

BỘ TÀI NGUYÊN VÀ MÔI TRƯỜNG
TRƯỜNG ĐẠI HỌC TÀI NGUYÊN VÀ MÔI TRƯỜNG TP. HCM
KHOA HỆ THỐNG THÔNG TIN & VIỄN THÁM



BÁO CÁO ĐỒ ÁN MÔN HQT CSDL
Designing an Accounting Information Management
System Using

Thành viên của nhóm:

1. Võ Nguyễn Đức Toàn
2. Nguyễn Nam Khởi

Lớp : 09_CNTT4

Thành phố Hồ Chí Minh, tháng 11, năm 2022

LỜI MỞ ĐẦU

Ngày nay, cùng với sự phát triển của xã hội, ngành CNTT cũng có nhiều bước phát triển đáng kể và đã dần đi vào cuộc sống của mọi người và được sử dụng hầu hết trong tất cả các ngành nghề trong xã hội. Với số lượng tài liệu trong các cơ quan, tổ chức tăng theo cấp số nhân theo từng năm, từng thời kỳ, thì việc lưu trữ số lượng tài liệu đó trở nên vô cùng khó khăn, đặc biệt khi nó là những tài liệu quan trọng mà lại được lưu trữ trên các thiết bị cứng. Yếu tố thời gian, các tác động bên ngoài có thể làm cho những tài liệu đó bị hỏng hóc, khó bảo quản. Do đó, nhu cầu sử dụng các phần mềm hỗ trợ khả năng lưu trữ các dữ liệu đảm bảo các yếu tố an toàn và tiện lợi trong thao tác với dữ liệu đó là vô cùng cần thiết. Và nhu cầu đó sẽ trở nên dễ dàng khi có sự vào cuộc của CNTT, đặc biệt là các hệ quản trị cơ sở dữ liệu (CSDL). Nói đến CNTT thì không thể không nói đến các hệ quản trị cơ sở dữ liệu. Đó là phần mềm hay hệ thống được thiết kế để quản trị một CSDL, nó hỗ trợ khả năng lưu trữ, sửa chữa, xóa và tìm kiếm trong tin trong một CSDL. Và có rất nhiều loại hệ quản trị CSDL khác nhau: từ phần mềm nhỏ chạy trên máy tính cá nhân cho đến những hệ quản trị phức tạp chạy trên một hoặc nhiều siêu máy tính. Đề tài báo cáo đồ án Designing an Accounting Information Management System Using bao gồm các nội dung:

Phần 1: Giới thiệu.

Phần 2: Tổng quan về quản lý thông tin hệ thống kế toán.

Phần 3: Thông tin quản lý kế toán.

Phần 4: Thiết kế AIMS.

Phần 5: Kiểm tra và mô phỏng hệ thống thông tin quản lý kế toán.

Phần 6: Kết Luận.

Phần 7: Demo.

LỜI CẢM ƠN

Lời đầu tiên, nhóm em xin gửi lời tri ân sâu sắc đến thầy **Phạm Trọng Huỳnh**. Trong quá trình tìm hiểu và học tập bộ môn **Hệ quản trị cơ sở dữ liệu**, nhóm em đã nhận được sự giảng dạy và hướng dẫn rất tận tình, tâm huyết của thầy. Thầy đã giúp nhóm em tích lũy thêm nhiều kiến thức hay và bổ ích. Từ những kiến thức mà thầy truyền đạt, nhóm em xin trình bày lại những gì mình đã tìm hiểu về đề tài đồ án Thiết kế hệ thống quản lý thông tin kế toán bằng cách sử dụng dữ liệu lớn và công nghệ đám mây gửi đến thầy. Mặc dù nhóm em đã có sự cố gắng, nhưng trong khoảng thời gian cho phép cũng như hạn chế về kiến thức nên đồ án này của em không thể tránh khỏi những kiến thức thiếu sót. Chính vì vậy, nhóm em rất mong nhận được sự góp ý của các thầy cô giáo và bạn bè. Kính chúc thầy hạnh phúc và thành công hơn nữa trong sự nghiệp “trồng người”. Kính chúc thầy luôn dồi dào sức khỏe để tiếp tục dìu dắt nhiều thế hệ học trò đến những bến bờ tri thức.

Sinh Viên Thực Hiện

Nhóm 10

MỤC LỤC

1.Giới thiệu	5
1.1 Những công việc liên quan đến dữ liệu lớn và công nghệ đám mây.	7
2.Tổng quan về quản lý thông tin hệ thống kế toán trên cơ sở dữ liệu lớn và điện toán đám mây(AIMS).....	9
2.1 Điện toán đám mây	9
2.2 Cấu trúc logic của Công nghệ đám mây.	10
3.Thông tin hệ thống quản lý kế toán.	11
3.1 Thông tin kế toán dựa trên dữ liệu lớn.....	11
3.2 Điện toán đám mây trong kỷ nguyên AIMS.....	14
3.3 Mô hình năng lượng máy Boltzmann bị hạn chế dựa trên thuật toán học sâu.	15
4.Thiết kế AIMS dựa trên Dữ Liệu lớn và Công nghệ đám mây.....	16
4.1 Cấu trúc chương trình của Quản lý thông tin kế toán.	16
4.2 Mô hình SAAS của hệ thống quản lý thông tin kế toán.....	17
4.3 Lưu trữ phân tán dựa trên nền tảng đám mây.....	17
5. Kiểm tra và mô phỏng hệ thống quản lý thông tin kế toán.....	18
5.1 Thời gian đáp ứng công việc.....	20
5.2 Phân tích vị trí nhiệm vụ hệ thống và tốc độ tăng tốc.	24
5.2.1. Phân tích vị trí nhiệm vụ.....	24
5.2.2 Phân tích tốc độ.	25
6. Kết Luận.....	26
7. ỨNG DỤNG	27
8.DEMO.....	28
Tài liệu tham khảo.....	37

1. Giới thiệu .

- Hiện nay, sự hội tụ của trí tuệ nhân tạo và học máy đã tạo ra máy tính như một công cụ quan trọng cho năng suất và cuộc sống hàng ngày của con người. Những đổi mới này đã đóng một vai trò quan trọng, đặc biệt là trong lĩnh vực xử lý dữ liệu, nơi chúng không chỉ có thể lưu trữ lượng dữ liệu khổng lồ mà còn thống kê và phân tích chúng cho đến khi điều tra thêm về giá trị thực tế của các ngành dữ liệu.
- Từ khi cơ sở dữ liệu ứng dụng ra đời, con người xử lý thông tin dễ dàng và chính xác hơn rất nhiều. Do đó, các nguồn dữ liệu ngày càng trở nên quan trọng trong các lĩnh vực kinh doanh, quản trị và điều tra.
- Với sự trưởng thành của công nghệ máy tính và việc sử dụng máy tính kể từ đầu thiên niên kỷ 21, nhiều lĩnh vực của cuộc sống đã tạo ra nhiều dữ liệu hơn mức có thể trong khi đạt được những tiến bộ to lớn. Do những điều trên, nhiều người sử dụng thuật ngữ “bùng nổ thông tin” để mô tả sự gia tăng theo cấp số nhân của dữ liệu. Sự phong phú của thông tin cũng gây ra nhiều thách thức trong các hoạt động và cuộc sống hàng ngày của mọi người, đặc biệt là những vấn đề liên quan đến cách tiếp cận thông tin liên quan một cách hiệu quả. Chỉ thông qua sự hiểu biết về động lực thị trường và nhu cầu của người dùng cũng như cung cấp các sản phẩm và dịch vụ đáp ứng các cơ hội thị trường cho khách hàng thì các doanh nghiệp hiện đại mới có thể thịnh vượng và phát triển. Trong một thị trường cạnh tranh cao, phải tìm kiếm thông tin có ý nghĩa để tận dụng tối đa các nguồn dữ liệu để theo dõi người dùng, nghiên cứu thị trường và ra quyết định phân tích. Trong quy trình trích xuất thông tin hữu ích, thông tin không liên quan đã trở thành vấn đề đau đầu đối với các doanh nghiệp, không chỉ ảnh hưởng đến hiệu quả xử lý thông tin mà còn đóng vai trò lừa đảo. Do đó, công nghệ thu thập dữ liệu có giá trị to lớn và cũng tạo ra sự tăng trưởng mạnh mẽ trong tương lai.
- Vì quản lý tài chính gắn liền với sự tồn tại và phát triển của doanh nghiệp, được coi là rất quan trọng trong quản lý doanh nghiệp. Ngoài ra, nó còn là một tiêu chí quan trọng để đánh giá thực trạng hoạt động của doanh nghiệp. Với tốc độ phát triển nhanh chóng của công nghệ mạng và công nghệ thông tin, ngày

càng có nhiều doanh nghiệp chú trọng đến việc thông tin hóa quản lý tài chính . Bộ phận quản lý tài chính là khía cạnh quan trọng nhất trong quản lý kinh doanh của doanh nghiệp. Tổ tài chính sẽ tổng hợp các quan điểm kiểm toán tất cả các nghiệp vụ của doanh nghiệp, nhân viên phòng tài chính sẽ hoàn thành việc kiểm toán thống nhất. Vì vậy, việc cải thiện quản lý thông tin tài chính doanh nghiệp là cấp thiết. Bài báo sử dụng công nghệ đám mây dữ liệu lớn để thiết kế hệ thống quản lý thông tin kế toán. Công nghệ đám mây dữ liệu lớn sử dụng công nghệ Internet để cung cấp cho mọi người nhiều tài nguyên khác nhau. Cấu trúc của nó chủ yếu có ba cấp độ bao gồm nền tảng, tài nguyên và ứng dụng. Nó sử dụng lưu trữ tài nguyên được chia sẻ và gọi tài nguyên để cải thiện tính khả dụng của đối tượng. Trong thời đại của những tiến bộ nhanh chóng trong công nghệ xử lý dữ liệu, các tổ chức không còn hài lòng với các lỗi tất như nghiên cứu ngẫu nhiên và phân tích mẫu. Chúng yêu cầu một khối lượng đáng kể dữ liệu quy mô lớn từ nhiều miền, cũng như quản lý kế toán trong hệ thống đào tạo dựa trên khung kế toán dựa trên đám mây. Khả năng xây dựng hiệu quả khung triển khai AMIS dựa trên kế toán đám mây ngày càng phụ thuộc vào khả năng truy cập dữ liệu khổng lồ. Cho phép thu thập dữ liệu lớn bên trong môi trường dữ liệu lớn đang trở thành một vấn đề rất quan trọng trong các cộng đồng công nghiệp và học thuật đương đại.

- Bằng cách kiểm tra các nghiên cứu hiện tại, người ta phát hiện ra rằng phần lớn nghiên cứu xây dựng AIMS hiện tại tập trung vào môi trường Internet thông thường, trong khi nghiên cứu thiết kế AIMS tích hợp dữ liệu lớn và đám mây là rất ít. Thật vậy, khi công nghệ dữ liệu lớn tiến bộ, nó đóng một vai trò thiết yếu trong quản lý doanh nghiệp và ra quyết định tài chính, và việc sử dụng thành công các khả năng dữ liệu lớn được liên kết chặt chẽ với AIS dựa trên kế toán đám mây. Do đó, nghiên cứu này chứng minh khung AIMS dựa trên kế toán đám mây trong kỷ nguyên dữ liệu lớn. Nó giải thích chi tiết về trách nhiệm của từng thành phần trong khuôn khổ dựa trên việc phân tích liệu kế toán đám mây có thể sử dụng thành công các ứng dụng hệ thống thông tin kế toán hay không. Ngoài những điều trên, bài báo này sử dụng công nghệ đám mây dữ liệu lớn để tích hợp tối đa và hiệu quả các tài nguyên máy tính và cải thiện

thông tin và trí thông minh của hệ thống thông qua một mạng học tập chuyên sâu.

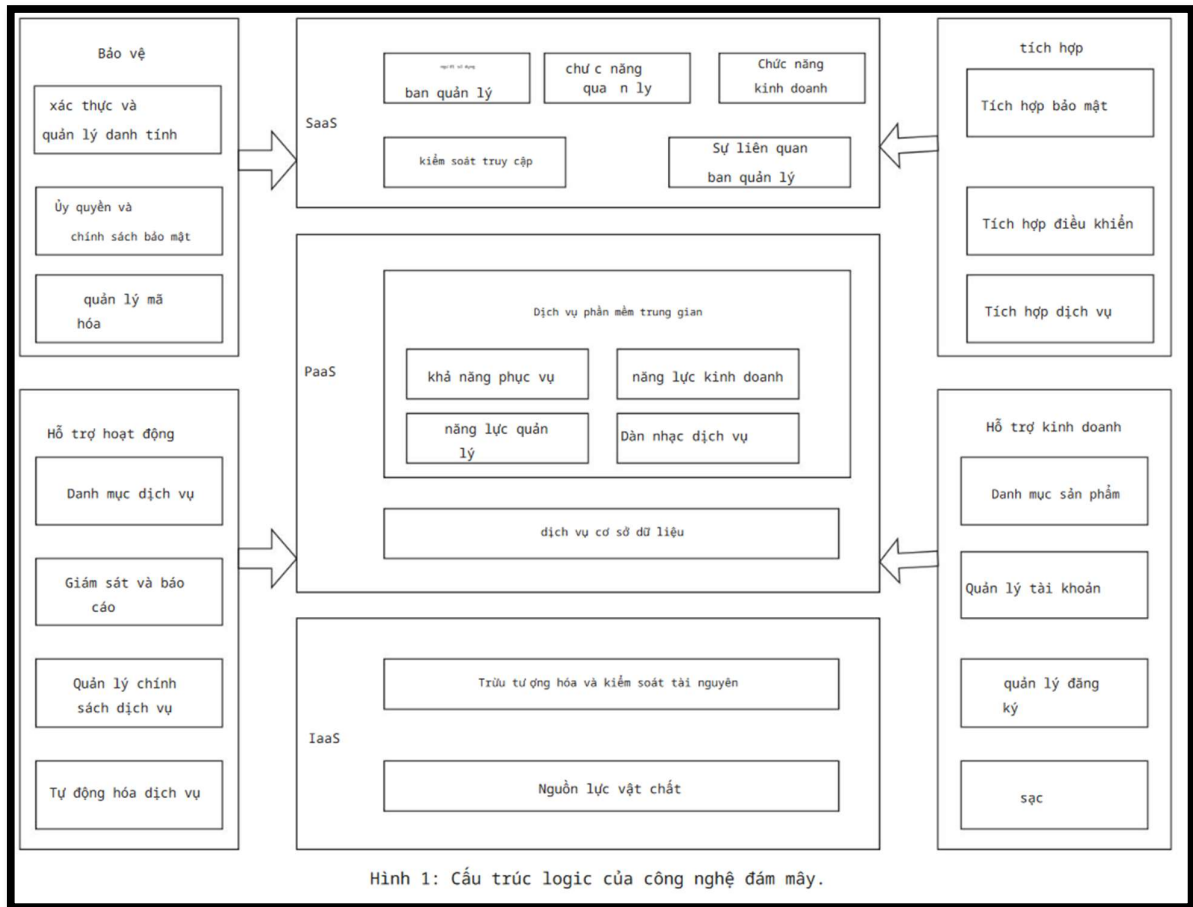
- Những đổi mới chính trong giai đoạn nghiên cứu của bài viết này như sau: Điện toán, cấu trúc logic của công nghệ đám mây và mô hình năng lượng RBM dựa trên thuật toán học tập. là thuật toán và lý thuyết đóng vai trò là cơ sở quan trọng cho việc thiết kế và phát triển hệ thống này. Dựa trên công nghệ đám mây dữ liệu lớn, kiến trúc tổng thể của hệ thống quản lý thông tin kế toán được thiết kế, mô hình SaaS của hệ thống được xây dựng và phương pháp lưu trữ phân tán được sử dụng để lưu trữ dữ liệu tài chính doanh nghiệp để hiện thực hóa văn phòng thông tin kế toán. Hình 1 làm nổi bật cấu trúc của bài viết này.
- Phần còn lại của bài viết này bao gồm 5 phần: Phần 1: Giới Thiệu. Phần 2 dựa trên các công trình liên quan của các học giả trong nước và quốc tế. Phần 3 trình bày tổng quan về công nghệ đám mây dữ liệu lớn cho Hệ thống quản lý thông tin kế toán (AIMS) và Phần 4 trình bày thiết kế AIMS dựa trên Công nghệ đám mây dữ liệu lớn. Phần 5 thử nghiệm và mô phỏng hệ thống. Phần 6 kết luận của bài báo.

1.1 Những công việc liên quan đến dữ liệu lớn và công nghệ đám mây.

- Với sự gia tăng nhanh chóng của dữ liệu đám mây, công nghệ đám mây đã trở thành một điểm nóng nghiên cứu mới trong và ngoài nước. Sub ramanian nghiên cứu sâu về các vấn đề bảo mật đám mây của ba thực thể chủ sở hữu dữ liệu, nhà cung cấp dịch vụ đám mây và người dùng đám mây ở cấp độ điện toán, giao tiếp, thỏa thuận cấp độ dịch vụ và dữ liệu. Các học giả trong đã đánh giá thực nghiệm tính thực tiễn của thông tin tài chính và dữ liệu kế toán trong quản lý chiến lược doanh nghiệp và thay đổi mô hình kế toán truyền thống dựa trên hệ thống hỗ trợ thông tin quản lý. Họ đã cố gắng mở rộng phạm vi thời gian kế toán và cải thiện quy trình lập ngân sách cũng như khả năng phân tích bối cảnh.
- Công việc ban đầu của đã phân tích vai trò của xã hội, môi trường và quản lý trong hệ thống kế toán và chọn dữ liệu ASG làm công cụ chính để kiểm tra sự phát triển bền vững. Về vấn đề này, nghiên cứu của đã chỉ ra rằng khả năng năng động của công ty, các điều kiện và môi trường thay đổi liên tục đều ảnh

hưởng đến hiệu quả của hệ thống quản lý kế toán, dẫn đến những thay đổi trong hệ thống thông tin quản lý kế toán. Các tác giả của đưa ra chiến lược điều chỉnh hệ thống thông tin kế toán quản trị hiệu quả.

- Thông tin kế toán đã được kiểm tra bằng phần mềm LISRELL và mô hình phương trình cấu trúc dựa trên phương sai cộng tác là dữ liệu tài chính quan trọng giúp các nhà quản lý doanh nghiệp hiểu được tình hình kinh doanh và xây dựng kế hoạch phát triển. Các phát hiện chỉ ra rằng phần mềm ứng dụng doanh nghiệp thành công có tác động trực tiếp đến chất lượng thông tin kế toán.
- Vẫn giữ quan điểm trên, công trình của đề xuất áp dụng mô hình điện toán đám mây trong quản lý thông tin kế toán của doanh nghiệp nhỏ. Nó cải thiện hiệu quả vận hành hệ thống bằng cách sử dụng xử lý phân tán trên đám mây, thiết kế hợp tác đám mây, mô phỏng đám mây và các công nghệ khác. Các học giả đề xuất quản lý thông tin kế toán giúp giảm chi phí dịch vụ quản lý chia sẻ tài chính, nâng cao chất lượng dịch vụ tài chính và tăng hiệu quả quản lý tổng thể. Đó là một mô hình thời đại con người mới.
- Trước khi thiết lập hệ thống chia sẻ tài chính, trước tiên doanh nghiệp cần xác định các yếu tố ảnh hưởng đến quản lý thông tin kế toán để đảm bảo an toàn cho dữ liệu tài chính doanh nghiệp và tính khả thi của quản lý thông tin kế toán. Tác giả đã sử dụng hệ thống quản lý thông tin kế toán tài chính trong các doanh nghiệp hóa chất của Trung Quốc để thay đổi hoàn toàn phương thức quản lý truyền thống bằng cách ứng dụng công nghệ thông tin trong quản lý kế toán tài chính. Nó tăng cường đáng kể hiệu quả quản lý kinh doanh bằng cách sử dụng đầy đủ công nghệ thông tin trong quản lý tài chính kế toán và nâng cao giá trị ứng dụng của công nghệ thông tin. Nghiên cứu này sử dụng mô hình và công nghệ trong thiết kế hệ thống quản lý thông tin kế toán bằng cách giải thích đơn giản ý tưởng về công nghệ đám mây và kiểm tra cấu trúc logic của công nghệ đám mây.



2. Tổng quan về quản lý thông tin hệ thống kế toán trên cơ sở dữ liệu lớn và điện toán đám mây (AIMS)

2.1 Điện toán đám mây

- Điện toán đám mây là sản phẩm quan trọng của sự phát triển nhanh chóng của xã hội và khoa học công nghệ. Tính năng chính của nó là nó có thể nhận ra các nền tảng đám mây khác nhau như điện toán đám mây và cung cấp dịch vụ cho người dùng một cách đơn giản và nhanh chóng. Về tài nguyên kỹ thuật, phân tích cũng có thể tối đa hóa việc sử dụng tài nguyên. Công nghệ điện toán đám mây có thể kết nối thông suốt các dữ liệu khác nhau trên mạng.
- Các doanh nghiệp và cá nhân là nhóm dịch vụ chính của điện toán đám mây. Nền tảng điện toán đám mây được thực hiện bằng cách sử dụng các công nghệ và thuật toán khác nhau, do đó chi phí phần cứng thấp. Có thể thực hiện nâng

cấp sau nếu công nghệ được thay thế và sản xuất mã. Không cần phải thêm hoặc thay thế phần cứng khác.

- Sau khi sử dụng công nghệ đám mây, hiệu quả hoạt động cũng được cải thiện đáng kể và sự phát triển của doanh nghiệp năng lực và trình độ phục vụ cũng nhanh chóng được nâng cao. Ở giai đoạn sau, các hoạt động bảo trì và nâng cấp có thể được hoàn thành trong môi trường mạng và sẽ không bị giới hạn bởi vị trí, thời gian, v.v.
- Hiện tại, mô hình dịch vụ đám mây bao gồm 3 loại: công khai, riêng tư và kết hợp. Các nhóm và cá nhân là những người dùng chính của đám mây riêng. Công chúng nói chung là người dùng chính của đám mây công cộng. Đám mây lai được tạo ra bằng cách kết hợp các đám mây công cộng và riêng tư. Các nhóm người dùng được phân loại dựa trên trạng thái người dùng của họ.

2.2 Cấu trúc logic của Công nghệ đám mây.

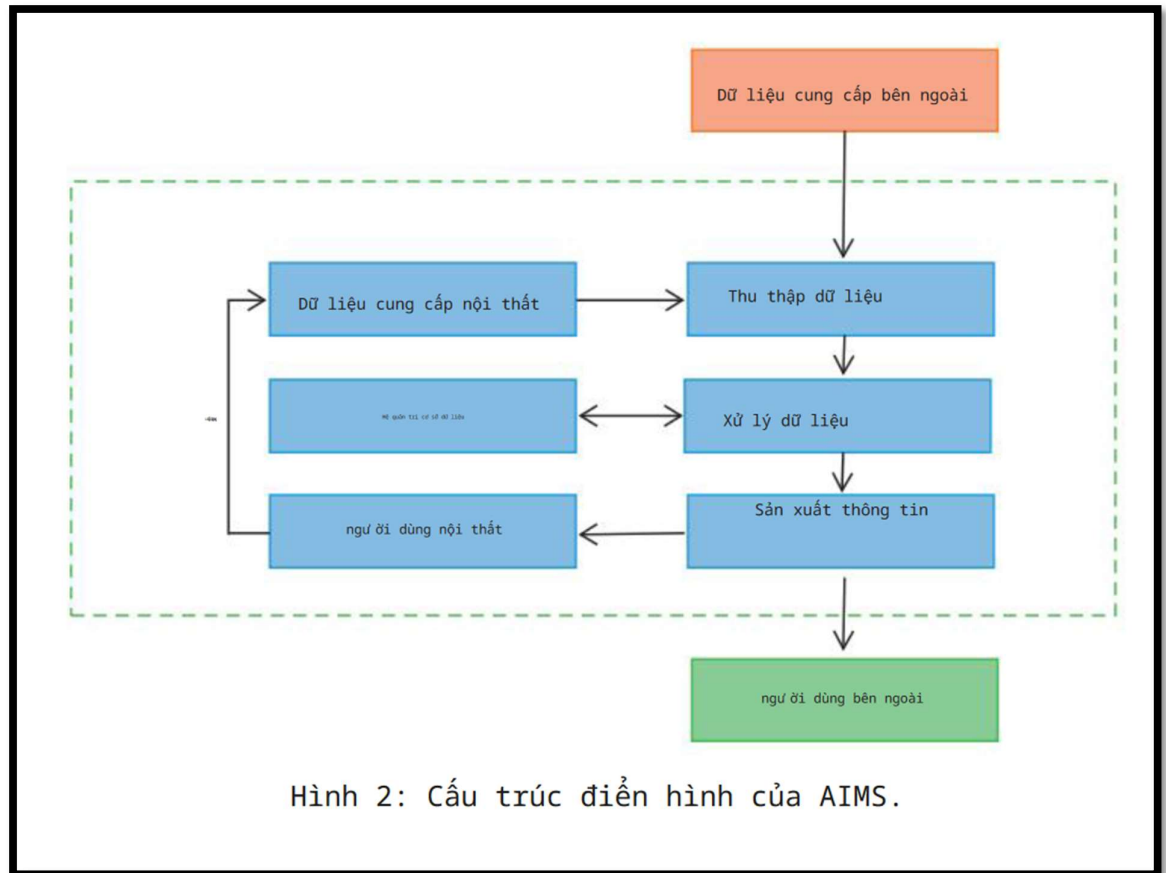
- NIST chia công nghệ đám mây thành nền tảng PaaS dưới dạng dịch vụ, phần mềm SaaS dưới dạng dịch vụ và cơ sở hạ tầng và dịch vụ IaaS. SaaS là một mô hình công nghệ đám mây được triển khai trên thiết bị của nhà cung cấp dịch vụ đám mây và được phân phối tới người tiêu dùng thông qua các ứng dụng web và giao diện trực tuyến. Người dùng không bắt buộc phải quản lý hay phát triển. PaaS có thể cung cấp cho các ứng dụng đang hoạt động và phát triển người dùng như một nền tảng phát triển thứ cấp. Người dùng có thể sử dụng các ứng dụng tích hợp sẵn để thực hiện các dịch vụ mà không cần quản lý phần cứng bên dưới. Mô hình IaaS cung cấp cho người dùng cơ sở hạ tầng ảo hóa, bao gồm mạng và máy chủ để người dùng có thể chạy ứng dụng đã triển khai. Các tổ chức trong và ngoài nước đưa ra nhiều sơ đồ cấu trúc công nghệ đám mây dựa trên ba cấp độ của công nghệ đám mây phát triển khung tham chiếu công nghệ đám mây bằng cách kết hợp ITU-T và SO/IEC JTC 1. Hình 1 cho thấy kiến trúc logic của công nghệ đám mây. Theo hình trên, sơ đồ xử lý cấu trúc công nghệ đám mây ở các quốc gia khác nhau, chủ yếu bao gồm lớp nền tảng PaaS, lớp tài nguyên IaaS, lớp ứng dụng SaaS,

hỗ trợ vận hành, bảo mật, hỗ trợ kinh doanh và lớp tích hợp cung cấp dịch vụ đa cấp.

3. Thông tin hệ thống quản lý kế toán.

3.1 Thông tin kế toán dựa trên dữ liệu lớn.

- Dữ liệu lớn đề cập đến một lượng lớn các nguồn dữ liệu với các nguồn khác nhau, phức tạp và nhiều loại, tốc độ xử lý cao hơn và giá trị của chúng có thể được tái tạo nhiều lần, với các đặc tính là khối lượng lớn, đa dạng, thông lượng nhanh và mật độ dân số thấp. Dữ liệu lớn bao gồm nhiều loại dữ liệu. Hiện nay, thông tin kế toán chủ yếu được thu thập thông qua dữ liệu có cấu trúc. Đồng thời, dữ liệu không có tổ chức rất quan trọng đối với các quyết định tài chính.
- Do đó, trong thời đại dữ liệu lớn, doanh nghiệp phải thu thập dữ liệu phi cấu trúc, nâng cao số lượng thu thập dữ liệu phi cấu trúc và phân tích, lĩnh hội dữ liệu này để nâng cao tính đúng đắn của việc ra quyết định.
- Hơn nữa, Hệ thống quản lý thông tin kế toán (AIMS) là một chương trình phần mềm được thiết kế để phân tích dữ liệu kế toán. Hệ thống quản lý thông tin kế toán (AIMS) là một sự đảm bảo bắt buộc để quản lý và phán đoán doanh nghiệp thành công, đây là một diễn đàn thiết yếu để phân tích thị trường doanh nghiệp. Hệ thống thông tin tài chính điện tử thu thập, lưu và đánh giá dữ liệu tài chính để phân tích hoạt động và kết quả kinh doanh của doanh nghiệp nhằm tạo ra dữ liệu tài chính. Sau khi hệ thống thông tin tài chính hoàn thành công việc kế toán, hệ thống thông tin quản lý bắt đầu đánh giá các dữ liệu quan trọng, cung cấp kiến thức thực tế và đưa ra các phán đoán thông minh.
- Thông tin tài chính do AIMS tạo ra được truyền tới hệ thống thông tin quản lý và được phân tích để giúp quản lý và phán đoán doanh nghiệp. Cấu trúc điển hình của Hệ thống quản lý thông tin kế toán (AIMS) có thể được trình bày trong Hình 2.



- i. Nguồn dữ liệu: Đây là các giao dịch tài chính được nhập vào bên trong hệ thống cũng như từ các nguồn từ bên ngoài. Các số liệu thống kê này là kết quả của các cân nhắc tài chính với các doanh nghiệp thương mại khác và các cá nhân bên ngoài công ty. Các giao dịch như bán sản phẩm và dịch vụ, mua và thu tiền thanh toán được xem là ví dụ. Các giao dịch tài chính nội bộ là các nguồn lực bên trong một tổ chức bao gồm các cân nhắc giao dịch. Ví dụ, chuyển động nguyên liệu thô, khấu hao phí cố định, sản xuất cổ phiếu, và lao động.
- ii. Thu nhập dữ liệu: Bước hoạt động ban đầu của AIMS là thu nhập dữ liệu. Trước khi nhập thủ tục, dữ liệu phải được làm sạch các lỗi quan trọng. Bước này là quan trọng nhất theo nhiều cách vì có khả năng cung cấp những thông tin sai nếu những sai lầm đáng kể không được sửa chữa. Những sự cố như vậy có thể dẫn đến những hành động và phán đoán không đúng đắn của người dùng.

- Trong việc thu thập dữ liệu, có hai nguyên tắc: tính phù hợp và tính hiệu quả. Chỉ những dữ liệu liên quan mới được nhập vào hệ thống đang hình thành. Trách nhiệm chính của người vận hành hệ thống là xác định xem dữ liệu có được chấp nhận hay không? Anh ấy hoặc cô ấy có được điều này bằng cách đánh giá nhu cầu của người dùng. Kết quả là, chỉ những dữ liệu liên quan mới được bổ sung vào thông tin. Bộ lọc cho dữ liệu không chính xác được tạo trong giai đoạn thu thập dữ liệu. Dữ liệu chỉ được thu thập một lần cho một bước thu thập dữ liệu hiệu quả. Sau đó, dữ liệu này có thể được sử dụng bởi nhiều người dùng, dữ liệu này có thể được lưu và sao lưu sau này. Khả năng của một hệ thống thông tin về thu thập, phân tích và lưu trữ dữ liệu bị hạn chế. Quá tải thiết bị trong khi sao lưu dữ liệu sẽ làm giảm hiệu quả của hệ thống.
 - i. Xử lý dữ liệu: Dữ liệu thu được phải được xử lý để tạo ra một mặt hàng. Các ví dụ bao gồm các thuật toán toán học được sử dụng trong quá trình lập kế hoạch sản xuất, các công cụ phân tích để dự báo doanh số bán hàng và các yêu cầu nhập kế toán.
 - ii. Hệ thống quản lý cơ sở dữ liệu (DBMS): Cơ sở dữ liệu tổ chức là các vị trí vật lý để lưu trữ dữ liệu tài chính. nơi lưu trữ có thể là tủ hồ sơ hoặc đĩa máy tính. Ba vai trò chính của quản trị cơ sở dữ liệu là lưu trữ, truy xuất lại và xóa. Trong quy trình lưu trữ, các mục mới được nhập và vận chuyển đến một vị trí thích hợp trên cơ sở dữ liệu. Chức năng tìm nạp được sử dụng để trích xuất một bản ghi hiện có từ cơ sở dữ liệu. Sau khi dữ liệu được xử lý, chức năng lưu sẽ tải dữ liệu mới lên, trong khi chức năng di chuyển lại sẽ loại bỏ dữ liệu cũ và không cần thiết khỏi cơ sở dữ liệu.
 - iii. Sản xuất thông tin: Quá trình tổ chức, cấu trúc và cung cấp thông tin cho người dùng được gọi là quá trình tạo thông tin. Là dữ liệu có thể là tài liệu vận hành, bao gồm doanh số đặc hàng, phân tích lý thuật hoặc thông báo đầu ra của máy tính. Một thông tin hữu ích có 5 yếu tố: năng ứng dụng, thực tế, độ chính xác, tính nhất quán, và tóm tắt. So với mạng thần kinh thể hệ thứ hai truyền thống, thuật toán học sâu giải quyết vấn đề hiệu quả học tập thấp.

3.2 Điện toán đám mây trong kỷ nguyên AIMS.

- Các mô hình hệ thống kế toán thông thường không thể phân tích và xử lý dữ liệu tài chính một cách hiệu quả trong thời đại dữ liệu lớn. Với mục đích này, các giải pháp dịch vụ điện toán đám mây được yêu cầu:
 - + Thứ nhất, điện toán đám mây có công nghệ xử lý và phân tích dữ liệu phi thường, cho phép nó xử lý khối lượng thông tin lớn trong thời gian ngắn đồng thời phân tích và tiết lộ chính xác các tính năng cũng như mối quan hệ ngầm giữa dữ liệu và thông tin kinh tế.
 - + Thứ hai, công nghệ dịch vụ đám mây có thể cho phép cộng tác giữa các doanh nghiệp. có thể thực hiện phân tích và giải thích toàn diện dữ liệu tài chính doanh nghiệp mà không cần tăng chi phí hoạt động và quản lý cũng như chia sẻ tài nguyên giữa nhiều cơ quan trong doanh nghiệp.
- Hơn nữa, công nghệ điện toán đám mây có thể tạo điều kiện thuận lợi cho việc chia sẻ tài nguyên của công ty cũng như đánh giá và thảo luận kỹ lưỡng về thông tin tài chính kinh doanh mà không làm tăng chi phí vận hành và quản lý. Ảo hóa thông tin tài chính và dịch vụ đám mây là mục đích chính của công nghệ điện toán đám mây. Ảo hóa dữ liệu tài chính được coi là một trong những công nghệ chính để triển khai thông tin kế toán trên đám mây. Nó kết hợp cả hoạt động xử lý và phân tích dữ liệu tài chính kinh doanh để tăng khả năng củng cố và phát triển phân tích dữ liệu. Ảo hóa tất cả các loại dữ liệu được lưu trong đám mây sẽ tối ưu hóa thông tin kế toán tài chính và nguồn cung cấp dữ liệu đồng thời cải thiện hiệu quả sử dụng tài nguyên máy chủ. Hơn nữa, ảo hóa thông tin tài chính cho phép người dùng nhanh chóng truy cập thông tin tài chính dựa trên các lệnh khác nhau và khám phá các mối liên kết logic ẩn sau dữ liệu khổng lồ.
- Bên cạnh những điều trên, AIMS thông thường không đủ mạnh trong đó việc xử lý dữ liệu không hiệu quả, do đó, đòi hỏi sự hợp tác thủ công của nhân viên tài chính ở mức độ lớn làm giảm đáng kể hiệu quả xử lý thông tin. Trong những trường hợp như vậy, thông tin kế toán Hệ thống quản lý chỉ có thể cung cấp cho người quản lý phân tích dữ liệu trong quá khứ và thực tế. Thật không may, việc không giành được quyền kiểm soát theo thời gian thực đối với quá trình quản lý

kinh doanh, lập kế hoạch tài chính và ra quyết định tài chính thường khiến thời điểm lý tưởng để giải quyết rủi ro bị hoãn lại. AIMS được xây dựng trên các dịch vụ điện toán đám mây, thúc đẩy sự hợp tác và thay đổi dữ liệu giữa nhiều phòng ban của công ty đồng thời giải quyết vấn đề về tính khả dụng của kế toán tài chính. Mục đích cuối cùng của nghiên cứu mô hình hệ thống vi tính hóa điện toán đám mây là đối phó với các kết quả đầu ra khác nhau của AIMS và cung cấp cho người tiêu dùng dữ liệu tài chính một nền tảng chính xác và đầy đủ để ra quyết định

3.3 Mô hình năng lượng máy Boltzmann bị hạn chế dựa trên thuật toán học sâu.

- Mạng nơ-ron là một điển hình công cụ toán học được sử dụng trong lĩnh vực AI và máy tính học. So với thể hệ thứ hai truyền thống, thuật toán học sâu giải quyết vấn đề hiệu quả thấp. Nó coi mô hình nhiều lớp mô hình xếp chồng nhiều mô hình đơn vị RBM hoặc tương tự khác nhau và sau đó học lớp RBM để thực hiện việc đào tạo mô hình nhiều lớp do đó, mô hình RBM là nội dung chính của thuật toán học sâu.
- Máy Boltzmann bị hạn chế (RBM) là một mạng thần kinh nhân tạo có xác suất sáng tạo, học tốt phân phối xác suất trên một tập hợp các đầu vào. RBM thuộc mô hình xác suất đồ thị vô hướng và được thực hiện dựa trên năng lượng. Ở đây, phân bố xác suất chung được xác định bằng cách kết hợp vector lớp ẩn h và hàm năng lượng vector lớp đầu vào x , như được tính toán trong.
- $$p(x, h) = \frac{e^{-\text{năng lượng}(x, h)}}{Z}.$$
- e phương trình trên $Z = \sum_{x, h} e^{-\text{năng lượng}(x, h)}$ đại diện cho hằng số chuẩn hóa hoặc hàm phân vùng và phân bố xác suất biên của dữ liệu đầu vào có thể quan sát được x có thể được tính bằng cách sử dụng.
- $$p(x) = \sum_h p(x, h) = \sum_h \frac{e^{-\text{năng lượng}(x, h)}}{Z}.$$
- Sau khi giới thiệu năng lượng tự do, phương trình trên có thể được thay đổi thành phương trình (3) dưới đây:
- $$p(x) = \frac{e^{-\text{năng lượng miễn phí}}}{Z},$$

- $Z = \sum_x e^{-\text{năng lượng miễn phí } x(x)}$. trong phương trình trên có thể là tính theo.
- $\text{năng lượng miễn phí}(x) = -\log \sum_h e^{-\text{năng lượng}(x,h)}$.
- Ở đây, B được đưa vào biểu thức của các tham số mô hình và đạo hàm của phương trình (4) ở trên được tính để thu được như sau:

$$-\frac{\partial \log(x)}{\partial \theta} = \frac{\partial \text{năng lượng}}{\partial \theta} + \frac{1}{z} \sum_x e^{-\text{năng lượng miễn phí}(x)} \frac{\partial \text{năng lượng miễn phí}(x)}{\partial \theta} = \frac{\partial \text{năng lượng miễn phí}(x)}{\partial \theta} + \sum_{\hat{x}} p(\hat{x}) \frac{\partial \text{năng lượng miễn phí}(\hat{x})}{\partial \theta}$$
- Hiện tại, phương pháp đào tạo xấp xỉ $z \log p(x)/z\theta$ ap được sử dụng để xử lý tính toán khó khăn của hàm phân vùng RBM và định nghĩa về quy tắc cập nhật tham số mô hình dựa trên gradient năng lượng tự do của các mẫu tuân theo phân phối mô hình p và các mẫu chịu phân phối dữ liệu $(x-p(x))$.
- $E_{\hat{p}} \left[\frac{\partial \log p(x)}{\partial \theta} \right] = E_{\hat{p}} \left[\frac{\partial \text{năng lượng miễn phí}(x)}{\partial \theta} \right] + E_p \left[\frac{\partial \text{năng lượng miễn phí}(\hat{x})}{\partial \theta} \right]$
- Trong phương trình trên, P đại diện cho phân phối xác suất mô hình, trong khi P đại diện cho phân phối xác suất thực nghiệm của tập dữ liệu huấn luyện. Tương tự, EP và E_p đại diện cho giá trị kỳ vọng theo xác suất phân phối tương ứng. Thuật ngữ đầu tiên trong công thức thường thay thế việc đào tạo giá trị kỳ vọng mẫu. Thuật ngữ thứ hai, có các mẫu thu được từ mô hình P và thuật toán MCMC thường được sử dụng để lấy mẫu dữ liệu mô hình.

4. Thiết kế AIMS dựa trên Dữ Liệu lớn và Công nghệ đám mây

4.1 Cấu trúc chương trình của Quản lý thông tin kế toán.

- Là thiết kế một hệ thống quản lý hiệu quả và đáng tin cậy. Hệ thống quản lý thông tin kế toán sử dụng điện toán đám mây dữ liệu lớn. Bản chất của nó là hoạt động trên đám mây của một hệ thống quản lý thông tin. Thiết kế nền tảng đám mây điện tử và triển khai kiến trúc đa cấp được hoàn thiện bằng cách hợp nhất các nhu cầu cơ bản của hệ thống quản lý thông tin kế toán với các đặc điểm của các loại phần mềm thông tin khác nhau trên nền tảng đám mây, như được minh họa trong Hình 3.

Ở trên Hình minh họa hệ thống quản lý thông tin kế toán dựa trên công nghệ đám mây dữ liệu lớn. Hệ thống điện tử bao gồm bốn lớp cụ thể là: lớp dữ liệu, lớp cơ sở hạ tầng, lớp ứng dụng và lớp nền tảng quản lý. Người dùng đăng nhập vào chức năng hệ điều hành hệ thống tại ứng dụng khách điện thoại thông minh hoặc máy tính PC. Các loại

người dùng chủ yếu bao gồm quản lý doanh nghiệp, công nhân doanh nghiệp và quản trị viên hệ thống có quyền cao nhất. Lớp cơ sở hạ tầng bao gồm nhiều thiết bị, thành phần hỗ trợ cơ bản cho hoạt động của hệ thống quản lý thông tin kế toán. Một lượng lớn tài nguyên dữ liệu được lưu trữ trong lớp dữ liệu làm cơ sở tài nguyên theo yêu cầu xoắn của hệ thống quản lý thông tin kế toán. Lớp nền tảng quản lý là trung tâm bảo trì của hệ thống quản lý thông tin kế toán và lớp ứng dụng là nhà cung cấp dịch vụ.

4.2 Mô hình SAAS của hệ thống quản lý thông tin kế toán.

- Trong bài báo này, lớp kiên trì được giới thiệu để xử lý hiệu quả của việc trích xuất dữ liệu hệ thống và chức năng ánh xạ ORM bên trong được sử dụng để tăng tốc độ đọc dữ liệu. Như trong Hình 4, mô hình SaaS cần thiết cho thiết kế này được hiển thị. lớp điều khiển sử dụng lớp kiên trì để ánh xạ dữ liệu trong quá trình đọc dữ liệu và sau đó đọc thông tin dữ liệu sau khi đi qua lớp dữ liệu.
- Phần chính của sơ đồ là ORM lớp kiên trì, được xử lý và xử lý thêm, dựa trên kiến trúc Dao, và lớp cơ sở dữ liệu và lớp logic nghiệp vụ được chọn để xây dựng một trung tâm tương tác. Dữ liệu hình xuyên của anh ấy được gửi đến hệ thống quản lý thông tin kế toán để phân tích thông qua lớp lưu trữ lâu dài và quy trình xử lý được tối ưu hóa. không cần truyền hoặc sao chép dữ liệu để xử lý nhằm nâng cao hiệu quả tổng thể của tương tác dữ liệu [17]. Các mô-đun cốt lõi của lớp kiên trì sử dụng khái niệm ánh xạ để trích xuất và truy xuất thông tin cơ sở dữ liệu. Quá trình đọc và ghi dữ liệu khác với chức năng đọc và ghi dữ liệu của các câu lệnh SQL. Nó có thể nhanh chóng nâng cao dữ liệu và đơn giản hóa quy trình trích xuất dữ liệu, để tăng cường đáng kể tất cả các quy trình của chương trình.

4.3 Lưu trữ phân tán dựa trên nền tảng đám mây.

- Dựa trên nền tảng đám mây của hệ thống quản lý thông tin kế toán đã được thiết lập, một lượng lớn dữ liệu tài chính và tài liệu của doanh nghiệp được lưu trữ theo mô hình lưu trữ phân tán . Nói chung, kế toán trước đây trong hệ thống văn phòng hình thành chủ yếu tập trung vào việc lưu trữ tập trung dữ liệu tài chính trong máy chủ, không thể đạt được mục đích lưu trữ một lượng lớn dữ liệu. Nó chỉ có thể trích xuất một số dữ liệu cục bộ và cung cấp các hoạt động

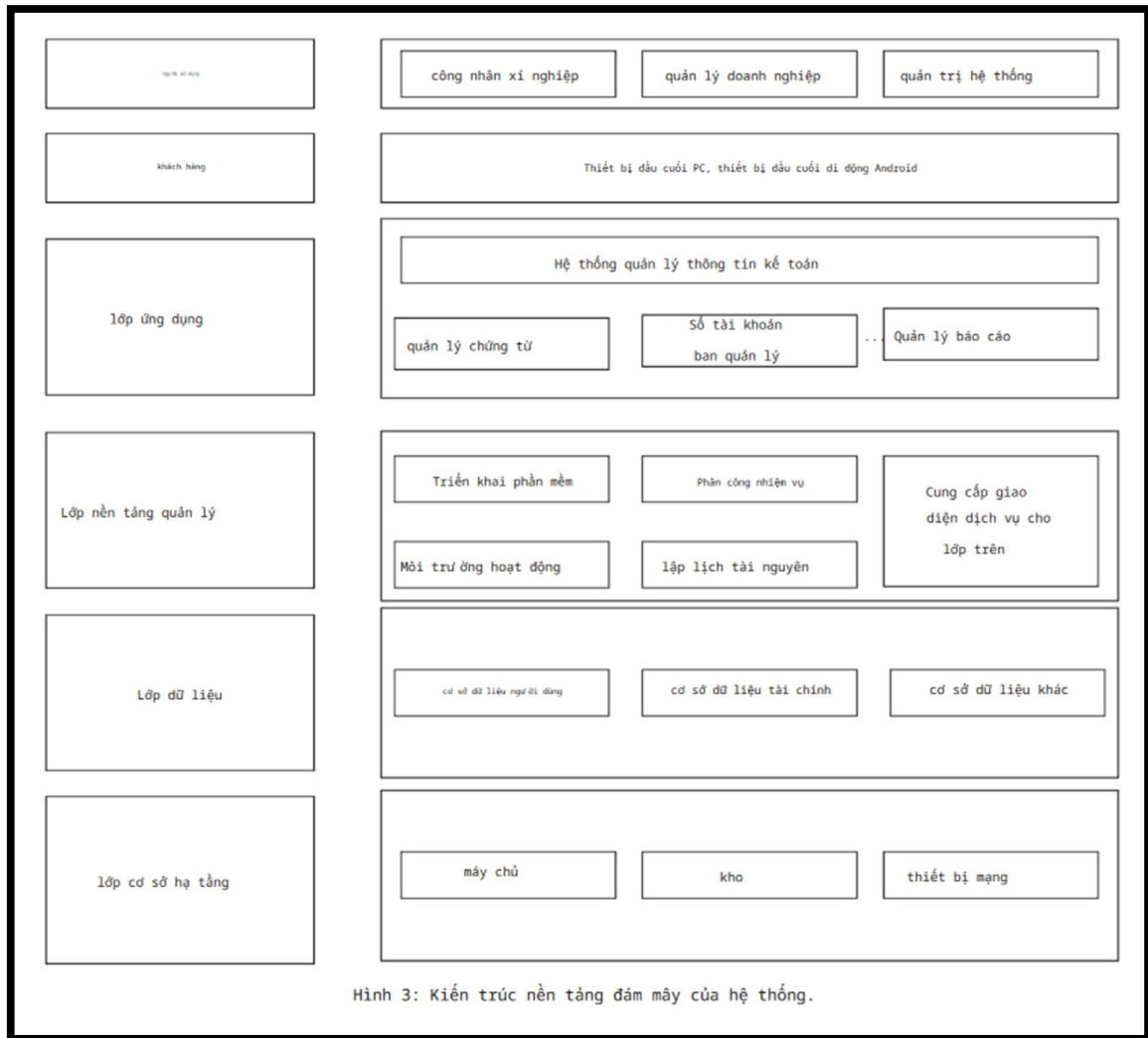
và dịch vụ kinh doanh cơ bản cho khách hàng địa phương, dẫn đến những hạn chế và giới hạn khu vực trong ứng dụng của hệ thống và sự hài lòng của người dùng công ty khi sử dụng hệ thống. Chiến lược lưu trữ phân tán được sử dụng trong nghiên cứu này có thể phá vỡ các giới hạn khu vực, tìm khu vực người dùng và bảo toàn dữ liệu tài chính từ mỗi công ty chi nhánh dựa trên các khu vực khác nhau. Do đó, hiệu quả lưu trữ dữ liệu của hệ thống đã được tăng lên rất nhiều, cũng như hiệu quả ứng dụng và tính bảo mật của hệ thống.

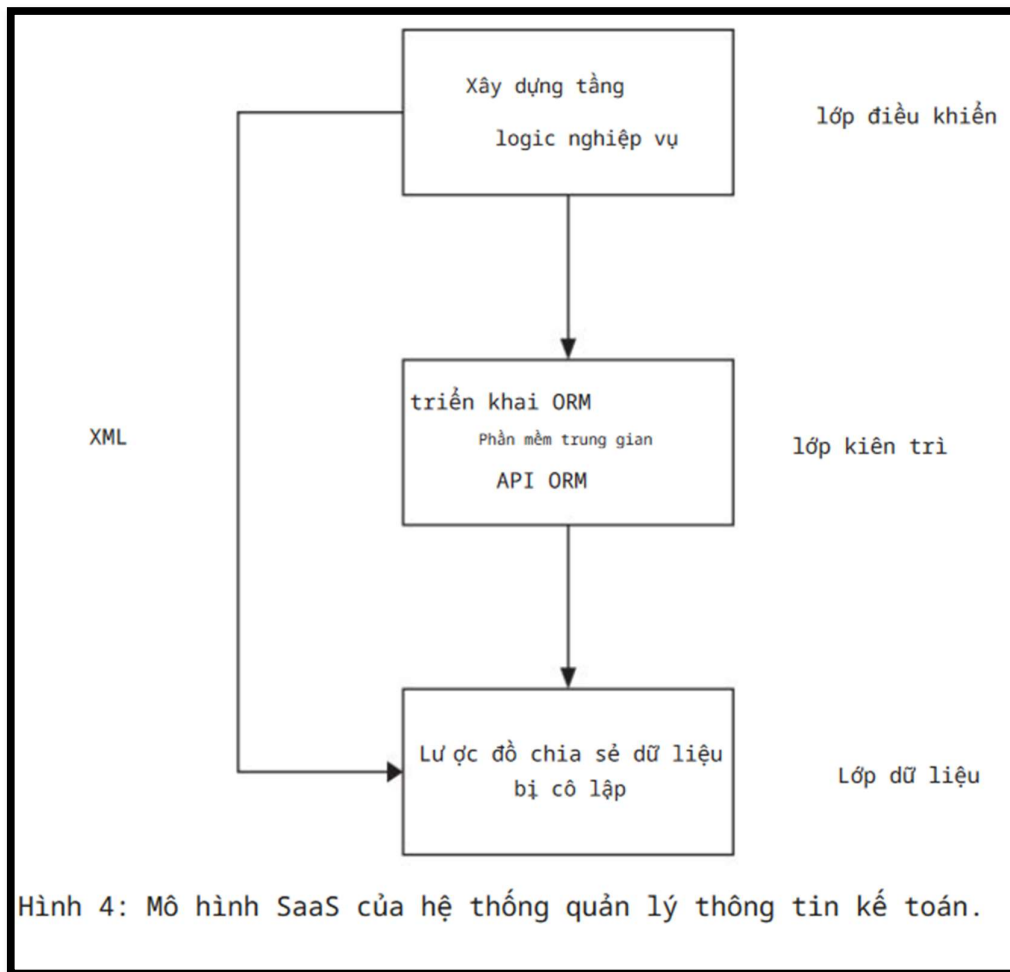
- Như thể hiện trong Hình 5, lược đồ lưu trữ phân tán dựa trên nền tảng đám mây. Hình trên cho thấy chế độ lưu trữ phân tán dựa trên nền tảng đám mây. Hệ thống điện tử có thể lưu dữ liệu kiểm đếm của các chi nhánh và phòng ban của các công ty tập đoàn lớn ở các vùng khác nhau của đất nước trong hệ thống thông tin kiểm đếm .
- Hình 1, 3 và 4 mô tả các hệ thống kế toán từ ba nhánh riêng biệt được truyền tới cơ sở dữ liệu đám mây thông qua nền tảng đám mây thông tin kế toán đã thiết lập. Dữ liệu điện tử trên nền tảng này cho phép các bộ phận tài chính của các chi nhánh khác nhau trao đổi tài nguyên dữ liệu của họ. Nhờ đó, các nhà quản lý doanh nghiệp và nhân viên có thể đăng nhập vào hệ thống và truy vấn dữ liệu tài chính của các chi nhánh khác nhau. Trong khi các bộ phận của tập đoàn nắm vững tình hình hoạt động của mình, nắm bắt đúng đắn dòng vốn chung của doanh nghiệp, ngăn chặn giới hạn kinh doanh của doanh nghiệp, rút ngắn chu kỳ vận hành hệ thống. Đồng thời, hệ thống thông tin kế toán của các chi nhánh được thiết lập ở các khu vực khác nhau sẽ lưu dữ liệu kế toán tương ứng vào cơ sở dữ liệu tài nguyên địa phương.
- Nơi nó sử dụng nền tảng đám mây để quản lý và phân tích nhất quán dữ liệu tài chính của từng khu vực nhằm tối đa hóa nguồn lực kinh doanh và đạt được sự chia sẻ tài nguyên và tài chính trong quá trình hình thành

5. Kiểm tra và mô phỏng hệ thống quản lý thông tin kế toán

- Là nghiên cứu đề xuất một hệ thống quản lý thông tin kế toán dựa trên công nghệ đám mây dữ liệu lớn để quản lý tài chính doanh nghiệp dựa trên thông tin hóa dữ liệu kế toán . Trong bài báo này, việc kiểm tra cụm hệ thống được thực

hiện bằng thực nghiệm. Nội dung kiểm tra bao gồm tính kịp thời của công việc, vị trí dữ liệu, cân bằng tải, v.v. Các hình thức công việc được áp dụng trong thử nghiệm được sửa đổi, yêu cầu và TeraSort. Công việc truy vấn điện tử được sử dụng để truy vấn thời gian hoàn trả hoặc chứng từ. Có rất nhiều hoạt động UO đã được yêu cầu trong giai đoạn hoạt động, có thể phát hiện chính xác.





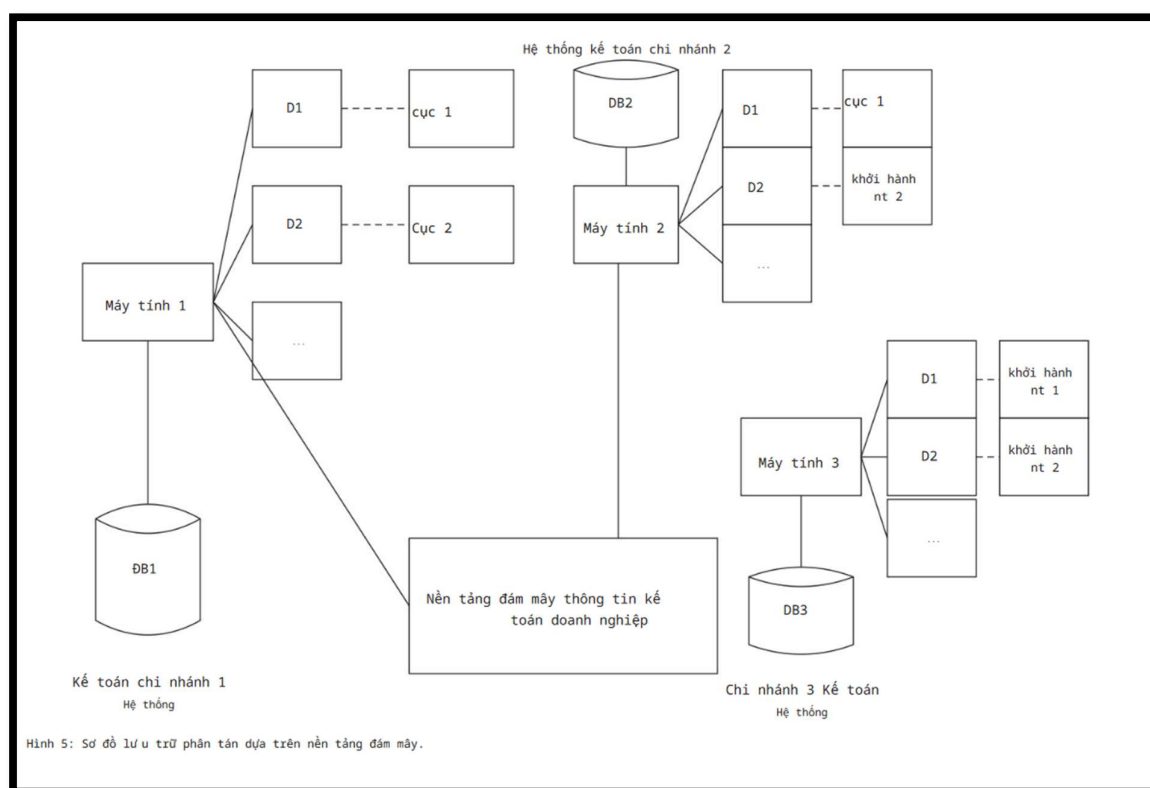
- Thông lượng nhiệm vụ nền tảng và trạng thái tải. TeraSort là công cụ chính để thực nghiệm nền tảng đám mây. Nó có thể nhập thông tin tài chính vào hệ thống quản lý thông tin kế toán mà không cần xử lý dữ liệu đầu ra.
- Công việc đã sửa đổi được sử dụng đại diện cho chứng từ cơ chế hoặc chứng từ thủ công. Sau khi thay đổi thông tin công việc, dữ liệu mới được ghi lại trong cơ sở dữ liệu HBase và thông tin trước đó sẽ bị xóa. Với tiền đề của cùng một dữ liệu đầu vào cho ba hình thức công việc khác nhau ở trên, trong khối lượng dữ liệu đầu ra, sửa đổi cao hơn TeraSort và TeraSort cao hơn yêu cầu. Chọn lịch trình nhiệm vụ tương ứng dựa trên thuật toán lập lịch trình hợp lý.

5.1 Thời gian đáp ứng công việc.

- Khi đánh giá thời gian đáp ứng công việc trong bài viết này, giả định rằng mỗi tác vụ có 10 lần giảm và 100 bản đồ. Tất cả các máy chạy các tác vụ cùng một

lúc. Đồng thời, có một số khác biệt nhất định về lượng dữ liệu trong quá trình thực hiện tác vụ, điều này có thể phản ánh đầy đủ tính công bằng của việc lập lịch trình. Thời gian phản hồi công việc của hệ thống được liệt kê trong Bảng 1.

- Theo dữ liệu trong bảng trước, giả sử rằng các tác vụ đầu vào hoàn toàn nhất quán, mối quan hệ phản hồi của các tác vụ khác nhau là sửa đổi vượt quá công tera và TeraSort vượt quá yêu cầu. Bởi vì các nhiệm vụ của teraport và reducer tương đối đơn giản, nên hiệu quả của các công việc đang chạy nhanh hơn hiệu quả của các công việc sửa đổi. số lượng tác vụ rút gọn và truy vấn ít nên chu kỳ yêu cầu ngắn. Một nghiên cứu dữ liệu chuyên sâu cho thấy rằng khi lượng dữ liệu ít, do đó hiệu quả của việc sử dụng nền tảng Hadoop trong hệ thống sẽ thấp hơn so với việc không sử dụng nền tảng Hadoop.



- Hiệu quả của nền tảng hệ thống này là rất đáng chú ý sau khi liên tục tăng lượng dữ liệu. Sau khi lượng dữ liệu đầu vào tăng lên 2560m, nhiệm vụ sửa đổi giảm đáng kể, từ 920 xuống 876 với mức tăng 4,78%. Nhiệm vụ teraport đã giảm từ 672 giây xuống 654 giây, tăng 2,68%, hoàn thành nhiệm vụ truy vấn điện tử đã tăng 6,8%, từ 247 giây lên 265 giây. Mục tiêu. Số lượng vị trí tác vụ

giảm trong cụm Hadoop ảnh hưởng trực tiếp đến kết quả tải cân bằng. Ở đây, tác vụ đầu vào được đặt thành 1280 M và có 8 khối lượng tác vụ bản đồ. Có thể nhận được các số vị trí nhiệm vụ khác nhau, ảnh hưởng đến thời gian phản hồi công việc của hệ thống thông tin kế toán. Trong bài báo này, ba chế độ nhiệm vụ khác nhau ở trên được sử dụng và có hai chế độ lập lịch trình. Dữ liệu thu được được thể hiện trong Bảng 2, trong đó OBDN thuộc thuật toán lập lịch cải tiến.

- Để cho phép các tác vụ chọn công việc một cách đồng đều và ngẫu nhiên, các công việc thuộc teraport nên có lượng dữ liệu đầu ra lớn. Ưu điểm của phương pháp này là nó có thể phát hiện tốt hơn trạng thái hoạt động thực tế của tất cả các nút. Dữ liệu đầu vào công việc 1280M, 8 tác vụ bản đồ và 8 tác vụ rút gọn được thiết lập ở đây được hiển thị trong Hình 6.
- Theo kết quả thời gian phản hồi của tác vụ giảm được hiển thị trong Hình 5 ở trên, việc thêm cân bằng tải có thể giữ cho chu kỳ xử lý của tất cả các tác vụ vị trí giảm ở trạng thái ổn định. Trong giai đoạn ra quyết định lập lịch trình của công bằng và OBDN, có sự khác biệt về độ lệch chuẩn và giá trị trung bình của chu kỳ hoạt động. Giá trị trung bình của giá trị hợp lý là 218 và độ lệch chuẩn là 30,79 tiêu chuẩn điện tử, độ lệch và giá trị trung bình của các quyết định lập kế hoạch công bằng lần lượt là 13,02 và 214.

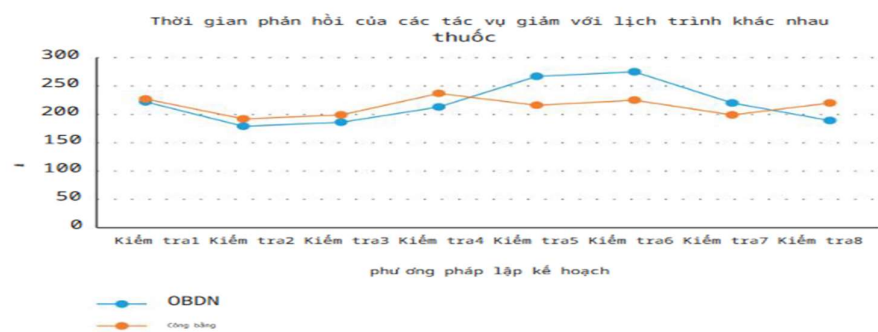
Bảng 1: Thời gian phản hồi của hệ điều hành trên các nền tảng có khối lượng dữ liệu khác nhau.

Có sử dụng khối lượng dữ liệu Hadoop hay không	320 M	640 M	1280 M	2560 M
sửa đổi N-Hadoop	11	224	139	453
sửa đổi Hadoop	236			441
N-Hadoop TeraSort	79		163	331
Hadoop TeraSắp xếp	86		171	330
Yêu cầu N-Hadoop	33		66	132
truy vấn Hadoop	31		61	120

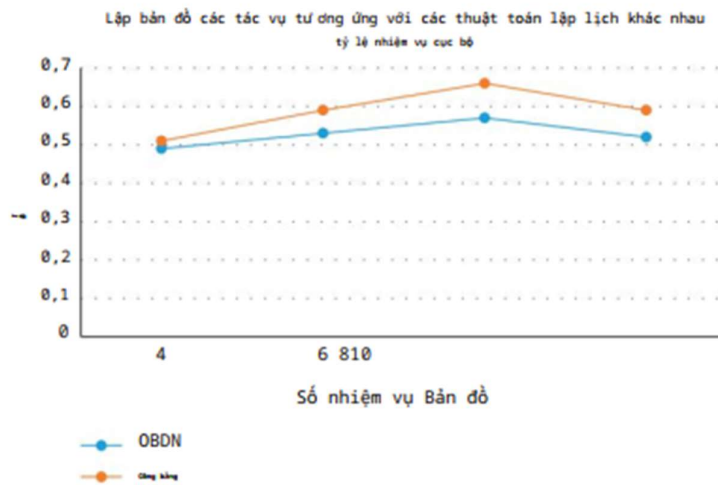
Bảng 2: Thời gian đáp ứng công việc của các thuật toán mức độ khác nhau.

Thuật toán lập lịch Số lượng khe giảm	2	4	8	16
Sửa đổi OBDN	762	593	450	393
Sửa đổi công bằng	801	612	414	381
TeraSort OBDN	630	448	327	267
Công bằng TeraSort	654	469	292	251
điều tra hội chợ	419	275	127	78
Yêu cầu OBDN	440	288	141	91

- Do đó, người ta kết luận rằng thời gian hoạt động trung bình của khe tác vụ đã giảm không thể tăng lên trong quá trình vận hành teraпорт, trong khi chu kỳ xử lý của tất cả các tác vụ tương đối cân bằng. Chu kỳ hoạt động của quá trình giảm được cân bằng và tốc độ di chuyển của nó tương đối đồng đều, có thể tối đa hóa việc sử dụng quá trình giảm.



Hình 6: Thời gian đáp ứng của các tác vụ giảm với các phương pháp lập lịch khác nhau.



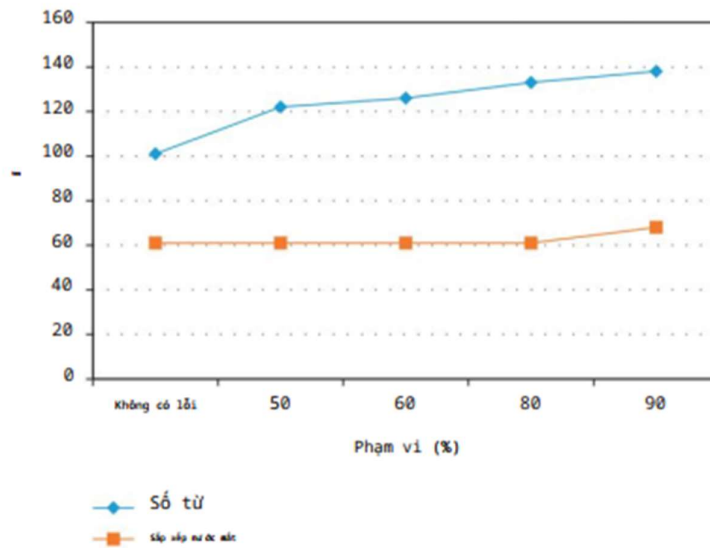
Hình 7: Lập bản đồ các tác vụ tương ứng với các tỷ lệ tác vụ cục bộ của thuật toán lập lịch trình khác nhau.

5.2 Phân tích vị trí nhiệm vụ hệ thống và tốc độ tăng tốc.

5.2.1. Phân tích vị trí nhiệm vụ

- Trong bài báo này, khi phát hiện các đặc tính cục bộ của dữ liệu cụm, kích thước của hàng dữ liệu cục bộ ảnh hưởng trực tiếp đến tốc độ hạch toán trong hệ thống quản lý hình thành. Trong bài báo này, teraPort, sửa đổi và truy vấn chạy cùng nhau và mỗi công việc được thiết lập để có 320 triệu dữ liệu đầu vào và 8 tác vụ được rút gọn. nhiều loại công việc khác nhau phải được gửi cùng một lúc bằng tập lệnh shell. Sau khi hoàn thành, tiếp tục chạy ba lần và sau đó đếm số lượng tất cả các tác vụ cục bộ. Phương trình (7) được sử dụng để tính tỷ lệ cục bộ của các nhiệm vụ.

$$\text{DataLocalityRate} = 1/n.$$



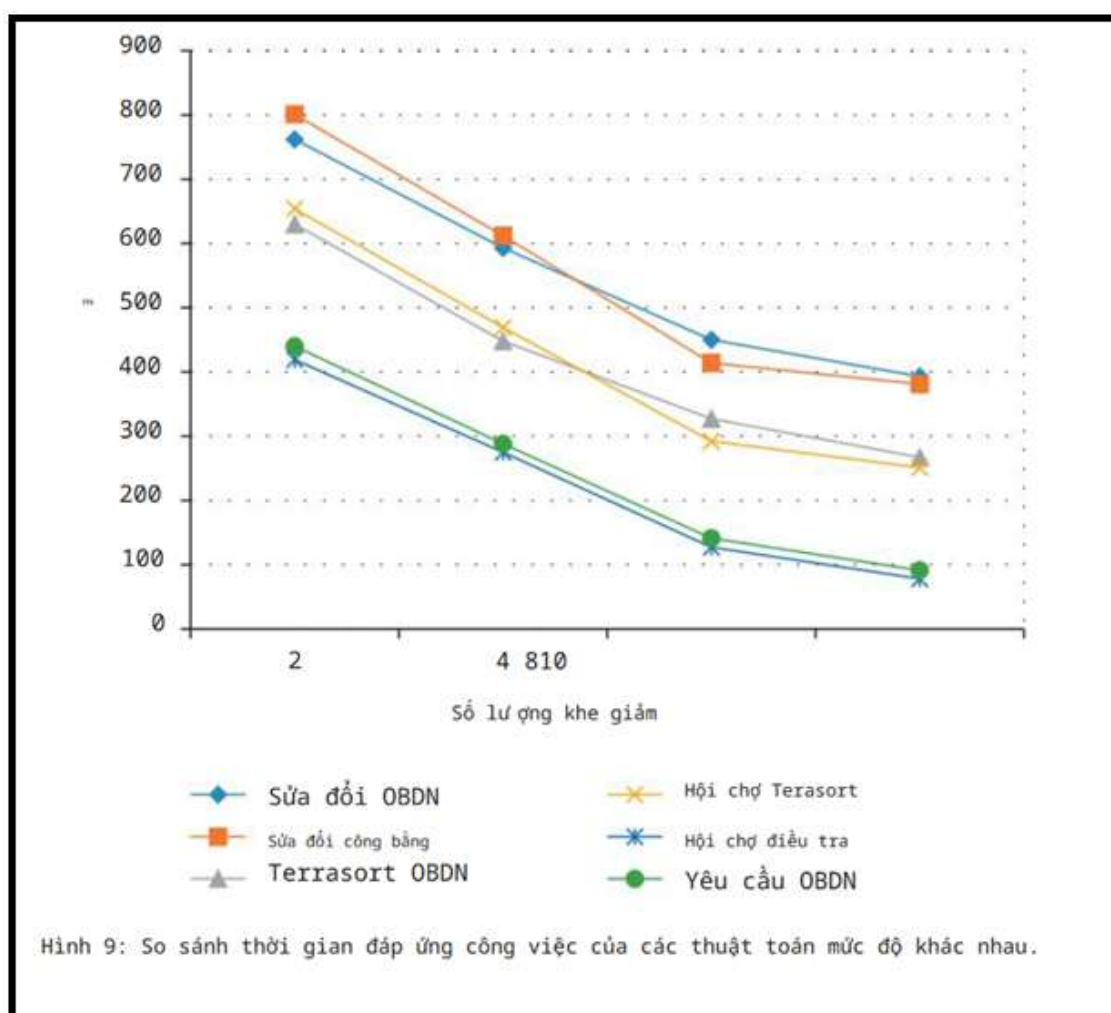
Hình 8: Mối quan hệ giữa thời gian thực thi và mức sử dụng CPU.

- Sau khi chạy chính sách lập lịch hợp lý ba lần, số tác vụ cục bộ là 20, 28, 40 và 49 và số tác vụ cục bộ của chính sách lập lịch OBDN là 19, 31, 46 và 52. Dữ liệu điện tử thu được được thể hiện trong Hình 7.
- Theo Hình trên, sau khi số lượng tác vụ bản đồ được điều chỉnh từ 4 thành 10, không có thay đổi đáng kể nào trong toàn bộ chiến lược lập lịch trình OBDN. Tốc độ cục bộ thấp hơn OBDN khi chính sách lập lịch hợp lý có bốn vị trí tác vụ. Sau khi số lượng nhiệm vụ giảm xuống, các vị trí tiếp tục tăng lên thì tỷ lệ cục bộ tăng lên một chút. Tuy nhiên, khi số lượng vị trí nhiệm vụ giảm tăng lên 10 thì tốc độ cục bộ sẽ giảm.

5.2.2 Phân tích tốc độ.

- Trong quá trình thử nghiệm này, Wordcount và TeraSort được sử dụng làm hai khối lượng công việc để phân tích tác động của quy mô tác vụ đối với tốc độ tăng tốc hệ thống. Đặt 1 g tác vụ đầu vào tại đây, bao gồm 6 tác vụ bản đồ và 12 tác vụ rút gọn. Mối quan hệ giữa thời gian thực thi và mức sử dụng CPU được thể hiện trong Hình 8. Chu kỳ thực thi tác vụ được kéo dài khi mức sử dụng CPU tăng lên, như có thể thấy bằng cách đánh giá dữ liệu trong Hình nói trên. Khối lượng công việc Wordcount và TeraSort nhạy cảm với việc sử dụng

CPU theo những cách khác nhau và Hadoop tăng tốc nhanh hơn.



- Hình 9 giải thích việc so sánh thời gian đáp ứng Công việc của các thuật toán mức độ khác nhau. Bằng cách phân tích con số này, nó có thể thấy rằng phương pháp lập lịch trình OBDN tốt hơn là hợp lý sau khi giảm số lượng vị trí đã giảm với tiền đề là chiến lược lập lịch trình công việc là hoàn toàn nhất quán. Sau khi liên tục tăng số lượng vị trí, tính năng của hội chợ có thể được phản ánh tốt hơn. Khi giảm số lượng 8 vị trí nhiệm vụ, thời gian nhiệm vụ được sửa đổi sẽ giảm xuống một mức nhất định, từ 450 giây xuống 414 giây và terasort từ 327 giây xuống 292 giây.

6. Kết Luận

- Ngày nay, sự mở rộng nhanh chóng của công nghệ thông tin và kiến thức mạng đã thúc đẩy sự phổ biến của tự động hóa, trí thông minh, thông tin hóa và các

công nghệ công nghệ cao khác trong cuộc sống của con người. Để cạnh tranh trên thị trường, các doanh nghiệp phải tăng cường các chiến lược hoạt động bên ngoài và thiết lập thông tin quản lý nội bộ. Quản lý thông tin hóa kiểm đếm đã trở thành cơ sở của tổ chức tài chính doanh nghiệp, có thể nâng cao hiệu quả khả năng thông tin hóa quản lý tài chính doanh nghiệp, giảm chi phí, đạt được công việc văn phòng hiệu quả và cải thiện độ chính xác của dữ liệu tài chính. Do đó, công việc này sử dụng công nghệ đám mây dữ liệu lớn để thiết kế và phát triển hệ thống quản lý thông tin kế toán và sử dụng hệ thống để hoàn thiện việc quản lý dữ liệu tài chính doanh nghiệp, để điều phối tốt hơn các nguồn lực và nhân sự của bộ phận tài chính. Lưu trữ dữ liệu phân tán của nền tảng đám mây được thông qua và hiệu suất của cụm hệ thống được đánh giá để xác định tác động của ứng dụng hệ thống, bằng cách phát triển kiến trúc nền tảng đám mây và mô hình SaaS của hệ thống quản lý thông tin kế toán dựa trên công nghệ đám mây. Bài viết này xem xét hiệu ứng ứng dụng của hệ thống từ ba khía cạnh: thời gian làm việc, hàng cục bộ dữ liệu và cân bằng tải. Theo kết quả, hệ thống có hiệu quả cao, tốc độ tăng tốc và tốc độ thực hiện tác vụ cao.

7. ỨNG DỤNG

- Ngày nay, tin học đã có những bước tiến nhanh chóng về ứng dụng của nó trong mọi lĩnh vực của cuộc sống trên phạm vi toàn thế giới nói chung và Việt Nam nói riêng. Tin học được người ta quan tâm và nhắc đến nhiều hơn bao giờ hết vì nó là một phần không thể thiếu trong cuộc sống văn minh, góp phần đẩy mạnh công cuộc công nghiệp hoá hiện đại hoá đất nước, tiến đến nền kinh tế tri thức. Máy vi tính cùng với những phần mềm là công cụ đắc lực giúp ta quản lý, tổ chức, sắp xếp và xử lý công việc một cách nhanh chóng và chính xác.
- Ở Việt Nam hiện nay, máy tính điện tử đặc biệt là máy vi tính trong nhiều năm qua đã được sử dụng rất rộng rãi. Sự phát triển của tin học, các công nghệ phần mềm, phần cứng, các tài liệu tham khảo đã đưa chúng ta từng bước tiếp cận với công nghệ thông tin trong mọi lĩnh vực nhằm đáp ứng nhu cầu của con người.
- Quản lý hệ thống thông tin kế toán là một đề tài không còn mới mẻ với các bài toán quản lý. Việc đưa tin học vào ứng dụng để quản lý là rất hữu ích, vì chúng

ta phải bỏ ra rất ít thời gian mà lại thu được hiệu quả cao, rất chính xác và tiện lợi nhanh chóng.

8.DEMO

Nội dung:

Những bước xây dựng phần mềm:

- Sử dụng công cụ Google Colab để chạy cơ sở dữ liệu.

Sau khi đã lên ý tưởng chúng em tiếp tục làm rõ các vấn đề liên quan đến ý tưởng đã đề ra.

- Xây dựng tools để chạy dự đoán chuỗi thời gian:

```
!pip install pycaret-ts-alpha

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting pycaret-ts-alpha
  Downloading pycaret-ts-alpha-1.0.0.dev1649817462-py3-none-any.whl (468 kB)
    |#####| 468 kB 4.9 MB/s
Collecting joblib==1.0.1
  Downloading joblib-1.0.1-py3-none-any.whl (303 kB)
    |#####| 303 kB 47.3 MB/s
Collecting category_encoders==2.4.0
  Downloading category_encoders-2.5.1.post0-py2.py3-none-any.whl (72 kB)
    |#####| 72 kB 846 kB/s
Requirement already satisfied: pandas<1.5.0,>=1.3.0 in /usr/local/lib/python3.7/dist-packages (from pycaret-ts-alpha) (1.3.5)
Requirement already satisfied: numpy==1.21 in /usr/local/lib/python3.7/dist-packages (from pycaret-ts-alpha) (1.21.6)
Requirement already satisfied: imbalanced-learn==0.8.1 in /usr/local/lib/python3.7/dist-packages (from pycaret-ts-alpha) (0.8.1)
Collecting scikit-plot==0.3.7
  Downloading scikit-plot-0.3.7-py3-none-any.whl (33 kB)
Collecting pmdarima==1.8.0
  Downloading pmdarima-2.0.2-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.manylinux_2_28_x86_64.whl (1.8 MB)
    |#####| 1.8 MB 47.9 MB/s
Collecting lightgbm==3.0.0
  Downloading lightgbm-3.3.3-py3-py37-cp37m-manylinux1_x86_64.whl (2.0 MB)
    |#####| 2.0 MB 38.8 MB/s
Collecting matplotlib==3.3.0
  Downloading matplotlib-3.3.3-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.whl (11.2 MB)
    |#####| 11.2 MB 37.6 MB/s
Requirement already satisfied: scipy==1.7.3 in /usr/local/lib/python3.7/dist-packages (from pycaret-ts-alpha) (1.7.3)
Requirement already satisfied: plotly==5.0.0 in /usr/local/lib/python3.7/dist-packages (from pycaret-ts-alpha) (5.0.0)
Requirement already satisfied: statsmodels==0.12.1 in /usr/local/lib/python3.7/dist-packages (from pycaret-ts-alpha) (0.12.2)
Collecting kaleido==0.2.1
  Downloading kaleido-0.2.1-py3-none-manylinux1_x86_64.whl (79.9 MB)
    |#####| 79.9 MB 143 kB/s
Requirement already satisfied: ipython==5.5.0 in /usr/local/lib/python3.7/dist-packages (from pycaret-ts-alpha) (7.9.0)
Collecting pyod==0.9.8
  Downloading pyod-1.0.6.tar.gz (141 kB)
    |#####| 141 kB 58.5 MB/s
```

```
!pip install prophet

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: prophet in /usr/local/lib/python3.7/dist-packages (1.1.1)
Requirement already satisfied: matplotlib==2.0.0 in /usr/local/lib/python3.7/dist-packages (from prophet) (3.5.3)
Requirement already satisfied: lunarcalendar==0.9 in /usr/local/lib/python3.7/dist-packages (from prophet) (0.6.0)
Requirement already satisfied: setuptools==42 in /usr/local/lib/python3.7/dist-packages (from prophet) (57.4.0)
Requirement already satisfied: tqdm==3.1 in /usr/local/lib/python3.7/dist-packages (from prophet) (4.64.1)
Requirement already satisfied: numpy==1.15.4 in /usr/local/lib/python3.7/dist-packages (from prophet) (1.21.6)
Requirement already satisfied: python-dateutil==2.8.0 in /usr/local/lib/python3.7/dist-packages (from prophet) (2.8.2)
Requirement already satisfied: wheel==0.37.6 in /usr/local/lib/python3.7/dist-packages (from prophet) (0.38.4)
Requirement already satisfied: convertdate==2.1.2 in /usr/local/lib/python3.7/dist-packages (from prophet) (2.4.0)
Requirement already satisfied: cmtstamps==1.0.4 in /usr/local/lib/python3.7/dist-packages (from prophet) (1.0.8)
Requirement already satisfied: pandas==1.0.4 in /usr/local/lib/python3.7/dist-packages (from prophet) (1.3.5)
Requirement already satisfied: setuptools-git==1.2 in /usr/local/lib/python3.7/dist-packages (from prophet) (1.2)
Requirement already satisfied: holidays==0.14.1 in /usr/local/lib/python3.7/dist-packages (from prophet) (0.17)
Requirement already satisfied: pytz==2019.1 in /usr/local/lib/python3.7/dist-packages (from prophet) (2022.5)
Requirement already satisfied: hijri-converter in /usr/local/lib/python3.7/dist-packages (from prophet) (2.2.4)
Requirement already satisfied: korean-lunar-calendar in /usr/local/lib/python3.7/dist-packages (from prophet) (0.3.1)
Requirement already satisfied: pytz in /usr/local/lib/python3.7/dist-packages (from lunarcalendar==0.9) (2022.5)
Requirement already satisfied: cython==0.29.21 in /usr/local/lib/python3.7/dist-packages (from lunarcalendar==0.9) (0.29.21)
Requirement already satisfied: pyparsing==2.2.1 in /usr/local/lib/python3.7/dist-packages (from matplotlib==2.0.0) (3.0.9)
Requirement already satisfied: packaging==20.9 in /usr/local/lib/python3.7/dist-packages (from matplotlib==2.0.0) (21.3)
Requirement already satisfied: cycler==0.10 in /usr/local/lib/python3.7/dist-packages (from matplotlib==2.0.0) (0.11.0)
Requirement already satisfied: pillow==6.2.0 in /usr/local/lib/python3.7/dist-packages (from matplotlib==2.0.0) (7.1.2)
Requirement already satisfied: fonttools==4.22.0 in /usr/local/lib/python3.7/dist-packages (from matplotlib==2.0.0) (4.38.0)
Requirement already satisfied: kiwisolver==1.0.1 in /usr/local/lib/python3.7/dist-packages (from matplotlib==2.0.0) (1.4.4)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.7/dist-packages (from kiwisolver==1.0.1) (4.1.1)
Requirement already satisfied: six==1.5 in /usr/local/lib/python3.7/dist-packages (from python-dateutil==2.8.0) (1.15.0)

from pycaret.datasets import get_data

from pycaret.datasets import get_data
```

```
[ ] dt = get_data('pycaret_downloads')
```

INFO:logs:Preparing display monitor

	Date	Total
0	7/8/2021	3773
1	7/9/2021	3172
2	7/10/2021	2079
3	7/11/2021	2725
4	7/12/2021	4542

```
dt.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ---
 0   Date    90 non-null      object
 1   Total   90 non-null      int64
dtypes: int64(1), object(1)
memory usage: 1.5+ KB
```

```
[ ] import pandas as pd
```

```
[ ] dt['Date'] = pd.to_datetime(dt['Date'])
```

```
[ ] dt.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ---
 0   Date    90 non-null      datetime64[ns]
 1   Total   90 non-null      int64
dtypes: datetime64[ns](1), int64(1)
memory usage: 1.5 KB
```

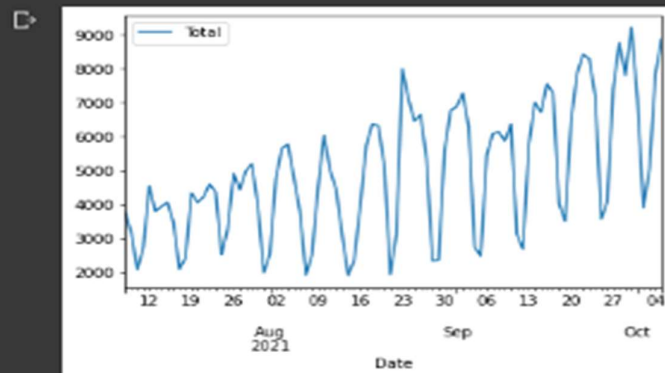
```
[ ] dt.set_index('Date',drop=True, inplace=True)
```

dt

Date	Total
2021-07-08	3773
2021-07-09	3172
2021-07-10	2079
2021-07-11	2725
2021-07-12	4542
...	...
2021-10-01	7115
2021-10-02	3903
2021-10-03	5038
2021-10-04	7843
2021-10-05	8872

90 rows x 1 columns

dt.plot();



```
[ ] from pycaret.time_series import *

ts=setup(dt,rb=15,fold_strategy='sliding')

INFO:logs:PyCaret TSForecastingExperiment
INFO:logs:logging name: ts-default-name
INFO:logs:ML Usecase: Multicase.TIME_SERIES
INFO:logs:version: 3.0.0
INFO:logs:initializing setup()
INFO:logs:self.OSI: 6744
INFO:logs:self.variable_keys: ('seed', 'seasonality_present', 'experiment_', 'OSI', 'exp_id', 'idx', 'all_models_internal', 'y_transformed', 'all_y_values', 'index_type', 'y', 'X_train', 'y_train', 'exp_name_log', 'X_test', 'y_test_transformed')
INFO:logs:checking environment
INFO:logs:python_version: 3.7.15
INFO:logs:python_build: ('default', 'Oct 12 2022 19:14:55')
INFO:logs:machine: x86_64
INFO:logs:platform: Linux-5.10.133-x86_64-with-Ubuntu-18.04-bionic
INFO:logs:Memory: svmem(total=13616353280, available=11897831424, percent=12.6, used=1525379872, free=8289635088, active=795693856, inactive=4258488128, buffers=416743424, cached=3464685696, shared=1208320, slab=266915840)
INFO:logs:Physical Core: 1
INFO:logs:Logical Core: 2
INFO:logs:checking libraries
INFO:logs:pd==1.3.5
INFO:logs:numpy==1.21.6
INFO:logs:sklearn==1.0.2
INFO:logs:lightgbm==3.3.3
INFO:logs:xgboost==0.90
WARNING:logs:catboost not found
WARNING:logs:mlflow not found
INFO:logs:Set Forecast Horizon.
INFO:logs:Set up Train-Test Splits.
INFO:logs:Preparing preprocessing pipeline...
INFO:logs:finished creating preprocessing pipeline.
INFO:logs:Pipeline: PyCaretForecastingPipeline(steps=[('forecaster',
                                                    TransformedTargetForecaster(steps=[('model',
                                                                    DummyForecaster()))]))

INFO:logs:Set up Seasonal Period.
INFO:logs:Set up whether Multiplicative components allowed.
INFO:logs:creating final display dataframe.
INFO:logs:Setup Display Container:
INFO:logs:

```

	session_id	Description	Value
0		4461	
1	Target	Total	
2	Approach	Univariate	
3	Exogenous Variables	Not Present	

INFO:logs:Creating final display dataframe.

INFO:logs:Setup Display Container:

	Description	Value
0	session_id	4461
1	Target	Total
2	Approach	Univariate
3	Exogenous Variables	Not Present
4	Data shape	(90, 1)
5	Train data shape	(75, 1)
6	Test data shape	(15, 1)
7	Fold Generator	SlidingWindowSplitter
8	Fold Number	3
9	Enforce Prediction Interval	False
10	Seasonal Period(s) Tested	7
11	Seasonality Present	True
12	Seasonalities Detected	[7]
13	Primary Seasonality	7
14	Target Strictly Positive	True
15	Target White Noise	No
16	Recommended d	1
17	Recommended Seasonal D	1
18	Missing Values	0
19	Preprocess	True
20	CPU Jobs	-1
21	Use GPU	False
22	Log Experiment	False
23	Experiment Name	ts-default-name
24	USI	67a4

INFO:logs:Preparing display monitor

	Description	Value
0	session_id	4461
1	Target	Total
2	Approach	Univariate
3	Exogenous Variables	Not Present
4	Data shape	(90, 1)

0	session_id	4461
1	Target	Total
2	Approach	Univariate
3	Exogenous Variables	Not Present
4	Data shape	(90, 1)
5	Train data shape	(75, 1)
6	Test data shape	(15, 1)
7	Fold Generator	SlidingWindowSplitter
8	Fold Number	3
9	Enforce Prediction Interval	False
10	Seasonal Period(s) Tested	7
11	Seasonality Present	True
12	Seasonalities Detected	[7]
13	Primary Seasonality	7
14	Target Strictly Positive	True
15	Target White Noise	No
16	Recommended d	1
17	Recommended Seasonal D	1
18	Missing Values	0
19	Preprocess	True
20	CPU Jobs	-1
21	Use GPU	False
22	Log Experiment	False

best_model=compare_models(n_select=3)

	Model	MAE	RMSE	MAPE	SMAPE	MASE	RMSSE	R2	TT (Sec)
arima	ARIMA	640.1899	879.3945	0.128	0.1305	1.0617	1.1031	0.734	0.0933
naive	Seasonal Naive Forecaster	652.0607	903.8883	0.1266	0.138	1.0762	1.1391	0.719	0.0267
auto_arima	Auto ARIMA	661.0579	1069.5014	0.1871	0.1831	1.405	1.3158	0.5965	2.1133
exp_smooth	Exponential Smoothing	940.4855	1082.2059	0.1915	0.1895	1.5344	1.3313	0.5867	0.0900
theta	Theta Forecaster	999.342	1202.9552	0.1895	0.2038	1.6595	1.5323	0.5065	0.0833
ets	ETS	930.877	1132.3046	0.2079	0.2083	1.5469	1.4264	0.5533	0.1233
dt_eds_dt	Decision Tree w/ Cond. Deseasonalize & Detrending	976.4834	1185.9875	0.2224	0.2088	1.8759	1.5816	0.5123	0.0567
lightgbm_eds_dt	Light Gradient Boosting w/ Cond. Deseasonalize...	1005.7502	1166.5013	0.2244	0.2107	1.9649	1.4752	0.531	0.0597
et_eds_dt	Extra Trees w/ Cond. Deseasonalize & Detrending	1016.6925	1186.2693	0.2264	0.2137	1.8963	1.519	0.5188	1.1700
omp_eds_dt	Orthogonal Matching Pursuit w/ Cond. Deseasona...	1026.1727	1201.388	0.2292	0.2144	1.6955	1.5159	0.5014	0.0500
knn_eds_dt	K Neighbors w/ Cond. Deseasonalize & Detrending	1030.8562	1222.8995	0.2341	0.2182	1.7064	1.5321	0.4808	1.0800
rf_eds_dt	Random Forest w/ Cond. Deseasonalize & Detrending	1027.3547	1180.3252	0.2317	0.219	1.7053	1.5016	0.5222	1.2367
ada_eds_dt	AdaBoost w/ Cond. Deseasonalize & Detrending	1051.9302	1218.3614	0.2346	0.2221	1.7212	1.5289	0.485	0.1500
gbr_eds_dt	Gradient Boosting w/ Cond. Deseasonalize & Det...	1078.87	1281.8604	0.2438	0.2270	1.8043	1.6581	0.4311	0.0767
br_eds_dt	Bayesian Ridge w/ Cond. Deseasonalize & Detren...	1101.1981	1285.3422	0.2473	0.2323	1.8388	1.8461	0.4365	0.0500
llar_eds_dt	Lasso Least Angular Regressor w/ Cond. Deseaso...	1377.8021	1586.5085	0.3251	0.3168	2.3236	2.0554	0.1438	0.0500
lr_eds_dt	Linear w/ Cond. Deseasonalize & Detrending	1434.7855	1657.7821	0.341	0.3351	2.4286	2.1615	0.0641	0.0467
ridge_eds_dt	Ridge w/ Cond. Deseasonalize & Detrending	1434.7851	1657.7815	0.341	0.3351	2.4286	2.1615	0.0641	0.0500
lasso_eds_dt	Lasso w/ Cond. Deseasonalize & Detrending	1434.7327	1657.7389	0.3409	0.3351	2.4286	2.1614	0.0641	0.0600
en_eds_dt	Elastic Net w/ Cond. Deseasonalize & Detrending	1434.7439	1657.7535	0.3409	0.3351	2.4286	2.1614	0.0641	0.0467
lar_eds_dt	Least Angular Regressor w/ Cond. Deseasonalize...	1434.7855	1657.7821	0.341	0.3351	2.4286	2.1615	0.0641	0.0500
polytrend	Polynomial Trend Forecaster	1556.9983	1849.4448	0.4369	0.3423	2.6526	2.3977	-0.1786	0.0267
croston	Croston	1721.7893	1857.5946	0.4099	0.3769	2.8502	2.3538	-0.1919	0.0233
grand_means	Grand Means Forecaster	1760.0607	1895.6314	0.4004	0.3861	2.9115	2.3952	-0.2424	0.0267

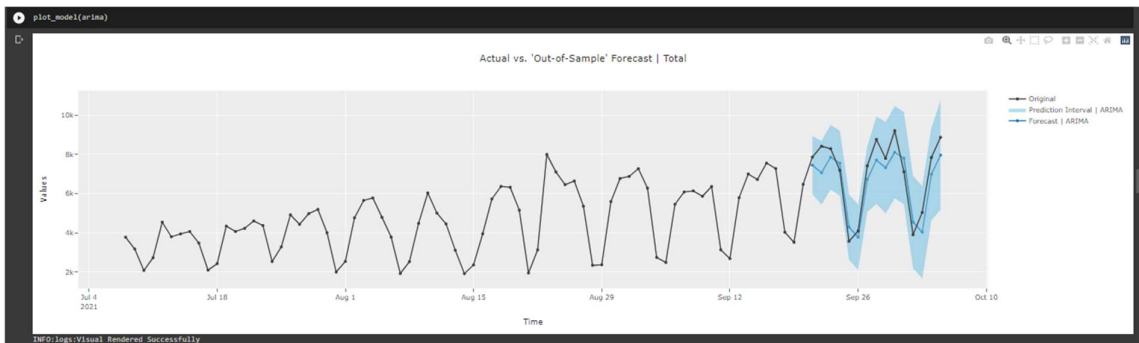
grand_means	Grand Means Forecaster	1760.0607	1895.6314	0.4004	0.3861	2.9115	2.3952	-0.2424	0.0267
huber_eds_dt	Huber w/ Cond. Deseasonalize & Detrending	1545.7512	1770.183	0.37	0.3871	2.6181	2.2722	-0.1354	0.0767
par_eds_dt	Passive Aggressive w/ Cond. Deseasonalize & De...	1901.9643	2327.1878	0.4213	0.4146	3.1291	2.8857	-1.1459	0.0600
naive	Naive Forecaster	2853.2444	3001.3012	0.5002	0.643	4.186	3.5173	-2.4115	0.0367

INFO:logs:master_model_container: 54
INFO:logs:display_container: 3
INFO:logs:[ARIMA(seasonal_order=(0, 1, 0, 7)), NaiveForecaster(sp=7), AutoARIMA(random_state=4461, sp=7, suppress_warnings=True)]
INFO:logs:compare_models() successfully completed.....

```
[ ] arima=create_model('arima')
```

	cutoff	MAE	RMSE	MAPE	SMAPE	MASE	RMSSE	R2
0	2021-08-06	545.3114	656.5094	0.1414	0.1344	1.1289	1.1749	0.8356
1	2021-08-21	788.7883	1234.1619	0.1411	0.1617	1.3651	1.5467	0.6181
2	2021-09-05	586.4699	747.5122	0.1017	0.0955	0.6912	0.5877	0.7483
Mean	NaT	640.1899	879.3945	0.1280	0.1305	1.0617	1.1031	0.7340
SD	NaT	106.4100	253.5946	0.0186	0.0272	0.2792	0.3048	0.0894

INFO:logs:master_model_container: 56
INFO:logs:display_container: 5
INFO:logs:ARIMA(seasonal_order=(0, 1, 0, 7))
INFO:logs:create_model() successfully completed.....




```

1 arima_tuned=tune_model(arima)

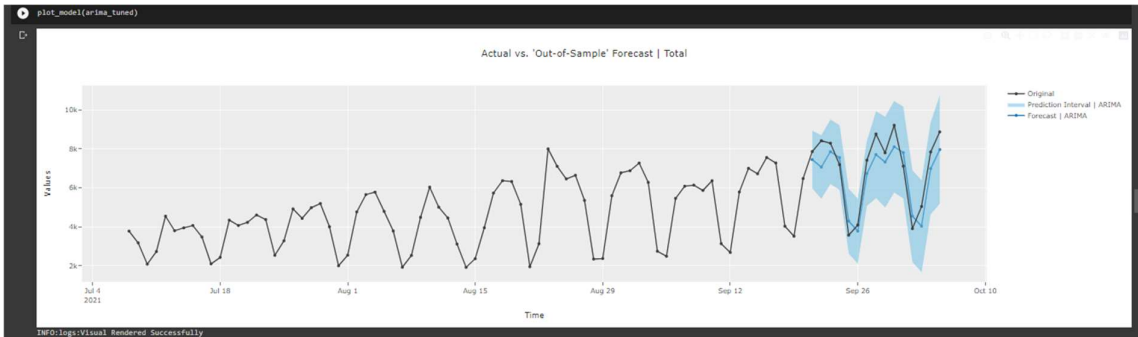
```

	cutoff	MAE	RMSE	MAPE	SMAPE	MASE	RMSSE	R2
0	2021-08-06	502.1217	673.8121	0.1075	0.1092	1.0395	1.2055	0.8269
1	2021-08-21	984.7737	1348.0216	0.1730	0.1996	1.7043	1.8894	0.5444
2	2021-09-05	746.8342	878.3243	0.1518	0.1572	0.8800	0.9005	0.8525
Mean	NaT	744.5099	988.8527	0.1441	0.1554	1.2079	1.1951	0.8748
SD	NaT	197.0476	282.3219	0.0273	0.0389	0.3570	0.4079	0.1164

```

INFO:logs:master_model_container: 59
INFO:logs:display_container: 7
INFO:logs:ARIMA(seasonal_order=(0, 1, 0, 7))
INFO:logs:tune_model() succesfully completed.....

```



```

1 prophet=create_model('prophet')

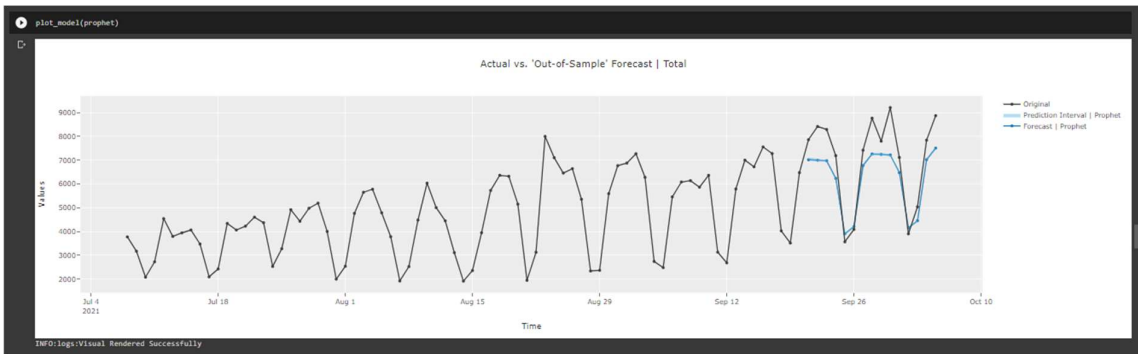
```

	cutoff	MAE	RMSE	MAPE	SMAPE	MASE	RMSSE	R2
0	2021-08-06	921.5944	1001.4911	0.3031	0.2521	1.9079	1.7622	0.8174
1	2021-08-21	1157.6705	1412.1821	0.1999	0.2232	2.0035	1.7698	0.5000
2	2021-09-05	854.9994	952.9632	0.1853	0.1407	1.0073	0.7492	0.5909
Mean	NaT	977.9881	1122.2054	0.2228	0.2083	1.8390	1.4371	0.5994
SD	NaT	129.9506	205.9853	0.0655	0.0431	0.4488	0.4885	0.0503

```

INFO:logs:master_model_container: 60
INFO:logs:display_container: 9
INFO:logs:ProphetPeriodPatched()
INFO:logs:create_model() succesfully completed.....

```



```

1 prophet_tuned=tune_model(prophet)

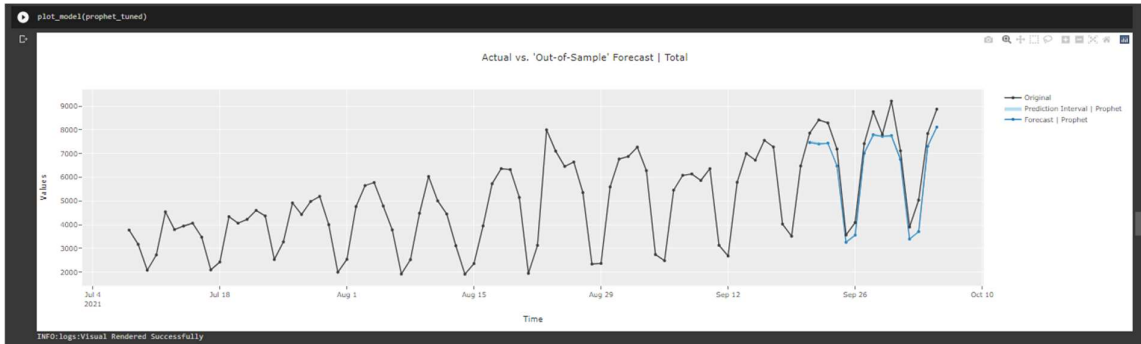
```

	cutoff	MAE	RMSE	MAPE	SMAPE	MASE	RMSSE	R2
0	2021-08-06	806.7449	909.4816	0.2533	0.2175	1.6701	1.6276	0.8845
1	2021-08-21	611.4835	898.0956	0.1385	0.1283	1.0582	1.1255	0.7978
2	2021-09-05	1151.5491	1339.9993	0.2019	0.1834	1.3572	1.0535	0.1911
Mean	NaT	856.5925	1049.1922	0.1979	0.1764	1.3819	1.2689	0.5578
SD	NaT	223.2805	205.6842	0.0470	0.0368	0.2498	0.2553	0.2634

```

INFO:logs:master_model_container: 66
INFO:logs:display_container: 12
INFO:logs:ProphetPeriodPatched(changepoint_prior_scale=0.12059742403082866,
                                holidays_prior_scale=0.12546259597149995,
                                seasonality_mode='multiplicative',
                                seasonality_prior_scale=2.1841289834490496)
INFO:logs:tune_model() succesfully completed.....

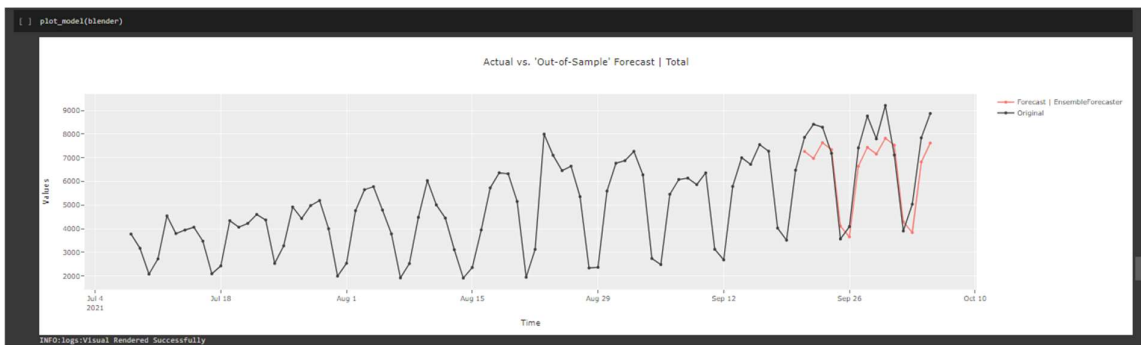
```



```
[ ] blender-blend_models(best_model)

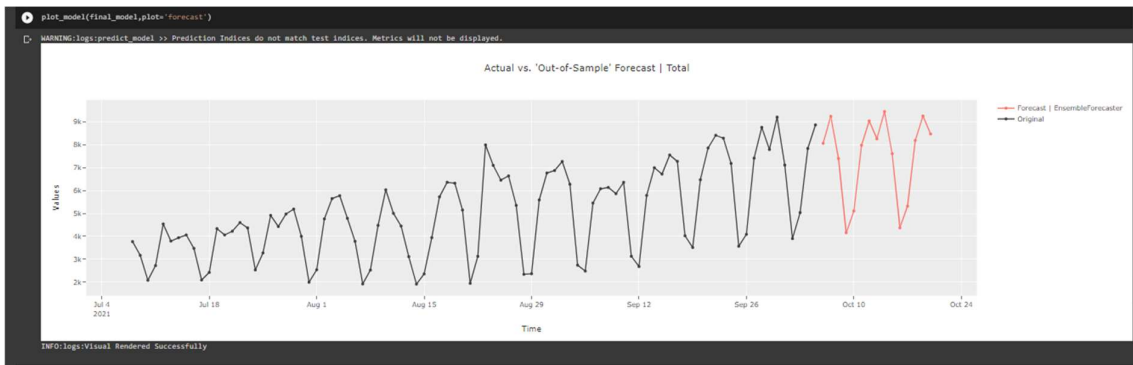
cutoff    MAE    RMSE    MAPE    SMAPE    MASE    RMSSE    R2
0  2021-08-06  581.0211  671.7482  0.1600  0.1475  1.2020  1.2021  0.8278
1  2021-08-21  803.3827  1292.9252  0.1546  0.1793  1.4942  1.6204  0.5809
2  2021-09-05  577.3002  723.7809  0.1020  0.0970  0.8804  0.5690  0.7640
Mean    NaT  673.9014  896.1448  0.1375  0.1412  1.1258  1.1305  0.7243
SD      NaT  133.9922  261.9685  0.0251  0.0339  0.3367  0.4322  0.1047

INFO:logs:master_model_container: 67
INFO:logs:display_container: 14
INFO:logs:_ensembleforecasterWithVoting(forecasters=[('ARIMA',
    ARIMA(seasonal_order=(0, 1, 0, 7))),
    ('Seasonal Naive Forecaster',
    NaiveForecaster(sp=7)),
    ('Auto ARIMA',
    AutoARIMA(random_state=4461, sp=7,
    suppress_warnings=True))),
    n_jobs=-1)
INFO:logs:blend_models() successfully completed.....
```



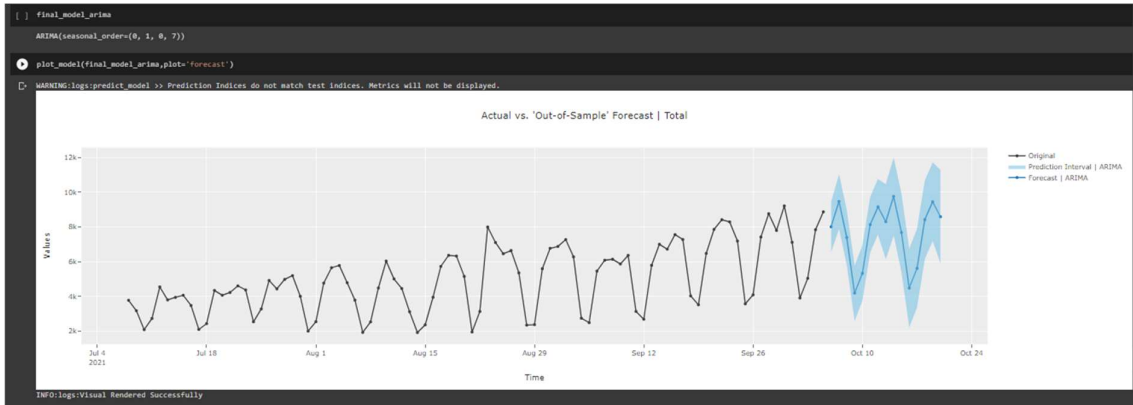
```
final_model=finalize_model(blender)

INFO:logs:Initializing finalize_model()
INFO:logs:finalize_model(self=pycaret.time_series_forecasting.exp.TSForecastingExperiment object at 0x77b648efb0b, estimator=_ensembleforecasterWithVoting(forecasters=[('ARIMA',
    ARIMA(seasonal_order=(0, 1, 0, 7))),
    ('Seasonal Naive Forecaster',
    NaiveForecaster(sp=7)),
    ('Auto ARIMA',
    AutoARIMA(random_state=4461, sp=7,
    suppress_warnings=True))),
    n_jobs=-1), fit_kwargs=None, groups=None, model_only=True, experiment_custom_tags=None, display=None)
INFO:logs:Finalizing _ensembleforecasterWithVoting(forecasters=[('ARIMA',
    ARIMA(seasonal_order=(0, 1, 0, 7))),
    ('Seasonal Naive Forecaster',
    NaiveForecaster(sp=7)),
    ('Auto ARIMA',
    AutoARIMA(random_state=4461, sp=7,
    suppress_warnings=True))),
    n_jobs=-1)
INFO:logs:Initializing create_model()
INFO:logs:create_model(self=pycaret.time_series_forecasting.exp.TSForecastingExperiment object at 0x77b648efb0b, estimator=_ensembleforecasterWithVoting(forecasters=[('ARIMA',
    ARIMA(seasonal_order=(0, 1, 0, 7))),
    ('Seasonal Naive Forecaster',
    NaiveForecaster(sp=7)),
    ('Auto ARIMA',
    AutoARIMA(random_state=4461, sp=7,
    suppress_warnings=True))),
    n_jobs=-1), fold=None, round=4, cross_validation=True, predict=True, fit_kwargs={}, groups=None, refit=True, probability_threshold=None, experiment_custom_tags=None, verbose=False, system=False, add_to_model_list=False)
INFO:logs:Checking exceptions
INFO:logs:Importing libraries
INFO:logs:Copying training dataset
INFO:logs:Defining folds
INFO:logs:Declaring metric variables
INFO:logs:Importing untrained model
INFO:logs:Declaring custom model
INFO:logs:_ensembleforecaster Imported successfully
INFO:logs:Starting cross validation
INFO:logs:Cross validating with SlidingWindowSplitter(fh=ForecastingHorizon([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]), dtype='int64', is_relative=True),
    step_length=15, window_length=30), n_jobs=-1
INFO:logs:Calculating mean and std
INFO:logs:Creating metrics dataframe
INFO:logs:Finalizing model
INFO:logs:Uploading results into container
INFO:logs:master_model_container: 67
INFO:logs:display_container: 16
```



```
final_model_arima=finalize_model(arima)
```

```
INFO:logs:Initializing finalizing_model()
INFO:logs:finalize_model(self:pycaret.time_series.forecasting.oop.TSForecastingExperiment object at 8a7f7b648efb0b, estimator=ARIMA(seasonal_order=(0, 1, 0, 7)), fit_huargs=None, groups=None, model_only=True, experiment_custom_tags=None, display=None)
INFO:logs:finalizing ARIMA(seasonal_order=(0, 1, 0, 7))
INFO:logs:Initializing create_model()
INFO:logs:create_model(self:pycaret.time_series.forecasting.oop.TSForecastingExperiment object at 8a7f7b648efb0b, estimator=ARIMA(seasonal_order=(0, 1, 0, 7)), fold=None, round=4, cross_validation=True, predict=True, fit_huargs=[], groups=None, ref=None)
INFO:logs:Checking exceptions
INFO:logs:Importing libraries
INFO:logs:Copying training dataset
INFO:logs:Defining folds
INFO:logs:Declaring metric variables
INFO:logs:Importing untrained model
INFO:logs:Declaring custom model
INFO:logs:ARIMA Imported successfully
INFO:logs:Starting cross validation
INFO:logs:Cross validating with SlidingWindowSplitter(fh=forecastinghorizon([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15], dtype='int64', is_relative=True), step_length=15, window_length=30), n_jobs=-1)
INFO:logs:Calculating mean and std
INFO:logs:Creating metrics dataframe
INFO:logs:finalizing model
INFO:logs:Uploading results into container
INFO:logs:master_model_container: 67
INFO:logs:display_container: 17
INFO:logs:ARIMA(seasonal_order=(0, 1, 0, 7))
INFO:logs:create_model() successfully completed.....
INFO:logs:master_model_container: 67
INFO:logs:display_container: 16
INFO:logs:ARIMA(seasonal_order=(0, 1, 0, 7))
INFO:logs:finalize_model() successfully completed.....
```



```
predict_model(final_model, fh=15)
```

WARNING:logs:predict_model >> Prediction Indices do not match test indices. Metrics will not be displayed.

```
y_pred
```

2021-10-06	8067.1000
2021-10-07	9247.4003
2021-10-08	7402.4460
2021-10-09	4180.3440
2021-10-10	5112.1593
2021-10-11	7982.8843
2021-10-12	9048.8743
2021-10-13	8284.6184
2021-10-14	9458.4056
2021-10-15	7613.4672
2021-10-16	4371.3720
2021-10-17	5323.1902
2021-10-18	8193.7164
2021-10-19	9259.9089
2021-10-20	8475.6492

```
predict_model(final_model_arima)

WARNING:logs:predict_model >> Prediction Indices do not match test indices. Metrics will not be displayed.

y_pred
2021-10-06 8010.3893
2021-10-07 8469.1729
2021-10-08 7391.2030
2021-10-09 4187.2898
2021-10-10 5325.7258
2021-10-11 8132.1880
2021-10-12 9101.8085
2021-10-13 8300.4595
2021-10-14 8759.3551
2021-10-15 7681.4329
2021-10-16 4477.5397
2021-10-17 5615.9845
2021-10-18 8422.4484
2021-10-19 9452.0704
2021-10-20 8590.7240
```

```
[ ] save_model(final_model,'pycaret_prediction')

INFO:logs:Initializing save_model()
INFO:logs:save_model(model:_ensembleforecasterwithvoting(forecasters=[('ARIMA',
ARIMA(seasonal_order=(0, 1, 0, 7))),
('Seasonal Naive Forecaster',
NaiveForecaster(sp=7)),
('Auto ARIMA',
AutoARIMA(random_state=4461, sp=7,
suppress_warnings=True))),
n_jobs=-1), model_name=pycaret_prediction, prep_pipe=None, verbose=True, use_case=MLUseCase.TIME_SERIES, kwargs={})
INFO:logs:Adding model into prep_pipe
WARNING:logs:Only Model saved. Transformations in prep_pipe are ignored.
INFO:logs:pycaret_prediction.pkl saved in current working directory
INFO:logs:_ensembleforecasterwithvoting(forecasters=[('ARIMA',
ARIMA(seasonal_order=(0, 1, 0, 7))),
('Seasonal Naive Forecaster',
NaiveForecaster(sp=7)),
('Auto ARIMA',
AutoARIMA(random_state=4461, sp=7,
suppress_warnings=True))),
n_jobs=-1)
INFO:logs:save_model() successfully completed.....
Transformation Pipeline and Model Successfully Saved
(_ensembleforecasterwithvoting(forecasters=[('ARIMA',
ARIMA(seasonal_order=(0, 1, 0, 7))),
('Seasonal Naive Forecaster',
NaiveForecaster(sp=7)),
('Auto ARIMA',
AutoARIMA(random_state=4461, sp=7,
suppress_warnings=True))),
n_jobs=-1), 'pycaret_prediction.pkl')

[ ] load_model('pycaret_prediction')

INFO:logs:Initializing load_model()
INFO:logs:load_model(model_name=pycaret_prediction, platform=None, authentication=None, verbose=True)
Transformation Pipeline and Model Successfully Loaded
_ensembleforecasterwithvoting(forecasters=[('ARIMA',
ARIMA(seasonal_order=(0, 1, 0, 7))),
('Seasonal Naive Forecaster',
NaiveForecaster(sp=7)),
('Auto ARIMA',
AutoARIMA(random_state=4461, sp=7,
suppress_warnings=True))),
n_jobs=-1)
```

Tài liệu tham khảo

- [1] A. Natorina, “the adaptive management system of marketin commodity policy,” *Baltic Journal of Economic Studies*, vol. 5, no. 1, pp. 131–136, 2019.
- [2] T. Muwema and J. Phiri, “the impact of integrated financial management information systems on procurement process in public sector in developing countries—a case of Zambia,” *Open Journal of Business and Management*, vol. 08, no. 02, pp. 983–996, 2020.
- [3] M. B. Tudose and S. Avasilcai, “Financial performance management and economic cycle variations. Evidence for

textile industry,” *IOP Conference Series: Materials Science and Engineering*, vol. 1169, no. 1, p. 012016, 2021.

[4] G. Cernius and L. Birškyte, “Financial information and management decisions: impact of accounting policy on financial indicators of the firm,” *Business: Theory and Practice*, vol. 21, no. 1, pp. 48–57, 2020.

[5] L. Hrytsenko and I. Boiarko, “the empirical evaluation of the usefulness of accounting and financial information for strategic management of enterprises in Ukraine,” *Financial and Credit Activity Problems of Theory and Practice*, vol. 3, no. 34, pp. 111–117, 2020.

[6] O. Kundrya-Vysotska and I. Demko, “Environmental social and management accounting information in accounting systems as a tool for verifying the concept of sustainable development,” *Financial and Credit Activity Problems of Theory and Practice*, vol. 2, no. 33, pp. 554–565, 2020.

[7] D. Wahyuningsih, R. A. Nuraliaty, D. C. Darma, J. Kasuma, and Sriwardani, “Why dynamic capacity influences the quality of management accounting Information systems in the public sector?” *International Journal of Psychosocial Rehabilitation*, vol. 24, no. 10, pp. 4032–4044, 2020.

[8] L. Puspitawati, “Strategic information moderated by effectiveness management accounting information systems: business strategy approach,” *Jurnal Akuntansi*, vol. 25, no. 1, p. 101, 2021.

[9] G. S. Tian and T. Li, “Application of accounting information technology in agricultural machinery modernization design—based on cloud computing system,” *Journal of Agricultural Mechanization Research*, vol. 42, no. 8, pp. 218–221, 2020.

[10] L. H. Zhu, “Strategy on accounting informatization management based on financial shared services,” *Journal of Nantong Textile Vocational Technology College*, vol. 18, no. 3, pp. 63–66, 2018.

[11] S. Y. Su, “Financial accounting information management system of petroleum enterprises in China,” *Yunnan Chemical Technology*, vol. 47, no. 1, pp. 29–30, 2020.

[12] A. Khaliq, A. Umair, R. Khan, S. Iqbal, and A. Abbass, “Leadership and decision making among SMEs: management accounting information and the moderating role of cloud computing,” *Business Ethics*

and Leadership, vol. 5, no. 2,
pp. 78–95, 2021.

[13] A. Alam, I. Ullah, and Y. K. Lee, *Video Big Data Analytics in the Cloud: A Reference Architecture, Survey, Opportunities, and Open Research Issues*, IEEE Access, vol. 99, p. 1, 2020.

[14] L. M. Akimova, O. O. Osadcha, V. V. Bashtannyk, N. M. Kondratska, and K. M. Fedyna, “Formation of the system of financial-information support of environmentally oriented management of the enterprise,” *Financial and Credit Activity Problems of Theory and Practice*, vol. 1, no. 32, pp. 434–443, 2021.

[15] X. Leng, “Discussion on accounting informatization construction under the background of big data Era,” *Journal of Chifeng University*, vol. 33, no. 4, pp. 95-96, 2017.

[16] R. Wang, “Impact of accounting informatization on financial management and countermeasures,” *Journal of Hunan Institute of Engineering (Social Science Edition)*, vol. 29, no. 3, pp. 25–28, 2019.

[17] S. Massicotte and J. F. Henri, “The use of management accounting information by boards of directors to oversee strategy implementation,” *British Accounting Review*, vol. 53, no. 3, p. 100953, 2021.

[18] Y. Ji, T. Xia, H. Zhang, and G. Chen, “Transaction data management system based on distributed storage architecture,” *Journal of Physics: Conference Series*, vol. 1550, no. 3, p. 032030, 2020. 10 Scientific Programming

[19] J. Liang and H. Chen, “Research on open source platform of digital power grid based on CEPH open source distributed storage technology,” *Journal of Physics: Conference Series*, vol. 2237, no. 1, p. 012029, 2022.

[20] Y. F. Zhao and Y. Y. Chen, “Research on enterprise environmental cost management system based on material flow cost accounting,” *Journal of Xi'an University of Finance and Economics*, vol. 31, no. 2, pp. 36–40, 2018.

[21] Y. H. Zhao, “the development trend of enterprise management informatization and its influence on accounting,” *HEILONGJIANG SCIENCE*, vol. 8, no. 18, pp. 90-91, 2017.