

ระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ

System of Book Recommendation in Library by Using Natural Language Processing

ณัฐสุดา คำเพราะ* และ มหศักดิ์ เกตุจำ

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

1518 ถนนประชากรราษฎร์ 1 แขวงวงศ์สว่าง เขตบางซื่อ กรุงเทพฯ 10800 E-mail: yui.kumpaor@gmail.com

Natsuda Kumpaor^{*} and Mahasak Ketcham

Faculty of Information Technology, King Mongkut's University of Technology North Bangkok

1518 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok 10800, THAILAND.

E-mail: yui.kumpaor@gmail.com

บทคัดย่อ

บทความนี้มีวัตถุประสงค์ 1) เพื่อพัฒนาระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ 2) เพื่อประเมินประสิทธิภาพระบบที่พัฒนาขึ้นและ 3) เพื่อประเมินความพึงพอใจของระบบที่พัฒนาขึ้น ดังนั้นผู้จัดทำจึงได้พัฒนาระบบขึ้นในรูปแบบเว็บแอปพลิเคชัน โดยใช้การเขียนโปรแกรมภาษา PHP และใช้ MySQL ในการจัดการฐานข้อมูลในส่วนของการทดสอบและการประเมินประสิทธิภาพของระบบที่พัฒนาขึ้น พบว่า การประเมินประสิทธิภาพ โดยการทดสอบหาค่าความแม่นยำจากการทดสอบระบบได้เท่ากับร้อยละ 98.72 ซึ่งมากกว่าเกณฑ์ของค่าความแม่นยำคือร้อยละ 97.5 และประเมินความพึงพอใจของระบบโดยผู้เชี่ยวชาญจำนวน 5 คนและผู้ใช้งานทั่วไป 50 คน มีค่าเฉลี่ยระดับคุณภาพเท่ากับ 4.53 (ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.507) ซึ่งสรุปได้ว่าผลในการประเมินความพึงพอใจที่มีต่อระบบที่พัฒนาขึ้นอยู่ในระดับดีมาก

คำสำคัญ: ระบบแนะนำหนังสือ ห้องสมุด การประมวลผลภาษาธรรมชาติ

Abstract

The purposes of this research are 1) to develop recommend in library by using natural language processing, 2) to evaluate the developed system efficiency and 3) to evaluate the satisfactory of the users to the developed system. Therefore, the system was developed as web application by using PHP programming language and MySQL as a database management. The accuracy of the developed system was tested by 5 experts and 50 general users. The result showed that the developed system provided 98.72

percent accuracy, which is more than the set up accuracy threshold of 97.5 percentages. The overall satisfaction of the system evaluated by 5 experts and 50 general users was in the highest level with the mean score of 4.53 (the standard deviation of 0.507)

Keywords: Book recommendation system, library, natural language processing

1. บทนำ

เทคโนโลยีในปัจจุบันการใช้งานทางอินเทอร์เน็ตเข้ามามีบทบาทในการใช้ชีวิตของคนเรามากขึ้นทุกวัน รวมไปถึงการทำกิจกรรมทางด้านต่างๆ ก็ใช้อินเทอร์เน็ตเป็นการสื่อสารข้อมูลในการหาความรู้และการสืบค้น ซึ่งการสื่อสารที่สามารถทำการติดต่อได้อย่างรวดเร็วในการเข้าถึงระบบการทำงานในด้านการทำงาน เช่น การสืบค้นการจัดหมวดหมู่ของข้อมูล หรือแม้กระทั่งการทำงานในระบบออนไลน์ เพื่อสนับสนุนการสืบค้นข้อมูลของผู้ใช้งานได้อย่างรวดเร็วและมีประสิทธิภาพมากขึ้น โดยการทำงานหลากหลายวิธี เช่น การค้นหาคำสำคัญ การสืบค้นข้อมูล การวิเคราะห์ข้อมูล การหาความหมาย เป็นต้น

ผู้วิจัยพบว่าเทคนิคการแนะนำหนังสือห้องสมุดที่จะศึกษา คือ Natural language processing [1] โดยมีการประยุกต์ใช้งานตัวอย่าง เช่น ในด้านการสืบค้นข้อมูลโดยภาษาธรรมชาติ การแปลภาษาด้วยคอมพิวเตอร์เป็นข้อมูลข่าวสารสำหรับการแนะนำหนังสือห้องสมุด ไม่ว่าจะอยู่ส่วนไหนของห้องสมุดก็จะสามารถเข้าถึงข้อมูลทรัพยากรได้ในเวลาไม่นาน แต่การที่จะเข้าถึงข้อมูลทรัพยากรที่ต้องการอย่างรวดเร็วและตรงกับความต้องการของผู้ใช้ไม่ใช่เรื่องง่ายเนื่องจากข้อมูลที่มีอยู่นั้นกระจัดกระจายอย่างไม่เป็นระบบ จึงต้องหาข้อมูลเกี่ยวกับชื่อหนังสือในการค้นหาและผลลัพธ์ที่ได้อาจจะไม่ตรงกับสิ่งที่ต้องการค้นหา ตัวอย่างเช่น ถ้าต้องการค้นหาหนังสือ “วัสดุวิศวกรรม” ผลการค้นหาอาจจะแสดงในส่วนของ “วิศวกรรม” ออกมาซึ่งไม่ตรงกับสิ่งที่เราต้องการสืบค้น หรือถ้าผู้ใช้งานต้องการหนังสือเรื่องหนึ่ง แต่ไม่สามารถคิดคำสำคัญออกมาได้ เช่น คิดชื่อเรื่องไม่ออก ระบบก็จะไม่สามารถแสดงข้อมูลของผู้ใช้ที่ต้องการได้ ซึ่งการพัฒนาที่ผู้วิจัยสนใจนั้น ผู้ใช้งานไม่ต้องใส่คำสำคัญโดยสมาชิกจะต้องให้คะแนนหนังสือแต่ละเรื่องที่มีความสนใจก่อนและหลัง นอกจากนั้นระบบก็จะทำการตรวจสอบความสำคัญของรายการหนังสือที่เราสนใจอันดับแรก และทำการเน้นข้อความหรือตำแหน่งของคำ เหมือนกับการค้นข้อมูล และทำการแนะนำในระบบต่อไป โดยระบบแนะนำหนังสือนั้นจะทำการให้โดยอัตโนมัติ และตรงกับความต้องการของผู้ใช้งานได้รวดเร็วและสะดวกโดยไม่ต้องเสียเวลาในการสืบค้นข้อมูลอีกด้วย ระบบแนะนำทรัพยากรห้องสมุดเป็นหนึ่งในเทคโนโลยีที่ทันสมัยใหม่ที่สามารถนำมาใช้ในการแนะนำข้อมูลอื่นๆ ได้อีกด้วย

ดังนั้นผู้วิจัยจึงนำเทคนิคการแนะนำหนังสือห้องสมุดที่ศึกษา คือ Natural language processing โดยมีการประยุกต์ใช้งานด้านการสืบค้นข้อมูล โดยใช้ภาษาธรรมชาติทำการวิเคราะห์จากข้อมูลว่ามีความสัมพันธ์กันกับข้อมูลของผู้ใช้งานหรือไม่

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบให้การแนะนำ

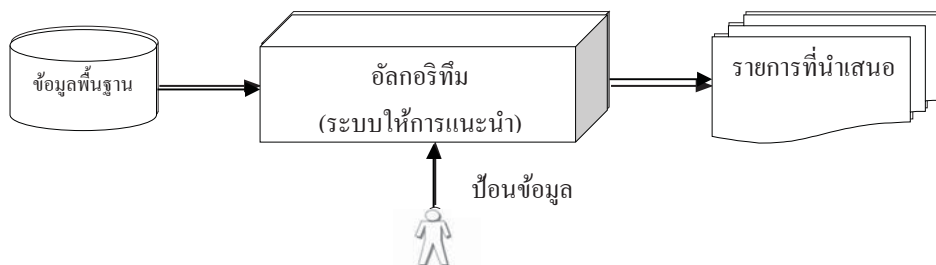
ระบบให้การแนะนำ (Recommender System) เป็นเทคโนโลยีสมัยใหม่ที่ถูกนำมาใช้ในการแนะนำข้อมูลต่างๆ ที่คาดว่าผู้ใช้น่าจะสนใจ หรืออาจจะเป็นข้อมูล que ผู้ใช้ต้องการและเหมาะสำหรับการดำเนินธุรกิจแบบพาณิชย์อิเล็กทรอนิกส์ (e-Commerce) [2] โดยทั่วไประบบให้การแนะนำประกอบด้วย 3 ส่วน ดังต่อไปนี้ คือ

- 1) ส่วนข้อมูลพื้นฐานที่จำเป็นต้องใช้ในการประมวลผล เช่น โปรไฟล์ของผู้ใช้แต่ละคน
- 2) ส่วนการป้อนข้อมูล เป็นข้อมูลที่ได้จากการป้อนข้อมูลเข้ามาของผู้ใช้ เช่น การให้คะแนนเรตติ้ง ซึ่งมีอยู่ 2 แบบ ดังนี้

- แบบชัดเจน (Explicit) ส่วนเรตติ้งแบบชัดเจนจะแสดงอยู่ในรูปของจำนวนตัวเลขตามระดับความนิยมตั้งแต่ 1 ถึง 5, 1 ถึง 10 หรือระดับอื่นๆ ขึ้นอยู่กับการใช้งาน

- แบบไม่ชัดเจน (Implicit) ส่วนเรตติ้งแบบไม่ชัดเจนได้มาจากพฤติกรรมการใช้งานของผู้ใช้ต่างๆ เช่น ประวัติการซื้อสินค้าหรือประวัติการเข้ามาใช้งานของผู้ใช้ในอดีตที่ผ่านมา

- 3) ส่วนอัลกอริทึมเป็นส่วนสำคัญที่สุดที่ใช้ประมวลผลข้อมูลเพื่อให้การแนะนำขึ้นข้อมูลออกมาแสดงไว้ในรูปที่ 1 แสดงระบบที่ช่วยในการแนะนำหรือนำเสนอข้อมูลให้แก่ผู้ใช้โดยระบบจะทำการวิเคราะห์และนำเสนอข้อมูลที่คาดว่าผู้ใช้น่าจะสนใจ หรือคาดว่าจะเป็นข้อมูลที่ผู้ใช้นั้นกำลังต้องการอย่างอัตโนมัติ



รูปที่ 1 สถาปัตยกรรมพื้นฐานของระบบให้การแนะนำ

(ที่มา: http://www.ctc.ac.th/mcu/%20MPU/clip_image20_000.jpg)

2.2 การประมวลผลภาษาธรรมชาติ (Natural Language Processing)

การประมวลผลภาษาธรรมชาติเป็นระบบที่จะช่วยทำให้คอมพิวเตอร์เข้าใจภาษาธรรมชาติของมนุษย์ กระบวนการของการประมวลผลภาษาธรรมชาติจะเริ่มจากการรับอินพุตที่เป็นข้อความภาษาธรรมชาติ จากนั้นระบบจะมีการวิเคราะห์ทั้งในเชิงโครงสร้างและความหมายทางภาษาโดยการทำงานร่วมกับฐานความรู้แล้วทำการประเมินค่า เพื่อหาคำตอบแล้วทำการแทนความหมายและสร้างเป็นเอาต์พุตที่เป็นภาษาธรรมชาติออกมาใน

ภาษาธรรมชาติคำที่เขียนหรือเปล่งเป็นคำพูดออกมาเมื่อนำมารวมกันจะเกิดเป็นประโยคที่มีความหมายขึ้น คอมพิวเตอร์จำเป็นต้องมีกระบวนการต่างๆ เพื่อจะได้เข้าใจความหมายของแต่ละประโยค

2.2.1 ความแตกต่างของภาษาคอมพิวเตอร์กับภาษามนุษย์

ภาษาคอมพิวเตอร์เป็นระบบที่มีการกำหนดขอบเขตไว้ในกรอบจำกัด มีการใช้คำจำกัด มีไวยากรณ์ที่ใช้จำกัดและการตีความหมายที่ชัดเจนจึงสามารถเรียกภาษาคอมพิวเตอร์อีกอย่างหนึ่งว่าภาษาที่มีรูปแบบ (Formal Language) ส่วนภาษาธรรมชาติเป็นภาษาที่มีขอบเขตกว้างมากจนยากที่จะหาวิธีแบบที่ตายตัวได้ กฎเกณฑ์ของภาษาธรรมชาติเป็นกฎเกณฑ์ที่เกิดขึ้นในการใช้ภาษาและเป็นที่ยอมรับของกลุ่มชนผู้ใช้นั้นๆ

2.2.2 การแบ่งแยกหน่วยย่อยของภาษา (Natural Language Entity)

การแบ่งแยกหน่วยย่อยของภาษานั้นจะประกอบด้วย

- 1) ตัวอักษร (Alphabet) คือ สัญลักษณ์ที่ใช้แทนเสียง
- 2) คำ (Word) คือ กลุ่มของตัวอักษรที่มาเรียงต่อกันเป็นคำ
- 3) ประโยค (Sentence) คือ กลุ่มของคำที่นำมาเรียงต่อกันเพื่อแทนความหมายประโยคจึงเป็นข้อความ

2.2.3 การวิเคราะห์ในเชิงโครงสร้าง (Syntactic Analysis)

การวิเคราะห์ในเชิงโครงสร้างเป็นการตรวจสอบโครงสร้างทางไวยากรณ์เกี่ยวกับการวางตำแหน่งของกลุ่มคำประเภทต่างๆ ที่รวมกันเป็นประโยค ในกรณีที่ประโยคอินพุตที่รับเข้ามาไม่ถูกต้องตามหลักไวยากรณ์ คอมพิวเตอร์ควรจะบอกได้ว่าประโยคที่ผิด เช่น ประโยค “The man old cried” เป็นประโยคที่มีโครงสร้างผิดหลักไวยากรณ์เนื่องจากลำดับที่ถูกต้องจะเป็น

ประโยค --> นามวลี + กริยาวลี

นามวลี --> คำนำหน้านาม + คำคุณศัพท์ + คำนาม

กริยาวลี --> กริยา

ดังนั้นประโยคที่ถูกต้องควรจะเป็น “The old man cried”

2.2.4 การวิเคราะห์ในเชิงความหมาย (Semantic Analysis)

การวิเคราะห์ในเชิงความหมายเป็นการตรวจสอบความถูกต้องในเชิงความหมายของประโยค โดยประโยคที่วางกลุ่มคำชนิดต่าง ๆ ตามโครงสร้างไวยากรณ์จะมีความหมายอย่างไรอย่างหนึ่งแน่นอน แต่ในบางครั้งประโยคที่กำลังพิจารณาอาจจะเขียนถูกต้องตามหลักไวยากรณ์แต่มีความหมายกำกวม หรือเป็นความหมายที่เป็นไปไม่ได้ หรือไม่ให้ความหมายอะไรเลย เช่น “The stones eat the boys” จะเห็นว่าประโยคนี้นี้โครงสร้างของประโยคถูกต้องตามหลักไวยากรณ์ คือ

ประโยค --> นามวลี + กริยาวลี

นามวลี --> คำนำหน้านาม + คำนาม

กริยาวลี --> กริยา + นามวลี

แต่เมื่อวิเคราะห์ดูความหมายแล้วเห็นว่าประโยคนี้มีความหมายที่เป็นไปไม่ได้ ในเมื่อหินเป็นสิ่งไม่มีชีวิต จึงทำกริยา “กิน” ไม่ได้

2.2.5 การวิเคราะห์ในเชิงตีความ (Pragmatic Analysis)

ประโยคที่เราพูดออกมาบางครั้งก็อาจจะไม่ได้มีความหมายตรงตามข้อความนั้นๆ ซึ่งจะต้องตีความตามสถานการณ์ที่เกิดขึ้น โดยที่ทั้งผู้ส่งข่าวสารและผู้รับข่าวสารจะต้องอยู่ในสถานการณ์เดียวกัน เช่น สมมติว่าตอนนี้เรายู่ที่สถานีรถไฟและกำลังกังวลว่าขณะนี้เวลาเท่าไรรถไฟใกล้จะออกหรือยังแต่เราไม่มีนาฬิกาพอคิหันไปเห็นคนข้างๆ กำลังดูตารางเวลาการเดินทางเช่นเดียวกัน เราเลยหันถามว่า “Do you have a watch?” ถ้าเราได้คำตอบว่า “yes” หรือ “no” แสดงว่าคำตอบที่ได้ผิด เพราะคำตอบที่เราต้องการจริงๆ คือ เวลา ณ ขณะนี้

2.2.6 การใช้งานการประมวลผลภาษาธรรมชาติ [3]

ในปัจจุบันได้มีการนำเรื่องของ การประมวลผลภาษาธรรมชาติมาประยุกต์ใช้ในงานด้านต่างๆ ดังตัวอย่างต่อไปนี้

1) การสืบค้นข้อมูลโดยใช้ภาษาธรรมชาติ

เป็นลักษณะของการสืบค้นข้อมูลในฐานะข้อมูลด้วยภาษาที่ใช้ในชีวิตประจำวันแทนการใช้ภาษาคอมพิวเตอร์ เช่น ถ้าต้องการหนังสือจากฐานข้อมูลแทนที่จะต้องมาเขียนคำสั่ง SQL ก็สามารถใช้คำถามที่เป็นภาษาธรรมชาติได้ โดยจะต้องมีการแปลงจากภาษาธรรมชาติที่ใส่เข้าไปให้อยู่ในรูปของภาษา SQL ที่เครื่องเข้าใจ โดยแนวคิดของการแปลงจากภาษาธรรมชาติให้เป็นภาษา SQL นั้นมีหลายแนวคิด เช่น Template-Based Approach แนวความคิดนี้จะเป็นการสร้างต้นแบบไว้ก่อน โดยไม่จำเป็นต้องรู้กฎไวยากรณ์ของภาษา เพียงแค่ข้อความที่เข้ามาสามารถจับคู่กับต้นแบบได้เท่านั้น จึงเป็นแนวความคิดที่ฟังปรารถนา แต่ในความเป็นจริงแล้วคงเป็นไปได้ยากที่จะสร้างต้นแบบไว้ให้ครอบคลุมทุกคำถามที่จะใช้ในการสืบค้นข้อมูล ดังนั้นแนวความคิดแบบนี้จึงอาจจะเป็นไปได้ในฐานข้อมูลที่มีขนาดเล็กเท่านั้น

2) การแปลภาษาด้วยเครื่องคอมพิวเตอร์

การแปลภาษาด้วยเครื่องคอมพิวเตอร์ (Machine Translation: MT) เป็นการนำเครื่องคอมพิวเตอร์เข้าใช้ในการแปลจากภาษาหนึ่งเป็นอีกภาษาหนึ่งตามที่ต้องการ เช่น แปลจากภาษาอังกฤษเป็นภาษาไทย เป็นต้น ในการพัฒนาระบบการแปลมีเทคนิคที่ใช้ในขั้นตอนการแปลหลายเทคนิค เช่น ระบบการแปลโดยตรง (Direct MT) วิธีการจะเป็นลักษณะของการแปลคำต่อคำหรือวลีต่อวลีโดยที่คุณภาพของการแปลขึ้นอยู่กับคุณภาพของพจนานุกรมของทั้งภาษาต้นฉบับและภาษาเป้าหมาย

3) การสรุปใจความสำคัญจากเอกสาร (Text Summarization)

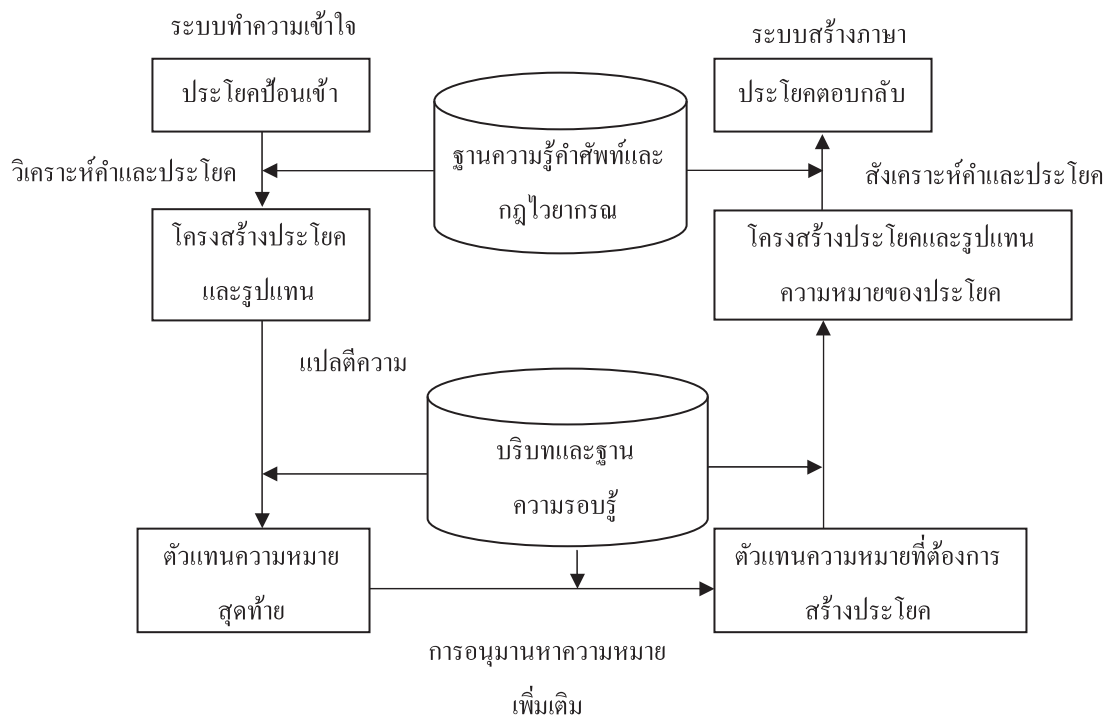
เป็นการสรุปใจความสำคัญจากเอกสารเอาเฉพาะส่วนที่สำคัญสร้างเป็นเอกสารใหม่ที่สั้นกว่าเดิมแต่มีเนื้อหาข้อมูลที่เหมือนต้นฉบับ โดยใช้วิธีการต่างๆ เช่น Text Extraction Technique วิธีนี้จะมีการให้คะแนนประโยคหรือข้อความโดยดูจากสำคัญของประโยคหรือข้อความและนำมาสร้างเป็นเอกสารใหม่โดยคะแนน

ที่ให้อาจจะพิจารณาจากความถี่ของคำที่ปรากฏในเอกสาร การเน้นข้อความหรือตำแหน่งของคำเสมือนกับการค้นข้อมูลโดยใช้สืบค้น โดยการคัดคำสำคัญและได้ข้อมูลที่พบตามคำสำคัญนั้นออกมาซึ่งเป็นการนำ Text Summarization มาใช้เป็นต้น

การเน้นข้อความหรือตำแหน่งของคำ เสมือนกับการค้นข้อมูลโดยใช้การสืบค้นโดยการคัดคำสำคัญและได้ข้อมูลที่พบตามคำสำคัญนั้นออกมาซึ่งเป็นการนำเอกสารเชิงความหมาย มาใช้เป็นต้น

2.2.7 สถาปัตยกรรมของระบบประมวลผลภาษาธรรมชาติ [4-5]

งานวิจัยทางการประมวลผลภาษาธรรมชาติในที่นี้ หมายถึงภาษาข้อความอักษรแบ่งเป็น 2 กลุ่มใหญ่ๆ คือระบบทำความเข้าใจภาษาและระบบสร้างภาษาทั้ง 2 ระบบต่างต้องใช้ฐานความรู้ขนาดใหญ่ในการประมวลผลเพื่อใช้ในการแปลความหมายที่แท้จริงของประโยคหรือวางลำดับในการสร้างประโยคดังแสดงในรูปที่ 2

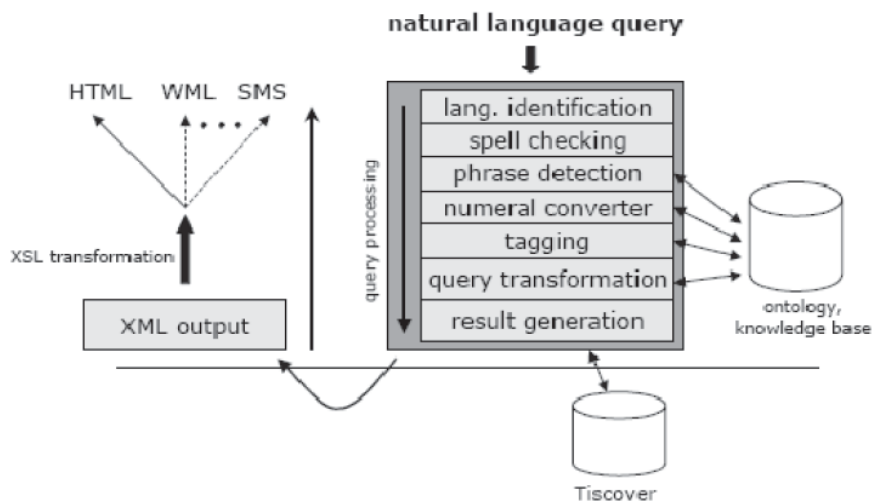


รูปที่ 2 สถาปัตยกรรมของระบบประมวลผลภาษาธรรมชาติ
(ที่มา : <http://www.viospeed.com/wp/uploads/2014/Raspi01.gif>)

2.2.8 ภาษาธรรมชาติกับการใช้ภาษาคอมพิวเตอร์ในการสืบค้นข้อมูล [6]

ภาษาธรรมชาติคือภาษาที่มนุษย์ใช้ในการติดต่อสื่อสารกันนั่นเอง ตัวอย่างเช่น ภาษาไทยภาษาอังกฤษ ภาษาญี่ปุ่น เป็นต้น ภาษาธรรมชาติต่างจากภาษาคอมพิวเตอร์ เพราะภาษาคอมพิวเตอร์มีไวยากรณ์แบบตายตัวแต่การประมวลแบบภาษาธรรมชาติมีไวยากรณ์ที่ซับซ้อนกว่ามาก คำว่าภาษาก็หมายถึง ภาษาคอมพิวเตอร์ต่างๆ เช่น ภาษาซี ภาษาฟอร์แทรน ภาษาโคบอล เป็นต้น ภายหลังได้มีความพยายามที่จะนำเอาคอมพิวเตอร์มาใช้ในการประมวลภาษามนุษย์ จึงทำให้มีการใช้คำว่า “ภาษาธรรมชาติ” เพื่อให้แตกต่างไปจากคำว่า “ภาษา” ซึ่งจะหมายถึง “ภาษาคอมพิวเตอร์”

Natural Language Query Interface คือ การใช้ภาษาคอมพิวเตอร์ในการสืบค้นข้อมูล ที่คอมพิวเตอร์สามารถเข้าใจภาษาธรรมชาติได้นั้นคอมพิวเตอร์ต้องสามารถแปลงภาษาธรรมชาติเป็น Natural Language Query Interface คือ การใช้ภาษาคอมพิวเตอร์ในการสืบค้นข้อมูลที่ภาษาเครื่องและสามารถแปลงภาษาเครื่องกลับมาเป็นภาษาธรรมชาติเพื่อโต้ตอบกับมนุษย์ ซึ่งจะเป็นการใช้ภาษาธรรมชาติมาสั่งการเรียกใช้ฐานข้อมูลแทนที่จะใช้ภาษาคอมพิวเตอร์ทั่วไป ซึ่งจะเป็นลักษณะของการถามหาข้อมูลด้วยภาษาธรรมชาติที่ใช้ในชีวิตประจำวัน โดยจะมีการแปลตีความและสร้างเป็นคำสั่ง (query) เพื่อเรียกใช้ฐานข้อมูล Natural Language Query Process คือ กระบวนการต่างๆ ที่เกิดขึ้นระหว่างการทำการประมวลผลคำสั่งภาษาธรรมชาติ ซึ่งสถาปัตยกรรมทางด้านซอฟต์แวร์ได้รับการออกแบบตามโครงสร้างแบบ Pipeline [7] ดังแสดงในรูปที่ 3



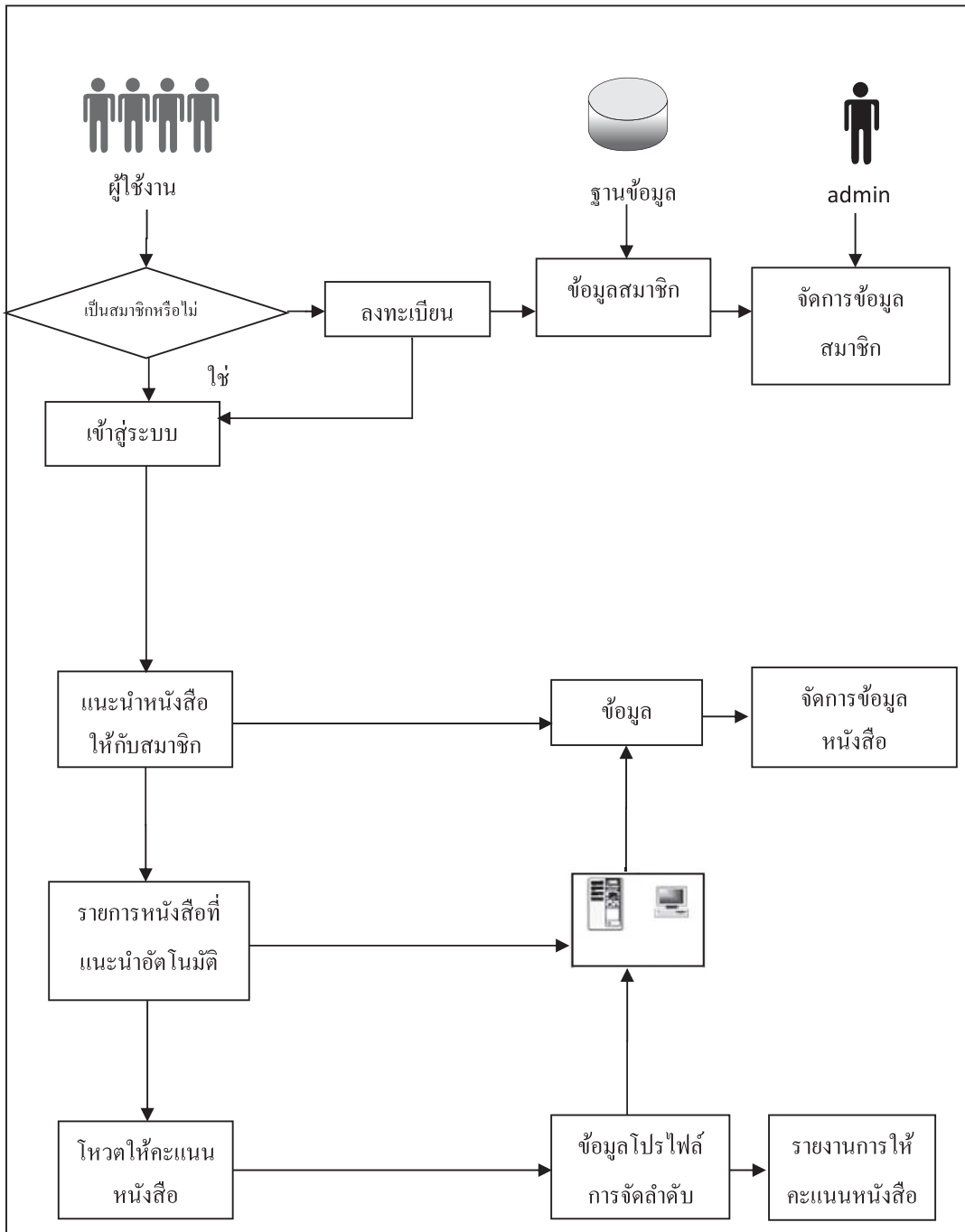
รูปที่ 3 สถาปัตยกรรมทางด้านซอฟต์แวร์ในการประมวลผลคำสั่งภาษาธรรมชาติ [4]

3. วิธีการดำเนินงานวิจัย

วิธีการดำเนินงาน ของระบบแนะนำหนังสือห้องสมุดโดยใช้หลักการการประมวลผลภาษาธรรมชาติ (Natural Language Processing) แบ่งเป็น 2 ส่วนดังนี้

3.1 การวิเคราะห์และออกแบบระบบ

ผู้วิจัยวิเคราะห์ภาพรวมการใช้งานระบบโดยมีการออกแบบการใช้งานของระบบแสดงในรูปที่ 4



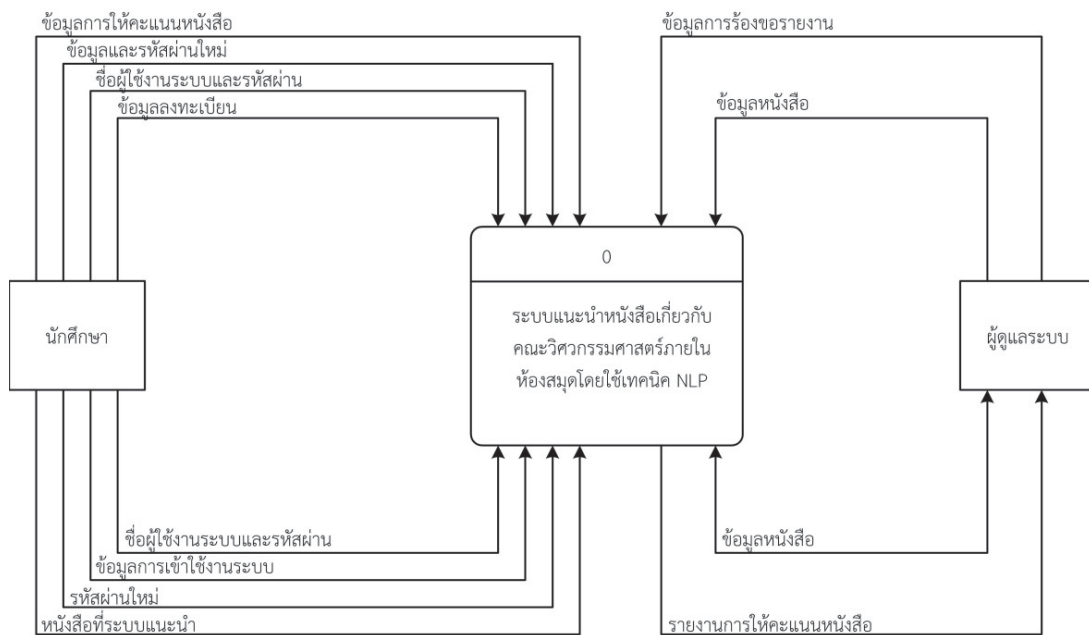
รูปที่ 4 ขั้นตอนการทำงานของระบบ

จากรูปที่ 4 อธิบายภาพรวมการใช้งานโปรแกรม โดยสมาชิกทำการเข้าระบบ และสมัครข้อมูลเพื่อใช้ในการสำรวจและเก็บข้อมูลรายละเอียดของสมาชิก เพื่อใช้แนะนำหนังสืออัตโนมัติให้ตรงกับผู้ใช้งานของระบบโดยเทคนิคการประมวลผลด้วยภาษาธรรมชาติ

3.2 การวิเคราะห์และการออกแบบระบบ

การออกแบบระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ ขั้นตอนนี้นับเป็นขั้นตอนที่สำคัญที่จะทำให้การพัฒนาระบบมีประสิทธิภาพ โดยการวิเคราะห์จะเกี่ยวข้องกับการออกแบบแผนภาพรายละเอียดต่างๆ ของการดำเนินงานและสร้างแผนภาพการทำงานต่างๆ เพื่อให้ง่ายต่อความเข้าใจ รวมทั้งการวิเคราะห์โครงสร้างของฐานข้อมูลและความสัมพันธ์ของแฟ้มข้อมูลต่างๆ ให้ระบบมีความสอดคล้องกันในการทำงาน และสามารถนำไปใช้ในการออกแบบ เพื่อให้การทำงานของระบบมีประสิทธิภาพสูงสุด

แผนภาพกระแสข้อมูลระดับสูงเป็นการออกแบบแผนภาพการไหลของข้อมูลระดับบนสุด ที่แสดงภาพรวมการทำงานของระบบที่มีความสัมพันธ์กับเอนทิตีที่เกี่ยวข้องโดยตรงกับระบบ ดังแสดงไว้ในรูปที่ 5



รูปที่ 5 แผนภาพกระแสข้อมูลระดับสูง (Context Diagram)

จากรูปที่ 5 แผนภาพ Context Diagram ข้างต้นสามารถอธิบายรายละเอียดได้ดังนี้ กลุ่มบุคคลที่มีส่วนเกี่ยวข้องกับระบบมี 2 กลุ่ม คือ ส่วนของนักศึกษา (ผู้ใช้งานระบบ) และส่วนของเจ้าหน้าที่ (ผู้ดูแลระบบ) ส่วนของนักศึกษา (ผู้ใช้งานระบบ) เป็นผู้เข้ามาใช้บริการของระบบ โดยสามารถจัดการข้อมูลส่วนตัว และแก้ไขรหัสผ่านลงคะแนนความชอบต่อหนังสือแต่ละเรื่อง หลังจากนั้นระบบจะแนะนำหนังสือให้โดยอัตโนมัติ ส่วนของ

เจ้าหน้าที่ (ผู้ดูแลระบบ) เป็นผู้ที่ มีหน้าที่ในการจัดการข้อมูลต่างๆ ของระบบ โดยสามารถเพิ่ม ลบ แก้ไข ข้อมูล หนังสือ และรายงานการให้คะแนนหนังสือแต่ละเรื่องได้

3.3 การพัฒนาระบบ

จากขั้นตอนศึกษาและรวบรวมข้อมูลระบบ วิเคราะห์ระบบ ออกแบบระบบ จนถึงขั้นตอนการออกแบบ ส่วนผู้ใช้ ทำให้ทราบถึงวิธีการและขั้นตอนต่างๆ ในการพัฒนาระบบ ซึ่งในการพัฒนาระบบแนะนำหนังสือ ห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ จะมีการดำเนินการพัฒนาตามขั้นตอนในแต่ละกระบวนการ โดยในกระบวนการศึกษาข้อมูลนั้นจะเป็นการศึกษาข้อมูลที่มีส่วนเกี่ยวข้องในการพัฒนาระบบและนำข้อมูล เหล่านั้นมาออกแบบกระบวนการทำงานของระบบได้อย่างถูกต้อง

ในการพัฒนาระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ พัฒนาโดยใช้ โปรแกรมภาษา PHP และใช้ระบบฐานข้อมูลมายเอสคิวแอล (MySQL) ในการเก็บข้อมูล ในส่วนของการออกแบบ หน้าจอเพื่อติดต่อกับใช้งาน (Graphic User Interface: GUI) โดยใช้โปรแกรม Adobe Photoshop CS3 และ Adobe Dreamweaver CS3 ในการออกแบบเพื่ออำนวยความสะดวกให้แก่ผู้ใช้และในส่วนของการแนะนำนั้น ผู้วิจัยได้นำ เทคนิคที่ใช้ในการให้คำแนะนำ คือ เทคนิคการประมวลผลภาษาธรรมชาติ (Natural Language Processing) พัฒนาเพื่อให้ระบบเป็นไปตามขอบเขตของการศึกษา

3.4 การหาคุณภาพของเครื่องมือ

ผู้วิจัยระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติได้นำแบบสอบถามที่ สร้างขึ้นไปทำการทดสอบหาค่าเที่ยงตรงของแบบสอบถาม (Validity) และความเชื่อมั่น (Reliability) ดังนี้

1) การหาค่าความเที่ยงตรงของแบบสอบถาม

ผู้วิจัยระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติได้ใช้วิธีการหาคุณภาพ ของเครื่องมือ โดยใช้วิธี IOC : Index of item objective congruence เพื่อหาค่าความเที่ยงตรงของแบบสอบถาม หรือค่าสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ของแบบประเมินคุณภาพของระบบ โดยผู้เชี่ยวชาญและแบบ ประเมินความพึงพอใจของระบบโดยผู้ใช้งานทั่วไป ในการหาค่าความเที่ยงตรงผู้พัฒนาระบบได้นำแบบสอบถาม ที่สร้างขึ้นไปตรวจสอบความเที่ยงตรงโดยผู้เชี่ยวชาญที่ถูกแต่งตั้ง จำนวน 5 ท่าน

หลังจากนั้นนำมาหาค่าดัชนีความสอดคล้อง (IOC) ดังสมการที่ 1

$$IOC = \frac{\sum R}{N} \quad (1)$$

เมื่อ IOC หมายถึงค่าดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์

R หมายถึงความคิดเห็นของผู้เชี่ยวชาญ โดยที่ ค่า +1 หมายถึง ข้อคำถามวัดได้ตรงตามวัตถุประสงค์ ค่า 0 หมายถึง ไม่แน่ใจว่าข้อคำถามวัดได้ตรงตามวัตถุประสงค์ และ -1 หมายถึง ข้อคำถามวัดได้ไม่ตรงตาม วัตถุประสงค์

N หมายถึงจำนวนผู้เชี่ยวชาญ

ทั้งนี้ผู้วิจัยระบบเลือกข้อความที่มีค่า IOC มากกว่า 0.5 มาใช้เป็นข้อความจากผู้เชี่ยวชาญทั้ง 5 คน ซึ่งได้ตรวจสอบแบบสอบถามแล้วเห็นว่าแบบสอบถามทุกข้อที่ผู้พัฒนาระบบสร้างขึ้นมีความเที่ยงตรงของเนื้อหาครอบคลุมในแต่ละด้านและครอบคลุมวัตถุประสงค์ของการพัฒนาระบบ โดยมีค่าดัชนีความสอดคล้องดังนี้ แบบประเมินคุณภาพของระบบโดยผู้เชี่ยวชาญ มีค่าดัชนีความสอดคล้องที่ 0.95 และแบบประเมินความพึงพอใจของระบบโดยผู้ใช้งานทั่วไป มีค่าดัชนีความสอดคล้องที่ 0.98

2) การหาความเชื่อมั่นของแบบสอบถามทั้งฉบับ (Reliability)

ผู้วิจัยระบบนำแบบสอบถามทั้ง 5 ส่วน ไปทดสอบให้กับผู้ใช้งานทั่วไป จำนวน 50 คน แล้วนำข้อมูลที่ได้มาวิเคราะห์หาค่าความเชื่อมั่น (Reliability) โดยหาค่าสัมประสิทธิ์อัลฟาของครอนบาค (Cronbach's alpha coefficient) ได้ค่าความเชื่อมั่นของแบบสอบถามทั้งฉบับเท่ากับ 0.90

3.5 การประเมินผลของระบบ

การประเมินผลของระบบประกอบด้วย

1) การทดสอบระบบ

2) การทดสอบแบบแอลฟา (Alpha Test) เป็นการทดสอบเพื่อหาข้อบกพร่องหรือปัญหาที่เกิดขึ้นกับระบบดำเนินการทดลองใช้ขั้นต้น โดยผู้ศึกษาเอง หลังจากนั้นจึงทำการแก้ไขระบบให้มีความสมบูรณ์ยิ่งขึ้น

3) การทดสอบแบบเบต้า (Beta Test) เป็นการทดสอบความสมบูรณ์ของระบบ โดยผู้วิจัยได้นำระบบไปทำการทดสอบประสิทธิภาพของระบบกับกลุ่มผู้ดูแลระบบ หลังจากนั้นผู้ศึกษาได้ทำการทดสอบ เพื่อหาความพึงพอใจของกลุ่มผู้ใช้งานจำนวน 50 คนและเมื่อทดสอบความพึงพอใจของกลุ่มผู้ใช้แล้ว ผู้ศึกษาได้นำข้อมูลไปแก้ไขปรับปรุงระบบให้ตรงกับความต้องการของกลุ่มผู้ให้มีความสมบูรณ์ยิ่งขึ้น

4) การประเมินระบบ เป็นการประเมินหลังจากที่ได้มีระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติเรียบร้อยแล้ว ผู้วิจัยระบบได้ทำแบบประเมินประสิทธิภาพของระบบโดยผู้ดูแลระบบหลังจากที่ได้ทดสอบระบบ โดยประเมินประสิทธิภาพของผู้ดูแลระบบจากการประเมินผลจะประเมินจากการอินพุตที่รับเข้าไปแล้วประมวลออกมาเป็นผลลัพธ์ได้อย่างถูกต้อง การวัดประสิทธิภาพจากระบบโดยต้องมีค่าความแม่นยำ (Accuracy) เท่ากับ 97.5% หรือ ค่าความคลาดเคลื่อน (error) ไม่เกิน 2.5%

โดยการคำนวณค่าความถูกต้อง/ความแม่นยำใช้สมการที่ (2)

$$\%Accuracy = 100 - \%Error \quad (2)$$

$$\%Error = Relative\ error \times 100$$

โดยที่ $Relative\ error = \left| \frac{X_{mea} - X_t}{X_t} \right|$

เมื่อ x_{mea} คือค่าที่ได้จากการวัด (measure value)
 x_t คือค่าจริง (true value)

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาระบบ

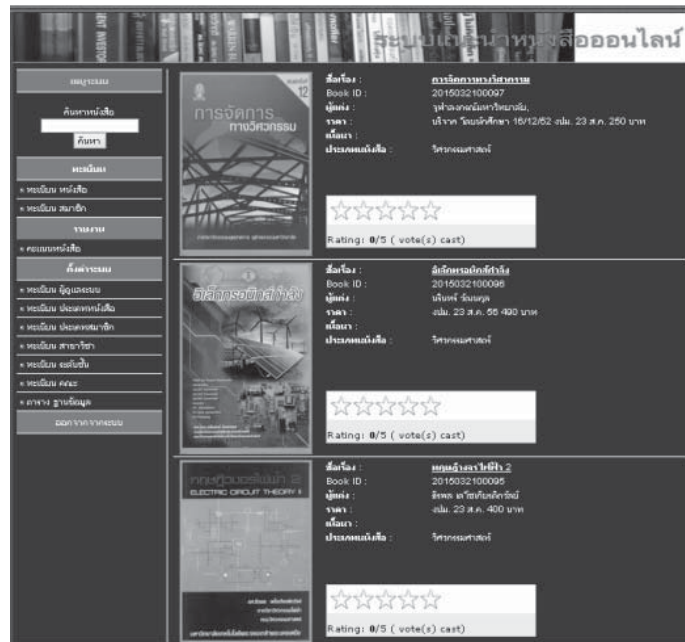
จากการที่ได้ออกแบบและพัฒนาระบบแนะนำหนังสือห้องสมุดโดยใช้เทคนิคการประมวลผลภาษาธรรมชาติ ตามขั้นตอนต่างๆ ที่ผู้พัฒนาได้วางแผนไว้ สามารถอธิบายผลจากการออกแบบและพัฒนาระบบตามสิทธิ์การเข้าใช้งานของระบบเป็นหัวข้อต่างๆ ดังนี้

1) หน้าจอเริ่มต้นเข้าสู่ระบบ ประกอบด้วยรายการหนังสือใหม่ที่แนะนำ ข้อมูลสถิติเยี่ยมชมเว็บไซต์ ส่วนของการป้อนชื่อผู้ใช้ระบบและรหัสผ่าน เพื่อทำการเข้าสู่ระบบ และสมัครสมาชิกลงทะเบียนเพื่อใช้ในการเก็บข้อมูลในการแนะนำหนังสือตามสาขาวิชา ภายในคณะวิศวกรรมศาสตร์ แสดงดังรูปที่ 6



รูปที่ 6 หน้าจอเริ่มต้นเข้าสู่ระบบ

2) หน้าจอแสดงรายการหนังสือแนะนำ เมื่อผู้ดูแลระบบทำการลงทะเบียนหนังสือรายการหนังสือทั้งหมดก็จะแสดงรายการหนังสือที่แนะนำ แสดงดังรูปที่ 7



รูปที่ 7 หน้าจอแสดงรายการหนังสือแนะนำ

3) หน้าจอแสดงระบบสมาชิก เมื่อทำการสมัครสมาชิกเรียบร้อยแล้วระบบจะทำการออกอัตโนมัติ เมื่อเข้ามาอีกกรอบระบบจะทำการแนะนำหนังสืออัตโนมัติ แสดงดังรูปที่ 8



รูปที่ 8 หน้าจอระบบสมาชิกที่เข้าใช้งานระบบ

4.2 ผลการประเมินระบบ

ตารางที่ 1 แสดงผลสรุปการประเมินประสิทธิภาพของระบบ โดยผู้ดูแลระบบจากหัวข้อการประเมินทั้งหมด 3 ด้าน ผลการประเมินประสิทธิภาพของระบบ แสดงดังตาราง ได้แก่ 1) ด้านความสามารถของระบบตรงกับความต้องการ คิดเป็นร้อยละ 99.16 2) ด้านความถูกต้องในการทำงานของระบบ คิดเป็นร้อยละ 98.5 และ 3) ด้านความปลอดภัยของระบบ คิดเป็นร้อยละ 98.5 สรุปผลการทดสอบระบบในการหาประสิทธิภาพ ซึ่งแสดงว่าระบบมีประสิทธิภาพอยู่ในระดับร้อยละ 98.72

การประเมินผลวัดจากประสิทธิภาพของระบบโดยต้องมีค่าความแม่นยำ(Accuracy) เท่ากับ 97.5% หรือมีค่าความคลาดเคลื่อน (error) เท่ากับ 2.5% ของผลลัพธ์จากการสืบค้น

ตารางที่ 1 สรุปผลการประเมินประสิทธิภาพของระบบ โดยผู้ดูแลระบบ

รายการประเมิน	การวัดประสิทธิภาพ			
	ทดสอบ (ครั้ง)	ผ่าน (ครั้ง)	ผิดพลาด	ร้อยละ
1. ด้านความสามารถของระบบตรงกับความต้องการ	100	99.16	0.84	99.16
2. ด้านความถูกต้องในการทำงานของระบบ	100	98.5	1.5	98.5
3. ด้านความปลอดภัยของระบบ	100	98.5	1.5	98.5
สรุปผล	300	296.16	3.84	98.72

ตารางที่ 2 แสดงผลสรุปการประเมินความพึงพอใจของระบบโดยผู้ใช้งาน (ค่าเฉลี่ยระดับคุณภาพเท่ากับ 4.53 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.507) ซึ่งแสดงว่าระบบที่ออกแบบและพัฒนาคุณภาพต่อการใช้งานระดับดีมาก โดยแยกเป็นด้านความสามารถของระบบตรงกับความต้องการ (ค่าเฉลี่ยระดับคุณภาพเท่ากับ 4.48 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.510) ซึ่งอยู่ในระดับดี ด้านความถูกต้องในการทำงานของระบบ (ค่าเฉลี่ยระดับคุณภาพเท่ากับ 4.48 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.510) ซึ่งอยู่ในระดับดี ด้านความสะดวกและง่ายต่อการใช้งานระบบ (ค่าเฉลี่ยระดับคุณภาพเท่ากับ 4.56 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.507) ซึ่งอยู่ในระดับดีมาก ด้านความปลอดภัยของระบบ (ค่าเฉลี่ยระดับคุณภาพเท่ากับ 4.60 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.500) ซึ่งอยู่ในระดับ ดีมาก

ตารางที่ 2 สรุปผลการประเมินความพึงพอใจของระบบโดยผู้ใช้งานจำนวน 50 คน

รายการประเมิน	ระดับความพึงพอใจ		
	\bar{X}	S.D.	เชิงคุณภาพ
ด้านความสามารถของระบบตรงกับความต้องการ	4.48	0.510	ดี
ด้านความถูกต้องในการทำงานของระบบ	4.48	0.510	ดี
ด้านความสะดวกและง่ายต่อการใช้งานระบบ	4.56	0.507	ดีมาก
ระบบด้านความปลอดภัยของระบบ	4.60	0.500	ดีมาก
สรุปผล	4.53	0.507	ดีมาก

5. สรุปและข้อเสนอแนะ

ระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ เป็นระบบที่สามารถช่วยแนะนำหนังสือสำหรับสมาชิกได้อัตโนมัติได้ สำหรับขั้นตอนการพัฒนากระบวนการนั้นจะใช้ภาษาที่ใช้ในการพัฒนาคือ Java, PHP เป็นภาษาหลักในการพัฒนาเว็บแอปพลิเคชัน และใช้ระบบจัดการฐานข้อมูล Microsoft SQL Server 2008 เป็นตัวช่วยบริหารจัดการฐานข้อมูล จากการออกแบบและพัฒนาระบบ ผู้วิจัยได้ทดสอบการใช้งานและประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญ จำนวน 5 คน รวมทั้งประเมินความพึงพอใจของระบบโดยผู้ใช้งานระบบ จำนวน 50 คน โดยการสุ่มอย่างง่ายจากประชากรที่สำรวจความพึงพอใจ

สรุปผลการทดสอบระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติได้ผลการสืบค้นและการแสดงข้อมูลให้ตรงกับผู้ใช้งาน โดยการทดสอบหาค่าความแม่นยำจากการทดสอบระบบได้ร้อยละ 98.72 ซึ่งไม่ต่ำกว่าเกณฑ์ในการคำนวณค่าความถูกต้อง/ความแม่นยำ คือ ร้อยละ 97.5 ดังนั้น การใช้งานของระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ จึงมีความถูกต้องและเชื่อถือได้

ผู้พัฒนาระบบแนะนำหนังสือห้องสมุดด้วยเทคนิคการประมวลผลภาษาธรรมชาติ มีข้อเสนอแนะในการออกแบบและพัฒนาระบบเพิ่มเติม ดังนี้

- 1) เพื่อให้การแนะนำหนังสือให้ตรงกับผู้ใช้งานมากยิ่งขึ้น ระบบควรมีการเก็บข้อมูลพฤติกรรม การสืบค้น และข้อมูลสถิติในการยืมหนังสือห้องสมุดอย่างละเอียด
- 2) ควรศึกษาและพัฒนาระบบโดยใช้เทคนิคหรือการวิเคราะห์ในการออกแบบเพื่อเปรียบเทียบคุณภาพของระบบในการพัฒนาให้มีคุณภาพที่ดียิ่งขึ้น
- 3) ระบบควรทำการเปรียบเทียบคุณภาพการแนะนำหนังสือให้ตรงกับความต้องการของผู้ใช้งานและมีการเปรียบเทียบการแนะนำระบบในการทดสอบระบบทุกครั้ง

เอกสารอ้างอิง

- [1] บุญเสริม กิจศิริกุล. (2545). ปัญญาประดิษฐ์. ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย.
- [2] ชัยวัฒน์ ตรีวีระจร. (2550). การทำนายข้อมูลโดยการรวม Content – based filtering with item – based collaborative filtering ด้วยกฎความสัมพันธ์. วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- [3] วรวิวี เกสร. (2553). การวัดความคล้ายคลึงของเอกสารภาษาไทยโดยใช้การประมวลผลภาษาธรรมชาติ (Similarity Measurement of Thai Document Using Natural Language Processing). สาขาวิชาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่ มหาวิทยาลัยเชียงใหม่.
- [4] วิโรจน์ อรุณมานะกุล. (2545). ภาษาศาสตร์คลังข้อมูล. กรุงเทพฯ: คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- [5] G. Adomavicius, and A. Tuzhilin, (2003). Recommendation Technologies: Survey of Current Methods and Possible Extensions. *Collections of Information Systems Working Papers*. Stern School of Business, New York University. Report No. IS-03-06.
- [6] เทคโนโลยีการประมวลผลภาษาธรรมชาติในปัจจุบัน, [Online] http://www.jaist.ac.jp/~ping/paper/NLP_tis.txt, Last modified: Unknown [October 15, 2014].
- [7] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, “Analysis of Recommendation Algorithms for E-commerce,” *Proceedings of the 2nd ACM Conference on Electronic Commerce*, New York, USA, 2000, pp. 158-167.