

ПРОГНОЗИРОВАНИЕ УСПЕВАЕМОСТИ ПЕРВОКУРСНИКОВ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ДОВУЗОВСКИХ ФАКТОРОВ

Садовникова Н.О., Черных Н.А.

Научный руководитель

к.т.н., доцент Золотарюк А.В.

Финансовый университет при Правительстве

Российской Федерации

В рамках проводимой реформы образования деятельность вузов направлена на совершенствование учебного процесса с целью реализации компетентностного подхода к формированию знаний, навыков и умений обучаемых, востребованных экономикой. Выпускники вузов должны в профессионально-ориентированных ситуациях комплексно исследовать возникающие проблемы, выбирать для их разрешения оптимальные технологические методы и средства, системно анализировать полученные результаты и на их основе предлагать эффективные научно-обоснованные решения.

В учебный процесс внедряются инновационные образовательные методики и технологии, активизируется НИР, развивается материально-техническая база. Вместе с тем, немаловажную роль в решении поставленной задачи играет контингент обучаемых, способность каждого зачисленного в вуз студента в полной мере реализовать свои интеллектуально-психологические особенности по усвоению требований образовательной программы.

В последние годы отбор в вузы осуществляется по результатам ЕГЭ, что, по задумке, должно обеспечить всем абитуриентам равные возможности, независимо от того, где и в какой школе они учились. Однако такой подход не лишен недостатков: порой подвергаются сомнению объективность результатов ЕГЭ в отдельных школах и регионах, затруднительным является учет творческих, индивидуальных особенностей школьника-выпускника.

Для устранения указанных нареканий, в ближайшее время планируется дополнительно учитывать при приеме в вуз данные паспорта школьника, в котором будут отражаться его физкультурно-спортивные показатели, творческие и образовательные достижения. Другими словами, прием будет осуществляться на основе интегральной формулы достижений абитуриента, что должно повысить объективность приема в вуз наиболее достойных, самых способных, творчески одаренных, физически выносливых, трудолюбивых студентов, которые с полной отдачей будут осваивать учебный материал, вести научно-исследовательскую и иную работу в вузе.

Исходя из сказанного, авторы задались несколькими вопросами. А нельзя ли каким-то образом спрогнозировать последующую вузовскую успеваемость абитуриентов? Есть ли взаимосвязь между довузовскими факторами абитуриента и его будущими успехами в вузе?

Сформулировав таким образом задачу, авторы провели исследование, взяв за основу информацию о студентах первого – третьего курсов Финансового университета, обучающихся по направлению 040100.62 «Социология».

В качестве довузовских (исходных) факторов были определены: $X(1)$ – школьная успеваемость; $X(2)$ – наличие медали; $X(3)$ – профиль класса; $X(4)$ – регион обучения; $X(5)$ – тип населенного пункта; $X(6)$ – пол; $X(7)$ – результат ЕГЭ по математике; $X(8)$ – результат ЕГЭ по русскому языку; $X(9)$ – условие зачисления в вуз; $X(10)$ – основа поступления в вуз.

Результативными (вузовскими) признаками стали рейтинговые баллы по итогам сдачи экзаменов на первом курсе студентами второго – третьего курсов, в частности: $Y(1)$ – по линейной алгебре; $Y(2)$ – по современным информационным технологиям в социальных науках.

В результате опроса студентов методом раздаточного анкетирования, а также использования информации из деканата, были получены требуемые сведения. В качестве инструментария была выбрана аналитическая платформа интеллектуальной обработки и извлечения знаний Deductor Studio.

Импорт исходной выборки в среду аналитической платформы обеспечил получение общих статистических данных (рис. 1). Используя метод корреляционного анализа, была исследована взаимосвязь между входными и выходными параметрами (рис. 2).

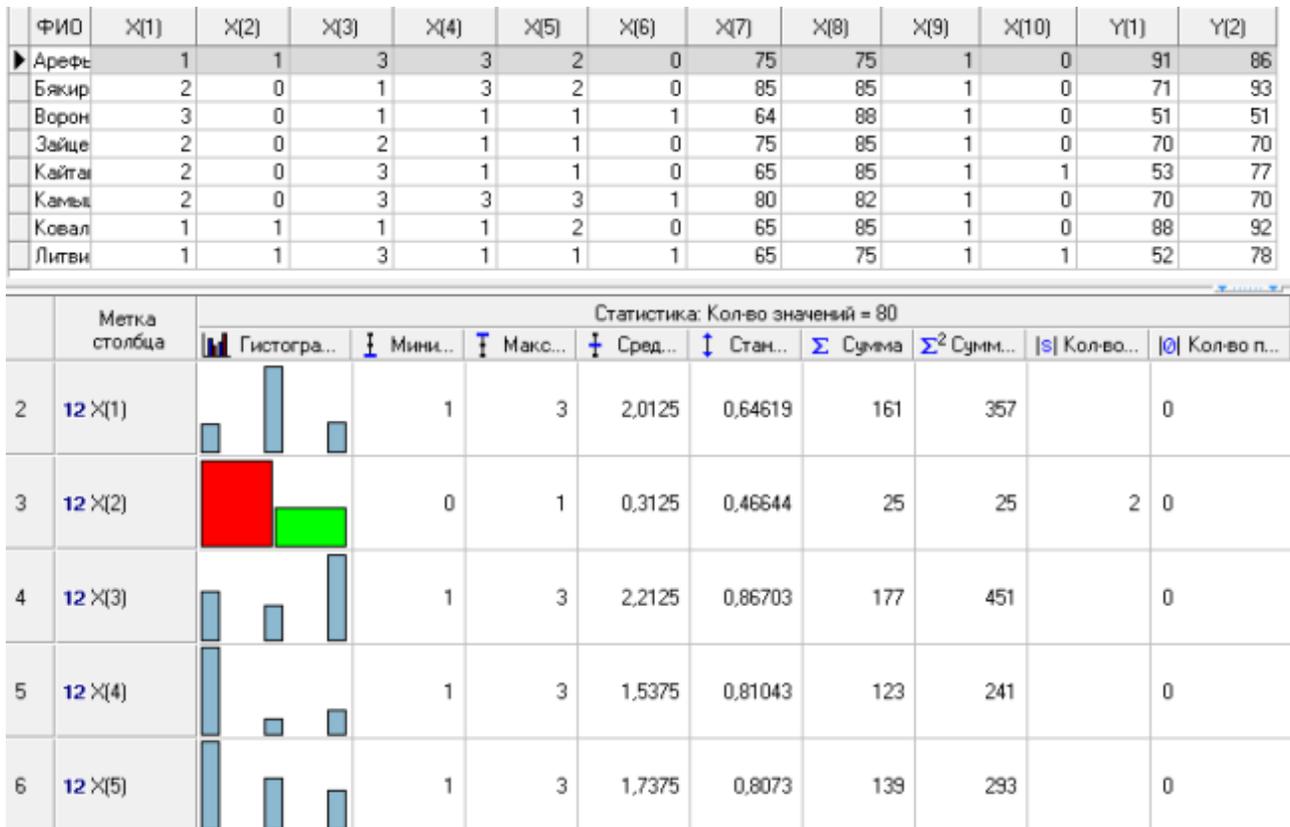


Рисунок 1. Фрагмент результатов импорта с данными статистики

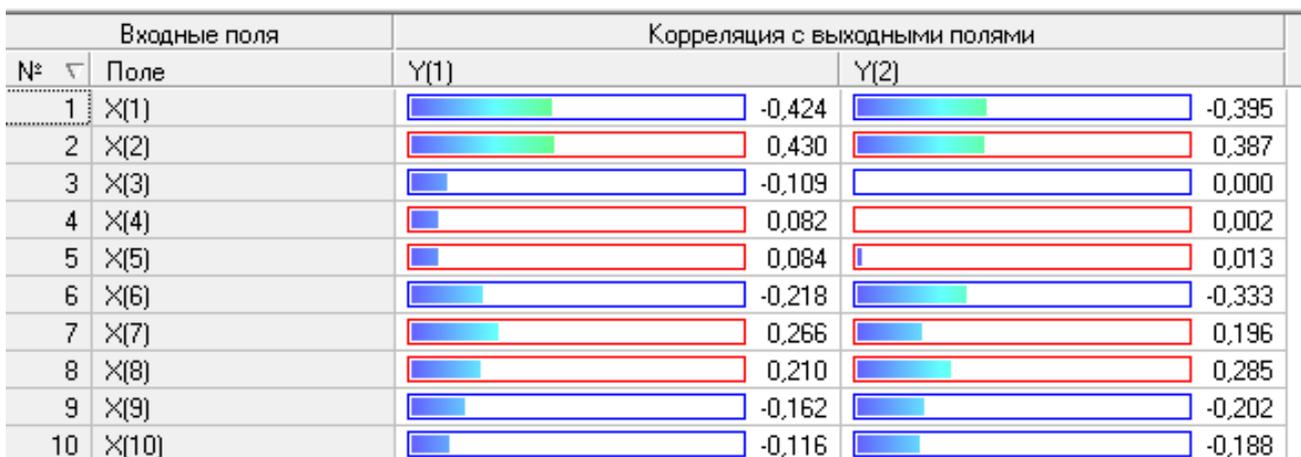


Рисунок 2. Корреляция входных и выходных параметров

Учитывая сложную форму зависимости результативных признаков от множества входных факторов, а в некоторых случаях полное отсутствие взаимосвязи, для решения поставленной задачи прогнозирования был использован метод многослойных нейронных сетей. На выборке данных студентов второго – третьего курсов было проведено обучение нейросети – эмпирическим путем выбрана ее оптимальная архитектура (количество скрытых слоев и нейронов в них), определены параметры обучения (объем тестовой выборки, вид активационной функции, максимальные значения допустимых ошибок, количество эпох (итераций) нахождения решения). В результате была построена нейронная сеть 10 x 7 x 4 x 2, а также сформированы выходные формы и таблицы. Таблица (рис. 3) содержит исходные данные, а также прогнозные (вероятностные) значения результативных признаков и соответствующие ошибки получения прогноза. Как видно, в большинстве случаев прогнозные параметры практически совпадают с фактическими, и только для отдельных строк ошибки присутствуют, что вполне согласуется с теорией.

X(1)	X(2)	X(3)	X(4)	X(5)	X(6)	X(7)	X(8)	X(9)	X(10)	Y(1)	Y(2)	Y(1)_OUT	Y(2)_OUT	Y(1)_ERR	Y(2)_ERR
1	1	3	3	2	0	75	75	1	0	91	86	92,6200	80,3747	0,0011	0,0132
2	0	1	3	2	0	85	85	1	0	71	93	71,2032	88,0518	0,0000	0,0102
3	0	1	1	1	1	64	88	1	0	51	51	51,0002	53,0992	0,0000	0,0018
2	0	2	1	1	0	75	85	1	0	70	70	69,3997	65,2765	0,0002	0,0093
2	0	3	1	1	0	65	85	1	1	53	77	55,6444	77,4842	0,0029	0,0001
2	0	3	3	3	1	80	82	1	0	70	70	51,0006	51,0278	0,1503	0,1499
1	1	1	1	2	0	65	85	1	0	88	92	88,5138	91,5887	0,0001	0,0001
1	1	3	1	1	1	65	75	1	1	52	78	52,1312	77,5030	0,0000	0,0001
1	1	1	1	1	0	65	85	1	0	71	86	67,7560	86,3662	0,0044	0,0001
2	0	3	3	3	0	55	75	1	0	75	86	76,3642	86,3920	0,0008	0,0001
3	0	3	1	1	1	75	85	3	0	51	51	51,0000	51,5000	0,0000	0,0001
2	1	1	1	1	0	65	75	1	1	75	86	78,3491	89,6161	0,0047	0,0054
2	0	3	3	1	0	75	85	1	0	56	79	52,7204	77,6292	0,0045	0,0008
3	0	2	3	2	1	73	88	1	0	76	61	77,9118	64,3970	0,0015	0,0048
2	0	3	2	3	0	95	75	1	0	73	70	75,5768	68,8971	0,0028	0,0005
2	1	3	1	3	0	75	85	3	0	77	88	83,9432	91,1413	0,0201	0,0041
1	1	3	1	1	0	85	75	1	1	97	77	97,1721	76,9999	0,0000	0,0000
2	1	3	1	3	0	66	88	1	0	92	86	96,7053	96,5059	0,0092	0,0460
2	0	1	2	2	1	75	75	1	0	57	51	55,3668	53,9923	0,0011	0,0037
2	0	3	1	1	0	65	95	1	0	93	91	93,8479	95,3978	0,0003	0,0081
1	1	2	1	2	0	75	85	1	0	94	100	95,8709	95,4852	0,0015	0,0085
2	0	3	3	2	0	75	85	1	0	53	77	52,7057	78,2098	0,0000	0,0006
2	1	2	2	3	0	65	95	3	1	52	57	51,1113	56,3836	0,0003	0,0002
2	0	3	1	1	0	85	75	2	0	86	83	100,0000	99,9165	0,0816	0,1192

Рисунок 3. Фрагмент таблицы с прогнозными значениями

Форма «Что-если» (рис. 4) позволяет прогнозировать значения результативных признаков в зависимости от заданных новых значений входных факторов – другими словами, прогнозировать успеваемость абитуриентов, а также отображать всевозможные зависимости в виде диаграмм различных типов

Дополнительно была оценена корреляция результатов сдачи ЕГЭ $X(7)$ – $X(8)$ с другими входными факторами $X(1)$ – $X(6)$. Не обнаружена зависимость от регионов. В то же время совсем незначительной является зависимость от школьной успеваемости и наличия медали, что может свидетельствовать о посредственной сдаче ЕГЭ некоторыми отличниками.

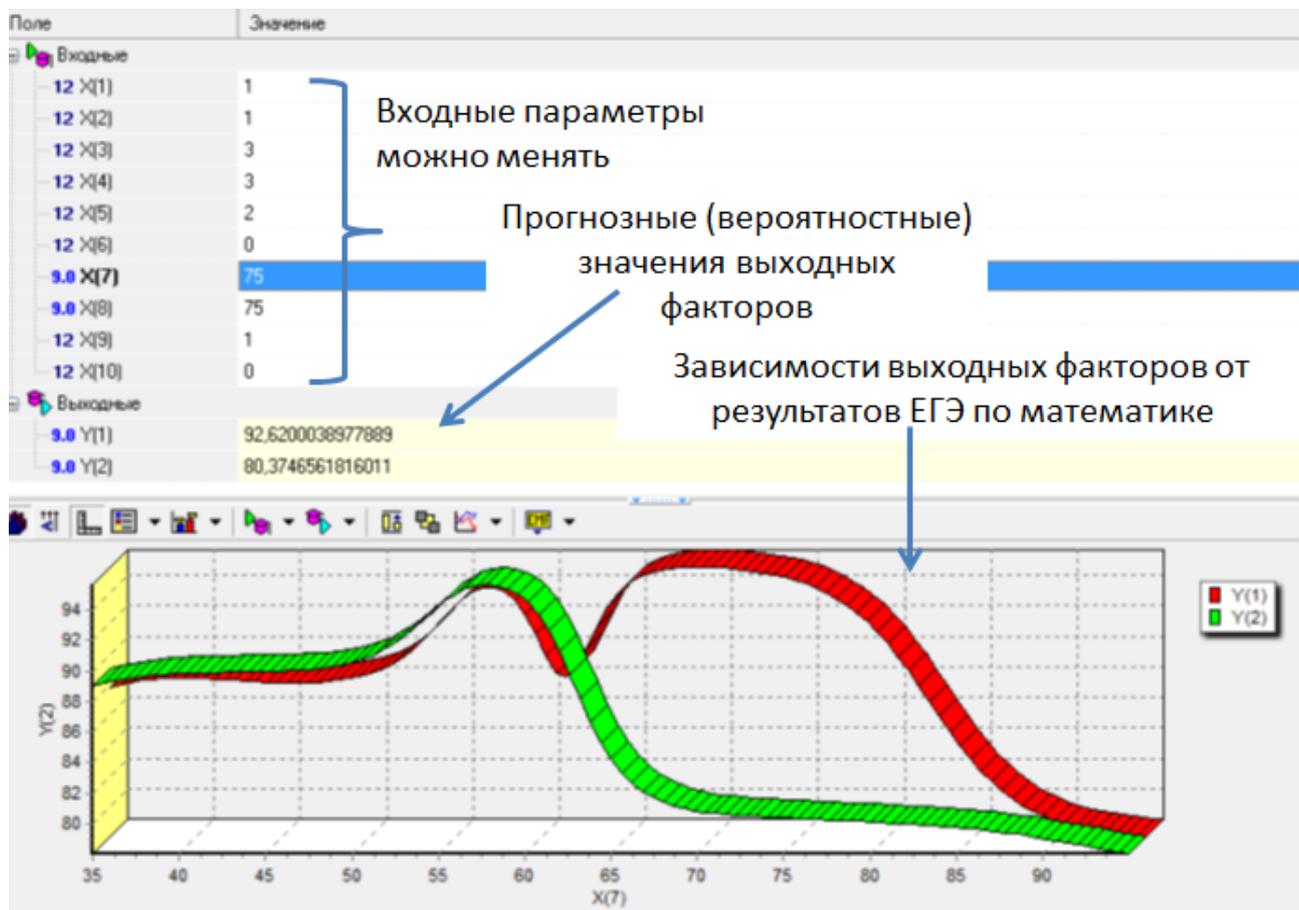


Рис. 4. Окно формы «Что-если» с рассчитанными вероятностными значениями и диаграммой изменения прогноза

В заключение отметим, что расширяя выборку входных данных и диапазон их значений, экспериментируя с параметрами нейросети, можно добиться более точных прогнозных результатов.

Таким образом, интеллектуальная обработка данных с помощью нейросетей позволяет формировать прогнозные вероятностные значения будущей успеваемости абитуриентов в вузе, что может способствовать улучшению качества отбора контингента первокурсников. Результаты прогнозирования могут использоваться приемными комиссиями в качестве информативно-рекомендательных. Кроме того, ППС на базе прогнозных значений может планировать форму и содержание индивидуальной работы со студентами с целью позитивного изменения прогноза.