Самойленко Володимир Анатолійович, студент

Харківський національний університет радіоелектроніки

**DATA SCIENCE ТА ОСНОВИ РОБОТИ З ІНФОРМАЦІЄЮ**

В сучасному світі, на фоні швидкого розвитку технологій та зростання обсягів даних, виникає все більша необхідність у вмінні працювати з інформацією. Одним з ключових напрямків в цій сфері є Data Science - інтердисциплінарне поле, яке поєднує математику, статистику, програмування та глибоке розуміння даних. У даному тексті ми розглянемо основи роботи з інформацією та роль Data Science у цьому процесі.

Першим кроком у роботі з інформацією є збір та збереження даних. Це може включати збір інформації з різних джерел, таких як бази даних, веб-сторінки, соціальні мережі тощо. Однак, важливо забезпечити якість та достовірність даних, виявляти й усувати помилки та неповноти.

Наступним етапом є очищення та попередня обробка даних. Це включає видалення дублікатів, заповнення пропущених значень, корекцію помилок та аномалій. Для цього можуть застосовуватися різні методи та алгоритми, зокрема статистичні техніки та машинне навчання. Метою цього етапу є отримання чистих, якісних та повних даних для подальшого аналізу.

Для ефективного розуміння та аналізу даних важливо використовувати візуалізацію. Це можуть бути графіки, діаграми, хмари слів та інші візуальні засоби, які допомагають візуалізувати залежності, розподіли та інші характеристики даних. Візуалізація сприяє кращому розумінню даних, виявленню патернів та інсайтів.

Після попередньої обробки та візуалізації можна перейти до аналізу даних та винесення рішень. Data Science використовує різні аналітичні методи, алгоритми машинного навчання та статистики для виявлення залежностей, класифікації, прогнозування та іншого аналізу даних. На основі цього аналізу можуть бути прийняті обґрунтовані рішення та розроблені стратегії для покращення бізнесу, науки та інших галузей.

Робота з інформацією та Data Science також поставляють перед нами виклики та перспективи. З одного боку, зростання обсягів даних та складність аналізу вимагають постійного оновлення навичок та знань. З іншого боку, необхідно бути усвідомленим щодо етичних питань, пов'язаних з використанням даних, таких як конфіденційність, безпека та відповідальне використання.

Data Science та основи роботи з інформацією стають все більш важливими в контексті сучасного світу, де дані стають цінним активом. Уміння збирати, очищувати, аналізувати та використовувати дані є ключовими для прийняття обґрунтованих рішень та досягнення успіху в різних галузях. Однак, варто пам'ятати про етичні аспекти та стежити за постійним розвитком у цій сфері.

**Література**

1. Ткачов В.М. Метод передачі даних в комп’ютерній мережі проміжного зберігання даних складної інформаційної системи / В.М. Ткачов// Системи управління, навігації та зв’язку. – Полтава: Полтавськийнаціональний технічний університет ім. Ю. Кондратюка, 2017. – № 3 (43). –С. 117-119.

2. Ткачев В. Н., Филимончук Т. В., Митин Д. Е. Использование информационной технологии распределения заданий при обработке больших массивов данных в виртуальных частных облаках. Информационные системы и технологии: Материалы 5-й международной научно-технической конференции (2-17 сентября 2016 г., Харьков). Х.: ДРУКАРНЯ МАДРИД, 2016. С. 333–334.

3. Tkachov V. Architecture of overlay network with nested vpn tunneling / M. Hunko, V. Tkachov, M. Bondarenko // "Сучасні напрями розвитку інформаційно комунікаційних технологій та засобів управління" : матеріали Дев'ятої міжнар. наук.-техн. конф., 9–10 квітня 2020 р. – Харків, 2020. – С. 36.