The effect of data mining on expert systems used for improving efficiency of correct speech E-learning systems

Zohreh Sedighian and Mahdi Javanmard

Payame noor University, Iran, Tehran

ABSTRACT

Analysis and storage of teachers' behavior regarding correct speech education as well as finding expert approaches and suggestion for speech correction by expert system will provide better feedback. Knowledge Based Expert Systems operate as the assistant of expert person by collecting related knowledge. On the other hand, converting human knowledge to the knowledge of an expert system is such challenging that is called the bottleneck of expert system implementation. The knowledge discovered by data mining can be utilized as a complementary knowledge source for expert system. Moreover, a data mining system which extracts knowledge needs expert guidance for successive decision making. This paper investigates cooperation between data mining and expert system to improve efficiency of correct speech E-learning systems.

INTRODUCTION

Expert knowledge and knowledge discovery are two powerful tools which could be combined. Utilizing them simultaneously provides higher quality. Expert system is executed on a knowledge database which includes knowledge extracted from expert. This knowledge base is demonstrated by a few forms such as rules, frames, Bayesian networks and so on. Afterwards, they are extracted by knowledge engineer and evaluated by expert. Obviously the system is subjective and it is limited by the amount of input knowledge shown in its knowledge base. The bottleneck of implementation of each expert system is knowledge extraction. This step is constrained by several limitations including the number of available experts or how the experts are professional in extracting actual knowledge. Recently, automatic knowledge discovery techniques have attracted considerable attention. Potentially it may help to overcome knowledge discovery bottleneck. Knowledge Discovery in Database (KDD), particularly data mining techniques, are exploited in automatic knowledge discovery. Data Mining Knowledge Discovery is implicitly inside the data and it can be in the form of patterns or models which are proper for data, timing data process, set of different data features and rules. The key point is that extracting knowledge from experts and knowledge discovery in database are complements. They might be applied simultaneously so that more efficient systems could be achieved. Data mining techniques might be utilized either to support various tasks involved in the Expert system or to develop Knowledge Base Systems (KBS). Experts’ knowledge could be exploited to improve results of each step of KDD process. This study aims to describe the key effects of interaction between ES and DML on efficiency of correct speech E-learning. First off, the existing literature consisting of expert systems, data mining and E-learning would be reviewed. Then, it is investigated how ES and DMK could be combined and how they affect the results. The cooperation of these two systems is bidirectional. In other words, both of them will modify the other one. The cooperation of these two systems might be employed in different cases; however, in our study the unidirectional cooperation of data mining with expert system is merely considered in an Electronic correct speech system.

Expert systems:

Expert systems are a specific group of softwares which try to help human experts or to be an alternative for them in limited professional fields. In fact they are primary samples of another technology called knowledge based systems. Expert systems have been utilized in diverse fields. These systems are able to perform different tasks rapidly and simply. These tasks include guidance, process, classification, consultant, design, detection, probing, prediction, creating concepts, justification, learning, managing, monitoring, planning, scheduling and test. Expert systems are employed as an alternative to expert person or as his assistant. The main difference
between expert system and other softwares is that they process the knowledge while other softwares process data and information (Darlington, 2000). The main issue of expert systems is decision making and problem solving. For this purpose, expert systems consist of two sections.

1. Information bank
2. Decision making engine

Information bank: it is referred to the mechanism of storing special information and regulations about a specific subject.

Decision making engine: it is a part of expert system which tries to find the results using the stored information.

To understand the expert system aforementioned sections might be divided into several components as follows.

User interface: it establishes a communication between user and the system. It enables the user to ask questions from the system. Furthermore, the system is capable of responding the user through this interface. This interface might be as simple as a simple menu for start and exit; whereas, it might be as complicated as conversation via natural language.

Knowledge base: it constitutes the main part of the system. It includes facts and rules associated with the field of expert system. It is usually stated in the form of “if-then” rules.

Control structure: it is also known as rule interpreter or inference engine. It is responsible for employing the existing knowledge for problem solving.

Short term memory: in addition to knowledge base which is considered as long term memory, short term memory is also needed. It is necessary to store different steps of solution finding and the path from question to answer (Mohsen Bayati, 1388) and (Abdoli et al 1387)

Data mining:

Information based society is defined as a society whose major part is involved in mental activities rather than physical ones. Such societies mostly focus on information activities including collecting, process, generation, registering, transmission and management of information and the large portion of budget is spent on information procedures. Development of base systems and large amount of stored data has necessitated tools so that the data could be processed and provide users with the information. Usually users introduce a hypothesis based on the observed reports; then, they reject or prove that. Nowadays, methods are needed which are able to discover the knowledge i.e. methods which state patterns and logical relations automatically with the minimum user interference. One of the most prominent methods by which useful patterns might be recognized inside data is data mining. This method provides users and analyzers with information so that they could make crucial and essential decisions based on provided information. It must be noticed that data mining is used when we are dealing with large amount of data (Gigabyte or Tera byte). In all data mining sources it is emphasized that the larger amount of data with more complicated relations the more complex extracting embedded information. In such cases the role of data mining as a knowledge discovery method would be more significant. While data mining products are powerful tools, they are not enough for applications. In order to succeed, data mining needs professional analyzers and expert people who are capable of analyzing and interpreting the output. Therefore, the limitations associated with data mining are mostly resulted from initial data or expert people rather than technology.

E-learning systems:

The current age is a combination of communication and information. Nowadays human kind needs information and communication to grasp necessity information more than previous decades. Today's people access facilities and information which enable them to communicate with each other. People are able to achieve required information wherever they are. Undoubtedly, information and communication technology have mostly affected educational environments. As a result of technology development, the educational environments tend to transform into virtual environments. Such changes necessitates modern facilities (such as computer devices and Internet) as well as rich information resources. Although education development is still a toddler, it has considerably affected global education. This effect on education all around the world will illustrate its demonstrative revolutionary face very soon. E-learning is called net-learning if it is performed by Internet and it is called virtual learning if it is done using information technology.

E-learning is the most prominent application of information technology presented in various forms including computer-based, online learning, offline learning, network-based learning and web-based learning. Nowadays, Internet have been an outstanding technology in educational, commercial and communication aspects owing to its benefits from various perspectives such as cost, time and space, high operation speed and high volume of information. In modern age electronic and technological revolution has changed our life drastically and necessitated remote leaning while making it richer and easier.

Definition of E-learning from different experts' point of view:
(Grasson, 2005) introduced E-learning for the first time. It referred to all educations which use Internet or intranet for education. (Cooper 2004) defined E-learning as a complex of educational activities performed by electronic tools including visual, vocal, computer-based, and network-based tools. (mayes, 2005) provides a conceptual definition of E-learning. Accordingly, it defines E-learning as active and smart learning which has revolved learning and educating procedure while playing a significant role in spreading and stabilizing information and communication technology culture.

(Howard Block): E-learning is being educated using Internet.

(elitttronden): E-learning employs network capabilities for people who are used to Internet so that they could learn simultaneously with working on the net.

Cisco systems officials believe that e-learning is an Internet-based learning which may include conveying the material in various forms, educational management, a networked set of students, a group of developers and a group of experts. Moreover, E-learning facilitates low cost and fast learning. It also provides education opportunity for all people.

**The effect of expert system on correct speech E-learning:**

The initial assumption in expert systems is that a minority of people (experts) are able to do some special tasks much better than rest of people (majority). Since the expertise of such people are unique, it would be beneficial to create expert systems which are able to take this expertise and provide it for all people so that decision making of non-expert people would improve. In case of E-learning this assumption will help people who are new in education. We are going to use database of an expert system to improve the performance of education. In fact, an electronic system which is able to solve problems of learners and teachers of correct speech in online manner. Firstly, we must know to what extent behavior of expert people could be simulated by computers. (Huichuan, 2011) reported two advantages for simulation. Since correct speech teachers in all fields (speech therapy, reading and foreign language) are not available all the time, they will be helped with education as well. These teachers are not experts in simulation software. Thus, this factor suggests that speech learners could be educated as they have more time. Most systems perform smart behaviors by extracting knowledge from a group of experts. Afterwards, an inference software is used to evaluate real situation reactions in created knowledge base (Huichuan, 2011). Knowledge based systems are designed so that they could act as an assistant for expert people. These smart assistants are designed based on expert systems technology. As a result their knowledge could be expanded in the future. Developing a smart assistant system is a prominent step in creating a complete expert system. Smart teacher is another application of expert systems. Despite previous versions of computer-based educational systems, the new ones are able to provide guidance or education dependent on field and concept (Ghazanfari and Kazemi13888). Most education assistance systems utilized expert systems based on Cases Base Reasoning. CBR is an inference method different from its counterparts. Instead of using general knowledge of a field, it establish connection between assumptions and results, in turn, it is able to achieve special knowledge from previous experiences. The achieved knowledge is applied instead of situations and conditions of main problem.

Knowledge based systems are designed so that they could be utilized as smart assistants for expert people. These smart assistants are designed based on expert systems technology. As a result their knowledge could be expanded in the future. Despite previous versions of computer-based educational systems, the new ones are able to provide guidance or education dependent on field and concept. These expert systems are employed in the following cases: human expertise is not available; expertise of several expert people is needed to be simulated; new expertise must be taught during the work; lack of capability to absorb and maintain expertise (Alireza Alidaei, 1380). Another characteristic distinguishing expert systems from other ones is their ability to learn. Some of them can learn without user's assistance through a process called induction rule. Numerous tools used for creating expert systems, create inference system which is able to create new rules in accordance with given samples (Darlington and Kase, 1380). Knowledge base of an expert system utilizes both factual and heuristic knowledge. Factual knowledge is a type of knowledge which could be shared and generalized in different fields because it is certainly correct. On the other hand heuristic knowledge is uncertain and based on personal interpretations. The better heuristic knowledge and guesses of the expert system, the higher its expert level i.e. it makes better decisions in special conditions. Heuristic structure based knowledge is essential in expert systems. It may accelerate problem solving procedure. The major problem associated with heuristic knowledge is that it cannot be used in all problems. In this project we need a heuristic expert system. Because stretching a letter during speech or pronouncing it with a particular accent should not misguide the system. To use expert systems we must integrate expertise of a speech therapist, a literature expert and a reader in one system. Solutions presented by a teacher for correct pronunciation of words and sentences as well as steps toward solving speech problem could be saved by involving E-learning system. The expert system uses the mentioned information to improve and accelerate the approach. Hence, it can help teachers during education.

**Cooperation between data mining and expert systems for improving efficiency of correct speech E-learning:**
Data mining is used for constructing expert system from extracted heuristic knowledge. It aims to optimize and maximize performance of KBS. The cooperation between knowledge discovery and expert system is designed to fulfill following goals. When expert's heuristic knowledge is basis of ES development and there is a consistent database including solved issues, DM is used for reinforcing ES in created database. An example of such cooperation might be found in (Kouk et al., 2000).

Expert's knowledge supports knowledge extraction using DM method and evaluation of this knowledge in a vast spectrum of types (behaviors), as explicated in the following. One of the problems associated with automatic knowledge extraction is validating and reviewing its proportionality. In this regard, Holmz and Kanigtham have described a sample of KBS construction where expert's knowledge can debug and validate a set of rules collected from DM. “Explora” is a statistical data mining program where knowledge extent and statistical measurement of database are applied to detect characteristic values and value composition occurring in the most or least number of databases. Knowledge discovery and construction is debugged and valued by expert in ES (Alonso & Martinez & perez&valente). Data mining procedure includes several steps and a knowledge excavator faces large amount of data and decisions. These decisions include model selection; whether it is simple Bayesian or neural network; discrete or continuous; if there is any sub samples for prune or not. Expert's knowledge could be exploited to guide users through these decision making processes. Conventional data mining is data based procedure including information such as existing knowledge, expert's experience, content and real limitations. The results could not be directly used for supporting business decision. Companies usually need interpretable, acceptable and valuable models which cannot be extracted without expert's knowledge. The way in which knowledge field and data mining methods are combined in knowledge discovery procedure is a prominent issue. Researchers who have focused on this subject (Huang, Zhang, Zhu, and Shi 2009) suggested several methods for a novel methodology called data mining unified with knowledge field. This methodology aims to practically discover knowledge resulted from composition of expert's knowledge and data mining knowledge. Expert's knowledge is used to check and debug the knowledge resulted from data mining. It results in constructing smart data mining systems including data mining knowledge and expert's knowledge. Such system is described by (Hu and Liu 2006). It consists of three layers: data mining system, expert system and expert user. Combining user's questions with expert's knowledge and generated rules, the smart data mining system generates an expert system. Then, it executes the expert system in an inference engine to answer user's questions. We investigate cooperation between DM and ES in correct speech E-learning system. The system is organized using knowledge in the field of correct speech. It is organized by composition of speech teacher's expertise in speech techniques and data mining techniques applied to existing data. Evaluating correct movement of speech organs is the main goal of speech teachers. The main goals are checking the effect of education, detecting speech disorders and evaluating effectiveness of speech therapy programs. At the beginning there is a huge amount of correct pronunciation data. A speech model is constructed using data mining. A common method is comparing and evaluating speech test of learner against a generated reference model. An import issue regarding this algorithm is selecting people for generating reference model. Distant points must be discarded and speech must be corrected from the beginning. It is done so that speech teachers could access a correct parameter. As mentioned before, the bottleneck of expert system is knowledge extraction. Here, data mining and knowledge discovery techniques are utilized to extract desired and useful patterns. Cooperation between expert system knowledge and KDD system (patterns and models) provides high level output which might be difficult for achieving use of one knowledge source. Knowledge obtained from these two sources is hardly achieved in case of using one knowledge source. For example, KBS is based on rule and KDD is based on expert functions for preparation and filtering. It must be considered that blind use of data mining without professional supervision of human results in wrong answers to wrong questions (Larose, 2004). Therefore, the process of using KDD should include experts in the whole process. They must be involved in real evaluation of input data mining. They play the most significant role in enriching generated results.

RESULTS AND DISCUSSION

Combing two systems; data mining and expert systems may increase quality of systems. Knowledge discovery which is the bottleneck of expert system design is simplified using data mining. In correct speech E-learning systems we need expert speech therapists, readers, literature experts and linguists. Extracting information from these experts and providing models for knowledge discovery in data mining systems considerably helps learners. It eliminates the need for extra cost, travels and therapist visits. Analyzing and storing teachers' behavior for correct speech education and finding expert solutions and suggestions for speech correction, the expert system would be able to provide the best feedback. Knowledge based expert systems operate as a smart assistant for expert people by collecting needed information in corresponding field. Using expertise gained by the system, teachers would be able to find the best solution for speech disorders. Furthermore, a comprehensive review on expert system, data mining, E-learning and their combination was
presented in this paper.

REFERENCES


Mayes, R., 2005; Modelling and Supporting Ict Implementation in Education; Oslo: Harvester.


Danubianu, M., Combining the Power of Data Mining with Expert System for Efficiently PersonalizeSpeech Therapy, Stefan cel Mare University of Suceava.