

การตรวจสอบอัตนัยภาษาไทยอัตโนมัติด้วยเทคนิคการวัดความคล้ายคลึง ของเอกสาร

บุญโชค ป้ายจุฬาลักษณ์¹ และ อีรพงษ์ สังข์ศรี²

^{1,2}โปรแกรมวิชาวิทยาการสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏนครราชสีมา นครราชสีมา

Emails: p-boy_po@hotmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบต้นแบบสำหรับการตรวจสอบอัตนัยภาษาไทยแบบอัตโนมัติโดยการประยุกต์ใช้เทคนิคการวัดความคล้ายคลึงกันของเอกสารแบบเชิงมุมโคไซน์ (Cosine Similarity Measure) ขั้นตอนการดำเนินงานประกอบด้วยทั้งหมด 4 ขั้นตอนคือ ขั้นตอนการเก็บรวบรวมข้อมูลโดยข้อมูลที่ใช้เป็นข้อมูลเอกสารการสอบ ขั้นตอนการจัดเตรียมข้อมูลโดยใช้เทคนิคทางด้านการประมวลผลภาษาธรรมชาติ ขั้นตอนการวัดความคล้ายคลึงกันของเอกสารจะใช้เทคนิคการวัดความคล้ายคลึงกันเชิงมุมโคไซน์ระหว่างสองเอกสารด้วยการให้น้ำหนักเชิงความถี่ และขั้นตอนสุดท้ายเป็นส่วนการพัฒนาระบบต้นแบบการตรวจสอบอัตนัยภาษาไทยแบบอัตโนมัติด้วยภาษาจาวาในรูปแบบการทำงานแบบผ่านเครือข่ายคอมพิวเตอร์ ผลการทดลองพบว่าค่าความถูกต้องในการตรวจสอบอัตนัยเท่ากับร้อยละ 69 และผลประเมินความพึงพอใจการใช้งานระบบอยู่ในระดับดี (ค่าเฉลี่ยเท่ากับ 3.92)

Abstract

This research aims to develop a prototype system of Thai essay scoring system using a similarity technique, Cosine similarity measurement. A data is the document of computer examination and so a natural language processing used in a data preparation step. For the similarity measuring step, a cosine similarity technique is used to do between two documents with the frequency weighting. And the last step is to develop the prototype of a Thai essay scoring system using a java language that works on a computer network. The accuracy of scoring is 69 percentage and the overall result of the satisfaction rating of this system is at a good level (3.92).

คำสำคัญ— ความคล้ายคลึงกันเอกสาร แบบจำลองเวกเตอร์สเปซ

1. บทนำ

การศึกษาในปัจจุบัน ในระดับอุดมศึกษามีการวัดและประเมินผลการเรียนรู้ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา การประกันคุณภาพบัณฑิต (TQF) 5 ด้าน ได้แก่ (1) ด้านคุณธรรมจริยธรรม (2) ด้านความรู้ (3) ด้านทักษะทางปัญญา (4) ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ (5) ด้านทักษะวิเคราะห์ตัวเลข การสื่อสาร และการใช้เทคโนโลยีสารสนเทศ เพื่อการวัดและประเมินผลการเรียนรู้ของนักศึกษาให้ได้ตรงตามกรอบการประกันคุณภาพบัณฑิต ด้านทักษะทางปัญญาซึ่งสามารถวัดผลสัมฤทธิ์ได้ด้วยข้อสอบแบบอัตนัยที่วัดความสามารถในการวิเคราะห์ สังเคราะห์และความเข้าใจในแนวคิด หลักการ ทฤษฎี ดังนั้นข้อสอบอัตนัยจึงมีบทบาทมากในการวัดและประเมินผลการเรียนรู้ของนักศึกษาในด้านทักษะทางปัญญามากกว่าข้อสอบแบบปรนัยหรือตัวเลือก

การใช้ข้อสอบอัตนัยในระดับอุดมศึกษามีน้อยกว่าข้อสอบปรนัยที่สามารถวัดและประเมินผลการเรียนรู้ด้านทักษะทางปัญญาของนักศึกษาได้น้อยกว่าข้อสอบอัตนัย เนื่องจากมีปัญหาบางประการในการดำเนินการได้แก่ ต้องใช้เวลาในการตรวจมาก ถ้าจำนวนชุดคำตอบมีมากและการตรวจอาจเกิดการเอนเอียงเนื่องจากผู้ตรวจมีอารมณ์และความรู้สึกต่างๆ ต่อผู้เรียนเข้ามาเกี่ยวข้อง โดยได้มีผลงานวิจัยต่างๆ เกี่ยวเนื่องดังต่อไปนี้ เช่น งานวิจัยของสมมาตร อังคเศรณีกุล ได้นำเทคนิคด้านการเรียนรู้ของเครื่องจักร ด้วยอัลกอริทึมแบบมีผู้สอน คือ อัลกอริทึม K-NN ในการตรวจสอบอัตนัยภาษาไทยแบบอัตโนมัติ ตัวอย่างงานวิจัยดังกล่าวมีจุดที่น่าสนใจในการตรวจสอบอัตนัยภาษาไทยแบบอัตโนมัติซึ่งเป็นประโยชน์สำหรับนักศึกษามหาวิทยาลัยราชภัฏนครราชสีมา ในการประยุกต์ใช้เทคนิค Cosine Similarity Measurement อีกทั้งข้อสอบอัตนัยมีลักษณะเป็นเอกสารแบบไม่มีโครงสร้างจึงมีความยากในการนำมาวิเคราะห์ความคล้ายคลึง เช่น งานวิจัยของวรารักษ์ [3] ได้

นำเสนอการสกัดสารสนเทศจากเอกสารโครงงานซึ่งเป็นเอกสารแบบไม่มีโครงสร้างที่มีการประยุกต์ใช้เทคนิคทางด้านการประมวลผลภาษาธรรมชาติเข้ามาช่วย เพื่อทำโครงงานชิ้นนี้

จากปัญหาที่กล่าวมานี้ โครงงานชิ้นนี้มีวัตถุประสงค์เพื่อพัฒนาระบบต้นแบบสำหรับการตรวจสอบข้อสอบอัตนัยภาษาไทยแบบอัตโนมัติโดยใช้เทคนิคความคล้ายคลึงกันของเอกสารด้วยคอมพิวเตอร์ขึ้น เพื่อช่วยสนับสนุนให้อาจารย์วัดและประเมินผลสัมฤทธิ์ในการเรียนการสอนด้วยข้อสอบอัตนัยมากขึ้น โดยระบบต้นแบบสามารถให้ผู้สอนสร้างรายวิชาและข้อสอบขึ้นได้ และผู้เรียนหรือนักศึกษาในแต่ละรายวิชาสามารถเข้าไปทำข้อสอบผ่านระบบเครือข่ายคอมพิวเตอร์ได้โดยตรงพร้อมกันทุกคน นอกจากนี้ระบบสามารถตรวจสอบข้อสอบอัตนัยให้ผู้สอนแบบอัตโนมัติซึ่งผู้สอนสามารถเลือกกระดานการให้คะแนนได้ผ่านการจัดการของระบบอีกด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยชิ้นนี้ได้มีการศึกษาในส่วนของทฤษฎีและงานวิจัยที่เกี่ยวข้อง ดังนี้

2.1. ข้อสอบอัตนัย

ข้อสอบอัตนัยเป็นการเขียนคำถามโดยกำหนดเป็นสถานการณ์หรือปัญหาในรูปใดรูปหนึ่ง เพื่อให้ผู้ตอบได้แสดงความรู้ ความเข้าใจ ความคิดเห็น ได้อย่างไม่จำกัดคำตอบของข้อสอบอัตนัยมีลักษณะ และปริมาณไม่แน่นอน การตอบข้อสอบอัตนัยจึงต้องจัดระเบียบคำตอบภายในเวลาที่กำหนด

2.1.1. ข้อสอบอัตนัยสามารถแบ่งประเภทได้สองประเภท คือ แบบไม่จำกัดคำตอบหรือหรือแบบขยายความ โดยให้ผู้ตอบแสดงความคิดเห็น อย่างอิสระ สามารถวัดสมรรถภาพด้านความคิดริเริ่มสร้างสรรค์ ทักษะคิด การประเมินค่าได้อย่าง กว้างขวาง ข้อสอบแบบนี้ส่งเสริมให้ผู้เรียนรู้จักการ รวบรวมความคิดต่าง ๆ และแบบจำกัดคำตอบ ข้อสอบแบบนี้จะถามแบบจำเพาะเจาะจง และต้องการ คำตอบ เฉพาะเรื่อง ซึ่งผู้ตอบต้องจัดเรียงความคิดให้เป็นระเบียบ เพื่อให้ตรงประเด็น ของคำถาม เพียงสั้น ๆ

2.1.2. วิธีการตรวจให้คะแนน การตรวจให้คะแนนข้อสอบอัตนัยที่ยอมรับกันทั่วไปแล้วมีอยู่ 2 วิธีคือ วิธีกำหนดค่าคะแนนและวิธีการตรวจให้คะแนนแบบแบ่งกลุ่ม

2.2. วิธีการวัดความคล้ายคลึงกันของเอกสาร [2]

การวัดความคล้ายคลึงกันของเอกสารเป็นกระบวนการวิธีการในการตรวจสอบความคล้ายคลึงกันของเอกสารสองเอกสารด้วยองค์ประกอบของคำที่อยู่ในเอกสารเป็นพื้นฐานสำคัญ โดยการวัดความคล้ายคลึงกันของเอกสารจากความถี่ของคำสำคัญที่พบในเอกสารเป็นการกำหนดให้เวกเตอร์แสดงถึงทุก ๆ คำสำคัญที่มีใน

เอกสาร ซึ่งเทคนิควิธีการที่ใช้ในการวัดค่าความคล้ายคลึงกันของเอกสารมีเทคนิควิธีที่ใช้วัดค่าความคล้ายคลึงของเอกสารอยู่ 4 เทคนิคคือ Inner Product, Dice, Cosine และ Jaccard

2.2.1. วิธีการวัดค่าความคล้ายคลึงของเอกสารแบบ Inner Product เป็นวิธีการวัดค่าความคล้ายคลึงของเอกสารสองเอกสารที่ถูกแทนที่ด้วยเวกเตอร์น้ำหนักของคำที่ปรากฏในเอกสาร แล้วนำมาหาค่า Dot Product ของเวกเตอร์สองเอกสาร ถ้ามีค่ามากแสดงว่าเวกเตอร์มีความคล้ายคลึงกันมาก โดยมีสูตรดังนี้

$$\sum_{k=1}^n w_{ik} \cdot w_{jk} \quad (1)$$

2.2.2. วิธีการวัดค่าความคล้ายคลึงของเอกสารแบบ Dice เป็นวิธีการวัดค่าความคล้ายคลึงของเอกสารสองเอกสารที่ถูกแทนที่ด้วยเวกเตอร์ น้ำหนัก ของคำที่ปรากฏในเอกสาร แล้วนำค่า Dot Product ของเวกเตอร์สองเอกสารมาเพิ่มค่า เป็น สองเท่าแล้วหารด้วยผลบวกของผลรวมค่าของเวกเตอร์สองเอกสารโดยมีสูตรดังนี้

$$\frac{2 \sum_{k=1}^n w_{ik} \cdot w_{jk}}{\sum_{k=1}^n w_{ik} + \sum_{k=1}^n w_{jk}} \quad (2)$$

2.2.3 วิธีการวัดค่าความคล้ายคลึงของเอกสารแบบ Cosine เป็นวิธีการเปรียบเทียบความคล้ายคลึงของเอกสารสองเอกสาร โดยแต่ละเอกสารจะถูกแทนด้วยเวกเตอร์ซึ่งจัดเก็บค่าน้ำหนักคำแต่ละคำในเอกสารนั้น (N-Dimensional Vector in Term Space) การเปรียบเทียบความคล้ายคลึงของเอกสารจะเปรียบเทียบโดยดูจากมุมโคไซน์ของมุมระหว่าง 2 เวกเตอร์ของเอกสาร หากเอกสารทั้งสองเอกสารคล้ายคลึงกันมากเวกเตอร์ของเอกสารทั้ง 2 จะมีค่าโคไซน์น้อยเนื่องด้วยมุมที่ทแยงระหว่างกันของเวกเตอร์เอกสารทั้งสองเกือบจะทับกัน

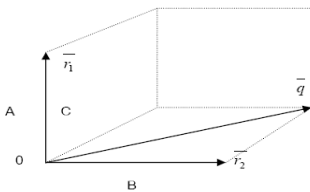
$$\frac{\sum_{k=1}^n w_{ik} \cdot w_{jk}}{\sqrt{\sum_{k=1}^n w_{ik}^2} \cdot \sqrt{\sum_{k=1}^n w_{jk}^2}} \quad (3)$$

2.2.4. วิธีการวัดค่าความคล้ายคลึงของเอกสารแบบ Jaccard เป็นวิธีการวัดค่าความคล้ายคลึงของเอกสารสองเอกสารที่ถูกแทนด้วยเวกเตอร์น้ำหนักของคำที่ปรากฏในเอกสาร แล้วนำค่า Dot Product ของเวกเตอร์สองเอกสารหารด้วยผลบวกของผลรวมค่าของเวกเตอร์สองเอกสารลบด้วยค่า Dot Product ของสองเวกเตอร์ มีสูตรดังนี้

$$\frac{\sum_{k=1}^n w_{ik} \cdot w_{jk}}{\sum_{k=1}^n w_{ik} + \sum_{k=1}^n w_{jk} - \sum_{k=1}^n w_{ik} \cdot w_{jk}} \quad (4)$$

2.3. แบบจำลองเวกเตอร์สเปซ [5]

แบบจำลองปริภูมิเวกเตอร์เป็นการจัดรูปแบบของเอกสารหรือตัวแทนของเอกสาร (Document Representation) เพื่อให้ง่ายต่อการนำไปประมวลผลและตีความหมายซึ่งรูปแบบจะมีลักษณะของการแสดงความสัมพันธ์ระหว่างคำในเอกสารทั้งหมดด้วยเวกเตอร์ 2 มิติซึ่งค่าที่ได้ต้องผ่านกระบวนการด้านประมวลผลภาษาธรรมชาติ (Natural Language Processing) ก่อนเสมอ เช่น การตัดคำ เป็นต้น รูปแบบเช่นนี้เรียกว่าการจัดรูปแบบแบบจำลองปริภูมิเวกเตอร์หรือรูปแบบของเอกสารที่ปรากฏในถุงของคำ (Bag of Words) รูปแบบการทำงานของแบบจำลองปริภูมิเวกเตอร์ โดยข้อดีของแบบจำลองปริภูมิเวกเตอร์คือการใช้รูปแบบคณิตศาสตร์ในการคิดจากความสัมพันธ์ของคำและสามารถจัดลำดับของเอกสารได้และสามารถใช้กับเอกสารที่มีข้อมูลมาก ๆ ได้ดี แบบจำลองเวกเตอร์เป็นการจับคู่แบบไม่เที่ยงตรง เช่นเดียวกับแบบจำลองความน่าจะเป็น (Probabilistic Model) ทำให้สามารถค้นหาแบบเป็นบางส่วน (Partial Search) ได้ ส่วนข้อเสียของแบบจำลองปริภูมิเวกเตอร์คือแบบจำลองปริภูมิเวกเตอร์จะไม่สนใจในความหมายของคำ วลี โครงสร้างของคำหรือคำที่มีความหมายเหมือนกัน เพราะแบบจำลองปริภูมิเวกเตอร์อยู่บนข้อสมมติที่ว่าคำทุกคำไม่มีความสัมพันธ์กันซึ่งในความเป็นจริงคำจะมีความสัมพันธ์หรืออ้างอิงกันในความหมาย ดังรูปที่ 1



รูปที่ 1. แบบจำลองเวกเตอร์สเปซสำหรับเอกสาร [5]

2.4 งานวิจัยที่เกี่ยวข้อง

สมมาตร อังคะระณีกุล [1] ได้เสนอเกี่ยวกับการตรวจสอบอัตโนมัติภาษาไทยโดยใช้ ชุดคำตอบจากลักษณะคำถามแบบปลายเปิดความยาวคำตอบแบบสั้น ทดลองกับอัลกอริทึมการเรียนรู้แบบมีผู้สอน คือ อัลกอริทึม K-NN ANN และ Bayesian และอัลกอริทึมการเรียนรู้แบบไม่มีผู้สอน คือ อัลกอริทึม EM และ Cobweb โดยใช้เทคนิคการขยายคำค้น (Query Expansion) และ เทคนิค LSA ร่วมด้วยในการวิเคราะห์ผล เพื่อเปรียบเทียบข้อดีข้อเสีย และสรุปผลที่ได้ในอัลกอริทึม การเรียนรู้แต่ละแบบ

พรเทพ จันทรเพ็ง [2] ได้เสนอการเปรียบเทียบเทคนิคการวัดความคล้ายคลึงของเอกสารแบบจำลองปริภูมิเวกเตอร์เพื่อการตรวจสอบอัตโนมัติ โดยนำหลักการของแบบจำลองปริภูมิเวกเตอร์ (Vector Space Model) มาใช้ร่วมกับการนับค่าความถี่ของคำที่ปรากฏในเอกสารและการให้น้ำหนักค่า แล้วทำการวัดความคล้ายคลึงด้วยเทคนิควิธีการวัดความคล้ายคลึง 4 เทคนิควิธีคือการวัดความคล้ายคลึงแบบ Inner Product, Dice, Cosine และ Jaccard ของเอกสารคำตอบกับเอกสาร คำเฉลย แล้วประเมินผลหาค่าความแม่นยำของแต่ละเทคนิควิธีโดยเปรียบเทียบค่ากับค่าความถูกต้องของผู้สอนตรวจให้คะแนนเป็นหลัก

บารมี โอสถิกรุล และรัชชัย งามสันติวงศ์ [4] ได้เสนอเกี่ยวกับระบบเทียบโอนรายวิชาโดยใช้ แบบจำลองเวกเตอร์สเปซ ทำการทดลองกับสาขาวิชาระบบสารสนเทศและ คอมพิวเตอร์ธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ มีวัตถุประสงค์เพื่อพัฒนาระบบเทียบโอนรายวิชาจากการวัดความคล้ายคลึงของรายวิชาโดยนำหลักการวัดความคล้ายคลึงเชิงมุมโคไซน์ และการให้น้ำหนักของคำมาใช้กับหลักการของแบบจำลองเวกเตอร์สเปซ เพื่อวิเคราะห์วิธีการวัดความคล้ายคลึงของภาษาไทยในการดำเนินงานประยุกต์ใช้การตัดคำและนำคำที่ได้นั้นไปเปรียบเทียบกับ Dictionary หลังจากนั้นนำคำที่มีความหมายมาเก็บไว้เพื่อเป็น Keywords ซึ่ง Model ที่นำมาใช้ในการเปรียบเทียบ คือ Vector Space Model (VSM) เป็นการเปรียบเทียบคำ โดยที่ไม่ต้องมีค่าที่ตรงกันทั้งหมดขึ้นอยู่กับความคล้ายคลึง และการวัดความคล้ายคลึงของเอกสารภาษาไทย แบ่งออกเป็น 2 ส่วน คือ ส่วนการเตรียมคำอธิบายรายวิชาก่อนการประมวลผลและส่วน การประมวลผลในการวัดความคล้ายคลึง (similarity) ระหว่างคำอธิบายรายวิชาระดับ(ป.ตรี) และคำอธิบายรายวิชาระดับ ประกาศนียบัตรวิชาชีพชั้นสูงถ้ามีค่าความคล้ายคลึงสูงย่อมแสดงว่ารายวิชา ระดับประกาศนียบัตรวิชาชีพชั้นสูงดังกล่าวนั้นมีโอกาสเทียบโอนหน่วยกิตกับระดับปริญญาตรีสูงสุด

3. ขั้นตอนการดำเนินงาน

การดำเนินงานในงานวิจัยชิ้นนี้ประกอบด้วยขั้นตอนการดำเนินงาน 4 ขั้นตอน ดังนี้

3.1 การเก็บรวบรวมข้อมูล

ข้อมูลที่ใช้ในงานวิจัยชิ้นนี้เป็นข้อมูลการสอบรายวิชาเหมือนข้อมูลและรายวิชาคอมพิวเตอร์กราฟิกส์จำนวน 200 เอกสารคำตอบและ 10 เอกสารเฉลย (จำนวนข้อสอบรายวิชาละ 5 ข้อ และนักศึกษาจำนวน 20 คน) โดยเป็นข้อมูลตัวอย่างจากคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา

แสดงข้อมูลตัวอย่างดังรูปที่ 2 นอกจากข้อมูลเอกสารคำตอบและเอกสารเฉลยแล้ว จะมีการจัดเก็บข้อมูลคำพิเศษประเภทคำหยุด (Stop-Word) เพื่อใช้สำหรับขั้นตอนการจัดเตรียมข้อมูล

ลำดับ	ข้อความ
เฉลย	เนื่องด้วยจอภาพหรือจอแอลซีดีในการแสดงผลภาพแบบ 2 มิติเท่านั้นไม่สามารถแสดงผลภาพแบบ 3 มิติได้ ดังเช่นการแสดงผลภาพในระบบ 3 มิติซึ่งต้องผ่านกระบวนการฉายภาพหรือโปรเจกต์เพื่อแปลงภาพจากระบบ 3 มิติให้อยู่ในระบบแบบ 2 มิติ
1	ภาพที่ฉายบนโลก 3 มิติ จะมีการแปลงโคออดิเนต เมื่อตกลงพื้นระนาบจะเป็นภาพ 2 มิติ เพราะการฉายภาพจะเป็นภาพ 3 มิติไม่ได้
2	เป็นสิ่งที่ต้องคำนึงก่อนที่จะมีการสร้าง Animation ต่างๆ ก่อนเสมอ เพื่อให้ได้รูปแบบว่าจะต้องเข้าไปในทิศทางใดและหาอย่างไร
3	สาเหตุที่กราฟิก 3 มิติต้องมีการทำกระบวนการไปเจ็ทอีกขึ้นเพราะว่า ต้องทำการภาพ 3 มิติไปแสดงในจอภาพได้ขึ้นเอง จอภาพไม่สามารแสดงผลภาพ 3 มิติได้ แสดงได้แค่ 2 มิติ จึงต้องทำโปรเจกต์ขึ้น คือ การแปลงภาพ 3 มิติให้อยู่ในรูป 2 มิติ
4	เพื่อที่จะช่วยให้เกิดความเหมือนจริงอย่างมากที่สุดและช่วยในการตัดสินใจเลือกสิ่งและภาพเคลื่อนไหว
5	ระบบ -D จะใช้แสดงและงาน ส่วนระบบ 2D ไม่มี คือ ระบบ 3D จะใช้แสดงเข้ามาช่วยเพิ่มเติมให้กับวัตถุ และเมื่อมีการใช้แสงจึงทำให้เกิดเงาตามภาพในภาพ 3D ขึ้นสวยงามและสมจริงกว่าภาพ 2D ที่ไม่มีแสงและเงา แต่จะใช้การใส่ระดับของพื้น
6	กระบวนการไปเจ็ทขึ้นเป็นขั้นตอนที่ 6 คือ การมองออกมาจากจุดหรือจุดตัดจากระยะใกล้เข้าไปหาวัตถุหรือวัตถุออกจากจุด โดยที่เราจะรับรู้ได้ว่าเมื่อมองออกไปก็จะเห็นวัตถุเล็ก ใกล้เข้ามาเป็นวัตถุใหญ่ขึ้น ก็จะเหมือนการดูเข้าหาลูกบอล

รูปที่ 2. ตัวอย่างข้อมูลเอกสารเฉลยและเอกสารคำตอบ

3.2 การจัดเตรียมเอกสาร

ในขั้นตอนนี้เอกสารทั้งหมดที่ถูกเก็บรวบรวมมาได้จะถูกนำมาผ่านกระบวนการก่อนประมวลผลด้วยเทคนิคด้านการประมวลผลภาษาธรรมชาติประกอบด้วยการตัดคำ การกำจัดคำหยุด เป็นต้น การตัดคำจากเอกสารเป็นการแบ่งแยกคำสำคัญต่างๆ ในเอกสารเพื่อใช้สำหรับนำไปสร้างเป็นตัวแทนของเอกสาร โดยงานวิจัยชิ้นนี้ได้ประยุกต์ใช้เทคนิคการตัดคำแบบอ้างอิงพจนานุกรมแบบยาวที่สุด (Longest Matching based on dictionary) ด้วยคลาส BreakIterator ในภาษาจาวาเข้ามาช่วยและใช้รายการคำหยุดที่นำมาช่วยในการตรวจสอบเพื่อกำจัดคำหยุดต่างๆ ก่อนนำไปสร้างเป็นแบบจำลองเวกเตอร์สเปซ ในการสร้างแบบจำลองเวกเตอร์สเปซจะกำหนดให้เวกเตอร์แรกของแบบจำลองเป็นส่วนเอกสารเฉลยและเวกเตอร์ถัดมาจะเป็นเอกสารคำตอบซึ่งมีการกำหนดค่าน้ำหนักของคำในแบบจำลองเป็นค่าความถี่ของคำต่างๆ ดังรูปที่ 3

VECTOR.csv	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	หัด	คล้ายคลึง	เอกสาร	ภาษา	ไทย	หลักการ	ประมวล	ประมวล	และ	วิธี	ลักษณะ	ระดับ	สูง	0
2	1	1	1	2	1	1	1	1	1	1	0	0	0	0
3	0	0	0	2	0	0	1	1	1	1	1	1	0	0
4	1	1	1	2	0	0	0	0	0	0	0	1	1	1

VECTOR.xls	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	หัด	คล้ายคลึง	เอกสาร	ภาษา	ไทย	หลักการ	ประมวล	ประมวล	และ	วิธี	ลักษณะ	ระดับ	สูง		
2	1	1	1	2	1	1	1	1	1	1	0	0	0	0	0
3	0	0	0	2	0	0	1	1	1	1	1	1	0	0	0
4	1	1	1	2	0	0	0	0	0	0	0	1	1	1	1

VECTOR.xlsx	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	หัด	คล้ายคลึง	เอกสาร	ภาษา	ไทย	หลักการ	ประมวล	ประมวล	และ	วิธี	ลักษณะ	ระดับ	สูง	
2	1	1	1	2	1	1	1	1	1	1	0	0	0	0
3	0	0	0	2	0	0	1	1	1	1	1	1	0	0
4	1	1	1	2	0	0	0	0	0	0	0	1	1	1

รูปที่ 3. ตัวอย่างแบบจำลองเวกเตอร์สเปซ

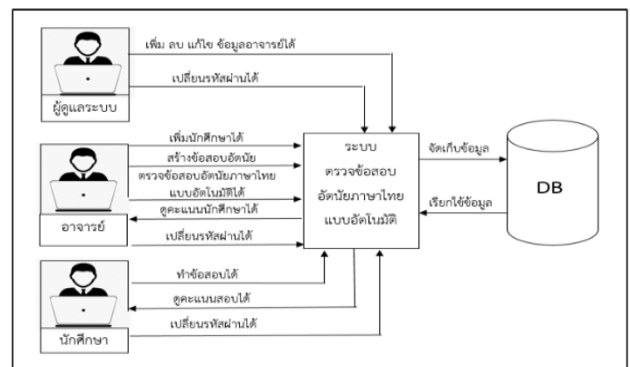
3.3 การวัดความคล้ายคลึงของเอกสาร

การวัดความคล้ายคลึงกันของเอกสารระหว่างเอกสารคำตอบและเอกสารเฉลยจะใช้เทคนิคการวัดความคล้ายคลึงกันเชิงมุมโคไซน์โดยการนำแบบจำลองเวกเตอร์สเปซมาเปรียบเทียบกับความคล้ายคลึงกันของเอกสารที่ละสองเอกสารด้วยสมการที่ 3 โดย

การคำนวณส่วนนี้จะถูกเขียนไว้ในส่วนของโค้ดเพื่อพัฒนาเป็นระบบต้นแบบการตรวจสอบข้อสอบอัตโนมัติแบบอัตโนมัติ

3.4 การพัฒนาระบบต้นแบบการตรวจสอบข้อสอบอัตโนมัติ

ในขั้นตอนนี้จะเป็นขั้นตอนการพัฒนากระบวนต้นแบบการตรวจสอบข้อสอบอัตโนมัติแบบอัตโนมัติ โดยใช้ภาษาโปรแกรมจาวาในการพัฒนาระบบและระบบจัดการฐานข้อมูลด้วยโปรแกรม MySQL สามารถแสดงผลแผนภาพการทำงาน ดังรูปที่ 4



รูปที่ 4. แผนภาพการทำงานของระบบต้นแบบการตรวจสอบข้อสอบอัตโนมัติแบบอัตโนมัติ

4. ผลการดำเนินงาน

งานวิจัยชิ้นนี้มีการทดลองการทำงานและแสดงผลการวิจัยได้ดังนี้

4.1 ผลการทำงานของระบบต้นแบบการตรวจสอบข้อสอบอัตโนมัติแบบอัตโนมัติ

การทำงานของระบบจะเริ่มจากผู้ดูแลระบบจะเพิ่มผู้ใช้งานระบบผู้สอนหรืออาจารย์เข้าระบบ ดังรูปที่ 5

รูปที่ 5. หน้าจอสำหรับการเพิ่มผู้ใช้งานระดับผู้สอน

จากนั้นผู้สอนเมื่อมีการสร้างรายวิชาขึ้นมาและเพิ่มรายชื่อผู้เรียนเข้ารายวิชา ผู้สอนจะทำการสร้างข้อสอบอัตโนมัติโดยแสดงได้ ดังรูปที่ 6

รูปที่ 6. หน้าจอสำหรับการสร้างข้อสอบอัตโนมัติ

ส่วนการทำงานของนักเรียนเมื่อมีรายชื่ออยู่ในระบบจะสามารถเข้าไปทำข้อสอบอัตโนมัติผ่านระบบได้ ดังรูปที่ 7

รูปที่ 7. หน้าจอแสดงการทำข้อสอบอัตโนมัติของนักเรียน

และระบบสามารถทำการตรวจข้อสอบอัตโนมัติที่ผู้เรียนได้ทำผ่านส่วนของการทำงานโดยกำหนดค่าความคล้ายคลึงกันของเอกสารให้เป็นรูปแบบของคะแนนและแสดงออกมาได้ ดังรูปที่ 8

รูปที่ 8. หน้าจอแสดงส่วนการให้คะแนนจากค่าความคล้ายคลึงกันของเอกสาร

4.2 ผลการประเมินประสิทธิภาพการตรวจข้อสอบอัตโนมัติ

สำหรับการประเมินประสิทธิภาพการตรวจข้อสอบอัตโนมัติของระบบต้นแบบจะใช้ค่าความถูกต้องในการให้คะแนนเทียบกับของผู้สอนที่ให้คะแนน โดยกำหนดช่วงคะแนนเป็นสามช่วง ดังตารางที่ 1 โดยใช้การสุ่มตัวอย่างเอกสารคำตอบจำนวน 100 เอกสาร

ตารางที่ 1 ค่าความถูกต้องในการตรวจข้อสอบของระบบ

สรุปค่าความถูกต้องแต่ละข้อ			
ข้อที่	ตอบถูก	ตอบผิด	คิดเป็นร้อยละ
1	17	3	70
2	10	10	50
3	14	6	70
4	12	8	60
5	18	2	70
ภาพรวม	69	31	69

จากตารางที่ 1 แสดงผลความถูกต้องในการตรวจข้อสอบอัตโนมัติโดยค่าความถูกต้องของระบบเท่ากับร้อยละ 69

4.3 ผลการประเมินความพึงพอใจของผู้ใช้

การประเมินความพึงพอใจของระบบใช้การประเมินความพึงพอใจด้วยมาตราอันดับ 5 อันดับ และแปลความหมายของค่าเฉลี่ยเป็นแบบ 5 ช่วงคะแนน ดังตารางที่ 2 โดยมีการทดสอบกับผู้ใช้จำนวน 11 คน ประกอบด้วยอาจารย์และนักศึกษา

ตารางที่ 2. เกณฑ์การแปลผลการประเมินความพึงพอใจ

เกณฑ์ค่าเฉลี่ย	การแปลความหมายของระดับคะแนน
4.51–5.00	เกณฑ์ดีมาก
3.51–4.50	เกณฑ์ดี
2.51–3.50	เกณฑ์พอใช้
1.51–2.50	เกณฑ์ปรับปรุง
1.00–1.50	เกณฑ์ไม่เหมาะสม

ซึ่งสถิติที่ใช้ในการวิเคราะห์ข้อมูลความพึงพอใจต่อระบบมี 2 สมการ คือ ค่าเฉลี่ยเลขคณิต (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (5)$$

$$S.D. = \sqrt{\frac{n \sum fx_i^2 - (\bar{x})^2}{n(n-1)}} \quad (6)$$

ผลการประเมินความพึงพอใจของผู้ใช้งาน มีความพึงพอใจต่อการใช้งานระบบต้นแบบการตรวจสอบอัตโนมัติแบบอัตโนมัติเท่ากับ 3.92 (อยู่ในระดับดี) ดังตารางที่ 3

ตารางที่ 3. ผลคะแนนการประเมินความพึงพอใจ

รายการประเมิน	ระดับความพึงพอใจ						s.d.	ผล ลัพธ์
	5	4	3	2	1	\bar{x}		
1. ความง่ายต่อการใช้งานของระบบ	2	2	2	-	-	4.00	0.89	มาก
2. ความเหมาะสมในการเลือกใช้นาฬิกาของตัวอักษรบนจอภาพ	1	4	1	-	-	4.00	0.63	มาก
3. ความเหมาะสมในการเลือกใช้นาฬิกาของตัวอักษรบนจอภาพ	1	5	-	-	-	4.17	0.41	ปานกลาง
4. ความเหมาะสมในการใช้สีของตัวอักษรบนจอภาพ	-	3	3	-	-	3.50	0.55	ปานกลาง
5. ความเป็นมาตรฐานเดียวกันในการออกแบบจอภาพ	1	3	2	-	-	3.83	0.75	มาก
6. ความเหมาะสมในการปฏิสัมพันธ์	3	1	1	1	-	4.00	1.26	มาก

โต้ตอบกับผู้ใช้								
7. คำศัพท์ที่ใช้ผู้ใช้มีความคุ้นเคยและปฏิบัติตามได้โดยง่าย	1	5	-	-	-	4.17	0.41	มาก
8. ความเหมาะสมในการนำไปใช้งานจริง	-	4	2	-	-	3.67	0.52	ปานกลาง
ค่าเฉลี่ยรวม						3.92	0.68	มาก

5. สรุปผลและข้อเสนอแนะ

ในการจัดทำวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสร้างระบบการตรวจสอบอัตโนมัติภาษาไทยแบบอัตโนมัติ โดยใช้เทคนิคการวัดความคล้ายคลึงกันของเอกสารที่ช่วยให้อาจารย์สามารถจัดการกับข้อสอบอัตโนมัติภาษาไทยได้ถูกต้องและรวดเร็ว โดยการนำเอกสารคำตอบและคำตอบมาวัดค่าความคล้ายคลึงกัน โดยใช้ อัลกอริทึม Cosine Similarity โดยให้น้ำหนักของคำด้วยความถี่ในการวัดความคล้ายคลึงกันของเอกสาร และใช้ BreakIterator ในการตัดคำซึ่งยังมีค่าผิดค่อนข้างมากทำให้ความแม่นยำในการตรวจข้อสอบลดลง และผลการประเมินประสิทธิภาพของโปรแกรมการตรวจสอบอัตโนมัติภาษาไทยแบบอัตโนมัติ โดยใช้การสุ่มข้อสอบจำนวน 100 คำตอบจากเอกสาร 200 เอกสาร มีค่าความถูกต้องเท่ากับร้อยละ 69 ซึ่งจากผลการทดลองพบว่าค่าความถูกต้องในการตรวจข้อสอบอัตโนมัติอยู่ในระดับที่ต่ำเนื่องด้วยลักษณะของคำตอบส่วนมากของนักศึกษาจะมีการใช้คำศัพท์ที่ตรงกับคำเฉลยแต่มีลักษณะการขยายความผิดตามความหมายที่ถูกต้องทำให้เกิดข้อผิดพลาดขึ้นมาได้ ซึ่งในอนาคตต่อไปจะมีการปรับให้ค่าน้ำหนักคำสำคัญมีค่าไม่เท่ากันต่อไป

จากการทดลองใช้ระบบโดยผู้ใช้งาน 11 คน แยกเป็นอาจารย์ 1 คน นักศึกษา 10 คน ผลการประเมินความพึงพอใจของระบบโดยผู้ใช้งาน เมื่อพิจารณาพบว่าทั้ง 8 รายการ มีความพึงพอใจระดับ “มาก” มีค่าเฉลี่ย (\bar{x} =3.92) ส่วนเบี่ยงเบนมาตรฐาน ($s.d.$ =0.68) สามารถเรียงตามค่าเฉลี่ยแต่ละรายการคือ ด้านความเหมาะสมในการเลือกใช้นาฬิกาของตัวอักษรบนจอภาพ (\bar{x} =4.17), ($s.d.$ =0.41) ด้านคำศัพท์ที่ใช้ผู้ใช้มีความคุ้นเคยและปฏิบัติตามได้โดยง่าย (\bar{x} =4.17), ($s.d.$ =0.41) ด้านความง่ายต่อการใช้งานระบบ (\bar{x} =4.00), ($s.d.$ =0.89) ด้านความเหมาะสมในการเลือกใช้นาฬิกาของตัวอักษรบนจอภาพ

(\bar{x} =4.00), ($s.d.$ =0.63) ด้านความเหมาะสมในการปฏิสัมพันธ์
โต้ตอบกับผู้ใช้ (\bar{x} =4.00), ($s.d.$ =1.26) ด้านความเป็น
มาตรฐานเดียวกันในการออกแบบจอภาพ (\bar{x} =3.83), ($s.d.$ =
0.75) ด้านความเหมาะสมในการนำไปใช้งานจริง (\bar{x} =3.67),
($s.d.$ =0.52) ด้านความเหมาะสมในการใช้สีของตัวอักษรบน
จอภาพ (\bar{x} =3.50), ($s.d.$ =0.55)

6. เอกสารอ้างอิง

- [1] สมมาตร อังคเศรณีกุล. (2555). การตรวจสอบข้อสอบอัตนัย
ภาษาไทยแบบอัตโนมัติ. วิทยาศาสตร์มหาบัณฑิต. วิทยาการ
คอมพิวเตอร์. มหาวิทยาลัยเกษตรศาสตร์
- [2] พรเทพ จันทรเพ็ญ. (2557). การเปรียบเทียบเทคนิคการวัด
ความคล้ายคลึงของเอกสารแบบจำลองปริภูมิเวกเตอร์เพื่อการ
ตรวจข้อสอบอัตนัย. วิทยาศาสตร์มหาบัณฑิต. วิทยาการ
คอมพิวเตอร์. มหาวิทยาลัยเชียงใหม่
- [3] วราภรณ์ คงสมพงษ์. (2557). การสกัดเนื้อหาสำหรับเอกสาร
โครงการนักศึกษาลงฐานข้อมูลแบบอัตโนมัติโดยใช้กฎ. วิทยา
ศาสตรบัณฑิต. มหาวิทยาลัยราชภัฏนครราชสีมา
- [4] บารมี โอษฐ์รกุล และธวัชชัย งามสันติวงศ์ (2555). ระบบ
เทียบโอนรายวิชาโดยใช้แบบจำลองเวกเตอร์สเปซ. วิทยาศาสตร์
มหาบัณฑิต. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
- [5] บังอร กลับบานเกาะ, “การค้นคืนสารสนเทศออนไลน์โดย
ใช้จีเนติกอัลกอริทึม”, วิทยานิพนธ์ ปริญญาวิทยาศาสตร
มหาบัณฑิต มหาวิทยาลัยมหิดล, 2548