

เทคโนโลยีปัญญาประดิษฐ์ AI กับการแบ่งกลุ่มข้อมูลผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 5 โดยใช้วิธีการเรียนรู้การแบ่งกลุ่มข้อมูลชนิด K-mean

บทคัดย่อ

การวิจัยครั้งนี้เป็นการวิจัยเพื่อศึกษาผลการจัดกลุ่มข้อมูลโดยใช้เทคโนโลยีปัญญาประดิษฐ์ AI แบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนกลุ่มตัวอย่างเป็น กลุ่มเก่ง กลุ่มปานกลาง กลุ่มอ่อน และเพื่อศึกษาผลการจัดกลุ่มข้อมูลโดยใช้เทคโนโลยีปัญญาประดิษฐ์ AI แบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของรายวิชาเป็น กลุ่มเด่น กลุ่มกลาง กลุ่มด้อย กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนชั้นประถมศึกษาปีที่ 5 ปีการศึกษา 2561 โรงเรียนสามเสนนอก (ประชากรราษฎร์อนุกุล) สำนักงาน เขตดินแดง กรุงเทพมหานคร จำนวน 33 คน ได้มาจากการสุ่มตัวอย่างอย่างง่ายด้วยวิธีจับสลาก เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ เทคโนโลยีปัญญาประดิษฐ์ AI โดยใช้วิธีการเรียนรู้การแบ่งกลุ่มข้อมูลชนิด K-mean ผลการวิจัย พบว่า การแบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนกลุ่มตัวอย่างเป็น กลุ่มเก่ง 11 คน กลุ่มปานกลาง 11 คน กลุ่มอ่อน 11 คน และการแบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของรายวิชาเป็น กลุ่มเด่น 4 รายวิชา กลุ่มกลาง 3 รายวิชา กลุ่มด้อย 3 รายวิชา โดยมีรายวิชา ที่เด่นสุดคือ วิชาประวัติศาสตร์ รองลงมาคือ วิชาการงานอาชีพ และรายวิชาที่ด้อยสุดคือ วิชาภาษาอังกฤษ

คำสำคัญ : เทคโนโลยีปัญญาประดิษฐ์ AI, ผลสัมฤทธิ์ทางการเรียน, การเรียนรู้การแบ่งกลุ่มข้อมูล, K-mean

บทนำ

ความก้าวหน้าและการพัฒนาทางเทคโนโลยีส่งผลให้ความเป็นอยู่ของชีวิตมนุษย์ดีขึ้น หนึ่งในเทคโนโลยีที่ถูกนำมาใช้คือ เทคโนโลยีปัญญาประดิษฐ์ (artificial intelligence: AI) ซึ่งได้เริ่มขึ้นหลังจากการประชุมวิชาการที่วิทยาลัยดาร์ตมัธ สหรัฐอเมริกา ในปี ค.ศ. 1956 ในครั้งนั้นมีผู้ร่วมประชุมหลายคน ได้แก่ จอร์จ แบริกคาร์ธีย์ เป็นประธานในที่ประชุม ร่วมด้วย มาร์วิน มินสกี, อัลเลน นิวเวลล์, อาเธอร์ ซามเอล, และเฮอร์เบิร์ต ไซมอน ต่อมาบุคคลเหล่านี้ได้กลายมาเป็นผู้นำทางสาขาปัญญาประดิษฐ์ แต่ยังไม่เป็นที่นิยมในสมัยนั้นเนื่องจากยากต่อการทำความเข้าใจของคนทั่วไป (Hutton, 2011) ปัจจุบันสถาบันการศึกษาหลายแห่งเปิดสอนหลักสูตรที่เกี่ยวข้องกับการนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ มีงานวิจัยเผยแพร่และข้อมูลจำนวนมากซึ่งถูกนำมาใช้ประโยชน์กับมนุษย์อย่างไร้ที่ติตามการใช้ประโยชน์ในเชิงกลยุทธ์จากข้อมูลที่รวบรวมได้ จำเป็นต้องมีระบบสารสนเทศที่สามารถเรียนรู้จากข้อมูลนั้น ๆ

การแบ่งกลุ่มข้อมูล (clustering data) เป็นวิธีการค้นหาข้อมูลโดยใช้หลักการคาดคะเน ด้วยการดำเนินการของการทำงานเครื่องจักรที่ไม่ต้องมีคนสอน (Sedigheh et al., 2015) เทคโนโลยีดังกล่าวถูกนำมาใช้เพื่อช่วยแก้ปัญหาและพัฒนาคุณภาพชีวิตของมนุษย์ รวมถึงสร้างความสะดวกสบายในด้านต่าง ๆ เช่น การแพทย์ ธุรกิจ การศึกษา (ชัยยงค์ พรหมวงศ์, 2521) ขอบข่ายของเทคโนโลยีการศึกษาสามารถกำหนดให้แสดงภาพเป็น 3 มิติ ได้แก่ 1) มิติแนวตั้งด้านสาระความรู้ทางเทคโนโลยี การศึกษา 2) มิติแนวนอนตามขอบข่ายบทบาทหน้าที่งาน คือ ด้านเครื่องมือในการบริหาร ด้านเครื่องมือด้านวิชาการ และเครื่องมือ ด้านบริการ และ 3) มิติเชิงลึกในบริบทของการศึกษา ทั้งนี้ ชัยยงค์ พรหมวงศ์ (2537) ได้ให้ความหมายในมิติแนวตั้งด้านสาระ ความรู้ทางเทคโนโลยี การศึกษาว่า หมายถึง ขอบข่ายสาระของเทคโนโลยีการศึกษา ได้แก่ สาระความรู้ การจัดระบบพฤติกรรม วิธีการ การจัดระบบสิ่งแวดล้อม การจัดการสื่อสาร และการประเมิน

เทคโนโลยี AI ในปัจจุบันถูกนำมาใช้กับการศึกษา แต่ยังไม่เป็นที่แพร่หลายและนิยมใช้โดยนักการศึกษาในประเทศไทย กิดานันท์ มลิทอง (2540) กล่าวว่า การนำเทคโนโลยีมาใช้ในการศึกษาหรือเรียกว่าเทคโนโลยีการศึกษา เป็นการประยุกต์เทคนิค วิธีการ วัสดุอุปกรณ์ และสิ่งต่าง ๆ จากเทคโนโลยี เทคโนโลยีปัญญาประดิษฐ์ AI เป็นการนำ AI มาใช้กับการทำงานของเครื่องจักร

บทคัดย่อ

- การวิจัยครั้งนี้เป็นการวิจัยเพื่อศึกษาผลการจัดกลุ่มข้อมูลโดยใช้เทคโนโลยีปัญญาประดิษฐ์ AI แบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนกลุ่มตัวอย่างเป็น กลุ่มเก่ง กลุ่มปานกลาง กลุ่มอ่อน
- และเพื่อศึกษาผลการจัดกลุ่มข้อมูลโดยใช้เทคโนโลยีปัญญาประดิษฐ์ AI แบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของรายวิชาเป็น กลุ่มเด่น กลุ่มกลาง กลุ่มด้อย
- กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนชั้นประถมศึกษาปีที่ 5 ปีการศึกษา 2561 โรงเรียนสามเสนนอก (ประชาราษฎร์อนุกุล) สำนักงาน เขตดินแดง กรุงเทพมหานคร จำนวน 33 คน
- ได้มาจากการสุ่มตัวอย่างอย่างง่ายด้วยวิธีจับสลาก
- เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ เทคโนโลยีปัญญาประดิษฐ์ AI โดยใช้วิธีการเรียนรู้การแบ่งกลุ่มข้อมูลชนิด K-mean
- ผลการวิจัย พบว่า การแบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนกลุ่มตัวอย่างเป็น กลุ่มเก่ง 11 คน กลุ่มปานกลาง 11 คน กลุ่มอ่อน 11 คน

- และการแบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของรายวิชาเป็น กลุ่มเด่น 4 รายวิชา กลุ่มกลาง 3 รายวิชา กลุ่มด้อย 3 รายวิชา โดยมีรายวิชา ที่เด่นสุดคือ วิชาประวัติศาสตร์ รองลงมาคือ วิชาการงานอาชีพ และรายวิชาที่ด้อยสุดคือ วิชาภาษาอังกฤษ
- คำสำคัญ : เทคโนโลยีปัญญาประดิษฐ์ AI, ผลสัมฤทธิ์ทางการเรียน, การเรียนรู้การแบ่งกลุ่มข้อมูล, K-mean

บทนำ

- ความก้าวหน้าและการพัฒนาทางเทคโนโลยีส่งผลให้ความเป็นอยู่ของชีวิตมนุษย์ดีขึ้น
- หนึ่งในเทคโนโลยีที่ถูกนำมาใช้คือ เทคโนโลยีปัญญาประดิษฐ์ (artificial intelligence: AI) ซึ่งได้เริ่มขึ้นหลังจากการประชุมวิชาการที่วิทยาลัยดาร์ตมัธ สหรัฐอเมริกา ในปี ค.ศ. 1956
- ในครั้งนั้นมีผู้ร่วมประชุมหลายคน ได้แก่ จอร์น แม็กคาร์ธี เป็นประธานในที่ประชุม ร่วมด้วย มาร์วิน มินสกี, อัลเลน นิวเวลล์, อาเธอร์ ซามเอล, และเฮอร์เบิร์ต ไซมอน
- ต่อมาบุคคลเหล่านี้ได้กลายมาเป็นผู้นำทางสาขาปัญญาประดิษฐ์ แต่ยังไม่เป็นที่นิยมในสมัยนั้นเนื่องจากยากต่อการทำความเข้าใจของคนทั่วไป (Hutton, 2011)
- ปัจจุบันสถาบันการศึกษาหลายแห่งเปิดสอนหลักสูตรที่เกี่ยวข้องกับการนำเทคโนโลยีปัญญาประดิษฐ์มาใช้
- มีงานวิจัยเผยแพร่และข้อมูลจำนวนมากทางศาสตร์ซึ่งถูกนำมาใช้ประโยชน์กับมนุษย์
- อย่างไรก็ตามการใช้ประโยชน์ในเชิงกลยุทธ์จากข้อมูลที่รวบรวมได้ จำเป็นต้องมีระบบสารสนเทศที่สามารถเรียนรู้จากข้อมูลนั้น ๆ
- การแบ่งกลุ่มข้อมูล (clustering data) เป็นวิธีการค้นหาข้อมูลโดยใช้หลักการคาดคะเน ด้วยการดำเนินการของการทำงานเครื่องจักรที่ไม่ต้องมีคนสอน (Sedigheh et al., 2015)
- เทคโนโลยีดังกล่าวถูกนำมาใช้เพื่อช่วยแก้ปัญหาและพัฒนาคุณภาพชีวิตของมนุษย์ รวมถึงสร้างความสะดวกสบายในด้านต่าง ๆ เช่น การแพทย์ ธุรกิจ การศึกษา (ชัยยงค์ พรหมวงศ์, 2521)
- ขอบข่ายของเทคโนโลยีการศึกษาสามารถกำหนดให้แสดงภาพเป็น 3 มิติ ได้แก่
- 1) มิติแนวตั้งด้านสาระความรู้ทางเทคโนโลยี การศึกษา

- 2) มิติแนวนอนตามขอบข่ายบทบาทหน้าที่งานคือ ด้านเครื่องมือในการบริหาร ด้านเครื่องมือด้านวิชาการ และเครื่องมือ ด้านบริการ
- และ 3) มิติเชิงลึกในบริบทของการศึกษา ทั้งนี้ ชัยยงค์ พรหมวงศ์ (2537) ได้ให้ความหมายในมิติแนวตั้ง ด้านสาระ ความรู้ทางเทคโนโลยีการศึกษาว่า หมายถึง ขอบข่ายสาระของเทคโนโลยีการศึกษา ได้แก่ สาระความรู้ การจัดระบบพฤติกรรม วิธีการ การจัดระบบสิ่งแวดล้อม การจัดการสื่อสาร และการประเมิน เทคโนโลยี AI
- ในปัจจุบันถูกนำมาใช้กับการศึกษา แต่ยังไม่เป็นที่แพร่หลายและนิยมใช้โดยนักการศึกษาในประเทศไทย
- กิดานันท์ มลิทอง (2540)
- กล่าวว่า การนำเทคโนโลยีมาใช้ในการศึกษาหรือเรียกว่าเทคโนโลยีการศึกษา เป็นการประยุกต์เทคนิค วิธีการ วัสดุอุปกรณ์ และสิ่งต่าง ๆ จากเทคโนโลยี
- เทคโนโลยีปัญญาประดิษฐ์ AI เป็นการนำ AI มาใช้กับการทำงานของเครื่องจักร

การวิเคราะห์ความคิดเห็นเกี่ยวกับร้านอาหารบนเว็บไซต์รีวิว

บทคัดย่อ

ในปัจจุบันจำนวนร้านอาหารมีมากมายและหลากหลายทั้งประเภทอาหารและการบริการทำให้การตัดสินใจของผู้บริโภคในการเลือกใช้บริการร้านอาหารยากขึ้น ผู้บริโภคจึงต้องการข้อมูลเพื่อช่วยในการตัดสินใจ และช่องทางหนึ่งที่ผู้คนในโลกออนไลน์นิยมใช้คือการเข้าไปหาข้อมูลในเว็บไซต์รีวิวออนไลน์ อย่างไรก็ตาม เนื่องจากความคิดเห็นส่วนใหญ่เป็นข้อความทำให้ผู้ใช้เว็บไซต์ดังกล่าวเสียเวลาในการอ่านและหาข้อสรุปเพื่อประกอบการตัดสินใจ ในส่วนของผู้ประกอบการเองก็ยากในการหาข้อดีข้อเสียสำหรับปรับปรุงร้านให้ดีขึ้น ดังนั้น งานวิจัยนี้ จึงเสนอระบบวิเคราะห์และสรุปความคิดเห็นเกี่ยวกับร้านอาหารจากเว็บไซต์รีวิวโดยอัตโนมัติ การวิเคราะห์จะใช้เทคนิคต่างๆ ในการประมวลผลภาษาธรรมชาติ เช่น เทคนิคการตัดคำ การวิเคราะห์ประเภทคำและรูปประโยค เพื่อหาความหมายเชิงบวกหรือเชิงลบของประโยค แล้วจึงคำนวณและแสดงผลออกมาเป็นค่าระดับความพึงพอใจสำหรับคุณลักษณะต่างๆ ในรูปแบบกราฟิกทั้งนี้เพื่อให้ผู้ใช้เข้าใจได้ง่าย ซึ่งจากผลการประเมินความพึงพอใจของผู้ใช้ ในด้านความสอดคล้องของผลสรุปจากระบบและผลสรุปจากผู้ใช้พบว่า การสรุปความคิดเห็นด้านบริการและอาหารอยู่ในระดับดี ขณะที่ด้านอื่นๆ อยู่ในระดับพอใช้

คำสำคัญ: การวิเคราะห์ความคิดเห็น, ความคิดเห็นเกี่ยวกับร้านอาหาร, การสรุปความคิดเห็น

บทนำ

ในปัจจุบันการรับประทานอาหารนอกบ้านโดยเฉพาะในภัตตาคารมีราคาค่อนข้างแพง ดังนั้นผู้บริโภคส่วนใหญ่ย่อมมีความต้องการการบริการที่ดีและคุณภาพอาหารที่คุ้มค่ากับเงินที่จ่ายไป ช่องทางหนึ่งที่ผู้บริโภคจะทราบถึงข้อมูลร้านอาหารเหล่านั้นนี้โดยการหาข้อมูลในเว็บไซต์ทางสังคมที่ให้บริการรีวิวร้านอาหาร เช่น เว็บไซต์ดวงใน (<http://www.wongnai.com>) และ TripAdvisor (<http://tripadvisor.com> เป็นต้น) เว็บไซต์รีวิวส่วนใหญ่นำเสนอความคิดเห็นจากผู้บริโภคที่เคยใช้บริการร้านอาหารในรูปแบบข้อความซึ่งถือเป็นข้อมูลที่ไม่มีโครงสร้าง หากร้านอาหารใดมีผู้แสดงความคิดเห็นมากก็จะปรากฏข้อความหลายร้อยข้อความ ทำให้ผู้ใช้ต้องเสียเวลาอ่านข้อคิดเห็นทั้งหมด แล้วทำการสรุปข้อดีข้อเสียในด้านต่าง ๆ สำหรับร้านที่ตนสนใจ แม้ว่าบางเว็บไซต์จะทราบถึงข้อจำกัดในเรื่องนี้และได้สร้างฟังก์ชัน rating เพื่อให้ผู้ใช้ให้คะแนนร้านอาหารในคุณลักษณะด้านต่างๆ แต่ความยุ่งยากนี้ก็กลับทำให้ผู้ที่เข้ามาแสดงความคิดเห็นมีจำนวนลดลง สุดท้ายหลายๆ เว็บไซต์จึงมีเพียงกล่องข้อความให้ผู้รีวิวเขียนข้อความ อัฟโหลด รูปภาพ และลดการ rating เหลือเฉพาะในภาพรวมของร้านเท่านั้น

จากปัญหาและที่มาข้างต้นผู้วิจัยได้นำเสนอระบบการวิเคราะห์ความคิดเห็นของผู้บริโภคบนเว็บไซต์รีวิวร้านอาหาร เพื่ออำนวยความสะดวกสำหรับผู้บริโภคในการรับรู้คุณภาพของร้านอาหารในด้านต่างๆโดยจะแบ่งรายละเอียดการสรุปออกเป็น 5 ข้อคือ (1) รสชาติอาหาร (2) บริการของร้าน (3) บรรยากาศของร้าน (4) ความสะอาด และ (5) ราคาอาหาร ระบบจะวิเคราะห์ข้อความด้วยเทคนิคการตัดคำ การวิเคราะห์ประเภทคำ และรูปประโยค เพื่อหาความหมายเชิงบวกหรือเชิงลบของประโยค แล้วจึงคำนวณและแสดงผลสรุปออกมาเป็นค่าระดับความพึงพอใจสำหรับแต่ละด้านในรูปแบบกราฟิกเพื่อให้ผู้บริโภคใช้ประกอบการตัดสินใจได้ง่ายยิ่งขึ้นนอกจากนี้ผู้พัฒนาคาดว่าผลสรุปจากความคิดเห็นจะเป็นประโยชน์สำหรับผู้ประกอบการร้านอาหารที่สามารถรับรู้ข้อบกพร่องของร้านตนเองเพื่อใช้ในการปรับปรุงร้านให้ดีขึ้น

บทคัดย่อ

- ในปัจจุบันจำนวนร้านอาหารมีมากมายและหลากหลายทั้งประเภทอาหารและการบริการทำให้การตัดสินใจของผู้บริโภคในการเลือกใช้บริการร้านอาหารยากขึ้น
- ผู้บริโภคจึงต้องการข้อมูลเพื่อช่วยในการตัดสินใจ และช่องทางหนึ่งที่ผู้คนในโลกออนไลน์นิยมใช้คือการเข้าไปหาข้อมูลในเว็บไซต์รีวิวออนไลน์
- อย่างไรก็ตามเนื่องจากความคิดเห็นส่วนใหญ่เป็นข้อความทำให้ผู้ใช้เว็บไซต์ดังกล่าวเสียเวลาในการอ่านและหาข้อสรุปเพื่อประกอบการตัดสินใจ
- ในส่วนของผู้ประกอบการเองก็ยากในการหาข้อดีข้อเสียสำหรับปรับปรุงร้านให้ดีขึ้น ดังนั้นงานวิจัยนี้
- จึงเสนอระบบวิเคราะห์และสรุปความคิดเห็นเกี่ยวกับร้านอาหารจากเว็บไซต์รีวิวโดยอัตโนมัติ
- การวิเคราะห์จะใช้เทคนิคต่างๆ ในการประมวลผลภาษาธรรมชาติ เช่น เทคนิคการตัดคำ การวิเคราะห์ประเภทคำและรูปประโยค เพื่อหาความหมายเชิงบวกหรือเชิงลบของประโยค แล้วจึงคำนวณและแสดงผลออกมาเป็นค่าระดับความพึงพอใจ
- สำหรับคุณลักษณะต่างๆในรูปแบบกราฟิกทั้งนี้เพื่อให้ผู้ใช้เข้าใจได้ง่าย ซึ่งจากผลการประเมินความพึงพอใจของผู้ใช้
- ในด้านความสอดคล้องของผลสรุปจากระบบและผลสรุปจากผู้ใช้พบว่าการสรุปความคิดเห็นด้านบริการและอาหารอยู่ในระดับดี ขณะที่ด้านอื่นๆ อยู่ในระดับพอใช้
- คำสำคัญ: การวิเคราะห์ความคิดเห็น ความคิดเห็นเกี่ยวกับร้านอาหาร การสรุปความคิดเห็น

บทนำ

- ในปัจจุบันการรับประทานอาหารนอกบ้านโดยเฉพาะในภัตตาคารมีราคาค่อนข้างแพง ดังนั้นผู้บริโภคส่วนใหญ่ย่อมมีความต้องการการบริการที่ดีและคุณภาพอาหารที่คุ้มค่ากับเงินที่จ่ายไป
- ช่องทางหนึ่งที่ผู้บริโภคจะทราบถึงข้อมูลร้านอาหารเหล่านั้นนี้โดยการหาข้อมูลในเว็บไซต์ทางสังคมที่ให้บริการรีวิวร้านอาหาร เช่น เว็บไซต์ดวงใจ (<http://www.wongnai.com>) และ TripAdvisor (<http://tripadvisor.com> เป็นต้น)
- เว็บไซต์รีวิวส่วนใหญ่นำเสนอความคิดเห็นจากผู้บริโภคที่เคยใช้บริการร้านอาหารในรูปแบบข้อความซึ่งถือเป็นข้อมูลที่ไม่มีโครงสร้าง หากร้านอาหารใดมีผู้แสดงความคิดเห็นมากก็จะปรากฏข้อความหลายร้อยข้อความ
- ทำให้ผู้ใช้ต้องเสียเวลาอ่านข้อคิดเห็นทั้งหมด แล้วทำการสรุปข้อดีข้อเสียในด้านต่าง ๆ สำหรับร้านที่ตนสนใจ
- แม้ว่าบางเว็บไซต์จะทราบถึงข้อจำกัดในเรื่องนี้และได้สร้างฟังก์ชัน rating เพื่อให้ผู้ใช้ให้คะแนนร้านอาหารในคุณลักษณะด้านต่างๆ แต่ความยุ่งยากนี้ก็กลับทำให้ผู้ที่เข้ามาแสดงความคิดเห็นมีจำนวนลดลง สุดท้ายหลายๆ เว็บไซต์มีเพียงกล่องข้อความให้ผู้รีวิวเขียนข้อความ อัฟโหลด รูปภาพ และลดการ rating เหลือเฉพาะในภาพรวมของร้านเท่านั้น
- จากปัญหาและที่มาข้างต้นผู้วิจัยได้นำเสนอระบบการวิเคราะห์ความคิดเห็นของผู้บริโภคบนเว็บไซต์รีวิวร้านอาหาร เพื่ออำนวยความสะดวกสำหรับผู้บริโภคในการรับรู้คุณภาพของร้านอาหารในด้านต่างๆ
- โดยจะแบ่งรายละเอียดการสรุปออกเป็น 5 ข้อคือ (1) รสชาติอาหาร (2) บริการของร้าน (3) บรรยากาศของร้าน (4) ความสะอาด และ (5) ราคาอาหาร ระบบจะวิเคราะห์ข้อความด้วยเทคนิคการตัดคำ การวิเคราะห์ประเภทคำ และรูปประโยค เพื่อหาความหมายเชิงบวกหรือเชิงลบของประโยค
- แล้วจึงคำนวณและแสดงผลสรุปออกมาเป็นค่าระดับความพึงพอใจสำหรับแต่ละด้านในรูปแบบกราฟิก เพื่อให้ผู้บริโภคใช้ประกอบการตัดสินใจได้ง่ายยิ่งขึ้น
- นอกจากนี้ผู้พัฒนาคาดว่าผลสรุปจากความคิดเห็นจะเป็นประโยชน์สำหรับผู้ประกอบการร้านอาหารที่สามารถรับรู้ข้อบกพร่องของร้านตนเองเพื่อใช้ในการปรับปรุงร้านให้ดีขึ้น

การวิเคราะห์ความรู้สึกทางอารมณ์ในบทความแนะนำสินค้าออนไลน์

บทคัดย่อ

ในงานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอการสร้างตัวแบบสำหรับการวิเคราะห์ระดับความรู้สึกของการแสดงความคิดเห็นต่อสินค้าและบริการออนไลน์ โดยใช้เทคนิคส่วนการเรียนรู้ของเครื่อง (Machine Learning) ด้านการทำเหมืองข้อความ (Text Mining) ร่วมกับเทคนิคการตัดคำ (Word Segmentation) 3 รูปแบบ และการสร้างคลังคำศัพท์ (Bag of words) จากนั้นนำเข้ากระบวนการแยกประเภทผลการวิเคราะห์ 4 เทคนิค ได้แก่ K-Nearest Neighbors, Random Forest, Logistic Regression และ Support Vector Machines (SVM) ในการสร้างตัวแบบมีกระบวนการทดลอง 5 ขั้นตอน ได้แก่ 1) การเก็บรวบรวมและการเตรียมข้อมูล (Data Preparation Phase) 2) การประมวลผลข้อความ (Text File Processing Phase) 3) การฝึกอบรมข้อมูล (Training & Streaming Phase) 4) การแยกประเภท (Classification Phase) และ 5) การประเมินตัวแบบ (Model Evaluation Phase) ทำการทดลองโดยใช้ข้อมูลตัวอย่างการแสดงความคิดเห็นต่อสินค้าและบริการออนไลน์บน Social Blog จำนวน 252 บทความ การแสดงความคิดเห็นจำนวน 1,412 ความเห็น จำนวน 83,670 คำ ใช้เป็นข้อมูลในการสร้างตัวแบบ เพื่อช่วยให้ผู้บริโภคตัดสินใจในการซื้อสินค้าและบริการ และช่วยให้ผู้ประกอบการมีข้อมูลสำหรับการพัฒนาสินค้าและบริการ ในอนาคตผลการทดลองสร้างตัวแบบการวิเคราะห์ระดับความรู้สึกทางอารมณ์ 3 ระดับ คือ ความรู้สึกเชิงบวก(Positive) เป็นกลาง (Neutral) และเชิงลบ (Negative) โดยแต่ละตัวแบบให้ความแม่นยำในการทำนายผล ดังนี้ K-Nearest Neighbors 52.6% Random Forest 71.6% Logistic Regression 77.2% และ Support Vector Machines (SVM) 79% สรุปการสร้างตัวแบบใช้เทคนิคการแยกประเภทแบบ Support Vector Machines (SVM) ร่วมกับเทคนิคการตัดคำแบบ Longest ให้คะแนนความแม่นยำดีที่สุดเมื่อเปรียบเทียบกับเทคนิคอื่นๆ เหมาะแก่การนำไปใช้ในการสร้างตัวแบบในการวิเคราะห์ระดับความรู้สึกสำหรับข้อความที่เป็นภาษาไทย

บทนำ

ปัจจุบันการซื้อขายสินค้าและบริการผ่านช่องทางออนไลน์หรือ อีคอมเมิร์ซ (E-Commerce) ในประเทศไทย มีสถิติการใช้อินเทอร์เน็ตเพิ่มขึ้นจากยอด 9.3 ล้านคนในปี พ.ศ. 2551 เป็น 45 ล้านคนในปีพ.ศ. 2560 จากยอดจดทะเบียนโทรศัพท์มือถืออีกกว่า 124 ล้านเลขหมาย นอกจากนี้พบว่า เวลาการใช้อินเทอร์เน็ตโดยเฉลี่ยคิดเป็นเวลา 10.05 ชั่วโมง/วัน/คน ข้อมูลเหล่านี้สะท้อนให้เห็นถึงการเข้าถึงและพฤติกรรมการใช้อินเทอร์เน็ตของคนไทยซึ่ง

เปลี่ยนไปตามยุคของเทคโนโลยีดิจิทัลที่ได้เข้ามามีบทบาทในชีวิตประจำวันมากขึ้น ซึ่งเป็นปัจจัยสำคัญที่ส่งผลให้มูลค่าอีคอมเมิร์ซในประเทศไทยมีการเติบโตขึ้นอย่างต่อเนื่อง โดยเฉลี่ยเพิ่มขึ้น 10% ต่อปี และการซื้อขายสินค้าและบริการออนไลน์เป็น 1 ใน 5 กิจกรรมยอดนิยมที่มีการใช้งานบนอินเทอร์เน็ตมากที่สุด (Electronic Transactions Development Agency (ETDA), 2019) ซึ่งในประเทศไทยมีเว็บไซต์ที่เปิดให้บริการในการซื้อขายสินค้าออนไลน์อยู่เป็นจำนวนมาก ในการตัดสินใจซื้อสินค้าและบริการผู้บริโภคจำนวนมากจะทำการสืบค้นข้อมูลต่างๆเกี่ยวกับตัวสินค้าและบริการ เช่น การสืบค้นข้อคิดเห็นต่อผลิตภัณฑ์ ข้อคิดเห็นต่อสินค้าและบริการ เป็นต้น ซึ่งเป็นช่องทางให้ผู้ที่สนใจใช้เป็นแนวทางในการตัดสินใจซื้อสินค้าและบริการ ซึ่งการแสดงความคิดเห็นดังกล่าวเป็นการแสดงถึงความรู้สึกทางอารมณ์ในรูปแบบข้อความที่มีต่อสินค้าและบริการ อีกทั้งข้อมูลการแสดงความคิดเห็นในปัจจุบันนี้มีปริมาณมหาศาลและเกิดขึ้นใหม่อยู่ตลอดเวลา ซึ่งทำให้ผู้บริโภคต้องใช้เวลาในการวิเคราะห์ข้อมูลเหล่านี้เพื่อใช้ในการตัดสินใจซื้อสินค้าและบริการมากขึ้น เนื่องจากยังไม่มีตัวช่วยในการวิเคราะห์คำและข้อความจากการแสดงความคิดเห็นต่อสินค้าและบริการแบบอัตโนมัติที่ช่วยในการตัดสินใจการเลือกซื้อสินค้าและบริการได้อย่างรวดเร็ว

การวิเคราะห์ความรู้สึกต่อความคิดเห็น (Sentiment Analysis) คือ กระบวนการวิเคราะห์และประเมินความรู้สึกต่อการแสดงความคิดเห็นบนสื่อสังคมออนไลน์ (Social Network) เช่น twitter.com, facebook.com, Instagram.com หรือ pantip.com เป็นต้น โดยมีเป้าหมายเพื่อแสดงทัศนคติทางอารมณ์ที่มีต่อข้อมูล บทวิจารณ์ โพสต์ กระทั่ง ที่เกี่ยวข้องกับการตลาด การประชาสัมพันธ์ การแสดงความคิดเห็นต่อผลิตภัณฑ์และบริการของลูกค้าบนโซเชียลมีเดีย (Sepandar and Jonathan, 2011) จากงานวิจัยที่ผ่านมาการวิเคราะห์ระดับความรู้สึก (Sentiment Analysis) ใช้กระบวนการประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP) และการทำเหมืองข้อความ (Text Mining) เป็นส่วนใหญ่ (Trupthi et al, 2017; Anurag and Vijay, 2015) อีกทั้งยังมีการใช้เทคนิค Bag-of-Words (CBOW) และ Skip-gram Model ร่วมด้วยเพื่อสร้างคลังคำศัพท์ในการวิเคราะห์ความรู้สึกที่เป็นเชิงบวกและเชิงลบ (Fan et al., 2016; Chumwatana, 2015) ในกรณีชุดข้อมูลที่เป็นภาษาไทยจะมีข้อจำกัดคือ การแสดงความคิดเห็นที่เป็นภาษาไทยจะเป็นข้อความที่ไม่มีการแบ่งส่วน แบ่งคำ ประโยคติดกัน อักขระที่ไม่มีขอบเขตคำ (Chumwatana, 2015; Netisopakul and Thong-iad, 2018) อีกทั้งคำศัพท์ในภาษาไทยหนึ่งคำยังสามารถสื่อความรู้สึกได้มากกว่าหนึ่งคำ (Netisopakul and Thong-iad, 2018) และในปัจจุบันการวิจัยเรื่องดังกล่าวในภาษาไทยยังมีไม่มากนัก อีกทั้งการประมวลผลภาษาไทยยังมีความยุ่งยากกว่าภาษาอังกฤษ เพราะโครงสร้างประโยคและไวยากรณ์มีความซับซ้อนมากกว่า (Sorjapo, 2019) และขาดแคลนแหล่งข้อมูล และเครื่องมือต่างๆ เช่น โปรแกรมแท็ก (Tag) ชนิดคำ เป็นต้น รวมไปถึงแนวโน้มของการเลือกใช้แบบจำลองในการวิเคราะห์ข้อมูล ซึ่งในปัจจุบันวิธีการเรียนรู้ด้วยเครื่อง (Machine Learning) เป็นที่นิยมมากกว่า

โดยอาศัยการสร้างชุดข้อมูลฝึกฝน (Training Set) (Lorattanachaiyong, 2017) และนำชุดข้อมูลทดสอบ (Test Set) เข้ากระบวนการจัดกลุ่มหรือแยกประเภท (Classification) และการประเมินคะแนนความแม่นยำของตัวแบบที่ได้โดยสรุปจากค่าความแม่นยำจาก Precision และ Recall นอกจากนี้การทำการวิเคราะห์ระดับความรู้สึกทางอารมณ์ส่วนใหญ่จะเป็นการตัดสินใจแบบ 2 ทิศทาง คือ แบบบวกและแบบลบ อย่างไรก็ตามจะมีความคิดเห็นที่เป็นกลาง ซึ่งถ้าหากถูกตัดสินเพียง 2 ระดับ อาจจะทำให้ข้อคิดเห็นโดยรวมเกิดอคติได้ ดังนั้น ถ้าหากมีการวิเคราะห์เป็นแบบ 3 ระดับ คือ มีระดับกลางเพิ่มขึ้นมา จะทำให้น้ำหนักของการตัดสินแบบบวกและแบบลบ มีความถูกต้องมากขึ้น

บทคัดย่อ

● ในงานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอการสร้างตัวแบบสำหรับการวิเคราะห์ระดับความรู้สึกของการแสดงความคิดเห็นต่อสินค้าและบริการออนไลน์ โดยใช้เทคนิคส่วนการเรียนรู้ของเครื่อง (Machine Learning)

ด้านการทำเหมืองข้อความ (Text Mining) ร่วมกับเทคนิคการตัดคำ (Word Segmentation) 3 รูปแบบ

- และการสร้างคลังคำศัพท์ (Bag of words) จากนั้นนำเข้ากระบวนการแยกประเภทผลการวิเคราะห์ 4 เทคนิค ได้แก่ K-Nearest Neighbors, Random Forest, Logistic Regression และ Support Vector Machines (SVM)
- ในการสร้างตัวแบบมีกระบวนการทดลอง 5 ขั้นตอน ได้แก่ 1) การเก็บรวบรวมและการเตรียมข้อมูล (Data Preparation Phase) 2) การประมวลผลข้อความ (Text File Processing Phase) 3) การฝึกอบรมข้อมูล (Training & Streaming Phase) 4) การแยกประเภท (Classification Phase) และ 5) การประเมินตัวแบบ (Model Evaluation Phase)
- ทำการทดลองโดยใช้ข้อมูลตัวอย่างการแสดงความคิดเห็นต่อสินค้าและบริการออนไลน์บน Social Blog จำนวน 252 บทความ
- การแสดงความคิดเห็นจำนวน 1,412 ความเห็น จำนวน 83,670 คำ ใช้เป็นข้อมูลในการสร้างตัวแบบ เพื่อช่วยให้ผู้บริโภคตัดสินใจในการซื้อสินค้าและบริการ และช่วยให้ผู้ประกอบการมีข้อมูลสำหรับการพัฒนาสินค้าและบริการ

- ในอนาคตผลการทดลองสร้างตัวแบบการวิเคราะห์ระดับความรู้สึกทางอารมณ์ 3 ระดับ คือ ความรู้สึกเชิงบวก(Positive) เป็นกลาง (Neutral) และเชิงลบ (Negative) โดยแต่ละตัวแบบให้ความแม่นยำในการทำนายผล ดังนี้ K-Nearest Neighbors 52.6% Random Forest 71.6% Logistic Regression 77.2% และ Support Vector Machines (SVM) 79%
- สรุปการสร้างตัวแบบใช้เทคนิคการแยกประเภทแบบ Support Vector Machines (SVM) ร่วมกับเทคนิคการตัดคำแบบ Longest ให้คะแนนความแม่นยำดีที่สุดเมื่อเปรียบเทียบกับเทคนิคอื่นๆ เหมาะแก่การนำไปใช้ในการสร้างตัวแบบในการวิเคราะห์ระดับความรู้สึกสำหรับข้อความที่เป็นภาษาไทย

บทนำ

- ปัจจุบันการซื้อขายสินค้าและบริการผ่านช่องทางออนไลน์หรือ อีคอมเมิร์ซ (E-Commerce) ในประเทศไทย มีสถิติการใช้อินเทอร์เน็ตเพิ่มขึ้นจากยอด 9.3 ล้านคนในปี พ.ศ. 2551 เป็น 45 ล้านคนในปีพ.ศ. 2560 จากยอดจดทะเบียนโทรศัพท์มือถือกว่า 124 ล้านเลขหมาย
- นอกจากนี้พบว่า เวลาการใช้อินเทอร์เน็ตโดยเฉลี่ยคิดเป็นเวลา 10.05 ชั่วโมง/วัน/คน ข้อมูลเหล่านี้สะท้อนให้เห็นถึงการเข้าถึงและพฤติกรรมการใช้เน็ตของคนไทยซึ่งเปลี่ยนแปลงไปตามยุคของเทคโนโลยีดิจิทัลที่ได้เข้ามามี
- บทบาทในชีวิตประจำวันมากขึ้น ซึ่งเป็นปัจจัยสำคัญที่ส่งผลให้มูลค่าอีคอมเมิร์ซในประเทศไทยมีการเติบโตขึ้นอย่างต่อเนื่อง โดยเฉลี่ยเพิ่มขึ้น 10% ต่อปี
- และการซื้อขายสินค้าและบริการออนไลน์เป็น 1 ใน 5 กิจกรรมยอดนิยมที่มีการใช้งานบนอินเทอร์เน็ตมากที่สุด (Electronic Transactions Development Agency (ETDA), 2019) ซึ่งในประเทศไทยมีเว็บไซต์ที่เปิดให้บริการในการซื้อขายสินค้าออนไลน์อยู่เป็นจำนวนมาก
- ในการตัดสินใจซื้อสินค้าและบริการผู้บริโภคจำนวนมากจะทำการสืบค้นข้อมูลต่างๆเกี่ยวกับตัวสินค้าและบริการ เช่น การสืบค้นข้อคิดเห็นต่อผลิตภัณฑ์ ข้อคิดเห็นต่อสินค้าและบริการ เป็นต้น ซึ่งเป็นช่องทางให้ผู้สนใจใช้เป็นแนวทางในการตัดสินใจซื้อสินค้าและบริการ
- ซึ่งการแสดงความคิดเห็นดังกล่าวเป็นการแสดงถึงความรู้สึกทางอารมณ์ในรูปแบบข้อความที่มีต่อสินค้าและบริการ

- **อีกทั้ง**ข้อมูลการแสดงความคิดเห็นในปัจจุบันนี้มีปริมาณมหาศาลและเกิดขึ้นใหม่อยู่ตลอดเวลา ซึ่งทำให้ผู้บริโภคต้องใช้เวลาในการวิเคราะห์ข้อมูลเหล่านี้เพื่อใช้ในการตัดสินใจซื้อสินค้าและบริการมากขึ้น
- **เนื่องจาก**ยังไม่มีตัวช่วยในการวิเคราะห์คำและข้อความจากการแสดงความคิดเห็นต่อสินค้าและบริการแบบอัตโนมัติที่ช่วยในการตัดสินใจการเลือกซื้อสินค้าและบริการได้อย่างรวดเร็ว
- **การวิเคราะห์ความรู้สึก**ต่อความคิดเห็น (Sentiment Analysis) คือ กระบวนการวิเคราะห์และประเมินความรู้สึกต่อการแสดงความคิดเห็นบนสื่อสังคมออนไลน์ (Social Network) เช่น twitter.com, facebook.com, Instagram.com หรือ pantip.com เป็นต้น
- **โดยมี**เป้าหมายเพื่อแสดงทัศนคติทางอารมณ์ที่มีต่อข้อมูล บทวิจารณ์ โพสต์ กระทั่ง ที่เกี่ยวข้องกับ การตลาด การประชาสัมพันธ์ การแสดงความคิดเห็นต่อผลิตภัณฑ์และบริการของลูกค้าบนโซเชียลมีเดีย (Sepandar and Jonathan, 2011)
- **จากงานวิจัย**ที่ผ่านมาการวิเคราะห์ระดับความรู้สึก (Sentiment Analysis) ใช้กระบวนการประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP) และการทำเหมืองข้อความ (Text Mining) เป็นส่วนใหญ่ (Trupthi et al, 2017; Anurag and Vijay, 2015)
- **อีกทั้ง**ยังมีการใช้เทคนิค Bag-of-Words (CBOW) และ Skip-gram Model ร่วมด้วยเพื่อสร้างคลังคำศัพท์ในการวิเคราะห์ความรู้สึกที่เป็นเชิงบวกและเชิงลบ (Fan et al., 2016; Chumwatana, 2015) ในกรณีชุดข้อมูลที่เป็นภาษาไทยจะมีข้อจำกัดคือ การแสดงความคิดเห็นที่เป็นภาษาไทยจะเป็นข้อความที่ไม่มีการแบ่งส่วน แบ่งคำ ประโยคติดกัน อักษรที่ไม่มีขอบเขตคำ (Chumwatana, 2015; Netisopakul and Thong-iad, 2018)
- **อีกทั้ง**คำศัพท์ในภาษาไทยหนึ่งคำยังสามารถสื่อความรู้สึกได้มากกว่าหนึ่งคำ (Netisopakul and Thong-iad, 2018) และในปัจจุบันการวิจัยเรื่องดังกล่าวในภาษาไทยยังมีไม่มากนัก อีกทั้งการประมวลผลภาษาไทยยังมีความยุ่งยากกว่าภาษาอังกฤษ เพราะโครงสร้างประโยคและไวยากรณ์มีความซับซ้อนมากกว่า (Sorjapo, 2019) และขาดแคลนแหล่งข้อมูล และเครื่องมือต่างๆ เช่น โปรแกรมแท็ก (Tag) ชนิดคำ เป็นต้น
- **รวมถึง**แนวโน้มของการเลือกใช้แบบจำลองในการวิเคราะห์ข้อมูล ซึ่งในปัจจุบันวิธีการเรียนรู้ด้วยเครื่อง (Machine Learning) เป็นที่นิยมมากกว่า โดยอาศัยการสร้างชุดข้อมูลฝึกฝน (Training Set) (Lorattanachaiyong, 2017)

- และนำชุดข้อมูลทดสอบ (Test Set) เข้ากระบวนการจัดกลุ่มหรือแยกประเภท (Classification) และการประเมินคะแนนความแม่นยำของตัวแบบที่ได้โดยสรุปจากค่าความแม่นยำจาก Precision และ Recall
- นอกจากนี้การทำการวิเคราะห์ระดับความรู้สึกทางอารมณ์ส่วนใหญ่จะเป็นการตัดสินใจแบบ 2 ทิศทาง คือ แบบบวกและแบบลบ
- อย่างไรก็ตามจะมีความคิดเห็นที่เป็นกลาง ซึ่งถ้าหากถูกตัดสินเพียง 2 ระดับ อาจจะทำให้ข้อคิดเห็นโดยรวมเกิดอคติได้ ดังนั้น ถ้าหากมีการวิเคราะห์เป็นแบบ 3 ระดับ คือ มีระดับกลางเพิ่มขึ้นมา จะทำให้น้ำหนักของการตัดสินใจแบบบวกและแบบลบ มีความถูกต้องมากขึ้น

การพัฒนาระบบตรวจสอบข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึง

บทคัดย่อ

การวิจัยเรื่องการพัฒนาแบบตรวจสอบข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึง มีวัตถุประสงค์เพื่อ 1) พัฒนาระบบตรวจสอบข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึง และ 2) ประเมินผลการใช้ระบบตรวจสอบข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึงสถิติที่ใช้ในการวิเคราะห์ข้อมูลคือ ค่าเฉลี่ยเลขคณิต (\bar{X}) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) กลุ่มเป้าหมาย คือ อาจารย์ และนักศึกษา หลักสูตรวิทยาการคอมพิวเตอร์ จำนวน 20 คน และแบบทดสอบที่มีคำตอบ จำนวน 170 คำตอบ

ผลการวิจัยพบว่า 1) ระบบตรวจสอบข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึงได้แบ่งผู้ใช้งาน 2 กลุ่ม ได้แก่ ผู้ออกข้อสอบ และผู้ทำข้อสอบ ซึ่งมีขั้นตอนคือสร้างคำถาม คำตอบ บันทึกลงฐานข้อมูลแล้ว ประยุกต์คำตอบเข้าสู่กระบวนการทางเทคโนโลยีประมวลผลภาษาธรรมชาติ เริ่มต้นจากกระบวนการเตรียมข้อมูล แบ่งคำและตัดคำไม่สำคัญ คำนวณค่าความถี่ของคำและค่าน้ำหนักของคำที่เกิดขึ้นในคำตอบ และคำนวณค่าความ เหมือนความคล้าย ระหว่างคำตอบผู้ออกข้อสอบและคำตอบของผู้ทำข้อสอบ ผ่านรูปแบบเว็บแอปพลิเคชันที่พัฒนา ด้วยภาษาไพทอน พบว่า ระบบสามารถใช้งานได้และผู้เชี่ยวชาญประเมินอยู่ในระดับมาก และ 2) ผลการประเมิน ความพึงพอใจต่อการใช้งานระบบมีผลประเมินโดยรวมทุกด้าน อยู่ในระดับพึงพอใจมาก ($\bar{X}=4.05$, S.D.=0.81)

บทนำ

ปัจจุบันการใช้เทคโนโลยีปัญญาประดิษฐ์มาประยุกต์ในงานต่าง ๆ มีแพร่หลายมากขึ้น โดยเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) ที่ทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์เพื่อสื่อสารและวิเคราะห์ข้อมูล เลียนแบบสมองของมนุษย์ได้ เทคโนโลยีการประมวลผลภาษาธรรมชาติ(Natural Language Processing: NLP) เป็นวิทยาการแขนงหนึ่งในเทคโนโลยีปัญญาประดิษฐ์ ที่ศึกษา เลียนแบบพฤติกรรม หรือคุณลักษณะของภาษาที่ใช้ในสื่อสารของมนุษย์มาประมวลผลด้วยเครื่องคอมพิวเตอร์ ยกตัวอย่างภาษาพูดทั่วไปที่ใช้พูดคุยกันตามปกติในชีวิตประจำวัน การเขียนข้อความความคิดเห็นในรูปแบบต่าง ๆ หลักการทำงานของ การประมวลผลภาษาธรรมชาติจะมีลักษณะคล้ายกันกับภาษาของมนุษย์โดยมีเทคนิคบางอย่างที่ทำให้เครื่องคอมพิวเตอร์สามารถทำความเข้าใจและประมวลผลภาษาเหล่านั้นได้ซึ่งทางสถาบัน SAS [1] ได้กล่าวไว้ และจะช่วยให้ประสิทธิภาพการทำงานของมนุษย์ดีขึ้น

การประยุกต์ใช้งานเทคโนโลยีประมวลผลภาษาธรรมชาติสามารถนำไปใช้ในงานต่าง ๆ มากมาย การศึกษาก็เช่นกัน [2] ส่งผลทำให้การเรียนการสอนมีประสิทธิภาพมากขึ้น โดยเฉพาะงานการตรวจความถูกต้องของข้อมูล ซึ่งถือเป็นแนวทางการวัดและประเมินผลหนึ่งของการศึกษา การตรวจข้อสอบอัตนัยนั้นค่อนข้างมีความท้าทาย เนื่องจากต้องใช้วิจารณ์ญาณของครูอาจารย์ และความถูกต้องของคำตอบที่ค่อนข้างแม่นยำ และใช้เวลานานในการตรวจสอบ [3] โดยเฉพาะภาษาไทยจะพบปัญหาเรื่องการแบ่งคำ แบ่งประโยคของข้อความ ซึ่งถ้ามองในด้านเทคนิคการประมวลผลภาษาธรรมชาติแล้ว ค่อนข้างยากในการพัฒนาระบบ เพื่อให้เกิดความสะดวกสบายและการตรวจคำตอบที่ง่ายด้วยโปรแกรมที่ถูกพัฒนาขึ้น การตรวจข้อสอบอัตนัยออนไลน์จึงเป็นที่น่าสนใจ ที่จะพัฒนาเป็นนวัตกรรมด้านซอฟต์แวร์ และเผยแพร่ผลงานให้กว้างขวางมากยิ่งขึ้น

ดังนั้นทางผู้วิจัยเล็งเห็นว่าการพัฒนาระบบตรวจข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึง เริ่มต้นจากการสร้างคำถาม คำตอบ แล้วบันทึกหลักฐานข้อมูล แล้วนำคำตอบที่ได้ ผ่านกระบวนการทางเทคโนโลยีประมวลผลภาษาธรรมชาติซึ่งประกอบไปด้วยการเตรียมข้อมูล การตัดคำไม่สำคัญ [4] การคำนวณค่าความถี่ของคำและค่าน้ำหนักของคำที่เกิดขึ้นในคำตอบ [5] และการคำนวณค่าความคล้ายคลึง หรือความเหมือนความคล้าย [6] ระหว่างคำตอบของผู้ออกข้อสอบและคำตอบของผู้ทำข้อสอบในโครงสร้างทางคณิตศาสตร์ หรือที่เรียกว่าเวกเตอร์และนำแนวทางที่ได้ไปใช้งาน ทั้งนี้ได้เลือกใช้ภาษาไพทอน รองรับการทำงานผ่านเว็บไซต์ แอปพลิเคชันที่เขียนขึ้นสำหรับภาษาไพทอนที่ทำงานร่วมกับเว็บเซิร์ฟเวอร์ได้[7]

บทคัดย่อ

- การวิจัยเรื่องการพัฒนาระบบตรวจข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึง มีวัตถุประสงค์เพื่อ
- 1) พัฒนาระบบตรวจข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึง
- และ 2) ประเมินผลการใช้ระบบตรวจข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึงสถิติที่ใช้ในการวิเคราะห์ ข้อมูลคือ ค่าเฉลี่ยเลขคณิต (\bar{X}) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) กลุ่มเป้าหมาย คือ อาจารย์ และนักศึกษา หลักสูตรวิทยาการคอมพิวเตอร์ จำนวน 20 คน และแบบทดสอบที่มีคำตอบจำนวน 170 คำตอบ
- ผลการวิจัยพบว่า 1) ระบบตรวจข้อสอบอัตนัยภาษาไทยบนพื้นฐานการวัดค่าความคล้ายคลึงได้แบ่งผู้ใช้งาน 2 กลุ่ม ได้แก่ ผู้ออกข้อสอบ และผู้ทำข้อสอบ

- ซึ่งมีขั้นตอนคือสร้างคำถาม คำตอบ บันทึกหลักฐานข้อมูล แล้ว ประยุกต์คำตอบเข้าสู่กระบวนการทางเทคโนโลยีประมวลผลภาษาธรรมชาติ
- เริ่มต้นจากกระบวนการเตรียมข้อมูล แบ่งคำและตัดคำไม่สำคัญ คำนวณค่าความถี่ของคำและค่าน้ำหนักของคำที่เกิดขึ้นในคำตอบ และคำนวณค่าความ เหมือนความคล้าย ระหว่างคำตอบผู้ออกข้อสอบและคำตอบของผู้ทำข้อสอบผ่านรูปแบบเว็บแอปพลิเคชันที่พัฒนา ด้วยภาษาไพทอน
- พบว่า ระบบสามารถใช้งานได้และผู้เชี่ยวชาญประเมินอยู่ในระดับมาก
- และ 2) ผลการประเมิน ความพึงพอใจต่อการใช้งานระบบมีผลประเมินโดยรวมทุกด้าน อยู่ในระดับพึงพอใจมาก ($\bar{X}=4.05$, S.D.=0.81)

บทนำ

- ปัจจุบันการใช้เทคโนโลยีปัญญาประดิษฐ์มาประยุกต์ในงานต่าง ๆ มีแพร่หลายมากขึ้น โดยเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) ที่ทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์เพื่อสื่อสารและวิเคราะห์ ข้อมูล เลียนแบบสมองของมนุษย์ได้
- เทคโนโลยีการประมวลผลภาษาธรรมชาติ(Natural Language Processing: NLP) เป็นวิทยาการแขนงหนึ่งในเทคโนโลยีปัญญาประดิษฐ์ ที่ศึกษา เลียนแบบพฤติกรรม หรือคุณลักษณะของภาษาที่ใช้ในสื่อสารของมนุษย์มาประมวลผลด้วยเครื่องคอมพิวเตอร์ ยกตัวอย่างภาษาพูดทั่วไปที่ใช้พูดคุยกัน ตามปกติในชีวิตประจำวัน
- การเขียนข้อความความคิดเห็นในรูปแบบต่าง ๆ หลักการทำงานของ การประมวลผล ภาษาธรรมชาติจะมีลักษณะคล้ายกันกับภาษาของมนุษย์โดยมีเทคนิคบางอย่างที่ทำให้เครื่องคอมพิวเตอร์สามารถ ทำความเข้าใจและประมวลผลภาษาเหล่านั้นได้ซึ่งทางสถาบัน SAS [1] ได้กล่าวไว้ และจะช่วยให้ประสิทธิภาพการทำงาน of มนุษย์ดีขึ้น
- การประยุกต์ใช้งานเทคโนโลยีประมวลผลภาษาธรรมชาติสามารถนำไปใช้ในงานต่าง ๆ มากมาย การศึกษาก็เช่นกัน [2] ส่งผลทำให้การเรียนการสอนมีประสิทธิภาพมากขึ้น
- โดยเฉพาะงานการตรวจความถูกต้อง ของข้อมูล ซึ่งถือเป็นแนวทางการวัดและประเมินผลหนึ่งของ การศึกษา การตรวจข้อสอบอัตนัยนั้นค่อนข้างมีความ ทำหาย เนื่องจากต้องใช้วิจารณ์ญาณของครู อาจารย์ และความถูกต้องของคำตอบที่ค่อนข้างแม่นยำ และใช้เวลานานในการตรวจสอบ [3]

- โดยเฉพาะภาษาไทยจะพบปัญหาเรื่องการแบ่งคำ แบ่งประโยคของข้อความ ซึ่งถ้า มองในด้านเทคนิคการประมวลผลภาษาธรรมชาติแล้ว ค่อนข้างยากในการพัฒนาระบบ เพื่อให้เกิดความ สะดวกสบายและการตรวจคำตอบที่ง่ายด้วยโปรแกรมที่ถูกพัฒนาขึ้น
- การตรวจสอบข้อสอบอัตโนมัติออนไลน์จึงเป็นที่น่าสนใจ ที่จะพัฒนาเป็นนวัตกรรมด้านซอฟต์แวร์ และเผยแพร่ผลงานให้กว้างขวางมากยิ่งขึ้น
- ดังนั้นทางผู้วิจัยเล็งเห็นว่าการพัฒนาระบบตรวจสอบข้อสอบอัตโนมัติภาษาไทยบนพื้นฐานการวัดค่าความ คล้ายคลึง เริ่มต้นจากการสร้างคำถาม คำตอบ แล้วบันทึกหลักฐานข้อมูล
- แล้วนำคำตอบที่ได้ ผ่านกระบวนการทางเทคโนโลยีประมวลผลภาษาธรรมชาติซึ่งประกอบไปด้วยการเตรียมข้อมูล การตัดคำไม่สำคัญ [4] การคำนวณ ค่าความถี่ของคำและค่าน้ำหนักของคำที่เกิดขึ้นในคำตอบ [5]
- และการคำนวณค่าความคล้ายคลึง หรือความเหมือนความคล้าย [6] ระหว่างคำตอบของผู้ออกข้อสอบ และคำตอบของผู้ทำข้อสอบในโครงสร้างทางคณิตศาสตร์หรือที่เรียกว่าเวกเตอร์และนำแนวทางที่ได้ไปใช้งาน
- ทั้งนี้ได้เลือกใช้ภาษาไพทอน รองรับการทำงานผ่านเว็บไซต์แอป ฟลิเคชันที่เขียนขึ้นสำหรับภาษาไพทอนที่ทำงานร่วมกับเว็บเซิร์ฟเวอร์ได้[7]

Thai CAPTCHA : Construction and Analysis

บทคัดย่อ

เว็บไซต์ต่างๆ ในปัจจุบันที่เปิดโอกาสให้ผู้ใช้สามารถส่งข้อมูลไปยังเครื่องแม่ข่ายได้ เช่นเว็บไซต์ที่มีการสมัครสมาชิก หรือ เว็บไซต์ที่มีเว็บบอร์ด เป็นต้น เว็บไซต์เหล่านี้มักตกเป็นเป้าหมายของการโจมตี โดยการใช้บ็อตเพื่อสร้างข้อมูลที่ผู้ควบคุมบ็อต ต้องการไปยังเครื่องแม่ข่ายของเว็บไซต์นั้นๆ ซึ่งอาจจะเป็นการโพสต์ข้อความขายของ หรือการสมัครสมาชิกปลอม เป็นต้น จากการสำรวจพบว่าบ็อตที่โจมตีตามเว็บไซต์ต่างๆ มาจากการภายนอกประเทศไทยเป็นหลัก ดังนั้นบทความนี้จึงนำเสนอเทคนิคของการสร้าง CAPTCHA โดยใช้ตัวอักษรภาษาไทย เพื่อที่จะแบ่งแยกผู้ใช้ที่เป็นคนไทยและผู้ใช้ที่เป็นบ็อต นอกจากนั้นยังมีการใส่การทำงานทางคณิตศาสตร์อย่างง่าย และใส่การรบกวนเพื่อเพิ่มประสิทธิภาพในการแบ่งแยกระหว่างผู้ใช้ที่เป็นคนและเครื่องคอมพิวเตอร์ และได้มีการวิเคราะห์ผลลัพธ์ของการใช้โปรแกรม OCR ที่ชื่อว่า Amthai เพื่อที่จะกำหนดคุณสมบัติของ CAPTCHA ที่นำเสนอให้มีประสิทธิภาพมากขึ้น

คำสำคัญ - แคปช่า, บ็อต

บทนำ

ในปัจจุบัน เว็บไซต์ต่างๆ มีหน้าเว็บที่ต้องการให้ผู้ใช้กรอกข้อมูลเพื่อส่งไปยังเครื่องแม่ข่าย เพื่อประมวลผลหรือจัดเก็บต่อไป อย่างไรก็ตามให้มีความพยายามที่จะหาวิธีการที่ผู้ใช้ไม่ต้องทำการสร้างข้อมูลเอง โดยอาศัยโปรแกรมที่ทำงานบนเครื่องคอมพิวเตอร์ ทำการสร้างข้อมูล แล้วทำการจัดส่งไปยังเครื่องแม่ข่ายโดยอัตโนมัติ โปรแกรมที่ทำงานบนเครื่องคอมพิวเตอร์เหล่านี้จะถูกคิดตั้งอย่างตั้งใจหรือไม่ตั้งใจ (ไวรัสหรือเวิร์ม) บนเครื่องคอมพิวเตอร์หนึ่งๆ และถูกควบคุมหรือสั่งการจากเครื่องคอมพิวเตอร์อีกเครื่องหนึ่งที่เป็นผู้สั่งการ เรามักเรียกเครื่องคอมพิวเตอร์ที่ถูกควบคุมการทำงานนี้ว่า บ็อต โดยปกติแล้ว บ็อต ถูกใช้อย่างแพร่หลายในหลายๆเรื่อง เช่น การช่วยค้นหาข้อมูลของการทำงานของเว็บไซต์กูเกิ้ล [1] การใช้บ็อตทำการเล่นเกมต่างๆ[2] แต่ปัญหาที่เกิดขึ้นส่วนใหญ่ก็คือการใช้บ็อตเพื่อปลอมตัวในการสมัครสมาชิกทางเว็บต่างๆ หรือ การใช้บ็อตสร้างข้อมูลบนเว็บบอร์ดหรือบล็อกเพื่อล่อลวงให้เชื่อมโยงไปยังเว็บไซต์ที่ผู้ควบคุมบ็อตต้องการ หรือการใช้บ็อตในการโจมตีแบบปฏิเสธการให้บริการ (Denial Of Service)

ดังนั้น CAPTCHA [3] ได้ถูกพัฒนาขึ้นเพื่อป้องกันการดำเนินงานของบ็อต ในการส่งข้อมูลไปยังเว็บไซต์ต่างๆ เช่น Google, Ebay เป็นต้น อย่างไรก็ตามงานวิจัย [4] ได้แสดงถึงจุดอ่อนของเทคนิคต่างบนแคปช่า ที่มีการใช้งาน

อยู่ในขณะนี้ อย่างไรก็ตามเมื่อสำรวจการทำงานของเว็บไซต์ต่างๆ ของประเทศไทยที่มีผู้ใช้เป็นคนไทย โดยส่วนใหญ่จะนำCAPTCHA ที่สร้างจากต่างประเทศ ซึ่งมีความเป็นไปได้ที่บ็อต จะสามารถแก้ปัญหาได้ตามที่ได้กล่าวถึงใน [4] ดังนั้นบทความนี้จะนำเสนอการสร้างและวิเคราะห์ CAPTCHA ที่ใช้ตัวอักษรภาษาไทยเพื่อป้องกันการโจมตีจากบ็อตที่มาจากต่างประเทศ

การจัดเรียงเนื้อหาของบทความเป็นดังต่อไปนี้ หัวข้อที่ 2 กล่าวถึงการทำงานพื้นฐานของ CAPTCHA และงานวิจัยที่เกี่ยวข้อง หัวข้อที่ 3 อธิบายแนวคิดและ โครงสร้างของวิธีการที่นำเสนอ หัวข้อที่ 4 อธิบายถึงวิธีการทดลอง หัวข้อที่ 5 เสนอผลการทดลองและวิเคราะห์ หัวข้อที่ 6 เป็นบทสรุปและงานวิจัยในอนาคต

บทคัดย่อ

- เว็บไซต์ต่างๆ ในปัจจุบันที่เปิดโอกาสให้ผู้ใช้สามารถส่งข้อมูลไปยังเครื่องแม่ข่ายได้ เช่นเว็บไซต์ที่มีการสมัครสมาชิก หรือ เว็บไซต์ที่มีเว็บบอร์ด เป็นต้น
- เว็บไซต์เหล่านี้มักตกเป็นเป้าหมายของการโจมตีโดยการใช้บ็อต เพื่อสร้างข้อมูลที่ผู้ควบคุมบ็อต ต้องการไปยังเครื่องแม่ข่ายของเว็บไซต์นั้นๆ ซึ่งอาจจะเป็นการโพสต์ข้อความขายของ หรือการสมัครสมาชิกปลอม เป็นต้น
- จากการสำรวจพบว่าบ็อตที่ โจมตีตามเว็บไซต์ต่างๆ มาจากการภายนอกประเทศไทยเป็นหลัก
- ดังนั้นบทความนี้จะนำเสนอเทคนิคของการสร้าง CAPTCHA โดยใช้ตัวอักษรภาษาไทย เพื่อที่จะแบ่งแยกผู้ใช้ที่เป็นคนไทยและผู้ใช้ที่เป็นบ็อต
- นอกจากนั้นยังมีการใส่การทำงานทางคณิตศาสตร์อย่างง่ายและใส่การรบกวนเพื่อเพิ่มประสิทธิภาพในการแบ่งแยกระหว่างผู้ใช้ที่เป็นคนและเครื่องคอมพิวเตอร์
- และได้มีการวิเคราะห์ผลลัพธ์ของการใช้โปรแกรม OCR ที่ชื่อว่า Amthai เพื่อที่จะกำหนดคุณสมบัติของ CAPTCHA ที่นำเสนอให้มีประสิทธิภาพมากขึ้น

คำสำคัญ - แคปช่า, บ็อต

บทนำ

- ในปัจจุบัน เว็บไซต์ต่างๆ มีหน้าเว็บที่ต้องการให้ผู้ใช้กรอกข้อมูลเพื่อส่งไปยังเครื่องแม่ข่าย เพื่อประมวลผลหรือจัดเก็บต่อไป

- **อย่างไรก็ตาม** ให้ความพยายามที่จะหาวิธีการที่ผู้ใช้ไม่ต้องทำการสร้างข้อมูลเอง โดยอาศัยโปรแกรมที่ทำงานบนเครื่องคอมพิวเตอร์ ทำการสร้างข้อมูล แล้วทำการจัดส่งไปยังเครื่องแม่ข่ายโดยอัตโนมัติ
- **โปรแกรม** ที่ทำงานบนเครื่องคอมพิวเตอร์เหล่านี้จะถูกติดตั้งอย่างตั้งใจหรือไม่ตั้งใจ (ไวรัสหรือเวิร์ม) บนเครื่องคอมพิวเตอร์หนึ่งๆ และถูกควบคุมหรือสั่งการจากเครื่องคอมพิวเตอร์อีกเครื่องหนึ่งที่เป็นผู้สั่งการ
- **เรามักเรียก** เครื่องคอมพิวเตอร์ที่ถูกควบคุมการทำงานนี้ว่า บ็อต
- **โดยปกติแล้ว** บ็อต ถูกใช้อย่างแพร่หลายในหลายๆเรื่อง

เช่นการช่วยค้นหาข้อมูลของการทำงานของเว็บไซต์กูเกิ้ล [1] การใช้บ็อตทำการเล่นเกมต่างๆ[2]

- **แต่ปัญหา** ที่เกิดขึ้นส่วนใหญ่ก็คือการใช้บ็อตเพื่อปลอมตัวในการสมัครสมาชิกทางเว็บต่างๆ หรือ การใช้บ็อตสร้างข้อมูลบนเว็บบอร์ดหรือบล็อกเพื่อล่อลวงให้เชื่อมโยงไปยังเว็บไซต์ที่ผู้ควบคุมบ็อตต้องการ หรือการใช้บ็อตในการโจมตีแบบปฏิเสธการให้บริการ (Denial Of Service)
- **ดังนั้น** CAPTCHA [3] ได้ถูกพัฒนาขึ้นเพื่อป้องกันการดำเนินงานของบ็อต ในการส่งข้อมูลไปยังเว็บไซต์ต่างๆ เช่น Google, Ebay เป็นต้น
- **อย่างไรก็ตาม** งานวิจัย [4] ได้แสดงถึงจุดอ่อนของเทคนิคต่างบนแคปช่า ที่มีการใช้งานอยู่ในขณะนี้ อย่างไรก็ตามเมื่อสำรวจการทำงานของเว็บไซต์ต่างๆ ของประเทศไทยที่มีผู้ใช้เป็นคนไทย โดยส่วนใหญ่จะนำCAPTCHA ที่สร้างจากต่างประเทศ ซึ่งมีความเป็นไปได้ที่บ็อต จะสามารถแก้ปัญหาได้ตามที่ได้กล่าวถึงใน [4]
- **ดังนั้น** บทความนี้จึงนำเสนอการสร้างและวิเคราะห์ CAPTCHA ที่ใช้ตัวอักษรภาษาไทยเพื่อป้องกันการโจมตีจากบ็อต ที่มาจากต่างประเทศ
- **การจัดเรียง** เนื้อหาของบทความเป็นดังต่อไปนี้ หัวข้อที่ 2 กล่าวถึงการทำงานพื้นฐานของ CAPTCHA และงานวิจัยที่เกี่ยวข้อง
- **หัวข้อที่ 3** อธิบายแนวคิดและ โครงสร้างของวิธีการที่นำเสนอ
- **หัวข้อที่ 4** อธิบายถึงวิธีการทดลอง
- **หัวข้อที่ 5** เสนอผลการทดลองและวิเคราะห์
- **หัวข้อที่ 6** เป็นบทสรุปและงานวิจัยในอนาคต

การประเมินคุณภาพการให้บริการของท่าอากาศยานสุวรรณภูมิจากความคิดเห็น ในสื่อสังคมออนไลน์ด้วย Importance-performance Analysis The service quality assessment of Suvarnabhumi Airport based on social media opinion with Importance-performance Analysis

บทคัดย่อ

ท่าอากาศยานถือเป็นหนึ่งในอุตสาหกรรมบริการที่มีบทบาทสำคัญในการเพิ่มการเติบโตทางเศรษฐกิจของประเทศ ท่าอากาศยานสุวรรณภูมิเป็นหนึ่งในจุดหมายปลายทางยอดนิยมสำหรับนักเดินทางและนักธุรกิจ และเพื่อส่งมอบบริการที่ตอบโจทย์ความต้องการของนักเดินทางและนักธุรกิจเหล่านี้ คุณภาพการให้บริการจึงถูกนำมาใช้ในการวัดผลการดำเนินงานของท่าอากาศยาน เป้าหมายหลักของการศึกษานี้ เพื่อประเมินคุณภาพการให้บริการของท่าอากาศยานสุวรรณภูมิจากความคิดเห็นทางสื่อสังคมออนไลน์ ข้อมูลในการศึกษานี้ถูกจัดเก็บโดยการใช้ความคิดเห็นทางสื่อสังคมออนไลน์ โดยมีความคิดเห็นภาษาไทยและภาษาอังกฤษที่ผ่านการคัดกรอง 660 ความคิดเห็น การวิเคราะห์ความคิดเห็น (Sentiment Analysis) และการวิเคราะห์ความสำคัญ-ผลการดำเนินงาน (Importance-performance Analysis) ถูกนำมาใช้เป็นเทคนิคสำหรับการวิเคราะห์ข้อมูลเชิงคุณภาพ เพื่อให้สอดคล้องกับปัจจัยคุณภาพการให้บริการท่าอากาศยาน (Airport Service Quality) ผลการศึกษาพบว่า มี 4 จตุภาค (Quadrant) อันประกอบด้วย จตุภาคที่ 1 การให้ความสำคัญกับสิ่งที่อยู่ในจตุภาคนี้ จตุภาคที่ 2 รักษาการทำงานที่ดีต่อไป จตุภาคที่ 3 ลำดับความสำคัญที่ต่ำ และจตุภาคที่ 4 การดำเนินงานที่มากเกินความจำเป็น ทั้งนี้ มี 10 ประเด็นที่ผู้ใช้บริการให้ความสำคัญ คือ การขนส่งทางบก ระยะเวลารอคอยตรวจหนังสือเดินทาง ความง่ายในการค้นหาเส้นทาง การช่วยเหลือของเจ้าหน้าที่ ร้านอาหาร ร้านค้าต่างๆ ความสะอาดของห้องสุขา ความสะอาดของท่าอากาศยาน การตรวจหนังสือเดินทางขาเข้า และบรรยากาศภายในท่าอากาศยาน โดยพบว่า 8 ประเด็นที่ถูกจัดให้อยู่ในจตุภาคที่ผู้ใช้บริการให้ความสำคัญมาก และผลการดำเนินงานสูง ในขณะที่มี 2 ประเด็น (ระยะเวลาการรอในส่วนการตรวจหนังสือเดินทางและการตรวจหนังสือเดินทางขาเข้า) ถูกจัดอยู่ในจตุภาคที่ผู้ใช้บริการให้ความสำคัญมาก แต่ผลการดำเนินงานต่ำ การศึกษานี้ได้มีการอภิปรายถึงผลทางทฤษฎีและทางปฏิบัติ

บทนำ

ท่าอากาศยานจัดเป็นหนึ่งในธุรกิจบริการที่สำคัญในการขับเคลื่อนเศรษฐกิจของประเทศ ในการกระตุ้นการเจริญเติบโตของประเทศทั้งจากนักท่องเที่ยวทั่วโลก และจากการขนส่งสินค้าระหว่างประเทศ ดังนั้นเพื่อเพิ่มขีดความสามารถในการแข่งขันและตอบสนองความต้องการของผู้ใช้บริการได้อย่างมีประสิทธิภาพ การประเมินคุณภาพการให้บริการและความพึงพอใจของผู้ใช้บริการ จึงนับเป็นสิ่งสำคัญสำหรับการแข่งขันในธุรกิจที่เกี่ยวข้องกับการบิน ทั้งนี้การประเมินความพึงพอใจของผู้ใช้บริการที่ท่าอากาศยานส่วนใหญ่จะมาจากการสำรวจของหน่วยงานภายใน และหน่วยงานภายนอก โดยวิธีการประเมินคุณภาพการให้บริการของท่าอากาศยานที่ได้รับความนิยม

เชื่อถือและยอมรับอย่างกว้างขวาง คือ Airport Service Quality program (ASQ) ของ Airports Council International (ACI) ที่เริ่มต้นตั้งปี พ.ศ. 2549 และในปัจจุบันมีท่าอากาศยานทั่วโลกมากกว่า 320 แห่งที่ใช้วิธีการประเมินของ ASQ-ACI ผ่านการสัมภาษณ์และแบบสอบถาม (Airports Council International, 2021; Culotta & Cutler, 2016; Lee & Yu, 2018; Martin-Domingo et al., 2019) อย่างไรก็ตาม การเติบโตของสื่อสังคมออนไลน์ที่ให้ผู้ใช้งานแสดงความคิดเห็นและประเมินความพึงพอใจของผู้ใช้งาน และข้อมูลที่แสดงนั้นเป็นความคิดเห็นที่ผู้ใช้บริการเป็นผู้สร้างเนื้อหาด้วยตนเองตามประสบการณ์ที่ตนเองได้รับ จากข้อมูลดังกล่าวหากมีวิธีการจัดการและนำข้อมูลมาวิเคราะห์ จะทำให้สามารถทราบได้ถึงความต้องการของผู้ใช้บริการและก่อให้เกิดประโยชน์แก่ผู้ประกอบการ โดยช่องทางที่เป็นที่นิยมสำหรับนักเดินทางและนักท่องเที่ยวในการแสดงความคิดเห็นการใช้บริการ คือ Google reviews ที่ผู้ใช้บริการมักจะแสดงความคิดเห็นของแหล่งท่องเที่ยว ร้านอาหาร สายการบิน (Lee & Yu, 2018) ท่าอากาศยานสุวรรณภูมิ จัดเป็นหนึ่งในท่าอากาศยานระหว่างประเทศที่ใหญ่และได้มาตรฐานของโลกรวมถึงปริมาณผู้ใช้บริการ (ผู้โดยสาร) ที่มีอัตราเติบโตมาโดยตลอด การประเมินความพึงพอใจของผู้ใช้บริการ (ผู้โดยสาร) นับเป็นอีกหนึ่งเครื่องมือที่ช่วยในการพัฒนาคุณภาพการให้บริการของท่าอากาศยานสุวรรณภูมิให้ตรงความต้องการของผู้ใช้บริการมากที่สุด อย่างไรก็ตามแม้ว่าจะมีงานวิจัยหลากหลายที่ศึกษาคุณภาพการให้บริการของท่าอากาศยานสุวรรณภูมิ แต่งานวิจัยที่วิเคราะห์ในลักษณะที่เลือกใช้การเก็บข้อมูลและวิเคราะห์จากความคิดเห็นของผู้ใช้บริการในสื่อสังคมออนไลน์ยังพบน้อยมาก และยังขาดการวิจัยในเชิงลึก แม้ว่าข้อมูลที่ได้มานั้นจะสะท้อนมุมมองของผู้ใช้บริการโดยตรงโดยปราศจากการชี้นำ ด้วยผู้ใช้บริการสามารถแสดงความคิดเห็นได้อย่างอิสระ ดังนั้นงานวิจัยนี้จึงมีวัตถุประสงค์หลัก คือ เพื่อประเมินคุณภาพการบริการของท่าอากาศยานสุวรรณภูมิจากความคิดเห็นในสื่อสังคมออนไลน์ ด้วยวิธี Importance-performance Analysis เพื่อจัดลำดับความสำคัญในการพัฒนาการให้บริการของท่าอากาศยานสุวรรณภูมิ

บทคัดย่อ

- ท่าอากาศยานถือเป็นหนึ่งในอุตสาหกรรมบริการที่มีบทบาทสำคัญในการเพิ่มการเติบโตทางเศรษฐกิจของประเทศ
- ท่าอากาศยานสุวรรณภูมิเป็นหนึ่งในจุดหมายปลายทางยอดนิยมสำหรับนักเดินทางและนักธุรกิจ และเพื่อส่งมอบบริการที่ตอบโจทย์ความต้องการของนักเดินทางและนักธุรกิจเหล่านี้ คุณภาพการให้บริการจึงถูกนำมาใช้ในการวัดผลการดำเนินงานของท่าอากาศยาน
- เป้าหมายหลักของการศึกษานี้ เพื่อประเมินคุณภาพการให้บริการของท่าอากาศยานสุวรรณภูมิจากความคิดเห็นทางสื่อสังคมออนไลน์ ข้อมูลในการศึกษานี้ถูกจัดเก็บโดยใช้ความคิดเห็นทางสื่อสังคมออนไลน์ โดยมีความคิดเห็นภาษาไทยและภาษาอังกฤษที่ผ่านการคัดกรอง 660 ความคิดเห็น

- การวิเคราะห์ความคิดเห็น (Sentiment Analysis) และการวิเคราะห์ความสำคัญ-ผลการดำเนินงาน (Importance-performance Analysis) ถูกนำมาใช้เป็นเทคนิคสำหรับการวิเคราะห์ข้อมูลเชิงคุณภาพ เพื่อให้สอดคล้องกับปัจจัยคุณภาพการให้บริการท่าอากาศยาน (Airport Service Quality)
- ผลการศึกษาพบว่า มี 4 จตุภาค (Quadrant) อันประกอบด้วย
- จตุภาคที่ 1 การให้ความสำคัญกับสิ่งที่อยู่ในจตุภาคนี้
- จตุภาคที่ 2 รักษาการทำงานที่ดีต่อไป
- จตุภาคที่ 3 ลำดับความสำคัญที่ต่ำ
- และจตุภาคที่ 4 การดำเนินงานที่มากเกินไปจนความจำเป็น
- ทั้งนี้ มี 10 ประเด็นที่ผู้ใช้บริการให้ความสำคัญ คือ การขนส่งทางบก ระยะเวลารอคอยตรวจหนังสือเดินทาง ความง่ายในการค้นหาเส้นทาง การช่วยเหลือของเจ้าหน้าที่ ร้านอาหาร ร้านค้าต่างๆ ความสะอาดของห้องสุขา ความสะอาดของท่าอากาศยาน การตรวจหนังสือเดินทางขาเข้า และบรรยากาศภายในท่าอากาศยาน
- โดยพบว่า 8 ประเด็นที่ถูกจัดให้อยู่ในจตุภาคที่ผู้ใช้บริการให้ความสำคัญมาก และผลการดำเนินงานสูง ในขณะที่มี 2 ประเด็น (ระยะเวลาการรอในส่วนการตรวจหนังสือเดินทางและการตรวจหนังสือเดินทางขาเข้า)
- ถูกจัดอยู่ในจตุภาคที่ผู้ใช้บริการให้ความสำคัญมาก แต่ผลการดำเนินงานต่ำ การศึกษานี้ได้มีการอภิปรายถึงผลทางทฤษฎีและทางปฏิบัติ

บทนำ

- ท่าอากาศยานจัดเป็นหนึ่งในธุรกิจบริการที่สำคัญในการขับเคลื่อนเศรษฐกิจของประเทศในการกระตุ้นการเจริญเติบโตของประเทศทั้งจากนักท่องเที่ยวทั่วโลกและจากการขนส่งสินค้าระหว่างประเทศ
- ดังนั้นเพื่อเพิ่มขีดความสามารถในการแข่งขันและตอบสนองความต้องการของผู้ใช้บริการได้อย่างมีประสิทธิภาพ
- การประเมินคุณภาพการให้บริการและความพึงพอใจของผู้ใช้บริการ จึงนับเป็นสิ่งสำคัญสำหรับการแข่งขันในธุรกิจที่เกี่ยวข้องกับการบิน
- ทั้งนี้การประเมินความพึงพอใจของผู้ใช้บริการที่ท่าอากาศยานส่วนใหญ่จะมาจากการสำรวจของหน่วยงานภายใน และหน่วยงานภายนอก
- โดยวิธีการประเมินคุณภาพการให้บริการของท่าอากาศยานที่ได้รับความนิยมและยอมรับอย่างกว้างขวาง คือ Airport Service Quality program (ASQ) ของ Airports Council International (ACI)

- **ที่เริ่มต้น**ตั้งแต่ปี พ.ศ. 2549 และในปัจจุบันมีท่าอากาศยานทั่วโลกมากกว่า 320 แห่งที่ใช้วิธีการประเมินของ ASQ-ACI ผ่านการสัมภาษณ์และแบบสอบถาม (Airports Council International, 2021; Culotta & Cutler, 2016; Lee & Yu, 2018; Martin-Domingo et al., 2019)
- **อย่างไรก็ตาม** การเติบโตของของสื่อสังคมออนไลน์ที่ให้ผู้ใช้งานแสดงความคิดเห็นและประเมินความพึงพอใจของผู้ใช้งาน และข้อมูลที่แสดงนั้นเป็นความคิดเห็นที่ผู้ใช้บริการเป็นผู้สร้างเนื้อหาด้วยตนเองตามประสบการณ์ที่ตนเองได้รับ
- **จากข้อมูลดังกล่าว**หากมีวิธีการในการจัดการและนำข้อมูลมาวิเคราะห์ จะทำให้สามารถทราบได้ถึงความต้องการของผู้ใช้บริการและก่อให้เกิดประโยชน์แก่ผู้ประกอบการ
- **โดยช่องทางที่เป็นที่นิยม**สำหรับนักเดินทางและนักท่องเที่ยวในการแสดงความคิดเห็นการใช้บริการ คือ Google reviews ที่ผู้ใช้บริการมักจะแสดงความคิดเห็นของแหล่งท่องเที่ยว ร้านอาหาร สายการบิน (Lee & Yu, 2018)
- **ท่าอากาศยานสุวรรณภูมิ** จัดเป็นหนึ่งในท่าอากาศยานระหว่างประเทศที่ใหญ่และได้มาตรฐานของโลก รวมถึงปริมาณผู้ใช้บริการ (ผู้โดยสาร) ที่มีอัตราเติบโตมาโดยตลอด
- **การประเมินความพึงพอใจ**ของผู้ใช้บริการ (ผู้โดยสาร) นับเป็นอีกหนึ่งเครื่องมือที่ช่วยในการพัฒนาคุณภาพการให้บริการของท่าอากาศยานสุวรรณภูมิให้ตรงความต้องการของผู้ใช้บริการมากที่สุด
- **อย่างไรก็ตาม**แม้ว่าจะมีงานวิจัยหลากหลายที่ศึกษาคุณภาพการให้บริการของท่าอากาศยานสุวรรณภูมิ
- **แต่งานวิจัยที่วิเคราะห์**ในลักษณะที่เลือกใช้ในการเก็บข้อมูลและวิเคราะห์จากความคิดเห็นของผู้ใช้บริการในสื่อสังคมออนไลน์ยังพบน้อยมาก และยังขาดการวิจัยในเชิงลึก
- **แม้ว่า** ข้อมูลที่ได้มานั้นจะสะท้อนมุมมองของผู้ใช้บริการโดยตรงโดยปราศจากการชี้นำ ด้วยผู้ใช้บริการสามารถแสดงความคิดเห็นได้อย่างอิสระ
- **ดังนั้นงานวิจัยนี้จึงมีวัตถุประสงค์หลัก** คือ เพื่อประเมินคุณภาพการบริการของท่าอากาศยานสุวรรณภูมิ จากความคิดเห็นในสื่อสังคมออนไลน์ ด้วยวิธี Importance-performance Analysis
- **เพื่อจัดลำดับความสำคัญ**ในการพัฒนาการให้บริการของท่าอากาศยานสุวรรณภูมิ

การพัฒนาแอปพลิเคชันระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาผ่านสมาร์ทโฟน

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอการพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาผ่านสมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์ด้วยเทคโนโลยีการเรียนรู้เชิงลึกสำหรับการจำแนกภาพธนบัตรและเทคโนโลยี Text to speech สำหรับการแปลงข้อความเสียงผ่านสมาร์ทโฟน การดำเนินงานเริ่มจากการรวบรวมกลุ่มตัวอย่างธนบัตรไทยจำนวน 2,700 ภาพ โดยการนำภาพมาวิเคราะห์และสร้างโมเดลจำแนกภาพด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันที่มีการใช้อัลกอริธึม MobileNet ผ่านไลบรารี Tensor low จากนั้นนำโมเดลไปพัฒนาเป็นแอปพลิเคชันบนสมาร์ทโฟน แอปพลิเคชันนี้พัฒนาด้วยโปรแกรม Android Studio ภาษ JAVA และไลบรารี Text to Spcech ผลการวัดประสิทธิภาพโมเดลได้ค่าความถูกต้องเท่ากับ 95.00% ผลการทดสอบแอปพลิเคชันพบว่าสามารถจำแนกธนบัตรได้ถูกต้อง 84.00% ส่วนผลประเมินความพึงพอใจจากผู้ใช้งานพบว่ามีความพึงพอใจเฉลี่ยเท่ากับ 4.33 ส่วนเบี่ยงเบนมาตรฐาน 0.56 อยู่ในระดับดี สามารถสรุปได้ว่าแอปพลิเคชันนี้มีประสิทธิภาพสามารถช่วยให้ผู้พิการทางสายตาทราบถึงชนิดธนบัตรได้ง่ายขึ้น เนื่องจากสามารถระบุชนิดธนบัตรด้วยเสียงและอำนวยความสะดวกแก่ผู้พิการทางสายตาในการใช้งานได้ทุกที่ทุกเวลา ทั้งยังช่วยส่งเสริมการดำเนินชีวิตสำหรับผู้พิการทางสายตาให้สามารถดำรงชีวิตขั้นพื้นฐานได้สะดวกเพิ่มมากขึ้น

บทนำ

ข้อมูลของผู้พิการทางสายตาจากกรมส่งเสริมและ พัฒนาคุณภาพชีวิตผู้พิการที่ได้รับการออกบัตรประจำตัวคนพิการ ในประเทศไทยปัจจุบัน พ.ศ.2562 มีจำนวน 186,081 คน [1] เป็นลำดับที่ 3 ของคนพิการทั้งหมด ซึ่งจากข้อมูลสถานการณ์พิการเห็นได้ชัดว่ามีผู้พิการเพิ่มขึ้นอย่างต่อเนื่องทุกปี ปัญหาที่ผู้พิการทางสายตาพบส่วนใหญ่เป็นปัญหาด้านการใช้ชีวิตประจำวัน โดยผู้พิการทางสายตาไม่สามารถใช้ชีวิตประจำวันได้เหมือนคนสายตาปกติจึงทำให้เกิดข้อจำกัดต่างๆ มากมาย เช่น ปัญหาเรื่องการเลือกซื้อสินค้าในชีวิตประจำวัน การเลือกซื้ออาหารเสื้อผ้าและผลิตภัณฑ์ต่างๆซึ่งจำเป็นที่จะต้องใช้จ่ายด้วยธนบัตร โดยที่ธนบัตร ไทยในปัจจุบันยังไม่ออกแบบมาให้รองรับกับผู้พิการทางสายตา ทำให้ผู้พิการจำเป็นต้องอาศัยความคุ้นเคยหรือสอบถามจากผู้คนรอบข้างให้ช่วยบอกราคาธนบัตร ส่งผลให้ผู้พิการทางสายตาประสบปัญหาการถูกหลอกลวงจากมิจฉาชีพได้ง่ายจากที่กล่าวมาข้างต้นทำให้เห็นถึงความสำคัญของการใช้ธนบัตรในชีวิตประจำวันของผู้พิการทางสายตา ดังนั้นคณะผู้วิจัยจึงมีแนวคิดที่จะพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาบนสมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์ โดยมีวัตถุประสงค์เพื่อช่วยเหลือผู้พิการทางสายตาให้สามารถช่วยเหลือตนเองได้ โดยใช้หลักการประมวลผลภาพด้วยเทคโนโลยีการเรียนรู้เชิงลึกระบุค่าเงินในธนบัตร ผู้พิการทางสายตาสามารถใช้อุปกรณ์สมาร์ทโฟนถ่ายภาพธนบัตรเพื่อระบุราคาธนบัตรด้วยเสียงได้ ซึ่งจะทำให้ผู้พิการทางสายตาสามารถดำรงชีวิตขั้นพื้นฐานได้สะดวกมากขึ้น

บทคัดย่อ

- งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอการพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาผ่านสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ด้วยเทคโนโลยีการเรียนรู้เชิงลึกสำหรับการจำแนกภาพธนบัตรและเทคโนโลยี Text to speech สำหรับการแปลงข้อความเป็นเสียงผ่านสมาร์ตโฟน
- การดำเนินงานเริ่มจากการรวบรวมกลุ่มตัวอย่างธนบัตรไทยจำนวน 2,700 ภาพ
- โดยการนำภาพมาวิเคราะห์และสร้างโมเดลจำแนกภาพด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันที่มีการใช้อัลกอริทึม MobileNet ผ่านไลบรารี Tensor low
- จากนั้นนำโมเดลไปพัฒนาเป็นแอปพลิเคชันบนสมาร์ตโฟน แอปพลิเคชันนี้พัฒนาด้วยโปรแกรม Android Studio ภาษา1 JAVA และไลบรารี Text to Spcech ผลการวัดประสิทธิภาพโมเดลได้ค่าความถูกต้องเท่ากับ 95.00%
- ผลการทดสอบแอปพลิเคชันพบว่าสามารถจำแนกธนบัตรได้ถูกต้อง 84.00% ส่วนผลประเมินความพึงพอใจจากผู้ใช้งานพบว่ามีความพึงพอใจเฉลี่ยเท่ากับ 4.33 ส่วนเบี่ยงเบนมาตรฐาน 0.56 อยู่ในระดับดี
- สามารถสรุปได้ว่าแอปพลิเคชันนี้มีประสิทธิภาพสามารถช่วยให้ผู้พิการทางสายตาทราบถึงชนิดธนบัตรได้ง่ายขึ้น
- เนื่องจากสามารถระบุชนิดธนบัตรด้วยเสียงและอำนวยความสะดวกแก่ผู้พิการทางสายตาในการใช้งานได้ทุกที่ทุกเวลา
- ทั้งยังช่วยส่งเสริมการดำเนินชีวิตสำหรับผู้พิการทางสายตาให้สามารถดำรงชีวิตขั้นพื้นฐานได้สะดวกเพิ่มมากขึ้น

บทนำ

- ข้อมูลของผู้พิการทางสายตาจากกรมส่งเสริมและ พัฒนาคุณภาพชีวิตผู้พิการที่ได้รับการออกบัตรประจำตัวคนพิการ ในประเทศไทยปัจจุบัน พ.ศ.2562 มีจำนวน 186,081 คน [1] เป็นลำดับที่ 3 ของคนพิการทั้งหมด
- ซึ่งจากข้อมูลสถานการณ์การเห็นได้ชัดว่ามีผู้พิการเพิ่มขึ้นอย่างต่อเนื่องทุกปี
- ปัญหาที่ผู้พิการทางสายตาพบส่วนใหญ่เป็นปัญหาด้านการใช้ชีวิตประจำวัน โดยผู้พิการทางสายตาไม่สามารถใช้ชีวิตประจำวันได้เหมือนคนสายตาปกติจึงทำให้เกิดข้อจำกัดต่างๆ มากมาย เช่น ปัญหาเรื่องการเลือกซื้อสินค้าในชีวิตประจำวัน
- การเลือกซื้ออาหารเสื้อผ้าและผลิตภัณฑ์ต่างๆซึ่งจำเป็นที่จะต้องใช้จ่ายด้วยธนบัตร

- โดยที่ธนบัตร ไทยในปัจจุบันยังไม่ออกแบบมาให้รองรับกับผู้พิการทางสายตา
- ทำให้ผู้พิการจำเป็นต้องอาศัยความคุ้นเคยหรือสอบถามจากผู้คนรอบข้างให้ช่วยบอกราคาธนบัตร
- ส่งผลให้ผู้พิการทางสายตาประสบปัญหาการถูกหลอกลวงจากมิจฉาชีพได้ง่ายจากที่กล่าวมาข้างต้นทำให้เห็นถึงความสำคัญของการใช้ธนบัตรในชีวิตประจำวันของผู้พิการทางสายตา
- ดังนั้นคณะผู้วิจัยจึงมีแนวคิดที่จะพัฒนาแอปพลิเคชันช่วยระบุธนบัตรไทยด้วยเสียงสำหรับผู้พิการทางสายตาบนสมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์
- โดยมีวัตถุประสงค์เพื่อช่วยเหลือผู้พิการทางสายตาให้สามารถช่วยเหลือตนเองได้
- โดยใช้หลักการประมวลผลภาพด้วยเทคโนโลยีการเรียนรู้เชิงลึกระบุค่าเงินในธนบัตร
- ผู้พิการทางสายตาสามารถใช้อุปกรณ์สมาร์ทโฟนถ่ายภาพธนบัตรเพื่อระบุราคาธนบัตรด้วยเสียงได้
- ซึ่งจะช่วยให้ผู้พิการทางสายตาสามารถดำรงชีวิตขั้นพื้นฐานได้สะดวกมากขึ้น

การพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวภาษาไทย โดยใช้วิธีการประมวลผลภาษาธรรมชาติ

บทคัดย่อ

ปัจจัยที่กระทบต่อราคาของหุ้นในตลาดหลักทรัพย์แห่งประเทศไทยนั้นมีอยู่หลายปัจจัย ข่าวสารต่าง ๆ ก็เป็นปัจจัยหนึ่งที่มีผลกระทบต่อราคาของหุ้น ผู้วิจัยจึงเกิดแนวคิดที่จะพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวโดยใช้วิธีการประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP) เพื่อให้นักลงทุนสามารถคาดคะเนทิศทางของราคาหุ้นก่อนที่ตลาดหลักทรัพย์แห่งประเทศไทยปิดโดยศึกษาข้อความข่าวจากแหล่งข่าวต่างๆ และ ใช้การตัดคำ(tokenizer) จาก Librry pythainlp ในโปรแกรมภาษาไพธอน ve.:3.7.1 จากนั้นสร้างแบบจำลองโดยใช้ตัวแบบการจำแนก(classification model) เพื่อหาแบบจำลอง (model) และวิธีการตัดคำ (Tokenizer) ที่มีค่าความถูกต้องแม่นยำ (accuracy)สูงสุด เพื่อใช้พยากรณ์ทิศทางของราคาหุ้นรายวัน ซึ่งในงานวิจัยนี้ได้พยากรณ์ทิศทางของราคาหุ้นทั้งหมด 3 วัน คือวันที่5,6 และ 7 กุมภาพันธ์ 2563 โดยสุ่มหุ้นอย่างละ 1 ตัว ด้วยการสุ่มตัวอย่างอย่างง่าย (SRS) จากหุ้น 5 กลุ่มคือ กลุ่ม ICTกลุ่ม ENERG กลุ่ม HEALTH กลุ่ม COMM และ กลุ่ม BANK ผลวิจัยพบว่า กลุ่ม ICT สุ่มได้หุ้นของบริษัท อินทัช โฮลดิ้งส์ จำกัด (INTUCH) ตัวแบบ Gadiant Boosting Classifier เป็นตัวแบบที่มีความเหมาะสมมากที่สุดและเปรียบเทียบสถานะค่าพยากรณ์กับค่าจริงได้ความถูกต้องร้อยละ 10 ในส่วนของกลุ่ม ENERG สุ่มได้หุ้นของบริษัท ไทยออยล์ จำกัด (TOP) และกลุ่ม HEALTH สุ่มได้หุ้นของบริษัท โรงพยาบาลบำรุงราษฎร์ จำกัด (BH) นั้น ไม่สามารถสรุปตัวแบบที่มีความเหมาะสมมากที่สุดได้และเปรียบเทียบสถานะค่าพยากรณ์ กับค่าจริงได้ความถูกต้องร้อยละ 66.67 ในกลุ่ม COMM สุ่มได้หุ้นของบริษัท ซีพี ออลล์ จำกัด (CPALL) และกลุ่ม BANK คือธนาคารกสิกรไทย จำกัด (มหาชน) (KBANK) ตัวแบบK-Ncighbors Classifier เป็นตัวแบบที่มีความเหมาะสมมากที่สุดและเปรียบเทียบสถานะค่าพยากรณ์กับค่าจริงจะได้ความถูกต้องร้อยละ 66.67

บทนำ

ตลาดหลักทรัพย์เป็นแหล่งซื้อขายแลกเปลี่ยนหลักทรัพย์ระยะยาว ทำหน้าที่เป็นตลาดทุน เพื่อให้บริษัทมหาชน จำกัด ซึ่งถือว่าเป็นตลาดรอง (Secondary Market) สามารถระดมเงินทุนเพิ่มเติมจากสาธารณะได้ โดยหลักทรัพย์ระยะยาวจะประกอบไปด้วยตราสารหนี้ และตราสารทุนซึ่งประกอบไปด้วย หุ้นสามัญ หุ้นบุริมสิทธิ ใบสำคัญแสดงสิทธิแบบต่าง ๆ ใบสำคัญแสดงสิทธิอนุพันธ์ หุ้นกู้และ หน่วยลงทุน เป็นต้น ในปัจจุบันการซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยเป็นที่นิยมและรู้จักกันอย่างกว้างขวาง เนื่องจากเป็นแหล่งระดมเงินทุนที่มีผลตอบแทนสูงกว่าการลงทุนประเภทอื่น เช่น การซื้อที่ดิน การซื้อทองคำแท่ง หรือ การฝากเงินกับธนาคารพาณิชย์ เป็นต้น จึงทำให้มีผู้สนใจเข้ามาลงทุนเพื่อสร้างความมั่งคั่งให้กับตนเองเป็นจำนวนมากแต่การลงทุนในตลาดหลักทรัพย์แห่งประเทศไทยมีอ่อนไหวสูง ไม่ว่าจะเป็นเหตุการณ์ใดๆ เกิดขึ้นก็จะส่งผลกระทบต่อสถานะตลาดได้อย่าง

รวดเร็ว เช่น ปัญหาเศรษฐกิจ ปัญหาการแพร่ระบาดของโรค ปัญหาการเมือง ความมั่นคง และการทหาร เป็นต้น ซึ่งปัญหาต่างๆ เหล่านี้จะส่งผลให้ราคาหลักทรัพย์มีความผันผวนตลอดเวลา ทำให้ผลตอบแทนไม่เป็นไปตามที่นักลงทุนคาดหวัง ดังนั้นการลงทุนในหลักทรัพย์ควรมีหลักเกณฑ์ในการพิจารณาเพื่อเป็นแนวทางในการตัดสินใจของนักลงทุนปัจจัยที่กระทบต่อราคาของหุ้นในตลาดหลักทรัพย์แห่งประเทศไทยที่นักลงทุนนิยมพิจารณานั้นมีอยู่หลายปัจจัย เช่น ผลการดำเนินงานของบริษัท จำนวนการซื้อขายหุ้นในแต่ละวัน รวมทั้งข่าวสารต่างๆ ที่มีผลกระทบต่อจิตวิทยาของตลาดหุ้นและราคาของหุ้น อาทิ ข่าวการลงทุนของบริษัท ข่าวการระบาดของโรค ข่าวการชุมนุมทางการเมือง เป็นต้น ด้วยความสำคัญของข่าวสารต่างๆ นี้ ผู้วิจัยจึงเกิดแนวคิดที่จะพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวภาษาไทย โดยใช้วิธีการประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP) เพื่อให้ นักลงทุนสามารถคาดคะเนทิศทางของราคาหุ้นก่อนที่ตลาดหลักทรัพย์แห่งประเทศไทยจะปิดได้ โดยจะทำการศึกษาข้อความข่าวของหุ้นจากตลาดหลักทรัพย์แห่งประเทศไทย (SET) จำนวน 5 ตัว จากหุ้น 5 กลุ่ม ด้วยโปรแกรมภาษาไพธอน ver.3.7. บนโปรแกรม Visual Studio Code และใช้โปรแกรม Power BI ช่วยในการแสดงผล

บทคัดย่อ

- ปัจจัยที่กระทบต่อราคาของหุ้นในตลาดหลักทรัพย์แห่งประเทศไทยนั้นมีอยู่หลายปัจจัย ข่าวสารต่าง ๆ ก็เป็น
- ปัจจัยหนึ่งที่มีผลกระทบต่อราคาของหุ้น
- ผู้วิจัยจึงเกิดแนวคิดที่จะพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวโดยใช้วิธีการประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP)
- เพื่อให้ นักลงทุนสามารถคาดคะเนทิศทางของราคาหุ้นก่อนที่ตลาดหลักทรัพย์แห่งประเทศไทยจะปิดโดยศึกษาข้อความข่าวจากแหล่งข่าวต่างๆ และ ใช้การตัดคำ (tokenizer) จาก Library pythainlp ในโปรแกรมภาษาไพธอน ver.:3.7.1
- จากนั้นสร้างแบบจำลองโดยใช้ตัวแบบการจำแนก (classification model) เพื่อหาแบบจำลอง (model) และวิธีการตัดคำ (Tokenizer) ที่มีค่าความถูกต้องแม่นยำ (accuracy) สูงสุด
- เพื่อใช้พยากรณ์ทิศทางของราคาหุ้นรายวัน ซึ่งในงานวิจัยนี้ได้พยากรณ์ทิศทางของราคาหุ้นทั้งหมด 3 วัน คือวันที่ 5, 6 และ 7 กุมภาพันธ์ 2563
- โดยสุ่มหุ้นอย่างละ 1 ตัว ด้วยการสุ่มตัวอย่างอย่างง่าย (SRS) จากหุ้น 5 กลุ่มคือ กลุ่ม ICT กลุ่ม ENER กลุ่ม HEALTH กลุ่ม COMM และ กลุ่ม BANK

- ผลวิจัยพบว่า กลุ่ม ICT สุ่มได้หุ้นของบริษัท อินทัช โฮลดิ้งส์ จำกัด (INTUCH) ตัวแบบ Gadiant Boosting Classifier เป็นตัวแบบที่มีความเหมาะสมมากที่สุดและเปรียบเทียบสถานะค่าพยากรณ์กับค่าจริงได้ความถูกต้องร้อยละ 10
- ในส่วนของกลุ่ม ENGER สุ่มได้หุ้นของบริษัท ไทยออยล์ จำกัด (TOP)
- และกลุ่ม HEALTH สุ่มได้หุ้นของบริษัท โรงพยาบาลบำรุงราษฎร์ จำกัด
- (BH) นั้น ไม่สามารถสรุปตัวแบบที่มีความเหมาะสมมากที่สุดได้และเปรียบเทียบสถานะค่าพยากรณ์กับค่าจริงจะได้ความถูกต้องร้อยละ 66.67
- ในกลุ่ม COMM สุ่มได้หุ้นของบริษัท ซีพี ออลล์ จำกัด (CPALL)
- และกลุ่ม BANK คือธนาคารกสิกรไทย จำกัด (มหาชน) (KBANK)
- ตัวแบบ K-Ncighbors Classifier เป็นตัวแบบที่มีความเหมาะสมมากที่สุดและเปรียบเทียบสถานะค่าพยากรณ์กับค่าจริงจะได้ความถูกต้องร้อยละ 66.67

บทนำ

- ตลาดหลักทรัพย์เป็นแหล่งซื้อขายแลกเปลี่ยนหลักทรัพย์ระยะยาว ทำหน้าที่เป็นตลาดทุน
- เพื่อให้บริษัทมหาชนจำกัด ซึ่งถือว่าเป็นตลาดรอง (Secondary Market) สามารถระดมเงินทุนเพิ่มเติมจากสาธารณะได้
- โดยหลักทรัพย์ระยะยาวจะประกอบไปด้วยตราสารหนี้ และตราสารทุนซึ่งประกอบไปด้วย หุ้นสามัญ หุ้นบุริมสิทธิ ใบสำคัญแสดงสิทธิแบบต่าง ๆ ใบสำคัญแสดงสิทธิอนุพันธ์ หุ้นกู้และ หน่วยลงทุน เป็นต้น
- ในปัจจุบันการซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยเป็นที่นิยมและรู้จักกันอย่างกว้างขวาง
- เนื่องจากเป็นแหล่งระดมเงินทุนที่มีผลตอบแทนสูงกว่าการลงทุนประเภทอื่น เช่น การซื้อที่ดิน การซื้อทองคำแท่ง หรือ การฝากเงินกับธนาคารพาณิชย์ เป็นต้น
- จึงทำให้มีผู้สนใจเข้ามาลงทุนเพื่อสร้างความมั่งคั่งให้กับตนเองเป็นจำนวนมากแต่การลงทุนในตลาดหลักทรัพย์แห่งประเทศไทยมีอ่อนไหวสูง ไม่ว่าจะเป็นเหตุการณ์ใดๆ เกิดขึ้นก็จะส่งผลกระทบต่อสถานะตลาดได้อย่างรวดเร็ว เช่น ปัญหาเศรษฐกิจ ปัญหาการแพร่ระบาดของโรค ปัญหาการเมือง ความมั่นคง และการทหาร เป็นต้น
- ซึ่งปัญหาต่างๆ เหล่านี้จะส่งผลให้ราคาหลักทรัพย์มีความผันผวนตลอดเวลา ทำให้ผลตอบแทนไม่แน่นอนตามที่นักลงทุนคาดหวัง

- ดังนั้นการลงทุนในหลักทรัพย์ควรมีหลักเกณฑ์ในการพิจารณาเพื่อเป็นแนวทางในการตัดสินใจของนักลงทุนปัจจัยที่กระทบต่อราคาของหุ้นในตลาดหลักทรัพย์แห่งประเทศไทยที่นักลงทุนนิยมพิจารณานั้นมีอยู่หลายปัจจัย เช่น ผลการดำเนินงานของบริษัท จำนวนการซื้อขายหุ้นในแต่ละวัน
- รวมทั้งข่าวสารต่างๆ ที่มีผลกระทบต่อจิตวิทยาของตลาดหุ้นและราคาของหุ้น อาทิ ข่าวการลงทุนของบริษัท ข่าวการระบาดของโรค ข่าวการชุมนุมทางการเมือง เป็นต้น
- ด้วยความสำคัญของข่าวสารต่างๆ นี้ ผู้วิจัยจึงเกิดแนวคิดที่จะพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวภาษาไทย
- โดยใช้วิธีการประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP) เพื่อให้ให้นักลงทุนสามารถคาดคะเนทิศทางของราคาหุ้นก่อนที่ตลาดหลักทรัพย์แห่งประเทศไทยจะปิดได้
- โดยจะทำการศึกษาข้อความข่าวของหุ้นจากตลาดหลักทรัพย์แห่งประเทศไทย (SET) จำนวน 5 ตัว จากหุ้น 5 กลุ่ม ด้วยโปรแกรมภาษาไพธอน ver.3.7. บนโปรแกรม Visual Studio Code และใช้โปรแกรม Power BI ช่วยในการแสดงผล

ระบบแบ่งปันสูตรการทำอาหารและค้นหาสูตรการทำอาหารจากภาพวัตถุดิบ ด้วยเทคนิคการเรียนรู้เชิงลึก

บทคัดย่อ

ปัจจุบันการรับประทานอาหารเพื่อสุขภาพกำลังได้รับความนิยมเป็นอย่างมาก ผู้คนจึงเริ่มทำอาหารรับประทานกันด้วยตัวเองจากวัตถุดิบในการปรุงอาหารที่มีอยู่ ซึ่งบางครั้งไม่ทราบว่าสามารถนำไปทำอาหารอะไรได้บ้าง จึงทำได้แต่รายการอาหารเดิม ๆ ทำให้เกิดความจำเจในการรับประทานอาหาร และไม่สนุกกับการทำอาหาร บทความวิจัยนี้นำเสนอการออกแบบและพัฒนาระบบแบ่งปันสูตรการทำอาหาร และค้นหาสูตรการทำอาหารจากภาพวัตถุดิบด้วยเทคนิคการเรียนรู้เชิงลึก โดยใช้งานได้นับอุปกรณ์พกพา ซึ่งผู้ใช้งานสามารถแบ่งปันสูตรการทำอาหารและสามารถใช้อุปกรณ์พกพาถ่ายรูปวัตถุดิบที่ผู้ใช้อยู่แล้วในครัว เช่น กระเทียม เนื้อหมู ผัก ฯลฯ เพื่อส่งภาพเข้ามาในระบบให้ค้นหาสูตรอาหารจากวัตถุดิบที่มีอยู่ ทำให้เพิ่มความสะดวกสบายในการค้นหาสูตรอาหารให้แก่ผู้ใช้ ส่วนประกอบหลักของระบบประกอบด้วย (1) แอปพลิเคชันบนอุปกรณ์พกพาสำหรับผู้ใช้งานทั่วไปที่พัฒนาด้วย React Native ซึ่งผู้ใช้งานสามารถเพิ่มสูตรการทำอาหาร และค้นหาสูตรการทำอาหารได้โดยใส่ชื่อวัตถุดิบ รวมถึงการใช้อุปกรณ์พกพาถ่ายภาพวัตถุดิบเพื่อค้นหาสูตรการทำอาหาร, (2) เว็บแอปพลิเคชันพัฒนาบน MERN stack สำหรับผู้ดูแลระบบ เพื่อใช้ในการเพิ่มคำหลักในการค้นหาให้กับวัตถุดิบ รวมถึงทดสอบแบบจำลองที่ได้สร้างขึ้น และ (3) โครงข่ายประสาทเทียมการเรียนรู้เชิงลึกที่มีการใช้อัลกอริทึม YOLO ผ่าน ไลบรารี Darknet สำหรับการสร้างแบบจำลองในการรู้จำภาพแบบจำลองได้ฝึกสอนให้สามารถรู้จำวัตถุดิบได้จำนวน 20 ประเภท โดยสอนด้วยภาพวัตถุดิบประเภทละ 100 ภาพ หลังจากฝึกสอนแบบจำลองจำนวน 36,000 รอบ แบบจำลองมีค่าสูญเสียเฉลี่ยอยู่ที่ 0.0408 และมีค่า Precision, Recall และ F1-score อยู่ที่ 0.96, 0.98 และ 0.97 ตามลำดับ

บทนำ

Deep Learning หรือการเรียนรู้เชิงลึก ถือว่าเป็นหนึ่งในระบบการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งมีโครงสร้างและการประมวลผลคล้ายกับสมองของมนุษย์ที่เรียกว่า โครงข่ายประสาทเทียม (Artificial Neural Network) โดยอันที่จริงแล้ว Deep Learning คือโครงข่ายประสาทเทียมขนาดใหญ่ ซึ่งหมายถึงเป็นโครงข่ายประสาทเทียมที่มีชั้นในการประมวลผลหลายชั้นช่วงไม่กี่ปีที่ผ่านมาการเรียนรู้เชิงลึกได้รับความนิยมและมีงานวิจัยที่เกี่ยวข้องมากขึ้น เช่น ใน [1] ได้พัฒนาระบบการแนะนำ (Recommendation system) ด้วยการใช้เทคนิคการเรียนรู้เชิงลึก หรือแม้กระทั่งใน [2] ได้มีการใช้เทคนิคการเรียนรู้เชิงลึกในการพัฒนาระบบตรวจจับผู้บุกรุก (Intrusion Detection System) แต่ด้านการใช้งานที่ทำให้เทคนิคการเรียนรู้เชิงลึกได้รับความนิยมมากมาจากการด้านรู้จำภาพ โดยเป็นการใช้โครงข่ายประสาทเทียมประเภท Convolutional Neural Network ดังอธิบายใน [3] และมีบทความวิจัยที่ศึกษาและประยุกต์ใช้การรู้จำภาพอย่างมากมาย เช่น ใน [4] เป็น โครงข่ายประสาทเทียมสำหรับตรวจจับภาพอาหารและ [5], [6] และ [7] เป็นโครงข่ายประสาทเทียมสำหรับตรวจจับอารมณ์ในภาพ

ในปัจจุบันแนวโน้มการรับประทานอาหารเพื่อสุขภาพกำลังได้รับความนิยมเป็นอย่างมาก เนื่องจากโรคประจำตัวต่างๆ ซึ่งส่วนใหญ่แล้วมักเกิดจากการรับประทานอาหาร ผู้คนส่วนใหญ่จึงเริ่มทำอาหารรับประทานกันด้วยตัวเอง แต่สำหรับผู้เริ่มต้นทำอาหารอาจมีความลำบากและยุ่งยากในการค้นหาสูตรการทำอาหารจากหนังสือทำอาหารหรือจากอินเทอร์เน็ตอีกทั้งบางครั้งมีวัตถุดิบในการปรุงอาหารอยู่ที่บ้าน แต่คิดไม่ออกว่าสามารถนำไปทำอาหารอะไรได้บ้าง จึงทำได้แต่รายการอาหารเดิม ๆ จึงเกิดความจำเจในการรับประทานอาหาร และไม่สนุกกับการทำอาหาร ดังนั้นบทความวิจัยนี้จึงต้องการนำเสนอการออกแบบและพัฒนาระบบแบ่งปันสูตรการทำอาหารและค้นหาสูตรการทำอาหารจากภาพวัตถุดิบด้วยเทคนิคการเรียนรู้เชิงลึก โดยใช้งานได้นับอุปกรณ์พกพา ซึ่งผู้ใช้งานสามารถแบ่งปันสูตรการทำอาหาร และสามารถใช้อุปกรณ์พกพารูปวัตถุดิบที่ผู้ใช้อยู่แล้วในครัว เช่น กระเทียม เนื้อหมู ผัก ฯลฯ เพื่อส่งภาพเข้ามาในระบบให้ค้นหาสูตรอาหารจากวัตถุดิบที่มีอยู่ เพื่อเพิ่มความสะดวกสบายในการค้นหาสูตรอาหารให้แก่ผู้ใช้

บทคัดย่อ

- ปัจจุบันการรับประทานอาหารเพื่อสุขภาพกำลังได้รับความนิยมเป็นอย่างมาก ผู้คนจึงเริ่มทำอาหารรับประทานกันด้วยตัวเองจากวัตถุดิบในการปรุงอาหารที่มีอยู่
- ซึ่งบางครั้งไม่ทราบว่าจะสามารถนำไปทำอาหารอะไรได้บ้าง จึงทำได้แต่รายการอาหารเดิม ๆ ทำให้เกิดความจำเจในการรับประทานอาหาร และไม่สนุกกับการทำอาหาร
- บทความวิจัยนี้นำเสนอการออกแบบและพัฒนาระบบแบ่งปันสูตรการทำอาหาร และค้นหาสูตรการทำอาหารจากภาพวัตถุดิบด้วยเทคนิคการเรียนรู้เชิงลึก
- โดยใช้งานได้นับอุปกรณ์พกพา ซึ่งผู้ใช้งานสามารถแบ่งปันสูตรการทำอาหารและสามารถใช้อุปกรณ์พกพารูปวัตถุดิบที่ผู้ใช้อยู่แล้วในครัว เช่น กระเทียม เนื้อหมู ผัก ฯลฯ เพื่อส่งภาพเข้ามาในระบบให้ค้นหาสูตรอาหารจากวัตถุดิบที่มีอยู่
- ทำให้เพิ่มความสะดวกสบายในการค้นหาสูตรอาหารให้แก่ผู้ใช้ส่วนประกอบหลักของระบบประกอบด้วย
- (1) แอปพลิเคชันบนอุปกรณ์พกพาสำหรับผู้ใช้งานทั่วไปที่พัฒนาด้วยReact Native ซึ่งผู้ใช้งานสามารถเพิ่มสูตรการทำอาหาร และค้นหาสูตรการทำอาหารได้โดยใส่ชื่อวัตถุดิบ รวมถึงการใช้อุปกรณ์พกพารูปวัตถุดิบเพื่อค้นหาสูตรการทำอาหาร
- (2) เว็บแอปพลิเคชันพัฒนาบน MERN stack สำหรับผู้ดูแลระบบ เพื่อใช้ในการเพิ่มคำหลักในการค้นหาให้กับวัตถุดิบ รวมถึงทดสอบแบบจำลองที่ได้สร้างขึ้น และ

- (3) โครงข่ายประสาทเทียมการเรียนรู้เชิงลึกที่มีการใช้อัลกอริทึม YOLO ผ่านไลบรารี Darknet สำหรับการสร้างแบบจำลองในการรู้จำภาพแบบจำลองได้ฝึกสอนให้สามารถรู้จำวัตถุได้จำนวน 20 ประเภท
- โดยสอนด้วยภาพวัตถุประเภทละ 100 ภาพ หลังจากฝึกสอนแบบจำลองจำนวน 36,000 รอบ แบบจำลองมีค่าสูญเสียเฉลี่ยอยู่ที่ 0.0408 และมีค่า Precision, Recall และ F1-score อยู่ที่ 0.96, 0.98 และ 0.97 ตามลำดับ

บทนำ

- Deep Learning หรือการเรียนรู้เชิงลึก ถือว่าเป็นหนึ่งในระบบการเรียนรู้ของเครื่อง (Machine Learning)
- ซึ่งมีโครงสร้างและการประมวลผลคล้ายกับสมองของมนุษย์ที่เรียกว่า โครงข่ายประสาทเทียม (Artificial Neural Network)
- โดยอันที่จริงแล้ว Deep Learning คือโครงข่ายประสาทเทียมขนาดใหญ่
- ซึ่งหมายถึงเป็นโครงข่ายประสาทเทียมที่มีชั้นในการประมวลผลหลายชั้นช่วงไม่กี่ปีที่ผ่านมาการเรียนรู้เชิงลึกได้รับความนิยมและมีงานวิจัยที่เกี่ยวข้องมากขึ้น เช่น ใน [1] ได้พัฒนาระบบการแนะนำ (Recommendation system) ด้วยการใช้เทคนิคการเรียนรู้เชิงลึก
- หรือแม้กระทั่งใน [2] ได้มีการใช้เทคนิคการเรียนรู้เชิงลึกในการพัฒนาระบบตรวจจับผู้บุกรุก (Intrusion Detection System)
- แต่ด้านการใช้งานที่ทำให้เทคนิคการเรียนรู้เชิงลึกได้รับความนิยมมากมาจากการงานด้านรู้จำภาพ โดยเป็น การใช้โครงข่ายประสาทเทียมประเภท Convolutional Neural Network
- ดังอธิบายใน [3] และมีบทความวิจัยที่ศึกษาและประยุกต์ใช้การรู้จำภาพอย่างมากมาย เช่น ใน [4] เป็นโครงข่ายประสาทเทียมสำหรับตรวจจับภาพอาหารและ [5], [6] และ [7] เป็นโครงข่ายประสาทเทียมสำหรับตรวจจับอารมณ์ในภาพ
- ในปัจจุบันแนวโน้มการรับประทานอาหารเพื่อสุขภาพกำลังได้รับความนิยมเป็นอย่างมาก เนื่องจากโรคประจำตัวต่างๆ
- ซึ่งส่วนใหญ่แล้วมักเกิดจากการรับประทานอาหาร ผู้คนส่วนใหญ่จึงเริ่มทำอาหารรับประทานกันด้วยตัวเอง แต่สำหรับผู้เริ่มต้นทำอาหารอาจมีความลำบากและยุ่งยากในการค้นหาสูตรการทำอาหารจากหนังสือทำอาหาร หรือจากอินเทอร์เน็ตอีกทั้งบางครั้งมีวัตถุดิบในการปรุงอาหารอยู่ที่บ้าน
- แต่คิดไม่ออกว่าสามารถนำไปทำอาหารอะไร ได้บ้าง จึงทำได้แต่รายการอาหารเดิม ๆ จึงเกิดความจำเจในการรับประทานอาหาร และไม่สนุกกับการทำอาหาร

- ดังนั้นบทความวิจัยนี้จึงต้องการนำเสนอการออกแบบและพัฒนาระบบแบ่งปันสูตรการทำอาหารและค้นหาสูตรการทำอาหารจากภาพวัตถุดิบด้วยเทคนิคการเรียนรู้เชิงลึก
- โดยใช้งานได้นับอุปกรณ์พกพา ซึ่งผู้ใช้งานสามารถแบ่งปันสูตรการทำอาหาร และสามารถใช้อุปกรณ์พกพารูปวัตถุดิบที่ผู้ใช้มีอยู่แล้วในครัว
- เช่น กระเทียม เนื้อหมู ผัก ฯลฯ
- เพื่อส่งภาพเข้ามาในระบบให้ค้นหาสูตรอาหารจากวัตถุดิบที่มีอยู่ เพื่อเพิ่มความสะดวกสบายในการค้นหาสูตรอาหารให้แก่ผู้ใช้

การแยกภาพตัวอักษรลายมือเขียนภาษาไทยแบบอัตโนมัติ โดยใช้การวิเคราะห์ถดถอยเชิงเส้นและ การเรียนรู้ด้วยต้นไม้ตัดสินใจ

บทคัดย่อ

บทความฉบับนี้นำเสนอหลักการสำหรับแยกภาพตัวอักษรลายมือเขียนที่อยู่ติดกันแบบสัมผัสใน เอกสารภาพตัวอักษรออกจากกัน ซึ่งเป็นกระบวนการเตรียมพร้อมสำหรับการรู้จำลายมือเขียน เนื่องจากลักษณะของการเขียนภาษาไทยมีความแตกต่างจากภาษาอังกฤษ ซึ่งสามารถแบ่งออกได้เป็น 4 ระดับ โดยสามารถติดกันได้ในระดับเดียวกันและข้ามระดับทั้งในแนวนอนและแนวตั้ง หลักการที่ใช้ในบทความนี้ประกอบด้วยการรับภาพเอกสารลายมือเขียนมาคัดแยกให้เป็นตัวอักษรเดี่ยวและตัวอักษรติดกัน จากนั้นจะวิเคราะห์ตัวอักษรด้วยคุณลักษณะต่างๆ ของตัวอักษรไทยเพื่อแยกตัวอักษรที่ติดกันในแนวนอนและแนวตั้ง โดยมีการใช้เส้นการวิเคราะห์การถดถอยสำหรับตัดแบ่งระดับพยัญชนะกับสระ ขั้นตอนนี้ทำการตัดแบ่งตัวอักษรก่อนการรู้จำตัวอักษรตามหลักของการรู้จำตัวอักษรไทย ผลการทดลองพบว่าความถูกต้องของการแยกตัวอักษรลายมือเขียนภาษาไทยเป็นร้อยละ 90.44

บทนำ

การรู้จำภาพตัวอักษรลายมือเขียน (handwriting recognition) เป็นงานวิจัยขั้นสูงต่อจากการรู้จำภาพตัวอักษรพิมพ์ (optical character recognition) โดยมีวัตถุประสงค์เพื่อช่วยให้ผู้ใช้สามารถจัดเก็บเอกสารที่มีอยู่เข้าสู่ระบบคอมพิวเตอร์ได้สะดวกขึ้น ซึ่งนอกจากจะเป็นเอกสารพิมพ์แล้วยังมีเอกสารจำนวนมากที่อยู่ในรูปของเอกสารลายมือเขียน เช่น ข้อมูลที่ผู้ใช้กรอกใบสมัครงาน รายงานการประชุม ข้อความจากการบรรยาย เป็นต้น งานวิจัยทางด้านการรู้จำภาพตัวอักษรลายมือเขียนในภาษาไทย (Phokharatkul and Kimpan, 2002) เสนอวิธีการรู้จำตัวอักษรลายมือเขียนในภาษาไทยโดยใช้ตัวอธิบายรูปร่าง แบบฟูริเยร์ (Fourier descriptor) เป็นตัวอธิบายลักษณะเด่นของตัวอักษร จากนั้นฝึกฝนและรู้จำโดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น ซึ่งค่าน้ำหนักของโหนดถูกพิจารณาโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (genetic algorithm) ผลการรู้จำเมื่อทดลองกับตัวอักษรลายมือเขียนไทยแบบบรรจง จำนวน 13,500 ตัวอักษร จากผู้ทดลอง 60 คน เท่ากับร้อยละ 99.12 แต่อย่างไรก็ตามตัวอักษรลายมือเขียนที่ใช้เป็นแบบเขียนบรรจงซึ่งไม่สอดคล้องกับการใช้งานจริง ต่อมา Methasate และ Sae-tang (2004) เสนอวิธีการจัดกลุ่มตัวอักษร ลายมือเขียนด้วยการพิจารณาโครงสร้างของตัวอักษรที่คล้ายกัน ลักษณะที่นำมาพิจารณา ได้แก่ เส้นตัวอักษรในแนวตั้ง (vertical stroke) และการกระจายของพิกเซลภาพตัวอักษร (pixel distribution) จากนั้นนำมาฝึกฝนและรู้จำโดยใช้โครงข่ายประสาทเทียมแบบแพร่กระจายกลับ (back propagation neural network) ผลการวิจัยสามารถแบ่งกลุ่มตัวอักษรออกได้เป็น 21 กลุ่ม มีระดับความถูกต้องที่ร้อยละ 97.60 นอกจากนี้ Nopsuwanchai และคณะ (2006) ได้ใช้ส่วน ประกอบสำคัญ (principal

component) ของตัวอักษรแต่ละตัวร่วมด้วยรูปแบบต่าง ๆ ของตัวอักษร ได้แก่ ภาพกลับซ้ายของตัวอักษร (polar transformed image) และภาพตัวอักษรหมุนขวา 90 องศา เป็นลักษณะเด่นของตัวอักษร จากนั้นนำไปฝึกฝน และรู้จำโดยใช้ แบบจำลองฮิดเดินมาร์ คอฟ โดยเปรียบเทียบระหว่างคลังข้อมูล 2 ชุด ได้ผลการรู้จำถูกต้องร้อยละ 95.98 และ 95.13 ตามลำดับ แต่อย่างไรก็ตามถึงแม้งานวิจัยที่ผ่านมาจะให้ความถูกต้องในการรู้จำสูงแต่ภาพ ตัวอักษร ลายมือเขียนที่นำมาทดสอบจะมีการแบ่งช่องว่าง ระหว่างตัวอักษรที่ชัดเจน ไม่มีการสัมผัส และ ทับ ซ้อน กัน ซึ่งการนำเอาวิธีการเหล่านี้มาใช้งานจริง อาจจะไม่สามารถรู้จำภาพลายมือเขียนที่มีการ สัมผัส (ขวา) และทับ ซ้อนกัน (ซ้าย) ได้อย่างถูกต้อง รูปที่ 1 แสดงถึงตัวอย่างของลายมือเขียนที่อาจทำ ให้เกิดข้อผิดพลาดในการรู้จำ มี นักวิจัยหลายท่านเห็นว่าถ้ามีการแยกตัว อักษรที่สัมผัสกันได้อย่างถูกต้อง เมื่อนำไปรู้จำแล้วจะได้ผลของการรู้จำที่ สูงขึ้น (Chatchinarat, 2009) โดยเทคนิคของแอ่งน้ำซึ่งเป็นการเติมน้ำลงบน ตัวอักษร โดยพบว่าตัวอักษรที่เชื่อม ติดกันจะมีแอ่งน้ำที่เกิดขึ้นระหว่างตัวอักษรคู่กันซึ่งสามารถเป็นจุดตัดได้สามารถแยกได้ถูกต้องร้อยละ 87.60 ซึ่ง ข้อผิดพลาดเกิดจากตัวอักษรที่ไม่สามารถหาแอ่งน้ำได้ และตัวอักษรที่มีจำนวนแอ่งน้ำเท่ากัน นอกจากนี้ขนาดของ ตัวอักษรแต่ละตัวจะต้องมี ขนาดเท่ากัน ดังนั้นงานวิจัยนี้จึงเสนอการแยกตัวอักษรลายมือเขียนที่เขียนในลักษณะ ธรรมชาติคือ มีการสัมผัสกันทั้งในแนวนอนและแนวตั้ง และมีขนาด ของตัวอักษรแตกต่างกันได้ โดยใช้การ วิเคราะห์ ถดถอยเชิงเส้น (linear regression analysis) สำหรับวิเคราะห์หากกลุ่มของตัวอักษรที่อยู่ติดกันใน แนวตั้ง และใช้ความกว้างมัธยฐานของตัวอักษร (median of character width) ในการวิเคราะห์กลุ่ม อักษรที่อยู่ ติดกันในแนวนอน จากนั้นแยกตัวอักษร โดยใช้แบบจำลองต้นไม้ตัดสินใจที่ได้มาจากการฝึกฝน

บทคัดย่อ

- **บทความ**ฉบับนี้นำเสนอหลักการสำหรับแยกภาพตัวอักษรลายมือเขียนที่อยู่ติดกันแบบสัมผัสใน เอกสาร ภาพตัวอักษรออกจากกัน
- **ซึ่ง**เป็นกระบวนการเตรียมพร้อมสำหรับการรู้จำลายมือเขียน เนื่องจากลักษณะของการเขียนภาษาไทยมี ความแตกต่างจากภาษาอังกฤษ
- **ซึ่ง**สามารถแบ่งออกได้เป็น 4 ระดับ โดยสามารถติดกันได้ในระดับเดียวกันและข้ามระดับทั้งในแนวนอน และแนวตั้ง หลักการที่ใช้ในบทความนี้ประกอบด้วยการรับภาพเอกสารลายมือเขียนมาคัดแยกให้เป็น ตัวอักษรเดี่ยวและตัวอักษรติดกัน
- **จากนั้น**จะวิเคราะห์ตัวอักษรด้วยคุณลักษณะต่าง ๆ ของตัวอักษรไทยเพื่อแยกตัวอักษรที่ติดกันใน แนวนอนและแนวตั้ง

- โดยมีการใช้เส้นการวิเคราะห์การถดถอยสำหรับตัดแบ่งระดับพยัญชนะกับสระ ขั้นตอนนี้ทำการตัดแบ่งตัวอักษรก่อนการรู้จำตัวอักษรตามหลักของการรู้จำตัวอักษรไทย
- ผลการทดลองพบว่าความถูกต้องของการแยกตัวอักษรลายมือเขียนภาษาไทยเป็นร้อยละ 90.44

บทนำ

- การรู้จำภาพตัวอักษรลายมือเขียน (handwriting recognition) เป็นงานวิจัยขั้นสูงต่อจากการรู้จำภาพตัวอักษรพิมพ์ (optical character recognition)
- โดยมีวัตถุประสงค์เพื่อช่วยให้ผู้ใช้สามารถจัดเก็บเอกสารที่มีอยู่เข้าสู่ระบบคอมพิวเตอร์ได้สะดวกขึ้น
- ซึ่งนอกจากจะเป็นเอกสารพิมพ์แล้วยังมีเอกสารจำนวนมากที่อยู่ในรูปของเอกสารลายมือเขียน เช่น ข้อมูลที่ใช้กรอกใบสมัครงาน รายงานการประชุม ข้อความจากการบรรยาย เป็นต้น
- งานวิจัยทางด้านการรู้จำภาพตัวอักษรลายมือเขียนในภาษาไทย (Phokharatkul and Kimpan, 2002) เสนอวิธีการรู้จำตัวอักษรลายมือเขียนในภาษาไทยโดยใช้ตัวอธิบายรูปร่าง แบบฟูริเยร์ (Fourier descriptor) เป็นตัวอธิบาย ลักษณะเด่นของตัวอักษร
- จากนั้นฝึกฝนและรู้จำโดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น ซึ่งค่าน้ำหนักของโหนดถูกพิจารณาโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (genetic algorithm) ผลการรู้จำเมื่อทดลองกับตัวอักษรลายมือเขียนไทยแบบบรรจง จำนวน 13,500 ตัวอักษร จากผู้ทดลอง 60 คน เท่ากับร้อยละ 99.12
- แต่อย่างไรก็ตามตัวอักษรลายมือเขียนที่ใช้เป็นแบบเขียนบรรจงซึ่งไม่สอดคล้องกับการใช้งานจริง
- ต่อมา Methasate และ Sae-tang (2004) เสนอวิธีการจัดกลุ่มตัวอักษร ลายมือเขียนด้วยการพิจารณาโครงสร้างของตัวอักษรที่คล้ายกัน ลักษณะที่นำมาพิจารณา ได้แก่ เส้นตัวอักษรในแนวตั้ง (vertical stroke) และการกระจายของพิกเซลภาพตัวอักษร (pixel distribution)
- จากนั้นนำมาฝึกฝนและรู้จำโดยใช้โครงข่ายประสาทเทียมแบบแพร่กระจายกลับ (back propagation neural network)
- ผลการวิจัยสามารถแบ่งกลุ่มตัวอักษรออกได้เป็น 21 กลุ่ม มีระดับความถูกต้องที่ร้อยละ 97.60 นอกจากนี้ Nopsuwanchai และคณะ (2006) ได้ใช้ส่วน ประกอบสำคัญ (principal component) ของตัวอักษรแต่ละตัวร่วมด้วยรูปแบบต่าง ๆ ของตัวอักษร ได้แก่ ภาพกลับหัวของตัวอักษร (polar transformed image) และภาพตัวอักษรหมุนขวา 90 องศา เป็นลักษณะเด่นของตัวอักษร

- จากนั้นนำไปฝึกฝนและรู้จำโดยใช้ แบบจำลองฮิดเดินมาร์ค คอฟ โดยเปรียบเทียบระหว่างคลังข้อมูล 2 ชุด ได้ผลการรู้จำถูกต้องร้อยละ 95.98 และ 95.13 ตามลำดับ
- แต่อย่างไรก็ตามถึงแม้งานวิจัยที่ผ่านมาจะให้ความถูกต้องในการรู้จำสูงแต่ภาพตัวอักษรลายมือเขียนที่นำมาทดสอบจะมีการแบ่งช่องว่าง ระหว่างตัวอักษรที่ชัดเจน ไม่มีการสัมผัส และ ทับ ซ้อนกัน
- ซึ่งการนำเอาวิธีการเหล่านี้มาใช้งานจริง อาจจะไม่สามารถรู้จำภาพลายมือเขียนที่มีการ สัมผัส (ขวา) และ ทับซ้อนกัน (ซ้าย) ได้อย่างถูกต้อง รูปที่ 1 แสดงถึงตัวอย่างของลายมือเขียนที่อาจทำให้เกิดข้อผิดพลาดในการรู้จำ
- มีนักวิจัยหลายท่านเห็นว่าถ้ามีการแยกตัว อักษรที่สัมผัสกันได้อย่างถูกต้อง เมื่อนำไปรู้จำแล้วจะได้ผลของการรู้จำที่สูงขึ้น (Chatchinarat, 2009)
- โดยเทคนิคของแอ่งน้ำซึ่งเป็นการเติมน้ำลงบน ตัวอักษร โดยพบว่าตัวอักษรที่เชื่อมติดกันจะมีแอ่งน้ำที่เกิดขึ้นระหว่างตัวอักษรคู่่นั้นซึ่งสามารถเป็นจุดตัดได้สามารถแยกได้ถูกต้องร้อยละ 87.60
- ซึ่ง ข้อผิดพลาดเกิดจากตัวอักษรที่ไม่สามารถหาแอ่งน้ำได้ และตัวอักษรที่มีจำนวนแอ่งน้ำเท่ากัน ข
- นอกจากนี้ขนาดของตัวอักษรแต่ละตัวจะต้องมี ขนาดเท่ากัน
- ดังนั้นงานวิจัยนี้จึงเสนอการแยกตัวอักษรลายมือเขียนที่เขียนในลักษณะธรรมชาติคือ มีการสัมผัสกันทั้งในแนวนอนและแนวตั้ง และมีขนาด ของตัวอักษรแตกต่างกันได้
- โดยใช้การวิเคราะห์ ถดถอยเชิงเส้น (linear regression analysis) สำหรับวิเคราะห์หากกลุ่มของตัวอักษรที่อยู่ติดกันใน แนวตั้ง และใช้ความกว้างมัธยฐานของตัวอักษร (median of character width) ในการวิเคราะห์กลุ่ม อักษรที่อยู่ติดกันในแนวนอน
- จากนั้นแยกตัวอักษร โดยใช้แบบจำลองต้นไม้ตัดสินใจที่ได้มาจากการฝึกฝน

คำนำหน้า	JN01	JN02	JN03	JN04	JN05	JN06	JN07	JN08	JN09	JN10	รวม
ผู้บริโภค	0	1	0	0	0	0	0	0	0	0	1
จึง	0	1	0	0	0	0	0	0	0	0	1
สำหรับ	0	1	0	0	0	0	0	0	0	0	1
ช่องทาง	0	1	0	0	0	0	0	0	0	0	1
ทำให้	0	1	0	0	0	0	0	0	1	0	2
แม้ว่า	0	1	0	0	0	1	0	0	0	0	2
จากปัญหา	0	1	0	0	0	0	0	0	0	0	1
แล้วจึง	0	1	0	0	0	0	0	0	0	0	1
ทำการ	0	0	1	0	0	0	0	0	0	0	1
สรุป	0	0	1	0	0	0	0	0	0	0	1
บทบาท	0	0	1	0	0	0	0	0	0	0	1
เนื่องจาก	0	0	1	0	0	0	0	1	0	0	2
จาก	0	0	1	0	0	1	0	0	0	0	2
รวมไปถึง	0	0	1	0	0	0	0	0	0	0	1
เริ่มต้น	0	0	0	1	0	0	0	0	0	0	1
พบว่า	0	0	0	1	0	0	0	0	0	0	1
โดยเฉพาะ	0	0	0	1	0	0	0	0	0	0	1
แล้ว	0	0	0	1	0	0	0	0	0	0	1
ทั้งนี้	0	0	0	1	0	1	0	0	0	0	2
จากการ	0	0	0	0	1	0	0	0	0	0	1
ในปัจจุบัน	0	0	0	0	1	0	0	1	0	0	2
โปรแกรม	0	0	0	0	1	0	0	0	0	0	1
เรา	0	0	0	0	1	0	0	0	0	0	1
แต่	0	0	0	0	1	1	0	0	2	0	4
ท่าอากาศยาน	0	0	0	0	0	4	0	0	0	0	4
เป้าหมายหลัก	0	0	0	0	0	1	0	0	0	0	1
การวิเคราะห์	0	0	0	0	0	1	0	0	0	0	1
ผลการศึกษา	0	0	0	0	0	1	0	0	0	0	1
จตุภาคที่1	0	0	0	0	0	1	0	0	0	0	1

คำนำหน้า	JN01	JN02	JN03	JN04	JN05	JN06	JN07	JN08	JN09	JN10	รวม
จดุภาคที่2	0	0	0	0	0	1	0	0	0	0	1
จดุภาคที่3	0	0	0	0	0	1	0	0	0	0	1
และจดุภาคที่ 4	0	0	0	0	0	1	0	0	0	0	1
โดยพบว่า	0	0	0	0	0	1	0	0	0	0	1
ถูกจัดอยู่ใน	0	0	0	0	0	1	0	0	0	0	1
การประเมิน	0	0	0	0	0	1	0	0	0	0	1
ที่เริ่มต้น	0	0	0	0	0	1	0	0	0	0	1
อย่างไรก็ตาม	0	0	0	0	0	2	0	0	0	0	2
เพื่อ	0	0	0	0	0	1	0	0	1	0	2
งานวิจัย	0	0	0	0	0	0	1	0	0	1	2
จากนั้น	0	0	0	0	0	0	1	1	0	5	7
ผลการ ทดสอบ	0	0	0	0	0	0	1	0	0	0	1
สามารถสรุป ได้ว่า	0	0	0	0	0	0	1	0	0	0	1
เนื่องจาก	0	0	0	0	0	0	1	0	0	0	1
ทั้งนี้ยัง	0	0	0	0	0	0	1	0	0	0	1
ข้อมูล	0	0	0	0	0	0	1	0	0	0	1
ปัญหาที่	0	0	0	0	0	0	1	0	0	0	1
โดยที่	0	0	0	0	0	0	1	0	0	0	1
ส่งผลให้	0	0	0	0	0	0	1	0	0	0	1
ผู้พิการทาง สายตา	0	0	0	0	0	0	1	0	0	0	1
ปัจจัย	0	0	0	0	0	0	0	1	0	0	1
ผู้วิจัย	0	0	0	0	0	0	0	1	0	0	1
เพื่อให้	0	0	0	0	0	0	0	3	0	0	3
ผลวิจัย	0	0	0	0	0	0	0	1	0	0	1
ในส่วน	0	0	0	0	0	0	0	1	0	0	1

[illegible]

บทความที่ 1

บทคัดย่อ

[การวิจัยครั้งนี้เป็นการวิจัยเพื่อศึกษาผลการจัดกลุ่มข้อมูลโดยใช้เทคโนโลยีปัญญาประดิษฐ์ AI แบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนกลุ่มตัวอย่างเป็น กลุ่มเก่ง กลุ่มปานกลาง , 'กลุ่มอ่อน และเพื่อศึกษาผลการจัดกลุ่มข้อมูลโดยใช้เทคโนโลยีปัญญาประดิษฐ์ AI แบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของรายวิชาเป็น กลุ่มเด่น กลุ่มกลาง กลุ่มด้อย , 'กลุ่มตัวอย่าง'ที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนชั้นประถมศึกษาปีที่ 5 ปีการศึกษา 2561 โรงเรียนสามเสนนอก (ประชาราษฎร์อนุกุล) สำนักงาน เขตดินแดง กรุงเทพมหานคร จำนวน 33 คน ได้มาจากการสุ่มตัวอย่างอย่างง่ายด้วยวิธีจับสลาก , 'เครื่องมือ'ที่ใช้ในการวิจัยครั้งนี้ คือ เทคโนโลยีปัญญาประดิษฐ์ AI โดยใช้วิธีการเรียนรู้การแบ่งกลุ่มข้อมูลชนิด K-mean ผลการวิจัย พบว่า การแบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของนักเรียนกลุ่มตัวอย่างเป็น กลุ่มเก่ง 11 คน กลุ่มปานกลาง 11 คน กลุ่มอ่อน 11 คน และการแบ่งกลุ่มผลสัมฤทธิ์ทางการเรียนของรายวิชาเป็น กลุ่มเด่น 4 รายวิชา กลุ่มกลาง 3 รายวิชา กลุ่มด้อย 3 รายวิชา โดยมีรายวิชา ที่เด่นสุดคือ วิชาประวัติศาสตร์ รองลงมาคือ วิชาการงานอาชีพ และรายวิชาที่ด้อยสุดคือ วิชาภาษาอังกฤษ]

- แบ่งประโยคด้วยมนุษย์ : 8 ประโยค
- แบ่งด้วย crfcut : 4 ประโยค

บทนำ

['ความก้าวหน้าและการพัฒนาทางเทคโนโลยีส่งผลให้ความเป็นอยู่ของชีวิตมนุษย์ดีขึ้น ' , 'หนึ่งในเทคโนโลยีที่ถูกนำมาใช้คือ เทคโนโลยีปัญญาประดิษฐ์ (artificial intelligence: AI) ซึ่งได้เริ่มขึ้นหลังจากการประชุมวิชาการที่วิทยาลัยดาร์ตมัธ สหรัฐอเมริกา ในปี ค.ศ. 1956 ในครั้งนั้นมีผู้ร่วมประชุมหลายคน ได้แก่ จอร์น แม็กคาร์ธี เป็นประธานในที่ประชุม ร่วมด้วย ' , 'มาร์วิน มินสกี, อัลเลน นิวเวลล์, อาเธอร์ ซามเอล, และเฮอร์เบิร์ต ไซมอน ต่อมาบุคคลเหล่านี้ได้กลายมาเป็นผู้นำทางสาขาปัญญาประดิษฐ์ แต่ยังไม่เป็นที่นิยมในสมัยนั้นเนื่องจากยากต่อการทำความเข้าใจของคนทั่วไป (Hutton, 2011) ปัจจุบันสถาบันการศึกษาหลายแห่งเปิดสอนหลักสูตรที่เกี่ยวข้องกับการนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ ' , 'มีงานวิจัยเผยแพร่และข้อมูลจำนวนมากมหาศาลซึ่งถูกนำมาใช้ประโยชน์กับมนุษย์ ' , 'อย่างไรก็ตามการใช้ประโยชน์ในเชิงกลยุทธ์จากข้อมูลที่รวบรวมได้ ' , 'จำเป็นต้องมีระบบสารสนเทศที่สามารถเรียนรู้จากข้อมูลนั้น ']

- แบ่งประโยคด้วยมนุษย์ : 16 ประโยค
- แบ่งด้วย crfcut : 6 ประโยค