COMPARISON AND GENERATION OF A POEM IN ARABIC LANGUAGE USING THE LSTM, BILSTM AND GRU

Abdelhamid. Atassi, LARI Laboratory, Ibn Tofail University
Ikram El Azami, LARI Laboratory, Ibn Tofail University

ABSTRACT

In this paper, we will demonstrate how we can automatically generate a new poem in Arabic language from a dataset containing a number of poems written by renowned Arabic poets and writers based on classical Arabic language and from colloquial Arabic. A comparison was also made between the different LSTM (Long Short-Term Memory), BiLSTM (Bi-directional Long Short-Term Memory) and GRU (Gated recurrent unit) generation networks in order to obtain a satisfactory result, keeping in mind the main one. The difference between a text in Arabic language and another in English or French is the orientation of the text, which is from right to left. For this reason, we tried to use the Bi-LSTM to have the added value of this network, but the latter did not lead to correct results neither at the sentence level nor at the word level, contrary to the GRU to give good results in terms of performance.

Keywords: RNN, Convolutional Neural Network, CNN, Gated Recurrent Unit, GRU, SemEval, Bi-directional Long Short-Term Memory, BiLSTM, Keras, convnets, William Shakespeare, Kaggle, Early Stoping, Over Fitting, NLP, BERT.

INTRODUCTION

This paper investigates one of the deep learning models that can process text as a sequence of words or characters. The two deep learning algorithms for sequence processing are Recurrent Neural Networks (RNNs) and 1D convNets which is the one-dimensional version of 2D convNets, which we will not cover in this paper. However, we will discuss three variants of recurrent neural networks which are: LSTM, Bi-LSTM and GRU, respectively.

Concerning the fields of application, we can find: The classification of documents - such as the identification of the subjects of an article or the author of a book - and the classification of time series such as the estimation of the degree of correlation between two documents or two stock charts.

In addition, the lack of text generation work in different languages and especially in Arabic motivated us to carry out this process in order to have an idea of the results received by the different networks.

Indeed, numerous types of work have been done such as the automatic generation of a new painting from Picasso's paintings (Bourached et al., 2019), or the automatic generation of a new musical piece from existing pieces. (Huang et al., 2016). While there is a scarcity of work devoted to the automatic generation of text from the works of Shakespeare for which the results obtained, so far, show that they are effective only from the form of point of view while the generated text is not meaningfully correct and the sentences are not comprehensible or consistent. However, deep learning has been shown to perform well in sentiment analysis, document classification, and detection of entities in text and document summary as well as the use of chat bots.

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.

Data Preparation

Our work is mainly based on a dataset of Arabic poems (Fahd, 2019) downloaded from the Kaggle web platform organizing competitions in data science. Although the dataset comprises of several columns such as (poem_id, poem_link, poem_style, poem_text, poem_title, poet_cat_ poet_id, poet_link, poet_name), the column we are concerned with most is poem_text, which contains the text of poems in Arabic by different famous poets and writers like “Al Mutamid ibn Abbad”, “Abu al-Hindi” and others.

It is believed that there are two reasons why authors choose the works of William Shakespeare. First, it is a large body of text, and second, it has a particular writing style. Moreover, it is generally recommended to have at least a source of one million characters in total to achieve realistic and optimal text generation.

On the other hand, our dataset contains more than 44 254 342 characters, which is highly sufficient in our case, especially after having combined the different texts of the poems into a single one and created a vocabulary of characters from all the characters text that is 152 characters long, see fig 1.

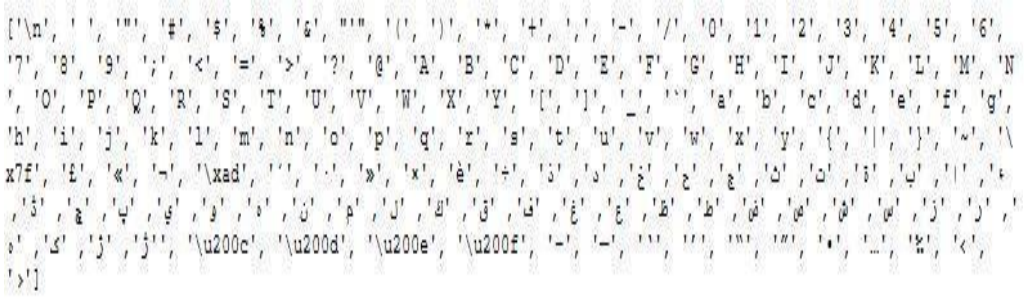

FIGURE 1
THE VOCABULARY TABLE GENERATED FROM THE TEXT

The text is vectorized and each character has been replaced by its index as it is entered in the vocabulary.

The Batch

Essentially our objective is to obtain the model that predicts the next most likely character, given a historical sequence of characters. The choice of the length of this historical sequence relies on the text itself, while using a relatively too short sequence will not produce us enough information.

For example, considering the letter "a", if we look for the following character, the sequence will be too long and the training will take too long and may over-train (being at the risk of obtaining a sequence of characters that are impertinent for more distant characters). While there is no correct choice of sequence length, we need to take the text itself into consideration, as well as the length of the normal sentences it contains, and have a reasonable idea of which characters or words are relevant to one another. In our framework, we take the average of the character lengths of the poems which is 761.72, and try to find the multiplication of the number of characters of the closest vocabulary so as to have better results, at which we attained 760 characters.

The actual text data will be the sequence of text shifted forward one character, see fig 2.

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.

FIGURE 2
SAMPLE OF A TEXT SEQUENCE SHIFTED FORWARD ONE CHARACTER

Subsequently, we shift the individual character calls to sequences that can be batch-fed, we use seq_len + 1 due to zero indexing.

Once we have our sequences, we will proceed to perform the following steps for each of these to create our target text sequences, see fig 3:

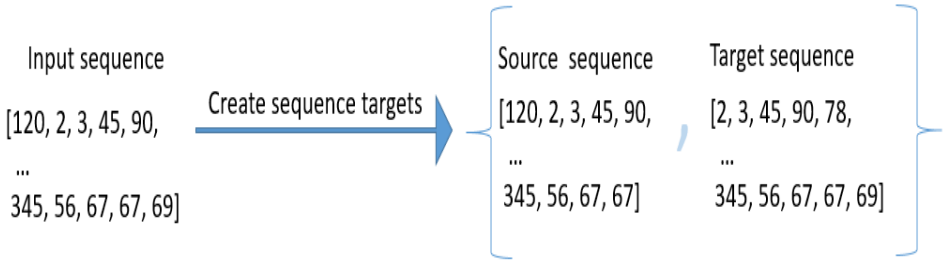

FIGURE 3
SAMPLE OF A SEQUENCE OF TEXT SHIFTED FORWARD ONE CHARACTER IN NUMERIC FORM

After getting the actual sequences, we will create the batches. We would like to shuffle these sequences in random order, so that the template does not fit into any section of the text, but rather, we will be able to generate characters from any seed text, see fig 4.

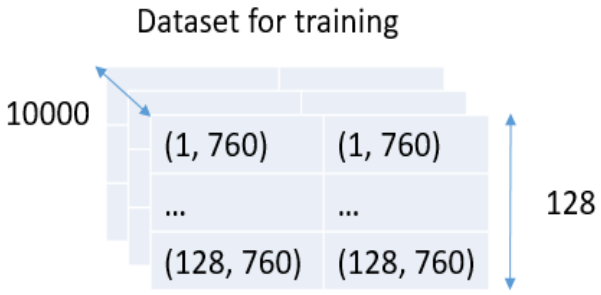

FIGURE 4
SAMPLE OF A BATCH DATASET

Architecture

Machine learning and deep learning algorithms are the up-and-coming approaches to solving prediction problems in time series. These techniques have been proved to produce more accurate results than conventional regression-based modeling.

As a matter of fact, it has been reported that artificial recurrent neural networks (RNNs) with memory such as long-term memory (LSTM) (Hochreiter et al., 1997) which are models based on the LSTM incorporate gates to store longer sequences of input data. However, the

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.

main question is whether the gates incorporated into the LSTM architecture already provide convenient prediction and whether further training of the data would be needed to further improve the prediction.

In contrast, bidirectional LSTMs (BiLSTMs) allow additional learning by scanning the input data in two steps (first from left to right, then from right to left). The results reveal that the additional training of the data, and therefore the modeling based on BiLSTM (Siami-Namini et al., 2019) offers better predictions than classical models based on the LSTM. Precisely, it has been observed that the BiLSTM models provide better predictions compared to the LSTM model. Not to mention that it has also been recorded that BiLSTM models reach equilibrium much more slowly than models based on LSTM.

GRU (Chung et al., 2014) is related to LSTM since both models use gate information in different ways to avoid the disappearance of the gradient problem. GRUs train faster and perform more efficiently than LSTMs on less training data if you carry out language modeling (but not other tasks). LSTMs should theoretically memorize longer sequences than GRUs and surpass them in tasks requiring modeling of long-distance relationships (Yin et al., 2017).

With the use of the Early Stopping function (Raskutti et al., 2011), the Keras library will stop training when a monitored metric has ceased to improve, assuming the goal of training is to minimize loss. With that, the metric to watch would be “loss”, and the mode would be “min”. A model. Fit () learning loop will verify at the end of each epochs if the loss no longer decreases, taking into account the min_delta and the patience if necessary (In this case, the value of the patience parameter is 1 by experience) . By the time it stops decreasing, model.stop_training is marked “True” and the training ends.

The LSTM Model

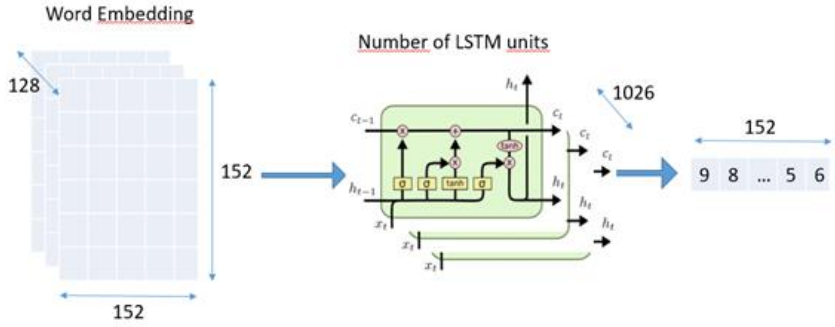

FIGURE 5
SAMPLE OF THE LSTM MODEL WITH 1026 NEURONS

The BiLSTM Model

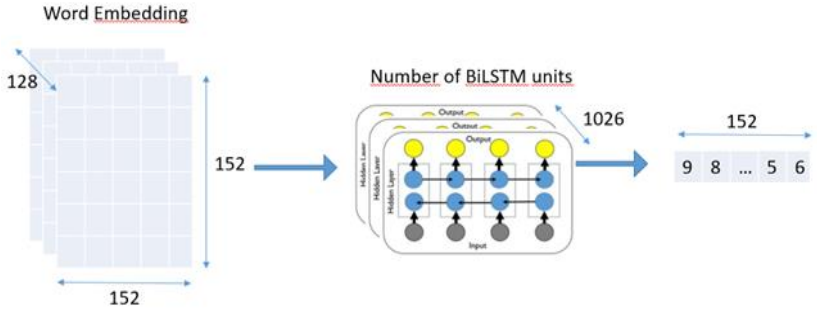

FIGURE 6
SAMPLE OF THE BILSTM MODEL WITH 1026 NEURONS

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.

The GRU Model

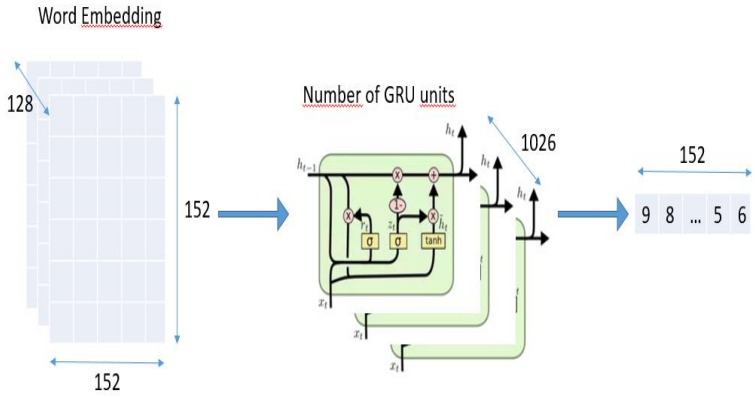

FIGURE 7
EXAMPLE OF THE GRU MODEL WITH 1026 NEURONS

Evaluation

The evaluation was performed with a DELL XPS i9-9980HK computer, 32GB of RAM, graphics card Model: NVIDIA GTX 1650, 4GB, 896 CUDA cores and a separate graphics card in a Razer Core X case, model: NVIDIA GeForce GTX TITAN X, 12GB, 3072 CUDA Cores. The most significant element for us is the number of CUDA Cores, which are required to be used by Tensor flow when processing.

After training the three models on the same dataset, with the same parameters as well as using the early stopping function of the Keras library to save time, we obtain the following results:

Regarding LSTM, see fig 5.

The time that each epoch takes is between 626s and 633s and starts with an error rate of 27.353% and can reach 17.029%, see fig 8.`

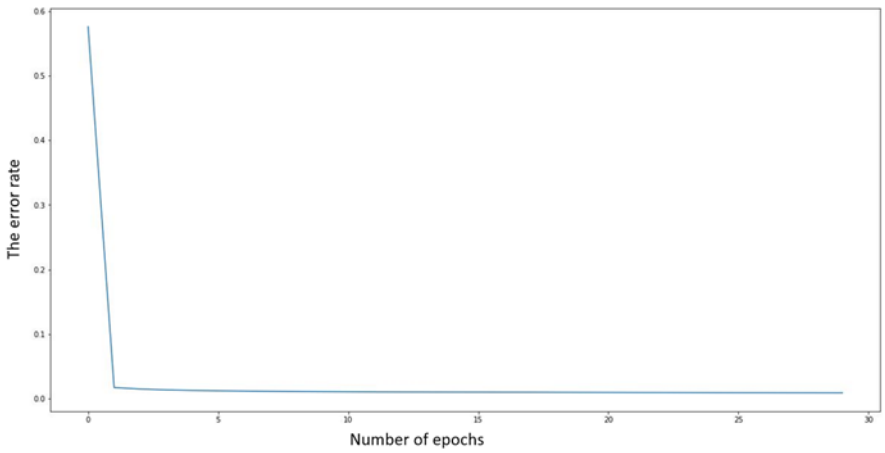

FIGURE 8
THE DECREASE IN THE ERROR RATE WITH THE USE OF THE BILSTM MODEL

The results obtained after the use of LSTM model to generate an Arabic text:

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.

ساحر واحمق اقصاري القارءه بهذي وجوع من تضع شلالي وبعد ان اعصي يا حقول ساحرف
يا ايها الحزن ارتواء الشرق احلامي يدير الازهار فالمدينة * لي اشداء المصاءد والسلام لا وعد لا
يجد او ندمت بالانسان اسمع شبابي وحدك في ساحة البدء انتظر المر شكله حتي رمثت افـءده القلب
من الامال والصمت اقطف علي صدرك الغاضب قدماي وقبض الغاءب في حقاءبه لا تركتني منك
الذي يذهله * حالي عن الحسب في زمن وما يغنيك من من ظل رسالة الحيوان تموز في فتورك انك
في الصعصار سيان – يلعب في العمر - كانت مثل سابقه الشعر لامراه تطل بسبابة وءوسعي شيمتك
الاسود المحزين يناطون عيونا طويلا لاعب الرشفة لا تفصلنا فما يعود الي بلاد في اليالي القصيده
في موحده مولوده وينام علي اختلاج نيران النهار الشعر تشكو يا ويحفظ الطعنات ويسمع حنا ويغسل
الكف شطرين الحزن حافية تعارضه السراب واين ماسين حالمني اصبحت تاسيني ه

FIGURE 9
TEXT GENERATED USING THE LSTM MODEL

We find out that there are correctly written words and sometimes, in addition to the form, there is even a meaning in the text, see fig 9.

Regarding BiLSTM, see fig 6.

The time that each epoch takes is between 252s and 248s and starts with an error rate of 15.162% and attains 00.088%, see fig 10.

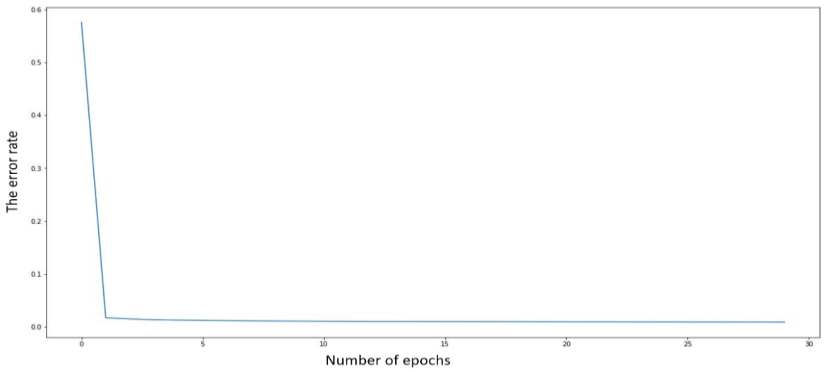

FIGURE 10
THE DECREASE IN THE ERROR RATE WITH THE USE OF THE BILSTM MODEL

After we performed a regeneration of the text, we observe that there is a repetition of a character or a series of characters, which implies that the BiLSTM model has passed to overfit ting, see fig 11.

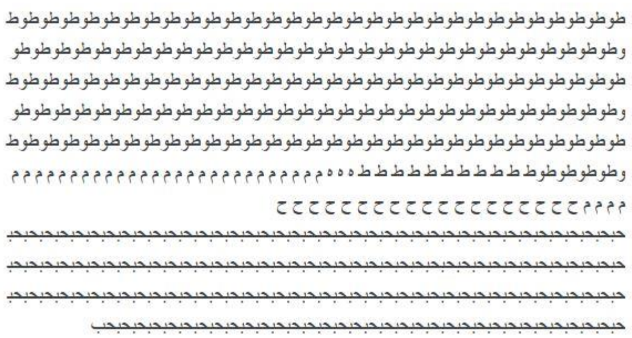

FIGURE 11
TEXT GENERATED USING THE LSTM MODEL

Concerning GRU see fig 7.

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.

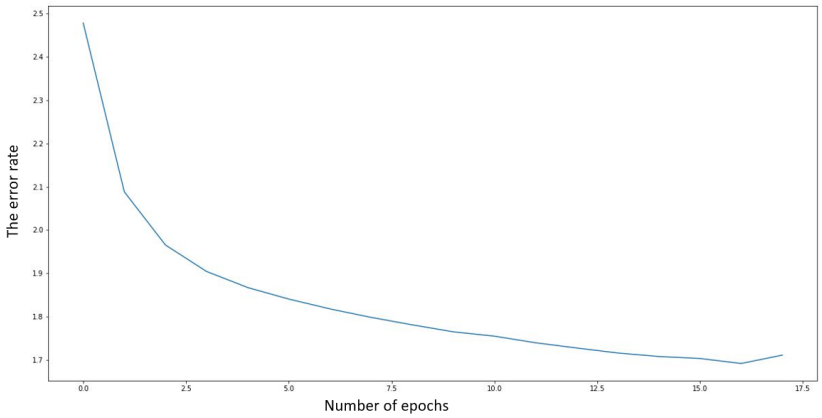

FIGURE 12
THE DECREASE IN THE ERROR RATE WITH THE USE OF THE GRU MODEL

The time that each epoch takes is between 239s and 240s and starts with an error rate of 28.412% and ends at 17.099%, see fig 11.

العشق انه صرخات في سمع الدهشة يشكو لاعم الجبليني قلنا استوردتهم سلم لعلكم قدمتم لم تستقبلين
اغنية وشنياك ما هذا الازمن شجر شعرك ينمو زيارته فيراد الماء الثاني حين خوارك يا انت البداية
عندما انتسبوا صلاه امنطق هل ثوره احبك في الظلام ستظل كلماتي اني وتسرب من جسد حتي اخر
هذا شاعر لا قتال لا يميع 31 ايها الرقصات لو انك تدري اين وقلت ارخصت غلالة تركتني والطهر
قالا باني الفت مشت في كل الامسيات وبتور ينصح بالشرخ الحر ومن حاك بليل ضاع غفلته ما
اخفتنا بارعاش الخطباء لسهل حكايتنا لارتعاش ونحن وفينا من فلكوكنة وذا بعد من عين مفتوران
لفجر واليل والجندان هبطت بنا قتل يوم اجمل من صوت رمانة روح بمنتهي الاسرن ومتفياد وكنا
علي نسق كل حب لا يموت قتل ورد فقولي لهم هذا الياس اتعناهم يدك انا حزن اخي نهديك فوقنا قبل
ان اترك العينين بلا ديدك انا (5) هاجمني دمعي , وانا اب

FIGURE 13
TEXT GENERATED USING THE GRU MODEL

We see that there are correctly written words and sometimes there is even a meaning of the text as in the case of the LSTM model, see fig 12.

Figure 14
A COMPARISON BETWEEN LSTM AND GRU MODELS

	Time (s)		Error rate (%)	
	1st epochs	Last epochs	1st epochs	Last epochs
LSTM	626	633	27.353	17.029
GRU	239	240	28.412	17.099

We notice that the GRU is more efficient than the LSTM model in terms of the duration of the processing of each epoch (approximately 3 times faster), see fig 13.

CONCLUSION AND FUTURE IMPLICATIONS

Following the data results obtained, we can conclude that the BiLSTM model (Siami-Namini et al., 2019) falls into over fitting despite processing texts from right to left. Nevertheless, the GRU model has proved to be more efficient than the LSTM model.

Besides, recent advancements in NLP have shown that transfer learning assists in achieving more remarkable results for new tasks by adjusting pre-trained models instead of starting from scratch. The Transformers have made a significant improvement by achieving new

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.

cutting edge results for many NLP tasks including, but not limited to, text classification, text generation, and sequence labeling. Most of these accomplishments actually rely on large data sets table 1.

Yet, it is questionable whether GRU models can achieve significantly higher results than a BERT model (Devlin et al., 2019) for a small data set, and that these simple models are trained in much less time than the BERT model (Devlin et al., 2019), adjusting their pre-trained counterparts. This subject will be the highlight of a new work in a prospective paper.

REFERENCES

Bourached, A., & Cann, G. (2019). Raiders of the lost art.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networkson sequence modeling.

Devlin, J., Chang, M., Lee, K., Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding.

Fahd. (2019). Arabic poetry dataset.

Huang, A., & Wu, R. (2016). Deep learning for music.

Hochreiter, S., & Schmidhuber, J. (1997). Long short term memory.

Raskutti, G., Wainwright, M.J., & Yu, B. (2011). Early stopping for non-parametric regression: An optimal data-dependent stopping rule.

Siami-Namini, S., Tavakoli, N., & Namin, A.S. (2019). The performance of LSTM and BiLSTM in forecasting time series.

Yin, W. (2017). Comparative study of CNN and RNN for natural language processing.

Received: 30-Nov-2021, Manuscript No. JMIDS-21-8875; **Editor assigned**: 02- Dec -2021, PreQC No. JMIDS-21-8875 (PQ); **Reviewed**: 11-Dec -2021, QC No. JMIDS-21-8875; **Revised**: 17-Dec-2021, Manuscript No. JMIDS-21-8875 (R); **Published**: 05-Jan-2022.

Citation Information: Atassi, A., & Azami, I.E. (2022). Comparison and generation of a poem in Arabic language using the LSTM, BILSTM and GRU. *Journal of Management Information and Decision Sciences, 25*(S2), 1-8.