

СОВРЕМЕННЫЕ МЕТОДЫ И АЛГОРИТМЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ ДЛЯ СОВЕРШЕНСТВОВАНИЯ ОБСЛУЖИВАНИЯ РАБОТЫ БАНКА

Павлов И. С.¹, Константинов П.К.², Файзуллин Р.В.³

¹*Павлов Илья Сергеевич - магистрант, РТУ МИРЭА, Институт Технологий
Управления*

²*Константинов Павел Константинович - магистрант, РТУ МИРЭА,
Институт Технологий Управления*

³*Файзуллин Ринат Васильевич - к.э.н., доцент кафедры «Информационные
технологии в государственном управлении», РТУ МИРЭА
г. Москва, Российская Федерация*

Аннотация: в статье поднимается вопрос эффективности современных методов и алгоритмов интеллектуального анализа данных в банковской сфере. Также, рассматриваются теоретические основы и существующие методы интеллектуального анализа данных. Кроме этого, раскрывается специфика ИАД, его технологии, выявляются типы знания, которые могут быть получены с помощью интеллектуального анализа данных.

Ключевые слова: интеллектуальный анализ данных, Data Mining, Big Data, Задачи Data Mining, методы интеллектуального анализа данных.

STATE-OF-THE-ART DATA MINING METHODS AND ALGORITHMS TO IMPROVE SERVICE

Pavlov I.S.¹, Konstantinov P.K.², Fayzullin R.V.³

¹*Pavlov Ilya Sergeevich - Master student, RTU MIREA, Institute of Management
Technologies*

²*Konstantinov Pavel Konstantinovich - Master student, RTU MIREA, Institute of
Management Technologies*

³*Fayzullin Rinat Vasilovich - Candidate of Economic Sciences, Associate Professor
of the Department of Information Technologies in Public Administration, RTU
MIREA*

Moscow, Russian Federation



***Abstract:** the article raises the issue of the effectiveness of modern methods and algorithms for data mining in the banking sector. Also, the theoretical foundations and existing methods of data mining are considered. In addition, the specifics of the IAD, its technologies are revealed, the types of knowledge that can be obtained using data mining are identified.*

***Keywords:** data mining, Data Mining, Big Data, Data Mining tasks, data mining methods.*

УДК 004.03

Введение

Объемы накопленных данных в настоящее время настолько внушительны, что человеку просто не по силам проанализировать их самостоятельно, хотя необходимость проведения такого анализа вполне очевидна, ведь в этих "сырых данных" заключены знания, которые могут быть использованы при принятии решений. Для того чтобы провести автоматический анализ данных, используется технология Data Mining.

В данной статье рассматриваются возможности и перспективы использования интеллектуального анализа данных для больших объемов данных. Кроме этого, выявляются типы знаний, которые могут быть получены с помощью интеллектуального анализа данных. В качестве методов исследования были изучены и применены методы интеллектуального анализа данных для больших объемов данных. Было произведено их сравнение и выявлены наиболее подходящие, которые бы позволили наиболее полно и точно проанализировать такой большой объем данных, как клиентская база современного банка.

Общее понятие Big Data и определение Data Mining

С каждым днем, количество накопленной информации растет экспоненциальными темпами. Структурированные и неструктурированные объемы данных называют Big Data или, иными словами, большие данные. Кроме самих объемов данных, Big Data представляет собой технологии обработки и использования, поиска данных в больших массивах.



Большие объемы данных обрабатываются для того, чтобы человек мог получить уже обработанные, конкретные результаты для применения в необходимой сфере деятельности.

Современные объемы поступающей информации невозможно обработать традиционными инструментами. Анализ данных позволяет увидеть незамеченные закономерности, которые не доступны взгляду человека, а также оптимизировать все сектора экономики, начиная от государственного управления и до производства и телекоммуникации [1].

Для анализа больших данных используют одну из техник анализа Big Data — Data Mining. Понимание сущности анализа данных изменялось с развитием статистических методов обработки данных, и, в настоящее время, чаще всего говорят об интеллектуальном анализе данных, или еще его называют Data Mining [2, 3]. Интеллектуальный анализ данных (ИАД) - процесс обнаружения новых, интерпретируемых знаний в исходных “сырых” данных, которые необходимы для принятия решений в сферах человеческой деятельности. Иными словами, можно сказать, что ИАД предназначен для обработки и анализа больших объемов данных.

Data Mining – мультидисциплинарная область, возникшая и развивающаяся на базе таких наук как статистика, распознавание образов, искусственный интеллект, теория баз данных и другие. Процесс нахождения скрытых закономерностей в существующих данных.

Основу Data Mining составляет концепция шаблонов, представляющих собой ряд закономерностей. Цель поиска закономерностей – представление данных в виде, отражающем искомые процессы. В результате обнаружения скрытых закономерностей, решаются задачи Data Mining. Одной из сфер применения интеллектуального анализа данных являются временные ряды [4, 5].

Тем не менее, технология Data Mining не может заменить аналитика, она дает мощный инструмент для улучшения и облегчения его работы.

Выделяют 5 типов закономерностей, которые еще называют задачами Data Mining [6]:



1. Ассоциация. При поиске решения задачи отыскиваются закономерности между связанными событиями в наборе данных.
2. Последовательность. Последовательность подобна ассоциации, но главной ее целью является установление закономерностей между событиями, связанными во времени или, другими словами, событиями, происходящими с некоторым определенным интервалом во времени.
3. Классификация. Данная задача делит множество объектов на заданные группы, называемые классами, внутри каждой из которых они предполагаются похожими друг на друга, имеющими примерно одинаковые свойства и признаки.
4. Кластеризация. Кластеризация является продолжением классификации. Ее особенностью является то, что классы объектов изначально не predetermined. В результате кластеризации происходит разбиение объектов на группы. Зачастую удобнее вырабатывать подход к какому-то виду объектов (кластеру), чем к каждому объекту отдельно [7, 8].
5. Прогнозирование или регрессия. Методом прогнозирования решаются задачи на основе особенностей исторических данных оцениваются пропущенные или будущие значения числовых показателей. Прогнозирование применяется в математической статистике, нейронных сетях и другом.

Основываясь на использовании открытых данных, технология «Big Data» применяется в таких банках, как «Банк ВТБ (ПАО)», ПАО «Сбербанк России», АО «АЛЬФА-БАНК», ПАО «БАНК УРАЛСИБ», АО «ОТП Банк», АО «Газпромбанк», АО «Райффайзенбанк». В целом, поступающие и имеющиеся в наличии данные можно охарактеризовать как потенциальную «золотую жилу» любого банка, но, чтобы добраться до настоящего золота, нужно качественно обработать внушительный объем разнородного материала и извлечь из него информацию о закономерностях и тенденциях, которые наиболее важны для принятия решений, связанных с деятельностью банка.



Далее, необходимо рассмотреть применение технологий и ключевых закономерностей Data Mining в АО «АЛЬФА-БАНК»

Аналитики компании АО «АЛЬФА-БАНК» обрабатывают огромные массивы данных для получения аналитических сведений, на основании которых принимаются те или иные решения.

В стратегии на 2014-2018 годы, АО «АЛЬФА-БАНК» заявил об оптимизации работы и автоматизации процесса обслуживания банка, с целью перехода в цифровой бизнес. Оптимизация работы представляет собой сокращение некоторых должностей, в первую очередь, это сокращение слоев управления. По причине увеличения пользователей интернет-банка происходит и сокращение отделений и офисов АО «АЛЬФА-БАНК». Клиенты все активнее используют Мобильный банк, и все меньше обращаются за консультацией к сотрудникам банка.

Можно справедливо утверждать, что информационные системы позволяют облегчить внутриорганизационные процессы и повысить качество производства или услуг. Технологии Data Mining, как один из элементов цифровизации позволяет достичь совершенного нового уровня управления качеством, перенаправить человеческие силы человеческих ресурсов на более значимые задачи [9].

Далее, необходимо привести анализ ключевых методов интеллектуального анализа данных на примере совершенствования обслуживания корпоративных клиентов в АО «АЛЬФА-БАНК».

Ключевыми закономерностями Data Mining для анализа имеющейся базы в банковской сфере являются: метод классификации и метод прогнозирования.

1. Метод классификации.

Классификация - наиболее простая и, одновременно, наиболее часто решаемая задача Data Mining. Классификация – это системное распределение предметов, процессов, видов, типов по каким-либо признакам для удобства при использовании. Группировка исходных параметров и расположение в определенном порядке, отражающем степень сходства [10].



Классификация – это закономерность, позволяющая делать вывод относительно определения характеристик конкретной группы. Задача классификации – предсказание категориальной зависимости переменной на основании выборки непрерывных и/или категориальных переменных.

Метод классификации в АО «АЛЬФА-БАНК» применяется при составлении воронки продаж по продуктам. В базе присутствует информация об имеющихся продуктах клиентов. Классификация отлично подходит при отборе клиентов для продажи продукта – эквайринг.

Для наиболее эффективного использования технологи, АО «АЛЬФА-БАНК» необходимо предоставлять услугу эквайринга на аутсорсинг государственным компаниям и брать комиссию от 1% до 4% суммы платежей по терминалу. Организация оплачивает комиссию банку в конце месяца, а в связи с тем, что расчетный счет находится в казначействе, данный процесс занимает длительное время. Именно поэтому многие клиенты отказываются от данной услуги, прибегают к проверенному способу: оплатой наличными средствами. Однако, в современном мире, возможность оплаты во многих государственных предприятиях по всей России – необходима. Для того чтобы реализовать подобную технологию, необходим пересмотр текущих бизнес-процессов в государственных компаниях.

Таким образом, применив метод классификации, отделив организации с низким объемом выручки от остальных организаций, можно получить воронку продаж по эквайрингу для отработки клиентскими менеджерами. Данная воронка позволит, во-первых, привлечь новых клиентов в АО «АЛЬФА-БАНК», а во-вторых, повысить лояльность обслуживающихся в банке клиентов, тем самым увеличится уровень обслуживания корпоративных клиентов в целом.

2. Метод прогнозирования.

Далее, не менее значимая задача Data Mining - прогнозирование. Задачи прогнозирования решаются в различных областях жизни: наука, экономика, производство и другие. Развитие метода прогнозирования связано с ростом



объемов хранимой информации и усложнением методов и алгоритмов прогнозирования [11, 12], реализованных в методах Data Mining.

Прогнозирование в АО «АЛЬФА-БАНК» направлено на определение тенденции динамики события или объекта на основе имеющихся ретроспективных данных, иными словами, анализ состояния в прошлом и настоящем.

Так, используемая в АО «АЛЬФА-БАНК» интеллектуальная система безопасности CASIOPEA 3.0 на базе существующих у нее прогнозных и поведенческих моделей, выявляет аномальные ситуации не только в системах интернет-банкинга, но и в поведении клиентов, ситуациях с банкоматами и POS-терминалами.

Источниками данных являются данные о SLA, транзакциях, информации с серверов, от продавцов, от службы техподдержки, информация о технических операциях с банкоматами и POS-терминалами (например, сведения об их замене).

Также, в АО «АЛЬФА-БАНК», с 2020 года начинает реализовываться инициатива внедрения модели прогнозирования анализа продаж пакетов услуг банка на потребительском рынке. Такая модель учитывает всю историю продаж услуг-аналогов, сравнивая каждый критерий и выстраивая собственную оценку прибыльности услуги за определенный промежуток времени. Данная технология позволит упростить работу финансового и маркетингового отдела, сократив время на прогнозирование продаж и подборку вариантов пакетов услуг для клиентов.

Заключение

В ходе выполнения работы были изучены методы интеллектуального анализа данных. По результатам, было выбрано два наиболее эффективных метода для банковской сферы: классификация и прогнозирование. Были приведены актуальные данные об использовании АО «АЛЬФА-БАНК» этих методов Big Data.



В ближайшие годы банкам, не использующие современные методы анализа данных, будет тяжело удержать большую часть своих клиентов, так как не смогут в полной мере удовлетворить и определить наиболее выгодное предложение.

Именно поэтому, технологии Big Data являются вынужденной мерой для того, чтобы соответствовать мировым тенденциям, в свою очередь, для успешного внедрения и единообразного применения кредитными организациями рассматриваемой технологии.

Сочетая новые источники данных с аналитикой реального времени и поведенческой информацией, компании смогут разработать новое поколение приложений, способных оперативно адаптироваться и обучаться. Готовые комплексные решения по классификации и прогнозированию, интегрирующие такие сервисы, как аналитика, исследование, подготовка и интеграция данных, упрощают разработку новых продуктов.

В результате изучения методов интеллектуального анализа данных можно сделать вывод, что данные методы подходят для аналитики больших объемов данных в банковской сфере.

Список литературы

1. Курс: Анализ данных [Электронный ресурс]: [б.и.], 2013, - Электрон. текстовые дан. on-line.
2. Левитин А.В. Алгоритмы. Введение в разработку и анализ. М.: Вильямс, 2006. 576с.
3. Поручиков М.А. Анализ данных: учебное пособие - Самара: Изд- во Самарского университета, 2016. - 88 с.
4. Файзуллин, Р. В. Перспективы автоматизации выявления причинно-следственных связей между финансовыми временными рядами // Наука и инновации XXI века: сборник статей по материалам V Всероссийской конференции молодых ученых. - В 3 т. - 2018. - С. 56-59.
5. Тормозов, В. С. Структурная оптимизация персептрона с помощью аппарата эволюционного моделирования для задачи прогнозирования значений временных рядов / В. С. Тормозов, А. Л. Золкин, К. А. Василенко // Промышленные АСУ и контроллеры. - 2020. - № 8. - С. 40-45.
6. Тебайкина Н.И. Методическое пособие «Бизнес-модель компании» [Презентация]



/ УрФУ 2015. Слайды 10-35.

7. Матвеева И. В., Хоменко Е. Б. Современные тенденции трансформации институтов предпринимательства: кластеризация и информатизация // Вестник ИжГТУ имени М. Т. Калашникова. 2016. № 2(70). С. 36-40.
8. Пермякова Т.В., Файзуллин Р.В. Иерархическая кластеризация нефтеперерабатывающих заводов России // Научный журнал НИУ ИТМО. Серия: Экономика и экологический менеджмент. 2014. № 2. С. 7. [Электронный ресурс] - URL: <http://economics.open-mechanics.com/articles/1011.pdf>. (Дата обращения: 20.05.2021).
9. Дмитриев Н.Д. Роль цифровой трансформации в информатизации менеджмента качества // Информационно-телекоммуникационные системы и технологии: материалы научно-практической конференции. 2019. С. 80-81.
10. Чернышова Г.Ю. Интеллектуальный анализ данных: учеб. пособие для студентов специальности 080801.65 «Прикладная информатика (в экономике)» - Саратовский государственный социально-экономический университет. – Саратов, 2012. – 92 с.
11. Касаткина Е.В., Вавилова Д.Д. Информационно-аналитическая система прогнозирования обобщающих показателей социально-экономического развития // Проблемы управления. 2015. № 4. С. 25-34.
12. Чазов Е.Л., Грахов В.П., Криворотов В.В., Симченко О.Л. Повышение эффективности планирования как основа управления инвестиционной деятельностью промышленного предприятия // Энергетика. Известия высших учебных заведений и энергетических объединений СНГ. 2019. Т. 62. № 1. С. 88-100.

