

АНАЛИЗ ОБРАЗОВАТЕЛЬНЫХ ДАННЫХ ДЛЯ СОВЕРШЕНСТВОВАНИЯ БИЗНЕС-ПРОЦЕССОВ В ВУЗЕ

Стасышин В.М

к.т.н., доцент, директор Центра информатизации
Новосибирский государственный технический университет (НГТУ)

Аннотация. Представлены направления и результаты исследований в области анализа образовательных данных, полученные с использованием системы поддержки принятия решений в НГТУ. Полученные результаты использовались для принятия управленческих решений для совершенствования бизнес-процессов в различных сферах деятельности вуза.

Ключевые слова: бизнес-процессы, информационная система, анализ образовательных данных, поддержка принятия решений, алгоритмы классификации.

ANALYSIS OF EDUCATIONAL DATA TO IMPROVE BUSINESS PROCESSES AT UNIVERSITY

Stasyshin V.M.

Resume. Some directions and results of research in the field of educational data analysis in decision-making support system at NSTU, are presented. The results used to make management decisions to improve business processes in various areas of the University.

Keywords: business processes, information system, educational data analysis, decision-making support, classification algorithms.

Решая задачу повышения уровня высшего образования, государство ставит вузы в условия соревновательной борьбы, в которой побеждают вузы, сумевшие перестроиться в соответствии с требованиями времени и мобилизовавшие для этой цели все свои ресурсы. Цель, преследуемая государством, – избавиться от слабых вузов на рынке образовательных услуг.

Глобализация образования приводит к расширению возможностей студентов получать качественное образование в учебных заведениях по всему миру. Университеты сегодня, как и бизнес-организации, работают в динамичной и конкурентной среде. Они сталкиваются с жесткой конкуренцией между собой, пытаясь привлечь наиболее подходящих студентов, которые успешно пройдут через образовательный процесс университета, и прилагая усилия, чтобы справиться с удержанием студентов.

Все это требует от администрации университета принятия своевременных управленческих решений по повышению эффективности и организации учебного процесса. Процесс принятия решений в сфере образования очень сложен, так как в нем участвует широкий круг заинтересованных лиц, так или иначе связанных с высшим образованием (преподаватели, сотрудники, студенты, их родители, сотрудники министерства и др.). Этот процесс неизбежно содержит множество факторов субъективного характера. Грамотный анализ информации, накопленной в информационной системе университета, является одним из элементов, позволяющих повысить уровень объективности и качества принимаемых решений. Руководству университета приходится быстро принимать важные решения, а это требует своевременной и качественной информации.

Современные вузы собирают большие объемы данных, касающихся их студентов, организации и управления образовательным процессом, а также других бизнес-процессов. За 15 лет своего развития в информационной системе (ИС) Новосибирского государственного технического университета (НГТУ) накоплен огромный объем данных по всем сферам деятельности университета. Эти данные могут быть использованы для принятия

управленческих решений по совершенствованию бизнес-процессов университета [1].

Анализ образовательных данных для совершенствования бизнес-процессов в вузе является одной из современных тенденций системы образования. В начале 2000-х годов появилось даже направление «Интеллектуальный анализ образовательных данных» (EDM – Educational data mining), задачей которого является всесторонний анализ данных, накопленных в информационных системах, обслуживающих образовательный процесс, с целью принятия управленческих решений на уровне вуза.

Ниже приводятся результаты исследований в области анализа образовательных данных в Новосибирском государственном техническом университете.

1. Одним из направлений исследований являлся анализ данных в части контингента абитуриентов и лиц, поступивших в университет.

Комплексный анализ данных Информационной системы в этой сфере, проводился по заданию администрации университета с целью выработки рекомендаций по совершенствованию профессионально-ориентационной работы вуза.

Проведенные исследования позволили [2]:

- проанализировать распределение контингента абитуриентов по регионам в различных разрезах (по административным районам г. Новосибирска, по районам Новосибирской области и городам областного подчинения, по субъектам и городам Российской Федерации, по городам республики Казахстан, по другим странам зарубежья);
- выявить набор школ-поставщиков абитуриентов НГТУ, дающих наибольший процент абитуриентов в целом за период 2011-2018 гг. и отдельно за каждый год по г. Новосибирску, НСО, России, республики Казахстан;
- проанализировать качество подготовки выпускников школ-поставщиков абитуриентов НГТУ (средний балл ЕГЭ, процент поступивших в университет, процент успешно завершивших университет, процент отчисленных);
- выявить предпочтения выпускников тех или иных школ-поставщиков в выборе факультета и направления обучения;
- проанализировать тенденции увеличения (уменьшения) числа абитуриентов школ-поставщиков абитуриентов НГТУ.

Выполненный анализ позволил сформировать профиль абитуриента НГТУ. Это в подавляющем случае выпускник городской школы (95%). Если рассматривать весь контингент абитуриентов (~ 64 тыс. за последние 7 лет), то в 90% случаев (~ 58 тыс. абитуриентов) это выпускник одной из школ России, а в половине случаев это выпускник школы Новосибирского региона. Из абитуриентов Новосибирского региона это преимущественно выпускник школы (лицея, гимназии) города Новосибирска, ведущей профильное обучение по некоторому направлению (математическое, физическое, языковое, экономическое и т.д.). Вторую половину абитуриентов НГТУ преимущественно составляют выпускники российских школ близлежащих регионов: Кемеровская область, Алтайский край, Иркутская область, республика Саха (Якутия), республика Бурятия, Красноярский и Забайкальский край, республика Хакасия, республика Алтай, Ханты-Мансийский автономный округ (по мере уменьшения доли абитуриентов).

Абитуриент зарубежья в подавляющем большинстве случаев (~ 8 тыс. абитуриентов за последние 8 лет) представлен выпускниками школ республики Казахстан. Это опять же в большинстве своем выпускники школ (лицеев, гимназий) крупных городов северо-восточной и центральной частей Казахстана: Усть-Каменогорск, Павлодар, Семей, Зыряновск, Караганда.

Сопоставление данных за 2017-2018 годы с данными за 2011-2016 годов выявило тенденцию к уменьшению числа абитуриентов из ряда новосибирских школ. Администрацией университета сделаны организационные выводы в отношении тех руководителей факультетов, которые допустили ослабление профессионально-ориентационной работы.

Проведенный анализ лег в основу управленческих решений администрации университета по совершенствованию профессионально-ориентационной работы в вузе. Был разработан комплекс мероприятий, направленных на активизацию работы со школами-

поставщиками абитуриентов НГТУ, подготовлены рекомендации для более адресной работы с абитуриентами.

2. Следующее исследование касается анализа факторов, влияющих на отчисление студентов. Задача удержания студентов – это классическая задача EDM, и ее решению посвящено множество публикаций, например, [3-5].

Новосибирский государственный технический университет, как и любой вуз, крайне заинтересован в сохранении набранного контингента студентов, поэтому полезно было бы заранее знать «группу риска» студентов с высокой вероятностью отчисления с тем, чтобы своевременно предпринять меры и оказать студентам помощь (особые условия обучения, дополнительные консультации и т.д.).

Цель исследования – выяснить, существуют ли шаблоны (наборы входных переменных-факторов), которые могут быть полезны для предсказания отчисления студентов (переменная выходного класса). Факторы, которые анализировались с точки зрения их влияния на отчисление студентов:

- средний балл ЕГЭ при поступлении;
- категория среднего учебного заведения (лицей/гимназия/средняя школа областного центра/районная средняя школа/сельская средняя школа/колледж/техникум);
- место проживания (дома/в общежитии/на квартире);
- социальная обеспеченность (удовлетворительная/ низкая: справка из отдела социальной защиты, сирота, инвалид);
- зачисление (на общих основаниях/целевой набор);
- средний балл в первую сессию.

В исследовании использовались данные из Информационной системы университета: данные за 2011-2015 гг. использовались для обучения классификатора, данные за 2011-2018 гг. использовались для тестирования. Анализировались данные по всем факультетам и по всем образовательным программам. Для построения модели классификации применялись несколько различных алгоритмов из различных пакетов программной системы R.

По большинству факультетов и образовательных программ достигнута высокая точность классификатора на уровне 75-82%.

В целом, по результатам исследований можно сделать следующие выводы [6]:

- 1) Рассмотренные входные признаки по-разному влияют на отчисление студентов. Решающий фактор из числа рассмотренных, влияющий на результаты классификации, – «Средний балл в первую сессию». Остальные атрибуты, в том числе «Средний балл ЕГЭ при поступлении» существенного влияния на отчисления студентов не оказывают (рис. 1).

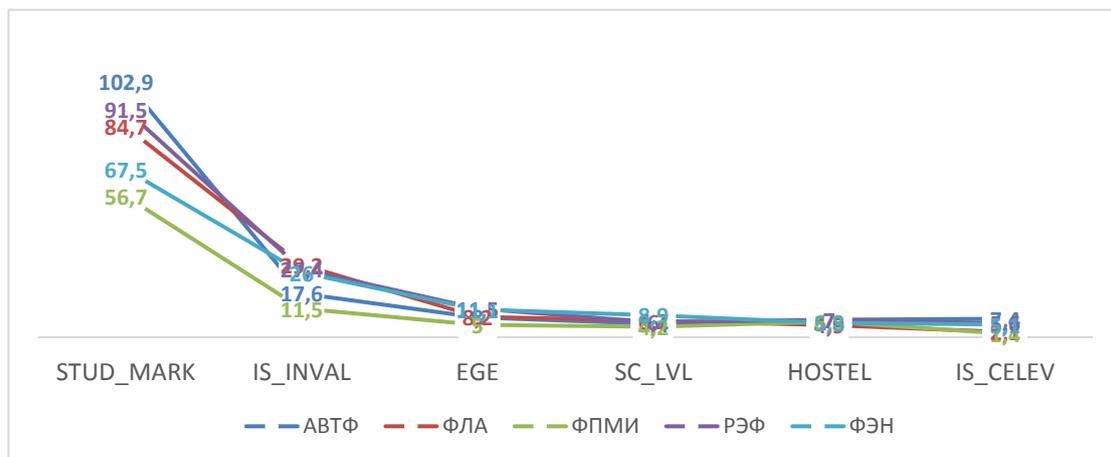


Рис. 1. Относительный вклад каждого из исследуемых атрибутов на результирующий фактор – отчисление студентов

- 2) Отсутствие, в большинстве случаев, атрибутов «Категория среднего учебного заведения» в построенных деревьях решений свидетельствует о том, что этот атрибут достаточно слабо влияет на успеваемость студентов в университете.
- 3) Низкое значение атрибута «Средний балл в первую сессию» с высокой долей вероятности свидетельствует о потенциальной угрозе будущего отчисления студента. Таким студентам с высокой угрозой будущего отчисления уже на ранней стадии (после первой сессии) следует предоставить меры дополнительной помощи (особые условия обучения, дополнительные консультации и т.д.) с целью их удержания в университете.

Помимо этого, была исследована динамика изменения значимости исследуемых факторов во временном диапазоне и степень влияния перехода от числовых данных к категориальным.

Полученные различными классификаторами деревья классификации (рис. 2) содержат от одного до пяти уровней иерархии. Атрибут «Средний балл в первую сессию» появляется на первом уровне дерева для всех без исключения факультетов, для факультетов Автоматики и вычислительной техники, факультета Прикладной математики и информатики, факультета гуманитарного образования его одного достаточно для проведения классификации с достаточной точностью. Для остальных факультетов атрибуты «Социальная обеспеченность», «Категория среднего учебного заведения», «Место проживания» и «Средний балл ЕГЭ при поступлении» отображаются со второго по пятый уровень дерева, что означает, что эти атрибуты влияют на большинство классификаций экземпляров, но это влияние значительно меньше, чем у атрибута «Средний балл в первую сессию». При этом атрибут «Зачисление» не появляется в дереве классификации ни для одного факультета и ни одного направления/специальности, что свидетельствует о том, что он не влияет на результаты классификации.

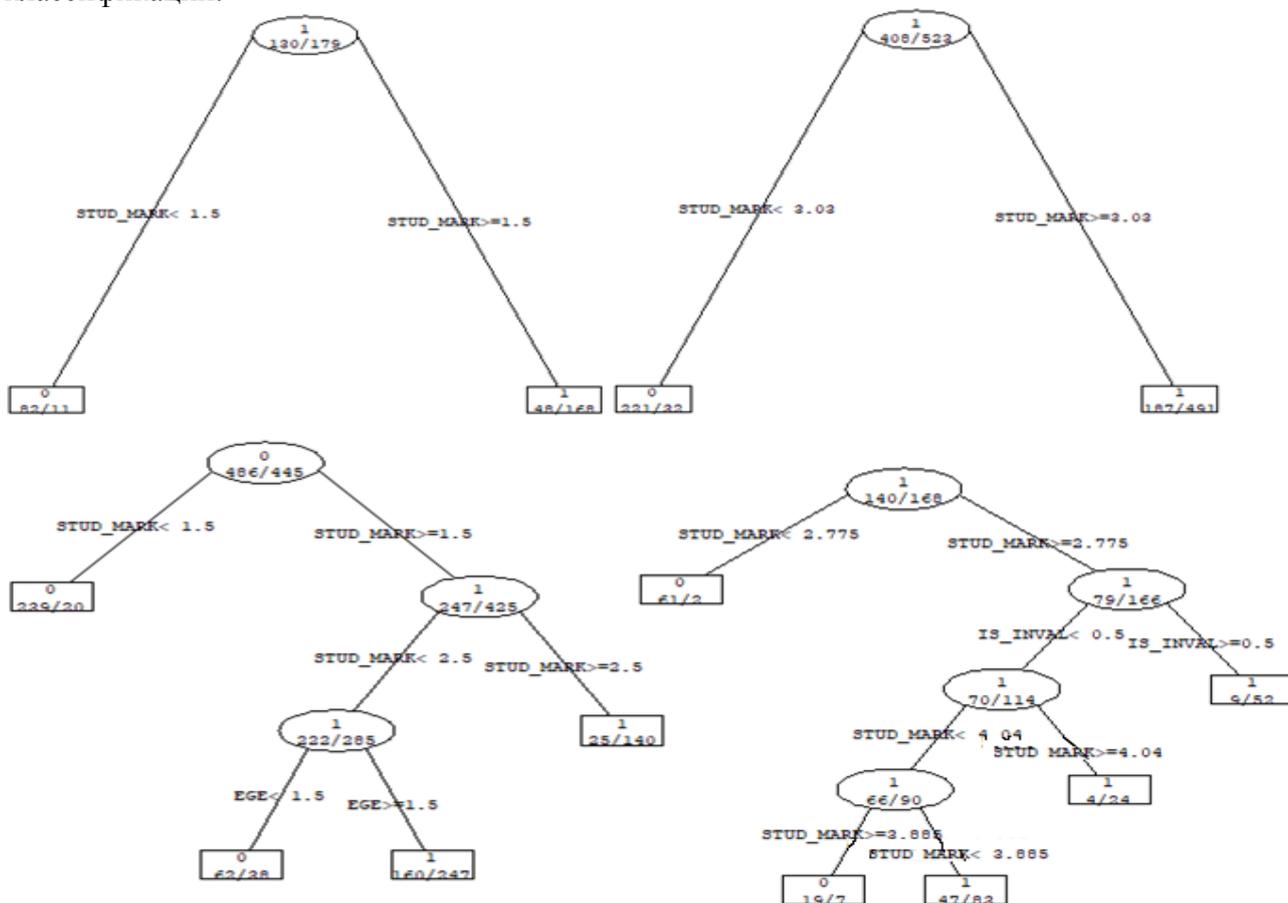


Рис. 2. Деревья классификации для различных факультетов

Точность предсказания можно было бы повысить, рассмотрев дополнительные входные факторы: факт работы студента параллельно с обучением, личные качества студента (мотивация, настойчивость и пр.), успеваемость по предметам в школе, относящимся к профилю обучения студента. К сожалению, этих данных в настоящее время нет в Информационной системе. Кроме того, полезны были бы такие факторы, как социальный статус семьи студента, квалификация, образование и семейный доход родителей, но сбор этих данных невозможен по Закону о персональных данных.

3. Создание электронных образовательных ресурсов (учебных пособий, конспектов лекций и т. д.) является одним из видов деятельности профессорско-преподавательского состава университета. Электронная библиотечная система НГТУ содержит несколько десятков тысяч электронных образовательных ресурсов, написанных сотрудниками университета. Однако уровень и качество этих ресурсов весьма различно. Так же как уровень научных публикаций оценивается индексом цитирования, логично оценивать качество образовательных ресурсов по их востребованности. Востребованность же образовательных ресурсов можно оценить по количеству загрузок этих ресурсов студентами и сотрудниками. Это позволяет получать рейтинги самых популярных электронных образовательных ресурсов и рейтинги авторов, ресурсы которых наиболее востребованы.

В НГТУ, как и в большинстве вузов, внедрена и на протяжении последних нескольких лет используется система эффективного контракта [7]. При этом среди показателей эффективного контракта доминируют два фактора с наибольшими весами: публикационная активность сотрудников и объем денег, заработанных сотрудниками для университета. Дважды в год публично подводятся итоги выполнения эффективного контракта факультетами, кафедрами и формируется рейтинг эффективного контракта научно-педагогическими работниками университета. При этом администрация университета постоянно отмечает, что около ста наиболее активных сотрудников университета выполняют 50% показателей эффективного контракта. Это свидетельствует, по мнению администрации, о неэффективной работе значительной части научно-педагогического коллектива университета.

В связи с этим было интересно соотнести показатели эффективного контракта научно-педагогических работников и востребованность электронных образовательных ресурсов этих сотрудников. На рис. 3 данные о количестве скачиваний авторских ресурсов наложены на данные рейтинга научно-педагогических работников. Мы видим, что спрос на образовательные ресурсы научно-педагогических кадров из первой тысячи рейтинга держится примерно на одном уровне, после чего они резко сокращаются.

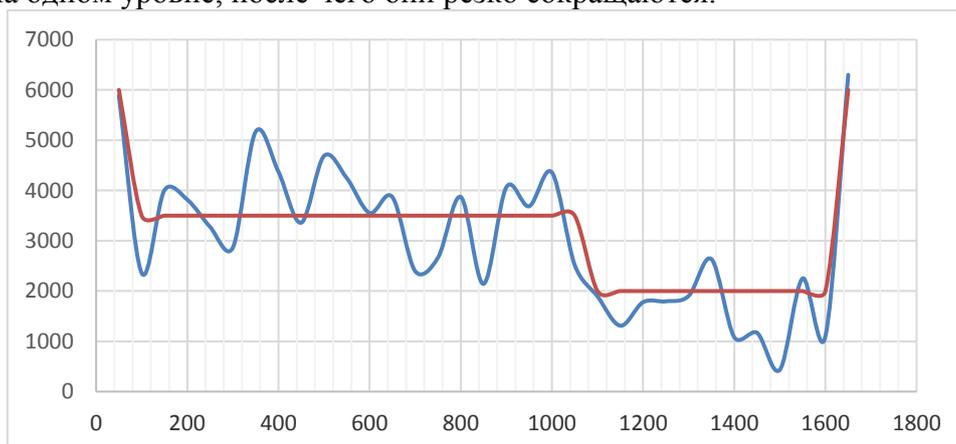


Рис. 3. Соотношение исполнения эффективного контракта и спроса на электронные образовательные ресурсы научно-педагогических работников

Более высокий спрос на образовательные ресурсы сотрудников, возглавляющих рейтинг эффективного контракта, можно объяснить тем, что их учебные материалы, как правило, готовятся в соавторстве с другими сотрудниками. При этом наблюдается резкий спрос на образовательные ресурсы тех сотрудников, которые замыкают рейтинг. Это свидетельствует

о том, что, не добившись успеха в научно-исследовательской деятельности, эти сотрудники очень полезны вузу в подготовке качественных электронных образовательных ресурсов.

4. Считается, что средний балл ЕГЭ при приеме является одним из показателей «эффективности» вуза. Общепринятая качественная модель зависимости успеваемости студента в вузе от его балла ЕГЭ (Модель 1) имеет вид, представленный на рис. 4.1.

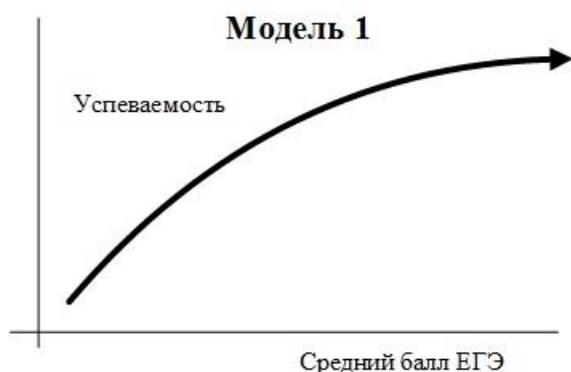


Рис. 4.1. Общепринятая качественная модель зависимости успеваемости студента в вузе от его балла ЕГЭ

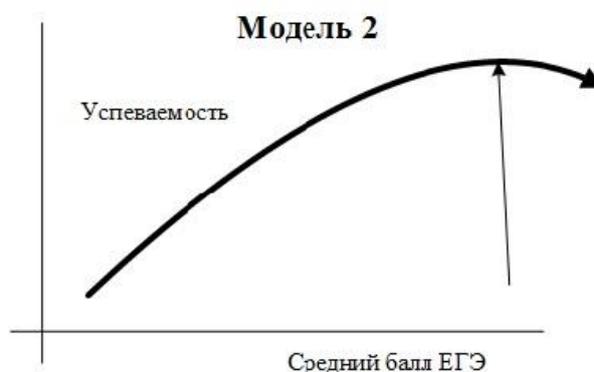


Рис. 4.2. Альтернативная качественная модель зависимости успеваемости студента в вузе от его балла ЕГЭ

В ходе исследований ставилась задача подтвердить или опровергнуть утверждение, что эта модель является единственно верной.

Анализ построенных отчетов и регрессионных моделей показал, что зависимость успеваемости от среднего балла ЕГЭ далеко не всегда имеет характер Модели 1, и лучшую успеваемость зачастую показывают студенты отнюдь не с самым высоким баллом ЕГЭ. Такой характер зависимости (Модель 2) представлен на рис. 4.2.

Несмотря на то, что Модель 1 является доминирующей и ее характер наблюдается в подавляющем числе случаев, по целому ряду направлений/специальностей отдельных факультетов, для отдельных семестров отдельных факультетов характер зависимости успеваемости от балла ЕГЭ соответствует Модели 2.

Ряд преподавателей, которые были ознакомлены с представленными данными, высказали предположение, что характер Модели 2 является признаком определенных проблем в учебном процессе: перегруженность студентов, несбалансированность учебных курсов и т.д.

5. Между Российской федерацией и республикой Казахстан существует межправительственное соглашение по взаимному признанию аттестатов средних школ. Новосибирская область граничит с республикой Казахстан. Следствием этого является то, что порядка десяти процентов абитуриентов – выпускники школ Казахстана, и в последние годы этот процент только растет. Сопоставительный анализ качества подготовки абитуриентов показывает, что средний балл ЕНТ выпускников школ Казахстана (Единое Национальное Тестирование – аналог российского ЕГЭ) составляет 75-85 баллов против среднего балла ЕГЭ (65-70 баллов) для выпускников российских вузов, поступающих в НГТУ. Следует отметить, что выпускники школ Казахстана поступают в НГТУ не на основе балла ЕНТ, а по результатам выездного тестирования, проводимого преподавателями НГТУ на территории Казахстана.

В этой связи вызывает интерес, действительно ли знания выпускников казахстанских школ выше знаний выпускников российских школ, как это следует из сопоставления баллов ЕГЭ и ЕНТ, и насколько коррелируют результаты ЕГЭ (тестирование для выпускников школ Казахстана) с успеваемостью студентов при обучении в вузе.

На рис. 5.1 отчетливо видно, что для выпускников российских школ успеваемость студентов в вузе тесно коррелирует со значением балла ЕГЭ при поступлении: для студентов, имеющих более высокий балл ЕГЭ, процент неудовлетворительных и удовлетворительных оценок относительно мал, а процент хороших и отличных оценок, наоборот, относительно велик. Для студентов, имеющих относительно низкий балл ЕГЭ, наблюдается обратная картина. У выпускников казахстанских школ, поступивших по результатам выездного

тестирования, корреляции между результатами тестирования выпускников школ и успеваемостью студентов не просматривается (рис. 5.2): получение любой оценки равновероятно независимо от балла тестирования.

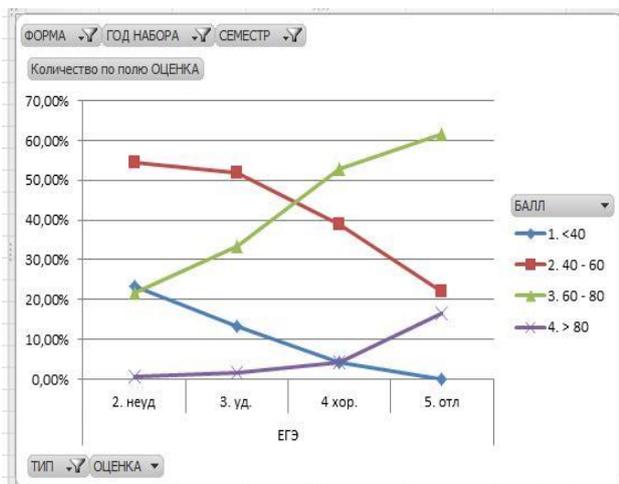


Рис. 5.1. Зависимость успеваемости от балла ЕГЭ для выпускников российских школ

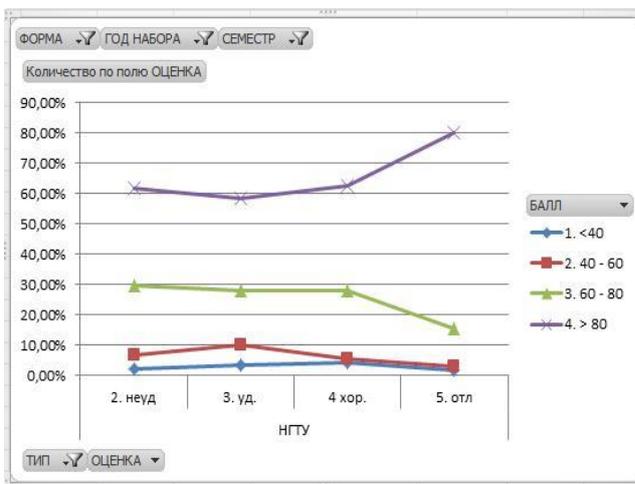


Рис. 5.2. Зависимость успеваемости от балла тестирования для выпускников казахстанских школ

Таким образом, можно сделать вывод, что приведенных цифры среднего балла ЕГЭ и ЕНТ отнюдь не свидетельствуют о более высоком уровне подготовки выпускников школ Казахстана. А причина увеличивающегося потока поступающих в НГТУ выпускников школ Казахстана связана с низким уровнем предлагаемых тестов и недостаточной частотой их обновления. Проведенный анализ позволил скорректировать методику проведения выездного тестирования на территории Казахстана.

6. Еще одно направление исследований заключалось в анализе подходов к повышению качества высшего образования, проводимое с использованием методов экспертных оценок.

В результате предварительных бесед с преподавателями университета, имеющими богатый опыт работы, было сформулировано восемь направлений (альтернатив) решения указанной проблемы:

- 1) Исключение прямой зависимости между численностью студентов вуза и объемом финансирования из министерства, что позволило бы избавляться вузу от нерадивых студентов.
- 2) Привлечение к процессу преподавания в вузе на условиях адекватной оплаты труда высококвалифицированных специалистов с производства, научно-исследовательских, коммерческих организаций.
- 3) Повышение планки уровня знаний при приеме абитуриентов, что позволило бы отсеять слабо подготовленных выпускников средней школы на вступительных испытаниях.
- 4) Повышение уровня оплаты труда преподавателей вуза.
- 5) Повышение уровня квалификации преподавателей вуза путем активного использования системы стажировки, переподготовки, повышения научной квалификации.
- 6) Участие студентов в научных, проектных и конструкторских работах, имеющих производственное или научное значение.
- 7) Отказ от слепого копирования системы высшего образования западного образца.
- 8) Повышение качества читаемых учебных курсов и их методического обеспечения путем предварительного обязательного обсуждения разработанных преподавателем учебных курсов внутри педагогического сообщества вуза.

Экспертная группа, участвовавшая в исследовании, состояла из 13 сотрудников НГТУ, имеющих стаж работы от 7 до 48 лет. Эксперты представляли три факультета (факультет прикладной математики и информатики, механико-технологического факультет и факультет

мехатроники и автоматизации), четыре кафедры (кафедры прикладной математики, теоретической и прикладной информатики, материаловедения и машиностроения, теоретических основ электротехники). Среди экспертов было 3 заведующих кафедрами, 4 доктора наук, 7 кандидатов наук, 5 профессоров, 6 доцентов, 1 старший преподаватель и 3 руководителя структурных подразделений университета.

Все эксперты были проранжированы рейтингом от 3,7 до 10,8 в зависимости от стажа работы, должности, звания.

Было проведено анкетирование, после чего предпочтения экспертов были обработаны пятью методами экспертных оценок: метод парных сравнений, метод взвешенных экспертных оценок, метод предпочтения, метод ранга, метод полного попарного сопоставления.

В работе [8] анализируются результаты проведенных исследований.

К числу наиболее неожиданных результатов следует отнести тот факт, что всеми пятью методами экспертных оценок предпочтение было отдано альтернативе «Повышение уровня оплаты труда преподавателей вуза», несмотря на то, что среди предлагаемых альтернатив, казалось бы, были альтернативы, заслуживающие большего внимания. При этом нельзя сказать, что победил меркантильный интерес, поскольку эксперты, в большинстве своем, не ставили эту альтернативу на первое место.

СПИСОК ЛИТЕРАТУРЫ

1. Стасышин В.М. Информационная система университета: опыт создания и текущее состояние / В.М. Стасышин, О.Е. Аврунев, Е.В. Афонина, К.Н. Лях // Открытое и дистанционное образование. – Томск, 2012. – № 2 (46). – С. 9-15.

2. Стасышин В. М. Принятие управленческих решений в НГТУ на основе анализа образовательных данных = Management decision making at NSTU based on the analysis of educational data [Электронный ресурс] / В. М. Стасышин, Т. В. Стасышин // Инновационные, информационные и коммуникационные технологии: сб. тр. 14 междунар. науч.-практ. конф., Сочи, 1–10 окт. 2017 г. – Москва: Ассоц. выпускников и сотр. ВВИА им. проф. Жуковского, 2017. – С. 179–184.

3. Kabakchieva D. Predicting student performance by using data mining methods for classification. – Cybernetics and Information Technologies. – 2013. – Vol. 13. – No. 1. – pp. 61-72.

4. Superby, J. Vandamme, J., Meskens, N. Determination of factors influencing the achievement of the first-year university students using data mining methods // Proceedings of the Workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems (ITS 2006). – Jhongli, Taiwan, pp. 37-44.

5. Kabakchieva, D., Stefanova, K., Kisimov, V. Analyzing University Data for Determining Student Profiles and Predicting Performance. Conference Proceedings of the 4th International Conference on Educational Data Mining (EDM 2011), 6-8 July 2011, Eindhoven, The Netherlands, pp.347-348.

6. Stasyshin V. M. Analysis of educational data in the decision-making support system of university / V. M. Stasyshin, T. V. Stasyshin // Актуальные проблемы электронного приборостроения (АПЭП–2018) = Actual problems of electronic instrument engineering (APEIE–2018): тр. 14 междунар. науч.-техн. конф., Новосибирск, 2–6 окт. 2018 г.: в 8 т. – Новосибирск: Изд-во НГТУ, 2018. – Т. 1, ч. 4. – С. 541-545 - ISBN (NSTU) 978-5-7782-3614-1.

7. Аврунев О. Е. Опыт внедрения системы эффективного контракта в Новосибирском государственном техническом университете / О. Е. Аврунев, Г. И. Расторгуев, В. М. Стасышин // Инновационные, информационные и коммуникационные технологии: сб. тр. 13 междунар. науч.-практ. конф., Сочи, 1–10 окт. 2016 г. – Москва: НИУ ВШЭ, 2016. – С. 120–123. - ISSN 2500-1248.

8. Stasyshin V. M. Analysis of educational data to improve business processes at university / V. M. Stasyshin, A. D. Svilpova // Information innovative technologies: materials of the intern. sci.-practical conf., Czech Republic, Prague, 22–26 Apr. 2019. – Moscow: Assoc. of graduates and employees of AFEA named after prof. Zhukovsky, 2019. – P. 246-252. - ISSN 2542-1824.