

ЭКОНОМИЧЕСКИЕ И СОЦИАЛЬНО-ГУМАНИТАРНЫЕ НАУКИ

МРНТИ 20.53.19

УДК 004.94

МЕТОДЫ И МОДЕЛИ ТЕХНОЛОГИИ МОНИТОРИНГА НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ ВУЗА

И. УВАЛИЕВА¹, Л. СУЛЕЙМЕНОВА¹, Ж. САГАНДЫКОВА²

¹*D. Serikbaev East Kazakhstan state technical university*

²*Regional specialized IT-School-Lyceum*

Аннотация: Для развития образования уже недостаточно влияния человеческого капитала. Возникает потребность изменения самой образовательной среды, т.е. не просто наращивать объёмы образования трудовых ресурсов, должны качественно измениться методы, инструменты и среды управления бизнес-процессами образования, необходимо использование инструментов самодиагностики и механизмы предсказания состояния бизнес-процессов, подключения «интеллектуальных» сервисов к главным системам образовательных учреждений.

В данной работе описана технология мониторинга научно-образовательной деятельности вуза, объединяющая методы и модели управления знаниями в вузе, алгоритм Staking-ансамблирования методов интеллектуального анализа данных образовательных бизнес-процессов, а также модель повышения эффективности управления деятельностью научно-педагогического персонала вуза.

Ключевые слова: обработка данных, smart технологии, информационные технологии в образовании, научные знания, интеллектуальный анализ данных

METHODS AND MODELS OF TECHNOLOGY FOR MONITORING RESEARCH AND EDUCATION IN UNIVERSITY

Abstract: The influence of human capital is no longer sufficient for the development of education. There is a need to change the educational environment and to change methods, tools and business management, to use self-diagnosis to predict business processes and to connect “intelligent” services to the main systems of educational institutions. This paper describes a technology for monitoring research and education in the University. It is achieved by combining methods and models of knowledge management at the University, algorithm of Staking-assembling methods of intellectual analysis of educational business processes data, as well as a model of effective personnel management.

Keywords: data processing, smart technologies, information technologies in education, scientific knowledge, data mining

ЖОҒАРЫ ОҚУ ОРНЫНЫҢ ҒЫЛЫМИ-БІЛІМ БЕРУ ҚЫЗМЕТІН МОНИТОРИНГТЕУ ТЕХНОЛОГИЯСЫНЫҢ ӘДІСТЕРІ МЕН МОДЕЛІ

Аңдатпа: Қазір білім беруді дамыту үшін адами капиталдың ықпалы жеткіліксіз. Білім беру ортасының өзін өзгерту қажеттілігі туындап отыр, яғни еңбек ресурстарының білім көлемін арттыру ғана емес, білім берудің бизнес-үдерістерін басқару әдістері, құралдары мен ортасы сапалы өзгеруі

тиіс, бизнес-үдерістердің жай-күйін болжау тетіктері мен өзін-өзі диагностикалау құралдарын пайдалану, білім беру мекемелерінің басты жүйелеріне “зияткерлік” сервистерді қосу қажет. Бұл мақалада ЖОО білімді басқару әдістері мен модельдерін, білім беру бизнес-үдерістерінің, деректердің, талдаудың зияткерлік әдістерінің Staking-ансамбльдеу алгоритмі біріктіретін ЖОО-ның ғылыми-білім беру қызметін мониторингтеудің технологиясын әзірлеу және енгізу, сонымен қатар ЖОО-ның ғылыми-педагогикалық құрамының қызметін басқарудың тиімділігін арттыру моделі ұсынылған.

Түйінді сөздер: деректерді өңдеу, smart технологиялар, білім берудегі ақпараттық технологиялар, ғылыми білім, деректерді интеллектуалды талдау

Введение

В настоящее время умные или Smart-технологии меняют нашу жизнь и общество на более новый качественный уровень, чем когда-либо прежде. Решения на основе смарт-технологий работают с использованием различных интеллектуальных информационных технологий. В связи с этим сегодня вузы согласно концепции государственной программы “Цифровой Казахстан”, стратегическому плану развития Республики Казахстан до 2025 г. требуют внедрения современных интеллектуальных технологических решений.

Прикладное решение проекта связано с исследованием по следующим направлениям: автоматизированная подсистема составления оптимального расписания учебных занятий в ВУЗе при организации учебного процесса; подсистема управления научными знаниями вуза, подсистема комплексного интеллектуального анализа данных научно-образовательной деятельности, а также подсистема управления деятельностью научно-педагогического персонала вуза.

В связи с совершенствованием системы управления ВУЗами на основе вычислительных и информационных систем все более актуальными становятся следующие вопросы: автоматизация составления оптимального расписания учебных занятий в вузе при организации учебного процесса; исследование проблем управления знаниями и разработка системы управления научными знаниями вуза.

Научная новизна исследования заключается в разработке концепции и вычислительной ИТ-инфраструктуры на основе технологии подсистем системы управления ВУЗами,

объединяющая «гибкие» методы и алгоритмы, гибридные алгоритмы, а также технологический подход к управлению научными знаниями вуза.

Анализ зарубежных и отечественных разработок в области научной деятельности вузов показал, что к настоящему времени отсутствует универсальная онтология научных знаний вуза, которую можно было бы положить в основу проектирования семантического портала научных знаний вуза [1]. Также не существует единого подхода к построению решений (семантических порталов), реализованных на базе технологий Semantic Web.

Для учебного заведения очень важно получать финансирование, которое обеспечит привлечение высококвалифицированных сотрудников и позволит создать развитую инфраструктуру вуза. Для этого требуется иметь высокие показатели эффективности деятельности вуза, которые обеспечиваются с большим трудом в конкурентной борьбе. В этом важную роль оказывает ППС, а механизмом управления является система материального стимулирования работников высшей школы. Традиционные подходы в системе стимулирования труда работников вуза не обеспечивают высокие достижения в профессиональной деятельности сотрудников и вуза в целом. Для реального повышения результативности их деятельности необходим эффективный инструментальный распределения стимулирующих выплат на основе комплексного учета всех показателей, отвечающий требованиям современных рыночных отношений [2].

Методологическую основу исследований составляют: теория управления знаниями; положения и методы теории принятия решений, системного анализа, онтологического инжиниринга; методы машинного обучения; методы модульного, структурного и объектно-ориентированного программирования; теория проектирования информационных систем.

1. Онтологическая модель научных знаний вуза

Онтологическая модель научных знаний вуза, включающая в себя основные классы и отношения объектов научной деятельности, обеспечивающие формирование научного профиля и мониторинг развития научных школ вуза [3]. Для разработки данной модели необходимо реализовать процедуру обработки информационных ресурсов вуза, которая является одним из инструментов управления научными знаниями вуза и включает этапы, представленные на рисунке 1.

Этап 1. Реализация методики извлечения терминологических коллокаций научных тек-

стов на основе статистических методов и морфологических шаблонов фильтров для выявления ключевых терминов по определённой предметной области. Для этого разработан и описан алгоритм извлечения терминологических коллокаций предметной области. В качестве метода для выявления коллокаций используется критерий χ^2 [2].

Этап 2. Разработан алгоритм выбора признаков для сокращения размерности пространства признаков и эффективности обучения классификаторов. В качестве метода определения признаков выбран метод взаимной информации (Mutual Information) [4].

Этап 3. Предложен метод формирования научного профиля на основе классификаций информационных ресурсов по научным направлениям. Классификация текстов по научным направлениям осуществляется методом kNN [3].

Этап 4. В результате обработки научных ресурсов вуза формируются профили документов. Профиль документа определяется как вектор всех его релевантных классов. В результате формируется профиль докумен-



Рис. 1 – Этапы процедуры обработки информационных ресурсов вуза

та посредством создания индивидов класса «Информационные ресурсы» онтологии научных знаний [3].

Этап 5. Следующим этапом в интеграции и обеспечении совместного доступа к научным знаниям вуза является разработка семантического портала научных знаний на основе онтологической модели и методов интеллектуальной обработки научных ресурсов вуза.

В качестве исходных данных выступают показатели научной деятельности ППС ВКГТУ им. Д.Серикбаева. Модель научной деятельности отдельного научного сотрудника определяется с помощью факторов научной активности. Эти факторы в онтологической информационной модели сгруппированы в следующие классы: события, проекты, публикации. Управление научными знаниями как процесс идентификации знаний и умений позволяет решать задачу подбора научных экспертов. Онтологическая модель может быть использована при подборе научных экспертов для оценки научных проектов, а также для подбора рецензентов научных работ.

2. Гибридный алгоритм

Stacking-ансамблирования методов интеллектуальной обработки данных научно-образовательного процесса

Гибридный алгоритм Stacking-ансамблирования методов интеллектуальной обработки данных научно-образовательного процесса используется для решения аналитических задач процесса и получения более точных результатов оценки бизнес-процессов научно-образовательной деятельности (как, например, выявление проблемных групп обучающихся, прогнозирование результатов сессии, определение причин академической задолженности, и т.д.) [5]. Алгоритм объединяет усилия нескольких интеллектуальных методов на основе создания ансамбля моделей-предикторов. Для построения ансамбля моделей на данных образовательной статистики был использован алгоритм Stacking, который агрегировал следующие 3 типа интеллектуальных моделей: алгоритм Байеса (БК), алгоритм дерева решений (ДР) и ней-

ронная сеть (НС). Обучение производилось на единственном наборе данных. Данные метауровня, используемые для обучения мета-модели, будут представлять собой результаты предсказания моделей для каждого из полей [6]. Для возможности использования в качестве мета-модели БК, включим в данные метауровня поля класса числом для БК, ДР и НС. Таким образом, алгоритм stacking-ансамблирования пытается обучить каждый классификатор, используя алгоритм метаобучения, который позволяет обнаружить лучшую комбинацию выходов базовых моделей.

Наиболее общим способом уменьшить перекосы, вызванные неудачным сэмплингом, является повторение всего процесса обучения и тестирования несколько раз при различных случайных выборках. На каждой итерации определенная доля данных (например, 2/3) случайно отбирается для обучения, а оставшиеся используются для тестирования. Для этого необходимо определить число итераций разбиения исходных данных. Далее задается вектор, в который собираются данные для обучения мета-модели.

Для реализации экспериментального исследования проверки алгоритма Staking-ансамблирования будут проанализированы показатели научно-образовательной деятельности для прогнозирования набора переводного балла GPA и трудоустройства выпускника вуза. Вычислительный эксперимент будет проведен на данных образовательного портала ВКГТУ им. Д.Серикбаева за 2009-2020 гг. по всем уровням и различным формам обучения (это более 1 млн. записей).

3. Модель повышения эффективности управления деятельностью научно-педагогического персонала вуза

Модель повышения эффективности управления деятельностью научно-педагогического персонала вуза основана на системе показателей эффективности деятельности научно-педагогического персонала, дополненная показателями деятельности, обеспечивающей дополнительное внебюджетное финансирование, и показателями деятельности,

обеспечивающей развитие вуза в перспективе [7]. Данная модель является гибридной и динамически изменяемой, основана на весовых коэффициентах показателей иерархической структуры и системе ограничений, обеспечивающей баланс между требованиями мониторинга эффективности вуза, целями и ресурсами. Реализация данной модели включает в себя следующие этапы.

Этап 1. Расчет диапазонов суммарных значений по каждому показателю.

Этап 2. Перерасчет плановых и рекомендуемых показателей на одного преподавателя, в зависимости от его типа.

Этап 3. Деятельность преподавателей по выполнению показателей.

Этап 4. Мониторинг выполнения показателей и информирование преподавателей об отклонениях от запланированных суммарных значений выполнения показателей с целью побуждения их к самостоятельной корректировке своих действий.

Этап 5. Оплата деятельности преподавателей за прошедший период.

Этап 6. Перерасчет размеров оплаты за единицу выполненных показателей, значения выполнения которых по результатам мониторинга отклоняются от рекомендованных.

Этап 7. Переход к шагу 3.

Этап 8. В случае экономии премиального фонда после завершения учебного года, расчет и выплата дополнительных разовых премий.

Достоверность проведенных исследований и полученных результатов будет обеспечена аргументацией выдвигаемых положений, обоснованным и корректным применением математического аппарата, результатами вычислительного эксперимента и результатами использования разработок.

Реализация ИТ-сервисов Smart-технологий должна быть реализована в виде динамической библиотеки. Реализация различ-

ных ИТ-сервисов в виде отдельной динамической библиотеки позволит использовать данную библиотеку в различных программных продуктах, а также впоследствии производить расширение её возможностей. Для реализации библиотеки будет использована среда разработки MSVisualStudio 2017, которая позволяет разрабатывать ПО на базе MS.NetFramework. В качестве языка программирования будет использован язык C#.

5. Программная реализация

Для построения метамодели будем использовать алгоритмы БК, ДР и НС. На рисунке 3 представлено окно подсистемы анализа данных образовательной статистики она основе алгоритма Staking-ансамблирования.

Использование нескольких моделей дает следующие преимущества:

- реализовывать предсказание выходного показателя образовательной статистики как в численном виде (НС) так и в виде классификации (ДР и БК), что определяет разнообразие возможных методов визуализации результатов;

- согласованность результатов по всем моделям позволяет сделать вывод об их достоверности [8].

Далее представлены результаты алгоритма Staking-ансамблирования, реализованного в подсистеме анализа данных. Обучение на основе базовых классификаторов для анализа данных первой задачи образовательного мониторинга показало следующие результаты. Окно программы с результатами обучения базового классификатора на основе алгоритма БК содержит следующие вкладки отображения взаимодействия между прогнозируемыми атрибутами и входными атрибутами для таблицы вариантов: сеть зависимостей; профили атрибутов; характеристики атрибута; сравнение атрибутов.

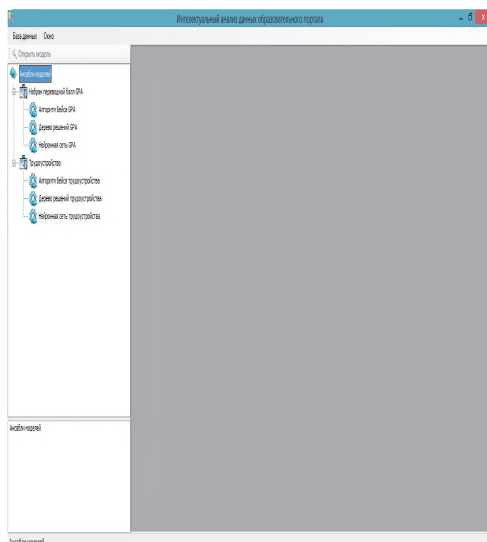


Рис. 3 – Главное меню подсистемы Staking-ансамблирования

На основе данных образовательной статистики были получены результаты обучения базового классификатора на основе алгоритма БК для анализа данных набора GPA.

Вкладка сеть зависимостей окна подсистемы системы отображает зависимости между входными показателями образовательной статистики вуза и прогнозируемым показателем, характеризующим набран ли GPA. Ползунок слева от средства просмотра выступает в качестве фильтра, привязанного к прочности зависимостей. При движении ползунка вниз отображаются только самые прочные из них. Условные обозначения в нижней части окна подсистемы связывают цветовые коды с типом зависимости на графе. Как видно из рисунка 4 на показатель «Набран GPA» влияют такие показатели образовательной статистики как число неосвоенных дисциплин за предыдущий курс обучения, общий балл GPA за предыдущий курс обучения.

Основой второго базового классификатора послужил алгоритм дерева принятия решений, результаты которого представлены на рисунке 4. Данный алгоритм поддерживает классификацию и регрессию, используется для прогнозирующего моделирования дискретных и непрерывных атрибутов.

Результаты обучения базового классификатора на основе алгоритма ДР показал,

что набор GPA имеет зависимость со следующими показателями образовательной статистики вуза: число неосвоенных дисциплин за предыдущий курс, общий балл GPA за предыдущий курс обучения, специальность и балл ЕНТ.

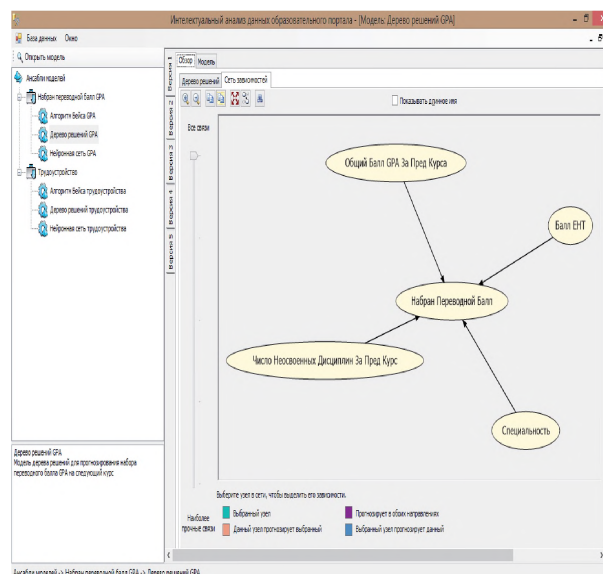
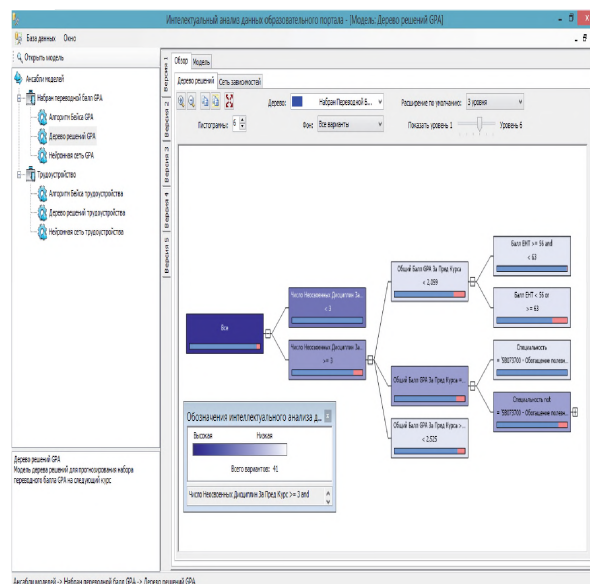


Рис. 4 – Окно с результатами обучения базового классификатора на основе алгоритма ДР для анализа данных набора GPA

Мета-классификатор Staking-ансамблирования был построен на основе алгоритма БК. Результаты точности прогноза реализации Staking-ансамблирования на основе матрицы классификации представлены на рис. 5.

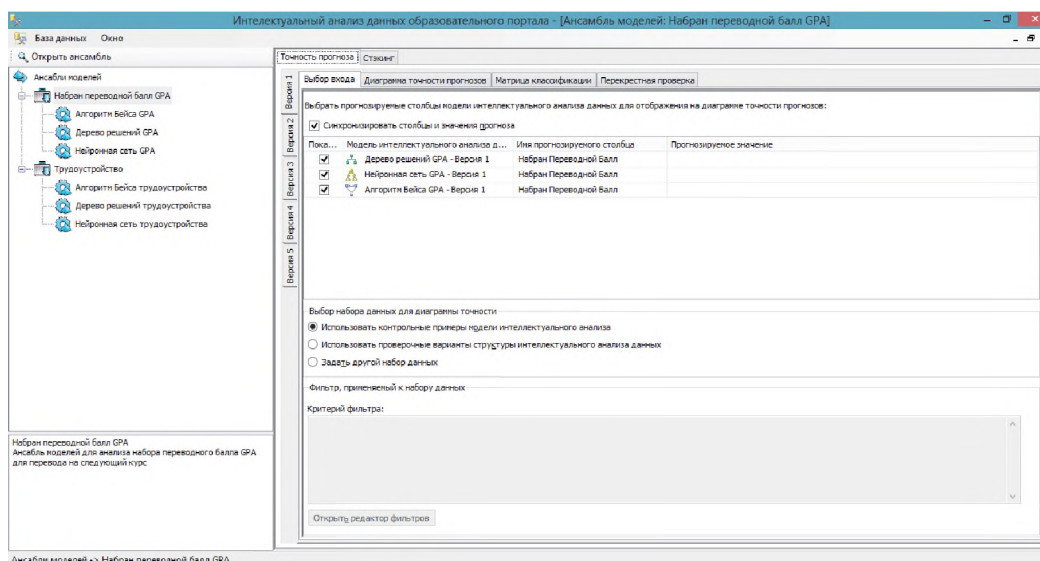


Рис. 5 – Окно с параметрами оценки точности реализации алгоритма Staking-ансамблирования

Результаты обучения мета-классификатора на вариационной выборке показали, что показатель образовательной статистики «Набран GPA» имеет зависимость с такими показателями как число неосвоенных дисциплин за предыдущий курс, общий балл GPA за предыдущий курс обучения, Базовое образование и Курс.

Заключение

Таким образом, данное научное исследование внесет значимый вклад по следующим направлениям: повышение эффективности реализации государственной программы

“Цифровой Казахстан” и стратегического плана развития Республики Казахстан до 2025 г.; развитие системы управления вузом на основе современных научных и технических решений, реализованные в Smart-технологии для управления качеством образовательной и научной деятельности вуза; повышение эффективности оценки научно-педагогической деятельности участников образовательного процесса; возможность выявления научных школ вуза и их использование в системе управления знаниями; формирование и укрепление научной школы по направлению исследования.

ЛИТЕРАТУРА

1. Жомарткызы Г., Балова Т.Г. Разработка онтологии образовательных ресурсов вуза» //Труды международной научной конференции студентов, магистрантов и молодых ученых «Ломоносов 2012». I-часть. – Астана: Казахстанский филиал МГУ им. В.М. Ломоносова, 2012. – С. 130-132.
2. Дюсекеев К.А. Методика формирования требований к системе показателей эффективности деятельности сотрудников вуза [текст] / Дюсекеев К.А., Нань Фэн, Шикунская О.М. // Материалы X Международной научно-практической конференции «Перспективы развития научно-технического сотрудничества стран-участниц Евразийского экономического союза». – г. Астрахань. – 9-11 ноября 2016 г. – С. 278-283.
3. Балова Т.Г., Жомарткызы Г. Автоматическое извлечение терминологических коллокаций научных текстов. //Вестник КазНТУ. – 2013. – №3(97). – С. 342-346.
4. Zakharov V. The distributive and statistical analysis as a tool to automate the formation of semantic fields (on the example of the linguocultural concept of “empire”) //Академия наук Республики Татарстан. – 2018. – С. 182.

5. Uvalieva I., Garifullina, Z., Utegenova, A., Toibayeva, S., Issin, B. Development of intelligent system to support management decision-making in education // e-Book of Papers. 6th International Conference on Modelling, Simulation and Applied Optimization. - Istanbul, Turkey. – 2015. pp.1-7.
6. Uvaliyeva I.M., G.Soltan, G.Sarsembayeva. Comprehensive monitoring of educational objects based on educational statistics // Herald of the Kazakh-British technical university. – Алматы. 2016. – Volume 13, Issue 2,3 – pp.166-173.
7. SolodovnikovaDarja, Niedrite Laila, Niedritis Aivars. Architecture Enabling Adaptation of Data Integration Processes for a Research Information System. Foundations of Computing and Decision Sciences, Volume: 43, Issue: 2, Pages: 129-149 Published: JUN 2018.
8. Uvaliyeva I.M. The architecture of information-analytical system of educational statistics // Herald of the Kazakh-British technical university. – Алматы. 2015. – Volume 12, Issue 2. – pp.66-76.