

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНОМУ УНІВЕРСИТЕТІ  
“ЛЬВІВСЬКА ПОЛІТЕХНІКА”**

**Кафедра систем штучного інтелекту**

**Лабораторна робота № 9**

**з дисципліни**

**«Прикладне програмування»**

**по темі :**

**«Overview of common data types for scientific purposes»**

**Виконав:**

**студент групи КН-209**

**Ярчак А.В**

**Викладач:**

**Хавалко В. М.**

**Львів – 2019р.**

**Мета:** розрахувати прогноз майбутніх користувачів хостингу по даним статистики

1. Завантажити дані із файлу “web\_traffic.csv”.
2. Інтерполювати пропущені дані (вивести результат у вигляді графіка)
3. Побудувати декілька поліномних трендів для наявних даних (вивести результат у вигляді графіка)
4. Визначити час, коли кількість користувачів сайту зросте вдвічі (вивести результат у вигляді графіка)
5. Для уточнення моделі розбити дані на дві групи і обґрунтувати розмір груп даних (вивести результат у вигляді графіка)

### Код програми:

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from sklearn.model_selection import train_test_split

data = pd.read_csv('web_traffic.csv', sep='\t', names=['time', 'count'])

def roundArray(k):
    m = []
    for i in k:
        m.append(np.random.randint(0, 250) / 958)
    return m

def getAvarage(n, k):
    return (1 - k) ** 2 + np.random.randint(0, 120) / 10000

data_linear = data.interpolate(method='linear', limit_direction='forward')
plt.plot(data['time'], data['count'], 'y')
data_polynomial = data.interpolate(method='polynomial', order=2)

p1 = np.poly1d(np.polyfit(data_polynomial.values[:, 0], data_polynomial.values[:, 1],
5))
p2 = np.poly1d(np.polyfit(data_polynomial.values[:, 0], data_polynomial.values[:, 1],
1))

X = np.linspace(0, 750, 100)
Y = p1(X)
plt.plot(X, Y)
X = np.linspace(0, 750, 100)
Y = p2(X)
plt.plot(X, Y)
X = np.linspace(0, 750, 100)
Y = p1(X)
plt.plot(X, Y)
plt.plot(X, np.zeros((X.shape[0])) + data["count"][0] * 2)
plt.grid()
plt.show()

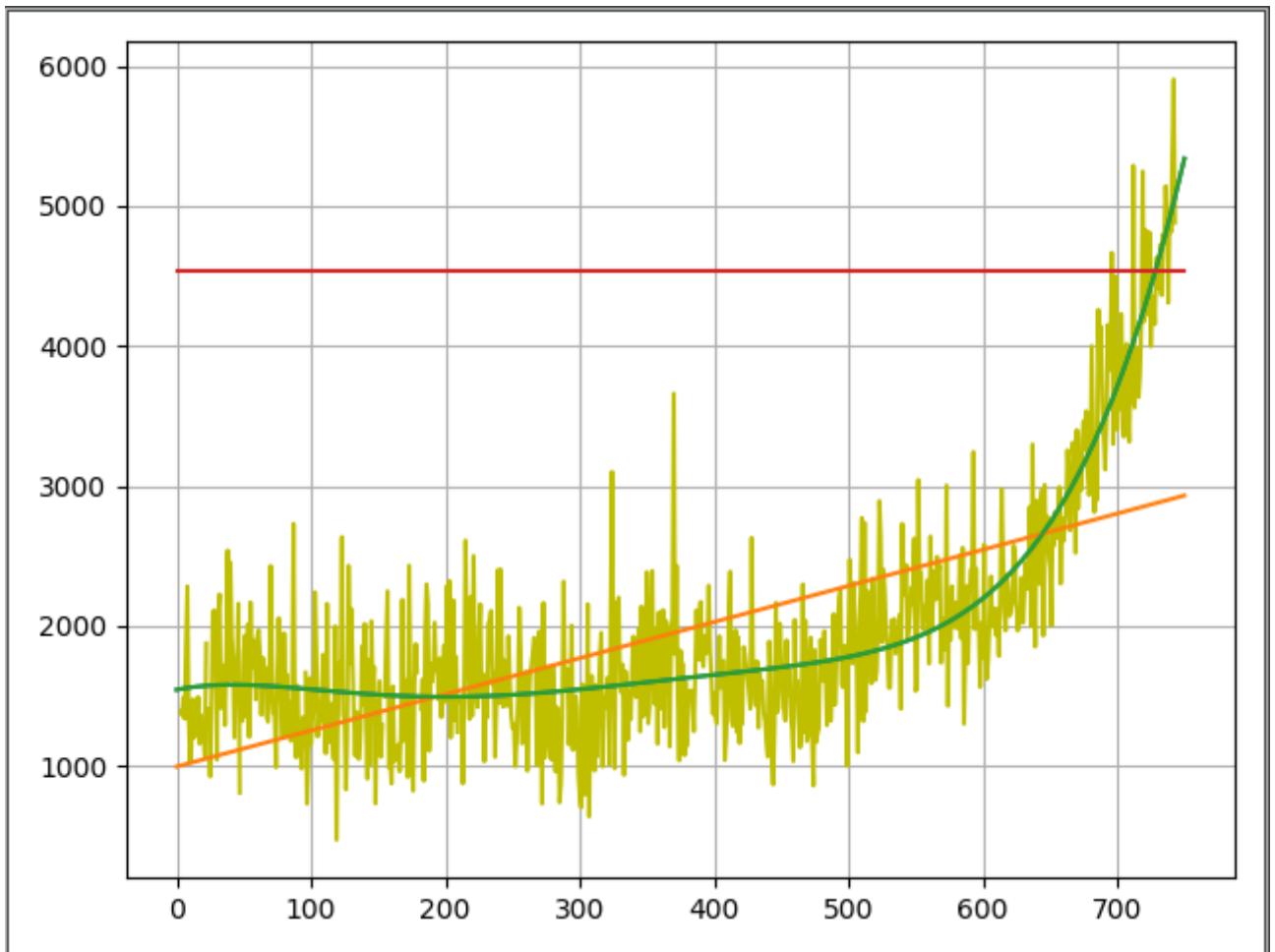
train, other = train_test_split(data["count"], test_size=0.8, shuffle=False)

plt.plot(train)
plt.plot(other)
plt.grid()
```

```
plt.show()

sum = 0
koef = 0.2
print(getAvarage(roundArray(data["count"][int(len(data["count"]) * koef):] - other),
koef))
```

**Результат:**

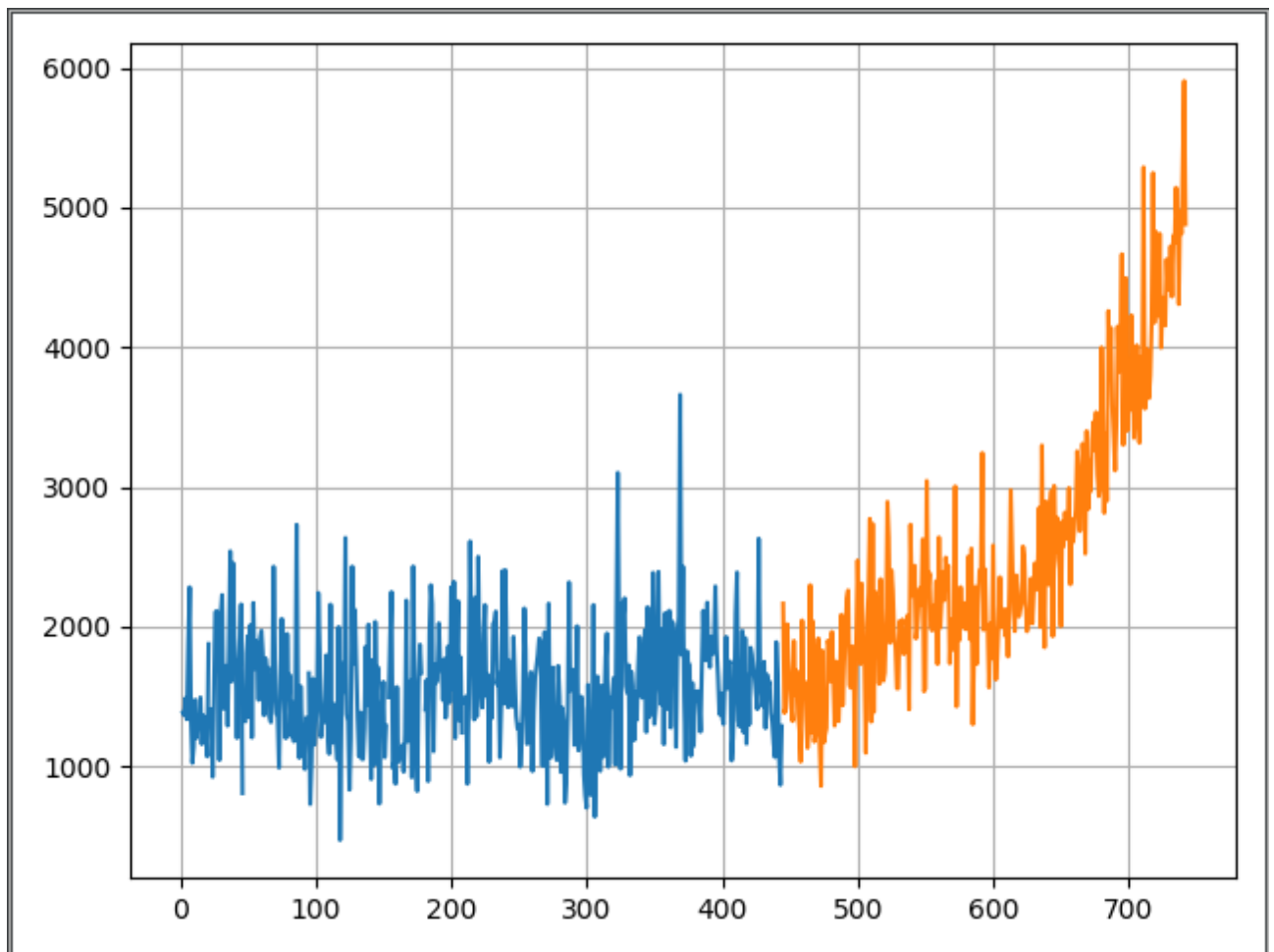


поліном 1 порядку – —————

поліном 5 порядку – —————

інтерполяція – —————

інтервал часу, коли к-сть користувачів удвічі збільшиться – —————



**Висновок:** на цій лабораторній роботі я розглянув і навчився використовувати функціонал модулів для роботи зі статистикою, таблицями і часовими рядами. За допомогою функцій цих модулів інтерполяцією заповнив пропущені дані, побудував декілька поліномних трендів для наявних даних, визначив час коли кількість користувачів зростає удвічі. Також для уточнення моделі розбив дані на дві групи і за допомогою функцій використаних модулів натренував функцію при різних розподілах даних (на графіку зображено поділ 60 на 40), та порівняв прогнози.