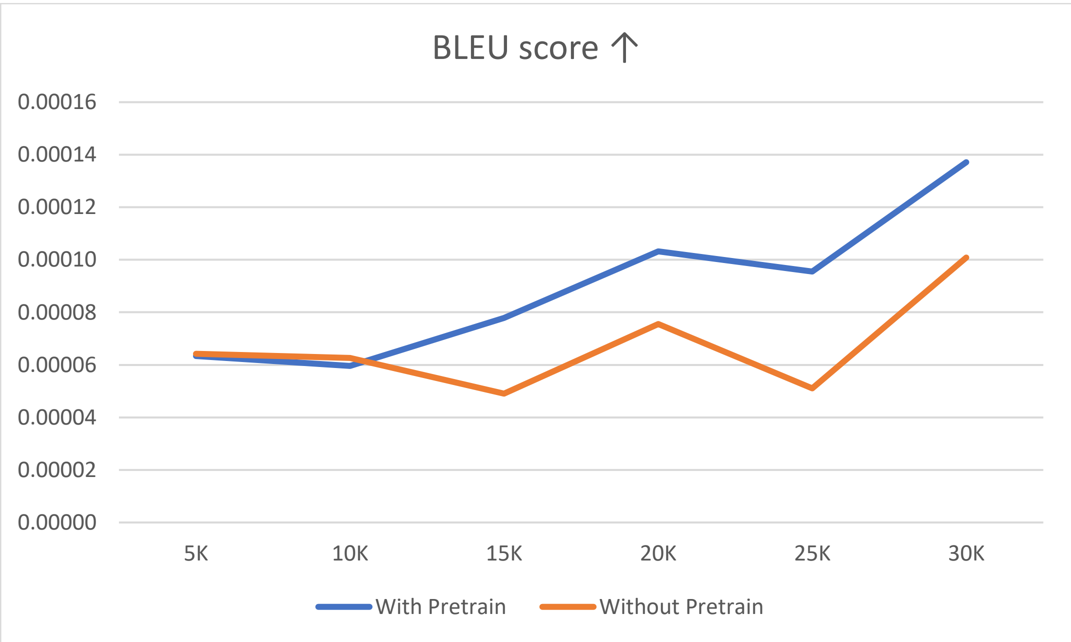
8.2. GAN Validation

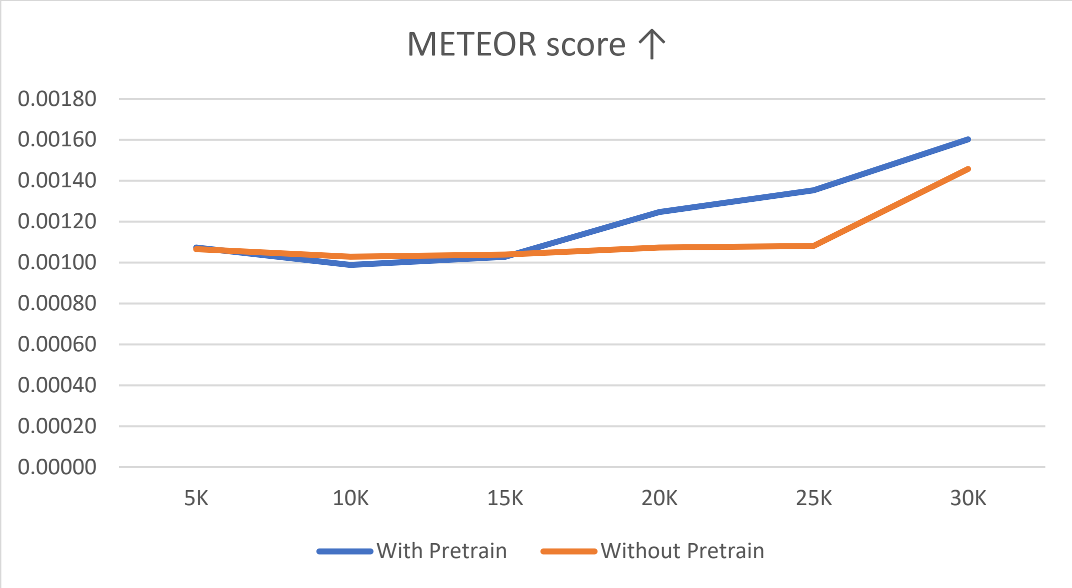


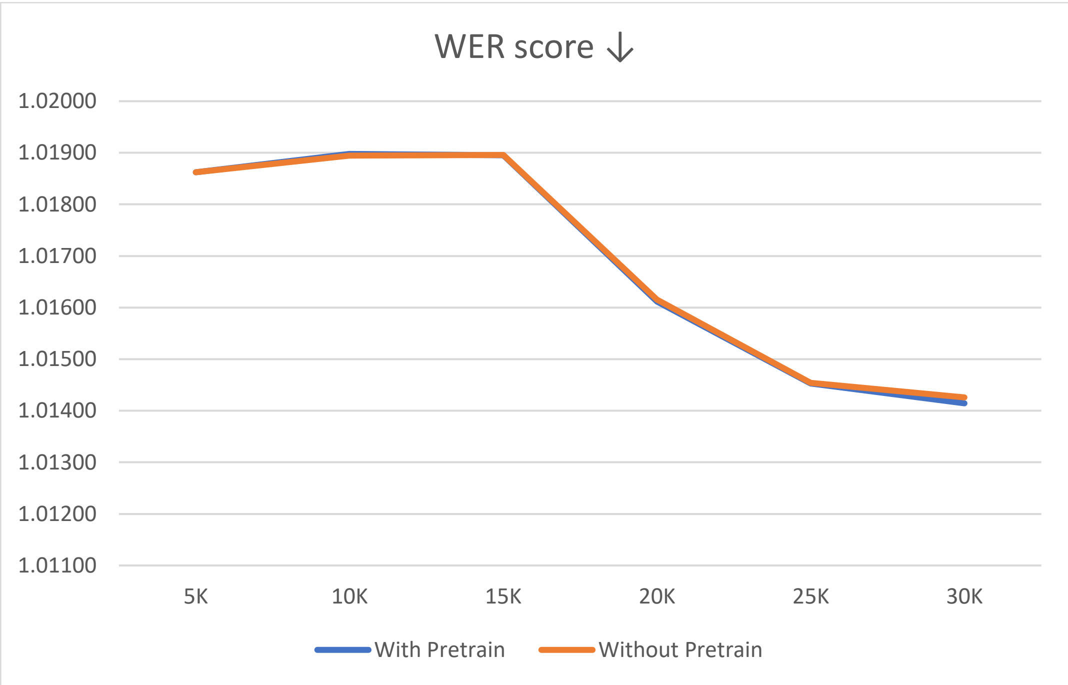
Key points from the data,

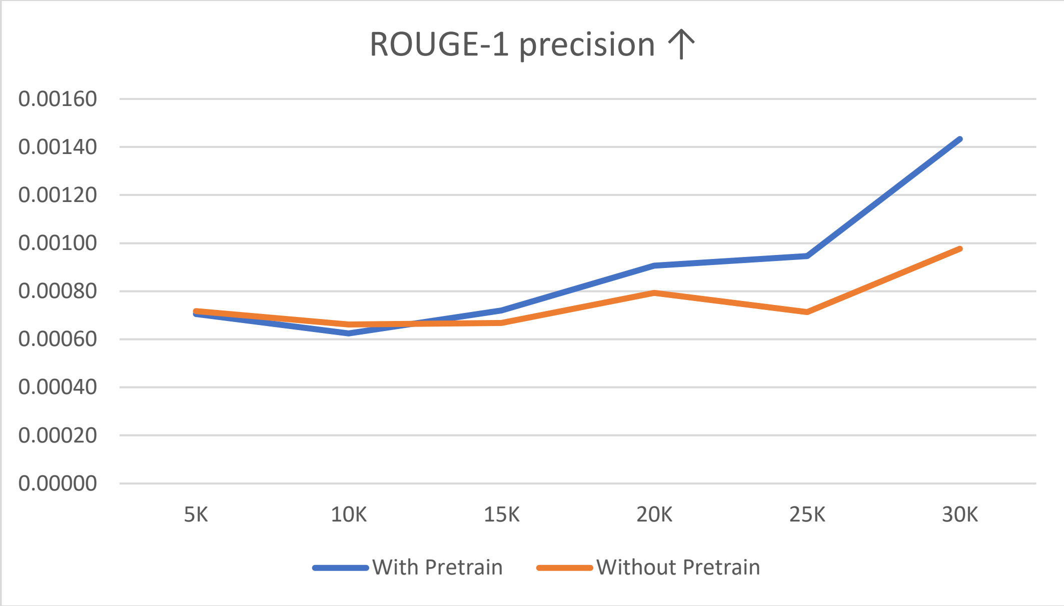
- Generator
    - At every testing milestone (5K, 10K, 15K, etc.), the pretrained generator, drawing from methodologies akin to sequence-to-sequence frameworks [5], consistently registers a lower value than the untrained generator. Such a trend serves as an empirical testament to its superior performance.
- Discriminator
    - As testing advances, the pretrained discriminator, informed by foundational principles of GANs as outlined by Goodfellow et al. [1], exhibits incrementing values. This upward trajectory might signify the escalating challenge it faces in differentiating genuine samples from those generated artificially.
    - In stark contrast, the untrained discriminator's metrics not only lag behind the pretrained discriminator's benchmarks but also depict a descending trend as testing continues. Such observations hint at the untrained discriminator's potential stagnation or even regression in performance [6].
- From a holistic perspective, the values associated with the pretrained GAN, enriched by prior training paradigms [3], exhibit fluctuations across the testing milestones. This variability might be indicative of its dynamic traversal through the solution landscape. Conversely, the untrained GAN's metrics, whether pertaining to the generator or discriminator, either plateau or descend consistently, suggesting potential hurdles in effective learning or optimization challenges [4].

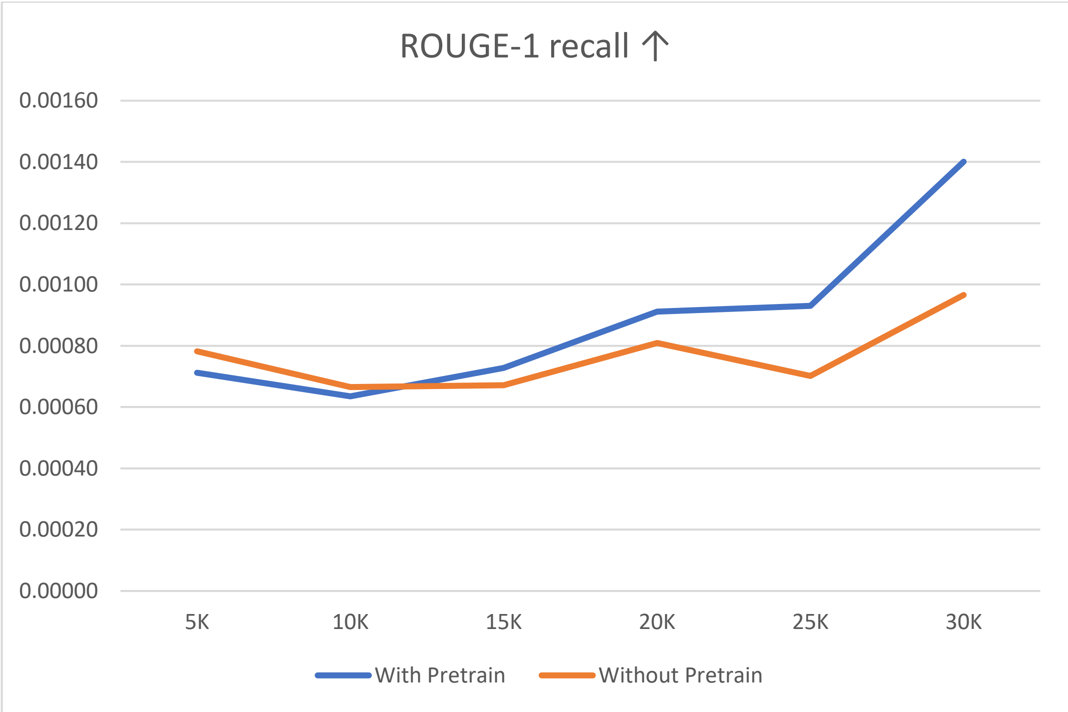
8.3. GAN Performance

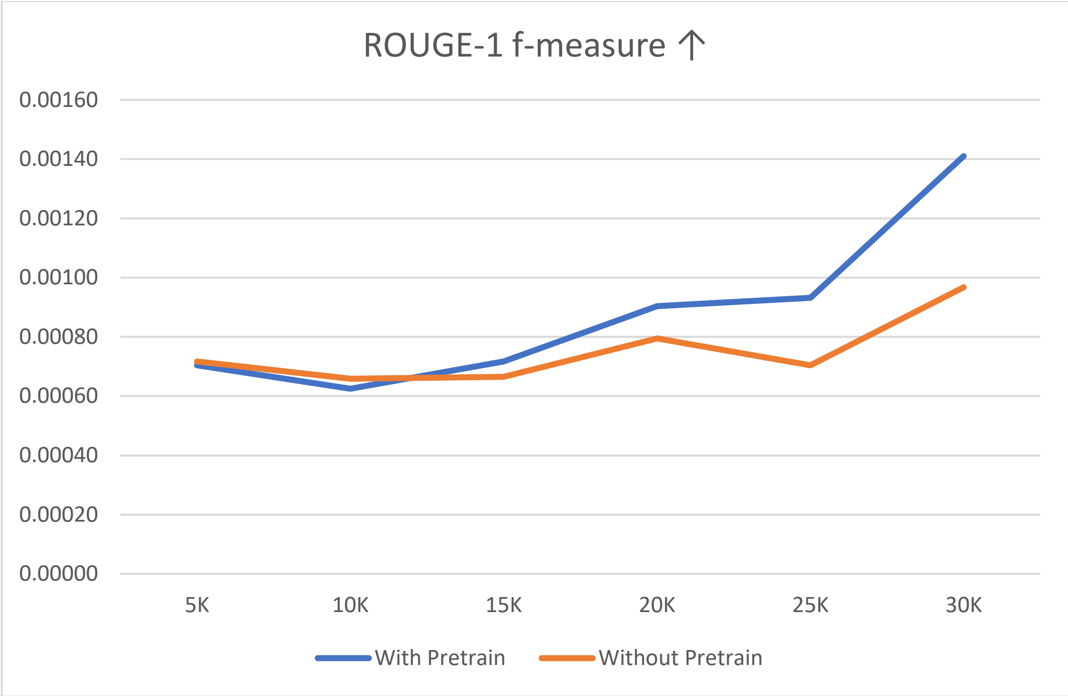


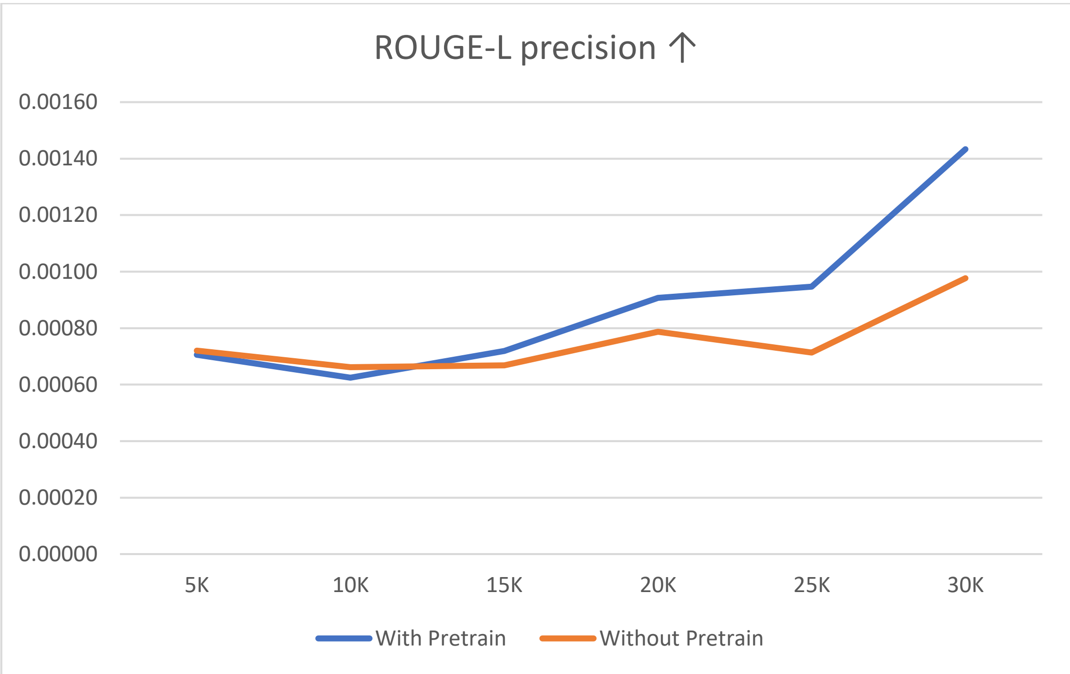

METEOR score ↑
0.00180
0.00160
0.00140
0.00120
0.00100
0.00080
0.00060
0.00040
0.00020
0.00000
5K
10K
15K
20K
25K
30K
With Pretrain
Without Pretrain

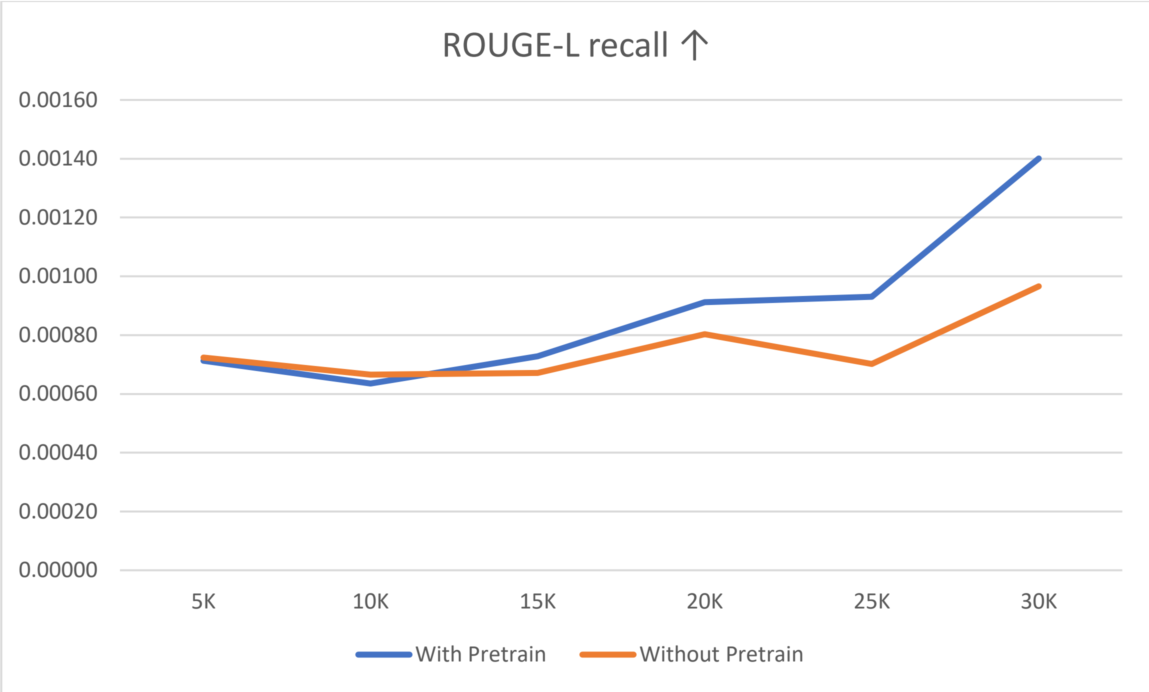

WER score ↓
1.02000
1.01900
1.01800
1.01700
1.01600
1.01500
1.01400
1.01300
1.01200
1.01100
5K
10K
15K
20K
25K
30K
With Pretrain
Without Pretrain

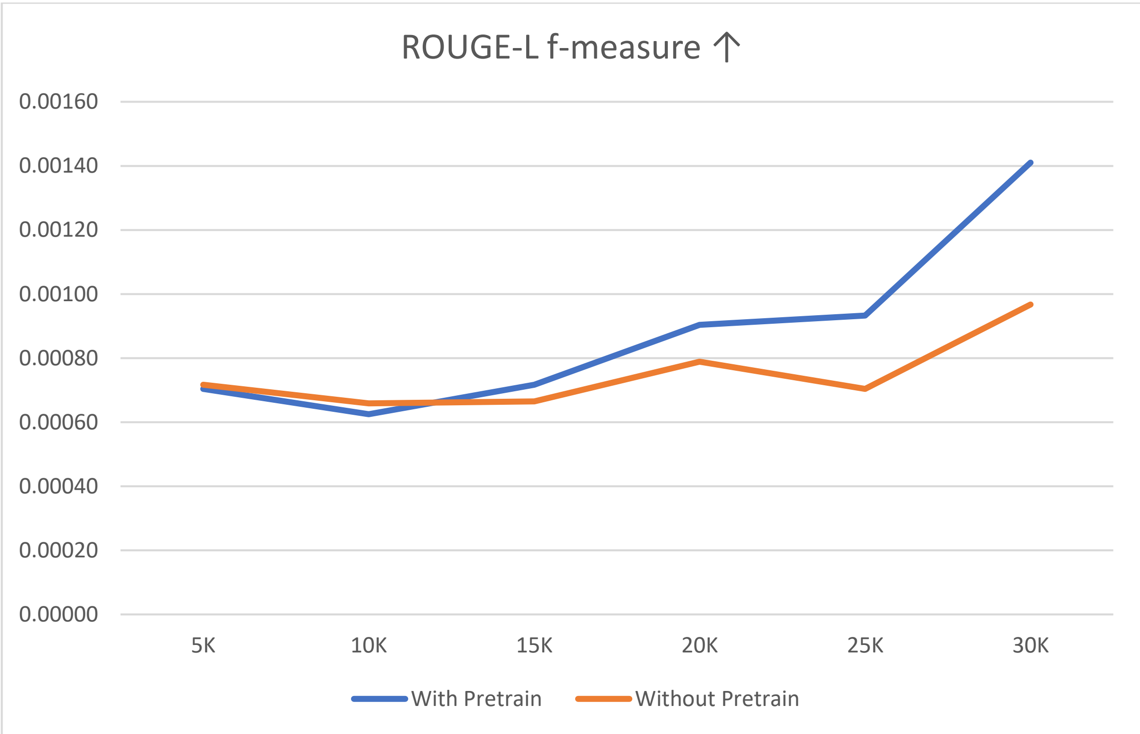

ROUGE-1 precision ↑
0.00160
0.00140
0.00120
0.00100
0.00080
0.00060
0.00040
0.00020
0.00000
5K
10K
15K
20K
25K
30K
With Pretrain
Without Pretrain


ROUGE-1 recall ↑
0.00160
0.00140
0.00120
0.00100
0.00080
0.00060
0.00040
0.00020
0.00000
5K
10K
15K
20K
25K
30K
With Pretrain
Without Pretrain


ROUGE-1 f-measure ↑
0.00160
0.00140
0.00120
0.00100
0.00080
0.00060
0.00040
0.00020
0.00000
5K
10K
15K
20K
25K
30K
With Pretrain
Without Pretrain


ROUGE-L precision ↑
0.00160
0.00140
0.00120
0.00100
0.00080
0.00060
0.00040
0.00020
0.00000
5K
10K
15K
20K
25K
30K
With Pretrain
Without Pretrain





8.4. GAN Output

plot	generated_original_summary
better times 1919 american silent comedy drama film directed marshall neilan follows woman forced choose faithless husband devoted suitor ultimately decides pursue happiness	woman brenda leonard saif chick happiness
1985 fantasy film directed rosemarie turko follows group people transported fantasy world must battle powerful force order return world group aided powerful wizard dungeonmaster enigmatic figure guides quest	nobility wizard evil sturgeon vsevolod ryder officials

log kya kahenge 1983 indian bollywood movie follows story young woman estranged family love lower class man despite facing opposition family society fights protect relationship prove love true	trance woman love ibbetson beth ramakrishnan
crime drama directed john cromwell follows tom mcquigg police captain determined take powerful crime boss mcquigg sets plan arrest boss backfires leads series events test mcquigg courage integrity	ones crime tuesdays forensic control belonging
movie first blood released 1982 follows john rambo vietnam veteran forced face harsh reality past small town sheriff tries arrest must fight survival sheriff forces national guard	rambo drilling mob chakoram tower fitri

9. *Conclusion*

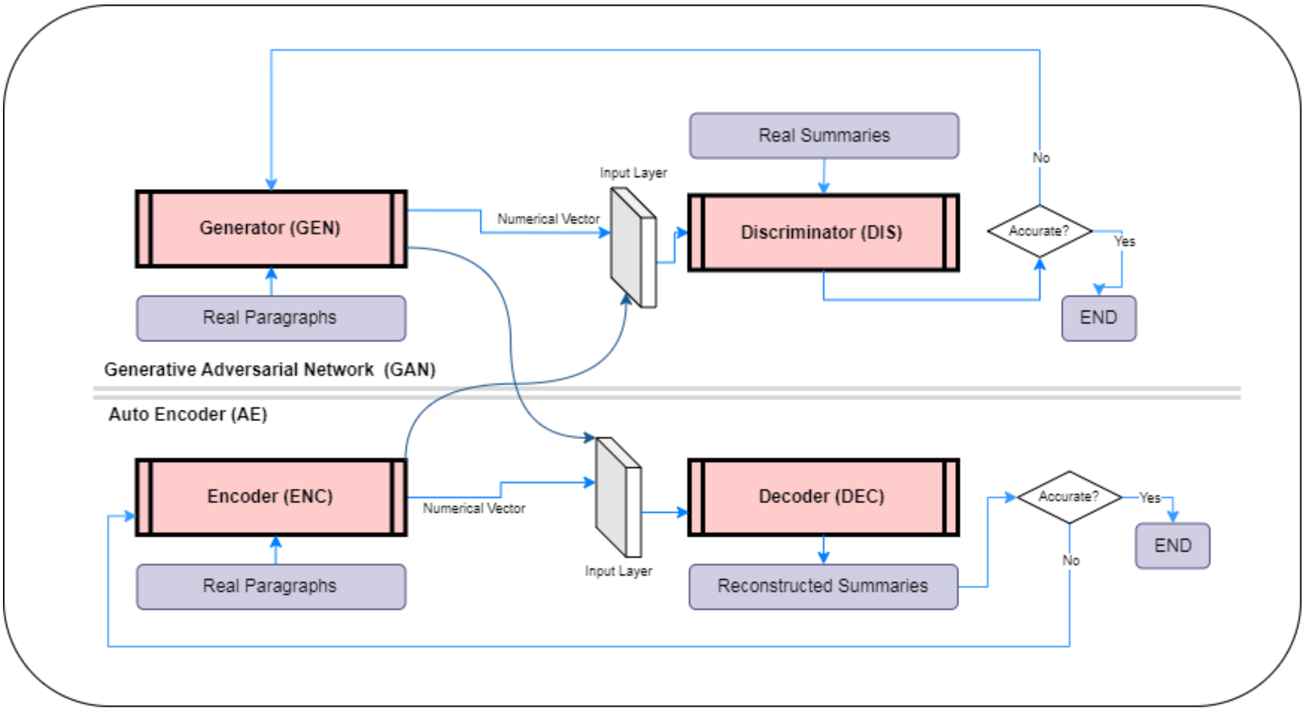
Efficient Information Processing: At the heart of the encoder-decoder architecture lies its ability to distill complex information into compact, latent representations and subsequently expand these into detailed outputs. This makes them especially adept at tasks like summarization, where capturing the essence of voluminous content succinctly is paramount.

Innovative Generation: GANs are intrinsically designed to generate novel data samples. Coupled with an encoder-decoder mechanism, this means they can create innovative summaries or ideas, extending beyond mere replication or rephrasing of existing content.

Advantages of Pretraining: Initiating the training process from a pretrained state provides models with foundational knowledge, potentially accelerating convergence and enhancing the quality of outputs. For tasks like summarization and idea generation, this can translate into coherent, contextually relevant, and high-quality results from the get-go.

Versatility and Customization: Given their pretrained nature, these GANs can be further fine-tuned on niche datasets or specific domains. This ensures that the generated content, be it summaries or ideas, is contextually aligned and tailored to the specific needs of a field or audience.

Refinement through Adversarial Training: The competitive dynamic inherent to GANs, where the generator and discriminator continuously challenge each other, ensures a consistent refinement of the generated content. This adversarial loop can lead to summaries and ideas that are not only accurate but also polished and of high quality.



10.Novelty

Combining Architectures: While encoder-decoder architectures and GANs have been explored individually in various contexts, merging them for the specific task of generating summaries and ideas is a fresh approach. Each structure brings its strengths: the encoder-decoder efficiently processes sequences, while the GAN ensures generation of novel content. The synergy of these two can lead to innovative outcomes.

Leveraging Pretraining: Pretraining has been widely adopted in deep learning to improve the efficiency and accuracy of models. Using pretrained models in the GAN context, especially with encoder-decoder structures for summarization and idea generation, means tapping into vast prior knowledge. This can lead to more refined, coherent, and contextually relevant outputs.

Adaptive Summarization: Traditional summarization techniques, even those using deep learning, often focus on extracting or rephrasing existing content. The adversarial nature of GANs, when applied to this task, can push the boundaries of what's possible, leading to the generation of summaries that are not just extracts but possibly innovative reinterpretations of the content.

Generating Ideas: While GANs have been used for generating images, music, and even text, the specific application of using them, combined with encoder-decoder architectures, to generate "ideas" is a forward-thinking approach. This can open doors to automated brainstorming or concept generation, which could be transformative in fields like content creation, research, and design.

Domain Adaptability: The potential to fine-tune these pretrained models on specific domains offers a level of customization that can cater to diverse fields and niches, making the approach versatile and widely applicable.

11.Societal & Market Impact

Social Impact:

Education: Automated summarization tools can make academic materials more accessible, allowing students to quickly grasp the essence of lengthy texts. This can be especially beneficial for students with learning disabilities or those who are overwhelmed with vast amounts of information.

Information Overload: In an age of digital information overload, having tools that can distill vast amounts of content into concise summaries can help individuals stay informed without feeling overwhelmed.

Language Barriers: Such models can potentially be used in tandem with translation systems, allowing content to be summarized and then translated, making knowledge more universally accessible.

Creativity Boost: Automated idea generation can serve as a brainstorming companion, aiding individuals in creative fields to break through blocks and find inspiration.

Bias and Misinformation: Like all AI models, GANs can unintentionally perpetuate biases present in their training data. If used widely for summarization and idea generation, there's a risk of reinforcing or spreading biased perspectives. This highlights the importance of ethically curating and auditing training datasets.

Market Impact:

Content Creation: Writers, journalists, and other content creators can leverage these tools for research, drafting, and brainstorming, potentially speeding up the content creation process.

Business Intelligence: Companies can use such tools to quickly summarize lengthy reports, market research, or competitor information, enabling faster decision-making processes.

R&D Innovation: In sectors like pharmaceuticals, technology, or design, automated idea generation can lead to novel solutions and products, potentially revolutionizing R&D processes.

Competitive Edge for AI Companies: Companies that successfully develop and deploy encoder-decoder pretrained GANs for summarization and idea generation will likely gain a significant competitive advantage, opening up new revenue streams.

Job Market Dynamics: On one hand, such automation tools might reduce the need for certain manual roles, such as basic content summarization. On the other hand, they could create opportunities for roles focused on model training, fine-tuning, and ethical oversight.

Customization and Niche Markets: The ability to fine-tune these models on specific domains can lead to a plethora of specialized applications. For instance, legal firms might use specialized versions for summarizing case law, while medical institutions might use them for distilling lengthy research papers.

12. Challenges and Risks

Challenges:

Quality Assurance: Ensuring the consistent generation of high-quality summaries and ideas is non-trivial. The outputs might sometimes miss key points or generate content that doesn't truly reflect the essence of the source material.

Training Data: Acquiring a diverse and comprehensive dataset that's representative of the vast range of content for which summarization or idea generation is needed can be difficult.

Computational Costs: GANs, especially sophisticated architectures like encoder-decoder models, can be computationally intensive, requiring significant resources for training and fine-tuning.

Model Stability: GANs are known for training instability. Achieving convergence and ensuring that the generator and discriminator reach a stable equilibrium can be challenging.

Domain Adaptation: While the models might be pretrained, fine-tuning them for specific niches or domains might not always yield satisfactory results.

Interpretable Outputs: Understanding why a particular summary or idea was generated can be challenging, given the black-box nature of deep neural networks.

Risks:

Bias and Ethical Concerns: If the training data contains biases, the generated summaries and ideas might perpetuate or even amplify those biases. This can lead to misrepresentations and potentially harmful outcomes.

Over-reliance: A heavy dependence on automated tools for summarization and idea generation might diminish critical thinking and analytical skills in users.

Intellectual Property Concerns: If the GAN generates ideas that closely resemble existing content, it could lead to copyright or patent infringements.

Misinformation: There's a risk of generating summaries that misrepresent the original content, leading to misinformation or misinterpretation.

Economic Impact: Over-automation in content generation and summarization sectors might result in job displacements.

Security Concerns: Like all AI models, GANs can be vulnerable to adversarial attacks. Malicious actors might exploit these vulnerabilities to manipulate generated content.

Validation and Trust: For critical applications, validating the accuracy and trustworthiness of automatically generated summaries or ideas becomes paramount. Building this trust, especially in sectors that traditionally rely on human expertise, can be challenging.

13.Relevant References

1. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," in Advances in Neural Information Processing Systems, vol. 3, 2014, [Online]. Available: https://doi.org/10.1145/3422622.
2. L. Liu, Y. Lu, M. Yang, Q. Qu, J. Zhu, and H. Li, "Generative Adversarial Network for Abstractive Text Summarization," in AAAI, vol. 32, no. 1, Apr. 2018.
3. F. Carrara, G. Amato, L. Brombin, F. Falchi, and C. Gennaro, "Combining GANs and AutoEncoders for efficient anomaly detection," in 2020 25th International Conference on Pattern Recognition (ICPR), 2021, pp. 3939-3946. [Online]. Available: https://doi.org/10.1109/ICPR48806.2021.9412253.
4. A. Rush, S. Chopra, and J. Weston, "A Neural Attention Model for Abstractive Sentence Summarization," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015, pp. 379-389.
5. I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in Advances in Neural Information Processing Systems, 2014, pp. 3104-3112.
6. A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, 2015.
7. C. Y. Lin, "ROUGE: A Package for Automatic Evaluation of Summaries," in Text Summarization Branches Out: Proceedings of the ACL-04 Workshop, 2004, pp. 74-81.
8. K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "BLEU: a method for automatic evaluation of machine translation," in Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, 2002, pp. 311-318.
9. A. Lavie and A. Agarwal, "METEOR: An Automatic Metric for MT Evaluation with High Levels of Correlation with Human Judgments," in Proceedings of the Second Workshop on Statistical Machine Translation, 2007, pp. 228-231.
10. L. Rabiner and B.-H. Juang, Fundamentals of Speech Recognition, Prentice-Hall, Inc., 1993.