

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Косенок Сергей Михайлович
Должность: ректор
Дата подписания: 15.06.2026 11:07:42
Уникальный программный ключ:
e3a68f3eaa1e62674b54f4998099d3d6bdfcf836

Оценочные материалы для промежуточной аттестации по дисциплине

Анализ данных

Код, направление подготовки	01.03.02, Прикладная математика и информатика
Направленность (профиль)	Технологии программирования и анализ данных
Форма обучения	очная
Кафедра-разработчик	Кафедра прикладной математики
Выпускающая кафедра	Кафедра прикладной математики

Типовые задания для контрольной работы

Вариант 1

- 1) Создать два одномерных массива на основе заданного csv-файла. Назвать созданные массивы x и y , т.е. x – 0-й столбец в наборе данных из файла, y – 1-й столбец.
- 2) Для данных массивов найти выборочные средние и стандартные отклонения.
- 3) Вычислить коэффициент корреляции.
- 4) Построить рисунок, на котором изобразить диаграмму рассеяния в координатах (x, y) и линию регрессии y на x .

Вариант 2

Пусть X_1, X_2, \dots, X_n – независимые случайные величины, равномерно распределенные на отрезке $[a, b]$, A – их математическое ожидание, D – их дисперсия, $S_n = X_1 + \dots + X_n$ – их сумма, $A_n = nA$, $D_n = nD$. Введем случайную величину X по формуле:

$$X = \frac{S_n - A_n}{\sqrt{D_n}}$$

Выполнить следующие действия:

- 1) провести N испытаний случайной величины X , т.е. получить набор из N псевдослучайных чисел соответствующих распределению X .
- 2) по данному набору чисел построить гистограмму из k столбцов с нормализацией.
- 3) на этом же рисунке построить график плотности вероятности стандартной нормальной случайной величины.

Типовые вопросы для экзамена

1. Характеристики случайных величин: плотность вероятности, функция распределения, математическое ожидание, дисперсия, стандартное отклонение, ковариация, коэффициент корреляции.
2. Нормальная случайная величина и ее свойства. Правило «трёх сигм». Центральная предельная теорема.
3. Основные понятия математической статистики: выборочное среднее, дисперсия, среднеквадратическое отклонение, коэффициент корреляции.

4. Основные понятия машинного обучения: признаки, категориальные признаки, ординальные и номинальные признаки, обучающие примеры, целевая переменная.
5. Основные типы машинного обучения: классификация, регрессия, кластеризация. Обучение с учителем и без учителя.
6. Модель линейной регрессии. Метод наименьших квадратов. Нелинейная регрессия. Регуляризация.
7. Бинарная классификация. Логистическая регрессия. Многоклассовая классификация.
8. Метод опорных векторов. Использование функций-ядер в методе опорных векторов.
9. Деревья принятия решений. Примеси. Случайный лес.
10. Метод k -ближайших соседей.
11. Предобработка данных. Заполнение пропущенных значений. Кодирование категориальных признаков.
12. Обучение без учителя. Основные методы кластеризации.
13. Визуализация данных. Линия среднеквадратической регрессии. Гистограмма. Boxplot.