

บทที่ 6

สรุปผลงานวิจัยและข้อเสนอแนะ

6.1 สรุปผลการวิจัย

งานวิจัยนี้เป็นการพัฒนาการแสดงผลภาพบิตแม็บสำหรับข้อมูลเอกสารดิจิทัล โดยการแปลงข้อมูลเอกสารเป็นภาพบิตแม็บ เพื่อช่วยให้ผู้ใช้สามารถพิจารณาความเหมือนและความแตกต่างของชนิดหรือหมวดหมู่ของเอกสารเป็นจำนวนมากในเบื้องต้นได้ ซึ่งผู้ใช้ไม่จำเป็นต้องเข้าไปพิจารณาในเนื้อความของเอกสาร โดยงานวิจัยนี้ได้แนวคิดมาจากทฤษฎีเคออสเกม ที่ถูกนำมาประยุกต์ใช้ในการแสดงผลภาพของข้อมูลดีเอ็นเอ และข้อมูลอนุกรมเวลา เพื่อทำการสรุปรวมข้อมูลที่ซับซ้อนและมีปริมาณมากให้สามารถแสดงผลเป็นภาพ ที่ง่ายแก่การพิจารณาเปรียบเทียบความเหมือนและความแตกต่างข้อมูล

ขั้นตอนการพัฒนาการแสดงผลภาพบิตแม็บสำหรับข้อมูลเอกสารดิจิทัล เริ่มจากการแปลงข้อมูลจากเอกสารดิจิทัลไปเป็นข้อมูลอนุกรมเวลาด้วยมาตรฐานแอลกี ซึ่งต้องมีการวิเคราะห์และปรับข้อมูลในเอกสาร แล้วทำการปรับข้อมูลอนุกรมเวลาที่ได้มา เพื่อลดความแปรปรวนของข้อมูล จากนั้นจะทำการแปลงข้อมูลอนุกรมเวลาไปเป็นอักขระ ด้วยวิธีการแบบแซ็ค (Symbolic Aggregate approXimation) แล้วทำการแปลงข้อมูลอักขระไปเป็นภาพบิตแม็บ ซึ่งภาพบิตแม็บที่ได้จะมีความสัมพันธ์กับลักษณะและเนื้อความในเอกสาร นอกจากนี้ยังได้ทำการทดลองกับข้อมูลเอกสารชุดตัวอย่างเพื่อหาค่าพารามิเตอร์ที่เหมาะสม เพื่อให้การแสดงผลภาพบิตแม็บ สามารถแสดงผลภาพบิตแม็บจากข้อมูลเอกสารได้ชัดเจนและมีประสิทธิภาพมากยิ่งขึ้น

งานวิจัยนี้ได้ทำการวัดผลการแสดงผลภาพบิตแม็บสำหรับข้อมูลเอกสารดิจิทัล โดยทำการวัดผล ทั้งกับข้อมูลดีเอ็นเอและข้อมูลอนุกรมเวลา เพื่อทดสอบความเป็นไปได้ของแนวทางการวิจัย และทำการทดสอบกับข้อมูลเอกสาร เพื่อทดสอบประสิทธิภาพของการแสดงผลภาพบิตแม็บ ซึ่งสรุปผลการทดสอบได้ดังนี้

6.1.1 ผลการทดสอบกับข้อมูลดีเอ็นเอและข้อมูลอนุกรมเวลา

ภาพบิตแม็บของข้อมูลดีเอ็นเอและข้อมูลอนุกรมเวลา ที่ได้จากการแสดงผลภาพบิตแม็บ สามารถสังเกตความเหมือนและแตกต่างจากลักษณะการแสดงสี และโครงสร้างของภาพบิตแม็บที่เกิดจากข้อมูลตามกลุ่มได้อย่างชัดเจน ทำให้สรุปได้ว่ามีความเป็นไปได้ที่จะพัฒนาการแสดงผลภาพบิตแม็บสำหรับข้อมูลเอกสารดิจิทัล

6.1.2 ผลการทดสอบกับข้อมูลเอกสารดิจิทัลโดยการจัดกลุ่มด้วยวิธีเคมีน

ผลจากการจัดกลุ่มภาพบิตแม็บจากข้อมูลเอกสารดิจิทัลด้วยวิธีเคมีน สามารถทำการจัดกลุ่มเอกสารจากภาพบิตแม็บได้ถูกต้องเฉลี่ย 78.18 เปอร์เซ็นต์ แต่หลังจากมีการปรับอัลกอริทึมของเคมีนในการเลือกตัวแทนกลุ่มเริ่มต้น โดยทำการตรวจสอบระยะห่างของตัวแทนกลุ่มเริ่มต้นในแต่ละกลุ่ม ให้มีระยะห่างกันไม่น้อยกว่าค่าระยะห่างที่มากที่สุดของเอกสารในกลุ่มเดียวกัน และเพิ่มขึ้นตอน การเปรียบเทียบสมาชิกที่ได้จากการจัดกลุ่มที่ต้องได้สมาชิกในกลุ่มเหมือนกันอย่างน้อย 2 ครั้ง ทำให้สามารถทำการจัดกลุ่มได้ถูกต้องเฉลี่ย 92.72 เปอร์เซ็นต์

สรุปผลจากการทดสอบการแสดงผลภาพบิตแม็บสำหรับข้อมูลเอกสารดิจิทัล พบว่าการแสดงผลภาพบิตแม็บนี้สามารถนำไปใช้เพื่อช่วยในการพิจารณา จัดกลุ่ม จำแนก และแยกแยะข้อมูลเอกสารดิจิทัลได้จริงและมีประสิทธิภาพ

6.2 ปัญหาที่พบจากการวิจัย

ปัญหาที่พบจากการทำการวิจัยบางประการ ที่น่าจะเป็นประโยชน์และสามารถนำไปเป็นแนวทางในการแก้ไขปัญหามานงานวิจัยที่ใกล้เคียงอื่นๆต่อไป ได้แก่

1) ปัญหาในการเลือกพิจารณาเนื้อความในเอกสารดิจิทัล ที่เกิดจากข้อมูลในเอกสารมีข้อความบางส่วนที่ไม่มีความสัมพันธ์กับเนื้อความหลักอยู่ โดยเฉพาะช่วงต้นของเอกสาร เช่น คำนำ สารบัญ หรือบทนำ เป็นต้น ซึ่งส่งผลให้การประมวลผลข้อมูลเอกสารของการแสดงผลภาพบิตแม็บเกิดความผิดพลาด ทำให้ได้ภาพบิตแม็บที่มีลักษณะโครงสร้างและการแสดงสีผิดแปลกออกไป ซึ่งการแก้ไขเบื้องต้นสามารถทำได้โดยเข้าไปทำการเลือกตัดข้อมูลในเอกสาร ที่ไม่เกี่ยวข้องของกับเนื้อความหลักของเอกสารด้วยมือ แต่ในกรณีทั่วไปสำหรับเอกสารที่มีความยาวมาก สามารถทำการแก้ไขปัญหานี้ได้ โดยการเลือกจุดเริ่มต้นการประมวลผลข้อมูลเอกสารด้วยการสุ่ม โดยคำนวณช่วงการสุ่มจากขนาดความยาวของเอกสาร ที่ต้องการประมวลผลกับขนาดความยาวทั้งหมดของเอกสาร

2) ปัญหาของการจัดกลุ่มภาพบิตแม็บของเอกสารด้วยวิธีเคมีน ที่ให้ผลการจัดกลุ่มเอกสารผิดพลาดอยู่บ้าง เนื่องจากได้สมาชิกกลุ่มที่ไม่แน่นอนจากการจัดกลุ่มในแต่ละครั้ง ซึ่งจากการศึกษาเพื่อหาสาเหตุดังกล่าวพบว่า ภาพบิตแม็บของเอกสารมีระยะทางที่แตกต่างกันไม่มาก และจำนวนเอกสารที่นำมาทดลองมีจำนวนน้อย ซึ่งเป็นข้อจำกัดอย่างหนึ่งของการจัดกลุ่มด้วยวิธีเคมีน ดังนั้นเพื่อให้จัดกลุ่มภาพบิตแม็บมีความถูกต้องมากขึ้น จึงทำการปรับอัลกอริทึมของเคมีนเพิ่มเติม โดยขั้นตอนการตรวจสอบระยะห่างของตัวแทนกลุ่มเริ่มต้นในแต่ละกลุ่ม ให้มีความแตกต่างกันไม่น้อยกว่าค่าระยะห่างที่มากที่สุดของเอกสารที่อยู่ในกลุ่มเดียวกัน เพื่อป้องกัน

ไม่ให้เกิดการเลือกตัวแทนกลุ่มเริ่มต้นที่ไม่ดี นอกจากนี้ยังเพิ่มขึ้นตอนในการเปรียบเทียบสมาชิกที่ได้จากการจัดกลุ่ม ที่ต้องมีสมาชิกในกลุ่มเหมือนกันจากการทำการจัดกลุ่มอย่างน้อย 2 ครั้ง เพื่อเป็นการตรวจสอบสมาชิกกลุ่มให้มีความถูกต้องมากยิ่งขึ้น

6.3 ข้อเสนอแนะ

1) อาจพัฒนาการแสดงผลภาพบิตแม็บต่อเนื่อง เพื่อให้รองรับการประมวลผลภาพบิตแม็บ จากข้อมูลเอกสารในหลากหลายรูปแบบ เช่น เอกสารเอชทีเอ็มแอล (HTML) เอกสารจากโปรแกรมไมโครซอฟต์เวิร์ด หรือ เอกสารอโครเบต (Acrobat Document)

2) หากสามารถรองรับการประมวลผลภาพบิตแม็บ จากข้อมูลรูปภาพที่ปรากฏอยู่ในเอกสารด้วย น่าจะทำให้ผลของภาพบิตแม็บจากเอกสารมีความชัดเจนมากขึ้น

3) สามารถนำเอาแนวทางการประมวลผลภาพบิตแม็บจากข้อมูลเอกสาร ไปพัฒนาในการประมวลผลภาพบิตแม็บเพื่อแยกแยะลักษณะหรือประเภท กับข้อมูลประเภทอื่นๆ ยกตัวอย่าง เช่น ข้อมูลแบบไบนารี และข้อมูลโปรแกรมที่ได้จากตัวแปลโปรแกรม (Compiler) เป็นต้น