Artificial intelligence inspired multilanguage framework for note-taking and qualitative content-based analysis of lectures

Munish Saini1 · Vaibhav Arora1 · Madanjit Singh2 · Jaswinder Singh2 · Sulaimon Oyeniyi Adebayo1

Received: 16 June 2021 / Accepted: 8 July 2022 / Published online: 18 July 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
With the advent of technology and digitization, the use of Information and Communication Technology (ICT) and its tools for the imperative dissemination of information to learners are gaining more ground. During the process of the conveyance of lectures, it is mostly observed that students (learners) are supposed to take notes (minutes) of the subject matter being delivered to them. The existence of different factors like disturbance (noise) from the environment, learner’s lack of interest, problems with the tutor’s voice, and pronunciation, or others, may hinder the practice of preparing (or taking) lecture notes effectively. To tackle such an issue, we propose an artificial intelligence-inspired multilanguage framework for the generation of the lecture script (of complete) and minutes (only important contents) of the lecture (or speech). We also aimed to perform a qualitative content-based analysis of the lecture’s content. Furthermore, we have validated the performance (accuracy) of the proposed framework with that of the manual note-taking method. The proposed framework outperforms its counterpart in terms of note-taking and performing the qualitative content-based analysis. In particular, this framework will assist the tutors in getting insights into their lecture delivery methods and materials. It will also help them improvise to a better approach in the future. The students will be benefited from the outcomes as they do not have to invest valuable time in note-taking/preparation.

Keywords Information and communication technology (ICT) · Natural language processing (NLP) · Summarization · Thematic analysis · Topic modeling · Speech-to-text

Madanjit Singh
madanjit.csc@gndu.ac.in

1 Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India
2 Department of Computer Science, Guru Nanak Dev University, Amritsar, India
1 Introduction

Modern learning and teaching embrace the usage of Information and Communication Technology (ICT) and other tools in education (Arkorful et al., 2021; Challa et al., 2021; Debis & Al-Edaini, 2021). The existence of the technological aspects helps in grooming and improving the interest of learners (Yokubov, 2021), improvising collaborative learning (Herrera-Pavo, 2021), developing creativity among learners (Collard & Looney, 2014; Mróz & Ocetkiewicz, 2021), and encourages the critical thinking of the learners (Allagui, 2021; Goodsett, 2020; Yulianti et al., 2020). The noticeable utility of technology in education effectively encourages the two modern forms of learning techniques a) Blended learning and b) Flipped learning. Blended learning is an interactive and advanced approach to education, incorporating online educational methods, materials, and techniques in conjunction with the conventional classroom pedagogy (Graham, 2006; Rasheed et al., 2020). Flipped learning approach focuses on engaging the students in active learning over class time (Birgili et al., 2021; Karabulut-Ilgu et al., 2018). The teacher provides the educational materials (lecture notes, presentations, audio lectures, and video lectures) to the students for referring outside the class (either at home or other places). This pedagogical strategy enables the students to realize their responsibility in self-learning and self-regulation (Park & Kim, 2021). The class time is efficiently employed to solve problems, assignments, and exploration or find knowledge (Yoon et al., 2021).

The implication of technical methods and methodology in education has touched new horizons of success. The students can learn in-depth theoretical, practical, and research-specific concepts of their subjects with the availability of different mobile apps (Kazemi et al., 2021; Ling et al., 2014; Liu et al., 2021), educational websites (or web pages) (Cengiz et al., 2021), educational blogs (Jimoyiannis & Angelaina, 2012), social media platforms (van Dijck & Poell, 2017), etc. One activity that lies in common for all learners (students) while using these education strategies (or tools) is the requirement of learners to create (make) the lecture notes (Isaacs, 1994; Luo et al., 2018; Peverly et al., 2014). It purposely facilitates the learners to have a clear understanding of the concepts and subject matter (Isaacs, 1994). The student can refer to these notes to explore further deeper into the context.

Piolat et al. (2005) pointed out that note-taking requires students’ cognitive efforts. Sometimes, it may become severe for the learners (specifically, learners with disabilities) to take notes during lectures (or speech) due to the presence of numerous obstructive components (Suritsky, 1992). The most prevalent challenges are difficulty in hearing the professor’s voice, disturbance in the teaching and learning environment, an issue in understanding the content, trouble with language of communication, unable to pay attention, and low writing speed, among other factors (Suritsky, 1992). To overcome all these issues, we have proposed an artificial intelligence-inspired automatic multilanguage note-taking framework known as LNT (i.e. Lecture Note-Taking) framework.

In particular, we aim to look answer for the following research objectives:
I. To design and create a multilanguage framework for the automatic generation of lecture notes using artificial intelligence.

II. To perform qualitative content-based analysis over the contents of the delivered lecture.

III. Introducing the performance rating metric to assess the quality of lecture contents.

The outcomes of this study will ultimately help the learners (students), to not get annoyed or feel frustrated if for one reason or another the anyhow learner had missed the content while taking notes. The learner can focus more on grasping and gaining knowledge. They can look forward to asking queries freely, and can critically evaluate the contents being delivered. Moreover, It will be a great asset for learners with any hearing impairments. Furthermore, the LNT framework will assist the teachers to explore and evaluate their performance. The teachers (or learners) can look forward to improving and enhancing the skills of delivering (or receiving) lecture content.

2 Related work

In this section, we discuss related research work conducted on explaining, exploring, and evaluating the lecture note-taking process and its strategies. We mentioned the literature on measuring the performance and quality of the contents. Further, we look forward to including details on the existing note-taking methods and comparing them with the proposed framework (LNT).

Peverly et al. (2014) emphasized the requirement of cognitive processes to take efficacious lecture notes. They investigated the relationship between quality of notes and transcription, articulateness, verbal working memory, and expertise in extracting the underlying idea. Kiewra et al. (1991) performed a study on analyzing the note-taking functions and techniques. They examined the three note-taking functions (i.e., encoding, encoding plus storage, external storage) in comparison to three note-taking methods (i.e., conventional, linear, matrix). The outcomes specified the facts on the supremacy of external storage over other note-taking functions. At the same time, they pointed out the efficiency of the matrix note-taking method over the other two methods.

Kim et al. (2009) studied the behavior of university students’ present note-taking. They evaluated the changes that are occurring due to the use of automated note-taking systems. Further, they specified the requirements of future note-taking systems. We extended the work of Kim et al. (2009) by proposing a multilanguage (Indian region-specific) framework for note-taking and measuring the quality of the content delivered by performing a qualitative content-based analysis of the delivered lecture.

In recent times, the use of ICT and other alternative tools are part of modern learning and teaching practices. Boyle and Joyce (2021) illustrated smartpens as the new technical device that ultimately helps students (with learning disabilities)
to take quality notes. Park et al. (2020) performed an extensive exercise of looking inside into the security features of mobile apps used for note-taking. They employed the use of reverse engineering to get insights into the forensic analysis of mobile apps. Huang et al. (2021) introduced the innovative perspective of having a new curriculum design that blends learning style concepts and peers’ learning to enhance students’ interaction and thinking abilities. The outcomes of their study illustrated the significant improvement in students’ academic performance and had an influential impact on student learning motivation.

The teachers (or speakers) in the process of learning and teaching practices, may prefer to speak in single or multi-language to disseminate the information (knowledge) (Kirkpatrick, 2012). Furthermore, many countries have opted for English as the first international language commonly used for teaching (Strevens, 1992). It is observed that students with English as a non-native language have faced problems in understanding and even taking notes in class (Crawford & Candlin, 2013). With the advent of technology and the improvisation of blended learning, the teaching pedagogy has been improved. The teaching methods like lecture capture supported pedagogy have evolved and adapted over the period (Shaw & Molnar, 2011). These techniques are widely used as part of education enrichment strategies.

A plethora of research has been performed by researchers and practitioners elaborating and explaining the concepts of speech-to-text or speech-speech conversion and summarization (Bansal et al., 2017; Furui et al., 2004; Reddy & Mahender, 2013; Wagner, 2005). Furui et al. (2004) applied the speech unit extraction and succession procedure for automatic speech-to-text (or speech-speech) summarizations. Ghadage and Shelke (2016) presented the multilanguage speech-to-text conversion system. They have implied the speech-to-text conversion based on the information contained in it. Specifically, they have explored the advanced feature extraction technique (Mel-frequency Cepstral Coefficient), minimum distance classifier, and support vector machine algorithm for speech classification. Gligorić et al. (2012) executed the research on evaluating the quality of lectures based on a number of metrics. They used the IoT infrastructure for capturing the scene, motion sensing, and audio recording. Uzelac et al., (2018) implemented the system for evaluating the student’s satisfaction by examining the parameters obtained from the physical environment (or surroundings). A summary of the related work in the domain of our study is portrayed in Table 1.

In the present study, we are proposing to design and implement the multilanguage framework (LNT) for examining the insights of the lectures delivered to the students. Specifically, we are targeting the 7 regional languages of the Indian country. The designed system will use the normalization and speech recognition module to suppress the surrounding voices and disturbances and would be converted to text. The LNT framework aims to provide the real-time complete script of the lecture, summarization based on an important idea, thematic statistics, and prominent topics discussed in the lecture, and will also measure the quality of the text delivered to students. This study will be of great importance to Indian students in particular as...
| Author/s Name          | Summary of related studies                                                                                                                                                                                                 | Methods/Techniques/Tools used                      | Relevance with the current study                                                                                                                                                                                                 |
|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Kim et al. (2009)     | Analyze the current state, and limitations of automatic note-taking systems and the changes needed in the future note-taking tools/frameworks                                                                                                                                  | Multilanguage framework                          | The proposed work is an extended version of this paper and focuses to cover the existing gaps                                                                                                                                       |
| Kiewra et al. (1991)  | Analyzed the note-taking functions and techniques. They examined the three note-taking functions (i.e., encoding, encoding plus storage, external storage) in comparison to three note-taking methods (i.e., conventional, linear, matrix)                                                                 | Comparative analysis                             | In present study, we extended the related work of note-taking with the improvisation of advanced methods like text summarization, thematic analysis, and content quality analysis |
| Bansal et al., 2017   | Analyze tools used for text-to-speech conversions and speech-to-text                                                                                                                                                                                                              | Analysis of multiple multi-language tools         | Our study used the methods that are best in the class of text and speech conversions as per the outcomes of this paper                                                                                                                   |
| Ghadage and Shelke (2016) | Explored the advanced feature extraction technique (Mel-frequency Cepstral Coefficient), minimum distance classifier, and support vector machine algorithm for speech classification                                                                          | Mel-frequency cepstral, minimum distance classifier, SVM | We used the prescribed techniques for the classification of spoken text                                                                                                                                                          |
| Gligorić et al. (2012) | Analyzed the quality of lectures based on a number of metrics. Further, they used the IoT infrastructure for capturing the scene, motion sensing, and audio recording                                                                                                                   | IoT infrastructure, Motion sensors, microphones | We refer to this study for performing the quality check of the lectures on several parameters                                                                                                                                         |
| Uzelac et al. (2018)  | They implemented the system for evaluating the student’s satisfaction by examining the parameters obtained from the physical environment (or surroundings)                                                                                                                                  | Comparative analysis of multiple metrics of quality checks | We refer to measure the parameters needed for the evaluating quality check of the lectures                                                                                                                                         |
most of them have trouble understanding English (non-native language). Moreover, persons with hearing impairments will also be benefited from the existence of this framework.

3 Data analysis and methodology

3.1 The design structure of the LNT framework

In this section, we represent the design layout of the proposed framework (LNT) (refer to Fig. 1). In the foremost step, the lecture is recorded and processed to handle the noise anomalies (Kochetov et al., 2018). The audio input is transformed into text using speech recognition library and Google API. LNT is a multilanguage framework, but to standardize its processing we are considering the conversion of different languages to the English language. Furthermore, this processed input is further explored to achieve the designated objectives (as specified in the previous section) of this research. More detailed analysis and explanation of all the components of the proposed framework are given in the proceeding sections.

3.2 Data collection

The data is collected in the form of audio files. For this purpose, we have performed the recording of the lectures, speeches, or other explanatory contents delivered by the teachers, professors, and other orators. The process of recording and collecting the lectures is not restricted to a single language (English only). The processed framework can work on multiple input languages. We collected data from more than 150 online lectures and around 67 offline lectures delivered on the university campus. Additionally, we tested the framework over many youtube videos and lectures as well. Further, we have processed the recorded inputs to remove or handle background noise or other disturbing sounds by employing the normalization module and audio chunking part. Additionally, the audio file on preprocessing is converted to the textual form, i.e. it is converted to a single standardized format (English). In case the collected content is not in English, LNT will convert it to the English text first using google text to speech (gtts) API.

3.3 Data preprocessing and filtration

Once we get an audio file in the specified language (see Fig. 2), we normalize the pitch of the audio file using the pydub library (refer to Fig. 3). This processing will
specifically handle certain noises caused by *applauses* from the audience or other *outlier noises* (Yang et al., 2010).

The audio file is chunked based on the *silence* or *pauses* into audio chunks by specifying arguments such as threshold value (in dfbs), and minimum silence duration (in milliseconds). After splitting, we move towards the audio-to-text mechanism over each chunk. In this, we use the *Speech recognition library*\(^4\) for recognizing the text from the audio file by specifying the language of the text and by appending the detected text into the file. Further, on converting a single chunk we additionally add ‘.’ at the end of each line. We applied *text pre-processing* and *cleaning* over the collected data. First, we translate the content of the collected data into the English language if it is not originally in English by using *googletrans*\(^5\) library. Afterward, we work on preprocessing of text by handling the text anomalies like removing extra free spaces, removing periods in multi-period abbreviations, removing punctuations, converting plural words to singular words, and converting text to lower case (Thanaki, 2017).

### 3.4 Natural language processing tasks

NLP is a branch of data science that includes the systematic processing of textual data with the end goal to retrieve information effectively (Jain et al., 2018). With this, we can do numerous tasks and solve a wide range of problems that deal with the handling of textual data like name entity recognition, sentiment analysis, speech recognition, thematic analysis, etc.

#### 3.4.1 Tokenization

The process of dividing or chopping the sentence into smaller parts (i.e., tokens) is called tokenization (Michelbacher, 2013). In most cases, the spaces in the sentence are taken as delimiters and on the appearance of space, the whole textual data is chopped into tokens (Setiawan et al., 2020). The tokens play a fundamental role in NLP tasks and are also called building blocks for natural language-related tasks (Grefenstette, 1999). In popular deep learning architectures like Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), and Long short-term memory (LSTM), tokenization is performed to process the raw text. In general, there are different variations of tokenization like White space tokenization, Dictionary-based tokenization, Rule-based tokenization, Penn tree tokenization, Spacy tokenization, Moses tokenization, and sub-word tokenization (Haruechayosasak & Kongthon, 2013). In specific, tokenization is significantly used for preparing feeding data for models, word count dictionary creation, and others. In our proposed methodology, we have employed the White space tokenization method.

---

\(^4\) [https://pypi.org/project/automatic-speech-recognition/](https://pypi.org/project/automatic-speech-recognition/)

\(^5\) [https://pypi.org/project/googletrans/](https://pypi.org/project/googletrans/)
with space as a delimiter. We create a *dictionary of the words* along with their *frequency* and *length of words*. It will help in generating the formative summarization and performing the topic modeling. For further processing, the *stop words* (Othman et al., 2015) and *noise words* (Li et al., 2018) are removed (or either

Fig. 1 Designing layout of the LNT framework

Fig. 2 Input audio waveform

Fig. 3 Normalized audio waveform
ignored) for getting effective outcomes. It will also help in further data processing and decision-making.

3.4.2 Lemmatization

We use lemmatization to convert the words into the root word (Lemma) from the dictionary of words having the same meaning or word structure (Halácsy & Trón, 2006). Unlike stemming, lemmatization not only considers the present word and removes the suffix or postfix but also deals with the proper usage of nouns and verbs. It also examines the context in which the word is used, and accounts for replacing it with a root word of the same meaning (Burns, 2020). It enables the making of a word dictionary of given text within limited and unique words. Moreover, it is mostly preferred while performing text preprocessing as its algorithm involves the trivial way to do lemmatization by dictionary lookup. It works well for straightforward inflected forms whereas handcrafted or automatically learned rules from the annotated corpse are needed. Specifically, we are using WordNetLemmatizer\(^6\) from nltk\(^7\) library that uses the inbuilt Morphy\(^8\) function to return the output of word based on WordNet.\(^8\)

3.4.3 Word2Vec

Word2Vec is based on the different model architectures and optimizations (Ma & Zhang, 2015). It is used for better word representation, as it allows the words of similar meaning to fall under a single umbrella. It is the popular way of learning word embedding, which preserves the relationship between words and gives much better results in deep learning applications (Ganguly et al., 2015). To maintain the association among words different methods exist but the most significant among them is the Continuous Bag-of-Words (CBOW) model and the Continuous Skip-gram model (Bérard et al., 2016) (shown in Fig. 4).

In CBOW, we predict the target center word from the context, as context contains the pool of various words. In this method, we do the one-hot encoding of all the words in the context and feed it into the model with a hidden layer of size equal to the context, and expect the outcome to be in the form of the probability distribution. For this, we have used softmax\(^9\) function that is further compared with the actual values. Further, based on the loss function, the weights are updated, and the algorithm proceeds ahead. In the case of Continuous Skip grams, we have to predict the context concerning the words. Consequently, we feed the input of the center word in the form of a one-hot encoded vector. It generates the vector of words in context. Furthermore, they are compared with the actual word vectors and are used to update the weights. In this way, both the algorithms contribute towards the act of updating

\(^6\) https://www.nltk.org/_modules/nltk/stem/wordnet.html
\(^7\) https://www.nltk.org/
\(^8\) https://www.nltk.org/howto/wordnet.html
\(^9\) https://towardsdatascience.com/softmax-function-simplified-714068bf8156
the word matrics based on the word and its context. Finally, the input weight matrics are taken and multiplied with the vector of the word. It gives an outcome in the form of a word vector and is known as an embedded word. The accuracy of algorithms could be increased by increasing the vector dimensions and training data. But it will result in the consumption of more time and computational power. Although, we can always use pre-trained models to fulfill our purpose to much extent.

3.4.4 Word frequency

It is the count of particular tokens (or words) occurring in the document (Larsen, 1999). According to Zipf’s Law, the number of occurrences of the Nth most frequently occurring word is proportional to 1/N. We use the frequency table for scoring each word that would be used while summarization and specifying the scores. Besides, we ignore the stop words. The Word frequency and word cloud (or dictionary) are significant while performing thematic analysis as we store the key in its root word form (Kumar & Rani, 2021). The unique word with different meanings is given as the key and the count is stored as the value.

3.4.5 Text summarization

To convert the teacher’s (or professor’s) lecture, speech, or explanation into a summarized form we have applied the summarization module. We have performed the text preprocessing, structuring, and then extracted the meaning full and important data from the delivered text. There are two basic forms of Text summarization: a) Extractive (or formative) summarization and b) Abstractive (or generative) summarization. In extractive summarization, the summary is generated from the text itself by looking at the word and phrase frequencies and then stacking them (Tas & Kiyani, 2007) While in abstractive summarization, the entire document is analyzed first and then the summary is generated automatically by using the NLP algorithm. The summary generated by this process may contain the words or set of words that may or may not be from the original document. In this study, we are using the extractive summarization technique. For this, we first clean the text by removing extra white spaces, and punctuations except '.', lowering the case of each word, etc. We do tokenization of each sentence and create the vectors of sentences by using the word2vec technique. Furthermore, we append it over the whole sentence and then take its average to have the sentence vector of fixed length. Afterward, we could create the similarity matrix using cosine similarity which measures the similarity between two vectors of an inner product space (refer to Eq. 1). Our framework uses cosine similarity because of its speed and usability over sparse data.
With the use of a similarity matrix, we represent data in the form of a graph having nodes as sentences and edges will represent the similarities. Advancing further, we arrange the sentences as per their ranks. These ranks are calculated based on the scores of each sentence. Scoring to text could be given by considering the various features such as Frequency of words or even sentence in some cases, Sentence position, Cue words in between the sentences, Similarity for other sentences, sentence length, proper noun, and, sentence reduction (Ferreira et al., 2013). This technique is much similar to the Google page ranking system in a certain context. After that, we print out the first $K$ sentences of ranking, where $K$ is the number of sentences in which we want to produce the summary.

### 3.4.6 Thematic analysis

The qualitative analysis of data is known as Thematic analysis (McGregor & Li, 2019; Yusoff et al., 2018). It is used in various approaches such as familiarization of datasets, generating themes, detecting patterns, reviewing themes, defining and naming themes, and producing reports (Clarke & Braun, 2014). It is highly significant while automating tasks such as note-taking, topic modeling, theme generation,

$$
\text{Similarity} = \cos(\theta) = \frac{A \cdot B}{|A| \cdot |B|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}
$$

![Fig. 4 Continuous bag-of-words (CBOW) model and continuous skip-gram model](https://www.google.com/search/howsearchworks/algorithms/)
and others depending upon the use cases which could save a lot of time, increase accuracy, and to automatizes the whole process. Specifically, we have used Meaning-Cloud\textsuperscript{15} as an add-on to Excel to perform topic modeling and to extract the themes from the delivered lecture. The extracted results help us to filter out the data of interest and do further analysis.

### 3.4.7 Topic modeling

Topic modeling is a type of statistical modeling technique used to classify text in a document to a particular topic (Curiskis et al., 2020). Topic modeling can be taken as a supervised machine learning task in which our program is capable of scanning the whole document, detecting word and phrase patterns which in return automatically clusters the word groups that could aptly describe the document. In the present study, we have used the Latent Dirichlet Allocation (LDA) algorithm for performing the topic modeling. The LDA algorithm follows a geometric approach in which if we have $N$ topics then, we will assume the $N$-dimensional complex has $N$ corners. Dirichlet’s distributions are fed into the multinomials distributions to generate the topic and word with a certain probability. The mathematical representation of the probability of generating the document is shown in Eq. 2 and also illustrated in Fig. 5 (Minka, 2000).

\[
P(W, Z, \theta, \varphi, \alpha, \beta) = \prod_{i=1}^{M} P(\theta_j; \alpha) \prod_{i=1}^{K} P(\varphi_i; \beta) \prod_{t=1}^{N} P(Z_{i,t}|\theta) P(W_{j,t}|\varphi Z_{j,t})
\]

Here $\alpha$ is the pre-document topic distribution, $\beta$ is pre-topic word distribution, $\theta$ is the topic distribution for document $a$, $\varphi$ is the word distribution for topic $c$, $Z$ is the topic for $b$th word in document $a$, and $W$ is the specific word (Ng et al., 2011).

Dirichlet’s distribution performs the arranging (or distribution) of documents (or texts) in the imaginary quadrilateral shape, mathematically represented as shown in Eq. 3 (Ng et al., 2011). The document with a higher probability of being similar to the text of the original document and with more context towards a single topic will be given a higher preference.

\[
f(x_1, ..., x_k; \alpha_1, ..., \alpha_k) = \frac{1}{B(\alpha)}
\]

### 3.5 Evaluation of content quality and its metrics

For evaluating the quality of the content delivered in the lecture, we have used the following quality metrics (as shown in Table 2). While analyzing the calculated metrics, it is observed that each metric has values on a different scale and carries distinct ranges. To measure the value of the quality score ($Q_i$), we need to remove

\textsuperscript{15} https://www.meaningcloud.com/products/exceladdin
heterogeneity among these metrics. For this purpose, we have applied the min–max normalization\(^\text{16}\) (refer to Eq. 4). It will transform the values of these metrics onto the same scale ranging from 0 to 1. Furthermore, we have assumed that each metric carries an equal weightage. We make this assumption as we do not have any such objective to evaluate the weightage of these parameters.

\[
V' = \frac{V - \min(a)}{\max(a) - \min(a)}(\text{newmax}(a) - \text{newmin}(a)) + \text{newmin}(A)
\]  

(4)

Additionally, these metrics are used to calculate the quality score (Qi) (refer to Eq. 5). The higher value of the quality index will signify the higher appropriateness and informativeness of the context delivered and vice-versa.

\[
\text{QualityScore}(Qi) = \text{Flesch}_\text{reading}_\text{ease} + \text{Cohesion} + \text{Coherence} + \text{Entropy}
\]

(5)

4 Results and analysis

4.1 Working of lecture note-taking (LNT) framework

LNT records the lecture (or speech) and feeds the input into the speech-to-text module (refer to Fig. 6). The audio file is normalized first and then chopped into chunks based on silence to mark the end of the sentence. Once the whole audio file is chopped and stored, the speech recognition library\(^\text{17}\) and Google API are used to convert each block of audio into text. Further, the text is appended one after the other in the text file. Then LNT will preprocess the textual data based on word frequency and sentence score and will generate the extractive summary. Furthermore, we aim to analyze the contents of the lecture (or speech) and measure the quality of the text delivered. For this purpose, the LNT framework applies the LDA\(^\text{18}\) algorithm to perform Topic modeling. It will give an output in terms of the prominent topics covered and delivered in the lecture (or speech). Moreover, this framework works on the processed input to extract the notable themes from the lecture using hapaxes,\(^\text{19}\) collocation,\(^\text{20}\) and bigrams techniques.\(^\text{21}\) Lastly, the quality of the delivered lectures is analyzed based on different metrics (as explained in Sect. 3.4) and the quality score (Qi) is generated from all the parameters after performing their normalization.

\(^{16}\) https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html
\(^{17}\) https://pypi.org/project/automatic-speech-recognition/
\(^{18}\) https://scikit-learn.org/stable/modules/generated/sklearn.discriminant_analysis.LinearDiscriminantAnalysis.html
\(^{19}\) https://www.nltk.org/book/ch01.html
\(^{20}\) https://towardsdatascience.com/collocations-in-nlp-using-nltk-library-2541002998db
\(^{21}\) https://towardsdatascience.com/collocations-in-nlp-using-nltk-library-2541002998db
Performance measures and validation of the LNT framework

LNT records the audio (speech), performs processing of voice, and then applies the voice-to-text conversion method to transform the input to the required text format (refer to Sect. 3 and its subsequent sections for more details).

For validation of the LNT framework, we have used two different approaches. Firstly, we evaluated the accuracies using the SpeechRecognition library which gives an accuracy of 92% (Approx). This matrix varies up to 10 to 15 percent while considering gender, origin, the language used, and other factors. Secondly, we consulted four experts: two educationists, one professor, and one Ph.D. researcher working in the field of education. We asked them to manually taking the notes on the same speech (lecture) by using similar criteria as used for the automated lecture note-taking (refer to Sects. 3.2 and 3.3). Further, we calculate the accuracy of notes taken by each of the experts, along with the automated generated lecture notes by LNT, with that of the actual speech (lecture). It is observed that LNT is more accurate in comparison to the manual experts for replicating the notes (see Table 3).

Table 2 Content quality metrics

| Quality Metric          | Description                                                                 |
|------------------------|-----------------------------------------------------------------------------|
| Flesch_reading_ease    | It indicates the level of ease to read English text                         |
| Cohesion               | It represents the use of vocabulary and grammatical structure to make the connection between ideas within the text. It is a vitally important characteristic of good academic writing because it promotes clarity. The sentences in a paragraph within the academic text should all be related to one another. The required cohesion could be achieved by the appropriate use of Pronouns, Lexical signposts, repeating Keywords, and anaphoric nouns |
| Coherence              | It specifies the contextual fitness of the text that contributes to understanding the meaning or message by promoting the thematic integrity of the text |
| Entropy                | It is a measurement of Randomness. Lowers the chaos or randomness lesser is the Entropy |

4.2 Performance measures and validation of the LNT framework

22 https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases#text=In%20202017%2C%20Google%20announced%20that%20significant%20race%20and%20gender%20biases.
We measured the quality of content by evaluating the *readability*, *cohesion*, *coherence*, and *entropy* metrics for the notes taken by four experts and LNT respectively (refer to Table 4). The higher values of the quality measures pinpoint the supremacy of the LNT in comparison to that of manual note-taking (Table 3).

| Table 3 | Comparison of accuracy in note-taking |
|---------|--------------------------------------|
| Expert  | Expert accuracy (in %) | LNT accuracy (in %) |
| Expert 1 | 89 | 89% |
| Expert 2 | 85.6 | 92% |
| Expert 3 | 87.9 | 85% |
| Expert 4 | 81.4 | 91% |

We measured the quality of content by evaluating the *readability*, *cohesion*, *coherence*, and *entropy* metrics for the notes taken by four experts and LNT respectively (refer to Table 4). The higher values of the quality measures pinpoint the supremacy of the LNT in comparison to that of manual note-taking (Table 3).

| Table 4 | Comparison of content quality (on sample lecture) |
|---------|--------------------------------------|
| Metrics | Expert 1 | Expert 2 | Expert 3 | Expert 4 | LNT |
| Readability | 0.82 | 0.84 | 0.73 | 0.77 | 0.92 |
| Cohesion | 0.32 | 0.30 | 0.30 | 0.41 | 0.42 |
| Coherence | 0.35 | 0.38 | 0.32 | 0.31 | 0.45 |
| Entropy | 0.21 | 0.18 | 0.25 | 0.26 | 0.15 |
4.3 Illustration with an example

For illustration, we have taken a sample lecture delivered by an expert in the field of education (refer to Fig. 7). The initialization of LNT has been done to record the lecture while it is being delivered. Meanwhile, LNT store the speech (as an audio file) at the specified location of the local machine. Once the lecture is completed, an audio preprocessing starts and it converts the audio file into textual form by using the SpeechRecognition library (refer to Sect. 3). After the generation of the text file for the delivered lecture, data preprocessing, summarization (using word2vec, word frequency, & sentence priority scoring), and Topic modeling & Theme extraction (refer to Fig. 8), and Content quality analysis (in Table 5) is performed. Moreover, a single score indicating the quality (Qi) of the text is also generated using four quality parameters (Readability, Cohesion, Coherence, and Entropy).

Furthermore, the recorded lecture text is processed (using MeaningCloud) to identify the most recurring themes and topics from the document. The outcome specifies the existence of a wide range of themes and topics in the delivered lecture, with education as the dominant area of discussion (refer to Table 5). Apart from education, there are some other themes like the economy, law, institutes, media, and politics that have also been spotted by the analyzer. Each of the themes comprises several topics. For instance, the education theme contains prominent topics such as

---

![Workflow of LNT on sample speech](image-url)
students, teachers, higher education, education policy, national education policy, etc. Similarly, the topics of newspaper, Television, and cinema are categorized under the theme of media (refer to Fig. 8).

The qualitative content-based analysis is performed on the processed text (of considered example) to extract the content-based metrics (as shown in Table 6).

Table 5 Extracted topics from the example lecture

| Theme  | Topics                                                                 |
|--------|------------------------------------------------------------------------|
| Education | Basic Education, Physics, Teachers, Students, Albert Einstein, Education Policy, Technical Education, National Education Policy, Cost of education, Quality of Education, Secondary Education, Primary Education, Physical Education, Medical Education, Public Education, literacy, Classes |
| Economy  | Employability, Education cost, Economically weaker, Socially backward, rural–urban divide, Tax economics, Employment employability, Gross enrolment |
| Institutes | Secondary School, School, Educational, Universities, Higher Education Commission, University Grants Commission, Curriculum Vitae, Department |
| Law      | Regulation Act, Articles, New education Policy, Rules                  |
| Linguistics | Pedagogy, Questions, Article, Vocation                                    |
| Media    | Program, Cinema, Broadcast, Newspaper                                     |
| Politics | Access to Education, Government, Authorities, India Council, Bill, Policy, Barrier |
| Technology | Technical Education, Digital divide, Digital device, Technology, Digital Access, coding |
| Sociology | Vocational skills, Society, Features, Humanities, Holistic manner         |
For more explanation on the text quality analysis metrics, refer to Sect. 3.4. The value of the calculated readability index depicts that 80% of the contents have a higher level of readability. The more the value of this metric, the more the context is understandable and readable (Hakim et al., 2021). Cohesion and coherence represent the thematic integrity and clarity within the text (Medve & Takač, 2013). The higher value of both metrics signifies the more appropriateness of the content. But generally, during lecture delivery, it is observed that the person delivering the lecture repeats many lines. It possesses a certain amount of duplication. So, an accordingly scale of 0.5 – 0.75 is pretty good (Crossley & McNamara, 2010). Entropy indicates the alignment of the whole content. The preferred value of entropy is either closer to 0 or closer to 1, the scale of 0.5 is considered as more chaos between themes and topics. With these parameters, we find a quality score (Qi) for the processed text (using the formula shown in Eq. 5). The value of Qi lies between 0 and 1. As it approaches near 1, it indicates the higher quality of the delivered text. Whereas if the value is close to 0, it specifies that the quality of the delivered text is not as good i.e. the text is more random, which means less information is present in the text. Moreover, it also points to the fact that the text has less readability, less thematic integrity, and less clarity. In this considered example, the value of Qi is 0.727. It illustrates the presence of less chaos, but more readability, clarity, and thematic integrity in the text.

### Table 6  Content quality metrics

| Metrics     | LNT     |
|-------------|---------|
| Readability | 0.8     |
| Cohesion    | 0.625   |
| Coherence   | 0.592   |
| Entropy     | 0.12    |
| Quality score | 0.727   |

5 Discussions

Education is considered a basic necessity for all human beings (Schulz, 2008). To make education reachable to every part of the nation, each country put forward some initiatives (Goodlad et al., 2004). India is not behind in promoting education and looking to cope with international standards (Thyagaraju, 2017). Indian government makes basic education free and compulsory for every child under the age of fourteen in its National Education Policy 2020 (Batra, 2020). The education sector is considered to be dynamic, where it has to adapt to every change that occurs in human life (Hogg & Hogg, 1995). Everyone knows the present situation where the whole world is suffering from the COVID-19 pandemic. The government forces schools, colleges, universities, and all other institutes to operate from home. Education faced a new revolutionary change, where every learner...
came to the digital world. Though online learning became a necessity to make education happen in this COVID-19 period it came with a few certain advantages and disadvantages. We are focused to overcome one of the disadvantages of online learning in this paper, which is preparing notes from the recorded lecture. Our purposed framework will also work effectively during offline classes where one has to record the lecture during the class and then process it using the LNT framework.

The framework job is divided into three phases; a) convert the recorded audio file into a text file, b) do the content-based analysis of the text, and c) measure the quality of the text. LNT framework has the facility to work in multiple languages. For this reason, firstly, we detect the language of the text and convert it into English if it is in other languages apart from English. The reason for doing this is that the used content analysis algorithms work in a single language. Moreover, to make algorithms work on the multi-language is a complex task and will degrade its performance. We scheduled this job for our future considerations. Once we got the final text file after all of the pre-processing, then the text is analyzed to extract the themes, topics, and summarization of the lecture. This step will very beneficial for the students to understand the concept of the lecture and it will save a lot of their time and energy. Lastly, the lecture content is checked for the quality of the delivered lecture. We mapped the lecture against four quality metrics; readability, cohesion, coherence, and entropy. The LNT framework provided satisfactory results as confirmed by the experts.

5.1 Discussion on the considered example

We processed the LNT framework against several online and offline, multidisciplinary, multilanguage, as well as a wide range of accents & pronunciations, and different voice pitches in lectures. In this paper, we illustrated the results of one such example. The example consists of a 30 min lecture delivered by the speaker. The lecture is recorded by our framework, first audio processing is performed, then text conversion and preprocessing have been carried out. The final version of the text file contains around 1947 words. Later, the text file is processed for its qualitative analysis, the results reveal the formation of 9 major themes and around 55 topics from the text. The extracted topics are education, economy, institutes, media, politics, technology, etc. Each of the themes consists of several topics with education as the core subject. On average there is one theme for every 217 words and a topic for every 36 words. The summarization contains about 284 words. These values vary from lecture to lecture and will depend upon the variation of the content in the lecture. Lastly, we give the quality checkup of the delivered lecture based on four quality metrics and further measure the quality score. The higher value of the quality score index represents the text in the lecture is less chaotic, has higher readability, high thematic integrity, and more clarity. This will help the lecturer to analyze his performance in the lecture and helps to identify his area of improvement.
6 Conclusions and recommendations

Much like all other systems, the education system is also suffering from the COVID-19 pandemic. The current operation of education is one way or other surviving on virtual setups. In this era of online teaching or the general classroom scenario, all students do not have the same potential to match the speed of the tutor. Many of them failed to make notes at the same time and later miss some important points. For such considerations, we introduce the LNT framework by incorporating the use of technology and advancement in the educational system to tackle the issue of creating notes. It enables students/attendees to solely focus on the listening and understanding part instead of making notes side by side. It can make summarized notes, analyze the content of the lecture, and can be operated in regional languages. It also gives constructive analysis to the lecturer/professor regarding his content delivery. LNT offers a quality analysis feature; in which based on readability, coherence, cohesion, and entropy the content is given a certain quality score. The LNT framework provides satisfactory results and it beat the accuracy of the experts in the note-making process. The work will help the students having hearing disabilities, or those who have language problems by preparing quality notes for them. The scope of this model could be extended in many spheres including higher-level meetings, seminars, and summits, among others.

Authors’ contributions All authors contributed to the data collection, analysis, and writing of the manuscript.

Availability of data and materials The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval The research meets all applicable standards concerning the ethics of experimentation and research integrity, and the following is being certified/declared true. As an expert scientist and along with co-authors of the concerned field, the paper has been submitted with full responsibility, following the due ethical procedure, and there is no duplicate publication, fraud, plagiarism, or concerns about animal or human experimentation.

Conflicts of interest/Competing interests It is to specifically state on behalf of all authors that “No Competing interests are at stake and there is No Conflict of Interest” with other people or organizations that could inappropriately influence or bias the content of the paper.

References

Allagui, B. (2021). TED talk comments to enhance critical thinking skills in an undergraduate reading and writing course. Education and Information Technologies, 26(3), 2941–2960.
Arkorful, V., Barfi, K. A., & Aboagye, I. K. (2021). Integration of information and communication technology in teaching: Initial perspectives of senior high school teachers in Ghana. Education and Information Technologies, 26(4), 3771–3787.
Bansal, S., Kamper, H., Lopez, A. and Goldwater, S., 2017. Towards speech-to-text translation without speech recognition. arXiv preprint arXiv:1702.03856.

Batra, P., 2020. NEP 2020: Undermining the Constitutional Education Agenda?

Bérard, A., Servan, C., Pietquin, O. and Besacier, L., 2016, May. MultiVec: a multilingual and multilevel representation learning toolkit for NLP. In The 10th edition of the Language Resources and Evaluation Conference (LREC).

Birgili, B., Seggie, F. N., & Öguz, E. (2021). The trends and outcomes of flipped learning research between 2012 and 2018: A descriptive content analysis. Journal of Computers in Education, 8(3), 365–394.

Boyle, J.R. and Joyce, R.L., 2021. Smartpens: Note-Taking Technology for Students with Learning Disabilities. Learning Disabilities: A Multidisciplinary Journal, 26(1).

Burns, P. J. (2020). Ensemble lemmatization with the classical language toolkit. Linguistic Studies and Essays, 58(1), 157–176.

Cengiz, B. C., Bal, N. G., & Ozkose-Biyik, C. (2021). An evaluation of the websites used by Turkish EFL teachers. Computer-Assisted Language Learning, 22(1), 308–332.

Challa, K., Sayed, A., & Acharya, Y. (2021). Modern techniques of teaching and learning in medical education: A descriptive literature review. MedEdPublish, 10(1), 18.

Clarke, V., & Braun, V., 2014 Thematic analysis. In Encyclopedia of critical psychology Springer, New York, NY(pp. 1947–1952).

Debis, R. M. K., & Al-Edaini, A. H. O. (2021). The effectiveness of a flipped learning-based educational unit in acquiring basic clothing finishing techniques knowledge and skills for female technical college trainees. International Design Journal, 11(1), 405–412.

Van Dijck, J., & Poell, T. (2017). Social media platforms and education. Van Dijck, J. & T. Poell (2018). Social media platforms and education. In The SAGE Handbook of Social Media, pp. 579–591.

Furui, S., Kikuchi, T., Shinnaka, Y., & Horii, C. (2004). Speech-to-text and speech-to-speech summarization of spontaneous speech. IEEE Transactions on Speech and Audio Processing, 12(4), 401–408.

Ganguly, D., Roy, D., Mitra, M. and Jones, G.J., 2015, August. Word embedding based generalized language model for information retrieval. In Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval (pp. 795–798).

Ghadage, Y.H. and Shelke, S.D., 2016, April. Speech to text conversion for multilingual languages. In 2016 International Conference on Communication and Signal Processing (ICCSP) (pp. 0236–0240). IEEE.

Grefenstette, G. (1999). Tokenization. In Syntactic Wordclass Tagging (pp. 117–133). Springer, Dordrecht.

Hakim, A. A., Setyaningsih, E., & Cahyaningrum, D. (2021). Examining the Readability Level of Reading Texts in English Textbook for Indonesian Senior High School. Journal of English Language Studies, 6(1), 18–35.
Halácsy, P., & Trón, V. (2006). September. Benefits of deep NLP-based Lemmatization for Information Retrieval. In CLEF (Working Notes).

Haruechaiyasak, C. and Kongthon, A., 2013, October. LexToPlus: A thai lexeme tokenization and normalization tool. In Proceedings of the 4th Workshop on South and Southeast Asian Natural Language Processing (pp. 9–16).

Herrera-Pavo, M. Á. (2021). Collaborative learning for virtual higher education. Learning, Culture and Social Interaction, 28, 100437.

Hogg, R.V., & Hogg, M.C. (1995). Continuous quality improvement in higher education. International Statistical Review/Revue Internationale de Statistique, pp. 35–48.

Huang, S. B., Jeng, Y. L., & Lai, C. F. (2021). Note-taking learning system: the use of the learning style theory and the peer learning method on computer programming course. Journal of Educational Computing Research, 59(5), 870–895.

Isaacs, G. (1994). Lecturing practices and note-taking purposes. Studies in Higher Education, 19(2), 203–216.

Jain, A., Kulkarni, G., & Shah, V. (2018). Natural language processing. International Journal of Computer Sciences and Engineering, 6(1), 161–167.

Jimoyiannis, A., & Angelaina, S. (2012). Towards an analysis framework for investigating students’ engagement and learning in educational blogs. Journal of Computer Assisted Learning, 28(3), 222–234.

Karabulut-Ilgu, A., Jaramillo Cherrez, N., & Jahren, C. T. (2018). A systematic review of research on the flipped learning method in engineering education. British Journal of Educational Technology, 49(3), 398–411.

Kazemi, S., Tavousi, M., & Zarei, F. (2021). A mobile-based educational intervention on STI-related preventive behavior among Iranian women. Health Education Research, 36(2), 212–223.

Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. Journal of Educational Psychology, 83(2), 240.

Kim, K., Turner, S. A., & Pérez-Quiones, M. A. (2009). Requirements for electronic note taking systems: A field study of note taking in university classrooms. Education and Information Technologies, 14(3), 255–283.

Kirkpatrick, A., 2012. English as an international language in Asia: Implications for language education. In English as an international language in Asia: Implications for language education (pp. 29–44). Springer, Dordrecht.

Kochetov, K., Putin, E., Balashov, M., Filchenkov, A. & Shalyto, A. (2018). Noise masking recurrent neural network for respiratory sound classification. In International Conference on Artificial Neural Networks (pp. 208–217). Springer, Cham.

Kumar, G. K., & Rani, D. M. (2021). February. Paragraph summarization based on word frequency using NLP techniques. In AIP conference proceedings (Vol. 2317, No. 1, p. 060001). AIP Publishing LLC.

Larsen, B. (1999). A trainable summarizer with knowledge acquired from robust NLP techniques. Advances in Automatic Text Summarization, 71.

Li, X., Wang, Y., Zhang, A., Li, C., Chi, J., & Ouyang, J. (2018). Filtering out the noise in short text topic modeling. Information Sciences, 456, 83–96.

Ling, C., Harnish, D., & Shehab, R. (2014). Educational apps: Using mobile applications to enhance student learning of statistical concepts. Human Factors and Ergonomics in Manufacturing & Service Industries, 24(5), 532–543.

Liu, C., Bano, M., Zowghi, D., & Kearney, M. (2021). Analysing user reviews of inquiry-based learning apps in science education. Computers & Education, 164, 104119.

Luo, L., Kiewra, K. A., Flanigan, A. E., & Peteranetz, M. S. (2018). Laptop versus longhand note taking: Effects on lecture notes and achievement. Instructional Science, 46(6), 947–971.

Ma, L., & Zhang, Y. (2015). Using Word2Vec to process big text data. In 2015 IEEE International Conference on Big Data (Big Data) (pp. 2895–2897). IEEE.

McGregor, K. A., & Li, J. (2019). 73. Fake Instagrams for real conversation: A thematic analysis of the hidden social media life of teenagers. Journal of Adolescent Health, 64(2), S39–S40.

Medve, V. B., & Takač, V. P. (2013). The influence of cohesion and coherence on text quality: A cross-linguistic study of foreign language learners’ written production. In Language in cognition and affect (pp. 111–131). Springer, Berlin, Heidelberg.

Michelbacher, L. (2013). Multi-word tokenization for natural language processing. Minka, T. (2000). Estimating a Dirichlet distribution.
Mróz, A., & Ocetkiewicz, I. (2021). Creativity for sustainability: how do polish teachers develop students’ creativity competence? *Analysis of Research Results. Sustainability*, 2021(13), 571.

Ng, K. W., Tian, G. L., & Tang, M. L. (2011). Dirichlet and related distributions: Theory, methods and applications.

Othman, M., Hassan, H., Moawad, R., & Idrees, A. M. (2015). Using NLP approach for opinion types classifier.

Park, S., & Kim, N.H. (2021). University students’ self-regulation, engagement and performance in flipped learning. *European Journal of Training and Development*.

Park, M., Kim, S., & Kim, J. (2020). Research on Note-Taking Apps with Security Features. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 11(4), 63–76.

Peverly, S. T., Garner, J. K., & Vekaria, P. C. (2014). Both handwriting speed and selective attention are important to lecture note-taking. *Reading and Writing*, 27(1), 1–30.

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*(3), 291–312.

Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the online component of blended learning: A systematic review. *Computers & Education, 144*, 103701.

Reddy, B. R., & Mahender, E. (2013). Speech to text conversion using android platform. *International Journal of Engineering Research and Applications (IJERA), 3*(1), 253–258.

Schulz, B. (2008). The importance of soft skills: Education beyond academic knowledge.

Setiawan, R., Budiharto, W., Kartowisastro, I. H., & Prabowo, H. (2020). Finding model through latent semantic approach to reveal the topic of discussion in discussion forum. *Education and Information Technologies, 25*(1), 31–50.

Shaw, G. P., & Molnar, D. (2011). Non-native english language speakers benefit most from the use of lecture capture in medical school. *Biochemistry and Molecular Biology Education, 39*(6), 416–420.

Strevens, P. (1992). English as an international language: Directions in the 1990s. *The Other Tongue: English across Cultures, 2*, 7–47.

Suritsky, S. K. (1992). Notetaking difficulties and approaches reported by university students with learning disabilities. *Journal of Postsecondary Education and Disability, 10*(1), 3–10.

Tas, O., & Kiyani, F. (2007). A survey automatic text summarization. *PressAcademia Procedia, 5*(1), 205–213.

Thanaki, J. (2017). Python natural language processing. Packt Publishing Ltd.

Thyagaraju, N. (2017). Promoting the quality of higher education in India through research & development. *International Journal of Research in Social Sciences, 7*(4), 204–211.

Uzelac, A., Gligorić, N., & Krsč, S. (2018). System for recognizing lecture quality based on analysis of physical parameters. *Telematics and Informatics, 35*(3), 579–594.

Wagner, S. (2005). Intralingual Speech-to-text conversion in real-time: Challenges and Opportunities. In *Challenges of Multidimensional Translation Conference Proceedings*.

Yang, X., Zhang, G., Lu, J., & Ma, J. (2010). A kernel fuzzy c-means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises. *IEEE Transactions on Fuzzy Systems, 19*(1), 105–115.

Yokubov, U. (2021). Mobile apps for learning English as a means of organizing independent work of students. *Academic Research in Educational Sciences, 2*(1), 447–454.

Yoon, M., Hill, J., & Kim, D. (2021). Designing supports for promoting self-regulated learning in the flipped classroom. *Journal of Computing in Higher Education, 33*(2), 398–418.

Yulianti, D., Rusilowati, A., & Nugroho, S.E. (2020). Student worksheets based on Science, Technology, Engineering and Mathematics (STEM) to facilitate the development of critical and creative thinking skills. *In Journal of Physics: Conference Series* (Vol. 1567, No. 2, p. 022068). IOP Publishing.

Yusoff, Z., Kamsin, A., Shamshirband, S., & Chronopoulos, A. T. (2018). A survey of educational games as interaction design tools for affective learning: Thematic analysis taxonomy. *Education and Information Technologies, 23*(1), 393–418.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.