

В.Ю.Кучерук, д.т.н., проф;
І. О Головащенко, к.ф.н. доцент;
М.В. Глушко, студент

Як Big Data (великі дані) впливають на буття людини

Вінницький національний технічний університет

Анотація

В даній роботі розглядаються великі дані як спосіб впливу на людину в теперішньому інформаційно-технічному світі та її буття на прикладі найвідоміших корпорацій світу.

Ключові слова: Big Data, великі дані, людина, буття.

Abstract

In this work the big data as a way to influence the people in the present information technology world and its being the example most famous companies.

Keywords: Big Data, big data, people, life.

У техніці використовується як термін Big Data, так і термін «великі дані». Великі дані не мають строгого визначення. Не можна провести чітку межу - це 10 терабайт або 10 мегабайт? Сама назва дуже суб'єктивна. Слово «велике» - це як «один, два, багато» у первісних племен.

Однак є усталена думка, що великі дані - це сукупність технологій, які покликані здійснювати три операції. По-перше, обробляти великі в порівнянні з «стандартними» сценаріями обсяги даних. По-друге, вміти працювати з даними, що швидко поступають в дуже великих обсягах. Тобто даних не просто багато, а їх постійно стає все більше і більше. По-третє, вони повинні вміти працювати зі структурованими і погано структурованими даними паралельно в різних аспектах. Великі дані припускають, що на вхід алгоритми отримують потік не завжди структурованої інформації, і що з нього можна витягти більше ніж одну ідею.

Типовий приклад великих даних - це інформація, яка надходить з різних фізичних експериментальних установок - наприклад, з Великого адронного коллайдера, сенсори якого постійно виробляють величезну кількість даних. Установка безперервно видає великі обсяги даних, а вчені з їх допомогою вирішують паралельно безліч завдань.

Багато розробників звикли працювати зі статичними об'єктами і мислити категоріями станів. У великих даних парадигма інша. Ми повинні вміти працювати з безперервним потоком даних, і це цікава задача. Вона зачіпає все більше і більше областей.

Проте є інша сторона цих даних – вплив на людину: її буття, думку, погляд, мислення, дії і т.д. Тобто по суті це маніпуляція людьми за допомогою великих даних. Проте як це відбувається?

Ми часто чуємо про чергові модні додатки та програми, які використовують якісь магічні штуки на кшталт нейромереж, великих даних або засновані на початках штучного інтелекту. Для багатьох це магія, тому що далеко не всі розуміють або усвідомлюють, що ж насправді ховається за цими словами. Найцікавіше, що Big Data (великі дані) вже багато років формує наш щоденний спосіб життя і впливає на наші вчинки та рішення. Великі дані - це спосіб обробки величезних і різноманітних масивів інформації, які надходять щосекунди. Такі дані важливо швидко обробити і

структурувати для того, щоб в майбутньому отримати щось корисне з них. Як ми стикаємося з Big Data кожен день? Людина - головний генератор і споживач великих даних. Щодня ми створюємо стільки нової інформації, скільки раніше створювали десятиліттями. Не важливо, що велика частина цього - наші Селфі. Зараз 90% наявної інформації було створено за останні два роки. Ми генеруємо інформацію не тільки за допомогою фотографій і соціальних постів. Це кожен наш пошуковий запит, крок, порахований фітнес-трекером, відео, переглянуте на YouTube. Майже кожна наша дія кимось записується і стає частиною Big Data.

Тільки в Google відбувається близько 40 000 пошукових запитів щомиті, що дає близько 1,2 трильйона пошукових запитів для великих даних щорічно. І з кожним днем кількість даних збільшується все швидше. Якщо сьогодні ми генеруємо 4,4 зетабайта даних, то до 2020 ми будемо створювати вже 44. До 2020 року майже третина всіх даних буде проходити через хмарні сервіси, а значить буде піддана аналізу.

Цікаво, що для аналізу великих даних не завжди використовуються комп'ютери корпорацій. Часто користувачі самі надають свої комп'ютери для вирішення різних наукових завдань. В цей же час 73% організацій проінвестували або збираються проінвестувати в розвиток великих даних.

Великі компанії типу Google, Facebook і навіть держави обробляють і використовують цю інформацію для поліпшення нашого способу життя. Ну, або для показу більш релевантної реклами. Якщо спростити, то великі дані збираються з декількох джерел:

- **Відкриті дані:** соціальна, економічна та інша публічна інформація про міста, країнах. Дані про законодавствах, статистика спортивних заходів. По суті, будь-яка відкрита довідкова інформація.

- **Соціальні мережі:** все те, що ми з вами добровільно розповідаємо Facebook і ВКонтакте. Навіть якщо пост не публічний, він стає частиною Big Data. У середньому Facebook пишуть близько 31,25 мільйонів повідомлень і переглядають 2,77 мільйонів відео кожену хвилину.

- **Інтернет речей** - це ще один цікавий термін. Якщо ваш холодильник має доступ в інтернет, значить ваші продукти - теж великі дані. Будь-які сенсори в телефоні, смарт-годинках, фітнес-трекерах передають найрізноманітнішу інформацію про вас і ваших заняттях. Наприклад, в 2016 році було продано близько 1,4 більйона (більше мільярда) смартфонів. У кожному з яких є безліч сенсорів для збору даних про їх власника. І з кожним роком з'являється все більше речей з інтернет-доступом, а ще вчора вони чудово працювали і без нього.

- **Особисті дані** теж стають частиною великих даних. Часто ці дані знеособлені: дані вашої медичної карти, списки справ і так далі.

- **Комерційні транзакції.** Мова не тільки про банківські транзакції, а й про будь-які платежі в інтернеті.

- **Будь-який інший контент**, створюваний нами: відео на YouTube, фотографії в Instagram. Кожну хвилину ми завантажуюмо 300 годин відео на YouTube, а в 2015 році завантажили близько трильйона фотографій, з яких мільйони доступні публічно. До кінця 2017 року 80% усіх фотографій буде зроблено на смартфон

Все це лякає і іноді хочеться сховатися під камінь і добувати вогонь за допомогою двох палиць. Однак, за Великими даними - майбутнє і доводиться миритися з тим, що Великий Комп'ютер знає про нас все. Хочемо ми цього чи ні. На основі цих даних ми можемо не тільки отримувати таргетовану рекламу, орієнтовану тільки на нас, але і значно поліпшити наше життя. Тільки більшість не розуміє, як відбувається це поліпшення і лякається будь-якої згадки про біг датою. Якщо вас вже вразили розміри того, скільки інформації обробляється, то ось ще один цікавий факт. Сьогодні обробці піддається лише 0,5% усіх доступних даних. Тому найцікавіше все ще попереду.

Часто Big Data використовують для передбачення майбутнього. Наприклад, спортивних заходів на кшталт Олімпійських ігор. Вчені в 2012 році за допомогою комп'ютерної моделі прогнали емуляцію Ігор і отримали прогноз, що США отримають 108 медалей, але по факту США отримали 104 медалі.

Звичайно, в спорті не останню роль відіграє людський фактор, але вже давно математичний аналіз впливає на покупки і продаж гравців в бейсболі і в інших спортивних іграх. З приходом великих даних і нейромереж провести цей аналіз стало ще простіше.

За допомогою них стає можливо запобігти епідемії. Аналізуючи пошукові запити і пости в соціальних мережах медики можуть відстежити збільшення ключових слів-симптомів хвороб і запропонувати користувачам звернутися до лікаря за профілактичної медичної допомогою.

У деяких містах вже сьогодні можна дістатися до роботи без пробок, завдяки великим даними. Наприклад, в Лондоні щомиті аналізується автомобільний трафік і світлофори перемикаються відповідно до завантаженості доріг. Іноді вони навіть змушують водіїв об'їхати пробку. Ви їдете на роботу і навіть не знаєте, що на сусідній вулиці сталася аварія.

Напевно, ви чули про автомобілі Tesla, які можуть їздити без допомоги водія. Мабуть, це перші безпілотні автомобілі, доступні будь-якому бажаючому. Tesla орієнтуються в просторі не тільки за допомогою супутника, карти, GPS та інших сенсорів.

Автомобіль постійно аналізує ваш стиль водіння: як часто ви набираєте швидкість, тиснете педаль гальма, як рухаєтеся по вашим щоденним маршрутами. Ці дані збираються з усіх автомобілів компанії і після аналізу повертаються назад. Вчора ви довірилися автопілоту Tesla, а на повороті він не скинув швидкість або взагалі звернув не туди. Сьогодні вже іншого власника автомобіль провезе по цій ділянці без помилок.

Звичайно, Tesla це тільки початок і скоро по всьому світу будуть їздити безпілотні Uber, Google і Яндекс таксі.

Якщо відволіктися від майбутнього, то здається, що Великі дані зовсім не впливають на нас і наше життя. Однак, це зовсім не так.

Наприклад, створення карт для GPS-навігації йде за допомогою Big Data. Тисячі знімків з супутника, аналіз руху автомобілів, дані про аварії. Все це враховується в будь-якому сучасному додатку для навігації і маршрут буде побудований на основі цих даних.

Новинні стрічки в соціальних мережах вже давно показують пости не за датою публікації, а на основі того, що думає про наших перевагах алгоритм мережі. Уже сьогодні складно знайти хоч одну область нашого життя, яка не схильна до великим даними.

Будь-яке наше дію так чи інакше вже кимось пораховано і вплине на наші дії в майбутньому.

Магазини все частіше будуть пропонувати товари, які нам можуть бути потрібні, ще до того, як ми про це подумаємо. Лікарі зможуть заздалегідь провести профілактичні заходи від майбутньої хвороби і запобігти епідемії. ВКонтакте нам підкаже, кого вибрати в президенти в цьому році, а поруч буде реклама холодильника, про який ми мріємо вже тиждень. І все це Біг дата, яку не варто боятися, і про яку варто знати. Вона змінює буття людини або із, або без її відома. Хтось не відчуває цей вплив керуючись відчуттям росту технологічного прогресу, а хтось розуміє, що інформаційно-технічний прогрес змінює нас.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Виктор Майер-Шенбергер Кеннет Кукьер. Большие данные (Big Data). Революция, которая изменит то, как мы живем, работаем и мыслим. // Издательство «Манн, Иванов и Фербер», Москва, 2014 г. ст. 48.
2. Rick Smolan, Jennifer Erwit. The Human Face of Big Data First Edition Edition //, USA 2014, p 234.
3. John W. Foreman. Data Smart: Using Data Science to Transform Information into Insight // 1st Edition, Kindle Edition, 2016, USA 2016 p. 132

ВІДОМОСТІ ПРО АВТОРІВ:

Кучерук Володимир Юрійович – доктор технічних наук, професор,
завідувач кафедри «Метрології та промислової автоматики»,

Вінницький національний технічний університет

Головащенко Ірина Олегівна - кандидат філософських наук,

доцент кафедри філософії та гуманітарних наук

Вінницький національний технічний університет

Глушко Михайло Васильович – студент-магістр, група ІВТ-16м