

## Лабораторна робота №2 ІАД

Виконав студент 3 курсу МІТ-31

Крижановський Олександр

**Тема:** Використання NumPy, Pandas та Matplotlib для обробки даних та візуалізації результатів

**Мета:** Ознайомитися з основами роботи з бібліотеками NumPy, Pandas та Matplotlib для аналізу даних, проведення базових операцій з даними та візуалізації результатів.

**Завдання:**

1. NumPy (створити набір даних (розмір за власним рішенням), провести CRUD операції над даними (скаляр, вектор, матриця)). Використати `arange`, `random`.
2. Pandas (створити `dataframe` різними способами, провести маніпуляції з даними (`head`, `describe`, `iloc`, `loc` ...))
3. Matplotlib (створити набір даних лінійної функції з урахуванням помилок вимірювань, побудувати графік функції та згенерованих даних, обчислити похибку з використанням метрик MAE, MSE). Записати отримані результати у csv файл (формат: X,Y, Y\_hat, mAE, mSE)
4. Відповіді оформити .ipynb документом
5. Викласти у зазначений репозиторій GitHub в окремій папці з назвою Lab2 де має бути .ipynb файл та необхідні файли з даними

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.metrics import mean_absolute_error, mean_squared_error

# 1. NumPy: создание данных и CRUD операции
# Создание скаляра, вектора и матрицы
scalar = 42
vector = np.arange(1, 11)
matrix = np.random.rand(3, 3)

# CRUD операции
# Create - уже создали
# Read
print("Vector:", vector)
print("Matrix:\n", matrix)

# Update
vector[2] = 99 # Обновляем третий элемент вектора
matrix[0, 1] = 5.5 # Обновляем элемент в матрице

# Delete
vector = np.delete(vector, 4) # Удаляем пятый элемент
print("Updated Vector:", vector)

# 2. Pandas: создание и манипуляция DataFrame
# Создание DataFrame с помощью разных методов
df1 = pd.DataFrame({
    'A': np.random.rand(5),
    'B': np.random.rand(5)
})

df2 = pd.DataFrame(np.random.rand(5, 3), columns=['X', 'Y', 'Z'])

# Операции с DataFrame
print("Head of df1:\n", df1.head())
print("Description of df2:\n", df2.describe())
print("iloc:\n", df2.iloc[0:2])
print("loc:\n", df1.loc[:, 'A'])

# 3. Matplotlib: построение графика и расчет ошибок
# Линейная функция  $y = 2x + 1$  с добавлением шума
X = np.arange(0, 10, 0.5)
Y_true = 2 * X + 1
Y_noisy = Y_true + np.random.normal(0, 2, len(X)) # Добавляем шум

```

```

# Побудова графіка
plt.plot(X, Y_true, label="True Function", color="blue")
plt.scatter(X, Y_noisy, label="Noisy Data", color="red")
plt.title("Linear Function with Noise")
plt.xlabel("X")
plt.ylabel("Y")
plt.legend()
plt.show()

# Визначення помилок (MAE і MSE)
mae = mean_absolute_error(Y_true, Y_noisy)
mse = mean_squared_error(Y_true, Y_noisy)
print("MAE:", mae)
print("MSE:", mse)

# 4. Запис в CSV файл
results = pd.DataFrame({
    'X': X,
    'Y': Y_true,
    'Y_noisy': Y_noisy,
    'MAE': [mae] * len(X),
    'MSE': [mse] * len(X)
})

results.to_csv('results.csv', index=False)

```

**Висновок:** Я Ознайомився з основами роботи з бібліотеками NumPy, Pandas та Matplotlib для аналізу даних, проведення базових операцій з даними та візуалізації результатів.