|  |  |
| --- | --- |
| **Gerb-BMSTU_01** | **Министерство науки и высшего образования Российской Федерации**  Калужский филиал  федерального государственного бюджетного  образовательного учреждения высшего образования  ***«Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»***  ***(КФ МГТУ им. Н.Э. Баумана)*** |

**ФАКУЛЬТЕТ** ***ИУК «Информатика и управление»***

**КАФЕДРА** \_\_***ИУК4 «Программное обеспечение ЭВМ, информационные технологии»***

**ЛАБОРАТОРНАЯ РАБОТА №8**

**«Аналитическая платформа Deductor»**

**ДИСЦИПЛИНА: «Технологии анализа данных»**

|  |  |  |
| --- | --- | --- |
| Выполнил: студент гр. ИУК4-82Б | | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ( Карельский М.К. )  (Подпись) |
| Проверил: | | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ( Ерохин И.И. )  (Подпись) |
| Дата сдачи (защиты):  Результаты сдачи (защиты): | | |
|  | - Балльная оценка:  - Оценка: | |

Калуга, 2024

**Цель:** ознакомление с архитектурой, основными частями и пользовательским интерфейсом Deductor Studio, освоение и закрепление навыков применения факторного и корреляционного анализа, ознакомление с методами трансформации данных в среде Deductor Studio.

**Задачи:**

1. Получение навыков импорта данных,
2. Изучение парциальной обработки,
3. Изучение восстановления данных,
4. Применение методов удаления аномалий в данных,
5. Изучение спектральной обработки данных,
6. Получение навыков удаления шумов в данных,
7. Изучить факторный и корреляционный анализ,
8. Получить навыки по работе с ними,
9. Научиться понимать разницу между ними,
10. Осознать области применения факторного и корреляционного анализа в data mining,
11. Изучение способов разбиения данных,
12. Применение методов квантования данных,
13. Получения навыков фильтрации данных.

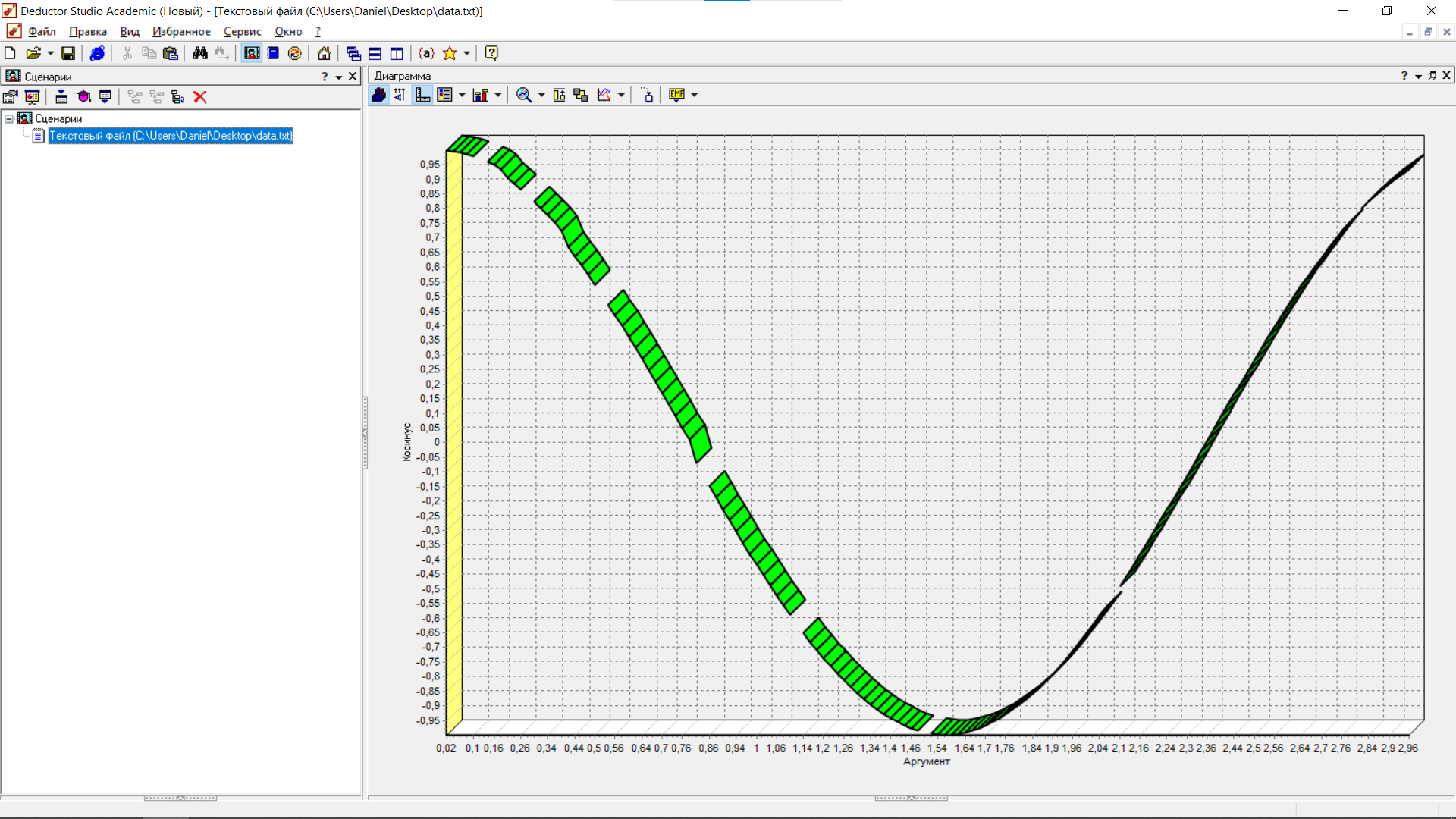
**Задание:**

Для всех вариантов необходимо провести классификацию с помощью наивного Байесовского классификатора. Составить модель для классификации. Сравнить данные, полученные моделью, с реальными данными. Данные для вариантов необходимо брать из набора sklearn.datasets

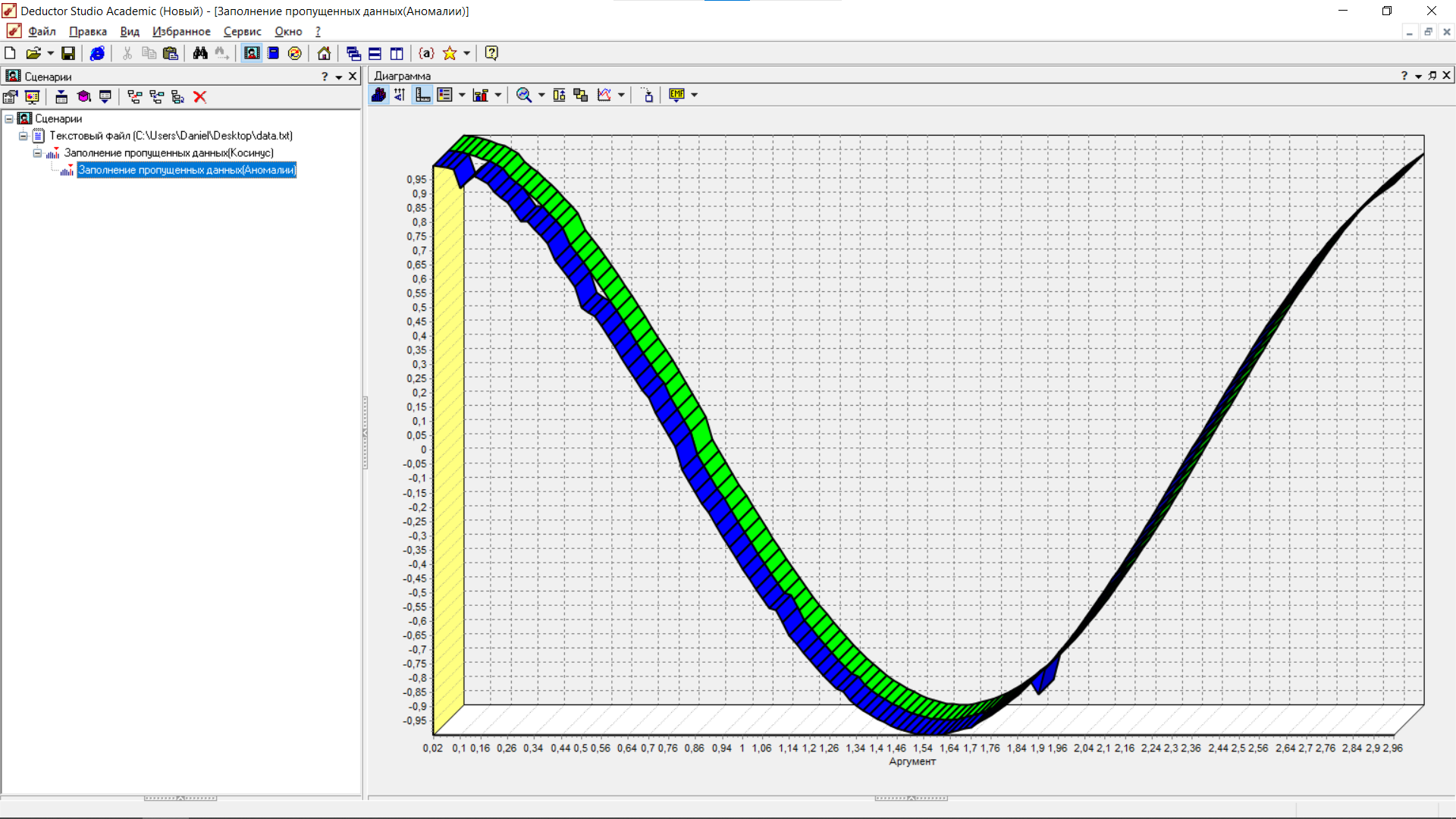
**Вариант 5**

Создать файл с данными аналогично файлу, приведенному в лабораторной работе. В качестве основной функции использовать: cos(2x). Выполнить импорт данных, созданного файла, обработку данных, восстановить пропущенные значения функции, выполнить парциальную обработку, удалить аномалии и шумы.

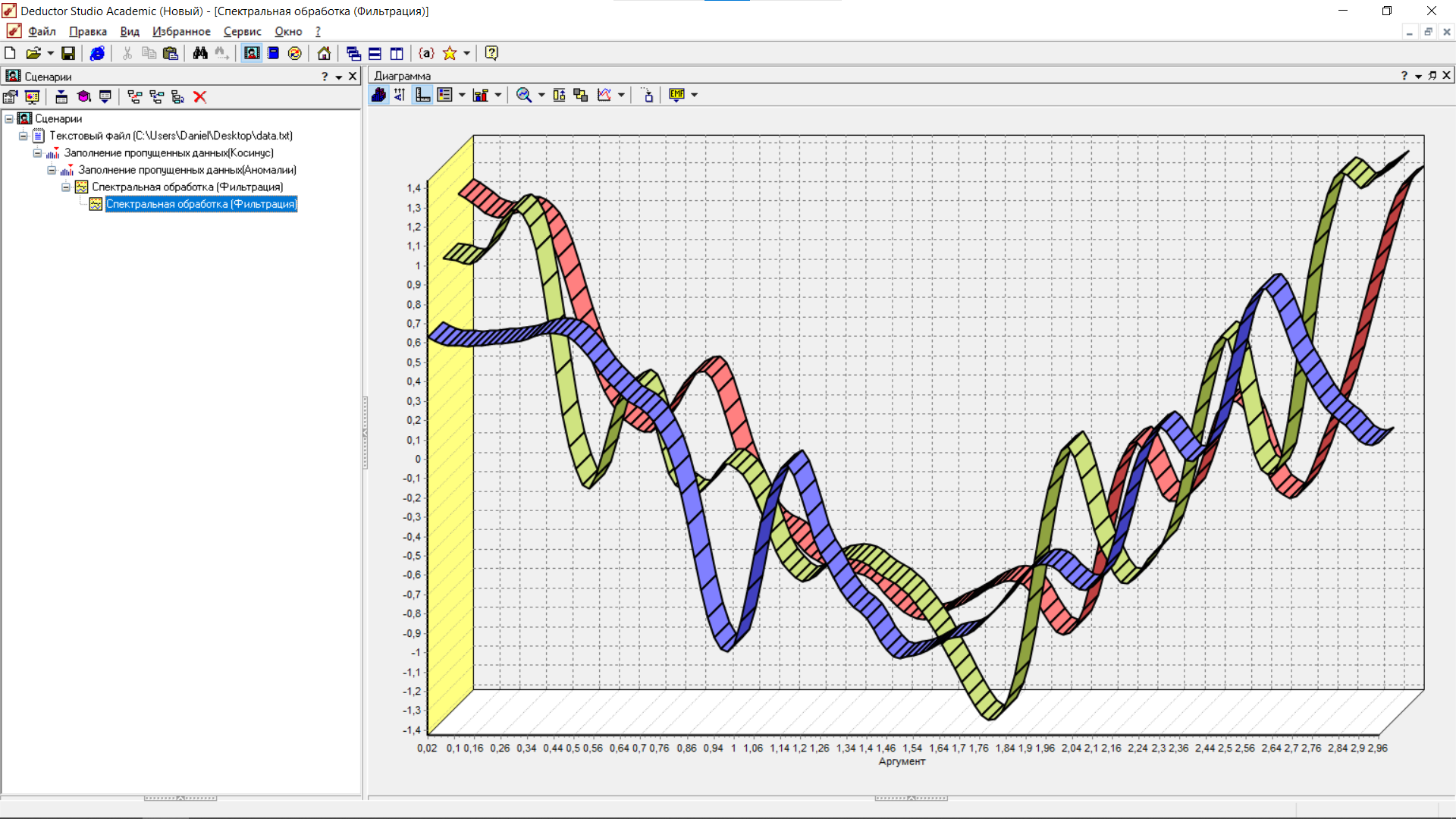
**Результат:**



**Рис. 1.** Импорт данных



**Рис. 2.** Восстановление пропущенных значений



**Рис. 3.** Удаление шумов

**Вывод:** в ходе выполнения лабораторной работы была изучена архитектура, основные части и пользовательский интерфейс Deductor Studio, получены практические навыки применения факторного и корреляционного анализа, изучены методы трансформации данных в среде Deductor Studio.