IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

|e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907090 |

A Survey on Automatic Question Generation Using Machine Learning

Prof. Aparna Hambarde, Kunal Paudmal, Rushikesh Bhagwat, Rohan Bhapkar, Krishna Jadhav,

Dept. of Computer Engineering, KJCOEMR, Pune, Maharashtra, India

ABSTRACT: Curiosity is the fuel for discoveries and learning, and we humans are always curious about learning novel things. Asking different types of questions is a tendency of a good learner and so the students ask a variety of questions from their teachers. Similarly to examine the students, the teachers also ask questions from the students and also evaluate their performances. Because of their inconsistent minds in certain situations, sometimes humans are not very skilled in asking good questions and because of that we have come with the idea of a system with the help of which we would be able to generate the questions from a text automatically. The system is known as the Automated Question Paper Generator System, which is a fast and secure system and which can also generate questions in a random way. The document or pdf file or simply text can be provided as an input to this proposed system. NLP technology is used for preparing the proposed system.

I. INTRODUCTION

Traditionally question papers were generated manually. Preparing any exam paper is a very challenging task for the educators because they have to check whether there is any repetition in the pattern or not, and other than that security is one of the major concerns for them , also due to lack of teaching staff in any institute, creation of paper is not at all an easy task.

So, here we are proposing an Intelligent Automatic Question Paper Generator System which provides storage of the data, fast operations, and high security for all its tasks. It can be helpful to many educational as well as NGO institutes. Tasks performed by this system are automated and because of that storage space, security is not a concern anymore. Proposed system works upon Natural Language Processing and is fast due to computer based automation, streamlined, randomized and unbiased, secure and generates unique questions and overcomes the problems of Human based paper generation.

II. LITERATURE SURVEY

Hochschule fur Technik Stuttgart Schellingstr(2017)

Empiricall y verify that Bloom's taxonomy, a standard tool for difficulty estimation during question creation. Question difficulty estimates guide test creation, but are too costly for small-scale testing. We empirically verify that Bloom's Taxonomy, a standard tool for difficulty estimation during question creation, reliably predicts question difficulty observed after testing in a short-answer corpus. We also find that difficulty can be approximated by the amount of variation in student answers, which can be computed before grading. We show that question difficulty and its approximations are useful for automated grading, allowing us to identify the optimal feature set for grading each question even in an unseen questionsetting. Testing is a core component of teaching, and many tasks in NLP for education are concerned with creating good questions and correctly grading the answers. We look at how to estimate question difficulty from question wording as a link between the two tasks. From a test creation point of view, knowing question difficulty levels is imperative: Too many easy questions, and the test will be unable to distinguish between the more able test-takers, who all achieve equally good results. Too many hard questions, and only the most able test-takers will be clearly distinguishable from the (low-performing) rest.

Neung Viriyadamrongkij and Twittie Senivongse (2017)

InOnline inquiry communities such as Question-Answer Communities (QAC) have captured interest of online users since they can share and search for any information from any place in the world. The number of questions and answers submitted to a popular community can increase rapidly, and that can make it difficult for users who look for the "right" questions to answer. That is, from the view of knowledgeable experienced users, they tend to look for hard challenging questions as an opportunity to share their knowledge and to build respect with the community. Hence it is desirable to distinguish difficult questions from easy ones. Current researches estimate complexity of questions based on the analysis of the features of the QAC without considering the contents of the questions.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907090 |

This paper presents a method to measure question difficulty levels based directly on the question contents. In particular, we analyze the difficulty of terms that appear in a JavaScript-related question, based on the proposed JavaScript concept hierarchy.

In an evaluation of the performance of the question difficultyestimation, our conceptbased measure gives similar performance to that of the existing measure based on the features of the QAC, but when they are used together, the performance can be enhanced.

Sasitorn Nuthong ,Suntorn Witosurapot(2017)

In This Automatic Quiz Generation system is utterly handy for reducing teachers' workloads in quiz creation. Nevertheless, by exploiting a coarsegranular concern inside difficulty ranking mechanism, only a few number of automaticgenerated quizzes can be obtained. In order to increase the number of usable quizzes, we suggest how a 5-level difficulty ranking score using a hybrid similarity measurement approach together with property filtering of the key data can be potential for serving this propose. Based on experiment results, our proposed similarity measure outperforms three other candidates.

Surbhi Choudhary, Abdul Rais Abdul Waheed,Shrutika Gawandi, Kavita Joshi (2015)

In this modern world e-book has become a basic requirement for the candidates to appear and prepare for their competitive exams within college premises. In this paper we are proposing a model system for smart question paper generation of universities. The mechanism behind this system is that many random question papers are generated along with the difficulty level of the questions in terms of percentage. After generation that particular question is then mailed to the respective university. In this system administration of the database inputs set of question paper with an option of check box to tick th e correct answer. More ever weightage of the particular question in terms of marks and hours and the complexity of the question is determined. After this process whole question paper along with the weightage is stored in the database. In order to make question paper for 100 marks, admin sets all the weightage and difficulties to solve the problem. As soon as the difficulty and weightage is specified a pre doc file as per selected format will be downloaded to the admin and an electronic mail will be triggered. Range of difficulty may vary from easy, medium and hard.

Pawel Jurczyk ,Eugene Agichtein (2007)

In Question-Answer portals such as Naver and Yahoo! Answers are quickly becoming rich sources of knowledge on many topics which are not well served by general web search engines. Unfortunately, the quality of the submitted answers is uneven, ranging from excellent detailed answers to snappy and insulting remarks or even advertisements for commercial content. Furthermore, user feedback for many topics is sparse, and can be insufficient to reliably identify good answers from the bad ones. Hence, estimating the authority of users is a crucial task for this emerging domain, with potential applications to answer ranking, spam detection, and incentive mechanism design. We present an analysis of the link structure of a general-purpose question answering community to discover 4 authoritative users, and promising experimental results over a dataset of more than 3 million answers from a popular community QA site. We also describe structural differences between question topics that correlate with the success of link analysis for authority discovery.

II. PROPOSED SYSTEM

The main task is to create questions automatically based on the given text, i.e. sentences or groups of sentences. The basic input is a positive or declarative sentence or group of sentences. The output should be different kinds of questions depending on the type of sentences. Basically we can say that the input would simply be the text whereas output would be a number of questions. The basic goal of this system is to offer a solution to the problem of first breaking down the paragraph and then turning it into questions

There are different types of problems that exist in the current manual system of question paper generation. These problemsare:-

- Low security as paper is not secured using anymechanism.
- Patterns or repetitions may occur in paper.
- Slow as human labour isinvolved.
- Less varietyofdifferent types of questions.

Therefore to overcome this we came up with 'Automatic Intelligent Question Paper Generator'.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907090 |

III. CONCLUSION AND FUTURE SCOPE

Establishing mechanisms to control and predict the difficulty of assessment questions is clearly a big gap in existing question generation literature. Our contributions have covered the deeper aspects of the problem, and proposed strategies, that exploit ontologies and associated measures, to provide a better difficulty-level predicting model, that can address this gap. We developed the difficultylevel model (DLM) by introducing three learnerspecific logistic regression models for predicting the difficulty of a given question for three categories of learners. The output of these three models was then interpreted using the Item Response Theory to assign high, medium or low difficultylevel. The overall performance of the DLM and the individual performance of the three regression models based on cross-validation were reported and they are found to be satisfactoryComparison with the state-of-the-art method shows an improvement of 8.5difficulty-levels of benchmark questions. The model proposed in this paper for predicting the difficulty-level of questions is limited to ABoxbased factual questions. It would be interesting to extend this model to questions that are generated using the TBox-based approaches. However, the challenges to be addressed would be much more, since, in the TBox-based methods, we have to deal with many complex restriction types (unlike in the case of ABox-based methods) and their influence on the difficulty-level of the question framed out of them needs a detailed investigation. For establishing the propositions and techniques stated in this paper, we have implemented a system which demonstrates the feasibility of the methods on medium sized ontologies. It would be interesting to investigate the working of the system on large ontologies.

REFERENCES

1. UlrikePad o “Question Difficulty How to Estimate Without Norming,How to Use for Automated Grading” Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications, Copenhagen, Denmark, September 8, 2017.
2. -Neung Viriyadamrongkij and Twittie Senivongse “Measuring Difficulty Levels of pythonScript Questions in Question-Answer Community Based on Concept Hierarchy”978-1-5090-4834-2/17/ 31.00 ©2017 IEEE
3. -Itziar Aldabe Edurne Martinez “ArikIturri: an Automatic QuestionGenerator Based on Corpora and NLP Techniques” Conference Paper in Lecture Notes in Computer Science · June 2006
4. - Sarah K. “Two Methods for Measuring Question Difficulty and Discriminationin Incomplete Crowdsourced Data ”Association for the Advancement of Artificial Intelligence 2013
5. - Asma Ben Abacha, Marcos Da Silveira, and Cedric Pruski. Medical ontology validation through question answering. In AIME, pages 196–205, 2013. 10.1007/978-3-642-38326-7 30.
6. -Maha Al-Yahya. Ontology-based multiple choice question generation. The Scientific World Journal, Vol 2014, page 9, ID: 10.1155/2014/274949, 2014.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details