Space Movie (Star Wars) Script Generator Using LSTM in Tensorflow

Randy Zakya Suchrady

Informatics Engineering, Institut Teknologi Bandung
Bumi Mutiara JD 1/1, Desa Bojong Kulur, Kec. Gn. Putri, Kab. Bogor, Jawa Barat, Indonesia
[1]13518061@std.stei.itb.ac.id

***Abstract*— In the modern days, the world runs fast with the help of technology. The current popular technology is artificial intelligence where computers can mimic human intelligence in certain tasks. There are some limits that computers can do, for example achieving creativity. There are some artificial intelligence that can be considered as creative, such as DALL-E 2 created by Open AI. Although it is considered creative, there are some flaws created by the AI. In this paper, a machine learning model will be given a task to generate a movie script referencing the Star Wars franchise as its training data. After the training, the model will be evaluated from a human's perspective, wether the model is considered as creative or not.**

***Keywords*— technology, computers, creativity, artificial intelligence, machine learning**

I. Introduction

A. Background

Filmography is one of the most popular types of creative industry. In creating a movie, there are several elements that should be prepared including the sets, actors and actresses, crews, and scripts that are prepared by the screenwriter. Creativity is a unique value proposition from human beings where so far there are no other beings that can achieve our creativity. In modern days, many human activities are assisted by artificial intelligence, including the creative industry. Script writing is one of the tasks that may be solved by artificial intelligence in the natural language processing domain. It is correct that artificial intelligence may not reach our level of creativity, but we are able to train a model with variants of data that the model might look like "creative" in their own way. In this paper, a machine learning model will be given a task to generate a script by training them with a dataset of another movie's script.

B. Questions of the Problems

1. Can an artificial intelligence agent replicate a human in writing scripts?
2. How can an artificial intelligence agent (machine learning model) solve script generating problems?

C. Objectives

1. To know if artificial intelligence can replicate human's creativity

D. Scope

This paper's scope only covers the definition of human creativity in the field of filmography especially writing scripts and also the theme of the movie is specified to a space themed movie (in this case, Star Wars).

II. Theoretical Study

A. Screenwriting

Screenwriting (or scriptwriting) is the process of writing stories in the screenplay medium, https://www.studiobinder.com/blog/what-is-script-writing/ . There are several elements in screenwriting such as movements, actions, expressions, and of course dialogues. In screenwriting, the script should describe well how those elements are delivered to the audience. The script also must be understandable by the actors and actresses, so they can perform their best act.

B. Natural Language Processing

Natural language processing is a popular problem domain in machine learning. The idea is to solve a certain task in the field of linguistics (writing, speaking, etc.) by creating a machine learning model. There are several common problems in neutral language processing such as sentiment analysis where a model can identify a certain text and what sentiment that is depicted from the text. There are also other types of problems including name entity recognition, emotion detection, question answering, etc.

C. Recurrent Neural Network

In machine learning, a type of machine learning model that is popular in solving natural language processing is the recurrent neural network. Recurrent neural network is a type of neural network model that considers data that have a sequential form. In identifying texts or other forms of natural language processing data, it is often that the sequence of the words must be calculated. In a normal neural network (feed forward neural network), the model is not considering the sequence of the model, thus it is lacking in solving sequential data.

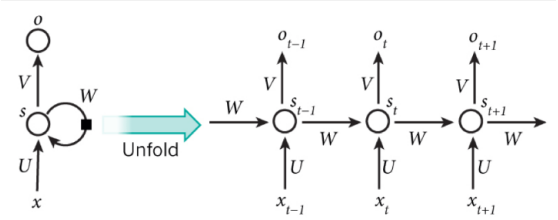

Fig. 1 Representation of recurrent neural network, Denny Britz, 2015

D. *Long Short Term Memory*

Long short term memory is one of the forms of recurrent neural network types of model. Long short term memory units are more sophisticated than the regular recurrent neural network unit, it handles remembering the past word that is fed into the unit. In the LSTM unit, the model calculates using functions such as sigmoid, tanh, and operate weight multiplications to solve recurring problems.

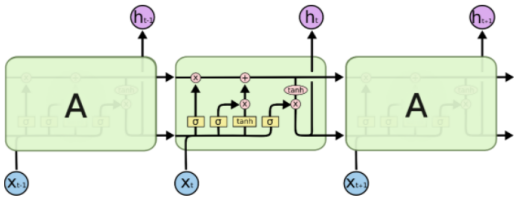

Fig. 2 Representation of LSTM layers

III. Methods

In creating a machine learning model, python and tensorflow module will be used. There are several steps in creating a natural language processing machine learning model.

A. *Data Collection*

Data collection is the step of gathering data for the model to train with. In this case, the dataset is taken from kaggle with the title "Star Wars movie Scripts" uploaded by Xavier (available at: https://www.kaggle.com/datasets/xvivancos/star-wars-movie-scripts). The data consists of three txt files covering the fourth, fifth, and sixth episode of the Star Wars franchise's movie scripts.

B. *Data Preprocessing*

Before training the model with the data, firstly the data must be clean. There are several steps to clean the data as follows:

1. Case Folding
   Case folding is a step of preprocessing where each of the letters in the text are converted to lowercase letters.
2. Trim and remove multiple spaces
   Sometimes there are texts that contain multiple spaces and a redundant space on the head of the text or trailing the text.
3. Drop empty strings
   After the two previous steps, sometimes there are texts that contain nothing.
4. Remove duplicates
   Removing duplicates may create a lighter dataset.
5. Append end of line token
   Because in this case the topic is about script generation, where there are changes of lines, so it must be decided when to end a line of dialogue. To achieve that each text is appended by an end of line token "[EOL]".

C. *Data Transformation*

Computers can not calculate using raw text, to make the model able to digest the data, the data should be transformed into a form that can be calculated, that is numbers. The words are tokenized using the tensorflow text Tokenizer.

D. *Modeling*

The model is consist of three main layers and structured as follows:

1. Embedding layer
   Embedding layer captures some of the semantics of the input by placing semantically similar inputs close together in the embedding space.
2. Bidirectional LSTM layer
   The LSTM layer consists of 128 units that consider the past word and also predict the future words.
3. Dense layer
   Output layer using softmax activation function. The output results are arrays of one hot encoded label.

For the optimizer, the model uses Adam optimizer with 0.01 learning rate and 20 epochs. To measure the performance, the model is compiled with the categorical cross entropy loss and accuracy as its metric.

IV. Result

The result for the model trained with the Star Wars scripts only achieved around 45% accuracy for 20 epochs. Previously the model was trained for 100 epochs but after the model went through 20 epochs, there are no signs of improvement.

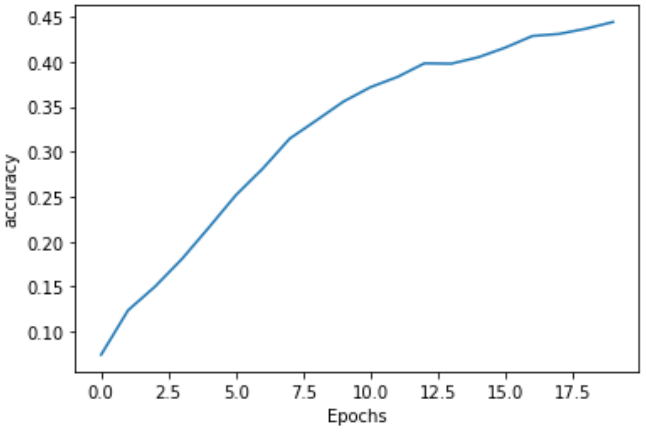

Fig. 3 Accuracy of the model during the training phase

After the first attempt of creating the model, it is decided that creating a second model with a lighter dataset may result in a better outcome. The second model is structured the same way as the first model, but the dataset that is fed to the model only consists of the fourth episode of the Star Wars movie. The result is a lot better, the model reached 64% training accuracy.

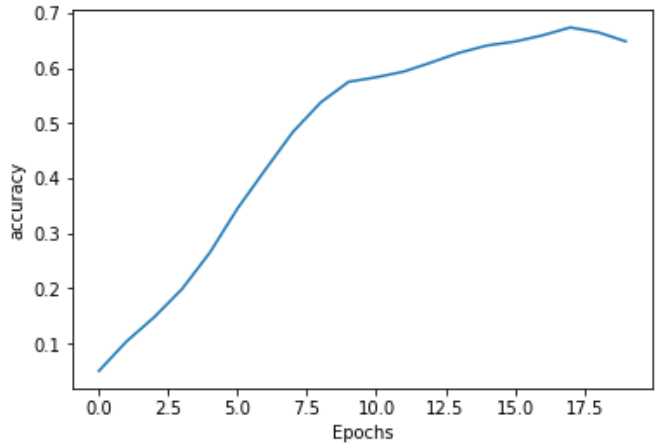

Fig. 4 Accuracy of the second model during the training phase

To make a script, it is needed for more than one sentence, there should be a dialogue. To make the script generator work, the model should predict what is the next dialogue. The idea is to make the first word in the next row as the label of the current row. Then that first word will be the seed text for the next dialogue. The model for creating the next line is structured as follows:

1. Embedding layer
2. One dimensional convolution layer
3. Global max pooling layer
4. ReLU dense layer
5. Softmax dense layer

Here is an example of the result script that is generated by the model.

```
thousand we don't serve their kind here better of any use the comlink oh the
empire to draft me for a thousand generations the planet is operational we are meant
on the exhaust port are you in the galaxy will be the only one we
will soon be back and in greater numbers for the rest of the season you
should have been now and my father die i found those plans and bring the
end of the empire to draft me into service the rebellion is spreading and i
can't shake him but i can't shake him i can't shake

shows well if we can go on him right away has seen luke take that
r2 unit we bought might soon be with this planet to alderaan one day whammo
you're going to make us go off in escape for the ease of yourself you'll
always be the boy in the force as i know what is that he has
in range in range for a good pilot like you to be my guest at
a ceremony that will make this battle station is operational no star system will dare
oppose the emperor now the data tapes of a person of

nice come on top of we're all you you big how you can repair him
can't you go but now i can trust him he's our extra he needs you
one long before we can just avoid any more female advice we ought to be
able to get out of here get your commission for the ease of yourself and
then you can go to the academy next year with it chewie but i can
do them if he gets himself killed come on on this year we'll make enough
on the harvest so he might be able to penetrate the
```

Fig. 5 Generated Star Wars script

V. Conclusions

With simple bidirectional LSTM, the model still can not handle script generation problems well, there are many flaws created by the model. The sentence frequently is not clear and doesn't have any meaning. This may be caused by the small dataset size and also the simple structure of the model. To access the project repository, go through this url: https://github.com/rdyzakya/StarWarsScriptGenerator

Acknowledgment

Thanks to Mr. Windy Gambetta as the lecturer in the Socio and Informatics subject. Also thanks to Mrs. Fariska Ruskanda as the supervisor in the GEMASTIK competition who helped me to learn new things, especially natural language processing. I thank Mr. Nugraha Priya Utama and Mrs. Nur Ulfa Maulidevi for giving lectures in the Artificial Intelligence class and Machine Learning class.

References

[1] Studiobinder. (2019) What is Script Writing? The Essentials to Writing Great Screenplays. [Online]. Available: https://www.studiobinder.com/blog/what-is-script-writing/

[2] Aditya Yanuar. (2018) Recurrent Neural Network (RNN). [Online]. Available: https://machinelearning.mipa.ugm.ac.id/2018/07/01/recurrent-neural-network-rnn/

[3] Christopher Olah. (2015) Understanding LSTM Networks. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/