

Công nghệ Big Data và xu hướng ứng dụng

Nguồn tài nguyên giá trị nhất của thế giới hiện nay không còn là dầu mỏ, mà là kho dữ liệu số đang tăng lên với cấp độ lũy thừa mỗi ngày. Trong cuộc cách mạng công nghiệp 4.0, Big Data là một yếu tố đóng vai trò then chốt. Vậy Big Data thực chất là gì, và nó đang được ứng dụng như thế nào? Đối với nhiều người, đó là một thuật ngữ mơ hồ về hình ảnh của những hệ thống máy chủ khổng lồ, hoặc sẽ liên hệ đến việc nhận được các loại quảng cáo từ một nhà bán lẻ. Nhưng về bản chất, Big Data sâu sắc hơn và rộng hơn thế. Trên cơ sở làm rõ những vấn đề này, bài viết đề cập đến một số lĩnh vực chính mà Big Data đang được sử dụng để tạo ra lợi thế tuyệt vời trong nhiều khía cạnh của sản xuất và đời sống.

Đôi nét về Big Data

Big Data (dữ liệu lớn) là thuật ngữ dùng để chỉ một tập hợp dữ liệu rất lớn và phức tạp, đến nỗi những công cụ lưu trữ, ứng dụng, xử lý dữ liệu truyền thống không thể nào đảm đương được. Big Data được khái quát là sự thu thập, quản lý và phân tích dữ liệu; những việc đó đã vượt xa dữ liệu cấu trúc tiêu biểu (typical), nó có thể được truy vấn với hệ thống quản lý dữ liệu quan hệ (thường với những tệp phi cấu trúc - unstructured files), video kỹ thuật số, hình ảnh, dữ liệu cảm biến, tệp lưu nhật ký, hay bất cứ dữ liệu nào không có trong hồ sơ với các phạm vi tìm kiếm khác. Từ năm 2001, nhà phân tích Doug Laney của META Group (nay là Công ty nghiên cứu Gartner) đã chỉ ra những thách thức và cơ hội trong việc tăng trưởng dữ liệu, được mô tả bằng ba chiều “3V”: tăng về số lượng lưu trữ (volume), tăng về tốc độ xử lý (velocity) và tăng về chủng loại (variety). Gartner cùng với nhiều công ty và tổ chức khác trong lĩnh vực công nghệ thông tin tiếp tục sử dụng mô hình “3V” này để định nghĩa Big Data. Đến năm 2012, Gartner bổ sung thêm, Big Data ngoài ba tính chất trên thì còn phải “cần đến các dạng xử lý mới để giúp đưa ra quyết định, khám phá sâu vào sự vật/sự việc và tối ưu hóa quy trình làm việc”.

Big Data chứa rất nhiều thông tin quý giá mà nếu trích xuất thành công, sẽ hỗ trợ đắc lực cho việc kinh doanh, nghiên cứu khoa học, dự đoán các dịch bệnh sắp phát sinh và thậm chí có thể đưa ra những khuyến nghị chính xác về điều kiện giao thông, thời tiết theo thời gian thực... Thuật ngữ Big Data đã được sử dụng từ những năm 1990, nhưng chỉ thực sự bùng nổ trong khoảng 10 năm trở lại đây. Báo cáo của Công ty Nghiên cứu thị trường IDC (Hoa Kỳ) cho thấy, mỗi ngày thế giới tạo ra khoảng 2,5 quintillion byte dữ liệu, và con số này vẫn không ngừng gia tăng, 90% dữ liệu trên thế giới hiện nay được tạo ra trong 3 năm gần đây. Doanh thu từ thị trường Big Data đã tăng lên 16,9 tỷ USD năm 2015 và tiếp tục tăng trưởng kép với tốc độ 27%, đạt 32,4 tỷ USD năm 2017. Big Data đang được ứng dụng rộng rãi trong nhiều lĩnh vực như: bán lẻ, ngân hàng, dịch vụ chăm sóc sức khỏe, viễn thông, giải trí, bảo hiểm, giao thông, giáo dục...

Những ứng dụng phổ biến

Phân tích mục tiêu của khách hàng

Đây là một trong những lĩnh vực lớn nhất của việc ứng dụng công nghệ Big Data, nó giúp các cơ quan, tổ chức hiểu rõ hơn về nhu cầu khách

hàng cũng như hành vi và sở thích của họ. Các công ty mở rộng bộ dữ liệu truyền thống của họ, kết hợp với dữ liệu truyền thông, xã hội, nhật ký trình, đồng thời phân tích văn bản và dữ liệu cảm biến... để có bức tranh hoàn chỉnh, toàn diện hơn về khách hàng. Mục tiêu lớn được hướng tới là tạo ra các mô hình tiên đoán. Ví dụ, nhà bán lẻ Target của Hoa Kỳ có thể dự đoán gần như chính xác thời điểm một khách hàng của họ mong muốn sinh em bé, để giới thiệu, cung cấp các sản phẩm phù hợp. Sử dụng Big Data, các công ty viễn thông có thể dự đoán tốt hơn việc thay đổi mong muốn của khách hàng; công ty bảo hiểm xe hơi hiểu rõ khách hàng của họ lái xe như thế nào... Ngay cả các chiến dịch bầu cử của chính phủ cũng có thể được tối ưu hóa bằng cách sử dụng phân tích dữ liệu lớn. Một số người tin rằng, chiến thắng của cựu tổng thống Hoa Kỳ Barack Obama trong chiến dịch bầu cử tổng thống vào năm 2012 là nhờ khả năng vượt trội của các công nghệ phân tích Big Data.

Hiểu và tối ưu hóa sản phẩm, quy trình kinh doanh

Big Data ngày càng được sử dụng nhiều trong việc tối ưu hóa quy trình kinh doanh. Các nhà bán lẻ có thể tối ưu hóa cổ phiếu của họ dựa trên các

dự đoán được tạo từ dữ liệu truyền thông xã hội, xu hướng tìm kiếm trên web; các nhà bán lẻ sử dụng rất nhiều phân tích dữ liệu lớn để tối ưu hóa chuỗi cung ứng, cũng như tuyến đường phân phối. Nhưng Big Data còn được ứng dụng cả trong việc đo lường văn hóa công ty và hiệu suất làm việc của nhân viên. Ví dụ, Công ty Sociometric Solutions (Hoa Kỳ) đặt cảm biến vào tấm huy hiệu tên nhân viên để có thể phát hiện các động lực xã hội tại nơi làm việc. Các cảm biến sẽ báo cáo về cách nhân viên di chuyển xung quanh nơi làm việc, nói chuyện, và thậm chí cả cách giao tiếp trong công việc của họ.

Mọi người có thể gắn thẻ RFID lên điện thoại, chìa khóa hoặc kính của bạn, để xác định vị trí của chúng, tránh bị mất. Nhưng ở tầm cao hơn, công nghệ này có thể mở ra một lĩnh vực hoàn toàn mới của “dữ liệu nhỏ”, nếu Big Data đang xem xét số lượng lớn thông tin và phân tích nó, thì dữ liệu nhỏ hướng tới từng sản phẩm riêng lẻ, ví dụ như việc theo dõi thùng sữa chua trong một lô hàng để nắm được khả năng bán hàng của các đại lý phân phối.

Chăm sóc sức khỏe cộng đồng

Sức mạnh tính toán của công nghệ Big Data cho phép giải mã toàn bộ chuỗi AND của người chỉ trong vài phút, hỗ trợ đắc lực các bác sỹ trong việc đưa ra phác đồ điều trị riêng cho từng bệnh nhân, phù hợp với đặc tính di truyền của họ. Hay những gì sẽ xảy ra khi tất cả dữ liệu cá nhân từ đồng hồ và thiết bị đeo thông minh được sử dụng để theo dõi sức khỏe từ xa cho hàng triệu người, giúp sớm phát hiện ra các bệnh họ mắc phải. Ví dụ vòng tay thông minh Jawbone giúp thu thập dữ liệu, từ đó nhận biết và đo đạc một cách chính xác mức độ tiêu thụ calo trong các hoạt động mà người đeo thực hiện, cũng như theo dõi giấc ngủ của họ vào ban đêm để phân tích những dấu hiệu

bất thường; ResearchKit - ứng dụng chăm sóc sức khỏe của Apple, đã biến điện thoại di động thành một thiết bị nghiên cứu y sinh học, phục vụ cho các nghiên cứu về sức khỏe thông qua việc theo dõi số bước đi thực hiện trong một ngày, hoặc yêu cầu bạn trả lời các câu hỏi về cảm giác sau khi hóa trị, hay theo dõi căn bệnh Parkinson đang tiến triển như thế nào... Các kỹ thuật Big Data còn được sử dụng để theo dõi trẻ sơ sinh trong một số trường hợp sinh non và ốm yếu. Bằng cách ghi lại và phân tích nhịp tim, nhịp thở của bé, các đơn vị cung cấp thiết bị đã có thể dự đoán nhiễm trùng trước 24 giờ khi có bất kỳ triệu chứng thể chất nào xuất hiện, giúp đưa ra những can thiệp sớm.

Nhưng ứng dụng nổi bật nhất của Big Data trong lĩnh vực y tế là phân tích dữ liệu, giúp theo dõi và dự đoán sự phát triển của dịch bệnh. Việc tích hợp dữ liệu từ hồ sơ y tế, phân tích truyền thông xã hội... cho phép theo dõi dịch cúm theo thời gian thực, chỉ đơn giản bằng cách lắng nghe những gì mọi người đang nói. Từ năm 2014, Google đã dự đoán dịch cúm dựa trên lưu lượng và vị trí của người tìm kiếm, tuy việc bạn tìm kiếm “các triệu chứng cúm” không có nghĩa là bạn đang bị bệnh, nhưng công nghệ Big Data hiện nay đã làm được nhiều hơn thế.

Nâng cao hiệu suất thể thao

Hầu hết các môn thể thao đang sử dụng công nghệ Big Data. Chẳng hạn như, công cụ IBM SlamTracker dành cho các giải đấu quần vợt; sử dụng phân tích video theo dõi hiệu suất của người chơi trong bóng đá hoặc bóng chày; công nghệ cảm biến trong thiết bị thể thao như bóng rổ hoặc gậy golf cho phép nhận phản hồi (qua điện thoại thông minh và máy chủ đám mây), giúp cải thiện những thao tác chưa đúng. Nhiều đội thể thao còn theo dõi vận động viên trong giờ nghỉ thông qua sử dụng

công nghệ Big Data để điều chỉnh chế độ dinh dưỡng, vận động và cả giấc ngủ, cũng như phân tích các cuộc trò chuyện trên phương tiện truyền thông xã hội để theo dõi mức độ hạnh phúc, tình cảm của các vận động viên.

Trong môn thể thao vua, những câu lạc bộ lớn thường sử dụng thông tin, thống kê chi tiết từ công nghệ Big Data để cải thiện năng suất của cầu thủ, giúp chuẩn bị chiến thuật và đánh giá những cầu thủ tiềm năng. Chúng ta có thể thấy ở kỳ World Cup năm 2014, nhà vô địch Đức đã thành công khi ứng dụng những hiểu biết về công nghệ Big Data vào bóng đá. Nhiều người hâm mộ bất ngờ khi biết rằng, các cầu thủ của đội tuyển Đức đã sử dụng bộ trang phục được trang bị hệ thống MiCoach của Adidas ở những buổi luyện tập trước và trong quá trình giải đấu diễn ra. Đây là dịch vụ giám sát sinh lý, thu thập và truyền tải thông tin trực tiếp từ cơ thể các cầu thủ (bao gồm nhịp tim, quãng đường đã di chuyển, tốc độ, gia tốc và năng lượng...), rồi gửi về một chiếc iPad của ban huấn luyện, phục vụ cho việc giám sát, phân tích, đánh giá điểm rơi phong độ. Điều thú vị là nó còn giúp tìm ra được những cầu thủ nào thích hợp để đá cùng nhau hơn là đá với những người khác.

Cải thiện an ninh và thực thi pháp luật

Big Data được áp dụng rất nhiều trong việc cải thiện bảo mật và cho phép thực thi pháp luật. Ví dụ Cơ quan An ninh quốc gia Hoa Kỳ (NSA) sử dụng các phân tích dữ liệu lớn dự đoán sớm về các cuộc khủng bố; một số đơn vị sử dụng công nghệ Big Data để phát hiện và ngăn chặn các cuộc tấn công trên mạng; lực lượng cảnh sát sử dụng các công cụ Big Data để bắt tội phạm và dự đoán hoạt động của các tổ chức tội phạm; các công ty tín dụng sử dụng Big Data để phát hiện các giao dịch

gian lận.

Trong môi trường cạnh tranh khốc liệt hiện nay, không chỉ các tổ chức chính phủ mà các doanh nghiệp cũng đang vận dụng rất nhiều yếu tố để bảo mật quyền lợi về thương hiệu, giảm thiểu rủi ro từ tác động bên ngoài. Thông qua phân tích dữ liệu, Big Data có thể giúp các tổ chức doanh nghiệp xác định các hoạt động khả nghi, và chỉ ra hành vi gian lận để giải quyết.

Xây dựng thành phố thông minh

Big Data được sử dụng để cải thiện nhiều khía cạnh của các thành phố và quốc gia. Trong tương lai, Big Data sẽ giữ vai trò chủ đạo đối với các thành phố, hướng tới xử lý dữ liệu ở những phạm vi, cấp độ siêu nhỏ và truyền tải thông tin tới cả chính phủ lẫn người dân trong khung thời gian giới hạn. Nó cho phép các thành phố tối ưu hóa hoạt động dựa trên lưu lượng truy cập thông tin giao thông theo thời gian thực, cũng như phân tích dữ liệu truyền thông để đánh giá các vấn đề về xã hội...

Một số thành phố hiện đang thử nghiệm các mô hình Big Data với mục đích chuyển đổi sang thành phố thông minh, nơi mà tất cả các quy trình, cơ sở hạ tầng giao thông và tiện ích đều được kết nối với nhau. Thành phố Long Beach, California (Hoa Kỳ) đang sử dụng đồng hồ nước thông minh để phát hiện mất trộm nước theo thời gian thực, giúp một số chủ nhà cắt giảm tới 80% lượng nước sử dụng. Điều đó rất quan trọng khi tiểu bang này thường phải trải qua những đợt hạn hán tồi tệ. Los Angeles sử dụng dữ liệu từ cảm biến và 4.500 camera lắp đặt trên các đường phố để kiểm soát đèn giao thông, giúp giảm tắc nghẽn giao thông lên đến 16%.

Startup công nghệ Veniam đang thử nghiệm một cách mới để tạo ra các điểm truy cập wifi di động trên

toàn thành phố Porto, Bồ Đào Nha. Hơn 600 xe buýt và taxi thành phố đã được trang bị máy phát wifi, tạo ra điểm truy cập wifi miễn phí lớn nhất thế giới. Đổi lại, thành phố nhận được một lượng lớn dữ liệu, để bù đắp chi phí về wifi; các dữ liệu này sẽ được cung cấp cho các bộ phận quản lý, giúp cảnh báo khi thùng rác đầy, hạn chế lãng phí thời gian di chuyển, tối ưu hóa giờ làm việc, hoặc cảnh báo về tắc đường, nắm bắt hoạt động của các tổ chức tội phạm...

Hoạt động tài chính - ngân hàng

Sự gia tăng của hoạt động giao dịch tần suất cao (High frequency trading - HFT) theo thời gian thực đã khiến các nhà quản lý trong lĩnh vực chứng khoán và thị trường chứng khoán phải làm việc rất nhiều để theo kịp sự mở rộng của dữ liệu cấu trúc và phi cấu trúc. Bằng cách sử dụng những công cụ và kỹ thuật mới của Big Data, những công ty hoạt động trong lĩnh vực thị trường vốn có thể cải thiện hoạt động giám sát giao dịch và phát hiện được những hành vi giao dịch bất thường và hành vi phạm pháp trước khi chúng được thực hiện. Hình thức HFT đang chiếm hơn 60% tổng số giao dịch trên thị trường Hoa Kỳ.

Bên cạnh đó, đặc thù của hoạt động ngân hàng tạo ra một lượng dữ liệu khổng lồ đến từ cả bên trong và bên ngoài. Trong nội bộ, đó là hệ thống xử lý giao dịch, hệ thống quản lý, hệ thống ATM, hệ thống giám sát hạ tầng... Yếu tố dữ liệu bên ngoài thường được xác định từ các tổ chức thông tin tín dụng, ngân hàng nhà nước, cơ quan thuế, tổng cục thống kê, bảo hiểm; mạng xã hội, các giao dịch trực tuyến qua mạng; hệ thống lưu trữ dữ liệu ngành, quốc gia... Giá trị của dữ liệu là độ tin cậy để các tổ chức ra quyết định, cung cấp câu trả lời nhanh, chính xác cho các câu hỏi kinh doanh; truy cập thông tin nhanh chóng, dễ dàng; hiểu biết sâu

sắc về hành vi của khách hàng; tạo ra tri thức mới... Từ những dữ liệu có cấu trúc như lịch sử giao dịch, hồ sơ khách hàng tới những dữ liệu phi cấu trúc như hoạt động của khách hàng trên website, ứng dụng mobile banking hay trên mạng xã hội..., Big Data sẽ đem lại những lợi thế cạnh tranh và hiệu quả to lớn cho lĩnh vực này.

Thay lời kết

Thế giới đang chứng kiến một cuộc cách mạng công nghiệp mới được thúc đẩy bởi các công nghệ Big Data, Internet kết nối vạn vật và tự động hóa... Sự giao thoa giữa các xu hướng công nghệ và các vấn đề trong phát triển kinh tế - xã hội, bao gồm cả việc sử dụng Internet ngày càng tăng; sự suy giảm chi phí thu thập, truyền tải, lưu trữ và phân tích dữ liệu... dẫn đến việc tạo ra những khối lượng dữ liệu khổng lồ, gọi chung là Big Data. Đây chính là nguồn lực để thúc đẩy hình thành các ngành công nghiệp, các quy trình sản xuất kinh doanh và tạo ra sản phẩm mới. Quan trọng hơn, Big Data đang mở ra những cơ hội lớn trong yêu cầu chuyển dịch sang mô hình phát triển kinh tế dựa trên dữ liệu. Trong mô hình này, dữ liệu là tài sản cốt lõi để tạo ra lợi thế cạnh tranh, chi phối hoạt động đổi mới sáng tạo, tăng trưởng..., đóng vai trò quyết định trong mục tiêu phát triển bền vững của mỗi quốc gia. Tuy vẫn còn những chỉ trích xoay quanh Big Data, nhưng lĩnh vực này đang còn rất mới, chúng ta hãy chờ xem trong tương lai Big Data sẽ biến đổi như thế nào ✍

Hồ Thị Hạnh

(lược dịch theo www.bernardmarr.com)