# บทที่ 1

# บทนำ

## ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันความก้าวหน้าทางเทคโนโลยีด้านคอมพิวเตอร์เป็นไปอย่างรวดเร็ว ทำให้มีโปรแกรมคอมพิวเตอร์ที่อำนวยความสะดวกให้แก่นักวิจัยเกิดขึ้นมากมาย ที่สำคัญคือโปรแกรมสำเร็จรูปทางสถิติ ซึ่งสร้างขึ้นมาเพื่อช่วยในการวิเคราะห์ข้อมูล แต่โปรแกรมเหล่านี้ไม่สามารถทำการวิเคราะห์ข้อมูลทางสถิติด้วยตัวของมันเองได้ เป็นแต่เพียงเครื่องมือช่วยคำนวณค่าสถิติให้เท่านั้น ผู้ใช้โปรแกรมสำเร็จรูปทางสถิติ จำเป็นต้องมีความรู้ทางสถิติอย่างเพียงพอ จึงสามารถใช้ประโยชน์จากโปรแกรมนี้ได้ โดยผู้ใช้ต้องมีความรู้ก่อนว่า ข้อมูลที่จะทำการวิเคราะห์นั้นต้องใช้วิธีการทางสถิติอะไรมาทำการวิเคราะห์จึงจะสรุปข้อสงสัยได้ แล้วจึงเขียนชุดคำสั่งเพื่อสั่งโปรแกรมสำเร็จรูปให้คำนวณค่าสถิติที่ต้องการออกมา หลังจากนั้นผู้วิจัยจึงนำไปตีความ และสรุปข้อสมมุติฐานต่อไป

ผู้ทำการวิจัยมักประสบปัญหาว่า จะวิเคราะห์ข้อมูลที่รวบรวมไว้ได้อย่างไร จะใช้สถิติตัวไหนมาวิเคราะห์ข้อมูล จึงจะถูกต้องและตอบข้อสงสัยได้ ผู้วิจัยมักไปปรึกษานักสถิติ แต่นักสถิติจะยังไม่สามารถตอบปัญหาได้ทันที ต้องใช้เวลาส่วนหนึ่งทำความเข้าใจปัญหา ความต้องการ วัตถุประสงค์ของการวิจัยและลักษณะของข้อมูลที่เก็บมา โดยพยายามหาสถิติที่เหมาะสมมาใช้วิเคราะห์ บางครั้งข้อมูลที่มีอยู่ยังไม่เพียงพอที่จะนำมาใช้ตอบวัตถุประสงค์ที่ตั้งไว้ได้ เพราะไม่มีการเตรียมเก็บข้อมูลให้ครอบคลุมเพียงพอต่อการวิเคราะห์ข้อมูลไว้ล่วงหน้า ทางแก้คือต้องมีการวางแผนการวิจัย ก่อนการเก็บข้อมูลจริง นักสถิติจะกำหนด ลักษณะ ชนิดของตัวแปร วิธีการเก็บข้อมูลไว้อย่างละเอียดเพียงพอต่อการวิเคราะห์ข้อมูล นักสถิติที่กล่าวถึงนี้ต้อง

เป็นผู้ที่มีความเชี่ยวชาญมีประสบการณ์ ทั้งการวิเคราะห์ข้อมูล และการวิจัยในเรื่องที่สนใจนั้น จึงจะให้คำปรึกษาที่ถูกต้องได้ เพราะมีวิธีการทางสถิติมากมาย แค่ข้อสมมุติ (Assumptions) ของสูตรสถิติอย่างเดียว ยังไม่พอเพียงต่อการเลือกใช้วิธีการทางสถิติให้ถูกต้องได้ ต้องมีประสบการณ์ ความรู้ความชำนาญในเนื้อหาของเรื่องที่ทำวิจัยด้วย บางครั้งก็ไม่สามารถหานักสถิติลักษณะนี้ได้ เนื่องจากขาดแคลน หรือไม่มีเวลาเพียงพอที่จะทำความเข้าใจเรื่องที่ทำการศึกษา หรือให้คำปรึกษาที่ ไม่ละเอียดเพียงพอต่อการวิเคราะห์ข้อมูล

ปัญหาเหล่านี้สามารถแก้ไขได้ ถ้าเราได้นำเอาความรู้และประสบการณ์ ของนักวิจัยที่มีความชำนาญในเรื่องที่เราสนใจศึกษา และของนักสถิติที่มีประสบการณ์ ในการวิเคราะห์ข้อมูลการวิจัย มารวบรวมไว้เป็นฐานความรู้ (Knowledge Base) เก็บด้วยเครื่องคอมพิวเตอร์ และสามารถดึงมาใช้ ให้คำปรึกษาในการ วางแผนการวิเคราะห์ข้อมูลเมื่อต้องการได้ จะช่วยให้การวิเคราะห์ข้อมูลที่เก็บมา สามารถตอบวัตถุประสงค์ของการวิจัยได้ดียิ่งขึ้น หลักการนี้เรียกว่า ระบบผู้เชี่ยวชาญ (Expert System) ผู้ใช้สามารถนำระบบนี้ ไปทดลองใส่ตัวแปรที่คาดว่าจะเก็บข้อมูลและสมมติฐานการวิจัย โปรแกรมจะแนะนำว่าข้อมูลของตัวแปรเหล่านี้ สามารถนำมาวิเคราะห์ทางสถิติเพื่อตอบสมมติฐานได้หรือไม่ ถ้าตอบได้ ระบบผู้เชี่ยวชาญจะแนะนำวิธีการทางสถิติที่จะนำมาใช้ คำสั่งโปรแกรมสำเร็จรูปทางสถิติเพื่อการคำนวณค่าสถิติที่ต้องการ และหลักการ สรุป ตีความ ผลลัพธ์ที่ออกมา แต่ถ้าไม่สามารถตอบสมมติฐานได้ จะมีข้อเสนอแนะเพื่อการแก้ไขและปรับปรุง ผู้ใช้สามารถทดลองกับระบบผู้เชี่ยวชาญนี้ได้หลายครั้ง จนเป็นที่แน่ใจว่า ข้อมูลที่จะเก็บรวบรวมนั้นสามารถนำ ไปใช้ตอบสมมติฐานที่สงสัยได้ จึงค่อยทำการเก็บข้อมูลจริง

**<u>วัตถุประสงค์ของการวิจัย</u>**

1. เพื่อสร้างระบบผู้เชี่ยวชาญที่ให้คำปรึกษาในการวางแผนการวิเคราะห์ข้อมูลทางสถิติ
2. เพื่อเป็นเครื่องมือให้คำแนะนำ การเลือกใช้สถิติอย่างถูกต้อง

## ขอบเขตของการวิจัย

การพัฒนาระบบผู้เชี่ยวชาญเพื่อการวางแผนการวิเคราะห์ข้อมูลทางสถิติ มีขอบเขตในเรื่องวิธีการแทนความรู้ วิธีการทางสถิติ วิธีการค้นหาในเครื่องอนุมาน และภาษาที่ใช้เขียนโปรแกรม ดังนี้

1. การจัดเก็บฐานความรู้ ทางสถิติ เป็นแบบฐานแห่งกฎ (Rule-based knowledge representation หรือ Production rules) แบบ IF-THEN-ELSE

2. ฐานความรู้ทางสถิติ เป็นข้อสมมุติการเลือกใช้วิธีการทางสถิติ ในวิธีการทดสอบสถิติต่อไปนี้

2.1 ความแตกต่างของข้อมูล 1 ชุด 2 ชุด และมากกว่า 2 ชุด ที่เป็นการทดสอบแบบมีพารามิเตอร์ (Parametric Test) และการทดสอบแบบไร้พารามิเตอร์ (Non-parametric Test)

2.2 การทดสอบความสัมพันธ์ของข้อมูลจัดกลุ่ม(Categories) ข้อมูลอันดับ(Rank) และข้อมูลปริมาณ การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression)

3. การจัดเก็บฐานความรู้งานวิจัย สรุปผลงานวิจัยที่ผ่านมา

4. เครื่องอนุมาน (Inference Engine) ใช้ทิศทางการค้นหาแบบลูกโซ่ไปข้างหน้า (Forward Chaining)

5. พัฒนาด้วยภาษา ซี บนเครื่องไมโครคอมพิวเตอร์พีซี

## วิธีการทางสถิติ (Statistical Methods)

วิธีการทางสถิติที่นำมาใช้ในการสร้างระบบผู้เชี่ยวชาญนี้ ได้นำมาเฉพาะวิธีการทดสอบทางสถิติที่มีการใช้ประจำ และเป็นสถิติพื้นฐานสำหรับนักวิจัยที่ควรทราบ มีการจำแนกตามวัตถุประสงค์ของการทดสอบได้ 2 ประเภทคือ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1. **การทดสอบความแตกต่าง**

การทดสอบความแตกต่าง เป็นการทดสอบข้อสงสัยหรือสมมุติฐานการวิจัยว่าจะมีความมากกว่า หรือน้อยกว่า หรือไม่เท่ากับ แบ่งตามจำนวนชุดข้อมูลเป็น 3 อย่างคือ

1.1 **ข้อมูล 1 ชุด**

เป็นการทดสอบข้อสงสัยว่าข้อมูลที่เก็บมาเป็นตัวอย่างมีการสุ่มมาจากประชากรที่มีการแจกแจงตามทฤษฎีที่สนใจหรือไม่ เช่น

1.1.1 การทดสอบโคลโมโกรอฟสไมร์นอฟ (Kolmogorov-Smirnov test) เป็นสถิติทดสอบที่นำมาใช้ทดสอบว่าข้อมูล ตัวอย่าง มีการสุ่มมาจากประชากรที่มีการแจกแจงปรกติ (Normal distribution) หรือการแจกแจงปัวส์ซอง (Poisson distribution) หรือการแจกแจงเอกรูป (Uniform distribution)

1.1.2 การทดสอบไบโนเมียล(Binomial test) เป็นสถิติทดสอบที่ใช้ทดสอบว่าข้อมูลตัวอย่าง สุ่มมาจากประชากรที่มีการแจกแจงทวินาม (Binomial distribution)

1.2 **ข้อมูล 2 ชุด**

เป็นการทดสอบข้อสงสัยว่า ข้อมูลที่เก็บมาเป็นตัวอย่าง 2 ชุดมีการสุ่มมาจากประชากรที่มีลักษณะของพารามิเตอร์ หรือลักษณะการแจกแจงที่แตกต่างกันหรือไม่ หรือมากกว่า หรือน้อยกว่าหรือไม่ ได้แก่

1.2.1 การทดสอบทีของตัวอย่างอิสระสองชุด (Two independent sample t-test) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของข้อมูลตัวอย่าง 2 ชุดที่สุ่มมาอย่างอิสระและมีการแจกแจงปรกติ

1.2.2 การทดสอบแมนวิทนียู (Mann-Whitney U -test) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่ามัธยฐานของข้อมูลตัวอย่าง 2 ชุดที่สุ่มมาอย่างอิสระ แต่ไม่มีการแจกแจงปรกติ

1.2.3 การทดสอบทีของตัวอย่างแบบคู่ (Paired t-test) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของข้อมูลตัวอย่าง 2 ชุดที่สุ่มมาอย่างมีความสัมพันธ์กันจัดเป็นคู่และมีการแจกแจงปรกติ

1.2.4 การทดสอบแมคเนมา (McNemar test) เป็นสถิติทดสอบที่ใช้ทดสอบความแตกต่างระหว่างลักษณะของข้อมูลตัวอย่างแบบทวิภาค (Dichotomous) 2 ชุดที่สุ่มมาจากคน หรือสัตว์ หรือสิ่งของอันเดียวกัน 2 ครั้ง

1.2.5 การทดสอบวิลคอกสันซายแลงค์ (Wilcoxon signed rank test) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างลักษณะของข้อมูลตัวอย่าง 2 ชุดที่สุ่มมาอย่างมีความสัมพันธ์กันแบบจัดเป็นคู่และไม่มีการแจกแจงปรกติ

1.3 **ข้อมูลมากกว่า 2 ชุด**

เป็นการทดสอบข้อสงสัยว่า ข้อมูลที่เก็บมาเป็นตัวอย่างมากกว่า 2 ชุดมีการสุ่มมาจากประชากรที่มีลักษณะของพารามิเตอร์ หรือลักษณะการแจกแจงที่แตกต่างกันอย่างน้อย 1 คู่ หรือไม่ และคู่ไหนแตกต่างกันบ้าง ได้แก่

1.3.1 การวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA หรือ Completely Random design) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของข้อมูลตัวอย่างมากกว่า 2 ชุดที่สุ่มมาอย่างอิสระและมีการแจกแจงปรกติ

1.3.2 การวิเคราะห์ความแปรปรวนแบบสองทาง (Two-way ANOVA หรือ Randomized block design) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของข้อมูลตัวอย่างมากกว่า 2 ชุดที่สุ่มมาอย่างมีความสัมพันธ์กันจัดเป็นบล็อก (Block) และมีการแจกแจงปรกติ

1.3.3 การทดสอบการทดลองแบบแฟกทอเรียล (Factorial Experiment) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของข้อมูลตัวอย่างมากกว่า 2 ชุดที่สุ่มมาอย่างอิสระและมีการแจกแจงปรกติ โดยมีทรีทเมนต์ตั้งแต่ 2 ประเภทขึ้นไปเกิดร่วมกันเป็นทรีทเมนต์ร่วม (Treatment combination)

1.3.4 การทดสอบครูสเคลวาลลิส (Kruskal-Wallis test) เป็นสถิติทดสอบที่ใช้ทดสอบความแตกต่างระหว่างค่ามัธยฐานของข้อมูลตัวอย่างมากกว่า 2 ชุดที่สุ่มมาอย่างอิสระแต่ไม่มีการแจกแจงปรกติ

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1.3.5 การทดสอบเฟลดแมน (Friedman test) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่ามัธยฐานของข้อมูลตัวอย่างมากกว่า 2 ชุด ที่สุ่มมาอย่างมีความสัมพันธ์ กันจัดเป็นบล็อก (Block) แต่ไม่มีการแจกแจงปรกติ

1.3.6 การวิเคราะห์ความแปรปรวนร่วม (Analysis of Covariance) เป็นสถิติทดสอบ ที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของข้อมูลตัวอย่างมากกว่า 2 ชุดที่สุ่มมาอย่างอิสระและมีการแจกแจงปรกติ ทีมีตัวแปรร่วม (Covariate variable)

## 2. การทดสอบความสัมพันธ์

เป็นการทดสอบความความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปรขึ้นไปโดยทดสอบข้อสงสัยหรือสมมุติฐานการวิจัยว่าจะมีความสัมพันธ์กันหรือไม่ มากน้อยแค่ไหน และสัมพันธ์กันในทิศทางแปรตามกัน หรือผกผันกัน มีการแบ่งตามลักษณะข้อมูลดังนี้

### 2.1 ข้อมูลจัดกลุ่ม

เมื่อข้อมูลเป็นข้อมูลคุณภาพที่จัดเป็นชั้น หรือข้อมูลปริมาณที่มีการจัดเป็นชั้น สนใจทดสอบว่าข้อมูลทั้งสองมีความสัมพันธ์กันหรือไม่ แบ่งสถิติทดสอบตามลักษณะข้อมูลได้ดังนี้

2.1.1 การทดสอบความสัมพันธ์ไคสแควร์ (Chi-square association test) เป็นการทดสอบความสัมพันธ์ระหว่างข้อมูลคุณภาพ หรือข้อมูลปริมาณที่จัดเป็นชั้น 2 ชุด โดยจำนวนชั้นของข้อมูลแต่ละชุด ไม่ใช่ขนาด 2 ชั้นทั้งสองตัวแปร

2.1.2 การทดสอบเยทไคสแควร์(Chi-square Yate's correction for continuty) เป็นการทดสอบความสัมพันธ์ระหว่างข้อมูลคุณภาพ หรือข้อมูลปริมาณที่จัดเป็นชั้น 2 ชุด โดยจำนวนชั้นของข้อมูลแต่ละชุด มีขนาด 2 ชั้นทั้งสองตัวแปร และมีค่าคาดคะเนตามทฤษฎี ของแต่ละช่องมากกว่าหรือเท่ากับ 5

2.1.3 การทดสอบฟิชเชอร์แอคแซก(Fisher Exact test)เป็นการทดสอบความสัมพันธ์ ระหว่างข้อมูลคุณภาพ หรือข้อมูลปริมาณที่จัดเป็นชั้น 2 ชุด โดยจำนวนชั้นของข้อมูลแต่ละชุด มีขนาด 2 ชั้นทั้งสองตัวแปร และมีค่าคาดคะเนตามทฤษฎีแต่ละช่องน้อยกว่า 5

2.2 **ข้อมูลอันดับ**

เมื่อข้อมูลเป็นข้อมูลคุณภาพที่จัดเป็นอันดับได้ หรือข้อมูลปริมาณที่มีการจัดอันดับ สนใจทดสอบว่าข้อมูลทั้งสองมีความสัมพันธ์กันหรือไม่ แบ่งสถิติทดสอบตามลักษณะข้อมูลได้ดังนี้

2.2.1 สหสัมพันธ์อันดับของสเปียร์แมน (Spearman rank correlation) เป็นการทดสอบความสัมพันธ์ระหว่างข้อมูลอันดับ หรือข้อมูลปริมาณที่จัดเป็นอันดับ 2 ชุด มีจุดประสงค์เพื่อทดสอบว่ามีความสัมพันธ์กันหรือไม่ มากน้อยระดับไหน และในทิศทางไหน

2.2.2 การทดสอบเคนเดิลเทา (Kendall tau )เป็นการทดสอบ ความสัมพันธ์ระหว่างข้อมูลอันดับ ที่เป็นคู่กัน มีจุดประสงค์เพื่อทดสอบว่ามีความสัมพันธ์กันหรือไม่ มากน้อยระดับไหน และในทิศทางแปรตามกัน หรือแปรผกผันกัน

2.3 **ข้อมูลปริมาณ**

2.3.1 สหสัมพันธ์เชิงเส้นแบบเดี่ยว (Simple Linear Correlation) เป็นการทดสอบความสัมพันธ์ระหว่างข้อมูลปริมาณที่มีการแจกแจงปรกติทวิ (Bivariate normal distribution) 2 ชุด จุดประสงค์เพื่อทดสอบว่ามีความสัมพันธ์กันหรือไม่ มากน้อยระดับไหน และในทิศทางแปรตามกัน หรือแปรผกผันกัน

2.3.2 การถดถอยเชิงเส้นแบบพหุ (Multiple Linear Regression) เป็นการทดสอบว่าการเปลี่ยนแปลงของตัวแปรตาม ขึ้นอยู่กับการเปลี่ยนแปลงของตัวแปรต้นหรือไม่ ตัวแปรตามต้องเป็นตัวแปรปริมาณเชิงสุ่ม 1 ตัวแปรที่มีการแจกแจงปรกติ และตัวแปรต้นเป็นตัวแปรปริมาณมีมากกว่า 1 ตัว เพื่อใช้ในการทำนายตัวแปรตามด้วยตัวแปรต้น

2.3.3 การถดถอยแบบโลจิสติก (Logistic regression) เป็นการทดสอบว่าการเปลี่ยนแปลงของตัวแปรตามที่เป็นสัดส่วน ขึ้นอยู่กับการเปลี่ยนแปลงของตัวแปรต้นหรือไม่ เพื่อใช้ในการทำนายตัวแปรตามด้วยตัวแปรต้น

2.3.4 การวิเคราะห์จำแนก (Discriminant analysis) เป็นการทดสอบว่าการเปลี่ยนแปลงของตัวแปรตามที่จัดชั้นขึ้นอยู่กับการเปลี่ยนแปลงของตัวแปรต้นหรือไม่ เพื่อใช้ในการทำนายตัวแปรตามด้วยตัวแปรต้น

**ข้อพิจารณาการเลือกใช้สถิติ**

การเลือกใช้สถิติให้ถูกต้องเหมาะสมต้องคำนึงถึงข้อสมมุติ (assumptions) การใช้สถิติแต่ละตัว เพราะข้อตกลงเบื้องต้นเป็นที่มาของสูตรสถิติ การใช้สูตรสถิติที่ถูกต้องทำให้มีอำนาจการทดสอบ (Power of the test)สูง ผลการทดสอบเป็นที่เชื่อถือ ข้อพิจารณาที่ใช้ได้แก่

1. วัตถุประสงค์ของการทดสอบ
2. มาตราของการวัดค่า (Scale of Measurement) ของแต่ละตัวแปร
3. จำนวนชุดข้อมูล
4. จำนวนตัวแปรที่สนใจศึกษา
5. ลักษณะตัวแปร :

   ตัวแปรต้น หรือ ตัวแปรอิสระ (Independent variables)

   ตัวแปรตาม (Dependent variables)

   ตัวแปรร่วม (Covariate variables)

6. ลักษณะการแจกแจงของตัวแปรปริมาณ เป็นแบบปรกติ (Normal Distribution) หรือไม่
7. ประเภทการทดสอบ : ทดสอบความแตกต่าง หรือ ทดสอบความสัมพันธ์
8. การทดสอบแบบมีพารามิเตอร์ (Parametric Test) หรือการทดสอบแบบไร้พารามิเตอร์ (Non-parametric Test)
9. ค่าความแปรปรวนต่างกัน หรือไม่แตกต่างกัน
10. ประเภทของตัวอย่าง:

    ตัวอย่างอิสระ (Independent Samples)

    ตัวอย่างสัมพันธ์ (Releated Samples)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**ขั้นตอนการวิจัย**

ในการสร้างระบบผู้เชี่ยวชาญทางสถิติมีขั้นตอนการดำเนินงานหลัก 5 ขั้นตอนคือ

1. สร้างรูปแบบฐานความรู้ทางสถิติในรูปของไฟล์ที่เป็นข้อความ ลักษณะเป็น IF-THEN-ELSE Clause เพื่อให้บันทึกเข้าระบบด้วยวิธีง่าย ๆ

2. สร้างฐานความรู้ ทางสถิติ (Statistical Knowledge Base) โดยทำการแปลงประโยค IF-THEN-ELSE ดังกล่าวให้อยู่ในรูปแบบที่เครื่องอนุมาน สามารถนำไปใช้ได้อย่างมีประสิทธิภาพ

3. สร้างรูปแบบฐานความรู้งานวิจัยในรูปไฟล์ข้อความ บอกสมมุติฐาน สถิติที่ใช้ทดสอบ ผลสรุปที่ได้จากการทดสอบ

4. สร้างโปรแกรมระบบผู้เชี่ยวชาญทางสถิติ ซึ่งประกอบไปด้วย

4.1 การอ่าน ไฟล์ฐานความรู้สถิติ เข้ามาในโปรแกรม

4.2 การอ่าน ไฟล์ฐานความรู้งานวิจัย เข้ามาในโปรแกรม

4.3 จอภาพติดต่อกับผู้ใช้ เพื่อรับข้อมูลลักษณะของตัวแปร สมมติฐานการวิจัยจากผู้ใช้ระบบ และการถามตอบกับผู้ใช้ระบบ เพื่อเพิ่มเติมข้อมูลที่ระบบต้องการ

4.4 เครื่องอนุมาน เพื่อนำข้อมูลจากผู้ใช้ระบบ ประกอบกับความรู้ในฐานความรู้ไปค้นหาคำตอบ

4.5 แสดงผลลัพธ์ พร้อมคำอธิบายเหตุผลการใช้สถิติ

4.6 คำแนะนำการใช้สถิติ

5. ทดสอบการทำงานของระบบผู้เชี่ยวชาญ โดยใช้ข้อมูลของงานวิจัย เพื่อการแก้ไขปรับปรุงระบบให้ดียิ่งขึ้น

ขั้นตอนที่ 1 และ 2 สามารถจะข้ามไปทำขั้นตอนที่ 3 ได้ หากใส่ความรู้ ในรูปแบบที่มีการจัดเตรียมความรู้ที่เกี่ยวข้องกันทั้งหมด เพียงครั้งเดียว และจัดให้อยู่ในรูปแบบที่เครื่องอนุมานสามารถนำไปใช้ได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

**งานวิจัยที่เกี่ยวข้อง**

Hijek และ Tvanek(1982) ได้สร้างโปรแกรมสำหรับ สร้างข้อสมมติฐานในการวิเคราะห์ ข้อมูล Pregibon และ Gale(1984) สร้างโปรแกรมเกี่ยวกับการสร้างโมเดลการถดถอย (Regression Model) สำหรับ ข้อมูลที่ได้จากการสังเกตุ โดยการหาวิธีการแปลง(Transform) ข้อมูลให้เหมาะสมกับโมเดลได้ดีที่สุด Smith, Lee และ Hand(1983) สร้างระบบผู้เชี่ยวชาญสำหรับการวิเคราะห์หลายตัวแปร(Multivariate Analysis) NelderและWolstenholme(1986)ได้สร้างฐานความรู้เสริมหน้า(Knowledge Base front-end) ให้กับโปรแกรม GLIM (General Linear Interactive Modeling) Nanta(1989)ได้ทำการศึกษาประสิทธิภาพของการเรียนวิชาสถิติเบื้องต้นที่ใช้ติวเตอร์เชิงปัญญา (Intelligent tutoring) กับไม่ใช้ ปรากฏว่าได้ผลทดสอบยังไม่แตกต่างกัน Sarah(1990) ได้นำระบบผู้เชี่ยวชาญไปใช้ฝึกเจ้าหน้าที่ ให้มีความเข้าใจในการใช้สถิติควบคุมกระบวนการ (Statistical process control) เพื่อการเพิ่มผลผลิต ผลการทดสอบ กรณีที่มีสถิติหลายตัวที่สามารถนำมาใช้วิเคราะห์ข้อมูลได้ ผู้จัดการที่ได้รับการฝึกจากระบบผู้เชี่ยวชาญจะมีการตอบคำถามที่ถูกต้องมากกว่าผู้จัดการที่ไม่ได้ใช้ระบบ

ในประเทศไทยมีการนำเอาระบบผู้เชี่ยวชาญมาใช้ในแง่ต่างๆดังนี้ ศิลา(2532)ทำวิทยานิพนธ์เรื่อง การใช้ระบบผู้เชี่ยวชาญควบคุมปฏิบัติการบิน ของสนามบินดอนเมือง แก้ปัญหาเมื่อเกิดกรณีผิดปกติของเที่ยวบิน เช่นเครื่องเสีย ภาวะอากาศแปรปรวน นำระบบผู้เชี่ยวชาญเพื่อเลือกทางแก้ปัญหาอันได้แก่การยกเลิกเที่ยวบิน การเลื่อนกำหนดเวลาของเที่ยวบิน การเปลี่ยนเส้นทางการบิน และการเปลี่ยนแบบเครื่องบิน มีการจัดรูปแบบ ความรู้แบบ Rule-Base เขียนด้วยภาษา LISP พิชโยทัย(2534) ทำวิทยานิพนธ์เรื่อง การพัฒนาระบบ ผู้เชี่ยวชาญเกี่ยวกับการวิเคราะห์หาสาเหตุขัดข้องของรถยนต์ ของรถยนต์ 12 ระบบ โดยใช้การจัดเก็บ ความรู้ในรู้ของ IF-THEN และการวินิจฉัยใช้วิธีการค้นแบบย้อนกลับ เขียนด้วยภาษาซีบนเครื่องไมโคร คอมพิวเตอร์ ไอบีเอม พีซี โอภาส(2534) ทำวิทยานิพนธ์เรื่องการพัฒนาต้นแบบเพื่อการสาธิตโครงระบบ ผู้เชี่ยวชาญ ใช้การจัดความรู้แบบ Rule-Base การควบคุมใช้วิธีการหาเหตุผลแบบไปข้างหน้าเป็นหลัก และ แบบย้อนกลับเป็นส่วนประกอบ เขียนด้วยภาษาซี

บนเครื่องไมโครคอมพิวเตอร์ ไอบีเอม พีซี องค์ประกอบของระบบผู้เชี่ยวชาญ สุชาติ(2534)ทำวิทยานิพนธ์เรื่องการพัฒนาโครงระบบผู้เชี่ยวชาญที่ใช้การอนุมานความรู้แบบหาเหตุผลไปข้างหน้า โดยการแทนความรู้แบบกฎและการอนุมานความรู้ที่มีความสามารถจะหยุดพักการปรึกษาไว้ชั่วคราวและบันทึกข้อเท็จจริงต่างๆเก็บไว้ในแฟ้มข้อมูลแล้วเรียกกลับมาใช้ปรึกษาต่อจากจุดเดิมโดยไม่ต้องเริ่มต้นใหม่ ใช้ความรู้บางส่วนของการตรวจจำแนกชนิดของเชื้อแบคทีเรียทางการแพทย์เป็นความรู้ตัวอย่างสำหรับการออกแบบและการทดสอบ เพชรรัตน์ (2536) ทำวิทยานิพนธ์เรื่องการพัฒนาระบบผู้เชี่ยวชาญสำหรับคำนวณต้นทุนแหวนประดับโดยดัดแปลงต้นแบบเพื่อการสาธิตโครงระบบผู้เชี่ยวชาญของ นายโอภาส น้านิรัติศัยให้เหมาะสมในการคำนวณต้นทุนแหวน ใช้วิธีการอนุมานความรู้แบบหาเหตุผลไปข้างหน้าและการแทนค่าความรู้เป็นแบบกฎ สุวิมล(2538) ทำวิทยานิพนธ์เรื่อง การพัฒนาต้นแบบระบบผู้เชี่ยวชาญการช่วยฟื้นฟูชีวิตทารกแรกเกิด โดยใช้การแทนความรู้แบบกรอบ และการอนุมานความรู้จากโครงสร้างลำดับชั้นของกรอบที่ใช้ในการแทนความรู้

**ประโยชน์ที่คาดว่าจะได้รับ**

1. ผู้ใช้สามารถนำเอาระบบผู้เชี่ยวชาญนี้มาทดสอบการเก็บข้อมูลว่าข้อมูลที่ได้จากตัวแปรที่จะเก็บ สามารถนำมาทำการวิเคราะห์ข้อมูล เพื่อตอบวัตถุประสงค์และสมมติฐานการวิจัยที่ตั้งไว้ได้หรือไม่ เพื่อจะได้ตัวแปรที่จำเป็นต้องมีการเก็บข้อมูลในการวิจัยจนเป็นที่แน่ใจ จึงค่อยไปเก็บรวบรวมข้อมูล

2. เพื่อนำไปใช้เป็นเครื่องมือในการสอนหรือให้คำแนะนำ การเลือกใช้สถิติอย่างถูกต้อง

3. ระบบผู้เชี่ยวชาญนี้สามารถไปใช้กับงานวิจัยเรื่องอื่นๆได้ เพียงเปลี่ยนฐานความรู้งานวิจัยไปเป็นงานวิจัยและ ทฤษฎีที่เกี่ยวกับเรื่องที่ต้องการศึกษาเท่านั้น

4. เพื่อเป็นแนวทางในการพัฒนาระบบผู้เชี่ยวชาญด้านอื่นๆ ที่เป็นฐานแห่งกฎ (Rule Base) และการค้นหาแบบลูกโซ่ไปข้างหน้า (Forward Chaining) ต่อไปได้

ในบทที่ 2 จะกล่าวถึง ทฤษฎี แนวคิด ของระบบผู้เชี่ยวชาญ ในแง่ของความหมาย ความเป็นมา โครงสร้างพื้นฐานของระบบผู้เชี่ยวชาญโดยทั่วไป การประยุกต์ใช้ระบบผู้เชี่ยวชาญในงานต่างๆ ขั้นตอนในการพัฒนาระบบผู้เชี่ยวชาญ หลักการแทนความรู้เข้าสู่ระบบผู้เชี่ยวชาญ วิธีการค้นหาคำตอบและกลยุทธ์การควบคุมวิธีการหาเหตุผล บทที่ 3 จะกล่าวถึงการออกแบบและพัฒนาระบบผู้เชี่ยวชาญ เริ่มต้นการอธิบายโครงสร้างทั้งหมดของระบบผู้เชี่ยวชาญและการทำงานร่วมกันของส่วนต่างๆ ก่อน ต่อจากนั้นจึงจะอธิบายในแต่ละส่วนในละเอียด ได้แก่การออกแบบฐานความรู้ทางสถิติ ฐานความรู้งานวิจัย เครื่องอนุมาน คำโต้ตอบกับผู้ใช้ระบบ และคำแนะนำการเลือกใช้สถิติ บทที่ 4 การใช้งานระบบผู้เชี่ยวชาญ ได้นำตัวอย่างจริงมาใช้อธิบาย เริ่มด้วยขั้นตอนการทำงานของระบบ วิธีการสร้างความรู้ทางสถิติ การทำงานของเครื่องอนุมาน และการให้คำแนะนำการเลือกใช้สถิติ บทที่ 5 เป็นการสรุปการพัฒนาระบบผู้เชี่ยวชาญและข้อเสนอแนะในการใช้ระบบและการพัฒนาระบบ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย