

**Козак Є.Б.**

GAN Inc.

## АНАЛІЗ ДАНИХ І МАШИННЕ НАВЧАННЯ У ХМАРНИХ І ТУМАННИХ ПЛАТФОРМАХ ЯК ОСНОВА ЕФЕКТИВНОЇ ПЕРЕДАЧІ ДАНИХ

*У статті досліджено аналіз даних і машинне навчання в хмарних і туманних платформах як основа ефективної передачі даних. Визначено напрями формування масивів даних, які є найактуальнішими в умовах сьогодення. Наголошено, що сучасні парадигми обчислень і зберігання містять хмарні, туманні й крайові обчислення, здійснено опис принципових особливостей кожного з видів. Підкреслено, що обчислення туману є середнім рішенням між обчислювальними парадигмами хмари й крайовими обчисленнями, за якого обробку й аналіз даних не потрібно розташовувати на краю мережі або навіть у централізованому сховищі. Усталений підхід, заснований на технологіях хмарних обчислень, не може забезпечити необхідну якість обслуговування в таких умовах, зокрема в питаннях зменшення часу затримки під час передачі великих масивів даних. Технологія туманних обчислень нині розглядається як багатообіцяюче рішення для обробки великого обсягу критично важливих і чутливих до часу даних. У статті розглянута технологія хмарних обчислень, а також аналізуються передумови до еволюційного розвитку такого підходу й появи концепції туманних обчислень як перспективного напрямку ефективної передачі даних. У рамках огляду ключових особливостей туманних обчислень проведений розбір критичних випадків з об'єднанням понять туманних і крайових обчислень. Наведено огляд технологій організації туманних обчислень – віртуалізації, контейнеризації та оркестрації, а також систематичний аналіз найпопулярніших платформ, що забезпечують підтримку туманних обчислень. Запропоновано узагальнену системну модель для спільної обробки хмарних платформ у різномірних мережах і схематично представлено злиття архітектури IoT та аналізу даних. Визначено поняття прогностичних моделей і розкрито їх види, наголошено, що сучасні методи прогнозування містять статистичні моделі на основі регресії, моделі дерева рішень й моделі, ґрунтовані на нейронній мережі або глибокій нейронній мережі.*

**Ключові слова:** машинне навчання, штучний інтелект, туман, хмара, обчислення, передача даних, платформа, аналіз даних.

**Постановка проблеми.** В умовах життя сучасного цифрового суспільства, існує великий ряд практичних завдань, в яких потрібно контролювати стан деякої безлічі датчиків або простих виконавчих пристроїв, розміщених на відносно невеликих відстанях і до яких важко або небажано підводити дротяні лінії. Однією з систем взаємозв'язку датчиків та виконавчих пристроїв є Інтернет речей (далі – IoT). Парадигма Інтернету речей є одночасно революційною, а також механізмом автоматизованого та зручного способу життя сучасних людей. Еволюцію IoT можна пояснити злиттям досягнень, що відбулися за останнє десятиліття в галузі обчислювальної техніки, зв'язку та дизайну додатків. Загальнозживані пристрої IoT для полегшення повсякденної діяльності включають смартфони, домашні помічники, такі як Google Play, розумні транспортні засоби, системи автоматизації будівель, що включають розумні ліфти та системи контролю температури, та безпілотні літальні апарати, такі як безпілотні літальні апарати для моніторингу навколишнього

середовища й відпочинку. Широкомасштабне розповсюдження пристроїв IoT поширюється за межі цих пристроїв у центрах зберігання даних, таких як серверні хмарні служби, які географічно розподілені. Як результат, великі обсяги даних генеруються пристроями IoT та їх підтримуючими платформами для передачі та подальшого зберігання та обробки в центральних хмарних центрах зберігання даних. Пристрої IoT генерують постійний потік необроблених даних, які неможливо розпізнати для значущих знань, якщо дані не обробляються за допомогою таких методів, як виявлення знань та штучний інтелект. Неоднорідність даних, що генеруються залежить від домену програми. Пристрої, які передають дані, розроблені за допомогою спеціальних протоколів, що враховують обмеженість ресурсів цих пристроїв, щоб зберегти споживання енергії, пов'язане з роботою пристрою. Пристрої реалізують великий обсяг даних, які обробляються локально, обмеженим чином і передаються в централізований обчислювальний вузол або хмарне сховище, де

їх можна обробити або проаналізувати для отримання знань. Машинне навчання визначається як сімейство методів аналізу даних, в яких процес побудови моделі на основі навчальних даних автоматизований, тобто практично не вимагає втручання людини. Отже, процес класифікації даних за різними класами повністю автоматизований. Роль аналітики для обробки даних не можна не враховувати, і машинне навчання дуже сильно сприяє швидкій обробці великих обсягів даних, що надходять з пристроїв, для створення моделей, що цікавлять аналітиків даних.

**Аналіз останніх досліджень і публікацій.** Хмарні та туманні обчислення є фундаментальною основою досліджень багатьох сучасних вчених. Д.Р. Колісник, К.С. Місевич і С.В. Коваленко [1] розглянули питання щодо системної архітектури IoT-Fog-Cloud, розглянули взаємодію між трьома рівнями IoT, Fog і Cloud для ефективного впровадження програм для аналізу великих даних і кібербезпеки.

О.О. Петренко [2] розкрив стратегії розвитку сервіс-орієнтованих систем у хмарному середовищі. Автором проведено аналіз сучасного стану сервісних технологій, базованих на композиції веб-сервісів з уніфікованими протоколами зв'язку (перше покоління SOA) і на поєднанні мікросервісів з контейнерами (друге покоління MSA) та вперше показано, як їх властивості підлягають конвергенції, оскільки кожен з них має свої переваги і недоліки.

Моделі і методи прогнозування навантаження в умовах хмарних обчислень з використанням математичного апарату нейронних мереж навів В. Давідов [3]. На основі проведеного аналізу результатів дослідження доведено доцільність вдосконалення аналітичних методів прогнозування навантаження. Удосконалено метод прогнозування навантаження обчислювальних ресурсів в системах хмарних обчислень, що дозволяє отримати більш точні результати оцінки і забезпечити переважань в системах хмарних обчислень. Б.І. Тимченко, П.О. Марченко й Д.В. Сподарець [4] дослідили принципи сегментації патернів організації хмар на супутникових зображеннях з використанням глибоких нейронних мереж.

Фактори продуктивності застосування моделей штучного інтелекту у хмарі з використанням GPU запропонували В.Г. Тульчинський, С.І. Лавренюк і В.Ю. Роганов [5]. Із зарубіжних авторів варто відзначити такі роботи: Alshammari, Hamoud and Abd ElGhany, Sameh and Shehab, Abdulaziz [6], J. Xie, F.R. Yu, T. Huang, R. Xie, J. Liu, C. Wang and Y. Liu [7], Yousefpour, Ashkan and Ishigaki, Genya and Gour, Riti

and Jue, Jason [8], Higinio Mora, María Teresa Signes-Pont, David Gil, Magnus Johnsson [9], M. Al-Khafajiy, L. Webster, T. Baker, and A. Waraich [10], Taneja, Mohit and Davy, Alan [11], M.A. Khan, M. Karim and Y. Kim [12], Revar, Dharmendrasinh and Talatiya, Tejalba and Hakani, Raj and Nayak, Hemal and Student, P and Prof, Asst [13], S. Din and A. Paul [14] та інші. Проте, враховуючи описані наукові набутки, за темою, питання аналізу даних і машинне навчання у хмарних і туманних платформах, як основа ефективної передачі даних залишається відкритим та потребує детального опрацювання.

**Постановка завдання.** Здійснити дослідження принципів ефективної передачі даних, з виділенням процесу аналізу даних і механізмів застосування машинного навчання у хмарних і туманних платформах.

**Викладення основного матеріалу дослідження.** Сучасні парадигми обчислень та зберігання включають: хмарні, туманні та крайові обчислення. Завдяки інтеграції цих парадигм з пристроями з'являється надійний механізм збору, зберігання, обробки та аналітики даних. Така структура має можливість надавати в реальному часі уявлення про моделі даних, а також полегшує застосування методів машинного навчання для реалізації інтелектуальної аналітики даних.

Хмарна парадигма – це централізована модель зберігання даних, яка надає різні послуги, такі як Програмне забезпечення як послуга (SaaS), Платформа як послуга (PaaS) та Інфраструктура як послуга (IaaS), щоб забезпечити обробку та аналіз даних, централізовано [15]. Крайові обчислення, з іншого боку, дозволяють обробляти та аналізувати дані ближче до мережі, на локалізованих обчислювальних вузлах, таких як базові станції. Як результат, вдається уникнути витрат, пов'язаних з передачею даних до централізованих вузлів. Обчислення туману є середнім рішенням між обчислювальними парадигмами хмари та крайових обчислень, при якому обробку та аналіз даних не потрібно розташовувати на краю мережі або навіть у централізованому сховищі. Швидше, парадигма туману висуває концепцію віртуальної платформи, яка не знаходиться виключно на межі мережі, для обробки та аналізу даних.

У парадигмі хмарних обчислень, віддалено розташовані обчислювальні засоби (сервери) використовуються за допомогою Інтернету для зберігання, збору, управління та обробки даних. В останні десятиліття, особливо нещодавно, хмарним обчисленням приділяли велику увагу завдяки новій інфраструктурі та архітектурі обробки, які вони надають

для підтримки різних служб. Хмара надає різноманітні послуги, що надаються через Інтернет.

В рамках парадигми IoT потрібно зберігати та аналізувати величезний обсяг сформованих даних. Архітектура хмари може відрізнятися залежно від сценарію використання, наприклад, публічна, приватна, гібридна архітектура.

Крайові обчислення стосуються аналітичних можливостей поблизу краю мережі пристроїв. Важливість крайових обчислень можна краще зрозуміти, усвідомивши виклики та обмеження хмарних обчислень. Хмарні обчислення є централізованим способом обробки даних і можуть збільшити накладні витрати на обробку масивних даних. Надсилання необроблених потоків даних до хмари збільшить вартість і вимагатиме високої пропускної здатності та потужності зв'язку. Натомість, обчислення краю забезпечує обчислення в межах «локальної» мережі джерела даних. Таким чином, це вирішує проблему пропускної здатності надсилання великих обсягів даних на центральний сервер. Більше того, обчислення краю зменшує ймовірність виникнення однієї точки відмови і підвищує аналітичну ефективність. Порівняно з величезними та централізованими серверами, які можна знайти у хмарних обчисленнях, у фронтових обчислень є менші сервери, які розподіляються. У роботі [16] було підкреслено, що хмарні обчислення придатні для стійкого до затримок та складного аналізу даних, тоді як обчислення краю призначені для операцій у режимі реального часу з низькою затримкою. Крім того, хмарні обчислення вимагають досить складних розгортань порівняно з обчисленнями краю.

На рисунку 1 показано архітектуру спільної моделі хмарних платформ для мереж IoT. На

рисунку кінцеві пристрої та хмара базової мережі з'єднані через шлюз IoT, де виконується обробка фронту. Як правило, розгортання крайових обчислень вимагає мінімального планування і в більшості випадків може носити спеціальний характер. Як результат, передбачається, що близько 45% даних IoT найближчим часом використовуватимуть крайові архітектури.

Концептуально туманні обчислення лежать десь посередині між хмарними та крайовими обчисленнями, і вони виконують роль моста між хмарними та крайовими ресурсами. Вперше ідея обчислень туману була запропонована Cisco [17] для вирішення проблем, пов'язаних із додатками, чутливими до затримок.

Згідно із [17], на відміну від обчислень краю, де обчислення відбуваються на межі мережі, обчислення туманом забезпечують мережеві послуги між хмарою та кінцевими пристроями, а також забезпечують обчислення та засоби зберігання через віртуальну платформу. Туманні вузли мають чутливі до часу можливості аналізу даних і добре підходять для додатків або служб із часом відгуку від мілісекунд до хвилин. Однак можливості зберігання даних вузлів туману дуже обмежені. З іншого боку, хмарні технології можуть зберігати дані протягом місяців і навіть років.

Рисунок 2 показує різні компоненти IoT та взаємозв'язок аналітики з цими компонентами. Через підвищені обчислювальні ресурси та можливості обробки більшість аналітичних підходів придатні на хмарному сервері установки IoT. З іншого боку онлайн-алгоритми краще підходять для крайових обчислень.

Системи IoT можуть збирати дані від кількох до тисяч розумних пристроїв і передавати їх у хмарне

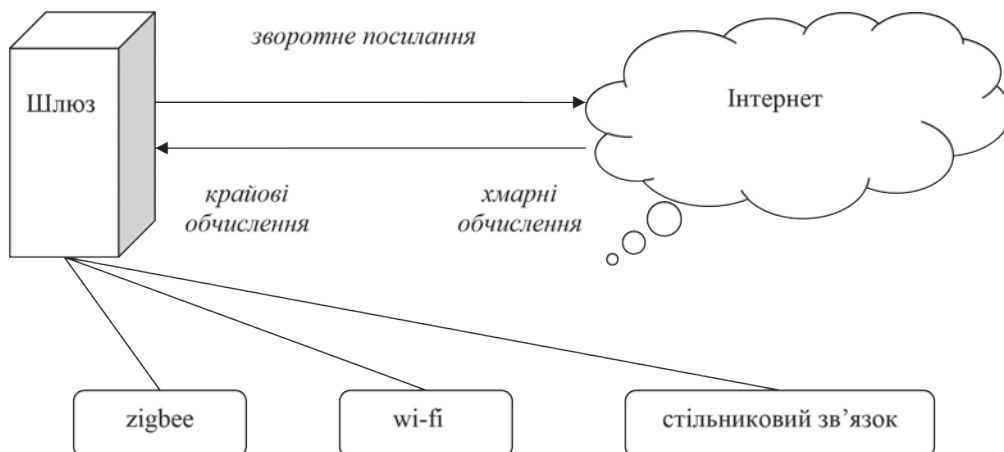


Рис. 1. Узагальнена системна модель для спільної обробки хмарних платформ у різнорідних мережах

середовище. На основі історичних даних завжди можна отримати детальну інформацію про минулі події, використовуючи передові методи машинного навчання. Ці групи алгоритмів, заснованих на машинному навчанні, які обробляють та узагальнюють вихідні дані та надають практичну інформацію, в основному складають область дослідження, яка називається описовою аналітикою. Агрегація даних, узагальнення даних, математичні логічні операції, видобуток даних (наприклад, алгоритми кластеризації) тощо – ось деякі приклади описової аналітики. Для описової аналітики потрібні великі обсяги даних. Останні технологічні досягнення продемонстрували, що хмарне сховище здатне зберігати величезні обсяги даних IoT, а хмарні сервери можуть обробляти складні завдання за допомогою високопродуктивних комп'ютерів та за допомогою хмарної аналітики IoT.

Прогностична аналітика покладається на історичні дані та використовує передові статистичні методи або методи машинного навчання для моделювання поведінки або шаблону, щоб можна було передбачити ймовірність можливих майбутніх тенденцій або закономірностей у даних. Підводячи підсумок, передбачається, що станеться в майбутньому, вивчаючи історичні закономірності та кореляцію існуючих даних. Прогностична аналітика широко застосовується для різних застосувань, включаючи прогнозне обслуговування, прогнозування ціни, тенденцію попиту та пропозиції або прогнозування ймовірності будь-якого результату. За даними SAS, яка є однією з провідних компаній з аналітики, існує два типи прогнозних моделей: класифікаційні моделі, що проводять аналіз прогнозування за членством у класі; моделі на основі регресії, які передбачають кількість про історичні спостереження та ймовірність. Сучасні методи прогнозування моделювання включають статистичні моделі на основі регресії, дерева рішень та моделі на основі нейронної мережі або глибокої нейронної мережі. Деякі інші широко використовувані алгоритми базуються на байєсівському аналізі, посиленні градієнта, аналізі на основі ансамблевої моделі тощо. Ці методи прогнозної аналітики залежать від даних для прийняття рішень. Парадигма IoT може полегшити процес збору

даних із розумних пристроїв IoT і може забезпечити аналітичну основу, використовуючи хмару або край мережі. Аналітика, що проводиться, пропонує підказки щодо реагування на будь-які майбутні події на основі аналізу даних. Цей клас аналізу не тільки передбачає майбутні області, але також містить рекомендації щодо прийняття результатів. Це більше схоже на методику майбутнього аналізу сценаріїв, яка використовує переваги як описової, так і прогнозної аналітики. Хоча прогнозована аналітика передбачає, що і коли відбуватиметься подія на основі майбутніх прогнозів, аналітична аналітика розширює можливості, надаючи уявлення про майбутні прогнози за допомогою аналізу впливу. Рецептивна аналітика широко використовується для оптимізації результатів бізнесу. Рецептивна аналітика підходить для установки, де рішення на основі бізнес-аналітики приймаються за допомогою можливостей хмарних / крайових обчислень, аналізу великих даних та машинного навчання. Послуги на платформі IoT-хмари можуть допомогти прийняти оптимальні рішення за допомогою розгортання інструментів бізнес-аналітики та аналітики.

Під час фактичної реалізації результат, отриманий в результаті прогностичної аналітики, повинен коригуватися за допомогою даних у реальному часі. Для цього використовується адаптивна аналітика для коригування або оптимізації результатів процесу на основі недавньої історії даних та перегляду їх співвідношень. Цей тип аналізу допомагає поліпшити продуктивність моделі й зменшити кількість помилок. Перевага адаптивної аналітики полягає в тому, що вона може коригувати результат рішення при отриманні нового

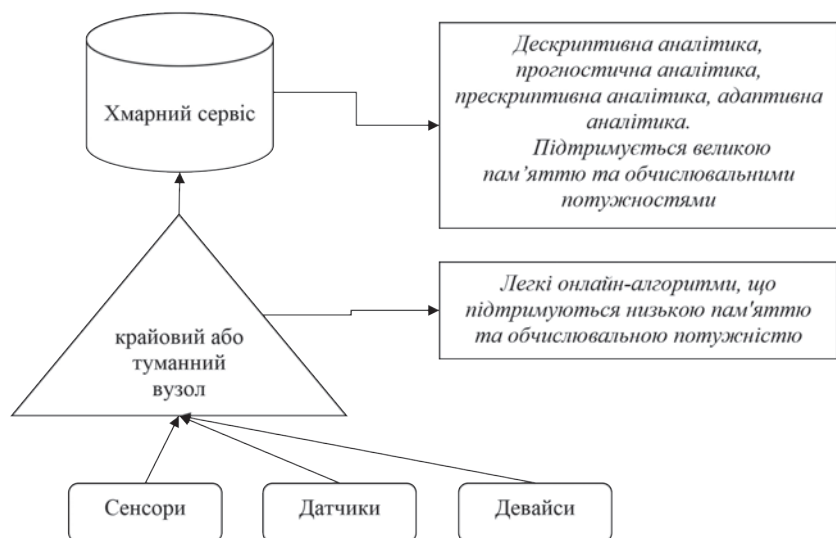


Рис. 2. Злиття архітектури IoT та аналізу даних



набору вхідних даних. Особливо в середовищі IoT адаптивна аналітика добре підходить для обробки поточкових даних у режимі реального часу. Оцінка в реальному часі потоків даних, що розвиваються, таких як ті, що знаходяться у програмному забезпеченні, також може бути предметом адаптивної аналітики для проведення аналізу даних.

Зближення машинного навчання та IoT відкриває шлях для перспективного прогресу в ефективності, точності, продуктивності та загальній економії коштів для обмежених ресурсів пристроїв IoT. Коли алгоритми машинного навчання та IoT працюють разом, є можливість досягти покращеної продуктивності для взаємозв'язку та обчислень, кращої керованості та кращого прийняття рішень. Завдяки вдосконаленому моніторингу від тисяч до мільярдів сенсорних пристроїв та вдосконалим комунікаційним можливостям, IoT має величезний потенціал для поліпшення якості людського життя та потенційних додатків для промислового зростання. Потенціал IoT значно покращився завдяки зближенню машинного навчання та штучного інтелекту. Передові методи машинного інтелекту дозволили видобувати величезний обсяг сенсорних даних IoT, щоб мати кращі уявлення про низку реальних проблем, а також здатність приймати критичні оперативні рішення. Отже, для вирішення реальних складних проблем та успішного задоволення обчислювальних та комунікаційних вимог IoT та машинне навчання повинні доповнювати одне одного. В останні роки аналіз даних IoT набув значного значення та уваги з наступних причин: Великий обсяг даних, що генеруються з розподілених пристроях IoT: За даними звіту про мобільність компанії Ericsson, прогноз показує, що до 2022 року у всьому світі буде 18 мільярдів підключених IoT-пристроїв<sup>1</sup>. Ця кількість з часом буде збільшуватися завдяки широкому впровадженню пристроїв IoT у широкому діапазоні критичних програм. Інтелектуальна аналітика даних відіграватиме важливу роль для виявлення та прогнозування майбутніх станів будь-якого процесу чи системи шляхом ефективного та розумного видобутку цієї величезної кількості даних. Висока мінливість типів даних від неоднорідних джерел даних: Завдяки широкому спектру застосувань та вимог існує велика різноманітність пристроїв IoT, що включають мобільні телефони, ПК / ноутбуки, планшети до пристроїв IoT короткого діапазону та широкого простору. Через неоднорідність даних функції, формати й атрибути даних відрізняються. Крім

того, на основі різних доменів додатків IoT джерела даних також різняться. Наприклад, пристрої IoT, що використовуються для медичних програм, будуть відрізнятися від IoT розумного будинку. Більше того, якість, обробка та зберігання даних також стали складним завданням через їх неоднорідність. Невизначеність у потоках даних IoT: невизначеність дуже поширена при практичному аналізі даних. Вона може виникнути в потоці даних IoT через відмову будь-якого пристрою IoT або каналу зв'язку під час передачі даних. Грубі помилки та відсутні дані є всюди в потоках даних IoT, що вимагає розширеної аналітики для попередньої обробки даних. Навіть кіберпроникнення може бути вагомою причиною невизначеності даних. Для підвищення точності під час прийняття рішень важливо забезпечити належну оцінку, розповсюдження та подання невизначеностей та розробити моделі та рішення, які можуть мати справу з цими факторами. Балансування масштабованості й ефективності: більшість аналітичних даних IoT виконуються у хмарі. Передача даних з пристрою IoT в хмару є дорогою (з точки зору затримки), що може бути складним для критично важливих для часу програм, особливо коли кількість пристроїв IoT велика. Наприклад, у підключеному транспортному середовищі може знадобитися великий пул автомобілів для прийняття рішень у режимі реального часу або майже в режимі реального часу. Тут важливо збалансувати швидкість і точність аналізу, коли кількість транспортних засобів збільшується. Пристрої IoT повинні обробляти дані для отримання знань. Як дані стають значущими для свого споживача, ілюструє архітектура, показана на рисунку 3. У цій архітектурі споживачем може бути машина або людина. Коли споживачі очікують отримати значущу інформацію, блок обробки даних складається з алгоритмів штучного інтелекту, таких як механізми машинного навчання. Передбачувальне та адаптивне навчання в Інтернеті речей можна найкраще описати через те, як програми обробляють дані для створення знань. Вищевказана архітектура використовується для організації оцінки того, як сучасні дослідження впроваджують штучний інтелект в парадигмі IoT. У зв'язку від машини до машини протокол рівня додатків потрібен для передачі даних у форматі, який може бути інкапсульований протоколами нижчого рівня. Як вже згадувалося раніше, широко використовуваними протоколами рівня IoT є протокол обмежених додатків, передача телеметрії в чергу масових повідомлень), розширений протокол

<sup>1</sup> <https://www.ericsson.com/en/mobility-report>.

черги повідомлень і протокол передачі HTTP. Всі ці протоколи дозволяють абстрагуватися від нижніх шарів.

Обмежені пристрої мають обмежені обчислювальні, комунікаційні й сховищні ресурси. Оскільки різні пристрої IoT мають різні вимоги до ресурсів, протоколи відрізняються своїми можливостями для задоволення цих вимог. З одного боку, протоколи додатків розглядають, як обмежені пристрої, які можуть надсилати дані. З іншого боку, аналіз машинного навчання вимагає великих обчислювальних ресурсів і великої кількості даних для обробки зібраних даних, щоб перетворити їх на значущі знання. Тому додатки IoT змінюють архітектуру створення знань, щоб перетворити дані на знання на основі структури та пропускну здатності мережі, доступної для передачі даних. Рисунок 4 ілюструє три моделі того, як можна застосувати архітектуру відкриття знань.

У межах першої моделі розумні пристрої мають достатньо обчислювальних ресурсів, щоб перетворити дані на знання локально або на межі. Це підходить для програм, чутливих до затримок. Тому адаптивне навчання обробляється на краю, або на машині обробки даних, або на пристрої IoT. У другій моделі обмежені пристрої можуть надсилати / приймати дані в дуплексному зв'язку. Віддалена система, напр. хмара, обробляє дані та повертає знання на пристрій. Цю модель можна побачити в багатьох додатках для смартфонів. Адаптивне навчання може оброблятися або на процесорі даних у хмарі, або на пристрої IoT. Третя модель – це коли обмежений пристрій надсилає дані у віддалену систему в симплексному режимі. Це має місце в процесі видобутку даних, коли дані використовуються або спільно використовуються з різними програмами. Кожен віддалений додаток може мати свій власний метод перетворення даних у знання. Адаптивне навчання відбувається в хмарі, оскільки зворотні зв'язки не можуть бути відправлені

назад на пристрій IoT. Іншими словами, пристрій IoT не може адаптивно навчатися.

**Висновки.** У роботі досліджено принципи ефективної передачі даних, з виділенням процесу аналізу даних і механізмів застосування машинного навчання у хмарних і туманних платформах.

З появою хмарних обчислень в інформатиці обробка стала простішою, а управління ресурсами ще кращим. Потім, з новою епохою обчислювальних програм із централізованим підходом хмарних обчислень, виникла потреба у подальшому скороченні часу відгуку у додатках, чутливих до затримок. Таким чином, у дію увійшли програми для обчислення туману. Зараз, за допомогою цієї нової технології, є ще певні труднощі щодо подальшої точної настройки

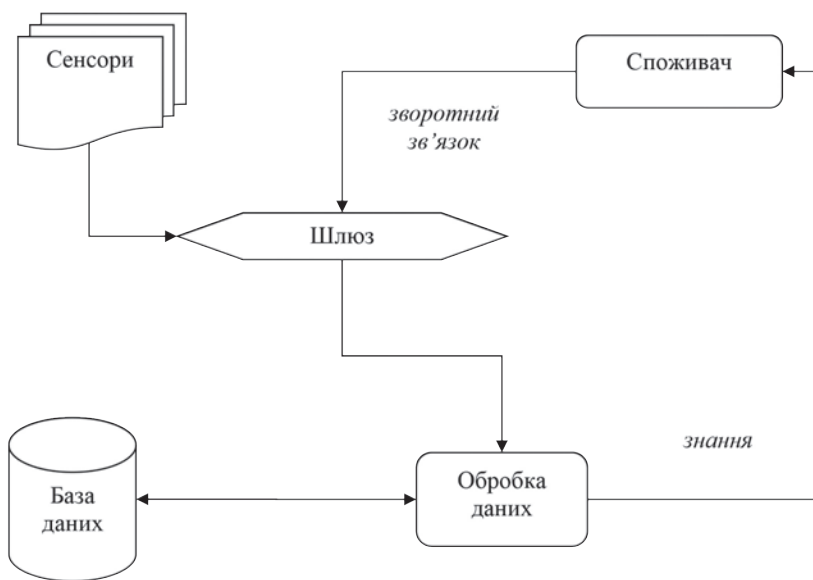


Рис. 3. Схема отримання знань

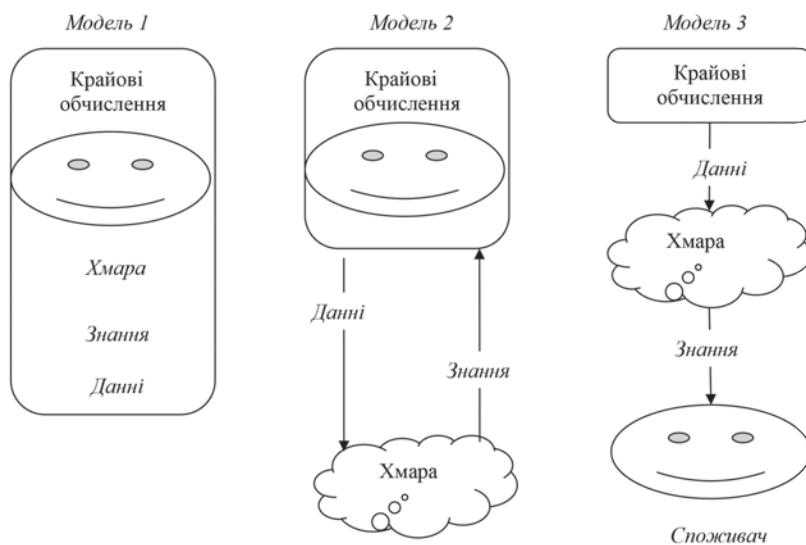


Рис. 4. Моделі отримання знань

системи, завдяки якій мережеві витрати будуть мінімізовані, а ефективність ще більше збільшена. Перспективи подальших досліджень ґрунтуються на розробці алгоритму інформаційного розміщення

платформи аналітики даних у архітектурі обчислень туману, яка адаптивно розгортає аналітичну платформу для роботи або в хмарі, або в тумані, тим самим зменшуючи мережеві витрати та час відгуку.

#### Список літератури:

1. Колісник Д.Р., Місевич К.С., Коваленко С.В. Системна архітектура IoT-Fog-Cloud для систем аналізу великих даних і кібербезпеки: огляд туманних обчислень, впровадження аудиту інтернету речей. *Сучасний захист інформації*. 2020. № 3. С. 34–38.
2. Петренко О.О. Стратегії розвитку сервіс-орієнтованих систем у хмарному середовищі : дис. ... канд. техн. наук. : 05.13.12. Київ, 2018. 239 с.
3. Davydov V., Hrebenuik D. Development the resources load variation forecasting method within cloud computing systems (Розробка методу прогнозування зміни навантаження ресурсів в системах хмарних обчислень). *Сучасні інформаційні системи*. 2020. Т. 4. № 4. Р. 128–135.
4. Tymchenko V.I., Marchenko P.O., Spodarets D.V. Segmentation of Cloud Organization Patterns from Satellite Images Using Deep Neural Networks (Сегментація патернів організації хмар на супутникових зображеннях з використанням глибоких нейронних мереж). *Вісник сучасних інформаційних технологій*. 2020. Т. 3. № 1. Р. 352–361.
5. Фактори продуктивності застосування моделей штучного інтелекту у хмарі з використанням GPU / В.Г. Тульчинський, С.І. Лавренюк, В.Ю. Роганов та ін. *Кібернетика та комп'ютерні технології*. 2020. Вип. 1. С. 74–82.
6. Alshammari, Hamoud, Sameh Abd El-Ghany, Abdulaziz Shehab. Big IoT healthcare data analytics framework based on fog and cloud computing. *Journal of Information Processing Systems*. 2020. No. 16.6. P. 1238–1249.
7. A survey of machine learning techniques applied to software defined networking (SDN): research issues and challenges / Xie, Junfeng, et al. *IEEE Communications Surveys & Tutorials*. 2018. Vol. 21. No. 1. P. 393–430.
8. Yousefpour, Ashkan and Ishigaki, Genya and Gour, Riti and Jue, Jason. On Reducing IoT Service Delay via Fog Offloading. *IEEE Internet of Things Journal*. 2018. P. 1–1. DOI: 10.1109/JIOT.2017.2788802.
9. Mora H., Signes-Pont M.T., Gil D., Johnsson M. Collaborative working architecture for IoT-based applications. *Sensors*. 2018. Vol. 18 (6). P. 1676.
10. Al-Khafajiy M., Webster L., Baker T., Waraich A. Towards fog driven IoT healthcare: challenges and framework of fog computing in healthcare. *Proceedings of the 2nd international conference on future networks and distributed systems*. 2018. P. 1–7.
11. Taneja M., Davy A. Resource aware placement of data analytics platform in fog computing. *Procedia Computer Science*. 2016. Vol. 97. P. 153–156. DOI: 10.1016/j.procs.2016.08.295.
12. Khan M.A., Karim M., Kim Y. A two-stage big data analytics framework with real world applications using spark machine learning and long Short-term memory network. *Symmetry*. 2018. Vol. 10. No. 10. P. 485.
13. Fog Node Possible Architectures and Propose the Changes in Mechanism to Architecture to Make Contiki OS Real Time / D.J. Revar, et. al. *International Journal of Scientific Research and Engineering Development*. 2020. Volume 3. Issue 3. P. 900–905..
14. Din S., Paul A. Smart health monitoring and management system: toward autonomous wearable sensing for internet of things using big data analytics. *Future Generation Computer Systems*. 2019. Vol. 91. P. 611–619.
15. Смірнова Т.В., Поліщук Л.І. Дослідження хмарних технологій як сервісів для системи інженерних розрахунків. *Кібербезпека та інформаційні технології* : монографія / М-во освіти і науки України, Центральнoукраїн. нац. техн. ун-т. Харків : ДІСА ПЛЮС, 2020. С. 106–121.
16. Усік П.С. Методи підвищення ефективності розподіленої обробки даних в комп'ютерних системах операторів стільникового зв'язку. Кваліфікаційна наукова праця на правах рукопису. 2021. URL: <https://er.chdtu.edu.ua/handle/ChSTU/2280> (дата звернення: 24.07.2021)
17. Janakiram MSV. Is Fog Computing The Next Big Thing In Internet of Things? URL: <https://www.forbes.com/sites/janakirammsv/2016/04/18/is-fog-computing-the-next-big-thing-in-internet-of-things/> (Last accessed: 17.03.2017).

#### **Kozak Ye.B. DATA ANALYSIS AND MACHINE LEARNING IN CLOUD AND FOG PLATFORMS AS A BASIS FOR EFFICIENT DATA TRANSFER**

*The article investigates data analysis and machine learning in cloud and fog platforms as a basis for efficient data transmission. The directions of formation of data sets, which are the most relevant in today's*

conditions, are determined. It is emphasized that modern paradigms of calculations and storage include: cloud, fog and boundary calculations, a description of the fundamental features of each species. It is emphasized that fog computing is a middle ground between cloud computing and edge computing paradigms, where data processing and analysis do not need to be located at the edge of the network or even in centralized storage. An established approach based on cloud computing technologies cannot provide the required quality of service in such conditions, in particular, in terms of reducing the delay time when transmitting large data sets. Foggy computing technology is now seen as a promising solution for processing large amounts of critical and time-sensitive data. This article considers the technology of cloud computing, as well as analyzes the prerequisites for the evolutionary development of this approach and the emergence of the concept of nebulous computing as a promising area of efficient data transmission. As part of the review of key features of nebulous computing, an analysis of critical cases with a combination of the concepts of nebulous and boundary computing. An overview of fog computing technologies is provided: virtualization, containerization and orchestration, as well as a systematic analysis of the most popular platforms that provide fog computing support. A generalized system model for joint processing of cloud platforms in heterogeneous networks is proposed and a fusion of IoT architecture and data analysis is schematically presented. The concept of predictive models is defined and their types are revealed, it is emphasized that modern methods of predictive modeling include statistical models based on regression, decision tree and models based on neural network or deep neural network.

**Key words:** machine learning, artificial intelligence, fog, cloud, computation, data transfer, platform, data analysis.