

KHAI PHÁ TRÍ TUỆ MARKETING TRONG THỜI ĐẠI DỮ LIỆU LỚN

Đàm Nguyễn Anh Khoa

Trường Đại học Bách Khoa – Đại học Đà Nẵng

Email: dnakhoea@dut.udn.vn

Nguyễn Văn Kỳ Long

Trường Đại học FPT

Email: longnvk@fe.edu.vn

Ngày nhận: 24/01/2020

Ngày nhận bản sửa: 20/4/2020

Ngày duyệt đăng: 05/5/2020

Tóm tắt:

Để sinh tồn và phát triển trong thời đại công nghệ ngày nay, chuyển đổi số là xu hướng tất yếu của mọi doanh nghiệp. Trong tiến trình này, xu thế khai thác trí tuệ marketing từ dữ liệu lớn là chìa khoá sống còn của doanh nghiệp. Việc ứng dụng các kỹ thuật khai phá dữ liệu để chuyển đổi dữ liệu lớn thành trí tuệ marketing giúp doanh nghiệp nâng cao hiểu biết về thị trường, khách hàng, sản phẩm và đối thủ cạnh tranh. Tuy nhiên, các nghiên cứu trước đây còn bỏ ngỏ vấn đề này, đặc biệt là việc ứng dụng các mô hình và kỹ thuật khai phá dữ liệu. Do đó, nghiên cứu này trình bày tổng quan các nghiên cứu về trí tuệ marketing trong thời đại dữ liệu lớn thông qua việc ứng dụng các mô hình và công cụ khai phá dữ liệu. Kết quả nghiên cứu đề xuất các mô hình, kỹ thuật khai phá dữ liệu cho từng thành phần của trí tuệ marketing. Bài viết này cũng đề xuất các hướng nghiên cứu khả thi trong tương lai.

Từ khóa: Trí tuệ marketing; Khai phá dữ liệu; Dữ liệu lớn; Tổng quan nghiên cứu.

Mã JEL: M31; M10.

Mining marketing intelligence in the age of big data

Abstract:

To survive in the age of digitalization, enterprises have no choice but to exploit marketing intelligence from big data. The application of data mining techniques to transform big data into marketing intelligence helps enterprises gain insights on market, customers, products, and competitors. However, literature has recognized a gap in how to apply data mining models and techniques to exploit marketing intelligence. To shed more light on this issue, this paper aims at conducting a literature review on marketing intelligence through data mining models and techniques. The findings of this paper discuss specific data mining techniques for each component of marketing intelligence on market, customers, products, and competitors. The study also suggests future research directions for scholars to enrich literature in this domain.

Keywords: Marketing intelligence, data mining, big data, literature review.

JEL code: M31; M10.

1. Giới thiệu

Trong thời đại bùng nổ dữ liệu ngày nay, chuyển đổi số là một tiến trình không thể tránh khỏi đối với doanh nghiệp (Dam & cộng sự, 2019, Verhoef & cộng sự, 2019). Theo đó, kỷ nguyên “Trí tuệ marketing” (*Marketing Intelligence* – từ đây viết tắt là MI) và

“Dữ liệu lớn” (*Big Data* - BD) đã mở ra những cơ hội tiềm năng cho doanh nghiệp (Sivarajah & cộng sự, 2017). BD được xem là nguồn lực to lớn cho MI với những đặc điểm về khối lượng, tốc độ, tính đa dạng và tính xác thực của dữ liệu (Janssen & cộng sự, 2017, Lau & cộng sự, 2016). Việc ứng dụng các

kỹ thuật Khai phá dữ liệu (*Data mining - DM*) từ BD để thu thập thông tin về khách hàng, thị trường, đối thủ cạnh tranh giúp doanh nghiệp có được MI và cải thiện chiến lược marketing trong thời đại số (Efrat & cộng sự, 2017, Huster, 2005).

Bên cạnh những cơ hội đầy tiềm năng trong kỷ nguyên dữ liệu lớn, việc khai thác MI từ BD sẽ phải đối mặt với một số thách thức. Trước hết, doanh nghiệp gặp khó khăn trong việc xác định các nguồn dữ liệu liên quan (Janssen & cộng sự, 2017, Baesens & cộng sự, 2016, Moges & cộng sự, 2013). Ngoài ra, doanh nghiệp phải đối mặt với các thách thức trong việc phân loại các thành phần khác nhau của MI với từng ứng dụng riêng biệt (Liang & Liu, 2018, Fan & cộng sự, 2015). Mặc dù việc ứng dụng BD trong marketing đã thu hút sự chú ý của rất nhiều các nhà học giả, các nghiên cứu đa phần bỏ ngỏ các nội dung liên quan đến việc sử dụng các mô hình DM vào DDL trong công tác marketing (Amado & cộng sự, 2018, Pan & cộng sự, 2013, Liao & cộng sự, 2012).

Theo đó, việc tổng quan các nghiên cứu về MI trong thời đại BD thông qua việc ứng dụng các mô hình và công cụ DM là thật sự cần thiết đối với các nhà học giả cũng như là đối với doanh nghiệp. Nghiên cứu này sẽ giúp bổ sung các khoảng trống lý thuyết về chủ đề MI trong kỷ nguyên bùng nổ thông tin hiện nay. Nghiên cứu này còn góp phần thu hẹp khoảng cách giữa hai lĩnh vực Marketing và Hệ thống Thông tin. Do đó, mục tiêu nghiên cứu của bài viết này hướng đến đánh giá tổng quan lý thuyết trong việc ứng dụng các mô hình DM từ BD để khai thác MI. Mục tiêu đầu tiên của nghiên cứu là phân loại MI. Mục tiêu thứ hai là xác định các nguồn BD thích hợp cho từng loại MI. Bên cạnh đó, bài viết còn tổng hợp các nghiên cứu trước nhằm đề xuất mô hình và công cụ DM phù hợp tương ứng cho từng loại MI với mục đích hỗ trợ cho việc nghiên cứu tổng quan các nghiên cứu trước đây về khai phá trí tuệ marketing.

Bài nghiên cứu được trình bày như sau: Tiếp theo phần đặt vấn đề là mục 2 đề cập cơ sở lý thuyết về MI và DM. Mục 3 trình bày phương pháp nghiên cứu bao gồm thu thập và thống kê dữ liệu. Kết quả phân tích, phân loại MI với các mô hình và kỹ thuật DM được trình bày ở mục 4. Cuối cùng, mục 5 nêu kết luận với phần thảo luận chuyên sâu về các đóng góp cũng như các hướng nghiên cứu quan trọng trong tương lai.

2. Tổng quan nghiên cứu

2.1. Trí tuệ marketing

Tổng quan lý thuyết trước đây chưa thực sự đưa ra một định nghĩa hay khái niệm thống nhất về MI (Lies, 2019). Điểm khác biệt cơ bản giữa định nghĩa truyền thống và định nghĩa hiện đại về MI nằm ở phương pháp thu thập thông tin (Huster, 2005, Lau & cộng sự, 2012). Theo cách tiếp cận truyền thống, MI phụ thuộc vào khảo sát thị trường và các nguồn thông tin nội bộ trong doanh nghiệp để thu thập thông tin về khách hàng, đối thủ cạnh tranh, thị trường và về ngành (Huster, 2005, Kohli & Jaworski, 1990, Xu & cộng sự, 2016). Ngày nay, định nghĩa về MI được hình thành từ việc áp dụng các mô hình và kỹ thuật DM nhằm phục vụ cho các quyết định chiến lược của doanh nghiệp (Chen & cộng sự, 2012, Lau & cộng sự, 2016). Theo cách tiếp cận này, MI được định nghĩa là quá trình thu thập thông tin về khách hàng, đối thủ, thị trường và ngành thông qua các mô hình DM và sau đó được áp dụng vào các kế hoạch marketing chiến lược của doanh nghiệp (Efrat & cộng sự, 2017, Huster, 2005).

Nghiên cứu này kế thừa quan điểm từ nhiều nghiên cứu trước đây để định nghĩa MI như là một ứng dụng của các mô hình và kỹ thuật DM để khám phá MI về thị trường, sản phẩm, khách hàng và đối thủ cạnh tranh (Chen & cộng sự, 2012, Lau & cộng sự, 2016). Các quan điểm này dựa trên lý thuyết Marketing hỗn hợp (*Marketing Mix*); vì vậy, định nghĩa này hầu như bao quát tất cả các khía cạnh quan trọng nhằm mục đích hỗ trợ các quyết định marketing (Fan & cộng sự, 2015, Lau & cộng sự, 2016). Tuy nhiên, quan điểm của *Marketing Mix* truyền thống, 4P bao gồm sản phẩm (*product*), giá cả (*price*), khuyến mãi (*promotion*) và địa điểm (*place*) bị phê bình là chỉ định hướng sản phẩm và thiếu sự tập trung vào khách hàng. Do đó, nghiên cứu này bổ sung P thứ năm (*people*) tập trung vào khách hàng (Fan & cộng sự, 2015, Lau & cộng sự, 2016).

2.1.1. Trí tuệ marketing về thị trường

MI về thị trường bao gồm thông tin về các yếu tố ngoại sinh có thể ảnh hưởng đến nhu cầu và sở thích khách hàng hiện tại và tương lai như công nghệ, sự cạnh tranh, pháp luật và các yếu tố khác tác động từ môi trường bên ngoài (Sorjonen & Uusitalo, 2003, Aggarwal & Singh, 2004). Có thể tóm lại, MI về thị trường bao gồm các loại MI từ chính trị - kinh tế đến văn hóa - xã hội (Lau & cộng sự, 2012, Navarro-García & cộng sự, 2016, Trim & Lee, 2008). Trước

đây, các nguồn thông tin truyền thống của MI trí tuệ thu thập được từ khảo sát, báo cáo kinh doanh, báo cáo thảo luận với khách hàng, nghiên cứu thị trường...(Sorjonen & Uusitalo, 2003, Aggarwal & Singh, 2004). Ngày nay, các nguồn thông tin mở thường được đề cập đến như nhật ký trực tuyến, sách trắng để thu thập thông tin MI về thị trường (Fleisher, 2008).

2.1.2. Trí tuệ marketing về sản phẩm

Hầu hết các định nghĩa MI về sản phẩm được định nghĩa theo quan điểm của sản phẩm thông minh (Gaham & Bouzouia, 2009, Rijdsdijk & cộng sự, 2007). Ngày nay, MI về sản phẩm ứng dụng các kỹ thuật DM để khai thác hiểu biết sâu sắc về sản phẩm nhằm tăng sự hài lòng của khách hàng và xác định các cơ hội kinh doanh (Amarouche & cộng sự, 2015, Fan & cộng sự, 2015). Một trong những cách tốt nhất để thỏa mãn nhu cầu của khách hàng là lắng nghe đánh giá, thảo luận, thái độ của khách hàng trên các diễn đàn, phương tiện truyền thông xã hội, blog và trang web (Abrahams & cộng sự, 2012, Gutt & cộng sự, 2019). Việc khai thác các nội dung do khách hàng tạo ra và nội dung về sản phẩm trên các website sẽ cho phép doanh nghiệp không chỉ đưa ra quyết định phát triển các sản phẩm phù hợp với nhu cầu của khách hàng mà còn giới thiệu sản phẩm tương ứng đến đúng khách hàng tiềm năng (Abrahams & cộng sự, 2013, Park & cộng sự, 2012).

2.1.3. Trí tuệ marketing về đối thủ cạnh tranh

MI về đối thủ cạnh tranh bao gồm các thông tin về các sản phẩm, giá cả, quảng cáo và kênh phân phối của đối thủ cạnh tranh (Navarro-García & cộng sự, 2016). Hơn thế nữa, MI giúp doanh nghiệp nắm bắt được điểm mạnh và điểm yếu của đối thủ cạnh tranh, từ đó dự đoán trước động thái và chiến lược của đối thủ (Wright & cộng sự, 2002). Để thu thập thông tin về đối thủ cạnh tranh, doanh nghiệp có thể thu thập nhật ký dữ liệu từ các trang web thương mại điện tử (Fan & cộng sự, 2015). Doanh nghiệp có thể dùng các thông tin trong nhật ký dữ liệu này như thứ hạng doanh số sản phẩm, giá niêm yết hay ngày phát hành để dự báo nhu cầu thị trường, ước tính chi phí và độ co giãn của giá (Fan & cộng sự, 2015). Ngày nay, không chỉ các văn bản mà cả hình ảnh với các thuộc tính như định dạng hiển thị, chất lượng hình ảnh, số lượt xem cũng có thể ảnh hưởng đến ý định của người mua (Di & cộng sự, 2014, Sukumaran & Sureka, 2006).

2.1.4. Trí tuệ marketing về khách hàng

MI về khách hàng bao gồm thông tin về nhu cầu, sở thích, văn hóa, lối sống, sức mua, hành vi mua sắm, phong tục và thói quen của khách hàng tiềm năng (Navarro-García & cộng sự, 2016). Trong thời đại kỹ thuật số, MI về khách hàng được khai thác đầu tiên dựa trên các trang web với các tìm kiếm giao thức hướng dữ liệu thông qua *cookie* và nhật ký *server* (Chen & cộng sự, 2012, Doan & cộng sự, 2011). Các nhà tiếp thị còn có thể phân tích lịch sử *click* chuột của khách hàng với thông tin về tần suất truy cập, các sản phẩm đã xem và thời gian truy cập trên một trang web để nắm bắt thói quen lướt web và hành vi mua hàng của khách hàng (Fan & cộng sự, 2015, Park & cộng sự, 2012). Doanh nghiệp có thể khai thác MI về khách hàng từ các nguồn nội bộ như hồ sơ thanh toán, nhật ký web của công ty, hệ thống Quản lý quan hệ khách hàng (Rygielski & cộng sự, 2002). Ngoài ra, còn có các nguồn thông tin bên ngoài như trên phương tiện truyền thông xã hội, trang web của đối thủ cạnh tranh, hay điểm tín dụng FICO (Fan & cộng sự, 2015, Liang & Liu, 2018, Rygielski & cộng sự, 2002).

2.2. Khai phá dữ liệu

Sáu mô hình DM điển hình với các kỹ thuật liên quan được trình bày như sau (Chen & cộng sự, 2012, Seng & Chen, 2010, Sivarajah & cộng sự, 2017):

Mô hình Phân loại được sử dụng để đưa ra dự báo về hành vi của khách hàng (Fan & cộng sự, 2015) hoặc xác định các thuộc tính của các cụm dữ liệu (Bose & Mahapatra, 2001). Các kỹ thuật phân loại thường được sử dụng như là *neural networks* (mạng nơ-ron), *Decision trees* (mô hình cây quyết định), *Naïve Bayes*, *Support Vector Machines*, *market basket analysis* (phân tích giỏ hàng), *genetic algorithms* (thuật toán di truyền), và điều kiện *if-then-else* (Ngai & cộng sự, 2011, Ngai & cộng sự, 2009).

Mô hình Kết hợp được sử dụng để tìm hiểu mối quan hệ giữa các sản phẩm mà khách hàng mua; vì vậy, các doanh nghiệp có thể xác định các sản phẩm có xu hướng bổ sung cho nhau (Bose & Mahapatra, 2001, Seng & Chen, 2010). Các kỹ thuật kết hợp thường được sử dụng như *association rules* (luật kết hợp), thống kê và thuật toán Apriori (Chen & cộng sự, 2012, Ngai & cộng sự, 2009).

Mô hình Phân cụm được sử dụng để phân khúc khách hàng và phân loại người dùng (Fan & cộng sự, 2015, Hosseini & cộng sự, 2010). Các kỹ thuật

Bảng 1. Thống kê số lượng bài báo khoa học theo tên tạp chí

Tên tạp chí khoa học	Số lượng	Tỷ lệ %
Expert Systems with Applications	10	18.18%
Decision Support Systems	7	12.73%
Journal of Business Research	4	7.27%
IEEE	3	5.45%
Information and Management	3	5.45%
MIS Quarterly	3	5.45%
European Journal of Marketing	2	3.64%
Marketing Intelligence & Planning	2	3.64%
Journal of the Academy of Marketing Science	1	1.82%
Journal of Marketing	1	1.82%
Others	19	34.55%
Total	55	100.00%

Nguồn: Tính toán của tác giả.

phân cụm phổ biến nhất là K-means, Naïve Bayes, mô hình phân tích RFM (tính chất mới xảy ra, tần suất và mức độ chi tiền), phân tích giỏ hàng, mạng nơ-ron và *self-organizing map* (hay còn gọi là kỹ thuật phân cụm SOM) (Hosseini & cộng sự, 2010, Seng & Chen, 2010, Sivarajah & cộng sự, 2017).

Mô hình Hồi quy được sử dụng để đưa ra dự báo hoặc tìm mối quan hệ nhân quả giữa các biến (Ngai & cộng sự, 2009). Các kỹ thuật hồi quy phổ biến được sử dụng là hồi quy tuyến tính và hồi quy logistic (Seng & Chen, 2010, Sivarajah & cộng sự, 2017).

Mô hình Dự báo được sử dụng để dự báo các giá trị trong tương lai dựa trên các ghi chép lịch sử (Bose & Mahapatra, 2001, Enke & Thawornwong, 2005). Các kỹ thuật dự báo phổ biến nhất được sử dụng là phân tích giỏ hàng, mạng nơ-ron, *survival analysis* (kỹ thuật phân tích sống còn), hồi quy tuyến tính và hồi quy logistic (Enke & Thawornwong, 2005, Seng & Chen, 2010).

Mô hình Khám phá chuỗi được sử dụng để xác định các sự kết hợp hoặc mô tả thứ tự các hành vi theo thời gian (Rygielski & cộng sự, 2002). Kỹ thuật khám phá chuỗi thường được sử dụng là kỹ thuật thống kê và lý thuyết tập hợp (Ngai & cộng sự, 2009).

3. Phương pháp nghiên cứu

3.1. Thu thập dữ liệu

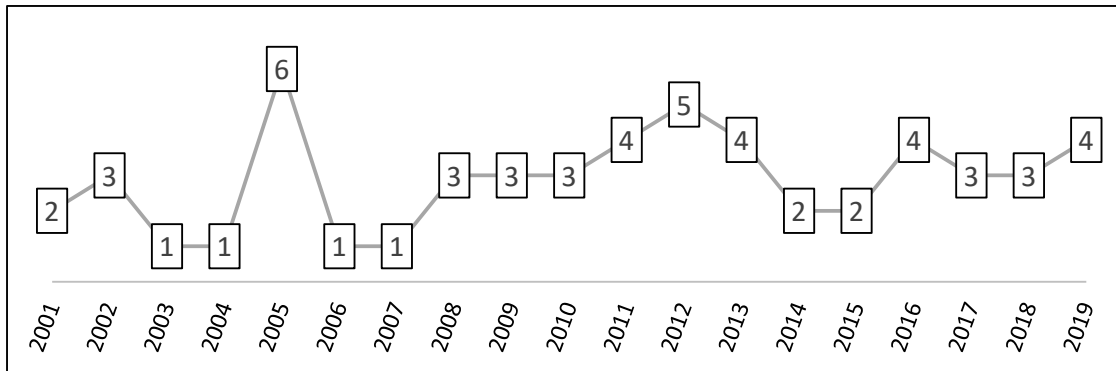
Các nghiên cứu khoa học về MI và DM nằm trong nhiều cơ sở dữ liệu khác nhau. Do đó, nghiên cứu

này sẽ xây dựng tổng quan tài liệu nghiên cứu từ các nguồn học thuật uy tín khác nhau như Science Direct, Emerald, Business Source Premier, EBSCOhost, ProQuest, Google Scholarship và IEEE Transaction (Ngai & cộng sự, 2009). Các từ khóa được sử dụng như “*marketing intelligence*” (trí tuệ marketing), hoặc “*data-driven marketing*” (*marketing* định hướng dữ liệu), và “*data mining techniques*” (kỹ thuật DM) để tìm kiếm các bài báo từ các nguồn dữ liệu đáng tin cậy trên. Các bài báo khoa học sẽ được thu thập từ các tạp chí Marketing và tạp chí Quản lý hàng đầu thế giới như Journal of Business Research, European Journal of Marketing, Journal of the Academy of Marketing Science, Journal of Marketing,... (Amado & cộng sự, 2018). Ngoài ra, các bài báo khoa học từ các tạp chí về Hệ thống thông tin cũng được nghiên cứu tổng hợp cùng như: Expert Systems with Applications, MIS Quarterly, Decision Support Systems, IEEE, Information and Management, Decision Support Systems,... (Chen & cộng sự, 2012). Để đảm bảo tính hợp lệ và độ tin cậy trong quá trình tìm kiếm tài liệu, kỹ thuật tìm kiếm *Forward-Backward* được tiến hành nhằm đảm bảo rằng 55 bài báo khoa học được chọn có thể đại diện cho nghiên cứu trong lĩnh vực này (Webster & Watson, 2002).

3.2. Thống kê dữ liệu

Tất cả các tạp chí đã lựa chọn ở trên đều nằm trong danh sách của Scimago Journal & Country Rank năm 2018. Bảng 1 cho thấy tỷ lệ của các bài báo khoa học được nghiên cứu đến từ tạp chí Expert

Hình 1. Thống kê bài báo theo năm xuất bản



Nguồn: Tính toán của tác giả.

Systems with Applications là khá lớn (18.18%), tiếp theo là tạp chí Decision Support Systems (12.73%), và Journal of Business Research (7.27%). Một số tạp chí khoa học khác cũng được tham khảo nhiều là IEEE, Information and Management, MIS Quarterly.

Nghiên cứu này cũng thống kê bài viết theo năm xuất bản và thể hiện theo Hình 1. Trong Hình 1, có thể thấy số lượng bài báo đã tăng từ năm 2001 đến 2019 với số lượng cao nhất vào năm 2005. Giai đoạn từ năm 2008 đến 2012 cho thấy số lượng bài viết trong lĩnh vực này tăng dần. Tuy nhiên, có sự biến động về số lượng bài báo từ 2013 đến 2019.

4. Kết quả và thảo luận

4.1. Khai phá Trí tuệ marketing về thị trường

MI về thị trường có thể được chia thành hai loại: kinh tế - chính trị và văn hóa – xã hội (Navarro-García & cộng sự, 2016, Trim & Lee, 2008). Có được MI về thị trường, doanh nghiệp có thể đưa ra các quyết định chiến lược như có nên thâm nhập vào thị trường mới hay không, sản phẩm hiện tại có phù hợp với yếu tố văn hóa của thị trường mục tiêu hay không... Các mô hình phân cụm, dự báo, kết hợp, phân loại và hồi quy thường được sử dụng để DM thông tin kinh tế - chính trị và văn hóa - xã hội (Chen & Zimbra, 2010, Chung & cộng sự, 2005). Một loạt các kỹ thuật khai phá cũng được áp dụng kèm với các mô hình này nhằm đáp ứng các mục đích của doanh nghiệp. Ví dụ: một số kỹ thuật được sử dụng như *sentiment and effect analysis* (phân tích tâm lý); cây quyết định, mô hình SVM và mô hình hồi quy logistic để dự báo và khám phá các mối tương quan ẩn; *genetic algorithms* để tối ưu hóa tìm kiếm thông tin về thị trường trên web (Bai, 2011, Chen & Zimbra, 2010, Chung & cộng sự, 2005).

4.2. Khai phá Trí tuệ marketing về khách hàng

Dựa trên công tác quản lý quan hệ khách hàng, MI về khách hàng bao gồm bốn cấp độ: Nhận diện khách hàng; Thu hút sự chú ý của khách hàng; Duy trì quan hệ khách hàng; Phát triển giá trị thặng dư (Hosseini & cộng sự, 2010, Ngai & cộng sự, 2009).

4.2.1. Nhận diện khách hàng

MI về khách hàng khởi đầu bằng việc xác định phân khúc nhóm khách hàng có cùng sở thích và khả năng sinh lợi tương đương (France & Ghose, 2018, Ngai & cộng sự, 2009). Theo đó, sử dụng mô hình phân cụm và phân loại sẽ rất hữu ích trong việc phân chia khách hàng thành các nhóm đồng nhất và từ đó xây dựng hồ sơ khách hàng (Amado & cộng sự, 2018, Fan & cộng sự, 2015). Cụ thể hơn, hồ sơ khách hàng sẽ chứa thông tin về các nhóm theo nhân khẩu học (tuổi, giới tính), hành vi mua (nhu cầu, sức mua, sở thích, lối sống), thuộc tính mua (tính chất mới xảy ra, tần suất, kích thước), danh mục sản phẩm, và giá trị vòng đời của khách hàng (Baars & Kemper, 2008, France & Ghose, 2018). Ở giai đoạn này, mô hình cây quyết định trong phân loại và kỹ thuật K-means trong phân cụm là phù hợp để phân khúc khách hàng có đặc điểm tương tự nhau (Hosseini & cộng sự, 2010). Sau đó, phương pháp phân tích khách hàng mục tiêu sẽ được sử dụng để chọn phân khúc sinh lợi nhất (Woo & cộng sự, 2005).

4.2.2. Thu hút sự chú ý của khách hàng

Với mục đích thu hút sự chú ý các phân khúc mục tiêu, ở bước này, mô hình phân loại là phù hợp nhất (Ngai & cộng sự, 2009). Mô hình hồi quy và mô hình phân cụm cũng thường được sử dụng cho mục đích này. Nếu xem xét kỹ hơn về mô hình phân loại, các kỹ thuật Naïve Bayes, cây quyết định, thuật toán

di truyền và mạng nơ-ron là những kỹ thuật phổ biến nhất được sử dụng (Baesens & cộng sự, 2016, Chen & cộng sự, 2012, Ngai & cộng sự, 2009). Thêm vào đó, phân tích RFM có thể được áp dụng để nắm bắt được hành vi mua của khách hàng và cải thiện chiến lược marketing để thu hút khách hàng (Hosseini & cộng sự, 2010).

4.2.3. Duy trì quan hệ khách hàng

Các chiến lược marketing được sử dụng phát triển hồ sơ khách hàng, nâng cao lòng tin của khách hàng, phát triển hệ thống đề xuất cho sản phẩm công ty hoặc chương trình khách hàng thân thiết nhằm mục đích tăng sự hài lòng của khách hàng và duy trì mối quan hệ với khách hàng lâu dài (Payne & Frow, 2005, Rygielski & cộng sự, 2002). Theo đó, các phương pháp DM khác nhau được áp dụng để hỗ trợ các hoạt động trên như phân loại, kết hợp, phân cụm, khám phá chuỗi và hồi quy (Ngai & cộng sự, 2009). Theo đó, luật kết hợp, cây quyết định, mạng nơ-ron, mô hình hồi quy logistic và thuật toán di truyền được ứng dụng nhiều (Liao & cộng sự, 2012, Ngai & cộng sự, 2009, Sivarajah & cộng sự, 2017).

4.2.4. Phát triển giá trị thặng dư

Với mục tiêu tối đa hóa việc tạo ra giá trị thặng dư cho doanh nghiệp, phát triển giá trị thặng dư bao gồm ba chiến lược chính: *up/cross-selling* (*Up-selling* là bán những hàng hóa có giá cao hơn hàng hóa mà khách hàng dự định mua, còn *cross-selling* là bán thêm cho khách hàng những hàng hóa có liên quan khác), giá trị vòng đời khách hàng và phân tích giỏ hàng (Ngai & cộng sự, 2009, Payne & Frow, 2005). Đối với hai chiến lược *up/cross-selling* và phân tích giỏ hàng, mô hình khám phá chuỗi và mô hình kết hợp được sử dụng (Baars & Kemper, 2008, Shaw & cộng sự, 2001), các kỹ thuật DM phổ biến kèm theo là luật kết hợp và mạng nơ-ron (Ngai & cộng sự, 2009). Ngoài ra, để ước tính giá trị vòng đời của khách hàng, các nhà khoa học thường áp dụng các mô hình DM khác nhau như phân loại, phân cụm, dự báo và hồi quy (Ngai & cộng sự, 2009, Seng & Chen, 2010). Theo đó, các kỹ thuật DM tương ứng là mạng nơ-ron, luật kết hợp, hồi quy tuyến tính, *survival analysis*, chuỗi Markov (Chen & Zimbra, 2010, Ngai & cộng sự, 2009).

4.3. Khai phá Trí tuệ marketing về sản phẩm

Dựa trên định nghĩa được đề xuất trong phần MI về sản phẩm, có thể thấy rằng MI về sản phẩm bao gồm hai phương diện: phát triển sản phẩm (Amarouche & cộng sự, 2015, France & Ghose,

2018) và đề xuất sản phẩm (Albadvi & Shahbazi, 2009, Park & cộng sự, 2012).

4.3.1. Phát triển sản phẩm

Các đặc tính sản phẩm có thể được trích xuất thông qua các phương pháp khai thác văn bản kết hợp với các mô hình DM như phân loại, kết hợp và phân cụm (Fan & cộng sự, 2015, France & Ghose, 2018). Đặc tính sản phẩm có thể là các dữ liệu về kích thước, trọng lượng, màu sắc, bao bì và chủng loại của sản phẩm (Albadvi & Shahbazi, 2009, Efrat & cộng sự, 2017). Các kỹ thuật khác nhau như kỹ thuật khai thác ý kiến (*Opinion mining*), mô hình hóa chủ đề (*Topic modeling*), hệ thống trả lời câu hỏi (*question-answering*), kỹ thuật trích xuất thông tin được thực hiện tùy thuộc theo các mục tiêu nghiên cứu khác nhau (Chen & cộng sự, 2012). Cụ thể hơn, *Topic modeling* phù hợp nhất để tìm chủ đề chính, trong khi *question-answering* là ứng dụng của quá trình xử lý ngôn ngữ tự nhiên để xây dựng mô hình sản phẩm (*Product ontologies*) nhằm mục đích hỗ trợ các tương tác giữa người và máy tính (Amarouche & cộng sự, 2015, France & Ghose, 2018). Các kỹ thuật còn lại như kỹ thuật khai thác ý kiến và phân tích tâm lý được áp dụng xác định thái độ và cảm xúc của khách hàng (Lau & cộng sự, 2014, Li & Li, 2013).

4.3.2. Đề xuất sản phẩm

Phương diện này của MI về sản phẩm nhằm mục đích làm hài lòng từng khách hàng bằng cách đưa ra các đề xuất riêng cho mỗi khách hàng thông qua việc phân tích dữ liệu từ lịch sử *click* chuột, hồ sơ khách hàng, nhật ký điện thoại và các giao dịch (Park & cộng sự, 2012). Trong công tác nghiên cứu, các mô hình DM khác nhau được sử dụng để xây dựng các hệ thống đề xuất sản phẩm như mô hình Kết hợp, Phân loại, Phân cụm và Hồi quy (Adomavicius & Tuzhilin, 2005). Theo đó, các kỹ thuật DM liên quan đi kèm là K-Recent Neighbor, phân loại Bayestian, luật kết hợp, cây quyết định, mạng nơ-ron, phân tích mối liên kết, hồi quy tuyến tính (Adomavicius & Tuzhilin, 2005, Park & cộng sự, 2012). Đặc biệt, *k-Nearest neighbor* là một trong những kỹ thuật quan trọng và phù hợp nhất trong việc xây dựng hệ thống đề xuất sản phẩm cho người tiêu dùng (Albadvi & Shahbazi, 2009).

4.4. Khai phá Trí tuệ marketing về đối thủ cạnh tranh

Như đã đề cập ở phần MI về đối thủ cạnh tranh, MI về đối thủ cũng bao gồm 4P của đối thủ theo *Marketing Mix* – sản phẩm, giá cả, khuyến mãi và

địa điểm.

4.4.1. Trí tuệ marketing về sản phẩm của các đối thủ cạnh tranh

Các nhà khoa học dữ liệu khai phá nội dung web với các mô hình phân cụm, kết hợp và các kỹ thuật xây dựng mô hình sản phẩm nhằm mục đích theo dõi thông tin về sản phẩm của các đối thủ cạnh tranh (Amarouche & cộng sự, 2015, France & Ghose, 2018). Thông qua việc ứng dụng luật kết hợp, không chỉ các tính năng, dịch vụ kèm theo của sản phẩm mà còn các mối đe dọa đến từ các sản phẩm thay thế có thể được thu thập từ đánh giá của khách hàng, xếp hạng sản phẩm và mô tả sản phẩm (Amarouche & cộng sự, 2015, France & Ghose, 2018, Navarro-García & cộng sự, 2016).

4.4.2. Trí tuệ marketing về giá cả của đối thủ cạnh tranh

MI giá cả của đối thủ cạnh tranh bao gồm chiến lược giá, chính sách chiết khấu của đối thủ cạnh tranh (Navarro-García & cộng sự, 2016). Mô hình hồi quy thường được sử dụng nghiên cứu chiến lược giá (Fan & cộng sự, 2015, Lau & cộng sự, 2016). Đặc biệt, kỹ thuật hồi quy đa biến được áp dụng để xác định các yếu tố quyết định giá đối thủ cạnh tranh (Fan & cộng sự, 2015). Các mô hình kết hợp cũng hữu ích để xác định các đối thủ tiềm năng, chiến lược giá và chính sách chiết khấu của đối thủ (Fan & cộng sự, 2015).

4.4.3. Trí tuệ marketing về khuyến mãi của đối thủ cạnh tranh

Để có được MI về chiến lược khuyến mãi của đối thủ cạnh tranh, doanh nghiệp cần lấy dữ liệu về thời gian khuyến mãi, kiểu loại khuyến mãi và doanh số của đối thủ cạnh tranh (Park & cộng sự, 2012). Mô hình phù hợp nhất để phân tích tìm hiểu mối quan hệ giữa các yếu tố trong chiến lược quảng cáo của đối thủ cạnh tranh chính là mô hình hồi quy, đặc biệt là các kỹ thuật hồi quy tuyến tính (Fan & cộng sự, 2015, Lau & cộng sự, 2016). Các kỹ thuật DM phổ biến khác để xây dựng các hệ thống đề xuất là *K-Nearest Neighbor*, luật kết hợp, phân tích mối liên kết (Park & cộng sự, 2012).

4.4.4. Trí tuệ marketing về địa điểm của đối thủ cạnh tranh

Ngày nay, doanh nghiệp có thể có được MI về địa điểm của đối thủ cạnh tranh bằng cách thu thập dữ liệu về các địa điểm khách hàng sử dụng thông qua dịch vụ định vị (Scellato & cộng sự, 2011). Dữ

liệu vị trí này có được thông qua các thiết bị di động có định vị GPS, WiFi, GSM, check-in từ các mạng xã hội (Pan & cộng sự, 2013, Scellato & cộng sự, 2011). Về mặt marketing dưới góc độ về địa điểm, mô hình phân loại, dự báo và hồi quy là các mô hình DM hữu hiệu nhất (Fan & cộng sự, 2015, Pan & cộng sự, 2013). Theo đó, các kỹ thuật DM khác nhau được thực hiện như hồi quy tuyến tính/phi tuyến tính, Naïve Bayes, *mạng nơ-ron* và SVM (Pan & cộng sự, 2013, Scellato & cộng sự, 2011).

5. Kết luận

Với mục đích đẩy mạnh chuyển đổi số và tăng lợi thế cạnh tranh của doanh nghiệp, nghiên cứu này trình bày tổng quan các nghiên cứu trước đây về chủ đề khai thác MI thông qua việc ứng dụng các mô hình và kỹ thuật DM. Theo cách tiếp cận này, các thành phần cơ bản của MI được phân loại cơ bản dựa trên lý thuyết Marketing hỗn hợp. Các nguồn dữ liệu liên quan, mô hình và kỹ thuật DM cũng đã được đề xuất cho từng thành phần của MI. Bài viết này xây dựng tổng quan lý thuyết chuyên sâu dựa trên 55 nghiên cứu khoa học liên quan từ các cơ sở dữ liệu khác nhau. Kết quả nghiên cứu đề xuất các mô hình và kỹ thuật khai phá cho từng loại MI tương ứng. Nghiên cứu này cũng cho thấy MI về sản phẩm, khách hàng và khuyến mãi dường như thu hút sự chú ý nhiều hơn so với MI về giá cả và địa điểm (France & Ghose, 2018). Do đó, các học giả có thể khai mở các hướng nghiên cứu về giá cả và địa điểm thêm trong tương lai.

Nghiên cứu này có những đóng góp nhất định đối với doanh nghiệp và giới học thuật. Ở góc độ doanh nghiệp, các kỹ thuật DM tương ứng với từng loại MI sẽ tăng năng lực cạnh tranh cho doanh nghiệp. Cụ thể, MI sẽ giúp doanh nghiệp tiếp cận thị trường tiềm năng, đối thủ cạnh tranh, sản phẩm và khách hàng một cách hiệu quả hơn (Efrat & cộng sự, 2017, Lau & cộng sự, 2016). Ở góc độ học thuật, nghiên cứu này góp phần thu hẹp khoảng cách giữa hai lĩnh vực Marketing và Hệ thống Thông tin. Nghiên cứu này sẽ là nguồn tham khảo cho cả doanh nghiệp và học giả nghiên cứu. Các học giả nghiên cứu có thể sử dụng nghiên cứu này để mở ra những hướng đi mới trong khai phá MI.

Tài liệu tham khảo:

- Abrahams, A. S., Jiao, J., Fan, W., Wang, G. A. & Zhang, Z. 2013, What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings, *Decision Support Systems*, 55, 871-882.
- Abrahams, A. S., Jiao, J., Wang, G. A. & Fan, W. 2012, Vehicle defect discovery from social media, *Decision Support Systems*, 54, 87-97.
- Adomavicius, G. & Tuzhilin, A. 2005, Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, *IEEE Transactions on Knowledge Data Engineering*, 734-749.
- Aggarwal, N. & Singh, R. 2004, Market orientation in Indian organizations: an empirical study, *Marketing Intelligence & Planning*.
- Albadvi, A. & Shahbazi, M. 2009, A hybrid recommendation technique based on product category attributes, *Expert Systems with Applications*, 36, 11480-11488.
- Amado, A., Cortez, P., Rita, P. & Moro, S. 2018, Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis, *European Research on Management and Business Economics*, 24, 1-7.
- Amarouche, K., Benbrahim, H. & Kassou, I. 2015, Product Opinion Mining for Competitive Intelligence, *Procedia Computer Science*, 73, 358-365.
- Baars, H. & Kemper, H.-G. 2008, Management Support with Structured and Unstructured Data—An Integrated Business Intelligence Framework, *Information Systems Management*, 25, 132-148.
- Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J. & Zhao, J. L. 2016, Transformational issues of Big data and Analytics in Networked business, *MIS Quarterly*, 40.
- Bai, X. 2011, Predicting consumer sentiments from online text, *Decision Support Systems*, 50, 732-742.
- Bose, I. & Mahapatra, R. K. 2001, Business data mining—a machine learning perspective, *Information & Management*, 39, 211-225.
- Chen, H., Chiang, R. H. L. & Storey, V. C. 2012, Business Intelligence and Analytics - From Big Data to Big impact, *MIS Quarterly*, 36, 1165-1188.
- Chen, H. & Zimbra, D. 2010, AI and opinion mining, *IEEE Intelligent Systems*, 25, 74-80.
- Chung, W., Chen, H. & Nunamaker Jr, J. F. 2005, A visual framework for knowledge discovery on the Web: An empirical study of business intelligence exploration, *Journal of Management Information Systems*, 21, 57-84.
- Dam, N. A. K., Le Dinh, T. & Menvielle, W. 2019, A systematic literature review of big data adoption in internationalization, *Journal of Marketing Analytics*, 7, 182-195.
- Di, W., Sundaresan, N., PIRAMUTHU, R. & BHARDWAJ, A. 2014. Is a picture really worth a thousand words?:-on the role of images in e-commerce. *Proceedings of the 7th ACM international conference on Web search and data mining*, 2014. ACM, 633-642.
- Doan, A., Ramakrishnan, R. & Halevy, A. Y. 2011, Crowdsourcing systems on the world-wide web, *Communications of the ACM*, 54, 86-96.
- Efrat, K., Gilboa, S. & Yonatany, M. 2017, When marketing and innovation interact: The case of born-global firms, *International Business Review*, 26, 380-390.
- Enke, D. & Thawornwong, S. 2005, The use of data mining and neural networks for forecasting stock market returns, *Expert Systems with Applications*, 29, 927-940.
- Fan, S., Lau, R. Y. K. & Zhao, J. L. 2015, Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix, *Big Data Research*, 2, 28-32.
- Fleisher, C. S. 2008, Using open source data in developing competitive and marketing intelligence, *European Journal of Marketing*, 42, 852-866.
- France, S. L. & Ghose, S. 2018, Marketing analytics: Methods, practice, implementation, and links to other fields, *Expert Systems with Applications*, 119, 456-475.
- Gaham, M. & Bouzouia, B. 2009. Intelligent product-driven manufacturing control: A mixed genetic algorithms and machine learning approach to product intelligence synthesis. *Information, Communication and Automation Technologies, 2009. ICAT 2009. XXII International Symposium on*, 2009. IEEE, 1-8.
- Gutt, D., Neumann, J., Zimmermann, S., Kundisch, D. & Chen, J. 2019, Design of review systems—A strategic instrument to shape online reviewing behavior and economic outcomes, *The Journal of Strategic Information Systems*.
- Hosseini, S. M. S., Maleki, A. & Gholamian, M. R. 2010, Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty, *Expert Systems with Applications*, 37, 5259-5264.
- Huster, M. 2005, Marketing intelligence: a first mover advantage, *Competitive Intelligence Magazine*, 8, 13-17.
- Janssen, M., van der Voort, H. & Wahyudi, A. 2017, Factors influencing big data decision-making quality, *Journal of Business Research*, 70, 338-345.
- Kohli, A. K. & Jaworski, B. J. J. T. J. o. M. 1990, Market orientation: the construct, research propositions, and managerial implications, 1-18.

- Lau, R. Y., Liao, S. S., Wong, K.-F. & Chiu, D. K. 2012, Web 2.0 environmental scanning and adaptive decision support for business mergers and acquisitions, *MIS Quarterly*, 1239-1268.
- Lau, R. Y. K., Li, C. & Liao, S. S. Y. 2014, Social analytics: Learning fuzzy product ontologies for aspect-oriented sentiment analysis, *Decision Support Systems*, 65, 80-94.
- Lau, R. Y. K., Zhao, J. L., Chen, G. & Guo, X. 2016, Big data commerce, *Information & Management*, 53, 929-933.
- Li, Y.-M. & Li, T.-Y. 2013, Deriving market intelligence from microblogs, *Decision Support Systems*, 55, 206-217.
- Liang, T.-P. & Liu, Y.-H. 2018, Research Landscape of Business Intelligence and Big Data analytics: A bibliometrics study, *Expert Systems with Applications*.
- Liao, S.-H., Chu, P.-H. & Hsiao, P.-Y. 2012, Data mining techniques and applications – A decade review from 2000 to 2011, *Expert Systems with Applications*, 39, 11303-11311.
- Lies, J. 2019, Marketing Intelligence and Big Data: Digital Marketing Techniques on their Way to Becoming Social Engineering Techniques in Marketing, *International Journal of Interactive Multimedia & Artificial Intelligence*, 5.
- Moges, H.-T., Dejaeger, K., Lemahieu, W. & Baesens, B. 2013, A multidimensional analysis of data quality for credit risk management: New insights and challenges, *Information Management*, 50, 43-58.
- Navarro-García, A., Peris-Oritz, M. & Barrera-Barrera, R. 2016, Market intelligence effect on perceived psychic distance, strategic behaviours and export performance in industrial SMEs, *Journal of Business Industrial Marketing*, 31, 365-380.
- Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y. & Sun, X. 2011, The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature, *Decision Support Systems*, 50, 559-569.
- Ngai, E. W. T., Xiu, L. & Chau, D. C. K. 2009, Application of data mining techniques in customer relationship management: A literature review and classification, *Expert Systems with Applications*, 36, 2592-2602.
- Pan, G., Qi, G., Zhang, W., Li, S., Wu, Z. & Yang, L. T. 2013, Trace analysis and mining for smart cities: issues, methods, and applications, *IEEE Communications Magazine*, 51, 120-126.
- Park, D. H., Kim, H. K., Choi, I. Y. & Kim, J. K. 2012, A literature review and classification of recommender systems research, *Expert Systems with Applications*, 39, 10059-10072.
- Payne, A. & Frow, P. 2005, A strategic framework for customer relationship management, *Journal of Marketing*, 69, 167-176.
- Rijsdijk, S. A., Hultink, E. J. & Diamantopoulos, A. 2007, Product intelligence: its conceptualization, measurement and impact on consumer satisfaction, *Journal of the Academy of Marketing Science*, 35, 340-356.
- Rygielski, C., Wang, J.-C. & Yen, D. C. 2002, Data mining techniques for customer relationship management, *Technology in Society*, 24, 483-502.
- Scellato, S., Noulas, A. & Mascolo, C. 2011. Exploiting place features in link prediction on location-based social networks. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2011. ACM, 1046-1054.
- Seng, J.-L. & Chen, T. C. 2010, An analytic approach to select data mining for business decision, *Expert Systems with Applications*, 37, 8042-8057.
- Shaw, M. J., Subramaniam, C., Tan, G. W. & Welge, M. E. 2001, Knowledge management and data mining for marketing, *Decision Support Systems*, 31, 127-137.
- 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.
- Sorjonen, H. & Uusitalo, L. 2003. Market orientation and the programming of performing arts organisations. *AIMAC 7th International Conference on Arts and Cultural Management, Milan*, 2003.
- Sukumaran, S. & Sureka, A. 2006, Integrating structured and unstructured data using text tagging and annotation, *Business Intelligence Journal*, 11, 8.
- Trim, P. R. & Lee, Y.-I. 2008, A strategic marketing intelligence and multi-organisational resilience framework, *European Journal of Marketing*, 42, 731-745.
- Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N. & Haenlein, M. 2019, Digital transformation: A multidisciplinary reflection and research agenda, *Journal of Business Research*.
- Webster, J. & Watson, R. T. 2002, Analyzing the past to prepare for the future: Writing a literature review, *MIS Quarterly*, xiii-xxiii.
- Woo, J. Y., Bae, S. M. & Park, S. C. 2005, Visualization method for customer targeting using customer map, *Expert Systems with Applications*, 28, 763-772.
- Wright, S., Pickton, D. W. & Callow, J. 2002, Competitive intelligence in UK firms: a typology, *Marketing Intelligence & Planning*, 20, 349-360.
- Xu, Z., Frankwick, G. L. & Ramirez, E. 2016, Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective, *Journal of Business Research*, 69, 1562-1566.