НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
март–апрель 2023 Том 23 № 2 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
March–April 2023 Vol. 23 No 2 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

doi: 10.17586/2226-1494-2023-23-2-299-303

Voice based answer evaluation system for physically disabled students using natural language processing and machine learning

Meenakshi Thalor[1]✉, Pradeep Mane[2]

[1,2] AISSMS Institute of Information Technology, Pune, 411001, India

[1] meenakshi.thalor@aissmsioit.org✉, https://orcid.org/0000-0001-7048-7289
[2] pbmane6829@gmail.com, https://orcid.org/0000-0002-3252-9983

Abstract

In the modern educational process, there is a need to automate response assessment systems. The task of the reviewer becomes more difficult when analyzing theoretical answers, because online assessment of answers is available only for questions with multiple choice answers. The teacher carefully examines the answer before giving the appropriate mark. The existing approach requires additional staff and time to study the responses. This article introduces a natural language processing and machine learning response-based app that includes a voice prompt for visually impaired students. The application automates the process of checking subjective responses by considering text extraction, feature extraction, and score classification. Evaluation measures, such as Term Frequency-Inverse Document Frequency (TF-IDF) similarity, vector similarity, keyword similarity, and grammar similarity, are considered to determine the overall similarity between teacher outcome and system evaluation. The conducted experiments showed that the system evaluates the answers with an accuracy of 95 %. The proposed methodology is designed to assess the results of exams for students who cannot write but who can speak. The application of the developed application allows reducing the labor costs and time of the teacher by reducing manual labor.

Keywords

cosine similarity, machine learning, Naive Bayes, natural language processing, speech to text conversion

For citation: Thalor M., Mane P. Voice based answer evaluation system for physically disabled students using natural language processing and machine learning. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 2, pp. 299–303. doi: 10.17586/2226-1494-2023-23-2-299-303

УДК 004.89

Голосовая система оценки ответов для учащихся с ограниченными физическими возможностями, использующих обработку естественного языка и машинное обучение

Минакши Талор[1]✉, Прадип Мане[2]

[1,2] AISSMS Институт информационных технологий, Пуне, 411001, Индия

[1] meenakshi.thalor@aissmsioit.org✉, https://orcid.org/0000-0001-7048-7289
[2] pbmane6829@gmail.com, https://orcid.org/0000-0002-3252-9983

Аннотация

В современном образовательном процессе возникает потребность в автоматизации систем оценки ответов. Задача проверяющего усложняется при анализе теоретических ответов, так как их онлайн-оценка доступна только для вопросов с несколькими вариантами ответов. Преподаватель тщательно изучает ответ, прежде чем поставить соответствующую оценку. Для изучения ответов существующий подход требует дополнительных сотрудников и времени. В работе представлено приложение, основанное на оценке ответов с использованием обработки естественного языка и машинного обучения, которое включает голосовую подсказку для слабовидящих учащихся. Приложение автоматизирует процесс проверки субъективных ответов, рассматривая извлечение текста и признаков, а также классификацию баллов. Мерами оценки являются сходства: Term Frequency-

Inverse Document Frequency (TF-IDF), векторов, ключевых слов и грамматики, которые рассматриваются для определения общего сходства между результатом учителя и оценкой системы. Полученные результаты показали, что система оценивает ответы с точностью 95 %. Предлагаемая методика предназначена для оценки результатов экзаменов учащихся, не умеющих писать, но умеющих говорить. Применение разработанного приложения позволит сократить затраты труда и времени преподавателя за счет сокращения ручного труда.

Keywords
косинусное сходство, машинное обучение, наивный байесовский анализ, обработка естественного языка, преобразование речи в текст

Ссылка для цитирования: Талор М., Мане П. Голосовая система оценки ответов для учащихся с ограниченными физическими возможностями, использующих обработку естественного языка и машинное обучение // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 2. С. 299–303 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-2-299-303

Introduction

Students' academic achievement is generally measured by their examination scores which might be subjective or objective. A variety of tools like Google Quiz Form, Microsoft Quiz Form exist that can quickly assess objective or multiple-choice questions. These tactics are tested in machines after producing pre-defined accurate responses. It is, however, only applicable to judging competitive or objective exams. All university and board level exams are built on the foundation of subjective examinations. Based on the detailed answer, the moderator will determine how much knowledge the student has acquired during his academic career and assign marks accordingly. Manually assessing subjective replies is a time-consuming and labor-intensive process that necessitates the participation of many people. Answer evaluation differs from moderator to moderator depending on their style of evaluation, mood at the time of evaluation, and interrelationship between student and moderator. The student's grade is affected because of this. Because of all abovementioned challenges, this paper proposes an automate system for evaluation process of descriptive answers, voice based answer evaluation system for physically challenged students where students can give the answer to descriptive questions by speech.

In [1–7] authors proposed systems which assess subjective question based on Natural Language Processing (NLP) and Machine Learning (ML) methodologies like calculate the length of the answer, keyword matching, grammar check, similarity measures, and contextual resemblance to the faculty model answer and the student's answer. In [8] author makes use of Google speech recognition software for audio to text conversion. In [9] author proposed a system with RNN-based model with a gated recurrent unit and achieved an accuracy of 87 percent. Similarly, in [10] the authors proposed a new system to determine the semantic meaning of student responses, considering that students can respond to questions in a variety of ways. While in [11, 12], the text mining approaches are used evaluation of descriptive answers.

This paper presents an application based on the evaluation of answers using NLP and ML which includes a voice aid for visually challenged students. This paper focuses on evaluation of disabled or physically challenged students who can speak but can't write. This work is extension of "Web app for quick evaluation of subjective answers using natural language processing" [13] which was proposed by us earlier in 2021 where authors developed web application for subjective answers checking and generates results through using NLP methods, like keyword matching semantic, lexical analysis and cosine similarity. In this paper a speech-to-text methodology is used in addition to approach used in [13] which provide exam evaluation for students who cannot write but can speak.

Methods and Materials

To analyze voice-based answers, the proposed method employs NLP and ML. As inputs, all the student responses are used as well as one standard answer for each question is taken from teacher. To begin, input is pre-processed to make it ready for evaluation. Student and standard responses have been tokenized, searched for synonyms, stopped, and stemmed. The proposed method is intended to assess exam outcomes for pupils who cannot write but can speak. Voice based answer assessment system is depicted in Fig. 1 as a series of phases.

There are two methods to the proposed system: Text Extraction and Feature Extraction.

Text Extraction

In this module, the speech-to-text conversion process is carried out by using Python Speech recognizer. The student response for descriptive answer is recorded and

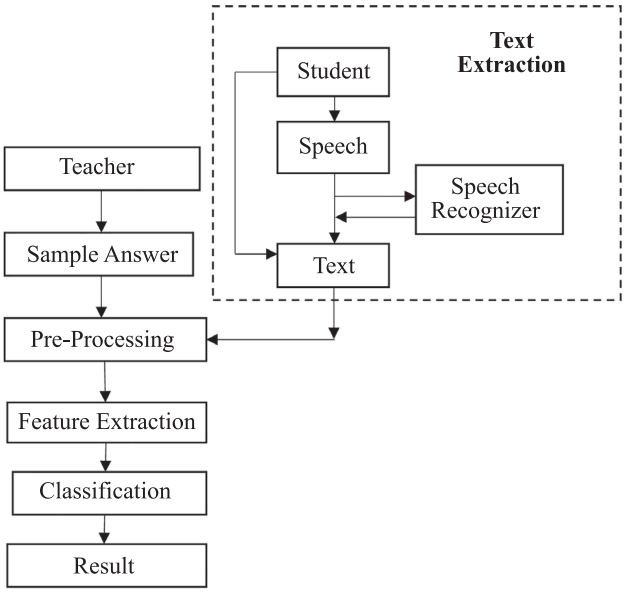

Fig. 1. Voice based answer assessment system

processed into text by using Application Programmable Interface (API). To convert speech to text, one can use one of following NLP speech-to-text methods:

1. ***DeepSpeech.*** DeepSpeech is a term that refers to a type of speech that DeepSpeech is a free, open-source voice-to-text engine that translates text into speech in real time. However, based on the work, we can infer that DeepSpeech produces good results, but at the expense of the module massive file size.
2. ***Python Speech Recognizer.*** The Python Speech Recognizer module is used to recognize voices using a number of engines and APIs. For recognizing speech for audio sources, the Recognizer class in Python uses various APIs from Microsoft, Google, IBM, etc.

In NLP, text pre-processing is a common stage. It transforms data into a more readable format, allowing ML algorithms to complete the remaining tasks more quickly. Sentence tokenization, word tokenization, stop word elimination, character set elimination as explained in [14] are employed in this stage.

Feature Extraction

Due to the inability to compute document data, it must be translated to numerical data such as a vector space model. This transformation is known as feature extraction which comprises of three steps as explained below.

Extraction of Keywords. The Rake-NLTK module [14] is used to extract relevant terms from the answer. This is done to shorten the answer and enable for faster and more efficient comparison. A sequence matcher from difflib is used to compare keywords — two strings — producing a ratio, such as the similarity ratio between the words "abc" and "abc123". If the similarity ratio is 0.8 or above, then it means that a keyword is identified. The fraction of keywords for each category is determined after the comparison.

Checking for similarity. Here the examination of the similarity of teacher and student responses and computation of the percentage of similarity is carried out. For examination of similarity, first Term Frequency-Inverse Document Frequency (TF-IDF) [14] of terms found in teacher and student responses is computed, and then cosine similarity is used to calculate percentage similarity [13]. Following equation is used to calculate the cosine similarity between the teacher's answer and the student's answer.

$$Cosine\ Similarity = \frac{Dot\ product\ (teacher, student)}{||teacher|| \times ||student||}.$$

A *Naïve Bayes classification.* ML algorithm is used to figure out which group students' answer belongs to and computation of his/her score. The Naive Bayes method is a supervised learning algorithm that uses the Bayes theorem to solve classification problems [15].

Results and Discussions

Every student's performance is evaluated in any educational system around the world through a series of tests. After doing some survey, we discovered that physically challenged students, especially those who can't write but speak, face lot of problems while appearing for exam. This application not only useful for descriptive exam assessment but also helpful for students who can't write because of disability or due to some accident happens just

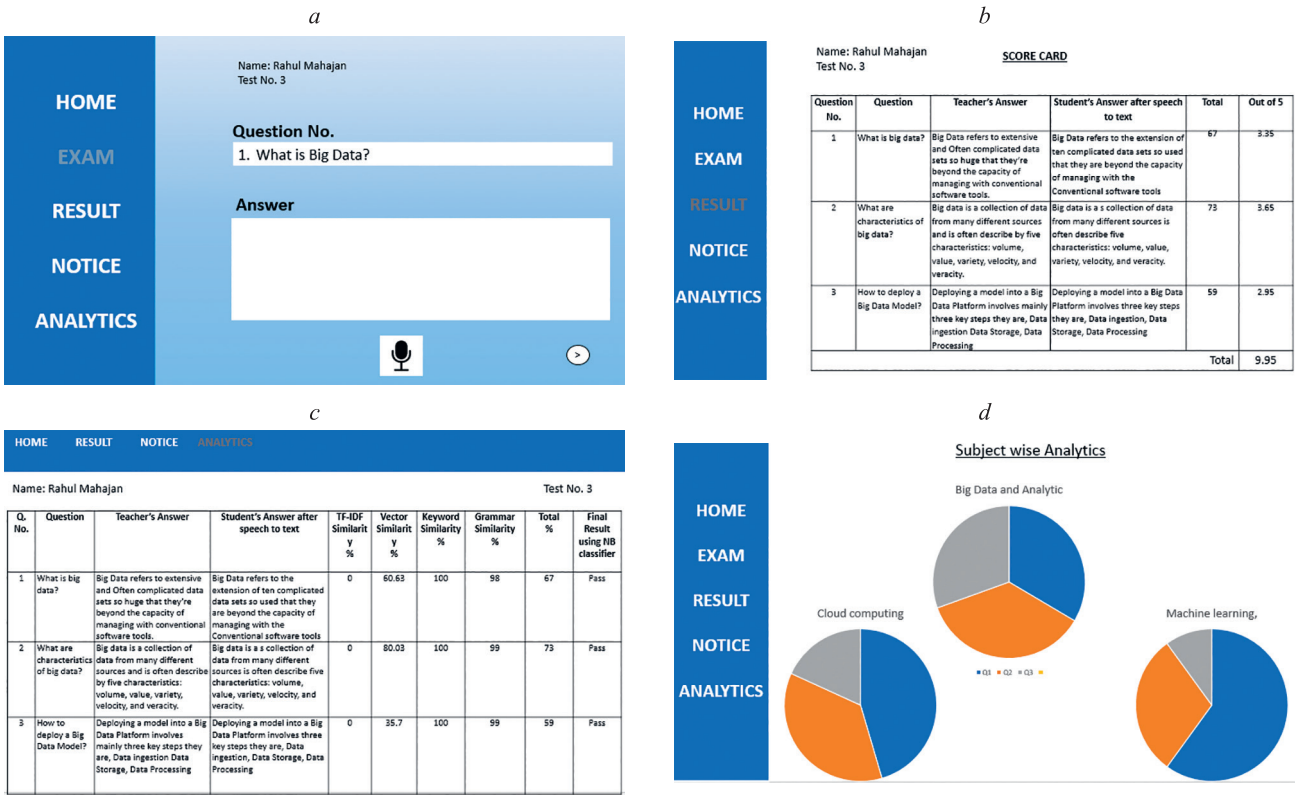

Fig. 2. Graphical User Interface: Question Panel with Audio Option (*a*); Result Page (*b*); Teacher's Result Page (*c*); Subject wise analytics (*d*)

Table. Sample Experimental Results of Voice based Answer Evaluation System

Question	What is Big Data?	What are characteristics of Big Data?	How to deploy a Big Data Model?
Teacher's answer	Big Data refers to extensive and often complicated data sets so huge that they're beyond the capacity of managing with conventional software tools	Big Data is a collection of data from many different sources that are often described by five characteristics: volume, value, variety, velocity, and veracity	Deploying a model into a Big Data Platform involves mainly three key steps: Data ingestion, Data Storage, Data Processing
Student's answer	Big Data refers to the extension of ten complicated data sets so used that they are beyond the capacity of managing with the conventional software tools	Big Data is a collection of data from many different sources that are often described by five characteristics: volume, value, variety, velocity, and veracity	Deploying a model into a Big Data Platform involves three key steps: Data ingestion, Data Storage, Data Processing
TF-IDF Similarity, %	0	0	0
Vector similarity, %	60.63	80.03	35.70
Keyword Similarity, %	100	100	100
Grammar Similarity, %	98	99	99
Total, %	67	73	59
Final Result using Naïve Bayes classifier	Pass	Pass	Pass

before few days of exam. In proposed application, voice recording feature to record answers orally is implemented and get evaluated by NLP and ML approach. This application extracts keywords from the responses given by students, and then compares the extracted and given keywords from the answer sheet using cosine similarity. Fig. 2 shows the graphical user interface of this application.

The scope of this work is to convert English speech to text and evaluate descriptive text of student. Experimental Test has been conducted in the college premises for three subjects: ML, cloud computing, and Big Data and Analytics. In total, three questions were given for subject, with five marks each. Ten engineering students have given the answer to the question by using speech for each subject. The result is generated in the form of a percentile score. Manual checking of the answers is done by the college professors to compare the results.

Table shows the experiment result for one subject, i.e., Big Data and Analytics where student has given the answer to the question by using speech. Speech Recognizer converts the English speech into text and display the text to the student; student can retake the question or can submit the response. After submission of response by the student, system will generate the question wise score based on similarity measures as shown in Table. To get the final score, each similarity indicator is assigned a weight which can be changed based on the admin's preference.

Based on experimental results, it has been observed that speech to text module affected by background noise which can be eliminated by using microphone and providing a quiet place to examiner. Secondly the pronunciation and speech speed also affect the performance of proposed system which can be improved by providing the speech to text screen to the student and student can re-speak the answer. The results of Table are recorded considering favourable conditions like isolated room with minimum background disturbance and can be improved further.

Conclusion

A voice-based answer evaluation method for physically challenged students that uses natural language processing and machine learning is proposed here. This method not only useful for descriptive exam assessment but also helpful for students who can't write because of disability or due to some accident happens just before few days of exam. This work is extension of "Web app for quick evaluation of subjective answers using natural language processing" [13] and it only focuses on voice-based evaluation and shows the accuracy of result considering unavoidable background noise. The proposed method is only capable of accepting English speech answers. Therefore, speech to text conversion for multilingual languages can be supported as future enhancement.

References

1. Abhishek Girkar, Mohit khambayat, Ajay Waghmare, Supriya Chaudhary. Subjective answer evaluation using natural language processing and machine learning. *International Research Journal of Engineering and Technology (IRJET)*, 2021, vol. 8, no. 4, pp. 5040–5043.
2. Shweta M. Patil, Sonal Patil. Evaluating student descriptive answers using natural language processing. *International Journal of Engineering Research & Technology (IJERT)*, 2014, vol. 3, no. 3, pp. 1716–1718.
3. Mercy R.A. Automated explanatory answer evaluation using machine learning approach. *Design Engineering (Toronto)*, July 2021.
4. Lakshmi V., Ramesh V. Evaluating students' descriptive answers using natural language processing and artificial neural networks. *IJCRT*, 2017, vol. 5, no. 4, pp. 3168–3173.
5. Mittal H., Devi M.S. Computerized evaluation of subjective answers using hybrid technique. *Advances in Intelligent Systems and Computing*, 2016, vol. 413, pp. 295–303. https://doi.org/10.1007/978-981-10-0419-3_35
6. Sakhapara A., Pawade D., Chaudhari B., Gada R., Mishra A. Bhanushali S. Subjective answer grader system based on machine learning. *Advances in Intelligent Systems and Computing*, 2019, vol. 898, pp. 347–355. https://doi.org/10.1007/978-981-13-3393-4_36
7. Mahmud T.A., Hussain M.G., Kabir S., Ahmad H., Sobhan M. A keyword based technique to evaluate broad question answer script. *Proc. of the 9th International Conference on Software and Computer (ICSCA)*, 2020, pp. 167–171. https://doi.org/10.1145/3384544.3384604
8. Bano S., Jithendra P., Niharika G.L., Sikhi Y. Speech to text translation enabling multilingualism. *Proc. of the IEEE International Conference for Innovation in Technology (INOCON)*, 2020. https://doi.org/10.1109/inocon50539.2020.9298280
9. Bhatia M.S., Aggarwal A., Kumar N. Speech-to-text conversion using GRU and one hot vector encodings. *Palarch's Journal of Archaeology of Egypt/Egyptology*, 2020, vol. 17, no. 9, pp. 8513–8524.
10. Johri E., Dedhia N., Bohra K., Chandak P., Adhikari H. ASSESS - automated subjective answer evaluation using semantic learning. *Proc. of the 4th International Conference on Advances in Science & Technology (ICAST2021)*, 2021. http://dx.doi.org/10.2139/ssrn.3861851
11. Deotare S., Khan R.A. Automatic online subjective text evaluation using text mining. *International Journal of Recent Technology and Engineering (IJRTE)*, 2019, vol. 8, no. 2, pp. 1238–1242. https://doi.org/10.35940/ijrte.I8725.078219
12. Bahel V., Thomas A. Text similarity analysis for evaluation of descriptive answers. *arXiv*, 2021, arXiv:2105.02935. https://doi.org/10.48550/arXiv.2105.02935
13. Meenakshi Anurag T., Pradeep B.M., Vishaka M. Web app for quick evaluation of subjective answers using natural language processing. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 3, pp. 594–599. https://doi.org/10.17586/2226-1494-2022-22-3-594-599
14. Mandge V.A., Thalor M.A. Revolutionize cosine answer matching technique for question answering system. *Proc. of the International Conference on Emerging Smart Computing and Informatics (ESCI)*, 2021, pp. 335–339. https://doi.org/10.1109/ESCI50559.2021.9396864
15. Rish I. An empirical study of the naïve bayes classifier. *IJCAI Work Empir Methods Artificial Intelligence. V. 3*, 2001.

Authors

Meenakshi Thalor — PhD, Associate Professor, AISSMS Institute of Information Technology, Pune, 411001, India, SC 57190340673, https://orcid.org/0000-0001-7048-7289, meenakshi.thalor@aissmsioit.org

Pradeep Mane — PhD, Principal, AISSMS Institute of Information Technology, Pune, 411001, India, SC 36450914400, https://orcid.org/0000-0002-3252-9983, pbmane6829@gmail.com

Received 12.07.2022
Approved after reviewing 08.02.2023
Accepted 14.03.2023

Литература

1. Abhishek Girkar, Mohit Khambayat, Ajay Waghmare, Supriya Chaudhary. Subjective answer evaluation using natural language processing and machine learning // International Research Journal of Engineering and Technology (IRJET). 2021. V. 8. N 4. P. 5040–5043.
2. Shweta M. Patil, Sonal Patil. Evaluating student descriptive answers using natural language processing // International Journal of Engineering Research & Technology (IJERT). 2014. V. 3. N 3. P. 1716–1718.
3. Mercy R.A. Automated explanatory answer evaluation using machine learning approach // Design Engineering (Toronto). July 2021.
4. Lakshmi V., Ramesh V. Evaluating students' descriptive answers using natural language processing and artificial neural networks // IJCRT. 2017. V. 5. N 4. P. 3168–3173.
5. Mittal H., Devi M.S. Computerized evaluation of subjective answers using hybrid technique // Advances in Intelligent Systems and Computing. 2016. V. 413. P. 295–303. https://doi.org/10.1007/978-981-10-0419-3_35
6. Sakhapara A., Pawade D., Chaudhari B., Gada R., Mishra A. Bhanushali S. Subjective answer grader system based on machine learning // Advances in Intelligent Systems and Computing. 2019. V. 898. P. 347–355. https://doi.org/10.1007/978-981-13-3393-4_36
7. Mahmud T.A., Hussain M.G., Kabir S., Ahmad H., Sobhan M. A keyword based technique to evaluate broad question answer script // Proc. of the 9th International Conference on Software and Computer (ICSCA). 2020. P. 167–171. https://doi.org/10.1145/3384544.3384604
8. Bano S., Jithendra P., Niharika G.L., Sikhi Y. Speech to text translation enabling multilingualism // Proc. of the IEEE International Conference for Innovation in Technology (INOCON). 2020. https://doi.org/10.1109/inocon50539.2020.9298280
9. Bhatia M.S., Aggarwal A., Kumar N. Speech-to-text conversion using GRU and one hot vector encodings // Palarch's Journal of Archaeology of Egypt/Egyptology. 2020. V. 17. N 9. P. 8513–8524.
10. Johri E., Dedhia N., Bohra K., Chandak P., Adhikari H. ASSESS - automated subjective answer evaluation using semantic learning // Proc. of the 4th International Conference on Advances in Science & Technology (ICAST2021). 2021. http://dx.doi.org/10.2139/ssrn.3861851
11. Deotare S., Khan R.A. Automatic online subjective text evaluation using text mining // International Journal of Recent Technology and Engineering (IJRTE). 2019. V. 8. N 2. P. 1238–1242. https://doi.org/10.35940/ijrte.I8725.078219
12. Bahel V., Thomas A. Text similarity analysis for evaluation of descriptive answers // arXiv. 2021. arXiv:2105.02935. https://doi.org/10.48550/arXiv.2105.02935
13. Meenakshi Anurag T., Pradeep B.M., Vishaka M. Web app for quick evaluation of subjective answers using natural language processing // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22. № 3. С. 594–599. https://doi.org/10.17586/2226-1494-2022-22-3-594-599
14. Mandge V.A., Thalor M.A. Revolutionize cosine answer matching technique for question answering system // Proc. of the International Conference on Emerging Smart Computing and Informatics (ESCI). 2021. P. 335–339. https://doi.org/10.1109/ESCI50559.2021.9396864
15. Rish I. An empirical study of the naïve bayes classifier // IJCAI Work Empir Methods Artificial Intelligence. V. 3. 2001.

Авторы

Талор Минакши — PhD, доцент, AISSMS Институт информационных технологий, Пуне, 411001, Индия, SC 57190340673, https://orcid.org/0000-0001-7048-7289, meenakshi.thalor@aissmsioit.org

Мане Прадип — PhD, руководитель, AISSMS Институт информационных технологий, Пуне, 411001, Индия, SC 36450914400, https://orcid.org/0000-0002-3252-9983, pbmane6829@gmail.com

Статья поступила в редакцию 12.07.2022
Одобрена после рецензирования 08.02.2023
Принята к печати 14.03.2023