

Министерство образования Республики Беларусь
Учреждение образования
«Брестский Государственный технический университет»
Кафедра ИИТ

Лабораторная работа №1
По дисциплине «Интеллектуальный анализ данных»
Тема: «PCA»

Выполнил:
Студент 4 курса
Группы ИИ-24
Крейдич А. А.
Проверила:
Андренко К. В.

Брест 2025

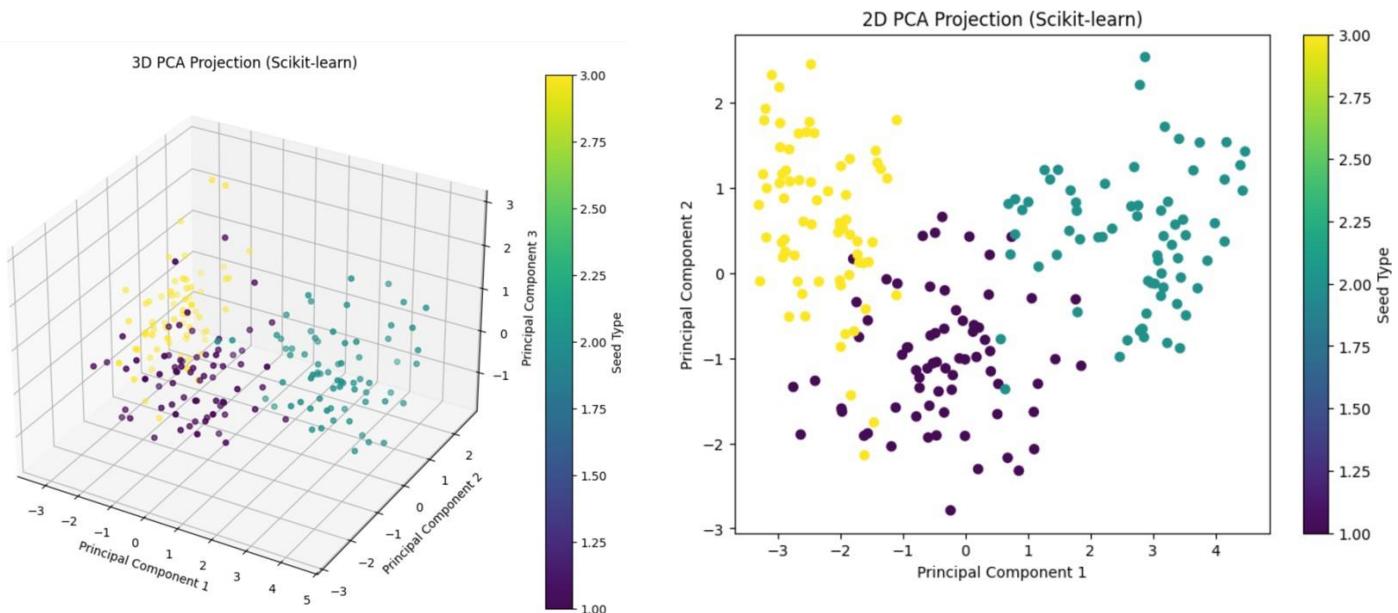
Цель работы: научиться применять метод РСА для осуществления визуализации данных.

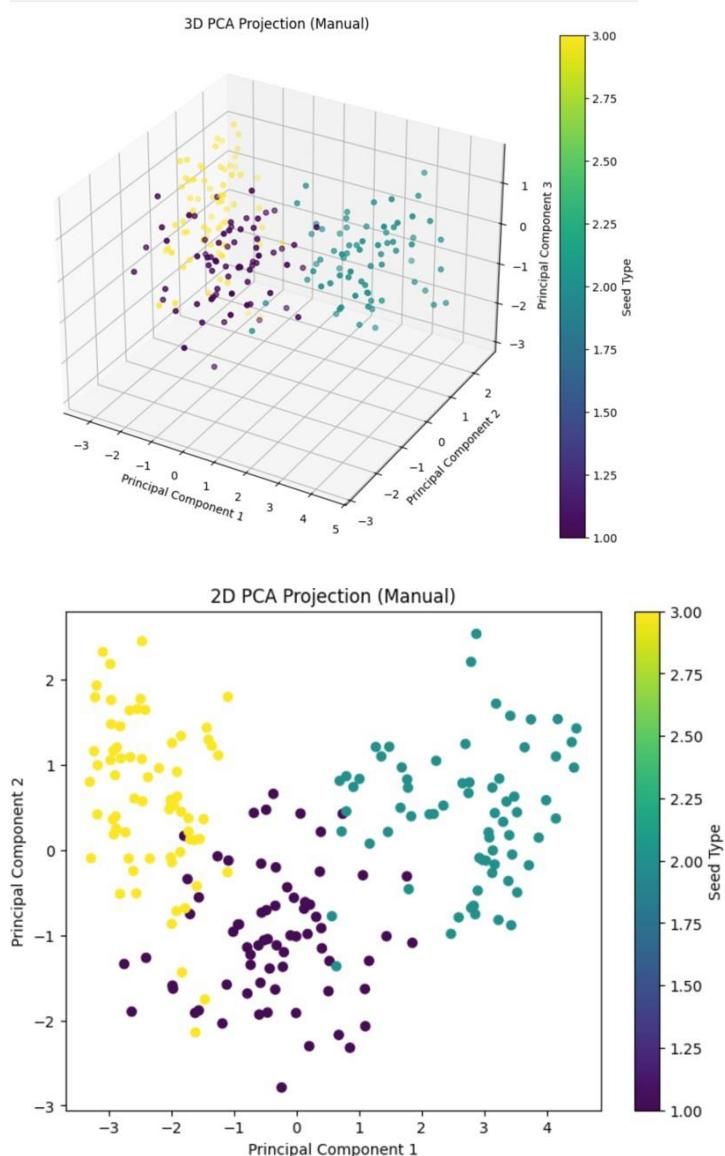
Общее задание

- Используя выборку по варианту, осуществить проецирование данных на плоскость первых двух и трех главных компонент (двумя способами: 1. вручную через использование `numpy.linalg.eig` для вычисления собственных значений и собственных векторов и 2. с помощью `sklearn.decomposition.PCA` для непосредственного применения метода РСА – два независимых варианта решения);
- Выполнить визуализацию полученных главных компонент с использованием средств библиотеки `matplotlib`, обозначая экземпляры разных классов с использованием разных цветовых маркеров;
- Используя собственные значения, рассчитанные на этапе 1, вычислить потери, связанные с преобразованием по методу РСА. Сделать выводы;
- Оформить отчет по выполненной работе, загрузить исходный код и отчет в соответствующий репозиторий на `github`.

№ варианта	Выборка	Класс
6	<code>seeds.zip</code>	Последняя колонка

Результаты программы:





Вывод: В ходе выполнения работы я изучил метод главных компонент (PCA) и применил его для снижения размерности и визуализации данных. Проецирование данных на первые главные компоненты было выполнено как вручную с использованием собственных значений и векторов, так и с помощью библиотеки `sklearn`, при этом полученные результаты совпали. Анализ собственных значений показал, что основные данные сохраняются в первых компонентах, а потери информации при уменьшении размерности являются незначительными.