
CÁC NHÂN TỐ ẢNH HƯỞNG ĐẾN VIỆC ÁP DỤNG PHÂN TÍCH DỮ LIỆU LỚN: NGHIÊN CỨU TẠI VIỆT NAM

Trần Bình Minh

Đại học Kinh tế Quốc dân

Email: minhbt@neu.edu.vn

Nguyễn Thị Quế*

Trường Đại học Công nghiệp Hà Nội

Email: nguyenthique@hau.edu.vn

Trương Thanh Hằng

Trường Đại học Công nghiệp Hà Nội

Email: truongthanhhang@hau.edu.vn

Mã bài: JED-2043

Ngày nhận bài: 04/10/2024

Ngày nhận bài sửa: 23/10/2024

Ngày duyệt đăng: 10/03/2025

DOI: 10.33301/JED.V1.2043

Tóm tắt

Phân tích dữ liệu lớn không chỉ giúp cải thiện hiệu suất và ra quyết định mà còn cung cấp các cơ hội mới để đổi mới và phát triển. Trong bối cảnh ngày càng có nhiều dữ liệu được tạo ra và thu thập, khả năng phân tích và tận dụng dữ liệu lớn sẽ trở nên ngày càng quan trọng. Nghiên cứu này được thực hiện nhằm xác định mức độ ảnh hưởng của các yếu tố đến việc áp dụng phân tích dữ liệu lớn trong các doanh nghiệp tại Việt Nam. Phân tích định lượng được thực hiện trên dữ liệu khảo sát gồm 330 công ty tại Việt Nam. Kết quả chỉ ra rằng, lợi thế tương đối, cơ sở hạ tầng công nghệ, khả năng hấp thụ thông tin, cạnh tranh trong ngành và sự hỗ trợ của các cơ quan nhà nước ảnh hưởng đến mức độ áp dụng phân tích dữ liệu lớn. Từ đó, nhóm tác giả đề xuất một số khuyến nghị nhằm tăng cường áp dụng phân tích dữ liệu lớn vào các doanh nghiệp.

Từ khóa: Dữ liệu lớn, doanh nghiệp, tiến bộ công nghệ, Việt Nam.

Mã JEL: C55, N75, O14

Determinants of the application of big data analytics: The case of Vietnam

Abstract

Big Data Analytics helps improve performance and decision-making and provides new opportunities for innovation and growth. In a context where more and more data is being generated and collected, the ability to analyze and leverage big data will become increasingly important. This research determines the extent to which various factors influence Big Data Analytics adoption in Vietnam's businesses. Data analysis was conducted using survey data from 330 companies in Vietnam. The results indicate that relative advantage, technological infrastructure, absorptive capacity, industry competition, and governmental agencies' support significantly affect Big Data Analytics adoption. Based on these findings, the authors propose several recommendations to enhance the adoption of Big Data Analytics in businesses.

Keywords: Big data analysis (BDA), firms, technological advancement, Vietnam.

JEL Codes: C55, N75, O14

1. Giới thiệu

Phân tích dữ liệu lớn đóng vai trò ngày càng quan trọng trong nhiều lĩnh vực, từ kinh doanh đến nghiên cứu khoa học. Đặc biệt, phân tích dữ liệu lớn, có thể cải thiện hiệu suất và hiệu quả bằng cách tối ưu hóa quy trình và quản lý hiệu quả như: Dữ liệu lớn giúp xác định các điểm yếu trong quy trình và hoạt động, từ đó tối ưu hóa quy trình để tăng cường hiệu suất và giảm chi phí. Phân tích dữ liệu lớn có thể phát hiện rủi ro và gian lận, giúp hiểu rõ hơn về sở thích và nhu cầu của khách hàng, và tạo ra các dịch vụ và sản phẩm cá nhân hóa hơn, cải thiện sự hài lòng của khách hàng. Ngoài ra, phân tích dữ liệu lớn có thể tăng cường bảo mật cho đơn vị, tối ưu hóa chi phí và phân tích được các hành vi của người dùng.

Zhang & cộng sự (2017), cho rằng phân tích dữ liệu lớn hỗ trợ các công ty quản lý và xử lý các cấu trúc dữ liệu phức tạp, đa dạng và quy mô lớn một cách hiệu quả. Phân tích dữ liệu lớn cho phép các doanh nghiệp trích xuất, nắm bắt, khám phá và phân tích dữ liệu song song với mục tiêu phát triển công nghiệp bền vững (Raut & cộng sự, 2019). Hơn nữa, các công ty không biết rằng họ đang xử lý khối lượng lớn dữ liệu tăng theo cấp số nhân mỗi năm. Tuy nhiên, khả năng xử lý dữ liệu hạn chế của họ có thể hạn chế các công ty này phát triển năng động (Williamson, 2014). Do đó, phân tích dữ liệu lớn là một trong những lựa chọn tốt nhất để theo dõi, quan sát và tạo ra thông tin cho các giao dịch kinh doanh có khối lượng lớn.

Việc tính xác thực đề cập đến tính xác thực, tính toàn vẹn và độ tin cậy của dữ liệu trong việc cung cấp thông tin đáng tin cậy cho người dùng (Fan & cộng sự, 2014). phân tích dữ liệu lớn là một cơ chế quan trọng để tích hợp, giải quyết và đẩy nhanh các quy trình kinh doanh trong ngành logistics. phân tích dữ liệu lớn mang lại lợi ích cho các công ty thông qua việc cải thiện hoạt động hàng ngày và trải nghiệm tổng thể của khách hàng (Gupta & George, 2016).

Nghiên cứu về việc áp dụng phân tích dữ liệu lớn tại Việt Nam là một lĩnh vực tương đối mới. Nghiên cứu này được thực hiện để các yếu tố ảnh hưởng đến việc áp dụng phân tích dữ liệu lớn trong các doanh nghiệp tại Việt Nam. Theo đó, nghiên cứu bao gồm cấu trúc sau: Giới thiệu chủ đề nghiên cứu (mục 1); Tổng quan nghiên cứu (mục 2); Phương pháp nghiên cứu (mục 3); Kết quả và thảo luận (mục 5); Kết luận và khuyến nghị (mục 5).

2. Tổng quan nghiên cứu

2.1. Việc áp dụng phân tích dữ liệu lớn

Kauffman & cộng sự (2012), một lượng lớn dữ liệu được phát triển và thu thập bởi các cửa hàng trực tuyến, trang mạng xã hội và các ứng dụng khác. Hiện tượng này chủ yếu là do sự phát triển nhanh chóng của công nghệ thương mại điện tử. Việc áp dụng quản lý và phân tích thông tin phù hợp là rất quan trọng trong kinh doanh ngày nay vì có thể hỗ trợ các công ty khám phá kiến thức mới và tạo điều kiện thuận lợi trong việc ứng phó với những cơ hội bất ngờ (Chen & cộng sự, 2013). phân tích dữ liệu lớn cũng được coi là điều quan trọng vì các tổ chức đang thay đổi mô hình kinh doanh (Gobble, 2013). Phân tích dữ liệu lớn trở thành một cách tiếp cận để các tổ chức đạt được lợi thế cạnh tranh trong những năm gần đây (Tan & cộng sự, 2015). Phân tích dữ liệu lớn cho phép các tổ chức đưa ra quyết định tốt hơn và nhanh hơn bằng cách hiểu được một tập dữ liệu khổng lồ khó truy cập bằng cơ sở dữ liệu truyền thống. Nó có thể hỗ trợ các tổ chức có được những hiểu biết mới về nguồn dữ liệu và đưa ra quyết định đúng đắn trong thời gian ngắn hơn.

Nghiên cứu của Gandomi & Haider (2015), các tổ chức phải đối mặt với các loại thông tin riêng biệt về cơ bản chứa nhật ký tìm kiếm và người dùng, tài liệu giao dịch của khách hàng và nội dung do khách hàng tạo. Do thông tin có khối lượng lớn, tốc độ cao và đa dạng được tạo ra thông qua các kênh riêng biệt, thuật ngữ 'dữ liệu lớn' được sử dụng để mô tả tính chất khổng lồ và phức tạp của thông tin được lưu trữ của tổ chức. Quy trình quản lý thông tin bao gồm quá trình thu thập, ghi, trích xuất, làm sạch và chú thích dữ liệu. Mặt khác, quy trình phân tích bao gồm mô hình hóa, phân tích và giải thích dữ liệu để phát hiện các mô hình, xu hướng và mối liên hệ liên quan đến hành vi và tương tác của con người. Zhang & cộng sự (2017) cho thấy lợi thế tương đối, cơ sở hạ tầng công nghệ, khả năng hấp thụ và sự hỗ trợ của chính phủ ảnh hưởng đến mức độ áp dụng phân tích dữ liệu lớn, trong khi cạnh tranh trong ngành dường như không có ảnh hưởng đáng kể.

Phân tích dữ liệu lớn giúp các tổ chức sử dụng đầy đủ dữ liệu và khám phá các cơ hội kinh doanh mới. Phân tích dữ liệu lớn hỗ trợ các công ty nâng cao hiệu quả hoạt động kinh doanh, đạt được tỷ suất lợi nhuận cao hơn và đạt được mức độ hài lòng của khách hàng cao hơn (Attanran & cộng sự, 2018). Một cuộc phỏng vấn với hơn 50 chủ doanh nghiệp trong năm 2018 cho thấy rằng các doanh nghiệp có thể đạt được mức độ

áp dụng phân tích dữ liệu lớn bao gồm giảm chi phí, ra quyết định kịp thời và chính xác cũng như quảng bá nhanh chóng các sản phẩm và dịch vụ mới. Nhờ đó, phân tích dữ liệu lớn giúp doanh nghiệp khám phá những hiểu biết sâu sắc hơn, đưa ra dự đoán, làm rõ các vấn đề tiềm ẩn và đưa ra những khuyến nghị hữu ích. Theo Van der Meulen & Woods (2015), 75% công ty đã có kế hoạch đầu tư vào công nghệ phân tích dữ liệu lớn trong hai năm tới. Các công ty này bao gồm nhiều lĩnh vực khác nhau như viễn thông, bán lẻ và chăm sóc sức khỏe, chuỗi cung ứng và lĩnh vực hậu cần. Các công ty đang trở nên tự tin hơn đối với công nghệ phân tích dữ liệu lớn được kỳ vọng sẽ thay đổi chiến lược kinh doanh của họ.

2.2. Lợi thế tương đối

Theo Williamson (2014), các doanh nghiệp có thể sử dụng dữ liệu bên trong và bên ngoài để phân tích các xu hướng kinh doanh hiện tại và những thay đổi về nhu cầu thông qua việc áp dụng phân tích dữ liệu lớn. phân tích dữ liệu lớn rất hữu ích trong việc hỗ trợ các công ty quản lý hiệu quả hoạt động kho hàng bằng cách giảm thiểu sự kém cỏi về thông tin bất cân xứng. Nghiên cứu của Agrawal (2015) cho rằng việc áp dụng phân tích dữ liệu lớn cho phép các công ty cải thiện hiệu quả hoạt động kinh doanh và giảm thời gian thực hiện cũng như chi phí lao động. Phân tích dữ liệu lớn cũng có thể hỗ trợ các công ty dự đoán rủi ro kinh doanh có thể xảy ra, theo dõi và theo dõi dòng hàng tồn kho, cải thiện việc ra quyết định và cung cấp thông tin theo thời gian thực để đạt được lợi thế cạnh tranh. Francisco & Swanson (2018) cũng nhấn mạnh rằng phân tích dữ liệu lớn cho phép minh bạch dữ liệu trong hoạt động kho hàng và toàn bộ quy trình chuỗi cung ứng.

Lợi thế tương đối là yếu tố quyết định quan trọng nhất trong bối cảnh áp dụng tiến bộ công nghệ (Wang & cộng sự, 2016). Wahab & cộng sự (2018) cho rằng, việc áp dụng tiến bộ công nghệ mới cho phép các công ty vận hành hoạt động kinh doanh tốt hơn, cung cấp các giải pháp ưu việt, tăng hiệu suất và hiệu quả hoạt động cũng như cải thiện mức năng suất. Vì vậy, nghiên cứu đưa ra giả thuyết như sau:

H1: Lợi thế tương đối của ảnh hưởng tích cực đến việc áp dụng phân tích dữ liệu lớn.

2.3. Cơ sở hạ tầng công nghệ

Theo Dang-Pham & Nkhoma (2013), cơ sở hạ tầng công nghệ được coi là nguồn lực kinh doanh chính để các doanh nghiệp duy trì lợi thế cạnh tranh trong dài hạn. Cơ sở hạ tầng công nghệ tích hợp cao có khả năng tận dụng thế mạnh của công ty và các đề xuất bán hàng độc đáo so với đối thủ cạnh tranh. Trong môi trường kinh doanh cạnh tranh hiện nay, cơ sở hạ tầng công nghệ tiên tiến giúp các doanh nghiệp có thể hành động kịp thời để nhu cầu của khách hàng. Cơ sở hạ tầng vững chắc cuối cùng sẽ cải thiện các giao dịch kinh doanh tổng thể (Tan & cộng sự, 2015). Hơn nữa, cơ sở hạ tầng công nghệ là điều cần thiết để tiếp thu những đổi mới và cải thiện các giao dịch vận hành (Kumar & Kumar, 2015).

Tan & cộng sự (2015) cho rằng sự tiến bộ của cơ sở hạ tầng công nghệ tạo nền tảng kỹ thuật cho việc phân tích dữ liệu, cho phép các công ty dễ dàng bắt đầu các hoạt động kinh doanh của mình. Do đó, người ta tin rằng cơ sở hạ tầng công nghệ ở mức độ cao hơn sẽ tăng cơ hội cho các công ty áp dụng phân tích dữ liệu lớn trong hoạt động kinh doanh của mình.

Ling & Wahab (2020), vì độ phức tạp của dữ liệu ngày càng tăng chắc chắn sẽ ảnh hưởng đến cơ sở hạ tầng công nghệ hiện có khả năng xử lý dữ liệu khá hạn chế. Điều này có nghĩa là cơ sở hạ tầng công nghệ phải có khả năng xử lý các tập dữ liệu khổng lồ và hệ thống phân tích tiên tiến có khả năng tạo ra các mô hình dự đoán và thuật toán thống kê. Cơ sở hạ tầng công nghệ không đủ có thể dẫn đến một số vấn đề công nghệ như lưu trữ không hiệu quả và không đáng tin cậy, xử lý chậm và tích hợp dữ liệu yếu. Những vấn đề này thường gặp phải bởi các công ty muốn trích xuất kiến thức mô tả và dự đoán từ các tập dữ liệu cực lớn của họ (Almoqren & Altayar, 2016). Các nghiên cứu gần đây đã chỉ ra mối quan hệ tích cực giữa cơ sở hạ tầng công nghệ và việc áp dụng công nghệ bao gồm phân tích dữ liệu lớn (Wahab & cộng sự, 2021). Vì vậy, nghiên cứu đưa ra giả thuyết như sau:

H2: Cơ sở hạ tầng công nghệ ảnh hưởng tích cực đến việc áp dụng phân tích dữ liệu lớn.

2.4. Khả năng hấp thụ thông tin

Theo Zhang & cộng sự (2017), khả năng hấp thụ được định nghĩa là khả năng và khả năng của công ty trong việc khám phá những thông tin có giá trị được thu thập từ môi trường bên ngoài, bao gồm cả các đối thủ cạnh tranh. Nó có thể được thực hiện bằng cách đồng hóa, cách mạng hóa và nâng cao các quy trình kinh

doanh hiện có theo hướng lợi thế cạnh tranh.

Anuradha (2015) coi khả năng hấp thụ là yếu tố quyết định quan trọng để các công ty tiếp thu kiến thức mới và quản lý chặt chẽ các luồng thông tin. Khả năng hấp thụ là hết sức cần thiết để các doanh nghiệp đạt được mức độ đổi mới cao hơn (Govindan & cộng sự, 2018). Nó cũng có thể được mô tả là mức độ mà các công ty có khả năng hoạt động. Hệ thống dữ liệu về mức độ sẵn sàng về mặt công nghệ (Hurwitz & cộng sự, 2013).

Vì vậy, việc áp dụng công nghệ phù hợp cũng như mức độ nỗ lực là những yếu tố tốt nhất tạo nên khả năng hấp thụ xứng đáng và hiệu quả. Nó có thể được tạo ra từ cả phương pháp định tính và định lượng bao gồm nghiên cứu và phát triển (R&D) và đầu tư (Neaga & cộng sự, 2015). Theo Wahab & cộng sự (2021), khả năng hấp thụ có thể được xác định rõ ràng thông qua mức độ kiến thức và kinh nghiệm trước đây mà các công ty sở hữu trước khi áp dụng công nghệ tiên tiến bao gồm phân tích dữ liệu lớn. Hơn nữa, kỹ năng và chuyên môn là một số thuộc tính quan trọng của khả năng hấp thụ. Nguồn nhân lực, đào tạo và giáo dục, kỹ năng và kinh nghiệm là những yếu tố quan trọng đặc trưng cho khả năng hấp thụ của một công ty. Các công ty có nhiều khả năng áp dụng công nghệ mới hơn khi có khả năng hấp thụ cao (Park & cộng sự, 2015). Điều này sẽ ảnh hưởng đáng kể đến sự thành công của việc áp dụng tiến bộ công nghệ. Vì vậy, nghiên cứu đưa ra giả thuyết như sau:

H3: Khả năng hấp thụ thông tin ảnh hưởng tích cực đến việc áp dụng phân tích dữ liệu lớn.

2.5. Cạnh tranh trong ngành

Cạnh tranh trong ngành đề cập đến sự cạnh tranh năng động giữa các đối thủ kinh doanh trong cùng một thị trường. Việc áp dụng phân tích dữ liệu lớn đóng vai trò là lợi thế cạnh tranh để một công ty đạt được thành công so với các đối thủ cạnh tranh (Wahab & cộng sự, 2021). Theo Chen & cộng sự (2013), với việc áp dụng tiến bộ công nghệ, các công ty có thể duy trì hoạt động kinh doanh lâu dài. Các học giả đã đồng ý rằng cạnh tranh trong ngành đóng vai trò là một khía cạnh quan trọng sẽ ảnh hưởng đến ý định áp dụng tiến bộ công nghệ mới của các công ty. Tương tự, việc áp dụng phân tích dữ liệu lớn mang lại lợi thế cạnh tranh to lớn cho công ty để đáp ứng tính chất khó lường; khối lượng giao dịch của khách hàng, giá trị của dữ liệu được trích xuất, tốc độ tạo dữ liệu mới, tính xác thực của dữ liệu được phổ biến và sự đa dạng về nhu cầu của khách hàng (Wang & cộng sự, 2016; Addo-Tenkorang & Helo, 2016; Ghobakhloo & cộng sự, 2011). Theo đó, việc áp dụng phân tích dữ liệu lớn được cho là giúp các công ty cải thiện hiệu quả kinh doanh tổng thể của chuỗi cung ứng (Gunasekaran & cộng sự, 2018). Phân tích dữ liệu lớn là một lựa chọn hoàn hảo để các công ty thiết lập lợi thế cạnh tranh (Sangari & Razmi, 2015), đạt được sự bền vững trong kinh doanh và cải thiện lợi tức đầu tư của họ (Sivarajah, Kamal & cộng sự, 2017). 40% công ty Malaysia áp dụng phân tích dữ liệu lớn được cho là đã đạt được lợi thế cạnh tranh ở mức độ cao. Vì vậy, nghiên cứu đưa ra giả thuyết như sau:

H4: Cạnh tranh trong ngành ảnh hưởng tích cực đến việc áp dụng phân tích dữ liệu lớn.

2.6. Hỗ trợ của chính phủ

Sự hỗ trợ của chính phủ có vai trò quan trọng đối với cả các doanh nhân trong việc mở rộng kinh doanh, đặc biệt là trong ngành logistics đầy cạnh tranh. Chính phủ có trách nhiệm cung cấp các gói thông tin bao gồm việc áp dụng công nghệ mới (Erind, 2015). Đặc biệt, mối quan tâm về môi trường đã thúc đẩy chính phủ điều chỉnh các hoạt động kinh doanh hỗ trợ kinh doanh bền vững thông qua việc áp dụng công nghệ xanh. Theo Raut & cộng sự (2019), một phần trách nhiệm của chính phủ đối với hoạt động kinh doanh bền vững là cung cấp các chính sách, thủ tục, sáng kiến và đầu tư rõ ràng để có thể thúc đẩy các doanh nghiệp áp dụng cơ sở hạ tầng công nghệ mới.

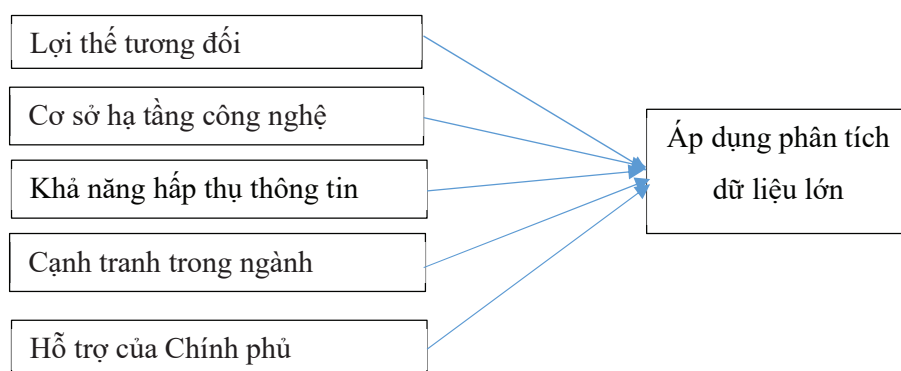
H5: Hỗ trợ của Chính phủ ảnh hưởng tích cực đến việc áp dụng phân tích dữ liệu lớn.

Từ đó, nhóm tác giả đề xuất mô hình nghiên cứu như tại Hình 1.

2.7. Các biến trong mô hình

Biến Áp dụng phân tích dữ liệu lớn được kế thừa Zhang & cộng sự (2017) bao gồm 3 biến quan sát: i/ Công ty chúng tôi dự định áp dụng phân tích dữ liệu lớn, ii/ Công ty chúng tôi dự định bắt đầu sử dụng phân tích dữ liệu lớn thường xuyên trong tương lai, iii/ Công ty chúng tôi rất muốn giới thiệu phân tích dữ liệu lớn để các công ty khác áp dụng.

Hình 1: Mô hình nghiên cứu đề xuất



Các biến độc lập bao gồm:

- Lợi thế tương đối: 5 biến quan sát;
- Cơ sở hạ tầng công nghệ: 5 biến quan sát;
- Khả năng hấp thụ thông tin: 4 biến quan sát;
- Cạnh tranh trong ngành: 5 biến quan sát;
- Hỗ trợ chính phủ: 5 biến quan sát.

3. Phương pháp nghiên cứu

3.1. Thu thập dữ liệu

Từ các nhân tố ảnh hưởng đã được khám phá trong giai đoạn nghiên cứu, bảng câu hỏi khảo sát định lượng được triển khai đến các đối tượng khảo sát như giám đốc, phó giám đốc, trưởng, phó phòng, các nhân viên đang làm việc tại các công ty: trách nhiệm hữu hạn, cổ phần, hành chính sự nghiệp tại Việt Nam. Các công ty này được tác giả gửi ngẫu nhiên thông qua các hình thức gửi khảo sát như sau: (i) trực tiếp, (ii) gửi thư, (iii) gửi email, (iv) trực tuyến (Google Docs), (v) khác. Kết quả khảo sát thu về 343 phiếu trong tổng số 398 phiếu phát ra. Sau khi lựa chọn các phiếu khảo sát không hợp lệ do có nhiều ô trống hoặc không đầy đủ thông tin, tác giả lựa chọn sử dụng 330 phiếu hợp lệ, đạt tỷ lệ 82,91%.

3.2. Xử lý dữ liệu

Tác giả sử dụng phần mềm SPSS 22, các chỉ tiêu để phân tích gồm: Giá trị trung bình (Mean), giá trị trung vị (Median), giá trị lớn nhất (Max), giá trị nhỏ nhất (Min) của các biến nghiên cứu và đo lường. Đo lường tính biến thiên của dữ liệu sử dụng giá trị độ lệch chuẩn (Standard deviation), Phân tích thống kê mô tả, Hệ số Cronbach's. Tiếp theo, nhóm tác giả phân tích EFA và phương pháp Principal Component Analysis, phép quay Varimax để rút trích các nhân tố chính. Trên cơ sở kết quả EFA, đề xuất mô hình nghiên cứu. Để ước lượng mức độ tương quan của các nhân tố đến việc áp dụng phân tích dữ liệu lớn, nhóm tác giả sử dụng mô hình phân tích hồi quy bội để tính toán các tham số của các nhân tố được sử dụng trong mô hình.

4. Kết quả nghiên cứu

4.1. Thống kê mô tả

Theo Bảng 1, các đối tượng tham gia khảo sát chủ yếu là Nữ, chiếm tỷ lệ là 60,3%, trong khi đó Nam chiếm 39,7%. Quan điểm của các đối tượng khảo sát tập trung nhiều ở nhóm nhân viên với số lượng là 134 người, chiếm tỷ lệ cao nhất 40.61%, nhóm có chức vụ trưởng phó phòng chiếm tỷ lệ cao thứ hai 36,67% và cuối cùng là nhóm chức vụ giám đốc/phó giám đốc tỷ lệ 22,73%.

Theo Bảng 1, Các đối tượng tham gia khảo sát chủ yếu là công ty cổ phần chiếm phần lớn 44,24%. các công ty trách nhiệm hữu hạn tham gia khảo sát chiếm 30,61%, các công ty, đơn vị thuộc hành chính sự nghiệp chiếm 25,15%.

4.2. Kiểm định thang đo

Kiểm định thang đo tác động của công nghệ thông tin đến kiểm toán nội bộ được thực hiện bằng hệ số tin

Bảng 1: Phân loại theo giới tính và chức vụ, công ty của đối tượng tham gia khảo sát

Giới tính	Chức vụ			Tổng cộng		Theo công ty khảo sát		
	Giám đốc/phó giám đốc	Trưởng, phó phòng tương đương	Nhân viên	Số lượng	Tỷ lệ	Các công ty tham gia khảo sát	Số lượng	Tỷ lệ
Nam	41	45	45	131	39,70%	Công ty cổ phần	146	44,24%
Nữ	34	76	89	199	60,30%	Công ty trách nhiệm hữu hạn	101	30,61%
Tổng cộng	75	121	134	330	100%	Hành chính sự nghiệp	83	25,15%
Tỷ lệ	22,73%	36,67%	40,61%	100%		Tổng cộng	330	100%

Nguồn: Dữ liệu do nhóm tác giả thu thập.

Bảng 2: Kết quả hệ số Cronbach's Alpha của các thang đo

	Số biến quan sát		Cronbach's Alpha	Hệ số tương quan tổng nhỏ nhất	Hệ số KMO và Bartlett	
	Trước	Sau			Chỉ tiêu	Mô hình
Lợi thế tương đối (RA)	5	5	,751	,746	Chỉ số KMO	0,873
Cơ sở hạ tầng công nghệ (TI)	5	5	,783	,674	Bartlett's	19,863,764
Khả năng hấp thụ thông tin (AOI)	4	4	,835	,771	Kiểm định Bartlett	0,000
Cạnh tranh trong ngành (CI)	5	5	,849	,822	có giá trị sig	
Hỗ trợ của Chính phủ (GS)	5	4	,877	,744	Giá trị tổng phương sai trích	55,642
Việc áp dụng phân tích dg values n (BDA)	3	3	,893	,874	Giá trị Eigenvalues nhỏ nhất	1,421

Nguồn: Kết quả phân tích dữ liệu của nhóm tác giả.

cây Cronbach's Alpha và phân tích EFA. Bảng 2 cho thấy, phần lớn nhân tố đều có hệ số Cronbach's Alpha lớn hơn 0,6; hệ số tương quan biến tổng của các thang đo đều lớn hơn 0,3. Sau khi loại trừ các biến quan sát bao gồm GS3.

4.3. Phân tích nhân tố khám phá và tương quan giữa các biến

4.3.1. Kiểm định KMO và Bartlett

Kết quả phân tích nhân tố (Bảng 3) cho thấy chỉ số KMO là 0,873 đều >0,5, điều này chứng tỏ dữ liệu dùng để phân tích nhân tố là hoàn toàn thích hợp.

Kết quả kiểm định Bartlett's là 1986,3764 với mức ý nghĩa đều là (p_value) sig = 0,00 < 0,05, (bác bỏ giả thuyết H0: các biến quan sát không có tương quan với nhau trong tổng thể). Như vậy, giả thuyết về ma trận tương quan giữa các biến là ma trận đồng nhất bị bác bỏ, tức là các biến có tương quan với nhau và thỏa mãn điều kiện phân tích nhân tố.

Kết quả cho thấy với các biến quan sát còn lại sau khi đã loại các thang đo không thỏa mãn độ tin cậy. Giá trị tổng phương sai trích là 55,642% đều đạt yêu cầu >50%; như vậy có thể nói rằng các nhân tố này giải thích được 55,642% sự biến thiên của dữ liệu. Giá trị hệ số Eigenvalues của các nhân tố đều cao (>1), nhân tố có Eigenvalues (thấp nhất) là 1,421 đều thỏa mãn >1.

Như vậy, phân tích EFA thích hợp với các dữ liệu và các biến quan sát có tương quan với nhau trong tổng thể, nên được sử dụng cho phân tích tiếp theo.

Phân tích EFA được thực hiện với phương pháp trích hệ số là Component Analysis và phép xoay Varimax, kết quả phân tích có 6 biến quan sát của thang đo tại Bảng 3.

4.3.2. Phân tích tương quan giữa các biến trong mô hình

Bảng 4 cho kết quả hệ số tương quan giữa các biến, mục đích của kiểm tra mỗi tương quan chặt chẽ giữa các biến độc lập và biến phụ thuộc để loại bớt các nhân tố có thể dẫn đến hiện tượng đa cộng tuyến trước khi chạy mô hình hồi quy. Kết quả cho thấy 05 nhân tố độc lập đều có hệ số Sig < 5% nên 05 nhân tố này đều

Bảng 3: Bảng phân tích nhân tố EFA của các biến

		Rotated Component Matrix ^a					
		Component					
		1	2	3	4	5	6
	RA1	0,621					
	RA2	0,512					
	RA3	0,624					
	RA4	0,647					
	RA5	0,787					
	TI1		0,634				
	TI2		0,556				
	TI3		0,659				
	TI4		0,667				
	TI5		0,698				
	AOI1			0,576			
	AOI2			0,511			
	AOI3			0,632			
	AOI4			0,675			
	CI1				0,552		
	CI2				0,643		
	CI3				0,667		
	CI4				0,786		
	CI5				0,745		
	GS1					0,534	
	GS2					0,723	
	GS4					0,654	
	GS5					0,550	
	BDA1						0,743
	BDA2						0,534
	BDA3						0,774

Nguồn: Kết quả phân tích dữ liệu của nhóm tác giả.

tương quan với biến phụ thuộc. Hệ số tương quan của 5 nhân tố là, RA: 0,321, TI: 0,365, AOI: 0,231, CT: 0,209, và GS: 0,215. Hệ số tương quan giữa 05 biến độc lập này trong mô hình không có cặp nào lớn hơn 0,8; do đó, khi sử dụng mô hình hồi quy, sẽ ít có khả năng gặp hiện tượng đa cộng tuyến. Điều này cho thấy biến phụ thuộc có sự tương quan tuyến tính với 5 nhân tố, các biến này có phân phối chuẩn.

Bảng 4: Ma trận tương quan giữa các thành phần

		BDA	RA	TI	AOI	CT	GS
BDA	Pearson Correlation	1,000	0,321	0,365	0,231	0,209	0,215
	Sig. (2-tailed)		0,000	0,000	0,000	0,000	0,000
	N	330	330	330	330	330	330
RA	Pearson Correlation	0,321	1,000	0,000	0,000	0,000	0,000
	Sig. (2-tailed)	0,000		1,000	1,000	1,000	1,000
	N	330	330	330	330	330	330
TI	Pearson Correlation	0,365	0,000	1,000	0,000	0,000	0,000
	Sig. (2-tailed)	0,000	1,000		1,000	1,000	1,000
	N	330	330	330	330	330	330
AOI	Pearson Correlation	0,231	0,000	0,000	1,000	0,000	0,000
	Sig. (2-tailed)	0,000	1,000	1,000		1,000	1,000
	N	330	330	330	330	330	330
CT	Pearson Correlation	0,209	0,000	0,000	0,000	1,000	0,000
	Sig. (2-tailed)	0,000	1,000	1,000	1,000		1,000
	N	330	330	330	330	330	330
GS	Pearson Correlation	0,215	0,000	0,000	0,000	1,000	0,000
	Sig. (2-tailed)	0,000	1,000	1,000	1,000		1,000
	N	330	330	330	330	330	330

Nguồn: Kết quả phân tích dữ liệu của nhóm tác giả.

4.4. Phân tích hồi quy tuyến tính đa biến

Để xác định, đo lường và đánh giá mức độ tác động của các nhân tố đến kiểm toán nội bộ, nhóm tác giả

sử dụng phương pháp hồi quy tuyến tính bội giữa 05 nhân tố ảnh hưởng thu được từ phân tích nhân tố khám phá và phân tích tương quan ở trên.

Bảng 5: Kết quả hồi quy đa biến

	Hệ số chưa chuẩn hoá		Hệ số chuẩn hoá	Giá trị t	Sig	Đa cộng tuyến	
	B	Độ lệch chuẩn	Beta			Độ chấp nhận	VIF
Hằng số	3,1901	0,043		7,247	0,000		
Lợi thế tương đối (RA)	0,216	0,023	0,201	4,415	0,000	0,731	1,128
Cơ sở hạ tầng công nghệ (TI)	0,305	0,042	0,289	4,243	0,000	0,621	1,311
Khả năng hấp thụ thông tin (AOI)	0,212	0,011	0,203	3,512	0,000	0,723	1,234
Cạnh tranh trong ngành (CI)	0,217	0,002	0,204	3,332	0,000	0,754	1,140
Hỗ trợ của Chính phủ (GS)	0,114	0,043	0,122	3,127	0,000	0,685	1,230
R ²							0,519
R ² hiệu chỉnh							0,483
Sig. F Change							0,000
Durbin-Watson							2,241

Nguồn: Kết quả phân tích dữ liệu của nhóm tác giả.

Theo kết quả hồi quy bảng 5, kết quả này cho giá trị R² hiệu chỉnh= 0,483; giá trị R² hiệu chỉnh cho biết rằng các biến độc lập trong mô hình có thể giải thích được 48,3% sự thay đổi của biến phụ thuộc. Đồng thời kết quả phân tích cho thấy hệ số phóng đại phương sai VIF rất nhỏ, đều nhỏ hơn 2, cho thấy các biến độc lập này không có quan hệ chặt chẽ với nhau nên không xảy ra hiện tượng đa cộng tuyến. Về kiểm định tính độc lập của phần dư chính là đại lượng thống kê Durbin –Watson của hàm hồi quy có giá trị 2,241 < 3 cho thấy không có hiện tượng tự tương quan chuỗi bậc 1 hay nói cách khác các phần dư ước lượng của mô hình độc lập không có mối quan hệ tuyến tính với nhau. Giá trị t tương ứng với Sig. của các biến độc lập đều nhỏ hơn 0,05 nên có ý nghĩa thống kê. Từ bảng 8 cho thấy 4 nhân tố tác động tích cực đến việc hoạt động kiểm toán. Phương trình hồi quy có dạng như sau:

$$BDA = 3,1901 + 0,216 \cdot RA + 0,305 \cdot TI + 0,212 \cdot AOI + 0,217 \cdot CI + 0,114 \cdot GS$$

Giả thuyết H1, H2, H3, H4, H5 được chấp nhận. Tức là lợi thế tương đối, cơ sở hạ tầng công nghệ, khả năng hấp thụ thông tin, cạnh tranh trong ngành, hỗ trợ của chính phủ đều có ảnh hưởng đến việc áp dụng phân tích dữ liệu lớn. Kết quả này tương đồng với các nghiên cứu trước như Williamson (2014), Agrawal (2015), Francisco & Swanson (2018), Dang-Pham & Nkhoma (2013), Tan & cộng sự (2015), Kumar & Kumar (2015), Ling & Wahab (2020), và Almoqren & Altayar (2016).

5. Kết luận và khuyến nghị

Trong 5 yếu tố ảnh hưởng đến việc áp dụng phân tích dữ liệu lớn thì yếu tố cơ sở hạ tầng công nghệ có ảnh hưởng lớn nhất. Trong khi đó, sự hỗ trợ của chính phủ có ảnh hưởng thấp nhất trong việc áp dụng phân tích dữ liệu lớn. Kết quả chỉ ra rằng 5 yếu tố cho thấy tác động tích cực đến việc áp dụng phân tích dữ liệu lớn.

Việt Nam đang sở hữu những lợi thế cho quá trình chuyển đổi số, bằng việc áp dụng dữ liệu lớn như: tỷ lệ dân số sở hữu điện thoại thông minh, máy tính cao. Mặt khác, các doanh nghiệp tại Việt Nam còn rất trẻ, họ có thể bắt đầu xây dựng cơ sở hạ tầng mà không bị ảnh hưởng từ hệ thống cũ tồn tại trước đó.

Tuy nhiên, bên cạnh một số ít doanh nghiệp đã triển khai và có những thành công nhất định trong việc ứng dụng dữ liệu lớn trong sản xuất – kinh doanh thì vẫn còn nhiều doanh nghiệp thương mại điện tử chưa thực sự tiếp cận được với công nghệ dữ liệu lớn, bởi nó đòi hỏi một nền tảng công nghệ thông tin mạnh, kho lưu trữ khổng lồ và các hỗ trợ của Chính phủ trong việc khai thác kho dữ liệu. Do đó, muốn sử dụng và khai thác dữ liệu lớn một cách hiệu quả, điều cần thiết phải xây dựng hạ tầng đủ để thu thập và lưu trữ dữ liệu, quản lý, tích hợp, phân tích dữ liệu, đồng thời, cung cấp quyền truy cập và bảo mật thông tin trong khi lưu trữ và chuyển tiếp.

Các doanh nghiệp tại Việt Nam nên có tư duy rõ ràng và hiểu biết sâu sắc hơn bằng cách cố gắng hiểu lợi ích của việc áp dụng phân tích dữ liệu lớn. Các doanh nghiệp phải thực hiện các chiến lược phù hợp nhất trong việc áp dụng phân tích dữ liệu lớn như một công cụ chiến lược, sau đó thúc đẩy hệ thống quản lý chuỗi cung ứng và hậu cần bền vững.

Các doanh nghiệp nên hợp tác với nhau thông qua các hiệp hội thương mại với tư cách là đại diện để tương tác với các cơ quan hữu quan. Những tương tác này liên quan đến việc tạo ra đầu vào cho chính phủ liên quan đến việc ra các chính sách kịp thời, thực tiễn.

Các doanh nghiệp nên tập trung vào việc đồng hóa cơ sở hạ tầng hiện tại với công nghệ phân tích dữ liệu lớn để cải thiện hoạt động kinh doanh. Việc chuyển sang phân tích dữ liệu lớn không chỉ là tổ chức và khai thác số lượng lớn tập dữ liệu; việc triển khai mô hình dự đoán bằng trí tuệ nhân tạo có thể cải thiện đáng kể năng suất của hoạt động của doanh nghiệp. Mặt khác, nó còn thúc đẩy tính hiệu quả như giảm chất thải và tiêu thụ năng lượng, đồng thời cho phép ngành công nghiệp tự thúc đẩy phát triển bền vững.

Tài liệu tham khảo

- Addo-Tenkorang, R., & Helo, P.T. (2016), 'Big data applications in operations/supply-chain management: A literature review', *Computers & Industrial Engineering*, 101, 528-543, DOI: <https://doi.org/10.1016/j.cie.2016.09.023>.
- Agrawal, K. (2015), 'Investigating the determinants of Big Data Analytics (BDA) adoption in Asian emerging economies', *Academy of Management Proceedings*, 2015(1), 11290-11290, DOI: 10.5465/AMBPP.2015.11290abstract.
- Almoqren, N., & Altayar, M. (2016), 'The motivations for big data mining technologies adoption in Saudi banks', In *2016 4th Saudi International Conference on Information Technology (Big Data Analysis) (KACSTIT)*, 1-8, DOI: 10.1109/KACSTIT.2016.7756075.
- Anuradha, J. (2015), 'A brief introduction on Big Data 5Vs characteristics and Hadoop technology', *Procedia Computer Science*, 48, 319-324, DOI: <https://doi.org/10.1016/j.procs.2015.04.188>.
- Attaran, M., Stark, J., & Stotler, D. (2018), 'Opportunities and challenges for big data analytics in US higher education: A conceptual model for implementation', *Industry and Higher Education*, 32(3), 169-182, DOI: <https://doi.org/10.1177/095042221877>.
- Chen, J., Chen, Y., Du, X., Li, C., Lu, J., Zhao, S., & Zhou, X. (2013), 'Big data challenge: a data management perspective', *Frontiers of Computer Science*, 7, 157-164, DOI: 10.1007/s11704-013-3903-7.
- Dang-Pham, Duy & Nkhoma, Mathews (2013), 'Contributing Factors of Cloud Computing Adoption: a Technology-Organisation-Environment Framework Approach', *International Journal of Information Systems and Engineering*, 1, 38-49, DOI: 10.24924/ijise/2013.04/v1.iss1/30.41.
- Erind, H. (2015), 'The technological, organizational and environmental framework of IS innovation adaption in small and medium enterprises. Evidence from research over the last 10 years', *International Journal of Business and Management*, 3(4), 1-14.
- Fan, J., Han, F., & Liu, H. (2014), 'Challenges of big data analysis', *National Science Review*, 1(2), 293-314, DOI: <https://doi.org/10.1093/nsr/nwt032>.
- Francisco, K., & Swanson, D. (2018), 'The supply chain has no clothes: Technology adoption of blockchain for supply chain transparency', *Logistics*, 2(1), 2, DOI: <https://doi.org/10.3390/logistics2010002>.
- Gandomi, A., & Haider, M. (2015), 'Beyond the hype: Big data concepts, methods, and analytics', *International Journal of Information Management*, 35(2), 137-144, DOI: <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>.
- Ghobakhloo, M., Arias-Aranda, D., & Benitez-Amado, J. (2011), 'Adoption of e-commerce applications in SMEs', *Industrial Management & Data Systems*, 111(8), 1238-1269, DOI: <https://doi.org/10.1108/02635571111170785>.
- Gobble, M.M. (2013), 'Big data: The next big thing in innovation', *Research-Technology Management*, 56(1), 64-67, DOI: <https://doi.org/10.5437/08956308X5601005>.
- Govindan, K., Cheng, T.E., Mishra, N., & Shukla, N. (2018), 'Big data analytics and application for logistics and supply chain management', *Transportation Research Part E: Logistics and Transportation Review*, 114, 343-349, DOI: <https://doi.org/10.1016/j.tre.2018.03.011>.
- Gunasekaran, A., Yusuf, Y.Y., Adeleye, E.O., & Papadopoulos, T. (2018), 'Agile manufacturing practices: the role of big data and business analytics with multiple case studies', *International Journal of Production Research*, 56(1-2), 385-397, DOI: <https://doi.org/10.1080/00207543.2017.1395488>.
- Gupta, M., & George, J.F. (2016), 'Toward the development of a big data analytics capability', *Information &*

Management, 53(8), 1049-1064, DOI: <https://doi.org/10.1016/j.im.2016.07.004>.

- Hurwitz, J.S., Nugent, A., Halper, F., & Kaufman, M. (2013), *Big data for dummies*, John Wiley & Sons.
- Kauffman, R.J., Srivastava, J., & Vayghan, J. (2012), 'Business and data analytics: New innovations in the management of e-commerce', *Electronic Commerce Research and Applications*, 11(2), 85-88, DOI: <http://doi.org/10.1016/j.elerap.2012.01.001>.
- Kumar, A., & Kumar, T.V. (2015), 'Big data and analytics: issues, challenges, and opportunities', *International Journal of Data Science*, 1(2), 118-138, DOI: <https://doi.org/10.1504/IJDS.2015.072412>.
- Ling, E.K., & Wahab, S.N. (2020), 'Integrity of food supply chain: going beyond food safety and food quality', *International Journal of Productivity and Quality Management*, 29(2), 216-232, DOI: <https://doi.org/10.1504/IJPM.2020.105963>.
- Neaga, I., Liu, S., Xu, L., Chen, H., & Hao, Y. (2015), 'Cloud enabled big data business platform for logistics services: A research and development agenda', Paper presented at the *Decision Support Systems V-Big Data Analytics for Decision Making: First International Conference*, ICDSSST 2015, Belgrade, Serbia, May 27-29, 2015, Proceedings 1, DOI: 10.1007/978-3-319-18533-0_3.
- Park, J.-H., Kim, M.-K., & Paik, J.-H. (2015), 'The factors of technology, organization and environment influencing the adoption and usage of big data in Korean firms', In *26th European Regional ITS Conference*, Madrid 2015 (No. 127173), International Telecommunications Society (ITS), <https://www.econstor.eu/bitstream/10419/127173>.
- Raut, R.D., Mangla, S.K., Narwane, V.S., Gardas, B.B., Priyadarshinee, P., & Narkhede, B.E. (2019), 'Linking big data analytics and operational sustainability practices for sustainable business management', *Journal of Cleaner Production*, 224, 10-24, DOI: <https://doi.org/10.1016/j.jclepro.2019.03.181>.
- Sangari, M.S., & Razmi, J. (2015), 'Business intelligence competence, agile capabilities, and agile performance in supply chain: An empirical study', *The International Journal of Logistics Management*, 26(2), 356-380, DOI: <https://doi.org/10.1108/IJLM-01-2013-0012>.
- Sivarajah, U., Kamal, M.M., Irani, Z., & Weerakkody, V. (2017), 'Critical analysis of Big Data challenges and analytical methods', *Journal of Business Research*, 70, 263-286, DOI: <https://doi.org/10.1016/j.jbusres.2016.08.001>.
- Tan, K.H., Zhan, Y., Ji, G., Ye, F., & Chang, C. (2015), 'Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph', *International Journal of Production Economics*, 165, 223-233, DOI: <https://doi.org/10.1016/j.ijpe.2014.12.034>.
- Van der Meulen, R., & Woods, V. (2015), 'Gartner survey shows more than 75 percent of companies are investing or planning to invest in big data in the next two years', *Gartner*, retrieved on Feb 10th 2025, from <<https://www.gartner.com/en/newsroom/press-releases/2014-09-17-gartner-survey-reveals-that-73-percent-of-organizations-have-invested-or-plan-to-invest-in-big-data-in-the-next-two-years>>.
- Wahab, S., Olugu, E., Lee, W., & Tan, S. (2018), 'Big data analytics adoption in Malaysia warehousing industry', In *The 32nd International Business Information Management Association Conference*, IBIMA 2018, 2349-2365, DOI: <https://doi.org/10.1016/j.jbusres.2016.08.009>.
- Wahab, S.N., Hamzah, M.I., Sayuti, N.M., Lee, W.C., & Tan, S.Y. (2021), 'Big data analytics adoption: an empirical study in the Malaysian warehousing sector', *International Journal of Logistics Systems and Management*, 40(1), 121-144, DOI: <https://doi.org/10.1504/IJLSM.2021.117703>.
- Wang, G., Gunasekaran, A., Ngai, E.W., & Papadopoulos, T. (2016), 'Big data analytics in logistics and supply chain management: Certain investigations for research and applications', *International Journal of Production Economics*, 176, 98-110, DOI: <https://doi.org/10.1016/j.ijpe.2016.03.014>.
- Williamson, J. (2014), *Getting a big data job for dummies*, John Wiley & Sons.
- Zhang, Y., Ren, S., Liu, Y., & Si, S. (2017), 'A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products', *Journal of Cleaner Production*, 142, 626-641, DOI: <https://doi.org/10.1016/j.jclepro.2016.07.123>.

*Tác giả liên hệ: Nguyễn Thị Quế. Email: nguyenthique@hau.edu.vn.