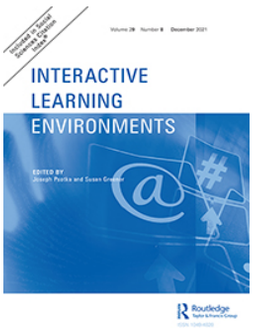

Interactive Learning Environments

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/nile20

Automatic text generation using deep learning: providing large-scale support for online learning communities

Hanxiang Du, Wanli Xing & Bo Pei

To cite this article: Hanxiang Du, Wanli Xing & Bo Pei (2021): Automatic text generation using deep learning: providing large-scale support for online learning communities, Interactive Learning Environments, DOI: 10.1080/10494820.2021.1993932

To link to this article: https://doi.org/10.1080/10494820.2021.1993932

INTERACTIVE LEARNING ENVIRONMENTS
https://doi.org/10.1080/10494820.2021.1993932

Routledge
Taylor & Francis Group

Automatic text generation using deep learning: providing large-scale support for online learning communities

Hanxiang Du, Wanli Xing and Bo Pei

Educational Technology, School of Teaching and Learning, University of Florida, Gainesville, FL, USA

ABSTRACT

Participating in online communities has significant benefits to students learning in terms of students' motivation, persistence, and learning outcomes. However, maintaining and supporting online learning communities is very challenging and requires tremendous work. Automatic support is desirable in this situation. The purpose of this work is to explore the use of deep learning algorithms for automatic text generation in providing emotional and community support for a massive online learning community, Scratch. Particularly, state-of-art deep learning language models GPT-2 and recurrent neural network (RNN) are trained using two million comments from the online learning community. We then conduct both a readability test and human evaluation on the automatically generated results for offering support to the online students. The results show that the GPT-2 language model can provide timely and human-written like replies in a style genuine to the data set and context for offering related support.

ARTICLE HISTORY
Received 3 March 2020
Accepted 11 October 2021

KEYWORDS
Deep learning; artificial intelligence; online learning; language models; text generation; online communities

1. Introduction

Online learning communities are well-supported by various theories of learning that emphasize the role of social interaction in knowledge building (Kožuh et al., 2015) and those suggesting that knowledge is constructed within a social milieu (Hemetsberger & Reinhardt, 2006). Many advantages are associated with learning in online communities including increased motivation, improved learning achievement, and perceived satisfaction (Dawson, 2006; Richardson et al., 2017; Zhao & Kuh, 2004). However, supporting and sustaining online learning communities is rather challenging because the asynchronous nature of online learning cannot guarantee in-time responses from peers or instructors. The lack of active interactions between participants makes them feel isolated and such feelings can lead to dropping out from the community (Lee & Choi, 2011; McInnerney & Roberts, 2004).

A body of literature has suggested the importance of interactions in online learning communities. Interaction is found positively related to learners' satisfaction and motivation in online learning environments (Dawson, 2006; Zhao & Kuh, 2004). Communication among learners is also highly valued by instructors due to their affordance to understand the class and activities (Stephens-Martinez et al., 2014). Many research has explored how teachers and/or forum moderators such as teaching assistants attended to the learning communities strategically to build relationships with students, facilitate communication and interaction among community members, and support the sustainability of the online communities (Swan, 2002; Tait, 2000; Vonderwell, 2003). However, many of these studies relyed on manual efforts and discussed the difficulty of providing timely support for online communities on a regular class (Brook & Oliver, 2003; Yukselturk, 2008), not to mention large-scale online communities with thousands of students.

CONTACT Wanli Xing wanli.xing@coe.ufl.edu

Automatic support, on the other hand, has the potential to achieve aforementioned benefits while managing instructors' workload and improving learning experience. Various tools were developed to identify knowledge gap, recommend learning materials and provide instructors with reference to the class learning process (Goradia & Bugarcic, 2019; Kulkarni et al., 2020; Lakkaraju et al., 2015; Lui et al., 2007; Rafaeli et al., 2008 Xing et al., 2019). Based on the students' participation behaviors in the online community, researchers developed prediction models using machine learning to identify at-risk students automatically so that proper help can be designed to improve their performance (Lakkaraju et al., 2015). Goradia and Bugarcic (2019) reviewed a number of recommender systems in health education and concluded that successfully identifying at-risk students could facilitate timely educational intervention. Aiming to facilitate instructor-student communication and track individual/community learning process, a few studies applied machine learning techniques to cluster and analyze online texts published by learners (Goggins & Xing, 2016; Lui et al., 2007; Trausan-Matu et al., 2014).

Another major area of automated online learning community support research is text generation. Text generation is the usage of computing techniques to automatically produce textual content that appears indistinguishable to human-written text. It can be used to generate reports, answer questions, collect information and prompt for feedback. Some task-based tutoring systems like CycleTalk, WrenchTalk used generated texts to automatically deliver support for learning tasks (Kumar & Rosé, 2011). Tang et al. (2016) also demonstrate the potential use of text generation in essay tutoring by providing automated next word suggestion. Given several words as "seeds" to start with, the language model will generate text accordingly. However, supporting online learning communities has not fully taken the advantage of the rapid development of deep learning on automatic text generation.

In this study, we explore the usage of state-of-art deep learning algorithms for text generation in providing automatic support for a massive online learning community, Scratch. Scratch is a large, online, text-based community for students to learn programming. We examined the power of Recurrent Neural Network (RNN) and GPT-2 for text generation using two million Scratch comments for providing automatic support in the online community. RNN-based language models are presumed to be state-of-the-art in text generation, demonstrating potential in providing writing support with automated word suggestions (Tang et al., 2016). GPT-2 is a powerful deep-learning-based language model released by Google OpenAI in 2019 (Radford et al., 2019). The overall goal of this research is to examine to what extent the deep learning-based text generation can offer efficient and meaningful textual support to the leaners in massive online communities.

2. Research background

2.1. Theoretical foundations

Social support theory is the leading framework for providing support in online communities, as shown in Figure 1. Social support is defined as the service and aid exchanged by individuals

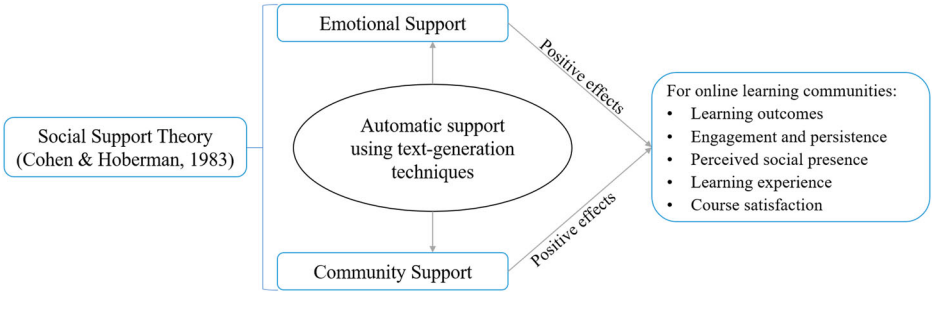

Figure 1. Theoretical framework.

(Cohen & Hoberman, 1983), emphasizing its interactive, helpful, and positive attributes. The positive relationship between social support and intended outcomes is commonly accepted in the research community (Hsu et al., 2018; Langford et al., 1997). Social support theory is conventionally applied in physical environments to explore the relation between individuals. However, as the digital tools and online community flourish, social support theory has been applied to much research in digital environments (Tsai et al., 2017). This study applies automatic text generation techniques to support online learning communities within the social support framework.

The importance of support has long been acknowledged by studies which examined the development of online learning communities. Lally and Barrett (1999) reported the development and analysis of an online learning community. They examined students' online interactions and argued that socio-emotional support is the important basis for academic work. Similarly, Charalambos et al. (2004) shared some interesting findings and valuable lessons in developing an online learning community, STAR-Online. When describing a successful online learning community, they mentioned the following characteristics: "There are capable moderators that provide facilitation, help, guidance and support as needed to the members of the community", and "There is mutual support among its members and sub-groups" (p. 139). The rest of this session focuses on emotional and community support and their influences on online learning communities.

Emotional Support The importance of emotions in online learning communities has long been studied in the past decades. Appropriate strategies to facilitate emotions are indispensable to success in online learning (Barbalet, 2002). In the community of inquiry (CoI) framework developed by Garrison et al. (1999), which has received significant attention and research efforts, emotional expression is identified as an essential component to be socially present in online communities. A body of studies focuses on students' motivation, anxiety, and confusion, among other achievement emotions related to performance in academic settings. In general, students' expression of different types of emotions often leads to different learning outcomes (Baumeister et al., 2007). In the context of large online learning communities such as Scratch, students' emotions such as happiness and anxiety are directly related to their engagement and persistence (Giannakos et al., 2014; Wilson & Moffat, 2010).

Emotional support, addressing emotions of members in online communities, has been found important to learners' learning and outcomes. Emotional support is described as affective assistance (Kahn, 1980), providing help with members in the form of empathy, caring, encouragement, and other sentimental reinforcement (Bambina, 2007; Klaw et al., 2000). It is assumed to be the most important factor to the participants' perceived social support (House, 1981). The importance of emotional support is also identified with online learning communities where it plays a critical role in student perceptions of social presence and collaborative learning (Cleveland-Innes & Campbell, 2012; So & Brush, 2008). Cleveland-Innes and Campbell (2012) examined the self-report data of 217 graduate students who attended online programs and concluded that socio-emotional support is essential online and presents in teaching and cognitive presence as well. Such support contributed significantly to perceived social presence and learning. Moreover, emotional support has been identified critical to student perceptions of collaborative learning, social presence, and satisfaction (So & Brush, 2008). Their work suggests that emotional support is necessary to reduce students' sense of distance and online learning environments shall be designed to provide socio-affective interaction among students. However, all aforementioned studies rely on interactions among students and instructors, while it remains unclear if such a task can be fulfilled through automatic support. By applying automatic text generation into online learning environments, this study aims to investigate to what extent automatic supporting can provide emotional support to online learners.

Community Support Another type of support in the social support framework is community support. Community support is to strengthen interaction and enhance a sense of community (Bambina, 2007), which describes a degree that one feels belonging in a group whose members share common characteristics and dependency with each other (Sarason, 1974). Community

support is the key to the community sustainability (Bruckman, 1998). One goal in community development is to provide mutual support through social activities which lead to a sense of belongings and increased social bonds (Liu et al., 2014; Wills, 1991). In online learning communities, shared social activities may take the form of collaboration, information sharing, and interaction through discussion forums.

Through enhancing the sense of community, community support helps learners with their learning in online learning environments. A number of studies have identified a positive relationship between sense of community and students' engagement, motivation, and learning outcomes (Kang et al., 2007; McInnerney & Roberts, 2004; Phirangee & Malec, 2017; Rovai, 2002; Tinto, 1994; Xing & Gao, 2018). Specifically, students who have strong feelings of community are found more likely to persist in learning with better learning experiences and performance (Rovai, 2002; Tinto, 1994), those who lack sense of community are related with feelings of isolation and distraction, poor learning experiences, and increasing risks of dropping out (Lee & Choi, 2011; McInnerney & Roberts, 2004; Phirangee & Malec, 2017). With all the benefits of community support, facilitating communication among online community members is highly valued. There is a possibility that members would stop visiting the online community if they fail to receive responses or communicate effectively with others. If an online learning community is able to actively prompt communication among members and contribute to their sense of community, learners are expected to benefit from it.

With all being said, emotional support and community support are able to support online learning communities and are expected to help learners persist in learning communities, as well as improve learners' learning experience. However, most existing recommendations focus on prompting interactions among members (Rovai, 2002; So & Brush, 2008; Yang et al., 2017), centering manual work. Guided by social support theory, this study explores the ways of providing emotional and community support using deep-learning-based language models.

2.2. Support in online learning community

Much supporting online communities research has centered manual work. Studies indicate that instructors spend much time and effort in articulating course materials and designing activities that prompt instructor-student, material-student, and student-student interactions in online communities (Dawson, 2006; Rovai, 2002). In order to build relationships with students and develop intervention activities aiming at prompting further interactions, instructors and teaching assistants need to check on the platform regularly, monitor the community development and students' behaviors, and provide timely feedback (Dawson, 2006; Vonderwell, 2003). Moderators are also highly recommended to facilitate and support online learning communities (Charalambos et al., 2004; Lally & Barrett, 1999). Recent work suggests that hub members can also scaffold the online learning community, as they spontaneously follow instructors closely and connect many small learning communities (Brown et al., 2015). However, all these supporting strategies require a large amount of time and manual efforts.

Correspondingly, various automatic methods were developed to support online learning communities. Graf et al. (2009) proposed an automatic approach to identify students' learning styles to enhance instructor–student understanding and facilitate support in an online learning environment. Piech et al. (2015) employed machine learning to model students' online learning as they interacted with coursework. Some language-based agents are also designed to support online learning communities specifically. Kumar et al. (2007) proposed an online learning environment using a dialogue agent to support collaborative learning between students as well as to provide instructions. Ferschke et al. (2015) applied two language-based agents in an online learning environment to facilitate community member help seeking and to prompt celebrative reflection. Their study demonstrated the combination of communication contexts provided by agents could benefit students' engagement in discussion.

2.3. Deep learning language models

Deep learning is one of the most popular topics in computer science. As technology advances quickly, deep learning has achieved outstanding performance in many fields like pattern recognition, machine translation, and natural language processing. Deep neural networks and transformers are the most popular structures for implementing language models (Io & Lee, 2017; Vinyals et al., 2015). Language models implemented through these sophisticated algorithms are able to generate human-written like texts, hold human–machine conversation, parse and analyze input texts, make recommendations based on textual data analysis results, and so on. Deep learning language models have been widely used in many industries. Given sufficient data, the language model can generate various reports on games, sports, weather, natural disasters, or traffic (Garcelon et al., 2018; Gkatzia et al., 2016; Io & Lee, 2017). This study aims to apply cutting-edge language models to provide conversational support for online learning communities.

3. Data

Scratch is a massive online programming learning community released in 2006. According to the platform, over 47 million registered users worldwide created 24 million studios leaving 221 million comments till November 2019. To benefit early programming, one of core design principles for Scratch is to be more social than other programming environments (Resnick et al., 2009). The concept of sharing is achieved by a series of learning activities. Based on the provided blocks or others' uploaded work, learners are able to create various projects like games, book reports, animation videos, music videos, greeting cards, and tutorials, etc. Besides viewing, downloading, and revising others' projects, users can leave comments or vote for a project as well. These activities generate a tremendous amount of data. This study employed a data set of 2 million comments collected from the Scratch online community as the training set (Hill & Monroy-Hernández, 2017). According to the website, the number of registered users on Scratch peaks at the age of 12. In comparison with random textual data, our data set will be able to generate conversation genuine to the Scratch learning context.

4. Methods and evaluation

Figure 2 overviews the methodology of this study. First, collected data are preprocessed for training purposes (see 4.1). As a result, redundant data and features are removed. Second, we compare the model performance between an RNN language model and a GPT-2 Small model using partial preprocessed data (see 4.2 and 4.3). Automatic measurements are applied during this step (see 4.4). Third, we use the whole data set to train a GPT-2 Small model. Then, we apply both automatic and human evaluation (see 4.4) to assess the performance of our model.

4.1. Data preprocessing

Some textual contents are supposed to be removed as they are unnecessary to the goal and may compromise the model performance. (1). Non-English texts. Although Scratch has set separate forums of different languages for its users all over the world, users may still post in any forum using any language. This study focuses on English only and removes all other texts. (2). Uniform Resource Locators (URLs), or web addresses. Numerous comments are associated with one or many URLs to share or distribute users' projects/interests. However, URLs' content does not make sense to human conversation, whether it is legitimate or not. Therefore, URLs are removed. (3). Other invalid characters and texts, such as non-ASCII (American Standard Code for Information Interchange) characters and empty strings. Computationally, invalid values are unacceptable to the model. These are associated with the way how data are stored in original files and decoding

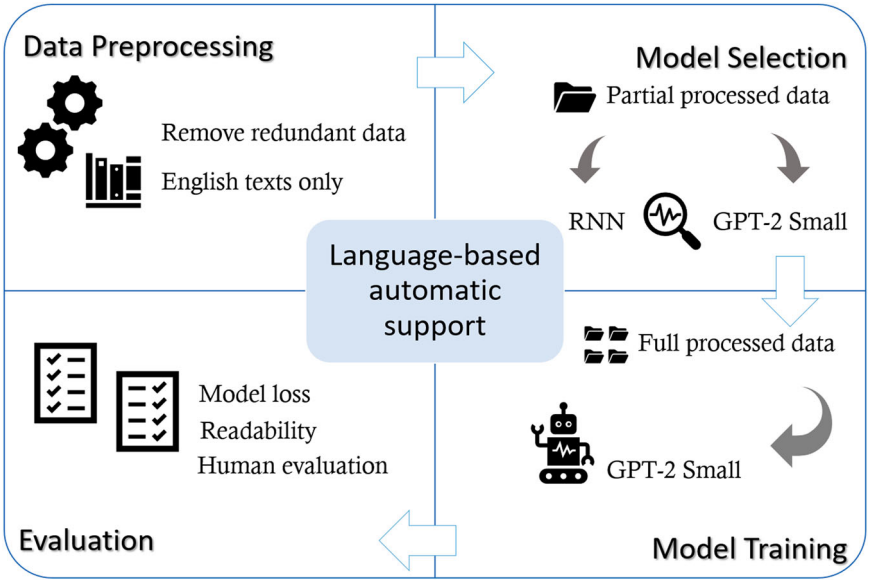

Figure 2. Methodology overview.

input data to train the model. Without proper data preprocessing, issues (1) and (2) result in a model generating texts with non-English terms and/or seemingly URLs, while issue (3) leads to experiment errors.

Several libraries are commonly used to detect language in Python 2, they are *TextBlob*, *langdetect* and *langid*. With the provided Google Translate API, library *TextBlob* can be used for language translation and detection. To detect one text is considered as one visit. To filter out all non-English texts of our data set needs about 2 million visits. The API, however, denies frequent visits in concerns of resources abuse, which makes it unsuitable to our study. Library *langdetect* is also undesirable as its performance on short texts is non-deterministic. In other words, the same text may end up with a different detection result each time. In comparison with *langdetect* and other language detectors, *langid* achieves a higher accuracy across all domains, including microblogs (Lui & Baldwin, 2012). This paper employs *langid* to address issue (1). Other irrelevant texts are removed via regular expressions and ASCII indexes in Python 2.

4.2. Recurrent Neural Network

A deep conventional neural network (CNN) typically has three layers: an input layer, a hidden layer with multiple layers, and an output layer. The sophisticated architecture in the hidden layer represents and learns from the data, tunes the parameters, or in other words, trains the model. As the multiple layers in the hidden layer always feed the data forward, CNN is also known as one of the feedforward neural networks. RNN is an adaption of a conventional feedforward neural network. As the name suggests, RNN has a recurrent hidden state in the hidden layer that integrates information from the current time and previous time. Figure 3 shows the structure of a standard RNN. The left-hand side is an overall structure and the right-hand side is the unfolded hidden layer. This structure enables RNN's capability for dealing with sequential data. The generative result of an RNN is a probability distribution over the next value following the given sequence.

For sequential data, the order of each input matters. The order of words is important to a sentence. If the order is changed, the meaning of the sentence may change greatly. The same rule applies to paragraphs and documents. Therefore, texts can be taken as sequential data in deep

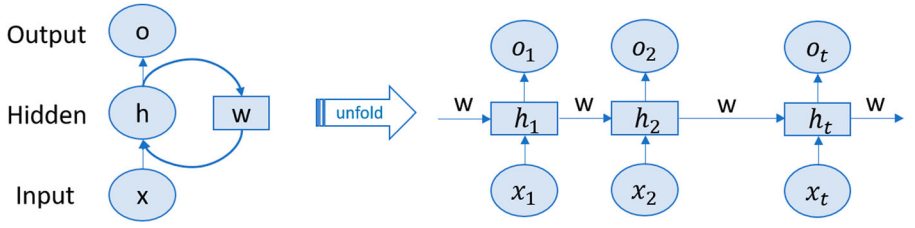

Figure 3. A standard RNN structure.

learning. RNN has been widely used to deal with sequential data, and various RNN-based language models has been developed accordingly. As a matter of fact, RNN-based language models are presumed to be state of art in many text generation fields (Vinyals et al., 2015).

4.3. GPT-2

GPT-2 is a transformer-based language model released by Google OpenAI in 2019. Transformer is an attention-mechanism-based architecture (Vaswani et al., 2017). It transforms one sequence into another one with the help of Encoder or Decoder without implying any Recurrent Networks. Encoder positionally encodes the sequence. Decoder decodes the positional sequence. Despite being extremely sophisticated, the goal of GPT-2 is rather clear: to predict the next following word to the previous one.

GPT-2 is trained on a data set of 40GB of Internet texts ending with 1.5 billion parameters. It has shown such outstanding performance on text generation that Google decides not to release the fully trained model in fear of malicious use such as fake news or e-mail composition. As a matter of fact, GPT-2 holds state of art results in 7 out of 8 testing languages (Radford et al., 2019). A small version with 124M parameters and a medium version of 345M parameters are released to the public for research and experiments. The small version has 12 stacking blocks and the medium version has 24 stacking blocks (Figure 4). In this study, we use a small data set of one month's comments to train a standard RNN language model and a GPT-2 Small, respectively. Using automatic measurement, we will select one language model and train it with the whole processed data set.

4.4. Evaluation

Natural Language Generation (NLG) evaluation methodologies are conventionally broken down into the following categories: intrinsic methods that involve subjects to read and rate the texts, like output quality and user rating; extrinsic methods to evaluate how the performance effects user or system task success (Belz & Reiter, 2006). These methods either employ automated measurements

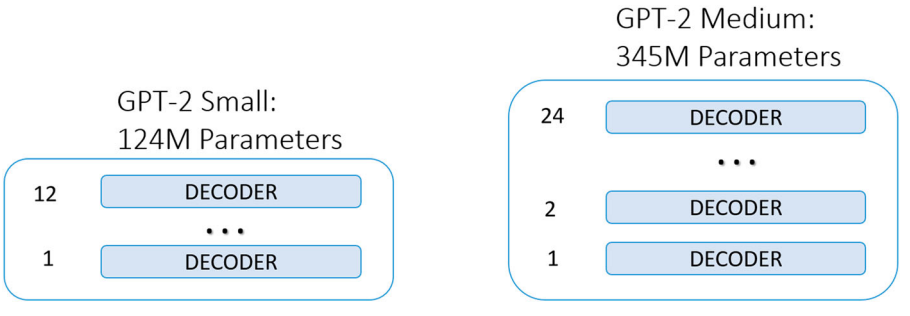

Figure 4. Stacking blocks of GPT-2.

coherent to the model or involve human effort to evaluate the performance (Gkatzia & Mahamood, 2015; Hastie & Belz, 2014). This research adopts both automatic measurement and human evaluation. The model loss which is used to track the training process helps to determine whether the training is completed. The Flesch-Kincaid Grade Level is an automated assessment of output quality which measures the understandability of an English text. In addition, human evaluation is used to assess community and emotional support delivered in generated texts.

Flesch-Kincaid (F-K) Grade Level Flesch-Kincaid Grade Level was first designed to test the understandability of an English text for the US Navy (Kincaid et al., 1975) and later became a military standard. It is also widely used in research, for example, to access readability of medical or educational publications (Dufty & Graesser, 2006; Williamson & Martin, 2010). Unlike some other readability tests, F-K Grade Level calculation does not require a minimal text length. Sentence length and word length are the building blocks of *F-K Formula*'s, which is described as:

$$\text{FK Grade Level} = 0.39*\left(\frac{\text{total words}}{\text{total sentences}}\right) + 11.8*\left(\frac{\text{total syllables}}{\text{total words}}\right) - 15.59$$

where the output is the approximate years of education or grade level needed to understand the text. This is implemented with the Python *textstat* library.

Human Evaluation In addition to F-K Grade Level, we also manually examine the automatic generated results to evaluate the magnitude of community and emotional support. Specifically, we compare the quality of support provided in the original replies with the support in the automatic replies generated by the deep learning language model. According to the power analysis results with a medium effect size, 79 comments with a reply are drawn randomly from the data set. Using each comment as a prompt, the trained model is able to generate an automatic reply accordingly. Then, coders rate the replies without knowing which reply is which. In the following of this study, we will use prompt, original reply, and generated results to describe the original comment, reply, and generated reply, respectively.

The studies of Lally and Barrett (1999) and Charalambos et al. (2004) have concluded a number of characteristics for successful online communities: mutual support among members; community-based facilitation, guidance, and support to members; and opportunities for socio-emotional discourse exchange. Adopting from their framework, several indicators are used for coders to examine the emotional support in a reply: (1) Whether there is an affective expression, including the use of acronyms such as "XD" and "LOL". (2) Whether the affective expression properly and positively addresses the current situation. An affective expression which indicates sarcasm shall not be considered supportive. (3) Whether the affective expression shows empathy, caring, encouragement, and other sentimental reinforcement with the member. Several other indicators are used for examining community support: (1) Whether the reply addresses the issues of a member. (2) Whether the reply is able to prompt interaction. (3) Whether the reply provides recognition, suggestions, and help, and increases the sense of belonging. Given a prompt, raters are asked to rate both the original reply and the generated reply without knowing which reply is which. Rating addresses the following questions:

- I found the reply provides emotional support.
- I found the reply provides community support.

The rating falls in a 1–3 scale, where 1 indicates *Weak*, 2 for *Medium*, and 3 for *Strong*. When coders can hardly or vaguely identify a type of support, the reply is thought to be *weak* in support. *Medium* support indicates that the type of support can be explicitly identified by meeting one or two indicators, yet support is not the theme of the reply. *Strong* support suggests one that can be explicitly identified and support is the theme. Examples are shown in Table 1. For each example, the first sentence is a prompt, the second one is a reply. The reply is either original or generated by the model. All texts are shown as what they are without grammatical correction.

Table 1. Examples for rating.

	Rate	Examples
Emotional support	*weak*	– "Why does he put EVERY SINGLE PROJECT I MAKE IN THIS GALLERY?! IT GETS … . SOOOOOOOO. ANNOYING!"
		– "If youre not going to let me, dont post anything."
	medium	– "stupid being a ghost … .."
		– "no idea. xD"
	strong	– "My stomach has been hurting reeally bad all day DX"
		– "hope it gets bettur!!"
Community support	*weak*	– "anyone here?"
		– "i might have to go soon … "
	medium	– "im not sure when it will come out i will go check now wait."
		– "29 of april i presume."
	strong	– "Fang frowned. He had a little sister in the lab … . but she was dead now."
		– "I wonder why Im here, though. Is there anything I can do?"

5. Results

5.1. Model selection

Both models are run on an i7-9700 CPU, 3.00 GHz, 32.0GB RAM computer in Python 3. The RNN model is implemented using TensorFlow with eager execution. GPT-2 is imported from the *gpt_2_simple* module. The selection experiment is executed with one month of data. The average running time for each epoch of the standard RNN model is 1700 s, 30 s for the GPT-2 model. Figure 5 shows the running time of these two language generation models for 20 epochs. For visualization purpose, the vertical axis shows the natural logarithm to the actual running time. GPT-2 runs much faster than a standard RNN model. This study proceeds with the GPT-2 model.

5.2. F-K Grade Level evaluation

We train a GPT-2 Small model using the entire preprocessed data set. After over 400 training steps, the average loss reaches its minimal value at 3.70 and continues to increase after that. Another parameter that affects the performance of generated results is the temperature, whose value controls how predictable the results are. In this study, we set it between 0.8 and 1.2 randomly to cover both predictable and surprising results. We randomly select 300 generated results and calculate their F-K

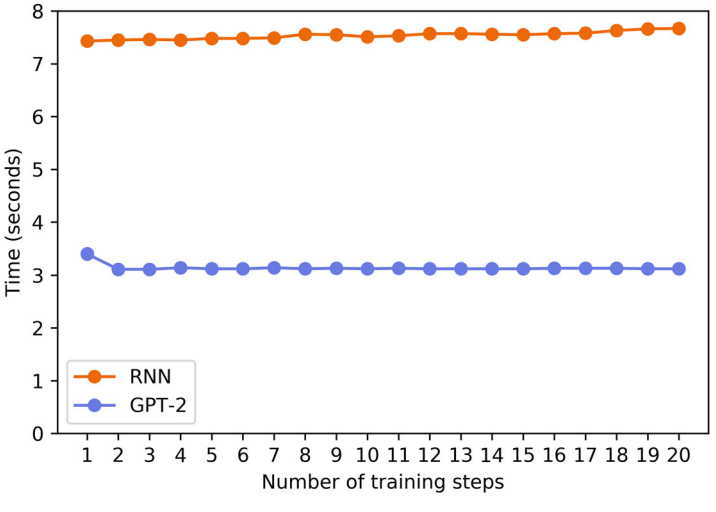

Figure 5. The natural logarithm of running time in seconds.

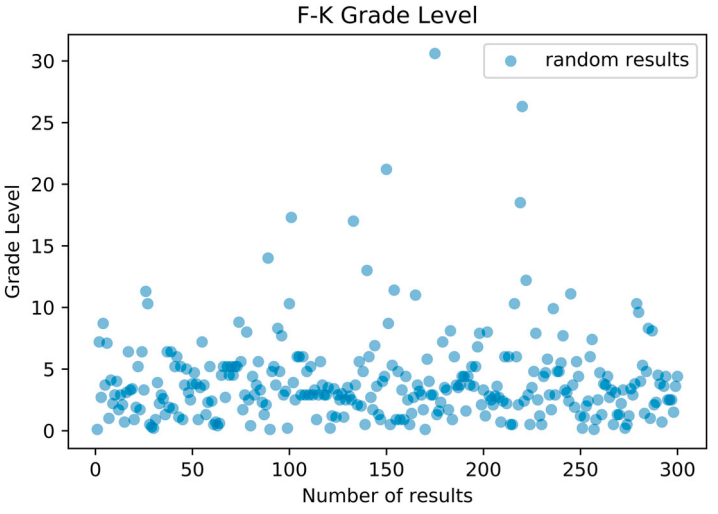

Figure 6. F-K Grade Level of randomly generated results.

Grade Level. Figure 6 shows the overall grade level distribution of the randomly generated results. In the figure, each dot represents one generated result. Most of the samples score between grade 0 and grade 10 and highly congregate under grade 5. This aligns with the user age distribution of the Scratch community, indicating readability of generated results fits the online learning community.

In addition to evaluate the readability of a text, F-K Grade Level has the ability to detect ungrammatical sentences and the outliers. In Figure 6, those few dots which lie away from the cluster can be seen as outliers. Table 2 lists generated results of different F-K Grade Level values. Generated results with a high F-K Grade Level value are indeed hard to understand. One speculation is that the model learns from the Internet expression practice which uses repeated letters or words to address one's opinion or feeling, such as "Soooooooo Annoying" (see Table 1). The usage of acronyms like "XD", "lol", and "btw" by generated results also indicates the model has adopted such an expression style. The average F-K Grade Level score of the randomly generated results is 4.02, with a standard deviation of 3.64. To summarize, most generated results congregate around grade level 5 with a small number of exceptions.

5.3. Human evaluation

We further qualitatively evaluate the results in terms of generating emotional and community support. Two advanced doctoral students with extensive experience in computing and education

Table 2. Examples of generated results with different F-K Grade Level scores.

F-K Grade Level	Randomly Generated Results
30.6	*Lol, basically, that means you were the person I used to call Ren and the dragon, who I go by much more than you, I make up~ I now call yourself the dragon that I call Ren and the dragon that I call Ren and the dragons that I call Ren and the dragons that I call Ren and the dragons that I call Ren and the dragons that I call Ren and the dragons that I call Ren and I believe*
21.2	*DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN DUN*
11.1	*"Yura was talking about her life, but she was thinking of everything else.// If I survive, I will be the leader of a clan that will protect the area"*
7.2	*"I think I was doing a bad job of my friends, because someone is NOT giving me a lot of stuff. XD"*
4.8	*AWESOME btw. im not that bad at math i do it on my laptop and i put it on*
4.8	*"The computer was running down all the rooms, trying to lock the door … "*
3.7	*"Okay, better look at the description. XD"*

Table 3. Human evaluation results.

	Supportive reply percentage		MANOVA		ANOVA
	Original reply	Generated result	Wilk's Λ	*F*	*F*
Emotional support	59.49%	58.23%	–	–	.13
Community support	39.24%	63.29%	–	–	12.28**
Overall	78.48%	79.75%	–	–	–
	–	–	.927	6.103*	–

*indicates significance at $p < .05$.
**indicates significance at $p < .001$.

independently rate the emotional and community support for original reply and generated results on a *Weak*, *Medium*, and *Strong* scale (see 4.4). Then, we calculate their inter-rater reliability for emotional support on original reply, and generated results, as well as community support on original reply, and generated results. The initial Krippendorff alpha coefficients are 0.71, 0.75, 0.70, and 0.64, respectively. After further discussion, raters reach 100% agreement on all ratings.

Table 3 reports results of human evaluation. A supportive reply is defined as a reply whose rating is *Medium* or *Strong*. A reply is considered non-supportive if its rating is *Weak*. Two types of percentage of supportive replies are calculated. First, we calculate the percentage of supportive replies which provide emotional or community support. Supportive reply percentage on emotional (or community) support is the total number of supportive replies that provide emotional (or community) support over the sample size. Then, we calculate an overall supportive replies percentage which includes both emotional and community support. In other words, a reply is considered as overall supportive regardless the type of support it provides. The overall percentage is the total number of overall supportive replies over the sample size.

Table 3 indicates the language model is able to provide emotional and community support. Measured by percentage, generated results provide an equivalent amount of emotional support to the emotional support in original replies ($p > .05$). Meanwhile, both raters find that our deep-learning-based language model provides more community support than the original reply does. There is a statistically significant difference between the amount of community support in original reply and generated result ($p < .05$). Although the difference between overall supportive reply percentage in original reply and generated result is small, MANOVA test results indicate a statistically significant difference between the amount of support provided by original reply and generated result ($p < .05$). To summarize, the deep-learning-based language model is able to provide online learning communities with more emotional and community support than community members do.

6. Discussion

Despite the various advantages of participating in online learning communities, supporting and facilitating online communities, especially massive online communities, are rather challenging (Brown et al., 2015; Swan, 2002; Tait, 2000). Due to their online asynchronous nature and scale, supporting online learning communities raises methodological questions for both researchers and practitioners as they strive to provide timely and quality support for community members. Traditional manual support methods used in small online classes (Rovai, 2002) may not meet the unique requirements posted by large-scale online communities. While there is some exploratory research on automatic support for online communities such as prediction mechanisms (Lakkaraju et al., 2015), and textual support using Quick Helper and Bazaar (Ferschke et al., 2015), these studies did not fully exploit the recent development of deep learning for building automatic support. This study proposed and examined the use of state-of-art deep learning language models to demonstrate their potential use in providing social support for a massive online learning community. The evaluation results show that deep-learning-based language models, especially the GPT-2 model, can provide stable and reliable support automatically to online learning community members.

Such an automatic nature and reliable performance have significant implications for support research in online learning communities. Interaction is the principal way for online learners to perceive social support whose importance has been addressed by a body of literature (e.g. Hsu et al., 2018; Phirangee & Malec, 2017; Rafaeli et al., 2008; Swan, 2002). Meanwhile, prompting interactions among learners has long been recognized as an essential strategy to develop and maintain online learning communities (Charalambos et al., 2004; Dawson, 2006; Lally & Barrett, 1999; Rovai, 2002). However, many comments and questions are left unanswered or have a tremendous delay in response in the online learning community. Even in our own examined 79 examples, the response delay ranges from several minutes to several days or longer. When learners post to seek help or express confusion, delayed responses or the absence of response will not only impact the learning experience, but also potentially lead to the feeling of isolation and drop out (Lee & Choi, 2011; McInnerney & Roberts, 2004). By employing our deep-learning-based automatic textual support in a brutal force manner, these concerns can be relieved significantly.

Our design can also be integrated with other commonly used machine-learning-based techniques to support online learning communities. A number of studies develop classifiers to identify posts of interest, such as urgent posts that need to be addressed immediately, posts that express confusion or seek help (e.g. Almatrafi et al., 2018; Ferschke et al., 2015). Integrated with these classifiers, our design can be deployed to provide in-time support to those members in need. Such an "identify & support" integration leads to a fine framework which provides the online learning community with characteristic support automatically and systematically. Meanwhile, our design can be used to manage the instructor's work load. In an online teaching scenario, deep-learning-based text generation can automatically generate a number of replies at instructors' disposal. Generated responses can either serve as the "seed" for instructors to work with, or be used to reply a comment.

This study has methodological implications as well. Although various frameworks have been used to study support in online learning communities (Richardson et al., 2017; So & Brush, 2008; Zhao & Kuh, 2004) highlighting the importance of social support, sense of community, and emotions to online learning (Bambina, 2007; Cleveland-Innes & Campbell, 2012; Kožuh et al., 2015; So & Brush, 2008;), limited work has actually applied these frameworks to guide the design and development of automatic language support specifically. Guided by the social theory, this study demonstrates deep-learning-based text generation language models can be used to support online learning communities. With sufficient data, members of online learning communities can receive genuine responses like human-written texts. Our work has demonstrated that the automatic generated results using deep language model offer not only a grammatic response, but also uses the Internet writing style. While this study is focused on an online learning community, the methodological framework can definitely be applied in other settings as well.

There are some limitations to this work as well. First, the design of this application is in a one-to-one manner: each comment gets one reply only. The evaluations on emotional and community support are based on such results as well. A richer context like continuous conversations may project different evaluation results. Second, each comment always gets a reply. In other words, each comment has a generated reply no matter what, even if it meant to close a conversation. Therefore, the application treats each comment equally, which may not suit the real world. Third, the readability of generated results cannot be guaranteed. As shown in Table 2, some of the results do not make sense. The randomness of such compromised results may impact users' experiences of a large group in actual use.

This is only an exploratory work where GPT-2 is applied to support a large-scale online learning community for emotional and community support. Much can be done to extend or improve this work. Future work can explore how to integrate a text classifier with a text generation language model to address a specific type of posts. The quality of generated results can be improved by testing the readability of the generated results before displaying it. It is also worth to investigate how to deploy the design, for example, initiating a private conversation page that prompts interpersonal communication, or leaving a response in its original page. Future studies can also replicate the

research findings in other online learning communities to examine their robustness and transferability.

7. Conclusion

This study took an initial step toward building deep-learning-based language models to provide automatic support to massive online learning communities. Specifically, this work employed the state-of-art deep learning algorithm GPT-2 in building such a language model and examined its effectiveness against the RNN algorithm. The results showed that the GPT-2 based text generation model not only outperformed the standard RNN model in execution time but also provided reliable emotional and community support to the learners. The findings have significant implications for large-scale support research and practice in online communities. By introducing deep-learning-based language models into the field, this study extends automatic support research in online learning communities.

Acknowledgements

This work is partially supported by the National Science Foundation (NSF) of the United States under grant number 1901704. Any opinions, findings, and conclusions or recommendations expressed in this paper, however, are those of the authors and do not necessarily reflect the views of the NSF.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Hanxiang Du is a PhD student in Educational Technology, College of Education at University of Florida. Her research interests are educational data analysis, learning analytics and STEM education.

Wanli Xing is an Assistant Professor of Educational Technology at University of Florida. His research interests are artificial intelligence, learning analytics, STEM education and online learning.

Bo Pei received the BS and MS degree in Computer Science in 2011 and 2015, respectively. He is currently a doctoral student in the School of Teaching & Learning, College of Education, University of Florida. His research interests focus on online learning, educational data mining, especially applying the machine learning approaches into the educational settings to identify the different learning patterns from the learning behaviors and build models to analyze the associations between the learning patterns and final learning performances to help the instructors to provide individualized interventions for students.

ORCID

Hanxiang Du http://orcid.org/0000-0002-9081-0706
Wanli Xing http://orcid.org/0000-0002-1446-889X
Bo Pei http://orcid.org/0000-0002-6328-6929

References

Almatrafi, O., Johri, A., & Rangwala, H. (2018). Needle in a haystack: Identifying learner posts that require urgent response in MOOC discussion forums. *Computers & Education*, *118*, 1–9. https://doi.org/10.1016/j.compedu.2017.11.002

Bambina, A.. (2007). *Online social support: the interplay of social networks and computer-mediated communication*. Cambria Press.

Barbalet, J. (2002). Introduction: Why emotions are crucial. *The Sociological Review*, *50*(2_suppl), 1–9. https://doi.org/10.1111/j.1467-954X.2002.tb03588.x

Baumeister, R. F., DeWall, C. N., & Zhang, L. (2007). *Do emotions improve or hinder the decision making process?* Russell Sage.

Belz, A., & Reiter, E. (2006, April 3–7). *Comparing automatic and human evaluation of NLG systems*. 11th Conference of the European Chapter of the Association for Computational Linguistics.

Brook, C., & Oliver, R. (2003). Online learning communities: Investigating a design framework. *Australasian Journal of Educational Technology*, *19*(*2*). https://doi.org/10.14742/ajet.1708

Brown, R., Lynch, C., Wang, Y., & Eagle, M. (2015, June 26–29). *Communities of performance & communities of preference*. 8th International Conference on Educational Data Mining (EDM) Workshops.

Bruckman, A. (1998). Community support for constructionist learning. *Computer Supported Cooperative Work (CSCW)*, *7* (1–2), 47–86. https://doi.org/10.1023/A:1008684120893

Charalambos, V., Michalinos, Z., & Chamberlain, R. (2004). The design of online learning communities: Critical issues. *Educational Media International*, *41*(2), 135–143. https://doi.org/10.1080/09523980410001678593

Cleveland-Innes, M., & Campbell, P. (2012). Emotional presence, learning, and the online learning environment. *The International Review of Research in Open and Distributed Learning*, *13*(4), 269–292. https://doi.org/10.19173/irrodl.v13i4.1234

Cohen, S., & Hoberman, H. M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology*, *13*(2), 99–125. https://doi.org/10.1111/j.1559-1816.1983.tb02325.x

Dawson, S. (2006). Online forum discussion interactions as an indicator of student community. *Australasian Journal of Educational Technology*, *22*(4), https://doi.org/10.14742/ajet.1282

Dufty, D. F., & Graesser, A. C. (2006). Assigning grade levels to textbooks: Is it just readability? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *28*, (28). https://escholarship.org/uc/item/44f184z9

Ferschke, O., Howley, I., Tomar, G., Yang, D., & Liu, Y. (2015). *Fostering discussion across communication media in massive open online courses*.

Garcelon, N., Neuraz, A., Salomon, R., Bahi-Buisson, N., Amiel, J., Picard, C., Mahlaoui, N., Benoit, V., Burgun, A., & Rance, B. (2018). Next generation phenotyping using narrative reports in a rare disease clinical data warehouse. *Orphanet Journal of Rare Diseases*, *13*(1), 1–11. https://doi.org/10.1186/s13023-018-0830-6

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in Higher education. *The Internet and Higher Education*, *2*(2-3), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Giannakos, M. N., Jaccheri, L., & Leftheriotis, I. (2014, June 22–27). *Happy girls engaging with technology: Assessing emotions and engagement related to programming activities*, International Conference on Learning and Collaboration Technologies, (Vol. 8523, pp. 398–409), Springer International. https://doi.org/10.1007/978-3-319-07482-5_38

Gkatzia, D., & Mahamood, S. (2015, September 10–11). *A snapshot of NLG evaluation practices 2005 - 2014*. Proceedings of the 15th European Workshop on Natural Language Generation (ENLG), 57–60. https://doi.org/10.18653/v1/W15-4708

Gkatzia, D., Rieser, V., & Lemon, O. (2016, July 24–29). *How to talk to strangers: Generating medical reports for first-time users*. 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). https://doi.org/10.1109/FUZZ-IEEE.2016.7737739

Goggins, S., & Xing, W. (2016). Building models explaining student participation behavior in asynchronous online discussion. *Computers & Education*, *94*, 241–251. https://doi.org/10.1016/j.compedu.2015.11.002

Goradia, T., & Bugarcic, A. (2019). Exploration and evaluation of the tools used to identify first year at-risk students in health science courses: A systematic review. *Advances in Integrative Medicine*, *6*(4), 143–150. https://doi.org/10.1016/j.aimed.2018.11.003

Graf, S., Kinshuk, & Liu, T. C. (2009). Supporting teachers in identifying students' learning styles in learning management systems: An automatic student modelling approach. *Journal of Educational Technology & Society*, *12*(4), 3–14.

Hastie, H., & Belz, A. (2014, May 26–31). *A comparative evaluation methodology for NLG in interactive systems*. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14).

Hemetsberger, A., & Reinhardt, C. (2006). Learning and knowledge-building in open-source communities. *Management Learning*, *37*(2), 187–214. https://doi.org/10.1177/1350507606063442

Hill, B. M., & Monroy-Hernández, A. (2017). A longitudinal dataset of five years of public activity in the Scratch online community. *Scientific Data*, *4*(1), 170002. https://doi.org/10.1038/sdata.2017.2

House, J. S. (1981). *Work stress and social support*. Reading, MA: Addison-Wesley.

Hsu, J.-Y., Chen, C.-C., & Ting, P.-F. (2018). Understanding MOOC continuance: An empirical examination of social support theory. *Interactive Learning Environments*, *26*(8), 1100–1118. https://doi.org/10.1080/10494820.2018.1446990

Io, H. N., & Lee, C. B. (2017, December). *Chatbots and conversational agents: A bibliometric analysis*. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). https://doi.org/10.1109/IEEM.2017.8289883

Kahn, R. L. (1980). Conboys over the life course: Attachment roles and social support. *Life-span Development and Behavior*, *3*, 253–286.

Kang, I., Lee, K. C., Lee, S., & Choi, J. (2007). Investigation of online community voluntary behavior using cognitive map. *Computers in Human Behavior*, *23*(1), 111–126. https://doi.org/10.1016/j.chb.2004.03.039

Kincaid, J. P., Fishburne, J., Robert, P. R., Richard, L., & Brad, S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*. *Defense Technical Information Center*. https://doi.org/10.21236/ADA006655

Klaw, E., Dearmin Huebsch, P., & Humphreys, K. (2000). Communication patterns in an on-line mutual help group for problem drinkers. *Journal of Community Psychology*, *28*(5), 535–546. https://doi.org/10.1002/1520-6629(200009)28:5<535::AID-JCOP7>3.0.CO;2-0

Kožuh, I., Jeremić, Z., Sarjaš, A., Bele, J. L., & Devedžić, V. (2015). Social presence and interaction in learning environments: The effect on student success. *Journal of Educational Technology & Society*, *18*(1), 223–236.

Kulkarni, P. V., Rai, S., & Kale, R. (2020, July 1–7). *Recommender system in eLearning: A survey*, Proceeding of International Conference on Computational Science and Applications: ICCSA 2019 (pp. 119–126), Springer Singapore. https://doi.org/10.1007/978-981-15-0790-8_13

Kumar, R., Gweon, G., Joshi, M., & Cui, Y. (2007). Supporting students working together on math with social dialogue. *Workshop on Speech … .*

Kumar, R., & Rosé, C. P. (2011). Architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies*, *4*(1), 21–34. https://doi.org/10.1109/TLT.2010.41

Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., & Addison, K. L. (2015, August 10–13). *A machine learning framework to identify students at risk of adverse academic outcomes*. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining – KDD' 15, 1909–1918. https://doi.org/10.1145/2783258.2788620

Lally, V., & Barrett, E. (1999). Building a learning community on-line: Towards socio-academic interaction. *Research Papers in Education*, *14*(2), 147–163. https://doi.org/10.1080/0267152990140205

Langford, C. P., Bowsher, J., Maloney, J. P., & Lillis, P. P. (1997). Social support: A conceptual analysis. *Journal of Advanced Nursing*, *25*(1), 95–100. https://doi.org/10.1046/j.1365-2648.1997.1997025095.x

Lee, Y., & Choi, J. (2011). A review of online course dropout research: Implications for practice and future research. *Educational Technology Research and Development*, *59*(5), 593–618. https://doi.org/10.1007/s11423-010-9177-y

Liu, L., Wagner, C., & Chen, H. (2014). Determinants of commitment in an online community: Assessing the antecedents of weak ties and their impact. *Journal of Organizational Computing and Electronic Commerce*, *24*(4), 271–296. https://doi.org/10.1080/10919392.2014.956609

Lui, A. K.-F., Li, S. C., & Choy, S. O. (2007, July 18–20). *An evaluation of automatic text categorization in online discussion analysis*. Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007), 205–209. https://doi.org/10.1109/ICALT.2007.59

Lui, M., & Baldwin, T. (2012, July 10). *Langid. py: An off-the-shelf language identification tool*. Proceedings of the ACL 2012 System Demonstrations, 25–30.

McInnerney, J. M., & Roberts, T. S. (2004). Online learning: Social interaction and the creation of a sense of community. *Educational Technology & Society*, *7*(3), 73–81.

Phirangee, K., & Malec, A. (2017). Othering in online learning: An examination of social presence, identity, and sense of community. *Distance Education*, *38*(2), 160–172. https://doi.org/10.1080/01587919.2017.1322457

Piech, C., Bassen, J., Huang, J., & Ganguli, S. (2015, December 7–12). *Deep knowledge tracing*. Proceedings of the 28th International Conference on Neural Information Processing Systems, Volume 1. 505–513.

Radford, A., Wu, J., Child, R., Luan, D., & Amodei, D. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, *1*(8), 9–9.

Rafaeli, S., Dan-Gur, Y., & Barak, M. (2008). Social Recommender systems: Recommendations in support of E-learning. In L. A. Tomei (Ed.), *Online and distance learning: Concepts, methodologies, tools, and applications* (pp. 2432–2448). IGI Global. https://doi.org/10.4018/978-1-59904-935-9.ch196

Resnick, M., Silverman, B., Kafai, Y., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., & Silver, J. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60. https://doi.org/10.1145/1592761.1592779

Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior*, *71*, 402–417. https://doi.org/10.1016/j.chb.2017.02.001

Rovai, A. P. (2002). Building sense of community at a distance. *The International Review of Research in Open and Distributed Learning*, *3*(1), 74–85. https://doi.org/10.19173/irrodl.v3i1.79

Sarason, S. B. (1974). *The psychological sense of community: Prospects for a community psychology*.

So, H.-J., & Brush, T. A. (2008). Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education*, *51*(1), 318–336. https://doi.org/10.1016/j.compedu.2007.05.009

Stephens-Martinez, K., Hearst, M. A., & Fox, A. (2014, March 4–5). *Monitoring moocs: Which information sources do instructors value?* Proceedings of the first ACM conference on Learning @ scale, 79–88.

Swan, K. (2002). Building learning communities in online courses: The importance of interaction. *Education, Communication & Information*, *2*(1), 23–49. https://doi.org/10.1080/1463631022000005016

Tait, A. (2000). Planning student support for open and distance learning. *Open Learning: The Journal of Open, Distance and e-Learning*, *15*(3), 287–299. https://doi.org/10.1080/713688410

Tang, S., Peterson, J. C., & Pardos, Z. A. (2016, April 25–26). *Deep neural networks and how they apply to sequential education data*. Proceedings of the Third (2016) ACM Conference on Learning @ Scale - L@S'16. https://doi.org/10.1145/2876034.2893444

Tinto, V. (1994). *Leaving college: Rethinking the causes and cures of student attrition*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226922461.001.0001

Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2014). PolyCAFe – automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning*, *9*(2), 127–156. https://doi.org/10.1007/s11412-014-9190-y

Tsai, H.-Y. S., Shillair, R., & Cotten, S. R. (2017). Social support and "playing around": an examination of how older adults acquire digital literacy with tablet computers. *Journal of Applied Gerontology*, *36*(1), 29–55. https://doi.org/10.1177/0733464815609440

Vaswani, A., Shazeer, N., & Parmar, N. (2017, December 4–9). *Attention is all you need. 31st Conference on Neural Information Processing Systems (NIPS 2017)*. 5998–6008.

Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015, June 7–12). *Show and tell: A neural image caption generator*. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3156–3164. https://doi.org/10.1109/CVPR.2015.7298935

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: A case study. *The Internet and Higher Education*, *6*(1), 77–90. https://doi.org/10.1016/S1096-7516(02)00164-1

Williamson, J. M. L., & Martin, A. G. (2010). Analysis of patient information leaflets provided by a district general hospital by the Flesch and Flesch-Kincaid method. *International Journal of Clinical Practice*, *64*(13), 1824–1831. https://doi.org/10.1111/j.1742-1241.2010.02408.x

Wills, T. A. (1991). *Social support and interpersonal relationships*.

Wilson, A., & Moffat, D. C. (2010). *Evaluating scratch to introduce younger school children to programming*. *PPIG*.

Xing, W., & Gao, F.. (2018). Exploring the relationship between online discourse and commitment in Twitter professional learning communities. *Computers & Education*, *126*, 388–398. https://doi.org/10.1016/j.compedu.2018.08.010

Xing, W., Tang, H., & Pei, B.. (2019). Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of MOOCs. *The Internet and Higher Education*, *43*, 100690. https://doi.org/10.1016/j.iheduc.2019.100690

Yang, X., Li, G., & Huang, S. S. (2017). Perceived online community support, member relations, and commitment: Differences between posters and lurkers. *Information & Management*, *54*(2), 154–165. https://doi.org/10.1016/j.im.2016.05.003

Yukselturk, E. (2008). Investigation of interaction, online support, course structure and flexibility as the contributing factors to students' satisfaction in an online certificate program. *Journal of Educational Technology and Society*, *11*(4), 51–65.

Zhao, C., & Kuh, G. D. (2004). Adding value: Learning communities and student engagement. *Research in Higher Education*, *45*(*2*), 115–138. https://doi.org/10.1023/B:RIHE.0000015692.88534.de