УНИВЕРСИТЕТ ИТМО
НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
май–июнь 2022 Том 22 № 3 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
May–June 2022 Vol. 22 No 3 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)
НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК УНИВЕРСИТЕТ ИТМО
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

doi: 10.17586/2226-1494-2022-22-3-594-599

Web app for quick evaluation of subjective answers using natural language processing

Meenakshi Anurag Thalor[1]✉, Pradeep B. Mane[2], Vishaka Mandge[3]

[1,2,3] 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
[3] mandge1999@gmail.com, https://orcid.org/0000-0002-1667-4814

Abstract

In current digital climate, education sector is evolving as the computer technology advances. Education is being digitized: online classes, online examination methods are conducted, etc. During examination, students are assessed by their answers having given for the question set by a teacher. Today many tools are available to assess the performance of a student using multi choice questions tools which provide instant evaluation, but there are available very limited and operational tools where subjective type answer of students are evaluated. This paper presents a web-based application to address this challenge. It automates the process of subjective answers checking and generates results through using natural language processing methods, like keyword matching semantic, lexical analysis and cosine similarity. Experiments show that appreciated by the teacher result and the system estimation does not have much difference which signifies that the system evaluates answers with a 97 % accuracy. The presented system not only reduces manpower but also eliminates the traditional method of conducting exclusively subjective exams using paper documents. It also eliminates the delays in the paper checking, result generation process. The cases of information leak are being reduced and the objectivity of the assessment is being increased.

Keywords

cosine similarity, information extraction, keyword matching, natural language processing

For citation: 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. doi: 10.17586/2226-1494-2022-22-3-594-599

УДК 004.89

Веб-приложение для быстрой оценки субъективных ответов с использованием обработки естественного языка

Талор Минакши Анураг[1]✉, Мане Прадип Б.[2], Мандж Вишака[3]

[1,2,3] AISSMS Институт информационных технологий, Пуна, 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
[3] mandge1999@gmail.com, https://orcid.org/0000-0002-1667-4814

Аннотация

В современном цифровом пространстве развитие образования связано с продвижением компьютерных технологий. Происходит цифровизация образования: проводятся онлайн-классы, онлайн-экзамены и т. д. В период проведения экзамена учащиеся оцениваются педагогом по ответам на заданные вопросы. На настоящий момент доступно множество инструментов оценки успеваемости учащегося с использованием вопросов с несколькими вариантами ответов. Такой подход обеспечивает мгновенную оценку, однако инструментов, дающих возможность оценки субъективных ответов учащихся еще недостаточно. В работе представлено веб-приложение

для решения этой проблемы. Приложение позволяет автоматизировать процесс проверки субъективных ответов и генерировать результаты с применением методов обработки естественного языка, таких как семантическое сопоставление ключевых слов, лексический анализ и косинусное сходство. Экспериментальные результаты показали, что ответы, оцененные педагогами, и предложенная система дают близкие результаты. Полученная точность ответов составила 97 %. Представленная система не только сокращает рабочее время, но и позволяет устранить традиционный метод проведения исключительно субъективных экзаменов с использованием бумажных документов. В результате сокращается время получения результатов и время проверки. Уменьшается утечка информации и повышается объективность оценки.

Ключевые слова

косинусное сходство, извлечение информации, сопоставление ключевых слов, обработка естественного языка

Ссылка для цитирования: Минакши Анураг Т., Прадип Б.М., Вишака М. Веб-приложение для быстрой оценки субъективных ответов с использованием обработки естественного языка // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 3. С. 594–599 (на англ. яз.). doi: 10.17586/2226-1494-2022-22-3-594-599

Introduction

In digital Era, most of the educational institutes conduct online examination in the form of multi choice questions using some tools/applications. Some of these tools/applications are also capable to accept subjective answers which need to be manually evaluated by a teacher since these tools are not trained enough to provide an accurate result for subjective type questions.

An example of an existing model that uses Natural Language Processing (NLP) [1–4] is Paper Pater. The application is designed to analyze the text and predict the plagiarism of the submitted text. It even helps improve the text grammatically using certain machine learning algorithms. E-rater Scoring Engine is a grading system that provides a grade for the essays submitted by students using NLP. Passage ranking system is used to rank a set of passages based on its relevance to a certain topic. N-gram and Support Vector Machines models are used for ranking these passages. Another example is Machine Reading and Comprehension (MRC) model; it retrieves the most relevant passage based on the question. These systems are good at parsing the textual data and they produce the required results. In literature, the system is found using Knowledge extraction [5], feature identification [6] and deep learning techniques [7] in Question Answer Matching [8–13].

Methods and Materials

The web based application system uses NLP to analyze the descriptive answers provided by a student as well as by the teacher. Fig. 1 shows the system architecture of Web-based Application for revolutionary assessment where student and teacher are two main users in addition to admin/superuser. Students and teacher can perform various activities as clearly depicted in pictorial representation. All the data once received will be stored in the MongoDB database and can be retrieved using Python to perform the further operations.

Web based application system includes revolutionary assessment and instant evaluation of following modules: Login Module, Preprocessing Module, Information Extraction, Score Generation.

Login module

This module is used for authentication purpose of the participants: superuser, teacher and student. Each person has his/her unique ID and password to login and work in the system after the system authenticates the user; after that the user can further access all the functionalities of the role.

1. Teacher login

The teacher is validated by using his/her unique ID and password. Once validation is done, the teacher gets logged in to his/her dashboard. Teacher has to define the character set for his model answers in order to exclude them during preprocessing. The teacher can set document where he/she can add questions and answers into the database. The teacher can also view results of all students and of each student separately as well as each subject individually. The key feature of the system is that the teacher can only view the answers of the student but cannot modify the answer. The questions will appear in the exam for the students and the answers stored in the database will be used as the model answer for comparison with the student's answers.

2. Student login

Student can login into the system by using his/her unique PRN number and password. After authentication the student gets logged in to the system and can see his/her dashboard. The students can view their previous test results but a student cannot view the current result until he/she has passed any test. To pass the test the student has to click the "Take test" option, select the subject from

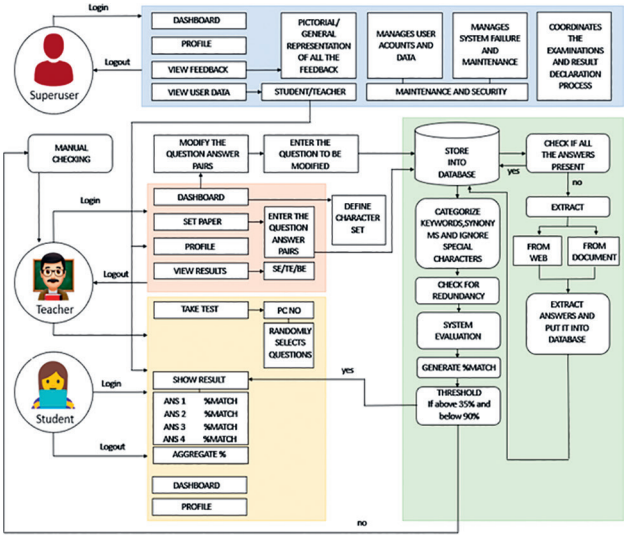

Fig. 1. Activity Diagram for Web Based Application System for Revolutionize Assessment and Instant Evaluation

the dropdown menu and then click on starting the test. As soon as the student has passed his/her test, the test results are automatically submitted into the database. And after completion of the test there is a feedback form for every student to which the superuser has access. The feedback form also gets stored into the database. The students can view their results once they are announced.

3. Superuser login

The superuser is an admin who handles all the database issues, failures, user registrations and any errors. He/she is responsible for the maintenance of the system and can respond to any kind of issues. The superuser has access to the feedback given by the students after passing the test.

Preprocessing module

During preprocessing, the sentence tokenization and work tokenization is applied to the descriptive answer of student and teacher as shown in Fig. 2. RAKE is Rapid Automatic Keyword Extraction algorithm available in Natural Language Toolkit (NLTK), the purpose of this algorithm is to determine key phrases in text using frequency of word and coexistence of words with other words. The individual tokens are parsed to eliminate the stop and repeated words, and keywords are extracted at this stage. The keyword matching of teacher and students is also performed here.

Information Extraction

Information extraction plays a significant role in NLP since only important keywords are considered after the irrelevant information is removed and keywords are ranked. The important and relevant keywords in descriptive answers of student and teacher are identified and matched using cosine similarity at this stage. Cosine similarity [14–15] is the evaluation measure to get the similarity between two documents. There exists a motivation behind selection of cosine similarity measure, it gives a ranking to documents corresponding to a given keywords irrespective of the size of the document. The larger the cosine angle, the less similarity exists between two documents, and vice versa. The cosine similarity is calculated for each answer proposed by the student. Following equation is used to calculate the cosine similarity between the teacher's answer and the student's answer.

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

The final result is calculated by aggregating these values which will then be converted into percentile. Cosine similarity is used to compare the answers of teacher and student in general. If the student's answer has obtained the maximum marks, it means that its cosine similarity was low.

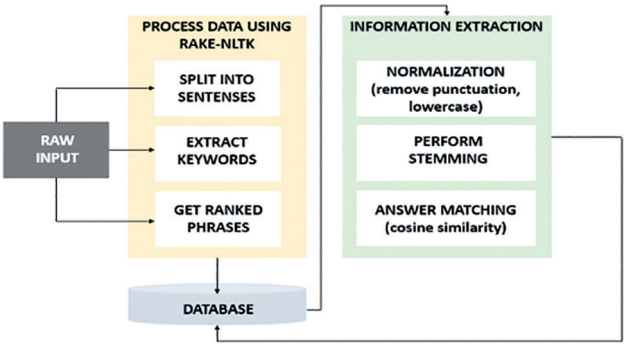

Fig. 2. Natural language processing modules or revolutionary assessment and instant evaluation

Score generation

The score of each student is shown to them in the form of pictorial representation, i.e., a graph. The graph represents the total percentage of the accuracy of the student's answers in comparison with the teacher's answers. Both the teacher and the student can view their result for a particular subject and also view the marks allotted for individual question.

Results and discussions

Test has been conducted in the college premises in two subjects: Machine Learning and Cloud Computing. In total, five questions were given for each subject, with five marks each. 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 1 shows that the answers assessed by the teachers and the answers assessed by the system doesn't have much variation. This result of comparison shows that the system evaluates answers with a great accuracy. Table 2 depicts the percentile score which is generated for individual question as well. Proposed system provides the accuracy of 97 % on the mentioned subjects.

Following screenshots represent the entire web portal created in order to revolutionize the traditional system of taking an exam using pen and paper.

Student registration

Fig. 3, *a* shows the student registration webpage. The student is required to do the registration before attending

Table 1. Results comparison, %

Results	System generated results for subjects		Teacher provided results for subjects	
	Machine Learning	Cloud Computing	Machine Learning	Cloud Computing
1	72.09	85.63	74.23	84.23
2	65.21	73.84	67.15	74.36
3	32.83	47.82	38.56	46.67
4	54.36	63.45	57.67	68.14
5	82.91	90.80	87.16	89.83

Table 2. Question wise results

Question	Total marks	Percentage obtained, %
What is Machine learning?	5	42.5
Mention the difference between Data Mining and Machine learning?	5	42.2
What is ‘Over-fitting’ in Machine learning?	5	21.0
How can you avoid over-fitting?	5	18.9
What is ‘Training set’ and ‘Test set’?	5	14.3

the exam. The Unique ID, i.e., the issued by the university PRN number, is used as the username.

Student login

After a student logs in, he/she is directed to the dashboard from where he/she can access several functionalities such as view profile, take test, view result and logout. The dashboard consists of some valuable information for the student such as time management, guidelines for the examination, etc. The student can even view his/her profile as shown in Fig. 3, *b*.

Take test

Once the student has successfully logged in to the system, he/she can now take test. In order to start the test as shown in Fig. 4, *a*, the student has to read carefully the instructions and to choose the subject. The time allotted for the exam is half an hour and, once the test is submitted, the student is directed to the feedback section where he/she is asked if he/she is satisfied with the new way of giving the subjective examinations and if he/she liked the procedure and feel that the entire system is developed properly.

View results

The student can view his/her results once they have been declared. The results won’t be visible if the student didn’t show up for the exam. The student can view the result of the individual subject as well as of the individual questions. Pictorial representations are used in order to make it easy for the student to view and understand the results as shown in Fig. 4, *b*.

Teacher registration and login

Fig. 5, *a* shows teacher’s login page where the teacher also registers in the system using a Unique ID and, after login, is directed to the Dashboard where he/she can access functionalities such as view profile, set document, view results and logout.

Set document and view results

Fig. 5, *b* shows the teacher document setting webpage. The teacher needs to specify the model file that the system evaluates with the students file. Once all the students have appeared for the examination, the teacher can view the overall results of the students.

Superuser and additional features

The superuser has the functionality of viewing the feedback of the students in order to improve the system as shown in Fig. 6. The system generates a false message whenever it discovers an invalid username and password pair.

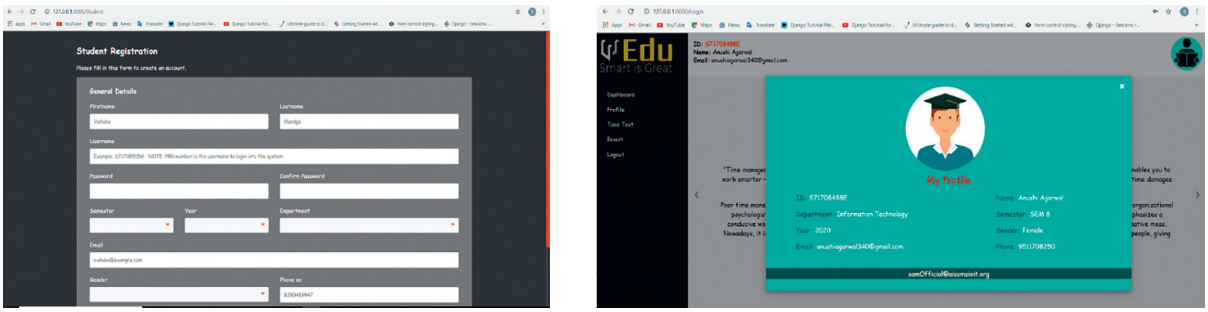

Fig. 3. Webpage: student registration (*a*), student logged in (*b*)

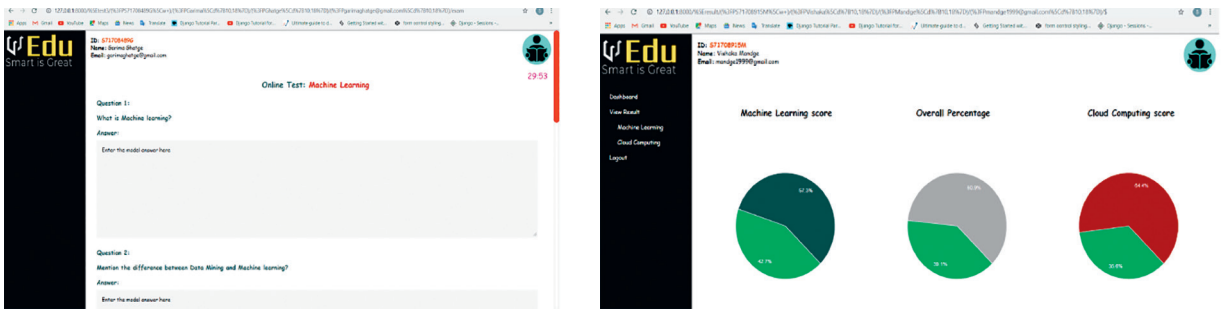

Fig. 4. Student’s webpage: test (*a*), results (*b*)

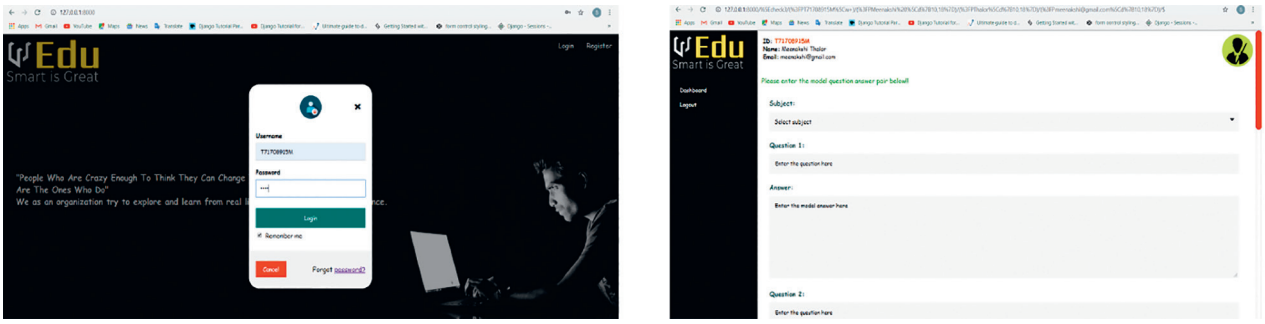

Fig. 5. Webpage: teacher's login (*a*), teacher document settings (*b*)

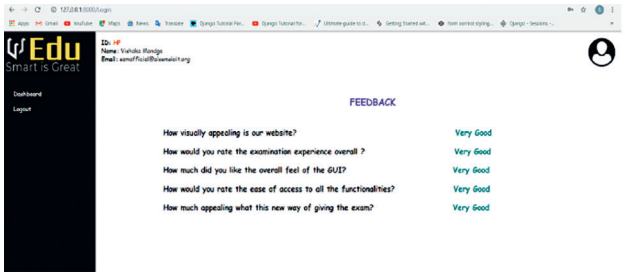

Fig. 6. Feedback webpage

Conclusion

The result is generated in the percentile form and is represented in the form of a graph. The proposed system also shows the individual marks allotted for each question by the student as well as by each teacher. The teacher can see the result of each student and also the answers of the student. The key feature of the system is that the teacher cannot modify the answer of the student. Although the system provides a percentile for the test taken by the student, the system is only fully suitable for subjective answers, which are purely theoretical. The proposed system evaluates answers written in English. The TF-IDF technique used in this way takes into account synonyms for the words that are the same, which results in a limitation for the algorithm so used. So, in future, we plan to include the synonyms in the TF-IDF technique, hence providing a completely new algorithm that gives a much more accurate result. The option of extracting questions and their answers from the web for incomplete data can be one of the future works. The current system is feasible for English language only so in future it can works for other languages as well.

References

1. Lakshmi V., Ramesh V. Evaluating students' descriptive answers using natural language processing and artificial neural networks. *International Journal of Creative Research Thoughts — IJCRT*, 2017, vol. 5, no. 4, pp. 3168–3173.
2. Patil S.M., Patil S. Evaluating student descriptive answers using natural language processing. *International Journal of Engineering Research & Technology (IJERT)*, 2014, vol. 3, no. 3, pp. 1716–1718.
3. Yin W., Kann K., Yu M., Schütze H. Comparative study of CNN and RNN for natural language processing. *ArXiv*, 2017, arXiv:1702.01923. https://doi.org/10.48550/arXiv.1702.01923
4. Kate A., Kamble S., Bodkhe A., Joshi M. Conversion of Natural Language Query to SQL Query. *Proc. of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2018, pp. 488–491. https://doi.org/10.1109/ICECA.2018.8474639
5. Thangarasu S., Sasikala D. Extracting knowledge from XML document using tree-based association rules. *Proc. of the International Conference on Intelligent Computing Applications (ICICA)*, 2014, pp. 134–137. https://doi.org/10.1109/ICICA.2014.37
6. Mati D.N., Ajdari J., Raufi B., Hamiti M., Selimi B. A Systematic mapping study of language features identification from large text collection. *Proc. of the 8th Mediterranean Conference on Embedded Computing (MECO)*, 2019, pp. 8760042. https://doi.org/10.1109/MECO.2019.8760042
7. Cai J., Li J., Li W., Wang J. Deeplearning model used in text classification. *Proc of the 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 2018, pp. 123–126. https://doi.org/10.1109/ICCWAMTIP.2018.8632592
8. Tan M., Santos C., Xiang B., Zhou B. Improved representation learning for question answer matching. *Proc. of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2016, pp. 464–473. https://doi.org/10.18653/v1/P16-1044

Литература

1. Lakshmi V., Ramesh V. Evaluating students' descriptive answers using natural language processing and artificial neural networks // International Journal of Creative Research Thoughts — IJCRT. 2017. V. 5. N 4. P. 3168–3173.
2. Patil S.M., Patil S. Evaluating student descriptive answers using natural language processing // International Journal of Engineering Research & Technology (IJERT). 2014. V. 3. N 3. P. 1716–1718.
3. Yin W., Kann K., Yu M., Schütze H. Comparative study of CNN and RNN for natural language processing // ArXiv. 2017. arXiv:1702.01923. https://doi.org/10.48550/arXiv.1702.01923
4. Kate A., Kamble S., Bodkhe A., Joshi M. Conversion of Natural Language Query to SQL Query // Proc. of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA). 2018. P. 488–491. https://doi.org/10.1109/ICECA.2018.8474639
5. Thangarasu S., Sasikala D. Extracting knowledge from XML document using tree-based association rules // Proc. of the International Conference on Intelligent Computing Applications (ICICA). 2014. P. 134–137. https://doi.org/10.1109/ICICA.2014.37
6. Mati D.N., Ajdari J., Raufi B., Hamiti M., Selimi B. A Systematic mapping study of language features identification from large text collection // Proc. of the 8th Mediterranean Conference on Embedded Computing (MECO). 2019. P. 8760042. https://doi.org/10.1109/MECO.2019.8760042
7. Cai J., Li J., Li W., Wang J. Deeplearning model used in text classification // Proc of the 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). 2018. P. 123–126. https://doi.org/10.1109/ICCWAMTIP.2018.8632592
8. Tan M., Santos C., Xiang B., Zhou B. Improved representation learning for question answer matching // Proc. of the 54th Annual Meeting of the Association for Computational Linguistics (ACL). 2016. P. 464–473. https://doi.org/10.18653/v1/P16-1044

9. Lan Y., Wang S., Jiang J. Knowledge base question answering with a matching-aggregation model and question-specific contextual relations. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2019, vol. 27, no. 10, pp. 1629–1638. https://doi.org/10.1109/TASLP.2019.2926125
10. Grappy A., Grau B., Falco M-H., Ligozat A-L., Robba I., Vilnat A. Selecting answers to questions from Web documents by a robust validation process. *Proc. of the IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, 2011, pp. 55–62. https://doi.org/10.1109/WI-IAT.2011.210
11. Lan Y., Wang S., Jiang J. Knowledge base question answering with a matching-aggregation model and question specific contextual relations. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2019, vol. 27, no. 10, pp. 1629–1638. https://doi.org/10.1109/TASLP.2019.2926125
12. Zhang W., Ming Z., Zhang Y., Liu T., Chua T. Capturing the semantics of key phrases using multiple languages for question retrieval. *IEEE Transactions on Knowledge and Data Engineering*, 2016, vol. 28, no. 4, pp. 888–900. https://doi.org/10.1109/TKDE.2015.2502944
13. Pakray P., Pal S., Bandyopadhyay S., Gelbukh A. Automatic answer validation system on English language. *Proc. of the 3rd International Conference on Advanced Computer Theory and Engineering*, 2010, pp. V6329–V6333. https://doi.org/10.1109/ICACTE.2010.5579166
14. Thalor M.A. A descriptive answer evaluation system using cosine similarity technique. *Proc. of the 4th International Conference on Communication information and Computing Technology (ICCICT)*, 2021. https://doi.org/10.1109/ICCICT50803.2021.9510170
15. 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

9. Lan Y., Wang S., Jiang J. Knowledge base question answering with a matching-aggregation model and question-specific contextual relations // IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2019. V. 27. N 10. P. 1629–1638. https://doi.org/10.1109/TASLP.2019.2926125
10. Grappy A., Grau B., Falco M-H., Ligozat A-L., Robba I., Vilnat A. Selecting answers to questions from Web documents by a robust validation process // Proc. of the IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology. 2011. P. 55–62. https://doi.org/10.1109/WI-IAT.2011.210
11. Lan Y., Wang S., Jiang J. Knowledge base question answering with a matching-aggregation model and question specific contextual relations // IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2019. V. 27. N 10. P. 1629–1638. https://doi.org/10.1109/TASLP.2019.2926125
12. Zhang W., Ming Z., Zhang Y., Liu T., Chua T. Capturing the semantics of key phrases using multiple languages for question retrieval // IEEE Transactions on Knowledge and Data Engineering. 2016. V. 28. N 4. P. 888–900. https://doi.org/10.1109/TKDE.2015.2502944
13. Pakray P., Pal S., Bandyopadhyay S., Gelbukh A. Automatic answer validation system on English language // Proc. of the 3rd International Conference on Advanced Computer Theory and Engineering. 2010. P. V6329–V6333. https://doi.org/10.1109/ICACTE.2010.5579166
14. Thalor M.A. A descriptive answer evaluation system using cosine similarity technique // Proc. of the 4th International Conference on Communication information and Computing Technology (ICCICT). 2021. https://doi.org/10.1109/ICCICT50803.2021.9510170
15. 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

Authors

Thalor Meenakshi Anurag — PhD, Associate Professor, AISSMS Institute of Information Technology, Pune, 411001, India, SC 57190340673, https://orcid.org/0000-0001-7048-7289, meenakshi.thalor@aissmsioit.org
Mane Pradeep B. — PhD, Principal, AISSMS Institute of Information Technology, Pune, 411001, India, SC 36450914400, https://orcid.org/0000-0002-3252-9983, pbmane6829@gmail.com
Mandge Vishaka — B.E., Student, AISSMS Institute of Information Technology, Pune, 411001, India, https://orcid.org/0000-0002-1667-4814, mandge1999@gmail.com

Авторы

Минакши Анураг Талор — PhD, доцент, AISSMS Институт информационных технологий, Пуна, 411001, Индия, SC 57190340673, https://orcid.org/0000-0001-7048-7289, meenakshi.thalor@aissmsioit.org
Прадип Б. Мане — PhD, директор, AISSMS Институт информационных технологий, Пуна, 411001, Индия, SC 36450914400, https://orcid.org/0000-0002-3252-9983, pbmane6829@gmail.com
Вишака Мандж — бакалавр, студент, AISSMS Институт информационных технологий, Пуна, 411001, Индия, https://orcid.org/0000-0002-1667-4814, mandge1999@gmail.com

Received 30.10.2021
Approved after reviewing 25.04.2022
Accepted 25.05.2022

Статья поступила в редакцию 30.10.2021
Одобрена после рецензирования 25.04.2022
Принята к печати 25.05.2022