

การวินิจฉัยคดีด้วยเทคนิคต้นไม้ตัดสินใจ

ชัชชัย แก้วตา และ อัจฉรา มหาวิวัฒน์

ภาควิชาคณิตศาสตร์ สถิติและคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

85 ถ.สดมภ์ ค.เมืองศรีโค อ.วารินชำราบ จ.อุบลราชธานี 34190

Emails: chutchai.ubu.ac.th@gmail.com, scatchma@ubu.ac.th

บทคัดย่อ

ปัจจุบันเทคโนโลยีคอมพิวเตอร์ถูกนำมาประยุกต์ใช้ในด้านต่างๆ มากมาย แต่สำหรับด้านนิติศาสตร์ยังมีข้อจำกัดอยู่มาก เช่น การวินิจฉัยคดีซึ่งถ้าไม่มีความชำนาญก็จะทำให้เกิดความยากลำบากในการค้นหาบทบัญญัติของกฎหมายเพื่อใช้วินิจฉัยกับข้อเท็จจริงที่เกิดขึ้น และถึงแม้จะมีงานวิจัยที่เสนอแนวทางเพื่อแก้ปัญหาดังกล่าวแต่ก็ยังมีข้อจำกัดเกี่ยวกับการวินิจฉัยคดีกับข้อเท็จจริงที่เกิดขึ้น บทความนี้จะนำเสนอวิธีการวินิจฉัยคดีโดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision tree techniques) เพื่อจำแนกความคิดซึ่งถูกอธิบายด้วยชุดของคุณลักษณะ (attributes) ออกเป็นมาตราต่างๆ ที่เหมาะสมกับคดีความนั้น และเปรียบเทียบความถูกต้องของการจำแนกข้อมูลด้วยขั้นตอนวิธี ID3 กับขั้นตอนวิธี C4.5 โดยจากผลการทดลองพบว่าต้นไม้ตัดสินใจที่สร้างจากขั้นตอนวิธี C4.5 สามารถจำแนกความคิดได้ถูกต้องมากกว่าขั้นตอนวิธี ID3 และสามารถนำไปประกอบการพิจารณาวินิจฉัยคดี

คำสำคัญ – ต้นไม้ตัดสินใจ; การวินิจฉัยคดี; เหมืองข้อมูล

1. บทนำ

ปัจจุบันเทคโนโลยีคอมพิวเตอร์ได้เข้ามามีอิทธิพลต่อการดำเนินชีวิตของมนุษย์เป็นอย่างมาก โดยถูกนำมาใช้เพื่อให้เกิดความสะดวกรวดเร็วในการสื่อสาร การค้นคว้าข้อมูลและในด้านอื่นๆ อีกมากมาย ยกตัวอย่างเช่น การนำคอมพิวเตอร์มาจำลองการรับแรงสั่นสะเทือนของอาคารเมื่อเกิดแผ่นดินไหว แต่สำหรับด้านนิติศาสตร์ยังมีข้อจำกัดอยู่มาก เนื่องจากกฎหมายของประเทศไทยเป็นระบบกฎหมายแบบลายลักษณ์อักษร (Civil law) ซึ่งถูกบัญญัติไว้ในรูปแบบของประมวลกฎหมาย พระราชบัญญัติ พระราชกำหนด เป็นต้น โดยในการวินิจฉัยคดีนั้นจะต้องยึดบทบัญญัติของกฎหมายเพื่อใช้ในการวินิจฉัยข้อเท็จจริงที่เกิดขึ้น ฉะนั้นหากไม่ใช่ว่าผู้ที่มีความรู้หรือผู้เชี่ยวชาญด้านกฎหมายแล้วจะทำให้เกิดความยากลำบากในการค้นหาบทบัญญัติของกฎหมายเพื่อใช้วินิจฉัยกับข้อเท็จจริงที่เกิดขึ้นได้อย่างถูกต้อง

ดังนั้นจึงมีการเสนองานวิจัยต่างๆ เพื่อแก้ปัญหาดังกล่าว เช่น การค้นหาคำหรือวลีแบบอัตโนมัติ (Full-Text-Search) [1] ซึ่งค้นหาคำหรือวลีจากคดีความเก่าๆ ที่เก็บไว้เพื่อนำมาตรวจสอบว่าสอดคล้องกับคดีในปัจจุบันที่กำลังพิจารณาอยู่หรือไม่ การจัดสรรกฎหมายที่เหมาะสมให้กับคดีความแบบอัตโนมัติ [2] ซึ่งใช้เทคนิคการจำแนกข้อมูล (Data Classification) เพื่อตัดคำจากคดีความแล้วนำคำที่ได้มาทำนายกฎหมายให้กับคดีความ และการตัดคำด้วยพจนานุกรม [3] โดยใช้การค้นหาความสัมพันธ์ (Association Rule Discovery) ในการสร้างพจนานุกรมสำหรับตัดคำที่มีความหมายเพื่อจัดสรรกฎหมายที่เหมาะสมให้กับคดีความ อย่างไรก็ตามงานวิจัยดังกล่าวยังมีข้อจำกัดทางด้านความสามารถในการจัดสรรหรือทำนายกฎหมายให้กับคดีความ เนื่องจากเป็นเพียงการจำแนกคำหรือวลีของคดีความออกมา เพื่อค้นหาว่าคำหรือวลีนั้นน่าจะอยู่ในกฎหมายฉบับใดและมาตราใดเท่านั้น จึงยังไม่ใช่วิธีของหลักการพิจารณาหรือวินิจฉัยคดีที่ถูกต้อง โดยหลักที่ถูกต้องคือควรพิจารณาถ้อยคำหรือวลีและข้อเท็จจริงทั้งหมดของคดีความนั้นด้วย เพราะถ้าเลือกอย่างใดอย่างหนึ่งแล้วจะทำให้ข้อเท็จจริงเปลี่ยนไปและเมื่อข้อเท็จจริงเปลี่ยนจะทำให้การวินิจฉัยหรือการพิจารณาเปลี่ยนไปด้วยจนส่งผลให้ฐานความคิดเปลี่ยนไปในที่สุด

บทความนี้จึงศึกษาปัจจัยที่มีผลต่อการพิจารณาข้อเท็จจริงสำหรับระบบที่จะได้รับของผู้กระทำความผิด เพื่อสร้างและนำเสนอแบบจำลองสำหรับสนับสนุนการวินิจฉัยคดีโดยจำแนกคดีออกเป็นกลุ่มตามมาตราด้วยเทคนิคต้นไม้ตัดสินใจ (Decision tree technique) และเชื่อมโยงความสัมพันธ์ของคุณลักษณะ (Attribute) ขององค์ประกอบความคิดเข้ากับบทบัญญัติของกฎหมายที่อยู่ในรูปของประมวลกฎหมาย

ส่วนต่อจากนี้ไปจะประกอบด้วยส่วนที่ 2 ซึ่งอธิบายถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 เป็นวิธีการดำเนินการวิจัยซึ่งจะประกอบด้วยขั้นตอนการเก็บรวบรวมข้อมูล การเตรียมข้อมูล จนถึงขั้นตอนการสร้างแบบจำลองการวินิจฉัยคดี ส่วนที่ 4 เป็นผลการทดลองและสุดท้ายคือส่วนที่ 5 ซึ่งเป็นบทสรุปและข้อเสนอแนะ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. ความรู้เกี่ยวกับการวินิจฉัยคดี

การพิจารณาว่าการกระทำของผู้ใดเป็นความผิดอาญาและผู้นั้นจะต้องรับผิดชอบอาญาหรือไม่นั้น จะต้องวินิจฉัยจากหลักเกณฑ์ 3 ประการซึ่งถือว่าเป็นองค์ประกอบในการวินิจฉัยคดีอาญา [4] ดังต่อไปนี้

1. การกระทำของผู้นั้น “ครบองค์ประกอบความผิดอาญา” หรือไม่ หลักเกณฑ์นี้เป็นไปตามหลัก “ไม่มีความคิด ไม่มีโทษ โดยไม่มีกฎหมาย” (nullum crimen, nulla poena sine lege) ซึ่งบัญญัติไว้ในประมวลกฎหมายอาญา มาตรา 2 กล่าวคือ บุคคลจะต้องรับโทษทางอาญาต่อเมื่อได้กระทำการอันกฎหมายที่ใช้ในขณะกระทำการนั้น บัญญัติเป็นความผิดและกำหนดโทษไว้

2. เหตุยกเว้นความผิด (Justification)

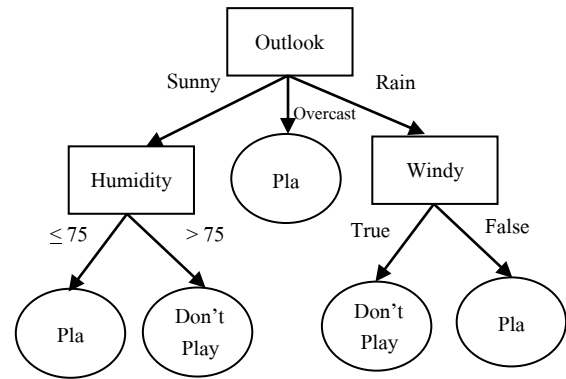
เหตุยกเว้นความผิด เป็นเหตุที่หากมีขึ้นแล้วจะทำให้การกระทำที่ครบองค์ประกอบความผิดแล้วไม่มีความผิดเลย หรืออีกนัยหนึ่งคือ ผู้กระทำการนั้นๆ มีอำนาจที่จะกระทำได้ โดยไม่มีความผิด (Justification)

3. เหตุที่ผู้กระทำไม่ต้องรับโทษ (Excuse)

การกระทำบางอย่าง แม้จะไม่มีเหตุยกเว้นความผิดก็ตาม แต่กฎหมายก็อาจไม่ลงโทษผู้กระทำก็ได้ในบางกรณีแม้จะเป็นความผิดก็ตาม

2.2. ต้นไม้ตัดสินใจ (Decision Tree)

กระบวนการทางด้านเหมืองข้อมูลนั้น (Data mining) ได้นำการตัดสินใจแบบโครงสร้างต้นไม้มาช่วยในการทำงานด้านการตัดสินใจต่างๆ ไม่ว่าจะเป็นทางด้านระบบธุรกิจหรือด้านอื่น ๆ โดยปกติมักประกอบด้วยกฎในรูปแบบ “ถ้า เงื่อนไข แล้ว ผลลัพธ์” เช่น “If Income = High and Married = No THEN Risk = Poor” โดยลักษณะของการตัดสินใจแบบโครงสร้างต้นไม้มีลักษณะคล้ายกับต้นไม้กลับหัวโดยโหนดแรกสุดจะเป็นรากของต้นไม้ (Root node) แต่ละโหนดแสดงคุณลักษณะ (attribute) แต่ละกิ่งจะแสดงค่าผลในการทดสอบ และโหนดใบ (Leaf node) แสดงคลาสที่กำหนดไว้ ดังรูปที่ 1



รูปที่ 1. ตัวอย่างต้นไม้ตัดสินใจ (Decision Tree)

2.2.1. ขั้นตอนวิธี ID3 (ID3 Algorithm)

ขั้นตอนวิธี ID3 [5] เป็นขั้นตอนวิธีที่ใช้ในการสร้างต้นไม้ตัดสินใจ โดยนำหลักการของทฤษฎีข่าวสารมาใช้ ค่าที่วัดได้จะนำมาตัดสินใจว่าจะใช้ตัวแปรใดในการแบ่งข้อมูลโดยวิธีการกำหนดโครงสร้างต้นไม้ตัดสินใจจะเป็นการเลือกข้อมูลตามลำดับของตัวชี้วัดหรือค่าเกน (Gain) สูงที่สุดเป็นข้อมูลเริ่มต้นและข้อมูลถัดไปที่มีค่าลดหลั่นกันตามลำดับ ตัวอย่างเช่น การพิจารณาจากกลุ่มข้อมูล 2 คลาสคือ P และ N โดยจำนวนตัวอย่างในคลาส P คือ p ตัว และจำนวนตัวอย่างในคลาส N คือ n ตัว ส่วนค่าของกลุ่มข้อมูลคือค่าคาดคะเนที่กลุ่มตัวอย่างต้องใช้จำนวนบิตในการแยกคลาส P และ N โดยนิยามตามสมการที่ 1

$$I(p,n) = -\frac{p}{p+n} \log_2 \left(\frac{p}{p+n} \right) - \frac{n}{p+n} \log_2 \left(\frac{n}{p+n} \right) \quad (1)$$

ค่าคาดคะเนของข้อมูล (Entropy) เป็นค่าที่แยกโดยการใช้ลักษณะประจำ A ซึ่งกำหนด A คือ ลักษณะประจำที่แบ่ง S ออกเป็น $\{S_1, S_2, \dots, S_v\}$ โดยให้ S_i มีตัวอย่างจากคลาส P จำนวน P_i และตัวอย่างจากคลาส N จำนวน n_i ดังสมการที่ 2

$$E(A) = \sum_{i=1}^v \frac{P_i + n_i}{p+n} I(P_i, n_i) \quad (2)$$

ดังนั้นค่าเกนข้อมูล (Data Gain) ที่ได้จากการแยกข้อมูลด้วยลักษณะประจำ A จะได้ดังสมการที่ 3

$$\text{Gain}(A) = I(p,n) - E(A) \quad (3)$$

2.2.2. ขั้นตอนวิธี C4.5 (C4.5 Algorithm)

ขั้นตอนวิธี C4.5 เป็นส่วนขยายของขั้นตอนวิธี ID3 ถูกพัฒนาโดย Ross Quinlan [6] ใช้ในการสร้างต้นไม้ประกอบการตัดสินใจสำหรับการจำแนกประเภทข้อมูล มีการใช้ค่าเกน (Gain) และค่าเอนโทรปีของข้อมูล (Entropy) เช่นเดียวกับ ID3 แต่จะมีส่วนที่เพิ่มเติมจากขั้นตอนวิธี ID3 ที่สำคัญดังนี้

1. สามารถใช้งานได้ทั้งข้อมูลที่มีคุณลักษณะแบบต่อเนื่อง (Continuous) และแบบไม่ต่อเนื่อง (Discrete) โดยในส่วนของข้อมูลแบบต่อเนื่อง ขั้นตอนวิธี C4.5 จะสร้างจุดเริ่ม (Threshold) และแยกคุณลักษณะนั้นออกเป็น 2 ส่วนคือส่วนที่มีค่ามากกว่าและน้อยกว่าและเท่ากับค่าที่ใช้ในการสร้างจุดเริ่ม
2. สามารถใช้กับข้อมูลฝึกสอน (Training Data) ที่ไม่มีค่าของคุณลักษณะได้ โดยจะทำการทำเครื่องหมายในคุณลักษณะนั้นเป็น “ ? ” และไม่นำค่านั้นมาคำนวณหาค่าเอนโทรปีของข้อมูล (Entropy)
3. สามารถใช้งานได้กับค่าที่มีความผิดปกติหรือมีความเสียหาย
4. สามารถทำการปรับแต่งต้นไม้ประกอบการตัดสินใจ (Pruning Trees) ในขณะสร้างได้

2.3. งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยที่ช่วยในการวินิจฉัยคืออย่างหลากหลาย โดยงานเหล่านี้มักเป็นการจัดสรรกฎหมายที่เหมาะสมให้กับคดีความโดยอัตโนมัติ [2] ด้วยเทคนิคของการทำเหมืองข้อมูลเช่น เทคนิคการจำแนกข้อมูล (Data classification) และการตัดคำเพื่อสร้างพจนานุกรมขึ้นมาแล้วสร้างตัวทำนายผลกฎหมายโดยใช้เทคนิคการค้นหากฎความสัมพันธ์ (Association Rule Discovery) [3] นอกจากนี้ยังมีการนำเทคนิคของการทำเหมืองข้อมูลมาประยุกต์ใช้ในงานอื่นเช่น การพัฒนาคุณภาพการศึกษานิสิตคณะวิศวกรรมศาสตร์ [7] ซึ่งใช้เทคนิคสำคัญ 3 ประการ ได้แก่ การค้นหากฎความสัมพันธ์ (Association rule discovery) การจำแนกข้อมูล (Data classification) การพยากรณ์ข้อมูล (Data prediction) เพื่อช่วยนิสิตเลือกสาขาที่เหมาะสม และทำนายเกรดแต่ละรายวิชาในภาคการศึกษาต่อไป โดยใช้เทคนิคต้นไม้ตัดสินใจ นอกจากนี้ยังมีการประยุกต์ใช้กับการพิจารณาอนุมัติให้สินเชื่อเพื่อการเช่าซื้อสินค้าอีกด้วย [8]

3. วิธีการดำเนินการวิจัย

3.1. ขั้นตอนการเก็บรวบรวมข้อมูล

ข้อมูลที่ใช้ในการทดลองคือข้อมูลจากประมวลกฎหมายอาญาลักษณะ 12 เรื่องความผิดเกี่ยวกับทรัพย์เนื่องจากเป็นคดีความสำคัญที่เกิดขึ้นบ่อยครั้งในชีวิตประจำวันและเพื่อศึกษาองค์ประกอบของการพิจารณาคดีความได้ครอบคลุมครบถ้วน โดยข้อมูลประกอบด้วยมาตราทั้งหมด 34 มาตราและ

ข้อมูลจากคำพิพากษาศาลฎีกา โดยคัดเลือกเฉพาะคดีที่ศาลฎีกาได้มีคำพิพากษาความคตินั้นแล้ว จากนั้นได้กำหนดคุณลักษณะ (attribute) ของข้อมูลที่ได้เพื่อใช้สำหรับอธิบายลักษณะของมาตราต่างๆ และสร้างต้นไม้ตัดสินใจ โดยสามารถกำหนดได้ทั้งหมด 55 คุณลักษณะและค่าข้อมูลที่เป็นไปได้ของคุณลักษณะจำนวน 187 ค่าข้อมูล นอกจากนี้ยังกำหนดรหัสแทนคุณลักษณะด้วยตัวอักษรภาษาอังกฤษเรียงตามตัวอักษรและค่าข้อมูลด้วยอักษรภาษาอังกฤษของคุณลักษณะกับตัวเลขเพื่อใช้สำหรับการประมวลผลข้อมูล ดังตัวอย่างในตารางที่ 1

ตารางที่ 1. คุณลักษณะและค่าที่เป็นไปได้เมื่อแทนด้วยรหัส

รหัส	คุณลักษณะ/ ค่าที่เป็นไปได้
A	ผู้กระทำความผิด
	A0 = ไม่ระบุ A1 = 1 คน A2 = 2 คน A3 = 3 คนขึ้นไป
B	ทรัพย์สินของผู้อื่นหรือผู้อื่นเป็นเจ้าของรวมอยู่ด้วย
	Yes = ใช่ No = ไม่ใช่
C	ลักษณะการครอบครองทรัพย์สิน
	C0 = ไม่ระบุ C1 = ได้ทรัพย์สินมาอยู่ในความครอบครองโดยเด็ดขาด C2 = ได้ทรัพย์สินมาอยู่ในความครอบครองโดยชั่วคราว
D	ลักษณะการกระทำความผิด
	D0 = ไม่ระบุ D1 = เอาทรัพย์สินไปโดยเจ้าของรู้สีกตัว D2 = เอาทรัพย์สินไปโดยเจ้าของไม่รู้สีกตัว D3 = หลอกลวงด้วยการแสดงข้อความอันเป็นเท็จ D4 = เข้าไป D5 = ปกปิดข้อความจริงซึ่งควรบอกให้แจ้ง
E	ลักษณะการเอาทรัพย์สินนั้นไป
	E0 = ไม่ระบุ E1 = ใช้กำลังประทุษร้าย E2 = ขู่ว่าจะใช้กำลังประทุษร้าย
AS	ผู้ถูกกระทำ
	AS0 = ไม่ระบุ AS1 = กระทำต่อบุคคลหรือกลุ่มคนโดยเฉพาะเจาะจง AS2 = กระทำต่อประชาชนหรือสาธารณชน AS3 = มีจิตอ่อนแอหรือเป็นเด็กเบาปัญญาและไม่สามารถเข้าใจในสาระสำคัญในการกระทำของตน

3.2. ขั้นตอนการเตรียมข้อมูล

หลังจากที่แปลงข้อมูลคุณลักษณะและค่าข้อมูลที่เป็นไปได้ของแต่ละคุณลักษณะให้อยู่ในรูปของสัญลักษณ์แล้ว ขั้นตอนต่อไปเป็นการจัดข้อมูลให้เข้ากับรูปแบบของคุณลักษณะที่ได้กำหนดไว้เพื่ออธิบายความสัมพันธ์ระหว่างคุณลักษณะกับมาตรา ดังตัวอย่างการจัดและการแปลงข้อมูลในตารางที่ 2 – 4

ตารางที่ 2. การกำหนดคุณลักษณะให้กับข้อมูล

คุณลักษณะ	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 1)	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 2)	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 3)
จำนวนผู้กระทำความผิด	2 คน	3 คน	ไม่ระบุ
ลักษณะการครอบครองทรัพย์สิน	ไม่ระบุ	ได้ทรัพย์สินมาอยู่ในความครอบครองโดยเด็ดขาด	ไม่ระบุ
ลักษณะการกระทำความผิด	เข้าไป	ไม่ระบุ	ปกปิดข้อความจริงซึ่งควรบอกให้แจ้ง
ลักษณะเอาทรัพย์สินไป	ไม่ระบุ	ใช้กำลังประทุษร้าย	ไม่ระบุ
มีเจตนา	ใช่	ใช่	ใช่
เพื่อแสวงหาประโยชน์	ไม่ใช่	ใช่	ใช่
ผู้ถูกกระทำ	ไม่ระบุ	ไม่ระบุ	กระทำต่อประชาชนหรือสาธารณชน
มีความผิดตามมาตรา	365 อนุ 2	340 วรรค 1	343 วรรค 1

ตารางที่ 3. การกำหนดสัญลักษณ์คุณลักษณะให้กับข้อมูล

คุณลักษณะ	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 1)	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 2)	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 3)
A	A2	A3	A0
C	C0	C1	C0
D	D12	D0	D3

ตารางที่ 4. การกำหนดสัญลักษณ์คุณลักษณะให้กับข้อมูล (ต่อ)

คุณลักษณะ	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 1)	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 2)	ค่าข้อมูลที่เป็นไปได้ (กรณีที่ 3)
E	E0	E1	E0
F	Yes	Yes	Yes
G	No	Yes	Yes
AS	AS0	AS0	AS2
Result	M365A2	M340W1	M343W1

3.3. การสร้างแบบจำลองสำหรับวินิจฉัยคดี

โปรแกรม Weka เวอร์ชัน 3.6.2 ถูกใช้เพื่อสร้างต้นไม้ตัดสินใจด้วยขั้นตอนวิธี ID3 และ C4.5 โดยมีการแบ่งข้อมูลแบบ cross-validation ซึ่งเป็นการแบ่งข้อมูลแบบสุ่มออกเป็นกลุ่ม กลุ่มละเท่าๆ กัน จากนั้นทำการดึงข้อมูลชุดหนึ่งออกมาสำหรับการทดสอบ (Test) และจะวนไปจนครบทุกกลุ่ม เช่น มีข้อมูลอยู่ 100 แถว ทดสอบแบบ Cross-validation 10 Folds จะเป็นการแบ่งข้อมูลออกเป็น 10 กลุ่มๆละ 10 ชุดข้อมูล แล้วทำการทดสอบ 10 รอบ โดยรอบที่ 1 เป็นการเอาข้อมูลกลุ่มที่ 1 ออกเพื่อใช้ทดสอบ (Test) และใช้กลุ่มข้อมูลที่ 2-20 ในการสอน (Train) รอบที่ 2 เป็นการเอาข้อมูลกลุ่มที่ 2 ออกเพื่อใช้ทดสอบ (Test) และใช้กลุ่มข้อมูลที่ 1, 3-20 ในการสอน (Train) โดยจะทำแบบนี้ไปจนครบ 10 รอบ ทั้งนี้ได้เลือกวิธี 10 Folds คือการแบ่งข้อมูลออกเป็น 10 กลุ่มๆ ละเท่าๆ กัน จากนั้นจะได้กฎสำหรับสร้างต้นไม้ตัดสินใจซึ่งอยู่ในรูปแบบของกฎ IF Then ในการระบุมตราจากกิ่งในต้นไม้ตัดสินใจ ดังตัวอย่างวิธีการระบุมตราจากกิ่งในต้นไม้ตัดสินใจด้วยกฎ IF Then ดังนี้

กฎที่ 1 IF G = Yes

AND D = D3

AND AS = AS3 THEN M343W1

หมายความว่า ถ้าเอาทรัพย์สินไปแสวงหาประโยชน์ เท่ากับ ใช่ และลักษณะการกระทำความผิด เท่ากับ หลอกลวงด้วยการแสดงข้อความอันเป็นเท็จ และ ผู้ถูกกระทำ เท่ากับ มีจิตอ่อนแอหรือเป็นเด็กเบาปัญญาและไม่สามารถเข้าใจในสาระสำคัญในการกระทำของตน คำตอบคือ มีความผิดตามมาตรา 343 วรรค 1

กฎที่ 2 IF G = No

AND H = H0

AND D = D0

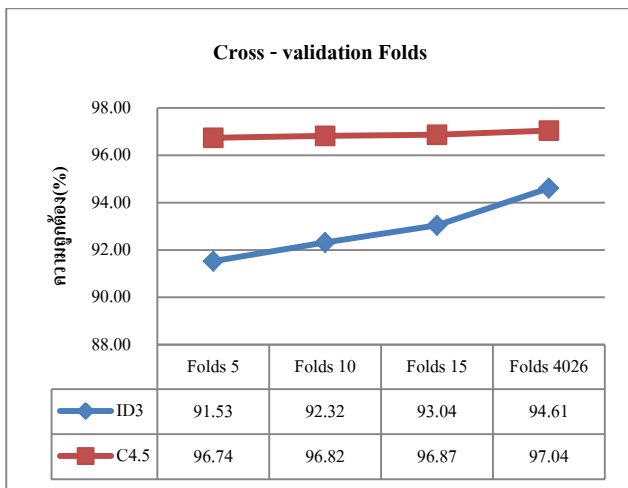
AND AE = Yes

AND E = E2 THEN M338

หมายความว่า ถ้าเอาทรัพย์สินไปแสวงหาประโยชน์เท่ากับ ไม่ใช่ และ
มูลเหตุจูงใจ ไม่ระบุ และ ลักษณะการกระทำความผิดเท่ากับ ไม่ระบุ และ
การข่มขืนใจผู้อื่น เท่ากับ ใช่ และลักษณะการเอาทรัพย์สินไป เท่ากับ ขู่ว่าจะ
ใช้กำลังประทุษร้าย คำตอบคือ มีความผิดตาม มาตรา 338

4. ผลการทดลอง

การสร้างแบบจำลองวินิจัยคดีได้ใช้ข้อมูลในการทดลองจำนวน 4026
ความคิดและเลือกวิธี Cross-validation จำนวน Folds 5, 10, 15, 4026
ตามลำดับ โดยเปรียบเทียบขั้นตอนวิธี ID3 และ C4.5 และให้ผลการ
ทดลองดังแสดงในรูปที่ 2 ซึ่งพบว่าขั้นตอนวิธี C4.5 ให้ค่าความถูกต้องใน
การจำแนกข้อมูลในชุดทดสอบสูงกว่าขั้นตอนวิธี ID3 โดยร้อยละความ
ถูกต้องของการจำแนกข้อมูลนี้จะเพิ่มขึ้นตามจำนวน Folds และถึงแม้ว่า
การแบ่งข้อมูลแบบ Folds 4026 จะให้ค่าความถูกต้องสูงที่สุด แต่ก็ต่างจาก
ค่าความถูกต้องในการจำแนกข้อมูลของการแบ่งข้อมูลแบบ Fold 10 และ
15 เล็กน้อยเมื่อเทียบกับระยะเวลาในการฝึกสอนสำหรับการแบ่งข้อมูล
แบบ Folds 4026 ที่มากกว่าแบบ Fold 10 และ 15 หลายเท่าตัว และ
เช่นเดียวกันเมื่อเปรียบเทียบระหว่างแบบ Fold 10 กับ Fold 15 ดังนั้นจึง
เลือกการแบ่งข้อมูลแบบ Fold 10 สำหรับสร้างต้นไม้ตัดสินใจ



รูปที่ 2. การเปรียบเทียบขั้นตอนวิธี ID3 กับ C4.5 ใน Cross-validation Folds

และเมื่อพิจารณาถึงโครงสร้างต้นไม้และผลการจำแนกคดีสามารถสรุปได้
ดังนี้

1. ลักษณะความคิดที่ขั้นตอนวิธี ID3 และ C4.5 ระบุฐานความคิดได้ถูกต้อง
จะพบว่าโครงสร้างต้นไม้ของลักษณะความคิดมีความคล้ายคลึงกันคือใช้

คุณลักษณะแบบเดียวกันมาสร้างโหนดเพื่อหาคำตอบ และมีบาง
คุณลักษณะที่ขั้นตอนวิธี C4.5 นำมาใช้หาคำตอบ แต่ขั้นตอนวิธี ID3 ไม่ได้
ใช้ ซึ่งทำให้กิ่งของต้นไม้จาก C4.5 ยาวกว่าและอาจส่งผลให้มีจำนวนครั้ง
ในการเปรียบเทียบมากกว่า แต่เมื่อพิจารณามาตราอื่นๆ แล้วพบว่า
คุณลักษณะหรือโหนดที่เพิ่มขึ้นของต้นไม้ C4.5 นั้นเป็นการสร้างความ
สัมพันธ์เพื่อเชื่อมโยงไปสู่คำตอบที่ถูกต้องในฐานความคิดอื่นๆ ซึ่ง
ต้นไม้ ID3 ยังให้คำตอบที่ผิด นอกจากนี้เมื่อพิจารณาในภาพรวมแล้วต้นไม้
ตัดสินใจ C4.5 มีขนาดเล็กกว่าต้นไม้ตัดสินใจ ID3

2. ลักษณะความคิดที่ทั้งขั้นตอนวิธี ID3 และ C4.5 ระบุฐานความคิดไม่
ถูกต้อง จะพบว่าโครงสร้างต้นไม้ ID3 มีขนาดยาวกว่าต้นไม้ C4.5 และเมื่อ
นำต้นไม้จากทั้งสองวิธีมาเปรียบเทียบกับองค์ประกอบความคิดตามหลัก
ของกฎหมาย ปรากฏว่าคุณลักษณะที่นำมาประกอบการพิจารณานั้นไม่
ครบองค์ประกอบความคิดตามหลักของกฎหมาย และเป็นสาเหตุของความ
ผิดพลาด

3. ลักษณะความคิดที่ขั้นตอนวิธี ID3 ระบุฐานความคิดได้ถูกต้องแต่ C4.5
ระบุฐานความคิดไม่ถูกต้อง จะพบว่ามี 5 มาตราที่ขั้นตอนวิธี ID3 ระบุได้
ถูกต้องแต่จะต้องใช้โหนดยาวถึง 21 โหนด แต่เมื่อเทียบกับหลักกฎหมาย
แล้วใช้แค่ 4 คุณลักษณะก็สามารถระบุฐานความคิดได้ ซึ่งแสดงให้เห็นว่า
ฐานความคิดเหล่านี้อาจมีความคลุมเครือจนทำให้ยากที่จะจำแนกข้อมูลได้

4. ลักษณะความคิดที่ขั้นตอนวิธี ID3 ระบุฐานความคิดไม่ถูกต้องแต่ C4.5
ระบุฐานความคิดได้ถูกต้อง จะพบว่าขั้นตอนวิธี C4.5 ใช้เพียง 2 โหนด
(หรือคุณลักษณะ) ก็สามารถระบุมาตราได้ เมื่อเทียบกับหลักกฎหมายแล้ว
ใช้ 6 คุณลักษณะก็จะระบุฐานความคิดได้ แสดงว่าขั้นตอนวิธี C4.5
สามารถจำแนกฐานความคิดเหล่านี้ได้ดีกว่าทั้งขั้นตอนวิธี ID3 และการ
เทียบกับหลักกฎหมาย

5. บทสรุปและข้อเสนอแนะ

บทความนี้ได้นำเสนอแบบจำลองสำหรับการวินิจัยคดี โดยประยุกต์ใช้
เทคนิคต้นไม้ตัดสินใจ (Decision tree techniques) ซึ่งได้เปรียบเทียบสอง
ขั้นตอนวิธีคือ ID3 และ C4.5 โดยพบว่าแบบจำลองที่ดีที่สุดคือต้นไม้
ตัดสินใจที่สร้างด้วยขั้นตอนวิธี C4.5 ซึ่งให้คำตอบที่ถูกต้อง 96.82% และ
เพื่อเพิ่มประสิทธิภาพของการวินิจัยคดีให้ดียิ่งขึ้น ในการวิจัยครั้งต่อไป
ควรดำเนินการดังนี้

1. กำหนดคุณลักษณะ (Attribute) เพิ่มขึ้นเพื่อให้สามารถจำแนกข้อมูลได้
อย่างถูกต้อง
2. ใช้ขั้นตอนวิธีอื่นมาปรับใช้ เช่น ขั้นตอนวิธี C5.0
3. เพิ่มข้อมูลเกี่ยวกับกฎหมายในลักษณะอื่นๆ

6. กิตติกรรมประกาศ

บทความนี้ได้รับการสนับสนุนด้านข้อมูลเกี่ยวกับกฎหมายจาก อาจารย์
อภิรักษ์ ศรีศิริ และอาจารย์มณฑิรา เดียประเสริฐ อาจารย์ประจำคณะ
นิติศาสตร์ มหาวิทยาลัยอุบลราชธานี

เอกสารอ้างอิง

- [1] Witten, I.H., A. Moffat and T.C. Bell. **Managing Gigabyte :
Compressing and Indexing Documents and Images**. Morgan
Kaufmann Publishers, San Francisco. 600 p, 1999.
- [2] T. Pongsiripreeda and K. Waiyamai. **Using Data Mining technique
to select the laws and articles for lawsuits (in thai language)**. The
National Computer Science and Engineering Conference (NCSEC'2000).
Thailand : Bangkok, November.
- [3] ชีระวัฒน์ พงษ์ศิริปริดา. **การจัดสรรกฎหมายที่เหมาะสมให้กับคดีความ
โดยอัตโนมัติ**. วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต ภาควิชาวิศวกรรม
คอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์, 2544.
- [4] สุรศักดิ์ ลิขสิทธิ์วัฒนกุล. **คำอธิบายความผิดเกี่ยวกับทรัพย์ ตาม
ประมวลกฎหมายอาญา**. พิมพ์ครั้งที่ 4. กรุงเทพฯ : สำนักพิมพ์วิญญูชน,
หน้า 13-24, 2548
- [5] J. R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol.
1, no. 1, pp. 81-106, 1986.
- [6] J. R. Quinlan. *C4.5: program for machine learning*, California :
Morgan Kaufmann, 1993.
- [7] กฤษณะ ไวยมัย, ชิดชนก ส่งศิริ และธนาวินท์ รักธรรมานนท์. **การใช้
เทคนิคดาต้าไมนิ่งเพื่อพัฒนาคุณภาพการศึกษานิติคณะ
วิศวกรรมศาสตร์**. The Nectec Technical Jurnal.2001; Vol. III (NO.11).
- [8] มลธิดา ฤทธิ์สมบูรณ์ และสุชา สมานชาติ. "การพัฒนาระบบสนับสนุน
การพิจารณาอนุมัติให้สินเชื่อเพื่อการเช่าซื้อสินค้าโดยใช้เทคนิคต้นไม้
ตัดสินใจ." **เทคโนโลยีสารสนเทศ** ปีที่ 4, ฉบับที่ (7 มกราคม-มิถุนายน
2551).