МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

(ФГБОУ ВО «КубГУ»)

Экономический факультет

Кафедра экономики и управления инновационными системами.

**КУРСОВАЯ РАБОТА**

**МЕТОДОЛОГИЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ**



Работу выполнила \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Колесникова А. Ж.

(подпись, дата)

Направление 27.03.03 Системный анализ и управление\_\_\_\_\_\_\_\_\_\_

Направленность (профиль) Системный анализ и управление экономическими процессами\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Научный руководитель:

канд. экон. наук, доц.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Библя Г.Н.

(подпись, дата)

Нормоконтролер:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Библя Г. Н.

(подпись, дата)

Краснодар

2020

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

(ФГБОУ ВО «КубГУ»)

Экономический факультет

Кафедра экономики и управления инновационными системами.

**ЗАДАНИЕ**

на курсовую работу

по теме: «Методология интеллектуального анализа данных»

Студента Колесниковой Алене Жановне группы 114 направление подготовки 27.03.03 Системный анализ и управление, направленность (профиль) Системный анализ и управление экономическими процессами

**Цель:** Изучить предметную область. Рассмотреть возможности применения методики интеллектуального анализа данных для предмета исследования.

**Основные вопросы, подлежащие разработке (исследованию):**

1. Теоретический обзор современных подходов, методов и алгоритмов исследования проблемы;
2. Анализ предметной области, спецификации интеллектуальных методов;
3. Реализация изученных методов.

**Основная литература:**

1. Кориков, А. М. Теория систем и системный анализ : учеб. пособие / А.М. Кориков, С.Н. Павлов. — М. : ИНФРА-М, 2018. — 288 с. — (Высшее образование: Бакалавриат).

Режим доступа: <http://znanium.com/catalog.php?bookinfo=935445>

2. Библя,Г.Н. Теория системного анализа и управления [Текст] : методические указания по выполнению курсовой работы / [Г. Н. Библя] ; М-во образования и науки Рос. Федерации, Кубанский гос. ун-т. - Краснодар : [КубГУ], 2016.

3. Библя,Г.Н. Теория системного анализа и управления [Текст] : учебно-методическое пособие; направление 27.03.03 Системный анализ и управление/ [Г. Н. Библя] ; ФГБОУ КубГУ. - Краснодар : [КубГУ], 201

Срок представления законченной работы 20 мая 2020 г.

Дата выдачи задания 06 февраля 2020 г.

Руководитель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ /Г.Н. Библя /

Задание получила 06 февраля 2020 г.

Студент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ /Колесникова А. Ж./

СОДЕРЖАНИЕ

[Введение 4](#_Toc41944153)

[1 Теория и методы интеллектуального анализа данных 6](#_Toc41944154)

[1.1 Данные и требования к их обработке 7](#_Toc41944155)

[1.2 Виды данных и критерии их оценивания 9](#_Toc41944156)

[1.3 Классификация стадий Data Mining 12](#_Toc41944157)

[1.4 Классификация методов интеллектуального анализа 13](#_Toc41944158)

[1.5 Задачи интеллектуального анализа 14](#_Toc41944159)

[2 Системный анализ деятельности компании 16](#_Toc41944160)

[2.1 Общая характеристика 16](#_Toc41944161)

[2.2 Анализ внутренней среды предприятия 18](#_Toc41944164)

[2.3 Анализ внешней среды предприятия 22](#_Toc41944165)

[3 Разработка проекта по внедрению системы интеллектуального анализа данных на предприятии 24](#_Toc41944166)

[3.1 Разработка сетевого графика 24](#_Toc41944167)

[3.2 Анализ резервов времени 28](#_Toc41944168)

[3.3 Перспективы развития систем интеллектуального анализа данных 29](#_Toc41944169)

[Заключение 30](#_Toc41944170)

[список использованной литературы 31](#_Toc41944171)

# ВВЕДЕНИЕ

Сегодня количество данных, доступных по всему миру, удваивается, что также дает нам возможность сказать, что информация, релевантная для использования, быстро сокращается. Быстрое развитие всех типов технологий и сбор данных о них в огромных массивах, которые необходимо регулярно анализировать, создает огромное бремя для человеческих возможностей, а объемы данных уже настолько велики, что все это оправдывает непрерывный поиск возможностей и путей извлечь полезную информацию из огромного потока данных. Исследования в области искусственного интеллекта как заменителя человека зашли далеко и позволили разработать методологию интеллектуального анализа данных (ИАД), которая не зависит от конкретных областей и может использоваться как во всей области, так и в отдельных случаях. Осознание несравненных преимуществ использования инструментов и методов ИАД и понимание того, что предлагает интеллектуальный анализ данных, помогает людям выработать необходимые рекомендации для улучшения работы систем. Вот почему интеллектуальные системы анализа данных (ИСАД) были спроектированы таким образом, чтобы минимизировать усилия некоторых лиц, принимающих решения (ЛПР), на протяжении всего процесса анализа данных, а также при создании специальных алгоритмов анализа. Многие ИСАД способны идентифицировать причинно-следственные связи, скрытые паттерны, которые также анализируются. Фактически, интеллектуальный анализ данных - это обработка базы данных и выявление тенденций, которые помогают принимать решения. Принципы ИАД известны уже много лет, но начали распространяться только с появлением огромного потока данных. Они вызвали взрыв популярности конкретного анализа - Data Mining.

*Целью курсовой работы* является изучение и анализ функционирования финансов и реализации компании и принятие решений с помощью дерева решений.

*Объектом исследования* данной работы является ПАО «Аэрофлот».

*Предмет исследования* является система финансов и организация рабочего процесса предприятий.

В качестве *информационной базы* исследования были использованы труды известных отечественных и зарубежных ученных по данной теме.

*Теоретической базой* работы послужили научные работы отечественных и зарубежных ученых и практиков в области системного анализа, экспертных методов и информационных технологий. Исследование проведено на основе изучения мирового и отечественного опыта развития и функционирования систем управления.

*Методологическая база* исследования включает в себя систематизацию данных, классификацию, логический и сравнительный анализ, метод обобщения и структуризацию.

*Структура работы* определена характером исследуемых в ней вопросов. Курсовая работа содержит: введение, три раздела, девять подразделов, заключение, список использованных источников.

Во введении обоснована актуальность работы, поставлена цель и задачи, объект и предмет данной работы.

Первая глава включает в себя полноразмерное описание всех стадий подготовки к интеллектуальному анализу, введение терминологии и первичное ознакомление с методами анализа.

Вторая глава включает результаты системного анализа организационной структуры управления ПАО «Аэрофлот».

Третья глава включает в себя разработку проекта по внедрению системы интеллектуального анализа данных на предприятии.

В заключении сделан вывод по проделанной работе.

# 1 Теория и методы интеллектуального анализа данных

Data Mining - это междисциплинарная область, которая была создана и развивается на основе прикладной статистики, распознавания образов, искусственного интеллекта, теории баз данных, визуализации, машинного обучения и других дисциплин. Давайте рассмотрим некоторые из них более подробно.

Статистика - это наука о том, как собирать, анализировать и обрабатывать данные для выявления тенденций, присущих определенному явлению. То есть это набор методов для планирования сбора и анализа данных. Работа с данными.

Машинное обучение - это процесс приобретения новых знаний по программе. Самый очевидный пример - нейронные сети.

Искусственный интеллект является научной областью, в которой определяются и решаются задачи различных типов моделирования человеческой деятельности, которые традиционно считаются интеллектуальными. Другими словами, интеллектуальные системы обладают способностью выполнять творческие функции, которые изначально считались ориентированными на человека.

Технология Data Mining - это поиск неподготовленных данных, ранее неизвестных, а также полезных и доступных для использования знаний для принятия решений. Поскольку эта технология предназначена для работы в глобальных базах данных, она в основном имеет несколько моделей, а именно: неочевидная, объективная и практически полезная.

Неочевидные стандарты это те, которые не могут быть обнаружены стандартным методом анализа.

Объективно - эти модели полностью соответствуют действительности, чего нельзя сказать о мнении эксперта, которое носит субъективный характер.

Практические модели говорят нам, что определенные практические ценности могут применяться напрямую.

Прежде чем приступить к обзору уже существующих методов, мы перейдем к объекту обращения ИАД — данные.

## Данные и требования к их обработке

Данные - это факты, текст, графика, числовые выражения, документы и другие типы сегментов. Все это может быть достигнуто путем практической деятельности. Данные должны быть конвертируемыми и могут изменять форму для требуемого типа хранилища. Вы можете сказать, и проще говоря, данные представляют собой набор необработанных данных.

Данные включают в себя 2 категории: объекты и атрибуты. Объект может быть просто серией чисел. Атрибуты могут быть ресурсами, отличными от объекта, то есть это тип переменной. В результате перехода из неопределенного состояния в конкретное состояние получается набор переменных изучаемой концепции. Обратите внимание на понятие переменной - это общее свойство всех объектов, которое может меняться от одного перехода к другому. Основным значением переменной является наличие ресурса.

При анализе базы данных, как правило, невозможно принять во внимание всю совокупность объектов в свете огромного потока данных, но необходимо учитывать только некоторые из них. Размер этой части должен зависеть от разнообразия доступных объектов, представленных в общей популяции. Население в целом - все элементы, необходимые для исследователя. Выборка является частью общей совокупности, которая была выбрана на основе определенных параметров. А параметры являются числовыми характеристиками.

Также, наряду с закономерностями существует и определенная специфика требований к обработке данных:

* неограниченный объем;
* разнородность;
* корректные результаты;
* простые инструменты для обработки.

Традиционный процесс сбора данных состоит из:

* анализ предметной области;
* постановка задачи;
* подготовка данных;
* построение моделей;
* проверка и оценка;
* выбор;
* применение;
* коррекция и обновление.

Анализ области исследования включает в себя разделение значимых и второстепенных факторов, которые влияют на текущее исследование. Существенность выбранных данных зависит от исходной задачи. В процессе ее исследования мы составляем ее элементарную модель. Она может включать в себя специализированные графические обозначения, блок-схемы и различные структурные анализы. Модель, в свою очередь, описывает процессы, происходящие в области и данные, используемые для этих процессов. Успех функционирования приложения Data Mining зависит от того, насколько хорошо составлена ​​модель.

Постановка задачи включает в себя: формулировку и формализацию задачи. Она также содержит описание статической и динамической сводки объектов. Описание статистки требует описание объектов в целом и их свойств. Динамка подразумевает особенности, которые напрямую влияют на поведение объекта. Иногда анализ и постановка объединяются в один шаг, если спектр данных не так велик.

Подготовка данных - это важный шаг, который влияет на конечный результат всех процессов. Этот этап включает в себя несколько шагов:

1. Определение и анализ требований к данным. Это как раз создание модели данных, необходимой для работы Data Mining. В этом случае изучаются все проблемы распространения данных, доступа к ним, необходимости дополнительных источников и общих аналитических характеристик системы.
2. Сбор данных. Наличие централизованного хранилища данных значительно облегчает работу с ними. Кроме того, это значительно дешевле, чем использование отдельных витрин и баз данных. Конечно, использование централизованного хранилища не всегда предоставляется возможным, тогда используют архивную систему базы данных, что также значительно облегчает работу с отдельными частями. Если данные уже упорядочены, то следует определить, имеют ли они сезонность или циклические компоненты. В противном случае, есть компоненты не связанные временным диапазоном и не имеют хронологического порядка, следует иметь в виду, что: недостаточное количество записей в наборе может привести к сбоям в скомпилированной модели, возможно, что некоторые данные устарели и должны быть исключены из доступного списка, алгоритмы, используемые для моделирования должны быть масштабируемыми, должна быть определенная пропорция входных переменных и наблюдений, а количество записей должно значительно их превышать. Сам набор данных должен быть разнообразным для всех типов ситуаций, чтобы иметь свойство приспособляемости.

Предварительная обработка данных включает в себя анализ данных. Анализировать можно как высококачественные, так и некачественные данные, но результат достигается в обоих случаях. Существует критерий качества для разграничения данных. Это важный шаг в процессе интеллектуального анализа данных (ИАД).

**1.2 Виды данных и критерии их оценивания**

Качество данных — критерий, определяющий целостность и точность предоставляемых данных, а также возможность их интерпретации. Высококачественные данные - это полные и точные данные, которые позволяют исследователю составить полностью функционирующую модель, способную к интерпретации и принятию решений. Данные низкого качества, соответственно, не имеют таких характеристик, включая неверный формат.

Наиболее распространенные виды так называемых «грязных» данных:

* пропущенные значения;
* дубликаты;
* шумы и выбросы.

Пропущенные значения. Некоторые значения могут быть проигнорированы из-за того, что данные не были просто собраны или что некоторые атрибуты не могут быть применены априори к данным.

Дублирование данных. Сами дубликаты являются записями, которые содержат одинаковую информацию для всех параметров. Наличие дубликатов в базе данных может указывать пользователю, что данные имеют определенный приоритет над другими. Но чаще всего можно столкнуться с тем, что это просто ошибка в их подготовке. Существует только два варианта работы с дубликатами: удаление всех дубликатов (только в случае обесценивании информации) или замена группы в одну уникальную запись.

Шумы и выбросы представляют собой особую группу характеристик данных низкого качества. Выбросы — это совершенно разные объекты от остальных данных с точки зрения типа данных и формата. Они могут быть независимыми, так и иметь целую объединенную группу. Важно не только обнаружить их, но и оценить их дальнейшее влияние на данные. Различные методы интеллектуального анализа имеют различную чувствительность к шумам и выбросам. С ними нужно быть особенно осторожными, поскольку обработка таких данных не может быть ни полезной, ни надежной.

Некачественные данные подлежат очистке, то есть, происходит выявление и удаление ошибок. Метод должен удовлетворять определенным критериям:

* + удаление всех основных ошибок;
	+ поддержание определенных инструментов;
	+ не затрагивать основные данные;
	+ функции очистки данных должны быть адаптивными и к другим базам данных;
	+ надёжность выполнения всех этапов.

Этапы очистки данных включают в себя 5 шагов: анализ данных, определение порядка и правил преобразования, подтверждение, преобразования и противоток очищенных данных.

Анализ данных необходим для выявления и удаления поверхностных и легкодоступных ошибок.

Определение порядка и правил преобразования. В зависимости от объема данных, степени неопределенности и загрязнения данные могут потребовать дальнейшего изучения на ошибки. Для этого необходимо составить порядок их анализа.

Подтверждение. На данном этапе определяется эффективность крайнего этапа. Это происходит путем тестирования копии источника.

Преобразования. Обычное обновление хранилища данных при ответе на запросы большей части источников.

Обратный поток удаленных данных. После очистки данных вы должны заменить данные во всех возможных источниках и в самой памяти новыми высококачественными. Этот процесс требует больших объемов метаданных.

Основная особенность интеллектуального анализа — это широкий математический инструментарий. К методам и алгоритмам относятся искусственные нейронные сети, деревья решений, символьные правила, метод ближайшего соседа и k-ближайшего соседа, метод опорных векторов, байесовские сети, линейная регрессия, корреляционно-регрессионный анализ, методы поиска ассоциативных правил, визуализации данных, генетические алгоритмы и ограниченного перебора.

Что же касается аналитических методов, то они несут в себе характер автоматизированных математических методов. К вышеперечисленным методам мы вернемся далее.

**1.3 Классификация стадий Data Mining**

Data Mining может включать в себя от двух до трёх стадий: свободный поиск и выявление закономерностей, прогностическое моделирование и анализ исключений.

Первая стадия «Свободного поиска» включает в себя поиск скрытых шаблонов. Кроме того, система определяет предварительные шаблоны, которые особенно полезны в сверхбольших базах данных

Вторая стадия состоит из анализа первой стадии и прогнозирования неизвестных значений, обеспечивающих развитие процессов. Решаются задачи классификации и прогнозирования. Прогнозирующее моделирование является дедуктивным. Закономерности, сформированные на этой стадии, формируются от общего к частному и единичному.

Третья стадия анализируются исключения или аномалии, выявленные в неопределенностях. Также, здесь появляется понятие нормы и пределы отклонений от неё. Возможны две интерпретации: логическое объяснение в форме правила и фактические ошибки в исходных данных.

**1.4 Классификация методов интеллектуального анализа**

У интеллектуального анализа есть 2 больше группы методов, по которым осуществляется его функционирование.

* + 1. Прямое использование или подробное архивирование данных, которые используются на 2 и 3 стадии. Основная проблема - несопоставимость с большим объемом данных. Этот метод включает в себя: кластерный анализ, метод ближайшего соседа, рассуждение по аналогии.
		2. Определение и использование формализованных шаблонов. Здесь используется шаблон (образец) для правильного анализа данных и преобразования в формальные конструкции, вид которых зависит от метода, используемого ИАД. Методы в этой группы: визуализация, уравнения, кросс-табуляция и логические методы.

Логические методы включаются в себя нечеткие запросы и анализы, деревья решений и генетические алгоритмы. Они чаще всего интерпретируются в базах данных.

Методы кросс-табуляции. Они не отвечают одному из требований к интеллектуальному анализу — независимому поиску шаблонов, но они действительно находят и предоставляют шаблоны полностью для пользователя.

Методы, основанные на уравнениях, представляют собой исключительно математическую группу. Однако они могут работать только с числовыми значениями при наличии переменных, которые должны быть закодированы определенным образом.

Кибернетические методы. Это компьютерная математика и использование искусственного интеллекта. Сюда относят: искусственные нейронные сети (прогноз), оптимизация, деревья решений, нечеткая логика.

**1.5 Задачи интеллектуального анализа**

Задачи в интеллектуальном анализе зачастую называют закономерностями. Есть несколько видов задач: классификация, кластеризация, ассоциация, последовательность, прогнозирование, оценивание, визуализация, подведение итогов.

Классификация является самой простой и наиболее распространенной из задач. В результате решения проблемы аналогичные функции обнаруживаются в базе данных и происходит распределение по классам.

Кластеризация — логическое продолжение классификации. Особенность данной задачи заключается в том, что данные изначально не предопределены на классы.

Ассоциация показывает шаблоны между связанными событиями. Разница между этой и предыдущими задачами заключается в том, что поиск по шаблону основан не на свойствах объекта, а на различных событиях.

Последовательность или последующее сопоставление находит только временные шаблоны. Это означает, что шаблоны ищутся между событиями, которые происходят не одновременно, как в ассоциации, а своевременно. Другими словами, задача состоит в том, чтобы найти последовательные шаблоны.

Прогнозирование. Здесь оценивается характеристика исторических данных или будущих значений целевых числовых показателей.

Оценка и визуализация - это прогнозирование непрерывных значений знака и создание графического изображения анализируемых данных.

Подведение итогов — описание определенных групп объектов из предмета исследования.

**2 Системный анализ деятельности компании**

**2.1 Общая характеристика**

Группа «Аэрофлот» – современный авиационный холдинг, крупнейший авиаперевозчик России, входит топ-20 авиаперевозчиков мира по пассажиропотоку. По итогам 2019 года Группа «Аэрофлот» занимает 41,3% российского рынка по пассажиропотоку. В течение года авиакомпании Группы перевезли 60,7 млн пассажиров, в том числе 37,2 млн пассажиров рейсами авиакомпании «Аэрофлот».

Головной компанией Группы является ПАО «Аэрофлот». В состав Группы «Аэрофлот» входят сетевые авиакомпании «Аэрофлот», «Россия», «Аврора», а также лоукостер «Победа». Мультибрендовое предложение позволяет максимально охватить все основные сегменты рынка на внутренних и международных направлениях. Группа «Аэрофлот» стремится к лидерству среди глобальных перевозчиков. Помимо авиакомпаний в Группу «Аэрофлот» также входят профильные сервисные компании.

Авиакомпания «Аэрофлот» входит в число лидеров глобальной отрасли по качеству обслуживания пассажиров, что подтверждается высокими позициями в признанных международных рейтингах. Аэрофлот обладает рейтингом «четыре звезды» от британской консалтинговой компании Skytrax, восемь раз был удостоен премии Skytrax World Airline Awards в категории «Лучшая авиакомпания Восточной Европы», а также завоевал рейтинг «пять звезд» от американской авиационной ассоциации APEX. В 2019 году авиакомпания «Аэрофлот» получила высшую оценку «пять звезд» от британского агентства OAG за пунктуальность выполнения полетов, а также стала самой пунктуальной авиакомпанией в мире на основе глобального рейтинга On-Time Performance Review от аналитического холдинга Cirium.

Одно из ключевых преимуществ Группы «Аэрофлот» – современный парк воздушных судов. Парк авиакомпании «Аэрофлот» является одним из самых молодых в мире: средний возраст самолетов составляет пять лет (по Группе – семь лет). По состоянию на конец 2019 года парк Группы «Аэрофлот» насчитывал 359 воздушных судов, 245 из которых эксплуатировались авиакомпанией «Аэрофлот».

В 2019 году сеть маршрутов Группы «Аэрофлот» охватывала более 340 регулярных направлений в 57 стран мира. Авиакомпания «Аэрофлот» является участником глобального альянса SkyTeam, что позволяет расширить сеть маршрутов и предложить своим клиентам уникальный продукт глобального альянса.

У ПАО «Аэрофлот» 51,17% акций принадлежит Российской Федерации в лице Федерального агентства по управлению государственным имуществом. В свободном обращении у институциональных и розничных инвесторов находится 40,9% капитала.

Акции ПАО «Аэрофлот» торгуются на Московской бирже (тикер – AFLT) и включены в котировальный лист первого уровня. За пределами Российской Федерации акции Аэрофлота обращаются в виде глобальных депозитарных расписок на внебиржевом рынке Франкфуртской фондовой биржи.

Списочная численность персонала Группы «Аэрофлот» по состоянию на 31 декабря 2018 года составила 41 299 человек, увеличившись по сравнению с предыдущим годом на 5,8% (на 31 декабря 2017 года – 39 051 человек). Рост численности обусловлен увеличением парка воздушных судов и расширением сети маршрутов.

Списочная численность персонала ПАО «Аэрофлот» по состоянию на 31 декабря 2018 года выросла на 5,5%, до 24 261 человека (на 31 декабря 2017 года – 22 991 человек).

1. 1.

## Анализ внутренней среды предприятия

Организационно-управленческая структура аптеки представляет собой состав и соотношение ее внутренних звеньев, где каждое звено — это отдельный руководитель, либо другое лицо, занимающее какую-либо должность.

Рисунок 1 - Организационно – функциональная структура ПАО «Аэрофлот»

*Совет директоров* — это избираемый на определенный срок собранием акционеров орган управления, который осуществляет руководство деятельностью акционерного общества в период времени между ежегодными собраниями акционеров согласно своей компетенции, предоставляемой *совету директоров* по закону и по уставу.

Руководство текущей деятельностью Общества осуществляется единоличным исполнительным органом Общества — *Генеральным директором*. Исполнительные органы подотчетны Совету директоров и общему собранию акционеров. Генеральный директор Общества избирается Советом директоров на срок не более 5 лет. Правление действует на основании Устава Общества, а также Положения о Правлении, утвержденного общим собранием акционеров Общества.

*Генеральный директор*организовывает выполнение решений общего собрания акционеров и Совета директоров Общества, а также утверждает штатное расписание и представляет Общество во всех учреждениях.

*Комитет по кадрам и вознаграждениям* Совета директоров является консультативно совещательным органом, создается на основании решения Совета директоров и в своей деятельности руководствуется действующим законодательством, Кодексом корпоративного управления, Уставом Общества, Положением о Совете директоров Общества, настоящим Положением и иными внутренними документами Общества.

*Комитет по аудиту* — рабочий комитет публичной компании. Его члены из состава Совета директоров компании.

*Комитет по стратегии* — это консультативно-совещательный орган компании, который изучает и разрабатывает предложения по стратегии развития организации для совета директоров. Комитет возглавляется Председателем комитета, которого выбирают из числа членов комитета.

*Аппарат совета директоров* - подразделение, осуществляющее организацию работы совета директоров общества, а также осуществляющее контроль за исполнением решений совета директоров.

*Заместитель генерального директора по стратегии и маркетингу* - это специалист, который отвечает за маркетинговую политику компании, проводит маркетинговые исследования, разрабатывает программы сбыта, строит рекламную стратегию, управляет финансовыми ресурсами, называется директором отдела маркетинга.

*Заместитель генерального директора по коммерции и финансам* - это топ менеджер, который руководит закупками и продажами, маркетингом и логистикой компании.

*Главный бухгалтер* **-** лицо, возглавляющее в организации бухгалтерскую службу (бухгалтерию).

Анализируя работу ПАО «Аэрофлот» можно определить следующие цели функционирования. Для примера построили дерево целей.

Рисунок 2 – Дерево целей ПАО «Аэрофлот»

Данное «дерево» целей было построено в процессе анализа внутренней структуры авиакомпании, благодаря чему можно выявить следующие функциональные задачи.

Таблица 1 – Функциональные задачи авиакомпании

|  |  |
| --- | --- |
| Номер функциональной задачи | Содержание функциональных задач |
| 1 | Выполнение авиационных перевозок по международным и внутренним авиалиниям на регулярной и чартерной основе в соответствии с действующими международными договорами и лицензиями на эксплуатацию авиалиний, выдаваемых в установленном порядке |
| 2 | Обслуживание пассажиров, предоставление им различного рода услуг |
| 3 | Обслуживание грузов отправителей и получателей |
| 4 | Техническое обслуживание и ремонт воздушных судов |
| 5 | Продажа пассажирских и грузовых перевозочных документов |
| 6 | Издание, оформление и продажа пассажирских и грузовых перевозочных документов |

На основании таблицы 1 построим организационно-функциональную модель ПАО «Аэрофлот», в которой отобразим роль и обязанности сотрудников, по какой-либо функциональной задачи.

Таблица 2 – Матрица распределения зон ответственности предприятия

|  |  |
| --- | --- |
| Исполнители | Название функциональной задачи |
| 1 | 2 | 3 | 4 | 5 | 6 |
| Бортпроводники | + | • |  |  |  |  |
| Аэропортовые службы | - | - | + |  | • | + |
| Прочий персонал |  |  | • | + | - | - |
| Персонал технического обслуживания и ремонта |  |  |  | • |  |  |
| Летный состав | • | + | - |  |  |  |
| Продажа билетов и рекламы |  |  |  | - |  | • |

+ Ответственный за процесс;

• Основное участие в процессе;

- Частичное участие в процессе.

Можно сделать вывод, что каждый сотрудник в ПАО «Аэрофлот» играет важную роль во внутренней организационной структуре авиакомпании.

## Анализ внешней среды предприятия

Данный анализ будет проведет с помощью маркетингового инструмента, предназначенного для выявления политических, экономических, социальных и технологических аспектов внешней среды, который имеет название PEST-анализ.

Таблица 3 - PEST-анализ

|  |  |
| --- | --- |
| Политические факторы. | Экономические факторы. |
| Из-за политических конфликтов между Россией и зарубежными странами могут ввести санкции.Государственное финансирование способствует развитию компании.Государство устанавливает фиксированные тарифы на авиаперевозки для всей территории России. | Изменение основных издержек организации (затраты на энергоносители, услуги связи, коммунальные услуги, затраты на сырье и материалы и т.д.).Население России в зимний период чаще посещает теплые страны.Уровень цен и заработной платы в России способствует изменению пассажиропотока.Инфляция влечет за собой повышение цен на авиабилет. |
| Социально-культурные факторы. | Технологические факторы. |
| Демографические изменения напрямую влияют на авиакомпанию, так как большая плотность населения увеличивает пассажиропоток.СМИ может составлять рейтинги авиакомпаний, повышая или понижая конкурентоспособность. | Появление новых самолетов с большим количеством мест влияет на себестоимость авиабилета.Влияние интернета обеспечивает быстрый доступ для покупки авиабилета.Новый вид топлива позволяет уменьшить расход и выброс вредных веществ в окружающую среду. |

Данный рынок является потенциально располагающим для продукции компании. Особое влияние стоит обратить на технологические факторы, они в большей степени способны повлиять на развитие компании.

# Разработка проекта по внедрению системы интеллектуального анализа данных на предприятии

## Разработка сетевого графика

Ниже приведены этапы создания информационно-аналитической системы. Также предоставлены желаемые результаты и продолжительность работ.

На стадии «Организация проекта» присутствуют такие этапы работы, как заключение контракта на разработку системы, согласование устава проекта и окончательное формирование команды. Примерная общая продолжительность данных работ 24 дня.

На стадии «Формирование требований и разработка технического задания» проводится интервью с функциональными специалистами заказчика, то есть происходит сбор информации о возможных системах-источников данных и сетевом окружении. Далее анализируется и обрабатывается полученная информация для утверждения технического задания. Примерная общая продолжительность данных работ 52 дня.

Поскольку на стадии «Эскизный проект» происходит выбор программного обеспечения и разработка предварительного регламента взаимодействия информационных систем, то результатом будет утверждение плана формирования технической инфраструктуры. Примерная общая продолжительность данных работ 31 день.

Проектирование интерфейсов пользователя и оформление технического проекта-основные этапы работы на стадии «Технический проект» и как следствие согласование и утверждение технического проекта. Примерная общая продолжительность данных работ 53 дня.

Предпоследняя стадия «Рабочая документация» содержит в себе разработку или адаптацию программ, то есть создание необходимых объектов базы данных, разработку процессов ETL и процессов обеспечения качества данных и реализацию витрин данных. Примерная общая продолжительность данных работ 82 дня.

В заключительной стадии «Ввод в действие и сопровождение» проводится подготовка объекта автоматизации к вводу системы в действие и предварительные испытания. По завершению работ подписывается акт. Далее в течении гарантийного года выявляют недостатки и устраняют их. Примерная общая продолжительность данных работ 396 дней.

Таблица 4 − Разработка проекта по внедрению интеллектуального анализа данных в ПАО «Аэрофлот»

|  |  |  |
| --- | --- | --- |
| **Наименование работ** | **События**  | **Продолжительность работ, дни** |
| Заключение контракта | 1,2 | 11 |
| Согласование процедур управления | 1,3 | 7 |
| Сбор команды проекта | 1,4 | 6 |
| Системно-аналитическое обследование объекта автоматизации | 2,5 | 21 |
| Анализ и обработка полученной информации | 3,5 | 19 |
| Согласование и утверждение | 4,5 | 12 |
| Обоснование выбора программного обеспечения и технической инфраструктуры | 5,6 | 8 |
| Разработка предварительного регламента взаимодействия информационных систем | 5,7 | 23 |
| Разработка и согласование регламентов взаимодействия информационных систем | 6,8 | 10 |
| Проектирование интерфейсов пользователя | 7,8 | 16 |
| Продолжение таблицы 4 |
| Оформление технического проекта | 8,9 | 21 |
| Согласование и утверждение | 9,1 | 6 |
| Разработка или адаптация программ | 10,11 | 66 |
| Согласование и утверждение | 10,12 | 16 |
| Подготовка объекта автоматизации к вводу системы в действие | 11,13 | 14 |
| Проведение предварительных испытаний | 12,13 | 10 |
| Завершение работ | 13,14 | 6 |
| Выполнение работ в соответствии с гарантийными обязательствами | 14,15 | 366 |

Для наглядного примера был составлен сетевой график, который схематично показывает длительность предстоящих работ и количество событий.



Рисунок 3 – Сетевой график

На графике видно, что самый трудоемкий и долгий этап разработки проекта это событие под номером 15 «Выполнение работ в соответствии с гарантийными обязательствами».

## Анализ резервов времени

Для того, чтобы понять насколько можно сократить время некоторых работ или точно понять где требуется больше времени, чем запланировано, составили таблицу резерва.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Работы | Количество предшествующих работ | Продолжительность | Ранние сроки начало | Ранние сроки окончание | Поздние сроки начало | Поздние сроки окончание | Резервы времени полный | Независимый резерв времени | Свободный резерв времени |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | -  |
| 1,2 | 0 | 11 | 0 | 11 | 0 | 11 | 0 | 0 | 0 |
| 1,3 | 0 | 7 | 0 | 7 | 6 | 13 | 6 | 6 | 6 |
| 1,4 | 0 | 6 | 0 | 6 | 14 | 20 | 14 | 14 | 14 |
| 2,5 | 1 | 21 | 11 | 32 | 11 | 32 | 0 | 0 | 0 |
| 3,5 | 1 | 19 | 7 | 26 | 13 | 32 | 6 | 6 | 6 |
| 4,5 | 1 | 12 | 6 | 18 | 20 | 32 | 14 | 14 | 14 |
| 5,6 | 3 | 8 | 32 | 40 | 53 | 61 | 21 | 21 | 21 |
| 5,7 | 3 | 23 | 32 | 55 | 32 | 55 | 0 | 0 | 0 |
| 6,8 | 1 | 10 | 40 | 50 | 61 | 71 | 21 | 21 | 21 |
| 7,8 | 1 | 16 | 55 | 71 | 55 | 71 | 0 | 0 | 0 |
| 8,9 | 2 | 21 | 71 | 92 | 71 | 92 | 0 | 0 | 0 |
| 9,10 | 1 | 6 | 92 | 98 | 92 | 98 | 0 | 0 | 0 |
| 10,11 | 1 | 66 | 98 | 164 | 152 | 218 | 54 | 54 | 54 |
| 10,12 | 1 | 16 | 98 | 114 | 98 | 114 | 0 | 0 | 0 |
| 11,13 | 1 | 14 | 164 | 178 | 164 | 178 | 0 | 0 | 0 |
| 12,13 | 2 | 10 | 114 | 124 | 168 | 178 | 54 | 54 | 54 |
| 13,14 | 1 | 6 | 178 | 184 | 178 | 184 | 0 | 0 | 0 |
| 14,15 | 1 | 366 | 184 | 550 | 184 | 550 | 0 | 0 | 0 |

Таблица 5 – Резервы времени

Данные из этой таблицы показали, что самый большой резерв времени составляет 54 дня у работ 10,11 и 12,13. Из этого следует, что времени на этих этапах понадобиться меньше, а значит разработка проекта ускорится.

По ниже предоставленной диаграмме можно отследить даты начала каждой работы. А также ознакомиться с этапом каждой из работ.

Рисунок 4­ – Диаграмма Ганта

Разработка проекта по моим расчетам должна начаться 06 июня 2020 года и закончиться 27 мая 2022 года. Конечно, здесь не брались в расчет случайные ситуации, провалы и неудачи. Поэтому в настоящем времени, график следует адаптировать под реальную жизнь.

## Перспективы развития систем интеллектуального анализа данных

Основным фактором, определяющим развитие технологий искусственного интеллекта сегодня, является скорость роста вычислительной мощности компьютеров, поскольку принципы человеческой психики до сих пор неясны (на уровне детализации, доступной для моделирования). Таким образом, тема ИИ-конференций кажется достаточно стандартной, и с точки зрения состава она в течение некоторого времени практически не менялась. Но увеличение производительности современных компьютеров в сочетании с повышением качества алгоритмов периодически дает возможность применять на практике различные научные методы.

Временно забытые методы расчета простых опций снова будут интенсивно развиваться, что обойдется в чрезвычайно упрощенным описанием объектов. Но благодаря такому подходу, можно, как и ожидалось, решить множество самых разных проблем. Относительно простые, но ресурсоемкие алгоритмы адаптивного поведения помогут автономным устройствам уверенно работать в сложном мире. В то же время цель заключается в разработке систем, которые внешне не похожи на людей, но действуют как люди.

Область ИИ, ставшая зрелой наукой, развивается постепенно - медленно, но неуклонно вперед. В результате результаты вполне предсказуемы, хотя о внезапных прорывах, связанных со стратегическими инициативами, не может быть и речи. Например, национальная компьютерная инициатива Соединенных Штатов более 30 лет назад вывела из лаборатории много направлений искусственного интеллекта и оказала значительное влияние на развитие теории высокопроизводительных вычислений и ее применение во многих прикладных проектах.

# ЗАКЛЮЧЕНИЕ

Данная курсовая работа ориентирована на применение интеллектуального анализа. В наше время все чаще большие компании стараются использовать искусственный интеллект, не только для экономии на сотрудниках, но и для более быстрого и успешного функционирования, ведь интеллектуальный анализа данных адаптирован практически под любой запрос пользователя.

Во время сбора данных о ПАО «АЭРОФЛОТ», оказалось, что данная авиакомпания всегда была заинтересована внедрением новых технологий, которые помогут улучшить качество своих полетов. Для такой цели был разработан проект по внедрению системы интеллектуального анализа данных в ПАО «АЭРОФЛОТ».

В работе были применены некоторые из множества методов системного анализа, такие как дерево целей и сетевой график, которые были рассмотрены во 2 и 3 главах. Данные методы являются наиболее распространенными и актуальными по сегодняшний день. По итогу системного анализа мы имеем масштабные цели авиакомпании и достаточный резервы времени для выполнения разработанного проекта. ПАО «АЭРОФЛОТ» рекомендуется использовать разработанный проект в данной курсовой работе для внедрения системы интеллектуального анализа данных.

# СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Боровиков, В.П. Популярное введение в современный анализ данных в системе STATISTICA. Учебное пособие для вузов. +CD / В.П. Боровиков. — М.: РиС, 2015. — 288 c.
2. Кабаков, Р. R в действии. Анализ и визуализация данных в программе R / Р. Кабаков. - М.: ДМК, 2016. - 588 c.
3. Кабаков, Р. R в действии. Анализ и визуализация данных в программе R / Р. Кабаков. — М.: ДМК, 2016. — 588 c.
4. Козлов, А.Ю. Статистический анализ данных в MS Excel: Учебное пособие / А.Ю. Козлов, В.С. Мхитарян, В.Ф. Шишов. - М.: Инфра-М, 2018. - 80 c.
5. Куприянов, М.С. Анализ данных и процессов. Учебное пособие (+ CD-ROM) / М.С. Куприянов. - М.: БХВ-Петербург, 2015. - 567 c.
6. Лесковец, Ю. Анализ больших наборов данных / Ю. Лесковец, А. Раджараман. - М.: ДМК, 2016. - 498 c.
7. Маккинли, У. Python и анализ данных / У. Маккинли. - М.: ДМК, 2015. - 482 c.
8. Макшанов, А.В. Технологии интеллектуального анализа данных: Учебное пособие / А.В. Макшанов, А.Е. Журавлев. - СПб.: Лань, 2018. - 212 c.
9. Марина, Юрьевна Архипова Анализ данных. Учебник для академического бакалавриата / Марина Юрьевна Архипова. - М.: Юрайт, 2016. - 931 c.
10. Марчук, Г.И. Геронтология in silico. Становление новой дисциплины. Математические модели, анализ данных и вычислительные эксперименты / Г.И. Марчук. - М.: Бином. Лаборатория знаний, 2017. - 108 c.
11. Мастицкий, С.Э. Статистический анализ и визуализация данных с помощью R (черно-белые графики) / С.Э. Мастицкий. — М.: ДМК, 2015. — 496 c.
12. Миркин, Б.Г. Введение в анализ данных. учебник и практикум / Б.Г. Миркин. — Люберцы: Юрайт, 2016. — 174 c.
13. Официальный сайт компании ПАО «Аэрофлот» (электронный ресурс) <https://www.aeroflot.ru/ru-ru/>
14. Просветов, Георгий Иванович Анализ данных с помощью EXCEL. Задачи и решения. Учебно-практическое пособие / Просветов Георгий Иванович. - М.: Альфа-пресс, 2016. - 21 c.
15. Рафалович, В. Data mining, или интеллектуальный анализ данных для занятых. Практический курс / В. Рафалович. - М.: SmartBook, 2018. - 352 c.
16. Рафалович, В. Data mining, или интеллектуальный анализ данных для занятых. Практический курс / В. Рафалович. - М.: SmartBook, 2018. - 352 c.
17. Сирота, А.А. Методы и алгоритмы анализа данных и их моделирование в MATLAB / А.А. Сирота. - СПб.: BHV, 2016. - 384 c.
18. Тюрин, Ю. Н. Анализ данных на компьютере. Учебное пособие / Ю.Н. Тюрин, А.А. Макаров. - М.: МЦНМО, 2016. - 368 c.
19. Тюрин, Ю.Н. Анализ данных на компьютере / Ю.Н. Тюрин, А.А. акаров. - М.: МЦНМО, 2016. - 368 c.
20. Уэс, Маккинни Python и анализ данных. Руководство / Маккинни Уэс. - М.: ДМК Пресс, 2015. - 185 c.
21. Финков, М.В. Вычисления, графики и анализ данных в Excel 2013. Самоучитель / М.В. Финков. - М.: Наука и техника, 2015. - 326 c.
22. Форман, Д. Много цифр: Анализ больших данных при помощи Excel / Д. Форман. - М.: Альпина Паблишер, 2017. - 461 c.
23. Форман, Дж. Много цифр: Анализ больших данных при помощи Excel / Дж. Форман. - М.: Альпина Паблишер, 2019. - 461 c.
24. Intuit.ru (интернет-источник) интеллектуальный анализ https://www.intuit.ru/studies/courses/6/6/info
25. Studme.org (интернет-источники) <https://studme.org/66918/ekonomika/strategicheskie_aspekty_mezhdunarodnogo_biznesa>

