

УДК 004.85:004.7

МАШИННЕ НАВЧАННЯ З ВИКОРИСТАННЯМ ХМАРНИХ ОБЧИСЛЕНЬ

Кучков Б.О.

Науковий керівник – доцент. Хаханова Г.В.

Харківський національний університет радіоелектроніки, каф. АПОТ
м. Харків, Україна

тел. (057) 702-13-26)

The given work discusses the long-term machine learning model using cloud computing. Designing a machine learning methodology in the cloud. Solving problems related to memory and limiting script execution by established cloud technology providers. Method for Executing Sequential Lambda Functions on AWS with Intermediate Results Stored in AWS EFS

У світі поширюється комерційне використання штучного інтелекту, та хмарних обчислень. Штучний інтелект на базі машинного навчання дозволяє вирішувати великий спектр завдань, воно допомагає автоматизувати рішення складних професійних задач в самих різних областях людської діяльності. Хмарні обчислення, стрімко набирають популярність по-перше через економічні причини, централізація серверів в дата центри, суттєво зменшують затрати на обслуговування. А також через наявність централізованих комунікацій усередині хмарного провайдера. Але у провайдерів які надають хмарні технології, є певні обмеження щодо часу виконання скрипту, та певний об'єм даних. Якщо об'єм даних вирішується переходом на більш дорогий тариф, то ліміт на виконання скрипту є загальноприйнятим правилом, щоб не навантажувати увесь сервіс багато навантажувальними скриптами. У такій ситуації використання машинного навчання у хмарному середовищі не є реальним, так як, програма повинна «навчатись» багато часу. *Мета дослідження* – створення методології для можливого використання машинного навчання у хмарному середовищі. Вирішення сучасних проблем зі обмеженням часу використання скрипту. Впровадження підходу для уникнення додаткового серверу для машинного навчання. *Задача* – розробка моделі машинного навчання у хмарному середовищі, з використанням послідовних AWS Lambda функцій, зі зберіганням проміжного результату в AWS elastic file storage за допомогою смикання та комбінування подій.

Для створення моделі, використовуємо AWS Lambda — без серверна обчислювальна служба, яку пропонує Amazon Web Service. Обчислювальна потужність AWS Lambda збільшується зі збільшенням її конфігурація пам'яті. Отже, хмарна функція більшої пам'яті може отримати вигоду від вищої обчислювальної потужності в його розпорядженні. Щоб вирішити проблему зберігання, ми можемо створити систему з комбінацією EFS і лямбда. У підсумку для вирішення питання про сховище, ми будемо

використовувати комбінацію налаштування лямбда-визначення та додавання сховища за допомогою EFS. Щоб подолати часові обмеження, ми можемо використати комбінацію серійних і розподілених працівників. Після завершення процесу навчання працівники зберігатимуть градієнти в каталозі в EFS. Потім ці градієнти будуть підібрані лямбда-функцією, яка об'єднує градієнти та оновлює вагу. Після оновлення ваги об'єднана лямбда знову викликає робочих, і процес продовжується. Кількість робітників, які працюють паралельно, визначатиметься значеннями параметрів, які ми передамо на початку навчання. Для методології навчання ми будемо використовувати підхід розподіленого навчання з використанням міні-пакетного градієнтного спуску (mini BGD). Ми розподілимо набір даних «Машинне навчання з використанням без серверних обчислень» серед робочих вузлів, і кожен вузол оброблятиме та оновлюватиме вагу моделі. Ми плануємо включити частину сервера параметрів у кожен робочий вузол, таким чином усуваючи потребу в автономному сервері. Це зменшує мережеві накладні витрати на передачу та отримання параметрів і можуть бути безпосередньо отримані з підключеного сховища (Рис - 1).

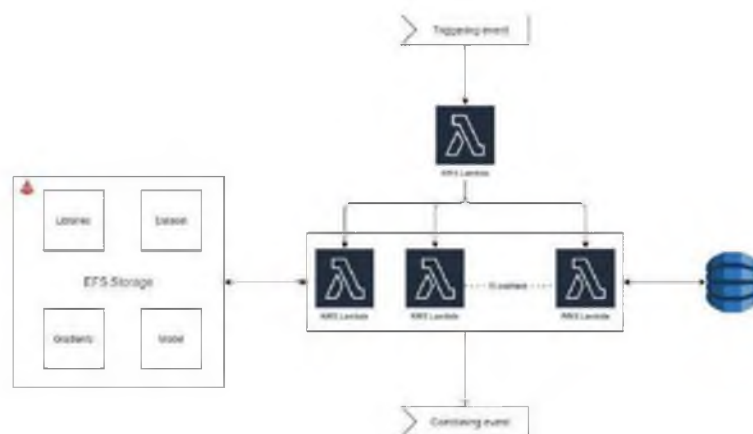


Рисунок 1 - Схема взаємодії компонентів у AWS

Наукова новизна – створення нового підходу до машинного навчання у хмарному середовищі, вирішення сучасних проблем із пам'яттю та часу виконання скрипту. Розробка методології використання машинного навчання у системі Amazon Web Service, з використанням AWS Lambda та AWS EFS.

Список джерел:

1. Warren E. Agin. A Simple Guide to Machine Learning [Text] // Business Law Today. 2017. – P. 1 – 5.
2. Хайкін С. Нейронні мережі: повний курс / С. Хайкін // М .: Діалектика, 2019. - 1104 с.
3. Каллан, Р. Нейронні мережі: Короткий довідник / Р. Каллан. - М .: Вільямс І.Д., 2017. - 288 с.