

บทที่ 5

สรุปผลการวิจัย อภิปรายการวิจัย และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยนี้เป็นการศึกษาและวิเคราะห์ข้อมูลการสำรวจด้านการทำงานของประชากร ของสำนักงานสถิติแห่งชาติ โดยนำความรู้ทางด้านการทำเหมืองข้อมูล มาใช้เพื่อสร้างแบบจำลองในการจำแนกประเภทข้อมูล เพื่อศึกษาปัจจัยที่มีผลต่อรายได้ของประชากรในประเทศไทย โดยใช้ต้นไม้การตัดสินใจอัลกอริทึม C4.5 มาใช้เพื่อสร้างแบบจำลองตามประเภทอุตสาหกรรม ซึ่งออกเป็น 4 กลุ่ม คือ กลุ่มที่ 1 การเกษตรและประมง กลุ่มที่ 2 การผลิต กลุ่มที่ 3 การค้าและบริการ และ กลุ่มที่ 4 อื่น ๆ (ข้าราชการ การศึกษา องค์กรระหว่างประเทศ เป็นต้น) โดยแอตทริบิวต์ที่ใช้ในการจัดกลุ่มข้อมูล 13 ตัวแปรประกอบด้วย 1) เพศ 2) อายุ 3) ระดับการศึกษา 4) อาชีพ 5) จำนวนชั่วโมงการทำงานต่อสัปดาห์ 6) ประเภทการจ้างงาน 7) รายได้อื่น ๆ 8) ภาค 9) เขตที่อยู่ 10) สถานะภาพสมรส 11) สถานะภาพการทำงาน 12) ขนาดขององค์กรหรือบริษัทที่ทำงาน และ 13) ช่วงรายได้ของประชากร โดยใช้ช่วงรายได้ของประชากรเพื่อเป็นคลาสในการสร้างแบบจำลองต้นไม้การตัดสินใจ ก่อนจะนำผลลัพธ์ที่ได้มาสร้างเป็นกฎความสัมพันธ์ เพื่อศึกษาปัจจัยที่มีผลต่อการมีรายได้ของประชากรในประเทศไทย

ในขั้นตอนการสร้างแบบจำลองต้นไม้การตัดสินใจนั้น พบว่าปัจจัยที่มีผลต่อรายได้ของประชากรในกลุ่มที่ 1 การเกษตรและประมง คือ อาชีพ ระดับการศึกษา ภาค อายุ และชั่วโมงการทำงาน หมายถึงว่า ประชากรที่จะมีรายได้ในช่วงมากคือ มีรายได้มากกว่า 30,000 บาทขึ้นไปนั้นจะต้องมีลักษณะดังนี้คือ มีอาชีพในกลุ่มที่ 1 มีระดับการศึกษาในระดับอุดมศึกษา อาศัยอยู่ในภาคกลาง ในแต่ละสัปดาห์จะทำงานค่อนข้างสูง คือ มากกว่า 45 ชั่วโมงขึ้นไป หรือ มีอายุอยู่ในช่วง 35-39 ปี จะมีโอกาสที่จะมีรายได้มากกว่า 30,000 บาท ถึงร้อยละ 80 โดยกลุ่มที่ 1 มีค่า Correction เท่ากับ 98.49 ค่า Precision, F-Measure เท่ากับ 0.98 และค่า Recall เท่ากับ 0.99

ปัจจัยที่มีผลต่อรายได้ของประชากรในกลุ่มที่ 2 การผลิต คือ ประเภทการจ้างงาน สถานะภาพการทำงาน อายุ อาชีพ ระดับการศึกษา จำนวนชั่วโมงการทำงาน เขตที่อยู่อาศัย สถานะภาพสมรส และการมีรายได้อื่นๆ โดยดูจากกฎความสัมพันธ์ที่สร้างจากต้นไม้การตัดสินใจ ความซับซ้อนของกฎความสัมพันธ์จะแปรผันโดยตรงกับขนาดของต้นไม้การตัดสินใจ โดยกลุ่มที่ 2 มีค่า Correction เท่ากับ 91.57 ค่า Precision, F-Measure เท่ากับ 0.90 และค่า Recall เท่ากับ 0.92

กลุ่มที่ 3 การค้าและบริการ มีปัจจัยที่มีผลต่อรายได้ของประชากร คือ อาชีพ อายุ ระดับการศึกษา การมีรายได้อื่นๆ จำนวนชั่วโมงการทำงาน และ เพศ โดยกลุ่มที่ 3 มีค่า Correction เท่ากับ 85.71 ค่า Precision, F-Measure เท่ากับ 0.85 และค่า Recall เท่ากับ 0.86

ส่วนกลุ่มที่ 4 อื่นๆ มีปัจจัยที่มีผลต่อรายได้ของประชากร คือ อายุ ระดับการศึกษา การมีรายได้อื่นๆ และ อาชีพ จะเป็นปัจจัยที่มีผลต่อการมีรายได้สูงของประชากร โดยกลุ่มที่ 4 มีค่า Correction เท่ากับ 75.98 ค่า Precision เท่ากับ 0.77 และ ค่า Recall, F-Measure เท่ากับ 0.76

ผลที่ได้จากการวิเคราะห์พบว่าอัลกอริทึมในการสร้างต้นไม้การตัดสินใจ C4.5 สามารถนำมาใช้ในการสร้างจำลองทางสังคมศาสตร์ได้อย่างดี

5.2 อภิปรายผลการวิจัย

จากขั้นตอนการสร้างแบบจำลองต้นไม้การตัดสินใจนั้น พบว่าปัจจัยที่มีผลต่อรายได้ของประชากรในทุกๆ กลุ่มจะมีลักษณะคล้ายกัน คือ ประเภทการทำงาน อาชีพ ระดับการศึกษา อายุ จำนวนชั่วโมงการทำงาน การมีรายได้อื่นๆ นอกเหนือ จากรายได้ประจำ

ซึ่งปัจจัยเหล่านี้เป็นปัจจัยที่มีความสัมพันธ์กัน การจะมีรายได้มากหรือในระดับสูงอาจจะต้องอาศัยปัจจัยเหล่านี้ควบคู่กันไป โดยสามารถอธิบายได้ดังนี้

ประชากรที่มีการทำงานในลักษณะรายเดือนจะมีรายได้สูงกว่าการทำงานแบบรายชั่วโมง รายวัน หรือรายสัปดาห์ เนื่องจากจะมีความมั่นคงมากกว่า และรายได้ค่อนข้างแน่นอนในทุก ๆ เดือน

อาชีพที่มีรายได้สูงจะอยู่ในกลุ่มที่ 1 และ 2 ซึ่งประกอบด้วยผู้บัญญัติกฎหมาย ข้าราชการระดับอาวุโส ผู้จัดการ ผู้ประกอบวิชาชีพด้านต่างๆ ช่างเทคนิคสาขาต่างๆ ผู้ประกอบวิชาชีพอื่นๆ ที่เกี่ยวข้อง ส่วนกลุ่มที่ 3 และ 4 ซึ่งประกอบด้วย เสมียน พนักงานบริการ พนักงานขายในร้านค้าและตลาด อาชีพขั้นพื้นฐานต่างๆ ผู้ปฏิบัติงานที่มีฝีมือในด้านการเกษตรและการประมง ผู้ปฏิบัติงานในธุรกิจด้านความสามารถทางฝีมือและธุรกิจอื่นๆ ที่เกี่ยวข้อง ผู้ปฏิบัติการเครื่องจักรโรงงานและเครื่องจักร และ ผู้ปฏิบัติงานด้านการประกอบ จะมีรายได้ค่อนข้างน้อยถึงปานกลาง เนื่องจากการที่จะได้รายได้มานั้น อาจจะขึ้นอยู่กับปัจจัยด้านอื่นๆ เช่น ลักษณะของภูมิอากาศ การเกิดโรคระบาด ภัยธรรมชาติ หรือ ปริมาณน้ำที่ต้องใช้ในการทำเกษตรกรรม โดยในแต่ละเดือนก็อาจจะได้รายได้ไม่เท่ากัน

ระดับการศึกษาจะมีผลต่อรายได้ของประชากร ผู้ที่มีการศึกษาสูงอยู่ในระดับอุดมศึกษาส่วนใหญ่จะมีรายได้สูงกว่า ประชากรที่มีการศึกษาอยู่ในระดับมัธยมศึกษา หรือระดับต่ำกว่า เนื่องจากเงินเดือนที่ได้บริษัทหรือองค์กรส่วนใหญ่จะขึ้นอยู่กับวุฒิการศึกษา ถ้ามีวุฒิการศึกษาสูง เงินเดือนก็จะสูงขึ้นตาม

อายุของประชากร ประชากรที่มีอายุค่อนข้างสูง จะมีรายได้สูงตามไปด้วย เนื่องจากมีประสบการณ์ในการทำงานมากกว่าประชากรที่มีอายุน้อย สำหรับประชากรที่มีอายุน้อย หรือเพิ่งเริ่มทำงาน อาจจะต้องทำการศึกษาเพิ่มเติม เพื่อให้ปัจจัยที่ส่งผลต่อการมีรายได้ปัจจัยอื่นๆ เพิ่มขึ้น เช่น เพิ่มเติมในเรื่องระดับการศึกษา หรือ หารายได้อื่น ๆ เพิ่มเติมนอกเหนือจากรายได้หลัก

การมีรายได้อื่น ๆ นอกเหนือจากรายได้ประจำ ก็จะเป็นอีกปัจจัยหนึ่งที่ทำให้ประชากรมีรายได้สูงตามไปด้วย เช่น การทำงานล่วงเวลา หรือ การทำอาชีพพิเศษเพิ่มเติม

5.3 ข้อเสนอแนะ

จากงานวิจัยฉบับนี้ใช้เทคนิคการจำแนกประเภทข้อมูล โดยให้ระดับความสำคัญของตัวแปรที่ใช้มีความสำคัญเท่ากัน แต่โดยความจริงแล้วตัวแปรแต่ละตัวจะมีความสำคัญไม่เท่ากัน ดังนั้นผู้วิจัยคาดว่าหากนำวิธีการให้ค่าน้ำหนักมาใช้เพื่อปรับน้ำหนักให้กับตัวแปรที่มีความสำคัญ ก็จะทำให้การทำงานในการจำแนกประเภทข้อมูลมีประสิทธิภาพดีขึ้น