

# МЕТОДЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ КАК ИНСТРУМЕНТ ПРИ СОЗДАНИИ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ

**Фот Наталия Павловна\***

К.т.н., доцент  
fotnp@mail.ru

**Колиниченко Владимир Васильевич\***

Магистрант  
orenvov@gmail.com

\*ФГБОУ ВО «Оренбургский государственный университет»,  
г. Оренбург, Россия

## Аннотация:

В работе проведен анализ существующих рекомендательных образовательных систем, представлены предложения по применению методов математического моделирования, разработана анкета, включающая расширенный перечень предъявляемых к преподавателю характеристик. Построены «логит-модели», позволившие выявить наиболее важные признаки для формирования успешной пары «учитель- ученик».

## Ключевые слова:

математическое моделирование, ученик, преподаватель, рекомендательная система, онлайн образовательная платформа

---

**УДК** 519.85

**Для цитирования:** Фот Н.П., Колиниченко В.В. Методы математического моделирования как инструмент при создании рекомендательных систем / Н.П. Фот, В.В. Колиниченко // Контентус. – 2023. – № 75. – Т.8. – С. 18 – 27.

---

Цифровые сервисы сегодня являются одним из главных инструментов маркетинга. Их стремительное развитие вывело на первый план необходимость создания рекомендательных систем, главной задачей которых является предоставление информации о товаре (услуге). При этом, формирование прибыли идет за счет предоставления данных информационных услуг.

Современные рекомендательные системы, в зависимости от целей и задач бизнеса, могут быть его основой или дополнительным сервисом. Однако, рост количества данных сервисов повлекли за собой

трудоемкость для рядовых пользователей выбора из этого многообразия той системы, которая максимально учитывала бы его потребности. Особенно это касается рынка образовательных услуг.

При выборе, предположим, курса или педагога, в рекомендательных системах в качестве источников рекомендации выступают схожие пользователи, экспертные оценки, рейтинг. При этом, допускается таргетинг по региону, но не учитываются личные предпочтения, такие как хобби, любимая музыка, фильм, которые, в свою очередь, могут стать основой для формирования критерия по выбору наиболее подходящего сервиса с учетом индивидуальных личных особенностей. Именно определенный набор личностных характеристик и типов поведения преподавателей оказывает существенное влияние на успех в обучении ученика [4, 5,7].

С другой стороны, для коммерческого успеха, онлайн платформе важно удерживать клиента, а для этого необходимо сформировать стабильную пару «учитель-ученик». Реализованные алгоритмы поиска «учителя» «учеником» базируются на фильтрах, заложенных сайтом. К примеру, на сайтах по изучению иностранного языка чаще всего используются следующие фильтры:

- носитель языка;
- дополнительные языки, которыми владеет преподаватель;
- стоимость часа обучения.

В результатах поиска, в этом случае, первыми выводятся преподаватели, оплатившие продвижение своих анкет. В связи с этим, целесообразным представляется добавление собственной оценки «учителя», выделяя ключевые слова из предоставленной им анкеты и добавляя новые фильтры для «ученика».

Кроме того, перспективным является добавление истории просмотра «учеником» для формирования дополнительных рекомендаций, а также информации о количестве уроков, уже проведенных «учителем» и продолжительность обучения каждого «ученика», если рассматривать в контексте персонального обучения.

Одна из важных составляющих качества обучающей рекомендательной системы – это оценка прогресса в обучении у «ученика», то есть оценивание его мотивации к обучению. На современных платформах по подбору репетитора, «ученик» сам формирует свой график, а учитель формирует контент для ученика. Поэтому, возникает необходимость выявления «слабых» сторон у «ученика» и на основе уже текущего уровня знаний, составление оптимального плана занятий на будущее время.

К широко распространённым рекомендательным алгоритмам относятся алгоритмы Summary-based (неперсональные), Collaborative Filtering (коллаборативная фильтрация), Matrix Factorization [1]. Данные

системы построены на основании кластеризации пользователей по оценкам пользователя и имеют проблему «холодного старта», т.е. когда ещё не накоплено достаточное количество данных для корректной работы рекомендательной системы [3,8,9]. Поэтому, для их применения в данной области образования, необходимо чтобы каждый ученик предварительно оценил определенное количество учителей для формирования на их базе соответствующих кластеров. Но логика взаимоотношений между преподавателем и учеником на онлайн платформах не подразумевает большого количества контактов для ученика с разными преподавателями и, в конечном итоге, так и не найдя нужного преподавателя, пользователь может покинуть онлайн платформу.

Именно поэтому для составления оптимальной пары «ученик-учитель» целесообразно применение методов прогнозирования [6,10], а учитывая, что в качестве критерия выступает вероятность создания удачной пары, наиболее предпочтительным в данном случае является построение бинарной логистической регрессионной модели.

Для выделения факторов, влияющих на успешность заключения пары «ученик-учитель» была составлена анкета (фрагмент анкеты для оценки учеником личных и учительских качеств приведены на рисунках 1 и 2) и произведено анкетирование 70 учеников средних школ в г. Оренбург через гугл-формы. В анкете ученику предлагалось пройти небольшой опрос и дать собственную оценку и оценку двух преподавателей своей школы. При этом, ученик должен был сам выбрать этих преподавателей из условия, что с одним из преподавателей у ученика «заключена» успешная пара и опрашиваемый связывает свой академический успех с деятельностью этого преподавателя, вторым преподавателем ученик выбирает учителя противоположным первым – выбирая учителя, с которым ученик связывает отсутствие академического успеха. Анкетирование и выбор оцениваемых преподавателей происходит полностью анонимно с целью исключения психологического давления на анкетированного ученика.

Анкета составлена таким образом, чтобы на основании ее анализа можно было произвести оценку пары «учитель-ученик» по критериям:

- профессиональные навыки;
- наличие общих интересов;
- психологические характеристики (тип темперамента, межличностное поведение);
- физиологические (пол, возраст).

1. Ваш пол?  
 Мужчина  
 Женщина
2. Оцените свое поведение  
 Я интроверт  
 Я экстраверт
3. Как Вы лучше усваиваете материал?  
 Я лучше усваиваю материал дома при самостоятельном изучении материала  
 Для усвоения материала мне достаточно информации, разобранный с преподавателем во время урока  
 Сначала должен объяснить учитель, а усвоение материала происходит при самостоятельном разборе
4. Какие качества преподавателя в большей степени влияет на качество обучения?  
 Личные качества (характер, поведение и т.п.)  
 Профессиональные качества (знание материала)  
 Личные и профессиональные

**Рисунок 1** – Фрагмент анкеты для оценки учеником личных качеств

5. Возраст преподавателя?  
 20 - 40  
 40 - 60  
 60 +
6. Пол преподавателя?  
 Мужской  
 Женский
7. Как Вы оцениваете профессиональные навыки преподавателя в его предмете?  
 Мастерски владеет всеми аспектами своего предмета и имеет знаниями за пределами изучаемой программы  
 Обладает достаточными знаниями и не обладает знаниями за пределами изучаемой программы  
 Не обладает достаточными знаниями
8. Данный преподаватель может поддержать разговор с Вами на темы отличные от учебной деятельности?  
 Да  
 Нет
9. Преподаватель оказывает Вам индивидуальное внимание во время занятий?  
 Да, внимание положительное  
 Да, внимание негативное  
 Нет
10. Имеется ли контакт с данным преподавателем вне учебного процесса?  
 Да  
 Нет
11. Проявляете ли Вы симпатию к преподавателю?  
 Да  
 Нет

**Рисунок 2** – Фрагмент анкеты для оценки учеником качеств учителя

Для дальнейшего анализа на основе результатов анкетирования для пар «учитель-ученик» выбраны следующие показатели:  $s\_t\_comp_i$  – значение результирующей переменной, где значение 1 соответствует успешной паре “учитель-ученик”, 0 соответствует не успешной паре “учитель - ученик”,  $s\_gender$  – пол ученика (0 - женщина, 1 - мужчина),  $s\_p\_type$  – психотип ученика (0 – интроверт, 1 – экстраверт),  $s\_ed\_type$  – тип лучшего усвоения материала учеником (0 – без преподавателя, 1 – самостоятельно после прохождения материала с преподавателем, 3 – только с преподавателем),  $s\_ness\_teacher$  – предпочтения ученика в качествах преподавателя (0 – профессиональные навыки, 1 – личные и

профессиональные, 2 – личные),  $t\_gender$  – пол преподавателя (0 – женщина, 1 – мужчина),  $t\_know$  – уровень профессиональных знаний преподавателя (0 – не достаточные знания, 1 – достаточные знания, 2 – знания, намного превышающие изучаемый материал),  $t\_small\_talk$  – способность преподавателя поддержать беседу на отстранённые темы (0 – не способен, 1 – способен),  $t\_atten$  – тип внимания, оказываемого учителем ученику (0 – внимание негативное, 1 – отсутствует внимание, 2 – внимание положительное),  $t\_contact$  – наличие контакта между и учителем и учеником вне учебного процесса (0 – отсутствует, 1 – присутствует),  $s\_teacher\_symp$  – наличие симпатии ученика к учителю (0 – отсутствует, 1 – присутствует),  $t\_age$  – возраст преподавателя (1 – 20/40 лет, 2 – 40/60 лет, 3 – 60+ лет). Все показатели являются качественными и для показателей, которые могут принимать более двух значений были введены фиктивные переменные.

Построена математическая модель, где в качестве результативной переменной выступила бинарная переменная  $s\_t\_comp$ , на основании которой произведен анализ факторов, влияющих на успешность подбора учителя для ученика с целью заключения успешной пары.

В общем виде бинарная модель подбора учителя имеет следующий вид:

$$s\_t\_comp_i^* = \beta_0 + t\_gender * \beta_1 + Dt\_know\_0 * \beta_2 + Dt\_know\_1 * \beta_3 + t\_small\_talk * \beta_4 + Dt\_atten\_0 * \beta_5 + Dt\_atten\_1 * \beta_6 + t\_contact * \beta_7 + Dt\_age\_0 * \beta_8 + Dt\_age\_1 * \beta_9 + z_i, \quad i = 1, \dots, n$$

$$s\_t\_comp_i = \begin{cases} 0, & \text{если } \Lambda(s\_t\_comp_i^*) \leq 0 \\ 1, & \text{если } \Lambda(s\_t\_comp_i^*) > 0 \end{cases}$$

где  $s\_t\_comp_i^*$  – значение латентной переменной для  $i$ -го наблюдения,

$\Lambda(s) = \frac{e^s}{1+e^s}$  – функция логистического распределения,

$Dt\_know\_0, Dt\_know\_1$  – бинарные фиктивные переменные для  $t\_know$  со значением 0 и 1 соответственно,

$Dt\_atten\_0, Dt\_atten\_1$  – бинарные фиктивные переменные для  $t\_atten$  со значением 0 и 1 соответственно,

$Dt\_age\_0, Dt\_age\_1$  – бинарные фиктивные переменные для  $t\_age$  со значением 0 и 1 соответственно,

$\beta_0, \beta_1, \dots, \beta_9$  – неизвестные коэффициенты, характеризующих влияние наблюдаемых характеристик на успешность,

$z_i$  – ошибка, характеризующая влияние всех неучтенных в модели факторов, влияющих на конечное наблюдение,

$n$  – количество объектов наблюдений в выборке.

Далее была оценена модель подбора учителя для ученика по всей выборке, игнорируя индивидуальные характеристики каждого ученика. Результаты оценки приведены в Таблице 1.

**Таблица 1** – Результаты расчета логистической регрессионной модели подбора (количество элементов выборки  $n = 140$ )

	Коэффициент	Ст. ошибка	z	p-значение	
const	22.3001	5200.49	0.004288	0.9966	
t_gender	0.126260	0.561238	0.2250	0.8220	
t_small_talk	0.905264	0.604546	1.497	0.1343	
t_contact	0.907668	0.559689	1.622	0.1049	
Dt_age_0	-0.385174	0.686216	-0.5613	0.5746	
Dt_age_1	0.0387444	0.708939	0.05465	0.9564	
Dt_know_0	-23.4880	5200.49	-0.004516	0.9964	
Dt_know_1	-21.5576	5200.49	-0.004145	0.9967	
Dt_atten_0	-3.10947	1.01487	-3.064	0.0022	***
Dt_atten_1	-2.49570	0.775371	-3.219	0.0013	***

Количество «корректно предсказанных» случаев составило 73.6%, что позволяет сделать вывод о корректности построенной модели. Анализ коэффициентов показал, что важным показателем при формировании успешной пары является внимание преподавателя к ученику (*Dt\_atten\_0*, *Dt\_atten\_1*).

Построенная модель была пересчитана с учетом индивидуальных параметров учеников, для чего был произведен расчет с данными, удовлетворяющими условиям индивидуальных параметров учеников. Сформированы выборки, удовлетворяющие индивидуальным параметрам учеников, которые ограничены условиями  $s\_p\_type = 0$ ,  $s\_p\_type = 1$ ,  $s\_ed\_type = 0$ ,  $s\_ed\_type = 1$ ,  $s\_ed\_type = 2$ . Результаты расчета моделей для сформированных выборок представлены в таблицах 2, 3, 4, 5, 6.

**Таблица 2** – Результаты расчета логистической регрессионной модели подбора для учеников интровертов ( $s\_p\_type = 0$ , количество элементов выборки  $n = 88$ )

	Коэффициент	Ст. ошибка	z	p-значение	
const	21.7762	6360.91	0.003423	0.9973	
t_gender	0.648984	0.668840	0.9703	0.3319	
t_small_talk	0.646788	0.715929	0.9034	0.3663	
t_contact	0.930711	0.645310	1.442	0.1492	
Dt_age_0	-0.328719	0.856771	-0.3837	0.7012	
Dt_age_1	0.245528	0.816205	0.3008	0.7636	
Dt_know_0	-23.0819	6360.91	-0.003629	0.9971	
Dt_know_1	-21.0527	6360.91	-0.003310	0.9974	
Dt_atten_0	-3.15198	1.16467	-2.706	0.0068	***
Dt_atten_1	-2.10090	0.868745	-2.418	0.0156	**

Количество «корректно предсказанных» случаев для модели, представленной в таблице 2 составило 73.6%, к значимым показателям отнесен коэффициент *Dt\_atten\_0*. Влияющим на формирование успешной пары для ученика интроверта, является показатель - негативное внимание со стороны преподавателя.

В результате расчета логистической регрессионной модели подбора для учеников экстравертов (таблица 3), количество «корректно предсказанных» случаев составило 84.6%, к значимым показателям отнесен коэффициент *Dt\_know\_1*. Показателем, влияющим на формирование успешной пары для ученика экстраверта, является наличие необходимых знаний у преподавателя.

**Таблица 3** – Результаты расчета логистической регрессионной модели подбора для учеников экстравертов (*s\_p\_type* = 1, количество элементов выборки *n* = 52)

	Коэффициент	Ст. ошибка	z	p-значение	
const	2.79988	1.70745	1.640	0.1010	
t_gender	-0.926575	0.754313	-1.228	0.2193	
t_small_talk	0.582458	0.755414	0.7710	0.4407	
t_contact	1.31898	0.835322	1.579	0.1143	
Dt_age_0	-0.954262	0.963394	-0.9905	0.3219	
Dt_age_1	-0.341522	0.903373	-0.3781	0.7054	
Dt_know_1	-2.55214	0.871231	-2.929	0.0034	***
Dt_atten_0	-1.75932	1.54113	-1.142	0.2536	
Dt_atten_1	-2.33166	1.34247	-1.737	0.0824	*

**Таблица 4** – Результаты расчета логистической регрессионной модели подбора для учеников эффективнее обучающихся самостоятельно (*s\_ed\_type* = 0, количество элементов выборки *n* = 32)

	Коэффициент	Ст. ошибка	z	p-значение	
const	20.6684	5652.01	0.003657	0.9971	
t_gender	0.319022	1.02249	0.3120	0.7550	
t_small_talk	0.337871	0.972426	0.3475	0.7283	
t_contact	-0.0984887	0.998524	-0.09863	0.9214	
Dt_age_0	-0.413540	1.27000	-0.3256	0.7447	
Dt_age_1	0.336975	1.11644	0.3018	0.7628	
Dt_know_0	-20.4404	5652.01	-0.003616	0.9971	
Dt_know_1	-19.3271	5652.01	-0.003420	0.9973	
Dt_atten_0	-2.68354	1.94214	-1.382	0.1671	
Dt_atten_1	-1.78527	1.54169	-1.158	0.2469	

Рассчитанная модель для учеников, которые предпочитают обучаться самостоятельно (таблица 4), не имеет значимых коэффициентов.

Рассчитанная модель для учеников, которые предпочитают обучаться самостоятельно, но после прохождения материала с преподавателем (таблица 5), также не имеет значимых коэффициентов.

**Таблица 5** – Результаты расчета логистической регрессионной модели подбора для учеников эффективнее обучающихся самостоятельно после предварительного прохождения материала с преподавателем ( $s\_ed\_type = 1$ , количество элементов выборки  $n = 42$ )

	Коэффициент	Ст. ошибка	z	p-значение	
const	160.392	58342.3	0.002749	0.9978	
t_gender	-68.0400	15536.6	-0.004379	0.9965	
t_small_talk	51.6339	12091.6	0.004270	0.9966	
t_contact	69.7732	15536.6	0.004491	0.9964	
Dt_age_0	-35.6487	13731.8	-0.002596	0.9979	
Dt_age_1	-87.2826	19955.5	-0.004374	0.9965	
Dt_know_0	-178.308	59874.8	-0.002978	0.9976	
Dt_know_1	-90.5311	51999.9	-0.001741	0.9986	
Dt_atten_0	-3.51739	45773.3	-7.684e-05	0.9999	
Dt_atten_1	-35.0789	9263.91	-0.003787	0.9970	

**Таблица 6** – Результаты расчета логистической регрессионной модели подбора для учеников эффективнее обучающихся самостоятельно с преподавателем ( $s\_ed\_type = 1$ , количество элементов выборки  $n = 66$ ).

	Коэффициент	Ст. ошибка	z	p-значение	
const	20.2457	6029.06	0.003358	0.9973	
t_gender	-0.312716	0.788031	-0.3968	0.6915	
t_small_talk	1.70710	0.878222	1.944	0.0519	*
t_contact	1.69726	0.904978	1.875	0.0607	*
Dt_age_0	-0.877315	0.954900	-0.9188	0.3582	
Dt_age_1	-0.417162	0.959884	-0.4346	0.6639	
Dt_know_1	-3.90915	1.04011	-3.758	0.0002	***
Dt_atten_0	-19.7620	6029.06	-0.003278	0.9974	
Dt_atten_1	-20.2240	6029.06	-0.003354	0.9973	

Количество «корректно предсказанных» случаев в модели для учеников, эффективнее обучающихся самостоятельно с преподавателем (таблица 6) составило 84.8%. К значимым показателям отнесены коэффициенты  $Dt\_know\_1$ ,  $t\_small\_talk$  и  $t\_contact$ . Наличие необходимых знаний у преподавателя, наличие контакта с преподавателем вне учебного процесса и возможность преподавателя поддержать разговор на не учебные темы влияют на формирование успешной пары для ученика, предпочитающего обучаться с преподавателем.

Анализ построенных моделей показал, что для моделей, рассчитанных с учетом индивидуальных параметров учеников, набор значимых факторов отличается от набора значимых факторов модели без учета этих характеристик. Также модели, построенные с учетом индивидуальных характеристик учеников, обладают более высокими

показателями точности. Таким образом, применение математического инструментария на основе методов прогнозирования может стать инструментом для повышения эффективности работы рекомендательных систем, учитывающих личные предпочтения клиента.

### Список использованных источников

1. **Ким Ф.** Рекомендательные системы на практике. / Под ред. Д.А. Мочан; – М.: ДМК-Пресс, 2020. – 448с
2. **Николенко С.А.** Рекомендательные системы. – СПб.: Изд-во Центр Речевых Технологий, 2012. – 52с.
3. **Собкин В.С.** Влияние отношений между учителем и учеником на академические достижения учащихся / В.С. Собкин, А.С. Фомиченко // Управление образованием: теория и практика. – 2015. – № 3. – С. 34-54. – ISSN 2311-2174
4. **Цыбов Н.Н.** Факторы, влияющие на эффективность процесса обучения в технических вузах // Бюллетень науки и практики. – 2019. – Т.5 №7. – С. 345- 357
5. **Кошевой О.С.** Модель логистической регрессии для прогнозирования использования населением портала государственных услуг // Государственное управление. Электронный вестник, Выпуск № 86. Июнь 2021. – С. 42-56.
6. Методы и модели эконометрики [Электронный ресурс] : учебное пособие для студентов, обучающихся по программам высшего образования по направлениям подготовки 01.03.04 Прикладная математика, 38.04.01 Экономика, 38.03.05 Бизнес-информатика / под ред. А. Г. Реннера; М-во образования и науки Рос. Федерации, Федер. гос. бюджет. образоват. учреждение высш. проф. образования "Оренбург. гос. ун-т". - Ч. 2. Эконометрика пространственных данных. - Электрон. текстовые дан. (1 файл: 10.59 Мб). - Оренбург : ОГУ, 2015. - 434 с.
7. **Таратухина Ю.В.** Машинное обучение модели информационной рекомендательной системы по вопросам индивидуализации образования / Ю.В. Таратухина, Т.В. Барт, В.В. Власов // Образовательные ресурсы и технологии. 2019. No 2 (27). – С. 7-14.
8. **Андрянов Н.А., Атаходжаева М.Р., Бородин Е.И.** Математическое моделирование рекомендательной системы и обработка данных телекоммуникационной компании с помощью моделей машинного обучения // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника», 2022. Т. 22, No 2. С. 17–28.
9. **Афанасьев В.В., Благий В.А., Воробьев А.А.** Алгоритм перераспределения неопределившихся респондентов на основе мультиномиальной логистической регрессии // Вестник Евразийской науки, 2019 No3, URL: <https://esj.today/PDF/30ITVN319.pdf> (доступ свободный)
10. **Сорокин А.С.** К вопросу валидации модели логистической регрессии в кредитном скоринге// Интернет-журнал «НАУКОВЕДЕНИЕ» Выпуск 2, март – апрель 2014.

# METHODS OF MATHEMATICAL MODELING AS A TOOL FOR CREATING RECOMMENDATORY SYSTEMS

**Phot Natalia Pavlovna\*\***

Ph. D, Associate Professor



**Kalinichenko Vladimir Vasilyevich\*\***

Student



\*\*Orenburg State University,  
Orenburg, Russia

## **Abstract:**

The article analyzes existing recommendatory educational systems, presents proposals for the application of methods of mathematical modeling, develops a questionnaire, including a more extended list of characteristics submitted to the teacher. «Logit-models» were built, which allowed to identify the most important signs for the formation of a successful pair «teacher-student».

## **Keywords:**

mathematical modeling, student, teacher, recommendatory system, online educational platform