

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ "ЛЬВІВСЬКА ПОЛІТЕХНІКА"

КАФЕДРА СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ



Звіт до лабораторної роботи №4
з дисципліни:
“Інтелектуальний аналіз даних ”

Підготувала:
студентка групи КН-309
Дипко Олександра
Викладач:
Виклюк Я.І.

Львів 2020

Візуальний аналіз даних методами Orange

1. Використовуючи Orange повторити розрахунки лабораторних робіт 1 та 2 з предмету візуалізація даних.
2. Оформити звіт з таблицею переваг та недоліків R/Python vs Orange.

	Переваги	Недоліки
Python	<ul style="list-style-type: none">• Оскільки Python - загальна мова програмування, вивчення дасть вам навички вийти за рамки простого аналізу даних. Програмування на Python широко використовується для веб-розробки, тестування автоматизації та ETL.• Програмісти вважають, що Python збігається з тим, як програмісти думають більше, ніж R, і тому він перекладається на інші мови легше.• Велика частина аналізу даних - це попереднє очищення даних. Приємно очищати дані мовою повного сервісу, такою як Python, тому що ви можете додавати нові функції та шари, щоб розділити ваші дані.• Python розвивається з часом. Вводиться новий код, який ламає старий код, що робить Python живою мовою.• Python рухається швидше, ніж R. Це тому, що R був розроблений, щоб зосередитись на зручності статистиків, а не на зручності комп'ютера.• Python набув широкої популярності, оскільки синтаксис зрозумілий для розуміння.	<ul style="list-style-type: none">• Python повільніший у порівнянні з іншими мовами програмування, оскільки він є інтерпретованою мовою.• Python вимагає ретельного тестування, оскільки помилки з'являються під час роботи.• Програмування на Python досі вважається слабким на мобільних обчислювальних платформах, оскільки існує мало програм, створених з Python як основною мовою.

<p>R</p>	<ul style="list-style-type: none"> • R чудово підходить для статистичного аналізу. • R також побудований навколо командного рядка, але багато людей працюють у таких середовищах, як RStudio або R Commander, які включають редактор даних, підтримку налагодження та вікно для зберігання графіки. • Мова R вважається найкращим інструментом для візуалізації даних. Візуалізовані дані можна зрозуміти краще, ніж вихідні цифри. • Програмування R дає найкращі результати візуалізації, які можуть бути використані в наукових роботах (довідкових документах). 	<ul style="list-style-type: none"> • Для користувачів, які не знають програмування, мова R буде трохи складною, оскільки вона має круту криву навчання. • Виведення належних рішень за допомогою мови програмування R можна вважати повільним, якщо код написаний погано. Щоб подолати цей недолік, обов'язковим є включення бібліотек для досягнення належного результату.
<p>Orange</p>	<ul style="list-style-type: none"> • За кілька кліків можна проаналізувати / візуалізувати мільйон записів за дуже короткий час. Рекомендовано для планування та прогнозування. • Orange забезпечує ефективний аналіз даних візуальними засобами, він має розумний інтерфейс для побудови моделей машинного навчання та заборон. • Корисний інструмент для аналізу великих даних, під час роботи з регресією та гіпотезою - це дуже корисно. • Ефективне використання передбачень та категоризованого машинного навчання в AI. 	<ul style="list-style-type: none"> • Недостатньо інструментів в Orange, оскільки для досягнення певних результатів потрібно використовувати Python Script. • Попередній аналіз не такий простий, він не призводить до помилок навіть при неправильних даних. Деякі рази сповільнюють роботу системи під час роботи з даними. • Помилка відстеження даних важка, оскільки ми повинні бачити її вручну. • Аналіз та прогнозування даних у реальному часі не полегшено Orange. Помилки не класифікуються та не підтверджуються номерами.

Висновок:

Отже, з усього вищеперечисленого можна зробити висновок, що кожен з цих інструментів є призначеним для якихось певних дій, для яких вони є найкращими. Усі вони мають свої переваги та недоліки, які я описала вище.

Наприклад, Python є хорошим для очищення даних перед їх аналізом, R чудово підходить для статистичного аналізу, а Orange є чудовим інструментом для візуалізації та аналізу даних, який розроблений для малого та великого бізнесу і допомагає керувати видобутком даних, попередньою обробкою, прогнозними моделюванням, оцінкою функцій тощо.