IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 10, 2022, accepted May 2, 2022, date of publication May 10, 2022, date of current version May 23, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3174108

A Systematic Literature Review on Text Generation Using Deep Neural Network Models

NOUREEN FATIMA[1], ALI SHARIQ IMRAN[2], (Member, IEEE), ZENUN KASTRATI[3], SHER MUHAMMAD DAUDPOTA[1], AND ABDULLAH SOOMRO[1]

[1]Department of Computer Science, Sukkur IBA University, Sukkur 65200, Pakistan
[2]Department of Computer Science, Norwegian University of Science and Technology (NTNU), 2815 Gjøvik, Norway
[3]Department of Informatics, Linnaeus University, 35195 Växjö, Sweden

Corresponding author: Ali Shariq Imran (ali.imran@ntnu.no)

This work was supported in part by the Department of Computer Science (IDI), Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology (NTNU), Gjøvik, Norway; and in part by the Curricula Development and Capacity Building in Applied Computer Science for Pakistani Higher Education Institutions (CONNECT) Project NORPART-2021/10502, funded by Diku.

ABSTRACT In recent years, significant progress has been made in text generation. The latest text generation models are revolutionizing the domain by generating human-like text. It has gained wide popularity recently in many domains like news, social networks, movie scriptwriting, and poetry composition, to name a few. The application of text generation in various fields has resulted in a lot of interest from the scientific community in this area. To the best of our knowledge, there is a lack of extensive review and an up-to-date body of knowledge of text generation deep learning models. Therefore, this survey aims to bring together all the relevant work in a systematic mapping study highlighting key contributions from various researchers over the years, focusing on the past, present, and future trends. In this work, we have identified 90 primary studies from 2015 to 2021 employing the PRISMA framework. We also identified research gaps that are further needed to be explored by the research community. In the end, we provide some future directions for researchers and guidelines for practitioners based on the findings of this review.

INDEX TERMS Systematic literature review, deep learning, text generation survey, natural langauge processing, quality metrics, neural network, GPT, LSTM.

I. INTRODUCTION

Text Generation is a field of study in Natural Language Processing (NLP) that combines computational linguistic and artificial intelligence to generate new text. It is a process of generating grammatically and semantically correct synthetic text. This process includes training a model that takes input data, learns the context from the input, and generates new text relating to the domain of input data. The generated text should satisfy the basic language structure and convey the desired message [1]. It is challenging to generate and evaluate grammatically, semantically, and synthetically correct text because text generation and its evaluation are open-ended. Thus, this Systematic Literature Review (SLR) discusses five research aspects associated with text generation. These include the deep learning approaches for text generation, quality metrics for evaluating generated text, training datasets used in the domain, languages on which the text generation is performed, and application areas for text generation.

Text generation can be performed at different granularity of text, i.e., character, word, and sentence level [2]. Text generation at a sentence level aims to analyze the entire text as a fine-grained and learn the relationship between the sentence and its context. Meanwhile, word-based text generation seeks to explore the structure of a sequence and predict the probability of the next word in a given text. Similarly, the model identifies the character rather than the entire document at character level text generation.

Automatic text generation was possible due to recent developments in computational resources coupled with advancements in deep learning techniques. Deep learning is a field of machine learning that uses artificial neural networks and representation learning. Text generation approaches can broadly

The associate editor coordinating the review of this manuscript and approving it for publication was Alicia Fornés.

be categorized into three types of deep learning models as given below:

1) Vector-Sequence Model – Input is a fixed-size vector, whereas output can vary. For instance, this model can be used for caption generation of images [3].
2) Sequence-Vector Model - Input is of variable size, and output is a fixed-size vector. Classification is an example of this model [4].
3) Sequence-to-Sequence Model - Input and output are variable sizes in this model type. It is the most widely used variant of text generation models. Language translation belongs to this type of text generation model [5]

Above all, deep learning has contributed immensely to different aspects of natural language generation for various tasks including, dataset balancing [6], [7], next word prediction & text suggestion in chatting, generation of answers to questions in question answering system, in chatbots [8], [9], machine learning translation [10], [11], text summarization [12]–[14], text classification [15], [16], text generation for topic modeling [17], dialogue generation [18], sentiment analysis [19], [20], poetry writing [21], script writing for movies [1], [22], and others.

Evaluating the quality of the generated text governs the model's performance and measures the diversity of generated and original text. The quality metrics are also known as evaluation methods. There are two ways to assess the quality of generated text: human-centric (HC) and machine-centric (MC) [23]. The human-centric evaluation method involves language and domain experts who evaluate the generated text. It is expensive in terms of both time and cost and is prone to human errors. On the other hand, the machine-centric evaluation method, as known as objective quality assessment, is widely adopted and found in the literature. It includes various evaluation metrics: Metric for evaluation of translation with explicit ordering (METEOR), bilingual evaluation understudy (BLEU), recall-oriented understudy for gisting evaluation (ROUGE), consensus-based image description evaluation (CIDEr), National institute of standards and technology (NIST), word error rate (WER), Word Perplexity, and BERTScore. The machine-centric method saves time and cost, but the quality of an objective evaluation metric is highly language-specific.

There are various deep learning architectural frameworks widely used in the literature to implement deep learning models. Recurrent Neural Networks (RNN) [24] is one of them. It is a class of neural networks that uses the output of previous states as input in future states. This is the first algorithm that preserves the outputs of past states. One problem with RNN is that it forgets the previous outputs over a period of time due to a vanishing gradient.

Bidirectional RNN [25] uses two RNN layers that look into the sequence in both directions, i.e., forward and backward, and combine their output. This is helpful when the current state is not only dependent on the previous state but also on the future state. One special class of RNN is Long Short-Term Memory (LSTM) [26] network that is used to retain the information of previous states over a very long period and forgets the irrelevant information. Gated Recurrent Unit (GRU) also overcomes the problem of vanishing gradient in RNN. GRU is a simplified version of LSTM.

Generative Adversarial Network (GAN) works on the concept of *minmax* game where the discriminator predicts if the sample is from the training set or is produced by a generative network, and the generator tries to maximize the mistakes of the discriminator.

GPT-2 was proposed by Radford *et al.* [27]. GPT-2 is a transformer-based model having 1.5 billion parameters. It is trained on 40GB of Internet text scrapped from eight million web pages. It is a revolutionary model in text processing. It has an exceptional human-like ability to generate long sequences.

In June 2020, OpenAI released the third version of GPT, which is 100 times larger than the previous model. GPT-3 is trained on 499 billion tokens of web data, and it has 175 billion parameters and 96 layers. It has more generative power that it may outperform in many different tasks like text generation, and zero-shot and one-shot learning [28]. However, the model is not publicly available; instead, API access was to be provided but to those who pay for it [29].

Usually, these pre-trained (LSTM and GPT-based) models are used to generate text in different domains. For example, it is possible to use the pre-trained GPT model for generating a movie script and customize its generation capability by fine-tuning using some movies' script datasets. Once the data is gathered and model learning is customized to generate domain-specific text, the next step is to assess the quality of the generated text. LSTM was introduced as a character-level text generation model.

A. BACKGROUND

The basic strategy of text generation is first to train any language model on lots of sequences of text data, and then the model is capable of generating the next character or multiple characters in the sequence given previous characters as input. For example, generate the next character '*k*' for given sequence '*Cat likes mil*', as shown in Figure 1. Looking over the example mentioned in this paper [6], conditioning text, i.e., initial text that is fed to LSTM network for predicting next character in the sequence is 'Cat likes m'. The LSTM model is trained on Wikipedia text or related domain of conditioning text. To predict the probability of the next character using the *Softmax* activation function at the output layer is defined in Equation 1.

$$Softmax(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \tag{1}$$

where, x_i is the LSTM score for character i to be the next character in conditioning text. Each of the x_i is not a probability score, therefore LSTM uses softmax to convert LSTM scores to probabilities score.

The actual magic of text generation is hidden in the sampling strategy. Text generation would be almost similar to

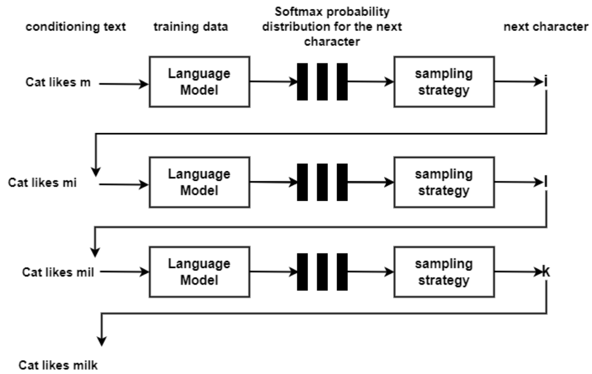


FIGURE 1. LSTM text generation process [6].

the original text if the next character is generated based on the highest score of probability taken from softmax output. Thus, some randomness was introduced in the generated text to introduce novelty and creativity to generated text. The sampling strategy introduces such randomness using temperature value.

Suppose, $P_{original}$ is original probability distribution at *Softmax*, the α term is defined as,

$$\alpha = \log \frac{(P_{original})}{temperature} \tag{2}$$

Once α is computed, the $P_{revised}$ is defined as,

$$P_{revised} = \frac{e^{\alpha}}{n} \tag{3}$$

where, n is the number of elements in the original distribution and temperature value is an arbitrary value ranging from any non-zero value up to 1.

B. RELATED SURVEYS

There are a handful of surveys published on the topic of text generation, as shown in Table 1. We have found that five research papers have worked on a single aspect. Li *et al.* in [30] have provided a systematic literature review on deep learning approaches along with its data type. The authors mainly focused on the encoder and decoder-based deep learning architecture. Besides that, different data types (unstructured input, structured, and multimedia) were discussed, along with best-fitted transformer-based models. Similarly, Gatt *et al.*, in [31], have worked on a systematic review of text generation-based applications. Lu *et al.* in [32], have provided a systematic literature review on the evolution matrix of SLR text generation. Lastly, both studies [23] and [33] have conducted a review study on text generation only based on quality metrics.

Besides that, few researchers have worked on two aspects of SLR text generation. Four research works [1], [31], [34], [36] have provided a literature review on quality metrics and deep learning approaches. Another review study focused on three aspects - quality metrics, approaches, and applications, and provided an overview of text generation [34]. Similarly,

TABLE 1. A summary of past related surveys.

#	Ref.	Year	QM	Approach	Application	Peer-review
1	[30]	2021		✓		
2	[34]	2020	✓	✓	✓	
3	[23]	2020	✓			
4	[35]	2020			✓	✓
5	[33]	2020	✓			
6	[36]	2020	✓	✓		
7	[1]	2020	✓	✓	✓	
8	[32]	2018		✓		
9	[31]	2018	✓	✓		✓

a study in [1] aimed to review multiple objectives of SLR on quality metrics, datasets, approaches, and application. There are two major limitations of these systematic literature reviews. First, seven articles were not peer-reviewed. Second, none of these attempts have worked on a comprehensive review of the dataset, quality metrics, languages, deep learning approaches, and trends in text generation in deep learning in a single study.

The limitations and findings shown in Table 1 provide a base for conducting a comprehensive review of text generation using deep learning. Therefore, our study focuses on articles published between 2015 to 2021. Ninety baseline articles are reviewed following the Preferred Reporting Items for Systematic Literature Review and Meta-Analysis (PRISMA) protocol for systematic literature review. We have investigated text generations on five different aspects, namely deep learning approach, quality metric, dataset, language, and application of text generation in deep learning, as shown in Figure 2.

The main contributions of this study are as follows:

1) A systematic map of 90 primary studies based on the PRISMA framework;
2) An analysis of the investigated text generations on five different aspects, namely deep learning approaches, quality metrics, datasets, languages, and applications on text generation in deep learning;
3) An overview of the challenges, opportunities, and recommendations of the field for future research exploration.

Additionally, this SLR provides an in-depth analysis, the most extensive and up-to-date body of knowledge of text generation based on five research aspects, and also focuses on the major challenges and future research directions in the text generation domain. To the best of our knowledge, there is no SLR on text generation that covers all these aspects.

The rest of the paper is organized as follows. Section II describes the research design of this SLR followed by Section III that covers the finding of RQs and provides the most relevant articles based on quality assessment criteria. Section IV provides the identified challenges and research gaps. Section V presents the recommendations and future research directions and finally Section VI summarizes the SLR.

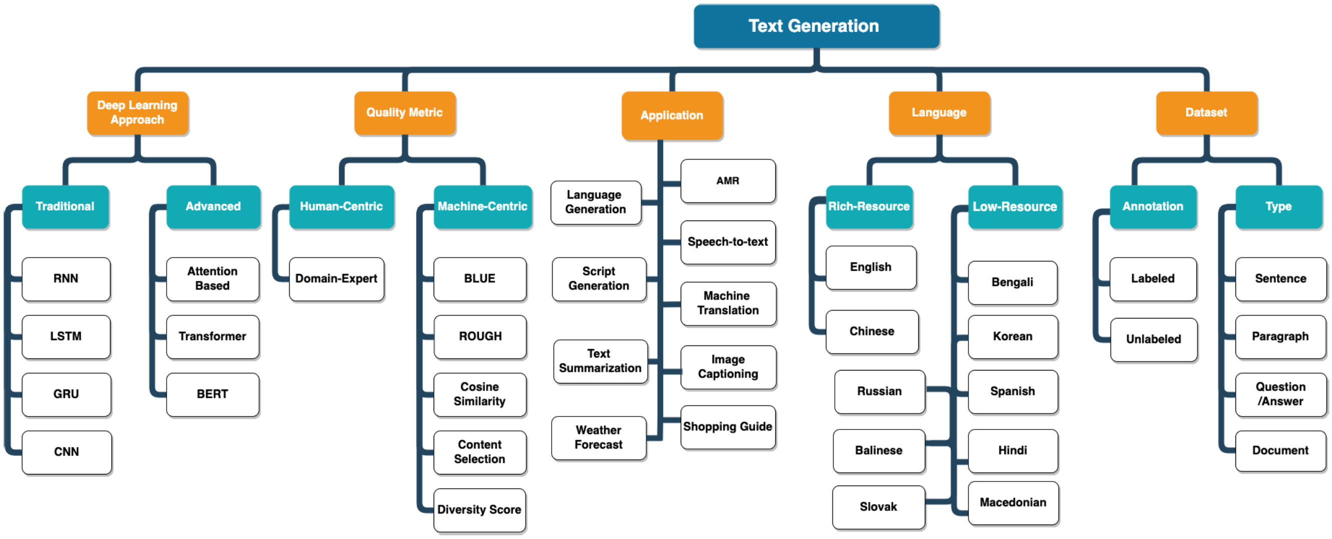


FIGURE 2. Taxonomy of the text generation.

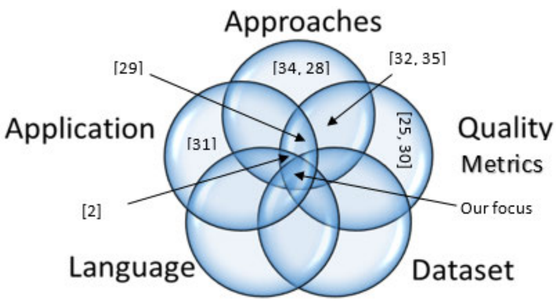


FIGURE 3. Related work and our focus.

II. RESEARCH DESIGN

In this study, we have applied systematic mapping as a research methodology for reviewing the literature. We have utilized the guidelines of PRISMA, given by [37]. This SLR consists of four major steps: planning and searching of primary studies, collection of studies, data extraction, and synthesis of data. The first step generally identifies research questions and objectives (stated in Section II-A). The search strategy step involves criteria for selecting studies, study selection procedure, keywords formulation for research and search queries, as well as the quality assessment criteria of extracted studies (which are addressed in Section II-B). The data extraction step involves strategies of data extraction from selected studies (see Section II-C and II-D for details). In addition, the final step involves Quality assessment (see Section II-E for more details).

A. RESEARCH QUESTIONS

The primary purpose of this SLR is to explore various techniques for text generation using deep learning. The following five research questions (RQs) were raised to achieve this aim, as shown in Table 2.

TABLE 2. Research questions of SLR.

RQ	Research question
RQ1	Which traditional and advanced deep learning approaches are used to generate text in the literature?
RQ2	What are the various metrics for evaluating the performance of text generation models?
RQ3	What are the major standard datasets for text generation in the literature?
RQ4	What are the application areas where text generation is extensively used?
RQ5	Which languages have been focused on text generation in deep learning?

B. RESEARCH OBJECTIVES

The following five research objectives of this study are given below:

- To investigate the existing traditional and advanced deep learning-based text generation approaches/techniques
- To explore various performance metrics used for evaluating text generation models
- To investigate various evaluation methods for measuring the quality of generated text
- To review the recent application domains where text generation is being applied
- To discuss the major challenges and future research directions in the text generation domain

C. SEARCH STRATEGY TO RETRIEVE PRIMARY STUDIES

The majority of studies have included text generation or automatic text generation as their data sources for text generation. Thus, various search keywords are formulated to retrieve the related literature from six reliable and high-quality academic databases, namely, Web of Science (WoS), Scopus, IEEE Xplore, Springer link, ScienceDirect, and ACM Digital Library. Five of the authors prepared a list of several relevant keywords to search the relevant literature on "text generation

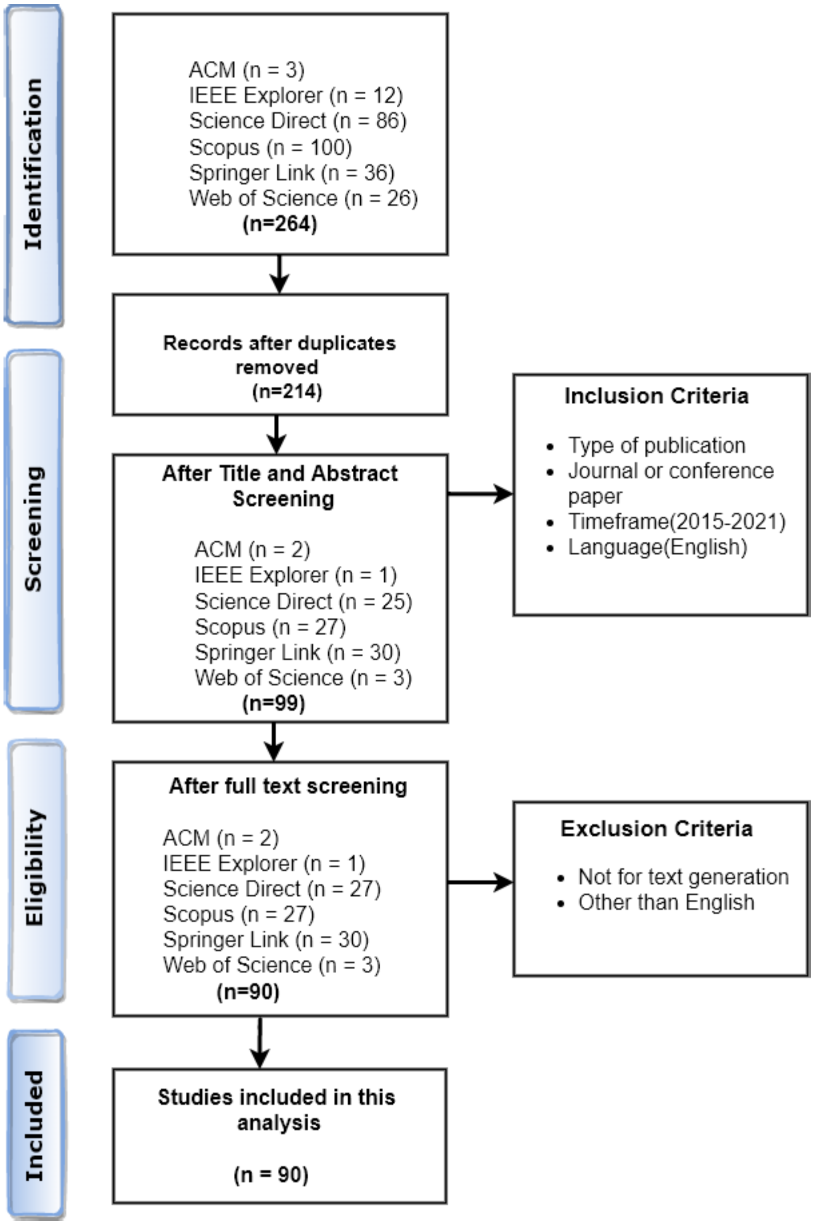


FIGURE 4. PRISMA search methodology.

techniques in deep learning'' from the selected databases. Table 3 shows the keywords used to perform queries. Each keyword within the group is paired using the OR operator, whereas the groups are paired using the AND operator (see Table 3) to form a search query. The last row of Table 3 shows how keywords from different groups are concatenated to form a query that was executed in all six bibliographic databases. The query was applied to the article title, article abstract, and article keywords to determine the relevant articles from the six selected bibliographic databases published in English from January 2015 to October 2021.

The search query identified 264 studies when applied to the five selected bibliographic databases, as shown in Figure 4. The identical studies from different databases were then extracted, and only distinctive copies were retained in EndNote for each primary sample. During removing of duplicate records, 50 studies were excluded.

TABLE 3. Selected keywords in the different groups.

Group 1: Deep Learning related Keywords	deep learning OR natural language processing OR NLP OR neural network OR RNN OR Recurrent OR Recursive OR LSTM OR GAN OR GPT-2 OR generative adversarial network
Group 2: Text Generation related Keywords	text generation OR language generation OR language modelling OR natural language generation OR neural language generation
Search Query	(Group 1) AND (Group 2)

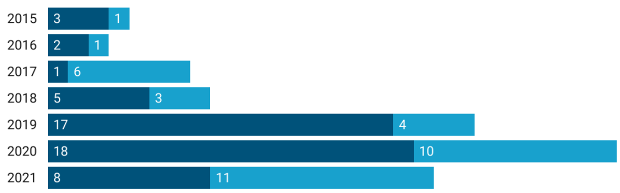


FIGURE 5. The number of collected conference and journal papers in 2015-2021.

D. ARTICLE SCREENING AND SELECTION CRITERIA

The remaining 214 studies were analyzed after the removal of duplicate records. The screening was done based on the title, abstract, and keywords of the articles retrieved. These studies were retrieved by four authors using inclusion and exclusions criteria. A majority vote was used to include or remove articles for all inconsistencies. Furthermore, a final decision was taken in the event of ties between all the authors. Figure 4 indicates the screening of all the articles based on the title, abstract, and keyword-based screening method. Moreover, only 90 out of 264 were selected for primary studies; the remaining articles were excluded. The distribution of conference and journal reviewed papers are shown in Figure 5.

There were established criteria for excluding 117 articles. First, the purpose of many excluded studies was to extract information other the text generation. Second, the majority of the studies were about text classification, which is out of our scope. Third, a number of articles were written other than in English. Lastly, studies that were not peer-reviewed were excluded from the analysis thus to maintain the quality of this SLR paper.

We use the following inclusion criteria:

- The article must be used to include a generative model for text only
- The article must be published from 2015 to 2021
- The article must be published in a journal or a conference
- The article must be published in the English Language

We use following exclusion criteria:

- The articles which used NLP or machine learning techniques but did not propose or used any text generation techniques are excluded
- The articles published in languages other than English are excluded

E. QUALITY ASSESSMENT

The quality assessment criteria (QAC) were used to assess the quality of the 90 selected studies. The QAC was used to assess whether a selected primary study could achieve our review objectives. To determine the consistency of selected primary studies, a variety of questions were asked by all the authors. Table 4 describes the list of 10 questions to check the quality of studies. Either Yes or No can be the answer to each question with weights of 1 and 0 respectively. A group of four authors reviewed the selected primary studies. Results were evaluated after the quality assessment of each primary study. Finally, each question is matched by all the authors of the current research for every study for the review process. However, the quality review process did not rule out any study as all the studies fit the quality assessment questions. This review, therefore, included all the 90 studies selected.

TABLE 4. Quality assessment questions.

Q1	Are the research objectives clearly stated?
Q2	Is the proposed methodology well defined?
Q3	Is the proposed text generation technique clearly explained?
Q4	Does the performance metrics for evaluating the text generation quality fully defined?
Q5	Is there enough information available for the dataset used?
Q6	Are the presented results clear and unambiguous?
Q7	Does the study perform comparison of proposed approach results with existing baseline approaches results?
Q8	Are the results properly interpreted and discussed?
Q9	Does the conclusion reflect the research findings?
Q10	Does the article contain future research directions and trends for text generation?

III. SYSTEMATIC MAPPING STUDY RESULTS

In this section, we critically analyze 90 primary studies from five different aspects, namely, deep learning approach, quality metric, dataset, language, and application.

RQ1: *Which traditional and advanced deep learning approaches are used to generate text in the literature?*

There are two different approaches found in the literature to generate the text: traditional deep learning approaches (TDLA) and advanced deep learning approaches (ADLA). In traditional approaches, many deep learning-based models and NLP techniques were employed to generate the text. The topmost text generation models are RNN [38], LSTM [39], and CNN [40]–[42]. The text generation domain has seen some limitations due to its discrete nature [42]. Language generation requires a lot of effort, domain knowledge, and skills to learn the different semantic and contextual meanings from the text. Every language has its own standard rules and regulations. Therefore, it is not possible that the generated CNN for the English model may perform well in the Urdu language. Moreover, the contextual meaning of the generated text was a major issue in traditional text generation deep learning-based algorithms. Thus, to overcome the problem in traditional approaches, many advanced deep learning models were introduced like transformer [43], [44], BERT [45], GPT2 [46] and GPT3 [28]. These latest models are content-dependent algorithms with attention mechanisms [46]. Moreover, many approaches and models have been employed to generate text in different languages, which can be categorized into three main groups. Table 5 shows the papers grouped based on traditional approaches of text generation, advanced deep learning, and a combination of both approaches. 47 papers out of 90 have employed advanced approaches to generate the text, 40 papers used the traditional approaches, and 3 papers have used both approaches.

Moreover, we found that after 2018 there was a drastic increase in using traditional as well as advanced text generation approaches, as shown in Figure 6. Many researchers

TABLE 5. Review of text generation approaches.

Approach	Papers
Advanced Deep Learning	[17], [47]–[86]
Traditional Deep Learning	[4], [18], [21], [21], [40], [87]–[92], [92]–[97] [3], [98]–[120]
Combination of both Approaches	[6], [81], [121]

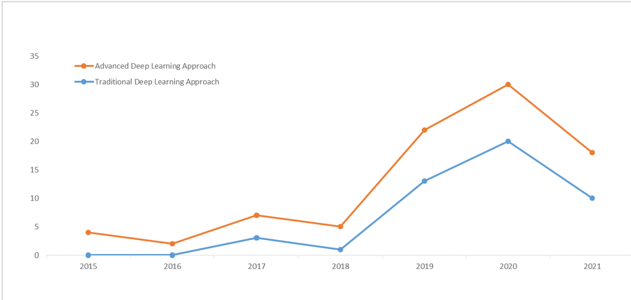


FIGURE 6. Text generations approaches used in the reviewed papers.

FIGURE 7. Summary of languages.

TABLE 6. Evaluation metrics applied in the reviewed papers.

Metrics Group	Studies
Human-centric (HC)	[49], [52], [97]
Machine-centric (MC)	[59], [62]–[66], [82], [83], [85] [3], [40], [68]–[70], [73], [96], [119] [71], [72], [74]–[79] [21], [81], [89]–[91], [93], [99], [122] [6], [47], [48], [50], [51], [54], [55], [123] [18], [56], [57], [112], [116], [124], [125] [82], [101], [109]–[111], [117] [102], [103], [105], [107], [113], [114], [120]
Both	[5], [53], [58], [121], [126] [60], [80], [84], [86], [95], [127] [98], [100], [104]

have been working on text generation in various languages. The most often used algorithms for text generations in the reviewed studies for traditional approaches are LSTM and RNN. On the other hand, for advanced approaches, we have found different versions of GAN and BERT.

RQ2: *What are the various metrics for evaluating the performance of text generation models?*

Two metrics, human-centric and machine-centric, can measure generated text. In this SLR, we categorize studies into three groups based on the approaches used to assess the quality of the generated text. As shown in Table 6, 64 out of 90 papers have evaluated generated text on the basis of a machine-centric approach, 3 papers have evaluated on the basis of human experts, and 14 studies have utilized both the approaches human- and machine-centric. However, we found 9 studies that have not performed any measures to evaluate the generated text.

As shown in Figure 8 BLEU score has been widely used to check the quality of the generated text. 80% of studies have

FIGURE 8. Evaluation metrics for text generation most commonly used in the literature.

used the BLEU score, 8% have used ROUGE and Perplexity, and 5% have used other metrics such as cosine similarity, content selection, diversity score, and word error rate.

RQ3: *What are the major standard datasets for text generation in the literature?*

The standard datasets for text generation based on their characteristic as mentioned below, have been extracted:

1) **Availability:** Private/Public
2) **Size:** Number of words, sentences, and reviews
3) **Type:** Sentence-, paragraph-, document-level and question/answer
4) **Format:** CSV, JSON, XML, files
5) **Annotation:** Labeled/unlabelled
6) **Quality:** Raw or pre-processed

The detail of datasets is given in Table 7. 9 out of 90 papers have used private datasets, and those are not publicly available [17], [49], [66], [67], [87], [89], [94]. 2 have used both public and private datasets [6], [84] and 79 studies have used the publicly available datasets, as shown in Table 7. In addition to this, 33 datasets were sentence-level, 14 were paragraph-level, 2 were document-level, 1 was question/answer type, and 1 study did not mention the type of dataset explicitly.

RQ4: *What are the application areas where text generation is extensively used?*

Text generation has wide range of applications in deep learning, which can be categorized into 18 groups as shown in Figure 9. We have found 10 out of 90 papers have main purpose to balance the dataset [6], [50], [68], [84], [85], [115], [118], [128], 8 out 90 papers have worked on data to text [47], [55], [64], [82], [82], [100], [102], [108], and speech to text [67], [83], [98], [103]–[105], [110], [113], respectively. 7 papers have worked on script writing [3], [17], [57], [60], [85], 5 papers have worked on machine translation [10], [11], [56], [87], [101], [120]. Apart from these, 4 papers have worked on text summarization [49], [87], [97], [126] and 2 papers [1], [90] have worked on abstract meaning representation (AMR)- AMR to text goal is to generate sentences from abstract meaning representation graphs and its seq2seq or graph2seq problem. In addition, we found 2 applications of product reviews [129], [129] of mobile devices that have worked in text generation. The single paper was found on C programming language code generation [106] and online shop guideline generation [84].

RQ5: *Which languages have been focused on text generation in deep learning?*

Many text generation models and approaches have been employed to generate text in different languages. There are eleven languages found in the literature: English,

TABLE 7. Detailed summary of datasets.

Name	Type	Size	Quality	Format	Lab	Private / Public	Link
Stanford Natural Language Inference (SNLI)	Sentence	30,000	Pre-process	Jsonl	Y	Public	https://nlp.stanford.edu/projects/snli/
Yelp Restaurant Review	Sentence	Not mentioned	Pre-process	Json	Y	Public	https://www.yelp.com/dataset
Alice in Wonderland.txt	Paragraph	1,63,780 characters	Raw	txt	N	Public	https://gist.github.com/phillipj/4944029
BRAD: Books Reviews in Arabic Dataset	Paragraph	510,600 book reviews	Pre-process	tsv	N	Public	https://github.com/elnagara/BRAD- Arabic-Dataset
AMADI_ LontarSet	Paragraph	No mentioned	Pre-process	txt	Y	Public	http://amadi.univ-lr.fr/ICFHR2016 Contest/index.php /download-123
Amazon Product Reviews Corpus (APRC)	Sentence	142.8 million reviews	Pre-process	txt	Y	Public	https://jmcauley.ucsd.edu/data/amazon/
Chinese Online- Shopping Reviews Corpus	Sentence	150 million words	Not mentioned	Not mentioned	Y	Private	https://passport.jd.com/uc/login?Return Url= http%3A%2F%2Fjddc.jd.com%2 Fauth_environment
Yelp review dataset	Paragraph	Not mentioned	Pre-process	Json	N	Public	https://www.kaggle.com/yelp-dataset/ yelp-dataset?select=yelp_academic_dataset_business.json
Stanford sentiment tree bank dataset	Paragraph	Not mentioned	Pre-process	tar	Y	Public	https://www.kaggle.com/atulanandjha/ stanford-sentiment-treebank-v2-sst2
NEWS dataset NYTimes	Paragraph	527595 articles	Raw	Json	N	Public	https://huggingface.co/datasets/ cc_news
AMR corpus (LDC2017T10)	Sentence	59,255 sentences	Raw	XML	Y	Public	https://catalog.ldc.upenn.edu/ LDC2017T10
Chinese E-commerce platform, Taobao	Not mentioned	Not mentioned	Not mentioned	Json	Y	Public	https://opendata.pku.edu.cn/ dataset.xhtml? persistentId= doi:10.18170/DVN/QJAFTE language=en
Amazon Product Review	Sentence	142.8 million reviews	Pre-process	txt	Y	Public	https://jmcauley.ucsd.edu/data/ amazon/
IMDB Movie Reviews	Sentence	Not mentioned	Pre-process	csv, txt	Y	Public	https://www.kaggle.com/lakshmi25 npathi/ imdb-dataset-of-50k- movie-reviews
Arabic Poetry	Sentence	58000 Poems	Pre-process	csv	Y	Public	https://www.kaggle.com/ahmedabelal/ arabic-poetry
alqasidah.com	Paragraph	I million words	Raw	Not mentioned	N	Private	https://alqasidah.com/
kitchen product reviews	Paragraph	4,253,926 reviews	Pre-process	csv	Y	Public	https://nijianmo.github.io/ amazon/index.html
Macedonian Storytelling	Paragraph	Not mentioned	Pre-process	txt	N	Public	https://raw.githubusercontent.com /Ivona221/ MacedonianStoryTelling /master/NewsGenerationData.txt
EMNLP2017 WMT News Dataset	Sentence	270,000 sentences	Pre-process	txt	Y	Public	http://statmt.org/wmt17/translation- task.html
COCO Image Captions	Sentence	20,734 words and 417,126 sentences	Pre-process	Not mentioned	Y	Public	https://cocodataset.org/#home
Chinese Poems	Sentence	Not mentioned	Pre-process	txt	N	Public	https://homepages.inf.ed.ac.uk/ mlap/ Data/EMNLP14/
ROCSTORY	Paragraph	45500 Sentence	Pre-process	csv	N	Public	https://github.com/snigdhac/ StoryComprehension_ EMNLP/tree/master/Dataset/ RoCStories
WMTNEWS	Paragraph	11.57M Token	Pre-process	XML	N	Public	https://opus.nlpl.eu/ WMT-News.php
Brown Corpus	Paragraph	10 million words	Pre-process	Json	N	Public	https://www.kaggle.com/nltkdata/ brown-corpus ?select=brown- meta.json
Penn Treebank Dataset	Sentence	10k unique words,	Pre-process	txt	N	Public	https://deepai.org/dataset/ penn-treebank
ROTOWIRE	Document	4853 summaries	Pre-process	.tar.bz2	N	Public	https://github.com/harvardnlp_ /boxscore-data
WikiBio	Sentence	728,321 biographies	Pre-process	.tar and .bz2	N	Public	https://rlebret.github.io/ wikipedia-biography-dataset/
Students Reviews	Sentence	5000 reviews	Pre-process	Not mentioned	N	Private	No link
Tweet motion	Sentence	No mentioned	Pre-process	.xlsx	Y	Public	http://saifmohammad.com/WebPages/ TweetEmotionIntensity-dataviz.html
MOOCs Lecture Transcripts	Sentence	878 thousand sentences	Pre-process	csv	Y	Public	https://www.kaggle.com/saurabhshahane/ mooc-lecture-dataset/version/1
Large Movie Review	Sentence	50000 Review	Pre-process	csv	Y	Public	https://ai.stanford.edu/ amaas/data/sentiment/
Obama Speech	Paragraph	Not mentioned	Pre-process	PDF, MP3	N	Public	https://www.americanrhetoric.com/ barackobamaspeeches.htm
Stanford WebNLG	Sentence	21,855 data/text file	Raw	XML, Json	Y	Public	https://gitlab.com/shimorina/webnlg-dataset
Restaurant Dataset	Sentence	Not mentioned	Pre-process	.tsv	Y	Public	https://www.kaggle.com/jurk06/ restaurant-dataset/code
finding a hotel, buying a laptop	Sentence	Not Mentioned	Raw	.dat	Y	Public	https://aclanthology.org/D15-1199/
Prothom Alo	Sentence	Not Mentioned	Pre-process	csv	Y	Public	https://www.kaggle.com/twintyone/prothomalo
MSCOCO	Questions/ Answers	1.7 million QA pairs	Pre-process	Json	N	Public	https://visualgenome.org/api/v0/api_home.html
ROBOCUP	Sentence	656 sentences	Pre-process	hrc	Y	Public	https://sites.google.com/unibas.it/ crowdsounddatasets
Kaggle Arabic Peoms	Sentence	58k Peoms	Pre-process	csv	Y	Public	https://www.kaggle.com/fahd09/ arabic-poetry-dataset-478-2017
USAToday	Sentence	1013569 headlines	Pre-prcoces	csv	Y	Public	https://www.kaggle.com/owen1226/usa-today
Reuters	Document	21,578 documents	Raw	txt	Y	Public	https://archive.ics.uci.edu/ml/datasets/ reuters-21578+text+categorization+collection
Amazon Fine Food review	Sentence	500,000 food reviews	Pre-process	csv	N	Public	https://www.kaggle.com/snap/ amazon-fine-food-reviews
Hindi dataset	Sentence	Not Mentioned	Pre-process	txt	Y	Public	https://github.com/skmalviya/RNNLG-Hindi
Google Sentence Compression(GSC)	Sentence	10,000 sentence	Pre-process	Json	Y	Public	https://github.com/google-research-datasets/ sentence-compression/tree/master/data
Cornell Movie Dialog	Sentence	304,713 utterances	Pre-process	txt	N	Public	https://www.cs.cornell.edu/ cristian/_ Cornell_Movie-Dialogs_Corpus.html
Beer Reviews	Sentence	11.5M reviews	Pre-process	csv	Y	Public	https://www.kaggle.com/rdoume/beerreviews
Data Euro-parliament speeches	Sentence	13000	Pre-process	Not mentiond	Y	Public	https://lindat.mff.cuni.cz/repository/ xmlui/handle/11858/00-097C-0000-0006-AAE0-A
Prothom Alo	Sentence	Not mentioned	Pre-process	csv	Y	Public	https://www.kaggle.com/twintyone/prothomalo
numericNLG	Paragraph	19.6K word	Raw	Json	Y	Public	https://github.com/titech-nlp/numeric-nlg
Prothom Alo	Sentence	Not mentioned	Pre-process	csv	Y	Public	https://www.kaggle.com/twintyone/prothomalo
LOGICNLG	Sentence	122K words	Raw	Json	Y	Public	https://github.com/wenhuchen/LogicNLG

Chinese, Bengali, Arabic, Russian, Korean, Slovak, Spanish, Czech, German, and Macedonian. This information can help researchers which languages lack research in this domain, which languages need to have focused more on, and what possible deep learning approaches could contribute to a specific language. As can be seen from Figure 7, 74% of the articles have worked in the English language, 7% of articles have worked in Chinese, and 4% of articles have worked in Bengali, 2% of the articles generated Arabic and Russian, and 1% of the articles found for rest of languages. Moreover, the detailed summary of languages according to deep learning approaches is shown in Figure 10.

Many resources are available for English and Chinese languages like Dataset, lexical, syntactic, and POS tagging and programming development support. Therefore, both languages are known as rich-resource languages. On the other hand, languages such as Bengali, Arabic, Russian, Korean, Slovak, Spanish, Czech, German, and Macedonian are known as low-resource languages because resource availability is scarce.

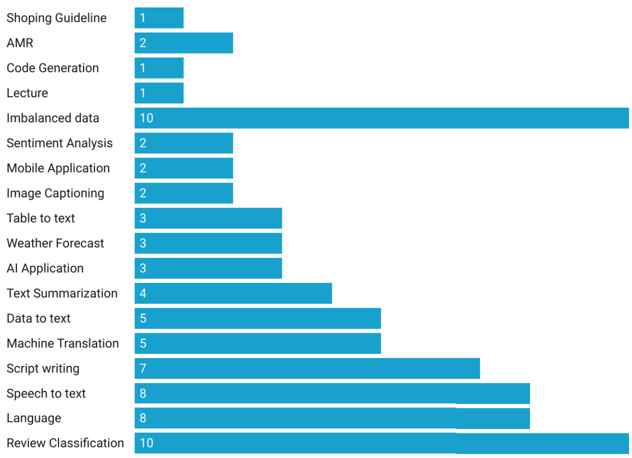


FIGURE 9. Applications of text generation found in the reviewed papers.

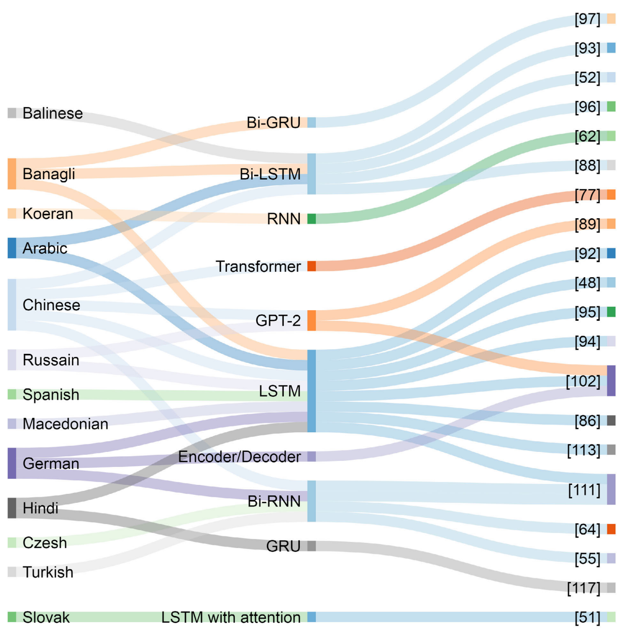


FIGURE 10. A brief summary of language on the basis of deep learning techniques.

TABLE 8. Text generation languages year-wise summary.

Language	Year							Tot.
	2015	2016	2017	2018	2019	2020	2021	
Arabic				1			1	2
Chinese	1		1		2	2		6
Bengali					1	1	1	3
Korean						1		1
Slovak					1			1
Spanish					1			1
Czech			1					1
German				1		1		2
Macedonian					1			1
Russian	1						1	2
English	2	3	4	4	18	21	10	62
Hindi						1	1	2
Balinese							1	1
Turkish			1					1

A brief summary of language on the basis of deep learning techniques is depicted in Table 8, Table 9 and in Figure 10. A variety of traditional and advanced deep learning-based approaches have been used for English text generation, as shown in Table 9. Yet, it needs more experiments with text generation in these languages like Turkish, Hindi, Russian, Macedonian, German, Czech, Spanish, Slovak, Korean, Bengali, and Arabic, as shown in Table 8. In addition to this, a brief summary of language year-wise is depicted in Table 8. The research work on English text generation has fast-growing after 2018 and found 74 studies. After English, we found that studies on Chinese language text generation are constantly growing and found 6 studies. A huge void is left for the other languages to benefit from in this domain.

TABLE 9. Text generation approaches for english language.

Language	Approaches
English	RNN,GRU, LSTM, Bi-RNN, Bi-LSTM (Encoder and Decoder), RNN, GRU, LSTM, Bi-RNN, Bi-LSTM (attention), SeqGAN, TranGAN, Rel-GAN, PAN, VAE GPT2, (RCAM SALSA-TEXT), Transformer, BERT

TABLE 10. Gaps linked to research question.

Identified Gaps	Research Question
Complex language constructs	RQ1
The demand for diversity	RQ1
Improper selection of quality metrics	RQ2
Limited resources	RQ3
Scarcity of datasets	RQ3
Un-standardized source of datasets	RQ3

IV. IDENTIFIED GAPS

In this section, we discuss the major gaps in some areas concerning text generations on the basis of five aspects that need further research and development. The following list shows some gaps which are mapped on our research questions, as shown in Table 10.

1) **Complex language constructs.** Language construct is a piece of language syntax, and every language has its own language constructs. Therefore, it may vary from language to language. For example, the language construct for English sentence is *Subject* + *Verb* + *Object*, whereas for Urdu language it is *Subject* + *Object* + *Verb*. There is no proper way to deal with complex language that requires construct morphology, delexicalised verbs, and abbreviations. Many researchers currently adopt a translation method where low-resource language is being converted to English. However, it mainly provides rigid word order and relatively poor morphology because the technique developed for English may not work for other low-resource languages [130].
2) **The demand for diversity.** Most of the generated text approaches found in the studies discussed in this SLR have a redundant and poor quality text generation problem [21], [49].
3) **Improper selection of quality metrics.** We observed in this survey paper that the selection of quality matrices was improperly employed in many of the studies

found in the literature, and the quality of the generated text was not properly evaluated. For example, the BLEU metric is used to measure the quality of two sentences; thus, it works well for the short sentence-based problem. However, it may not capture the semantic meaning and does not map well to human judgemental capacity. Keeping in view this point, AMR is a seq2seq generation or graph to sequence generation. It is known for semantics. However, many authors have validated the quality of generated text by using BLEU metrics [90].

4) **Limited resources.** Lack of resources such as dictionaries and POS taggers for low-resource languages such as Bengali, Arabic, Russian, Korean, Slovak, Spanish, Czech, German, Urdu, Hindi, Macedonian, etc. There are thirty language published dependency tree-bank reported in [131]. These languages have been working with Google Translate (support 80 different languages), as for the majority of languages, there is no support for NLP resources at all [132].
5) **Scarcity of datasets.** Lack of benchmark datasets for low-resource languages, including Bengali, Arabic, Russian, Korean, Slovak, Spanish, Czech, German, Urdu, Hindi, Macedonian, etc. For instance, we have found the three studies for Bengali text generation [87], [94], [133], and all these studies have used their own extracted dataset. The major reason for using their dataset was the lack of publicly available datasets for the Bengali language. Similarly, there is no dataset available for the Czech language; we have found only one paper for the Czech language in which a multilingual dataset is used [119].
6) **Un-Standardized sources of dataset.** We found that a wide variety of sources of datasets are available, like Quora, GitHub, Kaggle, own website. Sometimes, there is a considerable amount of noise available in datasets; therefore, researchers need to adopt a lot of prepossessing techniques to get the best results from the noisy datasets [134].

V. RECOMMENDATIONS AND FUTURE RESEARCH DIRECTIONS

In this section, we highlight various research directions for researchers in the field, which require considerable efforts to improve the performance of the text generation domain. These research directions are presented below.

A. STANDARDIZED DATASET

Research work is required to develop a few benchmark datasets for Arabic, Chinese, Bengali, Russian, Korean, Slovak, Spanish, Czech, German, Macedonian, and other low-resource languages. The Standardized dataset formation can be at the document level, question/answer form, and paragraph level. In addition to this, we recommend that researchers explore these benchmark datasets: Books3 Stack Exchange, PubMed Abstracts, and CC-2021-04 for various text generation applications, including automatic article summarization and generating synthetic samples to deal with the data imbalance problem.

B. QUALITY METRICS

Our study showed that researchers evaluated generated text using machine and human-based approaches. Nonetheless, a considerable number of research articles failed to evaluate the quality of generated text [17], [40], [49], [87], [115], [133], although they reported excellent results. These results may be biased, in which the experiments that obtained low results may not have been reported. To deal with this issue, we recommend a standard way to evaluate the generated text depending on the nature of the generated text. For instance, for text summarization, ROUGE quality metric is recommended.

Another issue found in the literature is the selection of an inappropriate metric for text generation quality assessment. For example, the BLEU metric is used to measure the quality of two sentences; thus, for a short sentence-based problem, it works well. However, it may not capture the semantic meaning and does not map well to human judgemental capacity. Keeping in view this point, AMR is a seq2seq generation or graph to sequence generation. It is known for semantics. However, many authors have validated the quality of generated text by using the BLEU metrics.

C. TEXT GENERATION IN LOW-RESOURCE LANGUAGES

We have observed a high demand and scope for text generation in low-resource languages in this SLR. A majority of studies have worked on English language text generation. Nonetheless, we have found that 23% of researchers have worked on low-resource languages. Low-resource languages such as Arabic, Spanish, Turkish, Slovak, Hindi, Russian, Macedonian, Czech, Bengali, Korean, Urdu, and alike require significant efforts in this domain. There exist loads of online text thanks to social media and news websites that such languages can benefit from in training language models for text generation. Thus, future research may explore and benefit from available resources for text generation through deep neural network models. Moreover, advanced deep learning approaches for text generation such as GPT-2, BERT, and ELMo should be further considered for exploration by researchers in this field as they have outperformed other methods in the English language [6].

D. USE OF GPT-3 FOR TEXT GENERATION

Existing studies either used traditional or advanced deep learning approaches for text generation. Thus, researchers can emphasize generating text using GPT-3, which is trained on 499 Billion tokens of web data and has 175 billion parameters and 96 layers. It has greater ability and generative power that it may outperform other algorithms in many different tasks like text generation [28].

E. NLP BASIC OPERATIONS IN LOW-RESOURCE LANGUAGES

Standard NLP operations like POS-tagging, tokenization, lemmatization, stemming, word meaning, and related tasks are extremely important in ensuring the quality of the generated text. In low-resource language, there exists an enormous scarcity of these standard basic tasks. Researchers are highly encouraged to come forward and contribute in these areas to further democratize the Internet with increased use of local languages alongside English and other resource-rich languages. It is essential to mention here that loads of mature algorithms are available in the field for these NLP tasks. Data availability is also not an issue. Only a few concentrated efforts are required to work on basic NLP tasks in low-resource languages. These efforts would certainly promote low-resource languages on the Internet.

VI. CONCLUSION

Text generation is the creative side of AI. For decades, computer scientists have been promising humanity to bring artificial intelligence equal to artificial general intelligence. We have to ensure AI can generate text that can pass the Turing test to fulfill these promises. In the past few years, we only recently observed that our dream of synthetic text generation is very close to reality, albeit it is only for a few resource-rich languages.

Text generation has gained wide popularity because a profusion of applications uses them, and there is abundant availability of text online thanks to social media, news outlets, and other sources with enormous usage of text. A few applications which are benefited from text generation include generating and predicting character/word/sentence while typing an email or chatting, chatbot, movie/drama scriptwriting, poetry generation, and many other applications. Moreover, text generation has also been attracting the attention of researchers in the application area of education, industry, and social networks to provide an insight view on different aspects of the approaches. In this context, this systematic literature review provides an analysis of the investigated 90 relevant papers (2015 to 2021) based on text generations in five different aspects; namely text generation approaches, quality metrics, dataset, languages, and applications on text generation in deep learning.

After thoroughly mapping the primary articles, we reviewed them critically to explore different aspects of text generation. For instance, diverse quality metrics are applied to evaluate the generated text. A myriad of approaches is proposed for text generation. A variety of datasets exist; our review concerned their size, format, and applications in which text generation has been applied.

We have provided an overall trend of publications investigating deep learning approaches for text generation throughout the studied years. We have noticed that there is a significant growth of articles published during the year 2018, where the advanced deep learning techniques were mostly represented. In addition to this, text generation in the English language has been more exploited in literature than in any other language.

This systematic literature review will help researchers, academicians, practitioners, and educators who are interested in text generation with data sources, approaches, trends, techniques, and languages.

REFERENCES

[1] H. Jin, Y. Cao, T. Wang, X. Xing, and X. Wan, "Recent advances of neural text generation: Core tasks, datasets, models and challenges," *Sci. China Technol. Sci.*, vol. 63, no. 10, pp. 1–21, 2020.

[2] D. Paper, "Automated text generation," in *TensorFlow 2.x in the Colaboratory Cloud*. Berkeley, CA, USA: Apress, 2021, doi: 10.1007/978-1-4842-6649-6_8.

[3] M. Toshevska, F. Stojanovska, E. Zdravevski, P. Lameski, and S. Gievska, "Explorations into deep learning text architectures for dense image captioning," in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, Sep. 2020, pp. 129–136.

[4] O. Abdelwahab and A. Elmaghraby, "Deep learning based vs. Markov chain based text generation for cross domain adaptation for sentiment classification," in *Proc. IEEE Int. Conf. Inf. Reuse Integr. (IRI)*, Jul. 2018, pp. 252–255.

[5] O. Dušek, J. Novikova, and V. Rieser, "Evaluating the state-of-the-art of end-to-end natural language generation: The E2E NLG challenge," *Comput. Speech Lang.*, vol. 59, pp. 123–156, Jan. 2020.

[6] S. Shaikh, S. M. Daudpota, A. S. Imran, and Z. Kastrati, "Towards improved classification accuracy on highly imbalanced text dataset using deep neural language models," *Appl. Sci.*, vol. 11, no. 2, p. 869, Jan. 2021.

[7] G. Tommi, "Natural language processing in adversarial settings and beyond: Benefits and risks of text classification, transformation, and representation," 2021.

[8] Y. Li, K. Li, H. Ning, X. Xia, Y. Guo, C. Wei, J. Cui, and B. Wang, "Towards an online empathetic chatbot with emotion causes," in *Proc. 44th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2021, pp. 2041–2045.

[9] J. Lee, B. Liang, and H. Fong, "Restatement and question generation for counsellor chatbot," in *Proc. 1st Workshop NLP Positive Impact*. Association for Computational Linguistics, 2021, pp. 1–7.

[10] B. Feng, D. Liu, and Y. Sun, "Evolving transformer architecture for neural machine translation," in *Proc. Genetic Evol. Comput. Conf. Companion*, Jul. 2021, pp. 273–274.

[11] D. Grechishnikova, "Transformer neural network for protein-specific de novo drug generation as a machine translation problem," *Sci. Rep.*, vol. 11, no. 1, pp. 1–13, Dec. 2021.

[12] J. Li, T. Tang, G. He, J. Jiang, X. Hu, P. Xie, Z. Chen, Z. Yu, W. X. Zhao, and J.-R. Wen, "TextBox: A unified, modularized, and extensible framework for text generation," 2021, *arXiv:2101.02046*.

[13] T. Shi, Y. Keneshloo, N. Ramakrishnan, and C. K. Reddy, "Neural abstractive text summarization with sequence-to-sequence models," *ACM Trans. Data Sci.*, vol. 2, no. 1, pp. 1–37, 2021.

[14] D. Khashabi, G. Stanovsky, J. Bragg, N. Lourie, J. Kasai, Y. Choi, N. A. Smith, and D. S. Weld, "GENIE: A leaderboard for human-in-the-loop evaluation of text generation," 2021, *arXiv:2101.06561*.

[15] J. Li, T. Tang, W. X. Zhao, Z. Wei, N. J. Yuan, and J.-R. Wen, "Few-shot knowledge Graph-to-Text generation with pretrained language models," 2021, *arXiv:2106.01623*.

[16] Z. Kastrati, A. S. Imran, and S. Y. Yayilgan, "The impact of deep learning on document classification using semantically rich representations," *Inf. Process. Manage.*, vol. 56, no. 5, pp. 1618–1632, Sep. 2019.

[17] O. Shatalov and N. Ryabova, "Towards Russian text generation problem using OpenAI's GPT-2," CEUR-WS.org, Aachen, Germany, Tech. Rep., 2021.

[18] M. Song, Y. Zhao, and S. Wang, "Exploiting different word clusterings for class-based RNN language modeling in speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 5735–5739.

[19] Z. Kastrati, A. S. Imran, and A. Kurti, "Weakly supervised framework for aspect-based sentiment analysis on Students' reviews of MOOCs," *IEEE Access*, vol. 8, pp. 106799–106810, 2020.

[20] Z. Kastrati, F. Dalipi, A. S. Imran, K. P. Nuci, and M. A. Wani, "Sentiment analysis of Students' feedback with NLP and deep learning: A systematic mapping study," *Appl. Sci.*, vol. 11, no. 9, p. 3986, Apr. 2021.

[21] H. D. Hejazi *et al.*, "Arabic text generation: Deep learning for poetry synthesis," in *Advanced Machine Learning Technologies and Applications*, vol. 1339, no. 1339. Egypt: Springer, 2021, pp. 104–116.

[22] N. Tintarev, E. Reiter, R. Black, A. Waller, and J. Reddington, "Personal storytelling: Using natural language generation for children with complex communication needs, in the wild," *Int. J. Hum.-Comput. Stud.*, vols. 92–93, pp. 1–16, Aug. 2016.

[23] J. Gu, Q. Wu, and Z. Yu, "Perception score, a learned metric for open-ended text generation evaluation," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 14, 2021, pp. 12902–12910.

[24] J. Koutnik, K. Greff, F. Gomez, and J. Schmidhuber, "A clockwork RNN," in *Proc. Int. Conf. Mach. Learn.*, 2014, pp. 1863–1871.

[25] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[27] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.

[28] T. Brown *et al.*, "Language models are few-shot learners," 2020, *arXiv:2005.14165*.

[29] R. Dale, "GPT-3: What's it good for?" *Natural Lang. Eng.*, vol. 27, no. 1, pp. 113–118, 2021.

[30] J. Li, T. Tang, W. X. Zhao, and J.-R. Wen, "Pretrained language models for text generation: A survey," 2021, *arXiv:2105.10311*.

[31] A. Gatt and E. Krahmer, "Survey of the state of the art in natural language generation: Core tasks, applications and evaluation," *J. Artif. Intell. Res.*, vol. 61, pp. 65–170, Jan. 2018.

[32] S. Lu, Y. Zhu, W. Zhang, J. Wang, and Y. Yu, "Neural text generation: Past, present and beyond," 2018, *arXiv:1803.07133*.

[33] A. Celikyilmaz, E. Clark, and J. Gao, "Evaluation of text generation: A survey," 2020, *arXiv:2006.14799*.

[34] C. Garbacea and Q. Mei, "Neural language generation: Formulation, methods, and evaluation," 2020, *arXiv:2007.15780*.

[35] S. Shahriar, "GAN computers generate arts? A survey on visual arts, music, and literary text generation using generative adversarial network," 2021, *arXiv:2108.03857*.

[36] T. Iqbal and S. Qureshi, "The survey: Text generation models in deep learning," *J. King Saud Univ. Comput. Inf. Sci.*, Apr. 2020.

[37] B. Takkouche and G. Norman, "PRISMA statement," *Epidemiology*, vol. 22, no. 1, p. 128, 2011.

[38] D.-H. Vu and A.-C. Le, "Topic-guided RNN model for Vietnamese text generation," in *Research in Intelligent and Computing in Engineering*. Germany: Springer, 2021, pp. 827–834.

[39] L. Li and T. Zhang, "Research on text generation based on LSTM," *Int. Core J. Eng.*, vol. 7, no. 5, pp. 525–535, 2021.

[40] N. I. Akhtar, K. M. I. Shazol, R. Rahman, and M. A. Yousuf, "Bangla text generation using bidirectional optimized gated recurrent unit network," in *Proc. Int. Conf. Trends Comput. Cognit. Eng.* Cham, Switzerland: Springer, 2021, pp. 103–112.

[41] L. Song, A. Wang, J. Su, Y. Zhang, K. Xu, Y. Ge, and D. Yu, "Structural information preserving for graph-to-text generation," 2021, *arXiv:2102.06749*.

[42] L. Logeswaran, H. Lee, and S. Bengio, "Content preserving text generation with attribute controls," *Adv. Neural Inf. Process. Syst.*, vol. 31, 2018, pp. 1–11.

[43] B. Guo, H. Wang, Y. Ding, W. Wu, S. Hao, Y. Sun, and Z. Yu, "Conditional text generation for harmonious human-machine interaction," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 2, pp. 1–50, Apr. 2021.

[44] S. S. Nadakuduti and F. Enciso-Rodríguez, "Advances in genome editing with CRISPR systems and transformation technologies for plant DNA manipulation," *Frontiers Plant Sci.*, vol. 11, p. 2267, Jan. 2021.

[45] J. H. Xu, K. Shinden, and M. P. Kato, "Table caption generation in scholarly documents leveraging pre-trained language models," 2021, *arXiv:2108.08111*.

[46] Y. Zhang, J. Wang, and X. Zhang, "Learning sentiment sentence representation with multiview attention model," *Inf. Sci.*, vol. 571, pp. 459–474, Sep. 2021.

[47] C. Rebuffel, L. Soulier, G. Scoutheeten, and P. Gallinari, "A hierarchical model for data-to-text generation," *Adv. Inf. Retr.*, vol. 12035, p. 65, Apr. 2020.

[48] C. Zhang, C. Xiong, and L. Wang, "A research on generative adversarial networks applied to text generation," in *Proc. 14th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Aug. 2019, pp. 913–917.

[49] Y. Qu, P. Liu, W. Song, L. Liu, and M. Cheng, "A text generation and prediction system: Pre-training on new corpora using BERT and GPT-2," in *Proc. IEEE 10th Int. Conf. Electron. Inf. Emergency Commun. (ICEIEC)*, Jul. 2020, pp. 323–326.

[50] G. Rizzo and T. H. M. Van, "Adversarial text generation with context adapted global knowledge and a self-attentive discriminator," *Inf. Process. Manage.*, vol. 57, no. 6, Nov. 2020, Art. no. 102217.

[51] K. Wang and X. Wan, "Automatic generation of sentimental texts via mixture adversarial networks," *Artif. Intell.*, vol. 275, pp. 540–558, Oct. 2019.

[52] D. Vasko, S. Pecar, and M. Simko, "Automatic text generation in Slovak language," in *Proc. Int. Conf. Current Trends Theory Pract. Informat.* Cyprus: Springer, 2020, pp. 639–647.

[53] H. Shao, J. Wang, H. Lin, X. Zhang, A. Zhang, H. Ji, and T. Abdelzaher, "Controllable and diverse text generation in e-commerce," in *Proc. Web Conf.*, 2021, pp. 2392–2401.

[54] J. Chen, Y. Wu, C. Jia, H. Zheng, and G. Huang, "Customizable text generation via conditional text generative adversarial network," *Neurocomputing*, vol. 416, pp. 125–135, Nov. 2020.

[55] H. Wang, W. Zhang, Y. Zhu, and Z. Bai, "Data-to-Text generation with attention recurrent unit," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8.

[56] C. Xu, Q. Li, D. Zhang, Y. Xie, and X. Li, "Deep successor feature learning for text generation," *Neurocomputing*, vol. 396, pp. 495–500, Jul. 2020.

[57] P. Cai, X. Chen, P. Jin, H. Wang, and T. Li, "Distributional discrepancy: A metric for unconditional text generation," *Knowl.-Based Syst.*, vol. 217, Apr. 2021, Art. no. 106850.

[58] O. O. Marchenko, O. S. Radyvonenko, T. S. Ignatova, P. V. Titarchuk, and D. V. Zhelezniakov, "Improving text generation through introducing coherence metrics," *Cybern. Syst. Anal.*, vol. 56, no. 1, pp. 13–21, Jan. 2020.

[59] Y. Wang, H. Zhang, Y. Liu, and H. Xie, "KG-to-text generation with slot-attention and link-attention," in *Proc. CCF Int. Conf. Natural Lang. Process. Chin. Comput.* China: Springer, 2019, pp. 223–234.

[60] M. A. Haidar, M. Rezagholizadeh, A. Do-Omri, and A. Rashid, "Latent code and text-based generative adversarial networks for soft-text generation," 2019, *arXiv:1904.07293*.

[61] Y. Oualil, M. Singh, C. Greenberg, and D. Klakow, "Long-short range context neural networks for language modeling," 2017, *arXiv:1708.06555*.

[62] D. Park and C. W. Ahn, "LSTM encoder-decoder with adversarial network for text generation from keyword," in *Proc. Int. Conf. Bio-Inspired Comput., Theories Appl.* China: Springer, 2018, pp. 388–396.

[63] V.-K. Tran and L.-M. Nguyen, "Neural-based natural language generation in dialogue using RNN encoder-decoder with semantic aggregation," 2017, *arXiv:1706.06714*.

[64] F. Nie, J. Wang, J.-G. Yao, R. Pan, and C.-Y. Lin, "Operations guided neural networks for high fidelity data-to-text generation," 2018, *arXiv:1809.02735*.

[65] M. Duan and Y. Li, "Penalty-based sequence generative adversarial networks with enhanced transformer for text generation," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–6.

[66] Y. Jianglin, G. Zhigang, and C. Gang, "Recurrent convolution attention model (RCAM) for text generation based on title," *J. Phys., Conf.*, vol. 1168, Feb. 2019, Art. no. 052049.

[67] T.-H. Wen and S. Young, "Recurrent neural network language generation for spoken dialogue systems," *Comput. Speech Lang.*, vol. 63, Sep. 2020, Art. no. 101017.

[68] Q. Liu, T. Li, W. Guan, and S. Li, "Refining data for text generation," in *Proc. China Nat. Conf. Chin. Comput. Linguistics*. China: Springer, 2020, pp. 81–94.

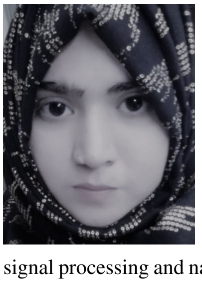
[69] W. Nie, N. Narodytska, and A. Patel, "RelGAN: Relational generative adversarial networks for text generation," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–20.

[70] S. Zhou, "Research on the application of deep learning in text generation," *J. Phys., Conf.*, vol. 1693, no. 1, Dec. 2020, Art. no. 012060.

[71] H. Cheng, J. Cai, and Y. Fang, "RL-Gen: A character-level text generation framework with reinforcement learning in domain generation algorithm case," in *Proc. Int. Conf. Neural Inf. Process.* Australia: Springer, 2019, pp. 690–697.

[72] J. Gagnon-Marchand, H. Sadeghi, M. A. Haidar, and M. Rezagholizadeh, "Salsa-text: Self attentive latent space based adversarial text generation," in *Proc. Canadian Conf. Artif. Intell.* Canada: Springer, 2019, pp. 119–131.
[73] J. Cao and C. Wang, "Social media text generation based on neural network model," in *Proc. 2nd Int. Conf. Comput. Sci. Artif. Intell.*, 2018, pp. 58–61.
[74] Q. Guo, X. Qiu, X. Xue, and Z. Zhang, "Syntax-guided text generation via graph neural network," *Sci. China Inf. Sci.*, vol. 64, no. 5, pp. 1–10, May 2021.
[75] H. Gong, X. Feng, B. Qin, and T. Liu, "Table-to-text generation via row-aware hierarchical encoder," in *Proc. China Nat. Conf. Chin. Comput. Linguistics*. China: Springer, 2019, pp. 533–544.
[76] Y. Yang, J. Cao, Y. Wen, and P. Zhang, "Table-to-text generation with accurate content copying," Sci. Rep., Nature, Tech. Rep. 11, 2021.
[77] H. Wang, Z. Qin, and T. Wan, "Text generation based on generative adversarial nets with latent variables," in *Proc. Pacific-Asia Conf. Knowl. Discovery Data Mining*. Australia: Springer, 2018, pp. 92–103.
[78] X. Chen, D. Lin, and D. Cao, "Text generation from triple via generative adversarial nets," in *Proc. CCF Conf. Comput. Supported Cooperat. Work Social Comput.* China: Springer, 2019, pp. 567–578.
[79] T. Hu and C. Meinel, "Text generation in discrete space," in *Proc. Int. Conf. Artif. Neural Netw.* Slovakia: Springer, 2020, pp. 721–732.
[80] S. Diao, X. Shen, K. Shum, Y. Song, and T. Zhang, "TILGAN: Transformer-based implicit latent GAN for diverse and coherent text generation," in *Proc. Findings Assoc. Comput. Linguistics (ACL-IJCNLP)*, 2021, pp. 4844–4858.
[81] V.-K. Tran, V.-T. Nguyen, K. Shirai, and M.-L. Nguyen, "Towards domain adaptation for neural network language generation in dialogue," in *Proc. 4th NAFOSTED Conf. Inf. Comput. Sci.*, Nov. 2017, pp. 19–24.
[82] J. Jang, H. Noh, Y. Lee, S.-M. Pantel, and H. Rim, "Narrative context-based data-to-text generation for ambient intelligence," *J. Ambient Intell. Hum. Comput.*, vol. 11, no. 4, pp. 1421–1429, Apr. 2020.
[83] V.-K. Tran and L.-M. Nguyen, "Natural language generation for spoken dialogue system using RNN encoder-decoder networks," 2017, *arXiv:1706.00139*.
[84] D. Liu, Y. Xue, F. He, Y. Chen, and J. Lv, "μ-forcing: Training variational recurrent autoencoders for text generation," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 19, no. 1, pp. 1–17, Jan. 2020.
[85] C. Li and W. Xing, "Natural language generation using deep learning to support MOOC learners," *Int. J. Artif. Intell. Educ.*, vol. 31, no. 2, pp. 1–29, 2021.
[86] D. Varshney, A. Ekbal, G. P. Nagaraja, M. Tiwari, A. A. M. Gopinath, and P. Bhattacharyya, "Natural language generation using transformer network in an open-domain setting," in *Proc. Int. Conf. Appl. Natural Lang. to Inf. Syst.* Germany: Springer, 2020, pp. 82–93.
[87] A. K. M. Masum, *A Bengali Text Generation Approach in Context of Abstractive Text Summarization Using RNN*. Singapore: Springer, 2020.
[88] M. W. A. Kesiman and I. M. D. Maysanjaya, "A model for posttransliteration suggestion for balinese palm leaf manuscript with text generation and lstm model," *J. Phys., Conf.*, vol. 1810, no. 1, Mar. 2021, Art. no. 012011.
[89] S. Xu, Z. He, J. Su, L. Zhong, Y. Xu, H. Gu, and Y. Huang, "A shopping guide text generation system based on deep neural network," in *Proc. Int. Conf. Wavelet Anal. Pattern Recognit. (ICWAPR)*, Jul. 2019, pp. 1–5.
[90] V. T. Sinh and N. Le Minh, "A study on self-attention mechanism for AMR-to-text generation," in *Proc. Int. Conf. Appl. Natural Lang. Inf. Syst.* U.K.: Springer, 2019, pp. 321–328.
[91] C. Gao and J. Ren, "A topic-driven language model for learning to generate diverse sentences," *Neurocomputing*, vol. 333, pp. 374–380, Mar. 2019.
[92] A. Souri, Z. El Maazouzi, M. Al Achhab, and B. E. El Mohajir, "Arabic text generation using recurrent neural networks," in *Proc. Int. Conf. Big Data, Cloud Appl.* London, U.K.: Springer, 2018, pp. 523–533.
[93] I. Milanova, K. Sarvanoska, V. Srbinoski, and H. Gjoreski, "Automatic text generation in Macedonian using recurrent neural networks," in *Proc. Int. Conf. ICT Innov.* Macedonia: Springer, 2019, pp. 1–12.
[94] S. Abujar, A. K. M. Masum, S. M. M. H. Chowdhury, M. Hasan, and S. A. Hossain, "Bengali text generation using bi-directional RNN," in *Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2019, pp. 1–5.
[95] N. Dethlefs and A. Turner, "Deep text generation–using hierarchical decomposition to mitigate the effect of rare data points," in *Proc. Int. Conf. Lang., Data Knowl.* Ireland: Springer, 2017, pp. 290–298.
[96] S. Yoon, H. Yun, Y. Kim, G.-T. Park, and K. Jung, "Efficient transfer learning schemes for personalized language modeling using recurrent neural network," in *Proc. Workshops 31st AAAI Conf. Artif. Intell.*, 2017, pp. 1–7.
[97] D. Tarasov, "Natural language generation, paraphrasing and summarization of user reviews with recurrent neural networks," in *Proc. Mater. Int. Conf. Dialog*, 2015, pp. 1–20.
[98] T.-H. Wen, M. Gasic, N. Mrksic, L. M. Rojas-Barahona, P.-H. Su, D. Vandyke, and S. Young, "Multi-domain neural network language generation for spoken dialogue systems," 2016, *arXiv:1603.01232*.
[99] N. Wang and R. R. Issa, *Natural Language Generation From Building Information Models for Intelligent NLP-Based Information Extraction*. Berlin, Germany: Universitätsverlag der TU Berlin, 2020.
[100] K. Chen, F. Li, B. Hu, W. Peng, Q. Chen, H. Yu, and Y. Xiang, "Neural data-to-text generation with dynamic content planning," *Knowl.-Based Syst.*, vol. 215, Mar. 2021, Art. no. 106610.
[101] C. Gulcehre, O. Firat, K. Xu, K. Cho, and Y. Bengio, "On integrating a language model into neural machine translation," *Comput. Speech Lang.*, vol. 45, pp. 137–148, Sep. 2017.
[102] N. Jiang, J. Chen, R.-G. Zhou, C. Wu, H. Chen, J. Zheng, and T. Wan, "PAN: Pipeline assisted neural networks model for data-to-text generation in social Internet of Things," *Inf. Sci.*, vol. 530, pp. 167–179, Aug. 2020.
[103] S. K. Barnwal and U. S. Tiwary, "RNN based language generation models for a Hindi dialogue system," in *Intelligent Human Computer Interaction*, vol. 11886. India: Springer, Dec. 2020, p. 124.
[104] T.-H. Wen, M. Gasic, N. Mrksic, P.-H. Su, D. Vandyke, and S. Young, "Semantically conditioned LSTM-based natural language generation for spoken dialogue systems," 2015, *arXiv:1508.01745*.
[105] T.-H. Wen, M. Gasic, D. Kim, N. Mrksic, P.-H. Su, D. Vandyke, and S. Young, "Stochastic language generation in dialogue using recurrent neural networks with convolutional sentence reranking," 2015, *arXiv:1508.01755*.
[106] S. Chakraborty, J. Banik, S. Addhya, and D. Chatterjee, "Study of dependency on number of lstm units for character based text generation models," in *Proc. Int. Conf. Comput. Sci., Eng. Appl. (ICCSEA)*, Mar. 2020, pp. 1–5.
[107] Y. Gao and C. Wang, "Symmetrical adversarial training network: A novel model for text generation," in *Proc. Int. Conf. Artif. Neural Netw.* Germany: Springer, 2019, pp. 269–280.
[108] A. V. Mota, T. L. C. da Silva, and J. A. F. De Macêdo, "Template-based multi-solution approach for data-to-text generation," in *Proc. Eur. Conf. Adv. Databases Inf. Syst.* France: Springer, 2020, pp. 157–170.
[109] J. A. Laura, G. Masi, and L. Argerich, "From imitation to prediction, data compression vs recurrent neural networks for natural language processing," 2017, *arXiv:1705.00697*.
[110] V.-K. Tran and L.-M. Nguyen, "Gating mechanism based natural language generation for spoken dialogue systems," *Neurocomputing*, vol. 325, pp. 48–58, Jan. 2019.
[111] W. Li, R. Peng, Y. Wang, and Z. Yan, "Knowledge graph based natural language generation with adapted pointer-generator networks," *Neurocomputing*, vol. 382, pp. 174–187, Mar. 2020.
[112] H. V. K. S. Buddana, S. S. Kaushik, P. Manogna, and S. K. P. S., "Word level LSTM and recurrent neural network for automatic text generation," in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2021, pp. 1–4.
[113] T.-H. Wen, M. Gasic, N. Mrksic, P.-H. Su, D. Vandyke, and S. Young, "Semantically conditioned LSTM-based natural language generation for spoken dialogue systems," 2015, *arXiv:1508.01745*.
[114] N. Tintarev, E. Reiter, R. Black, A. Waller, and J. Reddington, "Personal storytelling: Using natural language generation for children with complex communication needs, in the wild," *Int. J. Hum.-Comput. Stud.*, vols. 92–93, pp. 1–16, Aug. 2016.
[115] I. Dhall, S. Vashisth, and S. Saraswat, "Text generation using long short-term memory networks," in *Micro-Electronics and Telecommunication Engineering*. Singapore: Springer, 2020, pp. 649–657.
[116] M. A. Haidar and M. Rezagholizadeh, "TextKD-GAN: Text generation using knowledge distillation and generative adversarial networks," in *Proc. Can. Conf. Artif. Intell.* Canada: Springer, 2019, pp. 107–118.
[117] M. Sundermeyer, H. Ney, and R. Schlüter, "From feedforward to recurrent LSTM neural networks for language modeling," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 3, pp. 517–529, Mar. 2015.

[118] S. Akkaradamrongrat, P. Kachamas, and S. Sinthupinyo, "Text generation for imbalanced text classification," in *Proc. 16th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jul. 2019, pp. 181–186.
[119] H. Deng, L. Zhang, and X. Shu, "Feature memory-based deep recurrent neural network for language modeling," *Appl. Soft Comput.*, vol. 68, pp. 432–446, Jul. 2018.
[120] S. S. Amin and L. Ragha, "Text generation and enhanced evaluation of metric for machine translation," in *Data Intelligence and Cognitive Informatics*. Singapore: Springer, 2021, pp. 1–17.
[121] N. Dethlefs, A. Schoene, and H. Cuayáhuitl, "A divide-and-conquer approach to neural natural language generation from structured data," *Neurocomputing*, vol. 433, pp. 300–309, Apr. 2021.
[122] V. Srinivasan, S. Santhanam, and S. Shaikh, "Using reinforcement learning with external rewards for open-domain natural language generation," *J. Intell. Inf. Syst.*, vol. 56, no. 1, pp. 189–206, Feb. 2021.
[123] Y. Li, Q. Pan, S. Wang, T. Yang, and E. Cambria, "A generative model for category text generation," *Inf. Sci.*, vol. 450, pp. 301–315, Jun. 2018.
[124] Y. Yang, X. Dan, X. Qiu, and Z. Gao, "FGGAN: Feature-guiding generative adversarial networks for text generation," *IEEE Access*, vol. 8, pp. 105217–105225, 2020.
[125] J. Zhao, Z. Zhan, T. Li, R. Li, C. Hu, S. Wang, and Y. Zhang, "Generative adversarial network for Table-to-Text generation," *Neurocomputing*, vol. 452, pp. 28–36, Sep. 2021.
[126] J. Xu, X. Ren, J. Lin, and X. Sun, "Diversity-promoting GAN: A cross-entropy based generative adversarial network for diversified text generation," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2018, pp. 3940–3949.
[127] L. H. Suadaa, H. Kamigaito, K. Funakoshi, M. Okumura, and H. Takamura, "Towards table-to-text generation with numerical reasoning," in *Proc. 59th Annu. Meeting Assoc. Comput. Linguistics 11th Int. Joint Conf. Natural Lang. Process.*, 2021, pp. 1451–1465.
[128] Z. Chen and J. Ren, "Multi-label text classification with latent word-wise label information," *Int. J. Speech Technol.*, vol. 51, no. 2, pp. 966–979, Feb. 2021.
[129] A. P. N. Iyer, "A factorized recurrent neural network based architecture for medium to large vocabulary language modelling," in *Proc. IEEE 10th Int. Conf. Semantic Comput. (ICSC)*, Feb. 2016, pp. 293–300.
[130] O. Dušek and F. Jurčíček, "Neural generation for czech: Data and baselines," 2019, *arXiv:1910.05298*.
[131] D. Zeman, D. Marecek, M. Popel, L. Ramasamy, J. Stepánek, Z. Zabokrtskỳ, and J. Hajic, "HamleDT: To parse or not to parse?" in *Proc. LREC*, 2012, pp. 2735–2741.
[132] B. P. King, "Practical natural language processing for low-resource languages," Ph.D. dissertation, Dept. Comput. Sci. Eng., Univ. Michigan, Ann Arbor, MI, USA, 2015.
[133] A. K. M. Masum, M. M. Islam, S. Abujar, A. K. Sorker, and S. A. Hossain, "Bengali news headline generation on the basis of sequence to sequence learning using bi-directional RNN," in *Soft Computing Techniques and Applications*. Singapore: Springer, 2021, pp. 491–501.
[134] S. Shakeri and A. Sethy, "Label dependent deep variational paraphrase generation," 2019, *arXiv:1911.11952*.



NOUREEN FATIMA received the master's degree in computer science from Sukkur IBA University, Sukkur, Pakistan, in 2021. She is currently a Research Assistant with the Center of Excellence for Robotics, Artificial Intelligence, and Blockchain, Department of Computer Science, Sukkur IBA University. She is the author of several articles published in international journals. Her research interests include applied research in the field of artificial intelligence and its application to signal processing and natural language processing. She served as a Reviewer for IEEE ACCESS.

ALI SHARIQ IMRAN (Member, IEEE) received the master's degree in software engineering and computing from the National University of Sciences and Technology (NUST), Pakistan, in 2008, and the Ph.D. degree in computer science from the University of Oslo (UiO), Norway, in 2013. He is currently associated with the Department of Computer Science, Norwegian University of Science and Technology (NTNU), Norway, as an Associate Professor. He specializes in applied research with a focus on deep learning technology and its application to signal processing, natural language processing, and semantic web. He has over 65 peer-reviewed journals and conference publications to his name. He is a member of the Norwegian Colour and Visual Computing Laboratory (Colourlab) and the Norway Section. He served as a Reviewer for many reputed journals over the years, including IEEE ACCESS, as an Associate Editor.

ZENUN KASTRATI received the master's degrees in computer science through the EU TEMPUS Program developed and implemented jointly from the University of Pristina, Kosovo, the Université de La Rochelle, France, and the Institute of Technology Carlow, Ireland, and the Ph.D. degree in computer science from the Norwegian University of Science and Technology (NTNU), Norway, in 2018. He is associated with the Department of Informatics, Linnaeus University, Sweden. He is the author of more than 40 peer-reviewed journals and conferences. His research interests include artificial intelligence with a special focus on NLP, machine/deep learning, and sentiment analysis. He has served as a reviewer for many reputed journals over the years.

SHER MUHAMMAD DAUDPOTA received the master's and Ph.D. degrees from the Asian Institute of Technology, Thailand, in 2008 and 2012, respectively. He is currently working as a Professor in computer science with Sukkur IBA University, Pakistan. Alongside his computer science contribution, he is also a Quality Assurance Expert in higher education. He has reviewed more than 50 universities in Pakistan for quality assurance on behalf of Higher Education Commission in the role of educational quality reviewer. He is the author of more than 35 peer-reviewed journals and conference publications. His research interests include deep learning, natural language processing, and video and image processing.

ABDULLAH SOOMRO received the M.S. degree in computer science with specialization in data knowledge engineering from Sukkur IBA University, Sukkur, Pakistan, in 2019. From February 2017 to December 2019, he has worked as a Teaching Assistant with the Computer Science Department, Sukkur IBA University, for two years. His research interests include machine learning, deep learning, natural language processing, and computer vision.