

Применение интеллектуального анализа данных для поиска взаимосвязи между результатами ЕГЭ и академической успеваемостью студентов ВУЗа

Китаева Ольга Игоревна

Иркутский национальный исследовательский технический университет,

Россия, Иркутск, *koi_fk@mail.ru*

Аннотация. В статье рассматривается применение интеллектуального анализа образовательных данных для сравнения результатов успеваемости и баллов Единого государственного экзамена (ЕГЭ), студентов, обучающихся по ИТ-специальностям ВУЗа. Актуальность темы обусловлена необходимостью принимать объективные и обоснованные решения для повышения результативности образовательного процесса. Основным методом является получение интеллектуально улучшенных визуальных форм данных для более эффективного понимания скрытых в них закономерностей.

В работе рассмотрены возможности применения программного продукта с открытым кодом «Orange» для реализации интеллектуального анализа образовательных данных с целью определения взаимосвязи между показателями ЕГЭ и академической успеваемостью студентов. Рассмотрены задачи: оценка влияния результатов ЕГЭ на успешность учебы студента в ВУЗе, определение возможности использования результатов ЕГЭ как индикатора академической успеваемости студента, выявление студентов, находящихся в зоне риска по результатам ЕГЭ.

Визуальное программирование, реализуемое удобным графическим интерфейсом программного обеспечения Orange Data Mining, позволило получить интересные и полезные иллюстративные материалы, наглядно демонстрирующие связь баллов ЕГЭ с оценками, полученными студентами на первом курсе, и последующей результативностью обучения в ВУЗе. Использование виджета Box Plot обеспечило необходимую статистическую информацию.

Возможность оценивания студентов, исходя из имеющихся результатов ЕГЭ, в момент поступления в ВУЗ, очевидно, полезна и оправдана, т.к. на момент поступления в университет эта информация наиболее доступна и информативна. В ходе исследования использована классификация студентов, исходя из итогов обучения на младших курсах ВУЗа. Выявленные закономерности подтверждают целесообразность использования результатов ЕГЭ для прогнозирования успешности студента ВУЗа. Вместе с тем, полученные результаты показывают, что часть обучающихся, имеющих высокие баллы по ЕГЭ, впоследствии перестают учиться. Этот факт указывает на необходимость использования дополнительных показателей, в частности, постоянного мониторинга посещаемости и текущей успеваемости для своевременного выявления студентов, нуждающихся в дополнительной поддержке.

Ключевые слова: интеллектуальный анализ образовательных данных, Orange, анализ текущей успеваемости студентов, визуализация

Цитирование: Китаева О.И. Применение интеллектуального анализа данных для поиска взаимосвязи между результатами ЕГЭ и академической успеваемостью студентов ВУЗа / О.И. Китаева // Информационные и математические технологии в науке и управлении, 2026. – № 1(41). – С.68-78. – DOI:10.25729/ESI.2026.41.1.005.

Введение. Рассматриваемая тема привлекала и привлекает внимание авторов статей, являющихся преимущественно преподавателями ВУЗов, что говорит о большой заинтересованности в стремлении использовать все возможные приемы для повышения качества образования. Вместе с тем, приведенные результаты существенно разнятся. Так Рочев К.В., Куделин А.Г., Семяшкина А.В. указывают на явную корреляцию между результатами ЕГЭ и средней успеваемостью студентов [1]. В результате их исследования выявлена зависимость между диапазоном баллов ЕГЭ, полученных при поступлении, и последующими успехами, выраженными средним баллом и итоговыми дипломными оценками. В то время, как Бурухина Т. Ф., Винокуров Е. Г. считают [2], что результаты ЕГЭ обладают слабой прогностической способностью в отношении успешности обучения в высшей школе. Вместе с тем, авторы [2] считают, что «решение проблемы успешной адаптации первокурсников в масштабах каждого образовательного учреждения позволяет

повысить успеваемость обучающихся, уменьшить отсев, снизить пропуски занятий без уважительных причин, поднять уровень творческой, профессиональной и социальной активности студенчества и, в итоге, повысить качество образования до уровня, соответствующего запросам общественного развития». Вызывает интерес мнение Комковой О.Г., Переведенцевой С.Е., Ворсиной Е.В. [3] о том, что высокие баллы по результатам ЕГЭ могут служить лишь предикторами академической успеваемости студентов ВУЗа на начальных курсах. Авторы считают, что значение принадлежит таким компонентам личности обучающегося, как мотивация, интерес, система ценностей, ответственность, преодоление и продвижение себя. Сычев О.А. также считает [4], что для прогнозирования будущей успеваемости студентов и профилактики неуспеваемости полезно учитывать не только результаты ЕГЭ, но и учебную мотивацию студентов. Аналогичное мнение высказывает Нуриева Л.М. [5]: «Связь результатов ЕГЭ с успехами в вузе существует постольку, поскольку они измеряют одну и ту же характеристику студента: отношение к учебе. При этом перечень вступительных экзаменов почти не имеет значения, а прогностическая ценность профильных экзаменов (на примере математики) оказалась невысока». В статье «Результаты ЕГЭ и успеваемость студентов первого курса» [6] Щеголева Л. В. и Суровцова Т.Г. обращают внимание на то, что имеющееся в вузе большое количество данных, связанных с учебным процессом, практически не используется при формировании управленческих решений. Рассматривая проблему преемственности общеобразовательной и высшей школы, авторы [6] считают, что с точки зрения управления деятельностью вуза, введение комплексного мониторинга успеваемости с использованием баллов ЕГЭ позволит решить ряд задач по организации учебного процесса.

В данном случае целью является проверка гипотезы «существует значимая связь между баллами ЕГЭ и успеваемостью студентов», т.е. определение, насколько хорошо высокие баллы ЕГЭ предсказывают успешность обучения. Оценка уровня связи позволит, кроме повышения качества обучения и разработки программ поддержки слабых студентов, уделить внимание определению факторов, влияющих на изменение мотивации обучения студентов с высоким баллом ЕГЭ после поступления в вуз. Таким образом, поиск взаимосвязи между результатами ЕГЭ и академической успеваемостью студентов позволит указать пути решения одной из важных задач – повышение результативности образовательного процесса в ВУЗе.

Для проверки выдвинутой гипотезы используется Интеллектуальный анализ образовательных данных (от англ. Educational Data Mining) (ИАОД). ИАОД – совокупность методов и алгоритмов анализа данных, накапливаемых в учебном заведении в процессе его деятельности с целью выявления скрытых, неочевидных, практически полезных и интерпретируемых знаний об учебном процессе и его участниках для поддержки и принятия решений [7]. Это направление исследований, связанное с применением методов интеллектуального анализа данных (Data mining) [8-10], машинного обучения и статистики к информации, производимой образовательными учреждениями. В качестве инструмента исследования используется программный продукт Orange Data Mining

Основные задачи поиска взаимосвязи между результатами ЕГЭ и академической успеваемостью студентов ВУЗа. Основные задачи настоящего исследования приведены в таблице 1. Решение поставленных задач выполняется на базе результатов ЕГЭ и другой информации о студентах ИТ-специальностей Иркутского национального исследовательского технического университета (ИРНИТУ).

Таблица 1. Задачи исследования

| | Задача | Цель решения задачи |
|---|--|---|
| 1 | Определить, могут ли результаты ЕГЭ быть индикатором академической успеваемости студента | Определение целесообразности прогнозирования успешности студента ВУЗа, исходя из результатов ЕГЭ |
| 2 | Оценить влияние результатов ЕГЭ на успешность учебы студента в ВУЗе | Своевременная корректировка образовательной траектории студента |
| 3 | Выявление студентов, находящихся в зоне риска по результатам ЕГЭ | Выявление пороговых значений, при которых результаты ЕГЭ наиболее точно выявляют студентов, нуждающихся в дополнительной поддержке. |

Использование программного продукта Orange Data Mining для решения поставленных задач. Orange – свободно распространяемая библиотека, написанная на языке Python, основанная на принципе визуального программирования для наглядного доступа к алгоритмам Data mining. Разработана и поддерживается Bioinformatics Laboratory of the Faculty of Computer and Information Science Люблянского университета (Словения). Компоненты аналитической платформы называются виджетами и варьируются от минималистичной визуализации данных, выбора подмножеств и предварительной обработки до эмпирической оценки алгоритмов обучения и прогностического моделирования [11-14].

Основные этапы использования программы Orange:

1. Подготовка данных:
 - загрузка данных: импорт данных из файлов с помощью модуля "File",
 - обзор и визуализация: использование "Data Table", "Distributions" и "Box Plot" для предварительного анализа.
2. Предварительная обработка:
 - очистка,
 - масштабирование,
 - выбор признаков.
3. Построение моделей:
 - выбор алгоритма (например, «Logistic Regression», «Decision Tree», «k-NN»),
 - соединение модуля данных с модулем модели,
 - настройка параметров модели.
4. Обучение и тестирование модели:
 - разделение данных на обучающую и тестовую выборки (например, с помощью «Data Sampler»),
 - обучение модели и оценка её эффективности (например, с помощью «Test & Score»).
5. Визуализация и интерпретация результатов.
6. Применение модели.

Структура данных, использованная для решения Задачи 1 «Определить, могут ли результаты ЕГЭ быть индикатором академической успеваемости студента» приведена в таблице 2. Схема анализа данных в Orange Data Mining показана на рис.1.

Таблица 2. Структура данных

| Имя переменной | Назначение | Тип переменной |
|------------------|----------------------------|----------------|
| id | номер студента в списке | числовой |
| группа | название учебной группы | текстовый |
| курс | номер курса | числовой |
| Русский язык ЕГЭ | балл ЕГЭ по русскому языку | числовой |

| | | |
|---------------------|-----------------------------------|----------------|
| Математика ЕГЭ | балл ЕГЭ по математике | числовой |
| Физика ЕГЭ | балл ЕГЭ по физике | числовой |
| Информатика ЕГЭ | балл ЕГЭ по информатике | числовой |
| информатика/физика | дисциплина на выбор | текстовый |
| категория | категория студента | текстовый |
| Математика ВУЗ | оценка по математике в ВУЗе | числовой |
| Вовремя математика | признак сдачи математики вовремя | категориальный |
| Физика ВУЗ | оценка по физике в ВУЗе | числовой |
| Вовремя физика | признак сдачи физики вовремя | категориальный |
| Информатика ВУЗ | экзаменационная оценка | числовой |
| Вовремя информатика | признак сдачи информатики вовремя | категориальный |

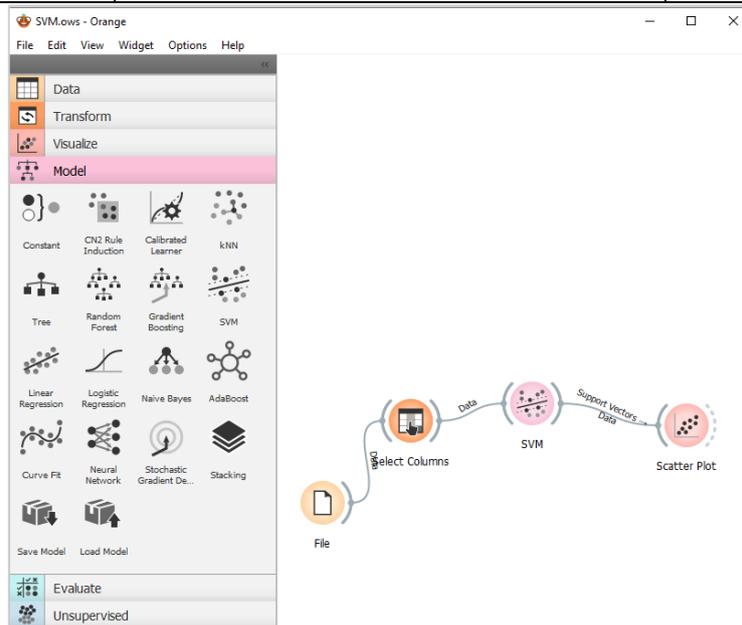


Рис. 1. Схема анализа данных в Orange Data Mining

На рис. 2 показаны результаты сравнения результатов ЕГЭ (информатика, математика,) с оценками в ВУЗе. Каждая точка на графике показывает информацию по конкретному студенту: горизонтальная ось – оценка по соответствующей дисциплине в ВУЗе, вертикальная – балл ЕГЭ, диапазоны цветов точек показаны в нижнем правом углу графиков. Значение «0» для оценки в ВУЗе соответствует несданному экзамену.

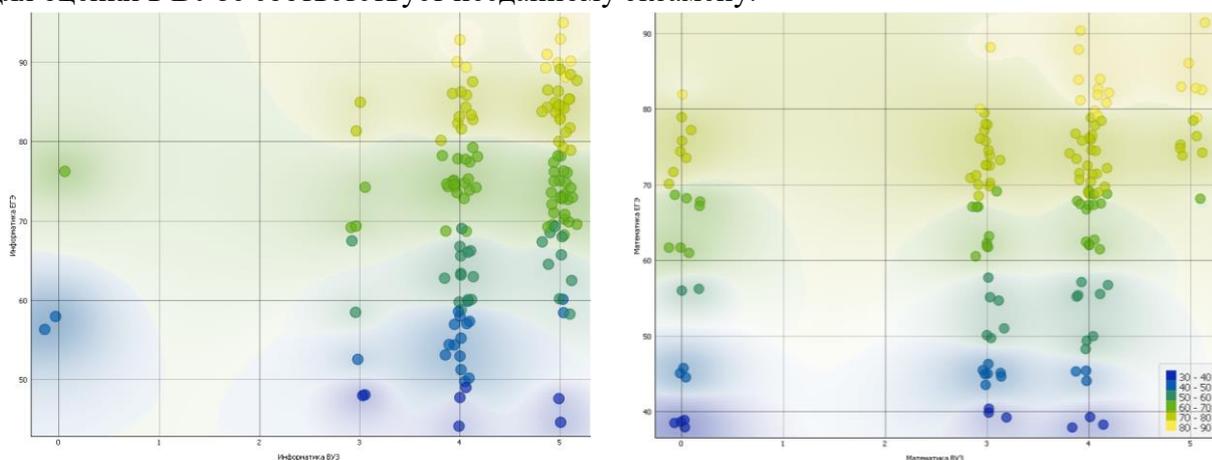


Рис. 2. Сравнение результатов ЕГЭ (информатика, математика,) с оценками в ВУЗе

Полученные графики (рис. 2) с точки зрения рассматриваемой задачи дают, на первый взгляд, разные выводы по выбранным дисциплинам. Первый из графиков, где оценки, полученные по информатике в ВУЗе, сравниваются с результатами ЕГЭ, несомненно подтверждает возможность использовать результаты ЕГЭ, как индикатор академической успеваемости студента. Вероятно, этому способствует и тот факт, что программа информатики в ВУЗе во многом близка школьной программе по этой дисциплине. Второй график позволяет сделать два неоднозначных вывода. Первый совпадает с предыдущим, как видно из рисунка 2, отличные оценки по математике получены студентами с высоким баллом ЕГЭ. Второй вывод касается студентов, не сдавшим математику во время сессии (точки на графике, соответствующие оценке 0), балл ЕГЭ у этих студентов находится в диапазоне от 39 до 82, этому может быть несколько причин, первая из которых – различие учебных программ и требований ВУЗа и школы. Следующие причины относятся скорее всего к проблемам адаптации к формату высшего образования. Таким образом, не отрицая целесообразность прогнозирования успешности студента ВУЗа, исходя из результатов ЕГЭ, необходимо находить и рассматривать и другие факторы.

Для решения задачи «Оценить влияние результатов ЕГЭ на успешность учебы студента в ВУЗе» в структуру данных, представленную в таблице 2, добавлена категориальная переменная «Успех», которая показывает, продолжает ли конкретный студент обучаться на момент анализа информации (в данном случае на 3 курсе).

Для сравнения балла ЕГЭ и успешного продолжения обучения использована схема, показанная на рисунке 3.

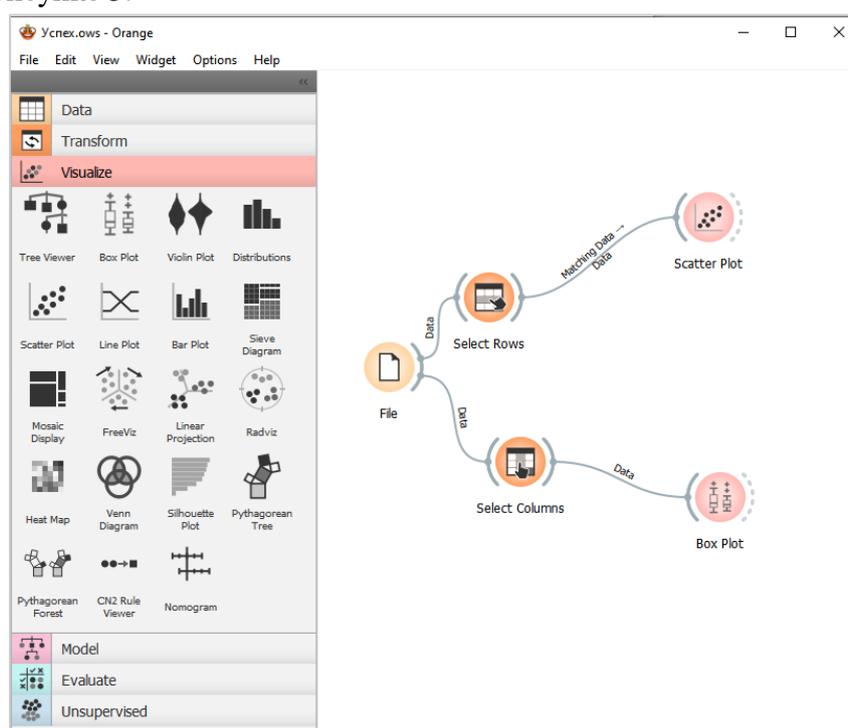


Рис. 3. Сравнение результатов ЕГЭ относительно переменной «Успех»

На рисунке 4 с левой стороны представлена визуализация результатов ЕГЭ по математике студентов, продолжающих обучение на 3 курсе, а с правой – тех, кто по каким-то причинам прервал обучение. Показанные на рисунке 4 результаты явно свидетельствуют о том, что балл ЕГЭ слабо влияет на то, продолжает ли студент заниматься непрерывно; разные интервалы результатов ЕГЭ, в том числе и весьма высокие, встречаются и в левой, и в правой частях рисунка.

На рисунке 5 показаны результаты сравнения баллов ЕГЭ студентов, прервавших и продолживших обучение, с использованием виджета Vox Plot. Vox Plot позволяет получить график, известный в описательной статистике под названиями «Ящик с усами», «Диаграмма размаха», «Усыковая диаграмма», «Коробчатая диаграмма». В этом случае верхний из графиков для каждой из дисциплин характеризует студентов, прервавших обучение, а нижний – продолживших. В контексте рассматриваемой задачи, синяя область («Коробка» или «Ящик») представляет межквартильный размах (IQR), распределение основных 50% студентов по баллам, и показывает диапазон, в котором сконцентрированы большинство результатов. Желтая вертикальная линия – медиана, средний или типичный балл, половина студентов набрала меньше этого значения, половина – больше. Среднее значение – тёмно-синяя вертикальная линия. Синяя линия – Усы (whiskers) – диапазон, охватывающий данные без выбросов. Тонкая пунктирная линия – Выбросы (outliers) – очень низкие или очень высокие результаты (например, студенты с необычно низкими или высокими баллами).

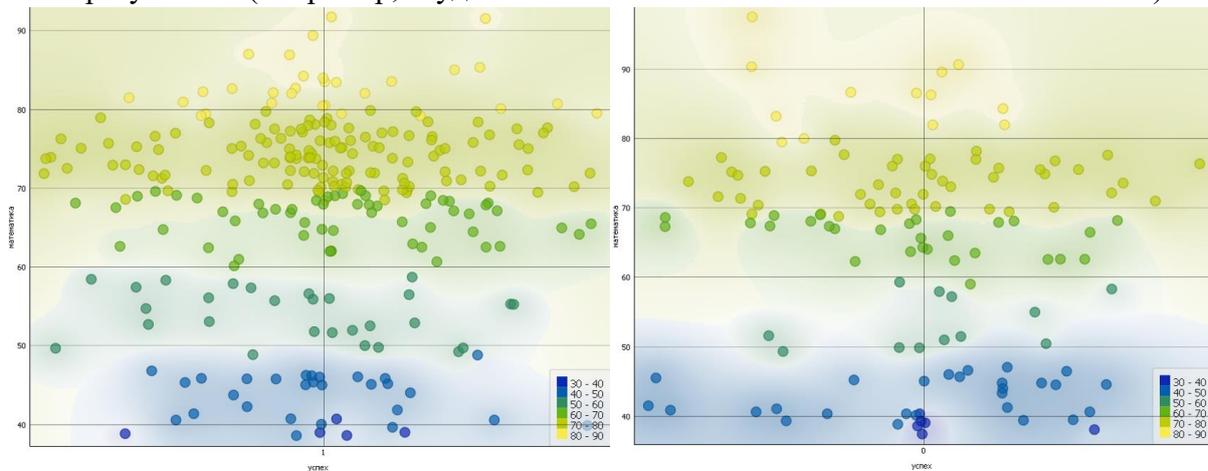
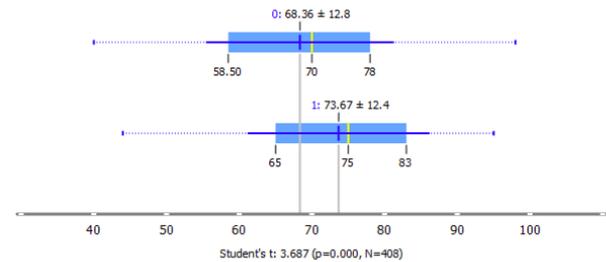
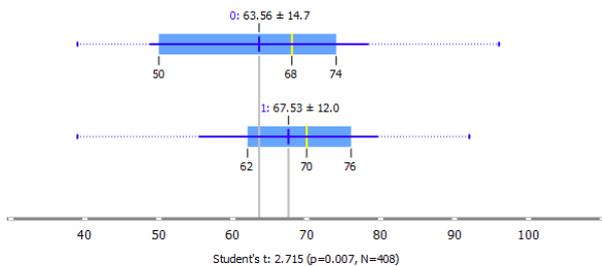


Рис. 4. Связь результатов ЕГЭ по математике с переменной «Успех» (1)

Математика

Информатика



Физика

Русский язык

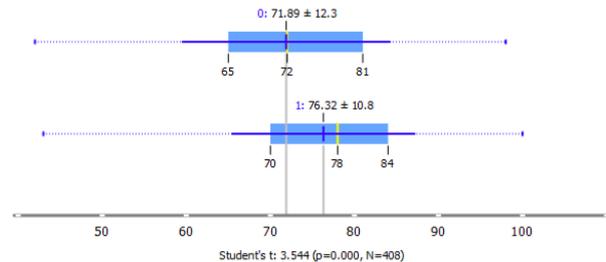
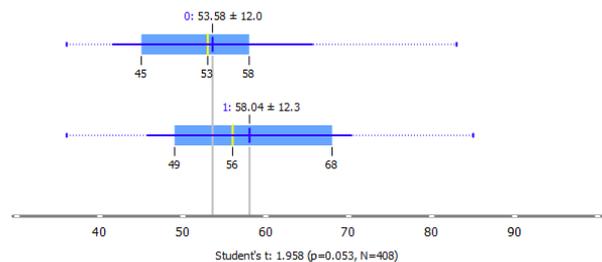


Рис. 5. Связь результатов ЕГЭ с переменной «Успех» (2)

При решении поставленной задачи наиболее логичным представляется использовать межквартильный размах (IQR) и диапазон баллов ЕГЭ, охватывающий данные без выбросов (whiskers); соответствующие значения приведены в таблице 3.

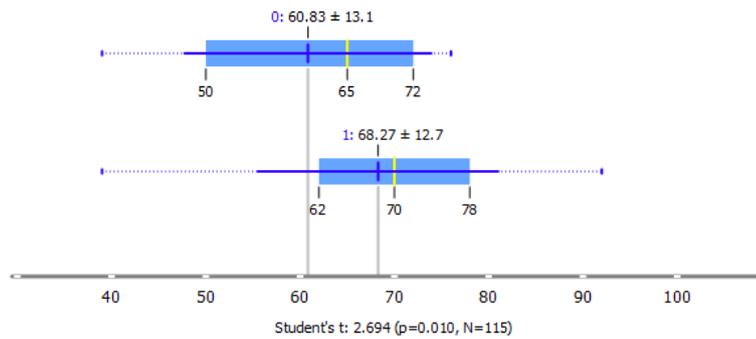
Таблица 3. Влияние балла ЕГЭ на продолжение обучения

| | Математика | | Русский язык | | Физика | | Информатика | |
|---------------------|------------|----------|--------------|----------|--------|----------|-------------|----------|
| | IQR | whiskers | IQR | whiskers | IQR | whiskers | IQR | whiskers |
| Продолжают обучение | 62-76 | 56-80 | 70-84 | 65-88 | 49-68 | 46-72 | 65-83 | 62-86 |
| Прекратили обучение | 50-74 | 48-79 | 65-81 | 60-85 | 45-58 | 42-66 | 59-78 | 56-82 |

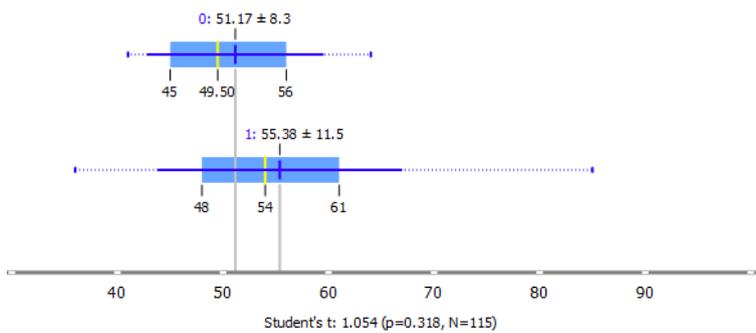
Значения, представленные в таблице, не противоречат выводам, сделанным выше, о том, что результат ЕГЭ не оказывает существенного влияния, на то, продолжает ли студент обучение. Исключение представляют результаты по математике, здесь, как видно из таблицы, прекращению обучения, вероятнее всего, соответствует низкий балл ЕГЭ.

Для определения пороговых значений для предсказания академических успехов по результатам ЕГЭ (задача «Выявление студентов, находящихся в зоне риска по результатам ЕГЭ») выполнено сравнение балла ЕГЭ и категориальной переменной, показывающей, была ли сдана вовремя соответствующая дисциплина в первом семестре обучения. Результаты сравнения показаны на рисунке 6, здесь вновь использован виджет Box Plot, т.е. результаты представлены в форме «Ящик с усами». Верхний график соответствует студентам, не сдавшим вовремя соответствующую дисциплину в первом семестре, нижний – сдали вовремя. Под пороговым значением будем понимать конкретное значение балла ЕГЭ, ниже которого студент попадает в зону риска.

Математика



Физика



Информатика

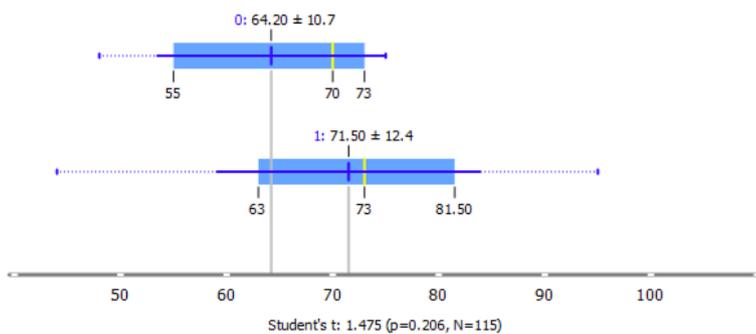


Рис. 6. Связь результатов ЕГЭ и своевременной сдачи дисциплин

Для сравнения результатов, представленных на рис. 6, все студенты были разделены на три группы: неуспевающие студенты, потенциально успешные (студенты с равной вероятностью успеха и неуспеха) и успешные (успевающие) студенты. В таблицу 4 сведены результаты ЕГЭ относительно предложенной выше классификации.

Таблица 4. Влияние балла ЕГЭ на успешность обучения

| Дисциплина | Баллы ЕГЭ, неуспевающих студентов | Баллы ЕГЭ, потенциально успешных студентов | Баллы ЕГЭ, успешных студентов |
|-------------|-----------------------------------|--|-------------------------------|
| Математика | 48-56 | 56-74 | 74-82 |
| Физика | 42-44 | 44-58 | 58-68 |
| Информатика | 53-58 | 58-75 | 75-84 |

Из вышесказанного можно сделать следующие выводы:

1. Отрицать влияние балла ЕГЭ на успешность обучения не верно, несмотря на то, что в целом ряде случаев связь результатов ЕГЭ с последующими итогами обучения не всегда прослеживается.
2. Возможность классификации студентов, исходя из имеющихся результатов ЕГЭ, в момент поступления в ВУЗ, очевидно полезна и оправдана, хотя бы потому, что на начальном этапе других сведений о поступивших или нет, или они не достаточны.
3. Тот факт, что обучающиеся, имеющие высокие баллы по ЕГЭ, впоследствии перестают учиться, говорит о необходимости постоянного мониторинга посещаемости и текущей успеваемости для своевременного выявления студентов, требующих поддержки.
4. Сравнение результатов, представленных в таблицах 3 и 4, говорит о необходимости постоянного внимания и своевременных корректирующих действий для всех студентов, вне зависимости от того, к какой группе из предложенной выше классификации они относятся.

Заключение. Интеллектуальный анализ образовательных данных позволяет использовать данные об учебном процессе и его участниках для выявления скрытых, неочевидных, практически полезных и интерпретируемых знаний [15]. В ходе исследования были проанализированы данные о студентах, обучающихся по программам ИТ-направления бакалавриата. Использован программный продукт с открытым исходным кодом Orange date mining. Визуальные данные, полученные программой Orange, не доказали и не опровергли связь балла ЕГЭ, имеющегося при поступлении, с успешностью обучения в ВУЗе. Значительная доля студентов, прекративших обучение, несмотря на высокий балл ЕГЭ, очевидно, обусловлена сложным сочетанием психологических, социальных, финансовых и академических факторов. Высокие результаты ЕГЭ, как правило, не гарантируют сохранение мотивации или устойчивого образовательного успеха в дальнейшем.

Предложено на базе результатов, полученных на основе интеллектуального анализа данных (сравнения баллов ЕГЭ и академической успеваемости) с применением программного продукта Orange, выявлять обучающихся, находящихся в зоне риска, для выполнения корректирующих действий с целью улучшения качества образовательного процесса и повышения академической успеваемости студентов. В этой связи представляется интересным продолжить исследование с учетом дополнительных факторов, таких, как интервал времени между сдачей ЕГЭ и поступлением в ВУЗ, а также сравнение успешности студентов, поступивших после окончания колледжа и сдававших ЕГЭ.

Список источников

1. Рочев К.В. Исследование комплексного влияния ЕГЭ на академическое развитие студентов / К.В. Рочев, А.Г. Куделин, А.В. Семяшкина. – DOI: 10.60797/IRJ.2024.147.49.
2. Бурухина Т.Ф. Анализ успеваемости студентов младших курсов и ее связи с результатами ЕГЭ / Т.Ф. Бурухина, Е.Г. Винокуров. – DOI: 10.31862/2218-8711-2021-2-139-147.

3. Комкова О.Г. Исследование взаимосвязи учебной успеваемости студентов-первокурсников медицинского вуза с высокими стартовыми показателями / О.Г. Комкова, С.Е. Переведенцева, Е.В. Ворсина // Психолого-педагогический журнал «Гаудеамус», 2024. – Т. 23. – № 3. – С. 38-47. – DOI: 10.20310/1810-231X-2024-23-3-38-47.
4. Сычев О.А. Результаты ЕГЭ и мотивация как предикторы долгосрочной успеваемости студентов педагогического вуза // Психологическое здоровье и психологическая культура в современном российском образовании: материалы VI Всероссийской научно-практической конференции, 2015. – С. 175-178.
5. Нуриева Л.М. Анализ связи результатов ЕГЭ с академической успеваемостью студентов-первокурсников // Познание и деятельность: от прошлого к настоящему: материалы II Всероссийской междисциплинарной научной конференции. – Омск, 2020. – С. 171-175.
6. Щеголева Л.В. Результаты ЕГЭ и успеваемость студентов первого курса / Л.В. Щеголева, Т.Г. Суровцова // Непрерывное образование: XXI век, 2015. – № 4 (12). – С. 33-41.
7. Оськин А.Ф. Применение интеллектуального анализа образовательных данных для прогнозирования успешности учебной деятельности / А.Ф. Оськин, Д.А. Оськин // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки, 2016. – № 4. – URL: <https://elib.psu.by/bitstream/123456789/17089/5/8-12.pdf> (дата обращения: 20.08.2025).
8. Дюк В. Data Mining: учебный курс / В. Дюк, А. Самойленко. – СПб.: Питер, 2001. – 368 с.
9. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям / Н.Б. Паклин, В.И. Орешков. – СПб.: Питер, 2013. – 704 с.
10. Агаев Ф.Т. Прогнозирование успеваемости студентов в электронном образовании с использованием методов data mining / Ф.Т. Агаев, Г.А. Мамедова, Р.Т. Меликова. – URL: <https://bik.sfu-kras.ru/ft/LIB2/ELIB/b74/free/i-489634964.pdf> (дата обращения: 20.08.2025).
11. Orange: Описание, Функции и Интерфейс // Soware. – 2025. – URL: <https://soware.ru/products/orange> (дата обращения: 21.08.2025).
12. Тюшняков В.Н. Интеллектуальный анализ маркетинговой деятельности на основе применения инструментов Orange data mining / В.Н. Тюшняков // Вестник Таганрогского института имени А.П. Чехова, 2024. – № 2. – С. 103-107.
13. Юсупов Н. Исследование методов классификации в программе Orange / Н. Юсупов, А. Савельева, О.Г. Леонова // Молодежная школа-семинар по проблемам управления в технических системах имени А.А. Вавилова, 2020. – Т. 1. – С. 27-30.
14. Юсупов Н. Исследование методов кластеризации в программе Orange / Н. Юсупов // Молодежная школа-семинар по проблемам управления в технических системах имени А.А. Вавилова, 2020. – Т. 1. – С. 35-37.
15. Китаева О.И. Интеллектуальный анализ образовательных данных учебной дисциплины с использованием программы Orange / О.И. Китаева // Информационные и математические технологии в науке и управлении, 2023. – № 1(29). – С. 190-200. – DOI: 10.25729/ESI.2023.29.1.016.

Китаева Ольга Игоревна. Старший преподаватель института информационных технологий и анализа данных Иркутского национального исследовательского технического университета, руководитель отдела по работе со студентами, сфера научных интересов: Интеллектуальный анализ образовательных данных. AuthorID: 688343, SPIN: 6638-7790.

UDC 378.14.015.62

DOI:10.25729/ESI.2026.41.1.005

The use of data mining to find the relationship between the results of the unified state exam and the academic performance of university students

Olga I. Kitaeva

Irkutsk National Research Technical University, Russia, Irkutsk, koi_fk@mail.ru

Abstract. The article discusses the use of intellectual analysis of educational data to compare the results of academic performance and scores of the Unified State Exam (USE), students studying in the IT specialties of the university. The relevance of the topic is due to the need to make objective and informed decisions to improve the effectiveness of the educational process. The main method is to obtain intelligently enhanced visual forms of data for a more effective understanding of the patterns hidden in them.

The paper considers the possibilities of using the open source software product "Orange" to implement the intellectual analysis of educational data in order to determine the relationship between the USE indicators and the academic performance of students. The tasks are considered: assessing the impact of the USE results on the student's academic success at the university, determining the possibility of using the USE results as an indicator of student academic achievement, identifying students at risk based on the results of the USE.

Visual programming, implemented by the user-friendly graphical interface of the Orange Data Mining software, allowed us to obtain interesting and useful illustrative materials that clearly demonstrate the relationship between USE scores, grades received by students in their first year and the subsequent effectiveness of their studies at the university. Using the Box Plot widget provided the necessary statistical information.

The possibility of evaluating students based on the available USE results at the time of admission to the university is obviously useful and justified, because at the time of admission to the university, this information is the most accessible and informative. In the course of the study, the classification of students was used based on the results of their studies at the junior courses of the university. The revealed patterns confirm the expediency of using the results of the Unified State Exam to predict the success of a university student. At the same time, the results show that some students who have high scores on the Unified State Exam subsequently stop studying. This fact indicates the need to use additional indicators, in particular, constant monitoring of attendance and current academic performance in order to identify students in need of additional support in a timely manner.

Keywords: intellectual analysis of educational data, Orange, analysis of current student performance, visualization

References

1. Rochev K.V., Kudelin A.G., Semyashkina A.V. Issledovaniye kompleksnogo vliyaniya YeGE na akademicheskoye razvitiye studentov [Study of the comprehensive impact of the Unified State Exam on students' academic development], DOI: 10.60797/IRJ.2024.147.49.
2. Burukhina T.F., Vinokurov E.G. Analiz uspevayemosti studentov mladshikh kursov i yeye svyazi s rezul'tatami YeGE [Analysis of junior students' academic performance and its connection with the results of the Unified State Exam], DOI: 10.31862/2218-8711-2021-2-139-147.
3. Komkova O.G., Perevedentseva S.E., Vorsina E.V. Issledovaniye vzaimosvyazi uchebnoy uspevayemosti studentov-pervokursnikov meditsinskogo vuza s vysokimi startovymi pokazatelyami [Study of the relationship between the academic performance of first-year medical university students and high starting indicators]. Psikhologo-pedagogicheskiy zhurnal «Gaudeamus» [Psychological and Pedagogical Journal "Gaudeamus"], 2024, vol. 23, no. 3, pp. 38-47, DOI: 10.20310/1810-231X-2024-23-3-38-47..
4. Sychev O.A. Rezul'taty YeGE i motivatsiya kak prediktory dolgosrochnoy uspevayemosti studentov pedagogicheskogo vuza [USE results and motivation as predictors of long-term academic performance of pedagogical university students]. Psikhologicheskoye zdorov'ye i psikhologicheskaya kul'tura v sovremennom rossiyskom obrazovanii: materialy VI Vserossiyskoy nauchno-prakticheskoy konferentsii [Psychological health and psychological culture in modern Russian education: proceedings of the VI All-Russian Scientific and Practical Conference], 2015, pp. 175-178.
5. Nurieva L.M. Analiz svyazi rezul'tatov YeGE s akademicheskoy uspevayemost'yu studentov-pervokursnikov [Analysis of the connection between the results of the Unified State Exam and the academic performance of first-year students]. Poznaniye i deyatel'nost': ot proshlogo k nastoyashchemu: materialy II Vserossiyskoy mezhdistsiplinarnoy nauchnoy konferentsii [Cognition and Activity: From Past to Present: proceedings of the II All-Russian Interdisciplinary Scientific Conference]. Omsk, 2020, pp. 171-175.
6. Shchegoleva L.V., Surovtsova T.G. Rezul'taty YeGE i uspevayemost' studentov pervogo kursa [Unified State Exam results and academic performance of first-year students]. Nepreryvnoye obrazovaniye: XXI vek [Lifelong Education: The 21st Century], 2015, no. 4 (12), pp. 33-41.
7. Oskin A.F., Oskin D.A. Primeneniye intellektual'nogo analiza obrazovatel'nykh dannykh dlya prognozirovaniya uspeshnosti uchebnoy deyatel'nosti [Application of educational data mining for predicting academic success]. Vestnik Polotskogo gosudarstvennogo universiteta. Seriya S. Fundamental'nyye nauki [Bulletin of Polotsk State University. Series C. Fundamental Sciences], 2016, no. 4. Available at: <https://elib.psu.by/bitstream/123456789/17089/5/8-12.pdf> (accessed: 08.20.2025).
8. Duk V., Samoylenko A. Data Mining: uchebnyy kurs [Data Mining: Training Course]. Saint Petersburg, Piter Publ., 2001, 368 p.
9. Paklin N.B., Oreshkov V.I. Biznes-analitika: ot dannykh k znaniyam [Business Analytics: From Data to Knowledge]. Saint Petersburg, Piter Publ., 2013, 704 p.
10. Agaev F.T., Mamedova G.A., Melikova R.T. Prognozirovaniye uspevayemosti studentov v elektronnom obrazovanii s ispol'zovaniyem metodov data mining [Predicting student performance in e-learning using data

- mining methods]. Available at: <https://bik.sfu-kras.ru/ft/LIB2/ELIB/b74/free/i-489634964.pdf> (accessed: 08.20.2025).
11. Orange: Opisanie, Funktsii i Interfeys [Orange: Description, Functions and Interface]. Soware, 2025. Available at: <https://soware.ru/products/orange> (accessed: 08.21.2025).
 12. Tushnyakov V.N. Intellektual'nyy analiz marketingovoy deyatel'nosti na osnove primeneniya instrumentov Orange data mining [Intelligent analysis of marketing activities based on the use of Orange data mining tools]. Vestnik Taganrogskego instituta imeni A.P. Chekhova [Bulletin of the Taganrog Institute named after A.P. Chekhov], 2024, no. 2, pp. 103-107.
 13. Yusupov N., Saveleva A., Leonova O.G. Issledovaniye metodov klassifikatsii v programme Orange [Research of classification methods in the Orange program]. Molodezhnaya shkola-seminar po problemam upravleniya v tekhnicheskikh sistemakh imeni A.A. Vavilova [A.A. Vavilov Youth School-Seminar on Control Problems in Technical Systems], 2020, vol. 1, pp. 27-30.
 14. Yusupov N. Issledovaniye metodov klasterizatsii v programme Orange [Research of clustering methods in the Orange program]. Molodezhnaya shkola-seminar po problemam upravleniya v tekhnicheskikh sistemakh imeni A.A. Vavilova [A.A. Vavilov Youth School-Seminar on Control Problems in Technical Systems], 2020, vol. 1, pp. 35-37.
 15. Kitaeva O.I. Intellektual'nyy analiz obrazovatel'nykh dannykh uchebnoy distsipliny s ispol'zovaniyem programmy Orange [Intellectual analysis of educational data of an academic discipline using the Orange program]. Informatsionnyye i matematicheskiye tekhnologii v nauke i upravlenii [Information and Mathematical Technologies in Science and Management], 2023, no. 1(29), pp. 190-200. DOI: 10.25729/ESI.2023.29.1.016.

Kitaeva Olga Igorevna. Senior Lecturer at the Institute of Information Technology and Data Analysis at Irkutsk National Research Technical University, Head of the Student Affairs Department. Her research interests include intelligent analysis of educational data. AuthorID: 688343, SPIN: 6638-7790.

Статья поступила в редакцию 28.08.2025; одобрена после рецензирования 17.11.2025; принята к публикации 02.12.2025.

The article was submitted 08/28/2025; approved after reviewing 11/17/2025; accepted for publication 12/02/2025.