

TRƯỜNG ĐẠI HỌC PHENIKAA



KHOA CÔNG NGHỆ THÔNG TIN

Báo cáo học phần : Điện toán đám mây

Đề tài: Triển khai 1 mô hình học máy trên nền tảng đám mây

Họ và tên sinh viên: Nguyễn Ngọc Hiếu

Mã sinh viên: 21011956

Lớp: Điện toán đám mây (N01)

Giảng viên hướng dẫn: Nghiêm Việt Cường

Hà Nội, 06/2024

Lời cam kết

Họ và tên sinh viên: Nguyễn Ngọc Hiếu

Điện thoại liên lạc: 0921218648. Email: 21011956@st.phenikaa-uni.edu.vn

Lớp: K15_KHMT

Hệ đào tạo: Cử nhân

Tôi cam kết Bài tập lớn (BTL) là công trình nghiên cứu của bản thân/nhóm tôi. Các kết quả nêu trong BTL là trung thực, là thành quả của riêng tôi, không sao chép theo bất kỳ công trình nào khác. Tất cả những tham khảo trong BTL – bao gồm hình ảnh, bảng biểu, số liệu, và các câu từ trích dẫn – đều được ghi rõ ràng và đầy đủ nguồn gốc trong danh mục tài liệu tham khảo. Tôi/chúng tôi xin hoàn toàn chịu trách nhiệm với dù chỉ một sao chép vi phạm quy chế của nhà trường.

Hà Nội, ngày tháng năm

Tác giả/nhóm tác giả BTL

Họ và tên sinh viên

Mục lục

Lời cam kết	2
Chương 1: Giới thiệu về ứng dụng	4
1.1 Đặt vấn đề	4
1.2 Giải pháp	4
Chương 2 : Thiết kế giải pháp điện toán đám mây	5
2.1 Tổng quan	5
2.2 Phân tích các thành phần được triển khai	5
Chương 3: Công nghệ deploy sử dụng	8
3.1. Flask	8
3.2 WinSCP	8
3.3 PuTTY	9
3.4 Quy trình hoạt động:	10
Chương 4 : Bảo mật ứng dụng	10
4.1 Bảo mật dữ liệu	10
4.2 Kiểm soát truy cập	10
Chương 5: Phân tích chi phí	11
Chương 6 : Triển khai ứng dụng, kiểm thử và đánh giá	11
6.1 Kết quả triển khai ứng dụng:	11
6.2 Kết luận	12
6.2 Đề xuất cải tiến, phát triển trong tương lai	13

Chương 1: Giới thiệu về ứng dụng

1.1 Đặt vấn đề

Trong những năm gần đây, trí tuệ nhân tạo (AI) đã chứng kiến sự phát triển vượt bậc và được ứng dụng rộng rãi trong nhiều lĩnh vực khác nhau, từ y tế, tài chính đến thương mại điện tử và nhiều lĩnh vực khác. Một trong những ứng dụng phổ biến của AI là tự động tính toán. Trong báo cáo này chúng tôi sẽ tập chung vào một mô hình MachineLearning: Linear Regression đơn giản để hỗ trợ việc tự động tính lương. Đây là một thuật toán phổ biến trong học máy và việc giải quyết nó không chỉ có giá trị học thuật mà còn mang lại nhiều ứng dụng thực tiễn trong đời sống. Tuy nhiên để mô hình machine learning có thể hoạt động được thì cần phải có phần cứng mạnh mẽ, đến thời điểm hiện tại, để sở hữu phần cứng mạnh mẽ là một vấn đề nan giải với nhiều cá nhân và doanh nghiệp ở mức trung bình trở xuống.

1.2 Giải pháp

Từ vấn đề trên và nhiều vấn đề khác liên quan, dịch vụ điện toán đám mây ra đời. Dịch vụ này cho phép các doanh nghiệp và nhà phát triển mở rộng hạ tầng công nghệ theo nhu cầu mà không cần đầu tư lớn vào phần cứng. Thay vì đầu tư vào phần cứng và phần mềm đắt đỏ, người dùng chỉ cần trả tiền cho tài nguyên họ sử dụng trên đám mây. Các dịch vụ đám mây cung cấp một loạt các công cụ và dịch vụ để hỗ trợ việc triển khai, huấn luyện và quản lý các mô hình AI.

Những nhà cung cấp nổi bật như Google Compute Engine, Microsoft (AWS), hay các doanh nghiệp tại Việt Nam như Viettel, VNPT đang là những lựa chọn nổi bật hiện nay. Trong khuôn khổ báo cáo này, dựa trên tài nguyên miễn phí từ Learner Lab của AWS. Tôi sẽ thực hiện một dự án đơn giản tính toán lương trên hệ thống máy ảo EC2 và mô hình được lưu trong S3 bucket.

Chương 2 : Thiết kế giải pháp điện toán đám mây

2.1 Tổng quan

Để triển khai mô hình LinearRegression trên hệ thống EC2 của AWS, chúng ta sẽ sử dụng các dịch vụ chính của AWS như EC2, S3 và IAM. Sau đây là 1 số chi tiết về cách thiết kế giải pháp điện toán đám mây

2.2 Phân tích các thành phần được triển khai

a. Amazon EC2 Instance



Hình: AWS EC2

Amazon Elastic Compute Cloud (Amazon EC2) là một cơ sở hạ tầng điện toán đám mây được cung cấp bởi Amazon Web Services (AWS) giúp cung cấp tài nguyên máy tính ảo hoá theo yêu cầu. EC2 cung cấp một hoặc nhiều máy chủ ảo có thể kết hợp với nhau để dễ dàng triển khai ứng dụng nhanh nhất và đảm bảo sẵn sàng cao nhất.

Dịch vụ này cung cấp linh hoạt và khả năng mở rộng, cho phép bạn triển khai ứng dụng và công việc tính toán của mình trên các máy ảo được quản lý một cách dễ dàng. Mỗi EC2 instance được cung cấp với tài nguyên

tính toán như CPU, bộ nhớ RAM, lưu trữ đĩa và băng thông mạng, và người dùng có thể chọn loại instance phù hợp với yêu cầu của ứng dụng hoặc công việc cụ thể của họ. EC2 là một trong những dịch vụ phổ biến nhất của AWS và được sử dụng rộng rãi trong các ứng dụng web, phát triển phần mềm, phân tích dữ liệu, máy học, và nhiều lĩnh vực công nghệ thông tin khác.

b. Amazon S3(Simple Storage Service)



Hình: Amazon S3

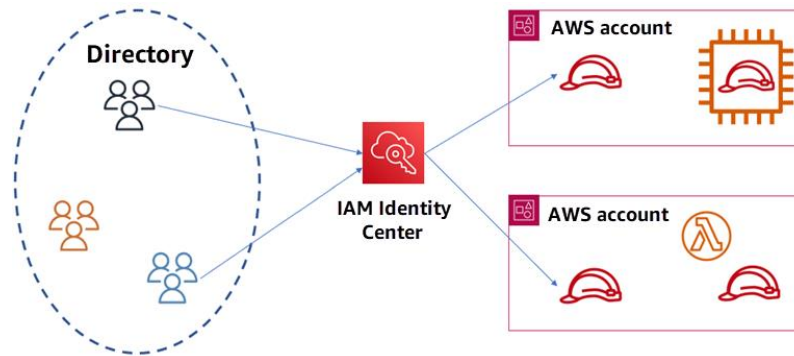
Amazon Simple Storage Service (Amazon S3) là một dịch vụ lưu trữ đối tượng cung cấp khả năng mở rộng, tính sẵn có của dữ liệu, bảo mật và hiệu suất hàng đầu trong ngành. Khách hàng thuộc mọi quy mô và ngành nghề có thể sử dụng Amazon S3 để lưu trữ và bảo vệ mọi lượng dữ liệu cho nhiều trường hợp sử dụng.

Chẳng hạn như kho dữ liệu, trang web, ứng dụng di động, sao lưu và khôi phục, lưu trữ, ứng dụng doanh nghiệp, thiết bị IoT và dữ liệu lớn phân tích. Amazon S3 cung cấp các tính năng quản lý để bạn có thể tối ưu hóa, sắp xếp và thiết lập cấu hình quyền truy cập vào dữ liệu của mình. Nhằm đáp ứng các yêu cầu tuân thủ, tổ chức và kinh doanh cụ thể của bạn. Dịch vụ

này cung cấp một nền tảng linh hoạt, bền bỉ và có khả năng mở rộng, giúp các tổ chức và cá nhân lưu trữ dữ liệu một cách an toàn và tiết kiệm chi phí.

Dự án này sẽ sử dụng S3 để có thể lưu mô hình và dataset

c. IAM (Identity and Access Management)



Hình: Amazon IAM

AWS IAM (Identity and Access Management) là một dịch vụ quản lý danh tính và quyền truy cập của người dùng vào các tài nguyên của Amazon Web Services (AWS). IAM cho phép bạn quản lý và kiểm soát ai có thể truy cập vào tài nguyên nào trong tài khoản AWS của bạn và cách họ có thể truy cập.

Trong dự án này, chúng ta sử dụng Learner Lab nên IAM đã được tạo từ trước, sử dụng cho việc triển khai và vận hành mô hình. Vai trò này cần quyền truy cập đến các dịch vụ EC2 và S3 , nhưng chỉ có quyền truy cập nhất định vào các tài nguyên.

Chương 3: Công nghệ deploy sử dụng

3.1. Flask

Flask là loại framework web phổ biến được viết bằng trình lập ngôn ngữ Python. Công nghệ thường được sử dụng để xây dựng trang web từ những ứng dụng đơn giản đến những hệ thống phức tạp hơn. Flask được thiết kế để hoạt động và mở rộng một cách hiệu quả, đồng thời nó cũng cung cấp các công cụ và thư viện cần thiết để phát triển ứng dụng web hiệu quả. Flask cũng có cộng đồng sáng tạo và hỗ trợ mạnh mẽ từ cộng đồng Python.

Flask được thiết kế để hoạt động và tùy chỉnh, không giới hạn chế độ người dùng trong cách tổ chức ứng dụng của họ. Công nghệ cung cấp công cụ tự động để lựa chọn các gói và thư viện bổ sung. Nó thích hợp cho các dự án nhỏ đến trung bình, nơi hoạt động và tùy chỉnh được yêu cầu chính.

3.2 WinSCP

WinSCP (Windows Secure Copy) là một ứng dụng SFTP và FTP Client miễn phí, mã nguồn mở, đặc biệt được thiết kế cho hệ điều hành Windows. Tích hợp chức năng truyền tải file an toàn và quản lý tệp tin, WinSCP là công cụ hiệu quả giúp người dùng kết nối và chuyển đổi dữ liệu giữa máy tính cá nhân của người dùng và máy chủ từ xa.

Chức năng chính của WinSCP bao gồm việc đảm bảo an toàn và bảo mật trong việc truyền tải dữ liệu qua các giao thức như FTP, SFTP và WebDAV. Giao diện của ứng dụng được chia thành hai phần, hiển thị cấu trúc file và thư mục của máy tính cá nhân ở phía bên trái và máy tính từ xa ở phía bên phải, tạo thuận lợi cho việc sao chép file giữa các hệ thống. WinSCP là một

công cụ truyền tải file nổi bật với nhiều ưu điểm đáng chú ý, làm cho nó trở thành lựa chọn phổ biến cho người dùng.

3.3 PuTTY



Hình PUTTY

PuTTY là phần mềm sử dụng để điều khiển server thông qua mạng internet, Nó hỗ trợ nhiều giao thức mạng, bao gồm SCP, SSH, Telnet, rlogin... PuTTY ban đầu được viết dành riêng cho hệ điều hành Windows, nhưng hiện nay nó đã được viết cho nhiều hệ điều hành khác như Unix, hệ điều hành MacOS, Symbian, Windows Mobile và android.

PuTTY hỗ trợ nhiều biến thể trên "các thiết bị đầu cuối" từ xa an toàn, và cung cấp cho người dùng trình điều khiển các SSH với khóa mã hóa, các giao thức, thuật toán mã hóa thay thế như 3DES, Arcfour, Blowfish, DES, và khóa công khai xác thực. Các lớp giao tiếp mạng hỗ trợ IPv6, và các giao thức SSH hỗ trợ các chương trình nén openssh bị trì hoãn. Nó cũng có thể được sử dụng với các kết nối cổng nối tiếp trong mạng LAN.

3.4 Quy trình hoạt động:

- + Sử dụng WinSCP để truy cập vào máy chủ cloud từ máy tính cá nhân và chuyển đổi các tệp tin của Flask application (ví dụ: mã nguồn, cấu hình, và các tệp tin liên quan) lên máy chủ.
- + Sử dụng PuTTY để truy cập vào máy chủ cloud và thiết lập môi trường cho Flask application. Điều này có thể bao gồm cài đặt Python, cài đặt các thư viện phụ thuộc, và cấu hình máy chủ web như Nginx hoặc Apache.
- + Sao chép các tệp tin của Flask application từ máy tính cá nhân lên máy chủ sử dụng WinSCP. Chạy Flask application trên máy chủ, thường thông qua một môi trường ảo.

Chương 4 : Bảo mật ứng dụng

4.1 Bảo mật dữ liệu

Amazon EC2 (Elastic Compute Cloud) sử dụng mã hóa RSA để đảm bảo rằng các kết nối SSH đến các phiên bản (instances) của nó là an toàn. Mã hóa RSA là một hệ thống mã hóa khóa công khai, giúp bảo vệ dữ liệu và xác thực danh tính giữa máy khách và máy chủ. Sử dụng các giao thức bảo mật như HTTPS, SSL/TLS để mã hóa dữ liệu khi truyền tải giữa người dùng và máy chủ.

Mã hóa dữ liệu khi lưu trữ (At-rest): Sử dụng các công cụ mã hóa như AWS KMS (Key Management Service) để mã hóa dữ liệu lưu trữ trên cơ sở dữ liệu, hệ thống tệp, hoặc các dịch vụ lưu trữ như AWS S3.

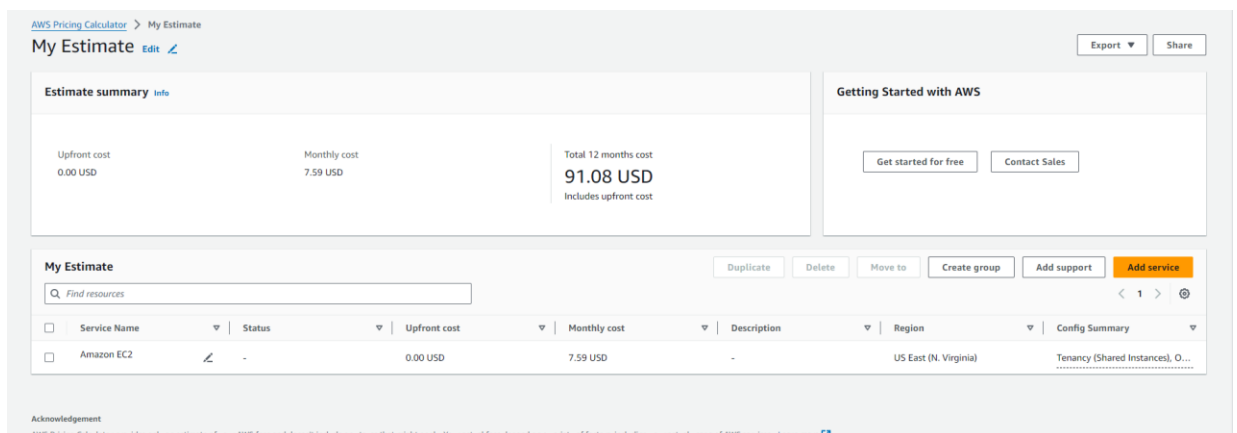
4.2 Kiểm soát truy cập

Quản lý danh tính và truy cập (IAM)

Trong dự án này bởi vì sử dụng Learner Lab nên IAM đã được thiết lập sẵn một labrole, không thể thực hiện quản lý người dùng và truy cập, tuy nhiên trong thực tế, với IAM có thể kiểm soát bằng những cách như phân quyền (Authorization), Kiểm soát truy cập dựa trên vai trò (Role-Based Access Control - RBAC) Kiểm soát truy cập dựa trên thuộc tính (Attribute-Based Access Control - ABAC).

Chương 5: Phân tích chi phí

Với cấu hình và tài nguyên sử dụng như trên, bằng việc sử dụng AWS Pricing Calculator tôi đã ước lượng được chi phí sử dụng của ứng dụng web như hình bên dưới

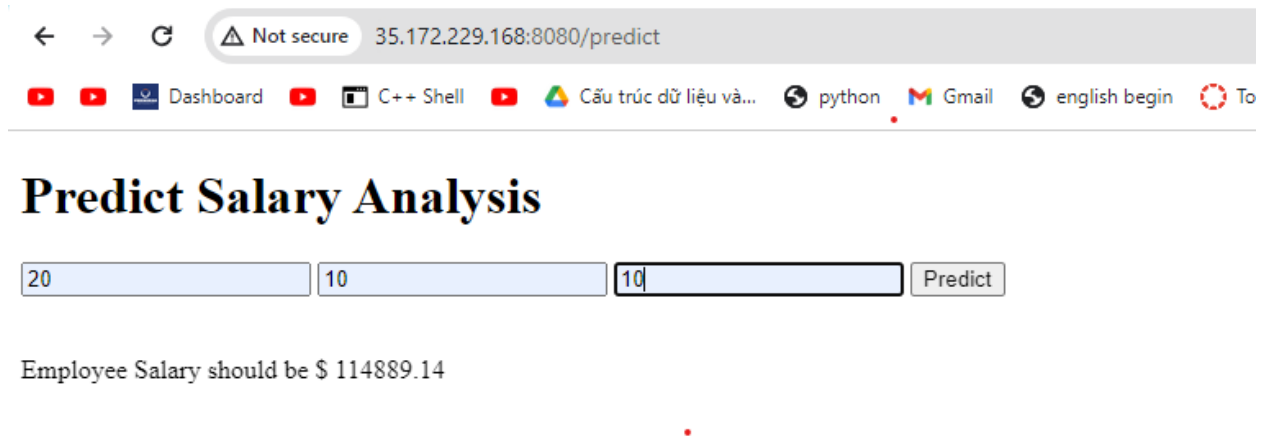


Chi phí

Chương 6 : Triển khai ứng dụng, kiểm thử và đánh giá

6.1 Kết quả triển khai ứng dụng:

Demo



← → ↻ Not secure 35.172.229.168:8080/predict

Dashboard C++ Shell Cấu trúc dữ liệu và... python Gmail english begin To

Predict Salary Analysis

Employee Salary should be \$ 114889.14

6.2 Kết luận

Trong quá trình triển khai mô hình trên hệ thống EC2 của AWS, chúng ta đã đạt được những kết quả sau:

Thiết Lập Hạ Tầng Điện Toán Đám Mây:

Triển khai thành công các instance EC2 với cấu hình phù hợp để huấn luyện mô hình AI, sử dụng loại instance P3 với GPU mạnh mẽ. Tạo và sử dụng AMI tùy 11 chỉnh chứa hệ điều hành và các công cụ cần thiết để đảm bảo môi trường triển khai nhất quán và dễ dàng tái tạo.

Quản Lý Dữ Liệu:

Sử dụng Amazon S3 để lưu trữ dữ liệu đầu vào và đầu ra một cách hiệu quả, với khả năng mở rộng và quản lý phiên bản giúp theo dõi và bảo vệ dữ liệu. Thiết lập quyền truy cập hợp lý để đảm bảo dữ liệu được bảo mật và chỉ những người/instance được phép mới có thể truy cập.

Quản Lý và Bảo Mật Hệ Thống:

Sử dụng IAM để quản lý quyền truy cập, đảm bảo rằng chỉ những người/instance cần thiết mới có quyền truy cập vào các tài nguyên quan trọng. Thiết lập bảo mật đa lớp với MFA và các chính sách bảo mật để giảm thiểu rủi ro bảo mật

6.2 Đề xuất cải tiến, phát triển trong tương lai

Tối Ưu Hóa Chi Phí

Tận dụng các chính sách thanh toán theo yêu cầu và spot instances để giảm chi phí vận hành. Sử dụng các công cụ theo dõi và quản lý chi phí của AWS để tối ưu hóa việc sử dụng tài nguyên.

Cải Thiện Quy Trình Quản Lý Dữ Liệu:

Sử dụng các công cụ như Amazon Athena để phân tích dữ liệu trực tiếp trên S3 mà không cần di chuyển dữ liệu. Áp dụng các chính sách quản lý dữ liệu tự động, chẳng hạn như lifecycle policies, để quản lý và tối ưu hóa dung lượng lưu trữ.

Bảo Mật và Tuân Thủ:

Đánh giá và cập nhật thường xuyên các chính sách bảo mật để đảm bảo hệ thống luôn tuân thủ các quy định và tiêu chuẩn bảo mật mới nhất. Sử dụng các dịch vụ bảo mật bổ sung như AWS Shield, AWS WAF để bảo vệ hệ thống trước các cuộc tấn công DDoS và các mối đe dọa khác.

