

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тольяттинский государственный университет»

Институт математики, физики и информационных технологий
(наименование института полностью)

Кафедра «Прикладная математика и информатика»
(наименование)

09.04.03 Прикладная информатика
(код и наименование направления подготовки)

Информационные системы и технологии корпоративного управления
(направленность (профиль))

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ)

на тему «Моделирование информационной системы управления предприятием на основе интеллектуального анализа данных»

Студент

С.А. Фадеев

(И.О. Фамилия)

(личная подпись)

Научный
руководитель

к.п.н, доцент, Е.А. Ерофеева

(ученая степень, звание, И.О. Фамилия)

Оглавление

Введение.....	4
Глава 1 Современное состояние проблемы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных	9
1.1 Понятие Data Mining.....	9
1.2 Исследование методов интеллектуального анализа данных применительно к CRM-системам	11
1.3 Обзор и анализ существующих CRM-систем	14
Глава 2 Методологические основы моделирования CRM-систем на основе интеллектуального анализа данных	20
2.1 Аналитический обзор методов и технологий интеллектуального анализа данных.....	20
2.2. Алгоритм k-средних.....	23
2.3 Алгоритм C4.5	25
Глава 3 Разработка модели системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных	31
3.1 Выбор технологии для анализа бизнес-процесса управления взаимоотношениями с клиентами	31
3.2 Концептуальное моделирование бизнес-процесса управления взаимоотношениями с клиентами	32
3.3 Обоснование выбора технологии моделирования CRM-системы.....	38
3.4 Обоснование выбора компонентов архитектуры CRM-системы.....	40
3.5 Моделирование CRM-системы.....	40
3.6 Обоснование выбора и описание основных средств реализации модели CRM-системы	47

Глава 4 Реализация системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных	50
4.1 Описание основного принципа работы CRM-системы	50
4.2 Реализация информационной безопасности CRM-системы.....	60
4.3 Расчет показателя эффективности CRM-системы.....	62
Заключение	64
Список используемой литературы и используемых источников.....	66
Приложение А. Диаграммы последовательности в нотации UML.....	71

Введение

В управлении предприятием одним из главных процессов является умение взаимодействовать с клиентами. Повышение объема продаж, сокращение материальных ресурсов на обслуживание, повышение прибыли предприятия за счёт увеличения лояльности клиента, реализация управления, направленного на качественную работу с клиентами, на решение их потребностей, осуществляются с помощью CRM-систем (Customer Relationship Management Systems).

Сбор клиентских данных, а также последующий поведенческий анализ этих данных и прогнозирование, благоприятно сказываются на возрастании количества продаж и оптимизации коммуникации с клиентами. Все это становится возможным и достигается благодаря использованию передовых информационных технологий при реализации CRM - концепции.

Современный рынок программных продуктов предоставляет возможность выбора готовых решений в области CRM-систем как для маленьких и средних компаний в качестве отдельных ИС для автоматизации управления заказами клиентов, так и для крупных компаний в виде КИС известных вендоров. Стоит обратить внимание, что в малых компаниях широкие возможности функциональности CRM-систем упускаются из виду в следствие того, что преимущественно компании используют программное обеспечение в качестве АИС управления контактами, в которой структурируются только клиентские данные.

Системы управления взаимоотношениями с клиентами можно классифицировать по уровню обработки информации на две группы:

- операционный уровень, ключевая миссия которого в процессе непосредственного контакта с клиентом - повысить его лояльность;
- аналитический уровень, направленный на глубокий анализ данных о клиентской базе предприятия в целях автоматизации управленческих процессов. Например, анализ успешности продаж

по периодам времени, по группам товаров (по услугам), сегментирование клиентов, классификация клиентов.

Для решения задач привлечения и удержания клиентов в ходе взаимодействия с ними предназначен операционный уровень CRM. Аналитические CRM оказывают содействие в разработке стратегии маркетинговой политики, основываясь на анализе имеющихся данных.

Сложные аналитические системы как отдельные ИС не осуществляются в связи с большими затратами на их разработку и внедрение. Такие возможности присутствуют лишь в функционале КИС. Data Mining лежит в основе таких аналитических систем.

Из вышесказанного можно сделать вывод, что задача разработки CRM-систем, которые при низкой себестоимости реализуют основные цели политики управления взаимодействием с клиентами на основе современных технологий интеллектуального анализа, является весьма **актуальной**.

Целью работы является разработка модели системы управления взаимоотношениями с клиентами на основе Data Mining.

Объект исследования - система управления взаимоотношениями с клиентами на основе Data Mining.

Предмет исследования - модель системы управления взаимоотношениями с клиентами на основе Data Mining.

Гипотеза исследования: использование предложенной в магистерской работе модели CRM-системы на основе интеллектуального анализа данных будет эффективно влиять на управление предприятием в целом.

Задачи исследования:

- исследовать методы и алгоритмы интеллектуального анализа данных;
- провести системный анализ методов и средств обработки информации в системах управления взаимоотношениями с клиентами;
- осуществить выбор методологических подходов к моделированию

системы управления взаимоотношениями с клиентами на основе Data Mining с помощью сравнительного анализа;

- проанализировать предметную область с целью выявления объектов автоматизации и аналогов существующих систем;
- разработать модель информационной системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных;
- оценить эффективность реализованной модели CRM-системы на основе Data Mining.

Теоретическую основу исследования составляют научные труды отечественных и зарубежных исследователей в области информационных систем управления взаимоотношениями с клиентами (Буянский С.Г., Митина О.А., Bourgeois D. и др.).

Для решения поставленных в исследовании задач планируется использовать методологический аппарат, включающий в себя такие **методы исследования**, как:

- системный анализ;
- структурный и объектно-ориентированный анализ;
- структурное и объектно-ориентированное моделирование.

Новизна исследования заключается в разработке новой модели информационной системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных.

Практическая значимость исследования заключается в возможности практического применения предлагаемой модели для построения CRM-системы на основе интеллектуального анализа данных, эффективно влияющей на управление предприятием в целом.

Исследование проводилось с 2019 года по 2021 год.

Основные этапы исследования:

На первом этапе (констатирующем этапе) была определена тема исследования, проведен обзор отечественных и зарубежных трудов по теме

исследования, выявлены проблема и актуальность данного исследования, была выдвинута гипотеза, поставлена цель и определены задачи для ее достижения, сформулированы предмет и объект исследования.

На втором этапе (поисковом этапе) был проведен анализ методологий моделирования систем управления взаимоотношениями с клиентами на основе интеллектуального анализа данных, была разработана модель CRM-системы на основе интеллектуального анализа данных, была написана и опубликована научная статья по теме исследования в международном научно-практическом журнале «Моя профессиональная карьера».

На третьем этапе (оценка эффективности) проводилась оценка эффективности и проверка адекватности предлагаемой модели системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных, были подведены итоги, сделаны выводы о результатах работы.

На защиту выносятся:

- модель CRM-системы на основе интеллектуального анализа данных;
- результаты проверки адекватности модели CRM-системы на основе интеллектуального анализа данных.

По теме исследования опубликована статья:

Фадеев С.А. Технология Data Mining как инструмент оптимизации взаимоотношений с клиентами // международный научно-практический журнал «Моя профессиональная карьера», 2020.

Диссертация состоит из введения, четырех глав, заключения и списка литературы.

Первая глава посвящена обзору актуального состояния проблемы на сегодняшний день. Изучены существующие методы интеллектуального анализа данных применительно к CRM-системам. Проведен обзор и анализ существующих CRM-систем.

Во второй главе сделан анализ методологических основ моделирования систем управления взаимоотношениями с клиентами на основе

интеллектуального анализа данных. Определен набор методов для проведения интеллектуального анализа данных на основе сравнительного анализа методов Data Mining с последующим описанием их алгоритма.

В третьей главе представлена разработка модели управления взаимоотношениями с клиентами на основе интеллектуального анализа данных.

Четвертая глава посвящена процессу реализации системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных и оценке ее эффективности.

В заключении работы представлены результаты проведенного исследования.

Работа изложена на 78 страницах и включает 49 рисунков, 7 таблиц, 41 источник.

Глава 1 Современное состояние проблемы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных

1.1 Понятие Data Mining

Проблеме управления взаимоотношениями с клиентами на основе Data Mining посвящены научные труды отечественных (Буянский С.Г., Митина О.А. и другие) и зарубежных исследователей (Bourgeois D. и другие).

«Технология интеллектуального анализа данных (Data Mining) является технологией, воплощающей аналитические функции систем управления взаимоотношениями с клиентами» [22].

Формированию больших массивов данных способствовали глобальная информатизация бизнес-процессов, быстрый темп образования данных, расширение возможностей Интернета, выступающего в качестве еще одного канала получения информации и способа коммуникации с клиентами.

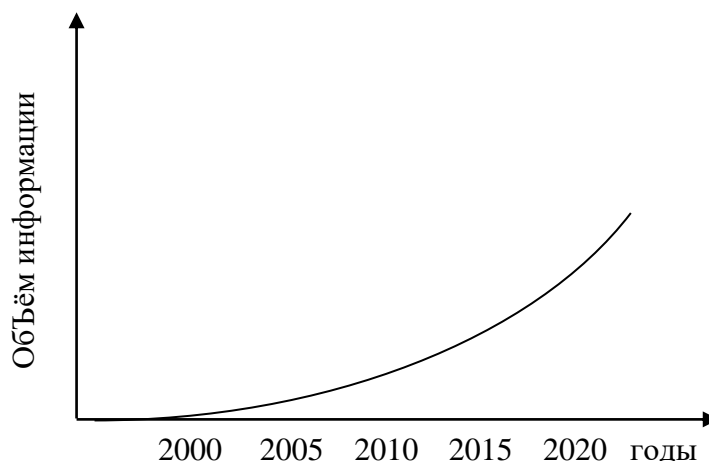


Рисунок 1 - Динамика изменения объема информации

В настоящее время «методы математической статистики не перспективны в задачах отыскания полезной информации в данных по причине усреднений по выборке, которые сводятся к операциям над

фиктивными величинами» [33] (например, средний балл по экзамену по русскому языку в Самарской области). «Математическую статистику рациональнее применять для подтверждения гипотез, сформулированных заранее. Основная идея паттернов (шаблонов), которые включают фрагменты многоаспектных взаимоотношений в данных (т.е. закономерности, принадлежащие подвыборкам данных, которые могут быть доступно интерпретированы), является базисом современной технологии Data Mining. Обнаружение шаблонов производится с помощью большого количества известных предположений о виде распределений значений анализируемых показателей и структуре выборки» [33].

На сегодняшний день в трудах отечественных и зарубежных исследователей встречается несколько определений понятия Data Mining:

- «Data Mining - это процесс выделения из данных неявной и неструктурированной информации и представления её в виде, пригодном для реализации» [40];
- «Data Mining - это процесс анализа, выделения и представления детализированных (detailed data) данных неявной конструктивной информации для решения проблем бизнеса» [35];
- «Data Mining - это процесс автоматического выделения действительной, эффективной, ранее не известной и совершенно понятной информации из больших баз данных и использование её для принятия ключевых бизнес-решений» [37].

В диссертации мы будем опираться на «определение Григория Пиатецкого-Шапиро, одного из основателей технологии интеллектуального анализа данных. Предложенная им трактовка содержит четыре ключевых момента, присутствующих в приведенных выше определениях» [19]:

«Data Mining - это процесс обнаружения в сырых данных:

- ранее неизвестных;
- нетривиальных;
- практически полезных;

- доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности» [36].

1.2 Исследование методов интеллектуального анализа данных применительно к CRM-системам

Цель разработки и внедрения системы взаимоотношений с клиентами - как накопление клиентской информации и реализация регламентных процедур, так и улучшение коммуникации с клиентами:

- выделение целевой аудитории;
- определение потребительских характеристик;
- анализирование алгоритма продаж;
- анализ и оценка эффективности маркетинговой деятельности предприятия.

Без последующего глубокого анализирования сбор информации, накапливающейся со временем, не имеет оптимальной эффективности, отсюда имеется противоречие: обилие информации, а на практике от нее - минимум выгоды. Только аналитическая деятельность приведет к преобразованию данных в знания.

Анализирование данных берет начало как с визуализации (использование графиков, гистограмм, диаграмм), так и с аналитической отчетности, применения нестандартных запросов и использования технологии OLAP (On-Line Analytical Processing). Эти практические инструменты направлены на изучение только явных связей.

В настоящее время предпринимательство характеризуется нетривиальными зависимостями, колоссальными объемами данных, стремительными преобразованиями. Инновационные инструменты анализа помогают обнаружить в больших объемах данных сложные закономерности.

По сути это является единственным вариантом получить пользу из накопленных сведений и преобразовать знания в превосходство перед конкурентами.

«Data Mining - это технология, которая ориентирована на поиск в базах данных неочевидных, объективных и полезных на практике закономерностей» [14].

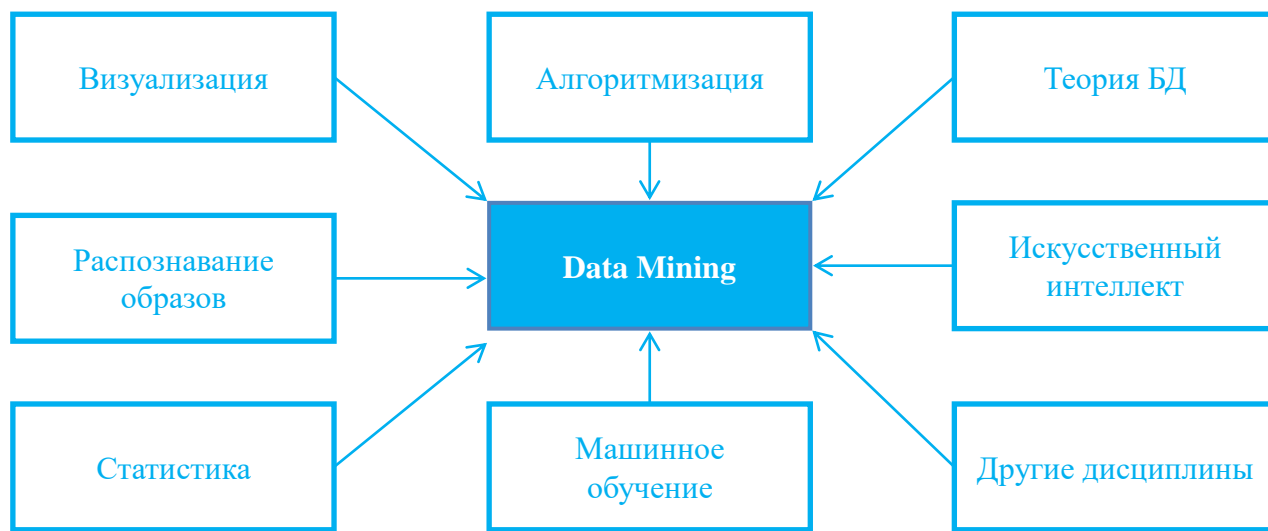


Рисунок 2 – Науки, входящие в Data Mining

Технология «Data Mining» сводится к решению пяти групп задач:

- ассоциация,
- классификация,
- кластеризация,
- регрессия,
- последовательность.

Ассоциация - это анализ транзакций, другими словами «выявление закономерностей между связанными событиями. Выявление зависимости, что из события А с определенной вероятностью следует событие В» [39]:

- прогнозирование действий покупателя и рекомендация продукта, которое вероятнее всего привлечет его внимание;
- «расположение продуктов на стеллажах, в каталогах товаров;

- перекрестные продажи – стимулирование продажи релевантной продукции за счет продажи другой;
- оптимизация имеющихся в наличии запасов продукция для дальнейшей реализации» [28].

Классификация - обнаружение функциональной зависимости между дискретным выходным атрибутом и входными атрибутами:

- предсказание успеха двустороннего договора;
- исследование рисков: предоставлять кредит или нет на продукцию;
- анализ бонусов: какую группу клиентов обеспечить скидками;
- оценка эффективности рекламной компании;
- анализ перспективности покупателей.

Кластеризация - разбиение объектов на группы схожих элементов. Данный метод заключается в анализе одних объектов по подобию поведения других:

- кластеризация продуктов, обнаружение релевантных продуктов;
- группировка клиентов по схожести поведения;
- анализ спроса в зависимости от запросов клиентов.

Регрессия - обнаружение функциональной зависимости между непрерывным выходным атрибутом и входными атрибутами. Дает возможность оценить вероятность появления события или его численное значение:

- оценка возможности повторных продаж;
- оценка реакции потребительского спроса на изменение цены товара;
- прогнозирование спроса клиентов;
- исследование факторов влияющих на спрос;
- вычисление загруженности склада, магазина.

Последовательность - анализ временных закономерностей между транзакциями. Выявление зависимости, что после события А спустя определенное время наступит событие В:

- прогнозирование вероятного поведения клиента;

- исследование нужд клиента, в дополнительных сопутствующих товарах или услугах;
- повторные продажи, оценка более возможного времени потребности в модификации.

Разрешение многих вопросов, касающихся взаимоотношений с клиентами, сводится к практическому применению методов Data Mining:

- поддержание продаж;
- анализ предпочтений клиента;
- оценка эффективности работ;
- прогнозирование потребительского спроса;
- Direct Mail (прямая почтовая рассылка);
- оценка эффективности управляющих.

1.3 Обзор и анализ существующих CRM-систем

Рассмотрим существующие CRM-системы.

Битрикс24 - данная CRM реализуется с 2009 года, целевая аудитория направлена на РФ, развивается стремительно и непоколебимо. Битрикс24 имеет в своем комплекте богатый функционал. Данная CRM позволяет оптимизировать работу не только с клиентами, но и также способствует ведению процессов внутри организации. Существуют как «коробочные» решения, так и облачные.

Перечень ключевых функциональных модулей в продукте Битрикс24 можно классифицировать на 4 основные группы, такие как: проектная деятельность (модуль «Задачи и проекты»), CRM, коммуникация и бизнес-процессы. Это, пожалуй, самые необходимые возможности Битрикс24, которыми пользуются более 75% от всех пользователей.

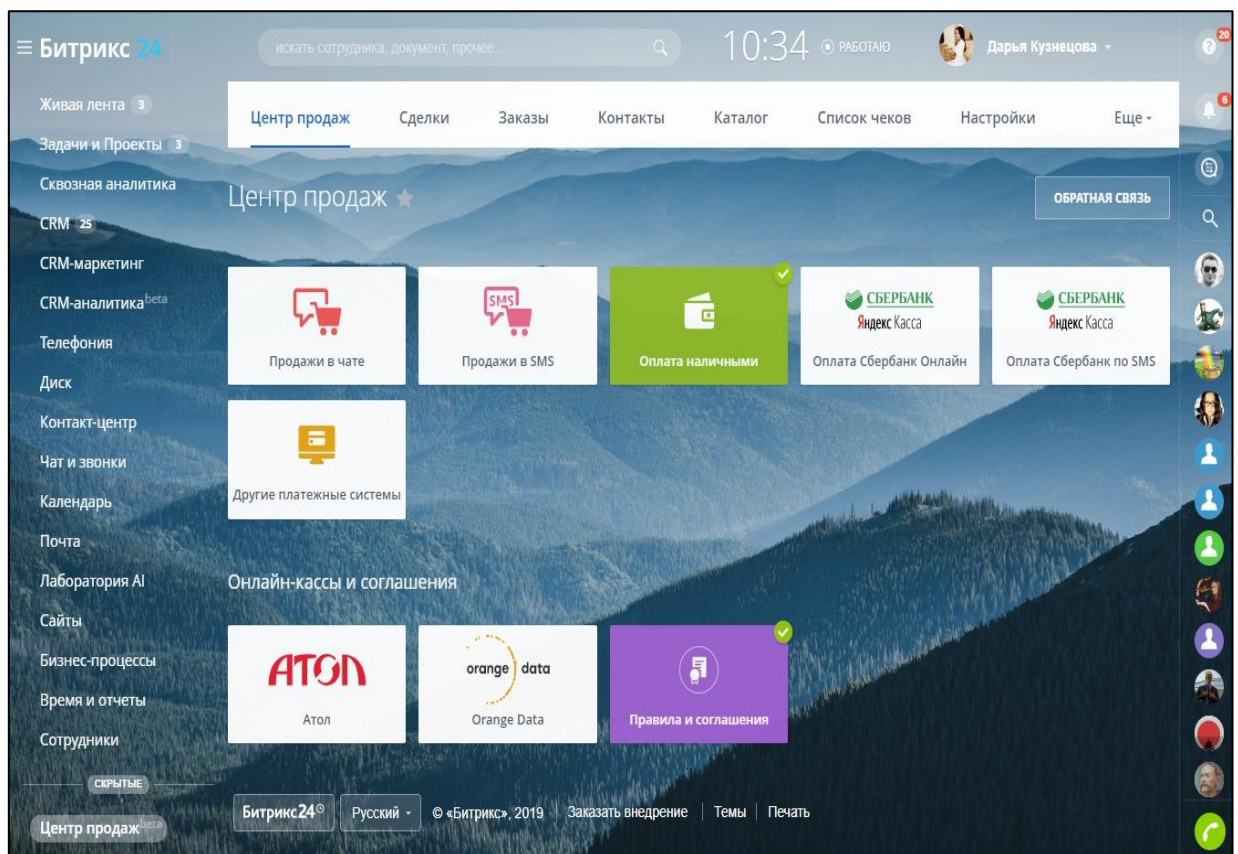


Рисунок 3 - Интерфейс Битрикс 24

Достоинства:

- самый широкий спектр функциональных возможностей;
- интуитивно понятная работа с файлами в форматах doc, pdf, xls и другими;
- возможность видеоконференций;
- присутствует возможность организации бесплатного рабочего почтового ящика;
- элементы игровых походов (бейджи);
- существует возможность интеграции с различными социальными сетями;
- интеграция с программными продуктами линейки 1С и другими известными вендорами;
- присутствует телефония.

Недостатки:

- в связи с тем, что весьма богат функционал, процесс обучения затруднителен;
- отсутствие дружелюбного интерфейса;
- недостаточная архитектура CRM: нет возможности ведения договоров, контрактов, формировать шаблоны документов.

CRM «Простой бизнес» реализуется с 2008 года. Изначально эксплуатировалась как частное ПО, но присутствовал спрос, и в 2011 году ИТ-команда выпустила новый релиз на CRM-рынок.

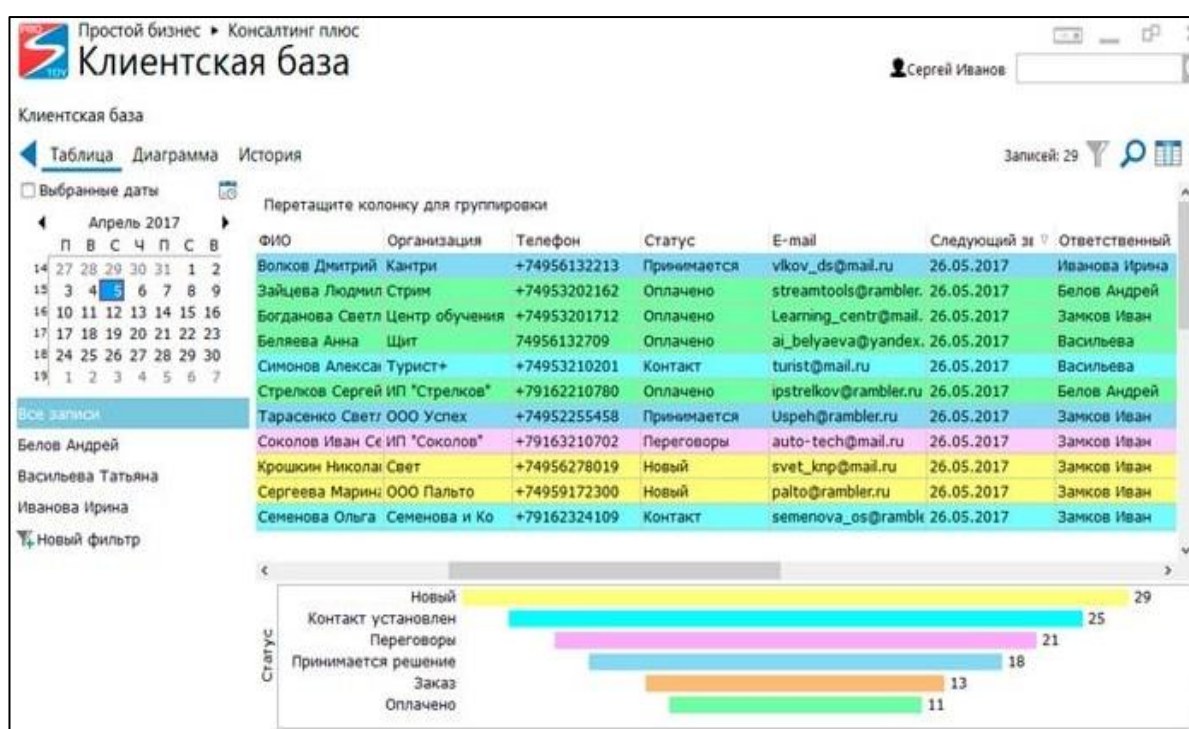


Рисунок 4 - Интерфейс CRM «Простой бизнес»

Достоинства:

- относительно приемлемая стоимость;
- все внимание сконцентрировано на основном бизнес-процессе без излишеств;
- присутствует IP-телефония;
- есть возможность демо-доступа в течении месяца;

- кроссплатформенная, имеет онлайн-версию;
- интегрированный конструктор веб-портала;
- достойная коммуникация (возможность видеоконференций);
- присутствует визуализация аналитики;
- походит для различных сфер бизнеса.

Недостатки:

- существует привязка к приложениям на Windows;
- нет возможности интегрироваться с 1С.

amoCRM реализуется на рынке уже 16 лет. По сравнению с Битрикс24, amoCRM не настолько многофункциональна, ей присуща простота. Есть интуитивно понятная воронка продаж, а также телефония.

Название сделки	Основной контакт	Телефон (контакт)	Статус сделки	Ближайшая за...	Источник	Бюджет, руб
Новая заявка	Виктор	+7 916 553 85 66	Проведена встреча		ВК Довжик	0
Новая заявка	Станислав Сергатский		Думают	27.10.2014 23:59	ВК Довжик	12000
Айти компания	Умид	+7 965 599 68 93	Думают	Завтра, 23:59	Introvert.bz	40000
Консалтинговая компания	Ален	+7 705 446 77 77	Думают	Завтра, 23:59	Introvert.bz	40000
Фотообои	Владимир	+7 928 110 88 11	Назначена встреча	Завтра, 11:00	Introvert.bz	70000
Новая заявка	Василий Батарейкин	+7 903 257 99 77	Думают	29.10.2014 23:59	Рекомендация	50000
Каркасные дома	Виктория Аксенова	+380 95 273 73 41	Проведена встреча	Завтра, 23:59	Рекомендация	0
Франшиза лапшичной	Снежанна	7 926 946 80 11	Думают	Завтра, 23:59	Introvert.bz	40000
Продажа окон	Юрий Михайлов		Проведена встреча	Завтра, 23:59	Биржа фриланса	1000
Кофе с собой	Павел Салтанов		Новая заявка		ВК Довжик	0
1дн Косметика для сал...	Евгения Замятина		Непонятно	21.10.2014 23:59	ВК Довжик	0
1дн Ритейл и недвижка	Руслан Морозов		Назначена встреча	21.10.2014 23:59	ВК Довжик	0
Интернет агентство	Илья Бабков	7 918 8 999 123	Думают	29.10.2014 23:59	ВК Довжик	50000
1дн Мебельные салоны	Илья Погонин		Назначена встреча	21.10.2014 23:59	ВК Довжик	40000
2дн Туризм	Имя не указано		Назначена встреча	20.10.2014 23:59	Introvert.bz	70000
Продажа конструкторов и...	Сергей	7 967 033 74 80	Думают	Завтра, 23:59	Introvert.bz	25000

Рисунок 5 - Интерфейс amoCRM

Достоинства:

- дружелюбный, эргономичный интерфейс;
- грамотно спроектированная архитектура CRM;
- присутствует телефония;
- существует интеграция с огромным спектром сервисов.

Недостатки:

- недостаточно богатый функционал;
- нет коробочного решения, только «облако»;
- тарифная политика подразумевает покупку лицензии минимум на пол года.

В таблице 1 представлен сравнительный анализ рассмотренных CRM-систем.

Таблица 1 - Сравнительный анализ CRM-систем

Критерии оценки	Битрикс24	CRM «Простой бизнес»	amoCRM
наличие бесплатной версии	-	-	-
возможность записи просмотра прайс-листа	+	+	+
возможность заказать обратный звонок	+	-	+
облачная версия	+	-	+
работа в режиме онлайн	+	+	+
возможность проверить статус заявки онлайн	-	+	-
Итого	5 из 7	4 из 7	5 из 7

Опираясь на результаты анализа в таблице 1, видно, что рассмотренные CRM-системы не в полном объеме удовлетворяют необходимым требованиям. Изучив аналоги CRM-систем, можно выделить, что они имеют специфический набор функций.

Таким образом, было принято решение о целесообразности разработки модели CRM-системы, удовлетворяющей требованиям.

Выводы к главе 1

Проблематика управления взаимоотношениями с клиентами на основе Data Mining рассматривается в трудах как отечественных, так и зарубежных ученых, при этом следует отметить дефицит работ, посвященных данной теме.

Технология «Data Mining» сводится к решению пяти групп задач: ассоциация, классификация, кластеризация, регрессия, последовательность.

Существующие CRM-системы не удовлетворяют всем требованиям эффективного управления взаимоотношениями с клиентами.

Для эффективного влияния на управление предприятием в целом необходимо внедрить CRM-систему на основе Data Mining, разработка модели которой является целью данной магистерской работы.

Глава 2 Методологические основы моделирования CRM-систем на основе интеллектуального анализа данных

2.1 Аналитический обзор методов и технологий интеллектуального анализа данных

Аналитический обзор и сравнительный анализ методов Data Mining в данной магистерской диссертации будут осуществляться «на основе распределения методов интеллектуального анализа данных по задачам. Такое распределение представляет собой деление методов на две группы» [15]:

- методы интеллектуального анализа данных, направленные на кластеризацию и классификацию (задачи сегментации);
- методы интеллектуального анализа данных, направленные на получение в результате описания и прогноза.

Далее будут рассмотрены детально данные группы.

К группе описательных методов, направленных на отыскание паттернов или образцов, описывающих данные, которые по мнению аналитика подвержены истолкованию, принадлежат такие инструменты, как:

- «алгоритмы k-медианы и k-средних - алгоритмы классификации без учителя (кластерный анализ), возможно использование для выявления аномалий в данных на основе меры расстояний;
- методы кросс-табличной визуализации - методы выявления закономерностей в данных;
- самоорганизующиеся карты Кохонена - нейронная сеть с обучением без учителя;
- иерархические методы кластерного анализа - алгоритмы соотнесения данных на основе графов» [31].

К группе «прогнозирующих методов, которые для прогнозирования неизвестных (пропущенных) или будущих значений одних (целевых) переменных используют значения других переменных» [38], принадлежат

такие инструменты, как:

- «деревья решений - прогнозный классифицирующий алгоритм;
- метод ближайшего соседа - простейший метрический классификатор, основанный на оценивании сходства объектов;
- нейронные сети - математические модели для прогнозирования данных;
- линейная регрессия - метод оценки зависимости между двумя переменными;
- метод опорных векторов - метод классификации данных, а также регрессионного анализа» [28].

В целях увеличения эффективности на предприятии необходимо определиться с набором средств для проведения кластеризации данных и прогнозирования. Для этого осуществлен анализ и сравнение часто используемых методов Data mining по трем показателям: быстрота, точность и трудоемкость, выбор которых определен представленными ниже требованиями:

- для достижения качественных решений поставленных проблем результат, получаемый с использованием интеллектуального анализа данных, должен быть точным;
- для сокращения времени обработки операций метод интеллектуального анализа данных должен достаточно быстро решать поставленную задачу;
- для легкости понимания кода другими пользователями, а также беря во внимание ограниченность времени на разработку, алгоритм интеллектуального анализа данных должен обладать малой трудоёмкостью.

«Каждая представленная характеристика оценивается по следующей возрастающей шкале:

- очень низкая - 1,
- низкая - 2,

- средняя - 3,
- высокая - 4,
- очень высокая - 5» [28].

Сравнительные результаты приведены в таблице 2.

Таблица 2 - Сравнение методов интеллектуального анализа данных

Метод	Быстрота	Трудоемкость	Точность	Популярность
Кластеризация				
k-medians (неиерархические методы кластерного анализа)	4	2	3	3
k-means (неиерархические методы кластерного анализа)	3	2	3	4
Самоорганизующи еся карты Кохонена	2	2	4	4
Метод опорных векторов	3	5	4	3
Метод ближайшего соседа	4	2	2	2
Прогнозирование				
Нейронные сети	2	3	4	4
Деревья решений	4	4	3	4

Проведенный анализ позволил определить следующий набор технологий для проведения интеллектуального анализа данных:

- метод k-means - для кластеризации клиентов, так как «данный метод работает быстро и выдает достаточно точный результат» [9];
- деревья решения (C4.5) - для классификации, как самый распространенный инструмент, используемый в этой области, выдающий достаточно точный результат и быстро работающий.

2.2. Алгоритм k-средних

«Метод k-средних представляет собой метод кластерного анализа. Назначение данного метода – разделение m наблюдений (из пространства R^n) на k кластеров. Важно: каждое наблюдение относится к тому кластеру, к центру (центроиду) которого оно ближе всего, мерой близости принимается Евклидово расстояние:

$$p(x, y) = \|x - y\| = \sqrt{\sum_{p=1}^n (x_p - y_p)^2} \quad (1)$$

где $x, y \in R^n$.

Рассмотрим ряд наблюдений $(x^{(1)}, x^{(2)}, \dots, x^{(m)})$, $x^{(j)} \in R^n$.

Метод k-средних разделяет m наблюдений на k групп (или кластеров) ($k \leq m$) $S = \{S_1, S_2, \dots, S_k\}$, чтобы минимизировать суммарное квадратичное отклонение точек кластеров от центроидов этих кластеров:

$$\min \left[\sum_{i=1}^k \sum_{x^{(j)} \in S_i} \|x^{(j)} - \mu_i\|^2 \right] \quad (2)$$

где $x^{(j)} \in R^n$, $\mu_i \in R^n$, μ_i -центроид для кластера S_i » [15].

Блок-схема алгоритма k-means представлена на рисунке 6.

Алгоритм

«Разбиение объектов на кластеры сводится к определению центроидов этих кластеров, если мера близости до центроида определена, при этом исследователем предварительно задается число кластеров k .

Рассмотрим стартовый набор k средних (центроидов) $\mu_1, \mu_2, \dots, \mu_k$ в кластерах S_1, S_2, \dots, S_k . Первоначально центроиды кластеров выбираются по

определенному правилу или случайным образом (например, выбрать центроиды, максимизирующие начальные расстояния между кластерами).

Сопоставляем наблюдения к тем кластерам, чье среднее (центроид) к ним ближе всего. Важно: даже если наблюдение можно отнести к двум и более кластерам, оно принадлежит только к одному кластеру» [18].



Рисунок 6 - Блок-схема алгоритма k-means

На следующем шаге «центроид каждого i -го кластера перевычисляется по следующему правилу:

$$\mu_j = \frac{1}{S_i} \sum_{x^{(j)} \in S_i} x^{(j)} \quad (3)$$

Следовательно, алгоритм *k*-средних заключается в перевычислении на каждом шаге центроида для каждого кластера, полученного на предыдущем шаге.

Алгоритм завершается, когда значения μ_j не меняются: $\mu_j^{\text{шаг } t} = \mu_j^{\text{шаг } t+1}$ [18].

«Важно: на некорректность результата может повлиять неправильный выбор первоначального числа кластеров *k*, следовательно, применяя метод *k*-средних необходимо осуществить первоначальную проверку подходящего числа кластеров для данного набора данных» [41].

Выделим ключевые особенности метода *k*-средних:

- в качестве меры близости принимается Евклидово расстояние;
- число кластеров предварительно неизвестно и задается исследователем;
- первоначальное разбиение влияет на качество кластеризации.

2.3 Алгоритм C4.5

Джоном Квинланом был разработан алгоритм для построения деревьев решений C4.5, для которого имеются необходимые условия к структуре обучающего набора данных и собственно к самим данным. Выполнение нижеприведенных условий позволит алгоритму C4.5 быть работоспособным, устойчивым и обеспечит точные результаты.

Данные должны иметь структуру в виде таблицы, в которой строки являются обучающими примерами, представляющие собой классифицируемые объекты с заданными метками класса (поскольку алгоритм использует обучение с учителем), а столбцы - атрибуты (признаки),

описывающие предметную область или бизнес-процесс. Набор атрибутов одинаков для всех строк.

Один из атрибутов должен быть представлен как атрибут класса (целевой). Метка класса должна быть для каждого обучающего примера. Входные атрибуты могут быть как дискретными т.е. принимать конечное число уникальных значений, так непрерывными, а атрибут класса - только дискретным.

Однозначно относиться к соответствующему классу должен каждый пример обучающего множества. Не применяются вероятностные оценки степени принадлежности примеров к классу (данное положение относится к нечётким деревьям решений). Число обучающих примеров должно быть намного больше числа классов в обучающем множестве.

Блок-схема алгоритма с4.5 представлена на рисунке 7.

Алгоритм

Имеется «обучающее множество S , содержащее m атрибутов и n примеров. Для множества S определено k классов C_1, C_2, \dots, C_k . Необходимо построить иерархическую классификационную модель в виде дерева решений на основе обучающего множества S » [27].

Построение дерева решений осуществляется от корневого узла к листьям (сверху вниз).

Изначально формируется «пустое» дерево, представляющее собой только корневой узел, включающий всё обучающее множество. Необходимо разделить корневой узел на подмножества. Из подмножеств будут разветвляться узлы-потомки. С этой целью «выбирается один из атрибутов и определяются правила, которые разбивают обучающее множество на подмножества. Число подмножеств соответствует количеству p уникальных значений атрибута» [27].

«Итогом разбиения получаются p (по числу значений атрибута) подмножеств и, соответственно, определяются p потомков корневого узла, к каждому из них сопоставляется свое подмножество. Данный процесс

рекурсивно применяется ко всем подмножествам до тех пор, пока не будет реализовано условие завершения обучения» [12].

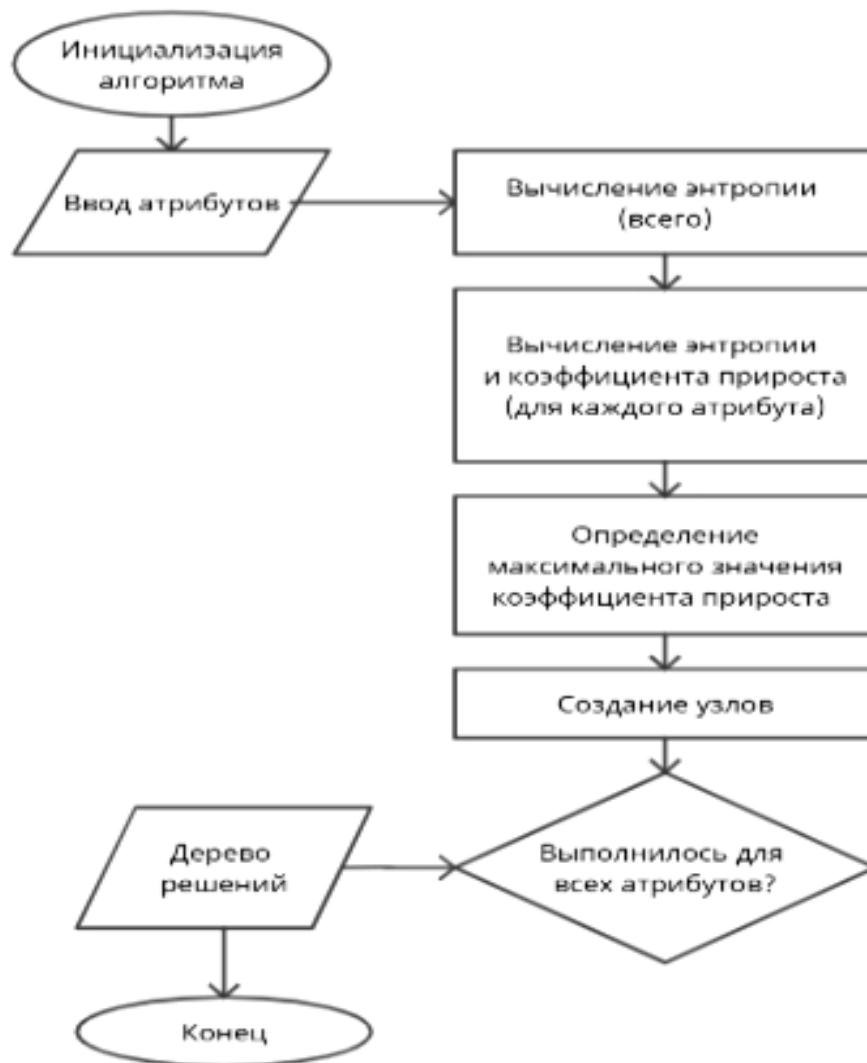


Рисунок 7 - Блок-схема алгоритма C4.5

Выбор атрибута, обеспечивающего оптимальное разбиение (полагаясь на некоторую меру качества) в текущем узле, является проблемой при обучении деревьев решений. В нашем случае каждый атрибут может применяться для разбиения произвольное число раз, в то время как некоторые алгоритмы обучения деревьев решений позволяют использовать каждый атрибут только один раз.

Пусть к обучающему множеству используется правило разбиения (рисунок 8), в котором применяется атрибут A , принимающий p значений a_1, a_2, \dots, a_p . В итоге получится p подмножеств S_1, S_2, \dots, S_p , куда будут входить примеры, в которых атрибут A принимает соответствующее значение.

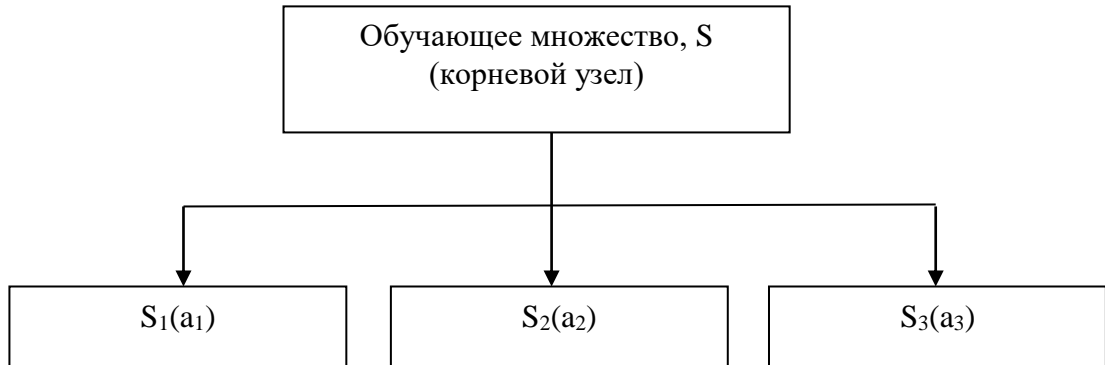


Рисунок 8 - Правило разбиения множества

Вопрос: можно ли получить лучшее разбиение, выбрав другой атрибут, или же разбиение по выбранному атрибуту является лучшим? Воспользуемся сведениями о числе примеров всех классов в обучающем множестве и в каждом полученном подмножестве.

Пусть $N(C_j S)$ - число примеров класса C_j в множестве S . Отсюда вероятность класса C_j в данном множестве будет вычисляться следующим образом:

$$P = \frac{N(C_j S)}{N(S)} \quad (4)$$

где $N(S)$ — общее число примеров в множестве S .

В теории информации величину рассчитываемую по формуле:

$$Info(S) = - \sum_{i=1}^m \frac{N(S_i)}{N(S)} \log \left(\frac{N(C_j S)}{N(S)} \right) \quad (5)$$

называют энтропией множества S , которая является средним количеством информации, важным для обозначения класса примера из множества S .

После разбиения множества S по атрибуту A , полученную данную оценку можно представить следующим образом:

$$Info_A(S) = \sum_{i=1}^k \frac{N(C_i S)}{N(S)} Info(S_i) \quad (6)$$

где S_i - i -й узел, полученный при разбиении по атрибуту A .

Тогда критерий прироста информации (от англ. gain - прирост, увеличение) вычисляемое по формуле (7), можно применять для выбора лучшего атрибута ветвления.

$$Gain(A) = Info(S) - Info_A(S) \quad (7)$$

Далее выбирается максимальный атрибут, предварительно произведя вычисление значения критерия для всех потенциальных атрибутов разбиения.

Пока значения критерия не перестанут существенно возрастать при новых разбиениях или будет реализовано другое условие остановки, представленный процесс используется к подмножествам S_i .

Формирование «пустого» узла (не попало ни одного примера) в процессе построения дерева приводит к образованию листа. Данный лист отождествляется с классом, в большинстве случаев встречающимся у прямого предка узла.

Свойство энтропии положено в основу критерия прироста информации: она наибольшая, когда все классы вероятны в равной степени, другими словами, имеет место быть максимально неопределённый выбор класса, а в случае, если примеры в узле принадлежат одному классу, она

равна 0 (в этом случае отношение под логарифмом равно 1, а его значение 0). Следовательно, увеличение классовой однородности результирующих узлов демонстрируется через прирост информации.

Описанный процесс относится к дискретным атрибутам. Алгоритм работает по-другому в случае непрерывных атрибутов. Определяется порог, относительно которого будет проводится сравнение всех значений. Итак, числовой атрибут X принимает конечное множество значений x_1, x_2, \dots, x_p . Отсортировав примеры по увеличению значений атрибута, имеем, что любое значение в диапазоне x_i и x_{i+1} разбивает на два подмножества все примеры. Первое подмножество будет представлять значения атрибута x_1, x_2, \dots, x_i , а второе - $x_{i+1}, x_{i+2}, \dots, x_p$.

Следовательно, порогом можно считать среднее следующим образом:

$$T_i = \frac{x_i + x_{i+1}}{2} \quad (8)$$

Таким образом, рассмотрение $n-1$ потенциальных пороговых значений T_1, T_2, \dots, T_{n-1} является задачей нахождения порога. Поочередно применяя формулы (5), (6) и (7) ко всем потенциальным порогам, выбираем то, значение которого будет максимальным по критерию (7). Далее это значение сравнивается со значением критерия (7), рассчитанным для других атрибутов. Данное значение выбирается в качестве порога для проверки, если оно будет максимальным среди атрибутов.

Выводы к главе 2

Проведен аналитический обзор и сравнительный анализ на основе классификации методов интеллектуального анализа данных по задачам.

Для кластеризации клиентов было решено использовать метод k-means, для осуществления классификации в данной работе будут использоваться деревья решения (метод C4.5).

Глава 3 Разработка модели системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных

3.1 Выбор технологии для анализа бизнес-процесса управления взаимоотношениями с клиентами

При моделировании больших информационных систем более оптимальным для использования является комбинированный метод моделирования. При этом выделяются два этапа:

- на первом этапе применяется структурный подход;
- на втором этапе «по всем предусмотренным работам структурного подхода осуществляется выработка спецификаций для определения объектов и перехода к объектно-ориентированному подходу» [35].

Для анализа процесса управления взаимоотношениями с клиентами будет использоваться структурный подход.

«Основная идея структурного подхода к разработке информационных систем представляет собой декомпозицию (разбиение) на автоматизируемые функции. Система делится на функциональные подсистемы, которые в свою очередь разбиваются на подфункции, подразделяемые на задачи, и так далее.

Наиболее популярные методологии моделирования бизнес-процессов и их анализа:

- моделирование бизнес-процессов в нотации IDEF0 предназначено для функционального описания бизнес-процессов;
- моделирование потоков данных в нотации DFD предназначено для описания потоков информации в процессе выполнения работ.» [12]

Таким образом, для анализа бизнес-процесса управления взаимоотношениями с клиентами будут использоваться технологии IDEF0 и DFD.

3.2 Концептуальное моделирование бизнес-процесса управления взаимоотношениями с клиентами

Первоначальный этап разработки и точное представление сферы деятельности менеджера лежат в основе разработки функциональной модели управления взаимоотношениями с клиентами.

Для полного определения и представления процесса управления взаимоотношениями с клиентами будет представлено несколько моделей бизнес-процессов, которые необходимы для определения круга проблем.

Целью модели является выделение и анализ бизнес-процесса управления взаимоотношениями с клиентами.

Ширина охвата данной модели представляется деятельностью одного менеджера, а глубина детализации обозначает степень подробности декомпозиции блоков, достигает 2-х уровней.

На рисунке 9 представлена модель «КАК ЕСТЬ», построенная по методологии IDEF0.

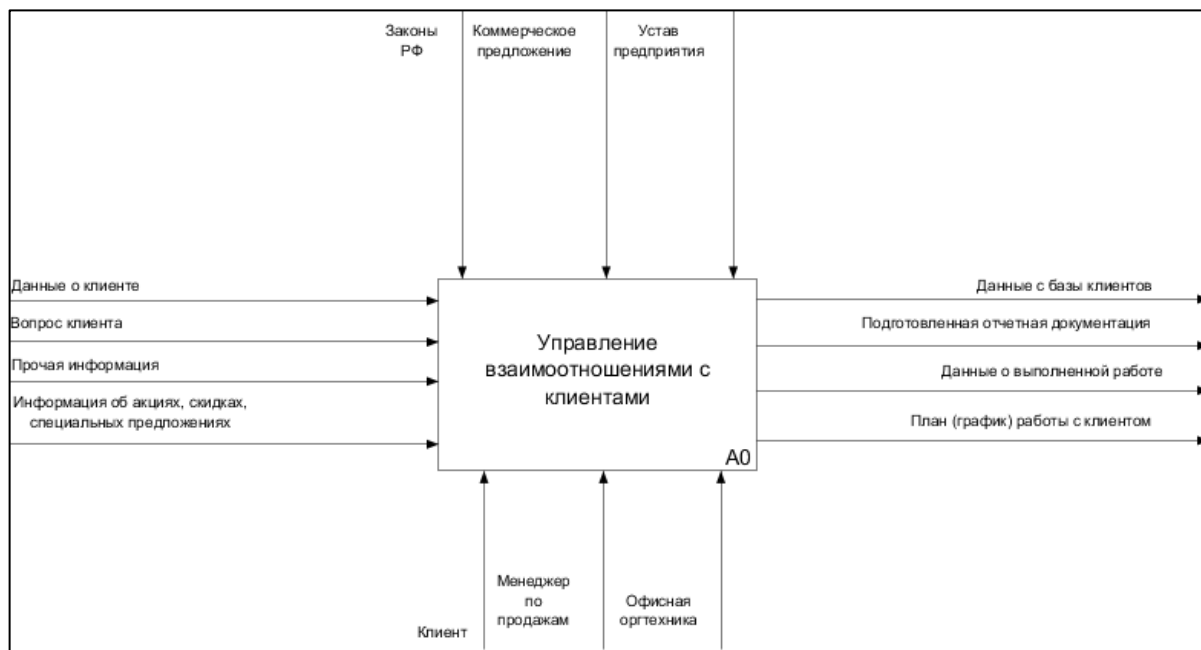


Рисунок 9 - А-0 Контекстная диаграмма бизнес-процесса управления взаимоотношениями с клиентами («КАК ЕСТЬ») в методологии IDEF0

Личные данные о клиенте, сформулированный вопрос клиента, сведения об имеющихся скидках, проводимых акциях, персональных специальных предложениях и прочие сведения выступают в качестве входных данных бизнес-процесса управления взаимоотношениями с клиентами. Отчетность, утвержденный график работы с клиентами, сведения о проделанной работе, рекомендуемая маркетинговая политика предприятия формируются на выходе бизнес-процесса. Клиент, менеджер и офисная оргтехника выступают в качестве исполнителей процесса, а управляющими документами являются принятые законы РФ, устав предприятия и коммерческое предложение.

На рисунке 10 представлена диаграмма декомпозиции главного бизнес-процесса А-0.

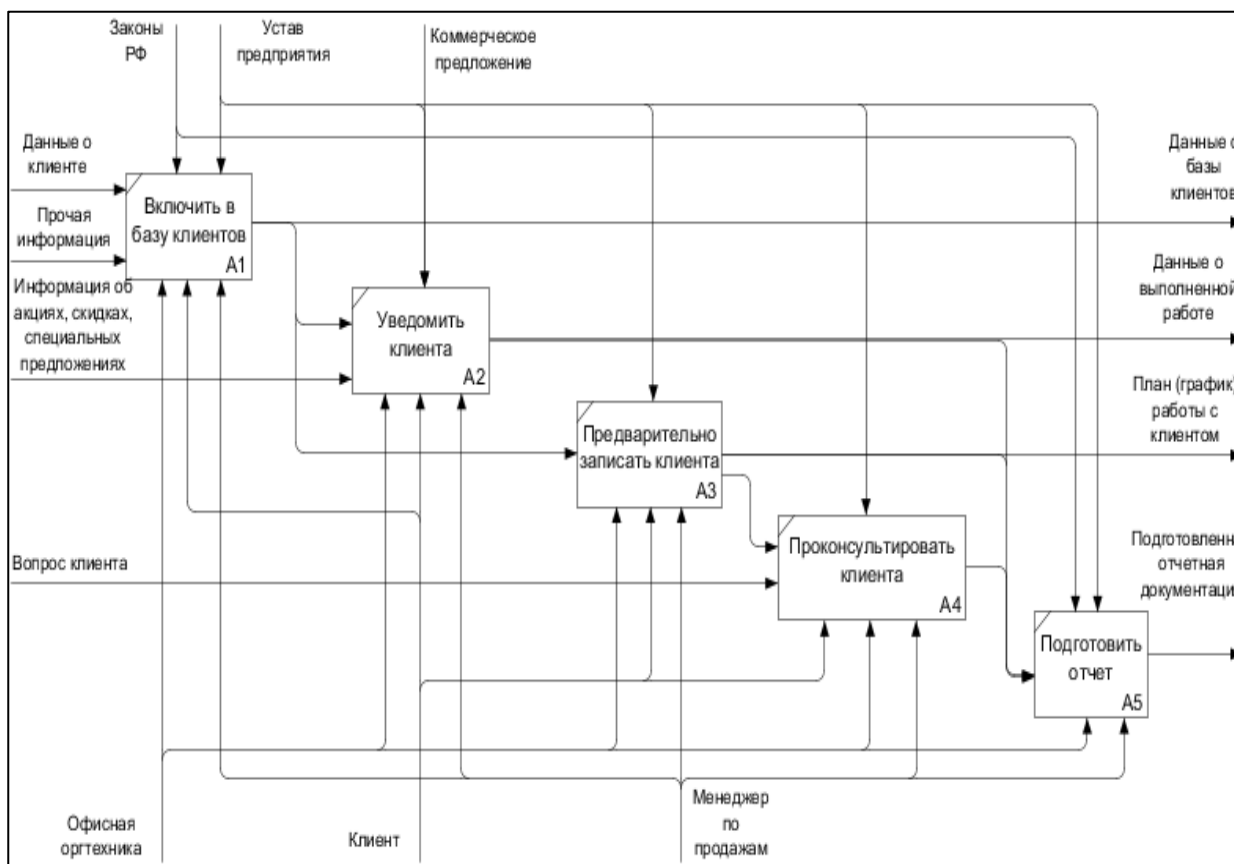


Рисунок 10 - А0 Декомпозиция бизнес-процесса управления взаимоотношениями с клиентами («КАК ЕСТЬ») в методологии IDEF0

Процесс «Управление взаимоотношениями с клиентами» включает в себя выполнение следующих процессов: «Уведомить клиента», «Включить в базу клиентов», «Подготовить отчётность», «Консультировать», «Предварительно записать клиента».

Проведя анализ деятельности менеджера, применив структурный подход моделирования, были отмечены следующие недостатки:

- формирование отчетной документации осуществляется в ручном режиме;
- информирование клиентов применяется без использования информационных технологий;
- процесс записи клиента осуществляется не автоматизировано;
- отсутствие автоматизированного процесса информационной поддержки клиента;
- процесс учета и внесения сведений клиентов в базу данных не автоматизирован, т.е. у клиентов отсутствует возможность зарегистрироваться самим.

Резюмируя вышесказанное и опираясь на выявленные недостатки, можно сформулировать основные аргументы необходимости системы управления взаимоотношениями с клиентами на основе Data Mining.

«Основопологающим техническим заданием на создание или модификацию системы является функциональная модель «КАК ДОЛЖНО БЫТЬ». Такая модель представляет собой концептуальную модель оптимизированного бизнес-процесса» [27].

Далее рассмотрим, как будет оптимизирован бизнес-процесс управления взаимоотношениями с клиентами с использованием CRM-системы.

На рисунке 11 продемонстрирована концептуальная модель процесса управления взаимоотношениями с клиентами после внедрения CRM-системы.

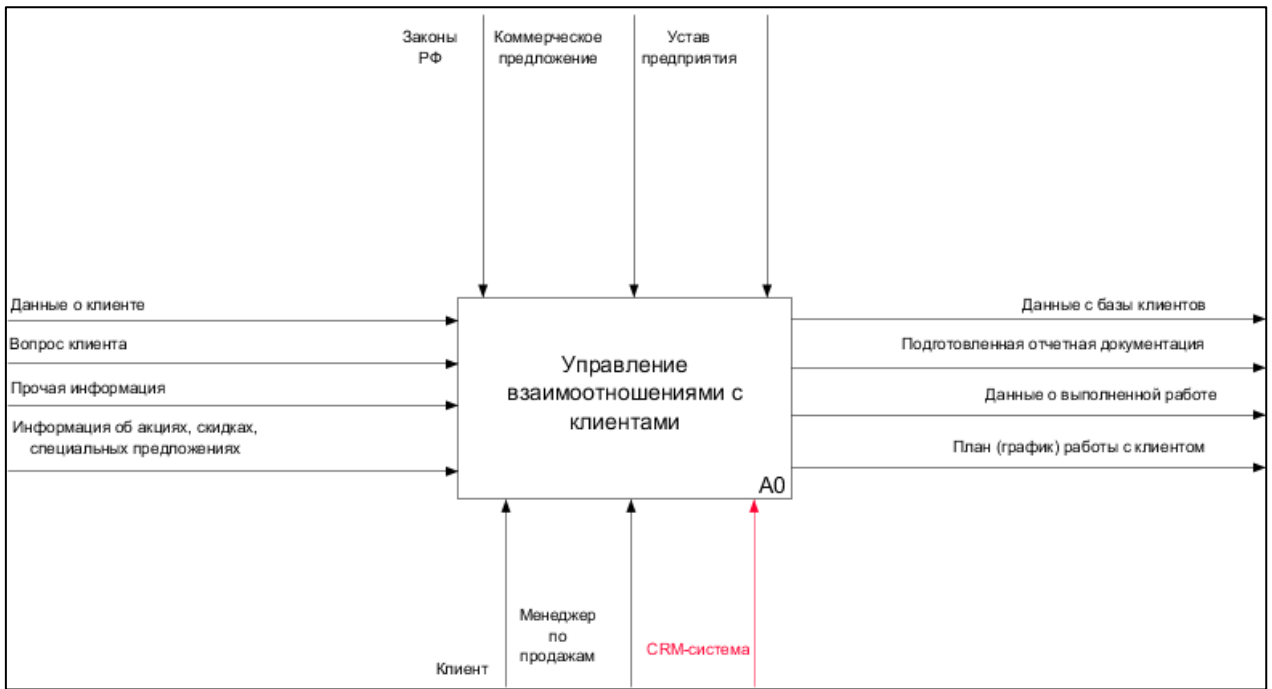


Рисунок 11 - A-0 Контекстная диаграмма бизнес-процесса управления взаимоотношениями с клиентами («КАК ДОЛЖНО БЫТЬ») в методологии IDEF0

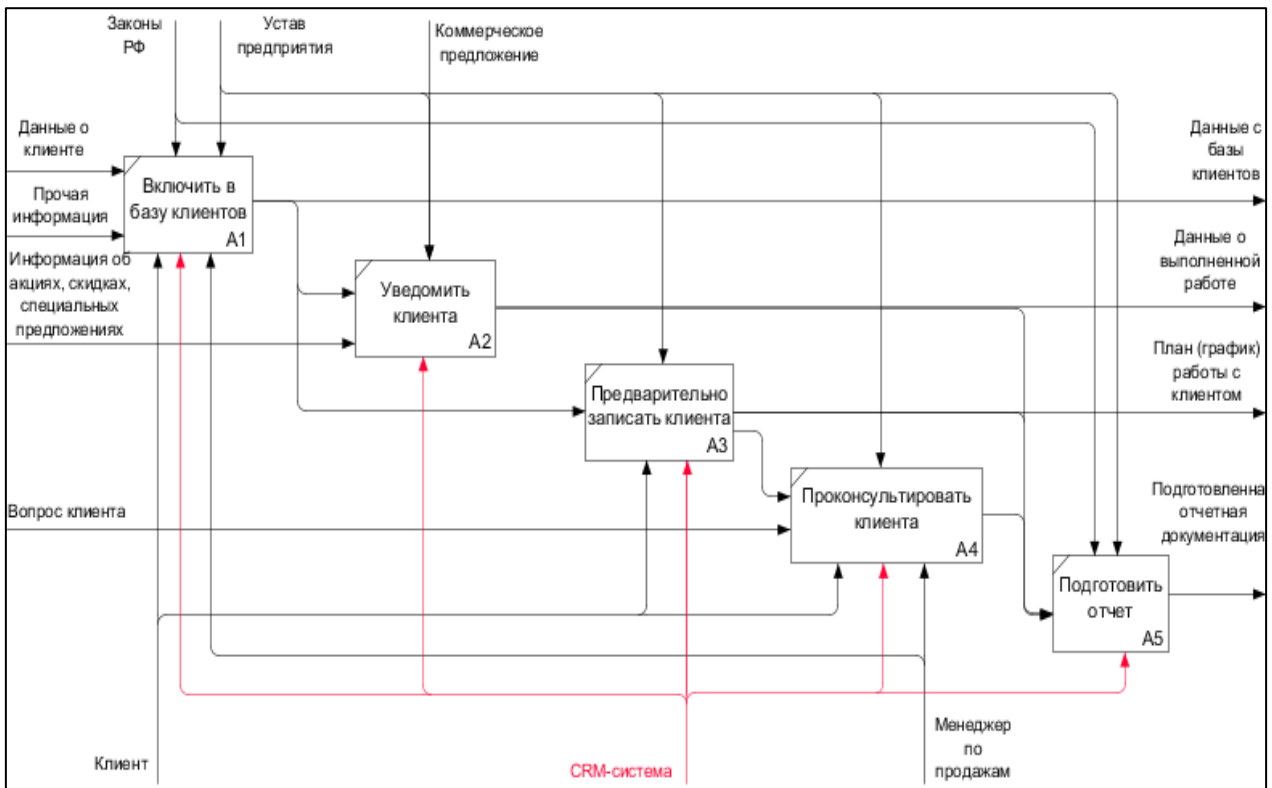


Рисунок 12 - A0 Декомпозиция бизнес-процесса управления взаимоотношениями с клиентами («КАК ДОЛЖНО БЫТЬ») в методологии IDEF0

На рисунке 12 приведена декомпозиция главного бизнес-процесса управления взаимоотношениями с клиентами после внедрения CRM-системы.

На данной диаграмме видно, что CRM-система выступает в качестве исполнителя следующих бизнес-процессов:

- уведомить клиента;
- предварительно записать клиента;
- проконсультировать клиента;
- подготовить отчет.

Таким образом, это поспособствует сокращению времени выполнения данных процессов, уменьшению возможных ошибок, а также, что немаловажно, снижению нагрузки менеджера предприятия.

На этапе концептуального моделирования применим «методологию графического структурного анализа DFD для определения необходимых хранилищ данных» [21] и потоков данных.

На рисунке 13 продемонстрирована контекстная диаграммы потоков данных.

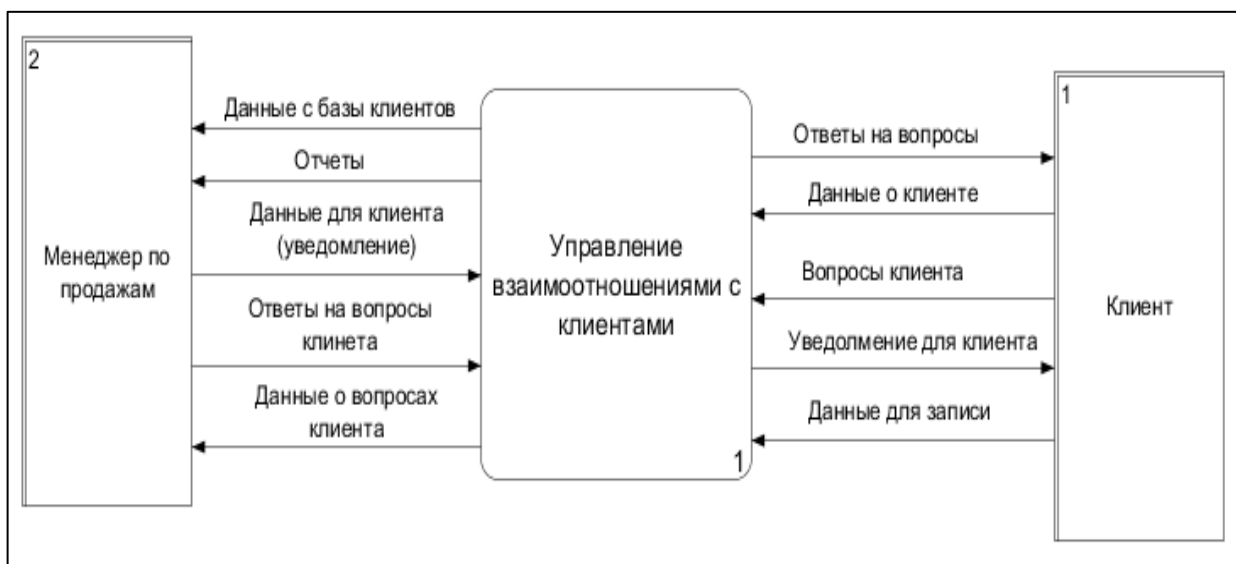


Рисунок 13 - А-0 Контекстная диаграмма бизнес-процесса управления взаимоотношениями с клиентами («КАК ДОЛЖНО БЫТЬ») в методологии DFD

В модели, изображенной на рисунке 13, были выделены две сущности: «менеджер по продажам» и «клиент». Также выделены хранилища данных: клиентская база, архив уведомлений, архив вопросов, архив записи.

На рисунке 14 представлена декомпозиция контекстной диаграммы, выполненная в методологии DFD.

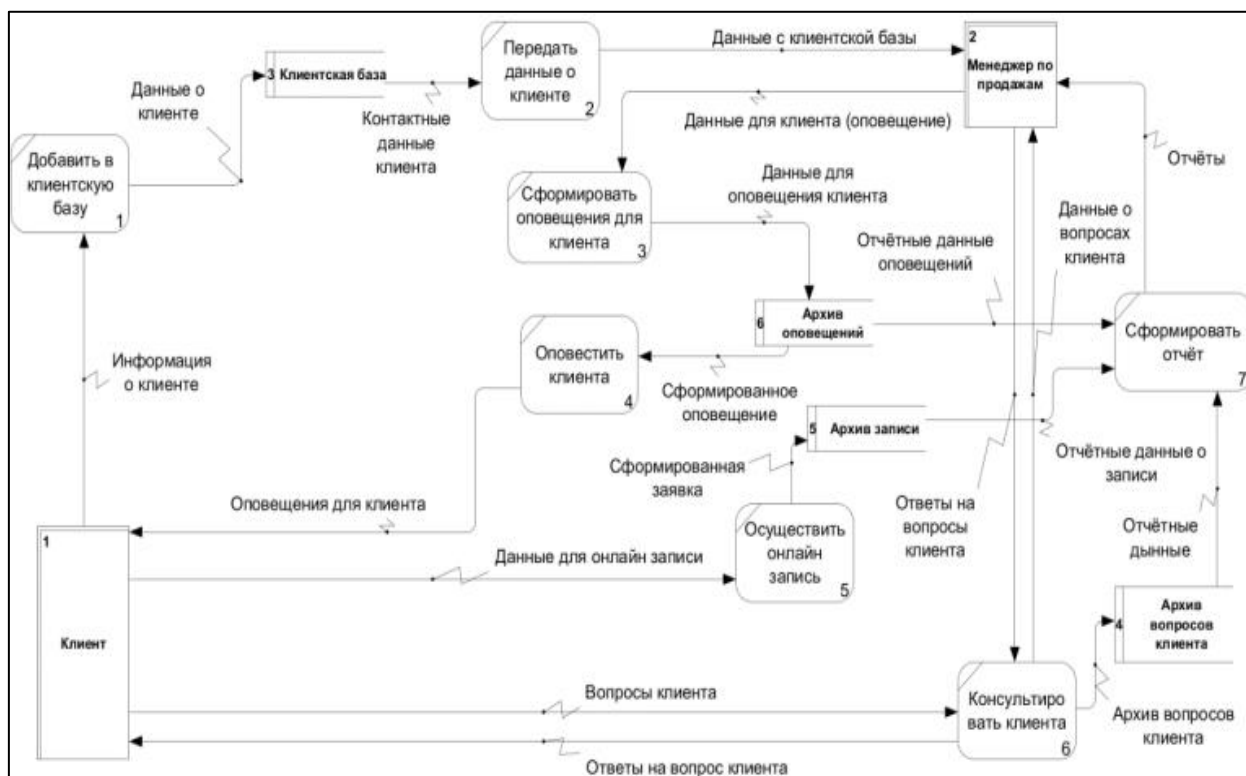


Рисунок 14 - А0 Декомпозиция бизнес-процесса управления взаимоотношениями с клиентами («КАК ДОЛЖНО БЫТЬ») в методологии DFD

На данной диаграмме видно, что «от внешней сущности «Клиент» передается поток данных «информация о клиенте», который обрабатывается в процессе «добавить в клиентскую базу», затем преобразуется в выходные данные «данные о клиенте» и далее помещаются в хранилище «клиентская база». Следующий шаг: из хранилища данные передаются в процесс «передать данные о клиенте» для предоставления «данных с клиентской базы» сущности «менеджеру». От сущности «клиент» передается поток «данные для онлайн записи» в процесс «осуществить онлайн запись» на

выходе для сохранения потока «сформированная заявка» в «архиве записи». От сущности «менеджер» передается поток «данные для клиента (уведомление)», обрабатываемый в процессе «сгенерировать уведомление для клиента», на выходе из которого образуются «данные для уведомления клиента», поступающие в хранилище «архив уведомлений» для оповещения клиента. От сущности «клиент» передается поток данных «вопросы клиента», который обрабатывается в процессе «консультировать клиента», формируемом поток «данные о вопросах клиента» для сущности «менеджер», от которого поступает поток «ответы на вопросы клиента». Этот поток обрабатывается в процессе «консультировать клиента» и отправляется к сущности «клиент». От хранилищ «архив уведомлений», «архив вопросов» и «архив записи» передаются данные в процесс «сформировать отчет», из которого транслируется поток данных «отчеты» к сущности «менеджер по продажам» [14].

В результате декомпозиция контекстной диаграммы, выполненная в методологии DFD, сыграла ключевую роль в подробном описании деятельности менеджера предприятия, и были выделены его главные функции, которые будут основополагающими для разработки CRM-системы.

Проанализировав деятельность менеджера, выявлены его основные функции. Данные функции будут основополагающими для разработки CRM-системы.

3.3 Обоснование выбора технологии моделирования CRM-системы

Для разработки модели данной системы будет использоваться объектно-ориентированный подход.

UML (с англ. Unified Modeling Language - унифицированный язык моделирования) является нотацией данного подхода. С помощью нее на достаточно содержательном и наглядном уровне можно описать предметную область. Одним из достоинств является то, что процесс модификации проекта

облегчается, т.к. при внесении изменений в один из объектов или процессов не надо изменять все остальные, например, как в структурном подходе.

На сегодняшний день «объектно-ориентированное моделирование пользуется большей популярностью на мировом уровне, чем структурное, т.к. основано на методологии объектно-ориентированного программирования.

Наиболее часто используемыми методологиями объектно-ориентированного моделирования являются:

- экстремальное программирование (XP);
- методология разработки программного обеспечения, предложенная корпорацией Microsoft (MSF);
- методология разработки программного обеспечения, созданная Rational Software (RUP);
- быстрая разработка приложений (RAD)» [22].

Был проведен сравнительный анализ указанных выше методологий, в результате которого обоснован выбор методологии моделирования CRM-системы.

Оценка производилась по пятибалльной шкале: где 0 - полное несоответствие критерию, а 5 соответствует полному его удовлетворению. Результаты оценки по критериям отображены в таблице 3.

Таблица 3 - Сравнительный анализ методологий объектно-ориентированного подхода к моделированию

Критерий	RUP	XP	RAD	MSF
Методология поддерживает мировые стандарты моделирования	5	1	3	4
Наглядность представленной модели	5	3	4	4
Возможность групповой работы	3	5	5	5
Полнота охвата ЖЦ разработки программного обеспечения	5	5	3	5
Гибкость проекта к требованиям	5	5	5	3
Короткий производственный график	3	1	3	3
Итого	21	19	20	20

В результате проведенного анализа видно, что наиболее предпочтительной является методология RUP. Данная методология будет положена в основу моделирования системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных.

3.4 Обоснование выбора компонентов архитектуры CRM-системы

При выборе серверного программного обеспечения был проанализирован результат тестирования популярных веб-серверов, реализованный экспертами компании WebConsult. Результат сравнения представлен в таблице 4.

Таблица 4 - Сравнительная оценка серверного ПО

Характеристики	Enterprise Web Server	Apache Tomcat	Microsoft IIS	iPlanet Web Server
Производительность	3	4	3	5
Доступность	4	5	4	4
Администрирование и управление	3	5	4	3
Разработка	4	5	4	4
Степень защиты	4	4	4	3
Итого	18	23	18	20

Видно, что итоговая оценка серверного программного обеспечения Apache Tomcat - самая наибольшая. Следовательно, для разработки системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных будет использоваться Apache Tomcat.

3.5 Моделирование CRM-системы

Рассмотрим проектируемую CRM-систему с точки зрения диаграммы вариантов использования. Обозначить внешние системы, коммутирующие с системой, ключевые бизнес-процессы и их корреляцию, возможно именно

«при помощи диаграммы вариантов использования. На ней визуально отображены континуум прецедентов и актеров, а также отношения (агрегация, обобщения, ассоциации) между ними. При помощи диаграммы вариантов использования можно определить функциональную структуру будущей CRM-системы, не вникая в детали ее реализации. Помимо этого, осуществляется предварительное выделение объектов CRM-системы и классификация этих объектов. На основании построенной модели составляется план разработки системы» [8].

Применительно к моделируемой CRM-системе требуется рассмотреть основной бизнес-процесс «Взаимодействие с CRM-системой». Диаграмма вариантов использования основного бизнес-процесса моделируемой CRM-системы изображена на рисунке 15. На данной диаграмме продемонстрированы основные прецеденты, которые используются при деятельности разных типов пользователей CRM-системы, такие как менеджер (по работе с клиентами) и менеджер (аналитик). Через механизмы аутентификации и авторизации посредством вводимых пользователем сведений логина и пароля, осуществляется разграничение прав доступа пользователей согласно типу учетной записи в системе безопасности базы данных.

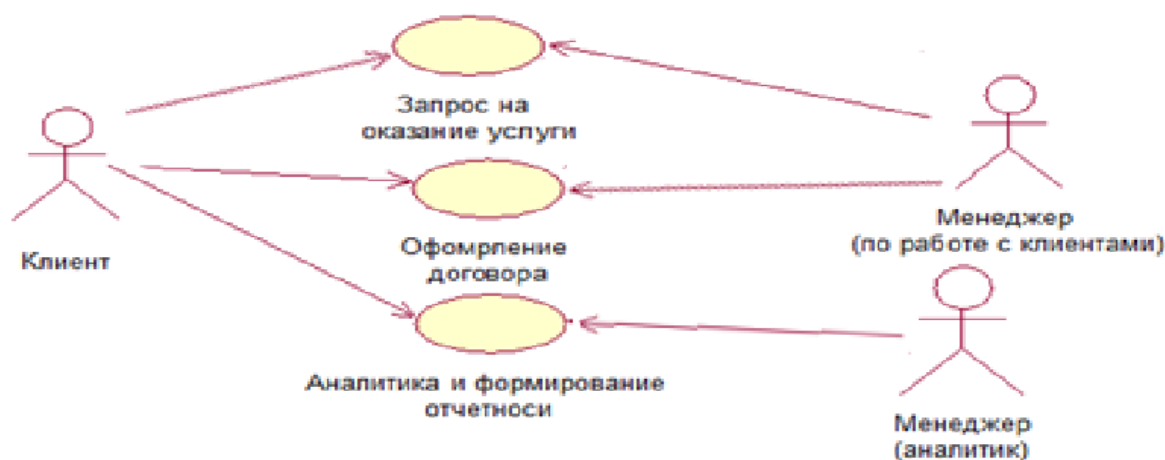


Рисунок 15 - Диаграмма вариантов использования бизнес-процесса «Работа с клиентами» в нотации UML

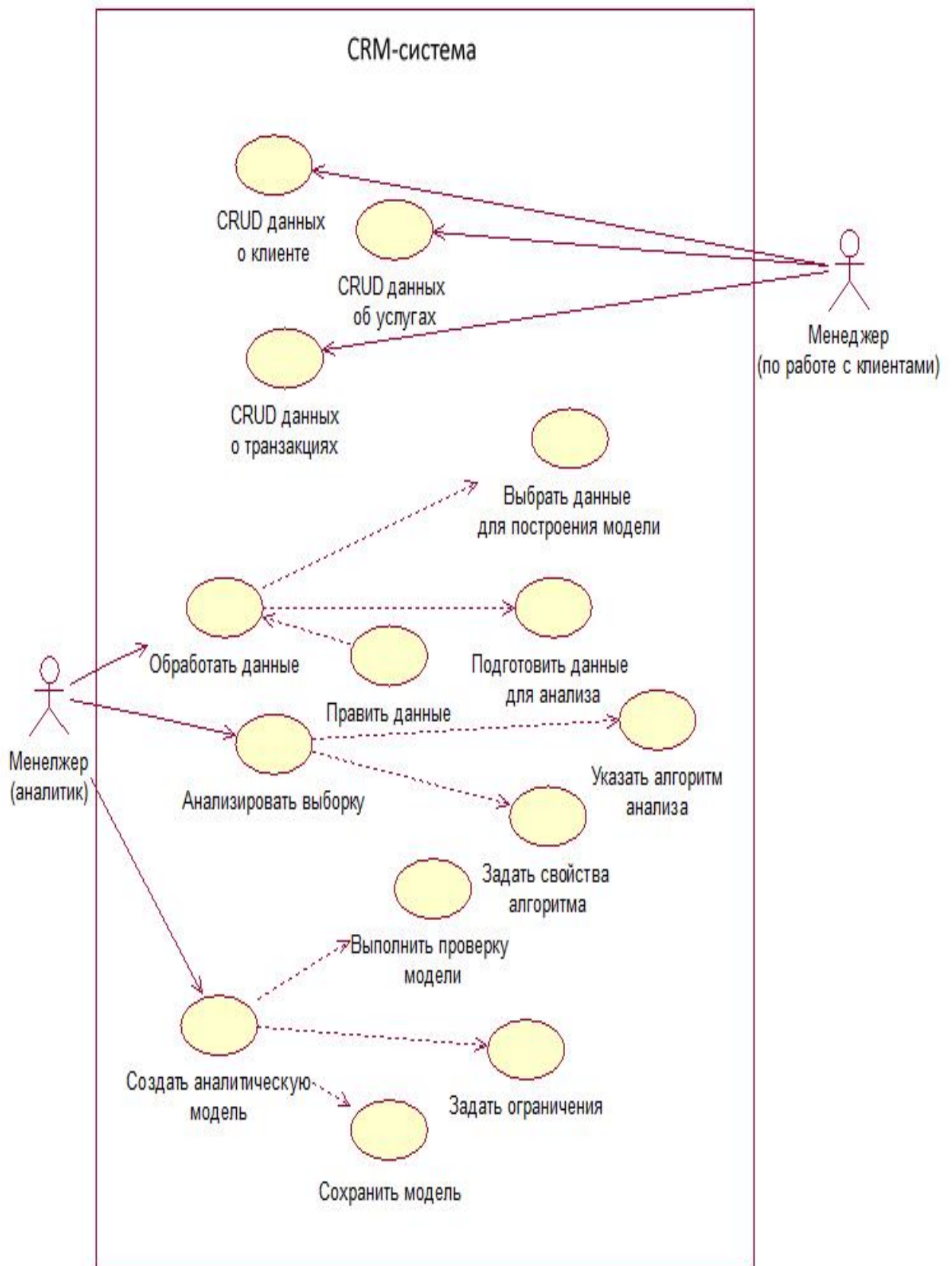


Рисунок 16 - Диаграмма вариантов использования бизнес-процесса «Взаимодействие с CRM-системой» в нотации UML

На диаграмме вариантов использования бизнес-процесса «Взаимодействие с CRM-системой», продемонстрированной на рисунке 16, видно, что для каждого типа пользователя характерны определенные действия по отношению к CRM-системе. С типом пользователя «Менеджер (аналитик)» соотносятся варианты: «Обработать данные», «Анализировать выборку», «Создать аналитическую модель». С типом пользователя «Менеджер (по работе с клиентами)» соотносятся варианты относящиеся к модулю сбора данных.

Следующим этапом приведем описание объектов и классов, комплектующих CRM-систему. Основываясь на том, что моделируемая CRM-система строится на базе клиент-серверной архитектуры, разобьем систему на серверную часть и клиентскую часть (рисунок 17).

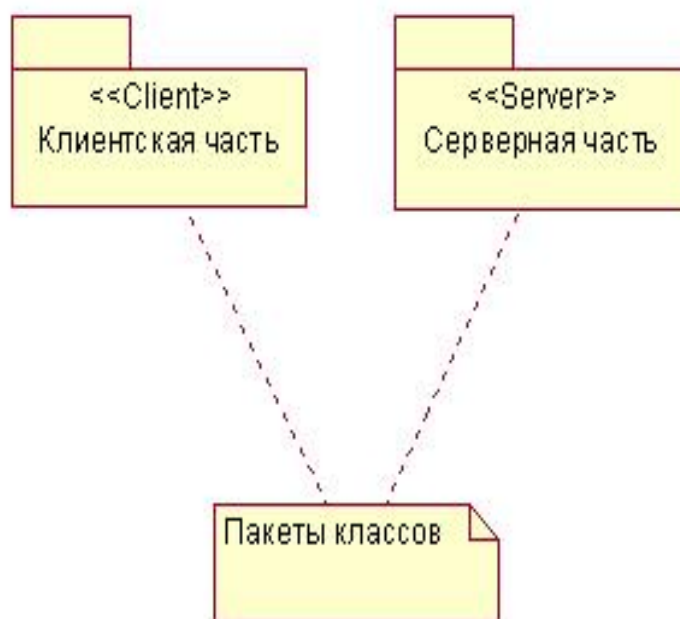


Рисунок 17 - Диаграмма пакетов классов в нотации UML

Приведем описание CRM-системы на основе диаграммы классов, на рисунке 18 и 19 представлены диаграммы классов модуля сбора данных и модуля интеллектуального анализа данных соответственно.

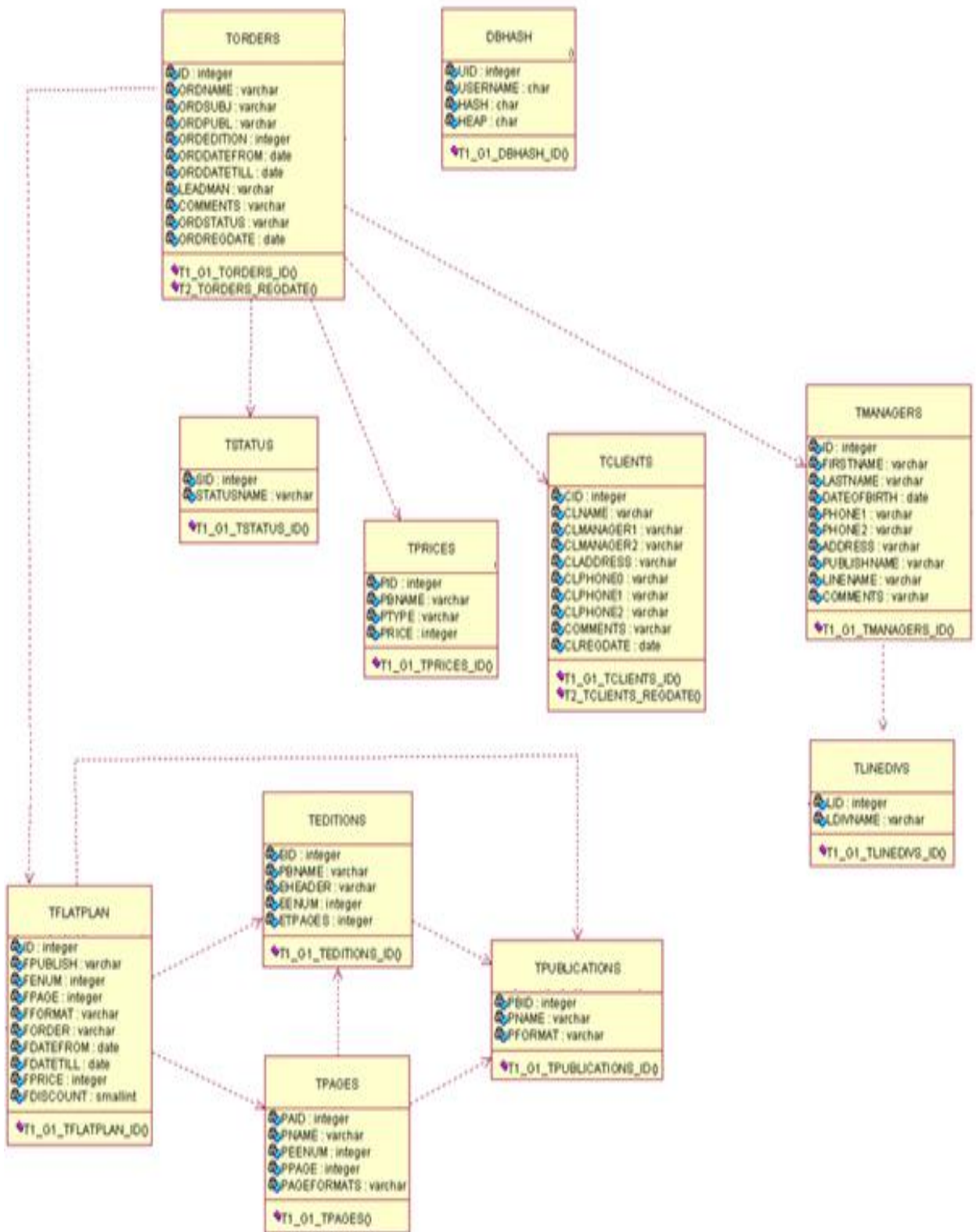


Рисунок 18 - Диаграмма классов «Модуля сбора данных CRM-системы» в нотации UML

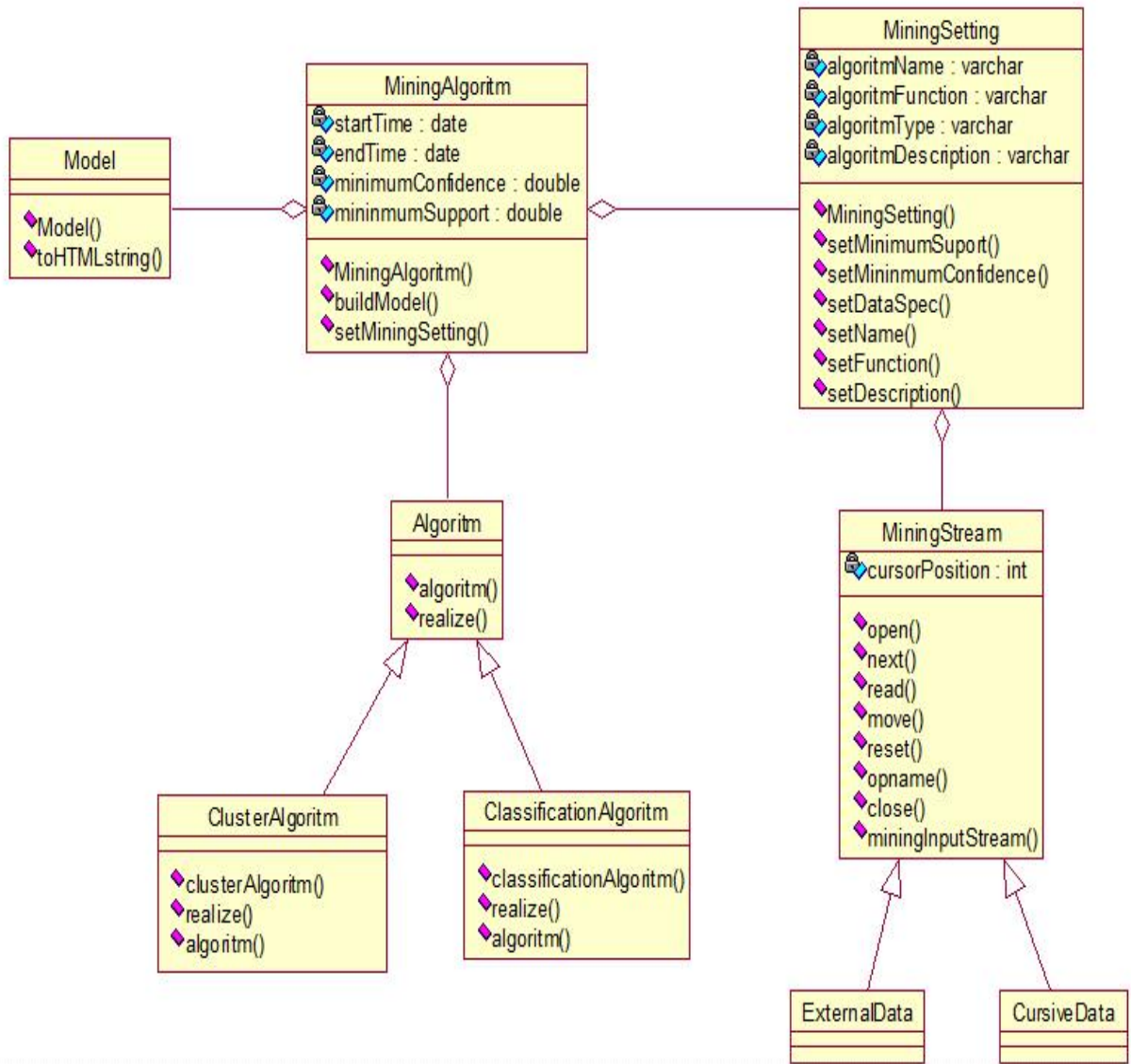


Рисунок 19 - Диаграмма классов «Модуля интеллектуального анализа данных CRM-системы» в нотации UML

Следующим шагом на описанных объектах и классах декомпозируем на уровень вверх каждый вариант использования взаимодействие с CRM-системой. Диаграмма последовательностей аналитического модуля представлена изображена на рисунке 20. На диаграмме последовательности представленная последовательность работы и взаимодействия выполнения прецедента «Создание аналитической модели». Результат декомпозиции для остальных вариантов представлен в виде диаграмм последовательностей в Приложении А на рисунках А.1-А.8.

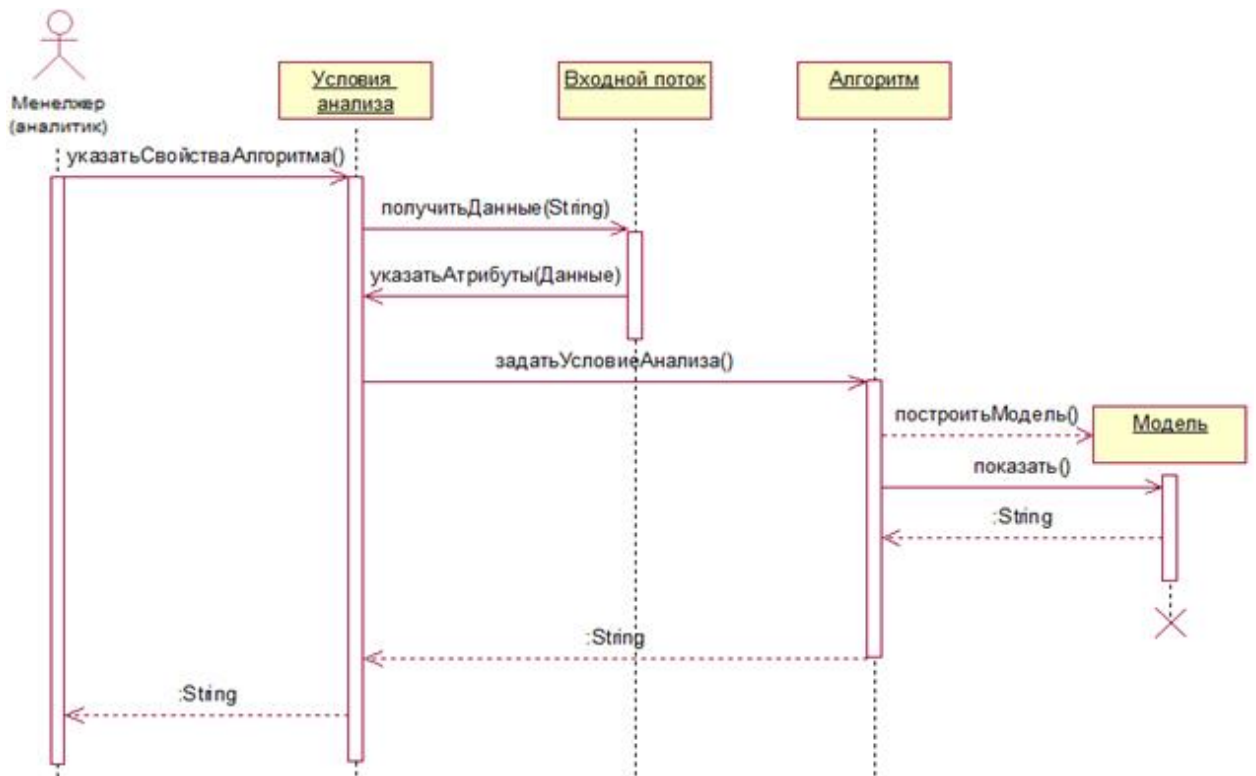


Рисунок 20 - Диаграмма последовательности для варианта использования «Создание аналитической модели»

После описания структуры объектов и классов системы приступим к описанию программных средств и размещению устройств CRM-системы.

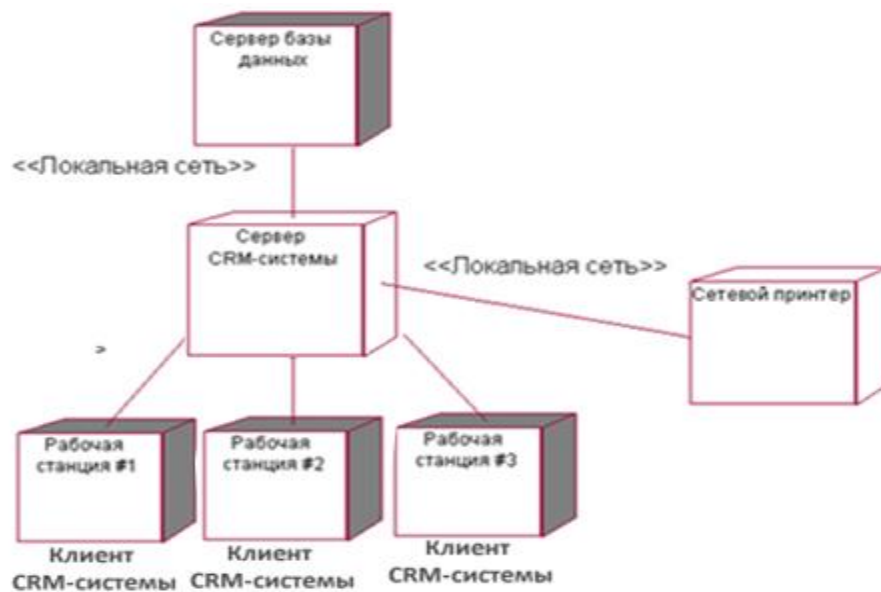


Рисунок 21 - Диаграмма размещения устройств и программных средств CRM-системы в узлах локально-вычислительной сети в нотации UML

На рисунке 21 продемонстрирована диаграмма размещения устройств и программных средств CRM-системы в узлах ЛВС.

Структурная схема модели CRM-системы представлена на рисунке 22.

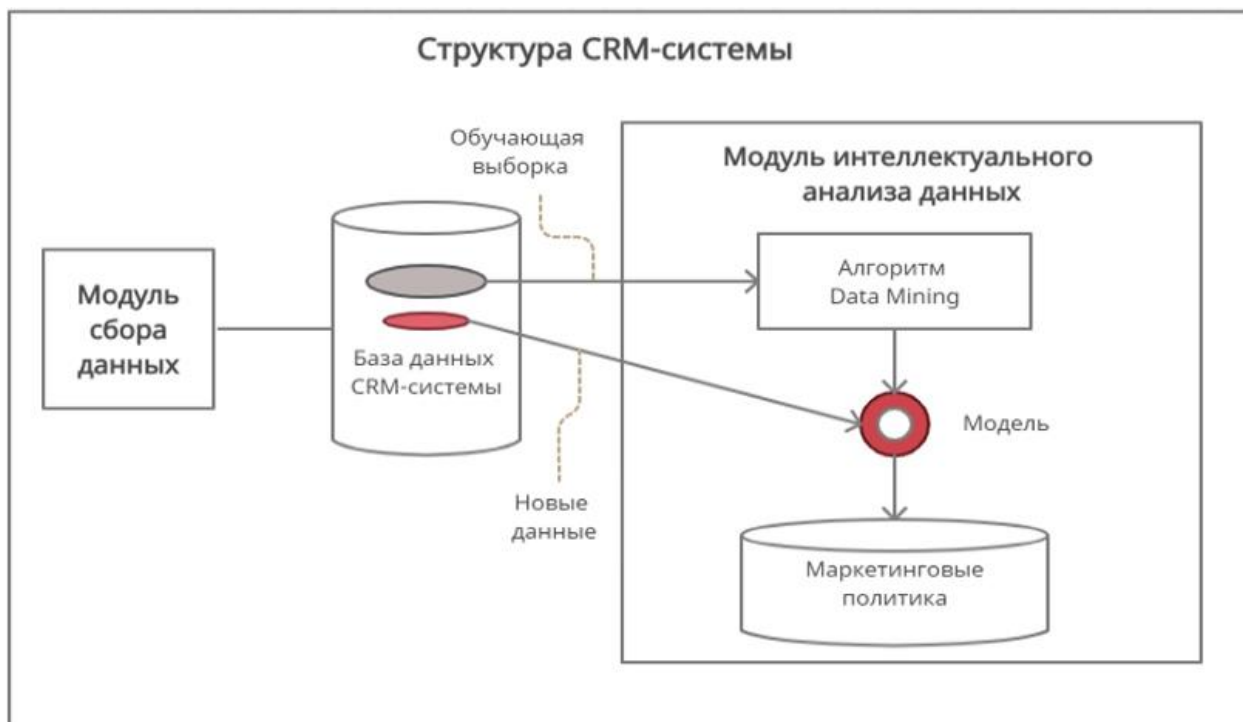


Рисунок 22 - Структурная схема модели CRM-системы

Модуль сбора информации предназначен для сбора сведений о клиентах, сведений об использовании имеющихся услуг, информации об сведениях удовлетворенности оказанной услугой. На основе собранной информации система может выполнять следующие операции над клиентом:

- сегментацию клиентов (метод k-means);
- классификация клиентов (метод С 4.5).

3.6 Обоснование выбора и описание основных средств реализации модели CRM-системы

На сегодняшний день среди существующих систем управления базами данных (СУБД) можно отметить самые популярные:

- MySQL,
- SQLite,
- ORACLE,
- PostgreSQL.

Сравнительный анализ СУБД в разрезе основополагающих характеристик представлен в таблице 5.

Таблица 5 - Сравнение основных СУБД

Сравнительная характеристика	MySQL	SQLite	PostgreSQL	ORACLE
Возможность осуществления клиент-серверной архитектуры	+	+	+	+
Реализация реляционной модели данных	+	+	+	+
Наличие механизмов защиты данных	+	+	+	+
Наличие встроенного интерпретирующего языка программирования	-	+	+	+
Возможность выполнения DML команд для созданных таблиц	+	-	-	+
Требования к ресурсам платформы невысокие	+	+	+	-
Присутствует опыт использования	+	-	-	-
Требование к дополнительным знаниям не предъявляется	+	+	+	+
Итого	8	7	7	7

Как видно из таблицы выше, система управления базами данных MySQL по результатам итоговой оценки является наиболее оптимальной. Вместе с тем, принимая во внимание имеющийся практический опыт работы с данной СУБД, для моделирования системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных была предпочтена СУБД MySQL.

«В настоящее время наиболее распространенными языками программирования для создания приложений являются: ASP, PHP и Java.»
[16]

Сравнение данных средств разработки представлено в таблице 6.

Таблица 6 - Сравнение языков программирования

Характеристика	PHP	ASP	JAVA
Способность работать с несколькими аппаратными платформами или операционными системами	Любая платформа	Win32, Solaris	Win32, Solaris, Linux
Имеет доступный, простой синтаксис кода	+	-	+
Имеется возможность взаимодействия с веб-сервером	IIS, Apache	IIS	IIS, Apache
Имеется возможность взаимодействие с FTP-сервером	+	-	+
Имеющийся опыт использования	-	-	+

Сравнив средства разработки с учетом наилучшей функциональности и опыта использования, выделим язык JAVA.

Выводы к главе 3

Для анализа бизнес-процесса управления взаимоотношениями с клиентами использовались технологии IDEF0 и DFD.

При проведении анализа бизнес-процесса управления взаимоотношениями с клиентами выявлены недостатки, и обоснована необходимость разработки модели CRM-системы.

Проведен сравнительный анализ технологий моделирования CRM-системы, в результате которого выбрана технология RUP.

Разработаны диаграммы в нотации UML: диаграмма вариантов использования, диаграмма классов, диаграммы последовательностей и диаграмма размещения устройств и программных средств

Приведено обоснование выбора и описание основных средств реализации модели CRM-системы.

Глава 4 Реализация системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных

4.1 Описание основного принципа работы CRM-системы

Одним из основных предназначений разработанной CRM-системы является сбор информации о клиентах, информации об использовании имеющихся услуг, информации об оценке удовлетворенности оказанной услугой. На основе собранной информации система может выполнять следующие операции над клиентом:

- расчет ценности клиента;
- сегментацию клиентов (метод k-means);
- классификация клиентов (метод С 4.5).

Сбор информации о клиентах, об использовании услуг.

На рисунках 23-26 продемонстрированы экранные формы CRM-системы для реализации просмотра, добавления, редактирования, удаления личной информации клиента и информации об оказанных услугах.

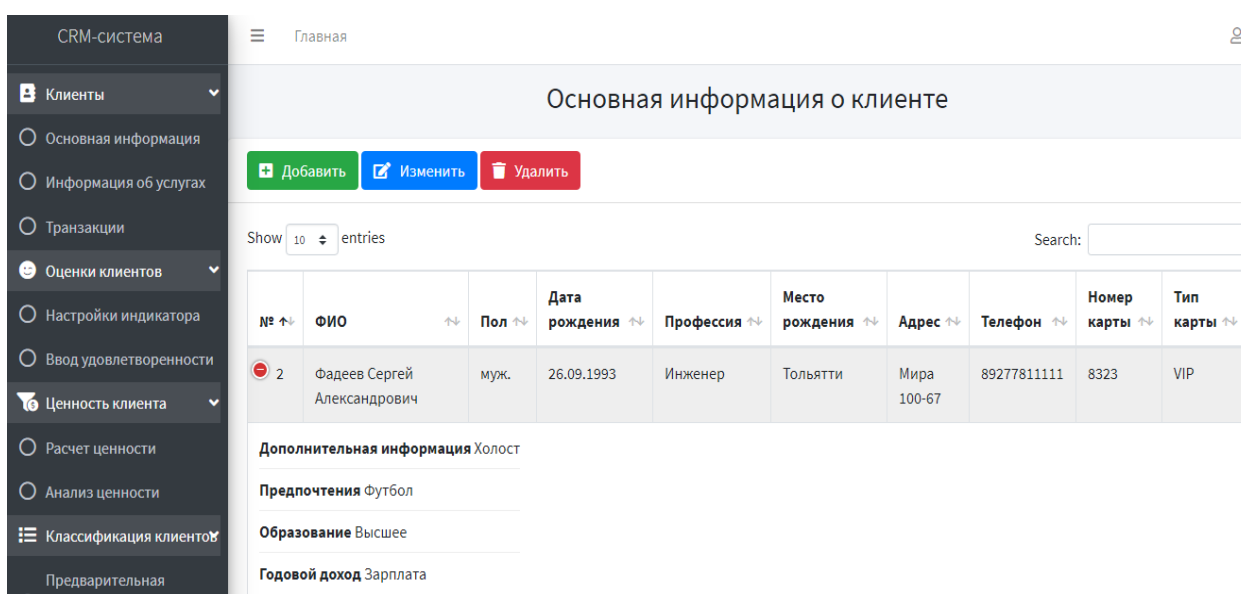


Рисунок 23 - Экранная форма «Управление основной информацией о клиенте»

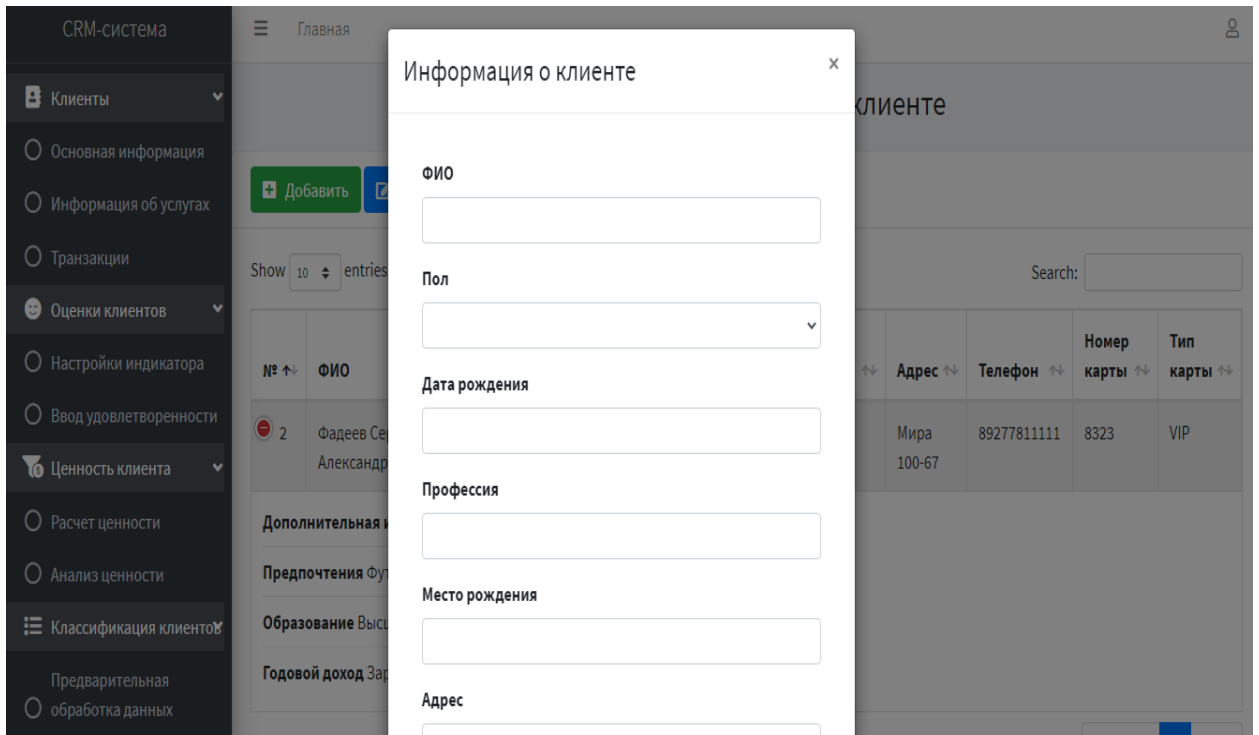


Рисунок 24 - Экранная форма «Добавление информации о клиенте»

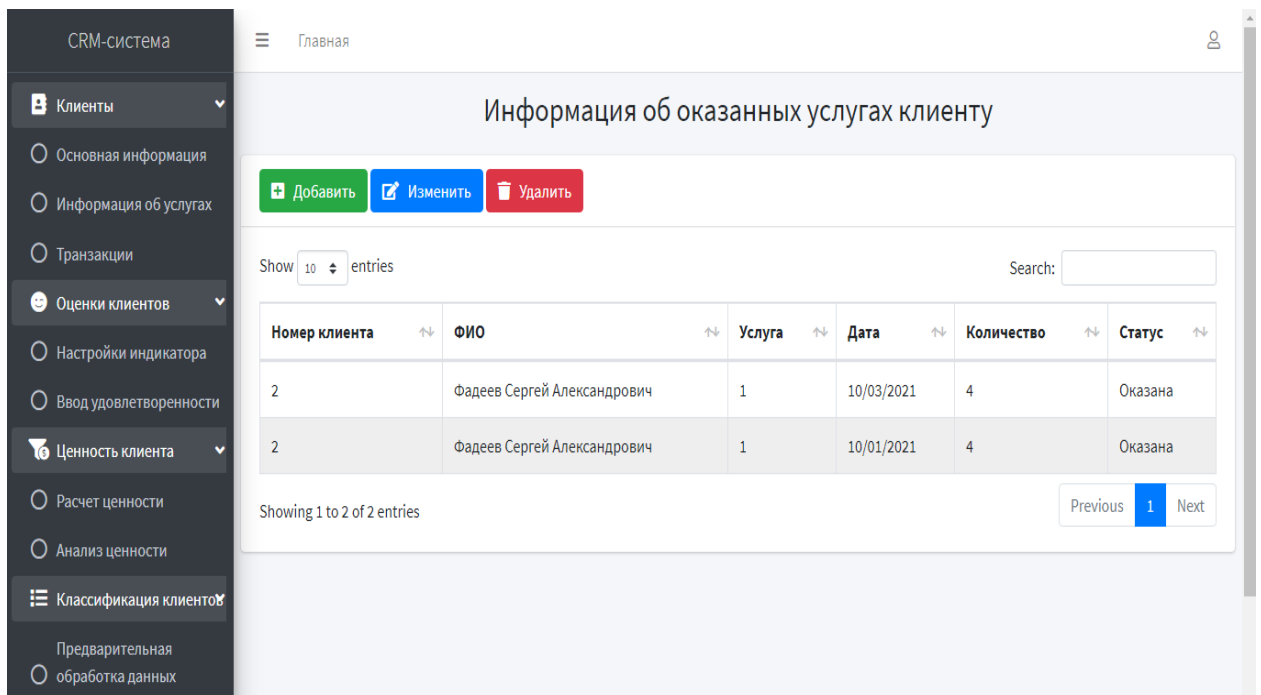


Рисунок 25 - Экранная форма «Информация об оказанных услугах»

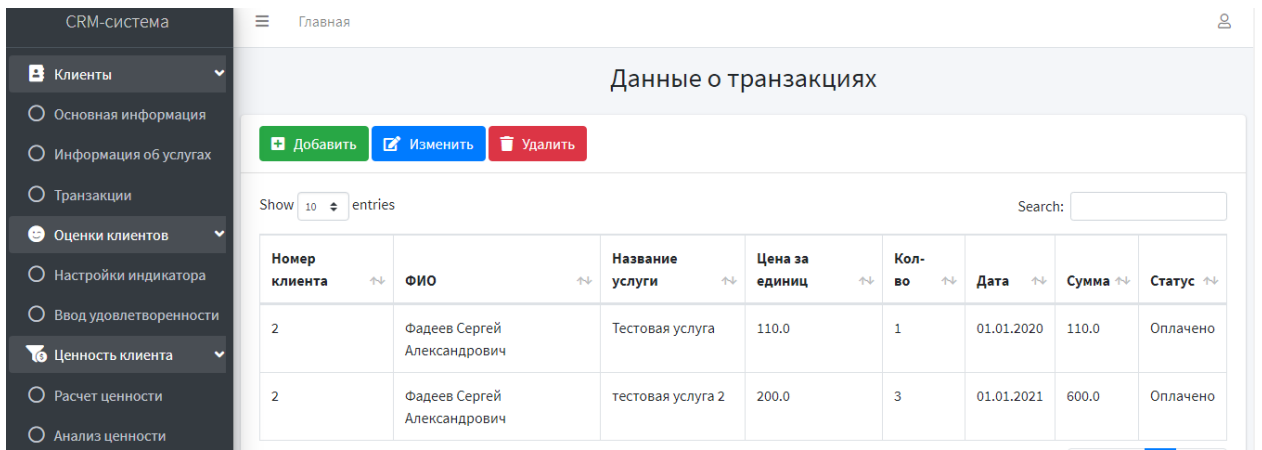


Рисунок 26 - Экранная форма «Данные о транзакциях»

Оценка удовлетворенности клиента.

Удовлетворенность клиента делится на 5 уровней:

- очень недовольный,
- недовольный,
- удовлетворенный,
- доволен,
- очень доволен.

Показатели удовлетворенности можно разделить на три уровня. Каждый уровень показателей оценки соответствует определенной доле предыдущих показателей оценки, их взаимосвязь представлена в таблице 7.

Таблица 7 - Уровни показателей удовлетворенности

Первый уровень	Второй уровень	Третий уровень
Общая оценка	Услуги	Доступность
		Надежность
		Своевременность
		Полнота
	Сервис	Индивидуальное обслуживание
		Обратная связь
	Персонал	Качество консультационных услуг
		Отношение обслуживающего персонала
	Потребительский опыт	Качество рекламы
		Рекомендации окружающих

Требуемые настройки можно сделать на вкладке «Настройки индикатора», как показано на рисунке 27.

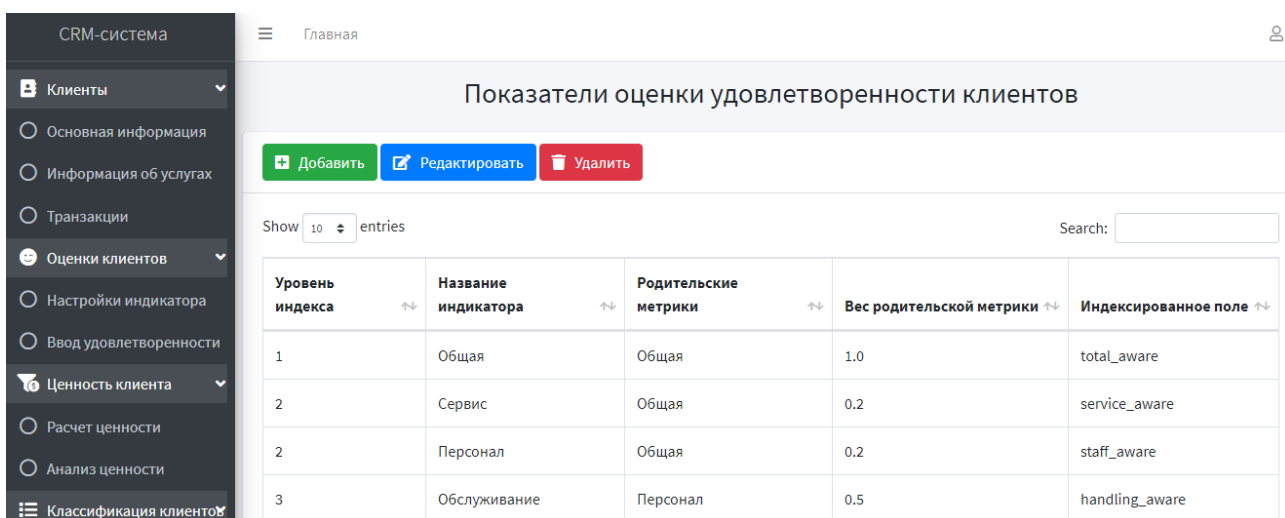


Рисунок 27 - Экранная форма «Настройки индикатора»

Клиенты могут использовать различные уровни показателей для оценки (рассмотрим в качестве примера вторичные показатели).

После успешного ввода данных будут рассчитаны результаты оценки удовлетворенности, как показано на рисунке 28.

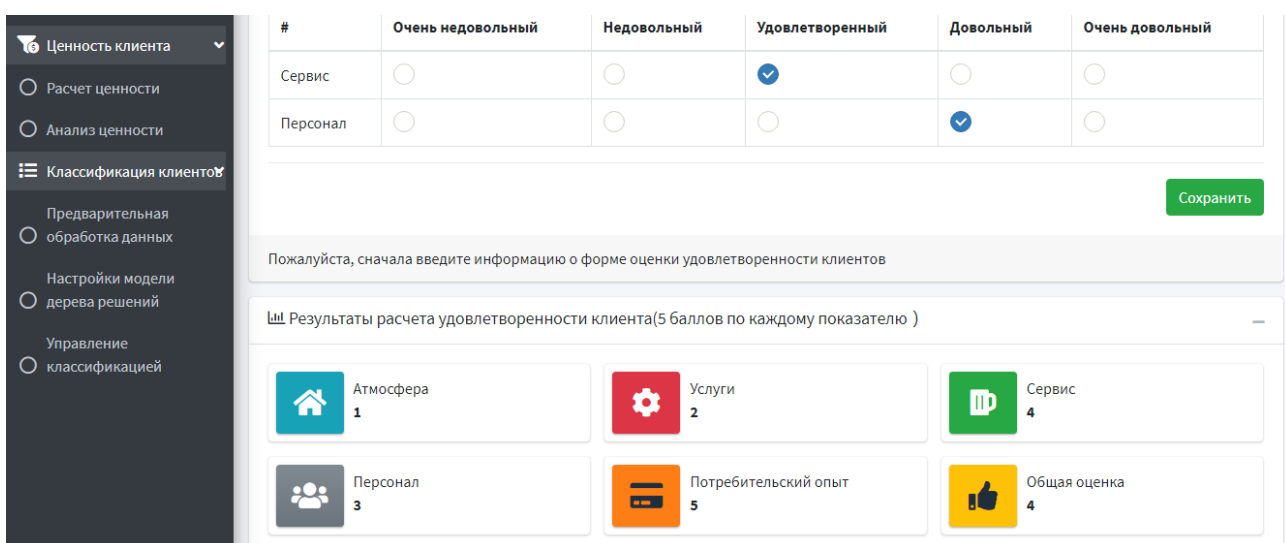


Рисунок 28 - Экранная форма «Ввод оценки удовлетворенности клиента»

Расчет ценности клиента.

Первоначально необходимо выбрать клиента, который может быть найден непосредственно через раскрывающийся список поиска, как показано на рисунке 29.

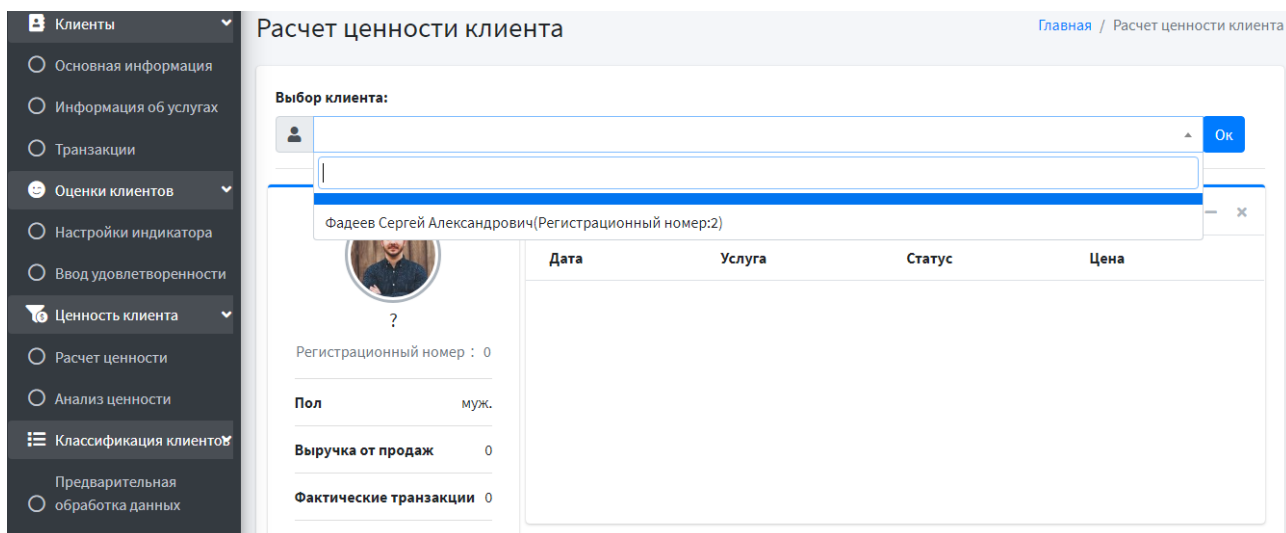


Рисунок 29 - Поиск клиента для расчета его ценности

Затем нажимаем кнопку «Ок», чтобы вывести данные о клиенте, как показано на рисунке 30.

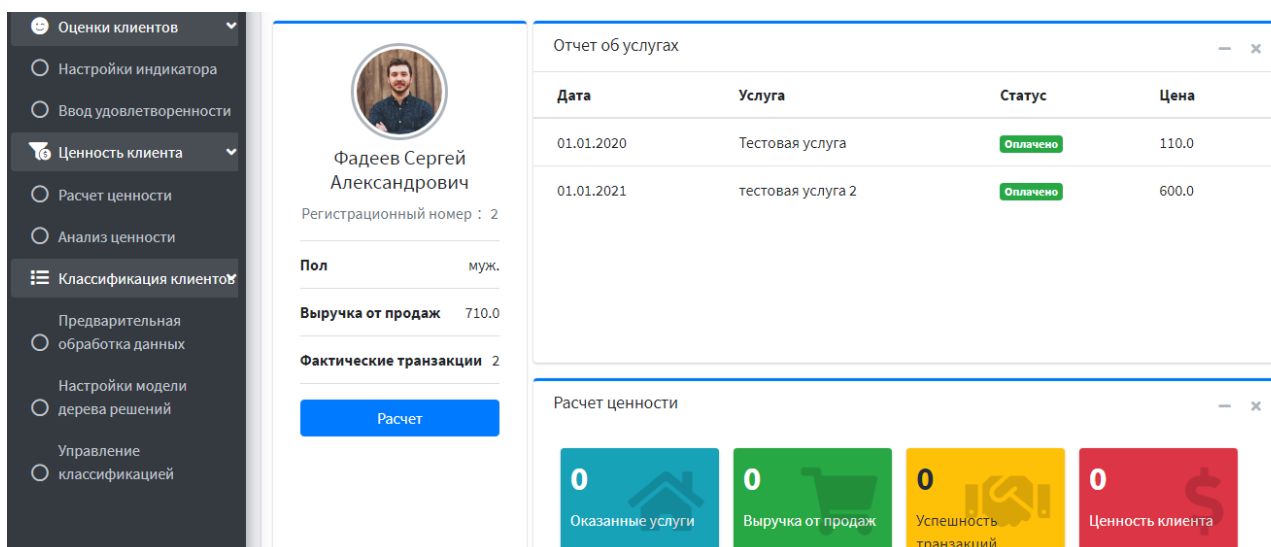


Рисунок 30 - Данные о клиенте

После нажимаем кнопку «Расчет». Получим ценность клиента по установленной формуле расчета, как показано на рисунке 31.

Оценки клиентов

- Настройки индикатора
- Ввод удовлетворенности
- ▶ Ценность клиента
- Расчет ценности
- Анализ ценности
- ☰ Классификация клиентов
- Предварительная обработка данных
- Настройки модели дерева решений
- Управление классификацией

Отчет об услугах

Дата	Услуга	Статус	Цена
01.01.2020	Тестовая услуга	Оплачено	110.0
01.01.2021	тестовая услуга 2	Оплачено	600.0

Расчет ценности

- 4,25 Оказанные услуги
- 2,727 Выручка от продаж
- 1 Успешность транзакций
- 2,742 Ценность клиента

Рисунок 31 - Экранная форма «Расчет ценности клиента»

Классификация клиентов.

Для классификации клиентов, необходимо на левой панели навигации перейти на вкладку «Предварительная обработка данных», как показано на рисунке 32.

Клиенты

- Основная информация
- Информация об услугах
- Транзакции
- ☺ Оценки клиентов
- Настройки индикатора
- Ввод удовлетворенности
- ▶ Ценность клиента
- Расчет ценности
- Анализ ценности
- ☰ Классификация клиентов
- Предварительная обработка данных

Классификация клиентов

Кластерный анализ (K-Means)

Выберите репрезентативную выборку, выполните кластерный анализ обучающей выборки, узнайте характер классификации клиентов и подготовьтесь к созданию дерева решений.

Период: 2020 Значение K: 7 **Ок**

Ценность Возраст Годовой доход Образование Удовлетворенность

Результаты интеллектуального анализа данных :

Рисунок 32 - Экранная форма «Классификация клиентов»

Затем из раскрывающегося списка выбрать «Период», чтобы выбрать источник данных о клиентах, используемый в этой кластеризации, который представляет собой записи клиентов за определенный год или данные клиентов за все время.

Отмечаем необходимые показатели. Система будет кластеризовать клиентов на основе значений этих выбранных показателей.

Вводим значение К, указывающее количество категорий клиентов в соответствии с выбранными показателями.

Нажимаем кнопку «Ок», чтобы выполнить кластерный анализ.

После будут представлены результаты кластеризации, как продемонстрировано на рисунке 33.

Классификация клиентов Главная / Классификация клиентов

Кластерный анализ (K-Means)

Выберите репрезентативную выборку, выполните кластерный анализ обучающей выборки, узнайте характер классификации клиентов и подготовьтесь к созданию дерева решений.

Период: 2020 Значение К: 7 **Ок**

Ценность Возраст Годовой доход Образование Удовлетворенность

Результаты интеллектуального анализа данных :

Удовлетворенность	Ценность	Образование	Возраст	Клиентов в кластере	Категория
1.2666666666666666	0.7655216378625501	3.7244444444444444	1.0444444444444444	45	
2.9791666666666665	0.8629461248145756	3.9249999999999994	2.1666666666666666	48	
3.8363636363636364	1.0241042283252635	4.07	4.2	110	
4	0.9084685412476794	4.072164948453608	2.49484536082474	97	
4.103448275862069	0.94192895554172	1	3.79310344827586	29	
5	0.9366719841152228	4.101449275362318	4.21739130434782	69	
5	0.909971451520689	4.020833333333333	2.7291666666666666	48	

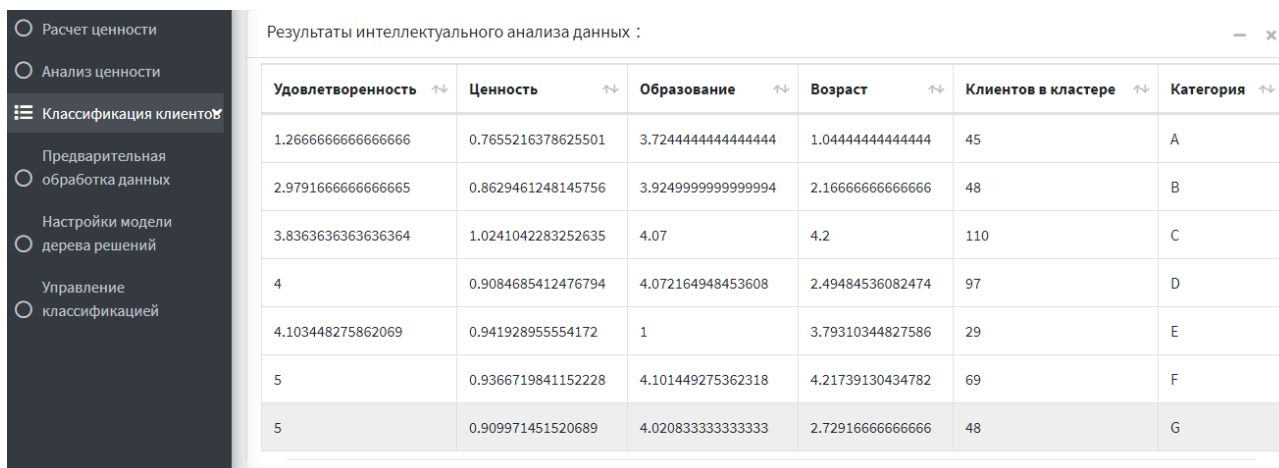
Интерпретация :
 Удовлетворенность клиентов делится на 5 уровней: очень недоволен, недоволен, удовлетворен, доволен и очень доволен. Соответствующие значения: 1, 2, 3, 4 и 5.
 Уровень образования клиента делится на 7 уровней: начальная школа, основное, среднее, средне-специальное, бакалавриат, магистратура и докторская степень. Соответствующие значения: 1, 2, 3, 4, 5, 6, 7.
 Годовой доход клиента делится на 5 уровней: менее 30 000, 30 000–50 000, 50 000–100 000, 100 000–200 000 и более 200 000. Соответствующие значения: 1, 2, 3, 4 и 5.
 Формула расчета ценности клиента: ценность клиента = ставка дохода * 0,35 + ставка дохода от услуги * 0,35 + ставка транзакции заказа * 0,3
 Возраст клиентов делится на 5 уровней: младше 20 лет, 20-30 лет, 30-40 лет, 40-50 лет и старше 50 лет. Соответствующие значения: 1, 2, 3, 4

Рисунок 33 - Результаты кластерного анализа

Каждая строка представляет кластер. В строке записываются выбранные значения индекса центра кластера и количество элементов в кластере.

Сохраняем этот результат.

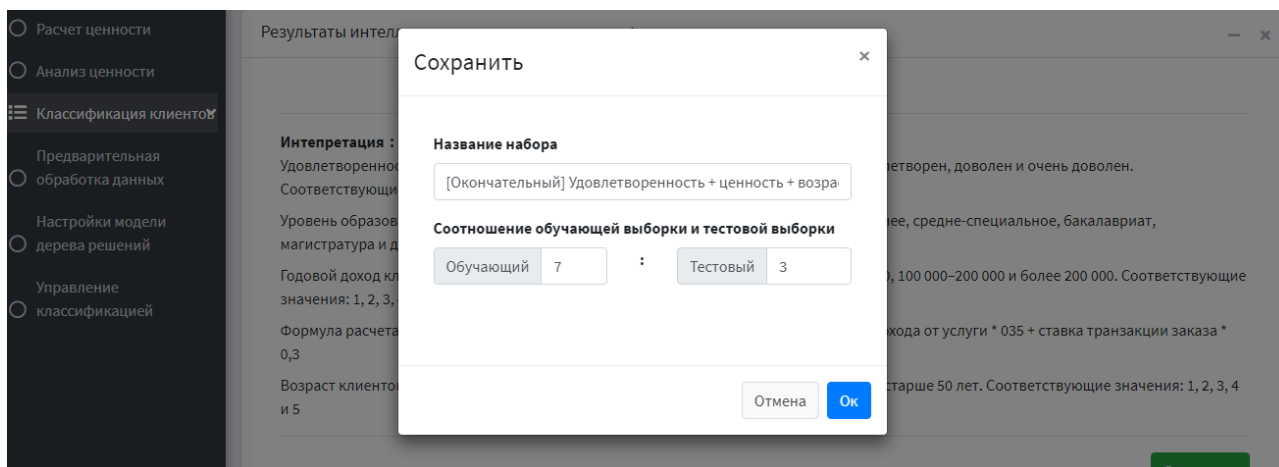
Заполняем отметки и определения пользовательских классификаций, сделанные в этой кластеризации, как показано на рисунке 34.



Удовлетворенность	Ценность	Образование	Возраст	Клиентов в кластере	Категория
1.2666666666666666	0.7655216378625501	3.7244444444444444	1.0444444444444444	45	A
2.9791666666666665	0.8629461248145756	3.9249999999999994	2.1666666666666666	48	B
3.8363636363636364	1.0241042283252635	4.07	4.2	110	C
4	0.9084685412476794	4.072164948453608	2.49484536082474	97	D
4.103448275862069	0.941928955554172	1	3.79310344827586	29	E
5	0.9366719841152228	4.101449275362318	4.21739130434782	69	F
5	0.909971451520689	4.020833333333333	2.7291666666666666	48	G

Рисунок 34 - Определение пользовательских классификаций

Нажимаем кнопку «Сохранить», указываем название набора и соотношение обучающей и тестовой выборки, как показано на рисунке 35.



Сохранить

Название набора

[Окончательный] Удовлетворенность + ценность + возра

Соотношение обучающей выборки и тестовой выборки

Обучающий 7 : Тестовый 3

Отмена Ок

Рисунок 35 - Сохранение результата кластеризации

Результаты кластеризации, сохраненные на предыдущем шаге, используются для построения классификационной модели, при помощи алгоритма дерева решений C4.5.

Нажимаем на левую навигационную панель «Настройки модели дерева решений», чтобы перейти в основной интерфейс, который продемонстрирован на рисунке 36.

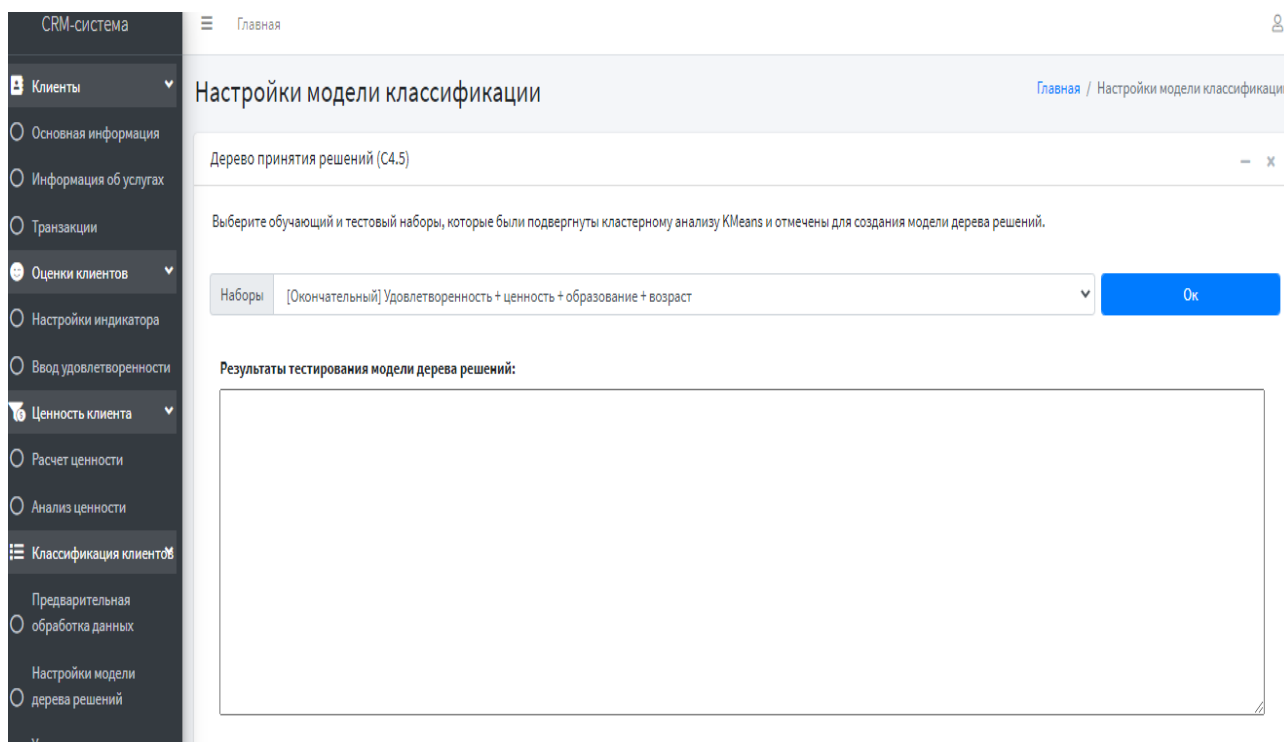


Рисунок 36 - Экранная форма «Настройка модели классификации»

Выбираем в раскрывающемся списке набор, используемый для построения классификатора, и нажимаем кнопку «Ок». Получим классификатор, построенный обучающим набором. Для проверки точности классификатора используем тестовый набор.

На рисунке 37 продемонстрирован результат, уровень точности хороший.

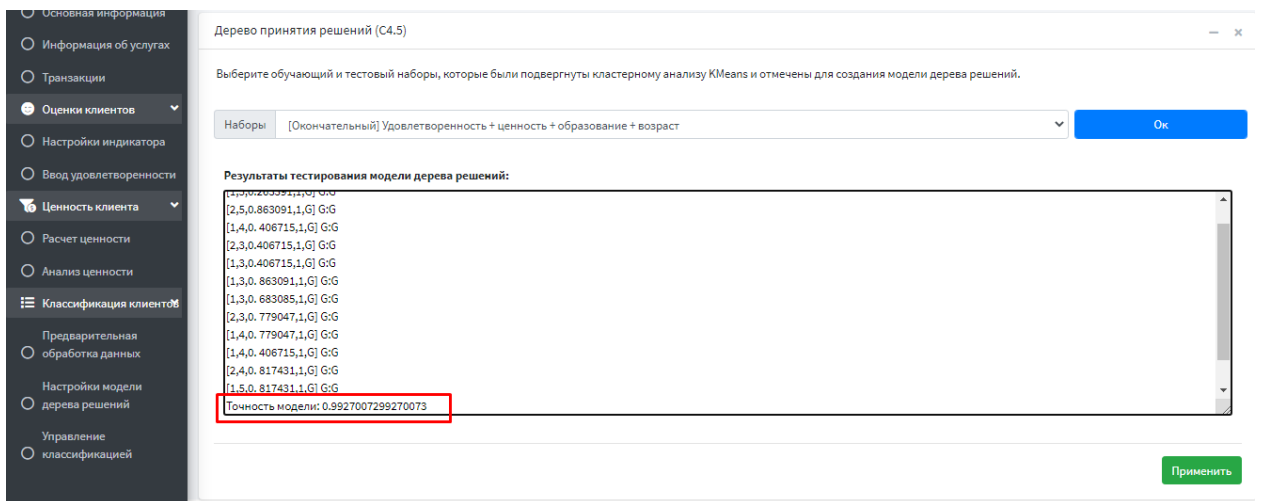


Рисунок 37 - Результат построения модели дерева решений

Нажимаем кнопку «Применить» для использования данной модели классификации в качестве классификатора клиентов системы.

Управление категориями.

Чтобы задать описание характеристик категорий клиентов и рекомендуемую маркетинговую политику для каждой категории в левой строке меню, выбираем вкладку «Управление классификацией», как показано на рисунке 38.

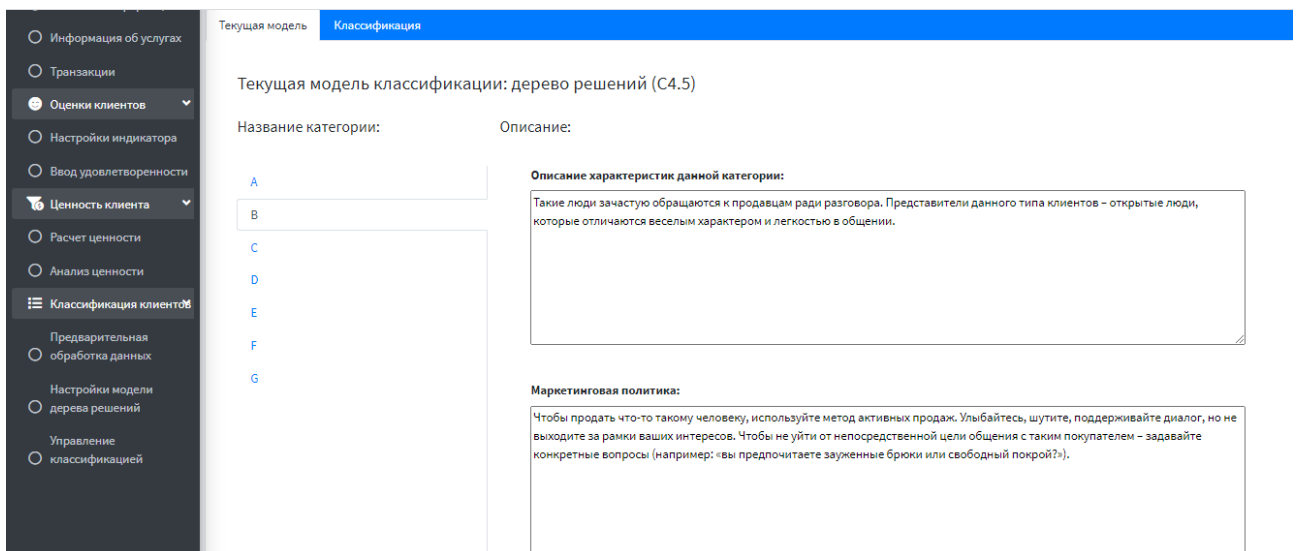


Рисунок 38 - Экранная форма «Управление классификацией»

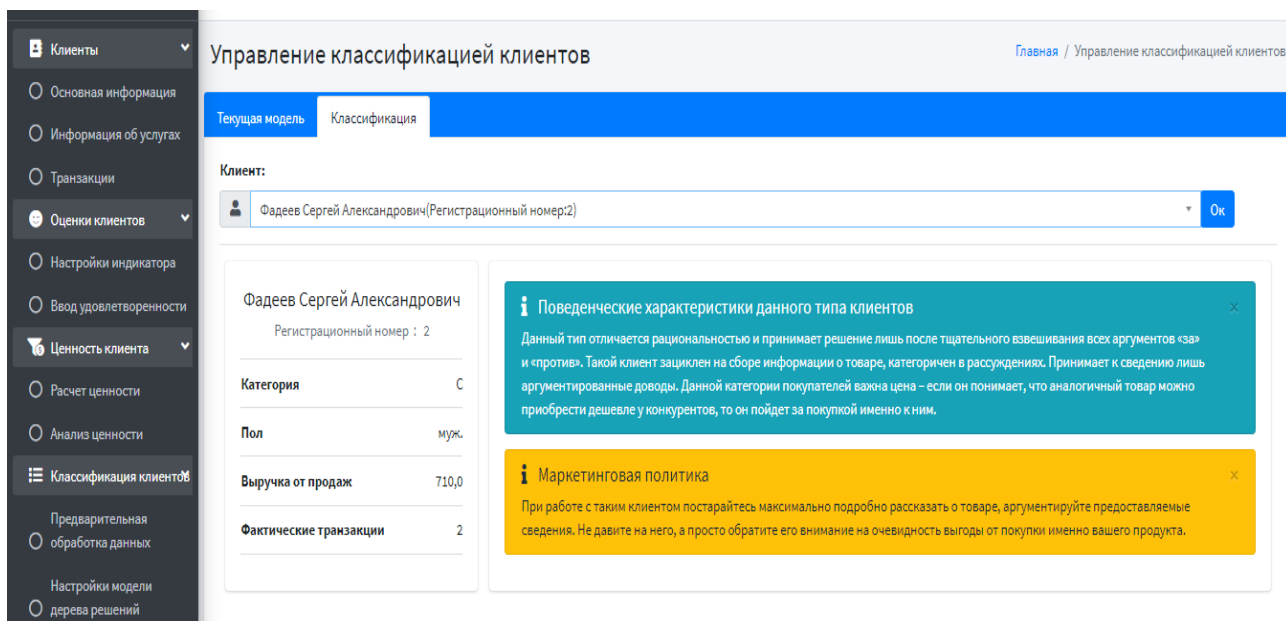


Рисунок 39 - Отчет по клиенту

Сделав запрос о клиенте на вкладке «Классификация», можно узнать, к какой категории клиент принадлежит, поведенческие характеристики данного типа клиентов и узнать рекомендуемую маркетинговую политику, как показано на рисунке 39.

4.2 Реализация информационной безопасности CRM-системы

С целью обеспечения информационной безопасности CRM-систем реализованы функции аутентификации для входа в систему и авторизации для предоставления определённому лицу прав на выполнение определённых действий, а также осуществления проверки данных прав при попытке выполнения этих действий.

В разработанной модели системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных реализована форма входа, изображенная на рисунке 40. Данная форма организует «доступ к функциям CRM-системы на основе указанной пары логина и пароля в соответствии с ролью пользователя» [7].

Рисунок 40 - Форма входа в CRM-систему

В целях исключения хранения в открытом виде паролей пользователей в базе данных системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных реализуется их хеширование. «В качестве алгоритма хеширования используется Md5, а для дополнительной безопасности реализовано хранение паролей в зашифрованном виде, но с добавлением случайных символов, уникальных для каждого пользователя» [9] системы. Фрагмент кода содержащего метод шифрования пароля представлен на рисунке 41.

```

public static String md5(String salt, String plainText)
    throws NoSuchAlgorithmException {
    MessageDigest md = MessageDigest.getInstance("MD5");

    if (salt != null) {
        md.update(salt.getBytes());
    }
    md.update(plainText.getBytes());

    byte byteData[] = md.digest();

    StringBuffer sb = new StringBuffer();
    for (int i = 0; i < byteData.length; i++) {
        sb.append(Integer.toString((byteData[i] & 0xff) + 0x100, 16)
            .substring(1));
    }
    return sb.toString();
}

```

Рисунок 41 - Фрагмент кода содержащего метод шифрования пароля

При использовании данного способа шифрования взломать пароль практически невозможно.

4.3 Расчет показателя эффективности CRM-системы

Показатель эффективности системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных в магистерской работе будет приравнен к показателю эффективности управления. Полезность и качество разработанной модели системы включены в данный показатель.

Существует несколько аспектов, рассматривающих понятие «эффективность управления»:

- «экономическая эффективность управления,
- целевая эффективность управления,
- функциональная эффективность управления» [34].

Функциональная эффективность управления, на наш взгляд, есть наиболее соответствующий показатель, формула расчета которого представлена следующим образом:

$$K_{\text{эу}} = \frac{\sum_{i=1}^n P_{yi}}{n} \quad (9)$$

где:

P_{yi} - вероятность выработки системой управления взаимоотношениями с клиентами на основе интеллектуального анализа данных эффективного управляющего воздействия при реализации i -й функции управления;

n - количество функций управления, реализуемых системой управления взаимоотношениями с клиентами на основе интеллектуального анализа данных.

В основу предложенной модели системы управления взаимоотношениями с клиентами на основе Data Mining положены три ключевых функции:

- управление взаимоотношениями с клиентами;
- классификация и кластеризация клиентов методами Data Mining;
- формирование отчетности.

Формирование отчетности является затратным и ответственным процессом. На его результат отрицательно может сказаться такая характеристика, как человеческий фактор. В этом случае значение показателя функциональной эффективности управления будет равно:

$$K_{\text{эу}} = \frac{2}{3} = 0,667.$$

Полученное из проведенных расчетов значение $K_{\text{эу}}$ больше 0,5. Данный факт позволяет сделать вывод о том, что представленная модель системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных имеет функциональную эффективность высокого уровня, что, в свою очередь, свидетельствует об ее адекватности.

Выводы к главе 4

Реализована система управления взаимоотношениями с клиентами на основе интеллектуального анализа данных, и описан основной принцип ее работы.

С целью обеспечения информационной безопасности CRM-систем реализованы функции аутентификации для входа в систему и авторизации для предоставления определённому лицу прав на выполнение определённых действий.

Коэффициент эффективности управления предлагаемой системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных имеет значение более 0,5, что позволяет говорить о высокой функциональной эффективности разработанной модели.

Заключение

Целью работы является разработка модели системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных.

В ходе проведенного исследования получены следующие основные результаты:

- обзор и анализ трудов как зарубежных, так и отечественных исследователей по предмету исследования подтвердил недостаточность работ, направленных на решение проблемы моделирования систем управления взаимоотношениями с клиентами на основе интеллектуального анализа данных, что обосновало актуальность представленной темы исследования;
- проведен аналитический обзор и сравнительный анализ на основе классификации методов интеллектуального анализа данных по задачам, в ходе которых оптимальными для кластеризации клиентов выбран метод k-means, для осуществления классификации - метод C4.5;
- для анализа бизнес-процесса управления взаимоотношениями с клиентами использовались технологии IDEF0 и DFD. Обоснована необходимость разработки модели системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных;
- разработана и реализована модель системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных. В процессе разработки модели использован объектно-ориентированный подход;
- выполнена оценка эффективности системы управления взаимоотношениями с клиентами на основе интеллектуального анализа для проверки адекватности разработанной модели.

Произведенный расчет свидетельствует, что функциональная эффективность системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных превышает 0,5, что соответствует требованиям к системам управления.

Таким образом, в работе решена актуальная научно-практическая проблема моделирования системы управления взаимоотношениями с клиентами на основе интеллектуального анализа данных.

Гипотеза исследования подтверждена.

Результаты магистерской диссертации представляют научно-практический интерес и предоставляют возможности применения предлагаемой модели для построения CRM-системы на основе интеллектуального анализа данных, эффективно влияющей на управление предприятием в целом.

Список используемой литературы и используемых источников

1. Абдикеев Н. М. Интеллектуальные информационные системы. Учебное пособие. М. : РЭА им. Плеханова, 2017. 205 с.
2. Аверкин А. Н. Нечеткие множества в моделях управления и искусственного интеллекта. М. : Наука, 2016. 104 с.
3. Аналитические возможности CRM [Электронный ресурс]. URL: <https://www.intuit.ru/studies/courses/1154/151/lecture/4214?page=3> (дата обращения: 10.03.2020).
4. Арлазаров В. Л. Теория и методы создания интеллектуальных компьютерных систем / В. Л. Арлазаров, Ю. И. Журавлев, О. И. Ларичев // Информационные технологии и вычислительные системы. 2018. №1. С. 47-48.
5. Барсягин А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / Барсягин А. А., Куприянов М. С., Степаненко В. В., Холод И. И. 2-е изд., перераб. и доп. СПб. : БХВ - Петербург, 2017. 384 с.
6. Батыршин И. З. К анализу предпочтений в системах принятия решений / И. З. Батыршин. М. : МЭИ, 2017. 423 с.
7. Волков А. А. Конспект лекций по Системам искусственного интеллекта. М. : МГСУ, 2015. 72 с.
8. Генкин Б. М. Анализ производительности труда / Б. М. Генкин. М. : НОРМАИНФРАМ, 2016. 506 с.
9. ГОСТ 19.701-90. Схемы алгоритмов, программ, данных и систем. Условные обозначения и правила выполнения (ИСО 5807-85). Введ. 1992-01-01. М. : Издательство стандартов, 1992. 14 с.
10. ГОСТ 34.003-90. Информационная технология. Комплекс стандартов на автоматизированные системы. Термины и определения. Взамен ГОСТ 24.003-84, ГОСТ 22487-77. Введ. 1992-01-01. М. : Издательство стандартов, 1992. 14 с.

11. ГОСТ 34.320-96. Информационная технология. Система стандартов по базам данных. Концепции и терминология для концептуальной схемы и информационной базы. Введ. 2001-07-01. М. : Издательство стандартов, 2001. 46 с.
12. ГОСТ 34.601-90. Информационные технологии. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Стадии создания. Введ. 1992-01-01. М. : Издательство стандартов, 1992. 6 с.
13. ГОСТ Р ИСО/МЭК 12207-99. Информационная технология. Процессы жизненного цикла программных средств. Введ. 2000-07-01. М. : Издательство стандартов, 2000. 30 с.
14. Грекул В. И. Проектирование информационных систем. Курс лекций [Электронный ресурс] : учебное пособие для студентов вузов, обучающихся по специальностям в области информационных технологий / В. И. Грекул, Г. Н. Денищенко, Н. Л. Коровкина. Москва, Саратов : ИнтернетУниверситет Информационных Технологий (ИНТУИТ), Вузовское образование, 2017. 303 с.
15. Грученков В. В. Микросервисная архитектура Web приложений / В. В. Грученков // Компьютерные системы и сети: материалы 52-й научной конференции аспирантов, магистрантов и студентов (Минск, 25-30 апреля 2016 года. Минск : БГУИТ, 2016. 4 с.
16. Истигечева Е. В. Моделирование логических схем бизнеспроцессов / Е. В. Истигечева, Т. Е. Григорьева // Информатика и системы управления. Том 2. № 48. С. 36-47.
17. Канер С. Тестирование программного обеспечения. Фундаментальные концепции менеджмента бизнес-приложений: Пер. с англ. / С. Канер, Д. Фолк, К. Нгуен. К. : Издательство «ДиаСофт», 2015. 544 с.
18. Киселева Т. В. Программная инженерия. Часть 1 [Электронный ресурс] : учебное пособие / Т. В. Киселева. Ставрополь : Северо-Кавказский федеральный университет, 2017. 137 с. URL: <http://www.iprbookshop.ru/69425.html> (дата обращения: 10.03.2020).

19. Классификация и специфицирование требований (RUP) [Электронный ресурс]. URL: <https://www.intuit.ru/studies/courses/2188/174/lecture/4726?page=2> (дата обращения: 10.03.2020).
20. Кравченко Т. К. Экспертная система принятия решений / Т. К. Кравченко, Г. И. Перминов. М. : ГУ-ВШЭ, 1999. 241 с.
21. Кулябов Д. С. Введение в формальные методы описания бизнес-процессов / Д. С. Кулябов, А. В. Королькова. М. : РУДН, 2017. 202 с.
22. Майоров С. И. Информационный бизнес: коммерческое распространение и маркетинг / С. И. Майоров. М. : Финансы и Статистика, 2015. 346 с.
23. Новиков Д. А. Теория управления организационными системами / Д. А. Новиков. Российская академия наук, 2016. 583 с.
24. Остроухов А. В. Интеллектуальные информационные системы и технологии / А. В. Остроухов, Н. Е. Суркова. Красноярск : Научно-инновационный центр, 2016. 370 с.
25. Певченко С. С. Методы анализа данных // Молодой ученый. 2015. №13. С. 167-169.
26. Саати. Т. Принятие решений. Метод анализа иерархий [Электронный ресурс]. URL: <http://www.pqmonline.com/assets/files/lib/books/saaty.pdf> (дата обращения 15.06.2020 г.).
27. Савенков П. А. Применение алгоритмов анализа данных в системах поддержки принятия управленческих решений [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/primenenie-algoritmov-analiza-dannyh-v-sistemah-podderzhki-prinyatiya-upravlencheskih-resheniy/viewer> (дата обращения 12.06.2020 г.).
28. Самуйлов С. В. Объектно-ориентированное моделирование на основе UML : учебное пособие / С. В. Самуйлов. Саратов : Вузовское образование, 2016. 37 с.
29. Севастьянова И. Г. Делегирование полномочий для принятия

эффективных решений // Власть. 2014. №1. С. 55-57.

30. Степанова Ю. Н. Основные принципы и этапы моделирования информационных систем управленческого учета // Молодой ученый. 2017. №10. С. 273-276.

31. Терентьев Н. Ю. Стили и процесс принятия решений // Проблемы учета и финансов. 2014. №1 (13). С.63-67.

32. Трахтенгерц Э. А. Компьютерная поддержка принятия решений / Э. А. Трахтенгерцу. М. : Наука, 2016. 365 с.

33. Фадеев С. А. Технология Data Mining как инструмент оптимизации взаимоотношений с клиентами // Международный научно-практический журнал «Моя профессиональная карьера», 2020. С. 45-50.

34. Фомичев А.Н. Исследование систем управления / А.Н. Фомичев. М. : Дашков и Ко, 2015. 348 с.

35. Aggarwal C. Data Streams: Models and Algorithms. Berlin. : Springer, 2016. 54 с.

36. CRM Models (Customer Relationship Management Model) [Электронный ресурс]. URL: <https://ninjaoutreach.com/crm-models/> (дата обращения: 14.03.2021).

37. Liu H. Social Computing: A Data Mining Perspective. N.Y. : Chapman & Hall/CRC, 2015. 110 с.

38. Miller H. Geographic Data Mining and Knowledge Discovery. N.Y. : CRC, 2017. 48 с.

39. Tan P. Introduction to Data Mining. N.Y. : Addison Wesley, 2018. 75 с.

40. Using CRM data: Modeling and measuring the effect of sales force knowledge on customer decision making [Электронный ресурс]. URL: http://gssi.world/wp-content/uploads/2018/11/2011-conference-023_Using-CRM-data-Modeling-and-measuring-the-effect-of-sales.pdf (дата обращения: 10.12.2020)

41. Visual Modeling with Rational Rose 2002 and UML [Электронный ресурс]. URL: https://www.oreilly.com/library/view/visual-modeling-with/0201729326/0201729326_ch01lev1sec8.html (дата обращения: 12.10.2020).

Приложение А

Диаграммы последовательности в нотации UML

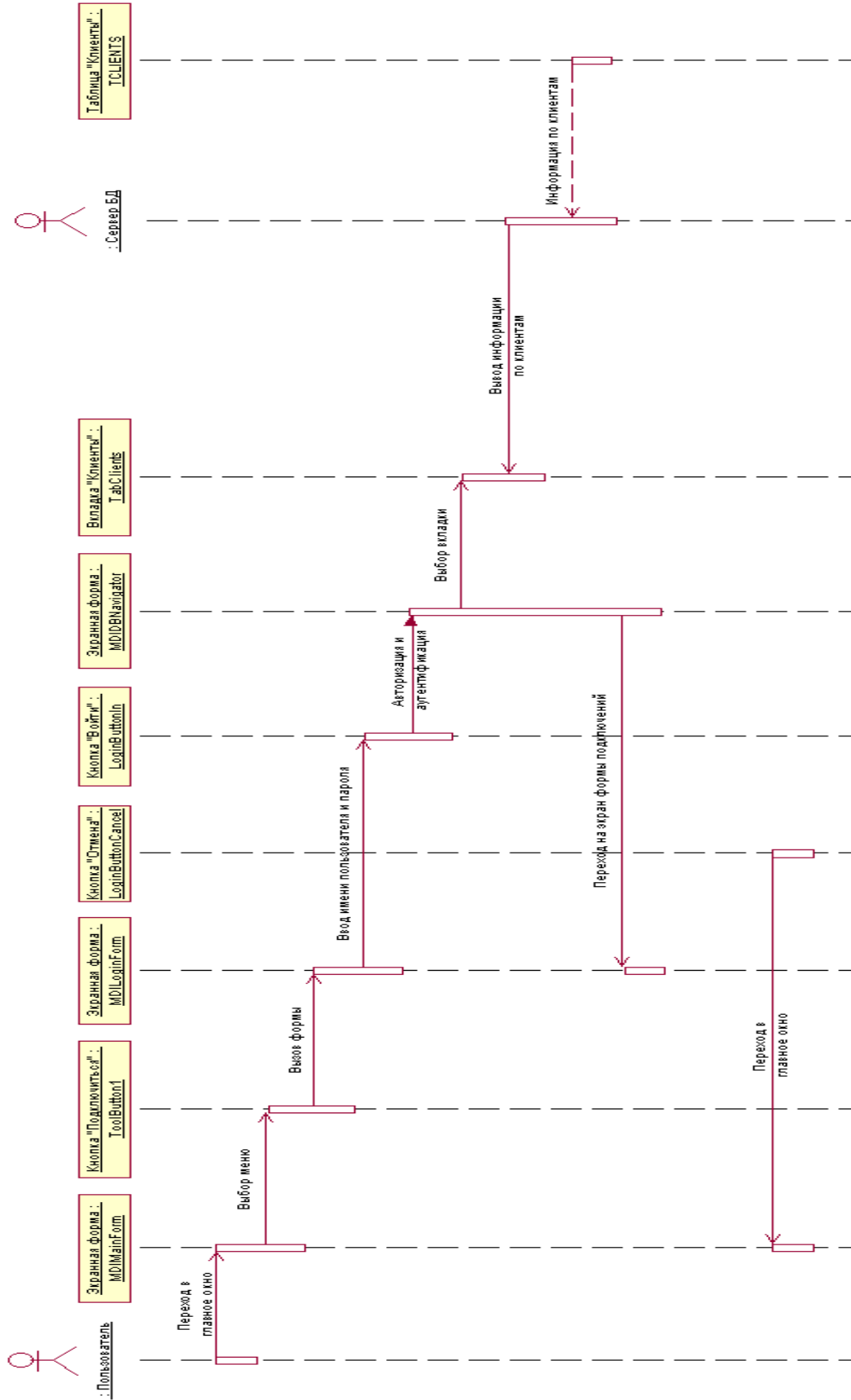


Рисунок А.1 – Диаграмма последовательности для варианта использования «Получение данных о клиенте»

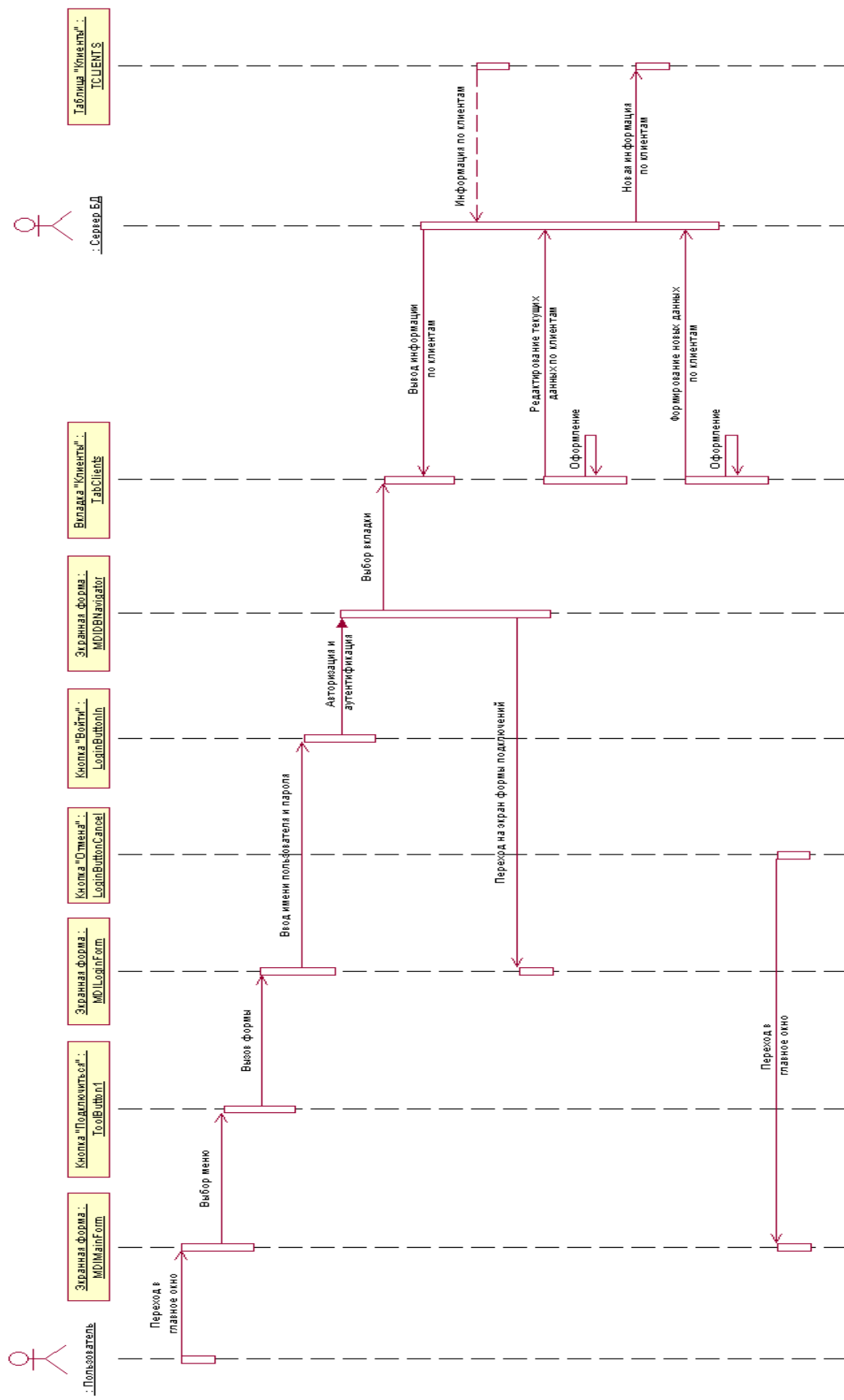


Рисунок А.2 – Диаграмма последовательности для варианта использования «Внесение данных о клиенте»

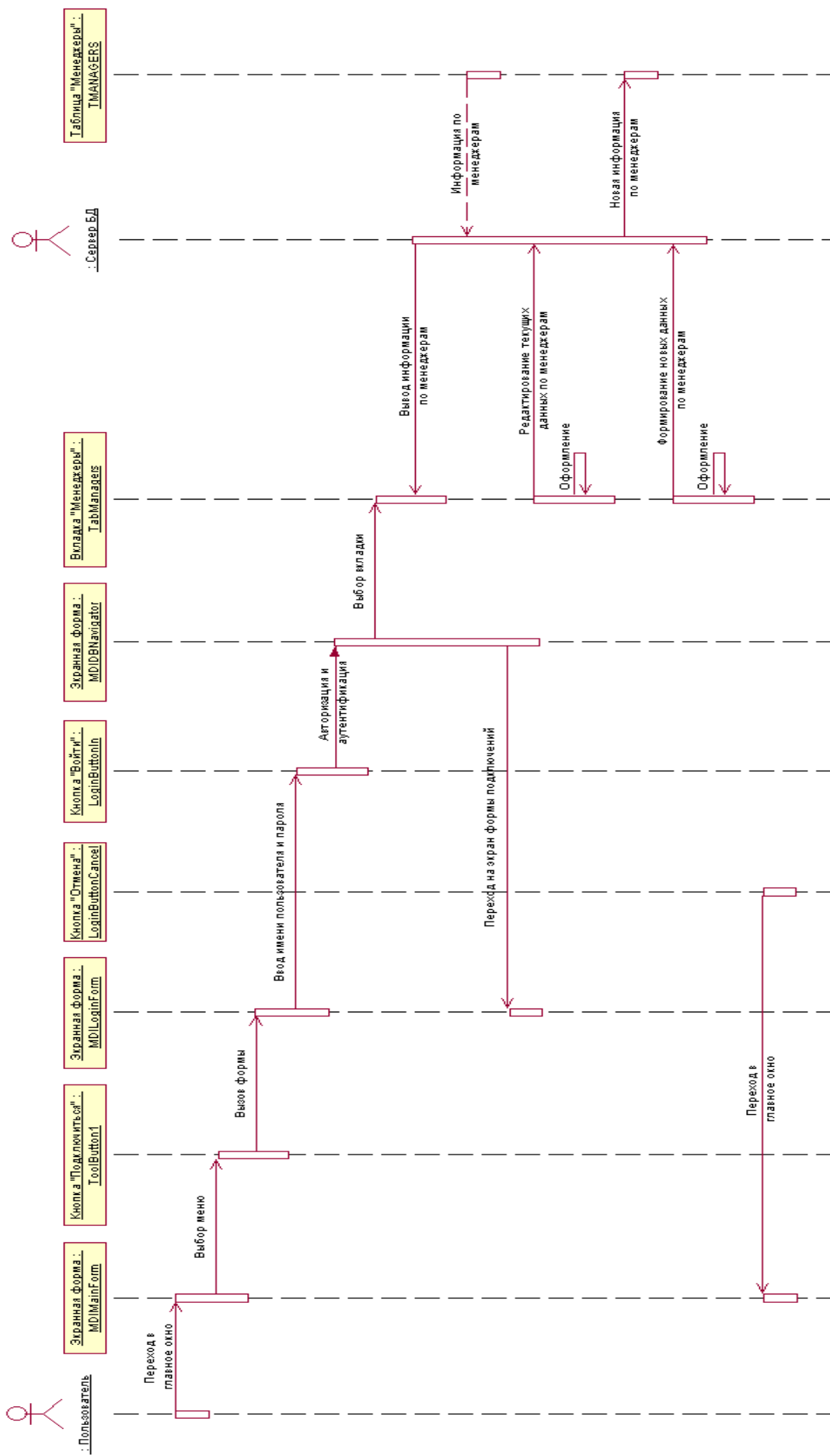


Рисунок А.3 – Диаграмма последовательности для варианта использования «Выбор направления для менеджеров»

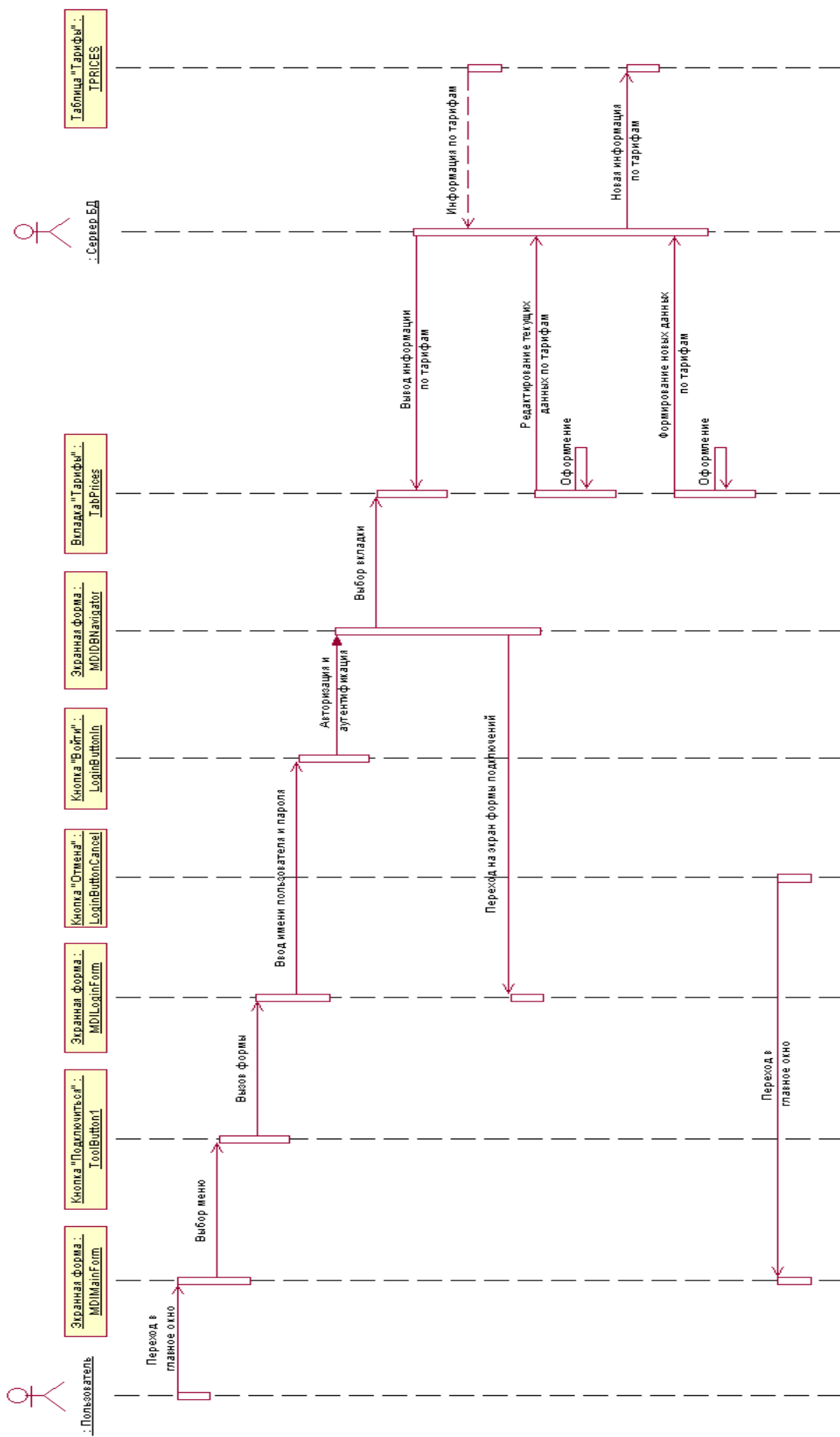


Рисунок А.4 – Диаграмма последовательности для варианта использования «Ведение тарифной политики»

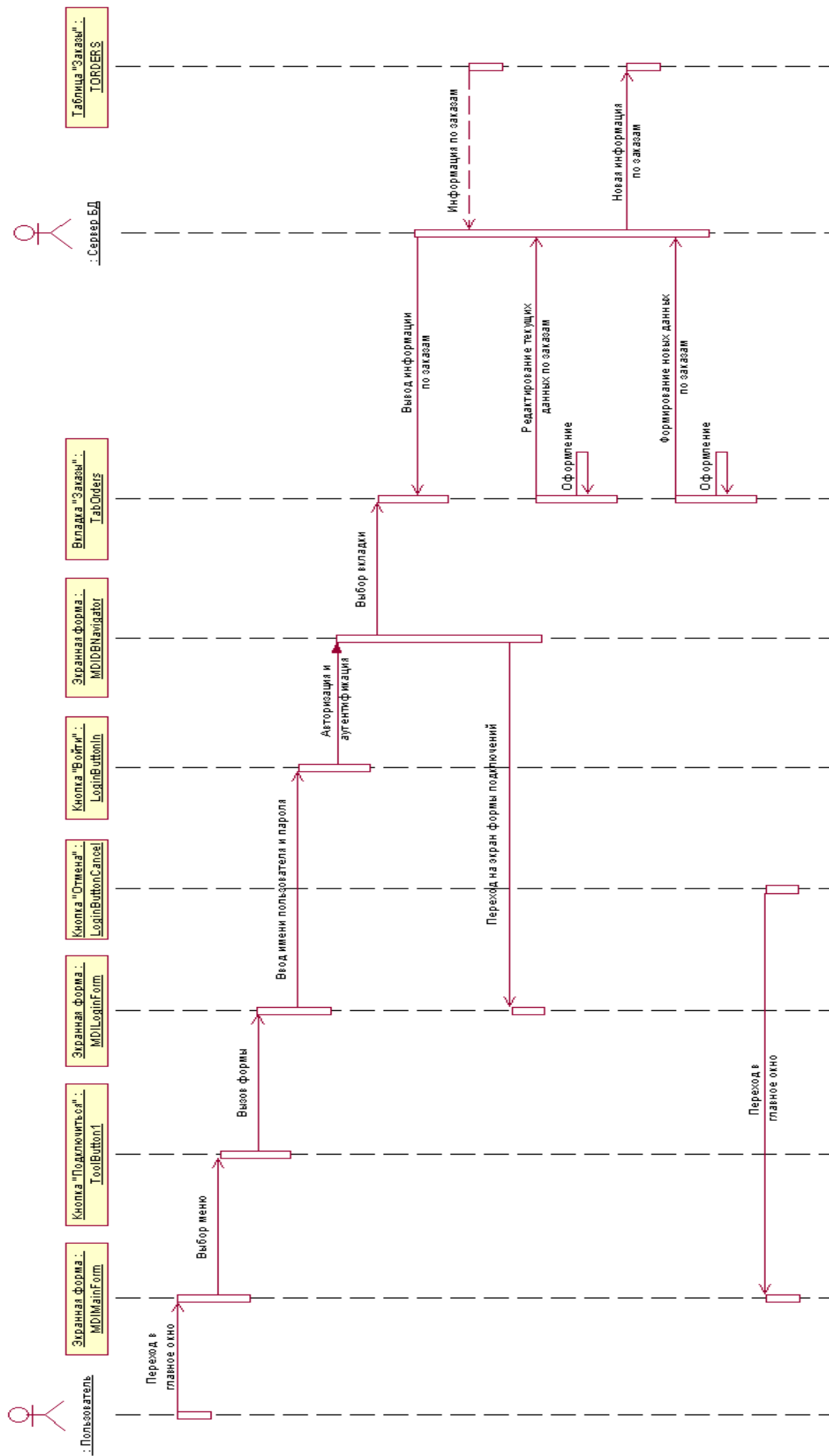


Рисунок А.5 – Диаграмма последовательности для варианта использования «Внесение данных об услугах»

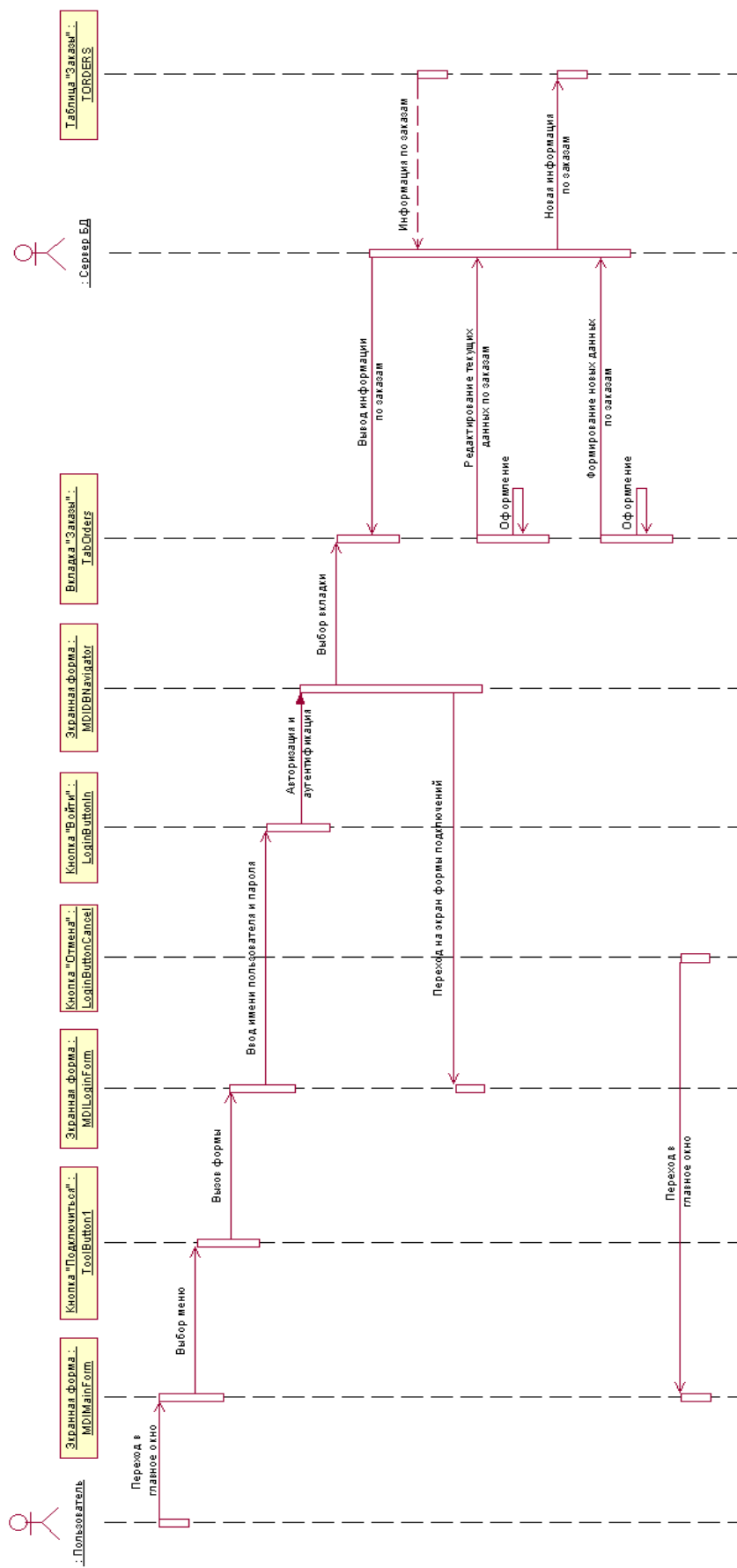


Рисунок А.6 – Диаграмма последовательности для варианта использования «Внесение данных об транзакциях»

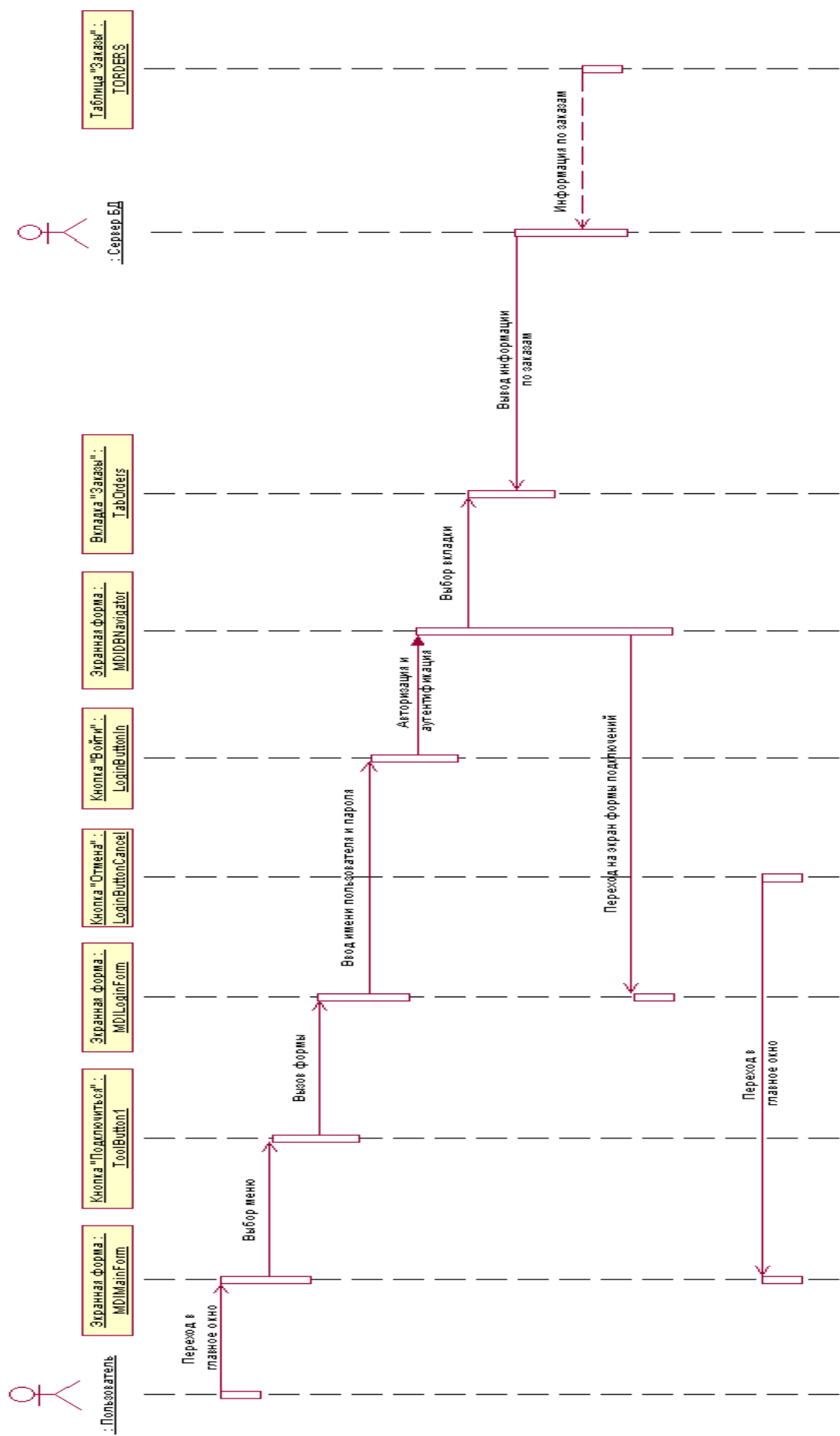


Рисунок А.7 – Диаграмма последовательности для варианта использования «Получение данных»

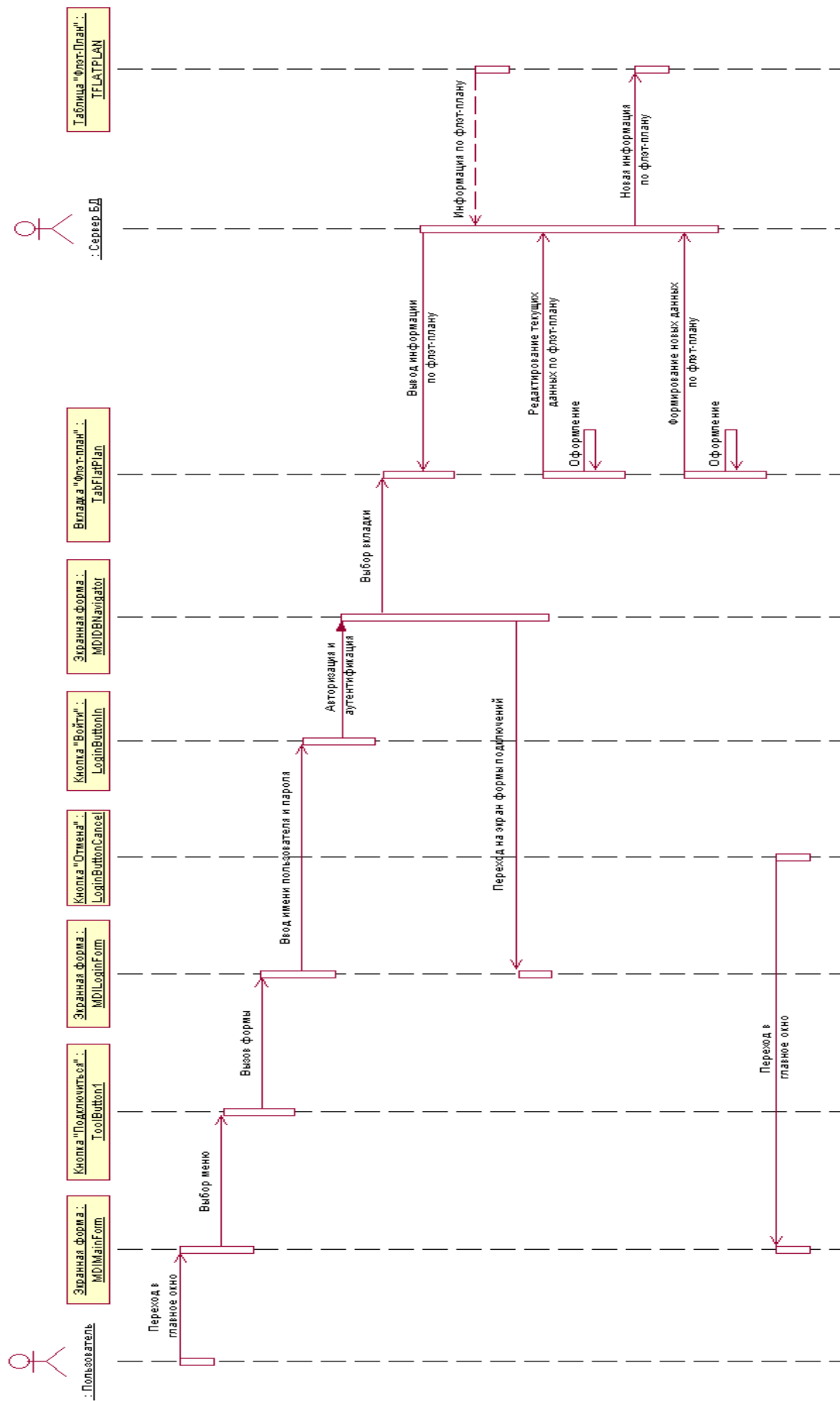


Рисунок А.8 – Диаграмма последовательности для варианта использования «Формирование плана» в нотации UML