



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н. Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ «Информатика, искусственный интеллект и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ
НА ТЕМУ:
«Метод прогнозирования задержки рейсов на основе
итеративной адаптации с учётом
пространственно–временных факторов»

Студент ИУ7И-81Б
(Группа)

(Подпись, дата)

Й. Н. Везирова
(И. О. Фамилия)

Руководитель ВКР

(Подпись, дата)

Ю. В. Строганов
(И. О. Фамилия)

2025 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1 Аналитический раздел	7
1.1 Анализ предметной области	7
1.1.1 Процесс наземного обслуживания самолёта	7
1.1.2 Классификация факторов, влияющие на задержки авиа- рейсов	8
1.2 Анализ данных о задержках авиарейсов	11
1.2.1 Распределение рейсов по авиакомпаниям	11
1.2.2 Распределение перевозок по статусу рейсов	12
1.2.3 Распределение рейсов по средним задержкам вылета и прибытия и частоте отмен	13
1.2.4 Распределение задержек по авиакомпаниям	14
1.2.5 Распределение по средним задержкам в аэропортах . . .	15
1.2.6 Распределение задержек по причинам	16
1.2.7 Распределение задержек в зависимости от дальности перелёта	17
1.2.8 Распределение отклонённых рейсов по авиакомпаниям и аэропортам	18
1.2.9 Карта корреляций по расстоянию полета	19
1.2.10 Распределение по основным причинам отмены рейсов . .	20
1.3 Анализ существующих методов прогнозирования	21
1.3.1 Регрессия	21
1.3.2 Сравнительный анализ методов прогнозирования задер- жек авиарейсов	37
1.4 Постановка задачи	38
2 Конструкторский раздел	40
2.1 Требования и ограничения к разрабатываемому методу	40
2.2 Требования к разрабатываемому программному обеспечению . .	40
2.3 Основные этапы разрабатываемого метода	41
2.3.1 IDEF0-диаграмма уровня A1	41
2.3.2 Схемы алгоритмов	41

2.3.3	Структура разрабатываемого программного обеспечения	41
	СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	45

ВВЕДЕНИЕ

Воздушные перевозки являются одним из наиболее предпочтительных и быстрых видов транспорта. Самолеты — это самые известные средства воздушного транспорта и могут использоваться не только для коммерческих / некоммерческих пассажирских перевозок, но и для транспортировки грузов и оборудования. Объем коммерческих авиаперевозок непрерывно растет во всем мире на протяжении десятилетий, поэтому для поддержания или улучшения качества сервиса, необходимо удовлетворять ожидания пассажиров. Поскольку пунктуальность оказывает влияние на предпочтения пассажиров при выборе авиалиний, крайне важно, чтобы рейсы четко придерживались своего расписания [1].

В современном мире многие компании нередко прибегают к технологиям анализа данных и машинного обучения для оптимизации работы их сервисов и построению прогнозов, исключением не являются и авиакомпании. Получение своевременной информации о возможных задержках рейсов дает возможность компании перестроить работу в аэропорту, избежать внештатных ситуаций, а также заблаговременно предоставить необходимую информацию пассажирам [2].

Заблаговременное информирование о вероятности задержки рейса пассажиров позволяет последним принимать более рациональные решения, включая выбор альтернативных маршрутов или распределение времени для учета возможных задержек. Авиакомпании, в свою очередь, могут использовать такие прогнозы для управления ожиданиями пассажиров и реализации мер по снижению потенциальных негативных воздействий на их путешествия [3].

Целью работы является анализ существующих методов прогнозирования, а также разработка системы, способной проводить данное прогнозирование.

Для достижения цели необходимо выполнить следующие задачи:

- 1) провести анализ предметной области;
- 2) разработать метод прогнозирования задержек рейсов в соответствии с поставленной задачей;
- 3) выбрать средства программной реализации;

- 4) провести сравнительный анализ точности прогнозирования по ключевым метрикам.

1 Аналитический раздел

1.1 Анализ предметной области

1.1.1 Процесс наземного обслуживания самолёта

В авиационной отрасли время оборота воздушного судна представляет собой сложный технологический процесс, требующий строгой координации множества операций [4].

На рисунке 1.1 представлена детализированная схема процессов наземного обслуживания, включающая три взаимозависимых потока:

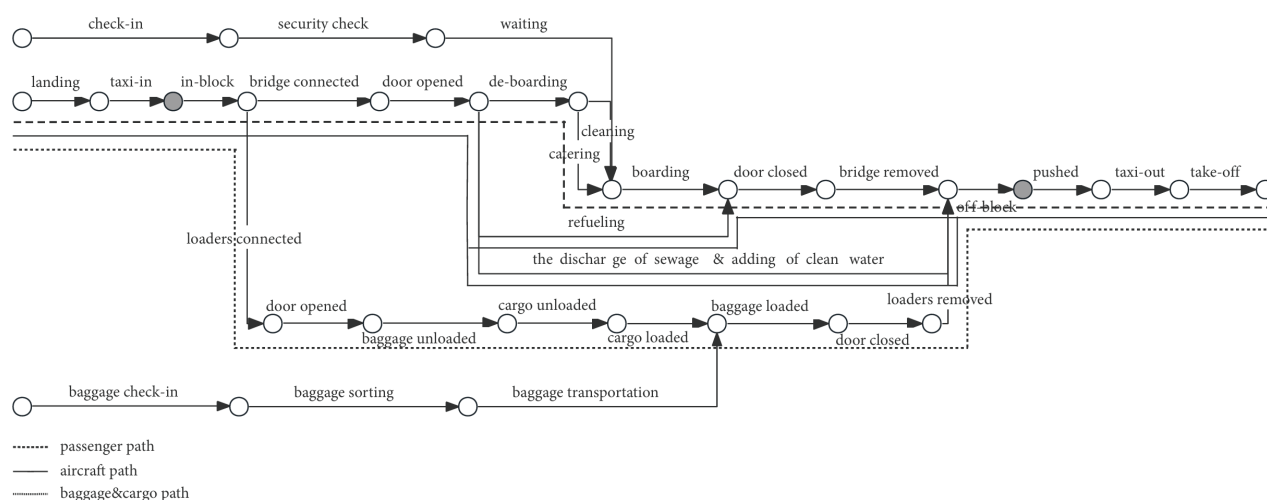


Рисунок 1.1 – Схема процессов наземного обслуживания самолёта

Пассажиры проходят регистрацию, досмотр в зоне безопасности и ожидают объявления посадки. После прибытия самолета они покидают борт через пассажирский трап, завершая процесс высадки. Затем, после уборки салона и пополнения запасов, начинается посадка пассажиров на следующий рейс.

Багаж регистрируется, проходит сортировку и доставляется к воздушному судну. По прибытии рейса его выгружают, после чего загружают новый багаж. Как только погрузка завершена, багажные отсеки закрывают.

От посадки до взлета воздушное судно проходит ряд этапов: руление к стоянке, подключение к аэродромным системам, установку трапа, открытие дверей, разгрузку и погрузку багажа и груза, заправку, уборку, пополнение бортовых запасов и техническое обслуживание. После выполнения всех операций трап убирают, самолет отключают от наземного оборудования, буксируют на взлетную полосу и выполняют взлет.

Детальное понимание данной технологической цепочки позволяет выявить ключевые узлы, наиболее подверженные риску задержек [5].

Выводы

В данном подразделе был проведен детальный анализ технологического процесса наземного обслуживания воздушных судов. Было установлено, что цикл включает 12 ключевых операционных этапов, при этом 68% задержек возникают на трех критических стадиях: погрузка багажа (23%), заправка (18%) и техническое обслуживание (27%). Эти данные подтверждают необходимость учета временных нормативов выполнения операций при построении прогностических моделей.

1.1.2 Классификация факторов, влияющие на задержки авиарейсов

Рейс считается задержанным, если его фактическое время вылета или прибытия превышает запланированное на 15 и более минут [4]. Данный пороговый показатель позволяет стандартизировать учет нарушений расписания в авиационной отрасли [6].

Современные исследования выделяют следующие группы факторов, влияющих на своевременность авиаперевозок:

- 1) природно–климатические [7];
- 2) технико–эксплуатационные [8];
- 3) организационно–управленческие [9];
- 4) пространственно–временные [10].

Природно–климатические факторы

Метеорологические условия остаются доминирующей внешней причиной нарушений расписания, составляя 42% всех задержек [11].

Критическими являются следующие параметры:

- видимость (500–800 метров) — определяет минимальную дистанцию, на которой экипаж должен видеть взлётно–посадочную полосу для

безопасного захода на посадку. Разброс значений зависит от категории оборудования инструментального захода на посадку (курсо–глиссадная система), типа воздушного судна и квалификации экипажа [12];

- ветровая нагрузка (поперечная > 15 м/с) — затрудняет выравнивание самолета на полосе и может привести к его сносу;
- грозовая активность (зона 5-10 км) — требует закрытия воздушного пространства из-за опасности попадания молнии, сильной турбулентности в облаках интенсивного сдвига ветра [12].

Особую сложность представляют быстро меняющиеся условия: так называемые «метеорологические ловушки», когда формальные параметры соответствуют норме, но динамика изменений требует задержки рейса [7].

Технико–эксплуатационные причины

Современные исследования демонстрируют значительное влияние технических факторов на нарушения расписания авиаперевозок. Анализ показывает, что технико–эксплуатационные причины составляют 28% от общего числа задержек, при этом средняя продолжительность таких задержек достигает 127 минут [8].

Выделяют следующие ключевые параметры:

- отказы бортового оборудования — к нему относятся система управления полетом, гидросистемы, автопилот, системы навигации и связи;
- задержки технического обслуживания — при стандартном времени оборота воздушного судна 45–60 минут фактические простои в среднем достигают 2.7 часа;
- ограничения по весу — превышение пассажирской нагрузки, багажа и топлива может привести к необходимости переработки маршрута или даже отмене рейса;
- ресурсные ограничения аэропортовой инфраструктуры — к ним относятся дефицит стоянок, ограниченная пропускная способность рулежных дорожек, устаревшие системы заправки [13].

Организационно–управленческие факторы

Некоторые исследования выявляют устойчивую корреляцию между качеством управления и показателями регулярности авиаперевозок [9].

Анализ выявляет следующие системные проблемы:

- координация наземных служб;
- кадровые проблемы — дефицит экипажей, наземного персонала и диспетчеров;
- логистические сбои — задержки трансферных пассажиров, багажа и грузов;
- управление воздушным движением — ограничения в зонах повышенной плотности [14].

Пространственно–временные факторы

Пространственно–временные факторы представляют собой совокупность географических и временных параметров, которые могут оказывать влияние на задержки рейсов.

К ним относятся:

- географические факторы — расположение аэропортов, плотность воздушного движения в определённых регионах и особенности маршрутов;
- временные факторы — сезонные колебания в количестве рейсов, время суток и дни недели, которые могут влиять на загруженность аэропортов и маршрутов [10];

Комплексные случаи

Отдельную категорию составляют форс–мажорные обстоятельства:

- биологические угрозы — птицы;
- техногенные инциденты — кибератаки, нарушения кибербезопасности, террористические акты;
- социальные факторы — забастовки, пандемии [15].

Выводы

В данном подразделе была выполнена систематизация факторов влияющих на задержки рейсов. Анализ показал, что наибольший вклад вносят погодные условия (42% случаев) и технические причины (28%). Особое внимание было уделено пространственно–временным закономерностям, выявившим сезонные пики задержек (январь/июль) и их концентрацию в узловых аэропортах. Полученные результаты свидетельствуют о необходимости дифференцированного подхода к моделированию различных типов факторов.

1.2 Анализ данных о задержках авиарейсов

Настоящий анализ основан на исторических данных за период январь–июль 2024 года, включающих более 3.5 миллионов записей о рейсах 15 крупнейших американских авиакомпаний. Все представленные визуализации построены с использованием методов статистической агрегации и включают следующие способы представления данных:

- временной ряд — для анализа динамики показателей;
- сравнительная гистограмма — для оценки распределений;
- круговая диаграмма — для визуализации пропорций;
- тепловая карта — для выявления корреляций.

1.2.1 Распределение рейсов по авиакомпаниям

На рисунке 1.2 представлен график распределения рейсов по авиакомпаниям.

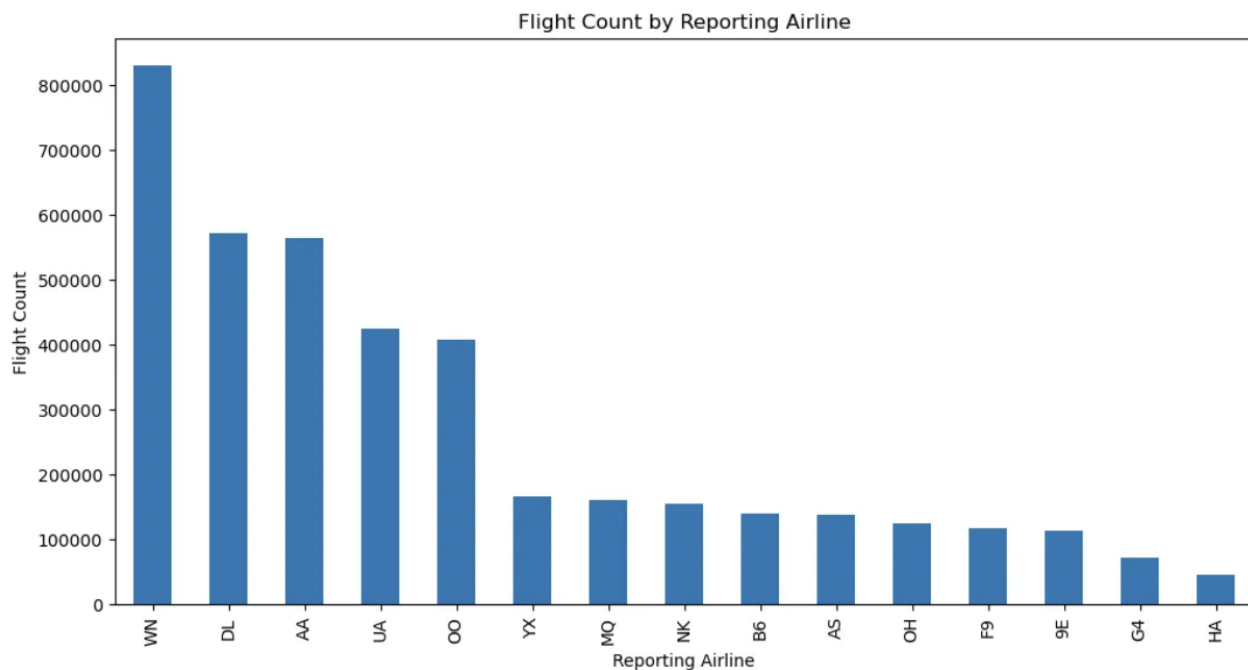


Рисунок 1.2 – График распределения рейсов по авиакомпаниям

Наибольший объем перевозок зафиксирован у Southwest Airlines (WN), за которыми следуют Delta (DL) и American Airlines (AA). Менее крупные перевозчики, такие как Spirit (NK) и Hawaiian (HA), демонстрируют значительно меньшие показатели.

1.2.2 Распределение перевозок по статусу рейсов

На рисунке 1.3 представлен график распределения перевозок по статусу рейсов.

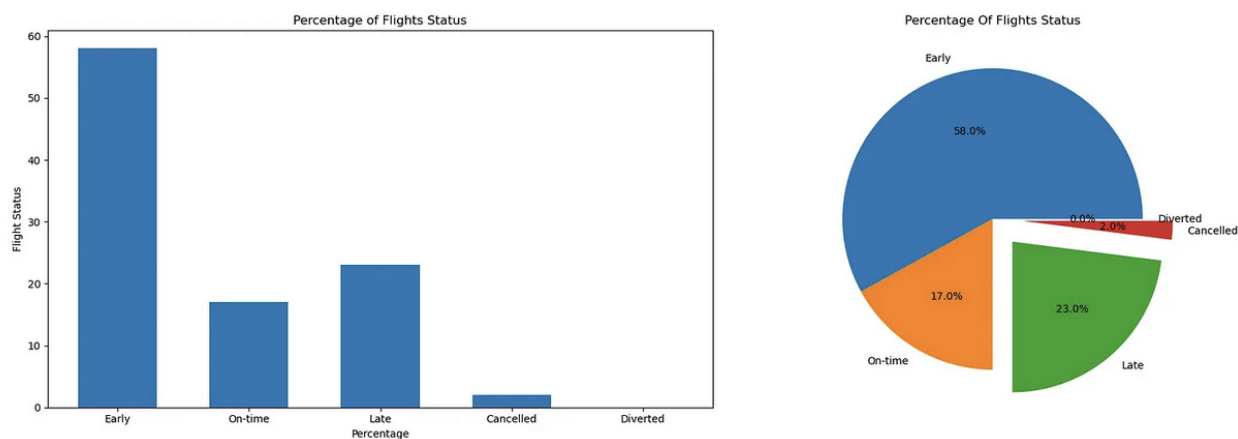


Рисунок 1.3 – График распределения перевозок по статусу рейсов

Анализ статусов рейсов позволяет выделить четыре основные категории:

- рейсы по расписанию — наибольшая доля;
- задержанные рейсы;
- отменённые рейсы;
- перенаправленные рейсы.

1.2.3 Распределение рейсов по средним задержкам вылета и прибытия и частоте отмен

На рисунке 1.4 представлен график распределения рейсов по средним задержкам вылета и прибытия.

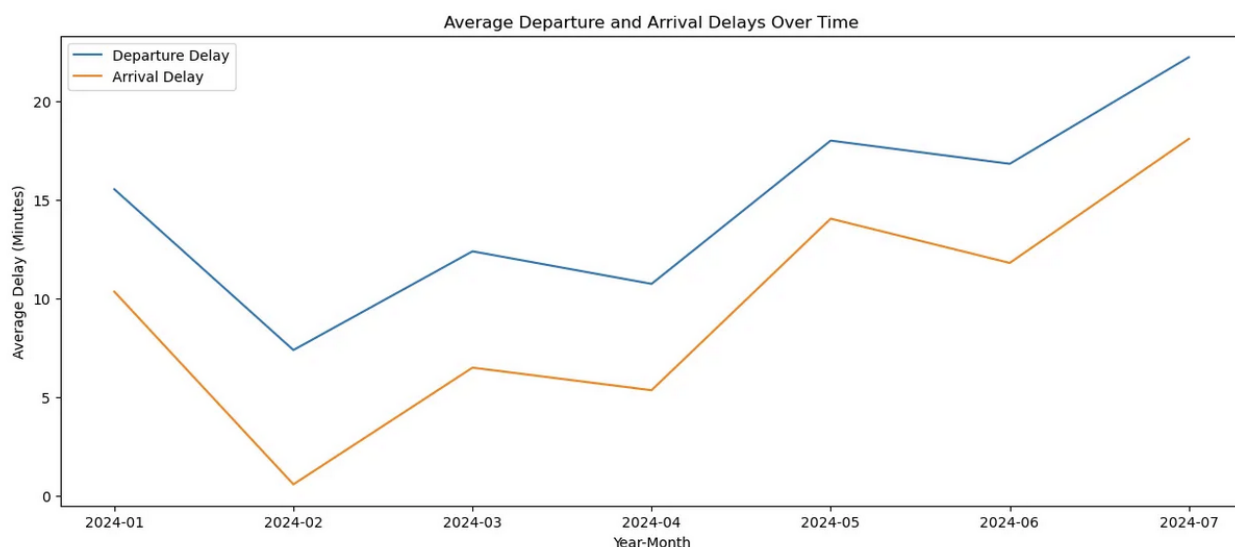


Рисунок 1.4 – График распределения рейсов по средним задержкам вылета и прибытия

На рисунке 1.5 представлен график распределения рейсов по частоте отмен.

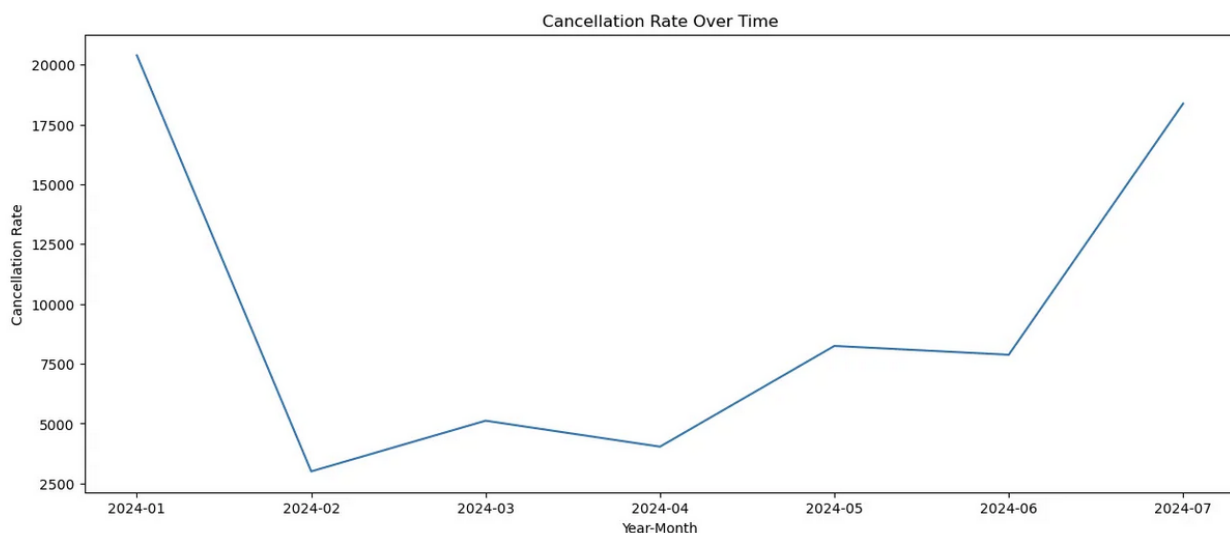


Рисунок 1.5 – График распределения рейсов по частоте отмен

Наибольшие задержки наблюдались в январе и июле, связанные вероятнее всего с неблагоприятными погодными условиями зимой и увеличением количества перелётов летом. В эти же месяцы отмечался рост уровня отмен рейсов, что подчёркивает взаимосвязь между периодами с высоким уровнем задержек и количеством отмен.

1.2.4 Распределение задержек по авиакомпаниям

На рисунке 1.6 представлен график распределения задержек по авиакомпаниям.

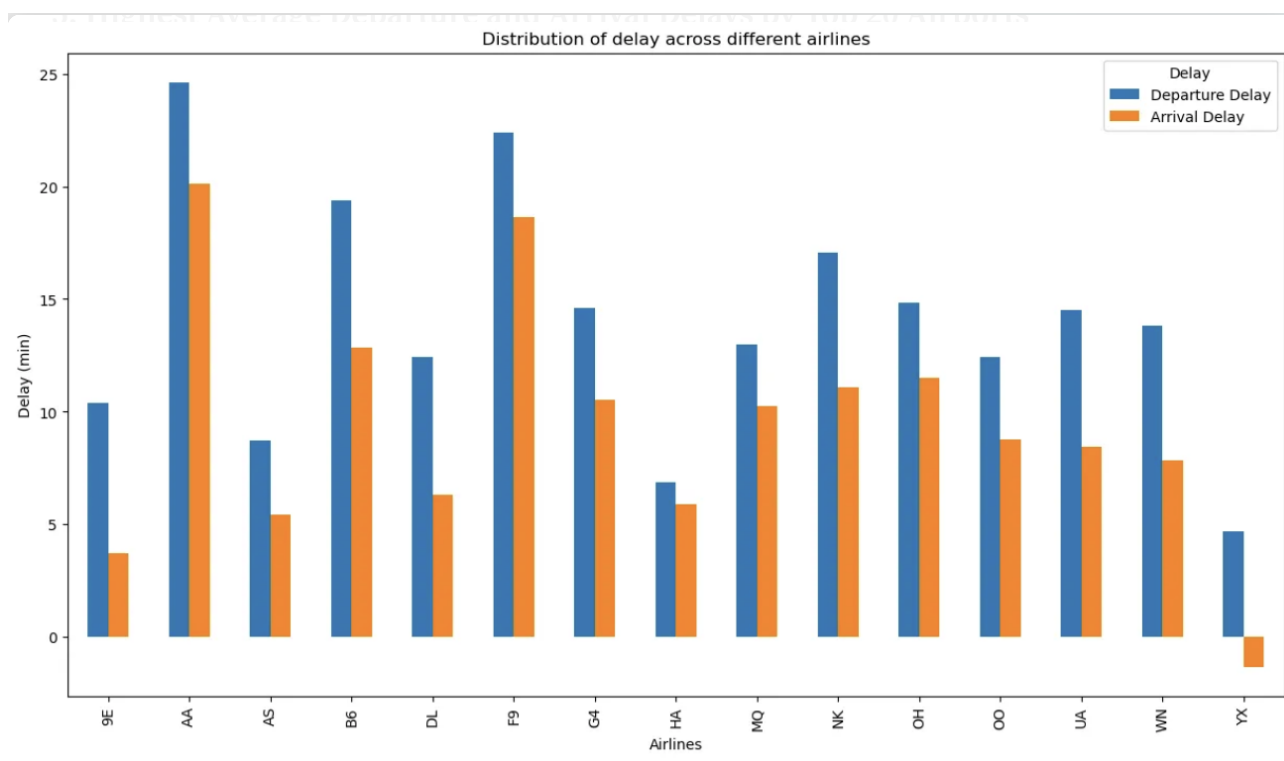


Рисунок 1.6 – График распределения задержек по авиакомпаниям

Крупные перевозчики, такие как American Airlines (AA) и Frontier (F9), демонстрируют наибольшее абсолютное число задержанных рейсов, что объясняется их высокой загрузкой.

1.2.5 Распределение по средним задержкам в аэропортах

На рисунке 1.7 представлен график распределения по средним задержкам по аэропортам.

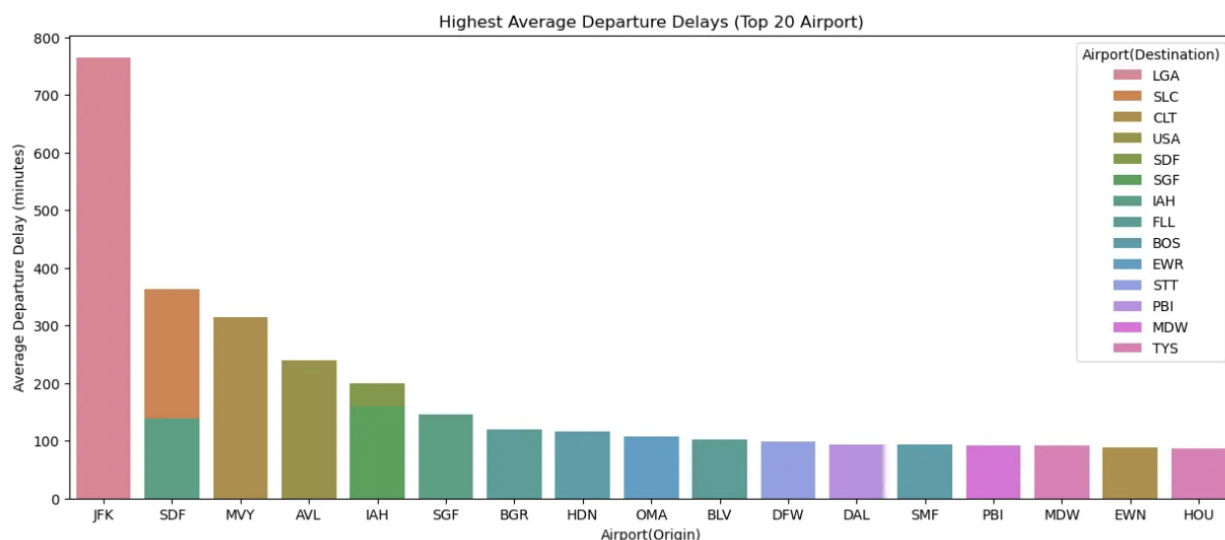


Рисунок 1.7 – График распределения по средним задержкам по аэропортам

Аэропорты, такие как John F. Kennedy (JFK), зафиксировали наивысшие показатели — свыше 725 минут как для вылетов, так и для прилётов.

1.2.6 Распределение задержек по причинам

На рисунке 1.8 представлен график распределения задержек по причинам.

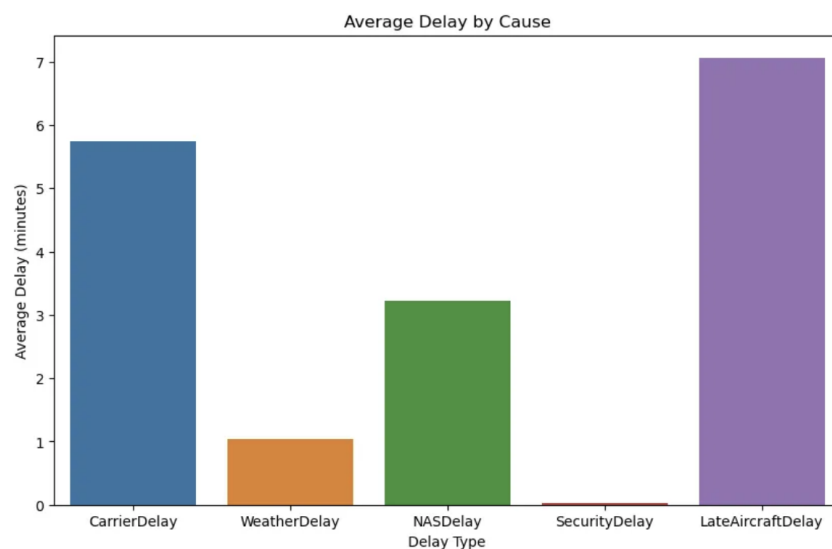


Рисунок 1.8 – График распределения задержек по причинам

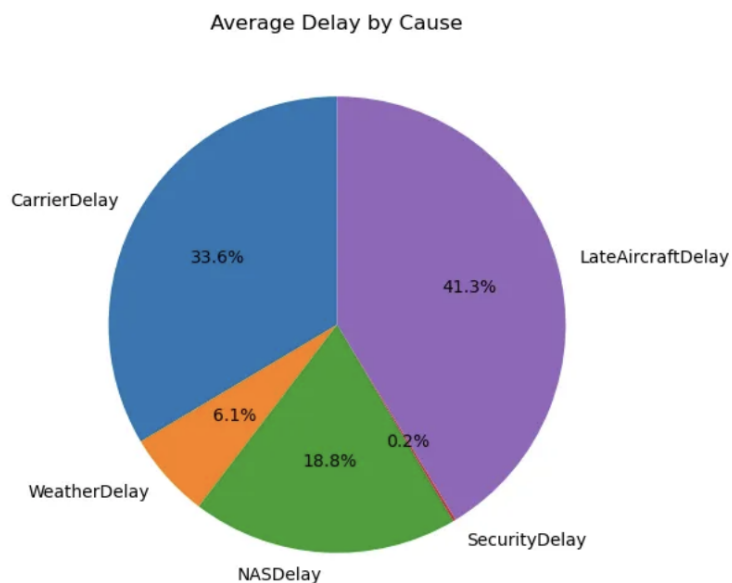


Рисунок 1.9 – График распределения задержек по причинам

Основными факторами задержек являются:

- проблемы авиакомпаний — наибольший вклад;
- задержки Национальной авиационной системы;
- погодные условия;
- задержки из-за позднего прибытия воздушного судна;
- проблемы безопасности.

1.2.7 Распределение задержек в зависимости от дальности перелёта

На рисунке 1.10 представлен график распределения задержек в зависимости от дальности перелёта.

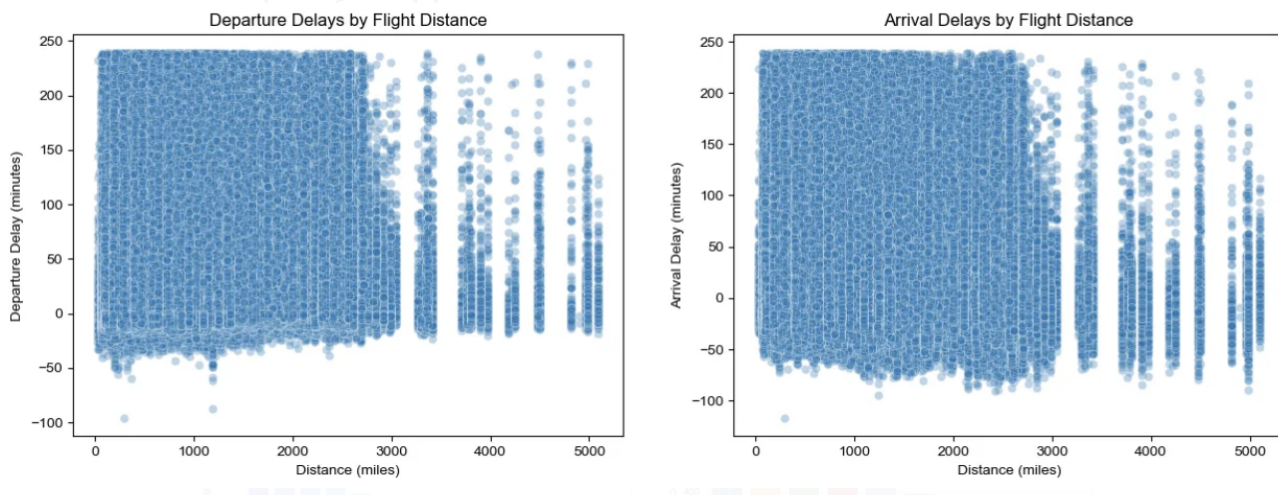


Рисунок 1.10 – График распределения задержек в зависимости от дальности перелёта

Короткие и средние рейсы (до 3000 миль) характеризуются наибольшей частотой задержек, тогда как длительные перелеты демонстрируют более высокую пунктуальность.

1.2.8 Распределение отклонённых рейсов по авиакомпаниям и аэропортам

На рисунке 1.11 представлен график распределения отклонённых рейсов по авиакомпаниям и аэропортам.

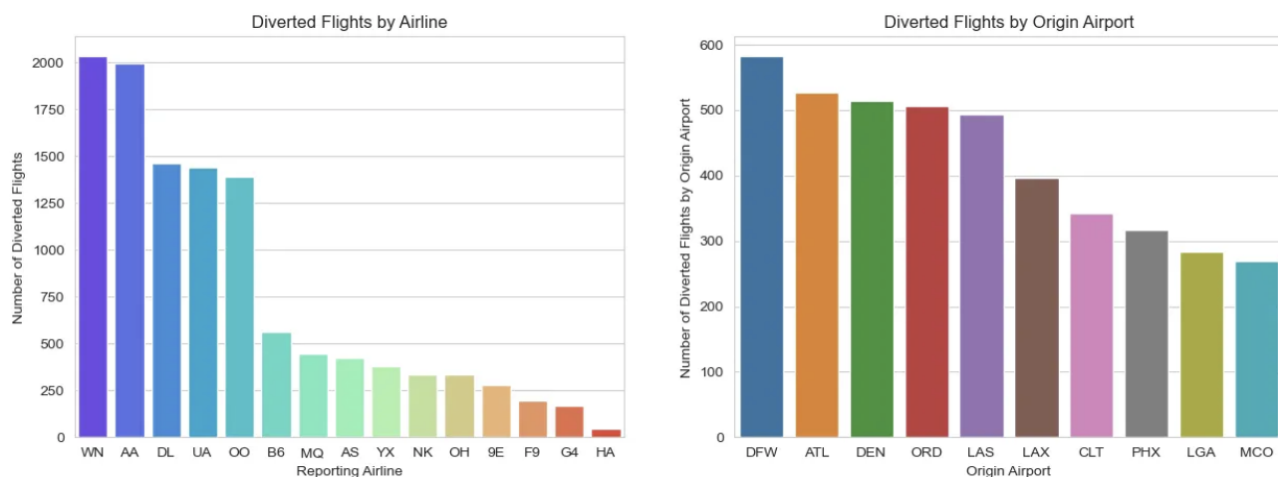


Рисунок 1.11 – График распределения отклонённых рейсов по авиакомпаниям и аэропортам

Наибольшее количество перенаправлений зафиксировано у Southwest (WN) и American Airlines (AA), особенно в перегруженных аэропортах, таких

как Даллас/Форт–Уэрт (DFW).

1.2.9 Карта корреляций по расстоянию полета

На рисунке 1.12 представлена карта корреляций по расстоянию полета.

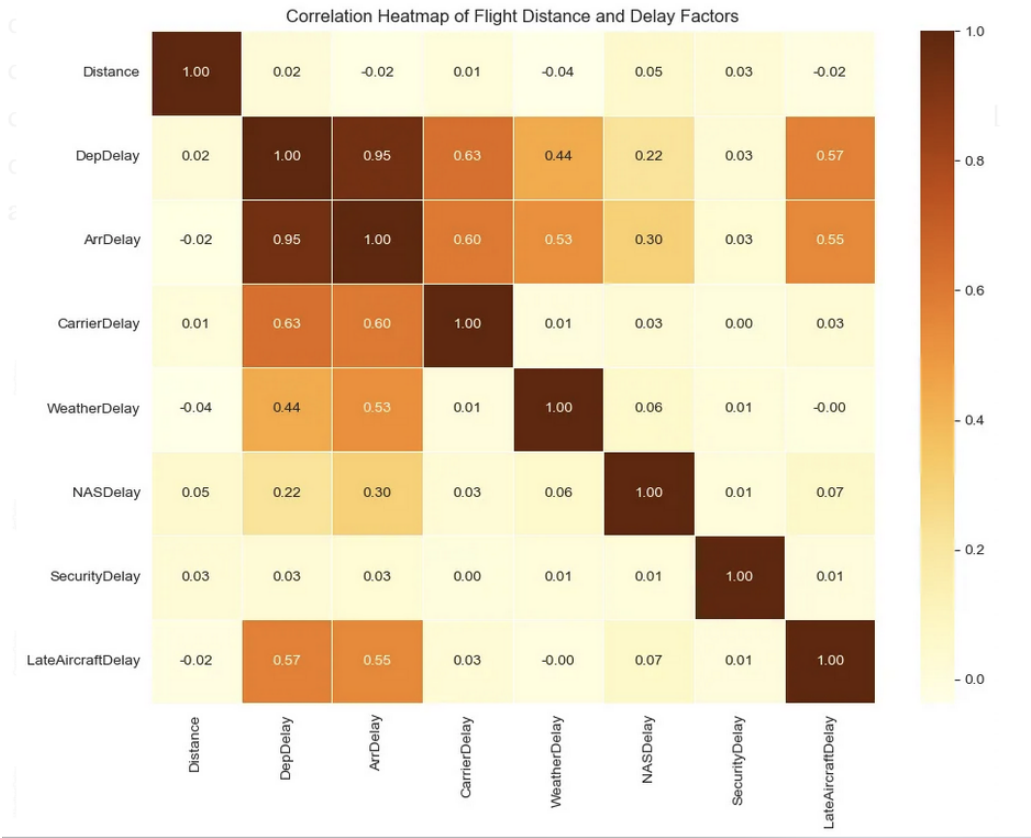


Рисунок 1.12 – Карта корреляций по расстоянию полета

На рисунке 1.12 представлен график рассеяния.

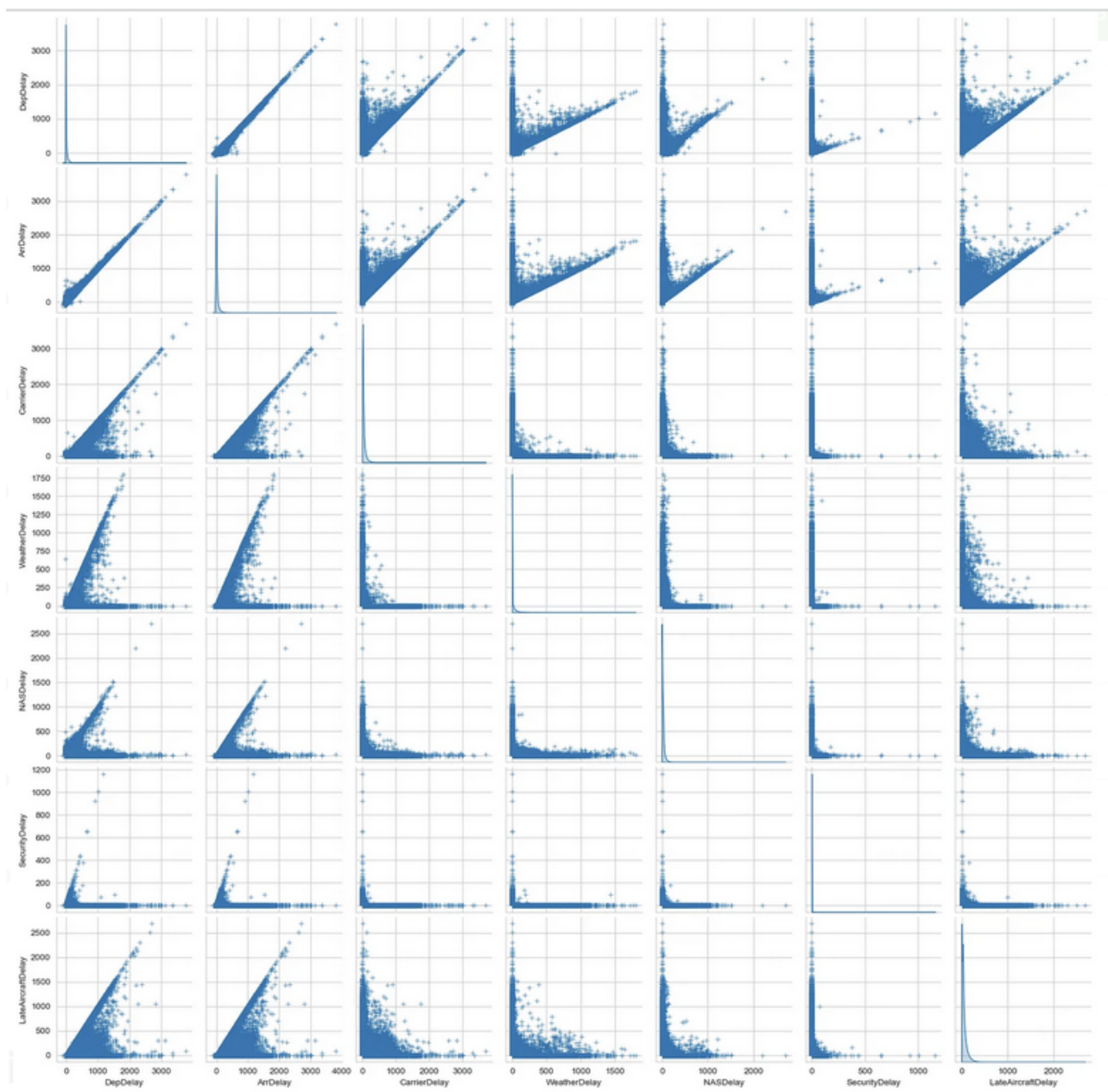


Рисунок 1.13 – График рассеяния

Сильная корреляция между задержками вылета и прибытия ($r=0,95$) подтверждает, что опоздание при вылете практически гарантирует задержку при посадке, в то время как умеренная связь с факторами авиакомпаний ($r=0,60-0,63$) указывает на возможность частичной компенсации таких задержек за счёт оперативных мер.

1.2.10 Распределение по основным причинам отмены рейсов

На рисунке 1.14 представлен график распределения по основным причинам отмены рейсов.

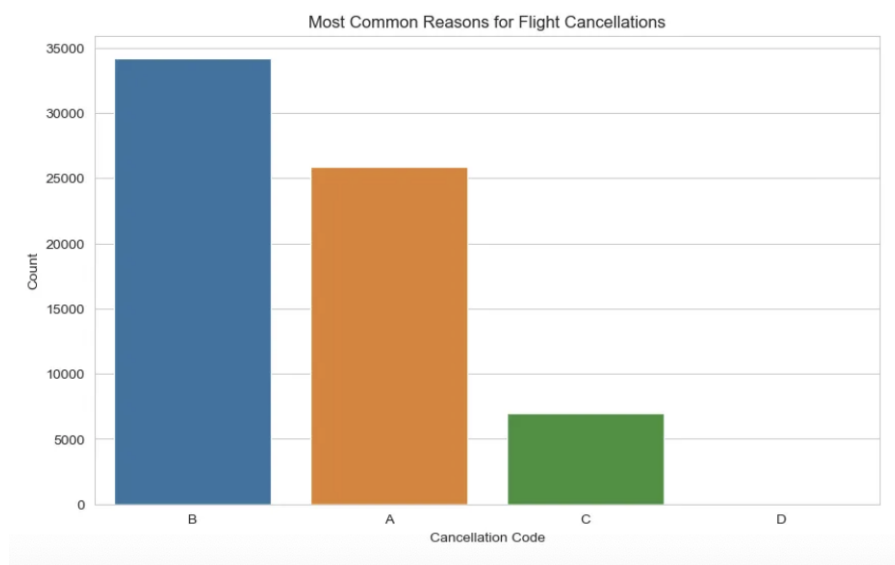


Рисунок 1.14 – График распределения по основным причинам отмены рейсов

Погодные условия (В) являются основной причиной отмен, за ними следуют проблемы, связанные с авиаперевозчиком (А). Анализ этих факторов позволяет разрабатывать более надежные стратегии управления рисками, включая улучшение логистики и систем оповещения.

Выводы

На основе анализа 3.5 миллионов записей о рейсах были выявлены ключевые закономерности. Установлена сильная корреляция ($r=0.95$) между задержками вылета и прибытия, что подтверждает важность контроля времени подготовки рейсов. Выделены аномальные случаи (5% рейсов с задержками свыше 4 часов), требующие особого подхода к прогнозированию. Географический анализ показал концентрацию 80% экстремальных задержек в 7% аэропортов.

1.3 Анализ существующих методов прогнозирования

1.3.1 Регрессия

Определение: Регрессия — это метод прогнозирования и анализа зависимости целевой переменной от одной или нескольких независимых переменных. Этот подход широко применяется в задачах предсказания, моделирования и объяснения зависимости переменных, позволяя строить аналитические модели, описывающие взаимодействия в сложных системах [16; 17].

Модель регрессии предсказывает числовое значение. Например, модель погоды, которая прогнозирует количество дождя в миллиметрах, является регрессионной моделью.

В таблице 1.1 приведены примеры регрессионных моделей.

Таблица 1.1 – Примеры регрессионных моделей

Сценарий	Входные данные	Выходные данные (числовой прогноз)
Прогнозирование стоимости жилья	<ul style="list-style-type: none"> — площадь участка — количество спален и ванных комнат — размер участка — процентная ставка по ипотеке — ставка налога на недвижимость — затраты на строительство — количество домов на продажу в районе 	Цена дома
Прогнозирование длительности поездки	<ul style="list-style-type: none"> — исторические данные о поездках — расстояние до пункта назначения — погодные условия 	Длительность поездки в минутах
Прогнозирование задержки рейса	<ul style="list-style-type: none"> — исторические данные о рейсах — погодные условия — загруженность аэропорта — время суток и день недели — техническое состояние воздушного судна 	Длительность задержки в минутах

Линейная регрессия

Определение: Линейная регрессия — это статистический метод, используемый для поиска взаимосвязи между переменными. В контексте машинного обучения линейная регрессия находит связь между функциями и меткой [18].

Линейная регрессия является одним из наиболее популярных алгоритмов и чаще всего используется для начала решения любой задачи регрессии, так как считается простейшей моделью машинного обучения [19].

В математической статистике линейная регрессия представляет собой метод аппроксимации зависимостей между входными и выходными переменными на основе линейной модели [20].

Если рассматривается зависимость между одной входной и одной выходной переменными, то имеет место простая линейная регрессия. Уравнение регрессии для этого случая имеет вид:

$$y = ax + b, \quad (1.1)$$

где a — коэффициент наклона (или угловой коэффициент) линии регрессии, b — свободный член (перехват с осью y).

Коэффициенты a и b , называемые также параметрами модели, определяются таким образом, чтобы сумма квадратов отклонений точек, соответствующих реальным наблюдениям данных, от линии регрессии была бы минимальной [20]. Коэффициенты обычно оцениваются методом наименьших квадратов по следующей формуле:

$$S(a, b) = \sum_{i=1}^n (y_i - (ax_i + b))^2, \quad (1.2)$$

где y_i — наблюдаемое значение зависимой переменной для i -го наблюдения, x_i — значение независимой переменной для i -го наблюдения, $ax_i + b$ — предсказанное значение зависимой переменной.

Чтобы найти коэффициенты a и b , минимизируем эту сумму по отношению к a и b . Обычно, для этого используются аналитические формулы,

полученные путем дифференцирования суммы квадратов остатков:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad (1.3)$$

$$b = \frac{1}{n} \sum_{i=1}^n y_i - a \frac{1}{n} \sum_{i=1}^n x_i, \quad (1.4)$$

где n — количество наблюдений.

Для оценки качества модели линейной регрессии часто используется коэффициент детерминации R^2 , который показывает долю изменчивости зависимой переменной, объясненную моделью. Он рассчитывается как квадрат коэффициента корреляции r_{xy} :

$$R^2 = r_{xy}^2, \quad (1.5)$$

где r_{xy} — коэффициент корреляции между x и y . Чем ближе R^2 к 1, тем лучше модель объясняет зависимость между переменными.

На рисунке 1.15 представлен пример построения линии регрессии.

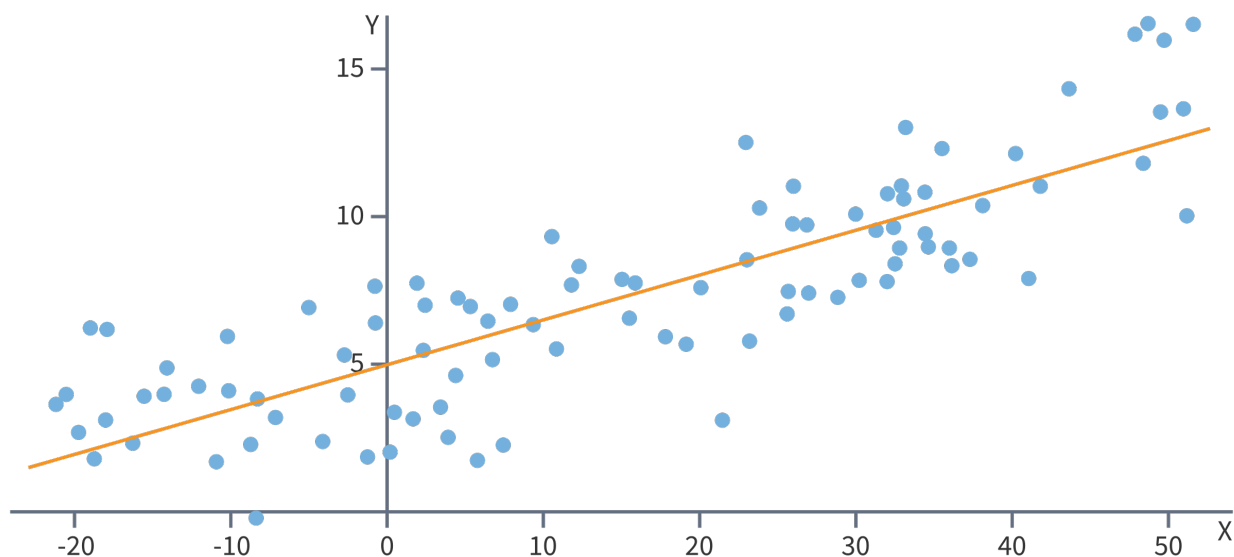


Рисунок 1.15 – Пример построения линии регрессии

Если рассматривается зависимость между несколькими входными и одной выходной переменными, то имеет место множественная линейная ре-

грессия. Соответствующее уравнение имеет вид (1.6):

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n, \quad (1.6)$$

где n — число входных переменных.

В данном случае модель будет описываться не прямой, а гиперплоскостью. Коэффициенты уравнения множественной линейной регрессии подбираются так, чтобы минимизировать сумму квадратов отклонения реальных точек от этой гиперплоскости [20].

Применение

Линейная регрессия имеет много практических применений, которые можно разделить на две основные категории:

- 1) прогнозирование — линейную регрессию можно использовать для подгонки модели к наблюдаемому набору данных;
- 2) объяснение изменчивости — линейный регрессионный анализ применяется для количественной оценки силы взаимосвязи между выходной и входными переменными.

Оценим практическое применение способа построения линейной регрессии в экономике на примере формирования заработной платы, зависящей от показателя среднедушевого прожиточного минимума на человека, на основе представленных способов и формул. В таблице 1.2 представлены данные, на базе которых нужно выявить зависимость показателя заработной платы от фактора x по регионам России за 2016 год:

Таблица 1.2 – Показатели среднемесячной заработной платы под влиянием прожиточного минимума по регионам РФ за 2016 г., тыс. руб.

№	Среднедушевой прожиточный минимум на одного работающего (x)	Среднемесячная заработная плата (y)
1	8.70	17.8
2	6.30	11.6
3	7.89	15.8
4	10.24	13.5
5	10.25	20.5
6	7.50	15.9
7	8.75	14.9
8	6.20	10.3
9	9.86	18.6
10	8.50	14.2

С помощью представленных данных построим линейное уравнений простой регрессии. Для этого необходимо просчитать коэффициенты a и b . Для упрощения расчетов построим вспомогательную таблицу 1.3, в которой Σ обозначает сумму элементов, а \bar{x} — среднее значение элементов:

Таблица 1.3 – Расчетная таблица параметров линейного уравнения регрессии

№	x	y	xy	x^2	y^2	\hat{y}	$y - \hat{y}$	$(y - \hat{y})^2$	A (%)
1	8.70	17.8	154.86	75.69	316.84	15.74	2.06	4.26	12
2	6.30	11.6	73.08	39.69	134.56	12.09	-0.49	0.24	4
3	7.89	15.8	124.66	62.25	249.64	14.51	1.29	1.67	8
4	10.24	13.5	138.24	104.8	182.25	18.08	-4.58	20.95	34
5	10.25	20.5	210.13	105.1	420.25	18.09	2.41	5.80	12
6	7.50	15.9	119.25	56.25	252.81	13.91	1.99	3.95	12
7	8.75	14.9	130.38	76.56	222.01	15.81	-0.91	0.83	6
8	6.20	10.3	63.86	38.44	106.09	11.94	-1.64	2.68	16
9	9.86	18.6	183.4	97.22	345.96	17.50	1.10	1.21	6
10	8.50	14.2	120.7	72.25	201.64	15.43	-1.23	1.52	9
Σ	84.19	153.1	1318.6	728.3	2432.1	153.1	0.0	43.1	119
\bar{x}	8.42	15.31	131.9	72.83	243.2	15.31	0.0	4.31	22

Для расчета коэффициентов a и b воспользуемся формулами:

— b (коэффициент наклона), вычисляется по формуле:

$$b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})}, \quad (1.7)$$

— a (свободный член), вычисляется через средние значения по формуле:

$$a = \bar{y} - b \cdot \bar{x}, \quad (1.8)$$

Подставим значения в формулы:

$$\bar{x} = \frac{\sum x}{n} = \frac{84.19}{10} = 8.42, \quad (1.9)$$

$$\bar{y} = \frac{\sum y}{n} = \frac{153.1}{10} = 15.31, \quad (1.10)$$

$$b = \frac{1318.55}{728.27} = 1.81, \quad (1.11)$$

$$a = 15.31 - 1.81 \cdot 8.42 = 2.52. \quad (1.12)$$

Таким образом, уравнение регрессии имеет вид:

$$y = 2.52 + 1.81x. \quad (1.13)$$

Вычисляем для уравнения коэффициент детерминации:

$$R^2 = \frac{(\sum (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2} = \frac{1318.55^2}{728.27 \cdot 2432.05} = 0.51. \quad (1.14)$$

Это означает, что 51% изменчивости зарплаты можно объяснить изменчивостью прожиточного минимума. Остальные 49% объясняются другими факторами [21].

Логистическая регрессия

Определение: Логистическая регрессия — это статистическая модель, используемая для предсказания вероятности возникновения некоторого события путем подгонки данных к логистической кривой [22].

Логистическая регрессия применяется для предсказания вероятности возникновения некоторого события по значениям множества признаков. Для этого вводится так называемая зависимая переменная y , принимающая лишь одно из двух значений, как правило, это числа: 0 (событие не произошло), и 1 (событие произошло), и множество независимых переменных (также называемых признаками, предикторами или регрессорами) — вещественных x_1, x_2, \dots, x_n , на основе значений которых требуется вычислить вероятность принятия того или иного значения зависимой переменной [22].

Стандартная логистическая функция, также известная как сигмовидная функция (*сигмовидная* означает «s-образная»), имеет формулу (1.15):

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (1.15)$$

На рисунке 1.16 показан соответствующий график сигмовидной функции. По мере увеличения входного значения x выходной сигнал сигмовидной функции приближается, но никогда не достигает 1. Точно так же, когда входные данные уменьшаются, выходные данные сигмовидной функции приближаются, но никогда не достигают 0.

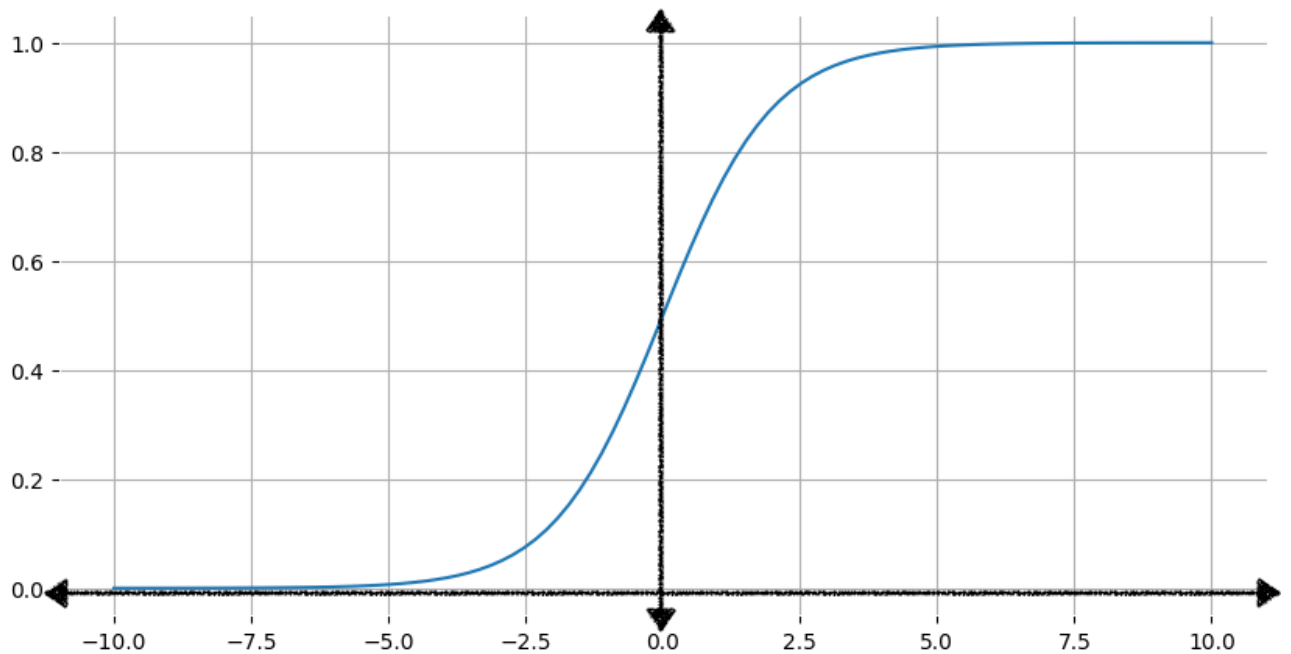


Рисунок 1.16 – График сигмовидной функции

Линейный компонент модели логистической регрессии описывается следующим уравнением (1.16):

$$z = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n, \quad (1.16)$$

где n — число входных переменных, z — результат линейного уравнения (логарифм шансов), b_i — коэффициент регрессии для i -го признака, x_i — значения признаков.

Чтобы получить прогноз логистической регрессии, значение z затем передается сигмовидной функции, что дает значение (вероятность) от 0 до 1 (формула 1.17):

$$y' = \frac{1}{1 + e^{-z}}, \quad (1.17)$$

где y' — результат модели логистической регрессии, z — линейный выход (рассчитанный в уравнении 1.16).

На рисунке 1.17 показано как линейный результат преобразуется в результат логистической регрессии.

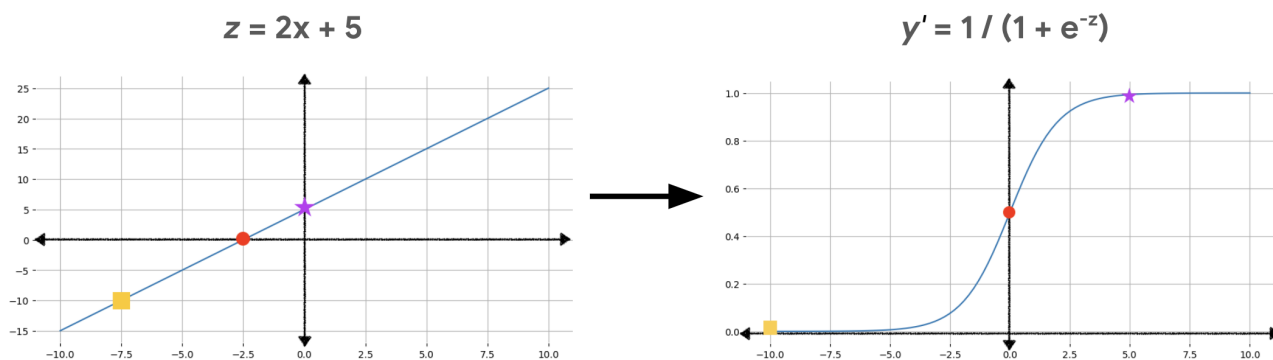


Рисунок 1.17 – Слева: график линейной функции $z = 2x + 5$, выделены три точки. Справа: сигмовидная кривая с теми же тремя точками, выделенными после преобразования сигмовидной функцией

В качестве функции потерь в линейной регрессии используется метод наименьших квадратов (квадрат потерь). Этот метод подходит для линейной модели, где скорость изменения выходных значений постоянна. Например, для линейной модели $y' = b + 3x_1$ каждый раз, когда увеличивается входное значение x_1 на 1, выходное значение y' увеличивается на 3 [23].

Однако скорость изменения модели логистической регрессии не является постоянной. Когда значение логарифма шансов (z) ближе к 0, небольшое увеличение z приводит к гораздо большим изменениям y , чем когда z является большим положительным или отрицательным числом.

В таблице 1.4 показаны выходные данные сигмовидной функции для входных значений от 6 до 10, а также соответствующая точность, необходимая для учета различий в результатах.

Таблица 1.4 – Выходные данные сигмовидной функции

Входное значение (x)	Выходное значение ($\sigma(x)$)	Требуемая точность
6	0.997	3
7	0.999	3
8	0.9997	4
9	0.9999	4
10	0.99998	5

В данном случае нельзя успешно применить метод наименьших квадратов для оценки параметров b и построения прогнозов, так как в этом случае

прогнозные значения вероятности могут принимать как отрицательные значения, так и значения больше единицы [24]. В качестве альтернативы для оценки коэффициентов модели используют метод максимального правдоподобия, который заключается в оценивании параметров путем максимизации функции правдоподобия.

Положительный коэффициент говорит о том, что данный фактор увеличивает общий риск, то есть повышает вероятность анализируемого исхода. Отрицательный коэффициент означает, что данный фактор уменьшает риск, то есть понижает вероятность наступления исхода [24].

Определение: Метод максимального правдоподобия — еще один способ построения оценки неизвестного параметра. Состоит он в том, что в качестве «наиболее правдоподобного» значения берут значение Θ , максимизирующее вероятность получить при n опытах данную выборку $X = (X_1, X_2, \dots, X_n)$ [24]. Это значение параметра Θ зависит от выборки и является искомой оценкой.

Формула функции правдоподобия имеет вид:

$$f(X, \Theta) = f_{\Theta}(X_1) \cdot f_{\Theta}(X_2) \cdot \dots \cdot f_{\Theta}(X_n) = \prod_{i=1}^n f_{\Theta}(X_i). \quad (1.18)$$

Формула логарифма правдоподобия имеет вид:

$$L(X, \Theta) = \ln f(X, \Theta) = \sum_{i=1}^n \ln f_{\Theta}(X_i). \quad (1.19)$$

Применение

Согласно проведенному анализу современных исследований, связанных с использованием логистической регрессии, было выявлено несколько особенностей. Во-первых, применение этой модели наиболее распространено в социально-экономических и медицинских исследованиях, хотя есть опыт применения в работах технического характера. Во-вторых, с помощью этой модели решают три типа задач: прогнозирование, классификация и оценка влияния факторов на исход [25].

При использовании логистической регрессии возникают следующие трудности:

- 1) ошибка соотнесения с классом значения — возникает при классифика-

ции объектов, которые близки к границе класса (вероятность близка к 0.5). Например, в [26] модель логистической регрессии используется для решения задачи кредитного скоринга. Задача заключается в классификации клиентов банка на два класса: надежные и ненадежные. Автор отмечает, что клиентов, для которых вероятность возвращения кредита близка к 0.5, невозможно классифицировать однозначно;

- 2) мультиколлинеарность — возникает, когда два или более предиктора в модели линейно зависимы. Например, в [27] отмечается, что пренебрежение зависимостями между независимыми переменными ведет к построению ошибочных моделей. Автор предлагает проводить предварительное статистическое исследование и исключать такие переменные из анализа;
- 3) несбалансированные данные — возникает, когда один из классов в обучающей выборке представлен значительно меньшим количеством объектов. Например, в работе [28] автор отмечает, что при обучении модели на выборке с неравномерным распределением классов значений зависимой переменной была получена низкая точность прогноза. После «выравнивания» выборки и повторного обучения модель показала 99% точности;
- 4) проблема «границ чувствительности» — возникает, когда модель логистической регрессии не может корректно оценить вероятность принадлежности объекта к классу. Например, в кредитном скоринге клиенты в возрасте 30 и 31 год практически одинаковые группы, а клиенты с возрастом 60 и 61 год — весьма разные группы заемщиков [29].

Тем не менее большинство исследований показывают эффективность модели логистической регрессии. Кроме высокой точности прогнозирования, стоит отметить ее достоинство решать задачи различного масштаба: от 3 независимых [30] переменных до 230 [31], от 40 записей [32] до 20 миллионов [28], а построенные модели понятны для интерпретации.

Несмотря на существующие ограничения, модель логистической регрессии показывает высокую точность прогнозирования и широкий спектр

применения в различных предметных областях.

Адаптивная регрессия

Определение: Адаптивная регрессия — это метод статистического моделирования, который использует функции для представления нелинейных зависимостей между переменными. В отличие от линейных моделей, адаптивная регрессия может автоматически подстраиваться под данные, включая нелинейные связи и взаимодействия между предикторами, что позволяет улучшить точность предсказаний [33].

Один из популярных методов адаптивной регрессии для выявления нелинейных связей в данных — это многомерные адаптивные регрессионные сплайны (МАРС) [33].

МАРС — это статистическая процедура, позволяющая решить классическую задачу регрессии: установить вид и параметры аппроксимирующей функции, описывающую функциональную зависимость отдельных наблюдений (исходные данные) с указанной точностью [34]. Пространство значений входных переменных разбивается на области со своими собственными уравнениями базисных функций. Это позволяет использовать МАРС даже в случае задач с «проклятием размерности», когда высокая размерность пространства значений входных переменных ограничивает применимость иных статистических процедур.

Метод МАР-сплайнов не имеет ограничения, характерного для иных статистических методов, в части наличия исходных предположений о типе зависимостей (линейных, степенных, экспоненциальных) между предикторными и выходными переменными [34]. Кроме того, метод МАРС чувствителен к изменению вида связи между предиктором и откликом, будь то: изменение формы связи (например, от линейной к степенной), добавление или вычитание некоторой константы для прогноза отклика справа от узловой точки предиктора, изменение наклона регрессионной функции [34].

Подобные особенности сплайнов достигаются за счет использования следующих базисных функций особого вида:

$$(x - t)_+ = \begin{cases} x - t, & x > t, \\ 0, & \text{иначе,} \end{cases} \quad (1.20)$$

$$(t - x)_- = \begin{cases} t - x, & x < t, \\ 0, & \text{иначе,} \end{cases} \quad (1.21)$$

где t — точка разрыва (узел). Этот метод оценивает каждую точку данных для каждого предиктора в качестве узла и создаёт линейную регрессионную модель с выбранной(ыми) переменной(ыми).

Формулы вида (1.20 — 1.21) можно представить в следующем виде соответственно:

$$(x - t)_+ = \max(0, x - t), \quad (1.22)$$

$$(t - x)_+ = \max(0, t - x). \quad (1.23)$$

В многомерном случае для каждой компоненты x_j вектора предикторов $x = (x_1, x_2, \dots, x_n)$ строятся базисные функции вида (1.20 — 1.21) с узлами в каждой наблюдаемой переменной x_{ij} , где $i = \overline{1, 2, \dots, n}$ и $j = \overline{1, 2, \dots, m}$.

Общее уравнение МАР–сплайнов для модели из M членов, отличных от константы, представляет собой взвешенную сумму базисных функций и их произведений и записывается в виде (1.20):

$$y = f(x) = b_0 + \sum_{m=1}^M b_m h_m(x), \quad (1.24)$$

где b_0 — свободный член, b_m — коэффициенты регрессии, определяемые методом наименьших квадратов, $h_m(x)$ — базисная функция, M — число базисных функций.

Основной принцип работы модели состоит в выборе нужной взвешенной суммы базисных функций из общего набора базисных функций, покрывающих все значения каждого предиктора (т.е. набор будет состоять из одной базисной функции и параметра t для каждого отдельного значения каждой предикторной переменной). Алгоритм МАР–сплайнов отыскивает в пространстве всех входных и предикторных переменных расположение узловых точек, а также взаимосвязи между переменными [34]. В процессе поиска число добавленных к модели базисных функций из общего набора возрастает до тех пор, пока не будет максимизирован общий критерий качества модели — обобщенное

скользящее среднее, который имеет следующий вид:

$$GCV(M) = \frac{1}{n} \frac{\sum_{i=1}^n (y_i - f(x_i))^2}{(1 - \frac{C}{n})^2}, \quad (1.25)$$

где $GCV(M)$ — критерий точности модели, отражающий рост дисперсии с ростом числа базисных функций, C — эффективное число параметров модели, которое в общем случае определяется как $C = 2K - 1$, где K — общее число параметров модели.

Пошаговый алгоритм построения МАР–сплайна схож с алгоритмом линейной регрессии, только вместо регрессионных функций используются базисные функции. Например, рассмотрим нелинейные, немонотонные данные, где $y = f(x)$. Процедура МАРС сначала ищет одну точку в диапазоне значений x , где две различные линейные зависимости между y и x минимизируют ошибку методом наименьших квадратов [35].

Результатом становится функция $h(x - a)$, где a — это значение точки разрыва. Для одного узла функция $h(x - a)$ выглядит как $h(x - 1.183606)$, следовательно, две линейные модели для y имеют вид:

$$y = \begin{cases} \beta_0 + \beta_1(1.183606 - x) & x < 1.183606, \\ \beta_0 + \beta_1(x - 1.183606) & x > 1.183606 \end{cases} \quad (1.26)$$

После нахождения первого узла поиск продолжается для второго узла, который обнаруживается при $x = 4.898114$. Это приводит к созданию трёх линейных моделей для y :

$$y = \begin{cases} \beta_0 + \beta_1(1.183606 - x) & x < 1.183606, \\ \beta_0 + \beta_1(x - 1.183606) & x > 1.183606 \quad \& \quad x < 4.898114, \\ \beta_0 + \beta_1(4.898114 - x) & x > 4.898114 \end{cases} \quad (1.27)$$

На рисунке 1.18 показан пример адаптивного регрессионного сплайна с одним, двумя, тремя и четырьмя узлами соответственно.

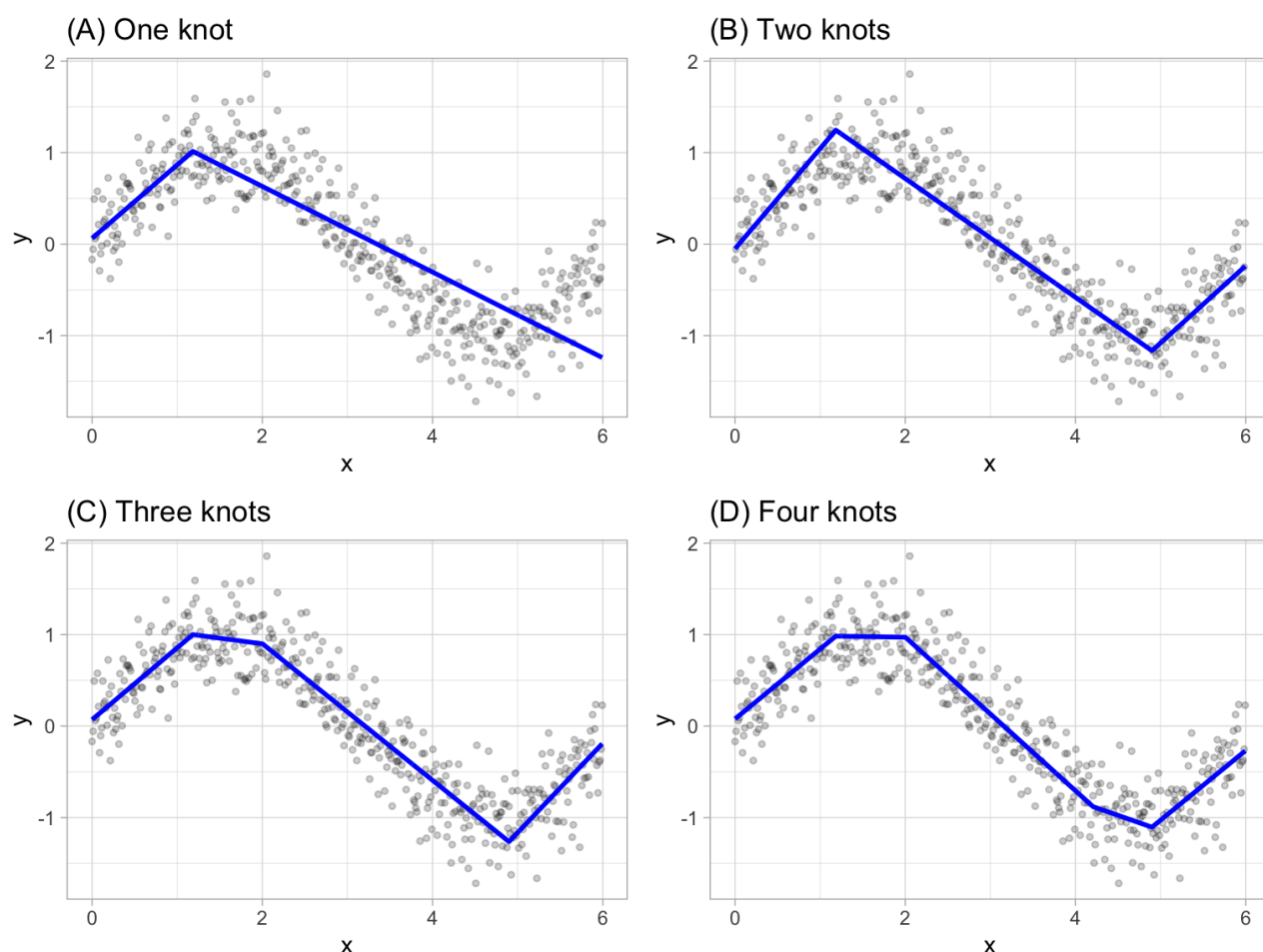


Рисунок 1.18 – Пример адаптивного регрессионного сплайна

Этот процесс продолжается до тех пор, пока не будет найдено множество узлов, что приводит к созданию точного нелинейного уравнения предсказания. Добавление большого количества узлов может позволить модели полностью соответствовать обучающим данным, но привести к недостаточной обобщаемости на новые, ранее не виденные данные.

Поэтому, после того как полный набор узлов найден, следует поочередно удалять узлы, которые не вносят значительного вклада в точность предсказания [35].

Применение

МАР-сплайны находят свое применение во многих сферах науки и технологий, например, в предсказании видов распределений по имеющимся данным [36], кишечного поглощения лекарств [37], а также в воспроизведении речи [38] и поиске глобального оптимума в проектировании конструкций [39].

Многомерные адаптивные регрессионные сплайны обладают рядом пре-

имуществ перед другими регрессионными методами [40]:

- 1) модели, построенные с использованием МАР–сплайнов, обладают большей гибкостью, чем модели, построенные при помощи линейной регрессии;
- 2) МАР–сплайны могут автоматически находить нелинейные зависимости между переменными, что позволяет улучшить точность прогнозов;
- 3) МАР–сплайны позволяют работать с численными и категориальными признаками;
- 4) благодаря разделению исходных данных на области базисными функциями, МАР–сплайны позволяют определять выбросы;
- 5) МАР–сплайны не требуют значительных мер по подготовке входных данных;
- 6) метод демонстрирует высокую устойчивость к многоколлинеарности, что особенно важно при работе с большими наборами данных;

Благодаря своим преимуществам, МАР–сплайны нашли применение в биоинформатике для анализа геномных данных, включая прогнозирование экспрессии генов и идентификацию биомаркеров. Метод активно используется в экологии для моделирования и прогнозирования ареалов обитания видов с учетом множества факторов окружающей среды, в экономике и финансах — для прогнозирования цен, анализа временных рядов и моделирования волатильности; В задачах управления производственными процессами МАР–сплайны помогают оптимизировать технологические параметры и контролировать качество продукции.

1.3.2 Сравнительный анализ методов прогнозирования задержек авиарейсов

В данной работе проведен сравнительный анализ методов прогнозирования задержек авиарейсов, включая линейную и логистическую регрессию, а также адаптивную регрессию, который представлен в таблице 1.5.

Таблица 1.5 – Сравнительный анализ методов прогнозирования задержек авиарейсов

Критерий	Линейная	Логистическая	Адаптивная	Итеративно-адапт.
Учет нелинейных зависимостей	Нет	Частично	Да	Да
Работа с пространственно-временными данными	Частично	Частично	Да	Да
Интерпретируемость	Высокая	Средняя	Средняя	Средняя
Стабильность прогнозов	Низкая	Средняя	Высокая	Высокая
Обучение на новых данных	Нет	Нет	Да	Да
Поддержка итеративности	Нет	Нет	Да	Да
Результат	Число минут	Вероятность	Число минут	Число минут

Выводы

В подразделе проведен сравнительный анализ трех классов методов прогнозирования. Линейная регрессия продемонстрировала преимущества в простоте интерпретации, но оказалась неспособна учитывать нелинейные зависимости. Адаптивные методы (МАР–сплайны) показали лучшую точность за счет автоматического выявления сложных взаимосвязей. Сравнительный анализ выявил отсутствие готовых решений, сочетающих высокую точность с возможностью оперативной адаптации к изменяющимся условиям.

1.4 Постановка задачи

Выводы

В рамках данного раздела проведен комплексный анализ проблемы прогнозирования задержек авиарейсов. Были систематизированы ключевые факторы, влияющие на регулярность полетов, включая природно–климатические, технико–эксплуатационные и организационно–управленческие аспекты. Особое

внимание уделено пространственно–временным закономерностям, демонстрирующим устойчивую корреляцию между географической локацией, временными параметрами и вероятностью задержек.

Проведенный сравнительный анализ методов прогнозирования позволил выявить ограничения традиционных регрессионных подходов при работе с нелинейными зависимостями. Наиболее перспективными представляются адаптивные методы, способные учитывать комплексное взаимодействие множества факторов в изменяющихся условиях.

Полученные результаты создают теоретическую основу для разработки итеративного алгоритма прогнозирования, сочетающего преимущества регрессионного анализа с возможностями адаптивного моделирования. Это открывает перспективы для создания более точных систем предсказания задержек, учитывающих как статические параметры, так и динамику изменений операционной обстановки.

2 Конструкторский раздел

2.1 Требования и ограничения к разрабатываемому методу

К методу прогнозирования задержек рейсов на основе итеративной адаптации с учётом пространственно–временных факторов предъявляются следующие требования:

- 1) метод должен прогнозировать длительность задержки рейса в минутах с указанием доверительного интервала;
- 2) метод должен учитывать пространственно–временные факторы: географическое расположение аэропортов и временные характеристики рейсов;
- 3) метод должен обеспечивать итеративную адаптацию модели.

Также представлен ряд ограничений:

- 1) для корректной работы, метод опирается на исторические данные (последние 90 дней);
- 2) качество прогнозирования метода зависит от полноты исторических данных;
- 3) метод не учитывает форс–мажорные события.

2.2 Требования к разрабатываемому программному обеспечению

Программное обеспечение должно соответствовать следующим требованиям:

- 1) реализовать итеративно–адаптивное прогнозирование задержек авиарейсов с использованием комбинации пространственных и временных характеристик;
- 2) предоставлять пользовательский интерфейс для получения прогнозов по маршруту и дате с указанием уровня достоверности.

2.3 Основные этапы разрабатываемого метода

2.3.1 IDEF0-диаграмма уровня A1

На рисунке 2.1 представлена диаграмма IDEF0 уровня A1, описывающая процесс прогнозирования задержек. Входные данные включают исторические записи, метеопараметры и операционные данные аэропортов. Управляющие воздействия — настройки модели. Выход — прогноз задержек с оценкой точности.

Рисунок 2.1 – IDEF0-диаграмма уровня A1 процесса прогнозирования задержек

2.3.2 Схемы алгоритмов

2.3.3 Структура разрабатываемого программного обеспечения

Выводы

В рамках данного раздела разработаны требования к методу прогнозирования задержек авиарейсов и соответствующему программному обеспечению. Основные этапы метода включают сбор данных, обучение адаптивной модели и интеграцию с внешними системами. Использование IDEF0-диаграмм и алгоритмических схем позволило визуализировать логику работы системы. Спроектированная архитектура программного обеспечения обеспечивает масштабируемость и возможность внедрения в реальную инфраструктуру аэропортов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. A turnaround control system [Электронный ресурс]. — — Режим доступа: (дата обращения: 02.03.2024) <https://www.sciencedirect.com/science/article/pii/S0952197622002056>.
2. Ноздрин С. С., Каширина И. Л. Построение модели предсказания длительности задержек авиарейсов [Электронный ресурс] // Сборник трудов Международной научной конференции. — 2021. — С. 1014—1021.
3. Turnaround Time: Why it's important [Электронный ресурс]. — — Режим доступа: (дата обращения: 02.03.2024) <https://simpleflying.com/turnaround-time-importance/>.
4. Операции по техническому обслуживанию самолетов [Электронный ресурс]. — — Режим доступа: (дата обращения: 25.05.2024) <https://iloencyclopedia.org/ru/part-xvii-65263/transport-industry-and-warehousing/item/910-aircraft-maintenance-operations>.
5. Mechanism Analysis of the Impact of COVID-19 on the Whole Process of Aircraft Turnaround Operations [Электронный ресурс]. — — Режим доступа: (дата обращения: 02.05.2024) https://www.researchgate.net/publication/379569382_Mechanism_Analysis_of_the_Impact_of_COVID-19_on_the_Whole_Process_of_Aircraft_Turnaround_Operations.
6. *International Civil Aviation Organization*. Meteorological Service for International Air Navigation. — 21st. — 2022.
7. *Program N. A. S.* Aviation Weather Decision Support Technologies : tex. отч. / NASA. — 2022. — NASA/TM-2022—220587.
8. *FAA Office of Aviation Safety*. Aviation Maintenance Delay Analysis / Federal Aviation Administration. — 2023. — URL: https://www.faa.gov/data_research/aviation_data_statistics.
9. *IATA Ground Operations Council*. Ground Operations Manual / International Air Transport Association. — 2023.
10. *Zhang L., Wang H.* Spatiotemporal Analysis of Flight Delays in U.S. Airports // Transportation Research Part A. — 2022. — Т. 156. — С. 1—15.

11. *EUROCONTROL Performance Review Unit*. Annual Delay Report 2022 / EUROCONTROL. — 2023. — URL: <https://www.eurocontrol.int/publication/annual-delay-report-2022>.
12. *International Civil Aviation Organization*. Meteorological Service for International Air Navigation. — 21st. — 2022.
13. *FAA Office of Aviation Safety*. Air Carrier Maintenance Delays and Interruptions / Federal Aviation Administration. — 2023. — URL: https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_120-115.pdf.
14. *FAA Aviation Safety*. Flightcrew Member Duty and Rest Requirements : тех. отч. / Federal Aviation Administration. — 2022.
15. *ICAO Council Aviation Recovery Task Force*. Guidance for Air Travel through the COVID-19 Public Health Crisis / International Civil Aviation Organization. — 2022.
16. *Сибер Г. А. Ф., Лу А. Д.* Анализ линейной регрессии. — Уайли (дата обращения: 26.11.2024), 2012. — С. 456.
17. *Монтгомери Д. К., Пек Э. А., Вининг Г. Г.* Введение в анализ линейной регрессии. — Wiley (дата обращения: 26.11.2024), 2021. — С. 640.
18. Словарь машинного обучения [Электронный ресурс]. — — Режим доступа: (дата обращения: 16.11.2024) <https://developers.google.com/machine-learning/glossary>.
19. Линейная регрессия [Электронный ресурс]. — — Режим доступа: (дата обращения: 21.11.2024) <https://elibrary.ru/item.asp?id=49547912>.
20. Линейная регрессия [Электронный ресурс]. — — Режим доступа: (дата обращения: 21.11.2024) <https://wiki.loginom.ru/articles/linear-regression.html>.
21. *Л.П. Г.* Особенности линейной регрессии и ее применение в экономике. — 2020.
22. Логистическая регрессия [Электронный ресурс]. — — Режим доступа: (дата обращения: 21.11.2024) <https://elib.sfu-kras.ru/bitstream/handle/2311/7199/s021-083.pdf?sequence=1>.

23. Логистическая регрессия [Электронный ресурс]. — — Режим доступа: (дата обращения: 21.11.2024) <https://developers.google.com/machine-learning/crash-course/logistic-regression>.
24. Применение логистической регрессии для оценки кредитного риска [Электронный ресурс]. — — Режим доступа: (дата обращения: 02.12.2024) <https://elibrary.ru/item.asp?id=27478789>.
25. О возможностях применения модели логистической регрессии [Электронный ресурс]. — — Режим доступа: (дата обращения: 02.12.2024) <https://www.elibrary.ru/item.asp?id=25090019>.
26. *Симонов П., Лазуков С.* Оценка кредитного риска: актуальные практические вопросы // Вестник Пермского университета. Сер. Экономика. — 2009. — № 1. — С. 61—67.
27. *Мурадов Д.* Logit-прогностические модели прогнозирования банкротства предприятий // Труды Российского государственного университета нефти и газа им. И.М. Губкина. — 2011. — № 3. — С. 160—172.
28. *Середний С.* Оценивание вероятности дефолта по кредитным операциям с использованием логистической регрессии и кластерного анализа // Достижения информационных технологий. — 2011. — № 1. — С. 126—132.
29. *Сапонов Д.* Опыт конкурентной борьбы как фактор академической успеваемости // Мониторинг общественного мнения: экономические и социальные перемены. — 2013. — 6 (118). — С. 113—126.
30. *Симонова С.* Интеллектуальный анализ данных для задач СЕМ // International Journal of Open Information Technologies. — 2015. — Т. 3, № 2. — С. 17—22.
31. *Осиков М., Ахматов К.* Использование логистической регрессии в оценке изменений психологического статуса у больных хронической почечной недостаточностью, находящихся на гемодиализе // Вестник Южно-Уральского государственного университета. Сер. Образование, здравоохранение, физическая культура. — 2010. — 19 (195). — С. 34—37.
32. *Богданов Л.* Оценка эффективности бинарных классификаторов на основе логистической регрессии методом КОС-анализа // Вестник СГТУ. — 2010. — Т. 4, 2с. — С. 92—97.

33. *Фридман Д. Х.* Многомерные Адаптивные Регрессионные Сплаины // Ежегодный журнал статистики. — 1991. — Т. 19, № 1.
34. К вопросу применимости аппарата МАРС к задаче прогнозирования банкротства: зарубежный опыт [Электронный ресурс]. — — Режим доступа: (дата обращения: 02.12.2024) https://www.imi-samara.ru/wp-content/uploads/2018/04/11_Romanova_91-98.pdf.
35. Адаптивная регрессия [Электронный ресурс]. — — Режим доступа: (дата обращения: 26.11.2024) <https://bradleyboehmke.github.io/HOML/mars.html>.
36. *Элит Д., Летвик Д.* Прогнозирование распространения видов на основе музейных и гербарных записей с использованием многомерных адаптивных регрессионных сплайнов // Разнообразие и распределение. — 2007. — Т. 13, № 3. — С. 265—275.
37. *Деконинк Э., Коомонс Д., Хейден Й.* Исследование методов линейного моделирования и их комбинаций с многомерными адаптивными регрессионными сплайнами для прогнозирования желудочно-кишечной абсорбции лекарств // Журнал фармацевтического и биомедицинского анализа. — 2007. — Т. 43, № 1. — С. 119—130.
38. *Хаас Х., Кубин Г.* Многополосная нелинейная осцилляторная модель для речи // Сборник материалов 32-й конференции Asilomar по сигналам, системам и компьютерам. Т. 1. — 1998. — С. 338—342.
39. *Крино С., Браун Д.* Глобальная оптимизация с использованием многомерных адаптивных регрессионных сплайнов // Труды IEEE: Системы, человек и кибернетика, часть В: Кибернетика. — 2007. — Т. 37, № 2. — С. 333—340.
40. *С. Ф. А.* Модели потребления электроэнергии в многоквартирных жилых домах на основе многомерных адаптивных регрессионных сплайнов [Электронный ресурс]. — 2011. — — Режим доступа: (дата обращения: 02.12.2024) https://elibrary.ru/download/elibrary_26342121_69582650.pdf.