

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

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

Кафедра «Прикладная математика и информатика»
(наименование кафедры)

09.03.03 Прикладная информатика
(код и наименование направления подготовки)

Бизнес-информатика
(направленность (профиль) / специализация)

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (БАКАЛАВРСКАЯ РАБОТА)

на тему «Разработка информационной системы для контроля и анализа
эффективности продвижения блога в социальной сети»

Обучающийся

Е.И. Санкович

(Инициалы Фамилия)

(личная подпись)

Руководитель

к.т.н, В.С. Климов

(ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)

Тольятти 2022

Аннотация

Тема бакалаврской работы – «Разработка информационной системы для контроля и анализа эффективности продвижения блога в социальной сети». В настоящее время ведение корпоративных блогов в социальных сетях используется различными компаниями для продвижения своих товаров и услуг. Для увеличения количества просмотров блога, а следовательно, и для роста количества потенциальных клиентов, компании заказывают мероприятия по их продвижению в маркетинговых агентствах. Однако не всегда работа маркетинговых агентств приводит к желаемому результату. Актуальность работы заключается в том, чтобы научиться оценивать эффективность продвижения корпоративного блога на начальном этапе запуска маркетинговой стратегии.

Объектом исследования бакалаврской работы является оценка эффективности продвижения корпоративного блога.

Предметом исследования бакалаврской работы является система для контроля и анализа эффективности продвижения блога.

Цель выпускной квалификационной работы – разработка информационной системы для контроля и анализа эффективности продвижения блога в социальной сети.

Методы исследования – технологии проектирования информационных систем, технологии программирования.

Данная работа состоит из введения, трех глав, заключения и списка используемой литературы.

В первой главе работы описываются критерии эффективности продвижения блога, во второй главе приводятся методы прогнозирования роста просмотров блога, в третьей главе описана практическая реализация системы для анализа эффективности продвижения блога.

Бакалаврская работа состоит из 43 страниц текста, 26 рисунков, 1 таблицы и 20 источников.

Оглавление

Введение.....	4
Глава 1 Анализ проблематики продвижения корпоративного блога	7
1.1 Обзор особенностей продвижения корпоративного блога.....	7
1.2 Моделирование процесса продвижения блога	11
Глава 2 Проектирование системы для прогнозирования эффективности продвижения корпоративного блога	17
2.1 Алгоритм использования разрабатываемой системы	17
2.2 Анализ временных рядов	19
Глава 3 Разработка приложения для классификации текста	22
3.1 Особенности реализации программного модуля.....	22
3.3. Результаты тестирования приложения	32
Заключение	38
Список используемой литературы и используемых источников.....	40

Введение

В данных отчета Global Digital 2022 говорится, что на протяжении последнего десятилетия от года к году растет количество пользователей сети Интернет, а также суммарное время, проводимое ими в социальных сетях. Популярность социальных сетей привела к тому, что многие компании рассматривают данные сервисы, как площадки для продвижения своих товаров и услуг. Поэтому большинство компаний открывает корпоративные блоги в социальных сетях, надеясь привлечь как можно большую аудиторию [4]. Распространение социальных сетей привело к тому, что большинство компаний для продвижения своих товаров и услуг создают в них корпоративные блоги [5]. Компании заинтересованы в том, чтобы размещаемый ими контент в социальных сетях получил как можно большее количество просмотров [6], [7].

С коммерческой точки зрения рост числа подписчиков и суммарного количества просмотров корпоративного блога являются необходимым условием для увеличения количества потенциальных клиентов предлагаемой продукции среди пользователей социальной сети [3].

Стандартным способом увеличения количества просмотров и подписчиков является процесс продвижения блога. Интенсивность этого процесса зависит от выделяемого бюджета и может включать в себя, например:

- выкуп рекламы по ключевым словам в поисковой строке;
- покупка рекламы в других блогах;
- продвижение посредством встроенных инструментов социальной сети;
- размещение баннеров.

Эффективность продвижения корпоративного блога оценивается в приросте количества подписчиков и просмотров. Неэффективное

продвижение блога приводит к трате бюджета и не обеспечивает прирост описанных выше показателей. Проблема заключается в том, что обычно эффективность продвижения блога оценивается по факту, когда весь выделенный на эти цели бюджет израсходован. На решение этой проблемы и направлены исследования, представленные в данной выпускной квалификационной работе.

В бакалаврской работе предлагается разработать информационную систему, которая на основе данных, полученных через некоторое время после старта продвижения корпоративного блога, могла бы прогнозировать конечный результат на заданную дату, выраженный в виде количества новых подписчиков и просмотров. И если прогноз такой системы будет свидетельствовать о низкой эффективности текущей стратегии продвижения, то ее можно будет скорректировать до момента расходования выделенного на продвижение бюджета.

Цель выпускной квалификационной работы – разработка информационной системы для контроля и анализа эффективности продвижения блога в социальной сети.

Для достижения данной цели необходимо выполнить следующие задачи:

- моделирование процесса продвижения корпоративного блога;
- проектирование системы анализа эффективности продвижения блога в социальной сети;
- реализация и тестирование системы анализа эффективности продвижения блога.

Методы исследования – методы и технологии проектирования информационных систем, технологии программирования.

Практическая значимость бакалаврской работы заключается в разработке программного продукта, используемого при маркетинговых исследованиях и способного прогнозировать изменение количества

подписчиков и просмотров корпоративного блога.

Данная работа состоит из введения, трех глав, заключения, списка используемой литературы и приложений.

В первой главе приводится описание исследований о росте роли социальных сетей, объясняется мотивация бизнеса при принятии решения об открытии корпоративного блога, описывается роль процесса продвижения при расширении аудитории корпоративного блога, проводится выбор оценки эффективности продвижения блога, моделируется процесс продвижения блога в нотации IDEF0.

Вторая глава посвящена проектированию системы для прогнозирования эффективности продвижения корпоративного блога и анализу алгоритмов прогнозирования.

В третьей главе представлен процесс разработки программного обеспечения, а также приведены результаты его тестирования на реальных данных, полученных из сервиса SocialBlade.

В заключении описываются результаты выполнения выпускной квалификационной работы.

Бакалаврская работа состоит из 43 страниц текста, 26 рисунков, 1 таблицы и 20 источников.

Глава 1 Анализ проблематики продвижения корпоративного блога

1.1 Обзор особенностей продвижения корпоративного блога

По данным отчета Global Digital 2022, составленным совместно агентством «We Are Social» и сервисом «Hootsuite» - более 62% мирового населения являются пользователями глобальной сети Интернет (рисунок 1). Прирост к предыдущему году составил 4%. Также прогнозируется, что пользователи в 2022 году проведут в глобальной сети около 12,5 триллионов часов.

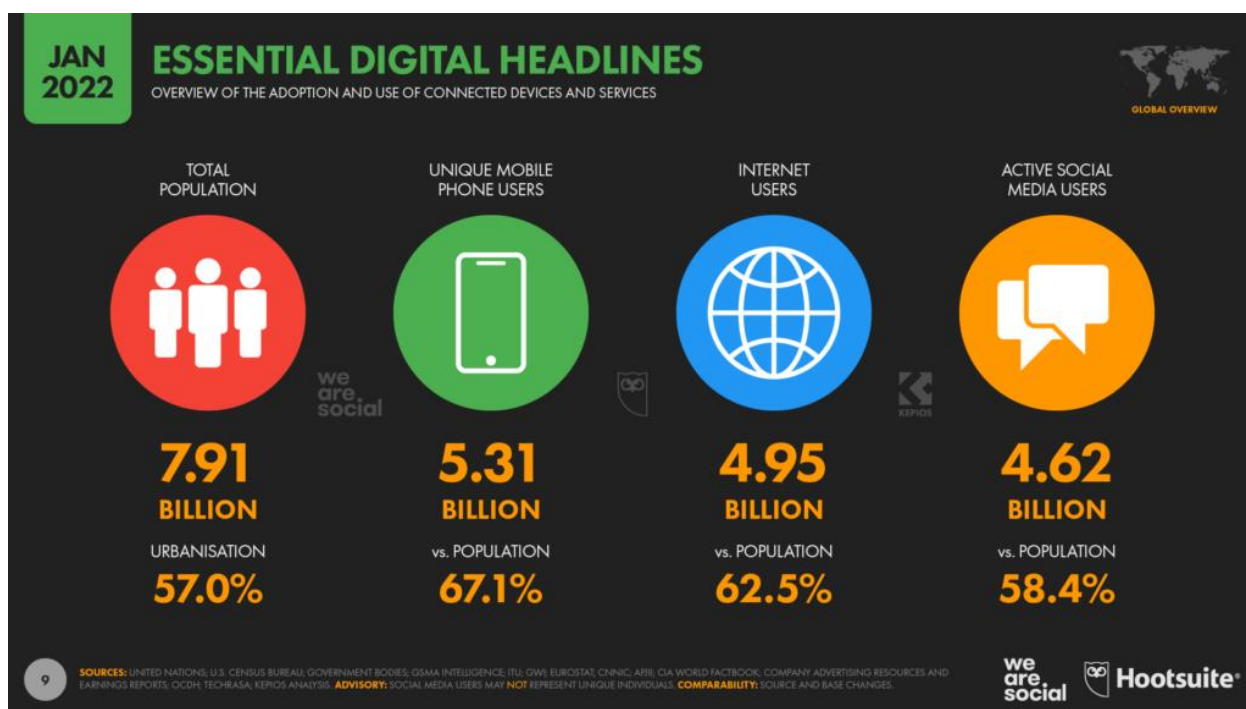


Рисунок 1 – Фрагмент отчета Global Digital 2022 с информацией о пользователях Интернет

Отдельно заслуживает внимание раздел отчета Global Digital 2022 посвященный росту популярности социальных сетей среди пользователей. Так по представленным данным видно, что начиная с 2012 года количество пользователей социальных сетей увеличилось более, чем в 3 раза (рисунок 2).

С учетом того, что в среднем каждый пользователь проводит в социальных сетях 147 минут в день, то суммарный охват и продолжительность использования социальных сервисов выглядят впечатляюще.

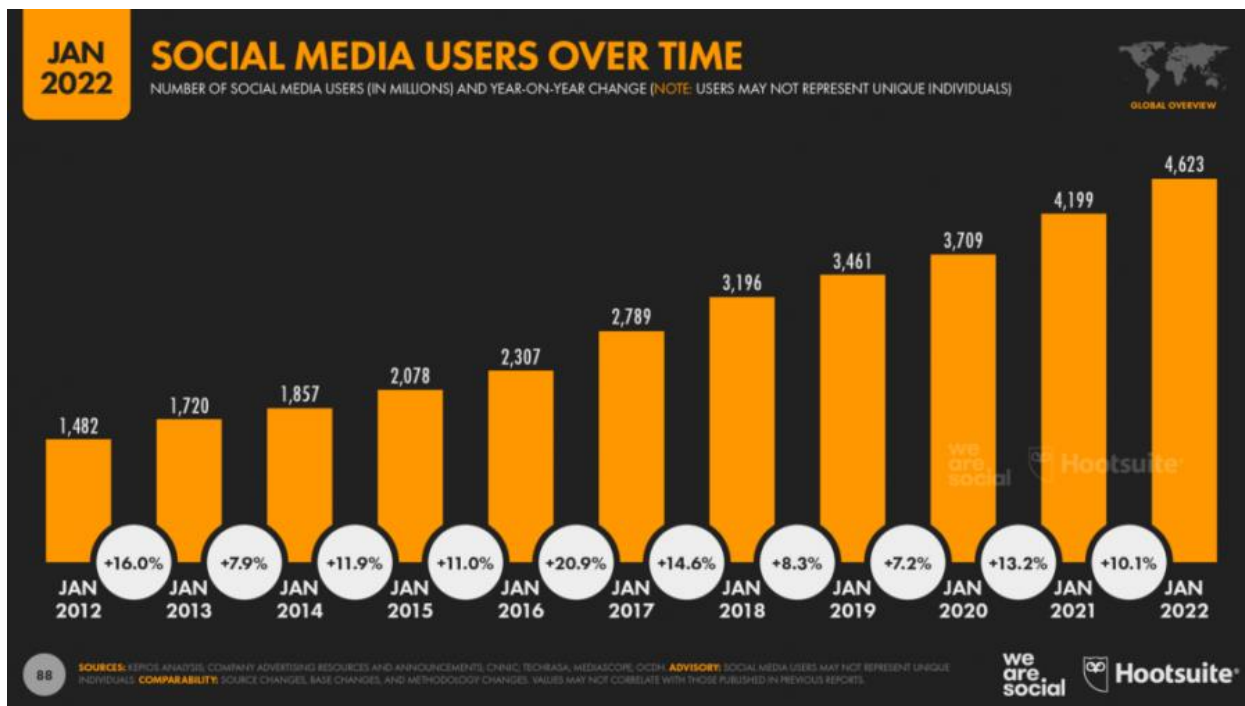


Рисунок 2 – Фрагмент отчета Global Digital 2022 изменение количества пользователей социальных сетей

Наиболее популярными в мире социальными сетями, не запрещенными на территории России являются YouTube, Tiktok, Telegram. В России наиболее популярная социальная сеть – ВКонтакте.

Бизнес-сообщество вынуждено реагировать на рост популярности социальных сетей путем размещения своих официальных каналов. Практически у каждой, даже небольшой, компании есть аккаунты в социальных сетях, которые используются для распространения информации о предоставляемых продуктах и сервисах.

Также как и в реальном мире, в социальных сетях компании борются за внимание пользователей. Рост подписчиков корпоративного блога компании необходим, так как каждый подписчик рассматривается компанией, как

потенциальный клиент будущих товаров и услуг [9]. Существует корреляция между количеством подписчиков на корпоративный блог компании и количеством ее клиентов. Также важным показателем корпоративного блога для компании является количество просмотров. По этому показателю компания может оценивать осведомленность аудитории о предлагаемых услугах и товарах.

Помимо увеличения охвата аудитории (количества подписчиков) существуют и другие уровни формирования целей при продвижении блога, показанные на рисунке 3. К целям первого уровня относится изменение положения компании на рынке, увеличение уровня продаж, увеличение прибыли. К целям продвижения второго уровня относится изменение потребительской базы, привлечение новых клиентов, изменение привычек использования товара и повышение лояльности к продукту. Цели третьего уровня – формирование потребностей в товаре, повышение узнаваемости товара или услуги, формирование отношения к товару или услуге, стимулирование приобретения товара.

Отдельно стоит отметить, что корпоративные блоги обязаны вести различные государственные органы. Это связано с тем, что председатель Правительства РФ Михаил Мишустин 2 сентября подписал распоряжение, согласно которому государственные органы, органы местного самоуправления, подведомственные организации и суды с 1 декабря 2022 года будут обязаны вести публичные блоги в социальных сетях. Таким образом, государство принимает важность социальных сетей и необходимость присутствия в них для обеспечения гласности. В этом случае показатели по количеству подписчиков и просмотров могут стать элементами КРІ (ключевых показателей эффективности) работы государственного органа.



Рисунок 3 – Детализация 4 уровней формирования целей при продвижении корпоративного блога

Теперь перейдем к моделированию процесса продвижения блога и рассмотрению его особенностей.

1.2 Моделирование процесса продвижения блога

Моделирование процесса продвижения блога будем осуществлять с использованием программы ERwin Process Modeler.

Входными параметрами процесса продвижения корпоративного блога являются:

- потребность в продвижении блога;
- бюджет на продвижение;
- данные блога.

Потребность в продвижении корпоративного блога описывается целями, детализация которых показана на рисунке 3. В зависимости от формулирования потребности выбираются инструменты продвижения, которые фиксируются в техническом задании. Размер бюджета влияет на то, какие методы продвижения будут использованы. Данные блога включают в себя описание его тематики и ключевые особенности.

Процесс продвижения блога регулируется законами РФ, а также правилами пользования социальной сети, в которой находится блог.

Выполнение процесса продвижения блога выполняет персонал, который может как входить в состав компании, владеющей корпоративным блогом, так и быть независимой отдельной компанией, выполняющей маркетинговые услуги. При выполнении продвижения блога используются различные материально технические средства, например компьютеры.

В результате выполнения продвижения блога на выходе мы получаем определенное количество новых подписчиков и увеличение количества просмотров.

Построенная модель AS-IS анализируемого процесса показана на рисунке 4.

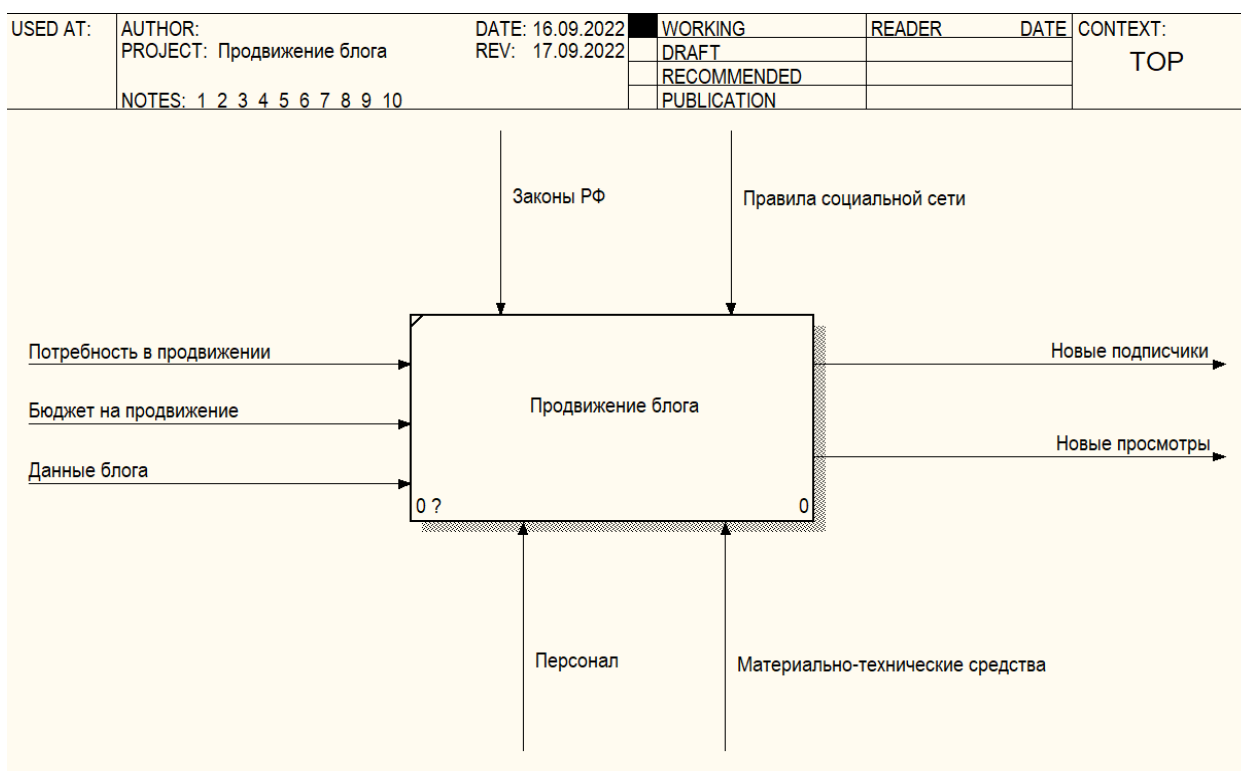


Рисунок 4 – Модель AS-IS, уровень А-0

Проведем декомпозицию процесса «продвижение блога». В общем случае, для продвижения блога необходимо выполнить следующие шаги:

- формирование технического задания;
- реализация маркетинговой стратегии.

В ходе формирования технического задания определяется длительность маркетинговой стратегии, реализуемой в рамках продвижения корпоративного блога, используемые инструменты, план расходования бюджета и цели продвижения. На данном этапе создается техническое задание, в соответствии с которым реализуется маркетинговая стратегия.

Результат декомпозиции процесса «продвижение блога» показан на рисунке 5.

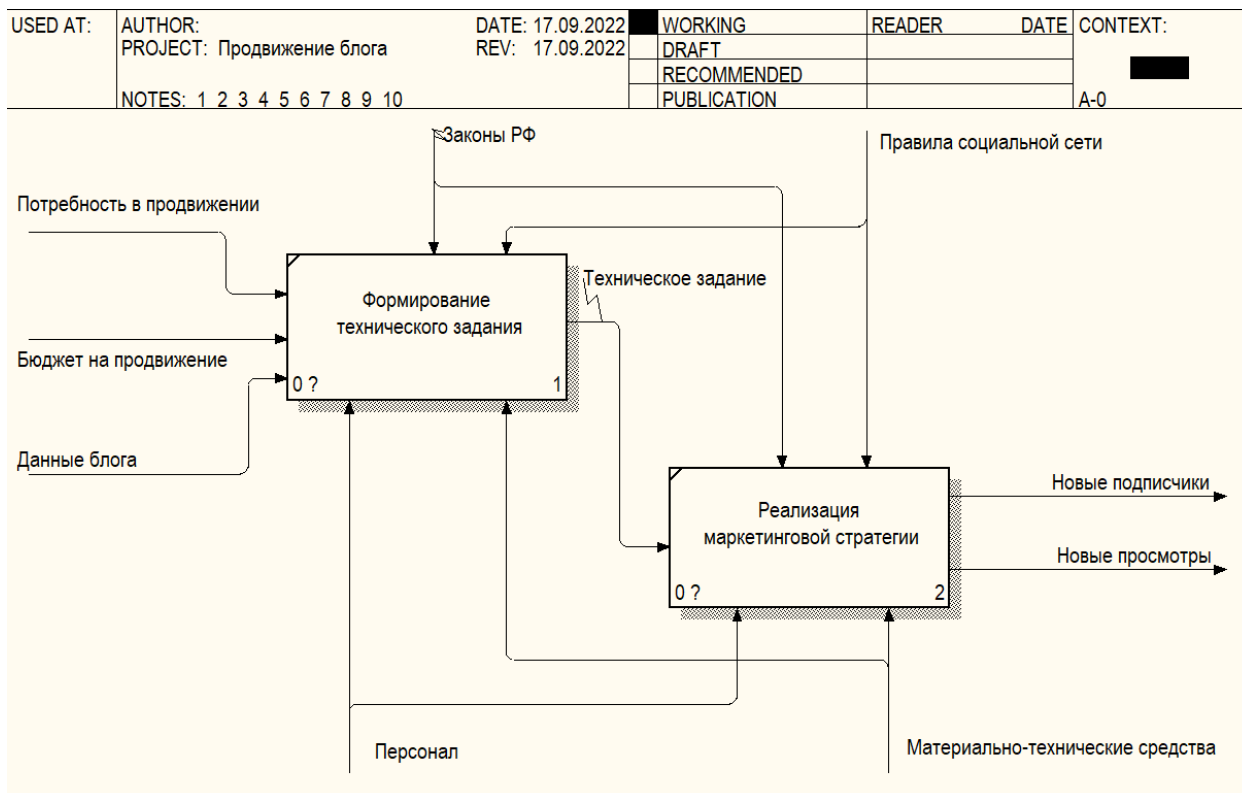


Рисунок 5 – Модель AS-IS, декомпозиция

При реализации процесса «продвижение блога» по описанной модели возникает следующая проблема: определение эффективности выбранной маркетинговой стратегии, выраженной в приросте количества подписчиков и просмотров, становится известно только после завершения продвижения блога.

Предложено модифицировать процесс «продвижение блога» добавив в него возможность контролирования промежуточных результатов реализации маркетинговой стратегии. Для предложено разработать информационную систему, которая могла бы прогнозировать конечное количество подписчиков и просмотров еще до окончания процесса продвижения блога.

С учетом выше сказанного модель TO-BE процесса «продвижение блога» будет выглядеть так, как это показано на рисунке 6.

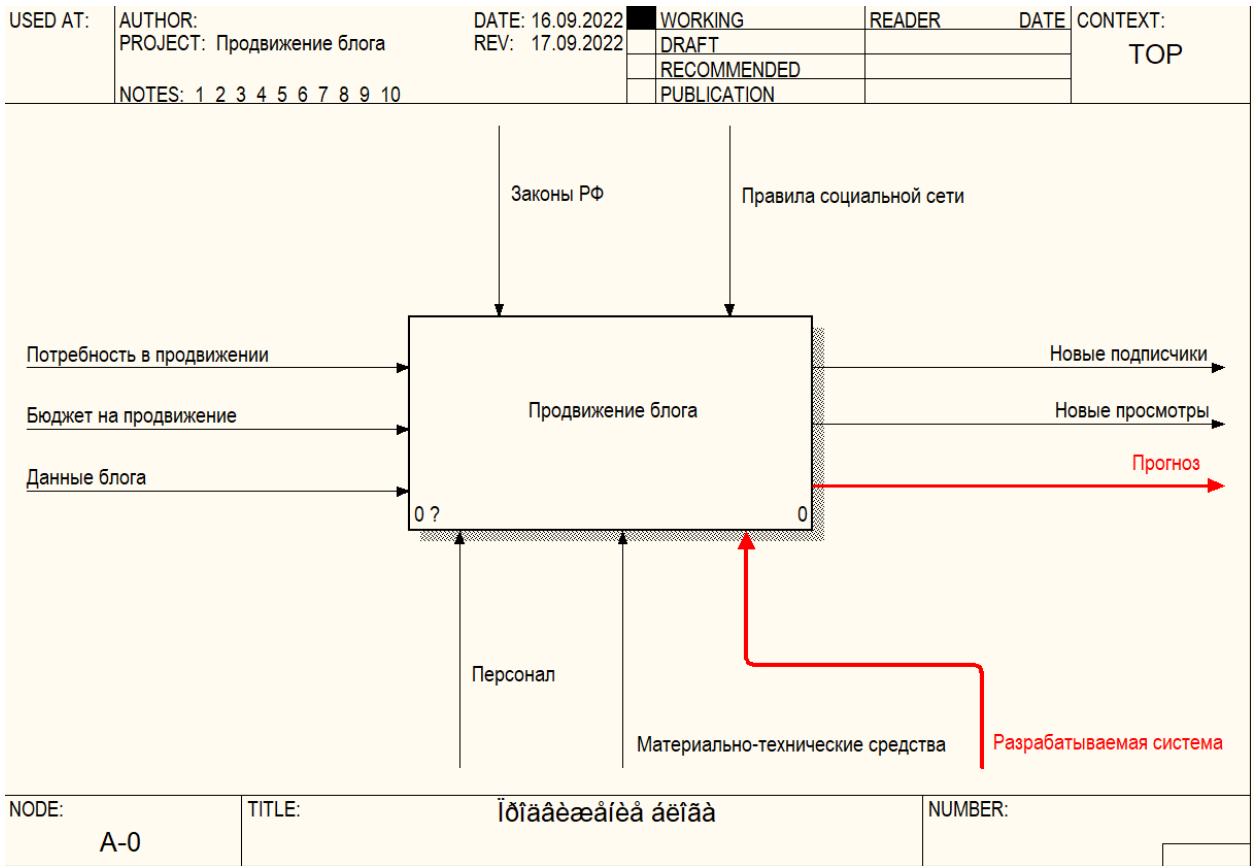


Рисунок 6 – Модель TO-BE, уровень А-0

Разрабатываемая система будет выполнять анализ результатов маркетинговой стратегии на основе промежуточных результатов по приросту новых просмотров и подписчиков, полученных от реализации маркетинговой стратегии. В результате анализа система будет прогнозировать конечный результат маркетинговой стратегии, выполняемой по текущему техническому заданию.

Если прогнозируемый результат текущей маркетинговой стратегии оказывается недостаточным, то производится корректировка технического задания. Таким образом текущая маркетинговая стратегия корректируется на основе промежуточных результатов по приросту подписчиков и просмотров.

Предложенный подход позволяет циклично выполнять корректировку технического задания до тех пор, пока прогнозируемый системой результат

продвижения корпоративного блога не начнет соответствовать ожиданиям. При этом прогноз выдаваемой системой будет обновляться при поступлении актуальных данных об изменении количества подписчиков и просмотров.

Декомпозиция Модель ТО-ВЕ с учетом предлагаемых изменений показана на рисунке 7.

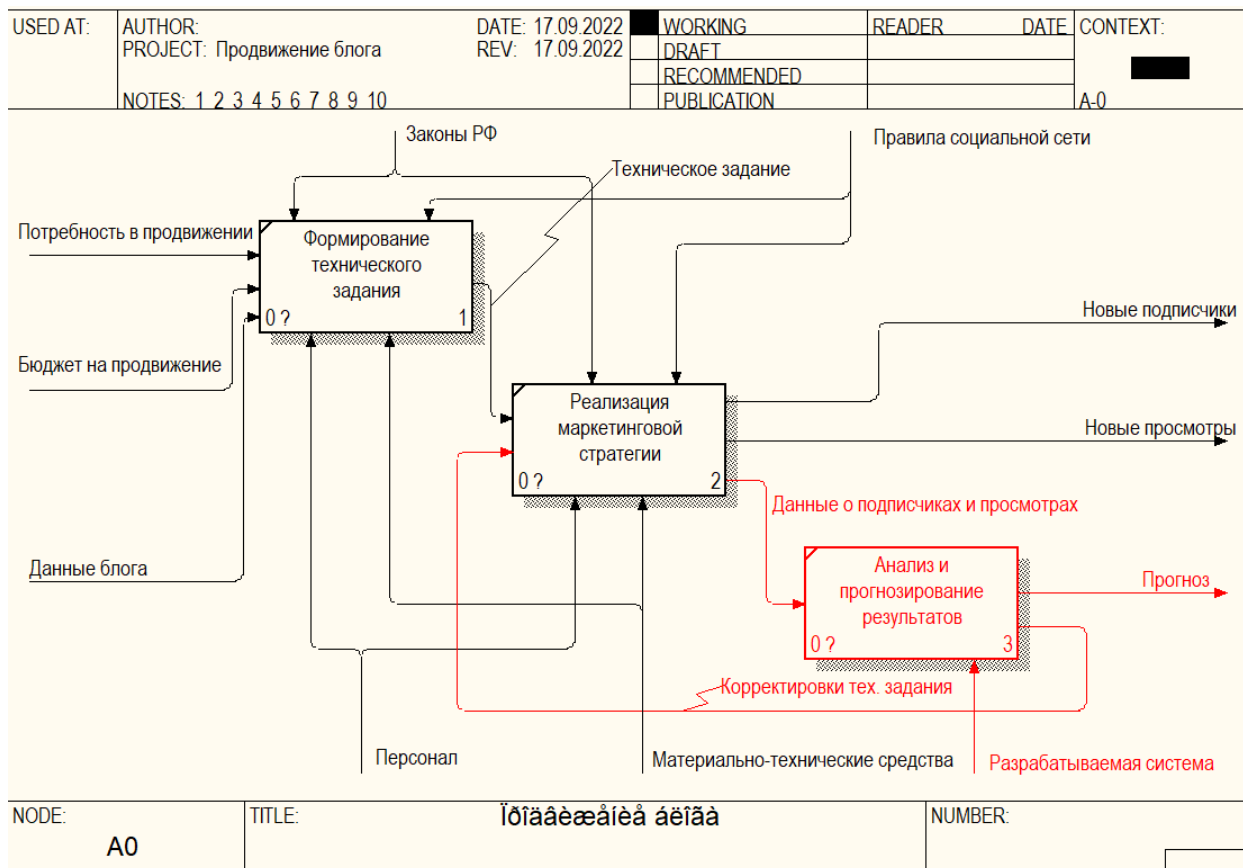


Рисунок 7 – Модель ТО-ВЕ, декомпозиция

После моделирования процесса «продвижение блога» и разработки способа по его улучшению перейдем к проектированию и реализации предложенной системы для анализа и прогнозирования результатов реализации маркетинговой стратегии.

Выводы по главе 1

Приведем выводы по первой главе бакалаврской работы:

- в ходе анализа литературных данных установлено, что за последние 10 лет количество пользователей социальных сетей увеличилось более чем в 3 раза и продолжает расти;
- увеличение роли социальных сетей привело к росту в них корпоративных блогов, цель ведения которых заключается в продвижении среди пользователей своих товаров и услуг;
- для продвижения своих товаров и услуг как можно большему числу пользователей владельцы корпоративных блогов стремятся увеличить количество своих подписчиков и число просмотров, для этого на основе выбранной маркетинговой стратегии выполняется продвижение корпоративного блога;
- с использованием методологии IDEF0 проведено функциональное моделирование процесса продвижения блога, в результате которого установлено, что данный процесс можно модернизировать за счет разработки и внедрения системы для анализа и прогнозирования результатов продвижения блога.

Глава 2 Проектирование системы для прогнозирования эффективности продвижения корпоративного блога

2.1 Алгоритм использования разрабатываемой системы

Как уже было сказано ранее – основными показателями корпоративного блога является количество подписчиков и просмотров. Для того чтобы произвести оценку эффективности продвижения блога необходимо получить численную оценку того, как продвижение повлияло на изменение этих показателей. С учетом вышесказанного, предложен следующий алгоритм оценки эффективности продвижения.

На первом этапе составляется прогноз на требуемую дату по естественному приросту количества просмотров и подписчиков, не связанных с реализацией маркетинговой стратегии.

На втором этапе компания приступает к реализации продвижения корпоративного блога, фиксируя фактические изменения количества подписчиков и просмотров. Затем рассчитывается второй прогноз на основе данных, полученных в результате продвижения блога.

На третьем этапе для оценки эффективности продвижения блога на будущую дату производится сравнение двух прогнозов, полученных на первом и втором этапе. Чем сильнее они отличаются, тем эффективней маркетинговая стратегия, выбранная для продвижения блога.

Чтобы реализовать предложенный алгоритм была спроектирована следующая схема работы программного обеспечения (рисунок 8). Пользователь разрабатываемой системы вводит данные блога в информационную систему анализа, система отправляет запрос на получение исторических данных по количеству подписчиков и просмотров в сервис Social blade. Сервис Social blade предоставляет запрошенную статистику в виде данных временных рядов.

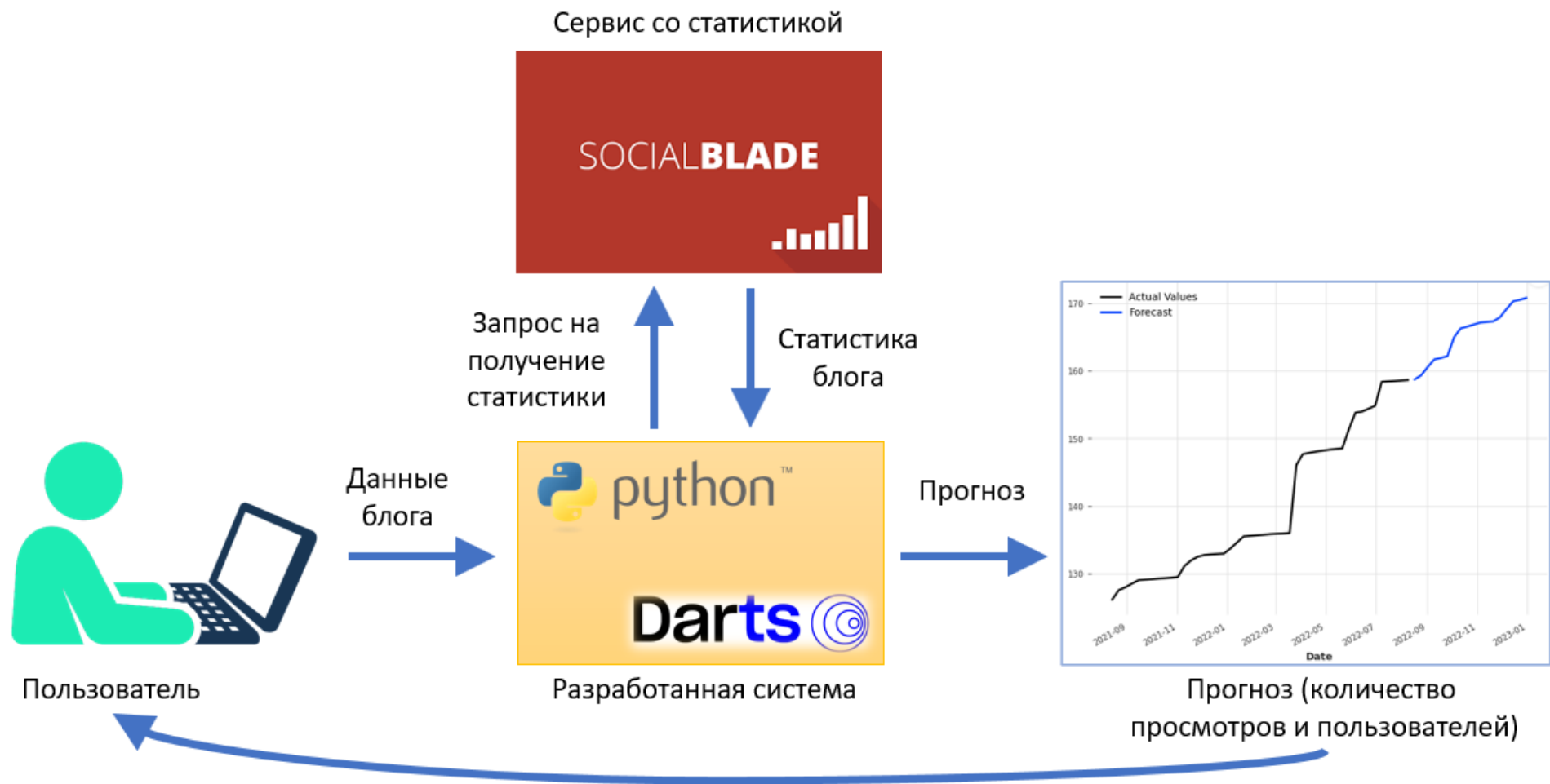


Рисунок 8 – Схема работы программного обеспечения

Затем информационная система анализирует полученные ряды и составляет прогнозы по изменению количества подписчиков и просмотров. Прогнозы визуализируются в виде графиков и отправляются пользователю системы. На основе полученных прогнозов пользователь оценивает эффективность продвижения корпоративного блога.

2.2 Анализ временных рядов

Для расчета прогнозов требуется выполнение анализа временных рядов.

«Анализ временных рядов — совокупность математико-статистических методов анализа, предназначенных для выявления структуры временных рядов и для их прогнозирования. Сюда относятся, в частности, методы регрессионного анализа. Выявление структуры временного ряда необходимо для того, чтобы построить математическую модель того явления, которое является источником анализируемого временного ряда. Прогноз будущих значений временного ряда используется для эффективного принятия решений» [10].

Временные ряды состоят из двух элементов:

- периода времени, за который или по состоянию на который приводятся числовые значения [16];
- числовых значений того или иного показателя, называемых уровнями ряда [17].

Временные ряды классифицируются по следующим признакам:

- по форме представления уровней [18];
- по количеству показателей, для которых определяются уровни в каждый момент времени: одномерные и многомерные временные ряды [19];
- по характеру временного параметра: моментные и интервальные временные ряды. В моментных временных рядах уровни характеризуют

значения показателя по состоянию на определенные моменты времени. В интервальных рядах уровни характеризуют значение показателя за определенные периоды времени. Важная особенность интервальных временных рядов абсолютных величин заключается в возможности суммирования их уровней. Отдельные же уровни моментного ряда абсолютных величин содержат элементы повторного счёта. Это делает бессмысленным суммирование уровней моментных рядов [20];

- по расстоянию между датами и интервалами времени выделяют равноотстоящие — когда даты регистрации или окончания периодов следуют друг за другом с равными интервалами и неполные (неравноотстоящие) — когда принцип равных интервалов не соблюдается [13];

- по наличию пропущенных значений: полные и неполные временные ряды [14];

- временные ряды бывают детерминированными и случайными: первые получают на основе значений некоторой неслучайной функции (ряд последовательных данных о количестве дней в месяцах); вторые есть результат реализации некоторой случайной величины [15];

- в зависимости от наличия основной тенденции выделяют стационарные ряды, в которых среднее значение и дисперсия постоянны, и нестационарные, содержащие основную тенденцию развития.

В разрабатываемом программном обеспечении анализ временных рядов будет применяться для прогнозирования изменения количества подписчиков и просмотров корпоративного блога.

Выбор методов для прогнозирования временных рядов зависит от применяемой в программном проекте библиотеки. В нашем случае, при разработке программного обеспечения будет использоваться библиотека Darts, реализует такие методы анализа временных рядов, как Naive Seasonal, Exponential Smoothing и Linear Regression [11], [12].

Выводы по главе 2

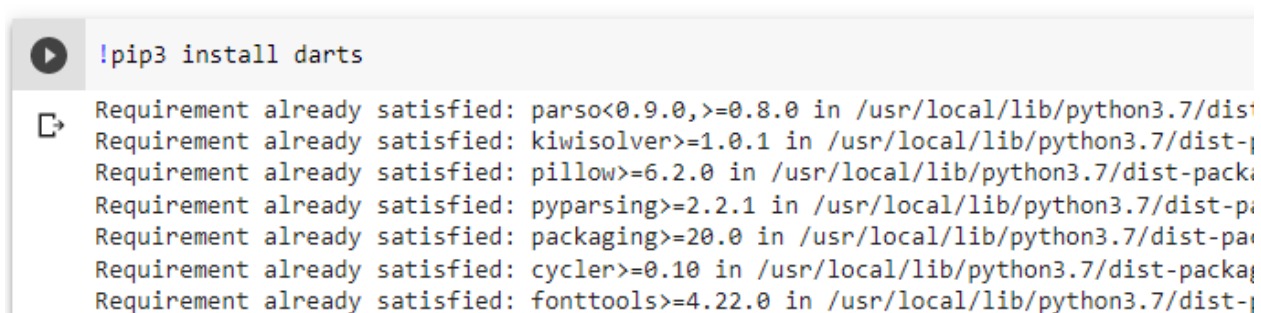
Приведем выводы по второй главе бакалаврской работы:

- предложен алгоритм оценки эффективности продвижения блога, который включает в себя: прогнозирование изменения количества подписчиков и просмотров за счет естественного прироста; прогнозирование изменения количества подписчиков и просмотров на основе промежуточных данных, полученных в начале продвижения блога; сравнение этих двух прогнозов друг с другом;
- предложена схема работы программного обеспечения, которая включает в себя загрузку исторических данных корпоративного блога с данными об изменении количества подписчиков и просмотров из сервиса Social blade и анализе полученных данных с использованием библиотеки Darts для составления прогноза по изменению показателей блога в результате его продвижения.

Глава 3 Разработка приложения для классификации текста

3.1 Особенности реализации программного модуля

Разработку программного обеспечения для контроля и анализа эффективности продвижения блога в социальной сети было принято реализовывать на основе готовых компонентов. Для языка python существует компонент darts, который содержит в себе необходимые методы по анализу временных рядов [1], [2], [8]. Интеграция компонента darts в программный проект осуществляется с помощью кода, показанного ниже (рисунок 9).



```
!pip3 install darts
Requirement already satisfied: parso<0.9.0,>=0.8.0 in /usr/local/lib/python3.7/dist-packages (from darts)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.7/dist-packages (from darts)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.7/dist-packages (from darts)
Requirement already satisfied: pyparsing>=2.2.1 in /usr/local/lib/python3.7/dist-packages (from darts)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.7/dist-packages (from darts)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.7/dist-packages (from darts)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.7/dist-packages (from darts)
```

Рисунок 9 – Интеграция библиотеки для анализа временных рядов в программный проект

Для табличного представления данных с количеством просмотров и подписчиков будет применяться библиотека Pandas. Для визуализации данных прогноза в виде графиков будут использоваться методы из библиотеки matplotlib. Обработка предупреждений будет осуществляться с помощью библиотеки logging.

Чтобы не захламлять проект интеграцией всех возможных методов из перечисленных выше библиотек осуществим подключение только тех методов, которыми мы будем пользоваться. Это осуществляется с помощью программного кода, показанного на рисунке 10.

```
[ ] import pandas as pd
import matplotlib.pyplot as plt
import matplotlib as mpl
from statsmodels.tsa.seasonal import seasonal_decompose
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from darts import TimeSeries
from darts.models import *
from darts.metrics import *
from darts.dataprocessing.transformers import Scaler
import logging
```

Рисунок 10 – Подключение необходимых модулей

Задать размер генерируемых графиков с прогнозами можно через свойство `figure.dpi`. Чем меньше это значение – тем меньше масштаб графика. Путем подбора было выбрано оптимальное значение равное 100. Также, с помощью команды `logging.disable` были отключены предупреждающие сообщения, встроенные в библиотеку `darts` (рисунок 11).

```
mpl.rcParams['figure.dpi'] = 100
logging.disable(logging.CRITICAL)
```

Рисунок 11 – Установка параметров отображения графиков и рисунков

Информация с данными временных рядов по количеству просмотров и пользователей для анализируемого корпоративного блога хранится в `xlsx` файлах (`excel`-файл). Для загрузки хранящихся там данных в программу используется метод `read_excel()`. Входными параметрами метода является расположение загружаемого `excel`-файла и названия столбца, содержащего в себе даты временного ряда.

Полный код для загрузки данных из `excel`-файла показан на рисунке 12.

```
▶ from google.colab import drive
drive.mount('/content/gdrive')
```

```
↳ Drive already mounted at /content/gdrive; to attempt to forcibly remount, call drive.mount('/content/gdrive', force_remount=True)
```

```
[ ] dfm = pd.read_excel('gdrive/MyDrive/Colab Notebooks/subs1.xlsx', index_col="Date")
```

Рисунок 12 – Обеспечение доступа к файлам

Загруженные из файла данные помещаются во временную переменную `dfm` типа `dataframe`. Данная переменная хранит в себе и временной ряд отвечающий за изменение количества подписчиков и временной ряд, отвечающий за изменение количества просмотров.

Для последующего анализа необходимо эти временные ряды разделить и поместить в отдельные переменные `df_subs` (данные по количеству подписчиков) и `df_views` (данные по количеству просмотров). Программный код, выполняющий данную задачу показан на рисунке 13.

```
[ ] df_subs = dfm.drop(['Views'], axis=1)
df_views = dfm.drop(['Subscribers'], axis=1)
```

Рисунок 13 – Формирование датафреймов

Теперь с помощью встроенной в библиотеку функции `plot()` пользователю выводятся графики, описывающие загруженные временные ряды:

- график изменения количества подписчиков (программный код и пример графика показан на рисунке 14);
- график изменения количества просмотров (рисунок 15).


```
[6] df_subs.plot(figsize=(8,5))
plt.title('Количества подписчиков (в тысячах)')

plt.show()
```



Рисунок 14 – Визуализация данных об изменении количества подписчиков

```
[7] df_views.plot(figsize=(8,5))
plt.title('Количество просмотров (в миллионах)')

plt.show()
```



Рисунок 15 – Визуализация данных об изменении количества просмотров

Для обеспечения наглядности, пользователю также выводится результат декомпозиции временного ряда, с помощью которого можно оценить влияние таких компонентов, как трендовая, сезонная и случайная составляющая ряда. Декомпозиция временного ряда выполняется с использованием функции `seasonal_decompose()`. Программный код и пример декомпозиции временного ряда показаны на рисунке 16

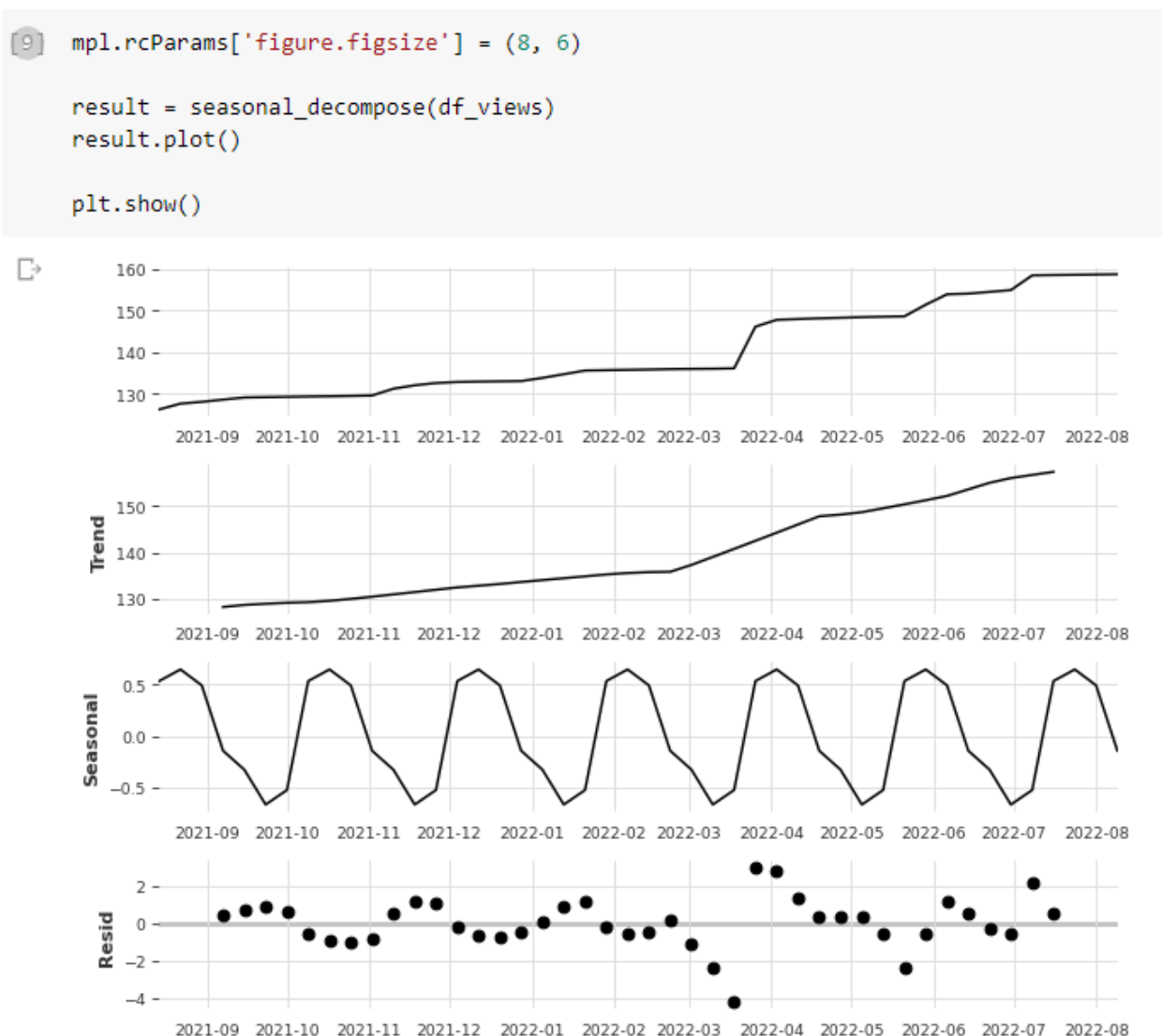


Рисунок 16 – Декомпозиция временного ряда с количеством подписчиков

Аналогичный программный код для выполнения декомпозиции временного ряда, связанного с количеством подписчиков корпоративного

блага показан на рисунке 17.

```
[8] mpl.rcParams['figure.figsize'] = (8, 6)

result = seasonal_decompose(df_subs)
result.plot()

plt.show()
```

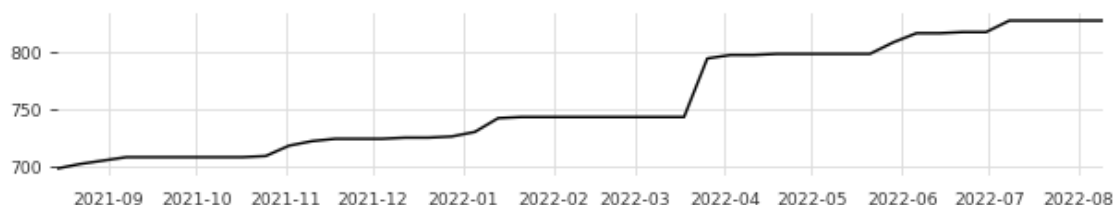


Рисунок 17 – Декомпозиция временного ряда с количеством просмотров

На следующем этапе работы программного обеспечения выполняется сравнение различных методов прогнозирования временных рядов. В программном обеспечении реализована поддержка следующих методов из библиотеки `statsmodels`:

- Naive Seasonal;
- Exponential Smoothing;
- Linear Regression.

Для того, чтобы подобрать лучший метод прогнозирования необходимо провести их сравнительный анализ. Сравнение методов осуществляется с использованием подхода «Back test», когда точность прогнозирования временного ряда проверяется на предыдущих исторических данных. Для удобства сравнения методов прогнозирования при проведении тестирования расчет и точностных показателей:

- ошибки MAE, RMSE, MAPE
- коэффициент детерминации R².

Чем меньше значение ошибок и больше значение коэффициента

детерминации, тем точнее метод прогнозирования. Программой код с функциями для тестирования методов и расчета точных показателей продемонстрирован на рисунке 18.

```
[10] df = df_views

[11] series = TimeSeries.from_dataframe(df)

start = pd.Timestamp('031022')
df_metrics = pd.DataFrame()

def plot_backtest(series, forecast, model_name):
    idx = -144
    series[idx:].plot(label='Actual Values')
    forecast[idx:].plot(label='Forecast')
    plt.title(model_name)
    plt.show()

def print_metrics(series, forecast, model_name):
    mae_ = mae(series, forecast)
    rmse_ = rmse(series, forecast)
    mape_ = mape(series, forecast)
    smape_ = smape(series, forecast)
    r2_score_ = r2_score(series, forecast)

    dict_ = {'MAE': mae_, 'RMSE': rmse_,
             'MAPE': mape_, 'SMAPE': smape_,
             'R2': r2_score_}

    df = pd.DataFrame(dict_, index = [model_name])

    return(df.round(decimals = 2))
```

Рисунок 18 – Функции для оценки качества прогнозных моделей

Программный код для тестирования каждого метода прогнозирования по отдельности на основе загруженных данных показан ниже:

- программный код для метода «Naive Seasonal» (рисунок 19)
- программный код для метода «Linear Regression» (рисунок 20)
- программный код для метода «Exponential Smoothing» (рисунок 21).

```
[12] model = NaiveSeasonal(K = 12)
      model_name = 'Naive Seasonal'

      plt.figure(figsize = (8, 5))

      forecast = model.historical_forecasts(series, start=start, forecast_horizon=1, verbose=True)
      plot_backtest(series, forecast, model_name)
      df_naive = print_metrics(series, forecast, model_name)
      df_metrics = df_metrics.append(df_naive)

      plt.show()
      df_naive
```

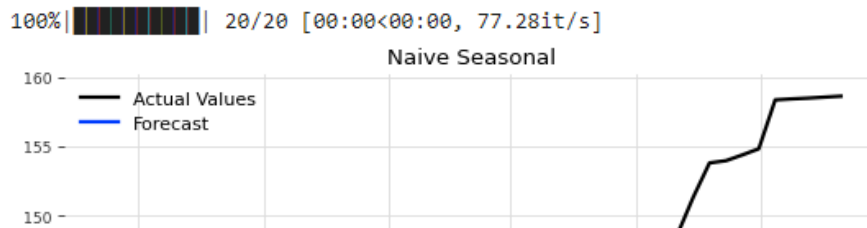


Рисунок 19 – Программный код для построения прогнозной модели методом «Naive Seasonal»

```
[13] model = LinearRegressionModel(lags = 12)
      model_name = 'Linear Regression'

      plt.figure(figsize = (8, 5))

      forecast = model.historical_forecasts(series, start=start, forecast_horizon=1, verbose=True)
      plot_backtest(series, forecast, model_name)
      df_lr = print_metrics(series, forecast, model_name)
      df_metrics = df_metrics.append(df_lr)

      plt.show()
      df_lr
```

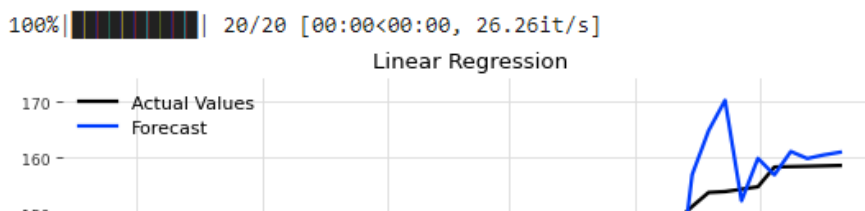


Рисунок 20 – Программный код для построения прогнозной модели методом «Linear Regression»

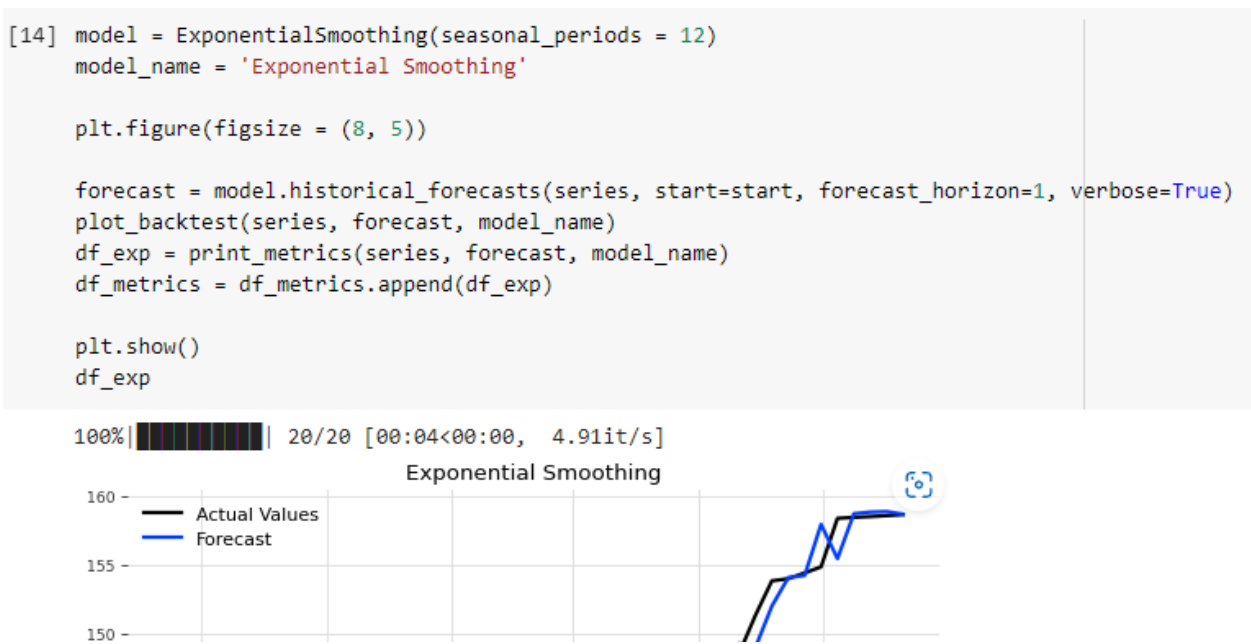


Рисунок 21 – Программный код для построения прогнозной модели методом «Exponential Smoothing»

В результате тестирования методов прогнозирования на имеющихся данных выбирается лучший по точности. Как показывают экспериментальные исследования, во всех случаях лучшим методом является «Exponential Smoothing».

Теперь с использованием лучшего метода производится прогнозирование изменения временного ряда в будущем. Прогнозирование реализуется с помощью метода `predict()`, которому в качестве параметра передается количество временных периодов для которых необходимо рассчитать прогноз. Временной период может составлять от одного до нескольких дней. Длина временного периода зависит от шага измерений в исходных данных.

Программный код для формирования результатов прогнозирования и их визуализации показан на рисунке 22.

```

▶ model = ExponentialSmoothing(seasonal_periods = 12)
  model_name = 'Exponential Smoothing'

  model.fit(series)
  #forecast = model.predict(13)
  forecast = model.predict(18)

  plot_backtest(series, forecast, model_name)
  print(forecast.pd_dataframe())

```

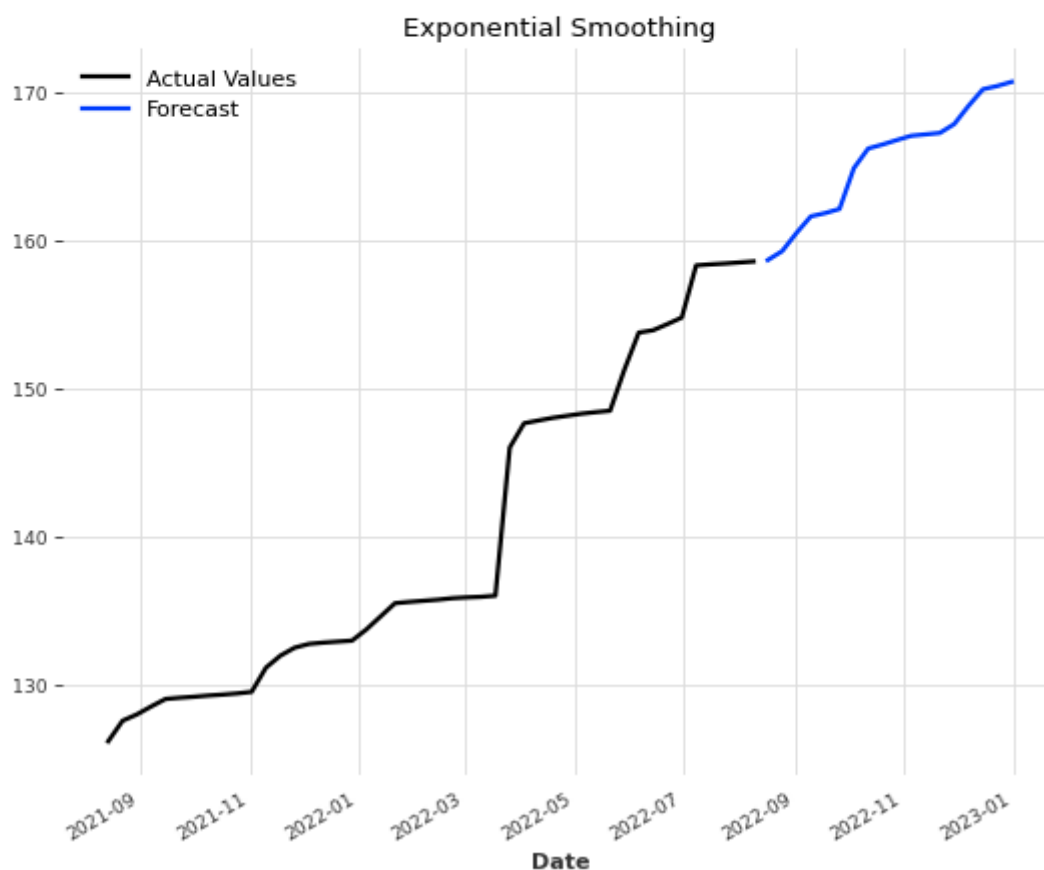


Рисунок 22 – Программный код для построения прогноза

На графике (рисунок 22) черная линия отвечает за фактические значения, а синяя за прогнозируемые значения. Для удобства работы с данными помимо графика в программе выводятся также и числовые значения прогноза. На рисунке 23 показан пример табличных данных прогноза по количеству просмотров.

component Date	Views
2022-08-17	158.704973
2022-08-25	159.306506
2022-09-02	160.539495
2022-09-10	161.660427
2022-09-18	161.880065
2022-09-26	162.160791
2022-10-04	164.919672
2022-10-12	166.248834
2022-10-20	166.519575
2022-10-28	166.821173
2022-11-05	167.111209
2022-11-13	167.205515
2022-11-21	167.301810
2022-11-29	167.903343
2022-12-07	169.136332
2022-12-15	170.257264
2022-12-23	170.476902
2022-12-31	170.757628

Рисунок 23 – Вывод прогноза по просмотрам в виде таблицы

Придем к тестированию программного обеспечения на реальных данных.

3.3. Результаты тестирования приложения

Приведем пример использования разработанного программного обеспечения при анализе данных корпоративного блога компании Nvidia. Проверяем, данные из каких социальных сетей доступны для просмотра по компании Nvidia через сервис Social Blade (рисунок 24).

Для анализа выберем данные корпоративного блога Nvidia из социальной сети YouTube.

The screenshot shows the Social Blade website interface. At the top, there is a navigation bar with a signal strength icon and the text 'SOCIAL BLADE'. Below this is a red navigation bar with links for 'SUPPORTED PLATFORMS', 'TOP LISTS', 'LIVE COUNTS / REALTIME', and 'CONSULTING'. A blue banner below the navigation bar reads 'Happy Fall! (in the Northern Hemisphere) Take 20% off all Business A'. The main content area is titled 'Search Results' and indicates '4 direct search results found'. The results are listed as follows:

- NVIDIA** (Twitter): 1,903,860 Followers, 12,317 Tweets.
- nvidia** (Twitch): 412,860 Followers, 15,192,255 Channel Views, Special Events.
- NVIDIA** (Instagram): Followers, Uploads.
- NVIDIA** (YouTube): 947,000 Subscribers, 189,967,261 Video Views, 2,363 Uploads.

Рисунок 24 – Данные корпоративных блогов компании Nvidia в различных социальных сетях в сервисе Social Blade

В разделе Detailed Statistics сервиса Social Blade для корпоративного блога компании Nvidia в социальной сети YouTube доступны следующие данные (рисунок 25):

- еженедельное изменение количества подписчиков за последний год (верхний график Total Subscribers);
- еженедельное изменение суммарного количества просмотров за последний год (верхний график Total Video Views).

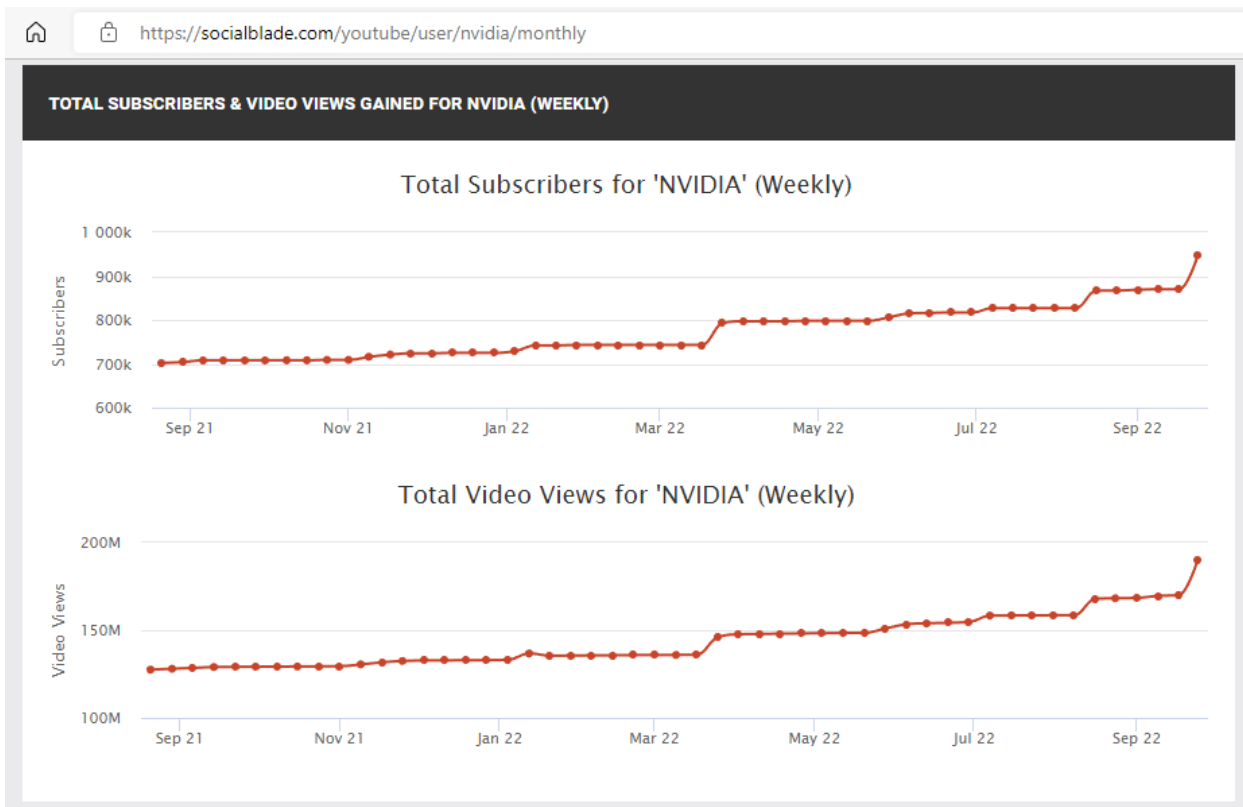


Рисунок 25 – Статистические данные по количеству подписчиков и количеству просмотров корпоративного блога Nvidia в социальной сети YouTube (сервис Social Blade)

Предположим, что компания Nvidia 17.08.2022 решила начать продвижение своего корпоративного блога в соответствии с маркетинговой стратегией, и хочет оценить ее эффективность. В качестве критерия эффективности продвижения в данном примере будет использоваться прирост суммарного количества просмотров на последний день 2022 года. Для оценки эффективности с помощью программного обеспечения сначала рассчитывается прогноз по естественному приросту количества просмотров, не связанных с реализацией маркетинговой стратегии. Затем компания начинает реализацию маркетинговой стратегии и получает новые данные по изменению количеству просмотров. Теперь строится второй прогноз на

основе данных, полученных в результате продвижения блога. Затем, чтобы оценить эффективность продвижения блога на будущую дату производится сравнение двух полученных прогнозов (рисунок 26).

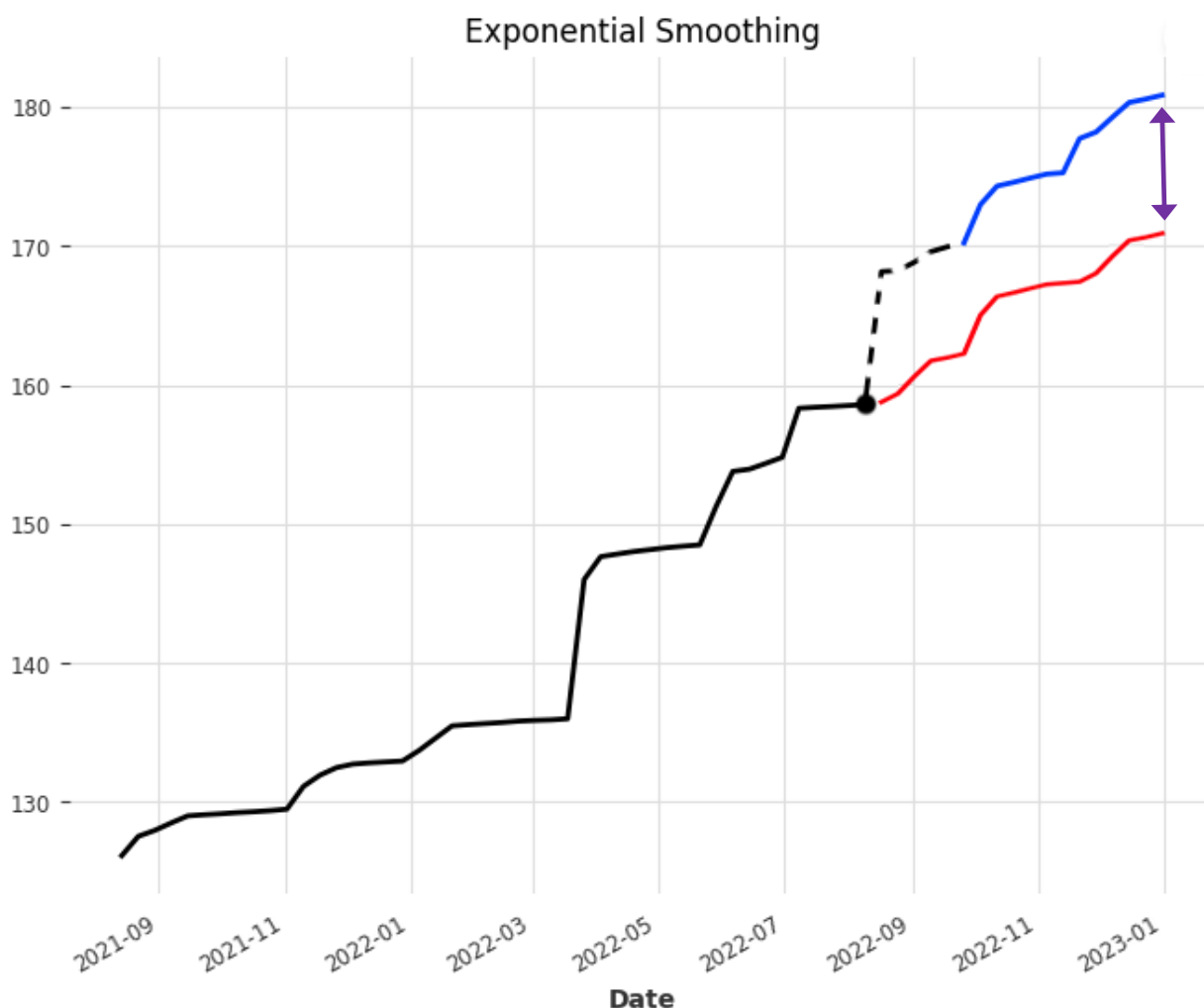


Рисунок 26 – Вывод прогноза по просмотрам в виде таблицы

На рисунке 26 приняты следующие обозначения:

- сплошной черной линией показаны фактические данные временного ряда;
- черной точкой показан момент времени, когда компания приняла решение о начале мероприятий по продвижению блога;

- красной линией показан рассчитанный прогноз по росту количества просмотров за счет естественных процессов (без использования продвижения);
- пунктирной черной линией показаны фактические значения после начала продвижения корпоративного блога;
- синей линией показан прогноз изменения количества просмотров на основе как исторических данных (черная линия), так и фактических значений, полученных после старта продвижения блога (пунктирная линия).

Результаты прогнозирования в числовом виде приведены в таблице 1.

Таблица 1 – Результаты прогнозирования изменения количества просмотров (в миллионах) на конец 2022 года без продвижения блога и с учетом промежуточных данных по продвижению

Без продвижения блога		После продвижения	
component Date	Views	component Date	Views
2022-08-17	158.704973	2022-09-26	170.240027
2022-08-25	159.306506	2022-10-04	172.998956
2022-09-02	160.539495	2022-10-12	174.328105
2022-09-10	161.660427	2022-10-20	174.598729
2022-09-18	161.880065	2022-10-28	174.900322
2022-09-26	162.160791	2022-11-05	175.190453
2022-10-04	164.919672	2022-11-13	175.284803
2022-10-12	166.248834	2022-11-21	177.747898
2022-10-20	166.519575	2022-11-29	178.216705
2022-10-28	166.821173	2022-12-07	179.293708
2022-11-05	167.111209	2022-12-15	180.341487
2022-11-13	167.205515	2022-12-23	180.587824
2022-11-21	167.301810	2022-12-31	180.868531
2022-11-29	167.903343		
2022-12-07	169.136332		
2022-12-15	170.257264		
2022-12-23	170.476902		
2022-12-31	170.757628		

Эффективность маркетинговой стратегии, реализуемой в результате продвижения, оценивается как разница между значениями из двух столбцов

таблицы 2 на конкретную дату. Так на 31.12.2022 продвижение корпоративного блога позволит увеличить количество просмотров до 180,96 миллионов. При этом продвижение обеспечивает прирост просмотров равный $180,86 - 170,75 = 10,11$ миллионов (фиолетовая стрелка на рисунке 25). Если такой прогнозируемый прирост является удовлетворительным, то текущая маркетинговая стратегия, реализуемая при продвижении, признается эффективной.

Выводы по главе 3

Приведем выводы по третьей главе бакалаврской работы:

- на языке программирования python разработан программный модуль для контроля и анализа эффективности продвижения блога в социальной сети;
- разработанное программное обеспечение протестировано на основе данных корпоративного блога Nvidia из социальной сети YouTube, результаты тестирования доказали работоспособность предложенного в бакалаврской работе решения.

Заключение

При выполнении бакалаврской работы были получены следующие результаты:

- в ходе анализа литературных данных установлено, что за последние 10 лет количество пользователей социальных сетей увеличилось более чем в 3 раза и продолжает расти;

- увеличение роли социальных сетей привело к росту в них корпоративных блогов, цель ведения которых заключается в продвижении среди пользователей своих товаров и услуг;

- для продвижения своих товаров и услуг как можно большему числу пользователей владельцы корпоративных блогов стремятся увеличить количество своих подписчиков и число просмотров, для этого на основе выбранной маркетинговой стратегии выполняется продвижение корпоративного блога;

- с использованием методологии IDEF0 проведено функциональное моделирование процесса продвижения блога, в результате которого установлено, что данный процесс можно модернизировать за счет разработки и внедрения системы для анализа и прогнозирования результатов продвижения блога;

- предложен алгоритм оценки эффективности продвижения блога, который включает в себя: прогнозирование изменения количества подписчиков и просмотров за счет естественного прироста; прогнозирование изменения количества подписчиков и просмотров на основе промежуточных данных, полученных в начале продвижения блога; сравнение этих двух прогнозов друг с другом;

- предложена схема работы программного обеспечения, которая включает в себя загрузку исторических данных корпоративного блога с данными об изменении количества подписчиков и просмотров из сервиса

Social blade и анализе полученных данных с использованием библиотеки Darts для составления прогноза по изменению показателей блога в результате его продвижения;

- на языке программирования python разработан программный модуль для контроля и анализа эффективности продвижения блога в социальной сети;

- разработанное программное обеспечение протестировано на данных корпоративного блога Nvidia из социальной сети YouTube, результаты тестирования доказали работоспособность предложенного в бакалаврской работе решения.

Список используемой литературы и используемых источников

1. Григорьев Е.А. Разведочный анализ данных с помощью Python / Григорьев Е.А., Климов Н.С. // E-Scio. 2020. №2 (41). URL: <https://cyberleninka.ru/article/n/razvedochnyy-analiz-dannyh-s-pomoschyu-python> (дата обращения: 22.09.2022).
2. Гришков, Д.Ю. Язык высокого уровня программирования Python / Гришков Данила Юрьевич, Аусилова Назерке Мырзабековна // НИР/S&R. 2022. №1 (9). URL: <https://cyberleninka.ru/article/n/yazyk-vysokogo-urovnya-programirovaniya-python> (дата обращения: 22.09.2022).
3. Евсюков, В.В. SMM - драйвер повышения эффективности маркетинга в современных условиях / Евсюков Владимир Васильевич, Райман Анастасия Константиновна // Известия ТулГУ. Экономические и юридические науки. 2018. №1-1. URL: <https://cyberleninka.ru/article/n/smm-drayver-povysheniya-effektivnosti-marketinga-v-sovremennyh-usloviyah> (дата обращения: 22.09.2022).
4. Ершов, В.Е. Тенденции развития рекламной деятельности в социальных сетях / Ершов Вадим Евгеньевич // Вестник евразийской науки. 2015. №5 (30). URL: <https://cyberleninka.ru/article/n/tendentsii-razvitiya-reklamnoy-deyatelnosti-v-sotsialnyh-setyah> (дата обращения: 22.09.2022).
5. Злобина, Н.В. Маркетинг в социальных сетях: современные тенденции и перспективы / Злобина Наталья Васильевна, Завражина Кристина Владимировна // π-Economy. 2015. №6 (233). URL: <https://cyberleninka.ru/article/n/marketing-v-sotsialnyh-setyah-sovremennye-tendentsii-i-perspektivy> (дата обращения: 22.09.2022).
6. Музыкант, В.Л. Опыт исследования SMM как вирусной стратегии в медийном пространстве / Музыкант Валерий Леонидович // Вестник РУДН. Серия: Социология. 2014. №4. URL: <https://cyberleninka.ru/article/n/opyt>

issledovaniya-smm-kak-virusnoy-strategii-v-mediynom-prostranstve (дата обращения: 22.09.2022).

7. Ольшевский, Д. SMM-продвижение как эффективный инструмент интернет-маркетинга / Ольшевский Дмитрий // Наука и инновации. 2017. №175. URL: <https://cyberleninka.ru/article/n/smm-prodvizhenie-kak-effektivnyy-instrument-internet-marketinga> (дата обращения: 22.09.2022).

8. Тарасова А.Н. Сентиментальный анализ постов в социальных сетях посредством Python / Тарасова А.Н., Иванов К.О. // Символ науки. 2022. №3-1. URL: <https://cyberleninka.ru/article/n/sentimentalnuy-analiz-postov-v-sotsialnyh-setyah-posredstvom-python> (дата обращения: 22.09.2022).

9. Тепляков, В.А. SMM-маркетинг. Особенности продвижения / Тепляков Владислав Алексеевич // Вестник науки и образования. 2018. №2 (38). URL: <https://cyberleninka.ru/article/n/smm-marketing-osobennosti-prodvizheniya> (дата обращения: 22.09.2022).

10. Чибирова, М.Э. Анализ данных и регрессионное моделирование с применением языков программирования Python и R / Чибирова Марина Эльбрусевна // Научные записки молодых исследователей. 2019. №2. URL: <https://cyberleninka.ru/article/n/analiz-dannyh-i-regressionnoe-modelirovanie-s-primeneniem-yazykov-programmirovaniya-python-i-r> (дата обращения: 22.09.2022).

11. Bettaiah, V. Alignment of Time Series for Subsequence-to-Subsequence Time Series Matching / Vineetha Bettaiah, Heggere S. Ranganath // SAI 2014 part of the Studies in Computational Intelligence book series: Intelligent Systems in Science and Information, 2014. - Springer Nature Switzerland AG, 2014. - pp.193–208

12. Higuchi, T. Special Section on Nonparametric Approach to Time Series Analysis / Tomoyuki Higuchi, Genshiro Kitagawa // Annals of the Institute

of Statistical Mathematics, 2002. - №54 (169). - Springer Nature Switzerland AG 2002. - pp.101-112

13. Kiefer, D. Univariate Time Series Forecasting: Machine Learning Prediction of the Best Suitable Forecast Model Based on Time Series Characteristics / Daniel Kiefer, Markus Bauer, Florian Grimm // Human Centred Intelligent Systems Proceedings of KES-HCIS 2021 Conference. - Springer Nature Switzerland AG, 2021. - pp.152–162

14. Kriegel, H.P. Analysis of Time Series Using Compact Model-Based Descriptions / Hans-Peter Kriegel, Peer Kröger, Alexey Pryakhin, Matthias Renz // International Conference on Database Systems for Advanced Applications DASFAA 2008: Database Systems for Advanced Applications. - Springer Nature Switzerland AG, 2008. - pp.698–701

15. Kriegel, H.P. Approximate Clustering of Time Series Using Compact Model-Based Descriptions / Hans-Peter Kriegel, Peer Kröger, Alexey Pryakhin, Matthias Renz, Andrew Zherdin // International Conference on Database Systems for Advanced Applications DASFAA 2008: Database Systems for Advanced Applications. - Springer Nature Switzerland AG, 2008. - pp.364–379

16. Nikolskii, I.M. Erratum to: Parallel Algorithm to Detect Structural Changes in Time Series / I.M. Nikolskii, K.K. Furmanov // Computational Mathematics and Modeling, 2016. - №27 (394). - Springer Nature Switzerland AG 2016. - pp.84-92

17. Paparoditis, E. Nonparametric Functional Time Series Prediction / Efstathios Paparoditis // Recent Advances in Functional Data Analysis and Related Topics - Conference proceedings. - Springer Nature Switzerland AG, 2011. - pp.251–254

18. Thuan, N.D. Using Fuzzy Time Series Model Based on Hedge Algebras and Relationship Groups Following Time Points for Forecasting Time Series / Nguyen Dinh Thuan, Hoang Tung // International Conference on Future Data and Security Engineering FDSE 2020: Future Data and Security Engineering.

Big Data, Security and Privacy, Smart City and Industry 4.0 Applications. - Springer Nature Switzerland AG, 2020. - pp.401–410

19. Truong, C.D. Time Series Prediction Using Motif Information / Cao Duy Truong, Duong Tuan Anh // International Workshop on Multi-disciplinary Trends in Artificial Intelligence MIWAI 2012: Multi-disciplinary Trends in Artificial Intelligence. - Springer Nature Switzerland AG, 2012. - pp.110–121

20. Widiputra, H. Multiple Time-Series Prediction through Multiple Time-Series Relationships Profiling and Clustered Recurring Trends / Harya Widiputra, Russel Pears, Nikola Kasabov // Pacific-Asia Conference on Knowledge Discovery and Data Mining PAKDD 2011: Advances in Knowledge Discovery and Data Mining. - Springer Nature Switzerland AG, 2011. - pp.161–172