YET ANOTHER TEXT-TO-TEXT GENERATION MODEL USING CHARACTER-BASED RECURRENT NEURAL NETWORK

Haris Ijaz
MSEE
319538

Table of Content

- Problem Statement
- Related Work
- Dataset
- Model Architecture
- Model Training
- Text Generation
- Evaluation
- Comparative Analysis
- Conclusion
- Future Work

Problem Statement

- Sentence creation is basic skill of a human brain
- If machine replicate this ability it will solve text generation problems
- Text Generation is very important task in NLG
- Used in Machine Translation, Text Summarization, Chatbots and Virtual Assistants
- Long term dependency problem in character-based text generation models
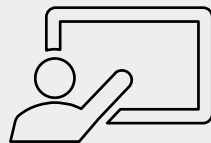
Related Work

- Yang, et al. (2020) [15] proposed a generative adversarial network (GAN) based model for text generation.

- Dialogue generation system is build using RNN by Wen et al (2015)

- Lifeng et al.(2015) [12] proposed RNN based model in encoder-decoder framework to generate response for conversation system.

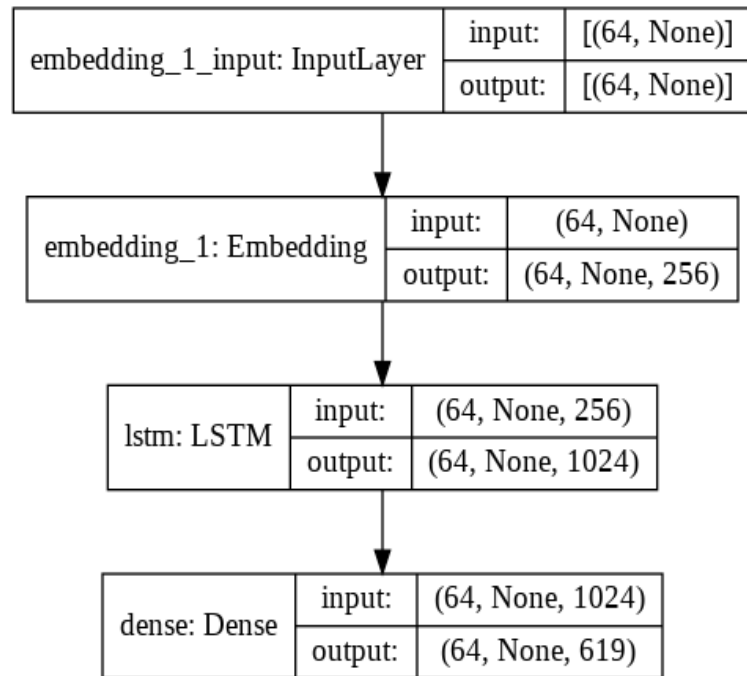- Ziyao, S. 2019 used LSTM model to generate Chinese text

Dataset

- Wikipedia Dataset is used

- Wikidump software is used

- Total 6,033,151 Articles

- Dataset Size is 33 GiB

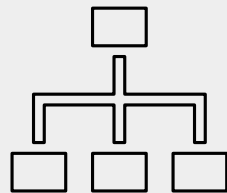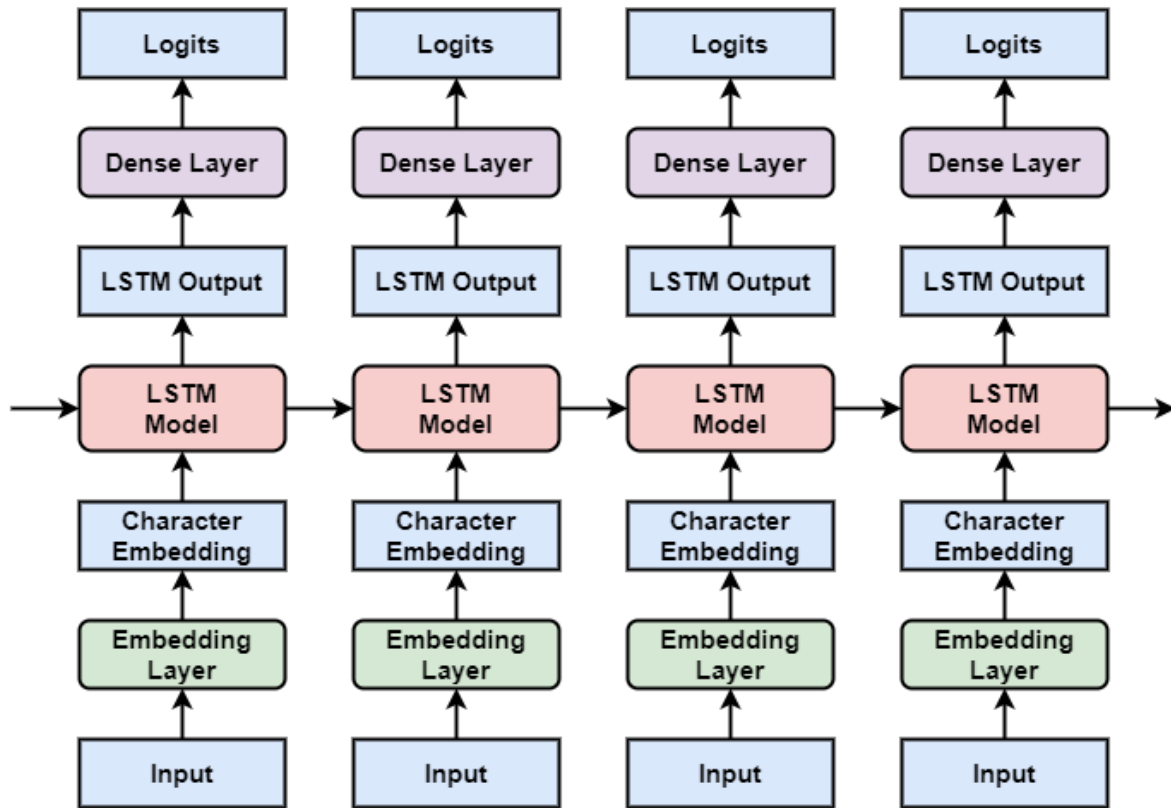- Dataset is converted into tfds format

Proposed Model Architecture
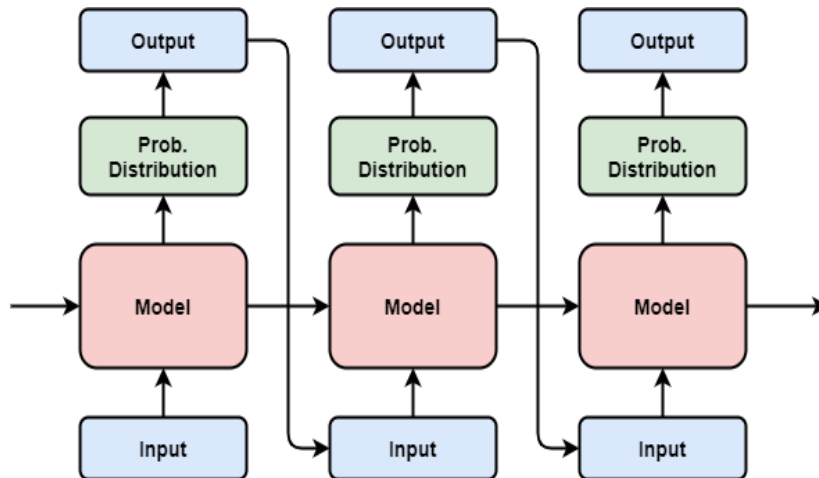
Use tf.keras.Sequential to define the model.

•tf.keras.layers.Embedding: The input layer. A trainable lookup table that will maps

• tf.keras.layers.LSTM: A type of RNN with size unit

•tf.keras.layers.Dense: The output layer, with vocab-size outputs.

embedding_1_input: InputLayer	input:	[(64, None)]
	output:	[(64, None)]

embedding_1: Embedding	input:	(64, None)
	output:	(64, None, 256)

lstm: LSTM	input:	(64, None, 256)
	output:	(64, None, 1024)

dense: Dense	input:	(64, None, 1024)
	output:	(64, None, 619)

# Model Training

Text Generation

| Input	Text Generation
At the beginning	At the beginning of the first season of the population of the state
Science is	Science is the constitution of the same controlled in the company of the control of the property of the store
Event	Eventer of the Armenian politician and a construction
Up to date	Up to date and the state of the control of the station of the

Evaluation

○ Model is evaluated on two metrics

○ BLEU

○ BERTScore

Temperature	BERTScore			BLEU
	P	R	F	
0.2	0.8609	0.8629	0.8619	0.74
0.4	0.6862	0.8652	0.8617	0.66
0.6	0.8438	0.8600	0.8518	0.66
0.8	0.8193	0.8443	0.8316	0.56
1.0	0.8010	0.8102	0.8204	0.54
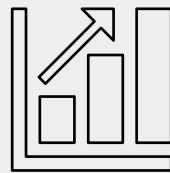
Comparative Analysis

Paper	BERTScore			BLEU
	P	R	F	
This Paper	0.8609	0.8629	0.8619	0.74
Ziyao, S. 2019	0.8690	0.859	0.8643	0.56
HAI,T. et al 2017	0.8339	0.8476	0.8407	0
Supha, A. et al.	0.8319	0.8559	0.8437	0.56
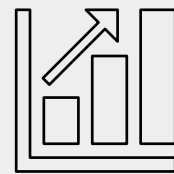
Conclusion

- LSTM Character Based Model is proposed

- Wikipedia Dataset is used

- Model is implemented on tensorflow framework

- CrossEntropy is used as loss function

- Adam Optimizer is used to optimize the model

- BLEU and BERTScore is used as evaluation metrics
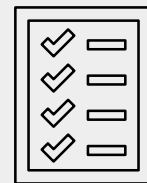
Future Work

- Fine Tune the hyperparameter

- Used more computing resources

- Improve Pre-processing on dataset