

ชื่อโครงการ	การประยุกต์ใช้เทคนิคการวิเคราะห์การจัดกลุ่มเพื่อทำความเข้าใจความหลากหลายของ Data Scientist
โดย	นางสาวหทัยวรัญญู จรุงดาวเรือง 66541207028-6
อาจารย์ที่ปรึกษา	อาจารย์ธีรภพ แสงศรี
หลักสูตร	ระบบสารสนเทศทางธุรกิจ สาขาบริหารธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์
ปีการศึกษา	2568

บทคัดย่อ

โครงการเรื่อง “การประยุกต์ใช้เทคนิคการวิเคราะห์การจัดกลุ่มเพื่อทำความเข้าใจความหลากหลายของ Data Scientist” มีวัตถุประสงค์เพื่อวิเคราะห์และจำแนกเงินเดือนของ Data Scientist โดยพิจารณาจากปัจจัยด้านประสบการณ์การทำงาน ตำแหน่งงาน ทักษะ ความเชี่ยวชาญเฉพาะด้าน และคุณลักษณะอื่น ๆ ที่เกี่ยวข้อง เพื่ออธิบายลักษณะเฉพาะของแต่ละกลุ่มที่ได้จากการจำแนก เช่น ระดับเงินเดือน ความสัมพันธ์กับประสบการณ์ และปัจจัยสำคัญที่ส่งผลต่อความแตกต่างของตลาดแรงงานสาย Data Scientist

โครงการนี้ใช้กระบวนการ CRISP-DM ในการวิเคราะห์ข้อมูลตั้งแต่การทำความเข้าใจธุรกิจ การเตรียมข้อมูล การสร้างโมเดล และการประเมินผล โดยทำการทำความสะอาดและแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสม เช่น การจัดกลุ่มเงินเดือน ตำแหน่งงาน และภูมิภาค จากนั้นสร้างโมเดลด้วย Decision Tree ผ่านโปรแกรม RapidMiner พบว่า Decision Tree เป็นโมเดลที่เหมาะสมที่สุด เนื่องจากมี Accuracy 89.17% สูงและอธิบายกฎการตัดสินใจได้ง่าย ผลลัพธ์สามารถจำแนกเงินเดือนได้เป็น Medium, Large, Small พร้อมวิเคราะห์ความสัมพันธ์กับประสบการณ์และตำแหน่งงาน และนำเสนอผ่าน เว็บไซต์ ในรูปแบบกราฟและตารางเพื่อให้ผู้ใช้งานเข้าใจข้อมูลได้ง่าย