

บทที่ 6

บทสรุป

ในบทนี้กล่าวถึงสรุปผลการศึกษาในการจำแนกชนิดข้อมูลสายลำดับด้วยสเปกตรัมเคอร์เนลโดยวิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีสเปกตรัมเคอร์เนลรวมกับวิธีของเค-เนียร์เรสต์เนเบอร์ รวมถึงปัญหา อุปสรรคระหว่างการศึกษาและข้อเสนอแนะในการค้นคว้าครั้งต่อไป ดังนั้นในบทนี้จึงได้แบ่งเนื้อหาออกเป็น 3 ส่วน ดังนี้

- สรุปผลการทดลอง
 - วิธีสเปกตรัมเคอร์เนลโดยวิธีซัพพอร์ตเวกเตอร์แมชชีน
 - วิธีสเปกตรัมเคอร์เนลรวมกับวิธีของเค-เนียร์เรสต์เนเบอร์
- ข้อจำกัด
- ข้อเสนอแนะ

6.1 สรุปผลการทดลอง

6.1.1 วิธีสเปกตรัมเคอร์เนลโดยวิธีซัพพอร์ตเวกเตอร์แมชชีน

เราได้แสดงให้เห็นว่าสเปกตรัมเคอร์เนลสามารถนำไปใช้ในการจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอได้ ซึ่งเป็นปัญหาของข้อมูลระดับโมเลกุล โดยลักษณะของข้อมูลที่ใช้ในการทดสอบจะอยู่ในรูปของดีเอ็นเอ ซึ่งวิธีการที่นำเสนอนี้สามารถใช้กับข้อมูลสายลำดับได้โดยตรง และรายละเอียดของข้อมูลทั้งหมดสามารถดูได้ตามตารางที่ 4.1 ซึ่งสามารถสรุปการเปรียบเทียบประสิทธิภาพในการจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอทั้งโมเลกุลเอชแอลเอกลุ่มที่ I และเอชแอลเอกลุ่มที่ II ได้ดังนี้

จากการทดลองในการใช้สเปกตรัมเคอร์เนลที่มีขนาดของความถี่ต่างๆ ได้แก่ เค เท่ากับ 3 ถึง เค เท่ากับ 8 ซึ่งขั้นต้นเราใช้สเปกตรัมที่ความถี่ เค เท่ากับ 3 ในการจำแนกชนิดข้อมูลสายลำดับของโมเลกุลเอชแอลเอ เนื่องจากสเปกตรัมเคอร์เนลที่ความถี่ เท่ากับ 3 หรือที่เรียกว่าโคดอน นั้นเป็นคุณลักษณะที่ใช้ในการศึกษาระดับโมเลกุล นอกจากนี้ยังมีรูปแบบของคุณลักษณะที่แตกต่างกันชัดเจนระหว่างชนิดและหน้าที่การทำงาน ซึ่งเป็นลักษณะเด่นของโมเลกุลแต่ละสปีชีส์ เป็นต้น

จากนั้นเราได้เพิ่มขนาดของช่วงการพิจารณาของสเปกตรัมเคอร์เนลที่มีความถี่เท่ากับ 6 หรือที่เรียกว่า ไค-โคดอน โดยวิธีการที่เราทำการทดลองนี้จะเห็นว่าได้ผลลัพธ์เป็นที่น่าพอใจ เมื่อเปรียบเทียบกับวิธีการทันสมัยอื่น ซึ่งได้แก่ วิธีการแบบเคเอสวีเอ็มและ โครงข่ายประสาทเทียม ดังตารางที่ 4.3 โดยข้อมูลที่ใช้ในการทดลองนั้นอยู่ในรูปของลำดับเบสใน โมเลกุลเอชแอลเอ ซึ่งแสดงให้เห็นว่าเมื่อใช้สเปกตรัมเคอร์เนลที่มีความถี่เท่ากับ 6 จะได้ผลลัพธ์คือค่าความถูกต้องเท่ากับ 99.6% ซึ่งผลลัพธ์ที่ได้ดีกว่าวิธีการของ โครงข่ายประสาทเทียมและเคเอสวีเอ็ม ที่ใช้ฟังก์ชันเคอร์เนลแบบเชิงเส้น, โพลีโนเมียลและ เกาส์เซียน นอกจากนี้เราได้เปรียบเทียบประสิทธิภาพการจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอ จากค่าสัมประสิทธิ์ความสัมพันธ์เมทริกซ์ ซึ่งเห็นว่าได้ผลลัพธ์ที่ได้ ออกมาก่อนข้างดี อย่างไรก็ตามค่าความถูกต้องและค่าความแม่นยำของโมเลกุลเอชแอลเอกลุ่มที่ II จะเห็นได้ว่าวิธีการของสเปกตรัมเคอร์เนลและเคเอสวีเอ็มที่ใช้ฟังก์ชันเคอร์เนลแบบเกาส์เซียนนั้นได้ผลลัพธ์ที่ใกล้เคียงกัน แต่ในท้ายสุดเราใช้สเปกตรัมเคอร์เนลในการจำแนกชนิดข้อมูลสายลำดับของโมเลกุลเอชแอลเอกลุ่มย่อย และได้ผลลัพธ์ของค่าความถูกต้องของการจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอกลุ่มย่อยเท่ากับ 99.4% และ 99.0% สำหรับโมเลกุลในกลุ่มย่อยของโมเลกุลเอชแอลเอกลุ่มที่ I และกลุ่มย่อยของโมเลกุลเอชแอลเอกลุ่มที่ II ตามลำดับ

6.1.2 วิธีสเปกตรัมเคอร์เนลรวมกับวิธีของเค-เนียร์เรสต์เนเบอร์

จากวิธีการที่เราได้นำเสนอคือการนำวิธีสเปกตรัมเคอร์เนลรวมกับวิธีของเค-เนียร์เรสต์เนเบอร์ ซึ่งเป็นแนวคิดใหม่สำหรับการจำแนกชนิดข้อมูลประเภทอื่น โดยวิธีการดังกล่าวนี้เป็นวิธีการที่ใช้ในการแก้ปัญหของการจำแนกชนิดข้อมูลสายลำดับของโมเลกุลเอชแอลเอโมเลกุลและใช้ข้อมูลที่อยู่ในรูปแบบดีเอ็นเอได้โดยตรง ซึ่งวิธีการที่เราได้นำเสนอนี้เป็นวิธีที่เข้าใจง่าย จากผลการทดลองพบว่า วิธีการนี้มีประสิทธิภาพดีที่สุดในการจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอ ซึ่งผลลัพธ์ที่ได้คือค่าความแม่นยำและค่าความไวเท่ากับ 99.7% และ 100.0% ตามลำดับ นอกจากนี้การจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอกลุ่มย่อยนั้น ได้ผลลัพธ์คือค่าความถูกต้องเท่ากับ 99.9% ซึ่งดีกว่าวิธีการทันสมัยอื่น ดังนั้นจากการศึกษาสามารถสรุปได้ว่าสเปกตรัมเคอร์เนลสามารถดึงคุณลักษณะต่างๆ ที่สามารถบ่งบอกถึงหน้าที่และชนิดของโมเลกุลเอชแอลเอกลุ่มที่ I ได้ดี ถึงแม้ว่าการทำนายของวิธีที่เราได้นำเสนอนั้นจะดีกว่าวิธีการทันสมัยอื่น เช่นในกรณีที่ทำนายชนิดข้อมูลของโมเลกุลเอชแอลเอกลุ่มที่ I ส่วนการทำนายด้วยวิธีเคเอสวีเอ็มนั้นจะได้ผลลัพธ์ของค่าความถูกต้องดีกว่าวิธีการที่เราได้นำเสนอเล็กน้อยสำหรับการจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอกลุ่มที่ II ถึงแม้ว่าวิธีการที่เราได้นำเสนอให้ผลลัพธ์ที่แม่นยำสูง แต่อาจต้องมีการปรับปรุงในส่วนของเคอร์เนลที่มีการจับกันอย่างไม่ถูกต้องในแต่ละตำแหน่งย่อยของข้อมูลสายลำดับ

6.2 ข้อจำกัด

กระบวนการวิธีในการจำแนกชนิดข้อมูลสายลำดับของโมเลกุลเอชแอลเอที่นำเสนอขึ้น ยังมีข้อจำกัดคือสามารถจำแนกข้อมูลที่อยู่ในรูปของดีเอ็นเอได้ เท่านั้นและข้อมูลที่น่ามาวิเคราะห์ต้องมีจำนวนมากเพื่อให้ได้ตัวแบบในการจำแนกข้อมูลสายลำดับของโมเลกุลเอชแอลเอที่ถูกต้องและเชื่อถือได้

6.3 ข้อเสนอแนะ

จากข้อจำกัดในการตรวจสอบที่กล่าวไว้ข้างต้น จึงนำไปสู่แนวทางการวิจัยได้ดังนี้

- 1) เนื่องจากยังไม่ได้นำเงื่อนไขทางด้านเวลาเข้ามาทำการพิจารณา จึงต้องทำการศึกษาเพิ่มเติมเพื่อปรับปรุงกระบวนการวิธีในการจำแนกชนิดข้อมูลของโมเลกุลเอชแอลเอให้มีประสิทธิภาพมากขึ้น
- 2) ในขั้นตอนการจำแนกข้อมูลนั้นควรมีการวิเคราะห์ความซับซ้อนทางด้านเวลาและเนื้อที่ที่ใช้ในการประมวลผล เพื่อนำไปสู่การปรับปรุงขั้นตอนวิธีให้มีสมรรถนะสูงขึ้นต่อไป
- 3) กระบวนการวิธีในการจำแนกชนิดข้อมูลสายลำดับที่เรานำเสนอขึ้น สามารถนำไปใช้กับข้อมูลตัวอย่างชุดหรือชนิดอื่นได้ แต่ประสิทธิภาพในการทำนายชนิดของข้อมูลนั้นอาจได้ผลลัพธ์ที่เพิ่มขึ้นหรือลดลง ซึ่งขึ้นอยู่กับชนิดและคุณลักษณะของชุดข้อมูลนั้นๆ