**ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ**

**УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА»**

**(СПбГУТ)**

Факультет Инфокоммуникационных сетей и систем

Кафедра Инфокоммуникационных систем

*Допустить к защите*

Заведующий кафедрой \_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_ .

*(подпись) (Ф.И.О.)*

«\_\_» июня 2020г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

«Нейросетевая самообучающаяся модель прогнозирования нагрузки контакт-центра»

*(тема ВКР)*

Вид выпускной квалификационной работы

*магистерская диссертация* \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_

*(бакалаврская работа, дипломная работа, дипломный проект, магистерская диссертация)*

Направление/специальность подготовки

11.04.02 Инфокоммуникационные технологии и системы связи \_\_\_\_\_\_\_\_\_\_\_\_\_

*(код и наименование направления/специальности)*

*\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

Направленность (профиль) Системы управления инфокоммуникациями \_\_\_

*(наименование)*

Квалификация магистр\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*(наименование квалификации в соответствии с ФГОС ВО / ГОС ВПО)*

Студент:

Васылив Назар Иванович, ИКТМ-83м

*(Ф.И.О. Группа)*

Руководитель:

К.т.н, доцент, кафедра ИКС Кисляков Сергей Викторович

*(Уч.степень, должность, кафедра,Ф.И.О.)*

Санкт-Петербург

2020

Текст ВКР размещен в электронно-библиотечной системе университета

Руководитель отдела комплектования библиотеки \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*(Ф.И.О.)*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*(дата) (подпись)*

ВКР выполнена в соответствии с календарным планом:

|  |  |  |  |
| --- | --- | --- | --- |
| №  п/п | Наименование этапов выпускной квалификационной работы | Срок выполнения этапов ВКР | Примечание |
| 1 | Знакомство с предметной областью и оформление задания по работе | 26.02.2018-  01.03.2018 |  |
| 2 | Развитие контакт-центров и возможности ит решения по прогнозированию нагрузки | 02.03.2018-  12.03.2018 |  |
| 3 | Анализ моделей | 13.03.2018-  05.04.2018 |  |
| 4 | Разработка модели | 06.04.2018-  19.04.2018 |  |
| 5 | Применение нейронной сети | 20.04.2018-  03.05.2018 |  |
| 6 | Оформление пояснительной записки и графического материала | 04.05.2018-  16.05.2018 |  |
| 7 | Подготовка доклада и презентации ВКР, | 17.06.2018-  19.06.2018 |  |

Коэффициент оригинальности ВКР \_\_\_\_\_\_\_\_ % .

Проверил руководитель ВКР: доцент, Кисляков Сергей Викторович\_\_\_\_\_\_\_\_\_\_\_

*(Должность, Ф.И.О.)*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*(дата) (подпись)*

**ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ**

**УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА»**

**(СПбГУТ)**

Факультет ИКСС  Кафедра ИКС\_\_\_\_\_\_\_\_\_\_

Направление (специальность) 11.04.02 Инфокоммуникационные технологии и системы связи \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(код и наименование)

**ЗАДАНИЕ**

**на выполнение выпускной квалификационной работы (ВКР)**

1. Студент Васылив Назар Иванович\_\_\_\_\_\_\_\_\_\_\_\_\_\_ № группы ИКТМ-83м

(фамилия, имя, отчество)

2. Руководитель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_ (фамилия, имя, отчество, должность, уч. степень и звание)

3. Квалификация \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(наименование в соответствии с ФГОС ВО/ ГОС ВПО)

4. Вид работы \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(бакалаврская работа, дипломный проект, дипломная работа, магистерская диссертация)

5. Тема ВКР \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

утверждена приказом ректора университета от «\_\_\_» \_\_\_ 20\_\_\_ г. № \_\_\_\_

6. Исходные данные (технические требования): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

7. Содержание работы (анализ состояния проблемы, проведение исследований, разработка, расчеты параметров, экономическое обоснование и др.)

1.ВВЕДЕНИЕ

2. РАЗВИТИЕ КОНТАКТ-ЦЕНТРОВ И ВОЗМОЖНОСТИ ИТ РЕШЕНИЯ ПО ПРОГНОЗИРОВАНИЮ НАГРУЗКИ

3.МЕТОДЫ ПРОГНОЗИРОВАНИЯ ПРИМЕНИТЕЛЬНО К КОНТАКТ-ЦЕНТРАМ

4. РАЗРАБОТКА АЛГОРИТМА

5 ЗАКЛЮЧЕНИЕ

6. СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

7.

8. 9.

8. Вид отчетных материалов, представляемых в ГЭК (пояснительная записка, перечень, графического материала, отчет о НИР, технический проект, образцы и др.): пояснительная записка, презентация

9. Консультанты по ВКР с указанием относящихся к ним разделов

|  |  |  |  |
| --- | --- | --- | --- |
| Раздел | Консультант | Дата | |
| Задание  Выдал | Задание  Принял |
| 1. Развитие контакт-центров и возможности ит решения по прогнозированию нагрузки | Кисляков С.В. |  |  |
| 2. Методы прогнозирования применительно к контакт-центрам | Кисляков С.В. |  |  |
| 3. Разработка алгоритма | Кисляков С.В. |  |  |
| 4. |  |  |  |

Дата подпись Васылив Н.И.(ФИО обучающегося)

**РЕФЕРАТ**

Магистерская работа содержит:

* 83 страниц;
* 23 рисунка;
* \_\_\_\_ таблиц;
* \_\_\_\_ приложение.

**Ключевые слова**: контакт-центр, нейронные сети, машинное обучение, deep learning, временные ряды, прогнозирование\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

В ходе выполнения магистерской работы был разработан алгоритм прогнозирования нагрузки нововведенного в работу контакт-центра \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ .

Работа написана мною самостоятельно и не содержит неправомерных заимствований

Дата подпись Васылив Н.И (ФИО обучающегося)

СОДЕРЖАНИЕ

[РЕФЕРАТ 6](#_Toc43860818)

[СОДЕРЖАНИЕ 7](#_Toc43860819)

[ВВЕДЕНИЕ 9](#_Toc43860820)

[Глава 1. РАЗВИТИЕ КОНТАКТ ЦЕНТРОВ И ВОЗМОЖНОСТИ ИТ РЕШЕНИЯ ПО ПРОГНОЗИРОВАНИЮ НАГРУЗКИ. 12](#_Toc43860821)

[1.1 Развитие контакт-центров 12](#_Toc43860822)

[1.1.1 Конвергенция и операторские центры 12](#_Toc43860823)

[1.1.2 Предпосылки в изменении архитектуры и функций операторских центров 12](#_Toc43860824)

[1.1.3 Функциональные особенности контакт-центра 22](#_Toc43860825)

[1.1.4 Доступ к услугам контакт центра 22](#_Toc43860826)

[1.2 Оптимизация расходов контакт-центра 24](#_Toc43860827)

[1.2.1 Workforce Management 25](#_Toc43860828)

[1.2.2 Plan and Forecast Workforce Management 27](#_Toc43860829)

[1.2.3 Manage Schedules & Appointments 28](#_Toc43860830)

[1.3. Возможности ИТ-решения для задачи прогнозирования. 28](#_Toc43860831)

[1.4 Вывод по главе 30](#_Toc43860832)

[Глава 2. МЕТОДЫ ПРОГНОЗИРОВАНИЯ ПРИМЕНИТЕЛЬНО К КОНТАКТ-ЦЕНТРАМ 32](#_Toc43860833)

[2.2. Архитектура и функциональные возможности Аргус решения 32](#_Toc43860834)

[2.2.1 Функциональные возможности системы 33](#_Toc43860835)

[2.2.2 Анализ пиков 36](#_Toc43860836)

[2.2.3 Учет трендовой и сезонных составляющих при формировании прогноза. 37](#_Toc43860837)

[2.2.4 Расчёт числа операторов 40](#_Toc43860838)

[2.3. Методы прогнозирования временных рядов. Сравнительный анализ. 42](#_Toc43860839)

[2.3.1 Классические модели. 43](#_Toc43860840)

[2.3.2 Модели машинного обучения 47](#_Toc43860841)

[2.3.2.1 Основные понятия в нейронных сетях 47](#_Toc43860842)

[2.3.2.2 Обычные нейронные сети и Deep Learning модели 50](#_Toc43860843)

[2.3.2.3 Регрессионные модели. 58](#_Toc43860844)

[2.4 Выбор и обоснование модели нейронной сети. 59](#_Toc43860845)

[2.5 Вывод 61](#_Toc43860846)

[Глава 3. РАЗРАБОТКА АЛГОРИТМА 63](#_Toc43860847)

[3.1 Выбор среды моделирования 63](#_Toc43860848)

[3.1.1 Google Colaboratory (так верно) 63](#_Toc43860849)

[3.1.3 NumPy 64](#_Toc43860850)

[3.1.4 Matplotlib 64](#_Toc43860851)

[3.1.5 Pandas 65](#_Toc43860852)

[3.1.6 Math 65](#_Toc43860853)

[3.1.7 Keras 65](#_Toc43860854)

[3.1.8 Scikit-learn 66](#_Toc43860855)

[3.1.9 Statsmodels 67](#_Toc43860856)

[3.1.10 SciPy 67](#_Toc43860857)

[3.2 Разработка модели 67](#_Toc43860858)

[3.2.1 Приведение временного ряда к стационарности 67](#_Toc43860859)

[3.2.2 Построение модели нейронной сети 71](#_Toc43860860)

[3.3 Моделирование и сравнительный анализ результатов 72](#_Toc43860861)

[3.3.1 Подготовка данных к подаче в нейронную сеть. 72](#_Toc43860862)

[3.3.2 Работа модели и результаты 74](#_Toc43860863)

[3.3.3 Выводы 76](#_Toc43860864)

[ЗАКЛЮЧЕНИЕ 78](#_Toc43860865)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 80](#_Toc43860866)

ВВЕДЕНИЕ

Торгово-рыночные отношения предполагают наличие конкурентной борьбы. Эта аксиома касается так же и предприятий, имеющих контакт-центры. Понимая это, любой владелец контакт-центра будет пытаться совершенствовать свою бизнес-модель уже на как на этапе проектирования, так и с первых дней работы. Существует множество стандартных решений, без которых нельзя обойтись при создании контакт-центра, множество решений, которые проверены теорией и практикой, что идеально подходят под бизнес. Некоторые из них не требуют совершенствования и без них контакт-центр не выдержит конкуренции. Однако все еще есть то, что можно совершенствовать.

Когда идеи о улучшении предоставляемых услуг исчерпаны, уникальность предприятия определена, можно обратить внимание на улучшение бизнес-процесса. Прогресс не стоит на месте и появляются новые технологии, способные выводить устоявшиеся процессы на новый уровень.

Подсистемы сбора, обработки информации, предоставления отчетности и прогнозирования работы контакт-центра являют одними из ключевых. Они помогают ответственным лицам эффективнее управлять и организовывать работу. Если инструменты сбора, обработки информации и предоставления отчетности реализованы универсально для множества контакт-центров, то с прогнозированием все не совсем так. Существуют стандартные методы расчетов и программное обеспечение. Но есть и куда двигаться, нельзя найти систему для разных поступающих потоков запросов разных контакт-центров с разной логикой обслуживания, что можно вводить в эксплуатацию “из коробки”, не требующие или требующие минимальных настроек в ходе работы.

При создании контакт-центра всегда необходимо знать предполагаемую нагрузку в виде количества поступающих вызовов за единицу времени. Это поможет грамотно распределить ресурсы, как сетевые, так и человеческие, в смысле количества операторов, обслуживающих вызовы. Исходя из нагрузки мы можем знать сколько операторов должно быть нанято в контакт-центр и выведено в работу. Чем точнее мы знаем какая нагрузка поступает на контакт-центр, тем большая вероятность задействовать необходимое количество операторов соответственно нагрузке, следовательно обслужить вызовы с заданным качеством и грамотно распределить денежные ресурсы на оплату труда. Это значит, что экономия денежного ресурса зависит от того, как точно мы предсказали нагрузку за определенный промежуток времени. Причем чем меньший промежуток времени и точней прогноз, тем мы лучше знаем сколько нужно операторов и сколько мы должны заплатить за часы их работы. Такая интересная задача уже реализуется на практике, однако в работе предлагается не просто построить прогноз на наборе уже известных данных, а параллельно реализовать алгоритм, при котором контакт-центр, якобы, только введен в эксплуатацию и только собирает данные для своего решения по прогнозированию. Так будет сымитирован процесс настройки модели прогнозирования в течении года за каждую неделю и получены результаты о том, когда модель достаточно настроена для введения в эксплуатацию.

Сейчас существует большое количество математических методов решения задач предсказания. Помимо классических статистических моделей для решения задач предсказания временных рядов применяются также и модели машинного обучения.

Современной теоретической альтернативой таким методам прогнозирования могут быть Deep Learning (глубокое обучение) методы прогнозирования, в частности нейронные сети.

Прежде всего нужно разобрать методы, помогающие определить какого вида последовательность данных будет разработана нейронной сетью. Нейронную сеть нужно сконструировать, добавив в нее алгоритм обучения.

Цель работы:

Построить нейросетевую модель и, обучая ее на известных данных, выяснить, какой объем исторических данных требуется для введения модели в эксплуатацию.

Цель разделим на следующие задачи:

1. Рассмотрение развития контакт-центров

2. Рассмотрение существующих моделей прогнозирования

3. Разработка модели нейронной сети

4. Применение нейронной сети для решения задачи предсказания поступающей нагрузки на контакт-центр

5. Анализ результатов применения

Глава 1. РАЗВИТИЕ КОНТАКТ ЦЕНТРОВ И ВОЗМОЖНОСТИ ИТ РЕШЕНИЯ ПО ПРОГНОЗИРОВАНИЮ НАГРУЗКИ

1.1 Развитие контакт-центров

1.1.1 Конвергенция и операторские центры

Необычайно быстрый рост Интернет в совокупности с появлением технологий и стандартов, обеспечивающих конвергенцию компьютерного и телекоммуникационного миров, может коренным образом изменить (и уже меняет) всю систему связи в мире. Сегодняшние тенденции развития обслуживания пользователей телекоммуникациями, а также бизнеса и коммерции в целом, пересекаясь с современными технологиями связи, формируют качественно новый уровень требований, предъявляемых к Call-центрам, к их архитектуре и функциональным возможностям.   Присутствие во «Всемирной паутине“ сегодня является не роскошью, а жизненной необходимостью для любой компании, претендующей на коммерческий успех.

1.1.2 Предпосылки в изменении архитектуры и функций операторских центров

Присутствие во «Всемирной паутине» сегодня является не роскошью, а жизненной необходимостью для любой компании, претендующей на коммерческий успех. Не менее значительное влияние на эволюцию операторских центров оказывает динамичное развитие электронного бизнеса и коммерции. Электронная коммерция позволяет избежать известных трудностей и неудобств, связанных, скажем, с какой-либо покупкой. По крайней мере, с точки зрения самого покупателя больше никто не заставляет клиента ехать через весь город, преодолевать тяжелые погодные условия, долго выбирать необходимый товар и выстаивать в очереди к кассе. Благодаря существованию электронной коммерции покупка необходимого товара в другом регионе страны может стать менее обременительной, чем покупка его в соседнем магазине. Кроме того, сама операция купли/продажи удобна для покупателя. По крайней мере, электронный магазин всегда открыт, и клиент может его посетить в удобное для него время, Таким образом, возможности привлечения новых клиентов — огромны. Но при этом нельзя не отметить, что наряду с открывающимися возможностями существует и определенная доля риска. Если клиент испытывает неудобства с Web-сайтом  например, он не может найти интересующую его информацию, не понимает, каким образом следует запомнить ту или иную форму, сайт не работает или не оправдывает его надежд — он просто покинет этот сайт и пойдет к конкурентам, На сколько клиенту проще совершить необходимые покупки, не выходя из дома. настолько же ему проще перейти на следующий сайт, если с предыдущим у него возникли какие-либо проблемы.

Итак, в условиях современного развития бизнеса и роста популярности Интернет практически любая компания, любая коммерческая или справочно-информационная служба имеет свой собственный Web-сайт. С другой стороны, процент посетителей этих сайтов, которые, в конечном итоге. пользуются предложенными услугами, пока еще очень мал. В первую очередь, это объясняется неуверенностью клиента в правильности своего выбора, недоверием к Интернет как к коммерческой структуре, излишней сложностью маршрута, который он должен пройти по сайту, чтобы получить желаемый результат.

Приведем характерный пример. Хотя сезон рождественской распродажи 1999 г на Западе принято считать началом эры электронной коммерции, по оценкам ряда экспертов суммы прибылей, недополученных многими компаниями в электронном бизнесе в этот период, огромны, Одно из исследований показывает, что более 60% покупателей, сделавших покупки в свою электронную корзину, не доводит этот процесс до конца. А другое исследование оглашает конкретную сумму недополученных за 1999 г, прибылей— 1.6 миллиардов USD. При этом, по данным психологов, если бы в момент сомнений при выборе покупки клиент имел возможность обратиться к оператору (т.е. воспользоваться привычным речевым способом общения с представителем фирмы), то процент совершенных сделок был бы гораздо выше.

Бурное развитие электронной коммерции и популярность технологии IP как единого транспорта для передачи информации любого вида дало возможность связать с ресурсами операторского центра и предложить клиентам персонализированное обслуживание и удобное средство общения. Интеграция традиционных операторских центров и Интернет позволяет превратить менеджмент взаимоотношений с клиентами в мощный инструмент, повышающий эффективность бизнеса компаний и удобства клиентов.

Покажем это на примерах, Турфирмы могут предложить персонализированную услугу планирования тура или привлечь большее количество потенциальных путешественников максимально заинтересовавших в предстоящей поездке. Web-страницы могут направляться на компьютер клиента - посетителя Web-сайта турфирмы — и этот клиент может увидеть картинку тропического острова, образец дорожного чемодана. отель, в котором его предполагается разместить, или другую информацию. причем все это  в ходе текстового или речевого общения с оператором центра в реальном времени, финансовые учреждения могут привлекать к себе потенциальных клиентов, предоставляя им легкодоступную полезную информацию в отношении сложных товаров и услуг. Посетитель Web-сайте компании может одним щелчком мыши установить речевую связь с оператором-агентом, чтобы получить детальную информацию о предоставляемых услугах или ответу на частные вопросы, не оговоренные на Web-cайте. Отделы технической поддержки и обслуживания заказчиков могут оказывать консультативные услуги, переключая постеливший от заказчика телефонный вызов к квалифицированному техническому специалисту, который может ответить на технические вопросы заказчика по телефону, часто — сопровождая ответ информацией передаваемой через Web. Менеджмент взаимоотношений с клиентами, который рассматривается сегодня как важнейший фактор, способствующий выживанию компании в условиях жесткой конкуренции, невозможен без эффективного использования всех технологий общения с реальным или потенциальным клиентом, Как следствие, конвергенция разнообразных инфраструктур, технологий и коммуникационных сред становится существенным средством повышения эффективности работы операторских центров и систем «самообслуживания» потребителей (self service customer services).

Одним из важных элементов конвергенции является отвод в Интернет части потока вызовов от систем и устройств, более дорогих в обслуживании , для быстрейшего сбора нужной информации и обеспечения высокого качества обслуживания потребителей, Происходит переход от операций, ориентированных, преимущественно, на речевую связь, к смешанным, включающим в себя, как важнейшие составляющие, электронную почту и Интегрированные операторские центры эффективно применяются и в корпоративных мультисервисных сетях, которые сегодня строятся на принципах пакетной коммутации. Это могут быть Интрасети или менее привычные Экстрасети (специализированные виртуальные частные сети VPN, аналогичные внутрикорпоративным сетям на базе IР, но разрешающие доступ к ним не одной компании, а группы таких компаний и их подразделений). Являясь расширением внутри корпоративной Интрасети, Экстрасеть позволяет деловым партнёрам обмениваться конфиденциальной информацией и совершать финансовые операции. Интегрированный центр поможет описанным выше корпоративным сетям наращивать возможности поддержки своих клиентов, в частности, благодаря эффективной организации на его базе справочных столов. Еще несколько лет назад ведущие мировые производители коммутационного оборудования рассматривали Саll-центре, только как дополнение к выпускаемым ими системам и возлагали на эти центры в основном, функции распределения вызовов. Сейчас эти производители (и не только они) осознали огромный потенциал решений на базе Саll-центров и Начали предлагать на телекоммуникационном рынке интегрированные продукты контакт центры, способные взаимодействовать с любыми телекоммуникационными средами.

Второй ключевой фактор, оказывающий влияние на развитие контакт центров - это развитие компьютерных технологий и рост вычислительных мощностей, фактор, который заставляет переоценить взгляды на архитектуру телекоммуникационных систем, выводит контакт центры из разряда узкоспециального оборудования, усиливает конкуренцию на рынке и позволяет разработчикам быстро создавать и предлагать покупателям новые системы. К тому же, стремительно развивающиеся IР технологии позволяют компаниям экономично использовать решения на основе Web-контакт центров и использовать (там. где это дает эффект) сеть Интернет, причем не только для обмена данными, совместного использования файлов и показа рекламы, но и для услуг, раньше считавшихся прерогативой телефонных сетей, Конвергенция систем переноса речи и данных по сетям открывает новые возможности предоставления дополнительных услуг, обеспечивает глобальный доступ к услугам web-контакт центра из любой точки сети и снижает производственные расходы за счет применения эффективных технологий пакетной коммутации , уменьшения полосы , используемой для передачи речевого трафика, и оптимизации программно-аппаратной архитектуры систем.

В контактцентрах интеграция услуг передачи речи/данных осуществляется на качественно новом уровне: архитектура таких систем обеспечивает возможность мультимедийного обмена сообщениями, т.е. приема, распределения и обработки по унифицированным алгоритмам сообщений и вызовов разного типа, приходящих из различных сетей. Интегрированный контакт центр нового поколения, упрощенная структура которого представлена на должен обеспечивать прием традиционных телефонных вызовов, телефонных вызовов. поступающих из сети Интернет с использованием технологий Voip, прием заявок, допускающих отложенную обработку, а также обработку в будущем запросов мультимедийной широкополосной связи. Должны также обеспечиваться поддержка режима чата (text chat) в реальном времени, доступ к динамической системе помощи, функции просмотра ЖО.страниц с сопровождением, возможности коллективного внесения отметок в страницы и совместного заполнения бланков заказов.

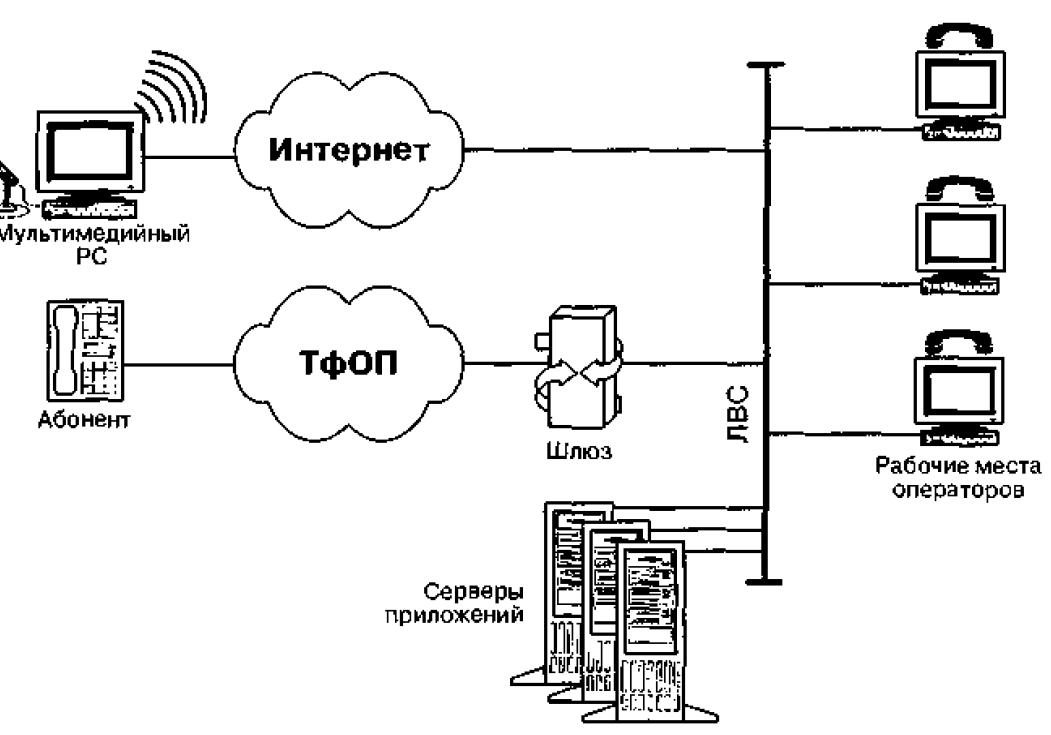


Рисунок 1 - Упрощенная структура контакт-центра

Функциональные возможности контакт центров, интегрированных с Интернет, позволяют компаниям, создающим такие центры, персонализировать отношения с посетителями обеспечивая доступ к агенту службы взаимодействия с потребителем а критический момент, когда у посетителя имеется важный вопрос, от ответа на который зависит, совершит он покупку или нет. Действительно, с точки зрения психологии очевидно, Что человек (клиент) почти всегда чувствует себя более комфортно, общаясь с другим человеком (оператором), чем просматривая безликие web-сайты с яркими картинками или прослушивая механический голос автоответчика. Иногда на сайте той или иной компании можно встретить раздел “наиболее часто задаваемые вопросы”, который фактически призван разрешать возникающие у клиента вопросы. Но нельзя не учитывать тот факт, что заранее предусмотреть жевания всех клиентов просто невозможно, другое дело, когда посетитель. не утруждая себя поиском нужной информации на web-сайте, просто нажимает нужную кнопку и лишних физических материальных затрат поучает всю необходимую информацию, причем  С Выгодной для компании стороны. Кроме того, информацию он может подавить в том виде, в каком ему удобно, или в том, который допускает имеющееся у него терминальное оборудование, Отсюда сразу становится очевидным второе преимущество. И контакт.центр должен сделать одинаково удобными для клиента такие разные способы обращения к персоналу центра, как телефонный вызов, вызов через Интернет, электронная почта или текстовый чат. При этом клиент должен быть уверен, что любой его вызов будет обработан с одинаковой тщательностью, и что если он, например, через два дня после отправки электронного письма или факса позвонит по поводу своего запроса в центр, ему не придется долго пересказывать оператору содержание этого запроса.

Абонент должен иметь возможность получить доступ к ресурсам центра, используя в качестве терминального оборудования либо телефонный аппарат, либо персональный компьютер. Сочетание традиционного операторского центра обслуживания вызовов с возможностями доступа из Интернет является мощнейшим катализатором для электронной коммерции и позволяет реализовать такие возможности взаимодействия с клиентом, которые выходят за рамки только использования информации, имеющейся на или только заочной телефонной беседы клиент-оператор». В настоящее время множество электронных коммерческих слепок расстраивается именно из-за того, что отсутствует общение клиента с живым сотрудником, который мог бы повлиять на решение клиента, В тоже время, только телефонного интерфейса часто уже недостаточно для того, чтобы клиент мог получить всю интересующую его информацию о товарах или услугах. Таким образом, интегрированные контакт-центры могут стать решающим фактором в превращении потенциальных покупателей в реальных. Важно, что благодаря возможностям технологий операторы контакт-центра, интегрированного с Интернет, получают возможность находиться территориально в любом месте. При этом создается виртуальный, работающий круглосуточно контакт-центр, не требующий использования дорогого сетевого ПО или оборудования дистанционных операторов. Облегчается поддержка многоязычности операторов контакт-центра, благодаря которой становится возможным обслуживать клиентов из разных уголков земного шара, Более того, учитывая наличие нескольких часовых поясов только на территории нашей страны, не говоря уже обо всем земном шаре, нетрудно обеспечить круглосуточное обслуживание клиентов.

Если  говорить о специфике использования глобальной сети Интернет в России, то надо отметить, что большинство пользователей предпочитают для работы в этой сети ночные часы, что обусловлено как наличием свободного времени именно ночью или поздним вечером, так и особенностями тарифов на услуги поставщиков. Интеграция операторских центров с Интернет должна позволить предприятиям сократить расходы на персонал благодаря переводу в Интернет (в режим, не требующий немедленного ответа) значительной части информационных запросов, сохранив при этом возможность того, чтобы клиент, при необходимости, в любой момент легко связаться с живым оператором. Обслуживая вызовы, поступающие из Интернет, предприятие предоставляет своим клиентам свободу просмотра любого оперативного материала (каталога продукции. технической информации). пока у ник не возникнут вопросы или комментарии, Введение этой возможности в может привлечь тех, например, кто предпочитает сообщать номера кредитных карт по телефону, а не пересылать их через компьютер, Все упомянутые выше преимущества. в совокупности с тенденциями, когда всё больше деловых операций осуществляется через Интернет, когда растет роль и место в жизни электронного бизнеса и электронной коммерции, а также с тем, что постоянное совершенствование технологии цифровой обработки речи и технологии ее передачи по сетям с пакетной коммутацией в последнее время существенно повысило качество речевой связи д пакетных сетях, ведут к тому, что в ближайшие годы прогнозируется настоящий бум в сфере операторских центров обслуживания вызовов, интегрированных с Интернет.

Крупные зарубежные производители, многие менее крупные (опять же в основном зарубежные) компании, а также некоторые отечественные разработчики (платформы ПРОТЕЙ, например) активно внедряются на рынок интегрированных операторских центров, Некоторые действующие Web-контакт центры уже сегодня предлагают возможность видеосвязи, если у клиента есть видеокамера и Н.З2З-совместимый Web-бpayзep. Правда, ее качество оставляет желать лучшего, но это поправимо — более важен прецедент. А операторские компании уже предоставляют услуги интернетовского контакт центра на условиях аутсорсинга фирмам, которые заинтересованы в таких услугах, Хочется верить, что в недалеком будущем подобным бизнесом займутся и отечественные телекоммуникационные операторы.

В настоящее время для интегрированного контакт-центра имеются варианты решений позволяющие предприятиям начать построение операторского центра с ёмкости всего в несколько портов и единиц рабочих мест операторов. Это дает небольшим контакт центрам возможность быстрее окупить капиталовложения. параллельно реализуя, в рамках общей стратегии продажи, высокопродуктивную и экономически выгодную программу электронной торговли, Тем компаниям, которые не уверены в том, насколько использование Интернет (подходит для их деятельности, такие системы дают средство проверить (причем с малым риском) возможности контакт центра, интегрированного с Web и разного рода дополнительные приложения, а также отладить процессы реорганизации бизнеса до заключения долговременных соглашений об использовании операторских центров или до приобретения оборудования. Разработаны методики оценки экономической целесообразности использования интегрированных контакт центров коммерческими компаниями. Так, если компания уже ведет электронную торговлю через Web, то выполняется анализ регистрационных файлов Web-сервера, по результатам которого компания может быстро определить, достаточны ли поступающий трафик и частота обращений к Web для того, чтобы оправдать использование контакт центра. Для компаний, которые, наоборот, сравнительно недавно занимаются электронной торговлей через Web и не знают, какие выгоды они могут получить от контакт центра интегрированного с Интернет, может выполняться, наряду с их обслуживанием на условиях аутсорсинга, экономический анализ, показывающий, как идут дела в течение года. Если наблюдается устойчивый рост числа операций, совершаемых через контакт центр, а доходы от обслуживания приближаются к закупочной цене системы, то приобретение системы в свое пользование становится экономически оправданным.

1.1.3 Функциональные особенности контакт-центра

“Call-Центр”, а также его менее часто употребляемые синонимы — «инфо-центр» или «операторский центр», Все эти термины относятся к структуре, оперирующей, в основном, с телефонными вызовами, Сегодняшний такого рода обрабатывает запросы клиентов и оптимизирует обработку этих запросов с учетом того, что телефонный разговор является отнюдь не единственным и даже, возможно, не основным средством доступа к информации. Отсюда — изменение названия, На смену терминам инфоцентр».  «операторский центр» или приходит новый термин —«контакт центр“ (или даже info-контакт центр»), более точно отражающий эволюцию понятий и функциональных возможностей подобных систем. Рассмотрим особенности новых, интегрированных с Интернет контакт центров, разбив эти особенности на функциональные группы, аналогичные по назначению группам функциональных возможностей традиционных Call-центров, а также отметим некоторые  специфические характеристики интегрированных контакт центров. Прежде всего это мультимедийность, понимаемая как способность обслуживать запросы разных типов, поступающие из разных телекоммуникационных сетей:

запросы речевой связи — из ТфОП;

запросы речевой связи — из Интернет, с использованием технологии  IP-телефонии;

запросы связи по факсу. электронной почте;

* запросы связи в режиме текстового чата — из Интернет;
* видеовызовы (в недалекой перспективе). Мультимедийность начинается с доступа к услугам контакт центров.

1.1.4 Доступ к услугам контакт центра

Рассмотрим некоторые из таких способов доступа к услугам, которые характерны именно для контакт центра, интегрированного с Web, Как правило, если мы говорим об интегрированном контакт центре, то предполагаем, что клиент получает доступ к ресурсам контакт центра со стороны корпоративного web-сайта и/или связанных с этим web-сайтом операторов, Таким образом. задачей контакт центра является обеспечение универсальности доступа с точки зрения абонента, свободы выбора метода доступа к услугам контакт центра. Полная прозрачность механизмов доступа приводит к тому, что абоненту достаточно просто нажать на соответствующую иконку на первой странице сайта компании, выбрав тем самым способ связи с центром.

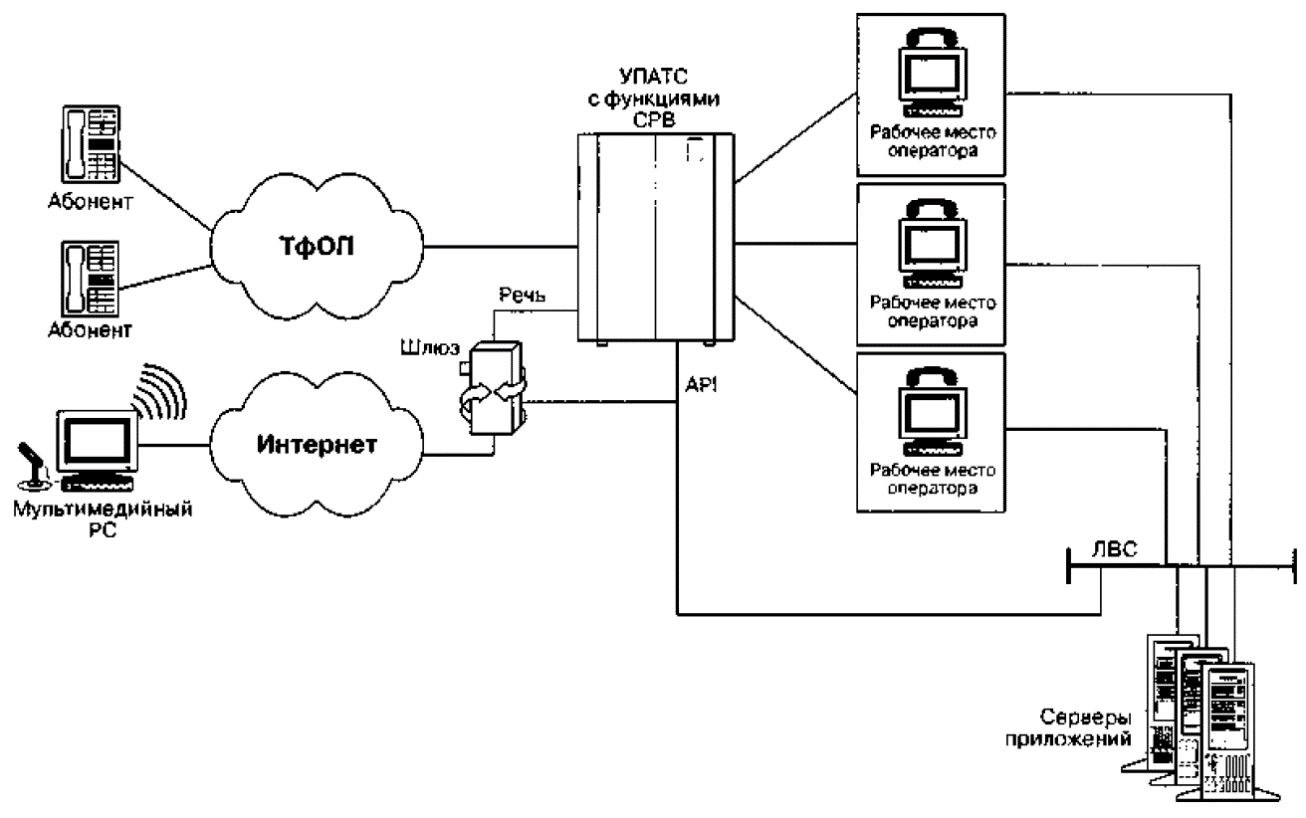
Существует целый ряд способов реализации такого универсального доступа. Одним из самых перспективных и экономически целесообразных способов (хотя несколько уступающим пока по качеству традиционной телефонной связи и требующим наличия у абонента хорошего Интернет Доступа) является доступ на базе технологии IР-телефонии Речевой диалог с клиентом проводится в виде сеанса VolP с использованием уже имеющегося соединения корпоративного web-сайта. При этом клиент и оператор контакт центра могут вести диалог и даже синхронно просматривать одни и те же web-страницы [6].

Рисунок 2 – Варианты доступа к контакт-центру

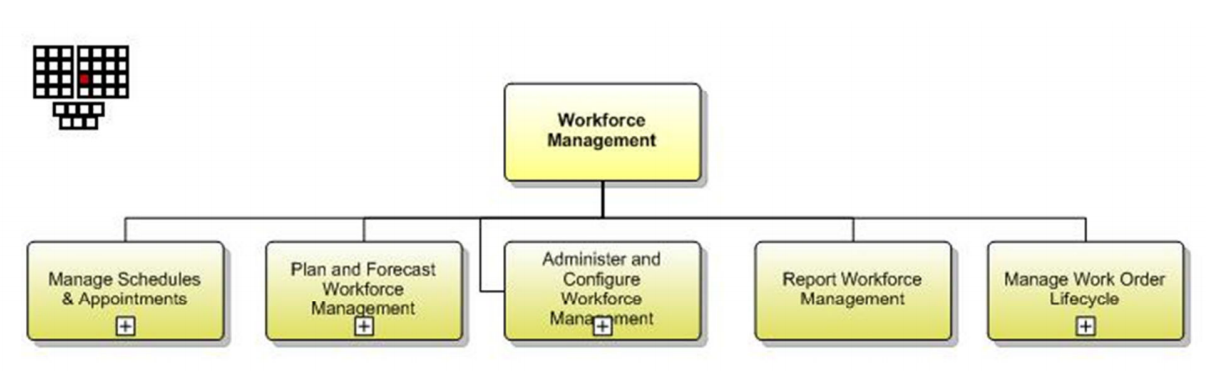
1.2 Оптимизация расходов контакт-центра

Одна из основных задач владельца контакт-центра – это увеличение качества предоставляемых услуг и уменьшение расходов. Существует весомая часть расходов на оплату работы сотрудников. В предыдущем разделе затрагивалась тема оптимизации расходов в связи с переходом на пакетную коммутацию и уменьшение клиентских заявок в связи с возможностью клиента самостоятельно ознакомится с требуемой ему информацией. Однако можно пойти дальше. В тарифный заработок работника включается заработная плата [5], начисленная за отработанное время, а также чтобы охарактеризовать затраты предприятия на рабочую силу обычно рассчитываются:

* величина затрат в среднем за месяц в расчете на одного работника;
* затраты на рабочую силу в расчете на один отработанный час;
* затраты на рабочую силу на один оплаченный час;

Следовательно, сократив затраты на рабочее время сотрудников, можно в целом сократить затраты организации на рабочую силу.

Сделать это можно с помощью гибкого менеджмента графиком работников.  Например, выводить определенное количество работников именно в тот промежуток времени, когда это требует нагрузка на контакт-центр и, соответственно, выводить меньше или больше, если нагрузка уменьшается или увеличивается. Для этого необходимо заранее составлять расписание операторов, а для составления расписания нужно знать, какая и в какое время нагрузка поступит в контакт-центр.  На карте eTOM (Enhanced Telecom Operations Map — многоуровневая модель бизнес-процессов управления производством) процессы, описывающие эту оптимизацию относятся к  процессам Workforce Management -  в них нас интересуют Manage Schedules & Appointments и Plan and Forecast Workforce Management (рис 3).

  
 рис. 1 Декомпозиция Workforce management на процессы третьего уровня)

1.2.1 Workforce Management

В обязанности процессов Workforce Management входит управление жизненным циклом рабочих заданий или рабочих заданий, выполняемых рабочей силой. Это включает управление персоналом (прямо или косвенно), нанятым или работающим на предприятии (т. е. Техническими специалистами, клерками, менеджерами и т. д.), Которые назначены, и выполнять указанную работу. Персонал, непосредственно управляемый Эти процессы включают всех сотрудников, подрядчиков и тех, кто оплачивает предприятие. Персонал, косвенно управляемый этими процессами, включает всех сотрудников, консультантов и подрядчиков, оплачиваемых третьими лицами, имеющими коммерческие соглашения с предприятием. В тех случаях, когда третьи стороны владеют и управляют инфраструктурой обслуживания и / или ресурсов, за которую отвечают процессы управления рабочей силой запрос на выполнение действий, а не непосредственное назначение конкретного персонала. Процессы управления персоналом также позволяют создавать отчеты и отслеживать назначенные действия. Обязанности этих процессов включают, но не ограничиваются:

• Создание и управление очередями рабочих заданий, через которые поступают запросы на рабочие действия от процессов eTOM;

• Создание списков персонала и управление ими, содержащих подробную информацию о назначаемых сотрудниках, таких как местонахождение, навыки, доступность для назначения и т. Д .;

• установление, управление и распределение индивидуальных рабочих заданий среди сотрудников с указанием ежедневных или других временных масштабов рабочих заданий;

• Создание и управление возможностями ускоренного и повторного назначения для возможности изменения рабочих заданий в соответствии с требованиями, касающимися опасности или других высокоприоритетных условий;

* Управление расписаниями встреч, включая просмотр и планирование встреч;
* Управление рабочими заказами, включая выдачу и закрытие;
* прогнозирование назначаемых кадровых потребностей на ежедневной, еженедельной и более длительной основе на основе исторических данных об объеме работы и прогнозируемых объемов деятельности;
* Определение оценок времени трудовой деятельности для всех известных видов трудовой деятельности на основе фактических исторических результатов или предварительных оценок, которые будут использоваться в качестве параметра для планирования рабочих списков;
* Создание и управление возможностями отзыва, позволяющими в нерабочее время отозвать сотрудников в случае непредвиденных обстоятельств;
* Управление процессами регистрации и контроля доступа, которые позволяют процессам создавать, изменять, обновлять, удалять и / или загружать данные о планировании и назначении работ в систему (ы) управления персоналом и из нее;
* Установление и управление передачей информации между системой (-ами) управления персоналом предприятия и системами внешних третьих сторон (когда инфраструктура принадлежит и управляется третьими сторонами);
* Обеспечение того, чтобы система (ы) управления персоналом точно собирала и записывала все детали назначения и планирования работ, используя автоматические или ручные проверки;
* Отслеживание и мониторинг использования и доступа к системе (системам) управления персоналом и связанных с ними затрат на процессы управления персоналом, а также отчетность по результатам; и Выявление любых технических недостатков системы (ов) управления трудовыми ресурсами и внесение вклада в процессы разработки и управления ресурсами для исправления этих проблем.

1.2.2 Plan and Forecast Workforce Management

Процессы планирования и прогнозирования рабочей силы отвечают за планирование и прогнозирование рабочей нагрузки, выполняемой вручную, которая должна быть назначена техническим специалистам на предприятии и / или третьим сторонам. Обязанности этих процессов включают, но не ограничиваются:

* Прогнозирование будущего спроса на рабочую силу на основе статистики рабочих заданий.
* Прогнозирование рабочей нагрузки и соответствующих назначаемых кадровых потребностей на ежедневной, еженедельной и более длительной основе на основе истории рабочей нагрузки, с учетом особых событий (внеочередных мероприятий), соображений об отпуске персонала и распределения имеющейся рабочей силы для наилучшего удовлетворения предполагаемой нагрузки.
* При необходимости вносить коррективы в расчетную нагрузку:
* Корректировка оценок продолжительности для всех известных видов работ, на основе фактических исторических результатов или предварительных оценок.

1.2.3 Manage Schedules & Appointments

В обязанности процессов управления расписанием встреч (Manage Appointment Schedule) входит возможность просмотра наличия соответствующего назначаемого персонала, а также возможность назначить встречу. И просмотр, и планирование могут быть основаны на наличии набора сотрудников, а не конкретного человека. Точно так же времена могут быть окнами времени (то есть - AM или PM) или определенным временем [7].

1.3. Возможности ИТ-решения для задачи прогнозирования.

Процесс Plan and Forecast Workforce Management для контакт-центра можно реализовать в виде ИТ-решения, которое предприятие будет разрабатывать под свои нужды, продавать другим организациям или все вместе. Оно может быть как самостоятельным решением, так и входить в состав приложения уровня выше, выполненное как составная часть, например, Workforce Management. Пример такого решения «WFM СС» от НТЦ Аргус[8], о котором будет отдельная часть в следующей главе.

Среди обязанностей процесса Plan and Forecast Workforce Management содержится обязанность по прогнозированию рабочей нагрузки. К теме данной работы подходит нагрузка звонков/обращений поступающих на контакт-центр. Интересна возможность предсказывать такую нагрузку для составления графиков работников. Такое решение требует осуществления статистического алгоритма. Современным решением в этом случае будет использование методов машинного обучения или его подкатегорию - глубокого обучения. Зависимость одного принципа от другого концептуально изображена на рисунке 4.

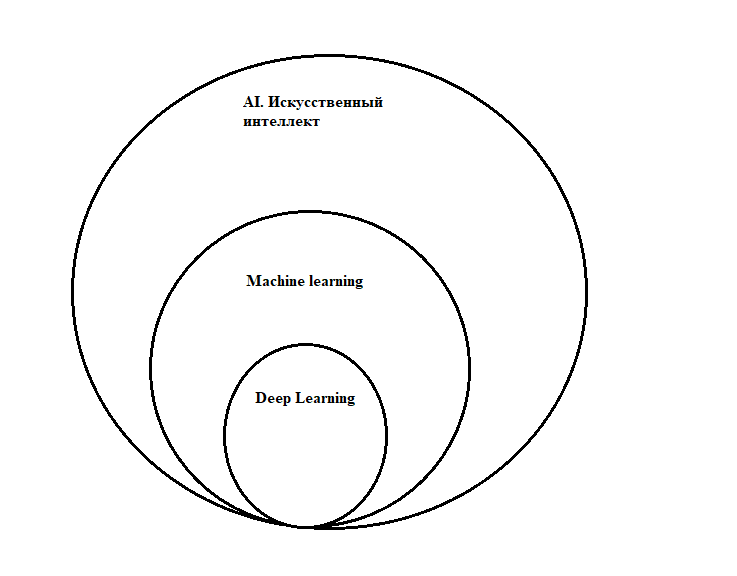


Рисунок 4 - зависимость между искусственным интеллектом, машинным и глубоким обучениями.

По сути, машинное обучение представляет из себя использование алгоритмов для извлечения информации из необработанных данных и представления ее в некоторой модели. Мы используем эту модель, чтобы сделать выводы о других данных, которые мы еще не смоделировали.

Глубокое обучение - нейронные сети с большим количеством параметров и уровней в одной из четырех основных сетевых архитектур:

* Предварительно обученные “без учителя” сети
* Сверточные нейронные сети
* Рекуррентные нейронные сети
* Рекурсивные нейронные сети

Нейронные сети являются одним из типов моделей для машинного обучения. Фундаментальной единицей нейронной сети является узел, который, в общих чертах, основан на биологическом нейроне в мозге млекопитающих. Связи между нейронами также смоделированы на биологическом мозге, так же, как эти связи развиваются с течением времени (с «обучением») [9].

Задача прогнозирования нагрузки на контакт-центр довольно интересна и уже реализуется с использованием нейронных сетей[10]. Однако это решение, как и любое другое, для прогнозирования требует наличие некоего объема исторических данных, которых может и не быть, при условии нового контакт-центра со своей спецификой клиентской нагрузки. В рамках работы предлагается построить, параллельно с прогнозом нагрузки, алгоритм, при котором происходит имитация работы контакт-центра изначально с небольшим или отсутствием исторических данных нагрузки, но имеющим в наличии модель прогнозирования, с постепенным обучением этой модели, для выяснения в какой момент из годового периода модель обучена достаточно, чтобы вводить ее в эксплуатацию.

1.4 Вывод по главе

Первая глава посвящена становлению контакт-центров, ответам на вопросы о том, какие предпосылки привели к современной практике реализации контакт-центра. Рассмотрены базовые определения характеризующие контакт-центр, его архитектурную составляющую, взаимосвязь с web-контакт центрами и call-центрами. Затронута тема функциональных особенностей контакт-центра и возможности доступа к услугам контакт-центра. В начале главы показан пример того, как со временем происходила оптимизация работы контакт-центра, а в следующей части рассказано о том, что существуют принципы Workforce Management на карте eTOM, указывающие на то, что контакт-центр можно оптимизировать еще больше, используя гибкое управление графиком работников. В конце главы был сделан вывод о том, что существует интересная задача по прогнозированию нагрузки на контакт-центр для возможности гибкого оптимизирования графиков сотрудников, тем самым проявляя возможность экономить на затратах на рабочую силу. Однако решение такой задачи не содержит достаточного оценочного функционала, который помог бы владельцам свежепостроенных контакт-центров определить, когда им вводить модель в работу.

Глава 2. МЕТОДЫ ПРОГНОЗИРОВАНИЯ ПРИМЕНИТЕЛЬНО К КОНТАКТ-ЦЕНТРАМ

2.2. Архитектура и функциональные возможности Аргус решения

Система АРГУС WFM CC (Workforce Management for Contact Center) предназначена для автоматизации управления контакт-центрами (КЦ), в частности бизнес-процессов прогнозирования нагрузки и планирования расписания работы операторов. Система ориентирована на операторов, супервайзеров (старших операторов) и высшее руководство контакт-центра.

Архитектура системы Система разработана под типовую структуру контакт-центра (рис. 5).

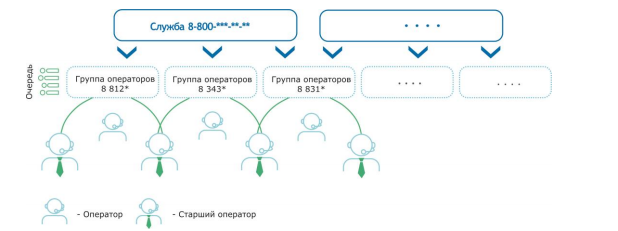


Рисунок 5 - Типовая структура контакт-центра

 В системе АРГУС WFM CC реализованы следующие основные роли (рис. 6):

Старший оператор (супервайзер):

 Оператор

 Руководитель направления

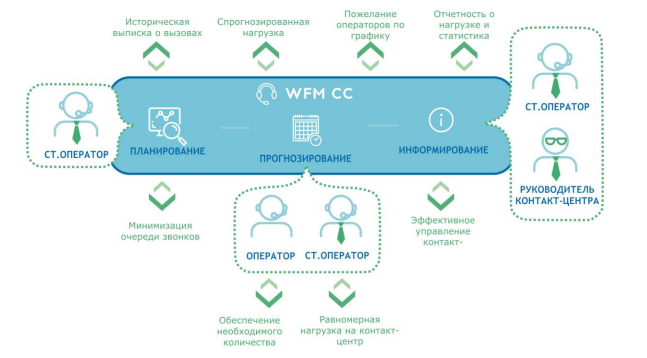


 Рисунок 6 - Роли в системе

2.2.1 Функциональные возможности системы

Богатая функциональность системы позволяет осуществлять все необходимые действия для достижения оптимального использования ресурсов КЦ:

* Управление персоналом – формирование структуры контакт-центра и управление доступным составом сотрудников;
* Прогнозирование нагрузки КЦ на основе исторических данных ЦОВ;
* Расчет необходимого количества операторов на основании прогнозируемой нагрузки с учетом регламентированного уровня обслуживания КЦ;
* Составление расписания работы операторов КЦ с учётом: прогнозируемой нагрузки; графиков смен, отпусков, больничных, отгулов; норм трудового законодательства, регламентов перерывов; запланированных регулярных мероприятий – совещаний, обучения и т.п;
* Оптимизацию расписания с учетом выбранных пользователем приоритетов – сокращения стоимости, повышения качества обслуживания;
* Анализ расписания в разрезе ключевых показателей – уровень сервиса, стоимость, загруженность операторов, выработка в часах, доля задействованных операторов;
* Оперативное обновление расписания с учетом непредвиденных обстоятельств - незапланированное отсутствие сотрудника, всплеск нагрузки на КЦ;
* Учёт пожеланий оператора – регистрация пожеланий по времени выхода на работу, обмен сменами, просмотр актуального расписания работы;
* Управление выработкой оператора – анализ статистики по выработке оператора.

Система АРГУС WFM CC формирует прогноз нагрузки на основе исторических данных ЦОВ. Загрузка исторических данных для анализа осуществляется в специальном окне (рис. 7).

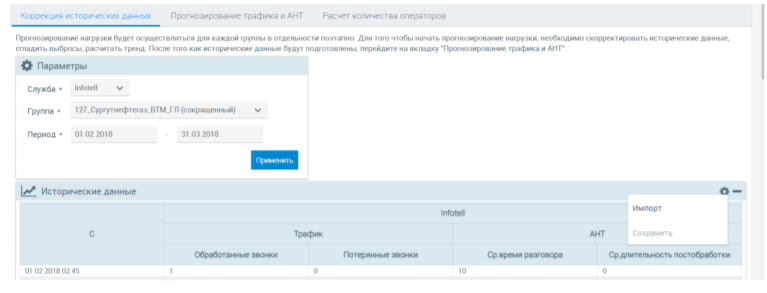


Рисунок 7 - Окно загрузки исторических данных для дальнейшего анализа

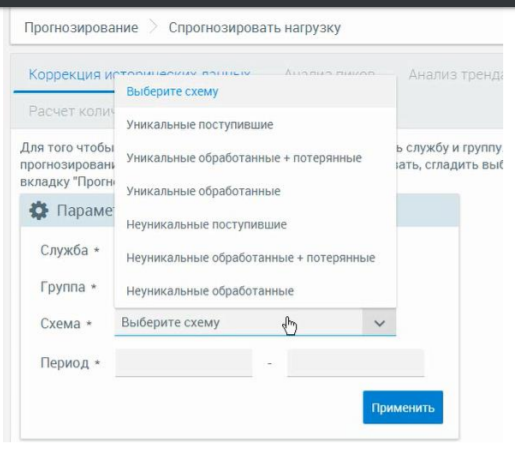
 Для повышения точности прогноза в системе предусмотрен ряд инструментов. Коррекция исторических данных Коррекция исторических данных может быть полезна, если при прогнозировании требуется учитывать или наоборот исключить некоторые категории вызовов. Настройку параметров прогнозирования можно сделать по каждой службе и группе операторов (рис. 8) АРГУС WFM СС.

Рисунок 8 - Коррекция исторических данных

2.2.2 Анализ пиков

 Этот инструмент необходим для исключения статистических выбросов данных из-за каких-то редких (разовых) событий. Такое возможно, если на определённом интервале происходят нетипичные события, приводящие к всплескам или провалам нагрузки: аварии, пожары, ввод новой услуги и т.п. Для повышения точности прогноза следует провести визуальную оценку нагрузки на периоде и при наличии выбросов исключить из расчёта тот или иной фрагмент данных (рис. 9).

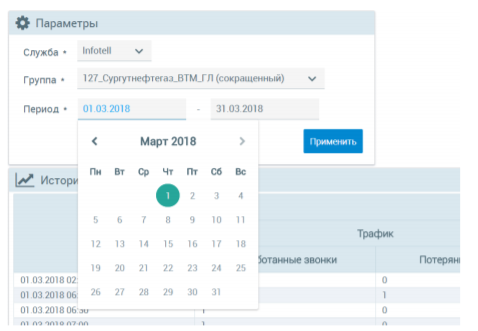


Рисунок 8 - Выбор исключаемого периода для повышения точности прогноза

2.2.3 Учет трендовой и сезонных составляющих при формировании прогноза.

Тренд – это усредненная динамика интенсивности нагрузки в рамках рассматриваемого периода. В качестве сезонных составляющих в системе учитываются следующие периоды: месяц, неделя в году, неделя в месяце, день в месяце, день недели, интервал внутри дня. Система позволяет учитывать или наоборот исключать при формировании прогноза тренды или сезонные составляющие, т.к. они могут снижать точность прогноза (рис. 9).

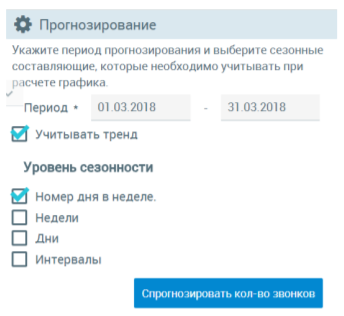


Рисунок 9 - Учёт тренда и сезонных составляющих

 Агрегированные прогнозы нагрузки

Учёт различных горизонтов планирования (рис. 10) необходим, когда требуется рассчитать количество персонала на год вперёд, либо осуществить набор дополнительных сотрудников на сезонный период, например, летние месяцы или предновогоднее время, когда нагрузка на контакт-центр может возрастать или ощутимо снижаться.



Рисунок 10 - Горизонты планирования в контакт-центре

 В системе предусмотрена возможность просмотра прогноза в режиме «внутригодовой профиль месячных периодов». В этом режиме можно проанализировать потребность в трудовых ресурсах в разрезе каждого месяца на основании прогнозируемого трафика. Например, в декабре нужно иметь 120 сотрудников, в январе – 90, в феврале – 70, и т.д. Располагая информацией, сколько сотрудников числится в штате на текущий момент, можно определить проблемные периоды и рассчитать, сколько дополнительных сотрудников потребуется и на какой период. Тактическое планирование Пользователю предоставляется возможность просматривать прогнозы в режиме «внутримесячный профиль дневных периодов» на интересующий его период. В этом режиме можно увидеть потребность в ресурсах на каждый день в рамках заданного периода. Пользователь видит, сколько сотрудников может быть задействовано в каждый из дней в рамках заданного периода с учетом графиков сменности, отпусков и других факторов. Пользователь также может увидеть, в какие дни сотрудников будет остро не хватать, и сможет ли компания покрыть нехватку штатными работниками, либо придётся нанимать временный персонал. Анализ точности прогнозов В системе предусмотрена возможность визуальной сравнительной оценки точности прогноза за выбранный период. Система даёт численное значение точности прогнозирования в процентах (рис. 11).

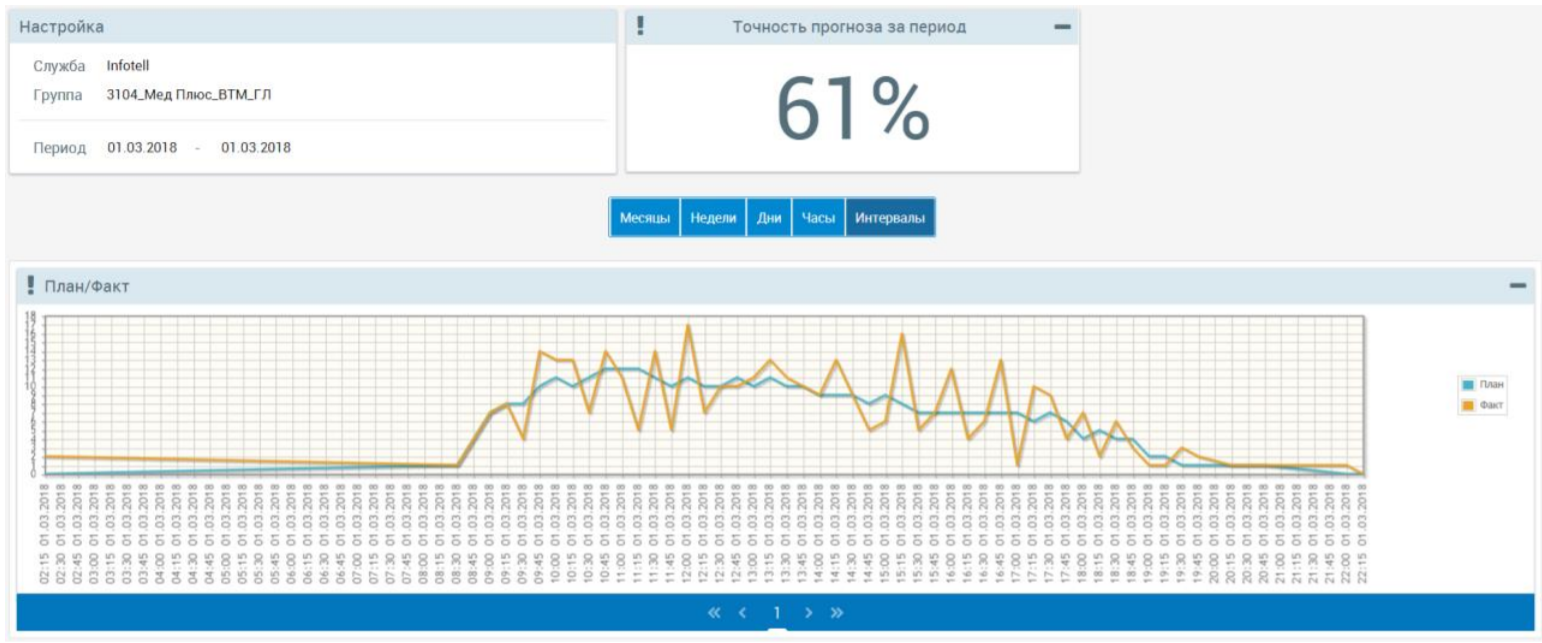


Рисунок 11 - Инструментарий оценки точности прогнозирования нагрузки.

В ряде случаев чисто автоматизированный анализ данных может давать слишком грубый или даже ошибочный прогноз. Это связано с неравномерностью нагрузки, её скачкообразностью. Периодически встречаются выбросы, приводящие к значимым отклонениям от «естественных» значений. В таких случаях наиболее целесообразно провести визуальную оценку и скорректировать данные. Система предоставляет пользователю обширный набор средств визуализации исторических данных по нагрузке и результатов прогнозирования. Возможна визуализация исторических данных и AHT на определенный период на уровне 15- минутного интервала, часа, дня, недели, месяца как в табличном виде, так и в виде графиков (рис. 12).



Рисунок 12 - Отображение данных о звонках за заданный период

2.2.4 Расчёт числа операторов

Для гарантии поддержания требуемого качества сервиса контакт-центры нередко выводят на линию большее количество операторов, чем требуется согласно прогнозу нагрузки. С учетом того, что количество операторов рассчитывается для каждой очереди звонков, это неминуемо приводит к росту ФОТ. В системе АРГУС WFM CC используется новая методика расчёта числа операторов, которая позволяет учитывать не только прогноз количества обращений в КЦ, но и загруженность оператора. В результате система рассчитывает минимально необходимое и при этом достаточное для поддержания заданного уровня обслуживания количество операторов, а значит трудовые ресурсы распределяются наиболее оптимально. Используемый в системе АРГУС WFM CC подход к расчёту числа операторов позволяет максимально эффективно распределять их трудовые ресурсы и в конечном итоге экономить на ФОТ. Управление параметрами расчёта количества операторов В системе реализован набор необходимых настроек для автоматического расчёта требуемого количества операторов с учётом ряда параметров (рис. 13): AWT (Average Waiting Time) - максимальное время, которое клиент будет ожидать ответа оператора (в секундах); OCC (Occupancy) - определяет максимальную занятость каждого оператора во время смены (в процентах); SL (Service Level), min/max – устанавливает пределы уровня качества обслуживания (в процентах); АНТ (Average Handling Time) - среднее время обработки звонка (в секундах).



Рисунок 13 - Параметры настройки расчёта количества операторов

С учетом полученного прогноза и установленных параметров система рассчитывает количество операторов, необходимое для работы в каждый 15- минутный интервал времени. Благодаря такому подходу обеспечивается требуемое качество сервиса [11].

2.3. Методы прогнозирования временных рядов. Сравнительный анализ.

Для начала следует определить, что такое временной ряд. Временной ряд — это последовательность упорядоченных во времени числовых показателей, характеризующих уровень состояния и изменения изучаемого явления. Всякий временной ряд включает два обязательных элемента: во-первых, время и, во-вторых, конкретное значение показателя, или уровень ряда.[12] Исходя из определения можно рассмотреть нагрузку запросов поступающую на контакт-центр как временной ряд.

Стационарность ряда обуславливается отсутствием таких характеристик временного ряда как тренд, сезонность и цикличность.

Цикличность существует, когда данные демонстрируют взлеты и падения, которые не имеют фиксированного периода. Продолжительность этих колебаний обычно составляет не менее 2 лет [13].

Тренд - содержится во временном ряду, значения которого увеличиваются или уменьшаются со временем, не обязательно линейно.

Сезонные колебания или сезонность - это циклы, которые регулярно повторяются с течением времени. Есть много видов сезонности. Некоторые очевидные примеры включают в себя; время суток, ежедневно, еженедельно, ежемесячно, ежегодно и т. д. Таким образом, определение наличия сезонности в проблеме временных рядов является субъективным [2].

Существует множество методов прогнозирования временных рядов. Глобально можно разделить методы на две категории. Классические статистические методы и современные методы машинного обучения.

2.3.1 Классические модели.

* Модель авторегрессии (AUTOREG)

Процедура AUTOREG [14] оценивает и прогнозирует модели линейной регрессии для данных временных рядов, когда ошибки автокоррелированы. Авторегрессионная модель регрессирует значение ряда в момент времени t (Yt) на значения в течение времени периоды (t - 1, t - 2, …, t - p. ) Математическая формула выражается следующим образом:

|  |  |
| --- | --- |
| , | (1) |

где , , , … , - коэффициенты линейной регрессии, - прогнозируемое значение в момент времени t, а - случайная ошибка. переменная и обычно предполагается, что имеет нормальное распределение со средним 0 и дисперсией .

* Модели авторегрессионной скользящей средней (ARMA)

Модель ARMA является одним из основных инструментов во временных рядах

 моделирование. Предположим, что временной ряд Y1, Y2,. , , Yt - стационарный случайный процесс, выражение ARMA (p, q) представляет модель с авторегрессионным порядком p и скользящая средняя порядка q. Эта модель представляет собой комбинацию моделей AR (p) и MA (q), где AR (p) записывается как и MA (q) записывается в виде .

Как и в модели AUTOREG, Yt является значением наблюдения в момент времени t. Процесс ARMA (p, q) обычно записывается как следующим образом:

|  |  |
| --- | --- |
|  | (2) |

где a, b и c - постоянные, - случайная ошибка переменная и обычно считается, что имеет нормальное распределение со средним 0 и дисперсией ; , , … , Arep являются авторегрессионные коэффициенты, подлежащие оценке, и , , … , являются скользящими средними коэффициентами, которые необходимо оценить.

* Интегрированная модель авторегрессии — скользящего среднего (ARIMA).

Процедура ARIMA функционирует как линейная комбинация ее текущих значений, прошлых значений, прошлых ошибок и прошлых значений других временных рядов (временных рядов предикторов) для прогнозирования будущего значения ответа во временном ряду.

При нестационарном поведении временных рядов вышеуказанная модель ARMA (p, q) может быть расширена и записана с использованием разности, которая определяется как: , где t - индекс времени , - временной ряд {: 1 ≤ t ≤ n} в момент времени t, а B - оператор обратного смещения, что означает, что B имеет эффект смещения данных на один период назад (т.е. B = ).

* Сезонная модель ARIMA.

Сезонная модель ARIMA записывается с общим выражением ARIMA (p, d, q) (P, D, Q) s. Символ p является порядком несезонной авторегрессионной составляющей, d является порядком разностного вычисления, q является порядком несезонного процесса скользящей средней, P является порядком сезонной авторегрессионной части, D является порядком сезонной разности , Q - порядок сезонного скользящего среднего, а s - продолжительность сезонного цикла. Пусть будет зависимым временным рядом {: 1 ≤ t ≤ n} в момент времени t, тогда математическая формула для сезонной модели ARIMA выражается следующим образом:

|  |  |
| --- | --- |
| , | (3) |

где μ - постоянное среднее значение, Bs - оператор сезонного обратного сдвига, – сезонная авторегрессионная составляющая, и - это составляющая сезонного скользящего среднего.

* Модель сезонного экспоненциального сглаживания.

В методе сезонного экспоненциального сглаживания (ESM) уравнение прогнозируемой величины в момент времени задается в журнале Healthcare Engineering

|  |  |
| --- | --- |
|  | (4) |

Уравнения сглаживания следующие:

|  |  |
| --- | --- |
|  | (5) |
|  | (6) |

где - наблюдение в момент времени t, а α и c - параметры уровня и сезонного сглаживания, соответственно, - оцененный компонент уровня в момент t, - предполагаемый сезонный компонент в момент t, а p - периоды, после которых сезонный цикл повторяется.

* Мультипликативная модель Холта-Винтерса.

Модель Хольта-Винтерса, также известная как тройное экспоненциальное сглаживание, применяется три типа экспоненциального сглаживания к временному ряду: значение, тренд и сезонность. Модельное уравнение для метода Хольта-Винтерса может быть либо аддитивной, либо мультипликативной моделью. Математическая формула, относящаяся к временному ряду с трендом и постоянной сезонной составляющей с использованием аддитивной техники Хольта-Винтерса, имеет прогноз в момент времени , определяемый следующим уравнением:

|  |  |
| --- | --- |
| , | (7) |

Уравнения сглаживания задаются с использованием следующих уравнений:

|  |  |
| --- | --- |
| , | (8) |
| , | (9) |
| , | (10) |

где - наблюдение в момент времени t, α, β и c -

уровень, тренд и сезонные соответствующие константы, соответственно, - расчетный уровень в момент времени t, - расчетный

тренд в момент времени t, - индекс сезонности в момент времени t, а p -

периоды, после которых сезонный цикл повторяется.

* Аддитивная модель Хольта-Винтерса.

Для аддитивной модели прогнозируемая оценка предложения для времени t + k дается следующим уравнением:

|  |  |
| --- | --- |
| . | (11) |

Оценки уровня, тренда и сезонных факторов для аддитивных модельных уравнений приведены с использованием следующих уравнений:

|  |  |
| --- | --- |
| , | (12) |
| , | (13) |
| . | (14) |

2.3.2 Модели машинного обучения

2.3.2.1 Основные понятия в нейронных сетях

Современные вычислительные системы, рассматриваемые в виде нейронных сетей, преобразуют информацию по подобию процессов в мозге человека. Информация представляется в численном виде, поэтому есть возможность использовать нейронную сеть, например, в качестве модели объекта с полностью неизвестными характеристиками. Некоторые типовые приложения нейронных сетей применяются для распознавания лиц на фотографиях и видеозаписях, а также анализа и преобразования образов.

Нервная система человека состоит из, как не странно догадаться, нервных клеток, называемых нейронами. Нейрон состоит из непосредственно тела, называемого сомой, отростков, по которым в нейрон поступает и передается информация – дендриты и аксоны. Через нервные стыки, называемыми синапсами передается возбуждение от одного нейрона другому. На Рисунке 14 представлена упрощенная модель нейрона [1].

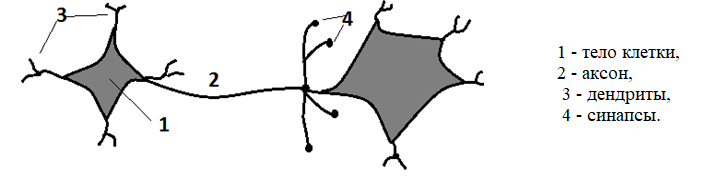


Рисунок 14 - Упрощенная модель нейрона и его соединения с соседним нейроном.

Рассмотрим модель нейрона, связанную с первыми попытками формализовать описание функционирования нервной клетки. Введем следующие обозначения:

   – входные сигналы данного нейрона, приходящие от других нейронов;

   - синаптические веса;

  y - выходной сигнал нейрона;

  v - пороговое значение.

Формула, описывающая функционирование нейрона, имеет вид;

|  |  |  |
| --- | --- | --- |
|  |  | (15) |

Модель (15) может быть представлена в виде;

|  |  |  |
| --- | --- | --- |
|  |  | (16) |

где

|  |  |  |
| --- | --- | --- |
|  |  | (17) |

а также,

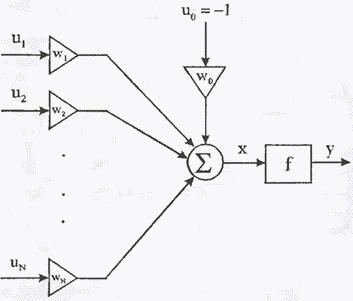
Формула (16) описывает модель нейрона, представленную на Рисунке 15

Рисунок 15 - Модель нейрона

Эта модель была предложена в 1943 г. МакКаллоком и Питтсом [1]. В качестве функции f может приниматься не только единичная функция (2.3), но и другие пороговые функции вида;

|  |  |  |
| --- | --- | --- |
|  |  | (18) |

или

|  |  |  |
| --- | --- | --- |
|  |  | (19) |

Нелинейную функцию f(x) называют активационной. Когда активационная функция одна и та же для всех нейронов сети, сеть называется однородной (гомогенной). При условии, что функция активации имеет зависимость от еще от одного или нескольких параметров с меняющимися значениями от нейрона к нейрону, сеть называется неоднородной (гетерогенной).

2.3.2.2 Обычные нейронные сети и Deep Learning модели

* Многослойный персептрон (MLP)

Классический вариант нейронной сети. MLP задается следующим образом:

|  |  |
| --- | --- |
|  | (20) |

где - входной вектор x, дополненный 1, т. е. , - весовой вектор для j-го скрытого слоя, , , … , - веса для выходных данных. узел, а - сетевой выход. Функция представляет собой выход скрытого слоя, и она задается в виде функции сжатия, например, логистической функции:

|  |  |
| --- | --- |
|  | (21) |

Связанной моделью в литературе по эконометрике является модель авторегрессии с плавным переходом, которая также основана на построении линейных функций и переходов логистических функций. MLP является сильно параметризованной моделью, и, выбирая количество скрытых нейронов, мы можем контролировать сложность модели. Прорыв, который предоставил возможности нейронных сетей, является свойством универсального приближения. При определенных мягких условиях на функции скрытых нейронов g любая заданная непрерывная функция на компактном множестве может быть аппроксимирована настолько близко, насколько это произвольно дано, используя сеть с конечным числом скрытых нейронов. Хотя это обнадеживающий результат, крайне важно избегать чрезмерной параметризации, особенно в приложениях прогнозирования, которые обычно имеют ограниченный объем данных с высоким уровнем шума. Поэтому выбор модели (путем выбора количества скрытых нейронов) вызвал большой интерес в литературе по нейронным сетям.

Для получения весовых коэффициентов определяется среднеквадратичная ошибка, а весовые коэффициенты оптимизируются с использованием методов градиента. Наиболее известным методом, основанным на концепции наискорейшего спуска, является алгоритм обратного распространения.

* Байесовская нейронная сеть (BNN)

Байесовская нейронная сеть (BNN) - это нейронная сеть, разработанная на основе байесовской вероятностной формулировки. Как таковые, BNN связаны с классической статистической концепцией оценки байесовских параметров, а также с концепцией регуляризации, такой как регрессия гребня. BNN широко применяются во многих областях, таких как экономика / финансы и инженерия. Идея BNN состоит в том, чтобы рассматривать сетевые параметры или веса как случайные величины, подчиняясь некоторому априорному распределению. Это распределение разработано таким образом, чтобы отдавать предпочтение моделям с низкой сложностью, то есть моделям, обеспечивающим плавные переходы. Как только данные наблюдаются, оценивается апостериорное распределение весов и может быть вычислено предсказание сети. Прогнозы затем будут отражать как аспект гладкости, наложенный через предыдущий, так и аспект точности точности, наложенный наблюдаемыми данными. Тесно связанной концепцией является аспект регуляризации, посредством которого следующая целевая функция строится и минимизируется

|  |  |
| --- | --- |
| *,* | (22) |

где - сумма квадратов ошибок в выходных данных сети, - сумма квадратов параметров сети (т. е. весов),   - параметр регуляризации. Для байесовского подхода типичным выбором априора является следующая нормальная плотность, которая придает больший вес меньшим значениям параметров сети

|  |  |
| --- | --- |
| *,* | (23) |

где L обозначает количество параметров (весов). Задняя часть тогда дается

|  |  |
| --- | --- |
|  | (24) |

где D представляет наблюдаемые данные. Предполагая нормально распределенные ошибки, плотность вероятности данных с учетом параметров можно оценить как

|  |  |
| --- | --- |
|  | (25) |

где M - это количество обучающих данных. Подставляя выражения для плотностей в (23) и (25) в (24), получим

|  |  |
| --- | --- |
|  | (26) |

где c - это некоторая нормализующая константа. Константа регуляризации также определяется с помощью байесовских понятий из

|  |  |
| --- | --- |
|  | (27) |

Оба выражения (26) и (27) должны быть максимизированы для получения оптимальных весов и параметра respectively соответственно. Член в (27) получается квадратичной аппроксимацией J по весам с последующим интегрированием весов.

* Нейронная сеть с функцией радиального базиса (RBF)

Сеть радиальных базисных функций аналогична архитектуре многослойной сети за исключением того, что нейроны имеют локализованную функцию активации. Чаще всего нейроны функции выбираются в качестве гауссовых функций, причем ширина гауссовой функции контролирует гладкость встроенной функции. Выходы нейронов объединяются линейно, чтобы получить конечный выход сети. В частности, выход дается

|  |  |
| --- | --- |
|  | (28) |

где , и обозначают соответственно вес объединения, ширину функции нейрона и центр функции нейрона для блока j. Из-за локализованного характера функций нейрона были разработаны другие более простые алгоритмы для обучения радиальных базисных сетей. Он основан на запуске с пустой сети и последовательном добавлении нейрона, пока не будет достигнута приемлемая ошибка в обучающем наборе.

* Обобщенная регрессионная нейронная сеть (GRNN)

Обычно его называют оценщиком Надарая-Ватсона или оценщиком регрессии ядра. В сообществе машинного обучения обычно используется термин обобщенная регрессионная нейронная сеть (или GRNN). Мы будем использовать этот последний термин. Модель GRNN является непараметрической моделью, в которой прогноз для заданной точки данных x задается как среднее из целевых выходных значений точек обучающих данных в окрестности заданной точки x. Локальное среднее строится путем взвешивания точек в соответствии с их расстоянием от x с использованием некоторой функции ядра. Оценка представляет собой просто взвешенную сумму наблюдаемых ответов (или целевых результатов), заданных

|  |  |
| --- | --- |
|  | (29) |

где веса определяются как

|  |  |
| --- | --- |
|  | (30) |

где my - целевой выход для обучающей точки данных xm, и это функция ядра. Типичное гауссово ядро. Параметр h, называемый шириной полосы, является важным параметром, так как он определяет гладкость изображения, так как увеличение или уменьшение его будет контролировать размер области сглаживания.

* Рекуррентная нейронная сеть (Recurrent neural network, RNN)

Рекуррентные нейронные сети [15] , или RNN, представляют собой семейство нейронных сетей для обработки последовательных данных. Поскольку сверточная сеть представляет собой нейронную сеть, которая специализируется для обработки сетки значений X, такой как изображение, рекуррентная нейронная сеть представляет собой нейронную сеть, которая специализируется для обработки последовательности значений x (1),. , , , х (τ). Подобно тому, как сверточные сети могут легко масштабироваться до изображений с большой шириной и высотой, а некоторые сверточные сети могут обрабатывать изображения переменного размера, так и рекуррентные сети могут масштабироваться до гораздо более длинных последовательностей, чем было бы целесообразно для сетей без специализации на основе последовательностей. Большинство повторяющихся сетей также могут обрабатывать последовательности переменной длины.

Чтобы перейти от многослойных сетей к рекуррентным сетям, нам необходимо воспользоваться одной из ранних идей, заложенных в модели машинного обучения и статистических моделях 1980-х годов: разделение параметров между различными частями модели. Совместное использование параметров позволяет расширить и применить модель к примерам различных форм (различные длины, здесь) и обобщать их. Если бы у нас были отдельные параметры для каждого значения временного индекса, мы не могли бы обобщать на длины последовательностей, не замеченные во время обучения, и не делить статистическую силу по разным длинам последовательностей и по разным позициям во времени. Такое совместное использование особенно важно, когда конкретная часть информации может встречаться в нескольких местах в последовательности. Например, рассмотрим два предложения: «Я отправился в Непал в 2009 году» и «В 2009 году я отправился в Непал». Если мы попросим модель машинного обучения прочитать каждое предложение и извлечь год, в котором рассказчик отправился в Непал, мы хотели бы, чтобы он признал 2009 год как соответствующую часть информации, независимо от того, указывается ли она в шестом слове или во втором Слово предложения. Предположим, что мы обучили сеть прямой связи, которая обрабатывает предложения фиксированной длины. Традиционная полностью подключенная сеть с прямой связью будет иметь отдельные параметры для каждой входной функции, поэтому ей нужно будет изучать все правила языка отдельно в каждой позиции предложения. Для сравнения, рекуррентная нейронная сеть имеет одни и те же веса на нескольких временных шагах.

Родственной идеей является использование свертки в одномерной временной последовательности. Этот сверточный подход является основой для нейронных сетей с задержкой во времени. Операция свертки позволяет сети обмениваться параметрами во времени, но является поверхностной. Вывод свертки представляет собой последовательность, в которой каждый элемент вывода является функцией небольшого числа соседних элементов ввода. Идея обменивания параметрами проявляется в применении одного и того же ядра свертки на каждом временном шаге. Рекуррентные сети используют параметры другим способом. Каждый элемент вывода является функцией предыдущих членов вывода. Каждый элемент выходных данных создается с использованием того же правила обновления, примененного к предыдущим выходным данным. Эта повторяющаяся формулировка приводит к совместному использованию параметров через очень глубокий вычислительный граф.

Рекуррентные нейронные сети могут быть построены разными способами. Поскольку почти любую функцию можно считать нейронной сетью с прямой связью, фактически любую функцию, включающую рецидив, можно считать рекуррентной нейронной сетью.

Некоторые примеры важных шаблонов проектирования для рекуррентных нейронных сетей включают в себя следующее:

* Рекуррентные сети, которые выдают результат на каждом временном шаге и имеют периодические соединения между скрытыми модулями.
* Рекуррентные сети, которые создают выход на каждом временном шаге и имеют рекуррентные соединения только от выхода на одном временном шаге к скрытым единицам на следующем временном шаге.
* Рекуррентные сети с повторяющимися соединениями между скрытыми модулями, которые читают всю последовательность, а затем выдают один вывод.
* Нейронная сеть долгой краткосрочной памяти (Long Short-Term Memory, LSTM)

Один из способов справиться с долговременными зависимостями - это разработать модель, которая работает в нескольких временных масштабах, чтобы некоторые части модели работали в точных временных масштабах и могли обрабатывать мелкие детали, тогда как другие части работают в грубых временных масштабах и передают информацию из далекого прошлого в настоящее более эффективно. Возможны различные стратегии построения как тонких, так и грубых временных масштабов. К ним относятся добавление пропущенных соединений во времени, «единиц утечки», которые объединяют сигналы с различными постоянными времени, и удаление некоторых соединений, используемых для моделирования точных временных шкал.

Наиболее эффективные модели последовательностей, используемые в практических приложениях, называются вентильными RNN. К ним относится долгая краткосрочная память и сети на основе вентильного рекуррентного блока. Вентильные RNN основаны на идее создания путей во времени, имеющих производные, которые не обнуляются и не стремятся резко вверх. Вентильные RNN обобщают это для весов соединений, которые могут изменяться на каждом временном шаге. Блоки с утечкой позволяют сети накапливать информацию (например, доказательства для определенной функции или категории) в течение длительного времени. Однако после использования этой информации нейронной сети может быть полезно забыть старое состояние. Например, если последовательность состоит из подпоследовательностей, и мы хотим, чтобы блок с утечкой накапливал свидетельства внутри каждой подпоследовательности, нам нужен механизм, чтобы забыть старое состояние, установив его в ноль. Вместо того, чтобы вручную решать, когда очищать состояние, мы хотим, чтобы нейронная сеть научилась решать, когда это делать. Это то, что делают вентильные RNN.

Методы машинного обучения, а особенно, нейронные сети, вызывают наибольший интерес у современных практиков в области изучения данных. Это сложные и интересные модели, однако их сложность не всегда может быть оправдана.

2.3.2.3 Регрессионные модели.

* Метод k-ближайших соседей (K Nearest Neighbor Regression)

Метод регрессии K ближайших соседей (KNN) является непараметрическим методом, который основывает свое прогнозирование на целевых выходах K ближайших соседей данной точки запроса. В частности, учитывая точку данных, мы вычисляем евклидово расстояние между этой точкой и всеми точками в тренировочном наборе. Затем мы выбираем ближайшие K обучающих точек данных и устанавливаем прогноз в качестве среднего значения целевых выходных значений для этих K точек. Количественно говоря, пусть  ℐ(x) - множество ближайших соседей точки x. Затем дается прогноз

|  |  |
| --- | --- |
|  | (31) |

где снова - целевой выход для точки данных обучения xm. Естественно, K является ключевым параметром в этом методе, и его следует выбирать с осторожностью. Большой K приведет к более плавной и, следовательно, более низкой дисперсии, конечно, за счет более высокого смещения, и наоборот, для небольшого

* Классификация и регрессия дерева решений (Classification and Regression Tree )

CART - это модель классификации или регрессии, основанная на иерархическом древовидном разделе входного пространства. В частности, входное пространство делится на локальные области, идентифицированные в последовательности рекурсивных разбиений. Дерево состоит из внутренних нейронов решений и конечных листьев. С учетом точки тестовых данных последовательность тестов вдоль нейронов решений, начиная с корневого нейрона, будет определять путь вдоль дерева до достижения конечного нейрона. В конечном нейроне делается прогноз в соответствии с локальной моделью, связанной с этим нейроном.

2.4 Выбор и обоснование модели нейронной сети.

В исследовании “Statistical and Machine Learning forecasting methods: Concerns and ways forward” рассказывается о том, что методы машинного обучения ошибочно принимаются как панацея в решении задач прогнозирования временных рядов. Авторы показывают, что классические модели решают задачи лучше и проще.

Авторы рассказывают нам о том, что для выяснения какой же метод в итоге лучший, множество исследований заканчиваются сравнением только лишь методов машинного обучения между собой. Результаты сравнения классических методов и методов машинного обучения доказывают это. Также существует исследование “Comparison of Time Series Methods and Machine Learning Algorithms for Forecasting Taiwan Blood Services Foundation’s Blood Supply” подтверждающее данное утверждение.   
  
Результаты обоих исследований показывают, что из представленных выше лучшие результаты дают методы  Хольт-Винтерс и ARIMA. Причем, такие популярные нейронные сети как LSTM  оказываются далеко позади. Авторы указывают на то, что этот факт удивителен, так как сеть LSTM хорошо подходит для таких задач, а также лучше классических моделей обучилась на тренировочной выборке. Добавляется, что это можно использовать в прогнозировании, совмещая классические методы и методы машинного обучения. И действительно, зачем тогда использовать сложные методы нейронных сетей, если есть гораздо более простые и изученные статистические методы?

Прежде всего следует посмотреть на данные, которые были использованы авторами. Максимальное количество значений временного ряда из представленных - 104 значения. В задаче, которая будет решена в рамках этой работы, содержится 35135 значений. Классические модели изначально были созданы для выборок, которые сейчас считаются небольшими. Наоборот методы машинного обучения используются для выборок с огромным количеством данных [17].

Далее исходя из описаний методов ARIMA и LSTM можно отметить, что в первом человек занимается настройкой параметров модели таких как  авторегрессионный порядок p и скользящая средняя порядка q. В отличие от LSTM, который предполагает, что нейронная сеть сама принимает решение о, например, очистке состояния. Эти факты наводят на мысль о том, что использование LSTM [18] нейронной сети более подходит для автоматизирует решение задачи. Менее требователен к участию человека в вычислениях. Такой метод легче запрограммировать для автономной работы, где модель нейронной сети самостоятельно обучается и по-сути являлась бы “черным ящиком”, в который подаются собранные данные для получения на выходе прогноза.

Сравнивая перечисленные методы и анализируя исследования противостояния машинного обучения и классических методов следующие факты приводят к выбору LSTM нейронной сети в качестве метода решения задачи прогнозирования контакт-центра:

* Метод LSTM это метод машинного и глубокого обучения, который позволяет работать с данным в задаче количеством данных
* Метод LSTM изначально разрабатывался, как метод решения задач с последовательностью данных, которые не должны “забываться” в ходе обучения.
* LSTM нейронная сеть разработана так, что насколько это возможно сама рассчитывает свои внутренние параметры, без нужды в человеческих решениях [19].

2.5 Вывод

Рассмотренное в начале главы решение НТЦ АРГУС “WFM CC” содержит в себе весомое количество функционала по расчету, составлению, управлению, оптимизации и учету в контексте применения к сотрудникам контакт-центра. Она составляет прогнозы с учетом множества параметров. Показывает возможности предоставления множества параметров для пользователя, чтобы составить более унифицированные и качественные прогнозы и графики работы. Однако указанного функционала, который был описан в конце предыдущей главы в решении нет. Для реализации предлагаемого алгоритма, прежде всего ставится задача, как задача прогнозирования временного ряда.

Такую задачу можно решать большим количеством способов. Представлены два глобальных, классические модели и модели машинного обучения, в частности, нейронные сети. Из классических показаны: Модель авторегрессии (AUTOREG), Модели авторегрессионной скользящей средней (ARMA), Интегрированная модель авторегрессии — скользящего среднего (ARIMA), Сезонная модель ARIMA, Модель сезонного экспоненциального сглаживания, Мультипликативная модель Холта-Винтерса, Аддитивная модель Хольта-Винтерса. Из методов нейронных сетей: Многослойный персептрон (MLP), Байесовская нейронная сеть (BNN), Нейронная сеть с функцией радиального базиса (RBF), Обобщенная регрессионная нейронная сеть (GRNN), Рекуррентная нейронная сеть (Recurrent neural network, RNN), Нейронная сеть долгой краткосрочной памяти (Long Short-Term Memory, LSTM).  Несмотря на то, что существуют исследования доказывающие, что классические модели строят в целом лучшие прогнозы, в отличии от моделей нейронных сетей, для практической части исследования выбраны именно нейросети класса LSTM, в связи со спецификой решаемой задачи.

Глава 3. РАЗРАБОТКА АЛГОРИТМА

3.1 Выбор среды моделирования

Класс приложений, построенных на нейронных сетях, обычно опирается на тяжелые вычисления с массивными наборами данных. Поэтому считается, что параллельные вычисления традиционно выполняют обучающий процесс модели нейронной сети в приемлемое время. Графические процессоры (GPU) являются массово параллельными устройствами-кандидатами для выполнения такой параллельной задачи. Этот тип ускорителя является вездесущим, доступным и обеспечивает высокую скорость GFlops / Dollar. Кроме того, основные рамки глубокого обучения запрограммированы для графических процессоров NVIDIA.

Аппаратные ресурсы увеличивают риски: ​​недостаточное использование и чрезмерное использование, износ оборудования и сбои. Есть также расходы, связанные с обслуживанием, энергией и человеческими ресурсами. В реальности исследовательской группы может быть трудно сохранить надежный компьютер с несколькими графическими процессорами для испытаний. Кроме того, для каждого члена команды стоит выделить рабочую станцию, оснащенную высокопроизводительным графическим процессором. В настоящее время облачные решения привлекательны тем, что предоставляют оборудование на лету и устраняют необходимость в обслуживании и настройке аппаратных ресурсов. Облачные платформы, такие как Amazon, Intel, Azure и Google Cloud, предоставляют графические процессоры с почасовой оплатой и среду выполнения, полностью настроенную для глубокого обучения.

3.1.1 Google Colaboratory (так верно)

В рамках представленной выше области, Google создал Colaboratory (a.k.a. Colab), облачный сервис для распространения образования и исследований в области машинного обучения. Среда выполнения, предоставляемая этим облачным сервисом, полностью сконфигурирована с ведущими библиотеками искусственного интеллекта (AI), а также предлагает надежный графический процессор. Эта служба Google связана с учетной записью Google Drive и является бесплатной [20].

Colab является хостом для среды разработки Jupyter notebook [21], в котором стандартно стандартно предоставляется возможность писать программный код на языке Python. У пользователя есть возможность использовать заранее установленные библиотеки Python для написания программ, а также устанавливать дополнительное ПО. В данной работе используются следующие библиотеки:

3.1.3 NumPy

NumPy - это фундаментальный пакет для научных вычислений на Python. Это библиотека Python, которая предоставляет объект многомерного массива, различные производные объекты (такие как маскированные массивы и матрицы) и набор подпрограмм для быстрых операций над массивами, включая математические, логические, манипуляции с формами, сортировку, выбор, ввод / вывод , дискретные преобразования Фурье, базовая линейная алгебра, основные статистические операции, случайное моделирование и многое другое. В основе пакета NumPy лежит объект ndarray. Он инкапсулирует n-мерные массивы однородных типов данных со многими операциями, выполняемыми в скомпилированном коде для повышения производительности [22].

3.1.4 Matplotlib

Matplotlib - наносит на график ваши данные на рисунках (т.е. окнах, виджетах Jupyter и т. Д.), Каждая из которых может содержать одну или несколько осей (т. Е. Область, в которой точки могут быть указаны в виде координат xy (или тета-r в полярной области). сюжет, или XYZ в 3D-графике и т. д.) [23].

3.1.5 Pandas

Pandas - это пакет Python, предоставляющий быстрые, гибкие и выразительные структуры данных, предназначенные для того, чтобы сделать работу с «реляционными» или «помеченными» данными как простой, так и интуитивно понятной. Он призван стать фундаментальным строительным блоком высокого уровня для практического, реального анализа данных в Python. Кроме того, он имеет более широкую цель стать самым мощным и гибким инструментом анализа / обработки данных с открытым исходным кодом, доступным на любом языке. Он уже хорошо продвигается к этой цели. Pandas хорошо подходят для самых разных видов данных:

* + Табличные данные с неоднородными столбцами, как в таблице SQL или электронной таблице Excel.
  + Упорядоченные и неупорядоченные (не обязательно фиксированные частоты) данные временных рядов.
  + Данные произвольной матрицы (однородно типизированные или неоднородные) с метками строк и столбцов
  + Любая другая форма наборов данных наблюдений / статистики. Данные на самом деле не должны быть помечены для размещения в структуру данных pandas [24]

3.1.6 Math

Math - математические функции. Этот модуль обеспечивает доступ к математическим функциям, определенным стандартом [25].

3.1.7 Keras

Keras - это API глубокого обучения, написанный на Python и работающий на платформе машинного обучения TensorFlow. Он был разработан с целью обеспечения быстрого экспериментирования. Быть способным перейти от идеи к результату как можно быстрее - это ключ к проведению хороших исследований. TensorFlow 2.0 - это комплексная платформа машинного обучения с открытым исходным кодом. Вы можете думать об этом как об уровне инфраструктуры для дифференцированного программирования. Он сочетает в себе четыре ключевых способности:

* Эффективно выполнять низкоуровневые тензорные операции на CPU, GPU или TPU.
* Вычисление градиента произвольных дифференцируемых выражений.
* Масштабирование вычислений для многих устройств (например, суперкомпьютер Summit в Oak Ridge National Lab, который охватывает 27 000 графических процессоров).
* Экспорт программ («графиков») во внешние среды выполнения, такие как серверы, браузеры, мобильные и встроенные устройства.

Keras - это высокоуровневый API TensorFlow: доступный, высокопродуктивный интерфейс для решения задач машинного обучения с акцентом на современное глубокое обучение. Он предоставляет необходимые абстракции и строительные блоки для разработки и поставки решений для машинного обучения с высокой скоростью итерации [26]

3.1.8 Scikit-learn

Scikit-learn - это модуль Python, объединяющий широкий спектр современных алгоритмов машинного обучения для средних и неконтролируемых задач среднего масштаба. Этот пакет направлен на привлечение машинного обучения к не специалистам, использующим язык высокого уровня общего назначения. Акцент делается на простоту использования, производительность, документацию и согласованность API. Он имеет минимальные зависимости и распространяется по упрощенной лицензии BSD, поощряя его использование как в академических, так и в коммерческих условиях [27].

3.1.9 Statsmodels

Statsmodels - это модуль Python, который предоставляет классы и функции для оценки множества различных статистических моделей, а также для проведения статистических тестов и исследования статистических данных. Для каждого оценщика доступен обширный список статистики результатов. Результаты проверяются на соответствие существующим статистическим пакетам, чтобы убедиться в их правильности. Пакет выпущен под лицензией Modified BSD с открытым исходным кодом [28].

3.1.10 SciPy

SciPy (произносится как «Sigh Pie») - это основанная на Python экосистема программного обеспечения с открытым исходным кодом для математики, науки и техники. SciPy относится к нескольким связанным, но отдельным объектам:

* + Экосистема SciPy, коллекция программного обеспечения с открытым исходным кодом для научных вычислений на Python.
  + Сообщество людей, которые используют и развивают этот стек.
  + Несколько конференций, посвященных научным вычислениям в Python - SciPy, EuroSciPy и SciPy.in.

Библиотека SciPy, один из компонентов стека SciPy, предоставляет множество числовых подпрограмм.

3.2 Разработка модели

3.2.1 Приведение временного ряда к стационарности

Перед тем, как составить модель для прогнозирования временной ряд следует привести к стационарности [29].

Для того чтобы понять стационарен ли ряд используется тест Дикки-Фуллера [2].

Статистические тесты могут использоваться только для информирования о степени, в которой нулевая гипотеза может быть отклонена или не отклонена. Результат должен быть интерпретирован для данной проблемы, чтобы быть значимым.

Тем не менее, они могут предоставить быструю проверку и подтверждающие доказательства того, что временной ряд является стационарным или нестационарным.

Дополненный тест Дики-Фуллера - это тип статистического теста, называемого тестом единичного корня. Интуиция за единичным корневым тестом заключается в том, что он определяет, насколько сильно временной ряд определяется трендом.

Существует целый ряд тестов единичного корня, и Дополненный тест Дики-Фуллера может быть одним из наиболее широко используемых. Он использует авторегрессионную модель и оптимизирует информационный критерий для множества различных значений запаздывания.

Нулевая гипотеза теста состоит в том, что временной ряд может быть представлен единичным корнем, что он не является стационарным (имеет некоторую зависящую от времени структуру). Альтернативная гипотеза (отвергая нулевую гипотезу) состоит в том, что временной ряд является стационарным.

Нулевая гипотеза (H0): если ее не удалось отклонить, она предполагает, что временной ряд имеет единичный корень, то есть он нестационарный. Он имеет некоторую зависящую от времени структуру. Альтернативная гипотеза (H1): нулевая гипотеза отвергается; он предполагает, что временной ряд не имеет единичного корня, то есть он является стационарным. Он не имеет временной структуры [2].

Далее над получившимся временном ряду было проведено преобразование в стационарный ряд. Для выполнения этого преобразования использовались следующие методы: дифференцирование, сезонное дифференцирование и преобразование Бокса-Кокса.

Дифференцирование - это метод преобразования набора данных временного ряда. Его можно использовать для устранения зависимости ряда от времени, так называемой временной зависимости. Это включает в себя такие структуры, как тренд и сезонность.

Дифференцирование может помочь стабилизировать среднее значение временного ряда, удаляя изменения уровня временного ряда и, таким образом, устраняя (или уменьшая) тенденцию и сезонность.

Дифференцирование выполняется путем вычитания предыдущего наблюдения из текущего наблюдения.

|  |  |
| --- | --- |
| y′t=yt−yt−1, | (32) |

где y’t - разница, yt - текущее наблюдение, yt−1 - предыдущее наблюдение.

Инвертирование процесса требуется, когда прогноз должен быть преобразован обратно в исходный масштаб. Этот процесс можно повернуть вспять, добавив наблюдение на предыдущем временном шаге к значению разницы.

Взятие разницы между последовательными наблюдениями называется разницей в лаг . Дифференцирование с лагом может быть скорректировано в соответствии с конкретной временной структурой. Для временных рядов с сезонным компонентом можно ожидать, что отставание будет периодом (шириной) сезонности. На рисунке 16 показан пример дифференцирования временного ряда, характеризующего количество пассажиров авиакомпании от продолжительности ее существования в месяцах.

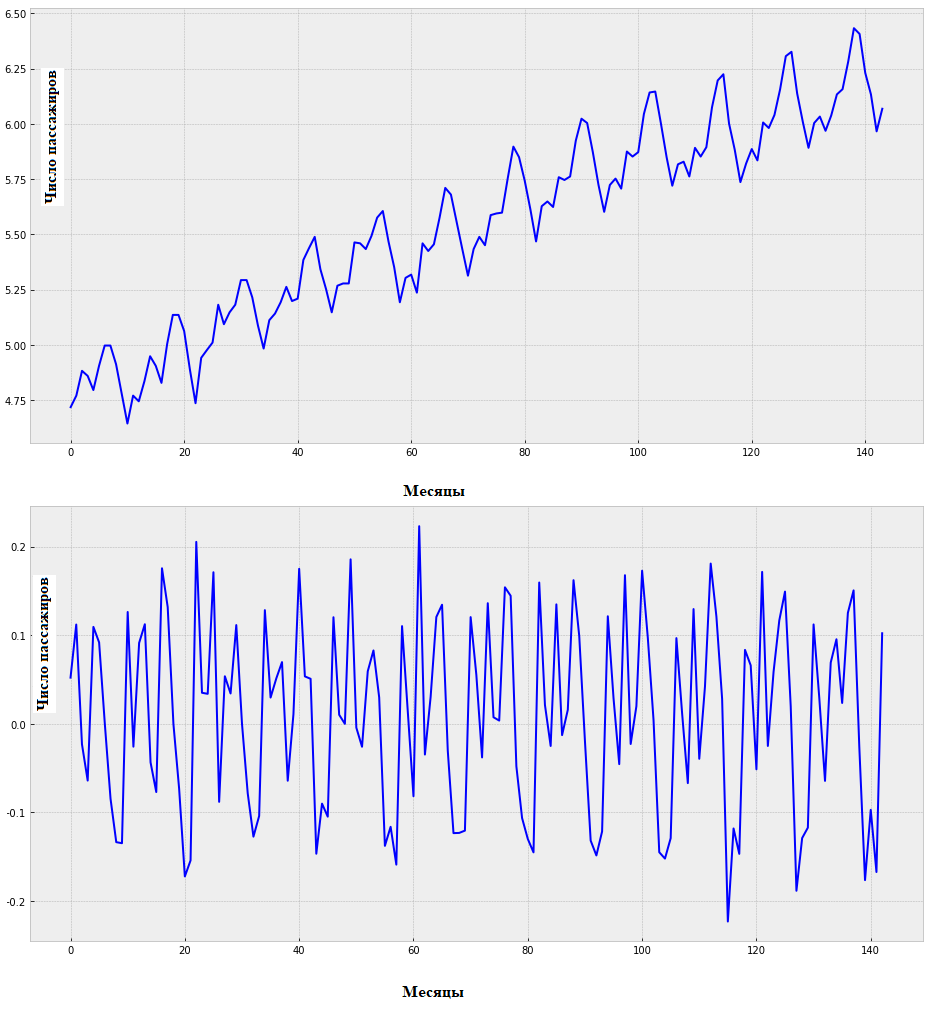


Рисунок 16 - Пример дифференцирования.

Верхний граффик отражает количество пассажиров до преобразования, нижний граффик отображает последствие преобразований.

Самый простой подход к определению того, есть ли аспект сезонности, состоит в том, чтобы построить и просмотреть данные, возможно, в разных масштабах и с добавлением линий тренда [2].

Преобразование Бокса-Кокса в свою очередь является силовым преобразованием. Силовое преобразование удаляет сдвиг из распределения данных, чтобы сделать распределение более нормальным (гауссовским). Пример преобразования на рисунке 17. В наборе данных временного ряда это может привести к удалению изменения дисперсии во времени. В библиотеке SciPy формула преобразования задается следующим образом:

(33)

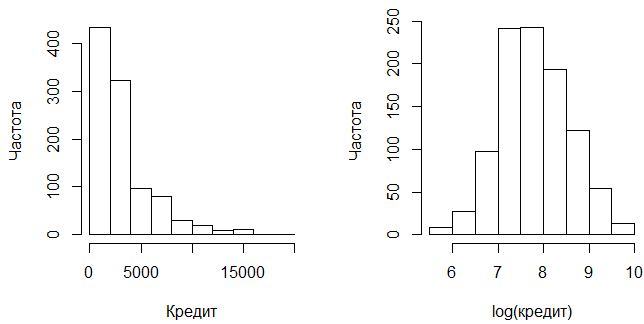


Рисунок 17 - Зависимость частоты выдачи кредитов от суммы кредита до(слева) и после(справа) преобразования Бокса-Кокса.

3.2.2 Построение модели нейронной сети

Модель архитектурно состоит из двух LSTM слоев с конечным полносвязным стандартным слоем для свертки в массив размерности (1:96) по числу пятнадцатиминутных интервалов в сутках. Используется активационная функция ReLu и оптимизатор Adam [30]. Особенность модели состоит в том, что используются слои dropout [4] и earlystopping [2] между слоями нейронов, и LSTM слои находятся в состоянии stateful [4].

В нейронную сеть для обучения и получения промежуточных результатов будут подаваться значения трех понедельников на обучение и одному на отклик со смещением в один понедельник. Это было сделано для того, чтобы имитировать накопление данных работы контакт-центра с его запуска. Соответственно, начиная с четвертой недели, модель начинает обучаться на реальных данных, настраивается, далее происходит смещение на одну неделю, модель обучается на значениях понедельников со второй по четвертую неделю со значениями отклика пятой недели и так далее (рисунок 1). Массив значений последнего понедельника будет взят для сравнения с предсказанием нейронной сети.

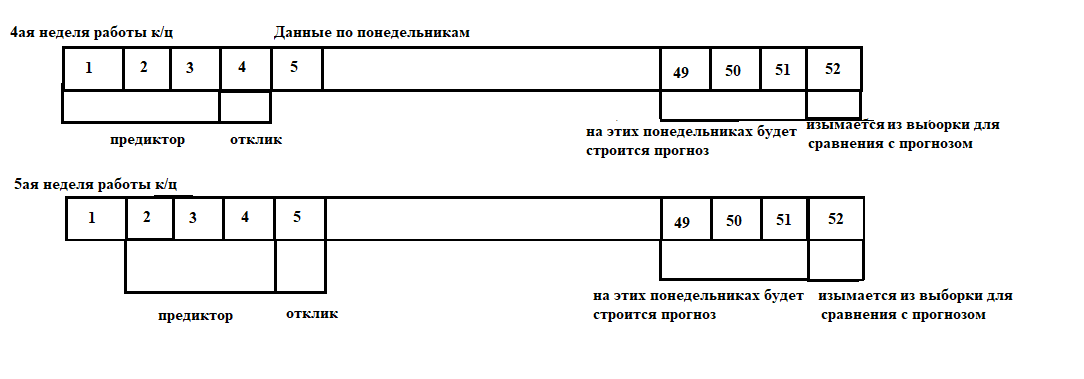


Рисунок 18 - Принцип обучения нейронной сети

По ходу обучения подсчитываются среднеквадратическая ошибка (Mean Squared Error, MSE) (1):

|  |  |
| --- | --- |
|  | (34) |

где тое значение из выборки,  тое значение, возвращенное моделью, n - количество значений.

Она нужна для подведения результатов в дальнейшем и для настройки весов нейронной сети.

3.3 Моделирование и сравнительный анализ результатов

3.3.1 Подготовка данных к подаче в нейронную сеть.

Исходный набор данных выглядит как представлено на рисунке 19. Данные представляют собой пятнадцатиминутные интервалы нагрузки на контакт центр за год. Для удобства работы с количеством данных и избежания недельной сезонности из данных были выбраны только понедельники. Всего получилось 4992 значения.  Требуется построить нейросетевую модель и, обучая ее на известных данных, выяснить, какой объем исторических данных требуется для введения модели в эксплуатацию. Идея такова, что нейронная сеть по данным о нагрузки на понедельники в течении года, будет предсказывать последний понедельник в году.

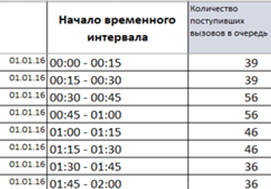


Рисунок 19 - Количество поступивших вызовов за первые 2 часа

В нейронную сеть будут поступать последовательно только значения нагрузки, то есть временные интервалы. Нейронная сеть будет настраивать веса исходя из подаваемой обучающей выборки, а на тестовой выборке мы выясним насколько хорошо работает модель. В итоге, построим предсказание и сравним с реальными данными.

Для того чтобы заранее убрать недельную сезонность, уменьшить время обучения модели, и удобства работы с данными, было решено выделить среди всей выборки только понедельники. Так задача немного изменилась. Теперь по понедельникам из выборки будет предсказан следующий понедельник.

Построение прогноза начинается с приведения временного ряда к стационарности. Прежде всего данные приводятся в нормальную форму, удаляется дисперсия с помощью преобразования Бокса - Кокса. Для удаления трендовой составляющей произведено дифференцирование ряда с лагом 1. Так как данные представляют из себя нагрузку на понедельники, предполагается, что имеет место годовая сезонность. Поэтому проводилось дифференцирование с лагом 12.

По итогу проведен тест Дики-Фуллера, выяснилось, что этого было достаточно для приведения ряда к стационарности, значение *p-value = 0.000000* (рисунок 1). Реализация теста не отображает значения менее 10^-6, однако такое отображение является достаточным для моделирования, так как ряд может считаться стационарным при значениях *p-value < 0,01* [2].

Приведенные к стационарности данные проходят преобразование с помощью функции MinMaxScaler [31]. Функция придает данным вид значений от 0 до 1. Данные подаются в нейронную сеть в виде тензора, содержащего 48 наборов пар “предиктор-отклик”, где массив “предикторов” содержит в себе по три массива значений нагрузки за три понедельника, итого 288 значений. А массив откликов один массив значений нагрузки за понедельник после трех понедельников предикторов итого 96 значений.

3.3.2 Работа модели и результаты

Реализация модели нейронной сети выглядит как показано на рисунке 2. Output Shape означает размерность выхода массива, после слоя нейронной сети. В конечном итоге мы получаем на выходе тензор, содержащий в себе последний понедельник, предсказанный нейронной сетью.

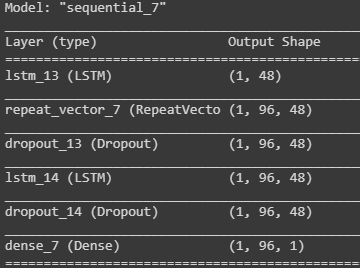


Рисунок 20 - Модель нейронной сети

Далее происходит процесс обучения нейронной сети (рис.21). На рисунке показано количество эпох обучения, то есть число поданных пар “предиктор-отклик”, результат среднеквадратической ошибки после на каждой эпохе(loss) и время, которое потребовалось на вычисления.

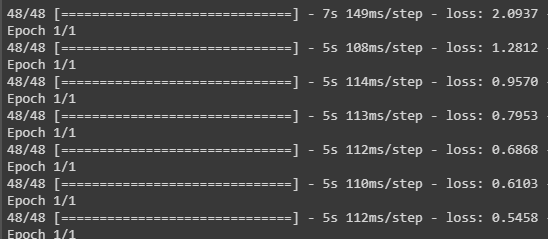
\

Рисунок 21 - Обучение нейронной сети

Нейронная сеть обучилась на тренировочных выборках, поэтому на нее подается тензор в виде трех понедельников предикторов для получения отклика в виде данных за один понедельнк.  После получения результатов и отображения их на графиках, можем заметить, что в ходе обучения, нейронная сеть не достигла минимальной ошибки, однако очень приблизилась к этому значению. Если ошибка и растет на последних эпохах, то не значительно - на графике заметить это получится с трудом (рис. 22). Тем не менее ошибка уменьшается по ходу всего обучении и нет точки, в которой можно сказать, что ошибка достигла минимума.

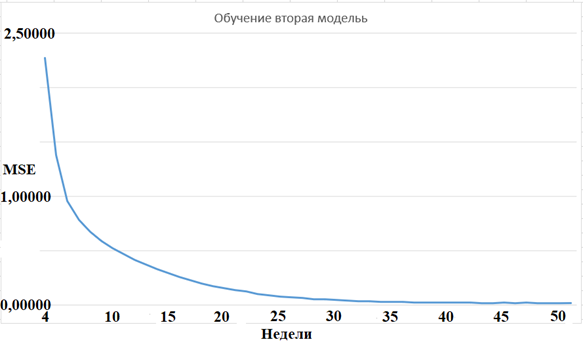


Рисунок 22 - Результаты обучения

Построим прогноз на последний понедельник в году и сравним с реальными значениями (рис. 23). Несмотря на то, что значения, полученные от нейронной сети, изменяются не так же линейно, как реальные, нейронная сеть верно отражает суточный тренд с определенной ошибкой. Среднеквадратическая ошибка, полученная от сравнения реальных значений понедельника и предсказанного, получилась равной 0,05868. Соотнеся со значениями ошибок, полученными на обучении, она находится между 24 (*MSE = 0,0685*) и 25 (*MSE = 0,057*) неделями.

3.3.3 Выводы

Из текущих результатов можно сделать вывод, что для ввода в эксплуатацию модели необходимы реальные значения за год. Тем не менее после определенного момента в обучении значение ошибки начинает снижаться слабее. Из соотнесения полученной среднеквадратической ошибки при предсказании и ошибками при обучении, можно сделать вывод о том, что с 25 недели модель можно вводить в эксплуатацию, не дожидаясь конца обучения, но и не заканчивая его. На данном этапе можно сказать о том, что при решении поставленной задачи более целесообразно использовать несколько *LSTM* слоев рекуррентной нейронной сети с stateful состоянием и dropout слоями (четвертая модель), а также метод early stopping. В дальнейших исследованиях следует ввести методику работы с большим количеством данных, для анализа не только на понедельниках, но и на всех значениях в году. Также следует понять, есть ли возможность и как сократить время на обучение модели, действительно ли необходимо ждать полгода для введения модели в работу.

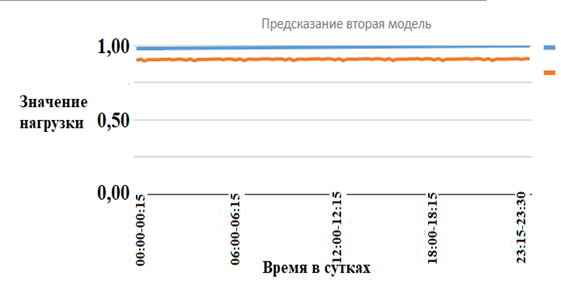


Рисунок 23 - Результаты предсказания. Верхняя линия отражает реальные значения нагрузки, нижняя - предсказание нейронной сети

ЗАКЛЮЧЕНИЕ

 В ходе развития контакт-центра как концепции новой структуры обработки клиентских вызов, была построена новая архитектура, введены новые понятия и функциональные особенности. Тем самым формат приема заявок был усовершенствован, была произведена оптимизация рабочих процессов. На сегодняшний день на карте eTOM можно найти процессы, указывающие на то, что есть возможность еще больше модернизировать подход к построению контакт-центра и управлению персоналом. Карта показывает процессы Workforce Management, в которых уделяется внимание управлению графиком работников параллельно с прогнозированием рабочей нагрузки, что в ключе контакт-центров будет нагрузкой поступающих заявок.  Однако решение такой задачи не содержит достаточного оценочного функционала, который помог бы владельцам свежепостроенных контакт-центров определить, когда им вводить модель в работу.

Решение НТЦ АРГУС “WFM CC” предоставляет большое число функций такие как: расчет, составление графиков, управление, оптимизация и учет многих аспектов деятельности контакт-центра. Оно составляет прогнозы с учетом различных параметров. Тем не менее решение не помогает владельцу нового контакт-центра, в котором еще нет накопленных исторических данных, определить, когда наиболее приемлемо вводить в работу составленные прогнозы. В работе предлагается составить алгоритм, который составит прогноз для нового контакт-центра через год накопления данных, делая промежуточные результаты, которые помогут решить такую задачу. В целом, такая задача является задачей прогнозирования временного ряда, ее можно решать большим количеством способов. Глобально методами решения могут быть классические модели и модели машинного обучения, в частности, нейронные сети. Автором работы был проведен анализ исследований, показывающих преимущество классических моделей, тем не менее было принято решении использовать нейросети, как подход, имеющий преимущества в реализации и работой с большим количеством данных.

Результаты работы показывают, что модель требует для максимального эффекта всего набора предоставленной ей выборки. Но отмечено, что после определенного момента в обучении значение ошибки начинает снижаться слабее. Из соотнесения полученной среднеквадратической ошибки при предсказании и ошибками при обучении, можно сделать вывод о том, что с 25 недели модель можно вводить в эксплуатацию, не дожидаясь конца обучения, но и не заканчивая его. Для дальнейшего исследования требуется более углубиться в понимание того, возможно ли быстрее прийти к наиболее приемлемому результату, зависит ли это от модели, реализованного алгоритма или же от имеющихся данных. Мир машинного обучения постоянно пополняется различными алгоритмами исследователей по всему миру, возможно существует более выгодные модели, но они не так хорошо изучены и про них мало информации.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Д.Рутковская. Нейронные сети, генетические алгоритмы и нечеткие системы; Горячая линия – Телеком. – Москва 2006. – 452с.
2. Brownlee J. Introduction to Time Series Forecasting With Python: How to Prepare Data and Develop Models to Predict the Future. Machine Learning Mastery, 2017. - 367 c.
3. Brownlee J. Deep Learning with Python: Develop Deep Learning Models on Theano and TensorFlow Using Keras. Machine Learning Mastery, 2017. - 245 c.
4. Brownlee J. Better Deep Learning: Train Faster, Reduce Overfitting, and Make Better Predictions. Machine Learning Mastery, 2018. - 575 c.
5. Суринов, А. Е. Экономическая статистика в страховании : учебник для академического бакалавриата / А. Е. Суринов. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2018. — 276 с.
6. Гольдштейн Б.С., Фрейнкман В.А Call-центры и компьютерная телефония СПб.: BHV – Санкт-Петербург, 2006. — 368 с
7. TM Forum. Business Process Framework (eTOM). [Электронный ресурс]. Путь доступа: http://www.elmayorportaldegerencia.com/Documentos/Telecomunicaciones/eTOM/
8. А.Гольдштейн, С. Кисляков, И. Садовский, Школа WorkforceManagement для контакт-центра. МТК (Мобильные телекоммуникации), 2017. – 13с.
9. Josh Patterson, Adam Gibson. Deep Learning: A Practitioner's Approach. O'Reilly Media, 2017. – 532с.
10. НТЦ «Аргус». Система для оптимизации работы контакт-центра Аргус WFM CC, 2017. – 4c. [Электронный ресурс]. Путь доступа: http://argustelecom.ru/listovka-wfm-cc-okt2017.pdf
11. НТЦ «Аргус». Аргус WFM CC. Управление Контактным Центром, 2019. – 32с. [Электронный ресурс]. Путь доступа: http://argus-wfmcc.ru/wp-content/uploads/2019/
12. Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование: Учебник. — М.: Финансы и статистика, 2001. — 228 с
13. Gerhard Bry, Charlotte Boschan. Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, National Bureau of Economic Research; distributed by Columbia University Press, 1971. – 216с.
14. Han Shih, Suchithra Rajendran. Comparison of Time Series Methods and Machine Learning Algorithms for Forecasting Taiwan Blood Services Foundation’s Blood Supply. 2019. – 7с.
15. Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning. MIT Press, 2016. – 781с.
16. The SciPy community. scipy.stats.boxcox, 2020. – 1с. [Электронный ресурс]. Путь доступа: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.boxcox.html
17. Danilo Bzdok, Naomi Altman & Martin Krzywinski. Points of Significance Statistics versus machine learning. Nature Methods, 2018. – 2с.
18. : Makridakis S, Spiliotis E, Assimakopoulos V Statistical and Machine Learning forecasting methods: Concerns and ways forward. PLoS ONE, 2018. – 26с.
19. Ahmed, Nesreen K. , Atiya, Amir F. , Gayar, Neamat El and El-Shishiny, Hisham. An EmpiricalComparison of Machine Learning Models for Time Series Forecasting', Econometric Reviews, 2010. - 594 — 621
20. Carneiro, Tiago et al. “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications.” IEEE Access 6.208. - 61677-61685.
21. Google. Colaboratory. Frequently Asked Questions. [Электронный ресурс]. Путь доступа: https://research.google.com/colaboratory/faq.html
22. NumPy. What is NumPy? [Электронный ресурс]. Путь доступа: https://numpy.org/doc/stable/user/whatisnumpy.html
23. John Hunter, Darren Dale, Eric Firing, Michael Droettboom. Matplotlib Release 3.2.2. 2020. – 2681c. [Электронный ресурс]. Путь доступа: https://matplotlib.org/Matplotlib.pdf
24. Wes McKinney. pandas: powerful Python data analysis Toolkit Release 1.0.5. Pandas. 2020 – 3091c. [Электронный ресурс]. Путь доступа: https://pandas.pydata.org/docs/pandas.pdf
25. Python. Math – Mathematical function. [Электронный ресурс]. Путь доступа: https://docs.python.org/3/library/math.html
26. Keras. About Keras. [Электронный ресурс]. Путь доступа: https://keras.io/about/#why-this-name-keras
27. Pedregosa, Fabian & Varoquaux, Gael & Gramfort, Alexandre & Michel, Vincent & Thirion, Bertrand & Grisel, Olivier & Blondel, Mathieu & Prettenhofer, Peter & Weiss, Ron & Dubourg, Vincent & Vanderplas, Jake & Passos, Alexandre & Cournapeau, David & Brucher, Matthieu & Perrot, Matthieu & Duchesnay, Edouard & Louppe, Gilles. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 2012. - 2825-2830 c.
28. Seabold, Skipper, and Josef Perktold. “statsmodels: Econometric and statistical modeling with python.” Proceedings of the 9th Python in Science Conference. 2010. [Электронный ресурс]. Путь доступа: https://www.statsmodels.org/stable/index.html
29. Kapoor Amita. Hands-On Artificial Intelligence for IoT: Expert machine learning and deep learning techniques for developing smarter IoT systems. 2019. – 363c.
30. Kingma, Diederik P. and Jimmy Ba. “Adam: A Method for Stochastic Optimization.” CoRR abs/1412.6980. 2015 – 15c.
31. Scikit-learn. sklearn.preprocessing.MinMaxScaler. [Электронный ресурс]. Путь доступа: https://scikit learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.htm