МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

Звіт

до лабораторної

роботи з дисципліни

**"Машинне навчання та аналіз даних"**

для студентів освітньо-кваліфікаційного рівня «бакалавр»

за спеціальністю 122 «Комп’ютерні науки та інформаційні технології», освітня програма «Інформатика»

Виконав студент группи ІТІНФ-18-2

Туркін Микита

ХАРКІВ 2020

**Лабораторна робота №1. Попередня обробка навчальних вибірок**

**Мета:** Засвоїти основні складові навчальної вибірки (приклади, атрибути, значення атрибутів, цільовий атрибут), повторити основні статистичні характеристики, необхідні для попередньої обробки навчальних вибірок.

**Хід роботы**

**Індивідуальне завдання:** Вибірка **iris.data**

В ході лабораторної роботи необхідно написати 4 процедури попередньої обробки вибірок:

• процедура доступу до довільного елемента вибірки;

• процедура випадкового сортування записів вибірки;

• процедура обчислення статистичних характеристик для значень кожного атрибуту (мат. очікування та дисперсії);

• процедура призначення номерів класів для значень цільового атрибуту, яка замінює символьні значення класів їх порядковими номерами.

Було розроблена скрипт на мові программуваняя python

За перший пункт відповідає функція

def read\_line(file\_name: str, line\_index: int) -> List[str]:

За другий

def random\_sort(file\_name: str, lines\_amount: int, attr\_index: int, amount=20) -> List[List[str]]:

Третій:

def mathematical\_expectation(

        attr\_list: Iterable[Union[str, float]],

        attr\_index: int) -> float:

def standard\_deviation(

        attr\_list: Iterable[Union[str, float]],

        attr\_index: int,

        math\_expectation: float) -> float:

def dispersion(

        attr\_list: Iterable[Union[str, float]],

        attr\_index: int,

        math\_expectation: float) -> float:

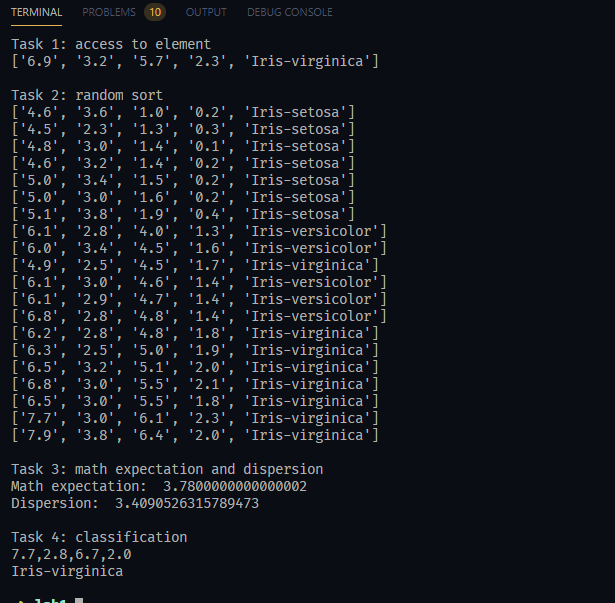
Четвертий:

def classification\_precalc(file\_name, class\_index=-1) -> By\_Class\_Dict:

def input\_set(attr\_len: int) -> List[float]:

def classify\_set(set\_item: List[float], by\_class\_dict: By\_Class\_Dict) -> str:

Виходні данні программи:



**Висновок:** Засвоїв основні складові навчальної вибірки, повторив основні статистичні характеристики, необхідні для попередньої обробки навчальних вибірок.