Educational and Information Technologies (2019) 24: 1925–1941
https://doi.org/10.1007/s10639-018-9813-4

RemedialTutor: A blended learning platform for weak students and study its efficiency in social science learning of middle school students in India

Dhawaleswar Rao CH[1] · **Sujan Kumar Saha**[1]

Received: 26 June 2018 / Accepted: 11 September 2018 / Published online: 18 September 2018

Abstract

Blended learning is one of the leading trends in education. Blended learning combines computer-assisted learning with traditional classroom learning. The literature shows that the blended learning often helps the students to achieve better learning outcome. However, a majority of the existing learning platforms do not focus on the problems of weak students. Here our objective is to develop a computer-assisted learning platform that focuses on performance improvement of weak students and study the efficacy of the system. This paper presents the proposed system, *RemedialTutor*, that assists the weak students in effective preparation for an examination. The learning platform performs several tasks on demand; for example, providing the meaning of unknown words, sentence simplification, identification of questionable sentences, extraction of summarized content on a specific topic, preparation of question paper and automatic evaluation, identification of less confident sections, etc. To study the effectiveness of the proposed system, it is tested using a blended learning framework. The system is provided to the students as a supplement to the traditional classroom activities and resources. During the comparative study, the experiment group students used this system during their exam preparation but the control group students relied only on their regular resources. It is found that the experiment group students perform better than the control group. The t-value is 2.3466 and p-value is 0.0243. These values indicate that the difference is statistically significant.

Keywords Computer assisted learning · Blended learning · Efficacy of CAL · Questionable sentence identification · Automatic question generation and evaluation · Natural language processing

✉ Sujan Kumar Saha
sujan.kr.saha@gmail.com

[1] Department of Computer Science and Engineering, Birla Institute of Technology Mesra, Ranchi, 835215, India

1 Introduction

The recent rapid advances in information technology and internet technology have had a significant impact on the education system. E-learning or computer-assisted learning or web-based learning have already become quite popular in many countries due to the wide availability of personal computers, internet and web, easier and cheaper access to the internet, and advanced information technology. However, the e-learning or distance learning alone cannot fulfill all the objectives of school education. Traditional face-to-face learning also has several advantages, such as, better able to focus on what is being taught, learning in a social interaction environment, learning from other participants, facilitates an exchange of ideas, lowers possibility of misunderstanding, the scope of customization etc. On the other hand, e-learning or distance learning offers convenience and flexibility in content, timing, duration and style; learning in a comfortable environment; more and more information on demand; and student-specific customization or demand fulfilling. As both the traditional classroom learning and computer-assisted learning (CAL) simultaneously offer strengths, these can be combined for a better learning outcome.

Blended learning is one of the leading trends in the modern era of education. Blended learning combines e-learning with traditional classroom learning to ensure maximum effectiveness (Bersin and Associates 2003; Dziuban et al. 2004; Graham 2006; Kim 2007; Ahmad 2010; Sharma 2017). In this context, the term e-learning refers to learning outside the traditional classroom using information technology for the delivery of the learning materials. That covers a wide range of information technology medium - computer-assisted learning, online learning resources, video streaming over the web, learning through e-mail or chat discussion, learning through an online community, self-assessment tools etc. Several studies were performed in various settings that often proved the superiority of blended learning over traditional classroom learning (Bayraktar 2000; Christmann and Badgett 2000; Dean et al. 2001; DeLacey and Leonard 2002; Rayne and Baggott 2004; Ahmad 2010; Yapici and Akbayin 2012).

However, traditional or face-to-face learning is still the most widely used and popular teaching method in school education in India. The main reason behind this is twofold: (i) the teacher community drew up in a traditional learning environment and (ii) insufficiency of technology support in India. Unavailability of the sufficient number of computers in schools, lack of exposure to computers, unavailability of tools that support language or subject-specific computer-assisted learning, insufficiency of high-speed internet, deficiency of expertise etc. are common facts in rural and semi-urban schools in India. A study is performed by Shashi Rekha (2013) on "blended learning in India: are teachers in India ready to go blended?" This study aimed to find out whether India has the basic computer skills and examines the attitudes of the teachers towards computer-based learning. The study reported that a large portion of the teachers does not prefer computer-assisted learning. However, the scenario is changing day by day and the interest, infrastructure and support are also increasing to implement e-learning or blended learning in school education in India. Development of various objective specific learning platforms and study the effectiveness of blended learning in Indian learning environment is essential to increase the overall trend.

In this scenario, we study the performance of blended learning in a middle school education setup in India. Here our objective of using computer-assisted learning is to leverage the examination performance of the students. Particularly we targeted the weak students. We label a student as 'weak' if his score in the previous examination is less than 50%. If the social science score in last class test or average score in last annual examination is below 50% then the student is included in this study. Every child must get the school education to the overall upliftment of the society. But when a middle-school student gets a 'fail' grade, it affects a lot to the student, the family, and the society. A set of remedial classes, giving suggestions for examination, identification of important portions from the textbook, more practices, trial examination etc. can help the student to perform better in the examination. Due to higher student-teacher ratio here, providing sufficient individual care is also difficult in school. We develop a computer-assisted learning platform to overcome this situation. The developed platform aims to assist the student in effective preparation for his examination. The CAL system performs a set of tasks: (i) sentence reformation and providing meaning of unknown words, (ii) automatic identification of questionable sentences from the textbook chapters, (iii) summarized description on a particular concept, (iv) providing solution of existing questions, (v) taking trial examination for growing confidence, (v) automatic assessment of the student answers, and (vi) identification of the portions where revision is required.

After the development of the platform, we study its performance in school education in India. The system is tested in examination preparation of class VII history. Here the strategy followed is to provide the needed materials to students just when these are needed. So, the students attend regular normal classes through which the course is taught and the system is provided only during their preparation for the examination. To study the effectiveness, two student groups are formed, each consisting of 20 students. The CAL system is given to one of these groups and the other group depends only on the traditional classroom resources. The comparison through pre-test and post-test performance of the students shows that the use of the developed system helps in significant improvement in performance.

The current paper contributes to the literature in several ways.

1. The primary objective of this paper is to develop a computer-assisted learning platform to assist the weak students.
2. The key feature of the developed CAL system is also novel. The system acts as a remedial tutor to provide customized resources for effective preparation for the examination.
3. The effectiveness of the CAL platform is studied in learning social science in school education in India. Such a study is also not available in the current literature.

2 Related work

Development of information and communication technology (ICT) infrastructure is one of the key priorities of education policies of many developing countries including

India. The core objective is on investing public resources in the implementation of ICT in education. Computer-based education and blended learning have been implemented and studied in various developed countries in the last decade. However, many of the developing countries still lagging. In India, the ICT infrastructure in schools have been increasing day-by-day in urban areas and privately funded schools but the overall scenario is still inadequate. Here a big challenge is to formulate the strategy on how to use ICT in teaching and learning so that it impacts the pupils' learning outcomes and attitudes.

Numerous studies are there in the literature that assess the effectiveness of computer-assisted instruction (CAI) or e-learning in education. Blended learning setup is also used in several studies. Here we mention a few from the literature.

Kulik (1994) performed one of the early meta-analysis to show that the computer based instruction are more effective that traditional classroom instruction. Sule Bayraktar (2000) performed a meta-analysis of the effectiveness of computer assisted instruction in science education. Effectiveness of computer assisted learning was found on student achievement in secondary school and college science education when compared to traditional instruction. An overall effect size of 0.273 was calculated from 42 studies yielding 108 effect sizes. The results of the study also indicated that some study characteristics such as student-to-computer ratio, CAI mode and duration of treatment were significantly related to the effectiveness of CAI. Cheung and Slavin (2012) reviewed 84 studies to investigate the effectiveness of computer based learning on the reading performance of K-12 students. They found positive significant effects of computer environments compared with traditional classroom environments.

Effectiveness of computer assisted instruction had been studied in specific subjects too. For example, several studies are there in the literature that shows the effectiveness of CAL in learning mathematical concepts of the students. Reimer and Mayer (2005) investigated the effectiveness of computer assisted instruction in learning mathematics. The results of the study showed that the use of computer based system in teaching mathematical concepts to pupils helps them in better understanding fractions. Several other studies also prove that computer assisted instruction can provide substantial benefits for student's mathematics learning (e.g., Roschelle et al. 2010; Koedinger et al. 2010; Pilli and Aksu 2013). Karaksha et al. (2011) studied the effectiveness of CAL platform in pharmacology education. Wang and Liao (2017) studied the effectiveness of CAL framework in students' English proficiency.

Effectiveness of CAL in a blended learning setup is also studied by many researchers. For instance, Christmann and Badgett (2000) studied the difference in achievement levels between pupils who were taught by traditional class-room approach and pupils who used computer based learning as a supplement to the traditional classes. They compiled data from 26 studies. The overall results suggested a mean effect size of 0.127. Hence, pupils who were taught via an educational software as supplement to traditional teaching performed better compared to the other pupils group. Rayne and Baggott (2004) performed a meta-analysis of 40 studies that scrutinized the differences in effectiveness of blended learning over traditional teaching approach. They concluded that the blended traditional-CAL

teaching approach was more effective as enables pupils to realize higher levels of achievement. Yapici and Akbayin (2012) presented a study that aims to determine the effect of the blended learning model on high school students' biology achievement. The research results revealed that the blended learning model contributed more to the students' biology achievement than traditional teaching methods did. Certain articles focus on development of CAL platform too. For example, an infrastructure for web-based computer assisted learning was proposed in Joy et al. (2002) and Lin et al. (2002) proposed a expert system framework for computer assisted learning.

However the literature is not so rich when the study on blended learning or CAL is focussed in Indian learning environment. Banerjee et al. (2007) studied the impact of supplementing classroom instruction with computer assisted learning in primary schools. They observed that CAL significantly improved student's scores in mathematics, but this was less cost-effective than the remedial tutor-based program. Ray (2010) studied the challenges of web based e-learning in India. They also discussed a few innovative methods to spread e-learning in India. However, they did not study the relative performance of e-learning over traditional learning. Sharma and Malhotra (2016) performed a comparison of computer assisted learning and practical animal experiment for undergraduate medical students in pharmacology curriculum. They conducted a questionnaire based study in a medical college of North India. The study concluded that all the students and teachers believed that CAL is better and more understandable than animal experiments done practically. Also, CAL is a better source of experimentation at the undergraduate level and provides dependable outcomes. However, the study was done purely through a questionnaire based approach; the comparison of performance of the CAL and traditional experiment based learning was not reported.

3 RemedialTutor: A computer assisted learning platform

The proposed computer assisted learning platform is described here. The developed platform acts as a remedial tutor. It primarily aims to assist the weak students in effective preparation for their examination. Therefore it is named 'RemedialTutor'. Figure 1 presents a sample page of the developed system.

3.1 Features of the CAL platform

RemedialTutor performs several responsibilities that help in the learning of the students. A text document, like, a chapter from the textbook, is taken as input to the system. Then it analyses the text and performs a set of tasks that aims to help the student in the effective study and examination-focussed preparation. As the system intends to help weak students as a remedial tutor, the primary features of the system aimed at that direction. The system performs the following tasks on demand:

– The system performs certain tasks to make the reading easier. A student might face difficulty with long and complex sentences. The system converts these

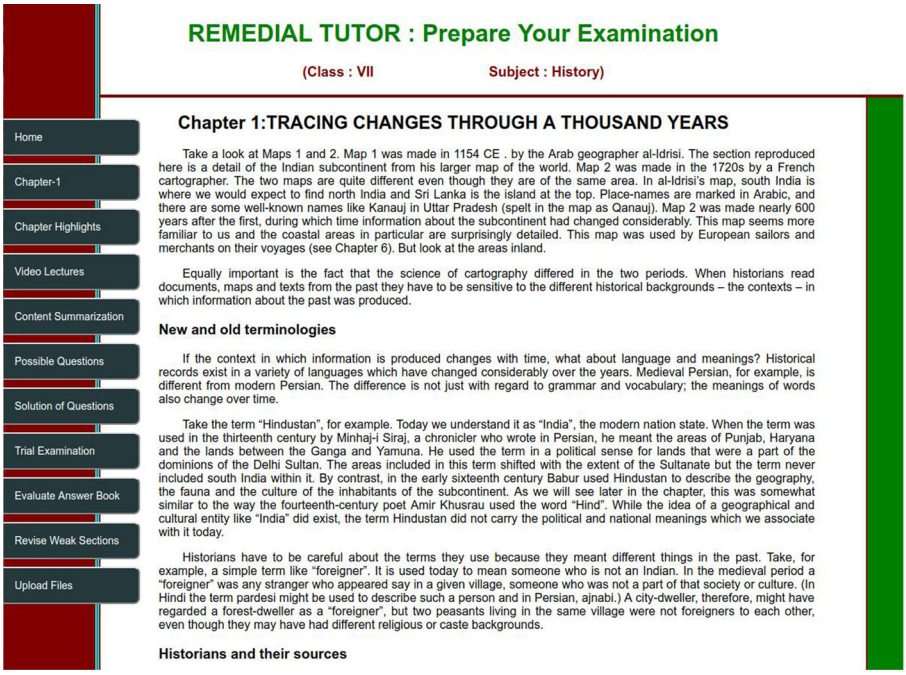

Fig. 1 A sample view of the developed RemedialTutor system

into small and simple sentences. It also provides the meaning of the words, if demanded by the student.

- Identification of factual sentences from the textbook chapter. Reading and remembering the whole chapter is difficult for a weak student but he feels comfortable if the content is less. Additionally, objective questions like fill-in-the-blanks (FIB), multiple-choice question (MCQ), short answer type question etc. are generally formed from the factual sentences.
- Long description on a particular topic often causes loss of interest of the student. The system provides a summarized and simple description on a topic the student is interested in. The system also helps in selecting the topics that are important for the examination by analyzing existing questions.
- Conducts trial examination. Appearing in a trial examination before the actual examination often helps in growing the confidence. The system can generate a test paper for the examination.
- The system is also capable of assessing the student answers instantly.
- Through this trial examination, the system also identifies the portions where the confidence level of the student is low. Then the student is suggested to revise those portions.

Figure 2 presents a high-level workflow of working of the overall RemedyFinder system. Working on the individual modules are discussed in the following section.

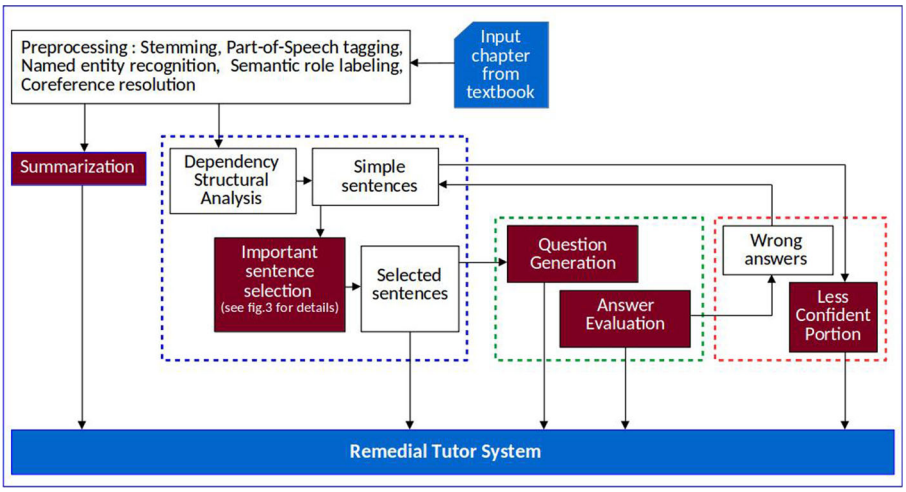

Fig. 2 Workflow of the RemedialTutor system

4 Development of the RemedialTutor system: Technical details

Now the strategies adopted for developing the key components of the system are discussed. The system contains several modules and each module uses multiple techniques; these are summarized below.

4.1 Sentence reformation

Students might face difficulty in understanding or memorize long sentences. Complex and compound sentences often contain multiple facts. However, one FIB or MCQ question normally deals with a single fact. Therefore, from the questioning point of view also, simple sentences are better. So, the system converts the complex and compound sentences into simple sentences. The conversion is done by analyzing the dependency parse structure of the sentences. The dependency structure is obtained from the Stanford CoreNLP suite[1] (Marneffe and Manning 2008). A simple sentence contains only one subject (nsubj or nsubjpass). For example, the *root* of the sentence "Kabir rejected all forms of ritual" is 'rejected' and the *nsubj* is 'Kabir'. So, this sentence is a simple sentence. Now consider the sentence, "Kabir rejected all forms of ritual and he tried to build a bridge between Hindu and Muslims". Dependency structure of this sentence shows that it contains a *conjunction* between two verbs: 'rejected' and 'tried' and contains two *nsubs* as 'nsubj(rejected, Kabir)' and 'nsubj(tried, he)'. When a sentence contains multiple subjects then the sentence is divided into simple sentences using the dependencies among the words.

The text often contains various pronouns. As the system aims to pick a subset of the sentences, these pronouns should be replaced by corresponding nouns to make

[1] https://stanfordnlp.github.io/CoreNLP/

the fact complete. So, the co-reference resolution is performed on the sentences using the ARKRef toolkit (O'Connor and Heilman 2013). If the meaning of any word is unknown to the student, he gets assistance from the system. The system uses the WordNet (Miller et al. 1990) to find the glosses and synonyms of the word.

4.2 Factual sentence identification

All the sentences in a text do not contain a questionable fact. This step aims to identify the sentences that can act as the basis of question formation. These sentences are treated as important.

Names play a major role in this domain. Most of the informative sentences contain one or more names. A sentence is likely to contain a fact if it contains a name. So, this is considered as the primary criteria for sentence selection. But the only occurrence of a named entity in a sentence does not guarantee that the sentence is questionable. So, additional checking is performed for effective sentence selection. Figure 3 presents a high-level workflow of the methodology used for sentence similarity computation. Two different approaches have been used for informative sentence selection.

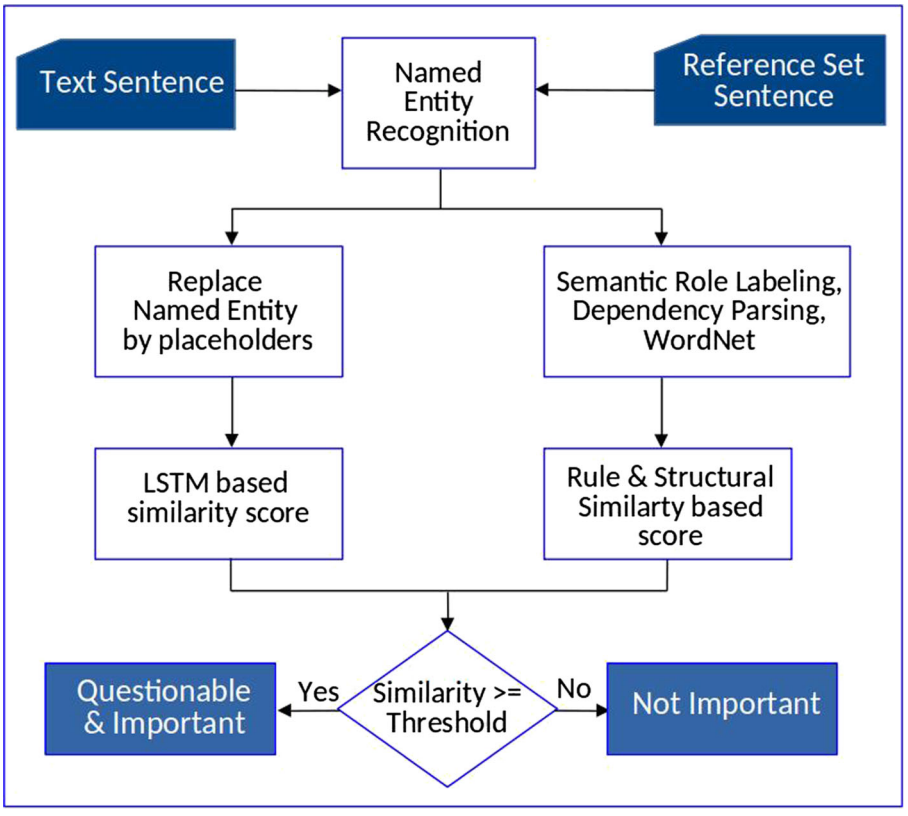

Fig. 3 A high-level overview of the strategy for selection of questionable sentences

The first approach of sentence selection is based on computing similarity between existing question sentences and sentences in the text. Existing fill-in-the-blank and multiple-choice questions act as a reference set. As the reference set is prepared from existing questions, these are questionable. Now, if a sentence from the text has a high similarity with any of the reference set sentences, then that sentence is also questionable. For the similarity computation, we use the Long-Short-Term-Memory (LSTM) based InferSent (Conneau et al. 2017). Recurrent Neural Networks are neural networks adapted for variable length sequence data ($x_1, \ldots, x_T$). RNNs uses a hidden vector whose activation is dependent on the immediate previous vector. Standard RNNs suffer from the vanishing gradient problem in which the backpropagated gradients become vanishingly small over long sequences. The Long Short Term Memory model was proposed as a solution to this problem. Detailed discussion on the working of LSTM can be found in Hochreiter and Schmidhuber (1997). InferSent is a sentence embeddings method that provides semantic sentence representations. It uses bi-directional LSTM with mean-max pooling to form a fixed-size vector. It is done either by selecting the maximum value over each dimension of the hidden units (max pooling) or by considering the average of the representations (mean pooling). After named entity recognition and replacement of the names by corresponding placeholders, InferSent is used to find the similarity between the reference set sentence and a text sentence. If the similarity is higher than a threshold then it is selected as informative. The threshold value is chosen through the experiments conducted during the development phase. The threshold obtained in our experiments is 0.75.

The second approach is based on the extraction of predicate-argument structure. For informative sentence selection, Majumdar and Saha (2014) proposed a parse-tree similarity-based approach. We adopt this technique. Additionally, we analyzed the relationship of the name with other words in the sentence is analyzed. This is done through predicate argument extraction. For the predicate-argument extraction, the SENNA toolkit (Collobert et al. 2011) is used. A subject-specific list of important predicates is prepared. For example, the predicates ‘defeated’, ‘conquered’, ‘born’, ‘succeeded’, ‘recruited’, ‘assassinated’ etc. are important in History subject. Availability of such predicates in a sentence indicates that the sentence may contain a questionable fact. Multiple arguments can be associated with these predicates. For example, the verb ‘defeated’ can have arguments like date, ruler1, ruler2, location etc. Some of these arguments are as essential and some are additional arguments. If a sentence contains an important predicate with the essential arguments, then the sentence is treated as important.

4.3 Summarized description on a topic

The system can also help the students during their preparation for descriptive questions. Given an input topic, the system locates all the sentences that discuss the topic. If the number of sentences is large, then the system provides a summarized description of the topic.

For the summarization, the TextRank algorithm (Mihalcea and Tarau 2004) is employed. TextRank builds a graph associated with the text, where the vertices are representative for the sentences. To build the graph, TextRank uses ‘similarity’

relation between the sentences, where similarity is a function of their content overlap. Such similarity is extracted as a process of 'recommendation'. A sentence that addresses certain concepts in a text, gives the reader a recommendation to refer to other sentences in the text that address the same concepts. Therefore, an edge can be drawn between any two such sentences that share common content. The high ranked sentences returned by the TextRank algorithm are then transformed into a simple form and provided to the student.

4.4 Conducting test: Question paper generation and answer evaluation

The system is also capable of conducting a trial examination. The test papers are primarily composed of a preloaded question bank. The textbook chapters often contain a set of practice questions. Also, the system can accept the previous year question papers when imported in a specified format. The question bank is compiled using these questions. The system randomly selects a specified number of questions from the question bank and generates a test paper.

Additionally, the system is capable of automatic generation of fill-in-the-blank questions. The system has already identified the questionable or factual sentences. These sentences contain one or more named entities. One of the named entities is replaced by a blank to form a fill-in-the-blank question.

The answers of the student are also evaluated by the system. Automatic evaluation of fill-in-the-blanks and multiple-choice questions are straightforward. In the case of multiple-choice questions, the evaluation task is to check whether the option number written by the student is correct. This is a simple character matching. Similarly, word or phrase matching is required to evaluate the fill-in-the-blank questions. But evaluation of descriptive question is a tricky task. The system uses a co-occurrence vector-based strategy for the task. For this, the system needs a model answer. A good answer should contain a similar set of words. To compute the co-occurrence vectors, a word list is composed that contains all the words in the corresponding chapter. Stopwords are removed from the list. The dimension of the vector is the same as the number of words in them. Initially, the vector for an answer is considered as a null vector. Then the words of the answer are scanned one-by-one from the beginning to the end and the value of the corresponding dimension is increased by one. So, if a word w_j occurs t times in the answer then the value of the jth position of the vector is t. Such vectors are formed for both the model answer and student answer. Then the similarity between the answers is taken as the cosine similarity[2] between the corresponding vectors and normalized to the required scale.

4.5 Identification of less confident portion

The system then analyses the score of the student to identify the portions where his preparation is not satisfactory. Then the student is suggested to revisit the corresponding sections. This is done by using section wise topic word extraction. The chapter

[2] https://en.wikipedia.org/wiki/Cosine_similarity

is divided into sections with the help of the headings and section number information. Then from each section, topics are extracted using Latent Discriminant Analysis (LDA) based approach (Blei et al. 2003). LDA is a probabilistic topic model. LDA is applied on the text with a number of topics as two for three iterations and take the intersection of these words as the final set of topics.

Now, the keywords of an unanswered or wrongly answered question are matched with the section specific topic words. Keywords of a question are the set of words in the corresponding sentence except for the question words, wh-words and stop-words. Such keyword extraction is done from the objective questions. But for a long or descriptive question, keywords are the topics retrieved by LDA from its model answer. The section having the highest match with the keywords is labeled as 'Requires Revision'. These sections are exposed to the student for a revision.

5 Study the effectiveness of the system

After the development of the RemedialTutor system, we study its effectiveness. The details of the experimental setup and results are discussed below.

5.1 Experimental setup

The effectiveness of the developed system is studied in school education in India. We have already mentioned that infrastructural support is not sufficient in Indian schools excepts International schools and a few private and government-aided schools in urban areas. However, the proposed RemedialTutor platform is focussed on providing support for examination preparation. So, the availability of computer and related infrastructure in school is not a necessity; rather, a computer should be available in their home. The proposed CAL platform is focussed to guide the weak students. Therefore, to study the effectiveness of the system, a group of students is selected who satisfy two criteria: (i) marks in last year final examination or last class test is below 50% and (ii) owns computer system and internet at home. A total of 20 students were selected with these criteria. As the objective of this study is to compare the effectiveness of the developed computer-assisted learning system in examination preparation, another group of students is chosen for the comparison purpose. This second group also contains 20 students. So, the total students are divided into two groups: 'experiment group' and 'control group'. Only the experiment group students studied with the RemedialTutor system. The setup is based on the strategy to provide the needed CAL support to the students just when needed. So, the CAL system was provided two days prior to the examination.

The study is done in the subject Social Science - History of class VII students. Both the experiment and control group students attended regular classes, studied from the same set of class teachers, participated in the same in-class activity, and appeared examination with the same question paper. The only difference is, the experiment group used the RemedialTutor system as a supplement to the traditional classroom resources. Performance of the students in the previous class test is recorded as their 'pre-test score'. So, the pre-test performance indicates their scores when they used

only the in-class materials and no CAL support is provided. During the preparation of the next class test, the RemedialTutor platform was provided to the experimental group students. Then their scores were recorded and it is referred to as the 'post-test score'. The pre-test and post-test scores are studied. Group-wise relative change in performance reflects the effectiveness of the developed CAL platform.

5.2 Performance comparison through students T-test

To compare the effectiveness of the CAL platform, a t-test is used Gosset (1908). To indicate the performance various metrics have been used. These are, mean score, standard deviation, t-value, and p-value. The t-value is the ratio between the difference between two groups and the difference between the groups. A larger t-value indicates more difference between the groups. A smaller t-value refers to a high amount of similarity between the groups. For instance, a t-value of x means that the groups are x times different from each other as they are within each other. The p-value is the probability that the results occurred by chance. It is denoted as a decimal value ranges from 0 to 1. A low p-value is good, which indicates that the result did not occur by chance.

As the students are divided into two groups, both the groups should contain the similar type of students. If one group is superior to the other, then the test result might not be conclusive. To ensure the similarity between two groups, their pre-test scores are compared. The mean of the scores obtained by the control group student is 7.55 and mean of the experiment group is 7.65. This scores are very close and indicate that the groups contain a similar type of students. We also find the t-value and p-value. The t-value is 0.1215 and the corresponding p-value is 0.904 with a degree of freedom of 38. These values also indicate that the difference between these groups is not statistically significant. The pre-test analysis and corresponding values are summarized in Table 1.

Then, the RemedialTutor platform is provided to the experiment group students. Another class test is conducted after providing the CAL support. In order to find the effectiveness of the learning platform, the *post-test* scores were analyzed similarly. When we compute the mean values, we found that the mean of the experiment group (10.4) is higher than the control group (8.2). We further compute the t-value. The post-test analysis is summarized in Table 2. The t-value is 2.3466 and the corresponding p-value is 0.0243. These values indicate that the difference between the groups is statistically significant.

From the results presented in Tables 1 and 2, it is clear that the developed RemedialTutor platform is effective to assist the students during their preparation for an

Table 1 Comparison of Pre-test scores of the students in the Experiment and Control Groups

Group	No. of students	Mean score	SD	SE_d	t-value	p-value
Experiment Group	20	7.65	2.62	0.823	0.1215	0.904
Control Group	20	7.55	2.58			

Table 2 Comparison of Post-test scores of the students in the Experiment and Control Groups

Group	No. of students	Mean score	SD	SE_d	t-value	p-value
Experiment Group	20	10.4	2.85	0.938	2.3466	0.0243
Control Group	20	8.2	3.07			

examination. The scores of the individual students in the pre-test and post-test examination is plotted in Figs. 4 and 5 respectively for the experiment group and control group. Performance improvement of the experiment group students is clearly visible from Fig. 3.

5.3 Discussions

The comparative analysis of performance shows the effectiveness of blended learning through the proposed RemedialTutor platform over the traditional classroom learning. The students performed better in their tests when the platform is provided as a supplement to the normal classroom resources.

When the minimum score of the experiment group students is analyzed, it is observed that the minimum score is increased to 6 in post-test, compared to 3 in the pre-test. When the corresponding answer books were analyzed, it was found that the students answered a larger amount of objective type questions. A subset of the sentences were identified as important by the system, the students read these carefully and it helped them to achieve better score. The highest score of the experiment group is also increased, 13 in pre-test and 16 in post-test. When the high-scoring answer books were analysed, it was observed that these students performed better in long questions too. Several students achieved much higher score in post-test (9 to 15, 10

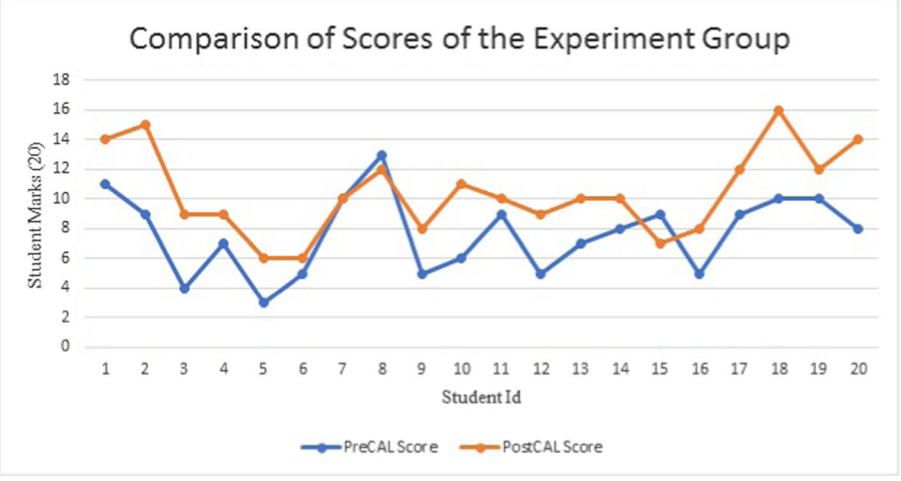

Fig. 4 Comparison of scores in pretest and posttest: experiment group students

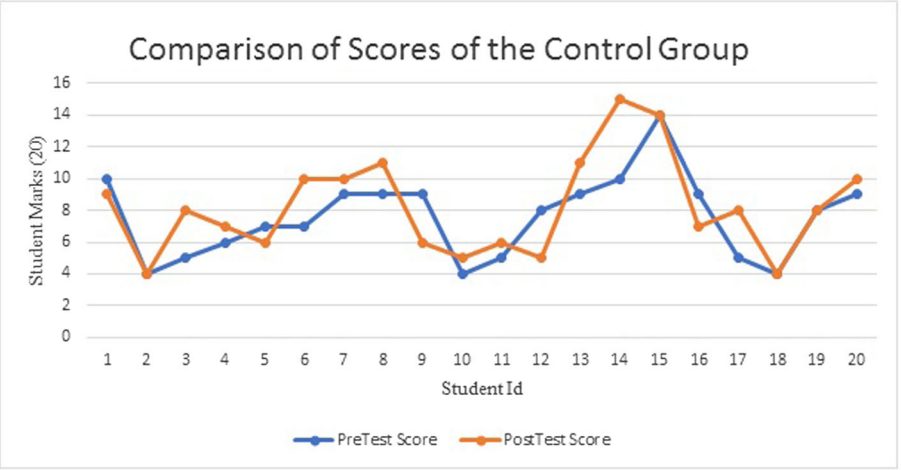

Fig. 5 Comparison of scores in pretest and posttest: control group student

to 16 etc.). These students were interviewed and they claimed that the developed system provided substantial help in faster preparation for objective questions and they got time to study long questions and previous years question papers that helped them in gaining higher marks.

6 Conclusion and future work

In this paper, we have presented a computer-assisted learning platform that assists the weak students in effective preparation for their examination. The proposed system provides the meaning of the unknown words, identifies questionable sentences, provides simple and summarized the discussion, conducts a trial examination, evaluates student answers, and identifies the portions where the student is not confident. These features of the system work like a remedial tutor and guide the students in various ways. The effectiveness of the system is tested in learning of class VII social science through a blended learning framework. The experiment group students were given the CAL framework as a supplement to the normal classroom activities. The control group students did not get the CAL support. To show the effectiveness of the system, their pre-test and post-test performance were compared. The t-test analysis showed that the performance of the experiment group students is significantly better than the control group.

There are a number of directions for future work. We feel that the proposed platform is generic and can be applied to other subjects and learning domains. However, it requires experimental verification. So, similar CAL system can be developed for other subjects and their effectiveness can be studied. Also, another study can be performed to see the effect of the platform on high scoring students. We hope, a similar system that retrieves a larger amount of relevant resources from the web, can fulfill the need of high scoring students. The proposed CAL platform can

be extended to provide more relevant e-resources. One can extend our system by incorporating more sophisticated approach for automatic generation of questions of various types including subjective questions and automatic generation of answers. Incorporation of a robust automatic answer evaluation module is another possible direction. Additionally, similar systems can be developed for the language-specific learning environment. In India, several state boards are of a non-English medium, where learning and evaluation are done through various Indian languages including Hindi, Bengali, Telugu, Tamil etc. A large portion of government schools in India uses Indian languages as the primary medium. Language-specific computer-assisted learning platform should be developed to implement e-learning or blended learning in such a scenario.

Funding Information This work is partially supported by the sponsored research project grant (project file no.: YSS/ 2015/ 001948) provided by the Science and Engineering Research Board (SERB), Govt. of India.

Compliance with Ethical Standards The authors declare that:

- This article does not contain any studies with human participants or animals performed by any of the authors.
- For this type of study formal consent is not required.

Conflict of interests The authors declare that they have no conflict of interest.

References

Ahmad, A.-H. (2010). *Constructivism based blended learning in higher education*. Master Thesis, Master of Management Information Systems. Universiteit Hasselt.

Banerjee, A., Shawn, C., Esther, D. (2007). Lindenleigh remedying education: evidence from randomized experiments in India. *The Quarterly Journal of Economics*, *122*(3), 1235–1264.

Bayraktar, S. (2000). *A meta-analysis of the effectiveness of computer-assisted instruction in science education*. Doctoral Dissertation, Athens: Ohio University. ISBN:0-599-86473-7.

Bersin and Associates (2003). Blended Learning: What Works? http://www.e-learningguru.com/wpapers/blended_bersin.doc.

Blei, D.M., Ng, A.Y., Jordan, M.I., Lafferty, J. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*.

Cheung, A.C., & Slavin, R.E. (2012). How features of educational technology applications affect student reading outcomes: a meta-analysis. *Educational Research Review*, *7*(3), 198–215. https://doi.org/10.1016/j.edurev.2012.05.002.

Christmann, E.P., & Badgett, J.L. (2000). The comparative effectiveness of CAI on collegiate academic performance. *Journal of Computing in Higher Education*, *11*(2), 91–103.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P. (2011). Natural Language Processing (Almost) from Scratch. *Journal of Machine Learning Research (JMLR)*.

Conneau, A., Douwe, K., Holger, S., Loic, B., Bordes, A. (2017). Supervised learning of universal sentence representations from natural language inference data. Volume 1. arXiv:1705.02364.

Dean, P., Stahl, M., Sylwester, D., Pear, J. (2001). Effectiveness of combined delivery modalities for distance learning and resident learning. *Quarterly Review of Distance Education*, *2*(3), 247–254.

DeLacey, B.J., & Leonard, D.A. (2002). Case study on technology and distance in education at the Harvard business school. *Educational Technology and Society*, *5*(2), 13–28.

Dziuban, C., Hartman, J., Moskal, P. (2004). Blended learning. Research bulletin, university of central Florida. *Educause Center for Applied Research*, *2004*(7), 1–12.

Gosset, W.S. (1908). The probable error of a mean. *Biometrika*, *6*(1), 1–25.

Graham, C. (2006). Blended learning systems: definition, current trends, and future directions. In *Handbook of blended learning: global perspectives, local designs* (p. 2006). San Francisco: Pfeiffer Publishing.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Joy, M., Muzykantskii, B., Rawles, S., Evans, M. (2002). An infrastructure for web-based computer-assisted learning. *ACM Journal of Educational Resources*, *2*(4), 1–19.

Karaksha, A., Grant, G., Davey, A., Dukie, S.A. (2011). Development and evaluation of computer-assisted learning (CAL) teaching tools compared to the conventional didactic lecture in pharmacology education. In *Proceedings of Edulearn11- international conference on education and new learning technologies* (p. 3580–3589).

Kim, W. (2007). Towards a definition and methodology for blended learning. In *Proceedings of the workshop on blended learning 2007* (p. 1–8). Edinburgh, United Kingdom.

Koedinger, K.R., McLaughlin, E.A., Heffernan, N.T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research*, *43*(4), 489–510.

Kulik, J.A. (1994). Meta-analytic studies of findings on computer-based instruction. In Baker, E.L., & O'Neill, H.F. (Eds.) *Technology assessment in education and training* (pp. 9–33). Hillsdale: Lawrence Erlbaum.

Lin, Y.T., Tseng, S.S., Tsai, C.J. (2002). The design and implementation of a computer-assisted learning expert system. *International Journal of Computer Processing of Languages*, *15*(1), 33–61.

Majumdar, M., & Saha, S.K. (2014). Automatic selection of informative sentences: the sentences that can generate multiple choice questions. *Knowledge Management and E-Learning: An International Journal.*, *6*(4), 377–391.

Marneffe, M.C.D., & Manning, C.D. (2008). *Stanford typed dependencies manual* (pp. 338–345). Technical Report, Stanford University. https://nlp.stanford.edu/software/dependencies_manual.pdf.

Mihalcea, R., & Tarau, P. (2004). TextRank: bringing order into texts. In *Proceedings the conference on empirical methods in natural language processing (EMNLP-04)*.

Miller, G.A., Beckwith, R., Fellbaum, C.D., Gross, D., Miller, K. (1990). Wordnet: an online lexical database. *International Journal of Lexicographics*, *3*(4), 235–244.

O'Connor, B., & Heilman, M. (2013). ARKRef: a rule-based coreference resolution system. arXiv:1310.1975.

Pilli, O., & Aksu, M. (2013). The effects of computer-assisted instruction on the achievement, attitudes and retention of fourth grade mathematics students in North Cyprus. *Computers and Education*, *62*, 62–71.

Ray, P.P. (2010). Web based e-learning in India: the cumulative views of different aspects. *Indian Journal of Computer Science and Engineering*, *1*(4), 340–352.

Rayne, R., & Baggott, G. (2004). Computer-based and computer assisted tests to assess procedural and conceptual knowledge. In *Proceedings of the 8th CAA conference*. Loughborough: Loughborough University.

Reimer, K., & Mayer, P.S. (2005). Third-graders learn about fractions using virtual manipulatives: a classroom study. *The Journal of Computers in Mathematics and Science Teaching*, *24*(1), 5–25.

Sharma, R. (2017). Computer assisted learning – a study. *International Journal of Advanced Research in Education and Technology (IJARET)*, *4*(2), 102–105.

Shashi Rekha, M. (2013). *Blended learning in India: are teachers in India ready to go blended? Technical report*. India: Jain University.

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., Gallagher, L. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: three large-scale studies. *American Educational Research Journal*, *47*(4), 833–878.

Sharma, D., & Malhotra, P. (2016). A comparison of computer assisted learning and practical animal experiment for undergraduate medical students in pharmacology curriculum - a questionnaire based study conducted in a medical college of North India. *International Journal of Basic and Clinical Pharmacology*, *5*(6), 2581–2584.

Wang, Y., & Liao, H.C. (2017). Learning performance enhancement using computer-assisted language learning by collaborative learning groups. *Symmetry*, *2017*(9), 141. https://doi.org/10.3390/sym9080141.

Yapici, I., & Akbayin, H. (2012). Umit the effect of blended learning model on high school students. *Biology Achievement and on Their Attitudes Towards the Internet the Turkish Online Journal of Educational Technology*, *11*(2), 228–237.