บทที่ 2

การทบทวนวรรณกรรม

บทนี้เป็นการศึกษาเกี่ยวกับความสัมพันธ์ระหว่างระเบียบวิธีแบบอไจล์ ทฤษฎีบุคลิกภาพ และเทคนิคการจัดกลุ่มข้อมูล โดยผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องในหัวข้อต่างๆ ประกอบด้วย 6 หัวข้อ ได้แก่ แนวคิดและทฤษฎีเกี่ยวกับระเบียบวิธีแบบอไจล์ การทดสอบบุคคลอไจล์ ลักษณะทางบุคลิกภาพ บุคลิกภาพห้าองค์ประกอบ แบบทดสอบบุคลิกภาพ และการทำนายและจัดกลุ่มข้อมูล โดยในแต่ละหัวข้อมีรายละเอียดดังนี้

2.1 แนวคิดและทฤษฎีเกี่ยวกับระเบียบวิธีแบบอไจล์

ระเบียบวิธีแบบอไจล์ (Agile Methodology) คือ ระเบียบวิธีการพัฒนาซอฟต์แวร์ ใช้ในการบริหารโครงการซอฟต์แวร [4] ระเบียบวิธีแบบอไจล์สนับสนุนให้ทีมพัฒนาซอฟต์แวร์พร้อมรับมือต่อการเปลี่ยนแปลงที่อาจเกิดขึ้นได้ระหว่างขั้นตอนการพัฒนาซอฟต์แวร์ ในปัจจุบันมีระเบียบวิธีที่ใช้หลักการของระเบียบวิธีแบบอไจล์ในการทำงาน เช่น สกรัม (Scrum) [14] เอกซ์ตรีมโปรแกรมมิ่ง (Extreme Programming) [15] และการพัฒนาซอฟต์แวร์โดยอาศัยการทดสอบเป็นตัวขับเคลื่อน (Test Driven Development : TDD) [16] เป็นต้น กลุ่มผู้ออกแบบระเบียบวิธีแบบอไจล์ได้ออกคำแถลงอุดมการณ์แห่งอไจล์ ซึ่งเป็นบทบัญญัติที่ถูกสร้างขึ้นมาเพื่อกำหนดคุณลักษณะของระเบียบวิธีแบบอไจล์ซึ่งมีเนื้อหาแสดงดัง ภาพที่ 2.1

เราค้นพบวิธีที่ดีกว่าในการพัฒนาซอฟท์แวร์ จากการลงมือทำจริงและช่วยเหลือผู้อื่น นั่นคือ เราให้ความสำคัญกับ: คนและการมีปฏิสัมพันธ์กัน มากกว่าการทำตามขั้นตอนและเครื่องมือ ซอฟต์แวร์ที่นำไปใช้งานได้จริง มากกว่าเอกสารที่ครบถ้วนสมบูรณ์ การร่วมมือทำงานกับลูกค้า มากกว่าการต่อรองให้เป็นไปตามสัญญา การตอบรับกับการเปลี่ยนแปลง มากกว่าการทำตามแผนที่วางไว้ นั่นคือ แม้เราเห็นความสำคัญในสิ่งที่กล่าวไว้ทางด้านขวา แต่เราให้ความสำคัญกับสิ่งที่กล่าวไว้ทางด้านซ้ายมากกว่า	We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value: Individuals and interactions over processes and tools Working software over comprehensive documentation Customer collaboration over contract negotiation Responding to change over following a plan That is, while there is value in the items on the right, we value the items on the left more

ภาพที่ 2.1 คำแถลงอุดมการณ์แห่งอไจล์

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

คำแถลงอุดมการณ์แห่งอไจล์ให้ความสำคัญกับปัจจัยอยู่ 4 ปัจจัยหลักได้แก่ การมีปฏิสัมพันธ์กันระหว่างบุคคล ซอฟต์แวร์ที่นำไปใช้งานได้จริง การร่วมมือทำงานกับลูกค้า และการตอบรับกับการเปลี่ยนแปลง

ในส่วนของการมีปฏิสัมพันธ์กันระหว่างบุคคล ผู้พัฒนาซอฟต์แวร์ควรให้ความสนใจกับการพูดคุย ระดมสมอง (Brain Stroming) เพื่อจัดการปัญหาที่เกิดขึ้น เนื่องจากการพูดคุยทำให้เกิดการแลกเปลี่ยนองค์ความรู้และความคิดเห็น ซึ่งข้อมูลเหล่านี้สามารถนำไปพัฒนาตัวบุคคลเพื่อวางแผนรับมือกับปัญหาในอนาคตต่อไปได้

ซอฟต์แวร์ที่นำไปใช้งานได้จริง เป็นปัจจัยที่สองในคำแถลงอุดมการณ์แห่งอไจล์ ระเบียบวิธีแบบอไจล์มุ่งเน้นให้ผู้พัฒนาซอฟต์แวร์พัฒนาและส่งมอบซอฟต์แวร์ให้แก่ลูกค้าได้โดยคาดหวังให้ลูกค้าเกิดความพึงพอใจสูงสุด ดังนั้นซอฟต์แวร์ที่ใช้งานได้จริงจึงเป็นสิ่งที่ถูกนำเสนอให้ลูกค้าเป็นอันดับแรก ในขณะที่เอกสารที่เกี่ยวข้องกับซอฟต์แวร์ถูกลดบทบาทให้เป็นส่วนเติมเต็มเพื่อความสมบูรณ์เท่านั้น อย่างไรก็ตามระเบียบวิธีแบบอไจล์ไม่ได้ชี้นำให้ผู้พัฒนาซอฟต์แวร์ยกเลิกการทำเอกสาร เพียงแต่ให้ความสำคัญกับซอฟต์แวร์ที่นำไปใช้งานได้จริงมากกว่าเท่านั้น

การร่วมมือทำงานกับลูกค้าก็เป็นอีกปัจจัยหนึ่งที่ระเบียบวิธีอไจล์ให้ความสำคัญด้วย เนื่องจากในการพัฒนาซอฟต์แวร์ ความต้องการของลูกค้าต้องมีความชัดเจนเพื่อให้ซอฟต์แวร์สามารถตอบสนองสิ่งที่ลูกค้าต้องการได้ครบถ้วน ยิ่งมีการสื่อสารระหว่างผู้พัฒนาซอฟต์แวร์และลูกค้ามากเท่าใด ความต้องการที่ได้มาจะยิ่งมีความชัดเจนมากขึ้น การร่วมมือทำงานกับลูกค้ามีความเกี่ยวข้องอย่างมากกับการตอบรับกับการเปลี่ยนแปลงที่เป็นปัจจัยต่อไป

ปัจจัยสุดท้ายในคำแถลงอุดมการณ์แห่งอไจล์คือ การตอบรับกับการเปลี่ยนแปลง เมื่อเข้าสู่กระบวนการพัฒนาซอฟต์แวร์ความต้องการของลูกค้าจะถูกกลั่นกรองให้เป็นรูปธรรมมากขึ้น และเมื่อลูกค้าเข้าใจความต้องการของตนเองมากขึ้น การเปลี่ยนแปลงด้านความต้องการย่อมเกิดขึ้น ดังนั้นระเบียบวิธีแบบอไจล์จึงมุ่งเน้นให้ผู้พัฒนาซอฟต์แวร์เตรียมพร้อมกับการเปลี่ยนแปลงดังกล่าวมากกว่าให้ผู้พัฒนาซอฟต์แวร์ยึดติดกับความต้องการในอดีตแรกที่ตกลงกับลูกค้าไว้

ทั้งนี้ความเหมาะสมของบุคลากรก็มีผลต่อการดำเนินการตามระเบียบวิธีแบบอไจล์ Madi และคณะ [5] ศึกษาเรื่องความสัมพันธ์ระหว่างบุคลากรกับระเบียบวิธีแบบอไจล์ โดยที่มีความเชื่อว่าหากผู้ปฏิบัติรู้จักถึงคุณค่า (Value) ที่แฝงอยู่ในระเบียบวิธีแบบอไจล์แล้ว ช่วยให้สามารถปฏิบัติตามระเบียบวิธีแบบอไจล์ได้ดียิ่งขึ้น ดังนั้น Madi และคณะ จึงได้วิเคราะห์หาคุณค่าที่ซ่อนอยู่นี้โดยวิธีการสกัดคำสำคัญที่อยู่ในหลักการเบื้องหลังคำแถลงอุดมการณ์แห่งอไจล์ และได้ค้นพบว่า คุณค่าที่สำคัญที่สุดในระเบียบวิธีแบบอไจล์คือ คุณค่าที่เกี่ยวข้องกับตัวบุคคล เช่น การทำงานร่วมกัน, การสื่อสาร,

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

การให้ความสำคัญกับลูกค้า, ความเคารพซึ่งกันและกันและความนอบน้อม, ความสนใจ, ความรับผิดชอบ หรือความคิดสร้างสรรค์ รองลงมาคือคุณค่าที่เกี่ยวข้องกับกระบวนการทำงาน เช่น การทำให้การทำงานซับซ้อนน้อยลง, ความยืดหยุ่นของกระบวนการ หรือ ความเร็วในการทำงาน นอกจากนี้ Madi และคณะได้ค้นพบว่า ระเบียบวิธีแบบอไจล์เป็นระเบียบวิธีที่มีความเป็นธรรมชาติ (Natural) อยู่ในตัว กล่าวคือ ผู้พัฒนาซอฟต์แวร์สามารถทำงานอย่างสอดคล้องกับระเบียบวิธีแบบอไจล์ได้เองแม้ไม่เคยศึกษาเรื่องระเบียบวิธีแบบอไจล์มาก่อน จากจุดนี้จึงสามารถตั้งสมมติฐานได้ว่ามีปัจจัยบางอย่างภายในตัวบุคคลที่ทำให้สามารถทำงานร่วมกับระเบียบวิธีแบบอไจล์โดยอัตโนมัติ

Silva และ Santos [17] ค้นพบปัจจัยสำคัญ 14 ปัจจัย ในการพัฒนาซอฟต์แวร์โดยใช้ระเบียบวิธีแบบอไจล์ โดยจัดกลุ่มปัจจัยเป็น 3 หมวด ได้แก่ ปัจจัยที่เกี่ยวข้องกับคน (People) ปัจจัยที่เกี่ยวข้องกระบวนการ (Process) และปัจจัยที่เกี่ยวข้องเทคโนโลยี (Technology) โดยปัจจัยที่เกี่ยวข้องกับคนประกอบด้วย ความสามารถและความเชี่ยวชาญ (Competence and expertise) การสนับสนุนของฝ่ายบริหาร (Executive support) แรงจูงใจของทีมและผู้ใช้งาน (Team and user motivation) ขนาดของทีม (Team size) และ การมีส่วนร่วมของผู้ใช้งาน (User participation) ในขณะที่ปัจจัยที่เกี่ยวข้องกับกระบวนการ ประกอบด้วย การปฏิบัติใช้งานระเบียบวิธีแบบอไจล์ (Agile practices) การเรียงลำดับการส่งงานโดยส่งมอบงานที่มีความสำคัญมากสุดเป็นอันดับแรก (Delivery most important features first) ความเหมาะสมของปริมาณเอกสาร (Right amount of documentation) และการสื่อสารที่ดี (Strong communication) และปัจจัยที่เกี่ยวข้องกับเทคโนโลยี ประกอบด้วย ความเหมาะสมของการอบรมด้านเทคนิค (Appropriate technical training) การทดสอบแบบบูรณาการ (Integration testing) ความเรียบง่ายของการออกแบบ (Simple design) การสนับสนุนของเครื่องมือ (Tool supports) และการกำหนดมาตรฐานการเขียนโค้ดที่ดี (Well defined coding standards) จากปัจจัยทั้งหมด Silva และ Santos ได้ระบุปัจจัยที่เกี่ยวข้องกับคนเป็นปัจจัยที่มีการอ้างอิงถึงมากที่สุด ดังนั้นจึงมีแนวโน้มว่าจะเป็นปัจจัยที่สำคัญที่สุดใน 3 หมวดที่กล่าวมา

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.2 การทดสอบบุคคลอไจล์

จากคำแถลงอุดมการณ์แห่งอไจล์และงานวิจัยข้างต้น จะเห็นได้ว่าปัจจัยที่เกี่ยวข้องกับระเบียบวิธีแบบอไจล์ มีความเกี่ยวข้องกับมนุษย์เป็นส่วนใหญ่ ผู้วิจัยจึงได้บัญญัติคำว่า “บุคคลอไจล์” (Agile Person) ขึ้นเพื่อใช้เรียกบุคคลที่มีแนวคิดสอดคล้องกับระเบียบวิธีแบบอไจล์ โดยในเบื้องต้นการค้นหาบุคคลอไจล์สามารถทำได้โดยใช้เครื่องมือวัด เช่น แบบทดสอบวัดความรู้พื้นฐานทางด้านอไจล์ [18] ซึ่งเป็นชุดคำถาม 60 ข้อ ซึ่งพัฒนาโดยสมาคม Driving Stratgy Deliver More (DSDM) มีจุดมุ่งหมายคือใช้ทดสอบความรู้พื้นฐานทางด้านระเบียบวิธีแบบอไจล์ ผู้ที่ตอบคำถามได้ถูกต้อง 42

ข้อขึ้นไปจะถูกรับรองจากสมาคมว่าเป็นผู้เชี่ยวชาญด้านระเบียบวิธีแบบอไจล์ ตัวอย่างของคำถามในแบบทดสอบดังกล่าวแสดงอยู่ในภาพที่ 2.2

1. Which of the following best describes the approach for determining the iteration (timebox) length?
   - A. Iterations (timeboxes) should always be 30 days
   - B. The team determines iteration (timebox) length by dividing the total number of story points by the average velocity of the team
   - C. Iterations (timeboxes) should always be two weeks
   - D. The team should agree on the length of the iteration (timebox), taking the size and complexity of the project into consideration

ภาพที่ 2.2 ตัวอย่างคำถามในแบบทดสอบวัดความรู้พื้นฐานทางด้านอไจล์

แบบทดสอบทางด้านอไจล์ประกอบไปด้วยคำถามวัดในเชิงองค์ความรู้ด้านอไจล์ อย่างไรก็ตามแบบทดสอบดังกล่าวอาจไม่เหมาะสมกับบุคคลที่ไม่เคยศึกษาหรือรู้จักระเบียบวิธีแบบอไจล์มาก่อน เนื่องจากเป็นคำถามที่ผู้ตอบต้องอาศัยประสบการณ์ทำงานเกี่ยวระเบียบวิธีแบบอไจล์เพื่อหาคำตอบที่ถูกต้อง Bishop และ Deoka [19] จึงเสนอแนวคิดว่าควรมีเครื่องมือบางอย่างที่สามารถวัดแนวคิดและทัศนคติเกี่ยวกับระเบียบวิธีแบบอไจล์ของผู้ตอบแบบทดสอบ กล่าวคือสามารถใช้เครื่องมือแบ่งได้ว่า ถ้าอ้างอิงจากความเชื่อที่ส่งผลต่อพฤติกรรม (Mindset) ของผู้ตอบแบบทดสอบแล้ว ผู้ตอบแบบทดสอบจะ “เลือกใช้” หรือ “ต่อต้าน” ระเบียบวิธีแบบอไจล์ ดังนั้นจึงได้พัฒนาเครื่องมือสำรวจความชื่นชอบในแนวคิดแบบอไจล์ (Agile Preference Instrument: API) ขึ้น เครื่องมือดังกล่าวสามารถสะท้อนแนวคิดและทัศนคติของผู้พัฒนาซอฟต์แวร์ที่สอดคล้องกับระเบียบวิธีแบบอไจล์ได้ดีกว่าใช้แบบทดสอบวัดความรู้ที่กล่าวมาในข้างต้น

ตัวอย่างของคำถามในเครื่องมือดังกล่าวแสดงอยู่ในภาพที่ 2.3 ผลลัพธ์จากการใช้แบบทดสอบและเครื่องมือที่กล่าวมาข้างต้นอยู่ในรูปแบบของคะแนน ซึ่งคะแนนเหล่านี้สามารถนำไปใช้แบ่งแยกบุคคลอไจล์ ออกจากผู้พัฒนาซอฟต์แวร์ได้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

Q1.	To satisfy the client, I like to...	A.	adapt to changing requirements
		B.	know that the requirements are firm and then build the program
Q2.	I prefer to get my information..	A.	by using documents and diagrams
		B.	by face-to-face communication
Q3.	I'd rather...	A.	build all the features and make sure they work well together before releasing the software
		B.	build small pieces of the total product, release it and then build more
Q4.	I prefer...	A.	working with the software users on a daily basis
		B.	working with written specifications and documents

ภาพที่ 2.3 ตัวอย่างเครื่องมือสำรวจความชื่นชอบในแนวคิดแบบอไจล์

จากภาพที่ 2.3 เครื่องมือสำรวจความชื่นชอบในแนวคิดแบบอไจล์ประกอบด้วยข้อคำถามจำนวน 40 ข้อ (ในรูปเป็นตัวอย่างที่นำมาเพียงส่วนหนึ่ง ชุดคำถามแบบสมบูรณ์แสดงอยู่ในภาคผนวก ก) ผู้ตอบแบบสอบถามต้องเลือกตอบหนึ่งข้อจากตัวเลือกสองข้อต่อหนึ่งคำถาม เช่น “To satisfy the client, I like to…” (ฉันชอบ….. เพื่อให้ลูกค้าพึงพอใจ) ตัวเลือกที่เลือกตอบได้มี “adapt to changing requirements” (ปรับตัวให้เข้ากับความต้องการที่เปลี่ยนแปลงไป) หรือ “know that the requirements are firm and then build the program” (ให้มีการกำหนดความต้องการให้เป็นที่แน่นอนก่อน จากนั้นค่อยลงมือสร้างโปรแกรม) หากผู้ตอบแบบสอบถามเลือก “adapt to changing requirements” จะถือว่าบุคคลนั้นมีแนวคิดสอดคล้องกับระเบียบวิธีแบบอไจล์ โดยอ้างอิงจากข้อความ “การตอบรับกับการเปลี่ยนแปลงมากกว่าการทำตามแผนที่วางไว้ (Responding to change over following a plan)” ในคำแถลงอุดมการณ์แห่งอไจล์ หรือข้อคำถาม “I prefer to get my information..” (ฉันต้องการที่จะรับข้อมูลสารสนเทศของฉัน…) ตัวเลือกที่เลือกตอบได้มี “by using documents and diagrams” (ผ่านเอกสารและแผนภาพ) หรือ “by face-to-face comunication” (ผ่านการสื่อสารโดยตรง) หากผู้ตอบแบบสอบถามเลือก “by face-to-face comunication” จะถือว่าบุคคลนั้นมีแนวคิดสอดคล้องกับระเบียบวิธีแบบอไจล์ โดยอ้างอิงจากข้อความ “คนและการมีปฏิสัมพันธ์กัน มากกว่าการทำตามขั้นตอนและเครื่องมือ (Individuals and interactions over processes and tools)” และ “ร่วมมือทำงานกับลูกค้ามากกว่าการต่อรองให้เป็นไปตามสัญญา (Customer collaboration over contract negotiation)” ในคำแถลงอุดมการณ์แห่งอไจล์ ในการตอบแบบสอบถาม ผู้ที่ได้คะแนนมากกว่าค่าเฉลี่ย (Mean) ของคะแนนจากผู้ตอบแบบสอบถามทั้งหมด (ในงานวิจัยใช้ 24.554 คะแนน) จะถือว่าเป็นผู้ที่ไม่ต่อต้านระเบียบวิธีอไจล์ ทั้งนี้กลุ่มตัวอย่างที่นำมาใช้ทดลองมีแต่ผู้ที่ทำงานเกี่ยวกับซอฟต์แวร์และอาศัยอยู่ในสหรัฐอเมริกาเท่านั้น นอกจากนี้ Bishop และ Deoka ศึกษาเพียงเรื่องความเกี่ยวข้องระหว่างปัจจัยส่งเสริมและต่อต้านระเบียบวิธีแบบอไจล์ แต่ไม่ได้ศึกษาในกรณีที่นำเครื่องมือที่คิดค้นขึ้นมาใช้ทำนายความเหมาะสมของบุคคลกับระเบียบวิธีอไจล์

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

เครื่องมือสำรวจความชื่นชอบในแนวคิดแบบอไจล์ของ Bishop และ Deoka เดิมทีถูกสร้างขึ้นเพื่อเป็นเครื่องมือช่วยค้นหาว่า มีปัจจัยใดบ้างที่ส่งผลให้บุคคลเลือกใช้หรือต่อต้านระเบียบวิธีแบบอไจล์ โดย Bishop และ Deoka นำทฤษฎีลักษณะทางบุคลิกภาพนำมาประยุกต์ใช้เพื่อจำกัดขอบเขตของปัจจัยที่ศึกษา

2.3.ลักษณะทางบุคลิกภาพ

ลักษณะของบุคลิกภาพ (Peosonaity Trait) แสดงถึงแบบแผนของความรู้สึก, ความนึกคิด, และพฤติกรรมของตัวบุคคล [20] ลักษณะของบุคลิกภาพ รวมถึงตัวบุคลิกภาพเองเป็นสิ่งที่เปลี่ยนแปลงได้ยาก จึงมักถูกใช้เป็นตัวแทนของกลุ่มคนในงานวิจัยต่างๆ เช่น การค้นหาความเข้ากันได้ของคนในทีมพัฒนาซอฟต์แวร์ [11] และ การพัฒนาแบบจำลองเพื่อพัฒนาความสามารถในการแข่งขันของผู้พัฒนาซอฟต์แวร์ [21] ตัวอย่างของทฤษฎีอธิบายลักษณะของบุคลิกภาพพื้นฐานของคน เช่น ทฤษฎีบุคลิกภาพห้าองค์ประกอบ ที่อธิบายถึงบุคลิกภาพพื้นฐานห้าด้านของมนุษย์ ได้แก่ บุคลิกภาพแบบเปิดรับประสบการณ์ (Openness to Experience: O), บุคลิกภาพแบบมีจิตสำนึก (Conscientiousness: C), บุคลิกภาพแบบแสดงตัว (Extraversion : E), บุคลิกภาพแบบประนีประนอม (Agreeableness : A) และ บุคลิกภาพแบบหวั่นไหว (Neuroticism : N) ทฤษฎีต้นทุนทางจิตวิทยา (Psychological Capital : PsyCap) [22] ที่อธิบายถึงตัวแปรด้านบวกของมนุษย์ ได้แก่ ความหวัง (Hope) การมองโลกในแง่ดี (Optimism) การรู้จักตนเอง (Self-efficacy) การฟื้นพลัง (Resilence) เป็นต้น

2.4 บุคลิกภาพห้าองค์ประกอบ

ทฤษฎีบุคลิกภาพห้าองค์ประกอบ (Big-Five Personality หรือ Five-Factor Model) เป็นทฤษฎีจำแนกลักษณะของคนออกเป็น 5 กลุ่มใหญ่ [23], [24] ได้แก่ บุคลิกภาพแบบเปิดรับประสบการณ์ (Openness to Experience: O), บุคลิกภาพแบบมีจิตสำนึก (Conscientiousness: C), บุคลิกภาพแบบแสดงตัว (Extraversion : E), บุคลิกภาพแบบประนีประนอม (Agreeableness : A) และ บุคลิกภาพแบบหวั่นไหว (Neuroticism : N) โดยแต่ละคนประกอบด้วยบุคลิกภาพทั้ง 5 รูปแบบนี้ในระดับที่แตกต่างกันออกไป บุคลิกภาพทั้ง 5 รูปแบบสามารถอธิบายในเชิงลักษณะของพฤติกรรมได้ ตารางที่ 2.1 แสดงบุคลิกภาพทั้ง 5 รูปแบบของทฤษฎีบุคลิกภาพห้าองค์ประกอบพร้อมลักษณะของพฤติกรรมที่โดดเด่น

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ตารางที่ 2.1ก บุคลิกภาพของทฤษฎีบุคลิกภาพห้าองค์ประกอบและลักษณะของพฤติกรรมที่โดดเด่น

บุคลิกภาพทั้ง 5 รูปแบบของทฤษฎีบุคลิกภาพห้าองค์ประกอบ	ลักษณะของพฤติกรรมที่โดดเด่น
บุคลิกภาพแบบเปิดรับประสบการณ์ (Openness to Experience: O)	-ช่างฝัน(Fantasy) -การมีอารมณ์สุนทรียภาพ (Aesthetics) -การเปิดเผยความรู้สึก (Feelings) -การปฏิบัติ (Actions) -การมีความคิด (Ideas) -การยอมรับค่านิยม (Values)
บุคลิกภาพแบบมีจิตสำนึก (Conscientiousness: C)	-การมีความสามารถ (Competence) -การมีระเบียบ (Order) -การมีความรับผิดชอบต่อหน้าที่ (Dutifulness) -การมีความต้องการสัมฤทธิผล (Achievement Striving) -การมีวินัยในตนเอง (Self-discipline) -การมีความสุขุม รอบคอบ (Deliberation)
บุคลิกภาพแบบแสดงตัว (Extraversion: E)	-ความอบอุ่น (Warmth) -การชอบอยู่ร่วมกับผู้อื่น (Gregariousness) -การกล้าแสดงออก (Assertiveness) -การชอบทำกิจกรรม (Activity) -การแสวงหาความตื่นเต้น (Excitement Seeking) -การมีอารมณ์เชิงบวก (Positive Emotions)
บุคลิกภาพแบบประนีประนอม (Agreeableness: A)	-การเชื่อใจผู้อื่น (Trust) -ความตรงไปตรงมา (Straightforwardness) -ความเอื้อเฟื้อ (Altruism) -การคล้อยตามผู้อื่น (Compliance) -ความสุภาพ (Modesty) -การมีจิตใจอ่อนไหว (Tender-mindedness)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ตารางที่ 2.1ข บุคลิกภาพของทฤษฎีบุคลิกภาพห้าองค์ประกอบและลักษณะของพฤติกรรมที่โดดเด่น

บุคลิกภาพทั้ง 5 รูปแบบของทฤษฎีบุคลิกภาพห้าองค์ประกอบ	ลักษณะของพฤติกรรมที่โดดเด่น
บุคลิกภาพแบบหวั่นไหว (Neuroticism: N)	-ความวิตกกังวล (Worry) -ความโกรธ (Angry Hostility) -ความท้อแท้ (Discouragement) -การคำนึงถึงแต่ตนเอง (Self-Consciousness) -การถูกกระตุ้น (Impulsiveness) -ความเปราะบาง(Vulnerability)

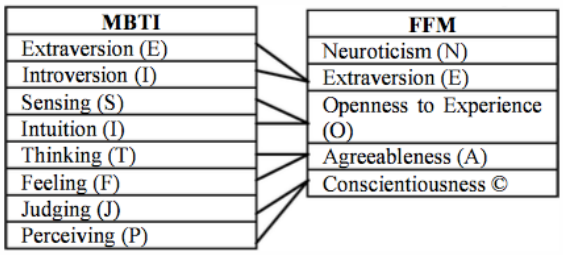
Kanij และคณะ [25] ได้ค้นพบว่าบุคลิกภาพแบบแสดงตัวมีความสัมพันธ์แบบแปรผกผันกับประสิทธิผลในการทดสอบซอฟต์แวร์ โดยมาตรวัดที่ใช้วัดประสิทธิในการทดสอบซอฟต์แวร์อยู่ 3 มาตรวัด ได้แก่ จำนวนของข้อผิดพลาดที่ถูกพบในเวลาต่อเวลาที่ใช้ทดสอบ (Bug Location Rate) ความหนาแน่นของข้อผิดพลาดเมื่อคำนวณรวมกับน้ำหนัก (Weighted Fault Density) และคุณภาพของการรายงานข้อผิดพลาด (Bug report quality) งานวิจัยนี้ค้นพบว่าบุคคลที่เปิดเผยตัว มีความกระตือรือร้น สื่อสารด้วยคำพูดได้ดี ไม่สามารถถูกอนุมานได้ว่าต้องเป็นผู้ที่ทำหน้าที่ทดสอบซอฟต์แวร์ได้ดีเสมอไป ซึ่งขัดกับสมมติฐานของงานที่ตั้งไว้ว่าบุคคลที่เปิดเผยตัวจะเป็นผู้ที่ทำหน้าที่ทดสอบซอฟต์แวร์ได้ดี งานวิจัยนี้แสดงให้เห็นว่าบุคลิกภาพห้าองค์ประกอบส่งผลต่อพัฒนาซอฟต์แวร์

Rehman และคณะ [26] ได้ศึกษาเรื่องตำแหน่งงานที่เกี่ยวกับวิศวกรรมซอฟต์แวร์ ทักษะที่ต้องการในตำแหน่งงานนั้นๆ และทฤษฎีบุคลิกภาพห้าองค์ประกอบ จากนั้นได้เชื่อมโยงแต่ละองค์ความรู้โดยใช้การวิเคราะห์และจับคู่ ตำแหน่งงานที่ถูกนำมาศึกษามีอยู่ 5 ตำแหน่งงาน ได้แก่ นักวิเคราะห์ซอฟต์แวร์ (Software Analyst), นักออกแบบซอฟต์แวร์ (Software Designeer), นักพัฒนาซอฟต์แวร์ (Software Developer), นักทดสอบซอฟต์แวร์ (Software Tester) และ นักซ่อมบำรุงซอฟต์แวร์ (Software Maintenance) ผลลัพธ์ของงานวิจัยพบว่า นักวิเคราะห์ซอฟต์แวร์ ควรมีบุคลิกภาพแบบแสดงตัวและแบบประนีประนอม ในขณะที่นักออกแบบซอฟต์แวร์ควรมีบุคลิกภาพแบบประนีประนอมและแบบเปิดรับประสบการณ์ นักพัฒนาซอฟต์แวร์ควรมีบุคลิกภาพแบบแสดงตัว แบบเปิดรับประสบการณ์และแบบประนีประนอม และนักทดสอบซอฟต์แวร์รวมทั้งนักซ่อมบำรุงซอฟต์แวร์ ควรมีบุคลิกภาพแบบเปิดรับประสบการณ์และแบบมีจิตสำนึก ซึ่งในส่วนของบุคลิกภาพที่

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

นักทดสอบซอฟต์แวร์ควรมี ไม่ขัดแย้งกับผลลัพธ์ของงานวิจัยของ Kanij และคณะ ที่กล่าวไว้ในข้างต้น

ทั้งนี้งานวิจัยของ Rehman และคณะในข้างต้นพัฒนามาจากงานวิจัยของ Capre และ Ahmed [27] ที่ใช้ทฤษฎีตัวชี้วัดของไมเออร์-บริกส์ (Myers-Briggs Type Indicator :MBTI) เป็นทฤษฎีอธิบายบุคลิกภาพ โดย Rehman และคณะ ให้เหตุผลในการเลือกใช้บุคลิกภาพห้าองค์ประกอบแทนทฤษฎีของไมเออร์-บริกส์ว่า ทฤษฎีบุคลิกภาพห้าองค์ประกอบเป็นที่ยอมรับมากกว่า อย่างไรก็ตาม Rehman และคณะได้เพิ่มเติมว่าทั้งสองทฤษฎีนี้มีความเชื่อมโยงและอาจเป็นตัวแทนซึ่งกันและกันได้ ความเชื่อมโยงระหว่างสองทฤษฎีแสดงในภาพที่ 2.4 โดยการเปิดเผย (Extraversion: E) และการเก็บตัว (Introversion: I) ของไมเออร์-บริกส์ มีความเชื่อมโยงกับบุคลิกภาพแบบแสดงตัวของบุคลิกภาพห้าองค์ประกอบ การสัมผัส (Sensing: S) และการใช้จิตใต้สำนึก (Intuition: I หรือ N) ของไมเออร์-บริกส์ มีความเชื่อมโยงกับบุคลิกภาพแบบเปิดรับประสบการณ์ของบุคลิกภาพห้าองค์ประกอบ ความคิด (Thinking: T) และความรู้สึก (Feeling: F) ของไมเออร์-บริกส์ มีความเชื่อมโยงกับบุคลิกภาพแบบประนีประนอมของบุคลิกภาพห้าองค์ประกอบ และความมีระเบียบ (Judging: J) และความยืดหยุ่น (Perceiving: P) ของไมเออร์-บริกส์ มีความเชื่อมโยงกับบุคลิกภาพแบบมีจิตสำนึกของบุคลิกภาพห้าองค์ประกอบ จะมีเพียงบุคลิกภาพแบบหวั่นไหว (Neuroticism: N) ของบุคลิกภาพห้าองค์ประกอบเท่านั้นที่ไม่มีความเชื่อมโยงกับทฤษฎีของไมเออร์-บริกส์



ภาพที่ 2.4 แสดงความเชื่อมโยงระหว่างทฤษฎีบุคลิกภาพของไมเออร์-บริกส์และบุคลิกภาพห้าองค์ประกอบ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

Chaiwon และคณะ [28] ได้ศึกษาถึงความสัมพันธ์ของบุคลิกภาพห้าองค์ประกอบและตำแหน่งงานที่เกี่ยวกับวิศวกรรมซอฟต์แวร์เช่นเดียวกับ Rehman และคณะ แต่สิ่งที่ต่างคือวิธีการวิเคราะห์ข้อมูลและตำแหน่งงานที่ศึกษา โดยงานวิจัยของ Chaiwon และคณะได้ใช้ตำแหน่งงานได้แก่ นักวิเคราะห์ซอฟต์แวร์ (Software Analyst), นักออกแบบซอฟต์แวร์ (Software Designeer), โปรแกรมเมอร์ (Programmer), นักทดสอบซอฟต์แวร์ (Software Tester) และตำแหน่งงานอื่นๆที่

นอกเหนือจากที่กล่าวมา Chaiwon และคณะได้นำเสนอวิธีการทำนายตำแหน่งงานดังกล่าวจากบุคลิกภาพห้าองค์ประกอบ โดยเก็บข้อมูลบุคลิกภาพจากผู้ที่ทำงานเกี่ยวกับการพัฒนาซอฟต์แวร์ด้วยแบบทดสอบบิกไฟว์อินเวนทอรี (Big Five Inventory: BFI) จากนั้นนำมาให้คะแนน หาคะแนนเฉลี่ยจากผู้ตอบแบบทดสอบแต่ละอาชีพในแต่ละข้อ และนำมาเรียงลำดับจากมากไปน้อยเพื่อค้นหาว่า ข้อคำถามใดมีความเกี่ยวข้องกับอาชีพนั้นๆ มากที่สุด 10 อันดับแรก ข้อคำถามเหล่านี้สามารถนำไปใช้หาอาชีพที่เหมาะสมสำหรับผู้ตอบแบบทดสอบได้โดยเปรียบเทียบกับคำถาม 10 อันดับแรกที่ผู้ตอบแบบทดสอบได้มากที่สุด วิธีที่นำเสนอมีความถูกต้อง (Accuracy) อยู่ที่ 70% อย่างไรก็ตามวิธีนี้อาจไม่เหมาะสมเมื่อนำมาประยุกต์ใช้กับแบบทดสอบทางจิตวิทยามีจำนวนข้อน้อยกว่าหรือเท่ากับ 10 เนื่องจากจะทำให้ผลลัพธ์การทำนายได้ออกมาเหมือนกันทุกกรณี เช่นแบบทดสอบวัดความเคารพในตนเองของโรเซนเบิร์ก (Rosenberg Self Esteem Scales :RSES) ที่ใช้วัดความเคารพในตนเองของบุคคล ประกอบด้วยข้อคำถามทั้งสิ้น 10 ข้อ เป็นต้น และนอกจากเรื่องการเปลี่ยนแบบทดสอบในภายหลังแล้ว งานวิจัยของ Chaiwon และคณะไม่ได้วิเคราะห์ต่อว่า ข้อคำถามใดบ้างในแบบทดสอบเกี่ยวข้องกับตำแหน่งงานที่นำมาศึกษา และไม่ได้ศึกษาถึงความเกี่ยวข้องกับระเบียบวิธีแบบอไจล์ อย่างไรก็ตาม จากงานของ Chaiwon และคณะได้แสดงให้เห็นว่า การที่จะวิเคราะห์ว่าบุคคลใดมีบุคลิกภาพเป็นอย่างไร จำเป็นต้องใช้เครื่องมือที่มีคุณภาพเพื่อใช้ในการทดสอบบุคลิกภาพของบุคคล

2.5 แบบทดสอบบุคลิกภาพ

แบบทดสอบบุคลิกภาพ (Personality Test) เป็นเครื่องมือที่ใช้วัดและจำกัดลักษณะของบุคคลออกเป็นประเภทตามทฤษฎีบุคลิภาพ โดยทฤษฎีบุคลิกภาพที่แตกต่างกันแบบทดสอบที่ใช้ย่อมแตกต่างกันด้วย ดังนั้นแบบทดสอบที่ถูกนำมาศึกษาในงานวิจัยนี้ได้แก่แบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรี เนื่องจากเป็นแบบทดสอบทางบุคลิกภาพที่ถูกสร้างขึ้นมาเพื่อใช้วัดทฤษฎีบุคลิกภาพห้าองค์ประกอบโดยเฉพาะ [29]

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

แบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรีเป็นเครื่องมือวัดลักษณะของบุคลิกภาพของมนุษย์ [12] แบบทดสอบดังกล่าวอยู่ในรูปแบบของชุดข้อความ ผู้ทดสอบต้องเลือกข้อความที่ตรงกับความเห็นหรือสิ่งที่ตนเองเป็นมากที่สุดในการทำแบบทดสอบ ผลลัพธ์ที่ได้จากการทดสอบออกมาในรูปแบบของคะแนนบุคลิกภาพในแต่ละหมวด แบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรี สามารถแบ่งคะแนนออกเป็น 5 หมวดตามบุคลิกภาพห้าองค์ประกอบ

แบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรี ประกอบด้วยข้อความที่เป็นคำถามทั้งหมด 44 ข้อ และมีตัวเลือกให้ผู้ทดสอบเลือกตอบอยู่ 5 ตัวเลือกในรูปแบบของลิเคิร์ทส

เกล (Likert Scale) [30] ผู้ทดสอบต้องแสดงความเห็นด้วยหรือไม่เห็นด้วยกับข้อความแต่ละข้อความ โดยใช้ตัวเลือกดังนี้ ไม่เห็นด้วยอย่างมาก (Disagree strongly), ไม่เห็นด้วยเล็กน้อย (Disagree a little), ไม่ทั้งเห็นด้วยและไม่เห็นด้วย (Neither agree nor Disagree), เห็นด้วยเล็กน้อย (Agree a little) หรือ เห็นด้วยอย่างมาก (Agree strongly) ตัวอย่างของแบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรีแสดงอยู่ในตารางที่ 2.2

ตารางที่ 2.2ก แบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรี

I See Myself as Someone Who.... (ข้อความส่วนที่ 1)			
ข้อ	ข้อความส่วนที่ 2	ข้อ	ข้อความส่วนที่ 2
1	Is talkative	2	Tends to find fault with others
3	Does a thorough job	4	Is depressed, blue
5	Is original, comes up with new ideas	6	Is reserved
7	Is helpful and unselfish with others	8	Can be somewhat careless
9	Is relaxed, handles stress well	10	Is curious about many different things
11	Is full of energy	12	Starts quarrels with others

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ตารางที่ 2.2ข แบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรี

I See Myself as Someone Who.... (ข้อความส่วนที่ 1)			
ข้อ	ข้อความส่วนที่ 2	ข้อ	ข้อความส่วนที่ 2
13	Is a reliable worker	14	Can be tense
15	Is ingenious, a deep thinker	16	Generates a lot of enthusiasm
17	Has a forgiving nature	18	Tends to be disorganized
19	Worries a lot	20	Has an active imagination
21	Tends to be quiet	22	Is generally trusting
23	Tends to be lazy	24	Is emotionally stable, not easily upset
25	Is inventive	26	Has an assertive personality
27	Can be cold and aloof	28	Perseveres until the task is finished
29	Can be moody	30	Values artistic, aesthetic experiences
31	Is sometimes shy, inhibited	32	Is considerate and kind to almost everyone
33	Does things efficiently	34	Remains calm in tense situations
35	Prefers work that is routine	36	Is outgoing, sociable
37	Is sometimes rude to others	38	Makes plans and follows through with them
39	Gets nervous easily	40	Likes to reflect, play with ideas
41	Has few artistic interests	42	Likes to cooperate with others
43	Is easily distracted	44	Is sophisticated in art, music, or literature

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

การคิดคะแนนของแบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรี สามารถถูกแบ่งออกไปตามบุคลิกภาพห้าองค์ประกอบ ตารางที่ 2.3 แสดงชุดของข้อความในแบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรี จำแนกตามบุคลิกภาพห้าองค์ประกอบ คะแนนที่ได้มาบ่งชี้ว่าผู้ทดสอบนั้นๆ ว่ามีความโดดเด่นในบุคลิกภาพด้านใดในบุคลิกภาพห้าองค์ประกอบ เงื่อนไขในการให้คะแนนมีดังนี้

ข้อที่ 1, 3, 4, 5, 7, 10, 11, 13-17, 19, 20, 22, 25, 26, 28-30, 32, 33, 36, 38-40, 42 และ 44

หากผู้ทดสอบ	เลือก	ไม่เห็นด้วยอย่างมาก	ได้ 1 คะแนนในข้อนั้น
	เลือก	ไม่เห็นด้วยเล็กน้อย	ได้ 2 คะแนนในข้อนั้น
	เลือก	ไม่ทั้งเห็นด้วยและไม่เห็นด้วย	ได้ 3 คะแนนในข้อนั้น
	เลือก	เห็นด้วยเล็กน้อย	ได้ 4 คะแนนในข้อนั้น
	เลือก	เห็นด้วยอย่างมาก	ได้ 5 คะแนนในข้อนั้น

ข้อที่ 2, 6, 8, 9, 12, 18, 21, 23, 24, 27, 31, 34, 35, 37, 41, และ 43

หากผู้ทดสอบ	เลือก	ไม่เห็นด้วยอย่างมาก	ได้ 5 คะแนนในข้อนั้น
	เลือก	ไม่เห็นด้วยเล็กน้อย	ได้ 4 คะแนนในข้อนั้น
	เลือก	ไม่ทั้งเห็นด้วยและไม่เห็นด้วย	ได้ 3 คะแนนในข้อนั้น
	เลือก	เห็นด้วยเล็กน้อย	ได้ 2 คะแนนในข้อนั้น
	เลือก	เห็นด้วยอย่างมาก	ได้ 1 คะแนนในข้อนั้น

ตารางที่ 2.3 ชุดของข้อความในแบบทดสอบวัดบุคลิกภาพห้าองค์ประกอบแบบบิกไฟว์อินเวนทอรีจำแนกตามบุคลิกภาพห้าองค์ประกอบ [29]

บุคลิกภาพของทฤษฎีบุคลิกภาพห้าองค์ประกอบ	ข้อ
บุคลิกภาพแบบเปิดรับประสบการณ์	5, 10, 15, 20, 25, 30, 35, 40, 41 และ 44
บุคลิกภาพแบบมีจิตสำนึก	3, 8, 13, 18, 23, 28, 33, 38 และ 43
บุคลิกภาพแบบแสดงตัว	1, 6, 11, 16, 21, 26, 31 และ 36
บุคลิกภาพแบบประนีประนอม	2, 7, 12, 17, 22, 27, 32, 37 และ 42
บุคลิกภาพแบบหวั่นไหว	4, 9, 14, 19, 24, 29, 34 และ 39

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ทั้งนี้คะแนนต่อบุคลิกภาพที่ได้จะถูกนำไปทำนายและจัดกลุ่มเพื่อหาความสัมพันธ์ระหว่างความเป็นบุคคลอไจล์และลักษณะทางบุคลิกภาพ

2.6 การทำนายและการจัดกลุ่มข้อมูล

การทำนาย (Prediction) คือ การคาดการณ์ถึงสิ่งที่เกิดขึ้นในอนาคตหรือเกิดขึ้นหลังจากเหตุการณ์หนึ่ง [31], [32] ในกรณีของการทำนายบุคคลอไจล์ หมายถึงการคาดการณ์ว่าผู้พัฒนาซอฟต์แวร์เป็นบุคคลอไจล์หรือไม่โดยใช้ข้อมูลบุคลิกภาพที่มีอยู่ ในปัจจุบันเทคนิคที่นิยมใช้มากที่สุดในการทำนายข้อมูลได้แก่ เทคนิคการจัดกลุ่มข้อมูล (Classification) [33]–[35]

เทคนิคการจัดกลุ่มข้อมูลเป็นวิธีการจัดการข้อมูลที่มีอยู่ออกเป็นกลุ่มๆ (Classes) ซึ่งการจัดกลุ่มข้อมูลสามารถทำได้แม้ไม่ทราบว่าแท้จริงแล้วข้อมูลนั้นถูกจัดอยู่ในกลุ่มใด (Unknown Data) โดยทั่วไปแล้วเทคนิคการจัดกลุ่มสามารถจัดกลุ่มข้อมูลได้จากการอ้างอิงความสัมพันธ์ของข้อมูลที่มีอยู่แล้วในอดีต ซึ่งแต่ละเทคนิคการจัดกลุ่มมีวิธีการหาความสัมพันธ์ของข้อมูลที่มีอยู่แตกต่างกันออกไป เทคนิคการจัดกลุ่มข้อมูลที่นิยมใช้ทำนายข้อมูล มีอยู่ 4 เทคนิค ได้แก่

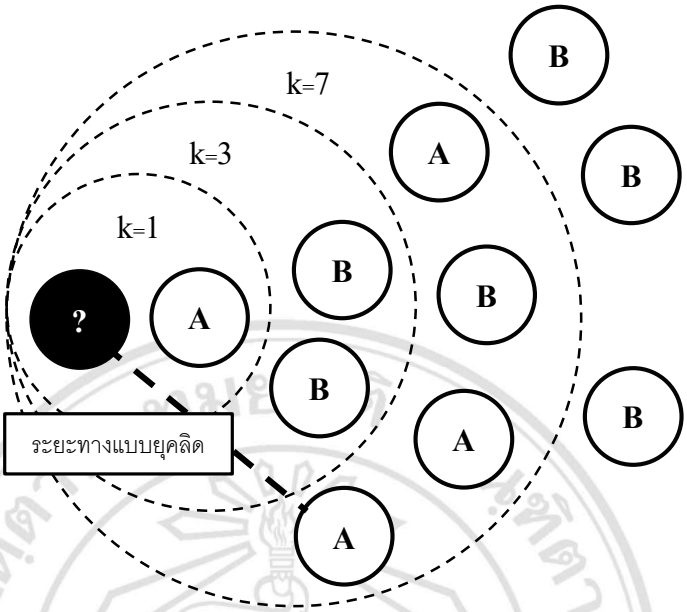
1. เทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด (k-Nearest Neighbor: k-NN)
2. เทคนิคการจัดกลุ่มข้อมูลแบบนาอีฟเบย์ (Naïve Bayes)
3. เทคนิคการจัดกลุ่มข้อมูลแบบต้นไม้ตัดสินใจ (Decision Tree)
4. เทคนิคการจัดกลุ่มข้อมูลแบบโครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

โดยแต่ละเทคนิคการจัดกลุ่มข้อมูลมีรายละเอียดดังนี้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.6.1 เทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด

เทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด(k-Nearest Neighbor: k-NN) เป็นเทคนิคที่ใช้การเปรียบเทียบระยะห่างระหว่างข้อมูลในการจัดกลุ่ม [36] โดยอาศัยแนวคิดที่ว่าข้อมูลที่จะจัดกลุ่มมีลักษณะเช่นเดียวกับข้อมูลเพื่อนบ้าน (Neighbor) ที่ใกล้ที่สุด ข้อมูลเพื่อนบ้าน คือข้อมูลที่มีอยู่แล้ว และสามารถถูกระบุได้ว่าถูกจัดอยู่ในกลุ่มใด เทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุดจะต้องกำหนดจำนวนข้อมูลเพื่อนบ้าน (k) เพื่อกำหนดชุดของการเปรียบเทียบ โดยจำนวนข้อมูลเพื่อนบ้านที่แตกต่างกันส่งผลให้ความถูกต้องในการทำนายแตกต่างกันได้ ภาพที่ 2.5 แสดงตัวอย่างของเทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด



ภาพที่ 2.5 ตัวอย่างของเทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด

จากภาพที่ 2.5 ข้อมูลเพื่อนบ้านถูกแบ่งออกเป็นสองกลุ่มแตกต่างกัน สมมติให้มีข้อมูลเพื่อนบ้านจำนวน 10 ข้อมูลถูกจัดกลุ่มให้อยู่ในกลุ่ม “A” หรือ กลุ่ม “B” เมื่อต้องการทราบว่าข้อมูลใหม่ถูกจัดอยู่ในกลุ่มใด จำนวนข้อมูลเพื่อนบ้านที่แตกต่างกันจะให้ผลลัพธ์ที่ต่างกัน ดังนี้

- ในกรณีที่กำหนดให้ข้อมูลเพื่อนบ้านที่ใกล้ที่สุดมีเพียงข้อมูลเดียว (k=1) ข้อมูลใหม่ถูกจัดอยู่ในกลุ่ม A เนื่องจากข้อมูลเพื่อนบ้านที่ใกล้ที่สุดถูกจัดอยู่ในกลุ่ม A
- ในกรณีที่กำหนดให้ข้อมูลเพื่อนบ้านที่ใกล้ที่สุดมี 3 ข้อมูล (k=3) ข้อมูลใหม่ถูกจัดอยู่ในกลุ่ม B เนื่องจากในข้อมูลเพื่อนบ้านที่ใกล้ที่สุด 3 ข้อมูล มีข้อมูล B อยู่จำนวน 2 ข้อมูลและมีข้อมูล A อยู่จำนวน 1 ข้อมูล
- ในกรณีที่กำหนดให้ข้อมูลเพื่อนบ้านที่ใกล้ที่สุดมี 7 ข้อมูล (k=7) ข้อมูลใหม่ถูกจัดอยู่ในกลุ่ม A เนื่องจากในข้อมูลเพื่อนบ้านที่ใกล้ที่สุด 7 ข้อมูล มีข้อมูล A อยู่จำนวน 4 ข้อมูลและมีข้อมูล B อยู่จำนวน 3 ข้อมูล

ระยะห่างระหว่างข้อมูลในเทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุดสามารถคำนวณได้โดยใช้สมการหาระยะทางแบบยุคลิด (Eucledian Distance) ซึ่งจะวัดระยะกระจัดระหว่างสองข้อมูลนี้ [37] สมการหาระยะทางแบบยุคลิดแสดงในภาพที่ 2.6

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$$□ = \sqrt{(□□_0 - □□_0)^2 + (□□_1 - □□_1)^2}$$

เมื่อ d คือ ระยะทางระหว่าง 2 เวกเตอร์, xA และ xB, จากจุดสองจุด

ภาพที่ 2.6 สมการหาระยะทางแบบยุคลิด

ทั้งนี้อัลกอริทึมเบื้องต้นของเทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุดแสดงใน ภาพที่ 2.7

k-NN Algorithm

1. Define k as the number of nearest neightbors and D as the set of training set examples
2. For each test examples Z
3. Find distance between Z and every example d, where d $\in$ D
4. Select D_z, where D_z is the set of k closest training examples to Z
5. Classified Z based on the majority class of D_z
6. end for

ภาพที่ 2.7 อัลกอริทึมเบื้องต้นของเทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด

จากภาพที่ 2.7 การใช้เทคนิคการจัดกลุ่มข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุดต้อง กำหนดจำนวนเพื่อนบ้านที่ใกล้ที่สุด (Number of nearest neightbors: k) เตรียมชุดของข้อมูลที่ใช้สร้างแบบจำลอง (Training set examples: D) และข้อมูลที่ต้องการทำนาย (Test example: Z) จากนั้นหาระยะห่างระหว่างข้อมูลที่ต้องการทำนายและข้อมูลที่ใช้สร้างแบบจำลองทุกตัวในชุด (แสดงดังบรรทัดที่ 3 ของอัลกอริทึม) เมื่อได้ระห่างของข้อมูลทั้งหมดในชุดแล้ว ในบรรทัดที่ 4 อัลกอริทึมจะเลือกเอาชุดของข้อมูลที่ใช้สร้างแบบจำลองบางส่วน (Set of closest training examples to Z: Dz) โดยเลือกจากข้อมูลที่ใช้สร้างที่มีระยะห่างจากข้อมูลที่ต้องการทำนายน้อยที่สุด (หรือใกล้ที่สุด) k ข้อมูล ในขั้นตอนสุดท้าย ข้อมูลที่ต้องการทำนายจะถูกจัดให้อยู่ในกลุ่มเดียวกับข้อมูลที่ใช้สร้างแบบจำลองที่เป็นเสียงส่วนใหญ่ (Majority) ในชุดของข้อมูลที่ใช้สร้างแบบจำลองที่เลือกไว้ ในกรณีที่มีข้อมูลที่ต้องการทำนายมากกว่า 1 ตัวให้ทำบรรทัดที่ 3-5 ซ้ำจนกว่าข้อมูลที่ต้องการทำนายทั้งหมดได้ถูกจัดกลุ่มเป็นที่เรียบร้อย

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.6.2 เทคนิคการจัดกลุ่มข้อมูลแบบนาอีฟเบย์

เทคนิคการจัดกลุ่มข้อมูลแบบนาอีฟเบย์ (Naïve Bayes) เป็นเทคนิคการจัดกลุ่มที่อาศัยทฤษฎีเรื่องความน่าจะเป็นของเบย์ มาประยุกต์ใช้ในการจัดกลุ่ม [38] ทฤษฎีความน่าจะเป็นของเบย์เสนอว่า เหตุการณ์หนึ่งๆ สามารถทำนายได้จากการเกิดของเหตุการณ์ต่างๆ ที่อาจมีความเกี่ยวข้องกันหรือไม่ก็ได้ ดังสมการในภาพที่ 2.8 และ 2.9 จุดเด่นของเทคนิคการจัดกลุ่มนาอีฟเบย์ คือ สามารถใช้ทำนายได้ดีแม้ว่ามีข้อมูลจำนวนน้อย

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$P(B|A)$ คือ ความน่าจะเป็นที่เกิดเหตุการณ์ A ก่อนแล้วจะมีเหตุการณ์ B ตามมา

$P(A \cap B)$ คือ ความน่าจะเป็นที่เกิดเหตุการณ์ A และ B ร่วมกัน

$P(A)$ คือ ความน่าจะเป็นที่เหตุการณ์ A จะเกิดขึ้น

ภาพที่ 2.8 สมการหาความน่าจะเป็นที่เกิดเหตุการณ์ A ก่อนแล้วจะมีเหตุการณ์ B ตามมา

$$P(C|A) = P(a_1|C) \times P(a_2|C) \times \ldots \times P(a_m|C) \times P(C)$$ เมื่อ

$P(C|A)$ คือ ความน่าจะเป็นที่เหตุการณ์ C จะเกิดขึ้นจากเหตุการณ์ A หรือข้อมูล A

a_m คือ จำนวนแอตทริบิวต์ที่อยู่ในเหตุการณ์ A หรือข้อมูล A

C คือ เหตุการณ์ที่จะเกิดขึ้น

ภาพที่ 2.9 สมการหาความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์ A จะมีคลาส C

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ภาพที่ 2.8 เป็นสมการเบื้องต้นที่ใช้หาเหตุการณ์ A ก่อนแล้วจะมีเหตุการณ์ B ตามมาตามทฤษฎีของเบย์ [39] ในการทำนายข้อมูลจำเป็นต้องใช้ ความน่าจะเป็นที่เหตุการณ์ A ก่อนจึงจะสามารถทำนายเหตุการณ์ B ที่เกิดขึ้นได้ ซึ่งสมการนี้จะถูกใช้ในในภาพที่ 2.9

ภาพที่ 2.9 ที่เป็นสมการที่ใช้ทำนายข้อมูลโดยแอตทริบิวต์ (Attribute) ของชุดของข้อมูลสร้างแบบจำลองทั้งหมดจะถูกคำนวณออกมาในรูปแบบของความน่าจะเป็น จากนั้นผลคูณของความน่าจะเป็นของทุกแอตทริบิวต์จะเป็นตัวบ่งชี้ว่า ข้อมูลที่ถูกทำนายควรจัดอยู่ในคลาสใด โดยมีหลักการว่าข้อมูลที่ถูกทำนายจะถูกจัดอยู่ในคลาสที่มีความน่าจะเป็นที่มากที่สุด

ทั้งนี้ เหตุการณ์ C ที่เกิดขึ้นได้ทั้งหมดจะถูกนำมาแบ่งเป็นคลาสต่างๆ ดังนั้นผลลัพธ์ที่ได้คือความน่าจะเป็นที่จะเกิดคลาสนั้นๆ ตัวอย่างของการประยุกต์ใช้เทคนิคการจัดกลุ่มแบบนาอีฟเบย์ เช่น การทำนายโอกาสที่เกิดฝนตกเมื่อมีเมฆดำโดยใช้สมการดังภาพที่ 2.10 และ 2.11

$$P(\text{ฝนตก}|\ A) = P(\text{เมฆดำ}\ |\ \text{ฝนตก}) \times P(\text{ฝนตก})$$

เมื่อ P(ฝนตก | A) คือ ความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์ A จะมีคลาส “ฝนตก”
P(เมฆดำ| ฝนตก) คือ ความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์ “เมฆดำ” จะมีคลาส “ฝนตก”
P(ฝนตก) คือ ความน่าจะเป็นที่ฝนตก
โดย

$$P(\text{เมฆดำ}|\text{ฝนตก}) = \frac{\square(\text{ฝนตก} \cap \text{เมฆดำ})}{\square(\text{ฝนตก})}$$

เมื่อ P(เมฆดำ|ฝนตก) คือ ความน่าจะเป็นที่เกิดเหตุการณ์ “ฝนตก” ก่อนแล้วจะมีเหตุการณ์ “เมฆดำ” ตามมา
n(ฝนตก ∩ เมฆดำ) คือ จำนวนเกิดเหตุการณ์ “ฝนตก” และ “เมฆดำ” ร่วมกัน
n (ฝนตก) คือ จำนวนที่เหตุการณ์ “ฝนตก” เกิดขึ้น

ภาพที่ 2.10 ตัวอย่างการทำนายฝนตกเมื่อมีเมฆโดยใช้ทฤษฎีของเบย์ (คิดกรณีฝนตก)

$$P(\text{ฝนไม่ตก}|\ A) = P(\text{เมฆดำ}\ |\ \text{ฝนไม่ตก}) \times P(\text{ฝนไม่ตก})$$

เมื่อ P(ฝนไม่ตก | A) คือ ความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์ A จะมีคลาส “ฝนไม่ตก”
P(เมฆดำ| ฝนไม่ตก) คือ ความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์ “เมฆดำ” จะมีคลาส “ฝนไม่ตก”
P(ฝนไม่ตก) คือ ความน่าจะเป็นที่ฝนไม่ตก
โดย

$$P(\text{เมฆดำ}|\text{ฝนไม่ตก}) = \frac{\square(\text{ฝนไม่ตก} \cap \text{เมฆดำ})}{\square(\text{ฝนไม่ตก})}$$

เมื่อ P(เมฆดำ|ฝนตก) คือ ความน่าจะเป็นที่เกิดเหตุการณ์ “ฝนตก” ก่อนแล้วจะมีเหตุการณ์ “เมฆดำ” ตามมา
n(ฝนไม่ตก ∩ เมฆดำ) คือ ความน่าจะเป็นที่เกิดเหตุการณ์ “ฝนไม่ตก” และ “เมฆดำ” ร่วมกัน
n (ฝนไม่ตก) คือ ความน่าจะเป็นที่เหตุการณ์ “ฝนไม่ตก” จะเกิดขึ้น

ภาพที่ 2.11 ตัวอย่างการทำนายฝนตกเมื่อมีเมฆโดยใช้ทฤษฎีของเบย์ (คิดกรณีฝนไม่ตก)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

จากภาพที่ 2.10 และ 2.11 แสดงความน่าจะเป็นที่เกิดเหตุการณ์ฝนตกเมื่อมีเมฆตามทฤษฎีของเบย์ สมมติให้ในหนึ่งปีมีฝนตก 256 วัน จาก 365 วัน และภายในปีเดียวกันนั้นเองมีเมฆดำอยู่ 160 วัน มีวันที่เมฆดำและฝนตกพร้อมกันอยู่ 90 วัน จากข้อมูลเหล่านี้ หากต้องการทราบว่าวันนี้หากมีเมฆดำแล้ว (เหตุการณ์ที่สนใจ: A) ฝนจะตกหรือไม่ ต้องเปรียบเทียบความน่าจะเป็นที่ ฝนตกเมื่อมีเมฆดำ และ ฝนไม่ตกเมื่อมีเมฆดำ ว่าความน่าจะเป็นใดที่มีค่ามากกว่า จึงจะสามารถทำนายได้ว่า วันนี้ฝนจะตกหรือไม่ โดยจากข้อมูลที่สมมติให้ สามารถเขียนเป็นสมการได้ดังนี้

$$P(\text{ฝนตก}|\ A) = \frac{\square(\text{ฝนตก} \cap \text{เมฆดำ})}{\square(\text{ฝนตก})} \text{ x } P(\text{ฝนตก})$$

$$= \frac{90}{256} \text{ x } \frac{256}{365}$$

$$= 0.34 \text{ x } 0.7$$

$$= 0.238$$

และ

$$P(\text{ฝนไม่ตก}|\ A) = \frac{\square(\text{ฝนไม่ตก} \cap \text{เมฆดำ})}{\square(\text{ฝนไม่ตก})} \text{ x } P(\text{ฝนไม่ตก})$$

$$= \frac{160-90}{365-256} \text{ x } (1-\frac{256}{365})$$

$$= \frac{70}{91} \text{ x } (1-0.7)$$

$$= 0.77 \text{ x } 0.3$$

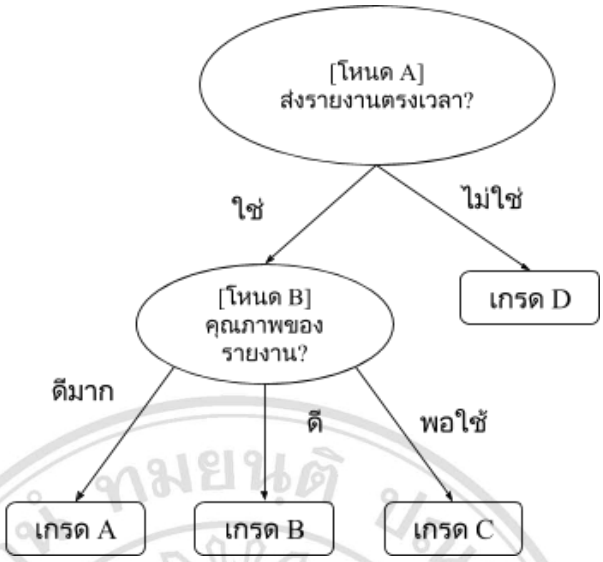
$$= 0.231$$

เมื่อเปรียบเทียบความน่าจะเป็นข้างต้นแล้ว โอกาสที่ฝนจะตกเมื่อมีเมฆดำมีมากกว่าฝนไม่ตก (0.238 > 0.231) ดังนั้นจึงสามารถทำนายได้ว่า วันนี้ฝนจะตก

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.6.3 เทคนิคการจัดกลุ่มข้อมูลแบบต้นไม้ตัดสินใจ

เทคนิคการจัดกลุ่มข้อมูลแบบต้นไม้ตัดสินใจ (Decision Tree) [35], [40], [41] เป็นเทคนิคสนับสนุนการตัดสินใจโดยใช้กราฟแบ่งเส้นทาง และนำไปสู่ผลลัพธ์ที่อาจเป็นได้ ต้นไม้ตัดสินใจ ประกอบด้วย โหนด (Node) และ กิ่ง (Branch) โดย โหนดทำหน้าที่กำหนดเงื่อนไขการตัดสินใจ เช่น “วันนี้เป็นวันจันทร์?” “ผลลัพธ์สุดท้ายที่คำนวณออกมาได้เป็นเลขคู่?” เพื่อให้การตัดสินใจดำเนินต่อไปตามกิ่งและเข้าสู่โหนดถัดไป



ภาพที่ 2.12 ตัวอย่างของเทคนิคการจัดกลุ่มข้อมูลแบบต้นไม้ตัดสินใจ
กรณีการทำนายเกรดจากรายงานที่ถูกส่งมา

ตัวอย่างของต้นไม้ตัดสินใจ แสดงอยู่ในภาพที่ 2.12 กราฟนี้สามารถใช้จัดกลุ่มและทำนายเกรดของรายงานที่ถูกส่งมาได้ กำหนดให้มี โหนดที่เป็นเงื่อนไขอยู่ 2 โหนดได้แก่ โหนด A และ B โดยแต่ละโหนดทำหน้าที่ดังนี้

- โหนด A ตรวจสอบเงื่อนไขว่ารายงานที่ถูกส่งมา ถูกส่งมาภายในเวลาที่กำหนดหรือไม่ สามารถแบ่งได้อีกเป็น 2 กรณี ดังนี้
  - กรณี “ไม่ใช่” การตัดสินใจเดินทางไปทางกิ่ง “ไม่ใช่” และรายงานฉบับนั้นถูกจัดอยู่ในกลุ่มของรายงานที่ได้เกรด D
  - กรณี “ใช่” การตัดสินใจเดินทางไปทางกิ่ง “ใช่” แทน และเข้าสู่โหนด B
- โหนด B ตรวจสอบว่าคุณภาพของรายงานอยู่ในระดับใด สามารถแบ่งได้อีกเป็น 2 กรณี ดังนี้
  - กรณี “รายงานนั้นมีคุณภาพอยู่ในระดับที่ดีมาก” รายงานฉบับนั้นจัดอยู่ในรายงานที่ได้เกรด A
  - กรณี “รายงานนั้นมีคุณภาพอยู่ในระดับที่ดี” รายงานฉบับนั้นจัดอยู่ในรายงานที่ได้เกรด B
  - กรณี “รายงานนั้นมีคุณภาพอยู่ในระดับที่พอใช้” รายงานฉบับนั้นจัดอยู่ในรายงานที่ได้เกรด C

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

เทคนิคการจัดกลุ่มข้อมูลต้นไม้ตัดสินใจสามารถใช้ทำนายข้อมูลที่สามารถจำกัดค่าได้ เช่น ข้อมูลที่เป็นได้เพียงสองค่าเท่านั้น (ใช่ หรือ ไม่ใช่, ถูกต้องหรือไม่ถูกต้อง) หรือข้อมูลที่ถูกจัดเป็นกลุ่มๆ (เกรด A, B, หรือ C) โดยต้องทราบตัวแปรทำนาย [41] เพื่อให้สามารถตรวจสอบความถูกต้องของเงื่อนไขได้ ซึ่งจากตัวอย่างข้างต้น ตัวแปรทำนายคือ เวลาที่ส่งงาน และ คุณภาพของรายงาน

2.6.4 เทคนิคการจัดกลุ่มข้อมูลแบบโครงข่ายประสาทเทียม

เทคนิคการจัดกลุ่มข้อมูลแบบโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) เป็นการจัดกลุ่มข้อมูลที่จำลองการทำงานจากเซลล์ประสาท และ เครือข่ายของเซลล์ที่ทำงานร่วมกัน [40], [41] หลักการทำงานของโครงข่ายประสาทเทียม คือ เมื่อข้อมูลที่เข้ามาในเครือข่ายมีน้ำหนักรวมเกินกว่าค่าตั้งต้น (Threshold) ที่ถูกกำหนดไว้ในแต่ละโหนด (Node) ข้อมูลเหล่านั้นถูกส่งต่อไปยังโหนดอื่นต่อไป การส่งต่อของข้อมูลดังกล่าวทำให้เกิดการเรียนรู้ขึ้นภายในเครือข่าย อัลกอริทึมเบื้องต้นของเทคนิคการจัดกลุ่มแบบโครงข่ายประสาทเทียมแสดงดังภาพที่ 2.13

Artificial Neural Network

1. Set all weight and node thresholds (by random)
2. Caculate active level (using Sigmoid fuction)
   Active level $(O_j) = F(\sum \square_{\square\square} \square_{\square} - \Theta_{\square})$,
   where $\square_{\square\square}$ is weight from an input $\square_{\square}$, $\Theta_{\square}$ is node thershold
   Sigmoid fucntion is $F(a) = 1 / 1+e^{-a}$
3. Adjust weights by function: $W_{ji}(t+1) = W_{ji}(t) + \Delta W_{ji}$,
   where $W_{ji}(t)$ = weight from unit i to j at time t
   ΔW_{ji} = weight adjustment, computed by (learning rate)$(error_j)(O_i)$,
   where $error_j = O_j (1- O_j)(T_j- O_j)$,
   where T_j is actual output
4. Repeat step 2-4 until reach stopping criteria, which can be:
   1. Number of trainning cycles
   2. Accepable error rate

ภาพที่ 2.13 อัลกอริทึมเบื้องต้นของเทคนิคการจัดกลุ่มแบบโครงข่ายประสาทเทียม

ในขั้นตอนแรก น้ำหนัก (weight) ของแต่โหนดจะถูกกำหนดขึ้น อาจโดยวิธีการสุ่มได้ จากนั้น Active level (O_j) ของแต่ละโหนดจะถูกคำนวณโดยใช้น้ำหนักที่กำหนดไว้ในขั้นตอนก่อนหน้าผ่านฟังก์ชันของซิกมอยด์ (Sigmoid fuction) เพื่อแปลงผลต่างของผลรวมข้อมูลภายในโหนด(ที่คูณด้วยน้ำหนักของโหนดแล้ว)และค่าตั้งต้นให้มีค่าอยู่ในช่วง [0,1] จากนั้นเมื่อ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ได้ค่าจากฟังก์ชันของซิกมอยด์แล้ว ค่านั้นจะถูกกำหนดให้เป็นค่าประจำโหนด ในขั้นตอนต่อไปน้ำหนักของแต่ละโหนดจะถูกปรับเพื่อให้เกิดการเรียนรู้ขึ้น โดยเป้าหมายของการเรียนรู้คือทำให้ค่าความคาดเคลื่อน (Error) จากค่าที่เป็นจริงลดลงจนมีค่าเข้าใกล้ 0 ดังนั้นจึงมีขั้นตอนการตรวจสอบย้อนกลับ (Backpropogation) เกิดขึ้นในทุกๆ รอบของการเรียนรู้ และในแต่ละรอบของการปรับค่าน้ำหนัก สามารถกำหนดความเร็วในการเรียนรู้ (Learning rate) ได้ โดยมีค่าอยู่ในช่วง (0,1] หากความเร็วในการเรียนรู้มีค่ามากจะทำให้การปรับน้ำหนักมีความหยาบมากขึ้น แต่จะทำให้ลดค่าความคลาดเคลื่อนหลังจากเรียนรู้ต่อรอบได้เร็วขึ้น การเรียนรู้จะหยุดเมื่อถึงเกณฑ์ที่กำหนดไว้ เช่น เมื่อเรียนรู้ครบตามจำนวนรอบการเรียนรู้ (Training Cycle) ที่กำหนด หรือ ค่าความคลาดเคลื่อนลดน้อยลงจนมีค่าถึงจุดที่ยอมรับได้

ตารางที่ 2.4 แสดงการเปรียบเทียบ จุดเด่น-จุดด้อย และข้อจำกัดของเทคนิคการจัดกลุ่มข้อมูล เทคนิคจัดกลุ่มข้อมูลแบบเพื่อนบ้านใกล้ที่สุด มีจุดเด่นเรื่องความแม่นยำสูง แม้ว่ามีค่าที่ผิดปกติอยู่ในข้อมูลที่ใช้ทำนายด้วย แต่เนื่องจากเทคนิคนี้จำเป็นต้องหาระยะทางระหว่างข้อมูลที่มีทั้งหมด เพื่อนำมาเปรียบเทียบกัน ทำให้เทคนิคนี้ใช้หน่วยความนำในการประมวลผลสูงด้วย ในขณะที่เทคนิคจัดกลุ่มข้อมูลแบบนาอีฟเบย์ มีจุดเด่นคือสามารถทำนายข้อมูลได้แม้มีข้อมูลที่มีอยู่น้อย แต่เนื่องจากข้อมูลที่มีอยู่น้อยนี้ คุณภาพของข้อมูลจึงเป็นปัจจัยสำคัญที่ส่งผลต่อการทำนายด้วย เทคนิคจัดกลุ่มข้อมูลแบบต้นไม้ตัดสินใจ มีจุดเด่นคือสามารถแสดงความสัมพันธ์ของข้อมูลที่ใช้ทำนายออกมาได้ในรูปแบบของเงื่อนไข ซึ่งทำให้เห็นความสัมพันธ์ระหว่างข้อมูลที่นำมาทำนาย แต่จุดด้อยคือหาใช้ตัวแปรควบคุมการทำนายที่ไม่เหมาะสม มีโอกาสที่ทำให้โมเดลสร้างฟังค์ชันออกมาใช้งานได้เฉพาะกับข้อมูลที่นำมาสร้างโมเดล ในขณะที่เทคนิคการจัดกลุ่มแบบโครงข่ายประสาทเทียมมีจุดเด่นที่สามารถนำไปใช้แก้ปัญหาได้หลากหลาย แต่เนื่องจากการทำงานที่ซับซ้อน ทำให้สามารถวิเคราะห์ความสัมพันธ์ระหว่างข้อมูลนำเข้าและส่งออกได้ยาก เทคนิคการจัดกลุ่มทั้งสี่ที่กล่าวมาสามารถทำนายตัวแปรแบบนามบัญญัติได้ ในกรณีที่ต้องการทำนายตัวแปรที่เป็นตัวเลข สามารถใช้เทคนิคการจัดกลุ่มแบบเพื่อนบ้านที่ใกล้ ต้นไม้ตัดสินใจ และโครงข่ายประสาทเทียมได้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ตารางที่ 2.4 แสดงการเปรียบเทียบจุดเด่น-จุดด้อย
และข้อจำกัดของเทคนิคการจัดกลุ่มข้อมูลที่ศึกษาในงานวิจัย [36], [38], [42], [43]

เทคนิคจัดกลุ่มข้อมูล	จุดเด่น	จุดด้อย	การทำนายตัวแปร	
			ตัวเลข	นามบัญญัติ
เพื่อนบ้านที่ใกล้ที่สุด	ความแม่นยำสูงและไม่อ่อนไหวต่อค่าผิดปกติ	ใช้หน่วยความจำในการประมวลผลสูง	ได้	ได้
นาอีฟเบย์	สามารถใช้ได้แม้มีข้อมูลน้อย	คุณภาพของข้อมูลมีผลกระทบต่อการใช้งานจริงอย่างมาก	ไม่ได้	ได้
ต้นไม้ตัดสินใจ	แสดงความสัมพันธ์ระหว่างข้อมูลในรูปแบบของเงื่อนไข ทำให้สะดวกต่อการวิเคราะห์ข้อมูล	มีโอกาสที่โมเดลสร้างฟังค์ชันออกมาใช้งานได้เฉพาะกับข้อมูลที่นำมาสร้างโมเดล	ได้	ได้
โครงข่ายประสาทเทียม	มีความยืดหยุ่นสูง ใช้แก้ปัญหาได้หลากหลายรูปแบบ	ถูกเปรียบเทียบว่าเป็นกล่องดำ (Blackbox) คือ หากต้องการทำความเข้าใจความสัมพันธ์ระหว่างข้อมูลนำเข้าและข้อมูลส่งออกจะทำได้ยาก	ได้	ได้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

จากการศึกษางานวิจัยเกี่ยวกับการทำนายและเทคนิคการจัดกลุ่ม ค้นพบว่า มีการเลือกใช้เทคนิคการจัดกลุ่มที่คล้ายกัน อย่างไรก็ตามแต่ละงานวิจัยก็ได้ใช้เทคนิคการจัดกลุ่มเพื่อวิเคราะห์และทำนายข้อมูลที่แตกต่างกันออกไป

Pratama และ Sarno [44] ได้ทดลองนำเทคนิคการจัดกลุ่ม 3 รูปแบบ ได้แก่นาอีฟเบย์, วิธีการเพื่อนบ้านใกล้ที่สุด และ ซัพพอร์ตเวกเตอร์แมชชีน มาใช้วิเคราะห์และทำนายข้อมูลบุคลิกภาพจากข้อความที่เขียนใน Twitter โดยข้อมูลบุคลิกภาพที่ทำนายได้ยังคงเป็นบุคลิกภาพห้าองค์ประกอบ และ

ได้ค้นพบว่า นาอีฟเบย์ ได้ให้ความถูกต้องของการทำนายมากที่สุดเมื่อเปรียบเทียบกับเทคนิคการจัดกลุ่มแบบวิธีการเพื่อนบ้านใกล้ที่สุด และ ซัพพอร์ตเวกเตอร์แมชชีนที่มีความแม่นยำใกล้เคียงกัน งานวิจัยนี้เกี่ยวข้องกับการทำนายข้อมูลโดยใช้ข้อมูลบุคลิกภาพ และเปรียบเทียบว่าเทคนิคการจัดกลุ่มใดให้ความถูกต้องของการทำนายผลมากที่สุด แต่ไม่ได้ทำนายข้อมูลที่เกี่ยวกับบุคคลอไจล์

Al-radaideh [45] ได้สร้างโมเดลเพื่อใช้ในการทำนายประสิทธิผลในการทำงานของพนักงาน โดยวิเคราะห์จากเทคนิคการจัดกลุ่มแบบ ต้นไม้ตัดสินใจ และ นาอีฟเบย์ คุณลักษณะของพนักงานที่นำมาใช้ในการทำนายมีอยู่ 19 คุณลักษณะ เช่น อายุ, เพศ, สถานะการแต่งงาน, จำนวนบุตร, ข้อมูลทางการศึกษา, และ ประสบการณ์ในการทำงาน ผลลัพธ์ที่ได้พบว่า ข้อมูลอาชีพมีอิทธิพลอย่างมากในการทำนายผล งานวิจัยนี้ได้ใช้การเปรียบเทียบความถูกต้องเพื่อหาเทคนิคการจัดกลุ่มที่เหมาะสม เช่นเดียวกับงานวิจัยของ Pratama และ Sarno แต่มีการวิเคราะห์ย้อนกลับเพื่อหาว่าปัจจัยใดส่งผลต่อการทำนายผลมากที่สุดด้วย

ในบทนี้เป็นส่วนทบทวนวรรณกรรมที่รวบรวมทฤษฎีที่จำเป็นต่อการทำนายบุคคลอไจล์โดยอาศัยข้อมูลบุคลิกภาพ และงานวิจัยอื่นที่แสดงให้เห็นถึงส่วนที่สามารถพัฒนาต่อยอดได้ โดยแนวคิดที่ได้นำเสนอเพื่อต่อยอดนี้แสดงอยู่ในบทที่ 3 วิธีการดำเนินงานวิจัยที่เป็นบทถัดไป

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved