

## บทที่ 5

### สรุปอภิปรายผลการศึกษาและข้อเสนอแนะ

ในบทนี้จะกล่าวถึง ข้อเสนอจากการดำเนินงานวิจัยปัญหาและอุปสรรคระหว่างการ วิจัย รวมทั้งข้อเสนอแนะ โดยมีรายละเอียดดังต่อไปนี้

#### 5.1 สรุปผลการดำเนินงานวิจัย

ในการดำเนินงานวิจัยการสกัดความสัมพันธ์แบบสตัฟฟ์จากเอกสารงานวิจัยทาง วิทยาศาสตร์ในวิทยานิพนธ์นี้เป็นการใช้หลักสถิติมาประยุกต์ใช้ในการสกัดความสัมพันธ์แบบ สตัฟฟ์ โดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ที่เป็นเนออีฟเบย์ (Naïve Bayes) ด้วย การเปรียบเทียบค่าต่างๆ ในเอกสารงานวิจัยทางวิทยาศาสตร์กับฐานข้อมูลความรู้เฉพาะทาง ด้าน วิทยาศาสตร์เคมีและชีววิทยา ซึ่งผลลัพธ์ของการสกัดความสัมพันธ์ที่ได้ขึ้นกับการเตรียมคลังข้อมูล (Corpus) ที่เหมาะสม ถูกต้องและรวมไปถึงการเลือกคุณสมบัติ ต่างๆที่ใช้พิจารณาให้เหมาะสม จะ ส่งผลให้ได้ผลลัพธ์จากระบบได้อย่างถูกต้องแล้วทำการประเมินผลเปรียบเทียบกับประเมินโดย ใช้ผู้เชี่ยวชาญเฉพาะด้านซึ่งปรากฏว่าผลลัพธ์ในส่วนของคุณค่าความถูกต้อง (precision)ยังไม่ค่อยดีนัก เนื่องจากงานวิจัยนี้ศึกษาบนแพทเทิร์นพื้นฐานแบบง่ายๆทางภาษาแบบเดียวคือ NP1 verb NP2 | NP2 verb NP1(เมื่อ NP1 เป็นนามวลีที่เป็น Chemical-Name Entity และ NP2 เป็นนามวลีที่เป็น Natural-Source-Name Entity) ดังนั้นในการสร้างคลังข้อมูลเพื่อฝึกฝนระบบจะแบ่ง คุณสมบัติต่างๆ เป็น3กลุ่มคือกลุ่มที่เป็นแนวคิดของสารเคมี กลุ่มที่เป็นแนวคิดของพืช และกลุ่มที่เป็นคำระหว่าง แนวคิดของสารเคมีและแนวคิดของพืช โดยไม่ได้สนใจกลุ่มคำที่อยู่ด้านนอกแนวคิดของสารเคมี และแนวคิดของพืช กลุ่มคำที่อยู่ระหว่างแนวคิดของสารเคมีด้วยกัน และกลุ่มคำที่อยู่ระหว่าง แนวคิดของพืชด้วยกัน

หากงานวิจัยนี้มา พัฒนาเพิ่มเติมจะสามารถนำผลลัพธ์ที่ได้ไปสร้างออนโทโลยี ของ สารผลิตภัณฑ์จากธรรมชาติ ซึ่งจะเป็นประโยชน์ต่ออุตสาหกรรมการสกัดสารเคมีจากพืช อาทิ อุตสาหกรรมยา อุตสาหกรรมผลิตภัณฑ์เสริมอาหาร เป็นต้น

## 5.2 ปัญหาและอุปสรรคจากการดำเนินงานวิจัย

ในการเตรียมคลังข้อมูลเพื่อฝึกฝนระบบ เนื่องจากจึงจำเป็นต้องแปลงข้อมูลในรูปแบบ pdf ให้อยู่ในรูปแบบข้อความก่อน ซึ่งจะมีปัญหาเรื่องสัญลักษณ์บางตัวที่ผิดเพี้ยนไป จำเป็นต้องใช้เวลา และแรงงานในการสร้างคลังข้อมูลมาก

## 5.3 ข้อเสนอแนะ

1. ระบบที่พัฒนาขึ้นนี้ เป็นการใช้เพียงรูปแบบโครงสร้างทางไวยากรณ์เดียว(NP1 verb NP2) ในการตัดสินใจเท่านั้น หากมีการศึกษาใช้รูปแบบโครงสร้างทางไวยากรณ์ที่แตกต่างออกไปมาช่วย ในการพิจารณาข้อมูลอาจทำให้ได้ประสิทธิภาพของระบบดีขึ้น
2. ในการสร้างคลังข้อมูลเพื่อฝึกฝนระบบควรนำคุณสมบัติอื่นมาพิจารณาร่วมด้วย เช่น กลุ่มคำ ที่อยู่ด้านนอกแนวคิดของสารเคมีและแนวคิดของพืช กลุ่มคำที่อยู่ระหว่างแนวคิดของสารเคมี ด้วยกัน และกลุ่มคำที่อยู่ระหว่างแนวคิดของพืชด้วยกัน อาจช่วยเพิ่มประสิทธิภาพของระบบได้
3. ควรเพิ่มขนาดของคลังข้อมูลให้มากขึ้น