

## РОССИЙСКИЙ МЕНЕДЖМЕНТ: СОВРЕМЕННЫЙ РАКУРС

Международный научно-практический журнал «Вестник МИРБИС» ISSN 2411-5703 <http://journal-mirbis.ru/>  
№ 2 (14) 2018 DOI: 10.25634/MIRBIS.2018.2

**Ссылка для цитирования этой статьи:** Ганявин В. А., Черников С. Ю., Макаров А. А. RFM-сегментация и классификация клиентской базы для группы 1R алгоритмом машинного обучения [Электронный ресурс] // Вестник Московской международной высшей школы бизнеса (МИРБИС). 2018. № 2 (14). С. 20-27. DOI: 10.25634/MIRBIS.2018.2.3

УДК 339.138 : 004.65

**Василий Ганявин<sup>1</sup>, Сергей Черников<sup>2</sup>, Александр Макаров<sup>3</sup>**

### RFM-СЕГМЕНТАЦИЯ И КЛАССИФИКАЦИЯ КЛИЕНТСКОЙ БАЗЫ ДЛЯ ГРУППЫ 1R АЛГОРИТМОМ МАШИННОГО ОБУЧЕНИЯ

**Аннотация.** Актуальность данной темы состоит в одной из самых больших проблем, стоящих перед маркетологами – персонализации обмена сообщениями с индивидуальными клиентами и клиентами, для достижения наилучшего отклика от получателя. Результаты релевантного маркетинга включают увеличение лояльности клиентов, вовлеченности и снижения непроизводительных расходов. Однако без машинного обучения сложно обработать огромное количество данных, поступающих из нескольких источников, необходимых при прогнозировании успешности маркетинговых предложений и стимулов для каждого отдельного клиента. Цель данной работы – проанализировать пример RFM-сегментации и машинная классификация клиентов компании, работающей на московском рынке FMCG косметических услуг. Методы исследования данной проблемы заключались в RFM сегментировании существующих данных методами машинного анализа. Результатами работы стало понимание наиболее практичного способа внедрения методик машинного обучения в практику маркетинга любой компании через модели сегментирования, которые оказались очень эффективны при классификации целевой аудитории на небольшие однородные группы клиентов с аналогичным поведением и предпочтениями. Практические результаты связаны с предложением наиболее популярных областей маркетингового применения данных технологий, в конечном счете, связаны с выявлением характеристик данных, связанных с желаемым для компании результатом взаимодействия с целевой аудиторией.

**Ключевые слова:** машинное обучение, большие данные, прогнозная аналитика, предикативный прогноз, RFM сегментация, FMCG косметических услуг.

JEL: M15, M31

1 **Ганявин Василий Александрович** – кандидат технических наук. ООО «Геолинк Ньютек». Россия, 125047, Москва, ул. Лесная, 1/2.  
E-mail: [vganyavin@geolink.ru](mailto:vganyavin@geolink.ru).

2 **Черников Сергей Юрьевич** – кандидат экономических наук, доцент кафедры маркетинга, экономический факультет. ФГАОУ ВО «Российский университет дружбы народов» (РУДН). Россия, Москва, 117198, ул. Миклухо-Маклая, 6. E-mail: [chernikov.rudn@gmail.com](mailto:chernikov.rudn@gmail.com).  
ORCID: 0000-0001-7524-4438; Scopus AuthorID : 57191344141.

3 **Макаров Александр Анатольевич** – доктор технических наук, профессор кафедры автоматизации и промышленной электроники. ФГБОУ ВО «Российский государственный университет им. А. Н. Косыгина» (РГУ им. А. Н. Косыгина). Россия, 117997, Москва, ул. Садовническая, 33/1.  
E-mail: [mak68@bk.ru](mailto:mak68@bk.ru).

#### Введение

В последние годы одним из главных трендов развития маркетинга в мире является использование прогнозной аналитики на основе механики машинного обучения. Это методика, объединяющая науку, статистику и компьютерное кодирование, целью которых является прогнозирование на основе закономерностей, обнаруженных в обрабатываемых данных. В отличие от систем принятия решений, основанных на правилах, которые заранее следуют явному набору заданных разработчиками инструкций, алгоритмы машинного обучения предназначены для относительно свободного анализа данных и обнаружения паттернов, которые человеческие исследователи либо не в состоянии найти сами, либо поиск которых требует гигантских трудозатрат. Иными словами, ма-

шинное обучение использует вычислительные возможности и объективность компьютеров, чтобы определить, как новые данные могут использоваться для точного прогнозирования результатов целевых процессов. Одна из самых больших проблем, стоящих перед маркетологами, заключается в персонализации обмена сообщениями с индивидуальными клиентами и клиентами, для достижения наилучшего отклика от получателя. Результаты релевантного маркетинга включают увеличение лояльности клиентов, вовлеченности и снижения непроизводительных расходов. Без машинного обучения сложно собрать и обработать огромное количество данных, поступающих из нескольких источников (например, поведение во время покупки, поток посещения веб-сайта, использование мобильных приложений и ответы на предыдущие

опросы в кампаниях), необходимых для прогнозирования успешности маркетинговых предложений и стимулов для каждого отдельного клиента.

**Материалы и методы.** Поскольку методика RFM сегментации методами машинного обучения требует объёмных данных, материалом данной работы являлись данные о клиентах реального участника рынка косметологических услуг.

Наиболее популярные области маркетингового применения данных технологий, в конечном счете, связаны с выявлением характеристик данных, связанных с желаемым для компании результатом взаимодействия с целевой аудиторией. Таким образом, определяется правильное время контакта с клиентом, структура продуктов, моделирование потребительского отклика, прогнозирование оттока клиентов. Открывая шаблоны в данных, генерируемых многими в прошлом остановившимися продажи контрагентов, механики машинного обучения могут точно предсказать, какие из текущих клиентов находятся в зоне наивысшего риска по негативному для компании поведению. Это позволяет маркетологам заниматься профилактической профилактикой, что является важным способом увеличения доходов.

Наиболее практичным способом внедрения методик машинного обучения в практику маркетинга любой компании является сегментация клиентской базы. Модели сегментирования очень эффективны при классификации целевой аудитории на небольшие однородные группы клиентов с аналогичным поведением и предпочтениями. В данной работе приведен пример RFM-сегментации и классификация клиентов компании, работающей на московском рынке FMCG косметических услуг.

**Результаты**

На данном рынке существуют следующие фундаментальные проблемы при организации стандартных маркетинговых исследований:

- неоднородность и обезличенность клиентской базы;
- непредсказуемость в поведении клиентов при покупке продукции компании;
- слабо структурированные прицельные (точечные) предложения для клиентов.

В силу данных ограничений была поставлена задача восстановления необходимой для моделирования выборки данных по сделкам и характеристикам клиентов (профиль и уровень клиента, количество сделок и их сумма по месяцам и т. д.), сегментированию клиентов по условию или событию, проведению анализа сегментов клиентской базы, прогнозирования поведения клиентов и выработки конкретных рекомендации для «прицельного» воздействия на сегменты. Ключевым в работе являлся анализ группы категории клиентов 1R на так называемое «увядание» (потерю мотивации совершать покупки) и «сползание» (постепенный понижающий переход из сегмента высокочастотных и высоко прибыльных групп в другие сегменты) и прогноз поведения клиента готового к совершению в будущем более серьезных сделок как по стоимости, так и по количеству.

Для сегментирования была выбрана методика Recency Frequency Monetary (RFM, «давность, частота, деньги»). «Давность» показывает время, прошедшее с последнего взаимодействия с клиентом в днях, неделях или месяцах и рассчитывается как разность между датами последнего заказа и настоящего момента. При этом подразумевается, что недавно совершившие у вас покупку клиенты более предрасположены к повторным заказам. «Частота» – показывает сколько взаимодействий в течение определённого периода времени было у вас с потребителем. Предполагается, что большое число покупок, совершённых клиентом выше вероятность их последующего повторения. Наконец, индикатор «деньги» показывает финансовый объём покупок за определённый период или количество взаимодействий и является в определённом смысле «стоимостью клиента» с точки зрения доходов и прибыльности.

**Обсуждение**

В исследуемом случае RFM сегментация клиентской базы была выполнена по всей выборке данных для более чем 65000 клиентов. В ходе исследования удалось разбить клиентов на группы на основе давности, частоте сделок, купивших на определённую сумму. Граничные условия показаны на рис. 1.

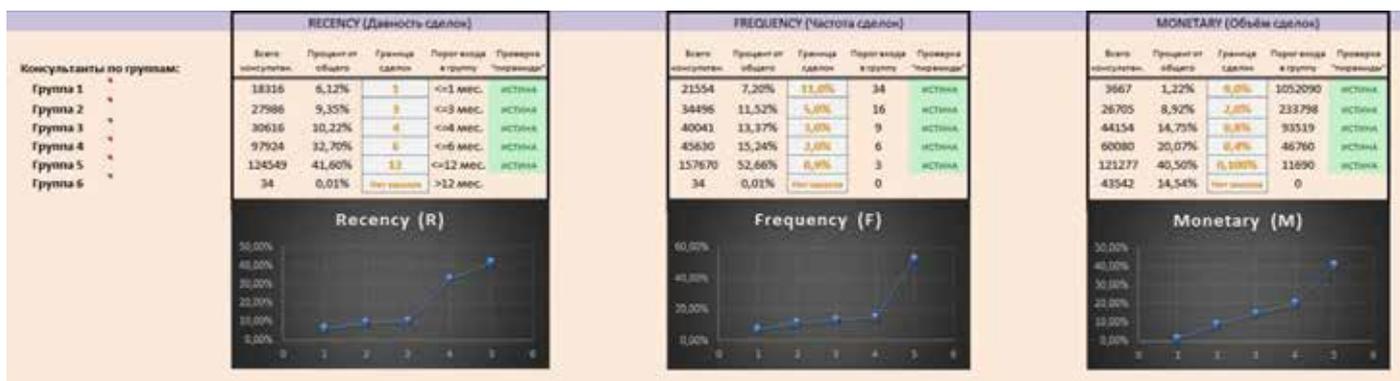


Рис. 1. Граничные условия RFM сегментации клиентов. Рисунок авторский.

По каждому RFM-сегменту автоматически формировался отчет с ключевыми маркетинговыми показателями: количество заказов, средний чек, сумма покупок, количество клиентов в сегменте, разбивка по уровням, отображение аномального цикла в поведении, проверка эффекта «губной помады»<sup>1</sup> по каждому сегменту (см. полную версию мнемосхемы – Рис. 8). Затем, когда клиенты были разбиты на отдельные группы, были выбраны три из категории 1R, которые приносят компании наибольшие продажи. По методике RFM эти группы требуют от компании наибольшего внимания в разрезе контроля поведения и предикативного прогноза по «сползанию» с высшей группы (1R) на нижнюю (2R).

При дальнейшем рассмотрении клиентской базы анализ позволил выявить и другие тенденции, требующие повышенного внимания со стороны корпоративных маркетологов. Так, выяснилось, что для 4 и 5 категорий свойственна ярко выраженная сезонность, что подразумевает разработку и ориентацию на них сезонные акции. В ходе сегментации проявилось, что низшие группы (4R–5R) являются более подверженные эффекту «губной помады» при условии двухмесячного смещения зависимости: количество сделок Nгр клиентов относительно курса национальной валюты (рубли к доллару). Это подразумевает необходимость для маркетологов компании

1 Увеличение объем продаж при кризисе (характерно для рынка косметических услуг).

предлагать руководству компромиссные подходы по предоставлению скидок для групп (4R-5R), если необходимо сохранить более-менее стабильные продажи в течение всего года. В противном же случае, в силу достаточно предсказуемого поведения данного сегмента, останется только предполагать, что в конечном итоге сезонность выровняет нагрузки за счёт пикового периода продаж и создаст предпосылки к повышению общего количества совершаемых сделок.

В качестве обучающей выборки для машинного обучения были выбраны группы:

- а) 1R1F1M – 3276 клиентов;
- б) 1R1F2M – 2612 клиентов;
- в) 1R2F2M – 648 клиентов.

Машинная классификация клиентской базы проводилась в программе MATLAB. В анализ попадали подготовленные данные 10460 клиентов категории 1R (~16% от общей выборки). Обучение проводилось для двух случаев: на общем наборе обучающей выборки, и на наборе случайной выборки размерностью 47 клиентов (для этой выборки из двух групп нормальных клиентов и трёх аномальных – общая размерность массива 235x24). Для классификации данных использовался алгоритм SVM (машина опорных векторов).

Результаты анализа представлены далее (в рисунках 2-5).

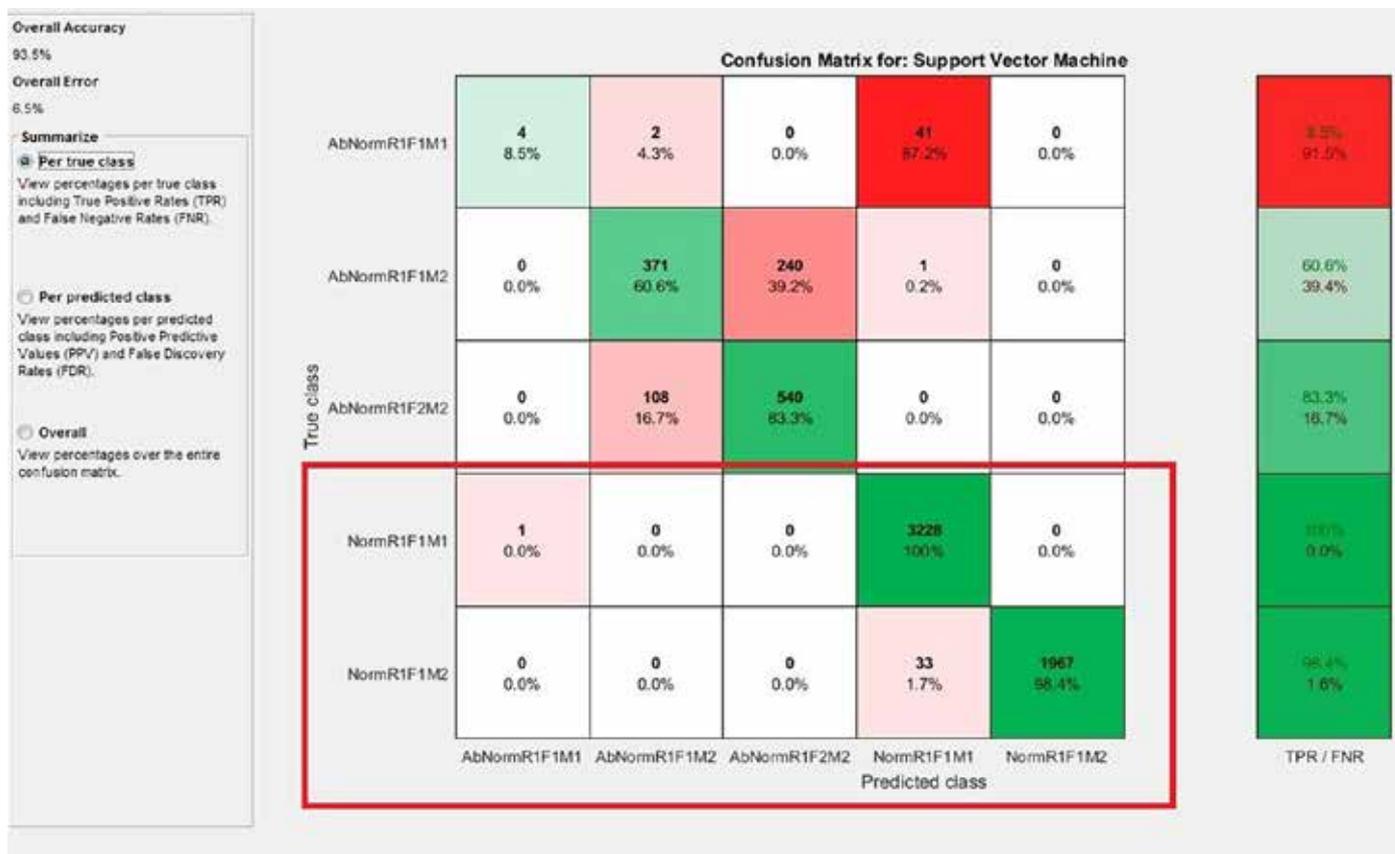


Рис. 2. Матрица принадлежности i-го клиента от суммы месячной сделки. Рисунок авторский.

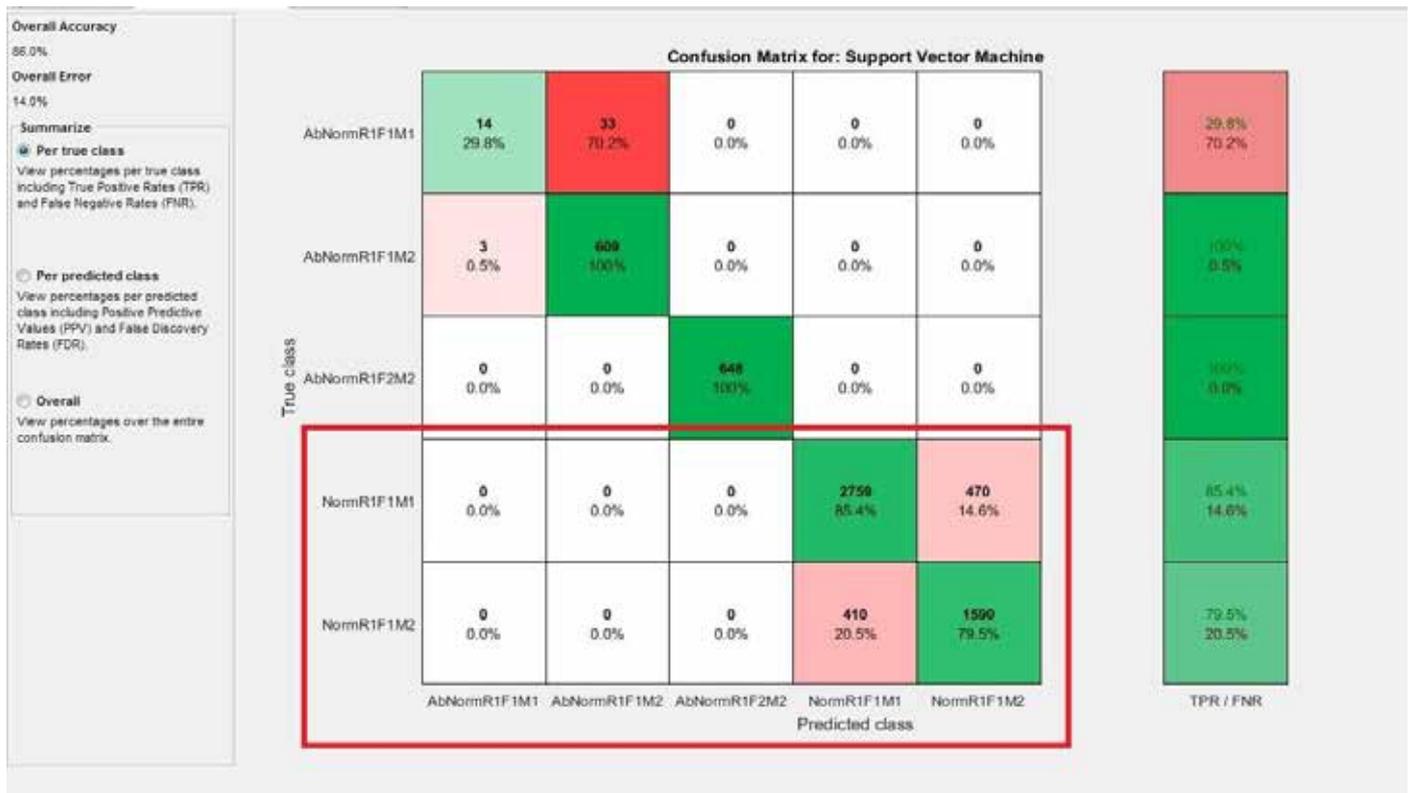


Рис. 3. Матрица принадлежности i-го клиента от общего количества сделок. Рисунок авторский.

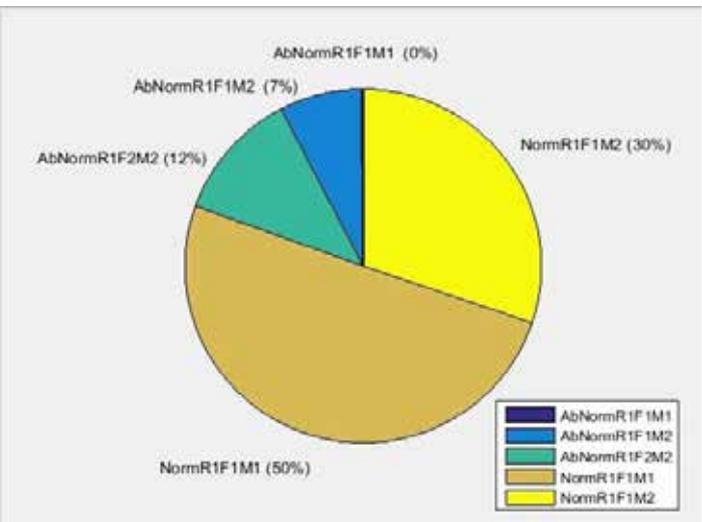
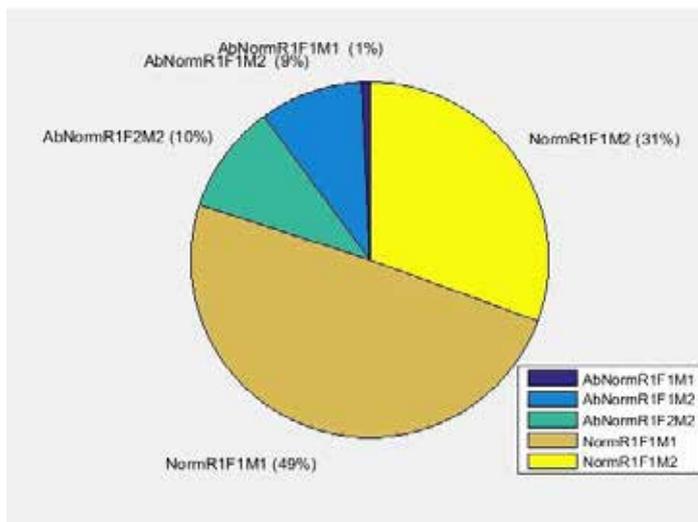


Рис. 4. Первоначальный вектор клиентов группы 1R. Рисунок авторский.

Рис. 5. Восстановленный вектор клиентов группы 1R. Рисунок авторский.

Восстановленные данные позволили сделать сортировку клиентов на следующие категории:

- а) идентифицированные клиенты находятся в «своей» группе, по ним отсеиваются маловероятные;
- б) клиенты, находящиеся на границе «отсева» на другую группу; выбраны по одному из двух возможных показателей (сумма месячной сделки или общее количество сделок);
- в) клиенты, которые практически перешли в другую группу; заранее классифицируются как «не свои» в группе

1R. При этом маркетологам необходимо провести дополнительную работу и определить намерения данной группы клиентов по «переходу» на уровень выше.

Вероятно, данная группа разделится на часть без необходимости дополнительного стимулирования и на часть, требующую релевантных маркетинговых предложений.

Результаты классификации клиентов для тестовой выборки (набор случайной выборки размерностью 47 клиентов) представлены на рис. 6 и рис. 7. На них представлены матрицы принадлежности i-го клиента к группе.



Рис. 6. Матрица принадлежности i-го клиента от суммы месячной сделки. Рисунок авторский.

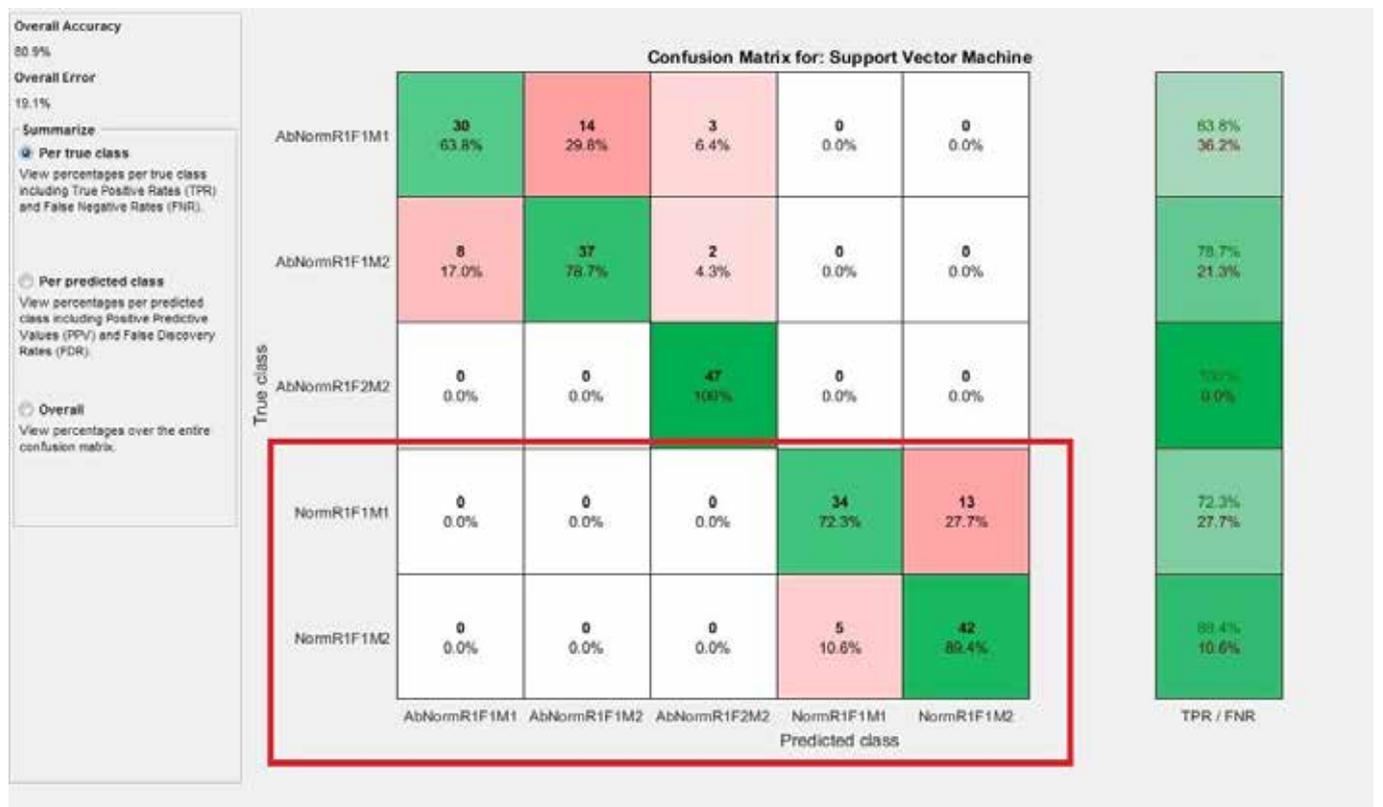


Рис. 7. Матрица принадлежности i-го клиента от общего количества сделок. Рисунок авторский.

**Заключение**

Данный анализ позволил сделать следующие выводы для отдела маркетинга компании:

1. Существует корреляция между общей выборкой данных и случайной выборкой – понижение размерности тестовой выборки на 60 % не приводит к существенному отличию в классификации клиентов;
2. Точность классификации выше для массива – зависимость «сумма месячной сделки» от профиля потенциального клиента (обучающая – 93,5 %; тестовая – 79,6 %);
3. Нормальная группа R1F1M1. Из 3276 клиентов – 470 (14,6 %) потенциальные клиенты, готовые «сползти» на группу ниже – R1F1M2 (частые, объёмные покупки

- с более низким общим чеком из-за смены ассортимента продукции);
4. Нормальная группа R1F1M2. Практически 23 % клиентов от общей численности группы готовы к дорогим покупкам; в ближайшее время часть из них может «войти» в группу R1F1M1 (идентификация по id профилю);
5. Аномальная группа R1F1M1. При анализе методика показала переход в аномальную R1F1M2 (около 75 %). Данная группа имеет нестабильные показатели и требует дополнительного внимания и исследования;
6. Аномальная группа R1F1M2. Существенное «сползание» в аномальную группу R1F2M2 (39,2 %).

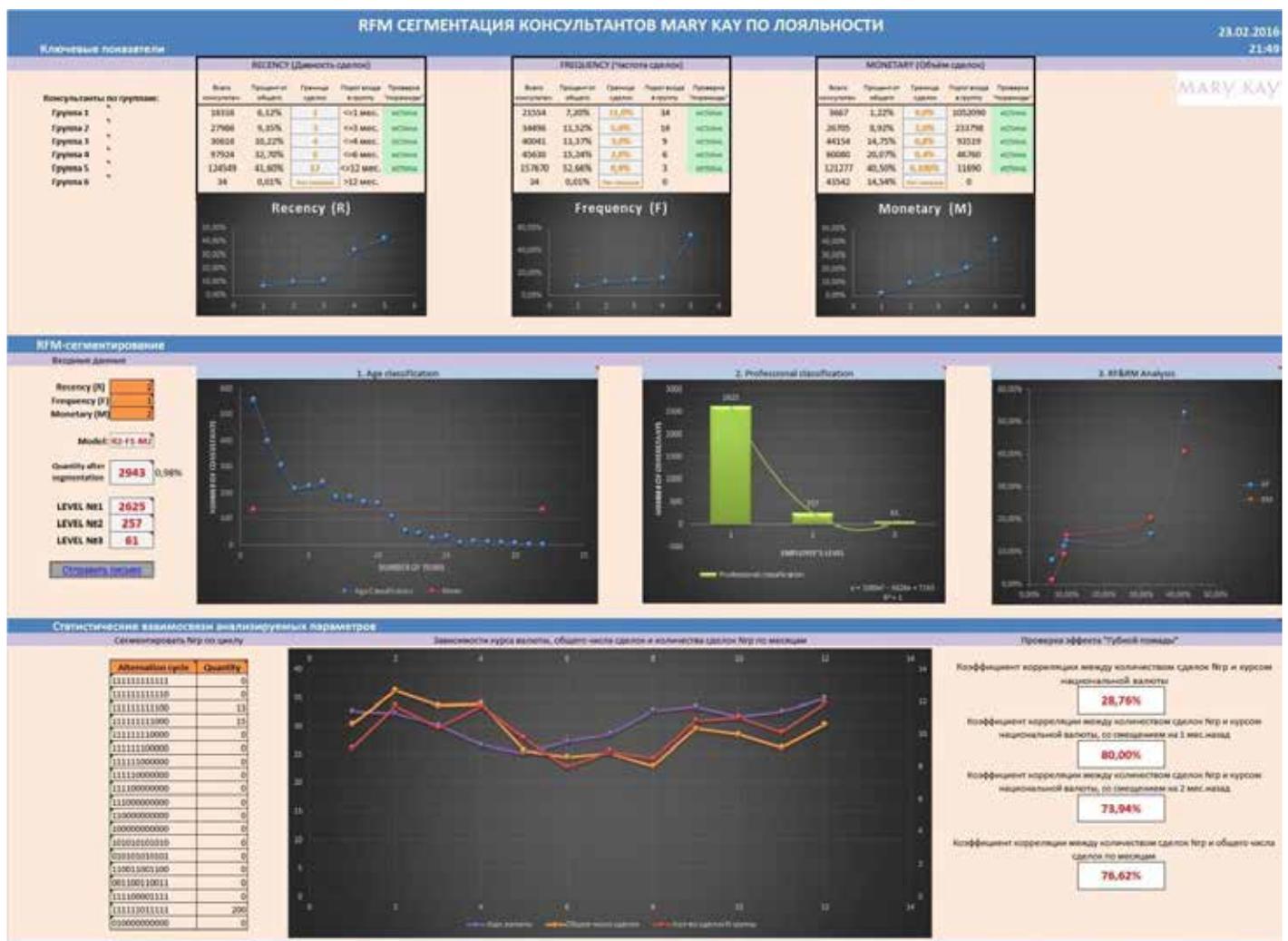


Рис. 8. Общая мнемосхема анализа. Рисунок авторский.

**Рекомендации**

Данная методика классификации применима и к другим группам клиентов модели RFM сегментации. Можно констатировать, что анализ клиентов позволит чётко понимать «климат» клиентской базы, точнее стимулировать клиентов на лояльность и совершение будущей покупки.

### Литература

1. *Бородин И. Д.* Рефлексия в машинном обучении на примере обучения деревьев решений // Аллея науки. 2017. Т. 4. № 9. С. 857-865.
2. *Орлов А.* Машинное обучение для больших данных // Открытые системы. СУБД. № 1. 2016. С. 26-27.
3. *Калыгина В. В.* Международные стратегии продуктовой адаптации на рынке потребительских товаров в условиях глобальной экономики // Наука и бизнес: пути развития. № 2 (32). 2014. С. 57-59.
4. *Чернышева А. М., Долгова А. В.* Организация маркетингового исследования в компании на современном этапе развития маркетинга // Наука и бизнес: пути развития. № 4. 2017. С. 41-43.
5. *Ямпольская Д. О.* Применения маркетинговых метрик для оценки деятельности компаний // Экономика и предпринимательство. № 2-1 (67-1). 2016. С. 969-972.
6. *Fokina E. N., Legotin F. Y., Chernova V. Y.* Evaluation of the IT outsourcing performance in the development of Bigdata systems // Journal of Theoretical and Applied Information Technology. Т. 95. № 21. 2017. С. 5795-5804.

## RUSSIAN MANAGEMENT: A MODERN PERSPECTIVE

Vasily Ganyavin<sup>1</sup>, Sergey Chernikov<sup>2</sup>, Alexander Makarov<sup>3</sup>

### RFM-SEGMENTATION AND CLASSIFICATION OF THE CLIENT BASE FOR THE 1R GROUP BY THE MACHINE LEARNING ALGORITHM

**Abstract.** The relevance of this topic is one of the biggest issues facing marketers - personalize messaging to individual customers and clients to achieve the best response from the recipient. Relevant marketing results include increased customer loyalty, engagement, and reduced non-productive costs. However, without machine learning, it is difficult to process the vast amount of data coming from multiple sources required to predict the success of marketing offerings and incentives for each individual customer. The purpose of this work is to analyze the example of RFM segmentation and machine classification of clients of the company operating in the Moscow market of FMCG cosmetic services. Methods of research of this problem consisted in RFM segmentation of the existing data by methods of the machine analysis. The results of this work have been in understanding the most practical way to implement techniques of machine learning to marketing practices any company using models of segmentation, which proved to be very effective in classifying the target audience on a small homogeneous group of customers with similar behavior and preferences. The practical results are related to the proposal of the most popular areas of marketing application of these technologies, ultimately related to the identification of data characteristics associated with the desired result of the company's interaction with the target audience.

**Key words:** machine learning, big data, Predictive Analytics, RFM segmentation, FMCG cosmetic services.

JEL: M15, M31

- 1 **Ganyavin Vasily Alexandrovich** – Candidate of Technical Sciences. Project Owner LLC Geolink Newtech. 1/2 Lesnaya st., Moscow 125047, Russia. Email: [vganyavin@geolink.ru](mailto:vganyavin@geolink.ru)
- 2 **Chernikov Sergey Yur'yevich** – Candidate of Economic Sciences, Associate Professor of the Department of Marketing, Faculty of Economics. Peoples' Friendship University of Russia (PFUR). 6 Miklouho-Maclay st., Moscow, 117198, Russia. E-mail: [chernikov.rudn@gmail.com](mailto:chernikov.rudn@gmail.com).  
ORCID: 0000-0001-7524-4438; Scopus AuthorID: 57191344141.
- 3 **Makarov Alexander Anatolievich** – Doctor of Technical Sciences, Professor of the Department of Automation and Industrial Electronics. RSU named after A. N. Kosygin (MSUDT). 33/1 Sadovnicheskaya st., Moscow 117997, Russia. Email: [mak68@bk.ru](mailto:mak68@bk.ru)

### References

1. Borodin I. D. Refleksiya v mashinnom obuchenii na primere obucheniya derev'yev resheniy [Reflection in machine learning on the example of learning decision trees]. *Alleya nauki* [Avenue of Scienc], vol. 4, no. 9, 2017, pp. 857-865 (In Russian).
1. Orlov A. Mashinnoye obucheniye dlya bol'shikh dannykh [Machine learning for large data]. *Otkrytyye sistemy. SUBD* [Open systems. DBMS], no. 1, 2016, pp. 26-27 (In Russian).
1. Kalygina V. V. Mezhdunarodnyye strategii produktovoy adaptatsii na rynke potrebitel'skikh tovarov v usloviyakh global'noy ekonomiki [International strategies of product adaptation in the market of consumer goods in the global economy]. *Nauka i biznes: puti razvitiya* [Science and business: ways of development], no. 2 (32), 2014, pp. 57-59 (In Russian).
1. Chernysheva A. M., Dolgova A. V. Organizatsiya marketingovogo issledovaniya v kompanii na sovremennom etape razvitiya marketinga [Organization of marketing research in the company at the present stage of marketing development]. *Nauka i biznes: puti razvitiya* [Science and business: ways of development], no. 4, 2017, pp. 41-43 (In Russian).
1. Yampol'skaya D. O. Primeneniya marketingovykh metrik dlya otsenki deyatelnosti kompaniy [The use of marketing metrics to assess the performance of companies]. *Ekonomika i predprinimatel'stvo* [Economics and Entrepreneurship], no. 2-1 (67- 1), 2016, pp. 969-972 (In Russian).
7. Fokina E. N., Legotin F. Y., Chernova V. Y. Evaluation of the IT outsourcing performance in the development of Bigdata systems // *Journal of Theoretical and Applied Information Technology*. Vol. 95. no. 21. 2017. Pp. 5795-5804.