บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ ผู้วิจัยได้กรอบแนวคิดจากการศึกษาค้นคว้าทฤษฎีและเอกสารที่เกี่ยวข้องซึ่งจะเสนอตามลำดับดังนี้

1. แนวทางการเทียบโอนผลการเรียนรู้
2. ระบบสนับสนุนการตัดสินใจ (Decision Support System)
3. การสร้างต้นไม้ตัดสินใจ (Decision Tree)
4. งานวิจัยที่เกี่ยวข้อง

แนวทางการเทียบโอนผลการเรียนรู้

สืบเนื่องจากพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และที่แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 มาตรา 15 กำหนดให้การจัดการศึกษามีสามรูปแบบ คือ การศึกษาในระบบ การศึกษานอกระบบ และการศึกษาตามอัธยาศัย และให้ผู้เรียนสามารถนำผลการเรียนที่สะสมไว้มาเทียบโอนในระหว่างรูปแบบเดียวกัน หรือต่างรูปแบบได้ไม่ว่าจะเป็นผลการเรียนจากสถานศึกษาเดียวกันหรือไม่ก็ตาม รวมทั้งจากการเรียนรู้นอกระบบ และอัธยาศัยจากการฝึกอาชีพ ประสบการณ์ชีวิต หรือจากประสบการณ์การทำงาน ประกอบกับปัจจุบันโลกไร้พรมแดน เป็นยุคข้อมูลข่าวสาร เทคโนโลยีสภาพของสังคมและเศรษฐกิจที่พัฒนาไปอย่างรวดเร็ว ทำให้ประชาชน นักเรียน นักศึกษา มีโอกาสหาความรู้ได้หลากหลายวิธี มีแหล่งให้ศึกษาหาความรู้ มีอุปกรณ์ช่วยในการศึกษาหลากหลาย การเรียนรู้เกิดได้ทุกสถานการณ์ เช่น การทำงาน การพบปะสังสรรค์ การเข้าร่วมกิจกรรม ชมรม กีฬา อาสาสมัคร กิจกรรมทางศาสนาและสังคม ซึ่งการเรียนรู้ที่เกิดนอกชั้นเรียนดังกล่าวจึงไม่เป็นมาตรฐานเดียวกัน การเทียบโอนความรู้และประสบการณ์เข้าสู่หน่วยกิตจะช่วยให้ผู้ประกอบอาชีพสามารถพัฒนาสมรรถนะตนจนได้วุฒิ ปวช. หรือ ปวส. ที่มีมาตรฐานแต่ละสาขาวิชาอย่างชัดเจนในระยะเวลาที่น้อยกว่านักเรียนนักศึกษาปกติที่ศึกษาในระบบเพียงอย่างเดียว

การเทียบโอนความรู้และประสบการณ์ เป็นการยอมรับผลการเรียนรู้ ที่ได้จากประสบการณ์ ดังนั้นการจัดการศึกษาตามหลักสูตรประกาศนียบัตรวิชาชีพ (ปวช.) และ

หลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ของสถาบันการอาชีวศึกษา สังกัดสำนักงานคณะกรรมการการอาชีวศึกษา จึงเปิดโอกาสให้นักเรียน นักศึกษา ประชาชนสามารถนำความรู้และประสบการณ์มาขอรับการประเมินเทียบโอน เพื่อยกเว้นการเรียนรายวิชาในสถานศึกษาที่เข้าเรียน และนับเป็นส่วนหนึ่งของผลการเรียนตามหลักสูตร ตามเจตนารมณ์ของพระราชบัญญัติการศึกษาแห่งชาติดังกล่าว

การเทียบโอนผลการเรียนรู้ สามารถดำเนินการได้ 2 ลักษณะ คือ การเทียบโอนความรู้และประสบการณ์จากการศึกษานอกระบบและตามอัธยาศัย และการเทียบโอนผลการเรียนรายวิชา/กลุ่มวิชาจากการศึกษาในระบบและนอกระบบ

การเทียบโอนความรู้และประสบการณ์เข้าสู่หน่วยกิตตามหลักสูตรประกาศนียบัตรวิชาชีพ (ปวช.) และหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) สามารถดำเนินการได้ทุกหมวดวิชา โดยเทียบโอนได้ไม่เกิน 2 ใน 3 ของจำนวนหน่วยกิตตามโครงสร้างหลักสูตร สำหรับรายวิชาที่เรียนจากหลักสูตรอื่นของกระทรวงศึกษาธิการที่อยู่ในระดับเดียวกันหรือเทียบเท่า และยังไม่จบหลักสูตร สามารถนำผลการเรียนมาเทียบโอนได้โดยปฏิบัติตามระเบียบกระทรวงศึกษาธิการว่าด้วยการจัดการศึกษาและการประเมินผลการเรียนตามหลักสูตรประกาศนียบัตรวิชาชีพ (ปวช.) และประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ที่กำหนดไว้

หลักเกณฑ์การขอประเมินเทียบโอนความรู้และประสบการณ์

1. คุณสมบัติของผู้ขอรับการประเมิน

1.1 เป็นนักเรียน นักศึกษา ในสถานศึกษาที่เปิดสอนตามหลักสูตรประกาศนียบัตรวิชาชีพ (ปวช.) หรือหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.)

1.2 มีพื้นความรู้และประสบการณ์ในรายวิชา หรือกลุ่มวิชาที่ขอประเมิน โดยมีหลักฐานที่แสดงถึงความรู้และประสบการณ์

2. เงื่อนไขการขอรับการประเมินเทียบโอนความรู้และประสบการณ์เข้าสู่หน่วยกิต ให้ปฏิบัติตามระเบียบกระทรวงศึกษาธิการ ว่าด้วยการจัดการศึกษาหลักสูตรประกาศนียบัตรวิชาชีพ(ปวช.) หรือหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ดังนี้

1.2.1 นักเรียน/นักศึกษาลงทะเบียนเพื่อประเมินเทียบโอนความรู้และประสบการณ์ได้ไม่เกิน 2 ใน 3 ของจำนวนหน่วยกิตตามโครงสร้างของหลักสูตร

1.2.2 นักเรียน/นักศึกษา ที่มีประสบการณ์ในงานอาชีพ หรือฝึกงานในสถานฝึกอาชีพ หรือทำงานในงานอาชีพนั้นอยู่แล้ว หรือมีความรู้ในรายวิชาตามหลักสูตรดังกล่าวก่อนเข้าเรียน หรือเข้าเรียนแล้วแต่ขอไปเรียน หรือฝึกปฏิบัติในสถานฝึกอาชีพ จะขอประเมิน

มหาวิทยาลัยราชภัฏพระนครศรีอยุธยา RAJABHAT MAHASARAKHAM UNIVERSITY

เทียบโอนความรู้และประสบการณ์ เพื่อยกเว้นการเรียนรายวิชานั้นได้ นักเรียน นักศึกษาที่ประเมินแล้วไม่ผ่านเกณฑ์ จะขอประเมินเทียบโอนความรู้และประสบการณ์ในภาคเรียนนั้นอีกไม่ได้ แต่สามารถลงทะเบียนเรียนในภาคเรียนนั้นได้หรือสามารถขอรับคำแนะนำจากคณะกรรมการเพื่อใช้เป็นข้อมูลในการขอรับการประเมินในภาคเรียนต่อไป

3. วิธีการประเมินเพื่อเทียบโอนความรู้และประสบการณ์ดำเนินการดังนี้

3.1 การประเมินเบื้องต้น

3.1.1 พิจารณาหลักฐานที่ทางราชการและ/หรือสถานประกอบการออกให้ เช่น ใบสำคัญวุฒิบัตร เกียรติบัตร หนังสือรับรองของสถานประกอบการ รางวัล ฯลฯ

หลักฐานที่นำมาแสดงพิจารณาเงื่อนไข ดังนี้

1) หลักฐานต้องมีอายุ ไม่เกิน 5 ปี นับถึงวันขอประเมินหรือให้อยู่ในดุลยพินิจของสถานศึกษา ถ้าเป็นหลักฐานที่ภาคเอกชนออกให้ต้องเป็นภาคเอกชนที่ประกอบการถูกต้องตามกฎหมาย

2) รายวิชา/กลุ่มวิชาที่เสนอขอให้ประเมินจะต้องมีจำนวนชั่วโมงปฏิบัติงานและ/หรือฝึกอบรมไม่น้อยกว่าจำนวนชั่วโมงของรายวิชาที่กำหนดไว้ในหลักสูตร และมีเนื้อหาสาระที่สอดคล้องกับรายวิชา/กลุ่มวิชา ไม่น้อยกว่าร้อยละ 60

3.1.2 พิจารณาข้อมูลการสัมภาษณ์เบื้องต้น

3.2 การประเมินความรู้และประสบการณ์ใช้วิธีการและเครื่องมือประเมินที่หลากหลาย ตามระเบียบกระทรวงศึกษาธิการว่าด้วยการประเมินผลการเรียนตามหลักสูตรที่ขอประเมิน โดยคณะกรรมการประเมินเป็นผู้ดำเนินการ

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

3.3 การตัดสินผลการประเมินและการให้ค่าระดับผลการประเมิน

3.3.1 ให้ตัดสินผลการประเมินเป็นรายวิชาตามรายวิชาในหลักสูตร

3.3.2 ต้องได้คะแนนจากการประเมินเพื่อเทียบโอนความรู้และประสบการณ์ เข้าสู่หน่วยกิต ไม่ต่ำกว่าร้อยละ 50 แต่ละรายวิชาที่ขอประเมินถ้ามีทั้งภาคทฤษฎีและภาคปฏิบัติ ให้ทำการประเมินทั้งภาคทฤษฎีและภาคปฏิบัติ บางส่วนอาจประเมินเช่นเดียวกับการประเมินผลการเรียนในสถานศึกษาปกติ และเวลาที่ใช้ในการประเมินต้องไม่น้อยกว่าเวลาเรียนต่อสัปดาห์ที่กำหนดไว้ในหลักสูตร หรืออยู่ในดุลยพินิจของสถานศึกษา

3.3.3 ให้ใช้ค่าตัวเลขแสดงค่าระดับผลการประเมินตามค่าระดับผลการเรียน ตามระเบียบกระทรวงศึกษาธิการ ว่าด้วยการประเมินผลการเรียนตามหลักสูตรที่ขอ

ประเมินเทียบโอนความรู้และประสบการณ์

4. ผู้ประเมิน

ให้สถานศึกษาที่ทำหน้าที่ประเมินแต่งตั้งคณะกรรมการเพื่อดำเนินการประเมินดังนี้

4.1 คณะกรรมการประเมินเบื้องต้น ประกอบด้วยผู้รับผิดชอบไม่น้อยกว่า 3 คน โดยมีผู้ช่วยผู้อำนวยการฝ่ายวิชาการ หัวหน้างานหลักสูตรและการสอน หัวหน้างานวัดผลและประเมินผลการศึกษาร่วมเป็นคณะกรรมการทำหน้าที่ประเมินจากหลักฐานข้อมูลที่ผู้ขอรับการประเมินนำมาแสดงไว้และสัมภาษณ์

4.2 คณะกรรมการประเมินความรู้และประสบการณ์ประกอบด้วยผู้รับผิดชอบไม่น้อยกว่า 3 คน โดยมีหัวหน้าแผนกวิชาหรือหัวหน้าคณะงานผู้สอนหรือครูฝึกในรายวิชานั้น และผู้เชี่ยวชาญด้านวิชาชีพ ร่วมเป็นคณะกรรมการทำหน้าที่ประเมินว่าผู้ขอรับการประเมินมีความรู้ ความเข้าใจทักษะ ความชำนาญ เจตคติ รวมทั้ง กิจนิสัยตามสมรรถนะ (Competencies) ในรายวิชาที่ขอประเมินผ่านเกณฑ์หรือไม่ การประเมินต้องดำเนินการให้มีมาตรฐานชัดเจน สมเหตุสมผลยุติธรรม โปร่งใส และตรวจสอบได้

ขั้นตอนการขอประเมินเทียบโอนความรู้และประสบการณ์ของนักเรียน/นักศึกษา

1. ศึกษาข้อมูลเกี่ยวกับการประเมินเทียบโอนความรู้และประสบการณ์เข้าสู่หน่วยกิตจากสถานศึกษา

2. ยื่นคำร้องขอประเมินความรู้และประสบการณ์ พร้อมหลักฐาน ได้แก่

2.1 บัตรประจำตัวนักเรียน/นักศึกษา

2.2 หลักฐานที่แสดงความรู้และประสบการณ์ เช่น

2.2.1 บันทึกประจำวัน ที่แสดงถึงสิ่งที่ผู้ขอรับการประเมินทำเป็นประจำวันทั้งที่บ้าน สถานประกอบการ ฯลฯ

2.2.2 แฟ้มสะสมงาน (Portfolio) โดยรวบรวมเอกสาร ภาพถ่าย สำเนาเกียรติบัตร วุฒิบัตร จากการศึกษาทั้งในหรือนอกประเทศ การอบรมระยะสั้น รางวัลต่าง ๆ ที่เคยได้รับตลอดจนการปฏิบัติงานอาชีพที่ชัดเจน

2.2.3 หนังสือรับรองจากนายจ้าง/ผู้ร่วมงาน/ผู้นำชุมชน ซึ่งแสดงถึงการมีความรู้ ทักษะ ความสามารถในด้านอาชีพ

2.2.4 บุคคลที่ผู้ขอรับการประเมินสามารถอ้างอิงได้

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

2.2.5 ชิ้นงานหรือผลงานดีเด่นด้านวิชาการ วิชาชีพ ฯลฯ

2.2.6 หลักฐานหรือเอกสารที่แสดงผลตอบแทนทางธุรกิจ (ถ้ามี)

3. เข้ารับการประเมินเบื้องต้น

4. รับทราบผลการประเมินเบื้องต้น และเตรียมความพร้อมเข้ารับการประเมินเทียบโอนความรู้และประสบการณ์

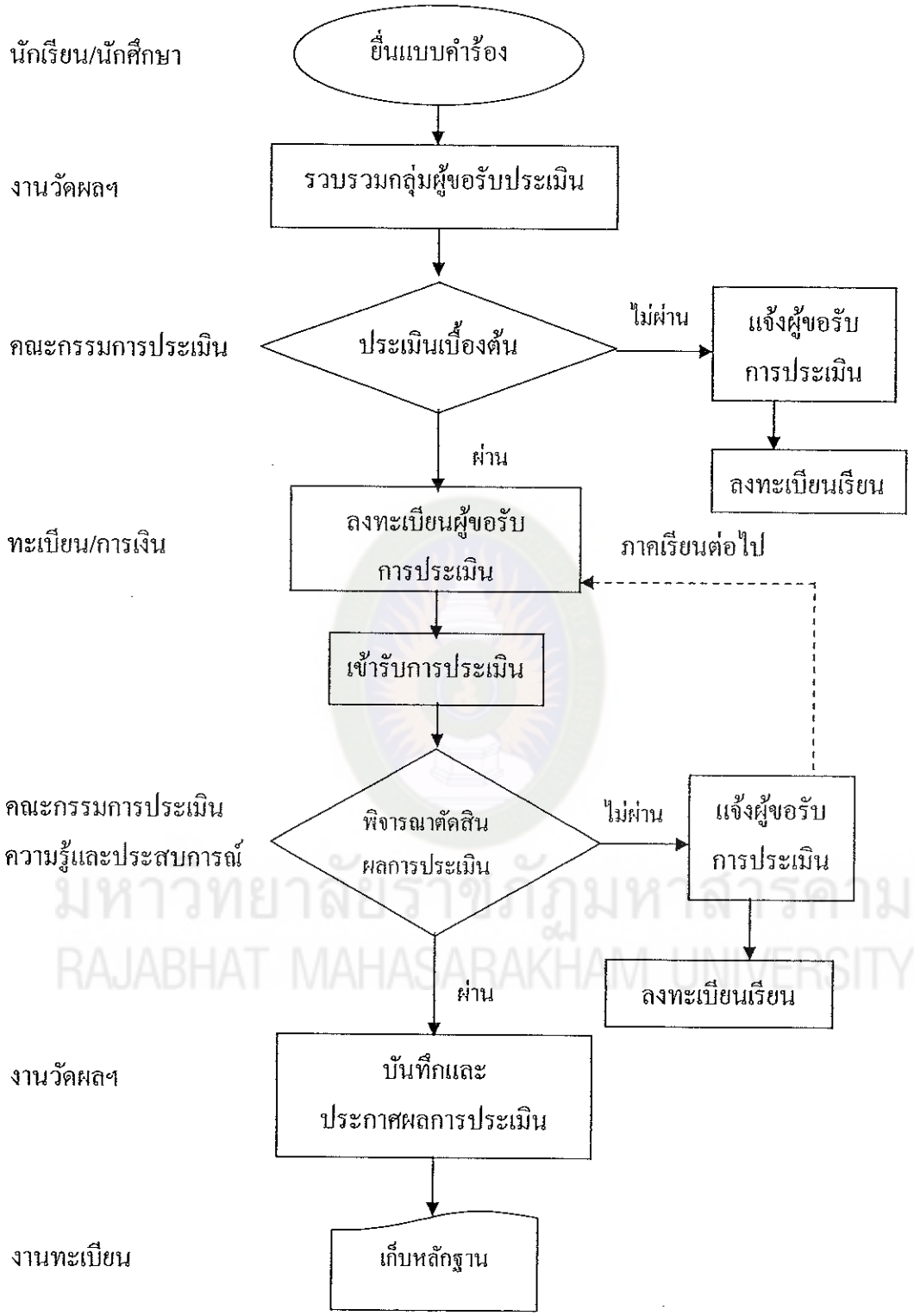
5. ลงทะเบียนขอประเมินตามระยะเวลาที่สถานศึกษากำหนด

6. เข้ารับการประเมินตามระยะเวลาและสถานที่ที่กำหนด

7. รับทราบผลการประเมิน

8. นำผลการประเมินไปเทียบโอนเพื่อยกเว้นการเรียนหรือลงทะเบียนเรียนรายวิชา

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY



แผนภาพที่ 2 แสดงขั้นตอนการประเมินเทียบโอนความรู้และประสบการณ์
(สำนักมาตรฐานการอาชีวศึกษาและวิชาชีพ : ม.ป.ป)

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

ระบบสนับสนุนการตัดสินใจ (Decision Support System)

เป็นระบบย่อยระบบหนึ่งที่อยู่ในระบบสารสนเทศที่นำเอาระบบคอมพิวเตอร์เข้ามาช่วยในการบริหารงานในแต่ละส่วนขององค์กร โดยสามารถใช้กับการตัดสินใจของบุคคลเพียงคนเดียวหรือช่วยในการตัดสินใจของบุคคลที่เป็นกลุ่มก็ได้ นอกจากนั้นระบบสนับสนุนการตัดสินใจยังช่วยในการตัดสินใจในปัญหาแบบกึ่งโครงสร้าง (Semistructured) และไม่มีโครงสร้าง (Unstructured) ได้อีกด้วย โดยที่ระบบสนับสนุนการตัดสินใจจะช่วยเพิ่มประสิทธิภาพในการตัดสินใจให้กับผู้บริหาร โดยมีการเสนอข้อมูลที่จำเป็นในการตัดสินใจให้กับผู้บริหาร เพื่อให้ผู้บริหารได้วิเคราะห์ และทำความเข้าใจกับข้อมูลเหล่านั้น ซึ่งสามารถใช้ในการตัดสินใจได้ นอกจากนั้นยังช่วยให้ผู้บริหารได้เลือกใช้ทางเลือกที่เหมาะสมกับปัญหานั้น ๆ มากที่สุด

ระบบการตัดสินใจได้เริ่มต้นขึ้นในปี ค.ศ. 1970 เนื่องมาจากการพัฒนาของเทคโนโลยีสารสนเทศ และการขยายตัวขององค์กรธุรกิจ ทำให้หลายหน่วยงานในสหรัฐอเมริกา ได้เริ่มพัฒนาระบบสารสนเทศที่มีขนาดและเทคโนโลยีที่เหมาะสม เพื่อใช้ในการรวบรวมข้อมูลและแบบจำลองในการตัดสินใจต่าง ๆ ตลอดจนพัฒนาให้ระบบสารสนเทศสามารถสื่อสารกับผู้ใช้ เพื่อช่วยในการตัดสินใจในปัญหาแบบกึ่งโครงสร้างและไม่มีโครงสร้าง จนก่อให้เกิดเป็นแนวคิดในการสร้างระบบการสนับสนุนการตัดสินใจ

คุณสมบัติของระบบสนับสนุนการตัดสินใจ

ระบบสนับสนุนการตัดสินใจที่ดีที่ช่วยจัดหาหรือจัดเตรียมเครื่องมือสำหรับผู้บริหาร และสามารถช่วยผู้บริหารในการตัดสินใจแก้ปัญหาต่าง ๆ ได้นั้น ควรมีคุณสมบัติดังต่อไปนี้

1. ควรเป็นระบบที่สนับสนุนการตัดสินใจทั้งแบบกึ่งโครงสร้างและแบบไม่มีโครงสร้าง

2. ควรเป็นระบบที่ง่ายต่อการเรียนรู้ และการใช้งาน เนื่องจากผู้ใช้บางคนอาจจะไม่ถนัดในการใช้งานในบางระบบ ดังนั้นระบบที่จะใช้งานดีและมีประสิทธิภาพควรเป็นระบบที่มีความสะดวกต่อผู้ใช้ในการใช้งานต่าง ๆ ภายในระบบนั้น

3. ควรเป็นระบบที่สามารถโต้ตอบ และสื่อสารกับผู้ใช้ได้อย่างรวดเร็ว เพื่อสนองตอบความต้องการของผู้ใช้ โดยเฉพาะการทำงานในปัจจุบันที่ต้องการความรวดเร็วในการทำงาน การแก้ปัญหาค่อนข้างสูง

4. ควรมีข้อมูลที่สนับสนุนการตัดสินใจได้อย่างเพียงพอ

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY

5. ควรมีความสามารถในการปรับปรุงแก้ไขข้อมูล เพื่อช่วยในการตัดสินใจได้

6. ควรมีข้อมูล และแบบจำลองสำหรับสนับสนุนการตัดสินใจที่เหมาะสมและสอดคล้องกับแต่ละลักษณะของปัญหา

7. ควรมีความยืดหยุ่นในการสนองความต้องการที่เปลี่ยนแปลงของผู้ใช้

8. ควรสนับสนุนการทำงานต่อผู้บริหารหลาย ๆ ระดับ

9. ควรสนับสนุนการทำงานเป็นกลุ่ม

10. ควรสนับสนุนการตัดสินใจที่เกี่ยวเนื่องกัน

ความแตกต่างระหว่างระบบสนับสนุนการตัดสินใจกับระบบสารสนเทศอื่น ๆ

1. ระบบสนับสนุนการตัดสินใจ เป็นระบบที่ได้รับการพัฒนาให้สามารถจัดการกับข้อมูลของปัญหากึ่งโครงสร้าง ซึ่งมักจะเป็นปัยหาของผู้บริหารระดับกลาง และระดับสูง ในขณะที่ระบบสารสนเทศในสำนักงานจะเกี่ยวข้องกับการทำงานประจำวันของพนักงานหรือหัวหน้างาน

2. ระบบสนับสนุนการตัดสินใจ เป็นระบบที่ได้รับการพัฒนาให้เหมาะสมกับการแก้ปัญหาของผู้ใช้ ซึ่งจะแตกต่างจากระบบสารสนเทศในการปฏิบัติงานที่เก็บรวบรวมและจัดการสารสนเทศทั่วไปขององค์กร(ณาตยา ฉาบนาค. 2548 : 184 – 185)

ประเภทของระบบสนับสนุนการตัดสินใจ

จำแนกตามแนวคิดของ Holsappleและ Whinstonเป็นรูปแบบที่ได้รับความนิยมมาก โดยได้จำแนกระบบสนับสนุนการตัดสินใจออกเป็น 6 ประเภทดังนี้

1. ระบบสนับสนุนการตัดสินใจโดยอาศัยข้อความ (Text-oriented DSS) เนื่องจากสารสนเทศ (รวมถึงข้อมูลและองค์ความรู้) จะถูกจัดเก็บในรูปแบบของข้อความและสามารถสืบค้นโดยผู้ตัดสินใจเท่านั้น จึงจำเป็นต้องนำเสนอ ประมวลผล และแยกประเภทข้อความเหล่านั้นอย่างมีประสิทธิภาพ โดยอาศัยวิธีการทางอิเล็กทรอนิกส์ของเทคโนโลยีสารสนเทศ เช่น การสร้าง การเรียบเรียง และการแสดงเอกสารทางอิเล็กทรอนิกส์ ผ่านทางเอกสารทางอินเทอร์เน็ต การสร้งภาพและการสร้างการเชื่อมโยงระหว่างข้อความ (Hypertext) และตัวแทนชาญฉลาด ซึ่งสามารถจัดอยู่ในประเภทระบบสนับสนุนการตัดสินใจโดยอาศัยข้อความได้

2. ระบบสนับสนุนการตัดสินที่อาศัยฐานข้อมูล (Data-oriented DSS) โดยฐานข้อมูลที่ใช้สำหรับระบบสนับสนุนการตัดสินประเภทนี้ ส่วนใหญ่เป็นฐานข้อมูลเชิง

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY

สัมพันธ์ซึ่งปริมาณสารสนเทศที่ทำการจัดการมีแนวโน้มที่จะเพิ่มปริมาณขึ้นเรื่อย ๆ และมีลักษณะของสารสนเทศเชิงบรรยาย ซึ่งมีโครงสร้างที่ตายตัว ไม่ยืดหยุ่น โดยระบบสนับสนุนการตัดสินใจประเภทนี้มีความสามารถในการสร้างรายงานและความสามารถในการสอบถามข้อมูลได้ดีและหน่วยงานที่มีบทบาทสำคัญสำหรับระบบสนับสนุนการตัดสินใจประเภทนี้ คือ หน่วยงานด้านการจัดการฐานข้อมูล

3. ระบบสนับสนุนการตัดสินใจแบบกระดาษคำนวณ (Spreadsheet-oriented DSS) โปรแกรมกระดาษคำนวณเป็นโปรแกรมที่ใช้ในการสร้างแบบจำลองเพื่อทำหน้าที่ในการวิเคราะห์ของระบบสนับสนุนการตัดสินใจแบบจำลองนี้ไม่เพียงแต่ทำหน้าที่สร้าง แสดง และปรับปรุงองค์ความรู้ที่เกี่ยวข้องกับกระบวนการแก้ไขปัญหา แต่ยังช่วยให้ระบบสารสนเทศสร้างส่วนให้คำแนะนำด้วยตนเอง โปรแกรมกระดาษคำนวณได้รับความนิยมในการนำมาใช้พัฒนาระบบสนับสนุนการตัดสินใจโดยผู้ใช้ และเครื่องมือที่ผู้ใช้มักนำมาพัฒนาระบบสนับสนุนการตัดสินใจคือ Microsoft Excel ซึ่งเป็นโปรแกรมที่รวบรวมฟังก์ชั่นทางด้านสถิติ คณิตศาสตร์ การเงิน วิศวกรรม และอื่น ๆ ไว้อย่างมากมาย

4. ระบบสนับสนุนการตัดสินใจเพื่อการแก้ไขปัญหา (Solver-oriented DSS) การแก้ไขปัญหาเป็นกระบวนการที่เขียนขึ้นในลักษณะเดียวกับโปรแกรมคอมพิวเตอร์ เพื่อแก้ไขปัญหาตามลักษณะโปรแกรมที่ใช้ โดยนักพัฒนาระบบอาจเขียนโปรแกรมภาษาต่าง ๆ เช่น C++ หรืออาจใช้เครื่องมือเพิ่มเติมจากโปรแกรมประเภทกระดาษคำนวณ เช่น Excel หรือใช้แบบจำลองด้านภาษาต่าง ๆ ในการพัฒนาระบบสนับสนุนการตัดสินใจ เช่น LINGO, LINDO ในการแก้ไขปัญหาที่มีความซับซ้อน เช่น โปรแกรมเชิงเส้น ซึ่งใช้สำหรับการหาจุด (คำตอบ) ที่เหมาะสมที่สุด ซึ่งมีประโยชน์ทางการค้าเป็นอย่างมาก

5. ระบบสนับสนุนโดยอาศัยกฎ (Rule-oriented DSS) ในการศึกษาถึงกระบวนการและการให้เหตุผลของกฎต่าง ๆ ในระบบสนับสนุนการตัดสินใจ อาจต้องอาศัยการอธิบายโดยใช้องค์ความรู้ ซึ่งเป็นองค์ประกอบหนึ่งของระบบสนับสนุนการตัดสินใจ โดยทั่วไปแล้วกฎเกิดจากการจำลองรูปแบบของกระบวนการคิดและให้เหตุผลของมนุษย์ผู้เชี่ยวชาญที่ใช้ในการตัดสินใจ

6. ระบบสนับสนุนการตัดสินใจแบบผสม (Compound DSS) เป็นระบบสนับสนุนการตัดสินใจที่มีการนำระบบต่าง ๆ มาใช้งานร่วมกัน ซึ่งระบบประเภทนี้อาจประกอบด้วยระบบสนับสนุนการตัดสินใจแบบต่าง ๆ ข้างต้น ตั้งแต่ 2 ประเภทขึ้นไป (กิตติ ภักดีวัฒนะกุล. 2550 : 25 - 26)

มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา PANSMET RAJABHAT UNIVERSITY

องค์ความรู้ที่ผู้ตัดสินใจจะสามารถค้นหาได้จากระบบสนับสนุนการตัดสินใจนั้น จะถูกจัดเก็บไว้ในฐานองค์ความรู้ (Knowledge Base) ซึ่งเป็นฐานข้อมูลที่จัดเก็บองค์ความรู้ที่ได้จากประสบการณ์ของผู้เชี่ยวชาญในการตัดสินใจแก้ไขปัญหามากมายหลายสาขา ดังนั้น องค์ความรู้จึงต้องมีระบบจัดการฐานองค์ความรู้ (Knowledge Base Management System : KBMS) เป็นส่วนช่วยในการจัดการองค์ความรู้ ทำให้ผู้ตัดสินใจสามารถเรียกใช้องค์ความรู้ที่ตรงกับปัญหาที่ต้องการแก้ไขได้อย่างรวดเร็วและถูกต้อง

ดังนั้นในบางครั้งทำให้ระบบสนับสนุนการตัดสินใจที่มีองค์ประกอบส่วนนี้รวมอยู่ถูกเรียกชื่อระบบได้อีกหลายชื่อว่า "ระบบสนับสนุนการตัดสินใจชาญฉลาด (Intelligent DSS)""ระบบสนับสนุนการตัดสินใจและระบบผู้เชี่ยวชาญ (DSS/ES)" "ระบบสนับสนุนการตัดสินใจด้วยฐานองค์ความรู้ (Knowledge-based DSS)" เนื่องจากอค์ประกอบส่วนนี้ ช่วยยกระดับความสามารถของระบบสนับสนุนการตัดสินใจที่มีองค์ประกอบเพียง 3 ส่วนแรกดังแผนภาพที่ 3

แผนภาพที่ 3 แสดงโครงสร้างของส่วนการจัดการองค์ความรู้

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

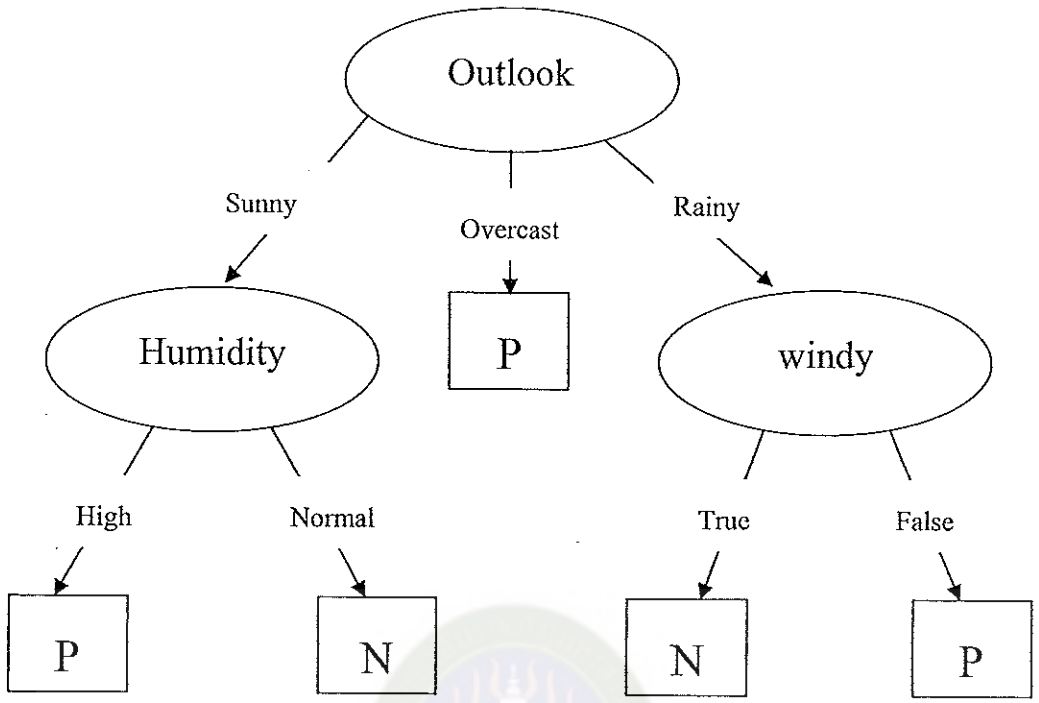
จากแผนภาพที่ 3 จะเห็นโครงสร้างการทำงานของของส่วนจัดการองค์ความรู้ พบว่ามีส่วนการทำงานย่อยที่สำคัญทั้งหมด 5 ส่วน ได้แก่ ส่วนที่ทำหน้าที่ในการรวบรวมองค์ความรู้จากแหล่งองค์ความรู้ เช่น ผู้เชี่ยวชาญหรือจากแหล่งอื่น ๆ เพื่อนำองค์ความรู้เหล่านั้นมาแปลงให้อยู่ในรูปที่โปรแกรมคอมพิวเตอร์สามารถเข้าใจได้ รวมแล้วจะเรียกกิจกรรมทั้งสองนี้ว่า "Knowledge Acquisition" จากนั้น องค์ความรู้จะเข้าสู่ส่วนที่ทำหน้าที่จัดรูปแบบองค์ความรู้ที่ระบบสามารถเข้าใจได้ (Knowledge Representation) ซึ่งอาจจะต้องมีการเข้ารหัสองค์ความรู้ เช่น เปลี่ยนเป็นสัญลักษณ์ต่าง ๆ เพื่อการประมวลผล ไปเก็บไว้ในฐานองค์ความรู้ที่จะต้องมีการตรวจสอบองค์ความรู้ (Knowledge Validation) ที่รับมาเก็บไว้ว่าถูกต้องหรือไม่ อย่างไร เมื่อผู้ใช้เรียกใช้ระบบ จะต้องมีส่วนที่ทำหน้าที่อธิบายความ (Explanation/ Justification) องค์ความรู้ต่าง ๆ ให้ผู้ใช้สามารถเข้าใจในแนวทางหรือคำตอบที่ระบบส่งให้กับผู้ใช้ด้วยส่วนที่ทำหน้าที่สำคัญที่สุดคือ "ส่วนวินิจฉัย/การสรุปความ (Inferencing)" (กิตติ ภักดีวัฒนะกุล. 2550 : 42 - 43)

การสร้างต้นไม้ตัดสินใจ (Decision Tree)

ต้นไม้ตัดสินใจ (Decision Tree) คือต้นไม้ที่ใช้สนับสนุนการตัดสินใจ มีลักษณะเป็นโครงสร้างคล้ายกับต้นไม้หัวกลับที่มีรากอยู่ด้านบน และใบอยู่ด้านล่างสุด โดยที่ภายในต้นไม้จะประกอบไปด้วยโหนด (Node) ซึ่งแต่ละโหนดนั้นจะแสดงถึงการทดสอบหรือการตัดสินใจบนข้อมูลของคุณสมบัติหรือแอททริบิวต์ต่าง ๆ กิ่งของต้นไม้แสดงถึงค่าหรือผลลัพธ์ที่ได้จากการทดสอบ และใบซึ่งเป็นสิ่งที่อยู่ล่างสุดของต้นไม้ตัดสินใจแสดงถึงกลุ่มข้อมูล (Class) หรือผลลัพธ์ที่ได้จากการทำนาย โหนดที่อยู่บนสุดของต้นไม้เรียกว่า โหนดราก (Root Node) แสดงโครงสร้างของต้นไม้ดังแผนภาพที่ 4 ซึ่งแสดงถึงต้นไม้ที่ใช้ในการตัดสินใจว่าจะออกไปเล่นกอล์ฟหรือไม่ (Quinlan. 1986 ; อ้างอิงใน นฤพนต์ ว่องประชานุกูล. 2548 : 4) โดยพิจารณาจากสภาพอากาศต่าง ๆ ประกอบการตัดสินใจโดยโหนดที่แสดงในรูปวงรี แสดงถึงการทดสอบค่าที่เป็นไปได้ของแอททริบิวต์นั้น ๆ และใบที่แสดงในรูปสี่เหลี่ยมจะแสดงการจำแนกกลุ่มของข้อมูล ซึ่งเป็นผลลัพธ์ของการทำนายว่าจะออกไปเล่นกอล์ฟ (P) หรือไม่ออกไปเล่น (N) จากการทดสอบตามเส้นทางของต้นไม้ตัดสินใจ

ในการจำแนกข้อมูลที่ได้รับเข้ามาใหม่นั้น ค่าของแอททริบิวต์ต่าง ๆ ของข้อมูลเหล่านั้นจะถูกทดสอบด้วยต้นไม้ตัดสินใจ โดยจะเริ่มต้นการทดสอบตั้งแต่โหนดรากไปจนถึงใบ โดยที่ใบจะแสดงถึงกลุ่มของการทำนายข้อมูลนั้น

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY



แผนภาพที่ 4 ต้นไม้ตัดสินใจที่ใช้ในการตัดสินใจการออกไปเล่นกอล์ฟ

โดยหลักการพื้นฐานของการสร้างต้นไม้ตัดสินใจ จะเป็นการสร้างในลักษณะจากบนลงล่าง (Top-down) คือจะเริ่มจากการสร้างรากของต้นไม้ก่อน แล้วจึงแตกกิ่งไปจึงถึงใบ โดยจะแสดงขั้นตอนการสร้างต้นไม้ตัดสินใจได้ดังนี้ (Han and Kamber. 2001 ; อ้างอิงในนฤพนต์ ว่องประชานุกูล. 2548 : 5)

1. ต้นไม้จะเริ่มต้น โดยมีโหนดเพียงโหนดเดียว แสดงถึงชุดข้อมูลฝึก (Training Set)

2. ถ้าข้อมูลทั้งหมดอยู่ในกลุ่มเดียวกันแล้ว ให้โหนดนั้นเป็นใบและตั้งชื่อตามกลุ่มของข้อมูลนั้น

3. ถ้าในโหนดมีข้อมูลหลายกลุ่มปะปนอยู่ จะต้องวัดค่าเกน (Gain) ของแต่ละแอททริบิวต์ เพื่อที่จะใช้เป็นเกณฑ์ในการคัดเลือกแอททริบิวต์ ที่มีความสามารถในการแบ่งแยกข้อมูลออกเป็นกลุ่มต่าง ๆ ได้ดีที่สุด โดยแอททริบิวต์ที่มีค่าเกนมากที่สุดจะถูกเลือกให้เป็นตัวทดสอบหรือแอททริบิวต์ที่ใช้ในการตัดสินใจ โดยแสดงในรูปของโหนดบนต้นไม้

4. กิ่งของต้นไม้จะถูกสร้างขึ้นจากค่าต่าง ๆ ที่เป็นไปได้ของโหนดทดสอบ และข้อมูลจะถูกแบ่งออกตามกิ่งต่าง ๆ ที่สร้างขึ้น

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

5. ทำการวนซ้ำเพื่อหาแอททริบิวต์ที่มีค่าเกนมากที่สุด สำหรับข้อมูลที่แบ่งแยกออกมาในแต่ละกิ่ง เพื่อนำแอททริบิวต์นี้มาสร้างเป็นโหนดตัดสินใจต่อไป โดยที่แอททริบิวต์ที่ถูกเลือกมาเป็นโหนดแล้ว จะไม่ถูกเลือกมาอีก สำหรับโหนดในระดับต่อ ๆ ไป

6. ทำการวนซ้ำเพื่อแบ่งข้อมูลและแตกกิ่งของต้นไม้ไปเรื่อย ๆ โดยการวนซ้ำจะสิ้นสุดก็ต่อเมื่อเงื่อนไขข้อใดข้อหนึ่งต่อไปนี้เป็นจริง

6.1 ถ้าข้อมูลทุกตัวในโหนดนั้นอยู่ในกลุ่มเดียวกัน ให้สร้างใบตามกลุ่มของข้อมูลนั้น

6.2 ถ้าไม่เหลือแอททริบิวต์ใดสำหรับใช้ในการแบ่งข้อมูลแล้ว ซึ่งในกรณีนี้จะใช้กลุ่มที่มีข้อมูลสนับสนุนมากที่สุดเป็นใบ

6.3 ถ้าไม่มีข้อมูลสนับสนุนสำหรับกิ่งนั้น ๆ แล้ว ให้สร้างใบตามกลุ่มที่มีข้อมูลสนับสนุนมากที่สุด

ในขั้นตอนการสร้างต้นไม้ตัดสินใจ อัลกอริทึม C4.5 เป็นอัลกอริทึมที่มีชื่อเสียงและเป็นที่รู้จักกันอย่างแพร่หลาย พัฒนาโดย Quinlan (1993) ที่ได้พัฒนาต่อมาจากอัลกอริทึม ID3 ที่เขาได้พัฒนาขึ้น (Quinlan. 1986 ; อ้างอิงในนฤพนต์ ว่องประชานุกูล. 2548 : 6) เป็นวิธีการเรียนรู้จากกลุ่มตัวอย่างที่เรียกว่า ชุดข้อมูลฝึก (Training Set) ที่อาศัยวิธีการจัดหมวดหมู่เพื่อสร้างต้นไม้ตัดสินใจ

ชุดข้อมูลฝึกจะมีลักษณะคล้ายกับข้อมูลในฐานข้อมูลสัมพันธ์ (Relation Database) แสดงในรูปของตารางที่ประกอบด้วย แถวแสดงข้อมูลหรือตัวอย่าง และคอลัมน์แสดงแอททริบิวต์ของข้อมูล ซึ่งแบ่งออกเป็น 2 ชนิด คือ

1. แอททริบิวต์ที่เป็นจุดมุ่งหมาย (Goal Attribute) ของการจำแนกกลุ่มข้อมูล เป็นแอททริบิวต์ที่กำหนดว่าตัวอย่างนั้น ๆ ถูกจัดอยู่ในกลุ่มไหน โดยจะมีเพียงแอททริบิวต์ในแต่ละชุดข้อมูล และข้อมูลจะเป็นชนิดข้อความเท่านั้น

2. แอททริบิวต์ประกอบการทำนาย (Predicting Attribute) เป็นแอททริบิวต์ที่บ่งบอกถึงคุณสมบัติต่าง ๆ ของตัวอย่างแต่ละตัวอย่าง โดยแต่ละแอททริบิวต์อาจมีข้อมูลเป็นชนิดข้อความหรือตัวเลขก็ได้

จากตารางที่ 1 เป็นตัวอย่างชุดข้อมูลฝึกที่ใช้ประกอบการตัดสินใจออกไปเล่นกอล์ฟโดยพิจารณาจากสภาพอากาศต่าง ๆ ประกอบการตัดสินใจ เมื่อนำมาสร้างเป็นต้นไม้ตัดสินใจ สามารถแสดงโครงสร้างของต้นไม้ดังแผนภาพที่ 3 โดยชุดข้อมูลฝึกนี้ประกอบด้วย

มหาวิทยาลัยราชภัฏพระนคร RAJABHAT MAHASARAKHAM UNIVERSITY

แอททริบิวต์ Class เป็นแอททริบิวต์ที่เป็นจุดมุ่งหมาย มีค่าที่เป็นไปได้ คือ P หรือ N แอททริบิวต์ Outlook, Temperature, Humidity และ Windy เป็นแอททริบิวต์ประกอบการทำนายของชุดข้อมูล

ประสิทธิภาพของต้นไม้ตัดสินใจไม่ได้อยู่ที่การสร้างต้นไม้ตัดสินใจ เพื่อให้สามารถจัดกลุ่มชุดข้อมูลฝึกได้อย่างถูกต้องเท่านั้น แต่ต้องสามารถจัดกลุ่มข้อมูลจากตัวอย่างใหม่ ๆ ที่นอกเหนือจากชุดข้อมูลฝึกได้อย่างถูกต้องด้วย ดังนั้นการสร้างต้นไม้ตัดสินใจจึงควรจะมีชุดข้อมูลทดสอบ (Test Set) ที่จะใช้ตรวจสอบความถูกต้องของต้นไม้ตัดสินใจด้วย

ตารางที่ 1 ชุดข้อมูลฝึกที่ใช้ประกอบการตัดสินใจเล่นกอล์ฟ

ID	Attribute				Class
	Outlook	Temperature	Humidity	Windy	
1	Sunny	Hot	High	False	N
2	Sunny	Hot	High	True	N
3	Overcast	Hot	High	False	P
4	Rainy	Mild	High	False	P
5	Rainy	Cool	Normal	False	P
6	Rainy	Cool	Normal	True	N
7	Overcast	Cool	Normal	True	P
8	Sunny	Mild	High	False	N
9	Sunny	Cool	Normal	False	P
10	Rainy	Mild	Normal	False	P
11	Sunny	Mild	Normal	True	P
12	Overcast	Mild	High	True	P
13	Overcast	Hot	Normal	False	P
14	Rainy	Mild	High	True	N

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

การคัดเลือกแอททริบิวต์เพื่อจำแนกกลุ่มของข้อมูล

ในการสร้างต้นไม้ตัดสินใจ ปัญหาสำคัญที่ต้องพิจารณาคือ ควรจะตัดสินใจเลือกแอททริบิวต์ใดมาทำหน้าที่เป็นโหนดราก ในแต่ละขั้นตอนของการสร้างต้นไม้และต้นไม้ย่อย

(Subtree) ของต้นไม้ตัดสินใจ เกณฑ์ที่ใช้ช่วยประกอบการเลือกแอททริบิวต์ คือการคำนวณค่าเกน (Gain) ซึ่งเป็นค่าที่บ่งบอกว่าแอททริบิวต์นั้นจะสามารถจำแนกกลุ่มข้อมูลได้ดีเพียงใด โดยทดลองเลือกแต่ละแอททริบิวต์ที่เป็นไปได้จากชุดข้อมูลมาทำหน้าที่เป็นโหนดราก ถ้าแอททริบิวต์ใดให้ค่าเกนที่สูงที่สุด แสดงว่าแอททริบิวต์นั้นสามารถจำแนกลุ่มข้อมูลได้ดีที่สุด หรือเป็นแอททริบิวต์ที่จัดกลุ่มของข้อมูลแล้ว ได้ข้อมูลในแต่ละใบของต้นไม้เป็นกลุ่มเดียวกันทั้งหมด หรือมีข้อมูลต่างกลุ่มปะปนมาบ้างเพียงเล็กน้อยเท่านั้น โดยค่าเกนสำหรับการเลือกแอททริบิวต์ที่สำคัญแสดงได้ดังนี้

1. ค่ามาตรฐานเกน(Gain Criterion)

วิธีการสร้างต้นไม้ตัดสินใจ โดยใช้อัลกอริทึม ID3 จะใช้ค่ามาตรฐานเกนในการตัดสินใจเลือแอททริบิวต์ที่จะใช้เป็นโหนดรากของต้นไม้ หรือของต้นไม้ย่อย โดยการคำนวณค่าเกนของแต่ละแอททริบิวต์ เมื่อใช้แบ่งกลุ่มตัวอย่าง และเลือกแอททริบิวต์ที่มีค่าเกนสูงที่สุดมาเป็นโหนดราก ซึ่งแอททริบิวต์นี้จะมีความสามารถในการจำแนกกลุ่มข้อมูลสูง โดยที่ต้องการข้อมูลจำนวนน้อยที่สุดในการที่จะระบุว่าข้อมูลนั้นอยู่ในกลุ่มใด และการคัดเลือกแอททริบิวต์นี้ทำให้สามารถแบ่งข้อมูลออกมาโดยที่มีการปะปนกันของกลุ่มที่ต่างกันเกิดขึ้นน้อยอีกด้วย ค่าเกนนี้คำนวณได้โดยใช้ความรู้จากทฤษฎีสารสนเทศ (Information Theory) ซึ่งมีสาระสำคัญคือ ค่าสารสนเทศของข้อมูลจะขึ้นอยู่กับความน่าจะเป็นของข้อมูล ซึ่งสามารถวัดอยู่ในรูปของบิต (Bits) เขียนเป็นสมการได้ดังนี้

ค่าสารสนเทศของข้อมูล = $-\log_2$ (ความน่าจะเป็นของข้อมูล) (1)

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

การใช้ค่า Information Gain จะช่วยลดจำนวนครั้งของการทดสอบในการแยกแยะข้อมูล อีกทั้งยังรับประกันว่าต้นไม้ตัดสินใจที่ได้จะไม่มีความซับซ้อนมากจนเกินไป ซึ่งค่า Information Gain นั้นสามารถคำนวณได้จากสมการดังต่อไปนี้ (Han and Kamber. 2001 อ้างอิงใน นฤพนต์ ว่องประชานุกูล. 2548 : 8) กำหนดให้

S เป็นเซตของข้อมูล ซึ่งประกอบด้วยข้อมูล sเรคคอร์ด

M เป็นจำนวนกลุ่มทั้งหมดที่ต่างกันของข้อมูลชุดนั้น

C_i แทนกลุ่มในลำดับที่ i โดยที่ iมีค่าระหว่าง 1 ถึง m

s_i แทนจำนวนข้อมูลที่เป็นสมาชิกของ Sและอยู่ในกลุ่มของ C_i

s_{ij} แทนจำนวนข้อมูลที่เป็นสมาชิกของ S ในกลุ่ม C_i จากการแบ่งข้อมูลด้วยค่าที่เป็นไปได้ j ของแอททริบิวต์ A โดยที่ j มีค่าระหว่าง 1 ถึง v

สำนักวิทยบริการฯ มหาวิทยาลัยราชภัฏมหาสารคาม วิทยานิพนธ์ งานวิจัย

s_i/s แทนค่าความน่าจะเป็นที่ข้อมูลจะอยู่ในกลุ่ม C_i

ค่า information gain ที่ต้องการสำหรับจำแนกข้อมูลออกเป็นแต่ละกลุ่มหาได้โดย

$$I(s_1, s_2, ..., s_m = -\sum_{i=1}^{m} \frac{s_i}{s} \log_2 \frac{s_i}{s} \quad (2)$$

ค่าคาดคะเนของข้อมูล (Entropy) ของแอททริบิวต์ A ซึ่งมีค่าของแอททริบิวต์เป็น (a_1 , a_2 , a_3 ,..., a_v) หาได้โดย

$$E(A) = \sum_{j=1}^{v} \frac{s_{ij} + ... + s_{mj}}{s} I(s_{ij}, ..., s_{mj}) \quad (3)$$

ค่ามาตรฐานเกนที่จะใช้พิจารณาเลือกแอททริบิวต์ A มาเป็นโหนดของต้นไม้มีค่าเท่ากับปริมาณข้อมูลที่ต้องการเพื่อให้สามารถจำแนกกลุ่มของข้อมูลได้ ลบด้วยปริมาณข้อมูลที่ต้องการเพื่อการจำแนกกลุ่มของข้อมูล โดยใช้แอททริบิวต์ A เป็นตัวตรวจสอบเพื่อจำแนกกลุ่มของข้อมูล เขียนเป็นสมการได้ดังนี้

$$Gain(A) = I(s_1, s_2, ..., s_m) - E(A) \quad (4)$$

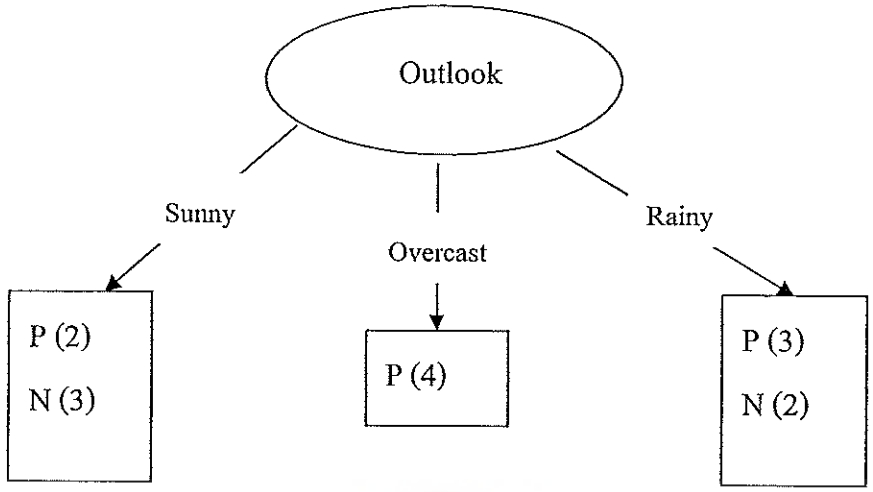
จากตัวอย่างข้อมูลสภาพอากาศ ประกอบด้วยการตัดสินใจเล่นกอล์ฟในตารางที่ 1 เซตของข้อมูลฝึก T ประกอบด้วยข้อมูลจำนวน 14 เรคคอร์ด แบ่งข้อมูลออกเป็น 2 กลุ่มคือ ข้อมูลที่ตัดสินใจออกไปเล่นกอล์ฟ (Class = P) จำนวน 9 เรคคอร์ด และตัดสินใจไม่ออกไปเล่นกอล์ฟ (Class = N) จำนวน 5 เรคคอร์ด การจะระบุว่าข้อมูลหนึ่งเรคคอร์ดอยู่ในกลุ่ม P หรือ N ต้องการปริมาณข้อมูลประกอบการตัดสินใจ เพื่อจำแนกกลุ่ม โดยใช้สมการที่ (2) ดังนี้

$$I(T) = -(9/14) \times \log_2(9/14) - (5/14) \times \log_2(5/14)$$
$$= 0.940 \text{ บิต}$$

การจะจำแนกกลุ่มของข้อมูลเพื่อตัดสินใจออกไปเล่นกอล์ฟหรือไม่นั้น ต้องใช้ข้อมูลจากแอททริบิวต์อื่นประกอบการตัดสินใจ ถ้าแบ่งข้อมูลชุดนี้โดยใช้แอททริบิวต์ Outlook จะสามารถจำแนกกลุ่มข้อมูลได้ดังภาพที่ 3 โดยได้แสดงจำนวนเรคคอร์ดของแต่ละกลุ่มข้อมูลไว้ในวงเล็บด้วย เมื่อแบ่งตามค่าที่เป็นไปได้จะต้องการปริมาณข้อมูลเพิ่มเพื่อ

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY

ประกอบการเลือกกลุ่ม และสามารถคำนวณค่า Entropy ของแอททริบิวต์ได้ โดยใช้สมการที่ 3



แผนภาพที่ 5 แสดงการจำแนกกลุ่มของข้อมูลโดยใช้แอททริบิวต์ Outlook

$$E(Outlook) = (5/14) \times (-(2/5 \times \log_2(2/5) - (3/5) \times \log_2(3/5)$$
$$+ (4/14) \times (-(4/4) \times \log_2(4/4) - (0/4) \times \log_2(0/4)$$
$$+ (5/14) \times (-(3/5) \times \log_2(3/5) - (2/5) \times \log_2(2/5)$$
$$= 0.693 \text{ บิต}$$

นั่นคือ ถ้าต้องการจำแนกกลุ่มของข้อมูลใหม่ โดยใช้แอททริบิวต์ Outlook เป็นตัวตรวจสอบ เพื่อจำแนกกลุ่มของข้อมูล การพิจารณาจากค่า Outlook ของข้อมูลใหม่นี้จะต้องใช้ข้อมูลเพิ่มอีก 0.693 บิต จึงจะบอกกลุ่มที่ถูกต้องข้อมูลใหม่นี้ได้

ดังนั้นสามารถคำนวณค่าเกนจากการเลือกแอททริบิวต์ Outlook เป็นแอททริบิวต์เพื่อใช้แบ่งข้อมูลได้จากสมการที่ (4) ดังนี้

$$\text{Gain (Outlook)} = I(T) - E(Outlook)$$
$$= 0.940 - 0.693$$
$$= 0.247 \text{ บิต}$$

แอททริบิวต์ที่เหลือที่สามารถถูกเลือกมาเป็นตัวทดสอบเพื่อจัดกลุ่มของข้อมูลฝึกคือแอททริบิวต์ Temperature, Humidity และ Windy สามารถคำนวณค่าเกนจากการเลือกแต่ละแอททริบิวต์ได้ดังนี้

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY

Gain (Temperature) = I (T) - E (Temperature)

= 0.940 - 0.911

= 0.029 บิต

Gain (Humidity) = I (T) - E (Humidity)

= 0.940 - 0.788

= 0.152 บิต

Gain (Windy) = I(T) - E (Windy)

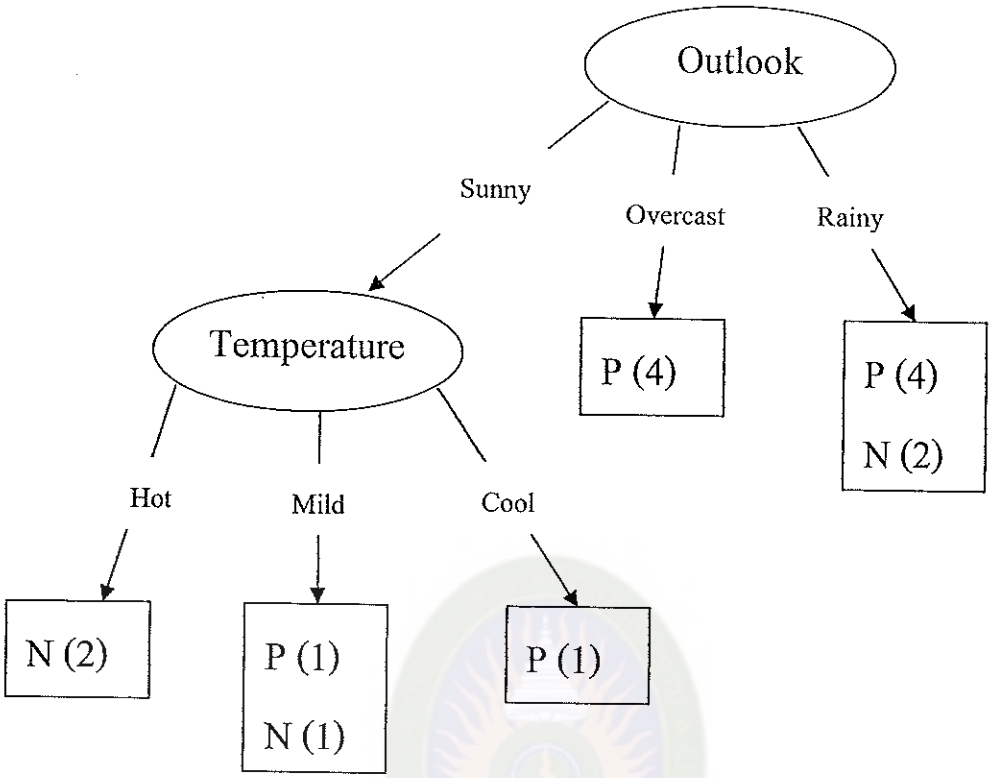
= 0.940 - 0.892

= 0.048 บิต

จะเห็นว่าแอททริบิวต์ที่ให้ค่าเกนสูงที่สุดคือ Outlook ดังนั้นแอททริบิวต์ Outlook จึงถูกเลือกเป็นโหนดรากของต้นไม้ตัดสินใจ แต่เนื่องจากยังไม่สามารถจัดกลุ่มของข้อมูลเดียวกันได้ทั้งหมด จึงต้องสร้างต้นไม้ตัดสินใจต่อไป โดยพิจารณาเลือกแอททริบิวต์ที่จะมาเป็นโหนดในระดับที่ 2 ต่อจากจากโหนดรากเพื่อจัดกลุ่มของข้อมูล ในกรณี Outlook = Overcast ไม่ต้องสร้างโหนดเพิ่มเติมอีกเนื่องจากสามารถจัดกลุ่มของข้อมูลที่เป็นกลุ่ม P ได้ทั้งหมดแล้ว

แอททริบิวต์ที่สามารถถูกเลือกเป็นโหนดในระดับที่ 2 ได้ประกอบด้วย Temperature, Humidity และ Windy โดยที่แอททริบิวต์ Outlook จะไม่ถูกเลือกมาอีกสำหรับโหนดในระดับต่อไป เมื่อพิจารณาการสร้างโหนดลูกทางด้านซ้ายมือ (Outlook = Sunny) ถ้าเลือกแอททริบิวต์ Temperature จะสามารถจำแนกกลุ่มของข้อมูลได้ดังแผนภาพที่ 6 และสามารถคำนวณค่าเกนได้ดังต่อไปนี้

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY



แผนภาพที่ 6 แสดงการจำแนกกลุ่มของข้อมูลโดยใช้แอททริบิวต์ Temperature เป็นโหนดระดับที่ 2

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

$$I(\text{Outlook} = \text{Wunny}) = -(2/5) \times \log_2(2/5) - (3/5) \times \log_2(3/5)$$
$$= 0.971 \text{ บิต}$$

$$E_{\text{temperature}}(\text{Outlook} = \text{Sunny}) = (2/5) \times (-(0/2) \times \log_2(0/2) - (2/2) \times \log_2(2/2)$$
$$+ (2/5) \times (-(1/2) \times \log_2(1/2) - (1/2) \times \log_2(1/2)$$
$$+ (1/5) \times (-(1/2) \times \log_2(1/2) - (0/1) \times \log_2(0/2)$$
$$= 0.4 \text{ บิต}$$

$$\text{Gain}(\text{Temperature}) = I(\text{Outlook} = \text{Sunny}) - E_{\text{temperature}}(\text{Outlook} = \text{Sunny})$$
$$= 0.971 - 0.4$$
$$= 0.571 \text{ บิต}$$

แอททริบิวต์ที่เหลือที่สามารถถูกเลือกมาเป็นตัวทดสอบเพื่อจัดกลุ่มของข้อมูลฝึก คือ แอททริบิวต์ Humidity และ Windy สามารถคำนวณค่าเกนจากการเลือกแต่ละแอททริบิวต์

ได้ดังนี้

$$\text{Gain(Humidity)} = I(\text{Outlook} = \text{Sunny}) - E_{\text{humidity}}(\text{Outlook} = \text{Sunny})$$
$$= 0.971 - 0$$
$$= 0.971 \text{ บิต}$$

$$\text{Gain(Windy)} = I(\text{Outlook} = \text{Sunny}) - E_{\text{windy}}(\text{Outlook} = \text{Sunny})$$
$$= 0.971 - 0.951$$
$$= 0.020 \text{ บิต}$$

จะเห็นว่าแอททริบิวต์ที่ให้ค่าเกนสูงที่สุดคือ Humidity ดังนั้นแอททริบิวต์นี้จึงถูกเลือกเป็นโหนดระดับที่ 2 ต่อจาก Outlook = Sunny และยังคงเหลือโหนดลูกทางขวาของโหนด Outlook (Outlook = Rainy) ที่ต้องพิจารณาเลือกแอททริบิวต์ และจากวิธีการคำนวณค่าเกนที่แสดงด้วยตัวอย่างก่อนหน้านี้ สามารถเลือกได้ว่าแอททริบิวต์ Windy จะให้ค่าเกนสูงที่สุด จึงถูกเลือกเป็นโหนดระดับที่ 2 ต่อจาก Outlook = Rainy กระบวนการสร้างต้นไม้ตัดสินใจ จะสิ้นสุดเมื่อโหนดใบเป็นกลุ่มของข้อมูลเดียวกันทั้งหมด และจะได้โครงสร้างของต้นไม้ตัดสินใจเป็นดังแผนภาพที่ 7

2. ค่ามาตรฐานอัตราส่วนเกน (Gain Ratio Criterion)

ในอัลกอริทึม ID3 จะใช้ค่ามาตรฐานเกนเป็นหลักในการเลือกแอททริบิวต์ที่จะใช้เป็นโหนดรากของต้นไม้ตัดสินใจหรือของต้นไม้ย่อย แต่ในอัลกอริทึม C4.5 ได้เพิ่มการใช้ค่ามาตรฐานอัตราส่วนเกนในการตัดสินใจเลือกแอททริบิวต์ที่จะใช้เป็นโหนดรากเข้ามาด้วย เนื่องจากค่ามาตรฐานเกนจะมีความลำเอียงอย่างมากกับข้อมูลที่ประกอบด้วยแอททริบิวต์ที่มีค่าที่เป็นไปได้จำนวนมาก ๆ เช่น ชุดข้อมูลที่ประกอบด้วยแอททริบิวต์หมายเลขประจำตัว ซึ่งมีค่าไม่ซ้ำกันในแต่ละตัวอย่าง ถ้าแบ่งข้อมูลตามแอททริบิวต์นี้จะทำให้ได้จำนวนตัวอย่างเพียง 1 ตัวอย่างต่อ 1 กิ่งของต้นไม้ และเมื่อคำนวณค่า Entropy จากการแบ่งตัวอย่างบนแอททริบิวต์นี้จะได้เท่ากับ 0 ทำให้ค่าเกนที่ได้ของแอททริบิวต์นี้มีค่าสูงที่สุด (ก้องศักดิ์ จงเกษมวงศ์. 2543 อ้างอิงใน นฤพนต์ ว่องประชานุกูล. 2548 : 12)

จากตัวอย่างข้อมูลการตัดสินใจเล่นกอล์ฟในตารางที่ 1 ถ้าใช้แอททริบิวต์ ID ในการจัดกลุ่มข้อมูลจะต้องการปริมาณข้อมูลประกอบการตัดสินใจเพื่อจำแนกกลุ่มดังนี้

$$E(ID) = (1/14) \times (-(0/1) \times \log_2(0/1) - (1/1) \times \log_2(1/1) + \ldots$$
$$+ (1/14) \times (-(0/1) \times \log_2(0/1) - (1/1) \times \log_2(1/1))$$
$$= 0 \text{ บิต}$$

เมื่อแบ่งตัวอย่างบนแอททริบิวต์นี้จะได้ค่า Entropy เท่ากับ 0 ดังนั้นค่ามาตรฐานเกนของแอททริบิวต์นี้จะเท่ากับปริมาณข้อมูลที่ต้องการจะระบุว่าข้อมูลหนึ่งเรคคอร์ดอยู่ในกลุ่ม P หรือ N ที่โหนดราก ซึ่งมีค่าเท่ากับ 0.940 บิต ทำให้ค่ามาตรฐานบิตเกนนี้มีค่าสูงกว่าแอททริบิวต์อื่น ๆ ดังนี้แอททริบิวต์ ID นี้จะถูกเลือกมาเป็นตัวทดสอบเพื่อจัดกลุ่มของข้อมูลฝึก

ดังนั้นจะเห็นว่า การวัดค่ามาตรฐานเกนจะได้ค่าสูงเมื่อแอททริบิวต์นั้นมีค่าที่เป็นไปได้จำนวนมาก ๆ ซึ่งไม่สามารถนำมาใช้เป็นโหนดของต้นไม้ เพื่อทำนายกลุ่มของข้อมูลใหม่ที่ไม่เคยเห็นได้อย่างถูกต้อง จึงต้องแก้ไขความลำเอียงนี้ โดยการปรับค่าเกนให้ถูกต้องโดยใช้ค่าสารสนเทศการแบ่งแยก (Split Information) ของแต่ละแอททริบิวต์ เพื่อใช้คำนวณค่ามาตรฐานอัตราส่วนเกน

ถ้ากำหนดให้ T แทนชุดของข้อมูลฝึก เมื่อแบ่งตัวอย่างโดยใช้แอททริบิวต์A จะได้ชุดของตัวอย่างย่อยในแต่ละกิ่งเป็น $\{t_1, t_2, \ldots, t_v\}$ จำนวน v ชุด ตามค่าที่เป็นไปได้ของแอททริบิวต์ A และสามารถคำนวณค่าสารสนเทศการแบ่งแยกได้ดังนี้

$$\text{ค่าสารสนเทศการแบ่งแยก} = -\sum_{i=1}^{v} \frac{|t_i|}{|T|} \log_2\left(\frac{|t_i|}{T}\right) \quad (5)$$

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

ค่าสารสนเทศการแบ่งแยกนี้จะแสดงถึงระดับการกระจายของข้อมูล เมื่อแบ่งข้อมูลตัวอย่าง T เป็น v ชุดย่อยตามค่าที่เป็นไปได้ของแอททริบิวต์ A โดยค่านี้จะมีค่าสูงสุด เมื่อ $|t_i|$ เป็น 1 เท่ากันในทุกกิ่ง และจะลดลงเมื่อค่า $|t_i|$ เพิ่มขึ้น เมื่อนำค่านี้ไปหารค่ามาตรฐานเกนจะได้ค่ามาตรฐานอัตราส่วนเกน ซึ่งช่วยแก้ไขความลำเอียงที่เกิดขึ้นของค่ามาตรฐานเกนได้ โดยทำให้ค่ามาตรฐานอัตราส่วนเกนของแอททริบิวต์ที่มีค่าที่เป็นได้จำนวนมากถูกปรับลดลง (ก้องศักดิ์ จงเกษมวงศ์, 2543 อ้างอิงใน นฤพนต์ ว่องประชานุกูล. 2548 : 13)

ค่ามาตรฐานอัตราส่วนเกน = ค่ามาตรฐานเกน / ค่าสารสนเทศการแบ่งแยก

จากตัวอย่างข้อมูลการตัดสินใจเล่นกอล์ฟในตารางที่ 1 สามารถคำนวณค่าอัตราส่วนเกนของแอททริบิวต์ Outlook ได้ดังนี้

ค่าสารสนเทศการแบ่งแยก (Outlook) = $-(5/14) \times \log_2(5/14) - (4/14) \times \log_2(4/14)$

$-(5/14) \times \log_2(5/14)$

= 1.577 บิต

Gain Ratio (Outlook) = 0.247 / 1.577

= 0.156

และเมื่อแบ่งข้อมูลตัวอย่างด้วยแอททริบิวต์ Temperature, Humidity และ Windy สามารถคำนวณค่า Gain Ratio ได้ดังนี้

Gain Ratio (Temperature) = 0.029 / 1.362

= 0.021

Gain Rratio (Humidity) = 0.152 / 1.000

= 0.152

Gain Ratio (Windy) = 0.048 / 0.985

= 0.049

จะเห็นว่าแอททริบิวต์ที่ให้ค่า Gain Ratio สูงที่สุดคือ Outlook เช่นเดียวกับการคำนวณค่า Information Gain ดังนั้นแอททริบิวต์ Outlook จึงถูกเลือกเป็นโหนดรากของต้นไม้ตัดสินใจ และจะสร้างต้นไม้ตัดสินใจต่อไปจนกระทั่งสามารถจัดกลุ่มของข้อมูลให้เป็นกลุ่มเดียวกันได้ทั้งหมด (นฤพนต์ ว่องประชานุกูล. 2548 : 4 – 14)

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

งานวิจัยที่เกี่ยวข้อง

1. งานวิจัยภายในประเทศ

ไพฑูรย์ จันทร์เรือง (2550 : 22 - 71) พัฒนาระบบสนับสนุนตัดสินใจเลือกสาขาการเรียนของนักศึกษาระดับปริญญาตรีโดยใช้เทคนิคการตัดสินใจ มีวัตถุประสงค์เพื่อพัฒนาตัวแบบใช้ในการทำนายแนวโน้มผลกาเรียนของนักศึกษา เมื่อเข้ามาศึกษาต่อในระดับปริญญาตรี คณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคล

ปริญญาตรี คณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ โดยใช้เทคนิคต้นไม้ตัดสินใจ ข้อมูลที่ใช้ในการวิจัยครั้งนี้ เป็นข้อมูลของนักศึกษาระดับปริญญาตรีของคณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ เก็บรวบรวมไว้ตั้งแต่ปีการศึกษา 2545 ถึงปีการศึกษา 2550 จำนวน 4,405 รายการ

ผลจากการวิจัย สร้างตัวแบบขึ้น 12 ตัวแบบ ให้ค่าความถูกต้องในการทำนายอยู่ในช่วง 78.46 % ถึง 92.31 % และมีจำนวนกฎการจำแนกอยู่ระหว่าง 60 – 336 กฎ โดยที่กฎส่วนใหญ่จะให้ผลลัพธ์อยู่ในช่วงกลางของข้อมูล คือ มีผลการทำนายเป็นระดับพอใช้และระดับปานกลาง ซึ่งน่าจะขึ้นกับจำนวนข้อมูลที่นำมาใช้ในการเรียนรู้ของอัลกอริทึม และสภาพความเป็นจริงของข้อมูลเป็นหลัก เมื่อนำมาสร้างเป็นโปรแกรมระบบสนับสนุนการตัดสินใจ ส่วนใหญ่จะอยู่ในระดับพอใช้และปานกลางเช่นกัน

สิทธิโชค มุกดาสกุลภิบาล (2551 : บทคัดย่อ) ศึกษาการวัดประสิทธิภาพของขั้นตอนวิธี ตัวจำแนก C4.5, ADTree และ Naïve Bayes ในการจำแนกข้อมูลการซุกซ่อนสิ่งเสพย์ติดสำหรับไปรษณีย์ระหว่างประเทศ เพื่อนำไปใช้กับข้อมูลการซุกซ่อนสิ่งเสพย์ติดสำหรับไปรษณีย์ระหว่างประเทศของเจ้าหน้าที่ศุลกากร ในการช่วยเหลือให้สามารถที่จะทำงานได้สะดวกสบายมากยิ่งขึ้น โดยงานวิจัยเริ่มตั้งแต่ขั้นตอนการเก็บรวบรวมข้อมูลทั้งหมดโดยวิธีการเอามาจากข้อมูลเก่า ๆ ที่เจ้าหน้าที่ได้ทำการเก็บไว้ และยังมีข้อมูลใหม่ที่ไปทำการสุ่มเก็บเพื่อเพิ่มเติมกับเจ้าหน้าที่ศุลกากร เพื่อให้ได้ปริมาณข้อมูลที่เพียงพอต่อการวิจัย แล้วยังมีการจัดการข้อมูลให้เหมาะสม, ปรับปรุงข้อมูล, จัดกลุ่มข้อมูล และแยกประเภท เพื่อให้พร้อมในการนำไปทดสอบกับตัวจำแนกทั้ง 3 ชนิด โดยที่ได้ทำการประเมินในแง่ของค่าความแม่นยำ, ค่าส่วนเบี่ยงเบนมาตรฐาน และระยะเวลา เพื่อประเมินว่าตัวจำแนกใดเหมาะสมกับข้อมูลที่ใช้ในงานวิจัยดังกล่าว ข้อมูลที่จะนำมำการวัดผลมีทั้งข้อมูลที่เป็นสองฉลาก และหลายฉลาก

ผลสรุปที่ได้ออกมาพบว่า ตัวจำแนกแต่ละตัวมีข้อดีข้อเสียที่แตกต่างกัน ซึ่งในแง่ของความถูกต้องโดยรวม ADTree สามารถทำงานได้ดี อยู่ในช่วงร้อยละ 94 – 97 แต่ในด้านความเร็วในการทำงานพบว่า ADTree ใช้เวลาในการประมวลผลนานที่สุดจากตัวจำแนกทั้ง 3 ซึ่งเบย์ส และ C4.5 ใช้เวลานการประมวลผลที่ใกล้เคียงกัน และ C4.5 มีความสามารถที่ทำนายข้อมูลที่ไม่มีสิ่งเสพย์ติดซุกซ่อนได้เป็นอย่างดี อยู่ที่ช่วงร้อยละ 99.0 – 99.9 แต่เบย์สนั้นให้ผลการทำนายข้อมูลสิ่งที่มีสิ่งเสพย์ติดซุกซ่อนได้ดี อยู่ที่ช่วงร้อยละ 21.0 – 84.4 แต่จะ

มากกว่าตัวจำแนกอื่นเมื่อมีขนาดของข้อมูลการเรียนรู้ที่สูงขึ้น ที่ร้อยละ 50 สำหรับข้อมูลแบบสองฉลาก และร้อยละ 70 สำหรับข้อมูลแบบหลายฉลาก

สุรเดช บุญลือ (2551 : 873-878) โครงงานวิจัยและพัฒนานี้ มีจุดประสงค์เพื่อการศึกษาและพัฒนาระบบจัดการคัดกรองผู้บริจาคโลหิตหมู่พิเศษ โดยใช้เทคนิคต้นไม้ตัดสินใจ ระบบสามารถคัดกรองรายชื่อผู้บริจาคที่มีความพร้อมให้บริจาค โดยการนำข้อมูลผู้บริจาคทั้งหมด ซึ่งเป็นฐานข้อมูลขนาดใหญ่มาผ่านตัวแบบต้นไม้ตัดสินใจ ซึ่งใช้อัลกอริทึม C4.5 ในท้ายที่สุดจะได้รายชื่อของผู้บริจาคที่มีคุณสมบัติตรงตามที่ต้องการ นอกจากนั้นเจ้าหน้าที่สามารถส่งข้อความสั้นผ่านระบบไปยังผู้บริจาคเพื่อเป็นการแจ้งเตือน เมื่อใกล้ครบกำหนดบริจาคครั้งต่อไป ซึ่งระบบบจะทำการคำนวณวันครบกำหนดครั้งต่อไปโดยอัตโนมัติ และสามารถจัดการข้อมูลของผู้บริจาค รวมไปถึงข้อมูลการติดต่อระหว่างเจ้าหน้าที่กับผู้บริจาค ซึ่งถือเป็นส่วนที่สำคัญในการใช้เป็นข้อมูลประกอบการตัดสินใจของตัวแบบต้นไม้ตัดสินใจ เครื่องมือที่ใช้ในการพัฒนา ได้แก่ Microsoft Visual Studio 2005 และ Microsoft SQL Server 2005 ภายหลังจากการพัฒนาระบบนี้ ได้ทำการประเมินคุณภาพโดยบุคคลสองกลุ่มคือ ผู้เชี่ยวชาญจากศูนย์บริการโลหิตแห่งชาติ และผู้เชี่ยวชาญด้านการพัฒนาซอฟต์แวร์ เพื่อหาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน พบว่าค่าเฉลี่ยของผลการประเมินของผู้เชี่ยวชาญซอฟต์แวร์อยู่ในระดับ 3.88 (SD = 0.81) ซึ่งอยู่ในเกณฑ์ดี ค่าเฉลี่ยของผลการประเมินของผู้เชี่ยวชาญอยู่ในระดับ 4.69 (SD = 0.26) ซึ่งอยู่ในเกณฑ์ดีมาก สามารถสรุปได้ว่าโครงงานนี้มีประสิทธิภาพ สามารถนำไปประยุกต์ใช้งานได้

ลภัส บุทธิจักร์ (2552 : 873 - 878) งานวิจัยนี้ได้พัฒนาระบบการตัดสินใจศึกษาต่อในวิทยาลัยประเทศออสเตรเลีย ซึ่งได้ประยุกต์ใช้เทคนิคการแบ่งกลุ่มแบบโครงสร้างต้นไม้ตัดสินใจ (Decision Tree) กับข้อมูลประวัติผู้สมัคร เพื่อสร้างโมเดล จากนั้นจึงนำโมเดลที่ได้ใช้ในการพัฒนาระบบ การพัฒนาระบบเริ่มจากการศึกษาความต้องการ รวบรวมข้อมูล วิเคราะห์ ออกแบบระบบ พัฒนาระบบ และทดสอบระบบโดยพัฒนาระบบเป็นลักษณะเว็บแอพพลิเคชั่น (Web Application) เพื่อใช้ในการติดต่อกับผู้ใช้เครื่องมือที่ใช้ในการพัฒนาคือภาษา ASP และฐานข้อมูล Microsoft SQL Server 2000

จากผลการประเมินความพึงพอใจต่อระบบ ของผู้เชี่ยวชาญ ค่าเฉลี่ยเท่ากับ 4.12ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.42 และผลประเมินของผู้ใช้งานทั่วไปค่าเฉลี่ยเท่ากับ 4.01 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.44 สรุปผลได้ว่าระบบที่พัฒนาขึ้นมีระดับความพึงพอใจในการใช้งานอยู่ในระดับดี สามารถนำไปประยุกต์ใช้ได้จริง

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY

ศุภชัย ประคองศิลป์ (2552 : 893-898) พัฒนาระบบสนับสนุนการตัดสินใจอนุมัติโครงการที่อยู่อาศัย โดยใช้เทคนิคต้นไม้ตัดสินใจ มีวัตถุประสงค์เพื่อสร้างโมเดลที่มีความเหมาะสมและถูกต้อง สำหรับการอนุมัติลูกบ้านเข้าโครงการ โดยใช้เทคนิคต้นไม้ตัดสินใจและอัลกอริทึม ID3 ให้ความแม่นยำคิดเป็นร้อยละมากที่สุด เมื่อเทียบกับอัลกอริทึมอื่นที่เป็นเทคนิคต้นไม้ตัดสินใจเหมือนกัน โดยรูปแบบโมเดลสำหรับการอนุมัติลูกบ้านเข้าโครงการนั้น ได้มาจากข้อมูลในอดีตของบุคคลที่เคยเข้าร่วมโครงการทั้งหมด 963 รายการ และวัดความแม่นยำของโมเดลคือ 93.67 % และวัดความถูกต้องของโมเดลที่ระบบสร้างขึ้นมาโดยใช้โปรแกรม Weka 6.5.8 โดยใช้ข้อมูลสำหรับการสอน (Training Data) ที่เหมือนกัน ผลปรากฏว่าสามารถสร้างโมเดลได้เหมือนกับระบบที่ทางผู้วิจัยพัฒนาขึ้นมา จึงสรุปได้ว่าโมเดลนั้นมีความถูกต้องและแม่นยำอยู่ในระดับดีมาก และสามารถช่วยสนับสนุนการตัดสินใจในการอนุมัติลูกบ้านเข้าโครงการ อีกทั้งยังสามารถประยุกต์ใช้โมเดลกับงานสารสนเทศอื่น ๆ และสามารถใช้อัลกอริทึมอื่น ๆ ที่เป็นการแบ่งกลุ่มเหมือนกัน เพื่อช่วยในการตัดสินใจในการอนุมัติลูกบ้านเข้าโครงการ

ชุติมา อุตมะมุณีย์ (2553 : 39 – 48) วิจัยการพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์สำหรับการเลือกสาขาวิชาของนักศึกษาระดับอุดมศึกษา มีจุดมุ่งหมายเพื่อสร้างตัวแบบสำหรับหาปัจจัยที่มีผลต่อการตัดสินใจในการเลือกสาขาวิชาเรียนของนักศึกษา และพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจอย่างอัตโนมัติแบบออนไลน์สำหรับแนะแนวทางการเลือกสาขาวิชาเรียนของนักศึกษาในระดับอุดมศึกษา งานวิจัยนี้ได้นำข้อมูลมาวิเคราะห์อัลกอริทึมที่มีประสิทธิภาพสูงสุด และศึกษาตัวแปรที่มีผลต่อการเลือกสาขาวิชาเพื่อเป็นแนวทางให้นักศึกษาในการใช้ประกอบกับการตัดสินใจ ใช้ข้อมูลจากกลุ่มตัวอย่างของนักศึกษาจากมหาวิทยาลัยรัฐบาลและเอกชน จำนวน 9 แห่ง และในการทดลองจะทำการแบ่งกลุ่มข้อมูลและวัดผลโดยการเปรียบเทียบประสิทธิภาพของแต่ละอัลกอริทึมด้วย Percentage Correct Procision Recall และ F-measure ผลการทดลองแสดงให้เห็นว่าอัลกอริทึมข่ายงานเบย์ (Bayesian Networks) มีประสิทธิภาพสูงกว่าอัลกอริทึมต้นไม้ตัดสินใจ (Decision Tree C4.5) ค่าเปอร์เซ็นต์ของการทำนายถูกต้อง (Percentage Correct) 91.35 % ค่าที่สามารถสืบค้นคำตอบสูงสุด (Precision) 0.94 ค่าที่ได้จากการตรวจพบข้อมูล (Recall) 0.89 และค่าที่แสดงความสัมพันธ์ระหว่างค่าที่สามารถสืบค้นคำตอบสูงสุดกับค่าที่ได้จากการตรวจพบข้อมูล (F-measure) 0.90 จึงเลือกเทคนิคข่ายงานเบย์มาใช้ในการวิจัยครั้งนี้

ผลการพัฒนาตัวแบบสนับสนุนการตัดสินใจในการเลือกสาขาวิชาเรียนของนักศึกษามีความถูกต้องอยู่ในระดับมากเนื่องจากมีความแม่นยำในการทำนายผลถึงร้อยละ 91.35 และมีค่าสัมประสิทธิ์การตัดสินใจ สามารถอธิบายการเปลี่ยนแปลงของตัวแปรตามได้ร้อยละ 89.50 ดังนั้นจึงถือว่าตัวแบบระบบสนับสนุนการตัดสินใจในการเลือกสาขาวิชาเรียนของนักศึกษามีความถูกต้องอยู่ในระดับมาก และจากการวเคราะห์ระดับความพึงพอใจการใช้งานระบบ จากผลการวิจัยพบว่า นักศึกษามีความพึงพอใจโดยรวมอยู่ในระดับมาก

ชัดชัย แก้วตา (2553 : 308 – 313) ประยุกต์ใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree Techmiques) เพื่อจำแนกความผิดเกี่ยวกับการวินิจฉัยคดี งานวิจัยนี้ศึกษาปัจจัยที่มีผลต่อการพิจารณาข้อเท็จจริงสำหรับระบุโทษที่จะได้รับของผู้กระทำความผิด เพื่อสร้างและนำเสนอแบบจำลองสำหรับสนับสนุนการวินิจฉัยคดี โดยจำแนกคดีออกเป็นกลุ่มตามมาตรา ด้วยเทคนิคต้นไม้ตัดสินใจ และเชื่อมโยงความสัมพันธ์ของคุณลักษณะ (Attribute) ขององค์ประกอบความผิดเข้ากับบทบัญญัติของกฎหมายที่อยู่ในรูปของประมวลกฎหมาย

ผลการทดลอง การสร้างตัวแบบจำลองวินิจฉัยคดีได้ใช้ข้อมูลทดลองจำนวน 4,026 ความผิด และเลือกวิธี Cross-validation จำนวน Folds 5, 10, 15, 4026 ตามลำดับ โดยเปรียบเทียบขั้นตอนวิธี ID3 และ C4.5 พบว่าขั้นตอนวิธี C4.5 ให้ค่าความถูกต้องในการจำแนกข้อมูลในชุดทดสอบสูงกว่าขั้นตอนวิธี ID3 ซึ่งให้คำตอบที่ถูกต้อง 96.82 %

นวรัตน์ ไวยาวัจมัย (2554 : 582 – 587) พัฒนาระบบประเมินค่ารักษาพยาบาลผู้ป่วยที่เข้ารับการผ่าตัด โดยใช้เทคนิคต้นไม้ตัดสินใจ กรณีศึกษาโรงพยาบาลพญาไท 2 มีวัตถุประสงค์ใช้เป็นแนวทางในการประเมินค่ารักษาพยาบาลผู้ป่วยที่เข้ารับการผ่าตัด และอำนวยความสะดวกแก่เจ้าหน้าที่ในการประเมินค่ารักษาพยาบาล โดยงานวิจัยนี้นำหลักการของต้นไม้ตัดสินใจและปัจจัยที่นำมาวิเคราะห์ได้แก่ อายุ เพศ จำนวนวันนอน รหัสโรค รหัสหัตถการและค่ารักษาพยาบาล พัฒนาระบบในรูปแบบเว็บแอปพลิเคชั่น เริ่มต้นจากการออกแบบอัลกอริทึม สำหรับเทคนิคการแบ่งประเภท เพื่อนำมาใช้ในการวิเคราะห์ สร้างและทดสอบโมเดลเพื่อคำนวณหาค่ารักษาพยาบาล จากนั้นจึงทำการออกแบบและพัฒนาระบบประเมินค่ารักษาพยาบาลผู้ป่วยที่เข้ารับการผ่าตัด โดยใช้เทคนิคต้นไม้ตัดสินใจ ผลการประเมินจากกลุ่มผู้เชี่ยวชาญ จำนวน 5 คน และกลุ่มผู้ใช้งานทั่วไป 10 คน พบว่าผู้เชี่ยวชาญมีความคิดเห็นอยู่ในระดับดี โดยได้ค่าเฉลี่ยเท่ากับ 3.95 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.49 กลุ่มผู้ใช้งานทั่วไปมีความพึงพอใจอยู่ในระดับดี ได้ค่าเฉลี่ยเท่ากับ 4.03 และส่วน

มหาวิทยาลัยราชภัฏมหาสารคาม RAJABHAT MAHASARAKHAM UNIVERSITY

เบี่ยงเบนมาตรฐานเท่ากับ 0.48 ผลลัพธ์ของงานวิจัยมีค่าความถูกต้องเท่ากับ 84.01 % สามารถสรุปได้ว่า ระบบที่พัฒนาขึ้นมีคุณภาพอยู่ในระดับดี

2. งานวิจัยต่างประเทศ

เอริค สเวอร์ และ เดวิด เจ มัลวานี (Erick Swere and David J Mulvaney. 2003 : 15–17) พัฒนาซอฟต์แวร์ระบบการนำทางหุ่นยนต์ด้วยสถาปัตยกรรมต้นไม้ตัดสินใจ ภายใต้ข้อจำกัดของระบบคือ หน่วยความจำที่มีจำกัด และดำเนินการในเวลาจริง ผลพัฒนา ระบบการนำทางของหุ่นยนต์สามารถที่จะเรียนรู้เพิ่มขึ้นในเวลาจริง ภายใต้หน่วยความจำที่มีขนาดจำกัดในการฝังตัว สถาปัตยกรรมต้นไม้ตัดสินใจที่เลือกใช้มีความสามารถในการทำให้ระบบการนำทางของหุ่นยนต์มีความน่าเชื่อถือและแม่นยำ ตอบสนองความต้องการต่อไปนี้ นำทางในสภาพแวดล้อมที่มีการเปลี่ยนแปลง รับมือกับการเปลี่ยนแปลงที่เกิดขึ้นอย่างไม่คาดคิด และปฏิบัติงานที่ได้รับมอบหมายโดยไม่ต้องอาศัยข้อมูลเดิม การทำงานของหุ่นยนต์จะเรียนรู้จากสิ่งแวดล้อมด้วยตัวหุ่นยนต์ และทำงานโดยเริ่มจากความรู้ที่มีหรือไม่มีจากสภาพแวดล้อมการทำงานจริง

คอนราด เอส. ทัคเกอ (Conrad S. Tucker. 2009 : 041004-1–041004-14) ศึกษาการจำแนกข้อมูลของสินค้าด้วยต้นไม้ตัดสินใจเพื่อเพิ่มประสิทธิภาพการออกแบบผลงาน การวิจัยนี้เป็นการพัฒนาวิธีการในการทำนายความต้องการของลูกค้าและนำมาออกแบบผลิตภัณฑ์ในสายพันธุ์ของผลิตภัณฑ์ โดยการทำเหมืองข้อมูลและการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของสินค้าที่ผู้ผลิตสร้างขึ้น โดยใช้ข้อมูลความต้องการของแต่ละบุคคลที่คาดการณ์ไว้ ตามแนวคิดผลิตภัณฑ์ในระดับการออกแบบทางวิศวกรรม กำหนดคุณลักษณะเฉพาะ (ทั้งชุดของแนวคิดผลิตภัณฑ์ที่เป็นไปได้) จะแคบลงถึง 46 แนวคิดของผลิตภัณฑ์ ซึ่งแนวคิดรวมอยู่ในการออกแบบผลิตภัณฑ์และกระบวนการคัดเลือก การใช้เทคนิคต้นไม้ตัดสินใจทำให้ผลิตภัณฑ์นั้นมีโอกาสประสบผลสำเร็จในตลาด องค์กรสามารถตัดสินใจเลือกกลุ่มผลิตภัณฑ์ที่ดีที่สุดในการผลิต และทำให้องค์กรเพิ่มผลกำไร รวมทั้งตอบสนองความต้องารของลูกค้าอย่างถูกต้องและมีประสิทธิภาพ

เดวิ ปราสาท บูคยา (Devi Prasad Bhukya. 2010 : 660–665) นำเสนองานวิจัยวิธีการจัดหมวดหมู่ข้อมูลโดยใช้ต้นไม้ตัดสินใจแบบ AVL ซึ่งจะช่วยเพิ่มประสิทธิภาพในการจำแนกระดับข้อมูลจำนวนมากจากฐานข้อมูลความรู้ที่มีขนาดใหญ่ การจัดหมวดหมู่ข้อมูลโดยทั่วไปจะใช้อัลกอริทึมทั่วไป โดยเรียงลำดับชั้นจากกฎของข้อมูลชุดฝึก ตามลักษณะของข้อมูล ซึ่งผลการใช้ต้นไม้ตัดสินใจแบบ AVL จะทำให้มีประสิทธิภาพดีกว่าอัลกอริทึมทั่วไป

คือ มีความยืดหยุ่น เช่น เราวางแผนที่ใช้การเรียนรู้ การปรับตัวของระบบเพื่อกำหนดความสำคัญของข้อมูล ต้นไม้ตัดสินใจแบบ AVL สามารถปรับความแตกต่างของข้อมูลให้มีความเชื่อมโยงจากระดับหนึ่งไปยังอีกระดับหนึ่งได้ ช่วยให้ผู้ใช้สามารถดูชั้นของข้อมูลได้และสามารถหยุดการค้นหาข้อมูลตามความต้องการของผู้ใช้ได้

เอนูจ ชาร์มา (Anuj Sharma. 2012 : 37–47) ศึกษาการตรวจสอบการทุจริตการบัญชีการเงินบนพื้นฐานของเทคนิคเหมืองข้อมูล ในการตรวจสอบการทุจริตการบัญชีการเงินเป็นการจัดการข้อมูลที่มีปริมาณมาก ข้อมูลที่มีขนาดใหญ่และมีความซับซ้อนมาก ในด้านการตรวจสอบการทุจริตการบัญชีการเงินจะตรวจสอบค่าความผิดปกติที่มีมาก เทคนิคการทำเหมืองข้อมูลเหมาะสำหรับการตรวจสอบการทุจริตการบัญชีการเงิน จะแยกความแตกต่างข้อมูลว่าเป็นข้อมูลจริงหรือหลอกลวง จึงจะสมควรได้รับการตรวจสอบมากขึ้น

บี วี ชอว์ดารี่ (B V Chowdary. 2012 : 475–479) พัฒนาวิธีการใหม่สำหรับต้นไม้ตัดสินใจ การจัดหมวดหมู่ของข้อมูล โดยใช้โครงข้อมูลที่เรียกว่า Peano Count Tree (P-tree) การศึกษาครั้งนี้ใช้อัลกอริทึม ID3 กับโครงสร้างข้อมูล P-tree เพื่อหาประสิทธิภาพการทำงาน ผลการทดลองพบว่าขนาดข้อมูลขนาดใหญ่ทำให้มีความเร็วในการสร้างข้อมูลเปรียบเทียบค่าในการสแกนข้อมูลมีความแตกต่างกัน สำหรับข้อมูลที่มีความแตกต่างกันสรุปได้ว่า การจัดหมวดหมู่ของข้อมูลโดยใช้ P-tree จะมีประสิทธิภาพ และมีความเร็วความยืดหยุ่นของการตัดสินใจตามลักษณะนามเหนี่ยวนำ

มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY