Received on (15-03-2022) Accepted on (15-08-2022)

Generating Attractive Advertisement Text Campaigns Using Deep Neural Networks

Atef Ahmed, Motaz Saad, and Basem Alijla
https://doi.org/10.33976/JERT.9.2/2022/2

Abstract—

Text generation task has drawn an increasing attention in the recent years. Recurrent Neural Networks (RNN) achieved great results in this task. There are several parameters and factors that may affect the performance of the recurrent neural networks, that is why text generation is a challenging task, and requires a lot of tuning. This study investigates the impact of three factors that affect the quality of generated text: 1) data source and domain, 2) RNN architecture, 3) named Entities normalization. We conduct several experiments using different RNN architectures (LSTM and GRU), and different datasets (Hulu and booking). Evaluating generated texts is a challenging task. There is no perfect metric judge the quality and the correctness of the generated texts. We use different evaluation metrics to evaluate the performance of the generation models. These metrics include the training loss, the perplexity, the readability, and the relevance of the generated texts. Most of the related works do not consider all these evaluation metrics to evaluate text generation. The results suggest that GRU outperforms LSTM network, and models trained on booking set is better than the ones that trained on Hulu dataset.

Index Terms— Deep learning, Recurrent Neural network, Advertisements campaigns, text generation.

I. INTRODUCTION

Online adverting is the process of marketing and advertising services and products over the internet Motaz . It has attracted the interest of investors and business owners. For instance, 77 % of EU businesses have a website and 26% of them use internet to advertise. In addition, 86 % of EU enterprises used at least one type of social media to build their image and to market their products [2]. The revenue of digital ads was worth $126 billion [3]. A successful advertising campaign is the one that has attractive ads, which are delivered to relevant and interested consumers (audience) with the precise, meaningful, and relevant contents. Generating attractive and successful ad campaign is beneficial and worthwhile, and it is subject to reach target customers at the right time [4]. Institutional advertisers use targeted advertisement method to generate attractive campaigns based on the requirements of advertising exchange system [5]. The very old methods of creating attractive contents of advertisement campaigns are either by hand of content writer or automatically base on fil-in-the blank" templates [6]. However, generating successful advertising campaigns that meet the customer's needs is very challenging, time-consuming and an expensive task. Significant advertising knowledge and good understanding of customers needs is required.

Machine learning and deep learning are successfully used in various applications, including machine translation [7, 8], text summarization [9, 10], text generation [11-15], speech-to-text and text-to-speech [16]. Deep learning has evolved many network architectures such as Recurrent Neural Networks (RNNs) [17], Long Short-Term Memory networks (LSTM) [18], and Gated recurrent Unit networks (GRU) [14]. Recent research showed impressive results of using deep learning techniques in NLP applications such as text generation and text summarization [11].

The work of [12] proposes a novel end-to-end model named to generate the AD post. The authors split the AD post generation task into two subprocesses: (1) select a set of products via the SelectNet (Selection Network). (2) generate a post including selected products via the MGenNet (Multi-Generator Network). Concretely, SelectNet first captures the post topic and the relationship among the products to output the representative products. Then, MGenNet generates the description copywriting of each product. Experiments conducted on a large-scale real-world AD post dataset demonstrate that their proposed model achieves impressive performance in terms of both automatic metrics as well as human evaluations.

The work of [19] proposed explore the possibility of collaboratively learning ad creative refinement via A/B tests of multiple advertisers. For generating new ad text, the authors used an encoder-decoder architecture with copy mechanism, which allows some words from the (inferior) input text to be copied to the output while incorporating new words associated with higher click-through-rate.

In[20], the authors proposed a query-variant advertisement text generation method that aims to generate candidate advertisement texts for different web search queries with various needs based on queries and item keywords. To solve the problem of ignoring low-frequency needs, they proposed a dynamic association mechanism to expand the receptive field based on external knowledge, which can obtain associated words to be added to the input. These associated words can serve as bridges to transfer the ability of the model from the familiar high-frequency words to the unfamiliar low-frequency words. With association, the model can make use of various personalized needs in queries and generate query-variant advertisement texts.

This paper proposes a method of using deep learning models (LSTM and GRU networks) to generate attractive text advertising campaigns that meet customer needs using pre-defined keywords. We investigate the generation of advertisements text campaigns mainly in two domains: hotel Booking and TV streaming (Hulu). In addition, two datasets in the domain mentioned earlier have been acquired and prepared for this research, to train the neural networks to generate attractive ads, based on a given keywords feed as a seed to the neural network. Besides the automatic evaluation metrics (perplexity and readability [21]), human annotators subjectively evaluated the readability and relevance of the generated ads.

The rest of this manuscript is organized as follows. Section II describes the methodology of advertisement text generation including: data acquisition, data Integration, data Pre-processing, Ads generation, and the evaluation. Experimental Studies and evaluation methods are presented in Section III. The discussion and experiments results are presented in Section IV, Finally, Section V presents Summary and Conclusions.

II. DEEP NEURAL NETWORKS TO GENERATE ADVERTISEMENT CAMPAIGNS

خطأ! لم يتم العثور على مصدر المرجع. shows the used methodology in this work. The methodology consists of five main steps: data acquisition, data integration, data pre-processing, text generation and evaluation for generating Advertisement text campaign using recurrent neural networks. These steps are described in detains in the following sub-sections. Although we use deep learning techniques but data preprocessing is needed because the data is noisy as it is collected from internet.

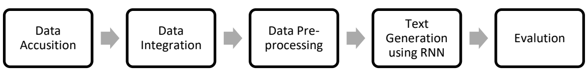

Figure 1: General Five Steps Metodology for Ads Generation

a. DATA ACQUISITION

The data is collected using SEMrush toolkit [22], it provides marketing information such as top ads, keyword analytics, and search tracking, etc.., for a particular website. The SEMrush retrieve and rank the Ads campaigns for the top-ranked website using Google and Bing search engines.

Table 1 describes the main charachteristics of the collected datasets. The collected data is limited to adversisment campagin for hotel and flights reservation, which is collected from Expedia.com and booking.com websites, and TV and movies streaming collected from Hulu.com websites. The data includes 42K text lines (campaigns) from Booking and 13K text lines campaigns from Hulu. The average campaign length is 67.53 and 227.07 for Booking and Hulu datasets respectively. The average number of words per line is 11.13 and 39.15 for Booking and Hulu datasets respectively. It is remarkable that Hulu campaigns length is shorter than Booking campaigns as shown in the table.

Table 1 : The main properties of collected dataset

Datasets	Size	Max Length	Min Length	Average Lengths	Average # of words
Booking	42k	85	19	67.53	11.13
Hulu	13k	368	19	227.07	39.15

b. DATASET INTEGRATION AND PRE-PROCESSING

Datasets were collected from two different sources. So, integrating data in a single and consistent representation is performed. Then the dataset is pre-processed to be suitable to be feeder to the neural networks. خطأ! لم يتم العثور على مصدر المرجع. depicts the main steps of data pre-processing, including data clearing and data normalization, and Name Entity (NE) normalization.

Data cleaning involves the processes of removing and correcting corrupt, unnecessary, or inaccurate records. So unnecessary HTML tags like , , and
 are removed. Moreover, all duplicated records in the data are deleted.

Data normalization involves the process of converting text to lower case and removing special characters and punctuations.

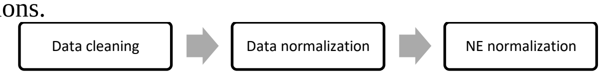

Figure 2: Pre-processing steps

Named Entity refers objects name such as person's name, location's name, and product's names [23]. To further normalize the text, name entities are replaced with tag name using GeoText [24] library and using static-NE list. Geotext [25] is A Python Library used to extract country and city from given text, and it is trained on data taken from geonames.org to recognize cities and countries names for another dataset or text. All cities and countries are replaced by ***i-city*** and ***i-coun*** labels respectively.

geotext may fail to recognize the names of some cities and countries because geotext depends on the training of data. So, static-NE list for cities and countries is proposed to overcome the limitation of geotext. Two lists of 4144 city names and 206 countries are collected from *geonames.org*. Consequently ***i-city*** and ***i-coun*** labels are proposed to replace city name and country name respectively.

III. Experimental Studies and Evaluation Methods

This section presents the proposed methods for advertisement text generation. Two implementations denoted as shakespear TensorFlow (TF)[1] and RNN TF Char[2] and Word[3] Levels are adopted to implement the Recurrent Neural Network (RNN) for GRU and LSTM encoding respectively. Both are sequence-to-sequence model that take keywords (i.e., seed text) as input to generate relevant text. For instance, Hotel, Reservations, Flights, Booking, and travel are general keywords that could be used for generating Ads related to Booking domain. keywords such as Series, TV, Movies, Channels, Episode, and Season could be supplied to the model for generating Ads related to Movie domain. The shake-spear TF only support the character level, while the RNN TF support both character level and word level encoding.

The following factors are considered in the application of series of experiments to investigates their impact on the quality of generated advertisements campaigns.

- Dataset domain: Datasets in Booking/Reservations and Movies (Hulu) domains are considered to train the NNs.
- Neural network architectures: LSTM and GRU neural networks are investigated to generate the text.
- Name entity replacement: the impact of replacing named entities with tags using GeoText and Static lists are used to investigate the impact on the quality of generated texts.
- Input / output encoding level: character level and word level encoding sequence are also explored.

The subsections present the experimental settings, and the evaluation metrics.

a. Parameters Settings

Table 2 describes the parameters settings of LSTM and GRU neural networks that are used in the experiments. The parameters settings of character-level GRU and both character-level and word-level LSTM are presented. The parameters values are the most recommended values, which are tunned after a series of experiments.

Table 2: Parameters setting values for LSTM and GRU NN

Parameter	Char-level LSTM	Word-level LSTM	GRU
RNN size	128	256	512
Hidden layers	2	2	3
Sequence length	50	25	30
Number of epochs	2000	2000	10
Learning Rate	0.002	0.002	0.001
Optimizer	Adam	Adam	Adam

The datasets that are used in our experiments are described in Table 1. The datasets are split into three subsets, 70% for Training subset, 15% for validation and 15% for testing. Experimental studies focus on character level encoding over the word level encoding, because character encoding does not suffer from out-of-vocabulary issues, and being able to model different and rare morphological variants of a word, and do not require segmentation [7].

b. Evaluation Metrics

The neural networks are trained on the forementioned datasets, and the evaluation metrics are the loss error and perplexity (PPL) criterion in order to judge the performance of learning models [26]. Moreover, readability and relevance of the generated text are subjectively assessed by human annotators, and also readability is objectively evaluated with statistical propertied using a python tool called textStat [21].

Text relevance refers to the match between the information inferred from the text and the reader's goal [27]. In other words, text relevance means the match between the gendered text and the keywords/domains/campaigns used to for generation. The more match between reader's goals and inferred information the more relevant to the supplied keyworks. In this study, A total of 54 Human annotators are hired from Amazon Mechanical Turk to evaluate the generated texts. The annotators are English native speaker and eligible to do "Human Intelligence Tasks" (HITs)[4]. They are distributed into 18 groups of three participant in a every group. Each group is provided with the generated campaigns and the same keywords, which are used in the generation process and asked to assess the relevance of the text by answering to two points: <u>rating scale **(R)** for relevance and **(I)** for irrelevance</u>. The majority answer of the three answers is consider as output of evaluation results.

Readability refers to the rate of easy of understanding the intended meaning of text. The Less complex, difficult, grammatical and linguistical errors text is the more readable text [28].

The groups of annotators are also asked to evaluate the generated campaign, and to assess readability in four-point rating scale (easy, normal, difficult, and confusing). Three annotators are asked to assess a given generated text and the final readability label is determined by voting their annotation as shown in table

Table 3.

Table 3 Votes to determine Readability

Evaluation Result	Vote Result
Easy	Easy
Easy	
Standard	
Difficult	Confused Vote
Easy	
Confusing	

[1] https://github.com/martin-gorner/tensorflow-rnn-shakespeare
[2] https://github.com/sherjilozair/char-rnn-tensorflow
[3] https://github.com/hunkim/word-rnn-tensorflow
[4] Selecting eligible workers - Amazon Mechanical Turk

The ***Textstat*** tool uses the Flesch Reading Ease Score (FRES) test to assess the overall readability of text based on the Flesch Reading Ease Formula [29]. FRES is a seven points diffculty scale, and human anotator in this study evaluate readability in four point scal as shown in Table 4. We make this mapping to convert FRES measure to four point scale to be consistent with human anotators evalautions.

Table 4: Normalize FRES to corresponding four-point scale

Score	Difficulty	Normalized 4-point Scale
90-100	Very Easy	Easy
80-89	Easy	
70-79	Fairly Easy	
60-69	Standard	Standard
50-59	Fairly Difficult	Difficult
30-49	Difficult	
0-29	Very Confusing	Confusing

IV. EXPERIMENTS AND EVALUATION RESULTS

The experiments are conducted on a dedicated root server with a minimal Ubuntu OS version. The server has RAM 64 GB DDR4, Hard drive SSD 500GB, graphics card GeForce® GTX 1080, CPU Intel® Core i7-6700 Quadcore processor built and connection speed with 1 GBit/s-Port. Python 3.5 and TensorFlow with enabled GPU is used to implement the proposed RNN architectures.

A series of experiments are conducted to investigate the factors mentioned in Section III (dataset domain, NN architecture, name entity normalization, and input/ output sequence level), which affect the quality of generated texts.

Experimental results are presented in the next section.

a. DATASET DOMAIN EXPERIMENTAL STUDY

To investigate the influence of dataset domain on the generated text, a total of 102 Ads were generated by two Shakespear TF character-level models. The first one is trained Hulu dataset, and the second one in trained on booking datasets. Table 5 Shows PPL and training loss of TF Shakespeare character-level models trained on Hulu and Booking datasets.

Table 5 PPL and training loss of TF Shakespeare character-level models trained on Hulu and Booking datasets

Datasets	Loss Error	PPL	Relevance
Hulu	0.115	80	99%
Booking	0.503	45	99%

The results show that loss Error is 0.115 and 0.503 for Hulu and Booking datasets respectively. The PPL values are 80 and 45 for Hulu and booking respectively. The results imply that campaigns generated on booking domain fits better than those that generated on Hulu domain. Human annotators are totally agreed that 99% of the Ads generated in Hulu and booking domains are relevant to the provided keywords.

presents the evaluation results of readability of Ads, which are generated by GRU in Hulu and Booking dataset domains.

Table 6: results of evaluating Readability for Ads generated by GRU in Hulu and Booking domains.

Datasets	Evaluator	Easy	Standard	Difficult	Confused vote
Hulu	Human	89%	8%	2%	1%
	Textstat	96%	4%	0%	0%
Booking	Human	98%	0%	0%	2%
	Textstat	92%	5%	3%	0%

The results show the percentage of evaluation Readability as rated by human annotators and TextStat Tool. In general, it can be noted from the results that the generated texts are mostly readable. In Hulu domain, 89% and 96% are rated as easy to read by human and textStat respectively. In the Booking domain, 98% and 92% are rated easy to read by human and textStat respectively. a very small percentage of Ads, 8%, 2%, and 1% are rated standard, difficult, and confused vote respectively.

Figure 3 Compares Human evaluation and TextStat tool evaluation results of readability for the Ads, which are generated by the GRU neural networks in Hulu and Booking domains.

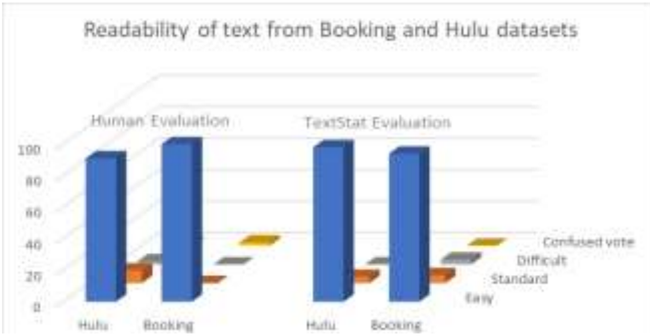

Figure 3: Readability Text from Booking and Hulu dataset.

The results imply that the two methods of evaluation (i.e., Human annotators and TextStat) are very compatible. Also, the GRU architecture extremely success in generating easy to read advertisement for both domains.

b. NEURAL NETWORKS EXPERIMENTAL STUDY

This study investigates the performance of the GRU and LSTM neural networks architectures which are implemented by shake-spear TF RNN and TF GRU respectively.

Table 7 presents the training loss and the evaluation relevance of Ads, generated by both GRU and LSTM neural networks using full character level encoding.

Table 7: Percentage of relevant Ads and training results (i.e., loss error and PPL) of GRU and SLTM neural network on Booking dataset

NN	Loss	PPL	Relevance
GRU	0.503	45	99%
LSTM	0.505	78	68%

We limit the dataset in this experiment to booking dataset (102 ads) because it showed betters results than the Movie domain dataset as shown in Table 5, and because the main point in this experiment is to compare LSTM and GRU for text generation.

The results of this experiments are shown in Table 7. The results show that the training loss of LSTM and GRU trained on the Booking dataset are 0.505 and 0.503 respectively. The loss errors for both are very close to each other's. On the other had, the PPL results are significantly different (78 and 45 for LSTM and GRU respectively). The results imply that Ads generated by the GRU has fits better than the Ads, which generated by the LSTM. In addition, the results show that human annotators rated Ads generated by GRU are more relevant than the ones by LSTM. It can be observed that there is a significant difference between in the performance the LSTM and GRU in terms of PPL and the relevance.

Table 8 presents the results of evaluating readability (Human and TextStat) of Ads, which are generated by GRU and LSTM at the character level encoding in Booking dataset domain. The results show that LSTM neural networks generates 43% and 45% easy to read Ads as rated by human and textStat respectively, while the GRU generates 98% and 92% easy to reads ads as rated by human and textStat respectively. More than 55% As generated by LSTM rated as difficult or confused Ads.

Table 8 Readability for Ads generated by GRU and LSTM in Booking domains

NN	Evaluator	Easy	Standard	Difficult	Confused vote
GRU	Human	98%	0%	0%	2%
	Textstat	92%	5%	3%	0%
LSTM	Human	43%	4%	26%	27%
	Textstat	45%	25%	30%	0%

The results imply that the Human evaluation is compatible and supporting the evaluation results of TextStat tool. In general, the result suggests that GRU network outperforms the LSTM network in generating easy to read and more relevance Ads text.

c. NAME ENTITY EXPERIMENTAL STUDY.

This experiment investigates the impact of NE normalization on the quality of generated texts. So, the NE normalization is applied on the training dataset (Booking).

Three experimental studies are conducted, NE normalization are applied by two different tools, i.e., geotext tool and a static list. We compare their impact on the DL model in Table 9, which presents the training loss and the PPL and the relevance results using NE normalization by GeoText library, static list, and without NE normalization for 102 ads. The texts are generated by GRU model trained on booking data.

Table 9: Percentage of relevant Ads and training results (i.e., loss error and PPL) of GRU on Booking dataset on three cases of NE removal

NE Library	Loss	PPL	Relevance
Geotext	0.437	43	99%
static-NE list	0.468	40	99%
Without NE Normalization	0.503	45	99%

The results in Table 9 suggest that NE normalization has no significant impact on the generated texts.
Table 10 presents the results of evaluating readability level of Ads by generated by GRU NNs trained on Booking dataset. The table includes the readability level of three cases (GeoText library, static list, and without NE normalization).

Table 10: Results of evaluating readability of Ads generated by GRU on Booking domains using different normalization techniques

NE Normalization	Evaluator	Easy	Standard	Difficult	Confused vote
Geotext	Human	98%	2%	0%	0%
	Textstat	77%	12%	11%	0%
static-list	Human	95%	2%	1%	2%
	Textstat	77%	14%	9%	0%
No NE Normalization	Human	98%	0%	0%	2%
	Textstat	92%	05%	03%	0%

The results show that, In the case of applying NE by geotext tool, Human rated 98% as easy to read and 2% Ads as standard, while TextStat is evaluated 77%, 12 % and 11% of Ads as easy to read, standard and difficult respectively.

In the case of performing NE using static-list, Human rated 95% of ads easy to read, 2% Ads standard, 1% difficult and 2% as confused vote. TextStat evaluated 77%, 14%, 9%, and 2% of generated ads as easy to read, standard, difficult, and confused vote respectively.

The results in Table 10 suggest that the application of NE normalization does not influence the human evaluation either for readability or relevance, while the textStat evaluation is negatively affected. The results show that the percentage of easy-to-read ads is degraded from 98% to 77 % and the percentage of standard and difficult to read Ads is increased to 12% and 11% respectively.

d. INPUT / OUTPUT TEXT SEQUENCE EXPERIMENTAL STUDY

This experiment investigates the influence of encoding level on the performance of LSTM neural networks. The experiment is limited to LSTM because the shake-spear TF only support character encoding, while RNN TF implementation supports both character-level and word-level encoding.

Table 11 presents the training loss, the PPL, and the relevance of Ads generated by the LSTM network trained on Booking dataset on both character-level and word-level encoding.

Table 11: Percentage of relevant Ads and training results (i.e., loss error and PPL) of LSTM on Booking dataset for Character-level and word level encoding.

Encoding level	Loss	PPL	Relevance
Character	0.505	78	83%
Word	1.232	86	89%

It can be noted from training loss and the PPL results in the table that the results of LSTM with character-level is better than LSTM with word-level encoding level. On the other hand, word level is more relevant than the character level, and this is because of the text generated using the character level scheme has some words that has some typos errors. The results also suggests that the character level model generates a text that fits to the target text, while the word level model generates more relevant texts.

Table 12 presents the results of evaluating readability of Ads, which are generated by LSTM with the character-level and word-level encoding. Human and textStat tool evaluation results is presented.

Table 12: Results of evaluating readability of Ads generated by LSTM on character and word level encoding in Booking domain.

Encoding level	Evaluator	Easy	Standard	Difficult	Confused vote
Character	Human	47%	06%	12%	35%
	Textstat	43%	29%	28%	0%
Word	Human	96%	04%	1%	9%
	Textstat	67%	14%	19%	0%

The results show that 47%, and 35% of Ads generated on for character-level are rated by human easy to read and confused vote. TextStat evaluate that 43% and 29%, 28% are evaluated as easy to read, standard and difficult respectively. In word-level encoding 96% of Ads are rated by human as easy to read, and 9% are rated as confused. TextStat evaluated 67% as easy to read and 14% standard and 19% difficult. The results show that word-level LSTM performs better than character-level LSTM in booking domain.

The results in Table 12 also suggest that the word level scheme is better than the character level because the texts generated by the character level have some types / spell errors. In addition, the results in Tables 11 and 12 support this conclusion.

V. CONCLUSION

This research proposed the application of GRU and LSTM deep neural networks in generating advertisement text campaigns. Two datasets' domains i.e., hotel Booking, and TV and movies streaming are included. Presrocessing including normalization, and Name Entity processing are performed to reduce the number of strange names and prepare the dataset for machine learning. The main contribution of this research is to investigate the influence of four factors including, neural network architecture, dataset domain, NE normalization, and input encoding (character / word levels), on generating Ads. Readability of the generated Ads is subjectively evaluated using human annotators and objectively assessed using TextStat tool, whereas the relevancy is only evaluated by human annotators. The implication is several factors could be tuned to improve the performance of neural network in generating attractive Ads. Several experiments have been conducted to investigate the impact of the factors mentioned earlier. An investigation has been conducted to determine the influence of every factor on the quality of generated text. In general, the results indicate that the GRU networks outperform the LSTM networking in generating easy to reads ads campaign. In addition, Training GRU NN on Booking domain has better performance compared to Hulu dataset domain.

It can be concluded from the results that the collected data and dataset domain and input/output encoding level are the most common factors influence the performance of the generated texts

For future work, generating advertisement campaign in Arabic language will be investigated. More experiments on other dataset domain including brands, shopping product need to be conducted too. Besides that, investigations pertaining to generate multiple ads campaigns for every keyword is required in divers Ads.

References

1. Handayani, W., S. Muljaningsih, and H.J.T.S.O.S.J. Ardyanfitri, *Online Marketing Supports Promotion And Advertising Sales In Communities Dolly Localization.* 2018. **2**(1): p. 75-87.
2. Eurostat, *Internet advertising of businesses-statistics on usage of ads Statistics Explained.* 2018.
3. IBA, *2019 Internet Ad Revenue Report.* 2020.
4. Kong, S., et al., *Web advertisement effectiveness evaluation: Attention and memory.* 2019. **25**(1): p. 130-146.
5. Bhatia, V. and V. Hasija. *Targeted advertising using behavioural data and social data mining*. in *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*. 2016. IEEE.
6. Bartz, K., C. Barr, and A. Aijaz. *Natural language generation for sponsored-search advertisements*. in *Proceedings of the 9th ACM Conference on Electronic Commerce*. 2008.
7. Lee, J., K. Cho, and T.J.T.o.t.A.f.C.L. Hofmann, *Fully character-level neural machine translation without explicit segmentation.* 2017. **5**: p. 365-378.
8. Singh, S.P., et al. *Machine translation using deep learning: An overview*. in *2017 international conference on computer, communications and electronics (comptelix)*. 2017. IEEE.
9. Allahyari, M., et al., *Text summarization techniques: a brief survey.* 2017.

10. El-Kassas, W.S., et al., *Automatic text summarization: A comprehensive survey.* 2021. **165**: p. 113679.
11. Iqbal, T., S.J.J.o.K.S.U.-C. Qureshi, and I. Sciences, *The survey: Text generation models in deep learning.* 2020.
12. Chan, Z., et al. *Selection and Generation: Learning towards Multi-Product Advertisement Post Generation*. in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020.
13. Hughes, J.W., K.-h. Chang, and R. Zhang. *Generating better search engine text advertisements with deep reinforcement learning*. in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019.
14. Taneja, P. and K.G. Verma, *Text Generation Using Different Recurrent Neural Networks*. 2017.
15. Yang, X., et al. *Advertising keyword recommendation based on supervised link prediction in multi-relational network*. in *Proceedings of the 26th International Conference on World Wide Web Companion*. 2017.
16. Hoobyar, T., T. Dotz, and S. Sanders, *NLP: The Essential Guide to Neuro-Linguistic Programming*. 2013: William Morrow.
17. Yang, Z., et al., *Review networks for caption generation.* 2016. **29**: p. 2361-2369.
18. Xiang, L., et al., *Novel linguistic steganography based on character-level text generation.* 2020. **8**(9): p. 1558.
19. Mishra, S., et al. *Learning to create better ads: Generation and ranking approaches for ad creative refinement*. in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2020.
20. Duan, S., et al. *Query-Variant Advertisement Text Generation with Association Knowledge*. in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 2021.
21. Bansal., S. *textstat Python Tool*. 2018 2018 [cited 2021 8/9/2021]; Available from: https://github.com/shivam5992/textstat,.
22. Vinayak, S., *SEMrush Toolkit For Bloggers: 8 Tools To Boost Blog Content For More Traffic And Revenue.* 2021.
23. Golikova, D.M.J.V.O., *Named Entities for Computational Linguistics.* 2018. **15**(1): p. 207-215.
24. Hu, Y.J.G.C., *Geo‐text data and data‐driven geospatial semantics.* 2018. **12**(11): p. e12404.
25. Palenzuela, Y.M. *geotext*. 2014 [cited 2021 6-9-2021]; Available from: https://geotext.readthedocs.io/en/latest/readme.html.
26. Klakow, D. and J.J.S.C. Peters, *Testing the correlation of word error rate and perplexity.* 2002. **38**(1-2): p. 19-28.
27. McCrudden, M.T., G. Schraw, and B. Hoffman, *Text Relevance*, in *Encyclopedia of the Sciences of Learning*, N.M. Seel, Editor. 2012, Springer US: Boston, MA. p. 3307-3310.
28. Deutsch, T., M. Jasbi, and S.J.a.p.a. Shieber, *Linguistic features for readability assessment.* 2020.
29. Farr, J.N., J.J. Jenkins, and D.G.J.J.o.a.p. Paterson, *Simplification of Flesch reading ease formula.* 1951. **35**(5): p. 333.

Atef Ahmed. Holds a master's degrees in data science from The Islamic University of Gaza, and he is a software engineer.

Motaz Saad is a computer scientist, he holds a Ph.D. degree in computer science from the University of Lorraine, France. His research interests include AI, NLP, and machine learning, and he published several papers in the field. He is currently working as an assistant professor at The Islamic University of Gaza, Palestine.

Basem O. Alijla received the Ph.D. degree in intelligent systems from The University Science Malaysia (USM), in 2015. He is currently Assistant Professor in Computer Science and deputy Dean Faculty of Information Technology, Islamic University of Gaza. He published several research papers in high impact factor journals and international conferences. His research interest includes evolutionary computing, optimization, machine learning, data mining and features extraction and selection.