

โปสเตอร์โครงงาน

Computer Science Department
Faculty of Informatics, Maharakham University

Extraction-Based Text Summarization for Cases Decided in the Supreme Court

การสรุปข้อความแบบการสกัดสาระสำคัญสำหรับ
เอกสารคดีแดงที่ถูกตัดสินในศาลฎีกา



MAHASARAKHAM
UNIVERSITY

ผู้พัฒนา : กนกวรรณ โพธิ์ชัย (Kanokwan Pochai)
จตุรรัตน์ ไทยสำโรง (Jutarat Thaisamrong)
อาจารย์ที่ปรึกษา : จันทิมา พลพิณิจ (Jantima Polpinij)



Intellect Laboratory สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
aimkanok.p@gmail.com , 61011212028@msu.ac.th , jantima.p@msu.ac.th

ที่มาและความสำคัญ

คดีความ (LAWSUIT) คือเรื่องพิพาทหรือกล่าวหากันในทางกฎหมาย ซึ่งต้องดำเนินการตามกระบวนการพิจารณาความตามที่กฎหมายกำหนด ซึ่งคดีความที่เกี่ยวข้องกับชีวิตประจำวันของผู้คนจะมีอยู่ 2 คดีคือ คดีแพ่งและพาณิชย์ (CIVIL AND COMMERCIAL CASE) และ คดีอาญา (CRIMINAL CASE) ตาม โดยทั่วไปแล้ว การตัดสินวินิจฉัยคดีความจะมี 3 ศาลคือ ศาลชั้นต้น ศาลอุทธรณ์ และศาลฎีกา คดีใดก็ตามหากมีการตัดสินในชั้นศาลฎีกาไม่ว่าผลจะออกมาเป็นเช่นไร ถือว่าคดีนั้นได้สิ้นสุดลงแล้ว สำหรับเป็นคดีที่รู้ผลแล้วว่าใครเป็นฝ่ายผิดฝ่ายถูก เนื่องจากศาลมีคำสั่งวินิจฉัยชี้ขาด หรือพิพากษา (JUDGE) แล้ว จะเรียกว่า “คดีแดง (DECIDED CASE)” สำหรับผู้คนที่สนใจด้านคดีความนั้น คดีแดงกลายเป็นแหล่งศึกษาความรู้และข้อมูลด้านกฎหมายที่สำคัญของประชาชน นักศึกษา หรือแม้แต่พนักงานเอง เพราะในการตัดสินคดีความในแต่ละคดีนั้น ผู้พิพากษาอาจจะมึมมองในตัดสินคดีความที่แตกต่างกันไป อย่างไรก็ตาม คดีที่ผ่านการตัดสินแล้วและมีการเผยแพร่ มักจะอยู่ในรูปแบบเอกสารข้อความที่ค่อนข้างยาว ทำให้ยากต่อการทำความเข้าใจ แม้ว่าเอกสารคดีแดงจะใช้คำศัพท์และรูปแบบที่เป็นมาตรฐานก็ตาม ซึ่งก็รวมถึงคดีแดงของประเทศไทยด้วย ดังนั้นโครงงานปริญญาโทฉบับนี้ จึงนำเสนอการประยุกต์เทคนิคการประมวลผลภาษาธรรมชาติ (NATURAL LANGUAGE PROCESSING: NLP) ในการสรุปประเด็นสำคัญที่อยู่ในคดีแดง

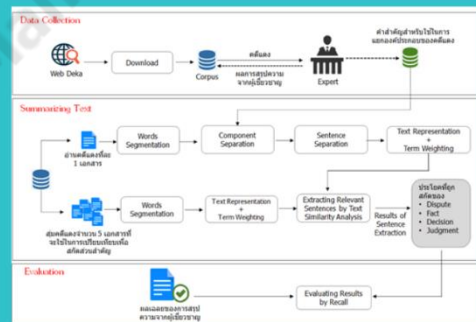
วัตถุประสงค์

นำเสนอการประยุกต์กระบวนการแบบ Extraction-Based Text Summarization ในการสรุปประเด็นสำคัญที่อยู่ในคดีแดงของคดีแพ่งและพาณิชย์ โดยเน้นคดีความเรื่องการซื้อขายที่ดิน

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้กระบวนการแบบ Extraction-Based Summarization ในการสรุปประเด็นสำคัญที่อยู่ในคดีแดงของคดีแพ่งและพาณิชย์ โดยเน้นคดีความเรื่องการซื้อขายที่ดินที่เป็นภาษาไทย
2. บุคคลทั่วไปและนักกฎหมายสามารถทำเข้าใจได้ง่ายยิ่งขึ้นจากการสรุปความภาษาไทย

กรอบการดำเนินงาน



ตัวอย่างหน้าจอโปรแกรม



สรุป

เนื่องจากการนำประโยคมาเปรียบเทียบแบบทีละ 1 ประโยคทำให้ค่า ROUGE-N และ ROUGE-L มีค่าค่อนข้างน้อย จึงได้ทำการปรับกระบวนการในการนำประโยคเข้ามาเปรียบเทียบโดยทำการเพิ่มจำนวนประโยคเป็น 5 ประโยค ซึ่งปรากฏว่าหลังจากที่ทำการเพิ่มจำนวนของประโยคแล้วทำให้ค่าของ ROUGE-N และ ROUGE-L อยู่ที่ 0.595 และ 0.513 ตามลำดับ ซึ่งทำให้เห็นว่าการทำงานของระบบมีประสิทธิภาพเพิ่มมากขึ้นหลังจากทำการปรับกระบวนการแล้ว