Прогнозирование оттока клиентов B2B/B2G на основании бинарных моделей

Важным направлением деятельности в ПАО Ростелеком является работа по удержанию корпоративных клиентов. Мероприятия по удержанию корпоративных клиентов применяются в отношении клиентов, имеющих явное или скрытое намерение по собственной инициативе прекратить пользоваться услугами, оказываемыми ПАО «Ростелеком».

Одной из задач ПАО Ростелеком является снижение оттока корпоративных клиентов, что непосредственно влияет на финансово-экономические результаты деятельности компании.

Одним из основных направлений работы по сокращению оттока являются преактивные мероприятия по предотвращению оттока в отношении корпоративных клиентов, склонных к оттоку по определенным признакам, указывающих на существование угрозы оттока, которые используются для формирования выборки корпоративных клиентов, по которым необходимо проведение профилактических мероприятий.

Одной из основных задач в рамках удержания корпоративных клиентов является прогнозирования оттока для качественного проведения преактивных мероприятий необходимо корректно выделить предикторы оттока – признаки, указывающие на существование угрозы оттока.

Выделенные предикторы используют для формирования выборки корпоративных клиентов, по которым необходимо проведение профилактических мероприятий, поэтому важно определить факторы, влияющие на отток для последующего его прогнозирования и проведения эффективной работы по их удержанию. Прогнозирование оттока дает возможность спрогнозировать такие показатели как абонентская база, выручка и потерянный доход.

В настоящий момент ПАО Ростелеком выделяет предикторы, однако нет информации о зависимости этих факторов на отток, а также ПАО Ростелеком выделяет только количественные предикторы. Поэтому возникает необходимость в построении качественной модели прогнозирования клиентов, которые рискую оказаться в оттоке с учетом качественных факторов, которая даст не только прогноз оттока клиентов, но и информацию о зависимости выделенных факторов влияния на отток, с целью выделения проблемных зон в обслуживании, сервисе и т.д.

Целью данной работы является прогноз количественного показателя оттока корпоративных клиентов ПАО Ростелеком и определение степени влияния факторов на отток корпоративных клиентов с целью их устранения.

Для достижения цели были поставлены следующие задачи:

– рассмотреть существующую процедуру по работе с оттоком в ПАО Ростелеком, и его прогнозом;

– выделить факторы, влияющие на отток, сформировать гипотезы об их влиянии на отток;

– собрать необходимые данные для построения модели прогнозирования оттока корпоративных клиентов сегмента B2B/B2G;

– построить качественную модель прогнозирования оттока;

– применить построенную модель на реальные данные.

Пусть зависимая переменная равна вероятности попадания корпоративного клиента в отток по услуге широкодоступного доступа в Интернет (ШПД) Приморского филиала ПАО Ростелеком.

В качестве функции распределения будут последовательно выбрана функция нормального, логистического распределения и распределение Гомпертца. В качестве факторов, влияющих на отток клиентов сегмента B2B/B2G по услуге ШПД в Приморском филиале ПАО Ростелеком были выбраны переменные, представленные в таблице 1 [27].

Таблица 1 *–* Факторы влияния на отток корпоративных клиентов

|  |  |  |
| --- | --- | --- |
| Переменная | Описание | Гипотеза |
| Количество месяцев в подключении (months) | Период пользования услугой ШПД, месяц | Рост показателя снижает риск  оказаться в оттоке |
| Технология Интернет (tehn) | xDSL, Оптическая технология (optika) | Клиенты на технологии xDSL  имеют большую вероятность  оказаться в оттоке. |
| Среднемесячная  абонентская плата (summ) | Среднемесячная  абонентская плата за период  пользования услугой ШПД | Рост показателя увеличивает  вероятность оттока клиента |
| Сегмент (segm) | Крупные клиенты (3К), cредние и малые предприятия (SMP) | Клиенты, обсуживающиеся в  офисах СМП имеют большую  вероятность оказаться в оттоке. |
| Макросегмент (makro) | Коммерческие организации (B2B), бюджетные организации (B2G). | Коммерческие клиенты B2G имею большую вероятность оказаться в оттоке. |
| Тип населенного пункта (geo) | Город (GST), прочие: села, ПГТ и т.д. (STS) | Клиенты ГТС имеют меньшую  вероятность оказаться в оттоке. |
| Миграция (migr) | Миграция на другую технологию Интернет за период пользования услугой (Да(Y)/Нет(N)) | Клиенты, переключившиеся на  другую технологию, имеют  меньшую вероятность оказаться в  оттоке. |

Как видно из таблицы 1 набор данных состоит из качественных и количественных (факторных) переменных. В таблице также выдвинуты предположения о влиянии факторов на отток клиентов.

Для построения модели оттока корпоративных клиентов случайным образом было отобрано 1522 наблюдения – клиенты Приморского филиала, подключившие услугу широкополосного доступа в Интернет с 2014 по 2016 год включительно, 609 клиентов из них в указанный период отключили услугу, 913 клиентов на дату 31 декабря 2016 года пользовались услугой ШПД.

Бинарная модель оттока корпоративных клиентов будет построена с помощью языка программирования R, который используется для статистической обработки данных и работы с графиками. R широко используется как статистическое программное обеспечение для анализа данных и фактически стал стандартом для статистических программ. Особенность языка R в том, что он поддерживает широкий спектр статистических и численных методов и обладает хорошей расширяемость с помощью пакетов. Пакеты представляют собой библиотеки для работы специфических функций или специальных областей применения. Еще одной особенностью R являются графические возможности, заключающиеся в возможности создания качественной графики, которая может включать математические символы.

Для удобства работы с R будет использован графический интерфейс RStudio – свободная среда разработки программного обеспечения с открытым исходным кодом для языка программирования R.

Описательная статистика набора данных представлена на рисунке 1.

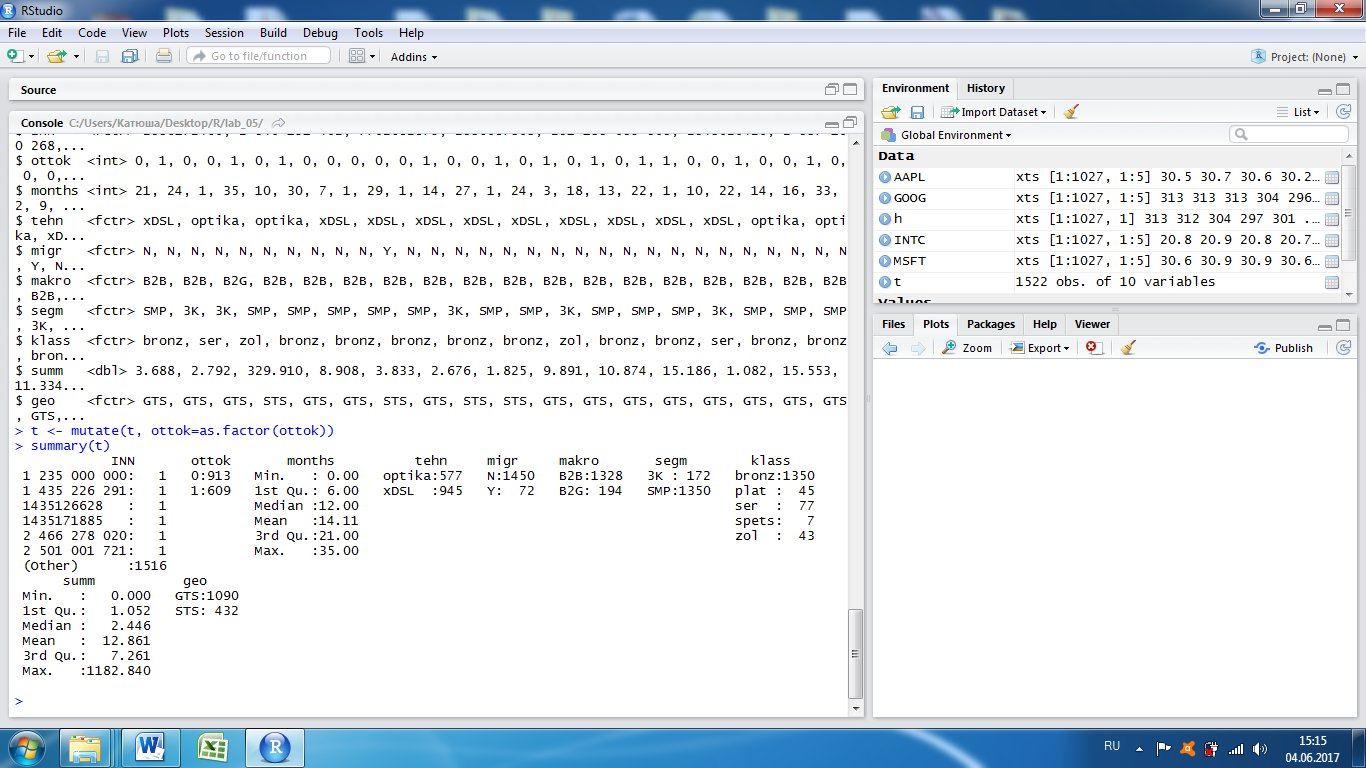


Рисунок 1 – Описательная статистика набора данных о клиентах B2B/B2G

Как видно из рисунка 1 в наборе данных для построения модели оттока 62% корпоративных клиентов пользовались услугой Интернет по технологии xDSL и 38% - по технологии оптика. Минимальная продолжительность пользования услугой ШПД за последние три года – менее месяца. Максимальная продолжительность 35 месяцев. Средняя продолжительность пользования услугой ШПД 14 месяцев. 89 % клиентов обслуживаются в отделе по работе с малым сегментом, 11% клиентов – в отделе по работе с крупными клиентами. 87% всех клиентов выборки относятся к коммерческим предприятиям, остальные 13% - бюджетные организации. 71% от всех клиентов ведут свою деятельность в городской местности, 29% - в сельской местности, 5% клиентов переключали технологию пользования ШПД.

Была оценена зависимость оттока от среднемесячных начислений от клиента, продолжительности пользования услугой ШПД, типа населенного пункта, офиса обслуживания, технологии Интернет, типа организации, совершения перехода с одной технологии Интернет на другую. Для оценки параметров была выбрано функция логистического распределения.

Коэффициенты оцененной логит-модели представлены на рисунке 2.

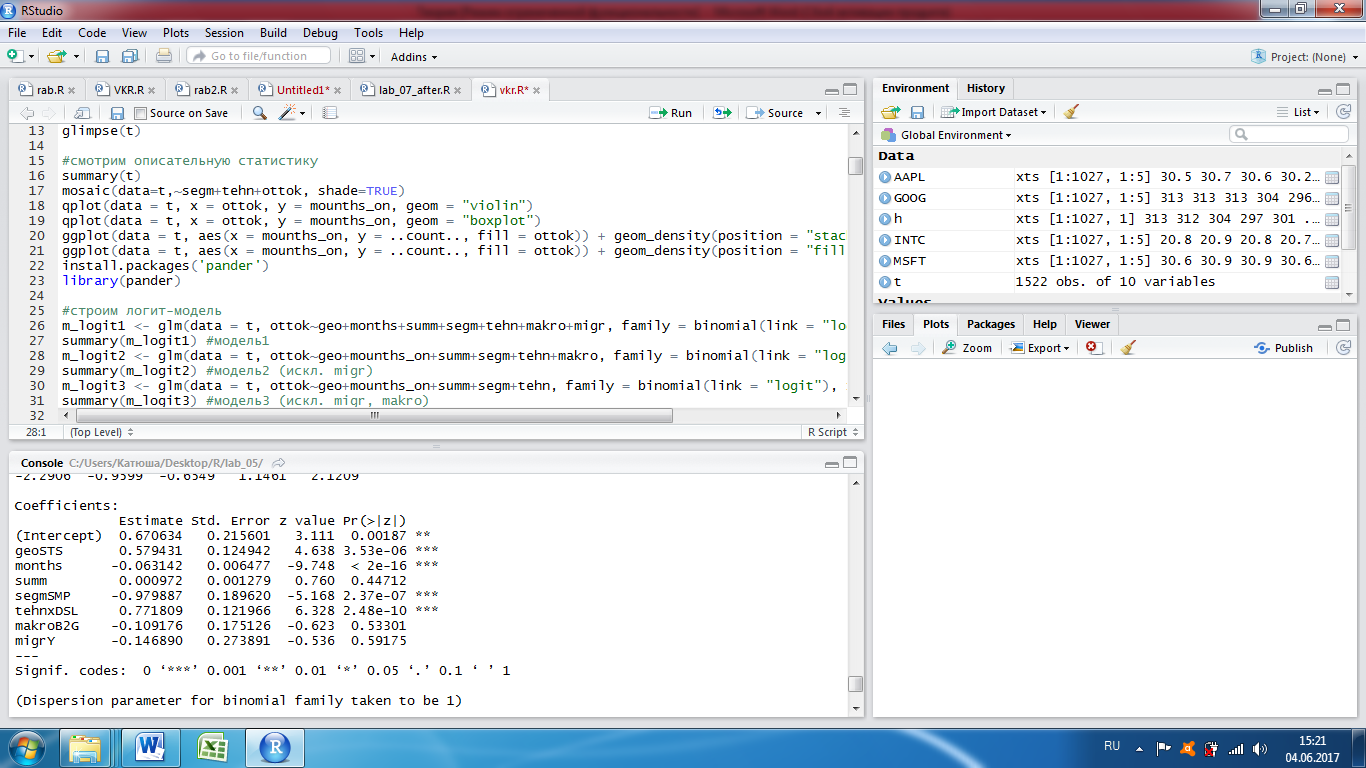


Рисунок 2 – Оценка коэффициентов логит-модели оттока КК

Как видно из рисунка 2 коэффициенты при свободной переменной, типа населенного пункта, продолжительности пользования услугой ШПД, офиса обслуживания и технологии Интернет являются значимыми (наличие звездочек в строке) согласно t-статистике [28].

Поочередно были оценены модели с различным набором зависимых переменных, представленных в таблице 2.

Таблица 2 – Логит-модель оттока

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Модель | Зависимые переменные | | | | | | | | Критерий Акаике | Критерий Байеса-Шварца |
| Logit1 | geo | months | summ | segm | tehn | makro | migr | intercept | 1883 | 1926 |
| Logit2 | geo | months | summ | segm | tehn | makro |  | intercept | 1882 | 1919 |
| Logit3 | geo | months | summ | segm | tehn |  |  | intercept | 1880 | 1912 |
| Logit4 | geo | months | summ | segm | tehn |  | migr | intercept | 1882 | 1919 |
| Logit5 | geo | months |  | segm | tehn |  | migr | intercept | 1880 | 1912 |
| Logit6 | geo | months |  | segm | tehn | makro | migr | intercept | 1882 | 1919 |
| Logit7 | geo | months |  | segm | tehn | makro |  | intercept | 1880 | 1912 |
| Logit8 | geo | months |  | segm | tehn |  |  | intercept | 1879 | 1905 |
| Logit9 |  | months |  | segm | tehn |  |  | intercept | 1898 | 1920 |
| Logit11 | geo | months |  |  | tehn |  |  | intercept | 1910 | 1931 |
| Logit12 | geo | months |  | segm |  |  |  | intercept | 1917 | 1939 |
| Logit13 | geo | months | summ | segm |  |  |  | intercept | 1919 | 1946 |
| Logit14 | geo |  |  | segm | tehn | makro | migr | intercept | 1993 | 2025 |

Как видно из таблицы 2 были вычислены коэффициенты информационных критериев Акаике и Байеса-Шварца для каждой оцененной логит-модели. С помощью данных критериев подтверждается, что переменные geo, months, segm, tehn действительно являются значимыми. И логит-модель logit8, содержащая именно эти факторы при отсутствии остальных, действительно является самой качественной из предложенных.

Результаты оценивания качественной модели logit8 изображены на рисунке 3.

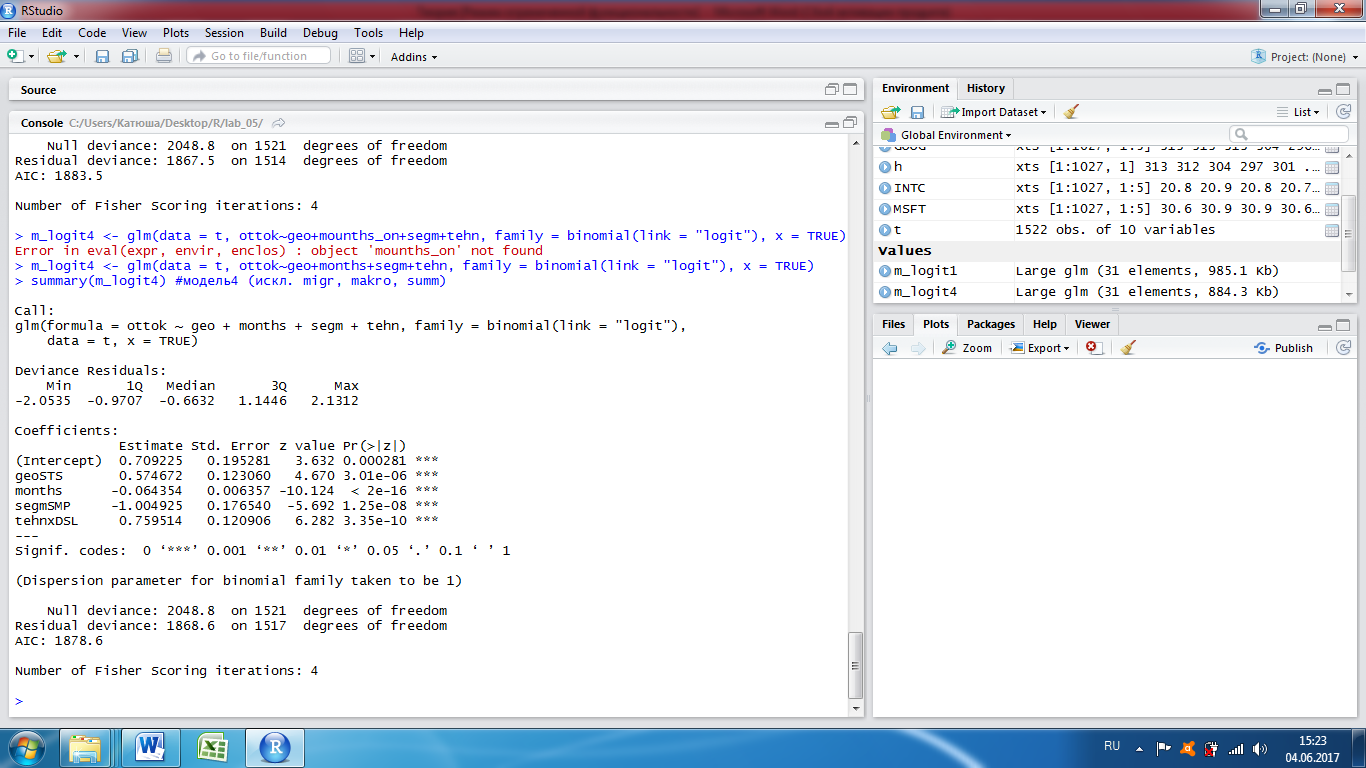


Рисунок 3 – Оценка logit8

Качественная логит-модель оттока клиентов B2B/B2G Приморского филиала по услуге ШПД имеет вид

. (1)

Аналогично были оценены пробит-модели с наборами факторами, представленными в таблице 3.  
Таблица 3 – Пробит-модель оттока клиентов B2B/B2G

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Модель | Зависимые переменные | | | | | | | | Критерий Акаике | Критерий Байеса-Шварца |
| Probit1 | geo | months | summ | segm | tehn | makro | migr | intercept | 1884 | 1926 |
| Probit2 | geo | months | summ | segm | tehn | makro |  | intercept | 1882 | 1919 |
| Probit3 | geo | months | summ | segm | tehn |  |  | intercept | 1880 | 1912 |
| Probit5 | geo | months |  | segm | tehn |  | migr | intercept | 1880 | 1912 |
| Probit6 | geo | months |  | segm | tehn | makro | migr | intercept | 1882 | 1919 |
| Probit7 | geo | months |  | segm | tehn | makro |  | intercept | 1880 | 1912 |
| Probit8 | geo | months |  | segm | tehn |  |  | intercept | 1879 | 1905 |
| Probit9 |  | months |  | segm | tehn |  |  | intercept | 1898 | 1919 |
| Probit10 | geo |  |  | segm | tehn |  |  | intercept | 1991 | 2012 |
| Probit11 | geo | months |  |  | tehn |  |  | intercept | 1909 | 1931 |
| Probit12 | geo | months |  | segm |  |  |  | intercept | 1917 | 1939 |
| Probit13 | geo | months | summ | segm |  |  |  | intercept | 1919 | 1946 |
| Probit14 | geo |  |  | segm | tehn | makro | migr | intercept | 1993 | 2025 |

Оцененные гомпит-модели с различным набором зависимых переменных и оцененные информационные критерии представлены в таблице 4.

Таблица 4 – Гомпит-модель оттока клиентов B2B/B2G

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Модель | Зависимые переменные | | | | | | | | Критерий Акаике | Критерий Байеса-Шварца |
| Gompit1 | geo | months | summ | segm | tehn | makro | migr | intercept | 1892 | 1935 |
| Gompit2 | geo | months | summ | segm | tehn | makro |  | intercept | 1890 | 1927 |
| Gompit3 | geo | months | summ | segm | tehn |  |  | intercept | 1889 | 1921 |
| Gompit4 | geo | months | summ | segm | tehn |  | migr | intercept | 1891 | 1928 |
| Gompit5 | geo | months |  | segm | tehn |  | migr | intercept | 1889 | 1920 |
| Gompit6 | geo | months |  | segm | tehn | makro | migr | intercept | 1890 | 1927 |
| Gompit7 | geo | months |  | segm | tehn | makro |  | intercept | 1888 | 1920 |
| Gompit8 | geo | months |  | segm | tehn |  |  | intercept | 1887 | 1913 |
| Gompit9 |  | months |  | segm | tehn |  |  | intercept | 1905 | 1927 |
| Gompit10 | geo |  |  | segm | tehn |  |  | intercept | 1994 | 2015 |
| Gompit11 | geo | months |  |  | tehn |  |  | intercept | 1913 | 1935 |
| Gompit12 | geo | months |  | segm |  |  |  | intercept | 1921 | 1943 |
| Gompit13 | geo | months | summ | segm |  |  |  | intercept | 1923 | 1950 |
| Gompit14 | geo |  |  | segm | tehn | makro | migr | intercept | 1996 | 2028 |

Как видно из таблицы 3 и таблицы 4 наиболее качественными согласно критерию Акаике и критерию Байеса-Шварца являются модели Probit8 и Gompit8 c аналогичным набором зависимых факторов.

Результаты оценивания качественной модели proit8 изображены на рисунке 3.

Качественная пробит-модель оттока клиентов сегмента B2B/B2G Приморского филиала по услуге ШПД имеет вид

*.*

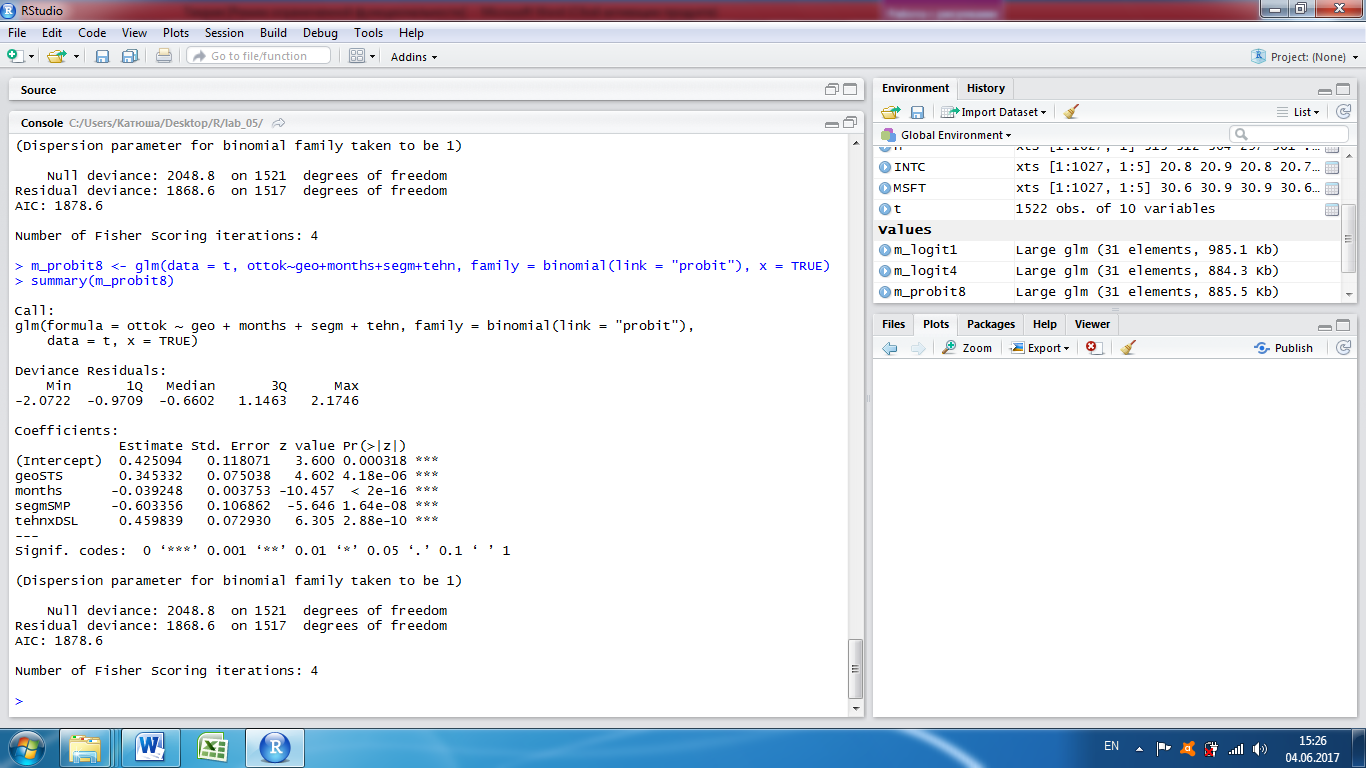


Рисунок 3 – Оценка probit8

Результаты оценивания качественной модели gompit8 изображены на рисунке 4.

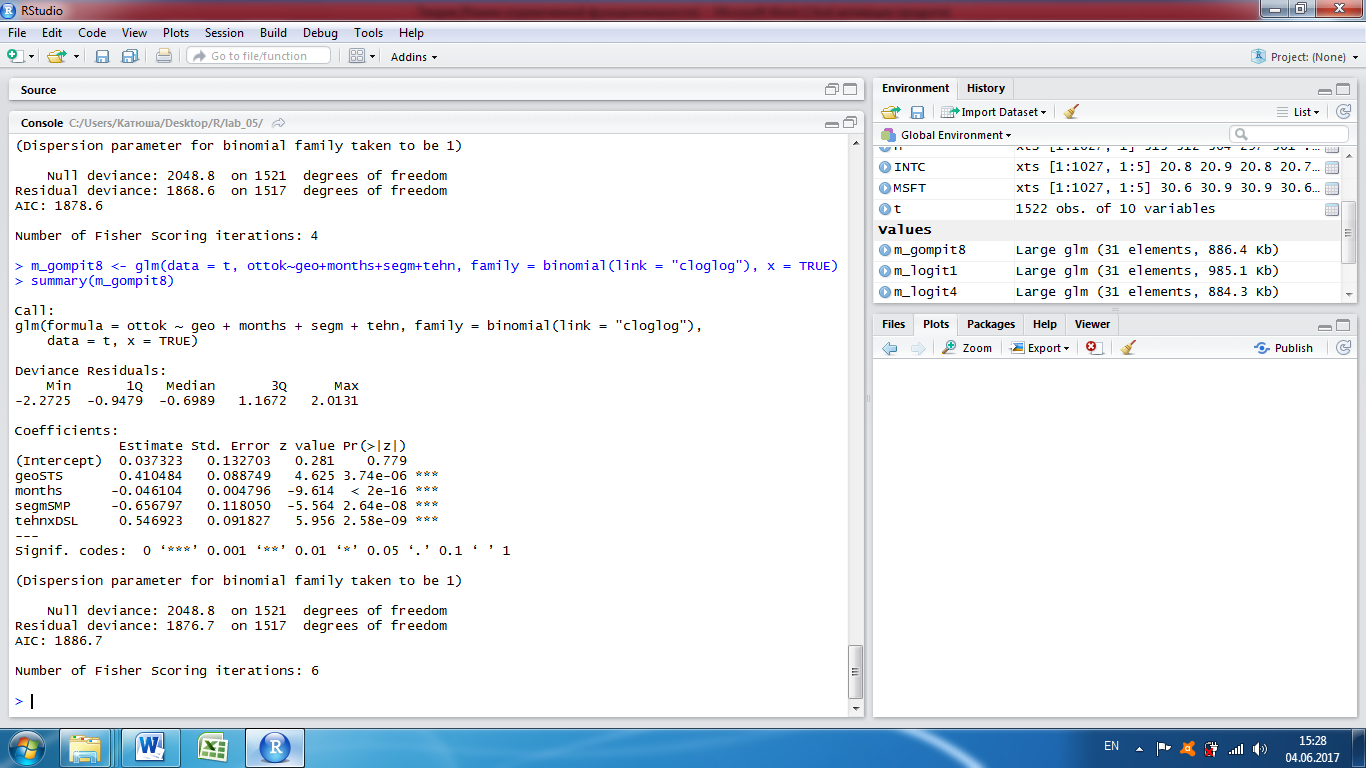


Рисунок 4 – Оценка gompit8

Качественная гомпит-модель оттока клиентов сегмента B2B/B2G Приморского филиала по услуге ШПД имеет вид

.

Для определения наиболее качественной модели для прогнозирования оттока корпоративных клиентов сравним по информационным критериям Акаике и Байеса-Шварца оцененные модели logit8, probit 8 и gompit8. Результаты сравнения представлены в таблице 6.

Как видно из таблицы 6 наиболее качественными моделями являются logit8 и probit8, однако бинарная модель на основе логистического распределения является немного лучше пробит-модели, поэтому именно логистическая модель будет взята для прогнозирования оттока клиентов B2B/B2G Приморского филиала по услуге ШПД.

Таблица 5 – Сравнение моделей

|  |  |  |
| --- | --- | --- |
| Модель | Критерий Акаике | Критерий Байеса-Шварца |
| logit8 | 1878,57 | 1905,20 |
| probit8 | 1878,60 | 1905,24 |
| gompit8 | 1886,67 | 1913,31 |

Таким образом, была получена модель зависимости оттока клиентов B2B/B2G Приморского филиала по услуге ШПД от типа населенного пункта клиента, офиса обслуживания, технологии Интернет, которой пользуется клиент и продолжительности пользования услугой ШПД (в месяцах).

Глядя на модель можно сделать выводы о степени зависимости каждого фактора, влияющего на отток.

Коэффициент при количестве месяцев в подключении равен -0,064. Это значит, что с ростом переменной на 1 месяц отношение шансов падает на 6,4%. Таким образом, при увеличении срока пользования услугой ШПД на 1 месяц вероятность отключения клиента уменьшается на 6,4%.

Качественные коэффициенты интерпретируются иначе [16].

Согласно соотношению (1) при коэффициенте

Если переменная меняется от 0 до 1, сравнивая два типа клиентов, один из которых обслуживается в офисе 3К, а другой в офисе СМП, то вероятность попасть в отток у клиентов из офиса обслуживания СМП в 2,7 раз ниже, чем у клиентов, которые обслуживаются в офисе 3К.

Аналогично, при коэффициенте

При изменении переменной от 0 до 1, сравнивая два типа клиентов: из города и сельской местности, то вероятность попасть в отток у клиентов из города в 1,8 раз ниже, чем у клиентов, ведущих свою деятельность в сельской местности.

При коэффициенте

При изменении переменной от 0 до 1, сравнивая клиентов которые пользуются услугой Интернет по технологии xDSL и клиентов на оптических технологиях, то у клиентов, которым предоставляется услуга ШПД на технологии xDSL вероятность попасть в отток в 2,14 раза выше [17].

В таблице 6 представлены заявленные гипотезы о значимости и влиянии фактора на отток и их сопоставимость с результатами согласно построенной логистической модели оттока.

Таблица 6 – Гипотезы и результаты о влиянии факторов на отток

|  |  |  |
| --- | --- | --- |
| Переменная | Гипотеза | Результат |
| Количество месяцев в подключении (months) | Рост показателя снижает риск оказаться в оттоке | Гипотеза подтвердилась. |
| Технология Интернет (tehn) | Клиенты на технологии xDSL имеют большую  вероятность оказаться в оттоке. | Гипотеза подтвердилась. |

Продолжение таблицы 6

|  |  |  |
| --- | --- | --- |
| Переменная | Гипотеза | Результат |
| Среднемесячная абонентская плата (summ) | Рост показателя увеличивает вероятность оттока клиента | Переменная не влияет на отток. |
| Сегмент (segm) | Клиенты, обсуживающиеся в офисах СМП имеют  большую вероятность  оказаться в оттоке. | Клиенты 3К имеют большую вероятность оказаться в оттоке |
| Макросегмент (makro) | Коммерческие клиенты B2G имею большую вероятность оказаться в оттоке. | Переменная не влияет на отток |
| Тип населенного пункта (geo) | Клиенты ГТС имеют  меньшую вероятность оказаться в оттоке. | Клиенты ГТС имеют большую вероятность оказаться в оттоке. |
| Миграция (migr) | Клиенты, переключившиеся на другую технологию,  имеют меньшую вероятность оказаться в оттоке. | Переменная не влияет на отток. |

Как видно из таблицы 7 только две гипотезы: о влиянии продолжительности подключения и технологии Интернет – подтвердились. 3 выбранных факторных переменных не влияют на отток клиентов. Теперь на основе оценённой логит-модели можно строить прогнозы оттока, задавая необходимые параметры.

Пусть есть потребность в прогнозировании вероятности оттока клиентов сегмента B2B/B2G Приморского филиала, которые пользуются услугой ШПД по технологии xDSL,обслуживаются в офисах 3К и базируются в городе. Продолжительность пользования услугами будет задаваться с 1 до 35 месяца с шагом в 1 месяц.

На рисунке 5 представлены результаты прогнозирования по искусственному набору данных.

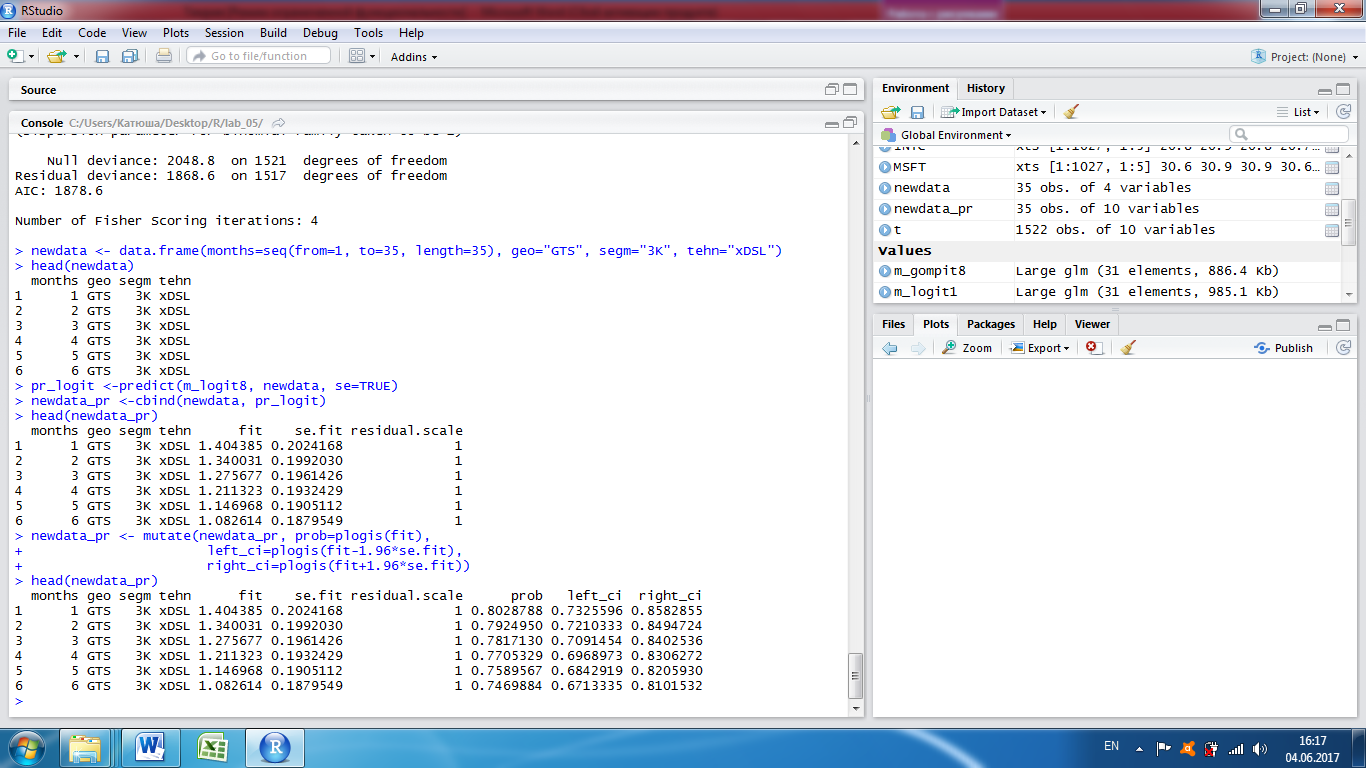


Рисунок 5 – Результаты прогнозирования вероятности оттока клиентов B2B/B2G

Как видно из рисунка 5 были спрогнозированы скрытые переменные fit, стандартные ошибки данной переменной se.fit, вероятность оттока клиента prob и 95%-е доверительные интервалы.

На рисунке 6 представлен график влияния продолжительности пользования услугой ШПД по технологии xDSL клиентами офиса обслуживания 3К, базирующимся в городе.

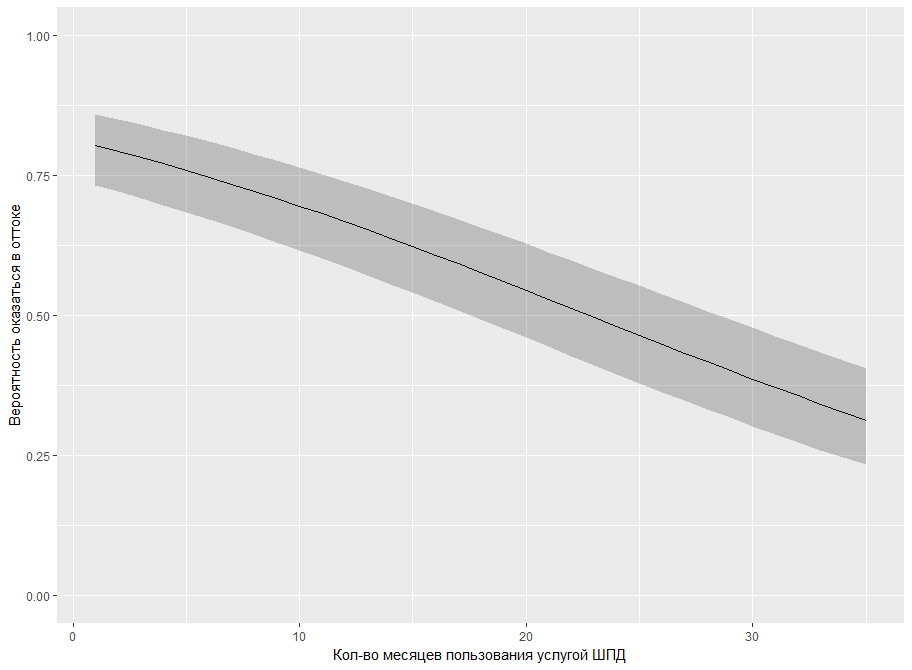


Рисунок 9

Как видно из рисунка 6, действительно, чем больше клиент пользуется услугой ПАО Ростелеком, тем вероятность снижается, при этом видно, что чем меньше клиент пользуется услугой, тем точнее прогноз [18].

На основе построенной логит-модели оценим предельные эффекты. Результаты оценивания представлены на рисунке 7.

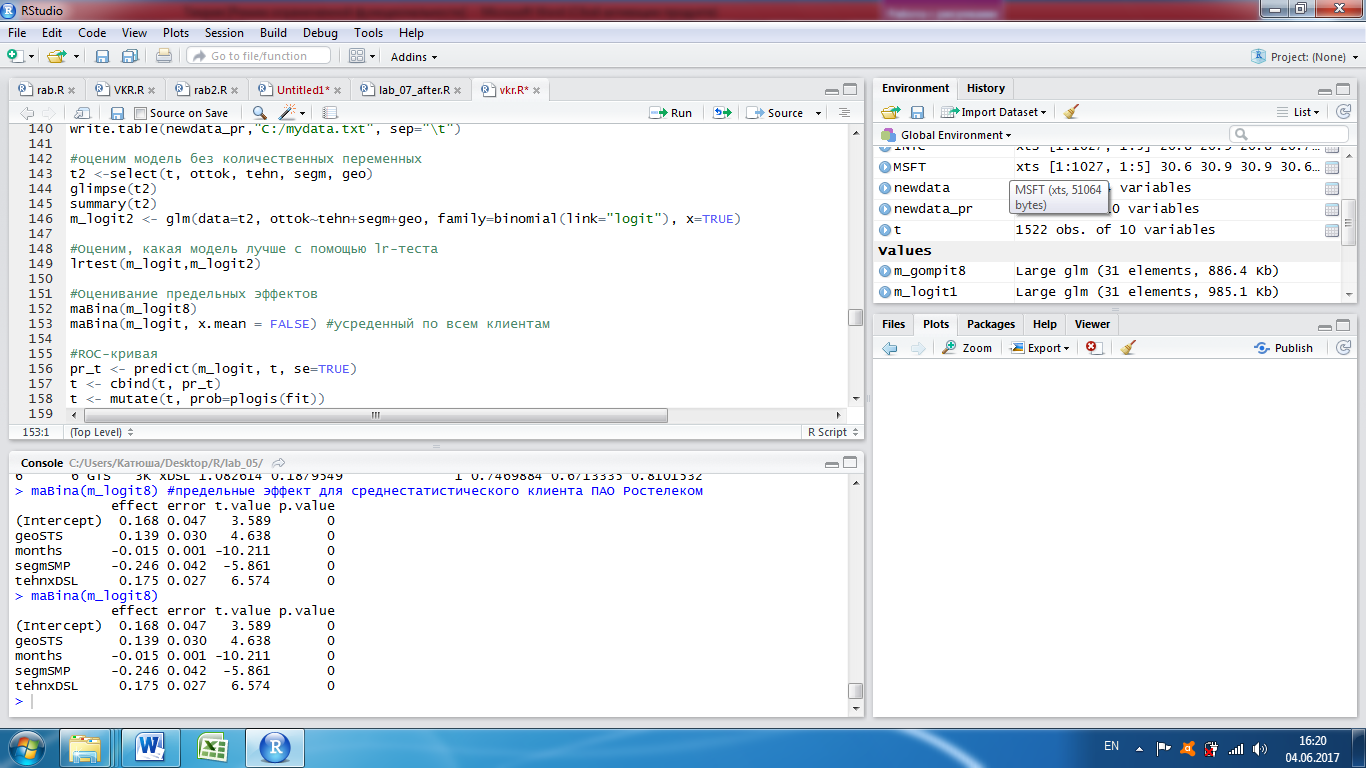


Рисунок 7 – Оценка предельных эффектов

Как видно из рисунка 7 получили результаты среднестатистического корпоративного клиента Приморского филиала ПАО Ростелеком, использующего услугу ШПД. Можно сказать, что увеличение продолжительности пользования услугой для среднестатистического клиента, то есть такого, который пользуется услугой ШПД 14 месяцев, который на 62% использует технологию xDSL и на 38% - оптическую технологию, который н 89% является клиентом СМП и на 11% клиентом 3К, и который на 71% базируется в городе и на 29% в сельской местности, снижает вероятность оттока на 0,015 [19].

Посчитаем средние эффекты, усреднив их по клиентам ПАО Ростелеком. Результаты представлены на рисунке 8.

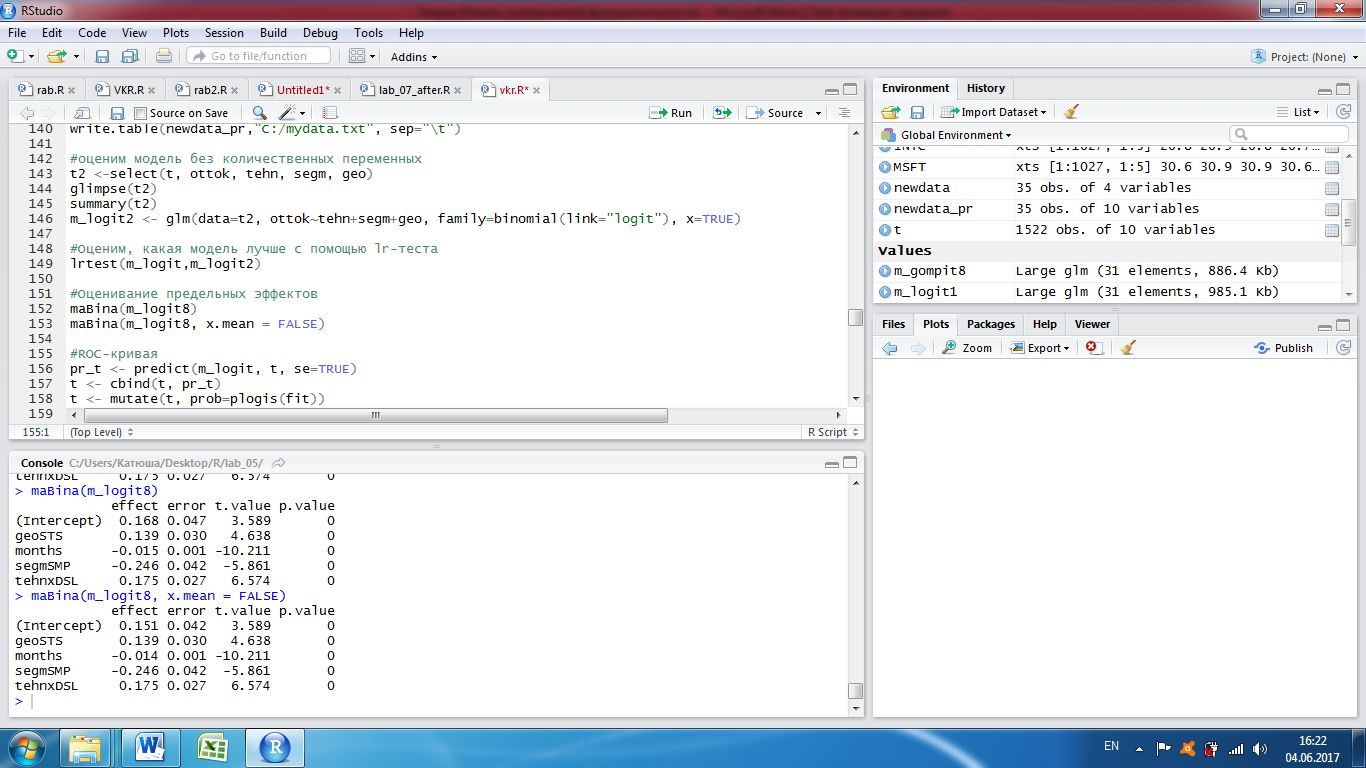


Рисунок 8 – Оценки средних предельных эффектов

Из рисунка 8 видно, что если каждому клиенту увеличить период пользования услугой ШПД на 1 месяц, то усреднив вероятность оттока по всем клиентам, получим, что она падает на 0,14. Был рассчитан предельный эффект по каждому клиенту, а затем усреднен по всем клиентам.

На основе оцененной бинарной модели оттока спрогнозируем вероятность оттока для корпоративных клиентов Приморского филиала, которые подключили услугу ШПД в 2017 году. Всего за период январь-май подключились 667 клиентов. Данные содержат информацию о технологии подключения, количестве месяцев в подключении, подразделении обслуживания и типе населенного пункта.

Применив оцененную модель к набору данных о клиентах, получили вероятность оттока клиента и 95 %-е доверительные интервалы, представленные на рисунке 9.

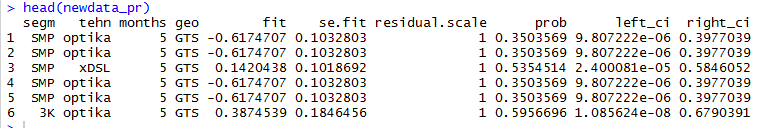


Рисунок 9 – Прогноз вероятности оттока корпоративных клиентов

Для того, что бы точно ответить на вопрос, попадет ли данный клиент в отток или нет, необходимо установить порог отсечения, учитывая, его значение влияет на вероятность разных ошибок, поэтому для выбора вероятности – порога отсечения будет использован анализ на основе ROC-кривой [20].

Спрогнозируем вероятность оттока для первоначального набора данных. На основании этих данных отберем разные пороги отсечения и посмотрим, как влияет порог отсечения на вероятность ошибочного отнесения клиента к тому или иному результату прогнозирования.

Данные для построения ROC-кривой представлены на рисунке 10.

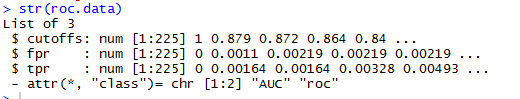


Рисунок 10 – Набор данных для построения ROC-кривой

На рисунке 10 представлены переменные:

– сutoffs – возможные пороги отсечения – граница за которой клиент признается как попавший в отток;

– fpr (False positive rate) – доля ошибочно классифицированных корпоративных клиентов, не попавших в отток;

– tpr (True positive rate) – верно классифицированные корпоративные клиенты, попавшие в отток.

На рисунке 11 представлен график зависимости доли верно классифицированных клиентов, попавших в отток от порога отсечения.

Как видно из рисунка 11 при выборе низкого порога отсечения большое количество клиентов классифицируется как не попавшие в отток. В таком случае доля верно классифицированных клиентов попавших в отток оказывается высокой. Выбирая высокий порог отсечения, существенно реже будут прогнозироваться клиенты как не попавшие в отток, и поэтому доля верно классифицированных клиентов попавших в отток будет ниже.

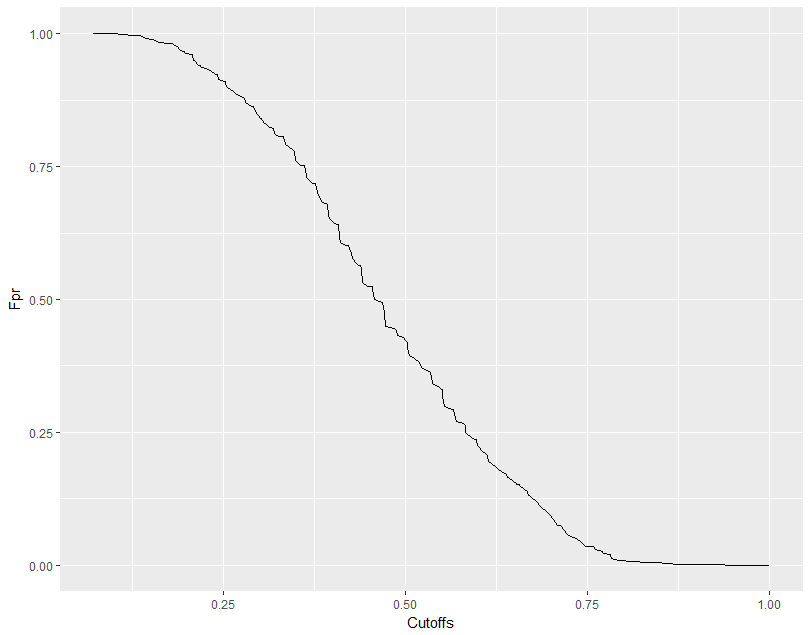


Рисунок 11

На рисунке 12 изображен график неверно классифицированных клиентов, не попавших в отток.

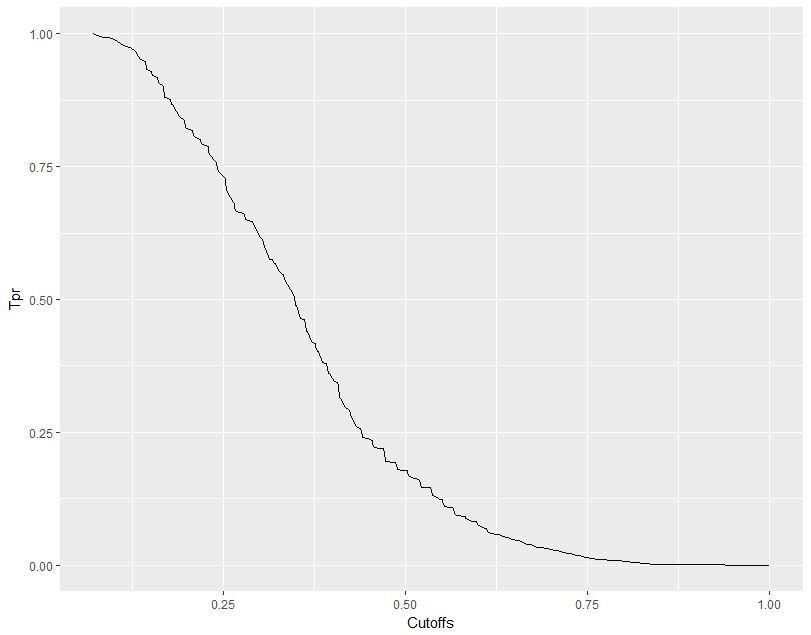


Рисунок 12

Как видно из рисунка 12, если выбирать очень маленький порог отсечения, то всех клиентов прогнозируем как попавших в отток, при этом доля неверно классифицированных попавших в отток клиентов будет достаточно высокой.

Сочетание двух графиков, изображенных на рисунке 11 и рисунке 12, называю ROC-кривой. На рисунке 13 представлена ROC-кривая.

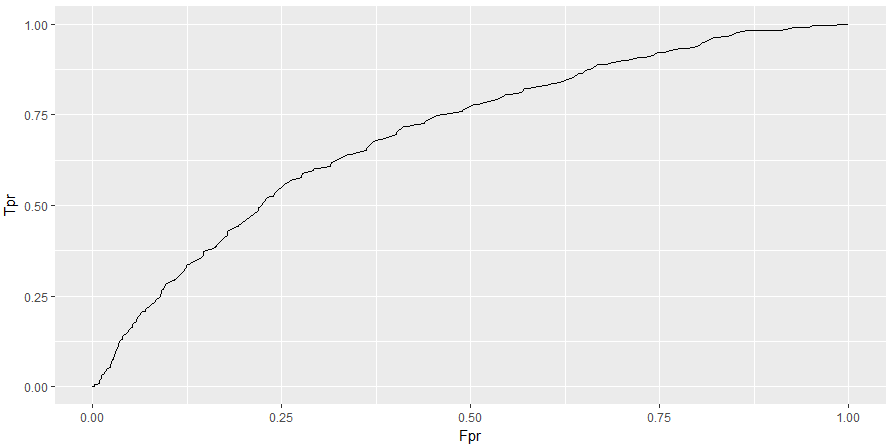


Рисунок 13 – ROC-кривая

Если установить высокий порог отсечения, то наблюдается низкая доля верно классифицированных клиентов в оттоке, но высокая доля верно классифицированных клиентов, которые не отключились. И наоборот, если установить низкий порог отсечения, то наблюдается высокая доля верно классифицированных клиентов, попавших в отток, но низкая доля неверно классифицированных клиентов, которые не отключились.

Площадь под ROC-кривой () равна 0,7. Так как >0,5 то модель признается рекомендуемой к прогнозу.

На основании ROC-кривой был выбран порог отсечения равный 0,6.

На основании прогнозной логит-модели были найдены вероятности оказаться в оттоке для клиентов Приморского филиала, подключивших услугу ШПД с января по апрель 2017 года. Описательная статистика набора данных о новых клиентах представлена на рисунке 13.

Как видно из рисунка 17 всего в 2017 году подключились 667 абонентов. Из них 23% клиентов обслуживаются в офисах 3К, 59% подключены на оптическую технологию,78% базируются в городской местности.

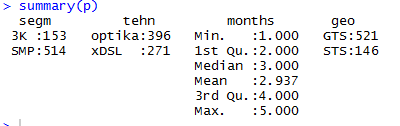


Рисунок 13 – Описательная статистика тестового набора данных

Применив выбранный порог отсечения к реальному набору данных получили вероятности попасть в отток. На основании ROC-анализа был выбран порог отсечения 0,6. Количество клиентов, имеющих вероятность оказаться в оттоке выше 0,6 равно 227.

В случае отключения данных клиентов потерянный доход в 2017 году составит 4 036 тыс. руб.

Таким образом, на вероятность оказаться в оттоке влияет длительность периода использовании услуги, офис обслуживания корпоративного клиента, технология ШПД и тип населенного пункта, в котором ведется деятельность организации-клиента. На основании построенной модели и выделенных факторов, необходимо выделить проблемные зоны в обслуживании и сопровождении корпоративных клиентов в разрезе офисов обслуживания корпоративных клиентов, типа населенного пункта, технологии Интернет, и провести соответствующие мероприятия по их устранению.

Каждые шесть месяцев следует формировать модель оттока, и, сравнив ее с предыдущей, дать оценку эффективности проведенным мероприятиям по решению существующих проблем и в последствии по сокращению оттока клиентов сегмента B2B/B2G.

В результате работы была сформирована логистическая модель оттока корпоративных клиентов Приморского филиала по услуге ШПД, определена степень влияния каждого фактора на отток. Также был спрогнозирован отток клиентов, подключивших услугу ШПД ПАО Ростелеком в 2017 году.

В процессе решения задачи по знакомству с процедурой по работе с оттоком в ПАО Ростелеком были рассмотрены основные мероприятия по предотвращению оттока корпоративных клиентов, предикторы, которые ПАО Ростелеком выделяет, как факторы, влияющие на отток, были рассмотрены основные показатели, которые применяют для отчетности по оттоку клиентов.

При третьей задачи были выделены факторы, влияющие на отток, а также бинарные признаки корпоративных клиентов, влияние на отток которых необходимо было рассмотреть. Были сформулированы гипотезы о влияния данных факторов на отток, а также собраны необходимые данные о клиентах Приморского филиала, которые подключили услугу ШПД в период с 2014 до 2016 года. Были выделены следующие факторы: среднемесячные начисления от клиента, продолжительность пользования услугой, офис обслуживания, тип предприятия (клиента), тип населенного пункта, технология Интернет, миграция на другую технологию.

В рамках задачи построении модели для прогнозирования модели с использованием инструментов языка R в среде RSudio были построены модели бинарного выбора на основе логистического, нормального распределения и распределения Гомпертца. Качественными переменными, влияющих на отток клиентов оказались: продолжительность пользования услугой ШПД, офис обслуживания, тип населенного пункта и технология Интернет. Таким образом были сделаны следующие выводы:

– чем больше период пользования услугой ШПД у клиента, тем меньше вероятность, что он окажется в оттоке;

– клиенты офиса обслуживания 3К имеет большую склонность к оттоку, чем клиенты офиса СМП;

– у корпоративных клиентов, использующих технологию Интернет xDsl вероятность оказаться в оттоке больше;

– клиенты, базирующиеся в сельских населенных пунктах, имеют большую вероятность оказаться в оттоке.

Также на основе построенной модели есть возможность прогнозировать вероятность оттока у клиента по заданным признакам.

При решении задачи прогнозирования оттока клиентов корпоративного и государственного сегмента были собраны данные по подключениям клиентов Приморского филиала в 2017 году к услуге широкополосного доступа в Интернет за четыре месяца с января 2017 года по апрель 2017 года, и построены прогнозы вероятности оттока для данных корпоративных клиентов. С помощью анализа ROC-кривой был выбран порог отсечения равный 0,6.

Таким образом при прогнозной вероятность большей 0,6 корпоративного клиента классифицируют как клиента, которых находится под угрозой оказаться в оттоке.

Таким образом из 667 клиентов, подключившимися в 2017 году 227 клиента рискуют оказаться в оттоке. На основе спрогнозированного оттока был посчитан возможный потерянный доход от корпоративных клиентов, ушедших в отток, за 2017 год в размере 4 036 тысяч рублей.

Также определив степень влияния факторов на отток, есть возможность определить существующие проблемы в сервисе в разрезе офисов обслуживания корпоративных клиентов, типов населенных пунктов, технологий предоставления Интернет. Исходя из полученных результатов необходимо инициировать мероприятия по устранению существующих проблем, при этом с помощью прогнозной модели возможно определить в какие направления более приоритетные для решения создавшейся проблемы.

Каждые шесть месяцев рекомендуется повторно формировать модель оттока, и, сравнив ее с предыдущей, дать оценку эффективности проведенным мероприятиям по решению существующих проблем и в последствии по сокращению оттока клиентов корпоративного и государственного сегмента.