

CÔNG CỤ THỨC ĐẨY HIỆU QUẢ KINH DOANH CHO CÁC DỊCH VỤ CÓ LƯỢNG NGƯỜI DÙNG LỚN

PGS.TS Phan Xuân Hiếu, TS Trần Mai Vũ, TS Lê Đức Trọng, TS Lê Hoàng Quỳnh

Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội

Dự án “Nghiên cứu và phát triển nền tảng tự động phân tích và hiểu khách hàng ứng dụng trong bán lẻ, thương mại điện tử và quảng cáo trực tuyến” là 1 trong 32 dự án được Quỹ Đổi mới sáng tạo (VINIF) thuộc Viện Nghiên cứu Dữ liệu lớn (VinBigData), Tập đoàn Vingroup tài trợ với tổng kinh phí 10 tỷ đồng. Đến nay, sau hơn 2 năm triển khai thực hiện, nhóm nghiên cứu của các nhà khoa học thuộc Trường Đại học Công nghệ (Đại học Quốc gia Hà Nội) đã hoàn thành các nội dung nghiên cứu và đóng gói thành công 2 sản phẩm công nghệ chủ lực gồm: “DSMiner - hệ thống phân tích và thấu hiểu khách hàng” và “DSWatcher - hệ thống lắng nghe mạng xã hội”.

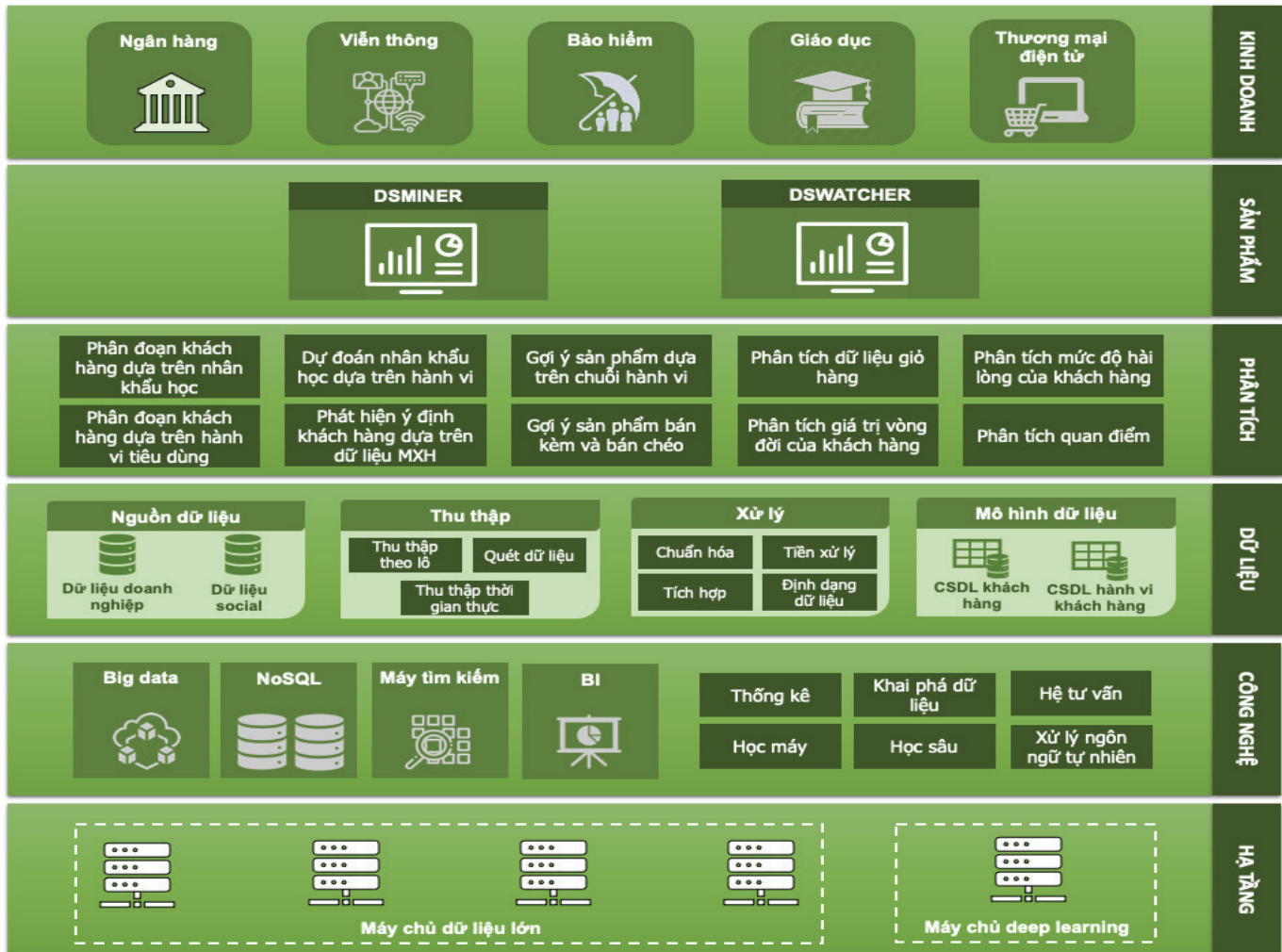
Việc mở rộng và giữ chân khách hàng ngày càng trở nên khốc liệt

Trong hơn một thập kỷ qua, Việt Nam là một trong các quốc gia có tốc độ tăng trưởng về sử dụng internet, dịch vụ viễn thông và mạng xã hội nhanh nhất trong khu vực và châu Á. Tính đến hiện tại, Việt Nam có khoảng 69 triệu người dùng internet, chiếm hơn 70% dân số. Hầu hết người dùng tham gia ít nhất một trong những mạng xã hội trực tuyến như Facebook, Tiktok, Instagram, Youtube, Zalo. Cùng với làn sóng số hoá và chuyển đổi số mạnh mẽ, không chỉ lĩnh vực thương mại, quảng cáo mà các dịch vụ truyền thống như ngân hàng, bảo hiểm, bán lẻ, giáo dục, y tế... cũng đã xem giao tiếp trực tuyến là kênh giao dịch quan trọng với người dùng và khách hàng.

Ngày nay các ứng dụng, dịch vụ thuộc những lĩnh vực kinh doanh khác nhau hoạt động và phát triển vừa cạnh tranh vừa cộng sinh. Giao dịch mua bán hoàn toàn trực tuyến hoặc theo hình thức bán trực tuyến (online to offline - O2O) xuất phát từ các kênh Facebook, Tiktok, Zalo... đã trở nên hết sức phổ biến. Hành trình khách hàng từ điểm tiếp xúc đầu tiên cho đến lúc thực hiện giao dịch không còn đơn thuần diễn ra trong lãnh địa của mỗi doanh nghiệp mà có thể đi xuyên qua các ứng dụng, dịch vụ thuộc nhiều lĩnh vực khác nhau. Vì thế, mỗi doanh nghiệp dù vẫn sở hữu dữ liệu riêng về người dùng, nhưng có thể không còn sở hữu toàn bộ dữ liệu hành trình và phễu khách hàng như cách marketing truyền thống. Điều này cản trở doanh nghiệp nắm bắt đầy đủ những chỉ

dấu, hành vi của người dùng. Sự kết nối, chia sẻ dữ liệu giữa các dịch vụ, doanh nghiệp gặp nhiều khó khăn do đặc thù kinh doanh khác nhau cũng như thiếu liên kết về mặt công nghệ. Thêm vào đó, nhiều lĩnh vực kinh doanh đã gần mức bão hoà nên việc cạnh tranh nhằm thu hút, mở rộng và giữ chân khách hàng ngày càng trở nên khốc liệt.

Trong bối cảnh đó, những câu hỏi cốt tử về khách hàng, thị trường tuy không mới nhưng giờ đây được đặt lại ở một mức độ cấp thiết và kỳ vọng cao hơn. Đó là làm sao để mở rộng tập khách hàng tiềm năng một cách hiệu quả? Làm sao để hiểu rõ sở thích, ý định, phản hồi của mỗi khách hàng? Làm sao để khuyến nghị những mặt hàng mà khách hàng quan tâm và có thể phát sinh giao dịch?... Hầu



Kiến trúc tổng quan của giải pháp phân tích dữ liệu và hiểu khách hàng.

hết các doanh nghiệp gặp khó khăn khi trả lời các câu hỏi này do thiếu khuyết dữ liệu; không kết nối được dữ liệu từ các điểm tiếp xúc trên hành trình khách hàng; không thể hợp nhất và khai thác từ đa nguồn dữ liệu mà đặc biệt là nguồn dữ liệu mở từ các mạng xã hội; thiếu năng lực tổ chức và xử lý dữ liệu lớn...

Dự án “Nghiên cứu và phát triển nền tảng tự động phân tích và hiểu khách hàng ứng dụng trong bán lẻ, thương mại điện tử

và quảng cáo trực tuyến” được thực hiện nhằm mục tiêu nghiên cứu, phát triển các giải pháp công nghệ phân tích dữ liệu lớn nhằm giải quyết các vấn đề bức thiết xoay quanh thị trường và khách hàng tại các doanh nghiệp. Đến nay, dự án đã hoàn thành tốt các nội dung nghiên cứu, phát triển và đóng gói thành công 2 sản phẩm công nghệ chủ lực (DSMiner và DSWatcher). Dự án đã được VINIF nghiệm thu và đánh giá cao.

Xây dựng kho dữ liệu với quy mô lớn dựa trên các công nghệ tiên tiến

Nếu như khách hàng là trung tâm của các bài toán kinh doanh, thì dữ liệu là cơ sở cho sự phân tích, thấu hiểu được nhu cầu và ý định của họ. Vì vậy, dự án hướng tới thu thập, xử lý và phân tích một cách đầy đủ nhất bức tranh tổng thể về dữ liệu liên quan tới các điểm tiếp xúc trên hành trình trải nghiệm của từng khách hàng từ đa nguồn dữ liệu đảm bảo các yếu tố về tính tin cậy và bảo mật.

Khoa học - Công nghệ và Đổi mới sáng tạo



Nhóm nghiên cứu của dự án “Nghiên cứu và phát triển nền tảng tự động phân tích và hiểu khách hàng ứng dụng trong bán lẻ, thương mại điện tử và quảng cáo trực tuyến”.

Cụ thể, 2 nguồn dữ liệu quy mô lớn với cấu trúc phức tạp được xem xét gồm: dữ liệu nguồn mở trên mạng xã hội và dữ liệu vận hành của doanh nghiệp. Việc khai thác dữ liệu đặc thù của doanh nghiệp giúp xác định sở thích, nhu cầu của tập khách hàng ở giai đoạn trước và sau khi phát sinh giao dịch. Trong khi đó, dữ liệu nguồn mở mạng xã hội không chỉ cung cấp khả năng đánh giá, định vị tập khách hàng tiềm năng trong giai đoạn trước khi phát sinh giao dịch, mà còn giúp làm giàu các thông tin bổ trợ, phục vụ các phân tích chuyên sâu về khách hàng ở những giai đoạn sau.

Các nguồn dữ liệu được tổ chức theo 2 nhóm mô hình chính: mô hình dữ liệu hồ sơ khách hàng (thông tin nhân khẩu được khách

hàng khai báo như tuổi, giới tính, nghề nghiệp, tình trạng hôn nhân) và mô hình dữ liệu hành vi khách hàng (thông tin mang tính thời gian được hình thành từ quá trình tương tác của khách hàng với các ứng dụng, dịch vụ của doanh nghiệp và các tín hiệu hành vi khác từ các nguồn mở mạng xã hội...). Để đáp ứng việc lưu trữ, xử lý các nguồn dữ liệu lớn trên, dự án đã thiết kế hạ tầng dữ liệu theo kiến trúc hồ dữ liệu (data lake). Đây là kiến trúc được nhiều công ty lớn trên thế giới và Việt Nam triển khai như Amazon, Microsoft, Viettel, Mobifone... Ưu điểm của kiến trúc này là tạo nên một kho dữ liệu cả thô lẫn tinh với quy mô lớn dựa trên các công nghệ tiên tiến về thu thập, lưu trữ, học máy, học sâu và phân tích dữ liệu cũng như các công cụ hỗ trợ

trực quan hoá, xây dựng báo cáo kinh doanh thông minh hay giám sát bảo mật.

Với mục tiêu nhận diện chân dung và nhu cầu của khách hàng trong giai đoạn trước khi có tương tác với các dịch vụ của doanh nghiệp, 3 vấn đề chính được xem xét gồm: dự đoán thông tin nhân khẩu học dựa trên hành vi khách hàng; phân đoạn khách hàng dựa trên thông tin nhân khẩu học và hành vi tiêu dùng; dự đoán và phát hiện ý định của khách hàng dựa trên hành vi mạng xã hội. Cả 3 vấn đề này đều dựa trên việc xác định những đặc điểm khác biệt trong hành vi của khách hàng khi tương tác với các dịch vụ mạng xã hội [1]. Tiếp đó, việc hiểu sở thích, ý định của khách hàng trong khi tương tác với các ứng dụng, dịch vụ được nghiên cứu bao gồm: gợi ý sản phẩm kế tiếp dựa trên chuỗi hành vi của khách hàng; gợi ý sản phẩm bán kèm, bán chéo, bán gia tăng; phân tích giỏ hàng. Các vấn đề này dựa trên sự phân tích các mẫu kết hợp hoặc tương hỗ [2], mẫu chuỗi thể hiện thói quen, hành vi khách hàng [3] từ các lịch sử giao dịch nhằm gợi ý sản phẩm mà khách hàng có thể quan tâm cao trong lần giao dịch tiếp theo. Cuối cùng, nhằm đánh giá những phản hồi và khả năng rời dịch vụ ở giai đoạn sau khi phát sinh giao dịch, dự án tập trung vào các vấn đề như phân tích mức độ hài lòng,

quan điểm, giá trị vòng đời và khả năng rời dịch vụ của khách hàng. Giải pháp bao gồm phân tích ngôn ngữ tự nhiên, hành vi tiêu dùng để dự đoán số lượng và giá trị giao dịch của mỗi khách hàng trong tuần, tháng tiếp theo dựa trên lịch sử tiêu dùng của họ [4] cũng như phân tích dấu hiệu rời dịch vụ [5].

Lắng nghe mạng xã hội để thấu hiểu khách hàng

DSMiner - hệ thống phân tích và thấu hiểu khách hàng, là sản phẩm đóng gói các mô hình tự động phân tích hành vi, thói quen của khách hàng dựa trên các kỹ thuật xử lý dữ liệu lớn, các phương pháp học máy thống kê và học sâu tiên tiến. DSMiner cung cấp một giải pháp toàn diện cho các doanh nghiệp trong bản đồ hành trình khách hàng, đồng hành trong các bài toán trải nghiệm khách hàng với doanh nghiệp từ sự tương tác đầu tiên và trong mối quan hệ dài hạn.

DSWatcher - hệ thống lắng nghe mạng xã hội, là giải pháp tự động thu thập và phân tích tín hiệu dữ liệu của người dùng/khách hàng từ nhiều nguồn dữ liệu khác nhau như mạng xã hội, báo chí trực tuyến, diễn đàn, blog... Từ đó, DSWatcher giúp doanh nghiệp hiểu được phản ứng của thị trường, hiểu được ý định của khách hàng, đánh giá hiệu quả chiến dịch truyền thông,

quản lý thương hiệu cũng như phân tích và cảnh báo sớm khủng hoảng truyền thông.

Để hoàn thiện và phát triển hai hệ thống DSMiner và DSWatcher, nhóm nghiên cứu đã khảo sát, tìm hiểu và đề xuất hợp tác nghiên cứu với các doanh nghiệp để có thể tiếp cận với nguồn dữ liệu cụ thể, nhiều nguồn dữ liệu được tiếp cận từ các nghiên cứu trước đây cũng như các nguồn dữ liệu mô phỏng nổi tiếng trên thế giới. Tính đến hiện tại, DSMiner và DSWatcher đã được triển khai ứng dụng tại hơn 30 doanh nghiệp, tổ chức thuộc nhiều lĩnh vực như ngân hàng, bảo hiểm, viễn thông, giáo dục, bán lẻ, thương mại điện tử... Kết quả thí điểm cho thấy, sự hiệu quả của các giải pháp trong việc nâng cao trải nghiệm quản lý khách hàng cũng như thúc đẩy hiệu quả kinh doanh dựa trên việc phân tích và hiểu rõ hơn về khách hàng. Tại mỗi đơn vị, dự án triển khai theo mô hình phù hợp với tính chất của từng lĩnh vực kinh doanh cũng như đảm bảo tối đa tính riêng tư của dữ liệu và an toàn thông tin khách hàng của doanh nghiệp.

Trong giai đoạn tiếp theo, nhóm nghiên cứu đang lên kế hoạch tiếp tục nâng cấp các tính năng phân tích chuyên sâu và thay đổi giao diện phù hợp hơn với các nhóm người dùng cuối, các sản phẩm sẽ được mở rộng thêm qua các nhóm ngành trong

lĩnh vực bán lẻ, thương mại điện tử. Kết hợp thêm các mô hình phân tích hình ảnh, âm thanh, video để đưa thêm các tính năng phân tích hành vi người dùng đa kênh, đa điểm chạm. Nhóm cũng sẽ thử nghiệm sản phẩm tại một số thị trường trong khu vực Đông Nam Á với cách tính dữ liệu, sản phẩm và người dùng khác trong năm 2023 ✍

TÀI LIỆU THAM KHẢO

- [1] B. Jang, M. Kim, G. Harerimana, S.U. Kang, J.W. Kim (2020), "Bi-LSTM model to increase accuracy in text classification: Combining Word2vec CNN and attention mechanism", *Applied Sciences*, **10(17)**, p.5841.
- [2] J.K. Tarus, Z. Niu, G. Mustafa (2018), "Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning", *Artificial Intelligence Review*, **50(1)**, pp.21-48.
- [3] D.T. Le, H.W. Lauw, Y. Fang (2019), "Correlation-sensitive next-basket recommendation", *Proceedings of The Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*, pp.2808-2814.
- [4] Y.T. Wen, P.W. Yeh, T.H. Tsai, W.C. Peng, H.H. Shuai (2018), "Customer purchase behavior prediction from payment datasets", *In WSDM'18*, pp.628-636.
- [5] A.K. Ahmad, A. Jafar, K. Aljoumaa (2019), "Customer churn prediction in telecom using machine learning in big data platforms", *Journal of Big Data*, **6(1)**, pp.1-24.