**МИНОБРНАУКИ РОССИИ**

**Санкт-Петербургский государственный**

**электротехнический университет**

**«ЛЭТИ» им. В.И. Ульянова (Ленина)**

Доклад

**Тема: Методы редукции данных**

Выполнили: Чепасов Д.В, Щедрин А.А.

Группа:2382

Санкт-Петербург

2023

Оглавление

[ВВЕДЕНИЕ 3](#_Toc130123050)

[PCA (Метод главных компонентов) 4](#_Toc130123051)

[**1.1)** **Математическое обоснование** 4](#_Toc130123052)

[**1.2)** **Роль в машинном обучении** 5](#_Toc130123053)

[**1.3)** **Реализация в прикладной задаче** 6](#_Toc130123054)

[**1.4)** **Плюсы и минусы метода** 7](#_Toc130123055)

[Корреляционный анализ 8](#_Toc130123056)

[**2.1) Математическая теория** 8](#_Toc130123057)

[**2.1.1) Основные понятия** 8](#_Toc130123058)

[**2.1.2) Различия корреляционных связей** 8](#_Toc130123059)

[2.2) Роль в машинном обучении 10](#_Toc130123060)

[**2.3) Реализация для прикладной задачи** 11](#_Toc130123061)

[2.4) Плюсы и минусы метода 12](#_Toc130123062)

[Метод максимального правдоподобия 13](#_Toc130123063)

[**3.1) Математическая теория** 13](#_Toc130123064)

[**3.2) Роль в машинном обучении** 14](#_Toc130123065)

[**3.3) Реализация для прикладной задачи** 15](#_Toc130123066)

[**3.4) Плюсы и минусы метода** 16](#_Toc130123067)

[Сравнение методов 17](#_Toc130123068)

[ЗАКЛЮЧЕНИЕ 18](#_Toc130123069)

[Использованные источники 19](#_Toc130123070)

**ВВЕДЕНИЕ**

**PCA (Метод главных компонентов)**

* 1. **Математическое обоснование**
  2. **Роль в машинном обучении**
  3. **Реализация в прикладной задаче**
  4. **Плюсы и минусы метода**

**Корреляционный анализ**

**2.1) Математическая теория**

Основной задачей корреляционного анализа состоит в количественном определении тесноты связи между двумя признаками при парной связи и между результативной и несколькими факторными признаками при многофакторной связи и статистической оценке надежности установленной связи.

**2.1.1) Основные понятия**

Корреляция – статистическая взаимосвязь двух или более переменных, в которой при изменении одной или нескольких величин, меняется другие величины или другая, но одна величина.

Корреляционный анализ – раздел математической статистки, которые изучает эти взаимосвязи.

Корреляционный связь – это согласованное изменение двух признаков, отражающее тот факт, что изменчивость одного признака находится в соответствии с изменчивостью другого. Т.е. каждой независимой переменной x соответствует определенное математическое ожидание значения y.

Группировочный признаки – признаки, по которым осуществляется разделение единиц изучаемой статистической совокупности на качественно однородные группы.

Факторный признаки – признаки, определяющие изменение показателей совокупности.

Результативные признаки – изменяются под влиянием факторных признаков.

Парная корреляция – связь между двумя признаками.

Частная корреляция – связь между двумя признаками, при фиксированных значениях других факторных признаков.

Множественная корреляция - связь между результативным и двумя или более факторными признаками, включенными в исследование.

Регрессия – зависимость среднего значения какой-либо величины от некоторый другой величины или от нескольких величин.

**2.1.2) Различия корреляционных связей**

Корреляционные связи могут различаться по трем свойствам:

1. *По форме.* По форме корреляционная связь может быть прямолинейном или криволинейной. Примером прямолинейной связи можно назвать связь между количеством тренировок и количеством решенных задач на контрольной сессии.
2. *По направлению.* Корреляционная связь может быть положительной и отрицательной. При положительной корреляционной связи более высоким значениям одного признака соответствуют более высокие значения другого, при более низких значениях одного признака соответствуют более низкие значения другого. При отрицательной связи обратные соотношения: более высоким соответствуют более низкие, более низким соответствуют более высокие.
3. *По тесноте (силе или степень).* Определяется по коэффициенту корреляционной связи.

**2.1.3) Парный корреляционно-регрессионный анализ.**

## 2.2) Роль в машинном обучении

**2.3) Реализация для прикладной задачи**

## 2.4) Плюсы и минусы метода

**Метод максимального правдоподобия**

**3.1) Математическая теория**

**3.2) Роль в машинном обучении**

**3.3) Реализация для прикладной задачи**

**3.4) Плюсы и минусы метода**

**Сравнение методов**

# ЗАКЛЮЧЕНИЕ

# Использованные источники