

เพิ่มมูลค่าข้อมูลทางธุรกิจด้วย Data mining



วิเศษคุชท์ เมาระพงษ์

ศึกษานิเทศก์กรมการสารสนเทศฯ หน่วยงานภาครัฐ
สังกัดสถาบันวิจัยและให้คำปรึกษาแห่งมหาวิทยาลัยธรรมศาสตร์

ต่อ จากฉบับที่แล้ว

ข้อมูลของลูกค้าดังกล่าวมานี้ ความจริงแล้วมีให้นำมาใช้ได้มากมายอยู่แล้ว แต่อาจจะอยู่ในรูปที่เป็นได้ไม่ชัดเจน อันได้แก่ ข้อมูลที่รวบรวมไว้จากการบันทึกใน log file ของการใช้บริการ Web site หรือข้อมูลจากการสมัครสมาชิกในรูปแบบต่างๆ เป็นต้น ข้อมูลเหล่านี้สามารถอำนวยความสะดวกในการติดตามดูผู้ใช้ (user tracking) ยิ่งผู้ใช้เข้าใช้บริการ Web site บ่อยและนานขึ้น เราก็จะยิ่งมีโอกาสทราบและรู้จักกับผู้ใช้มากขึ้นเท่านั้น สำหรับข้อมูลดังกล่าวเกี่ยวกับผู้ใช้จะมีการวิเคราะห์ออกมาใน 3 ลักษณะดังต่อไปนี้

1. **Demographics** เป็นข้อมูลเกี่ยวกับที่อยู่ หรือสถานที่ของผู้ใช้ในขณะที่ใช้บริการ Web site ซึ่งจะสามารถประมวลเป็นสถิติบริเวณที่อยู่อาศัยของกลุ่มผู้ใช้ส่วนมากได้

2. **Psychographics** เป็นข้อมูลด้านจิตวิทยา ซึ่งแสดงถึงพฤติกรรม หรือค่านิยมในด้านต่างๆ ของผู้ใช้ โดยสามารถแบ่งแยกกลุ่มผู้ใช้ตามข้อมูลการเข้าใช้บริการ web ทั้งในแง่ของเวลาและเนื้อหา

3. **Technolographics** เป็นข้อมูลที่แสดงถึงระดับความรู้และความสนใจในเทคโนโลยีด้านต่างๆ ของผู้ใช้ รวมถึงเครื่อง Computer ที่ติดต่อเข้ามาด้วย ซึ่งจะช่วยในการพัฒนาสินค้าและบริการให้น่าสนใจและเหมาะสมในแง่ของเทคโนโลยีได้ยิ่งขึ้น

เมื่อนำข้อมูลที่วิเคราะห์แล้วทั้ง 3 ลักษณะนี้มาพิจารณาโดยละเอียด จะเกิดประโยชน์อย่างมากในการศึกษาเกี่ยวกับสภาพและพฤติกรรมโดยรวมของประชากร ซึ่งจำนวนข้อมูลที่นำมาใช้วิเคราะห์มักจะมีจำนวนมากและให้ผลการวิเคราะห์ที่มีความแม่นยำสูง

การประยุกต์ Web mining ทางด้านการตลาดบน Internet

Data mining เป็นกระบวนการสืบค้นข้อมูลสำคัญอันจะเป็นประโยชน์ต่อการดำเนินงานธุรกิจออกจากกองข้อมูลขนาดมหาศาลที่จัดเก็บอยู่ภายในแหล่งฐานข้อมูลของแต่ละองค์กร การขุดค้นข้อมูลด้วยเทคโนโลยี Data mining ประกอบไปด้วยขั้นตอนเทคโนโลยีอันสลับ

Web mining กับการทำธุรกิจ e-Commerce

ในการทำธุรกิจต่างๆ ไม่ว่าจะในรูปแบบของห้างร้าน บริการส่งของทางไปรษณีย์ หรือการทำธุรกิจแบบอิเล็กทรอนิกส์นั้น ปัจจัยหนึ่งที่มีความสำคัญอย่างมากต่อความสำเร็จของธุรกิจก็คือ ความเข้าใจในตัวลูกค้า หรือกลุ่มลูกค้า ยิ่งรู้ข้อมูลมาก ทำให้เข้าใจลูกค้าได้อย่างแท้จริงมากขึ้น โอกาสที่จะทำธุรกิจให้ตรงกับความต้องการของตลาดก็จะมีมากขึ้นไปด้วย

ซับซ้อนมากมาย ไม่ว่าจะเป็นการเรียนรู้จุดจำแบบแผนมาตรฐานของข้อมูลเฉพาะกลุ่ม (pattern recognition technology) การใช้ประโยชน์จากอัลกอริทึม ซึ่งออกแบบมาเพื่อให้ระบบ Computer สามารถเรียนรู้และพัฒนาตัวเองได้ (machine learning & genetic algorithms) หรือการเชื่อมโยงแลกเปลี่ยนข้อมูลภายในเครือข่ายเน็ตเวิร์คแบบโพลีโนเมียลซึ่งเลียนแบบการทำงานของระบบประสาทของมนุษย์ (Netural & Polynomial networks) ฯลฯ

อย่างไรก็ตาม ไม่ว่าจะมีการใช้เทคโนโลยีทาง Computer ที่สลับซับซ้อนเพียงไร เป้าหมายของ Data mining ก็เป็นเรื่องง่าย ๆ คือ ทำอย่างไรถึงจะเอาชนะคู่แข่งทางธุรกิจให้ได้ ทำอย่างไรถึงจะรู้ว่าประชากรกลุ่มไหนคือ ลูกค้าเป้าหมาย (Who) และประชากรกลุ่มนี้ต้องการซื้ออะไร (What) ซื้อที่ไหน (Where) ซื้อเมื่อไหร่ (When) และที่สำคัญที่สุดก็คือ อะไรคือแรงขับเคลื่อนให้ลูกค้าเป้าหมายต้องการสินค้าที่ว่านั้น (Why) ซึ่งคำถามหลักๆ เหล่านี้เมื่อมาผนวกเข้ากับรูปแบบการดำเนินธุรกิจยุคดิจิทัลที่มีเครือข่าย Internet เป็นสมรรถภูมิการตลาดอันไร้พรมแดน ก็ส่งผลให้เทคโนโลยี Data mining กลายสภาพมาเป็นเครื่องมือการตลาดที่สำคัญมาก สำคัญขนาดที่ว่า นักธุรกิจรายใดไม่ให้ความสำคัญกับมัน ก็อาจจะถูกคู่แข่งทางการค้าของตนทิ้งห่างไปเลย

การเรียนรู้ลูกค้าผ่านเครือข่าย Internet

ในอดีตการขุดค้นข้อมูลด้วยเทคนิค Data mining มักจะพุ่งเป้าไปที่แหล่งข้อมูลสำคัญสามแหล่ง คือ คลังข้อมูล (Data warehouse) ศูนย์บริการข้อมูล (Data marts) และกลุ่มไฟล์ข้อมูลลูกค้า (Customer information file) แต่ในระยะหลัง เมื่อรูปแบบการค้ามีการปรับเปลี่ยนไปจากการซื้อขายกันด้วยเงินสดที่ต้องมีการออกไปสั่งซื้อ หรือออกเอกสารใบเสร็จจำนวนมาก ไปเป็นการติดต่อค้าขายด้วยข้อมูลอิเล็กทรอนิกส์ (electronic commerce) และใช้เครือข่าย Internet เป็นช่องทางดำเนินธุรกิจ วิธีการขุดค้นข้อมูลด้วยเทคนิค Data mining จึงต้องได้รับการปรับเปลี่ยนรูปแบบให้ทันสมัยมากขึ้นและมีการให้ความสำคัญกับวิธีการจัดการสร้าง Web site ขององค์กรธุรกิจมากขึ้น

สำหรับการศึกษาวิเคราะห์ข้อมูลของลูกค้ากลุ่มเป้าหมายด้วยเทคโนโลยี Data mining ผ่านเครือข่าย Internet เท่าที่ดำเนินการกันอยู่ทั่วไปในขณะนี้ มักจะมุ่งเน้นไปที่การศึกษารูปแบบวิธีการที่บรรดากลุ่มลูกค้าเป้าหมายเคยใช้ติดต่อ (หรือกำลังใช้ติดต่อ) กับ Web site ของตน เพียงแต่อาจจะแตกต่างกันออกไปบ้างในรายละเอียดของข้อมูลเป้าหมาย ยกตัวอย่าง Web site ซึ่งดำเนินธุรกิจค้าปลีกอาจจะให้ความสำคัญกับรหัส Barcode สินค้าที่ลูกค้าสนใจมากเป็นพิเศษ ในขณะที่ Web site ให้บริการด้านการเงินอาจจะให้ความสำคัญกับข้อมูลการโอนเงินสด หรือรหัสเลขหมายบัตรเครดิต ส่วน Web site ของหน่วยงานบริการด้านโทรคมนาคมอย่างองค์กรโทรศัพท์ ก็อาจจะมุ่งเน้นความสนใจไปที่บันทึกการติดต่อทางโทรศัพท์ของลูกค้าแต่ละคน ฯลฯ

เมื่อแต่ละองค์กรธุรกิจวิเคราะห์ได้แล้วว่า ข้อมูลอะไรกันแน่ที่ตนเองต้องการรับทราบจากลูกค้าของตนภารกิจถัดมาที่ผู้บริหารระบบสารสนเทศขององค์กรธุรกิจนั้นจะต้องกระทำก็คือ การ



ออกแบบ Web site ให้มีองค์ประกอบซึ่งเอื้อแก่การเก็บรวบรวมข้อมูลที่ต้องการจากผู้ที่จะแวะเวียนเข้ามาเยี่ยมชม ตัวอย่างขององค์ประกอบ Web site ที่ว่านี้ได้แก่ ไฟล์บันทึกการเรียกใช้บริการเครื่องแม่ข่าย ไฟล์คุกกี้ของ Web site (Site cookie file) แบบฟอร์มสมาชิก (Registration form) และรายงานสภาพการจราจรข้อมูลของ Web site (Web site traffic report) ฯลฯ

โดยในส่วนของการจัดสร้างระบบรวบรวมข้อมูลสภาพการจราจรบน Web site นี้ หากเกินขีดความสามารถของระบบ Computer ภายในองค์กร ผู้บริหารสารสนเทศขององค์กรก็อาจจะเรียกใช้บริการจากองค์กร Software ที่ให้บริการวิเคราะห์สภาพการจราจรของ Web site หรืออาจจะหา Software ประเภท web tracking software มาลองใช้กับข้อมูลรายงานจราจรที่ผ่านเข้า/ออก Web site ของตนเองด้วยเช่นกัน แต่รายงานจราจรข้อมูลที่ได้จากโปรแกรมกลุ่มนี้ยังให้รายละเอียดของลูกค้าได้ไม่ชัดเจนนัก แคบกว่ามีลูกค้า IP address อะไรติดต่อเข้ามาดูเมื่อไหร่และเปิดไปดูส่วนไหนของ Web site เท่านั้น

ที่สำคัญการได้รับข้อมูลลูกค้าแค่ที่ใช้ IP address อะไรติดต่อเข้ามาเยี่ยมชม Web site ก็ยังเป็นข้อมูลที่ค่อนข้างหยาบแทบจะไม่ได้บอกอะไรแก่เจ้าของ Web site เลย เพราะผู้ใช้บริการ Internet หลายๆ รายอาจจะใช้เครื่อง Computer เครื่องเดียวกัน เข้ามาเยี่ยมชม Web site ภายใต้เทคนิคการระบุค่า IP address แบบ dynamic Ips ยิ่งในกรณีของสถานศึกษา หรือหน่วยงานที่มีการติดตั้ง proxy server ไว้ในเครือข่ายเน็ตเวิร์คของหน่วยงานด้วยแล้ว รหัส IP address แทบจะไม่มีคามหมายอะไรเลย เนื่องจากผู้คนนับเป็นร้อยเป็นพันในสถาบันศึกษาแห่งนั้น อาจจะติดต่อเข้าสู่ Internet ภายใต้รหัส IP address เดียวกันหมด ถ้าเช่นนั้นแล้วเราจะใช้วิธีใด