
ISSN 2306-1561

Automation and Control in Technical Systems (ACTS)

2014, No 1.2(9), pp. 93-102.

DOI: 10.12731/2306-1561-2014-1-23



Approach to the development of the rule base for the expert selection system of adaptive learning strategies

Popov Dmitriy Ivanovich

Russian Federation, Doctor of Technical Sciences, Professor, Header of «Computer Science, Engineering and Automation in the Media Industry».

Moscow State University of Printings Arts (MSUPA), 127550, Russian Federation, Moscow, Pryanishnikova Str., 2A. Tel.: +7 (499) 977-80-03, <http://www.mgup.ru>

popov@hi-edu.ru

Demidov Dmitriy Grigorievich

Russian Federation, Ph. D., Associate Professor, Department of «Computer Science, Engineering and Automation in the Media Industry».

Moscow State University of Printings Arts (MSUPA), 127550, Russian Federation, Moscow, Pryanishnikova Str., 2A. Tel.: +7 (499) 977-80-03, <http://www.mgup.ru>

demidov@hi-edu.ru

Abstract. This paper proposes an approach to the construction of the rule base for the selection of the optimal trajectory of training specialists of enterprises. Rules are formed on the basis of data mining Open Education Institute MSUPA, which contain information on the learning outcomes students using distance learning technologies for 10 years. Based on this analysis identified as the most successful and negative learning paths that can be used by other students in determining their learning strategies.

Keywords: expert systems, data mining, strategy training, e-learning.

ISSN 2306-1561

Автоматизация и управление в технических системах (АУТС)

2014. – №1.2(9). – С. 93-102.

DOI: 10.12731/2306-1561-2014-1-23



УДК 004.8

Подход к разработке базы правил для экспертной системы выбора адаптивной стратегии обучения

Попов Дмитрий Иванович

Российская Федерация, доктор технических наук, профессор, заведующий кафедрой «Информатика, вычислительная техника и автоматизация в медиаиндустрии».

ФГБОУ ВПО «Московский государственный университет печати имени Ивана Федорова (МГУП им. И. Федорова)», 127550, Российская Федерация, Москва, ул. Прянишникова, д. 2А. Тел.: +7 (499) 977-80-03, <http://www.mgup.ru>

popov@hi-edu.ru

Демидов Дмитрий Григорьевич

Российская Федерация, кандидат технических наук, доцент кафедры «Информатика, вычислительная техника и автоматизация в медиаиндустрии».

ФГБОУ ВПО «Московский государственный университет печати имени Ивана Федорова (МГУП им. И. Федорова)», 127550, Российская Федерация, Москва, ул. Прянишникова, д. 2А. Тел.: +7 (499) 977-80-03, <http://www.mgup.ru>

demidov@hi-edu.ru

Аннотация. В статье предложен подход к построению базы правил для выбора оптимальной траектории обучения специалистов предприятий. Правила формируются на основе интеллектуального анализа данных Института открытого образования МГУП имени Ивана Федорова, которые содержат информацию о результатах обучения слушателей с применением дистанционных образовательных технологий за 10 лет. На основании данного анализа выявляются как наиболее успешные, так и негативные траектории обучения, которые могут использоваться другими слушателями при определении их стратегий обучения.

Ключевые слова: экспертная системы, интеллектуальный анализ данных, стратегия обучения, электронное обучение.

1. Введение

В настоящее время сеть Интернет проникает практически во все области человеческой деятельности, открывая перед нами новые возможности, в том числе и

для получения образования. На данный момент перспективным и наиболее популярным способом профессионального обучения персонала на предприятиях является обучение «без отрыва от производства». Такой подход обуславливается недостатком финансовых и человеческих ресурсов, спецификой ведения бизнеса, отсутствием гибкого расписания занятий. При таких исходных данных наиболее предпочтительной формой обучения является электронное обучение, которое активно развивается повсеместно [1, 11 – 25].

При поверхностном рассмотрении, электронное обучение состоит из самостоятельного изучения материалов и прохождения компьютерного тестирования – основной формы контроля знаний в системах электронного образования [8 – 25]. Принцип электронного обучения в настоящее время заключается в том, что студент получает список необходимых для изучения дисциплин и необходимый образовательный контент. Последовательность же изучения предложенных дисциплин выбирается студентом самостоятельно (рисунок 1).

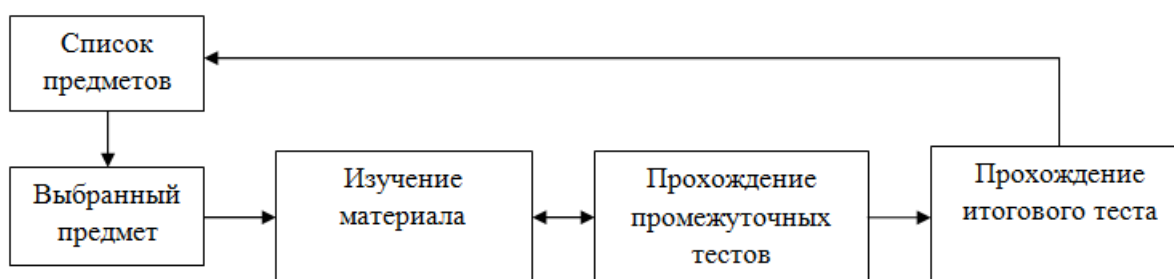


Рисунок 1 - Классический принцип электронного обучения

В связи с постоянно увеличивающимся потоком информации, развивающимися технологиями и углублением областей знаний, использование единого подхода к обучению для всех форм обучения и для всех студентов выливается в проблему постоянного ухудшения качества образования, и это в первую очередь связано с ухудшением усвоения информации студентом.

2. Описание экспертной системы

Одним из способов решения данной проблемы является предоставление студенту экспертной системы, позволяющей составлять наиболее удачную стратегию (последовательность) изучаемых дисциплин, используя заранее заложенную продукционную модель (рисунок 2).

Экспертная система – компьютерная система, способная частично заменить специалиста-эксперта в разрешении проблемной ситуации. Принцип функционирования экспертной системы состоит в том, что пользователь передает в экспертную систему факты или другую информацию и получает в качестве результата экспертные знания.

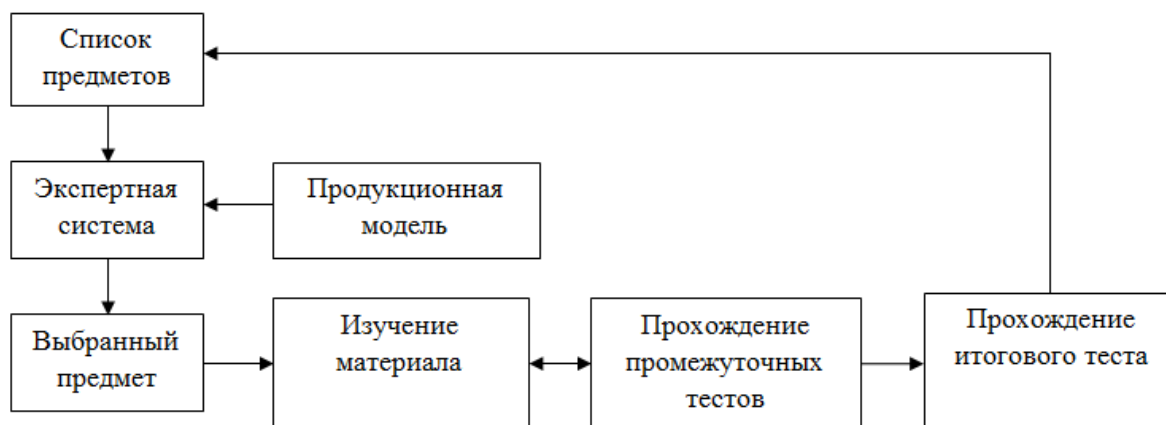


Рисунок 2 - Электронное обучение с использованием экспертной системы

По своей структуре (рисунок 3) экспертная система состоит из базы знаний и машины логического вывода. База знаний содержит правила, на основе которых машина логического вывода формирует заключения [4, 9, 10, 12, 13].

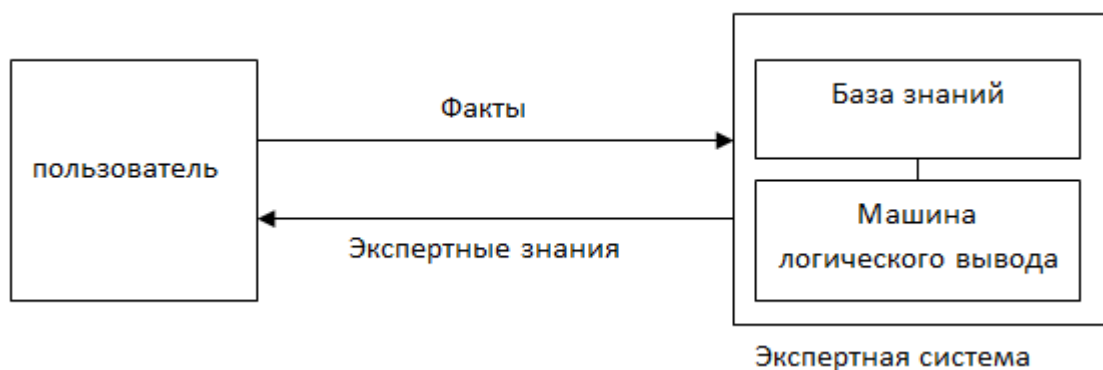


Рисунок 3 - Основные принципы функционирования экспертной системы

Главной задачей предложенной экспертной системы является предоставление студенту возможности ввода в качестве фактов списка дисциплин и иных параметров для их последующего анализа и выдачи заключения в виде рекомендаций по составлению стратегии обучения [7].

3. Разработка продукционной модели

В предложенной экспертной системе в качестве базы знаний выступает продукционная модель, которая основана на правилах, и которая позволяет представить знание в виде импликаций: «Если (условие), то (действие)» [2].

В случае с электронным обучением, источником получения продукционных правил является база данных сеансов тестирования Института открытого образования МГУП имени Ивана Федорова, насчитывающая более 650000 записей.

Для формирования продукционных правил, прежде всего, необходимо провести интеллектуальный анализ данных с использованием Data Mining [3] для получения знаний в удобном для понимания и классификации виде. Общие принципы анализа

представлены на рисунке 4 и заключаются в выделении целевых данных (учитываемых в последующем анализе), их предварительной обработке (очистка от «шума», приведение в удобную для анализа форму), самого анализа и выдачи результатов.

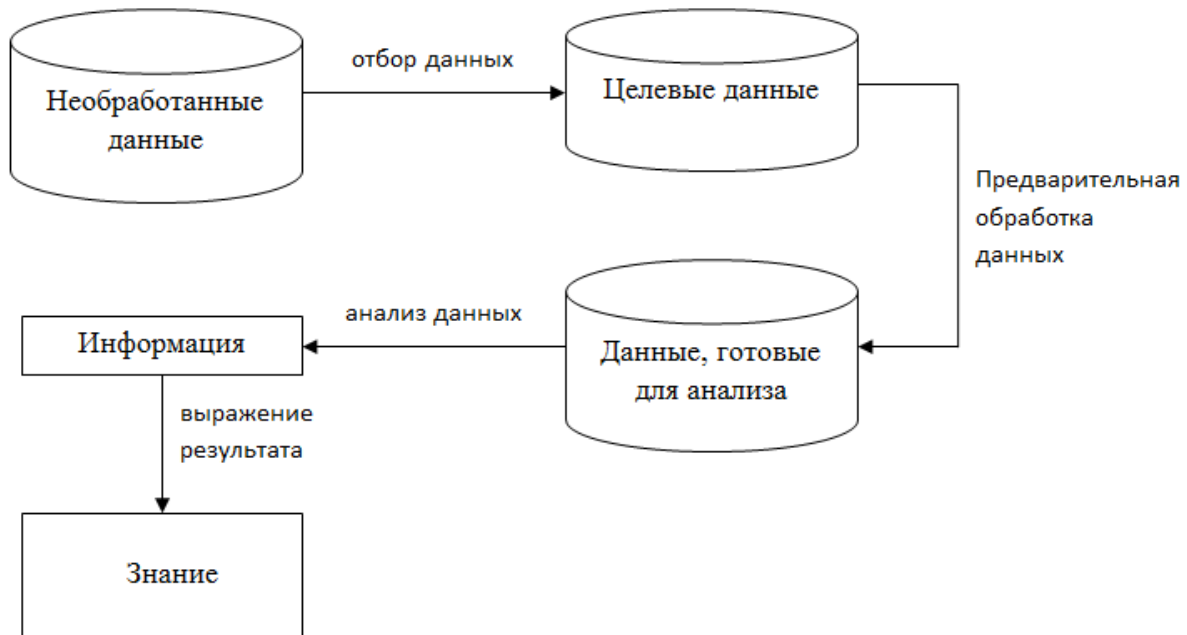


Рисунок 4 - Принцип интеллектуального анализа данных

Интеллектуальный анализ данных способен эффективно давать ответы на вопросы, используя методы классификации, кластеризации, взаимосвязей, регрессии, прогнозирования, анализа последовательностей, анализа отклонений [3]. В общем виде, задачи, решаемые Data Mining, можно разделить на описательные и предсказательные. В описательных задачах самое главное – дать наглядное описание имеющихся скрытых закономерностей, в то время как в предсказательных задачах на первом плане стоит вопрос о предсказании для тех случаев, для которых данных ещё нет [5]. Составление продукционных правил относится к описательной задаче.

Для решения задач описательного типа более эффективно использование методов кластеризации, классификации, взаимосвязей, отклонения и, в некоторых случаях, анализа последовательности [6]. Таким образом, использование алгоритмов анализа данных, типичных для решения описанных методов, приведет к получению явных знаний, которые смогут быть внесены в продукционную модель экспертной системы в виде продукционных правил.

Интеллектуальный анализ данных проводился с использованием СУБД Microsoft SQL Server 2008 R2. После предварительной обработки база данных, содержащая результаты тестирования студентов за 2004-2013 гг., насчитывала 474000 записей. Кроме результатов тестирования в базу данных была внесена информация о специальностях и направлениях подготовки, дисциплинах, группах и студентах. Процесс интеллектуального анализа данных проводился в среде Microsoft Business Intelligence Development Studio (BI Studio), специализированной для работы с SQL

Server. Анализ проводился в разных направлениях с использованием упрощенного алгоритма Байеса, алгоритма дерева принятия решений, кластеризации, кластеризации последовательностей и нейронной сети. Загруженная в BI Studio база данных представлена на рисунке 5.

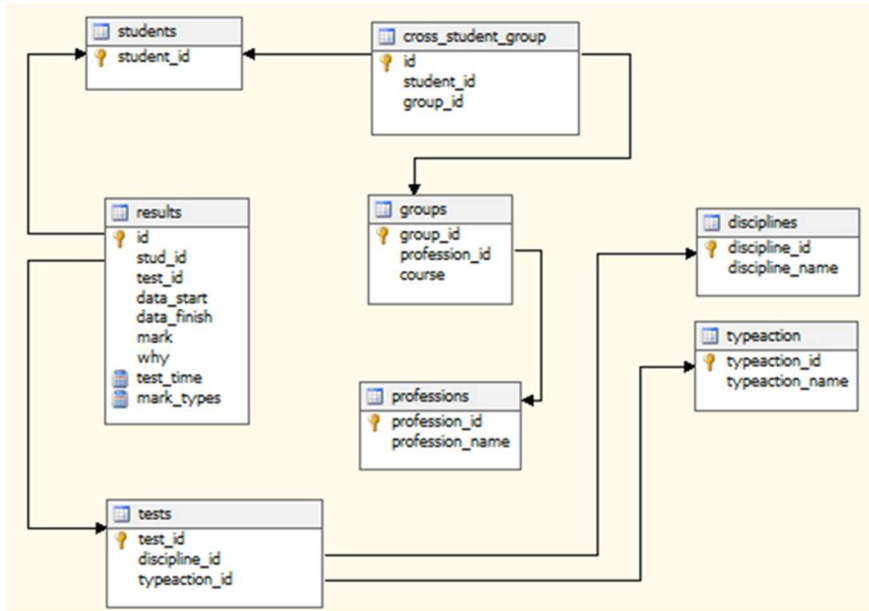


Рисунок 5 - Структура базы данных

4. Ограничения

Однако она не может быть использована для проведения анализа в представленном виде, т.к. существуют ограничения в реализации алгоритмов Data Mining. Кроме этого, существует проблема отображения и восприятия полученных результатов. Для решения указанных проблем использовались механизмы создания представлений BI Studio, которые позволяют создавать на основе исходных таблиц новые, а также добавлять атрибуты в существующие сущности, используя простые SQL-запросы. При этом исходная база данных не изменяется. Так, в BI Studio были внесены следующие изменения:

- создано представление, которое включает в себя полный список студентов с указанием группы, специальности и курса;
- объединены таблицы результатов тестирования, дисциплин и типов тестирования (итоговое, промежуточное), добавлен атрибут времени тестирования в минутах;
- в отдельное представление выделены результаты итоговых тестирований;
- создана таблица с периодами изучения дисциплин студентами в днях и содержащая множество дополнительных атрибутов, например, количество сеансов тестирования по дисциплине за указанный период, средняя оценка и т.д.

5. Результаты

Проведя анализ базы данных системы тестирования, были получены следующие результаты.

Ранжирование дисциплин (рисунок 6) по полученным оценкам, которое позволяет определить вероятность получения заданной оценки по всем дисциплинам. В анализе проводился с использованием алгоритма кластеризации, как на всех сеансах тестирования, так и исключительно на сеансах итогового тестирования.

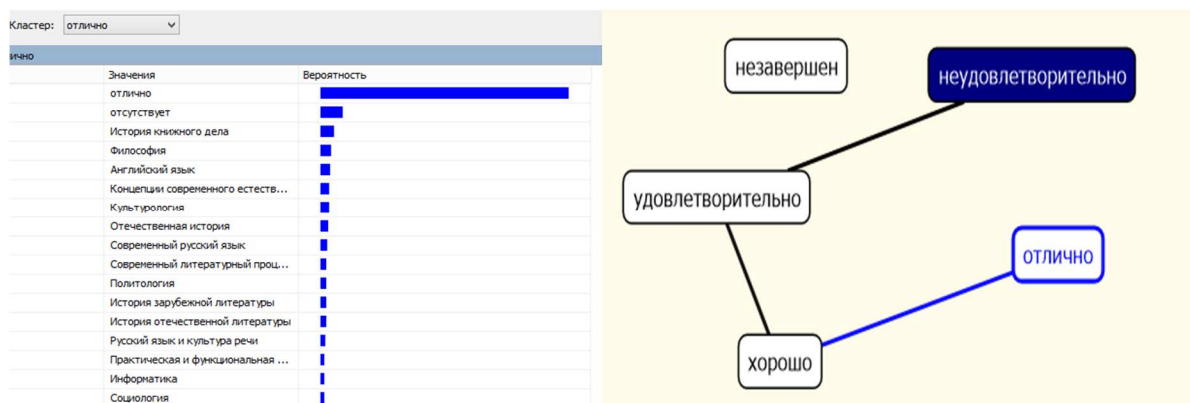


Рисунок 6 - Ранжирование дисциплин

Оценка затраченного на прохождение теста времени (рисунок 7). Данный результат позволяет оценить время, затраченное студентами на изучение конкретной дисциплины и выделить на основе указанной оценки кластеры. Анализ проводился, как на всех сеансах тестирования, так и исключительно на сеансах итогового тестирования.

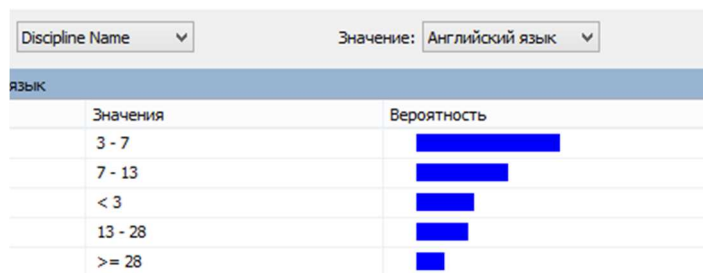


Рисунок 7 - Затраченное на прохождение теста время

Усредненная оценка времени, количества сеансов тестирования и полученных оценок по каждой дисциплине (рисунок 8). Данный результат позволяет оценить трудность дисциплины в целом.

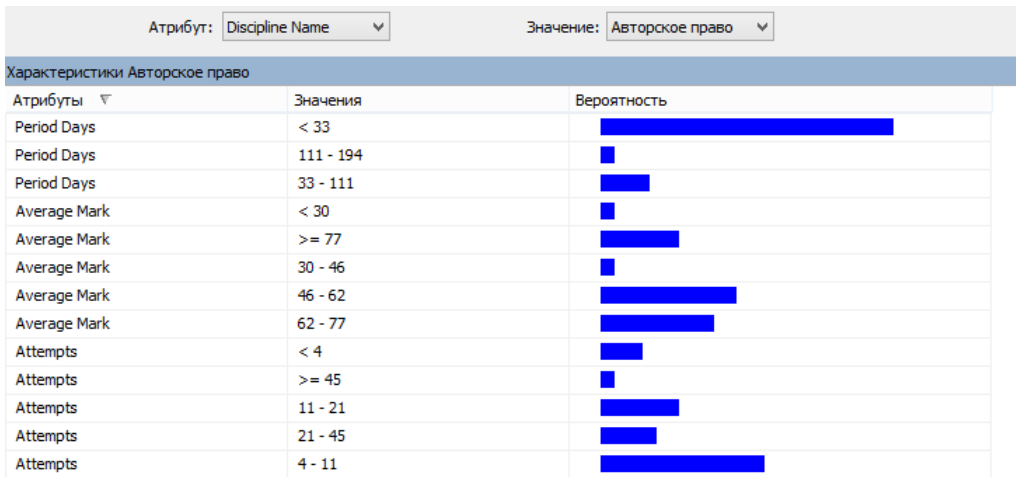


Рисунок 8 - Оценка трудности дисциплины

Зависимость оценки от среднего количества попыток (рис. 9), которая позволяет выявить случаи раскречевания банка тестовых заданий.

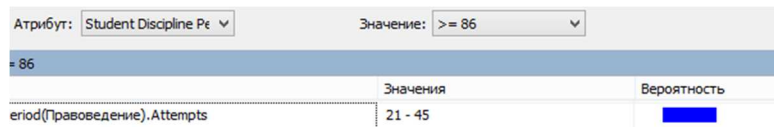


Рисунок 9 - Зависимость оценки от среднего количества попыток

Таким образом, для разработки продукционных правил предлагается использовать существующие последовательности изучения дисциплин, сформированные студентами самостоятельно и результаты их успеваемости. Выявляя схожесть последовательностей изучения дисциплин при относительно стабильной успеваемости студентов можно сформировать базу правил для экспертной системы, на вход которой будет подаваться список дисциплин. На выходе же будут формироваться рекомендации по выбору стратегии обучения. Используя алгоритм кластеризации, удалось выделить стратегии обучения, уже использованные студентами, которые представлены на рисунке 10.

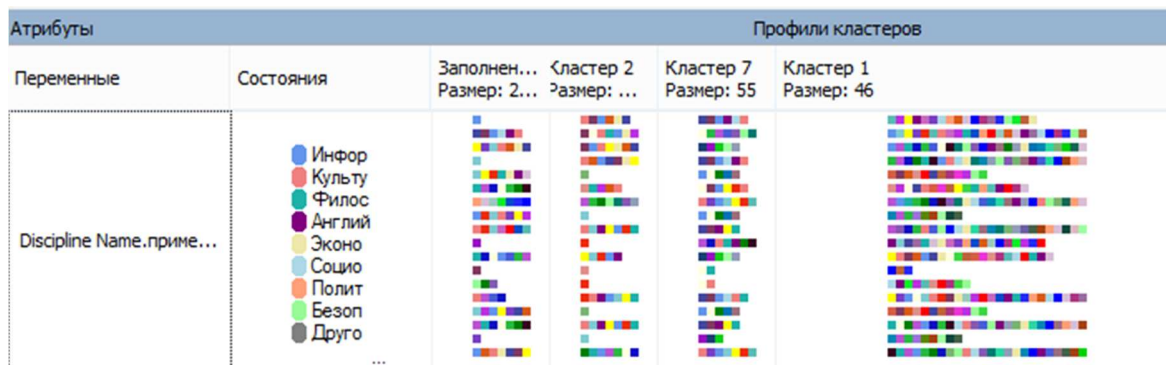


Рисунок 10 - Стратегии обучения

В результате, проведенная модификация базы данных, позволила более гибко использовать исходные данные, обойти ограниченность реализации алгоритмов Data Mining и решить проблему интерпретации результатов анализа.

6. Заключение

Проведенный анализ показал, что последовательность изучения дисциплин учебного плана напрямую связана с результативностью обучения слушателей. Поэтому экспертная система, созданная на основании полученных правил будет являться незаменимым инструментом тьютора и поможет автоматизировать процесс поддержки электронного образования.

Список информационных источников

- [1] Демидов Д. Г. Разработка моделей и алгоритмов автоматизации процессов адаптивного обучения специалистов для предприятия / Д.Г. Демидов // Диссертация. – М.: 2011. – 142 с.
- [2] Гаврилова Т. А. Базы знаний интеллектуальных систем. Учебник // Т.А. Гаврилова, В.Ф. Хорошевский // СПб.: Питер, 2000, 384 с.
- [3] Макленнен Д. Microsoft SQL Server 2008: Data Mining. Интеллектуальный анализ данных / Д. Макленнен, Ч. Танг, Б. Криват // БХВ-Петербург, 2009, 700 с.
- [4] Джерратано Д. Экспертные системы. Принцип разработки и программирования. Учебник / Д. Джерратано, Г. Райли // Издательский дом «Вильямс», 2007, 1152 с.
- [5] Свободная энциклопедия «Википедия» - http://ru.wikipedia.org/wiki/Data_mining
- [6] Попов Д.И. Проектирование интеллектуальных систем дистанционного образования //Д.И.Попов/ Известия Южного федерального университета. Технические науки – 2001. – Т. 22. № 4. – С. 325-332.
- [7] Попов Д.И. Адаптивная стратегия обучения персонала предприятий // Д.И.Попов, Д.Г.Демидов / В мире научных открытий. – 2011. –№ 9. – С. 65-71.
- [8] Попов Д.И. Выбор алгоритмов оценки результатов тестирования // Д.И.Попов, Д.А.Арсентьев / В мире научных открытий. – 2011. – № 9. – С. 60-65.
- [9] Остроух А.В. Основы построения систем искусственного интеллекта для промышленных и строительных предприятий: монография / А.В. Остроух. – М.: ООО «Техполиграфцентр», 2008. - 280 с. - ISBN 978-5-94385-033-2.
- [10] Остроух А.В. Системы искусственного интеллекта в промышленности, робототехнике и транспортном комплексе: монография / А.В. Остроух - Красноярск: Научно-инновационный центр, 2013. – 326 с. - ISBN 978-5-906314-10-9.
- [11] Исмоилов М.И. Подготовка и переподготовка персонала предприятий промышленного и транспортного комплексов с применением мобильных технологий: монография / М.И. Исмоилов, А.Б. Николаев, А.В. Остроух. – Saint-Louis, MO, USA: Publishing House Science and Innovation Center, 2013. – 166 с. – ISBN 978-0-615-67111-6.
- [12] Остроух А.В. Информационные технологии в научной и производственной деятельности / [ред. А.В. Остроух] - М: ООО "Техполиграфцентр", 2011. – 240 с. - ISBN 978-5-94385-056-1.

- [13] Остроух А.В. Ввод и обработка цифровой информации: учебник для нач. проф. образования / А.В. Остроух. – М.: Издательский центр «Академия», 2012. – 288 с. – ISBN 978-5-7695-9457-1.
- [14] Краснянский М.Н. Интеграция виртуальных тренажеров в процесс обучения операторов технических систем с использованием Интернет-технологий / А.В. Остроух, М.Н. Краснянский, С.В. Карпушкин, Д.Л. Дедов // Дистанционное и виртуальное обучение. - 2010. - №7. - С. 66-70.
- [15] Чурин В.В. Использование компьютерных тренажеров для подготовки рабочих дорожно-строительных профессий / В.В. Чурин, А.В. Остроух, А.А. Подберёзкин // Молодой ученый. - 2011. - №4. Т.3. - С. 28-29.
- [16] Остроух А.В. Опыт разработки электронных образовательных ресурсов нового поколения для дистанционной технологии обучения / А.В. Остроух // В мире научных открытий. - 2011. - №9 (21). - С.149-158.
- [17] Краснянский М.Н. Постановка задачи проектирования виртуальных тренажерных комплексов для обучения операторов технических систем /А.В. Остроух, М.Н. Краснянский, С.В. Карпушкин, Д.Л. Дедов // Информационные технологии в проектировании и производстве. - 2012. - №2. - С. 46-50.
- [18] Краснянский М.Н. Алгоритм проектирования виртуальных тренажерных комплексов для обучения операторов технических систем /А.В. Остроух, М.Н. Краснянский, С.В. Карпушкин, Д.Л. Дедов // Информационные технологии в проектировании и производстве. - 2012. - №3. - С. 68-75.
- [19] Остроух А.В., Петриков П.А. Использование дистанционных образовательных технологий для подготовки и переподготовки персонала промышленных предприятий // Автоматизация и управление в технических системах. – 2012. – № 1. – С. 94-100.
- [20] Остроух А.В. Разработка автоматизированной системы дистанционного обучения на основе программного обеспечения с открытым кодом / А.В. Остроух, М.Н. Краснянский, П.А. Петриков, Н.В. Михайлова // В мире научных открытий. Серия «Проблемы науки и образования». – 2012. – №12. – С.76-86.
- [21] Krasnynskiy M., Nikolaev A., Ostroukh A. Application of virtual simulators for training students in the field of chemical engineering and professional improvement of petrochemical enterprises personnel. International Journal of Advanced Studies. 2012. Vol. 2. No 3. pp. 4. DOI: 10.12731/2227-930X-2012-3-4.
- [22] Barinov K., Krasnynskiy M., Malamut A., Ostroukh A. Algorithm of Virtual Training Complex Designing for Personnel Retraining on Petrochemical Enterprise. International Journal of Advanced Studies. 2012. Vol. 2. No 3. pp. 6. DOI: 10.12731/2227-930X-2012-3-6.
- [23] Ostroukh A.V., Barinov K.A., Nikolaev A.B., Stroganov V.Yu. Interactive Game Modeling Concept for Personnel Training at the Industrial Enterprises // World Applied Sciences Journal (WASJ). 2013. Vol. 28, No 1. pp. 44-55. DOI: 10.5829/idosi.wasj.2013.28.01.1876.
- [24] Ostroukh A.V., Barinov K.A., Nikolaev A.B., Stroganov V.Yu. Formal methods for the synthesis of the organizational structure of the management through the personnel recruitment at the industrial enterprises // Journal of Applied Sciences (JAS). 2014. Vol. 14, No 5. pp. 474-481. DOI: 10.3923/jas.2014.474.481.
- [25] Ostroukh A.V., Pomazanov A.V. Development of information storage and retrieval system on nanomaterials and nanotechnology // International Journal of Advanced Studies (iJAS). 2013. Vol. 3, Issue 3, pp. 24-30. DOI: 10.12731/2227-930X-2013-3-5.