|  |
| --- |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА − Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ)**

**Кафедра прикладной математики (ПМ)**

**КУРСОВАЯ РАБОТА**

по дисциплине: «Языки программирования для статистической   
обработки данных»

**Тема курсовой работы**: «Разработка программы для прогнозирования временного ряда на основе линейных авторегрессионных моделей на основе данных статусов авиаперевозок»

|  |  |  |
| --- | --- | --- |
| Студент группы ИМБО-02-22 | Ким Кирилл Сергеевич | (подпись) |
| Руководитель  курсовой работы | старший преподаватель,  Юрченков И.А. | (подпись) |
| Работа представлена к защите | «\_\_\_»\_\_\_\_\_\_\_\_\_\_\_2024 г. |  |
| Допущен к защите | «\_\_\_»\_\_\_\_\_\_\_\_\_\_\_2024 г. |  |

Москва 2024 г.

|  |
| --- |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА − Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ)**

**Кафедра прикладной математики (ПМ)**

|  |  |  |
| --- | --- | --- |
|  |  | Утверждаю |
|  |  | И.о. заведующего кафедрой ПМ |
|  |  | \_\_\_\_\_\_\_\_\_\_\_\_\_\_Смоленцева Т.Е.  (подпись) |
|  |  | «2» февраля 2024 г. |

**ЗАДАНИЕ**

**на выполнение курсовой работы**

по дисциплине «Языки программирования для статистической обработки данных»

|  |  |  |  |
| --- | --- | --- | --- |
| Студент Ким Кирилл Сергеевич | | | Группа ИМБО-02-22 |
| **Тема** «Разработка программы для прогнозирования временного ряда на основе линейных авторегрессионных моделей на основе данных статусов авиаперевозок» | | | |
| **Исходные данные:** выбранная студентом задача и алгоритм её решения, а также набор данных | | | |
| **Перечень вопросов, подлежащих разработке, и обязательного графического материала:**  Описание решаемой задачи машинного обучения или статистической обработки данных (математическая формулировка, проблематика, существующие способы решения)  Анализ выбранного алгоритма или метода решения выбранной задачи  Выбор данных и описание набора данных, описание предикторов и целевых факторов  Построение сценария и логики обработки данных на основе выбранного алгоритма или метода Оценка качества решения задачи на основе метрик качества | | | |
| **Срок представления к защите курсовой работы: до** «24» мая 2024 г. | | | |
| **Задание на курсовую работу выдал** | *Подпись руководителя* | Юрченков И.А.  *(ФИО руководителя)* | |
| «2» февраля 2024 г. | | | |
| **Задание на курсовую работу получил** | *Подпись обучающегося* | Ким К.С.  *(ФИО обучающегося)* | |
| «2» февраля 2024 г. | | | |

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 4](#_Toc163818181)

[1 ТЕОРЕТИЧЕСКАЯ ЧАСТЬ 5](#_Toc163818182)

[1.1 Основные понятия временных рядов 5](#_Toc163818183)

[1.2 Авторегрессионная модель 7](#_Toc163818184)

[2 ПРАКТИЧЕСКАЯ ЧАСТЬ 8](#_Toc163818185)

[2.1 8](#_Toc163818186)

[2.2 8](#_Toc163818187)

[ЗАКЛЮЧЕНИЕ 9](#_Toc163818188)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 10](#_Toc163818189)

[Теоретическая часть 10](#_Toc163818190)

[Практическая часть 10](#_Toc163818191)

[ПРИЛОЖЕНИЯ 11](#_Toc163818192)

[Листинг 11](#_Toc163818193)

# ВВЕДЕНИЕ

В современном мире авиационная отрасль играет ключевую роль в международных и национальных перевозках. Эффективное управление авиаперевозками требует точного прогнозирования временных рядов статусов авиаперевозок, таких как задержки, отмены рейсов и другие изменения. Разработка программы для прогнозирования временного ряда на основе линейных авторегрессионных моделей позволит улучшить планирование и управление авиаперевозками.

Целью данного исследования является разработка программы для прогнозирования временного ряда на основе линейных авторегрессионных моделей на основе данных статусов авиаперевозок. Для достижения этой цели поставлены следующие задачи:

1. Изучить теоретические основы временных рядов и авторегрессионных моделей.
2. Провести анализ данных статусов авиаперевозок.
3. Разработать программу прогнозирования временного ряда на основе полученных данных.

После завершения данного исследования будет создана программа, способная предсказывать статусы авиаперевозок с высокой точностью, что поможет авиакомпаниям оптимизировать свою деятельность и повысить уровень обслуживания пассажиров.

# 1 ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

## 1.1 Основные понятия временных рядов

Для разработки программы прогнозирования временного ряда на основе данных статусов авиаперевозок необходимо разобраться с основными понятиями временных рядов. Временной ряд — это последовательность данных , измеренных в последовательные моменты времени . Он может быть стационарным или нестационарным, что необходимо учитывать при выборе модели.

Временной ряд называется стационарным

* в узком смысле, если для любых вектор совпадает по распределению с , то есть при сдвиге всех моментов времени на одно и тоже число совместное распределение значений временного ряда в эти моменты времени не поменяется;
* в широком смысле, если
  + для любого ;
  + не зависит от , то есть в среднем значение временного ряда постоянно;
  + для любых , то есть значение автокорреляции зависит только от длины отрезка времени между двумя значениями;
* для гауссовских распределений, то есть для случая, когда все векторы вида имеют нормальное распределение, определения эквивалентны. Это следует из того, что распределение гауссовского случайного вектора полностью определяется математическим ожиданием и ковариациями;

В нестационарных временных рядах статистические свойства меняются со временем. Они показывают сезонные эффекты, тренды и другие структуры, которые зависят от временного показателя.

Каждое измерение представляет собой наблюдение для конкретной переменной в определенный момент времени.

Компоненты временного ряда:

1. Тренд (T — trend) — долгосрочное изменение уровня значений временного ряда. Тренд может быть восходящим, нисходящим или стационарным.
2. Сезонность (S — seasonal) — циклическое повторение паттернов в данных с постоянным интервалом времени. Например, продажи игрушек могут иметь сезонное изменение в преддверии праздников.
3. Цикл (C — cyclic) — периодические колебания в данных, обычно с более длительным циклом, чем у сезонности. Например, экономические циклы имеют периодичность в несколько лет.
4. Шум (E — errors) — непредсказуемая случайная переменная, которая не может быть объяснена трендом, сезонностью или циклом. Шум включает в себя случайные флуктуации и ошибки измерения.

Включение компонентов в модель временного ряда:

1. Аддитивная модель: Аддитивная модель используется, если амплитуда колебаний более-менее постоянная.
2. Мультипликативная модель: Мультипликативная — если амплитуда колебаний зависит от значения сезонной компоненты.

Выбор между аддитивной и мультипликативной моделями зависит от того, как компоненты взаимодействуют друг с другом в конкретном временном ряде.

Анализ временных рядов позволяет выявлять закономерности в данных, делать прогнозы и принимать более обоснованные решения на основе исторических данных. Методы анализа временных рядов включают в себя статистические и математические модели, такие как ARIMA (среднее, интегрированное, скользящее среднее), экспоненциальное сглаживание и регрессионный анализ.

## 1.2 Линейная авторегрессионная модель

Линейная авторегрессионная модель (AR(p)) — это статистическая модель, которая используется для прогнозирования временных рядов. Она основана на предположении о том, что текущее значение временного ряда зависит от предыдущих значений этого же ряда.

Линейная авторегрессионная модель может быть представлена в виде уравнения:

где — стационарный ряд, а — гауссовский белый шум, то есть и независимы. Отметим, что, вообще говоря, для стационарности нужны некоторые условия на коэффициенты .

По сути, наш ряд выражается через сумму некоторого фиксированного числа , значения белого шума в текущий момент времени и не более предыдущих значений этого же ряда, домноженных на некоторые коэффициенты, которые являются параметрами модели.

# 2 ПРАКТИЧЕСКАЯ ЧАСТЬ

## 2.1

## 2.2

# ЗАКЛЮЧЕНИЕ

# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

## Теоретическая часть

## Практическая часть

# ПРИЛОЖЕНИЯ

## Листинг