

CHROMA - БАЗА ДАНИХ З ВІДКРИТИМ ВИХІДНИМ КОДОМ

Хоча штучний інтелект має довгу історію, зараз переживаємо бум штучного інтелекту, який розпочався з переходом до глибокого навчання у 2010 році. З того часу було пережито три етапи розвитку інфраструктури штучного інтелекту. Перша хвиля інфраструктури штучного інтелекту з 2010 по 2015 рік була зарезервована для обраних в академічних колах і великих технологічних компаній, що публічно торгуються на біржі. Друга хвиля інфраструктури штучного інтелекту, приблизно з 2016 по 2020 рік, призвела до появи постачальників інфраструктури штучного інтелекту, які демократизували доступ до інструментарію за допомогою продуктів, отриманих на основі гіперскалерів.

Зараз суспільство перебуває у третій хвилі розвитку інфраструктури штучного інтелекту. Інженери-програмісти стали творцями штучного інтелекту, масово збільшуючи кількість практиків, а фундаментальні моделі значно знизили бар'єр для впровадження штучного інтелекту. Генеративний штучний інтелект, здатний створювати контент на кшталт ChatGPT, захопив світ.

Критично важливим шаром у стеку штучного інтелекту є база даних вбудовувань - база даних, створена з нуля на основі робочого процесу штучного інтелекту з вбудовуваннями. Вбудовування - це щільне числове представлення об'єктів і зв'язків, яке виражається у вигляді вектора. Векторний простір кількісно визначає семантичну схожість між категоріями. Вектори включень, які є близькими один до одного, вважаються схожими. Вбудовування можна використовувати для точного представлення неструктурованих даних (наприклад, зображень, відео та природної мови) або структурованих даних (наприклад, потоків кліків і покупок в електронній комерції).

Вбудовані бази даних спочатку виникли для підтримки таких варіантів використання, як персоналізація, рекомендації, семантичний пошук і комп'ютерний зір, але генеративний штучний інтелект є величезним попутним вітром. Генеративні моделі є фантастичними, але вони обмежені централізованою та загальною базою знань, що ускладнює якісне виконання специфічних для домену або конкретного завдання висновків. Вбудовування баз даних дозволяє командам використовувати генеративні моделі з власними даними за допомогою генерації, доповненої пошуком, для створення індивідуальних, розширених додатків штучного інтелекту.

Вважається, що під час третьої хвилі штучного інтелекту з'явиться новий стек штучного інтелекту, орієнтований на розробників, який дозволить створювати додатки, що вражають уяву. Переможні рішення штучного інтелекту будуть зосереджені на оптимізації взаємодії з кінцевим користувачем, враховуючи простоту використання, ергономічність та продуктивність. Chroma - база даних для вбудовування з відкритим вихідним кодом на основі штучного інтелекту.

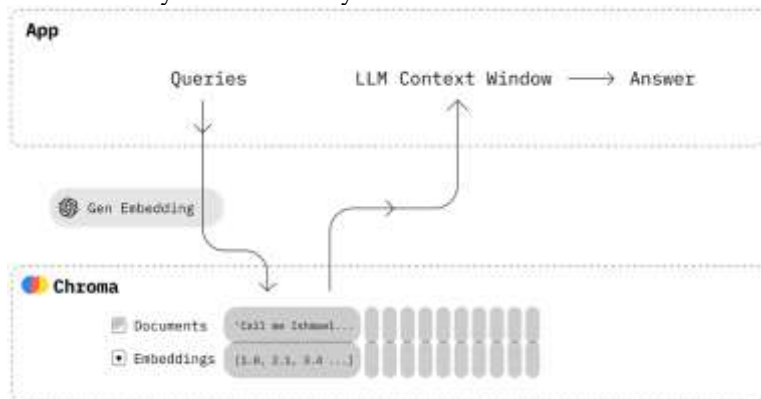


Рисунок 1 – Схематичне зображення Chroma

Джефф Хубер і Антон Тройніков, які мають безпосередній досвід роботи зі штучним інтелектом у Facebook, Nuro і Standard Cyborg, заснували Chroma з думкою про розробника. Використовуючи вбудовування, Chroma дозволяє розробникам додавати стан і пам'ять до своїх додатків зі штучним інтелектом. Chroma поставляється «в комплекті» з усім необхідним для зберігання, вбудовування та запиту даних з потужними функціями, такими як вбудована фільтрація, а незабаром з'явиться ще більше функцій, таких як автоматична кластеризація та релевантність запитів. Він має реалізацію як python, так і typescript API, а також вбудовану підтримку OpenAI та LangChain. Важливо, що він працює як у вбудованій конфігурації в пам'яті (як DuckDB), так і в клієнт-серверній версії. Незабаром з'явиться повністю керована хостингова версія.

Запущена в лютому 2023 року, Chroma була завантажена 35 тисяч разів за останній місяць. Його дуже легко встановити. Просто встановити chromadb за допомогою рір, щоб почати. Сьогодні оголошено, що Quiet Capital очолив посівний раунд Chroma на суму \$18 млн.