

***АНАЛИЗ КЛИЕНТСКОЙ БАЗЫ КОМПАНИИ ПАО
«РОСТЕЛЕКОМ» НА ОСНОВЕ МОДЕЛИ ЛИНЕЙНОЙ РЕГРЕССИИ С
ГЕОГРАФИЧЕСКИ ВЗВЕШЕННЫМИ КОЭФФИЦИЕНТАМИ***

Танкова А.В.,

Магистрант

ФГБОУ ВПО «Оренбургский государственный университет»

Оренбург, Россия

Бантикова О.И.

Кандидат экономических наук, доцент, доцент кафедры математических методов и моделей в экономике

ФГБОУ ВПО «Оренбургский государственный университет»

Оренбург, Россия

Аннотация: В статье реализован подход к анализу и моделированию клиентской базы компании ПАО «Ростелеком». Выявлены факторы, влияющие на длительность пользования услугами компании с учетом влияния местоположения абонентов. Предложены рекомендации по применению полученных моделей.

Ключевые слова: географические координаты, лояльность абонентов, длительность пользования услугами, контурная диаграмма.

***ANALYSIS OF CLIENT BASE OF THE COMPANY PJSC
“ROSTELECOM” ON THE BASIS OF A LINEAR REGRESSION MODEL
WITH GEOGRAPHICALLY WEIGHTED COEFFICIENTS***

Tankova A. V.

Graduate student

FGBOU VPO "Orenburg state University"

Orenburg, Russia

Bantikova O.I.

Candidate of economic Sciences, associate Professor, Department of mathematical methods and models in economy

FGBOU VPO "Orenburg state University"

Orenburg, Russia

Abstract: the paper implements an approach to the analysis and modeling of the customer base of the company PJSC "Rostelecom". The factors influencing the duration of use, taking into account the influence of the location of subscribers. Recommendations on the application of the obtained models.

Keywords: geographic coordinates, customer loyalty, duration of use, contour plot.

Для компании ПАО «Ростелеком» показатель удовлетворения потребителей в сфере предоставления услуг является основополагающим фактором. Если клиент уверен, что соотношение цены и качества услуги оптимальное, он с большой вероятностью заключит договор, но существует и ряд других факторов, которые влияют на принятие решения абонентом о подключении услуги и длительности ее пользования. Заключив договор, также ряд факторов будет влиять на длительность нахождения абонента в компании в статусе клиента. Выявление факторов, влияющих на длительность пользования услугами, позволит компании заблаговременно проводить ряд мероприятий по увеличению лояльности клиентов и сохранения их в компании в качестве клиентов. Одним из таких факторов выступают географические координаты, то есть местоположение клиента. Здесь под местоположением понимается фактический адрес клиента, где

была подключена услуга и где в дальнейшем она предоставляется клиенту. Поэтому данный анализ представляет собой важную составляющую, предшествующую формированию политики компании по сохранению и расширению клиентской базы [3,5]. Для проведения данного анализа были выбраны показатели, характеризующие сотрудничество компании ПАО «Ростелеком» с физическими лицами:

y - длительность пользования услугами, дни;

x_1 - сумма оплаты за весь период, руб.;

x_2 - количество подключенных услуг, ед.;

x_3 - сумма задолженности клиента, руб.;

$$x_4 = \begin{cases} 1, & \text{если договор расторгнут} \\ 0, & \text{если договор не расторгнут.} \end{cases}$$

$$x_5 = \begin{cases} 1, & \text{если клиент мужчина} \\ 0, & \text{если клиент женщина.} \end{cases}$$

Данные о характеристиках каждого объекта дополнены географическими координатами:

u_lat - северная широта (latitude), градусов;

v_long - восточная долгота (longitude), градусов.

С помощью модели линейной регрессии с географически взвешенными коэффициентами проанализировано влияние местоположения объектов на зависимость длительности пользования клиентами услуг компании от перечисленных факторов. С помощью оценки глобальной модели (1), представленной ниже, был сделан ряд выводов

$$\hat{y} = \underset{(286,61)}{2857,65}(u_i, v_i) + \underset{(0,01)}{0,02}(u_i, v_i) \cdot x_1 - \underset{(206,85)}{1013,96}(u_i, v_i) \cdot x_2 - \underset{(0,14)}{0,33}(u_i, v_i) \cdot x_3 - \underset{(203,9)}{416,07}(u_i, v_i) \cdot x_4 + \underset{(151,99)}{250,3}(u_i, v_i) \cdot x_5 \quad (1)$$

Все коэффициенты модели значимо отличаются от нуля, их знаки свидетельствуют о том, что: увеличение суммы оплаты за весь период и заключение договора об использовании услуг с мужчинами увеличивают

среднюю длительность пользования услугами, а количество подключенных услуг, задолженность клиента и расторжение договора уменьшают ее. В данной модели максимальное расстояние, между вошедшими в выборку объектами с запада на восток, составило 11,73 км, а с севера на юг – 14,99км.

Также был сделан вывод о том, что гипотеза о незначимости линейной модели географически взвешенной регрессии отвергается [1,2].

На основе результатов описательной статистики (Таблица 1), осуществленной по каждому локальному коэффициенту, получен вывод о влиянии выбранных факторов на длительность пользования услугами. В таблице 1 приведены оценки среднего, среднеквадратического отклонения, минимального и максимального значений размаха квантилей уровня 0,25, 0,50 и 0,75, межквартильного размаха.

Таблица 1 – Результаты описательной статистики модели линейной регрессии с географически взвешенными коэффициентами

Variable	Mean	STD	
Intercept	2726,519661	673,906572	
x1	0,016638	0,002608	
x2	-917,935765	351,842209	
x3	-0,320249	0,053336	
x4	-443,323481	166,072680	
x5	195,100539	85,446290	
Variable	Min	Max	Range
Intercept	1466,812179	3756,280447	2289,468267
x1	0,009273	0,021014	0,011741
x2	-1566,528736	-297,231958	1269,296778
x3	-0,411881	-0,240100	0,171782
x4	-862,435400	-201,193071	661,242329
x5	25,563911	331,767458	306,203548
Variable	Lwr Quartile	Median	Upr Quartile
Intercept	2224,826178	2791,924789	3257,156428
x1	0,015603	0,017266	0,018301
x2	-1186,492038	-906,016765	-616,749368
x3	-0,366099	-0,306074	-0,273167
x4	-520,304420	-452,613895	-284,815006
x5	125,678408	200,878751	269,152582

Так, например, среднее для коэффициента при переменной «сумма задолженности клиента» составляет минус 0,32 – это означает, что в среднем длительность пользования услугами при увеличении задолженности на 100рублей при прочих равных условиях сократится на 32 дня. Вместе с тем, для некоторых клиентов такой фактор снижает длительность нахождения в компании в качестве клиента на 41 день, а для некоторых – только на 24 дня. При прочих равных условиях для 25% рассматриваемых клиентов понижение длительности пользования услугами составит в среднем не менее чем 37 дней, для 50% - в среднем не менее чем 31 день, для 75% - в среднем не менее чем 27 дней.

Проведен сравнительный анализ между классической моделью регрессии и географически взвешенной регрессией, и принято решение в пользу модели ГВР. Гипотеза о совпадении коэффициентов географически взвешенной регрессии с коэффициентами классической регрессионной модели отвергается и делается вывод об адекватности применения модели ГВР.

На основе анализа оценок коэффициентов при всех регрессорах, построенных таблиц значений и контурных диаграмм, были сделаны выводы о влиянии каждого фактора на длительность пользования услугами в каждой части города Оренбург. Были выявлены районы, которые наиболее подвержены влиянию факторов, и районы, где влияние факторов незначительно [4].

К примеру, для оценки константы, которую в рассматриваемой модели можно интерпретировать как «базовую» длительность пользования услугами, которая будет увеличиваться уже в зависимости от характеристик услуги, результаты представлены в таблице 2 и на рисунке 1.

Таблица 2 – Значения оценок константы (intercept) модели линейной регрессии с географически взвешенными коэффициентами

	U														
	1,7 1	1,7 2	1,7 3	1,7 5	1,7 6	1,7 7	1,7 8	1,7 9	1,8	1,8 1	1,8 2	1,8 3	1,8 4	1,8 5	
5,05										815					
5,06							469								
5,08										901					
5,09					757		329		708		051				
5,1			828	655	587	615				181	165				
5,11	264	150			466	551									
5,12					662								670		
5,13						892		885				756		610	
5,14						992		921	090		530	674	693	641	
5,15							604	745		511		591			
5,16								809				299	579		
5,17						453	587			242	122	171			
5,18						615	717			257					
5,19						721	727								
5,22								013							

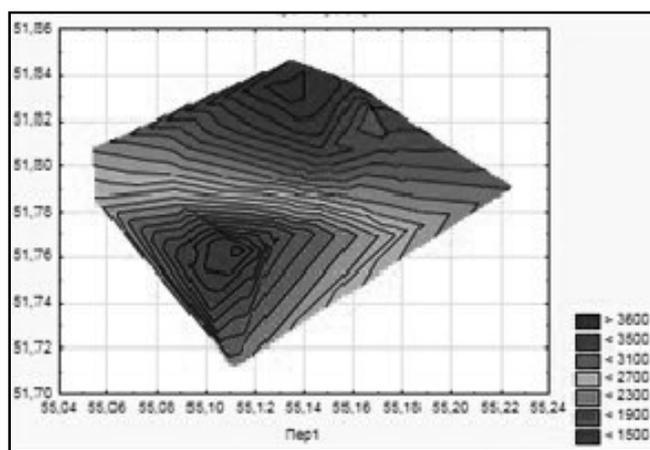


Рис. 1 – Контурная диаграмма для оценок константы (intercept)

Самая высокая «длительность пользования» около 3700 дней, также высокие показатели характерны для восточной части города (выделены в таблице темно-серым цветом). Более низкие значения отмечены в таблице светло-серым цветом. При движении с востока на запад данный показатель снижается до 2150 в западной части города. Самые низкие значения в таблице цветом не отмечены и более характерны для западно-центральной части города. При интерпретации графика следует обратить внимание, что темно-зеленым цветом на нем отображены зоны, по которым не было исходных данных.

Для регрессора «сумма оплаты за весь период» Зона наиболее высоких значений показателя расположена в квадратах ($U=51,8, V=55,09$; $U=51,81, V=55,08$ и $55,1$; $U=51,82, V=55,1$; $U=51,83, V=55,13$; $U=51,84, V=55,12$ и $55,14$; $U=51,85, V=55,13$ и $55,14$) – здесь при прочих равных условиях увеличение суммы оплаты на 100 руб. приводит к увеличению длительности пользования услугами в среднем на 2 дня. Чуть севернее (квадраты $U=51,82, V=55,14$; $U=51,83, V=55,14, 55,15, 55,16, 55,17$; $U=51,84, V=55,16$) аналогичное увеличение суммы оплаты за пользование услугами приводит к увеличению длительности пользования услугами на 18 дней. Дальнейшее движение от описанной зоны ведет к уменьшению продолжительности

нахождения клиента в компании. К наиболее малому увеличению продолжительности пользования услугами приведет увеличение суммы оплаты на 100 рублей в северной части города (U от 51,79; $V=55,22$) и западнее – менее 12 дней за каждый дополнительный рубль.

Для регрессора «количество подключенных услуг» данный показатель отрицательно влияет на результативный признак. То есть максимальное сокращение длительности пользования услугами при подключении каждой дополнительной услуги характерно для восточной части города. К примеру, для квадрата $U=51,85$; $V=55,14$, подключение еще одной услуги абоненту приведет к сокращению длительности пользования услугами на 1500 дней. Самые минимальные значения в таблице отмечены светло-голубым цветом, и характерны для юго-западной части города.

Для регрессора «сумма задолженности клиента» данный показатель также отрицательно влияет на результативный признак, то есть с увеличением задолженности клиента продолжительность пользования услугами сокращается. Максимальное сокращение длительности пользования услугами при увеличении задолженности у клиента характерно для восточной части города. Двигаясь западнее от квадратов, закрашенных в таблице темно-серым цветом, свидетельствующих о максимальном сокращении длительности пользования услугами при увеличении задолженности, данные значения сокращаются. И, к примеру, в квадрате $U=51,71$; $V=55,11$ достигает минимального значения. Здесь увеличение задолженности клиента на 100рублей сократит его пребывание в компании на 21 день. Для центральной части характерны значения близкие к минимальному. Северная и южная части города интервал изменения значений схож, от 29 до 34.

Для регрессора «договор» данный показатель также отрицательно влияет на результативный признак, то есть если клиент решил расторгнуть договор, то количество дней пользования услугами сокращается. Либо, если компания, решила отказать клиенту доступ к пользованию услугами, то

длительность пользования также сокращается. Максимальные значения сокращения результативного признака характерны для северной части города, а западная и центральная части города характеризуются средними значениями показателя. Южная и восточная части города менее чувствительными к изменению коэффициентов при данном регрессоре.

Для регрессора «пол» зона наибольшего увеличения продолжительности пользования услугами вследствие заключения договора с мужчинами находится в квадратах центрально-южной части города, а движение на запад, север, и восток, характеризуются снижением длительности пользования услугами. Самые минимальные значения можно обнаружить в квадратах восточной части города.

Можно отметить, что влияние объясняющих переменных на продолжительность пользования услугами наиболее чувствительно в северной и восточной части города. Так как ее максимальное изменение наблюдается именно в этих частях города по таким показателям как: оценки константы, сумма оплаты за весь период, количество подключенных услуг, сумма задолженности клиента, договор. И только влияние последнего показателя (пол) минимально влияет на северную и восточную части города, имея максимальные значения в центрально-южной части города.

На основе проведенного исследования можно сделать вывод о том, что анализ влияния местоположения клиентов компании на длительность пользования услугами, позволит компании осуществить мероприятия по улучшению сотрудничества в выделенных районах и для выделенных групп лиц по половому признаку. Среди таких мероприятий можно выделить: разработку оптимальных пакетов услуг, формирование привлекательной тарифной линейки, проведение акций и скидок.

Библиографический список:

1. Балаш, В.А. Особенности построения географически взвешенной регрессии для моделирования рынка недвижимости / В.А. Балаш, О.С. Балаш, А.В. Харламов // Вестник Саратовского государственного социально-экономического университета. – 2008. №5. – С. 125–127.
2. Седова, Е.Н. Моделирование стоимости вторичного жилья на региональном рынке: пространственный подход / Е.Н. Седова // Вестник ОГУ. – 2014. №14(175). – С. 458–464.
3. Скляр, Е.Н. Формирование и реализация программ потребительской лояльности для ОАО «Ростелеком» / Е.Н. Скляр, П.В. Яшкина // Маркетинг и маркетинговые исследования. – 2014. №2. – С. 126–138.
4. Стебунова, О.И. Линейная модель географически взвешенной регрессии : методические указания / О.И. Стебунова, Е.Н. Седова // Оренбургский гос. ун-т. – Оренбург: ОГУ, 2015. – 65с.
5. Хайрулина, Л.Р. Реализация корпоративной концепции сервисного поведения сотрудников офисов продаж ПАО «Ростелеком» / Л.Р. Хайрулина // Двадцать вторые апрельские экономические чтения. Материалы международной научно-практической конференции – 2016. – С. 401–406.