

การจำแนกบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงานด้วยวิธีการจำแนก
แบบหลายลาเบล

ปฐมภัทร คำตา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมข้อมูลขนาดใหญ่
วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์
มหาวิทยาลัยธุรกิจบัณฑิตย์

พ.ศ. 2563

**Multi-Label Classification of Employee Job Performance Prediction by DISC
Personality**

Patompat Kamtar

**A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering
Department of Big Data Engineering,
College of Innovative Technology and Engineering,
Dhurakij Pundit University**

2020

หัวข้อวิทยานิพนธ์	การจำแนกบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงาน ด้วยวิธีการจำแนกแบบหลายเวลาเบล
ชื่อผู้เขียน	ปฐมภัทร คำตา
อาจารย์ที่ปรึกษา	ผศ.ดร.ดวงใจ จิตคงชื่น
สาขาวิชา	วิศวกรรมข้อมูลขนาดใหญ่
ปีการศึกษา	2562

บทคัดย่อ

การพัฒนาของเทคโนโลยี และองค์ความรู้ใหม่ๆ ในปัจจุบัน ทำให้องค์กรธุรกิจต่างๆ เกิดการเปลี่ยนแปลงจากเมื่อก่อนเป็นอย่างมาก โดยเฉพาะการเปลี่ยนแปลงในเรื่องของบุคลากร ในองค์กร ที่หลายๆ องค์กร ได้ทำการลดขนาดของพนักงานในตำแหน่งที่ไม่จำเป็นลงไป และแทนที่ด้วยเทคโนโลยีที่สามารถสร้างประสิทธิภาพการทำงานให้ดีขึ้น หรือรวดเร็วขึ้นได้ ซึ่งจะทำให้เกิดปัญหาโดยตรงต่อตัวพนักงาน และผู้บริหารของบริษัท เพราะต้องทำงานเท่าเดิม หรือมากกว่าเดิม ด้วยจำนวนบุคลากรที่น้อยลง จนอาจจะก่อให้เกิดปัญหาต่างๆ ที่ส่งผลต่อความสำเร็จของหน่วยงานหรือองค์กรตามมาได้

ดังนั้น ในส่วนของพนักงานปัจจุบัน องค์กรเองก็ควรจะมีการปรับเปลี่ยนบุคลากรให้เหมาะสมกับงานที่ทำ รวมไปถึงพัฒนาความสามารถของพนักงานให้สูงขึ้น เพื่อที่พนักงานสามารถที่จะมีความรู้ ความรู้ความสามารถเหล่านั้นมาประยุกต์ใช้ เพื่อเพิ่มประสิทธิผลของการทำงานได้ และในส่วนของพนักงานใหม่ที่ต้องหามาเพื่อให้ตอบสนองต่อเป้าหมายขององค์กร ก็ควรจะต้องทำการสรรหาให้เหมาะสมกับตำแหน่งงานแต่ละประเภทด้วย โดยผู้ที่มีหน้าที่สรรหาพนักงานควรจะต้องทราบว่า คนประเภทใดเหมาะสมกับงานลักษณะไหน ซึ่งสามารถสังเกตได้จากบุคลิกภาพ อุปนิสัย ทักษะ และศักยภาพต่างๆ ของพนักงานที่อยู่ในองค์กร และเมื่อองค์กรสามารถระบุได้ว่า ต้องการคนประเภทใดเพื่อมาทำงานแบบไหน นอกจากจะสามารถวางแผนกำหนดแนวทางในการเสริมสร้างให้พนักงานทำงานได้อย่างเต็มศักยภาพแล้ว ก็ยังสามารถวิเคราะห์ได้ว่างานแต่ละแบบมีความต้องการใช้คุณสมบัติของพนักงานอย่างไร โดยถ้าจะให้ทำงานมีประสิทธิภาพและบรรลุตามเป้าหมายขององค์กร ก็ควรจะต้องเลือกใช้บุคลากรที่เหมาะสมกับงานนั้นๆ

Thesis Title	Multi-Label Classification of Employee Job Performance Prediction by DISC Personality
Author	Patompat Kamtar
Thesis Advisor	Asst. Prof. Dr. Duangjai Jitkongchuen
Department	Big Data Engineering
Academic Year	2019

ABSTRACT

The relationship between personality and job performance has been a frequently studied research topic in industrial-organizational psychology, the result of many research found that personality is one of the most importance key factor that affects to employee's job performance. Job performance generally relates to the positive things that people do to succeed at work, including task performance, discretionary behaviors, future-focused and improvement behaviors.

However, not every personality is suited for every job position, so it's important to recognize personality traits and pair employees with the job that fit their personalities the best. This can lead to increased productivity and job satisfaction, helping business function more efficiently.

To put the right man on the right job, understanding personality which suited for job position is crucial.

Past approaches toward personality and job performance conducted with only single-label of performance, but in fact, each employee can perform more than one different performance at the same time. Single-label study cannot apply to complicated problem.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยการให้ความช่วยเหลือแนะนำของ ผศ.ดร.ดวงใจ จิตคงชื่น ที่ได้กรุณาที่ให้คำแนะนำข้อคิดเห็นตรวจสอบ และแก้ไขร่างวิทยานิพนธ์มาโดยตลอด ผู้เขียนจึงขอกราบขอบพระคุณไว้ ณ โอกาสนี้

ผู้เขียนขอกราบขอบพระคุณอาจารย์ทุกท่านที่สละเวลาเพื่อมาเป็นกรรมการในการสอบวิทยานิพนธ์ และได้ให้คำแนะนำ แนวทางทราบเป็นประโยชน์ต่องานวิจัย ตลอดจน นางสาวกุลธิดา รอดบุญ รวมถึงเจ้าหน้าที่บัณฑิตมหาวิทยาลัยธุรกิจบัณฑิตฯ ทุกท่านที่ให้ความสะดวกด้านอำนวยความสะดวกและประสานงาน ในการทำวิทยานิพนธ์ให้ผู้เขียนตลอดมาตลอดจนค้นคว้าหาข้อมูลในการจัดทำวิทยานิพนธ์ของผู้เขียนครั้งนี้สำเร็จลุล่วงไปด้วยดี

ท้ายนี้ผู้เขียนขอโน้มรำลึกถึงอำนาจบารมีของคุณพระศรีรัตนตรัย และสิ่งศักดิ์สิทธิ์ทั้งหลายที่อยู่ในสากล โลก อันเป็นที่พึ่งให้ผู้เขียนมีสติปัญญาในการจัดทำวิทยานิพนธ์ให้สำเร็จลุล่วงไปด้วยดี ผู้เขียนขอให้เป็นกตเวทิตาแต่บิดา มารดา ครอบครัวของผู้เขียน ตลอดจนผู้เขียนหนังสือ และบทความต่าง ๆ ที่ให้ความรู้แก่ผู้เขียนจนสามารถให้วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยดี

ปฐมภัทร คำตา



สารบัญ

	หน้า
บทคัดย่อภาษาไทย	๗
บทคัดย่อภาษาอังกฤษ.....	๘
กิตติกรรมประกาศ.....	๑
สารบัญภาพ	๙
บทที่	
1. บทนำ.....	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 สมมติฐานของการวิจัย.....	2
1.4 ขอบเขตงานวิจัย.....	3
1.5 คำจำกัดความในการวิจัย.....	3
1.6 กรอบแนวคิดในการวิจัย.....	4
1.7 ประโยชน์ที่คาดว่าจะได้รับ.....	4
2. แนวคิด ทฤษฎี และผลงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ความหมาย แนวคิด และทฤษฎีเกี่ยวกับบุคลิกภาพ.....	5
2.2 แนวคิดแบบวัดบุคลิกภาพ DISC	10
2.3 แนวคิดและทฤษฎีเกี่ยวกับผลการปฏิบัติงานของพนักงาน.....	13
2.4 งานวิจัยด้านบุคลิกภาพที่ส่งผลต่อการทำงาน.....	17
2.5 วิธีการจำแนกข้อมูลแบบหลายตาเบล.....	18
2.6 การจำแนกประเภท	20
2.7 การเลือกคุณลักษณะ	24
3. ระเบียบวิธีวิจัย	26
3.1 แนวทางการวิจัย	26
3.2 เครื่องมือที่ใช้ในการวิจัย	31
4. ผลการวิจัย.....	33
4.1 ผลการทดสอบประสิทธิภาพและการอภิปรายผล.....	33
5. สรุปผลและข้อเสนอแนะ.....	41

สารบัญ (ต่อ)

5.1 สรุปผลการวิจัย.....	41
5.2 ข้อจำกัดและแนวทางการแก้ไขของงานวิจัย.....	42
บรรณานุกรม.....	43
ภาคผนวก	47
ประวัติผู้เขียน	54



สารบัญภาพ

ภาพที่	หน้า
2.1 การทำงานของ Binary Relevance.....	18
2.2 การทำงานของ Classifier Chains.....	19
2.3 การทำงานของ Label Powerset.....	19
2.4 แผนผังการวิธีการสแตกกิ่ง.....	24
3.1 แผนผังการดำเนินงาน โดยรวมของระบบ.....	26
3.2 แผนผังการสร้างชุดข้อมูลใหม่สำหรับการทดลองแบบที่ 1.....	28
3.3 แผนผังการสร้างชุดข้อมูลใหม่สำหรับการทดลองแบบที่ 2.....	29
3.4 แผนผังการสร้างชุดข้อมูลใหม่สำหรับการทดลองแบบที่ 3.....	30
4.1 ภาพเปรียบเทียบค่า Hamming Loss ของการทดสอบในแต่ละรูปแบบ.....	33
4.2 ภาพเปรียบเทียบค่า Accuracy ของการทดสอบในแต่ละรูปแบบ.....	35
4.3 ภาพเปรียบเทียบค่า F1 Score ของการทดสอบในแต่ละรูปแบบ.....	36
4.4 ภาพเปรียบเทียบค่า Precision ของการทดสอบในแต่ละรูปแบบ.....	37
4.5 ภาพเปรียบเทียบค่า Recall ของการทดสอบในแต่ละรูปแบบ.....	39

สารบัญตาราง

ภาพที่	หน้า
4.1 เปรียบเทียบค่า Hamming Loss ของการทดสอบในแต่ละรูปแบบ.....	34
4.2 เปรียบเทียบค่า Accuracy ของการทดสอบในแต่ละรูปแบบ.....	35
4.3 เปรียบเทียบค่า F1 Score ของการทดสอบในแต่ละรูปแบบ.....	38
4.4 เปรียบเทียบค่า Precision ของการทดสอบในแต่ละรูปแบบ.....	37
4.5 เปรียบเทียบค่า Recall ของการทดสอบในแต่ละรูปแบบ.....	40
4.6 เปรียบเทียบความสัมพันธ์ของแต่ละลาเบล.....	41



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การพัฒนาของเทคโนโลยี และองค์ความรู้ใหม่ๆ ในปัจจุบัน ทำให้องค์กรธุรกิจต่างๆ เกิดการเปลี่ยนแปลงจากเมื่อก่อนเป็นอย่างมาก โดยเฉพาะการเปลี่ยนแปลงในเรื่องของบุคลากรในองค์กร ที่หลายๆ องค์กร ได้ทำการลดขนาดของพนักงานในตำแหน่งที่ไม่จำเป็นลงไป และแทนที่ด้วยเทคโนโลยีที่สามารถสร้างประสิทธิภาพการปฏิบัติงานให้ดีขึ้น หรือรวดเร็วขึ้นได้ ซึ่งจะทำให้เกิดปัญหาโดยตรงต่อตัวพนักงาน และผู้บริหารของบริษัท เพราะต้องปฏิบัติงานเท่าเดิม หรือมากกว่าเดิม ด้วยจำนวนบุคลากรที่น้อยลง จนอาจจะก่อให้เกิดปัญหาต่างๆ ที่ส่งผลต่อความสำเร็จของหน่วยงานหรือองค์กรตามมาได้

ดังนั้น ในส่วนของพนักงานปัจจุบัน องค์กรเองก็ควรจะมีการปรับเปลี่ยนบุคลากรให้เหมาะสมกับงานที่ทำ รวมไปถึงพัฒนาความสามารถของพนักงานให้สูงขึ้น เพื่อที่พนักงานสามารถที่จะมีความรู้ ความรู้ความสามารถเหล่านั้นมาประยุกต์ใช้ เพื่อเพิ่มประสิทธิภาพของการปฏิบัติงานได้ และในส่วนของพนักงานใหม่ที่ต้องหามาเพื่อให้ตอบสนองต่อเป้าหมายขององค์กร ก็ควรจะต้องทำการสรรหาให้เหมาะสมกับตำแหน่งงานแต่ละประเภทด้วย โดยผู้ที่มีหน้าที่สรรหาพนักงานควรจะต้องทราบว่า คนประเภทใดเหมาะสมกับงานลักษณะไหน ซึ่งสามารถสังเกตได้จากบุคลิกภาพ อุปนิสัย ทักษะ และศักยภาพต่างๆ ของพนักงานที่อยู่ในองค์กร และเมื่อองค์กรสามารถระบุได้ว่า ต้องการคนประเภทใดเพื่อมาปฏิบัติงานแบบไหน นอกจากจะสามารถวางแผนกำหนดแนวทางในการเสริมสร้างให้พนักงานปฏิบัติงานได้อย่างเต็มศักยภาพแล้ว ก็ยังสามารถวิเคราะห์ได้ว่างานแต่ละแบบมีความต้องการใช้คุณสมบัติของพนักงานอย่างไร โดยถ้าจะให้งานมีประสิทธิภาพและบรรลุตามเป้าหมายขององค์กร ก็ควรจะต้องเลือกใช้บุคลากรที่เหมาะสมกับงานนั้นๆ

การศึกษาเรื่องบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงาน จึงมีความสำคัญอย่างยิ่งสำหรับองค์กร เพราะบุคลิกภาพเป็นลักษณะเฉพาะตัวของบุคคลทั้งที่สามารถมองเห็นได้ เช่น รูปร่าง หน้าตา กิริยาท่าทาง มารยาท การแต่งตัว การพูดจา การเคลื่อนไหว และที่ไม่สามารถมองเห็นได้โดยตรง เช่น สติปัญญา ความถนัด อารมณ์ หรือค่านิยม จึงทำให้คนแต่ละคนมีความแตกต่างกันไป

ซึ่งการปฏิบัติงานแต่ละประเภทเองก็มีความต้องการลักษณะของพนักงานที่มีความแตกต่างกันไป เช่น การมีไหวพริบปฏิภาณ การคิดวิเคราะห์ ความมีเหตุมีผล การพบปะและมีปฏิสัมพันธ์กับผู้คนทุกประเภท ความสามารถในการปรับตัวได้อย่างรวดเร็วตามสถานการณ์ ความสามารถในการยอมรับต่อการเปลี่ยนแปลงสิ่งใหม่ที่เกิดขึ้น และแรงจูงใจใฝ่สัมฤทธิ์ เป็นต้น ซึ่งบางคนอาจมีลักษณะดังกล่าวมาก แต่บางคนก็อาจจะมีน้อย ซึ่งส่งผลให้แต่ละคนมีผลการปฏิบัติงาน หรือปฏิบัติหน้าที่แต่ละแบบแตกต่างกันไป

และแม้จะมีงานวิจัยหลายงานในเรื่องบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงานในบริษัท แต่ส่วนใหญ่ก็จะเป็นการหาความแตกต่างระหว่างกลุ่มประชากร หรือหาความสัมพันธ์ (Correlation) ระหว่างตัวแปรเท่านั้น และบางงานวิจัยก็ได้ทำการศึกษาตัวแปรในเรื่องผลของการปฏิบัติงานเพียงแค่ตัวแปรเดียว ซึ่งอาจจะทำให้เกิดข้อจำกัด รวมถึงอคติต่างๆ จากการกำหนดตัวแปรขึ้นได้ ยกตัวอย่างเช่น พนักงานขายในธุรกิจการเงินการธนาคาร หากทำการชี้วัดเพียงแค่ยอดขายอย่างเดียวอาจจะไม่ได้บ่งบอกว่าพนักงานขายคนนั้นเป็นพนักงานขายที่ดีหรือไม่ เนื่องจากองค์กรยังคาดหวังผลการปฏิบัติงานในเรื่องอื่นๆ อีก ไม่ว่าจะเป็น การให้บริการที่ดีเยี่ยม การสร้างความสัมพันธ์ที่ดีต่อลูกค้า การป้องกันและบริหารความเสี่ยงต่างๆ (เพื่อไม่ให้เกิดการทุจริต หรือเกิดหนี้เสียขึ้น) เป็นต้น

ดังนั้นผู้วิจัยจึงมีความสนใจศึกษาในเรื่องของการจำแนกบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงานด้วยวิธีการจำแนกแบบหลายลาเบล

1.2 วัตถุประสงค์ของงานวิจัย

- 1.2.1 เพื่อศึกษาการจัดแบ่งประเภทบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานในด้านต่างๆ
- 1.2.2 เพื่อศึกษาวิธีการจำแนกแบบหลายลาเบล (Multi-label Classification) ของบุคลิกภาพต่อผลการปฏิบัติงาน

1.3 สมมุติฐานการวิจัย

- 1.3.1 วิธีการจำแนกแบบหลายลาเบลแบบ Classifier Chains และ Label Powerset ได้ผลลัพธ์ที่ดีกว่าวิธีการที่เป็นตัวชี้วัดอย่าง Binary Relevance
- 1.3.2 วิธีการสแตคกิ้ง (Stacking) สามารถสร้างแบบจำลอง (Model) ที่มีประสิทธิภาพกว่าการใช้ตัวแปรธรรมดาได้
- 1.3.3 การใช้วิธีการคัดเลือกคุณลักษณะ (Feature Selection) ช่วยทำให้แบบจำลองมีประสิทธิภาพเพิ่มมากขึ้น

1.4 ขอบเขตงานวิจัย

1.4.1 งานวิจัยนี้ใช้ข้อมูลจากการเก็บแบบสอบถามจากพนักงาน และฐานข้อมูลของบริษัทการเงินแห่งหนึ่งในประเทศไทย

1.4.2 งานวิจัยนี้เลือกใช้ค่าของ Binary Relevance เป็นตัวเปรียบเทียบพื้นฐาน (Baseline)

1.4.3 ผลการปฏิบัติงานอ้างอิงตามข้อมูลผลการปฏิบัติงานที่ได้จัดเก็บและมีการให้คะแนนของบริษัทที่ทำการศึกษาเท่านั้น

1.4.4 งานวิจัยนี้ใช้อัลกอริทึมสำหรับสร้างแบบจำลองการทำนายทั้งหมด 6 วิธี ดังนี้

- ต้นไม้การตัดสินใจ (Decision Tree)
- การสุ่มของป่า (Random Forest)
- การหาเพื่อนบ้านใกล้ที่สุด K ตัว (K-Nearest Neighbor)
- ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)
- ตัวแยกแยะเบย์อย่างง่าย (Naïve Bayes)
- การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression)

1.5 คำจำกัดความในการวิจัย

1.5.1 บุคลิกภาพ หมายถึง ลักษณะเฉพาะของบุคคลทั้งภายในและภายนอกที่เป็นตัวกำหนดพฤติกรรมการแสดงออก ตามแนวของ DISC

1.5.2 ตัวจัดแบ่งประเภทหรือวิธีการจัดแบ่งประเภท (Classifier) เป็นอัลกอริทึมของการเรียนรู้ด้วยเครื่องจักร สามารถนำมาใช้จัดแบ่งประเภทได้ เช่น ต้นไม้การตัดสินใจ, ซัพพอร์ตเวกเตอร์แมชชีน และตัวแยกแยะเบย์อย่างง่าย เป็นต้น

1.5.3 แบบจำลองหรือตัวแบบ (Model) เป็นแบบจำลองที่สร้างขึ้นเพื่อนำไปใช้ในการจัดแบ่งประเภท

1.5.4 การจำแนกแบบหลายลาเบล (Multi-label Classification) คือ แต่ละบุคลิกภาพสามารถมีค่าผลการปฏิบัติงานอยู่ในหลายลาเบล

1.5.5 การจัดแบ่งประเภทแบบไบนารี (Binary Classification) หรือการจัดแบ่งประเภทแบบครึ่งละสองกลุ่ม คือการจัดแบ่งประเภท 1 ครั้ง สามารถจัดแบ่งออกได้เป็น 2 กลุ่ม ผลการปฏิบัติงานของพนักงาน หมายถึง ผลการปฏิบัติงานที่ถูกประเมินออกมาเป็นตัวเลข โดยมีการชี้วัดทั้งหมด 4 ประเภท คือ ยอดขายสินค้า A, ยอดขายสินค้า B, คะแนนการให้บริการ และคะแนนยอดหนี้เสีย

1.6 กรอบแนวคิดในการวิจัย

การศึกษานี้จะทำการศึกษานุคลิกภาพที่ถูกวัดโดยแบบทดสอบบุคลิกภาพ DISC ที่ส่งผลต่อการปฏิบัติงาน โดยผลการปฏิบัติงานนั้นจะถูกนำมาใช้เพียง 4 ด้าน (ซึ่งเป็นข้อมูลที่บริษัทได้มีการเก็บเอาไว้) ดังนั้นการนำไปประยุกต์ใช้ในบริบทอื่น หรือใช้ในบริษัทอื่น ก็จำเป็นจะต้องนำไปปรับใช้อย่างระมัดระวัง

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 สามารถพัฒนาอัลกอริทึมในการจัดแบ่งข้อมูลบุคลิกภาพที่เหมาะสมสำหรับพนักงานขายของบริษัทได้

1.7.2 รู้จักใช้เครื่องมือในการจัดแบ่งประเภทของบุคลิกภาพ



บทที่ 2

แนวคิด ทฤษฎี และผลงานวิจัยที่เกี่ยวข้อง

ในการศึกษาวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาแนวคิด ทฤษฎี และผลงานวิจัยที่เกี่ยวข้องและใกล้เคียงเพื่อนำมาประมวลเป็นองค์ความรู้พื้นฐานและใช้ในการกำหนดแนวทางในการดำเนินการศึกษาวิจัย ดังนี้

- 2.1 ความหมาย แนวคิดและทฤษฎีเกี่ยวกับบุคลิกภาพ
- 2.2 แนวคิดแบบวัดบุคลิกภาพ DISC
- 2.3 แนวคิดและทฤษฎีเกี่ยวกับผลการปฏิบัติงานของพนักงาน
- 2.4 งานวิจัยด้านบุคลิกภาพที่ส่งผลต่อการทำงาน
- 2.5 แนวคิดการจำแนกแบบหลายตาเบล (Multi-label Classification)
- 2.6 แนวคิดการจำแนกประเภท (Classification)
- 2.7 แนวคิดการเลือกคุณลักษณะ (Feature Selection)

2.1 ความหมาย แนวคิด และทฤษฎีเกี่ยวกับบุคลิกภาพ

2.1.1 ความหมายของบุคลิกภาพ

Guilford (2010) กล่าวว่า บุคลิกภาพนั้นเป็นลักษณะเฉพาะเจาะจงของแต่ละบุคคล ซึ่งประกอบไปด้วย รูปร่างภายนอก ความถนัด สภาพทางอารมณ์ ทักษะสติ ความสนใจ ความต้องการ และร่างกาย

Allport (2009) ได้ให้ความหมายที่เกี่ยวข้องกับบุคลิกภาพเอาไว้ว่า เป็นหน่วยรวมของระบบทางกายและจิตใจ ซึ่งกำหนดลักษณะการปรับตัวให้เข้ากับสิ่งแวดล้อมภายนอกได้

Hilgard (2009) กล่าวว่า บุคลิกภาพเป็นลักษณะส่วนรวมของแต่ละบุคคล และจะเป็นแนวทางในการปรับตัวให้เข้ากับสิ่งแวดล้อมภายนอก โดยแต่ละคนจะมีรูปแบบของการแสดงออกที่แตกต่างกันไป

Bernard (2009) กล่าวว่า บุคลิกภาพเป็นผลรวมทั้งหมดของ ท่าทาง รูปร่าง ลักษณะทางกาย พฤติกรรมที่แสดงออก แนวโน้มการกระทำ รวมไปถึงขอบเขตของความสามารถ

Robbin (2003) กล่าวถึงบุคลิกภาพเอาไว้ว่า เป็นส่วนผสมขององค์ประกอบต่างๆ ภายในตัวบุคคล เช่น การรับรู้ แรงจูงใจ การเรียนรู้ ทักษะ และอารมณ์ เป็นต้น โดยสิ่งต่างๆ เหล่านี้จะสะท้อนออกมาให้เห็นในลักษณะภายนอกที่ปรากฏ เช่น รูปร่างหน้าตาที่ดูเฉยเมย การแต่งกายที่สุภาพเรียบร้อย และพฤติกรรมต่างๆ ที่แสดงออก รวมไปถึงบุคลิกภาพบางส่วนที่จะเป็นการตระหนัก หรือเข้าใจในตัวเอง

จากความหมายของคำว่าบุคลิกภาพที่ได้กล่าวไปข้างต้น สามารถสรุปได้ว่าบุคลิกภาพ หมายถึง คุณลักษณะทางกาย และพฤติกรรมที่แสดงออกต่อบุคคลอื่น ทั้งลักษณะกริยาท่าทาง การแสดงออกทางอารมณ์ ความคิด เจตคติ ตลอดจนพฤติกรรม และปฏิกิริยาตอบสนองทั้งทางกายและทางจิตใจ โดยบุคลิกภาพของแต่ละบุคคลนั้นมีลักษณะที่ค่อนข้างคงตัว หากใครมีลักษณะบุคลิกภาพที่เป็นอย่างไร ก็มีแนวโน้มที่จะเป็นแบบนั้นตลอดไป

2.1.2 แนวคิดและทฤษฎีบุคลิกภาพ

ทฤษฎีบุคลิกภาพ เป็นแนวทางที่นักจิตวิทยาใช้ในการอธิบายถึงธรรมชาติของบุคลิกภาพที่เกี่ยวข้องกับโครงสร้าง กระบวนการ และสาระสำคัญ เพื่อให้เข้าใจถึงความแตกต่างของบุคลิกภาพที่เกิดขึ้นในแต่ละบุคคล โดยมีทฤษฎีดังนี้

1) ทฤษฎีจิตวิเคราะห์ (Psychoanalysis Theory) ซิกมันด์ ฟรอยด์ จิตแพทย์ชาวออสเตรีย เป็นผู้ก่อตั้งทฤษฎีจิตวิเคราะห์นี้ขึ้น โดยอาศัยข้อมูลและประสบการณ์จากการรักษาผู้ป่วยในคลินิกของเขา โดยฟรอยด์ได้ให้ความสนใจในเรื่องของบุคลิกภาพการแสดงออกของแต่ละคนเป็นอย่างมาก และได้อธิบายทฤษฎีบุคลิกภาพของบุคคล (เดมส์คัต คทวนิช, 2547) ไว้ดังนี้

ทฤษฎีโครงสร้างบุคลิกภาพ (Structure of Personality) ฟรอยด์ พบว่าโครงสร้างบุคลิกภาพของบุคคลเกิดจากความขัดแย้งกันระหว่างพลังทางจิต 3 ส่วน ได้แก่ อิด (Id) อีโก้ (Ego) และซูเปอร์อีโก้ (Super Ego) โดยพลังทั้ง 3 ส่วนนี้จะอยู่ในจิตทั้ง 3 ระดับ

(1) อิด เป็นพลังงานทางจิตที่ซ่อนอยู่ภายในจิตใต้สำนึกเป็นส่วนใหญ่ โดยคำว่าพลังงานทางจิต หมายถึง ความอยาก ความต้องการ กิเลส และตัณหาทั้งหลาย ซึ่งเมื่อเกิดขึ้นแล้วจะพยายามหาทางออกโดยไม่สนใจว่าจะเป็นไปได้หรือไม่ในโลกของความเป็นจริง โดยฟรอยด์กล่าวว่าอิดของบุคคลจะเกิดจากสัญชาตญาณ 2 ประเภท ได้แก่

(1.1) สัญชาตญาณแห่งการดำรงชีวิตอยู่ (Life Instinct) ซึ่งเป็นสัญชาตญาณที่จะกระตุ้นให้บุคคลแสวงหาสิ่งที่ทำให้เกิดความสุข ความสบาย และความพึงพอใจแก่ตนเอง ซึ่งในบรรดาสัญชาตญาณแห่งการมีชีวิตอยู่นั้น ฟรอยด์ได้ให้ความสำคัญกับความต้องการทางเพศ (Sexual) มากที่สุด แต่ความต้องการทางเพศในทัศนะของฟรอยด์นั้นไม่ได้หมายถึงความต้องการความสุขจาก

การมีเพศสัมพันธ์เท่านั้น ยังครอบคลุมถึงความต้องการความสุข ความพึงพอใจ หรือความ สะดวกสบายทุกอย่าง

(1.2) สัญชาตญาณแห่งความตาย (Death Instinct) เป็นสัญชาตญาณที่กระตุ้นให้บุคคล เกิดการเอาชนะ ต่อสู้ ทำลาย ซึ่งฟรอยด์ได้ให้ความสำคัญกับความก้าวร้าว (Aggression) มากที่สุด ตัวอย่างความก้าวร้าว ได้แก่ ความต้องการที่จะได้รับการคัดเลือกขึ้นเป็นหัวหน้างาน จึงพยายาม ขยันทำทุกอย่างเพื่อให้ผู้บังคับบัญชาเลือกตน แม้กระทั่งการทะเลาะเบาะแว้ง ชกต่อย ทำร้ายร่างกาย รวมกระทั่งทำสงครามระหว่างกัน

แต่ในระหว่างความต้องการทางเพศกับความก้าวร้าว สัญชาตญาณที่มีอิทธิพลต่อมนุษย์ มาก คือ ความต้องการทางเพศ ด้วยเหตุผลนี้ฟรอยด์จึงอธิบายว่าบุคลิกภาพของมนุษย์ที่แสดง ออกมานั้น จะตอบสนองความต้องการทางเพศของตนเป็นส่วนใหญ่ สำหรับบุคคลที่มีบุคลิกภาพ แบบอิดนั้นมักจะแสดงอะไรตามสัญชาตญาณของตนโดยไม่สนใจอะไรทั้งสิ้น เพียงเพื่อให้ตนเอง ได้รับความสุขและความพอใจเพียงอย่างเดียว

(2) อีโก้ เป็นพลังงานที่จะอยู่ในจิตสำนึกและกึ่งจิตสำนึกเป็นส่วนใหญ่ เป็นพลังงาน ทางจิต ที่จะทำหน้าที่บริหารพลังจากฝ่ายอิด และฝ่ายซูเปอร์อีโก้ให้สมดุล และแสดงออกให้ สอดคล้องกับโลกแห่งความเป็นจริงที่สังคมยอมรับ หรือเหมาะสมกับเหตุผลในสถานการณ์นั้น แต่ ก็ต้องขึ้นอยู่กับความแข็งแกร่งหรือความอ่อนแอของอีโก้ในแต่ละบุคคลด้วย

สำหรับบุคคลที่มีบุคลิกภาพแบบอีโก้ (Ego Personality) นั้นมักจะแสดงออกมาตาม เหตุผลความเป็นจริงที่ตนพิจารณาแล้วว่าเหมาะสม และถูกต้องในโลกของความเป็นจริง

(3) ซูเปอร์อีโก้ เป็นพลังงานที่อยู่ภายในจิตสำนึกเป็นส่วนใหญ่ ที่รวมตัวขึ้นจากการ เรียนรู้ในระเบียบ กฎเกณฑ์ กติกา กฎของศีลธรรม และกฎหมายของสังคม ซูเปอร์อีโก้เป็นตัวบอก ให้รู้ว่าอะไรดี อะไรชั่ว อะไรถูก อะไรผิด ควรหรือไม่ควร จึงมีลักษณะตรงข้ามกับอิด โดยจะมี หน้าที่คอยควบคุมความต้องการทางเพศและความก้าวร้าวในอิดไม่ให้แสดงออก ซูเปอร์อีโก้จะ ควบคุมอิดได้ดีแค่ไหนนั้นก็ขึ้นอยู่กับว่าซูเปอร์อีโก้ของบุคคลนั้นแข็งแกร่งเพียงใด

สำหรับบุคคลที่มีบุคลิกภาพแบบซูเปอร์อีโก้ (Superego Personality) มักจะชอบทำตาม ระเบียบ กฎเกณฑ์ กติกา และกฎหมายอย่างเคร่งครัด ยึดมั่นในทฤษฎีและมีอุดมคติสูง

การที่บุคคลแสดงบุคลิกภาพในรูปแบบใดออกมานั้นย่อมขึ้นอยู่กับความขัดแย้งกัน ระหว่างพลังทางจิตทั้งสามส่วนนี้ว่าพลังงานทางจิตส่วนใดจะมีอำนาจเหนือกว่า บุคคลก็จะแสดง บุคลิกภาพออกมาตามอิทธิพลของพลังงานทางจิตฝ่ายที่มีอำนาจนั้น แต่ถ้าเมื่อใดที่พลังงานระหว่าง อิดกับซูเปอร์อีโก้มีความขัดแย้งกันอย่างรุนแรงมากเกินไป บางครั้งอีโก้จะหาทางประนีประนอม

เพื่อลดความขัดแย้งนั้นให้น้อยลง โดยใช้วิธีการปรับตัวที่เรียกว่า กลไกในการป้องกันตัวเอง (Defense Mechanisms) ถ้าทำสำเร็จจะช่วยให้บุคคลนั้นกลายเป็นโรคจิตและโรคประสาทในที่สุด

2) ทฤษฎีจิตวิทยาวิเคราะห์ (Analytical Psychology Theory) ผู้ก่อตั้งทฤษฎีนี้ได้แก่ คาร์ล กุสตาฟ จุง (Carl Gustav Jung) นักจิตวิทยาชาวสวิสเซอร์แลนด์ ได้เสนอแนวความคิดทฤษฎีบุคลิกภาพที่จำแนกออกได้เป็นส่วนสำคัญ 2 ส่วน (เดิมศักดิ์ คทาวณิช, 2547) ดังนี้

(1) โครงสร้างบุคลิกภาพ (Structure of Personality) บุคลิกภาพตามความหมายของจุงคือจิต (Psyche) ซึ่งประกอบด้วยระบบต่างๆ เป็นส่วนย่อยมาทำงานรวมกัน ได้แก่

(1.1) อีโก้ จุงเชื่อว่าอีโก้เป็นศูนย์กลางของบุคลิกภาพของบุคคลซึ่งอยู่ในส่วนของจิตสำนึก (conscious) ซึ่งประกอบไปด้วยความจำ ความรู้สึกลึกซึ้ง การตัดสินใจ และการมีสติ ซึ่งเป็นส่วนสำคัญที่บุคคลจะสามารถรับรู้เกี่ยวกับตนเองได้ตลอดเวลา จึงเท่ากับว่าอีโก้เป็นตัวกำหนดบทบาท หน้าที่ และความเป็นเอกลักษณ์เฉพาะของแต่ละบุคคล

(1.2) จิตใต้สำนึกส่วนบุคคล (Personal Unconscious) ส่วนนี้จะอยู่ถัดจากอีโก้ลงไปเป็นส่วนที่ประกอบไปด้วยประสบการณ์ต่างๆ ที่เคยอยู่ในจิตสำนึกมาก่อนแต่ได้ถูกกดลงสู่จิตใต้สำนึก (Unconscious) ด้วยกลไกทางจิต ทั้งนี้เนื่องจากความต้องการที่จะลืมประสบการณ์เหล่านั้น เพราะเป็นความเจ็บปวด เป็นทุกข์ หรือไม่พอใจเป็นต้น ต่อมาภายใต้สถานการณ์เหล่านั้น สิ่งแวดล้อมหรือได้รับสิ่งเร้าที่เหมาะสม ประสบการณ์เหล่านั้นอาจจะผลักดันขึ้นมาสู่จิตสำนึกที่รับรู้ได้อีกครั้ง ประสบการณ์ต่างๆ ภายในจิตใต้สำนึกส่วนบุคคลนี้ถ้าได้รับการรวบรวมให้เป็นกลุ่มหรือหมวดหมู่ของความรู้สึก (Constellation) ที่เป็นอันหนึ่งอันเดียวกันแล้ว จุงเรียกการเกิดสภาวะเช่นนั้นว่า ปม (Complex) ดังนั้นเท่ากับว่าจิตใต้สำนึกส่วนบุคคลจึงเป็นแหล่งรวบรวมปมของบุคคลไว้มากมาย เช่น ปมเกี่ยวกับแม่ (Mother Complex) เกิดจากการจัดกลุ่มหรือหมวดหมู่ของประสบการณ์ต่างๆ เกี่ยวกับแม่ เช่น ความรู้สึก ความจำต่างๆ ที่ได้รับจากแม่จนก่อตัวขึ้นเป็นปมเมื่อพลังจากปมนี้มีมากจะกลายเป็นจุดศูนย์กลางในการควบคุมบุคลิกภาพของบุคคลนั้นให้ทำตามสิ่งที่แม่พูด แม่สั่ง แม่คิด หรือสิ่งที่เป็นความประสงค์ของแม่ แม้กระทั่งการเลือกภรรยาที่จะเลือกบุคคลซึ่งมีลักษณะคล้ายคลึงกับแม่ของตน เป็นต้น อย่างไรก็ตาม ประสบการณ์ต่างๆ ที่รวมกลุ่มหรือหมวดหมู่ขึ้นเป็นปมนั้นอาจกลับขึ้นมาสู่จิตสำนึกได้อีกครั้งถ้าอยู่ในสถานะการณ์ที่เหมาะสม

(1.3) จิตใต้สำนึกส่วนที่สะสมประสบการณ์ในอติชชาติ (Collective Unconscious) จุงอธิบายว่าจิตใต้สำนึกส่วนนี้จะทำหน้าที่สะสมประสบการณ์ต่างๆ ที่ทุกคนได้รับเป็นมรดกสืบทอดมาจากบรรพบุรุษ ตั้งแต่เริ่มต้นมีมนุษย์เกิดขึ้นภายในโลกเป็นครั้งแรก ด้วยเหตุนี้จึงเชื่อว่ามนุษย์ทุกคน ทุกตระกูล ทุกเชื้อชาติ และทุกเผ่าพันธุ์ต่างก็มีประสบการณ์ในจิตใต้สำนึกส่วนที่สะสม

ประสบการณ์ในอดีตชาติที่เป็นต้นฉบับเดียวกันทั้งสิ้น โดยบันทึกเป็นข้อมูลอยู่ในสมองแล้ว ถ่ายทอดกันมาแต่ละรุ่นยาวนานตลอดจนถึงปัจจุบัน ด้วยเหตุนี้ในการแสดงพฤติกรรมบางอย่างของมนุษย์ทั่วโลก จึงอธิบายจากต้นฉบับที่ถูกบันทึกไว้ตั้งแต่อดีตชาติได้เหมือนกันหมด เช่น ทำไมเด็กทารกจึงเกิดมาด้วยความพร้อมที่จะรับรู้และสร้างสัมพันธภาพอันดีกับแม่เป็นอันดับแรก ถ้าอธิบายตามแนวคิดของจุงนั้นเป็นเพราะว่าเด็กได้รับประสบการณ์ความสัมพันธ์กับแม่มาตั้งแต่อดีตชาติภายใต้จิตใต้สำนึกส่วนที่สะสมประสบการณ์ในอดีตชาติ จึงทำให้มีพื้นฐานเช่นนี้มาตลอดจนถึงปัจจุบัน

(1.4) หน้ากาก (Persona) หมายถึง สภาวะของบุคคลที่จะต้องแสดงบทบาทไปตามความคาดหวังของสังคมและเป็นไปตามขนบธรรมเนียมประเพณีตามที่สังคมกำหนด หรือเป็นการแสดงออกเพื่อให้ได้รับการยอมรับและสร้างความประทับใจบุคคลอื่นๆ ดังนั้นในบางครั้งบุคลิกภาพของบุคคลที่เกิดจากการใช้หน้ากาก จึงอาจจะมี ความขัดแย้งกับบุคลิกภาพที่แท้จริงภายในตัวบุคคลนั้นได้ เท่ากับว่าหน้ากากจึงทำหน้าที่ควบคุมบุคลิกภาพส่วนที่ไม่ดีที่แท้จริงของบุคคลไม่ให้ปรากฏออกมาต่อสังคมภายนอก ความขัดแย้งระหว่างบุคลิกภาพเหล่านี้ถ้าเกิดบ่อยครั้งในหลายๆ เรื่องอาจจะทำให้บุคคลนั้นขาดความเป็นตัวของตัวเอง กลายเป็นคนที่สวมหน้ากากเข้าหาผู้อื่น หรือเป็นคนที่มีความขัดแย้งภายในจิตใจได้

(1.5) ลักษณะซ่อนเร้น (Anima or Animus) จุงเชื่อว่ามนุษย์มีลักษณะทั้งสองเพศอยู่ในคนๆ เดียวกัน โดยจะเห็นได้จากการที่เพศชายจะมีความนุ่มนวลและอ่อนโยน ซึ่งเป็นลักษณะของเพศหญิงอยู่ในตัว จุงเรียกลักษณะเช่นนี้ว่า แอนิมา (Anima) ส่วนผู้หญิงจะมีความเข้มแข็ง กล้าหาญ และเด็ดขาด ซึ่งเป็นลักษณะของเพศชายซ่อนเร้นอยู่ในตัวเช่นกัน จุงเรียกลักษณะเช่นนี้ว่า แอนิมัส (Animus) จากลักษณะทั้งสองเพศที่ซ่อนเร้นอยู่นี้จึงทำให้ผู้ชายเข้าใจธรรมชาติของผู้หญิง และผู้หญิงก็มีความเข้าใจธรรมชาติของผู้ชายได้ด้วยตัวของตัวเอง

(1.6) เงาแฝง (Shadow) เป็นภาพที่ก่อตัวมาจากสัตว์ก่อนที่จะมีวิวัฒนาการมาเป็นมนุษย์ ซึ่งกล่าวได้ว่า เงาแฝงเป็นสัญชาตญาณของสัตว์ที่จะส่งผลให้มนุษย์แสดงความชั่วร้าย ก้าวร้าว และป่าเถื่อน รวมทั้งพฤติกรรมต่างๆ ที่ขัดแย้งกับกฎหมาย ขนบธรรมเนียมประเพณีอันดีงามของสังคม โดยเงาแฝงเหล่านี้จะถูกควบคุม และปกปิดโดยหน้ากาก หรือเก็บกดไว้ในจิตใต้สำนึก

(2) ลักษณะบุคลิกภาพของแต่ละบุคคล จากโครงสร้างทางบุคลิกภาพที่จุงได้อธิบายไว้ จุงจึงแบ่งลักษณะบุคลิกภาพของบุคคลไว้ 2 ประเภทคือ

(2.1) แบบเก็บตัว (Introvert) เป็นบุคลิกภาพของบุคคลที่มีแนวโน้มเป็นพวกเก็บตัว ชอบความสงบเงียบไม่ชอบการเข้าสังคม ซื่อาย พอใจที่จะอยู่เบื้องหลัง ขาดความมั่นใจในตนเอง

ชอบใช้วิธีหนีปัญหามากกว่าเผชิญปัญหา ส่วนที่เป็นข้อดีของบุคลิกภาพแบบเก็บตัวนี้ มักจะเป็นบุคลิกของ นักประดิษฐ์และนักคิดค้นทั้งหลาย แต่ส่วนที่เป็นข้อเสียคือ มักจะเกิดอาการซึมเศร้า แยกตัว และไม่สนใจสังคม

(2.2) แบบแสดงตัว (Extrovert) เป็นบุคลิกภาพประเภทแสดงตัว ชอบเข้าสังคมรักความสนุกสนาน ชอบการเปลี่ยนแปลงไปตามสถานการณ์ที่เป็นจริง มีมนุษยสัมพันธ์ดี กล้าที่จะแสดงออก ชอบความเป็นผู้นำ ต้องการเป็นที่รู้จักของคนทั่วไป คบคนง่าย ชอบเผชิญปัญหามากกว่าการหนีปัญหา

อย่างไรก็ตาม ในความเป็นจริงจะพบว่าคนบางคนจะมีบุคลิกภาพแบบกลางๆ กล่าวคือ มักจะมีบุคลิกภาพเป็นไปตามสถานการณ์ เช่น ในสถานการณ์หนึ่งอาจมีบุคลิกภาพแบบแสดงตัว ในอีกสถานการณ์หนึ่งอาจเป็นแบบเก็บตัวก็ได้ จึงจัดคนประเภทนี้อยู่ในพวกแอมบิเวิร์ต

2.2 แนวคิดแบบวัดบุคลิกภาพ DISC

คาร์ล กุสตาฟ จุง (Carl Gustav Jung) ได้ทำการแบ่งรูปแบบพฤติกรรมมนุษย์ออกเป็น 4 รูปแบบ ได้แก่ D I S C (Mary Kay Slowikowski, 2005)

D - style (dominance) มีลักษณะเด่นคือ มีความกล้าตัดสินใจ เน้นผลลัพธ์มุ่งมั่น ทำอะไรรวดเร็ว พุดจาเสียงดัง ไม่ประนีประนอม กล้าเสี่ยง ชอบการแข่งขัน ยึดตนเองเป็นศูนย์กลาง ไม่ค่อยฟังผู้อื่น ตรงไปตรงมา บางทีอาจจะดูเป็นคนขวานผ่าซากไปบ้าง มีข้อเรียกร้องสูง ดื้อดึง มั่นใจในตัวเอง

I - style (influence) จะเป็นคนช่างพูด ชอบเข้าสังคม ดูเป็นมิตร เปิดเผย จุงใจคนเก่ง ชอบขายไอเดีย เป็นคนร่าเริง มีอารมณ์อ่อนไหว มักหลีกเลี่ยงรายละเอียด เวลาคุยกันมักพูดออกนอกประเด็นไปเรื่อย ไม่ค่อยรักษากฎ ไม่ชอบคนน่าเบื่อ หรือเข้มงวด

S - style (stediness) คนที่มีด้านนี้เด่นจะเป็นคนใจเย็น สงบ เสมอต้นเสมอปลาย ทำอะไรเป็นระบบ ละเอียดรอบคอบ ถ่อมตัว ประนีประนอม ชอบเป็นผู้ฟังมากกว่าผู้พูด จะพูดเมื่อถูกถาม มักแสดงความเห็นด้วยกับผู้อื่นตลอดเวลา ต้องการคำแนะนำจากผู้อื่น ไม่ชอบการเปลี่ยนแปลง

C - style (compliance) จะชอบความสมบูรณ์แบบ ยึดติดกับรายละเอียด มีเหตุผล มีหลักการ เน้นความชัดเจนถูกต้อง ให้ความสำคัญกับข้อเท็จจริง เป็นคนละเอียดถี่ถ้วน มีระเบียบ ปฏิบัติตามกฎเกณฑ์ ไม่ชอบเสี่ยง ทำอะไรอย่างระมัดระวัง

โดยบางคนอาจเป็นประเภทลักษณะเด่น โดดๆ ตัวเดียว เช่น D แต่บางคนอาจเป็นลักษณะผสม โดยมีลักษณะเด่นนำ 1 ตัว และตาม 1 ถึง 2 ตัว เช่น DC หรือ DCI เป็นต้น

จากแนวคิดของคาร์ล กุสตาฟ จุง ต่อมา วิลเลียม มาร์สตัน (William Moulton Marston) สนใจศึกษาเรื่องความแตกต่างทางพฤติกรรมของมนุษย์และการตอบสนองพฤติกรรมของแต่ละบุคคลในแต่ละสถานการณ์ที่เปลี่ยนไป โดยใช้หลัก การรับรู้ (Perception) ของมนุษย์ที่มีต่อตนเอง และต่อสิ่งแวดล้อม เป็นแนวทางอธิบายพฤติกรรมตอบสนองของบุคคล มิติการรับรู้ 2 แกนนี้ มีนัยต่อพฤติกรรมแสดงออก รูปแบบความรู้ตื้นึกคิด ลักษณะการตอบสนองสิ่งแวดล้อมและผู้อื่น ที่สั่งสมมาจากประสบการณ์ การเรียนรู้ การถูกป้อนข้อมูลตั้งแต่วัยเด็ก

แกนตั้ง : เป็นการรับรู้เชิงเปรียบเทียบพลังของตนเองที่มี มากกว่าหรือน้อยกว่าต่อสภาพแวดล้อมภายนอก บอกถึงระดับพลังงานในตัว ระดับความมั่นใจในการเปิดเผยตัวตน แสดงความคิดเห็น ระดับความกล้าเผชิญ เช่น ด้านพลังสูง มักกล้าใช้พลังความสามารถ โน้มน้าว ควบคุม บังคับ กล้าแตกต่าง จนถึงทำทนายการเปลี่ยนแปลง ประเมินความสามารถตนเองสูง ส่วนอีกด้านตรงข้าม ที่ประสบการณ์สั่งสมทำให้บุคคลระมัดระวังสิ่งที่แวดล้อม จึงมักคอยตั้งรับ เก็บตัว เก็บอาการ เช่น ชอบอยู่ในกลุ่ม ทำตามกติกากลุ่มหรือพวกพ้อง ค่อยตามเสียงส่วนใหญ่ อึดอัดใจกับการเปลี่ยนแปลงฉับพลัน ประเมินความสามารถตนเองต่ำ เป็นต้น

แกนนอน : เป็นการรับรู้ความพึงพอใจที่มีต่อสภาพแวดล้อมและ ผู้คน บอกถึงระดับความไว้วางใจ เชื่อใจ มุมมองต่อโลก การตั้งข้อสังเกต สงสัย เช่น ด้านพึงพอใจ ก็จะตอบสนองต่อสิ่งแวดล้อมและผู้คนอย่างเป็นมิตร ไว้วางใจและเชื่อใจได้ง่ายกว่า มองโลกในแง่ดี ไม่ค่อยระแวงสงสัย ส่วนอีกด้านตรงข้าม จะตอบสนองต่อสิ่งแวดล้อมและผู้อื่น อย่างระแวงระวัง ตั้งคำถาม สงสัย เพิ่มความรอบคอบและวางแผนรัดกุม

จากการรับรู้ใน 2 แกนนี้ สามารถจำแนกคุณลักษณะพฤติกรรมมนุษย์นี้ออกเป็น 4 หมวดคุณลักษณะ คือ Dominance, Inducement (ต่อมาเปลี่ยนเป็น Influence), Submission (ต่อมาเปลี่ยนเป็น Steadiness) และ Compliance (ต่อมาเปลี่ยนเป็น Conscientious) จึงเป็นที่มาของคำว่า DISC ซึ่งมาจากอักษรย่อของ 4 คุณลักษณะนี้

Dominance ใช้แรงขับ ความพยายามเพื่อเข้าควบคุม เข้าพิชิต มุ่งผลลัพธ์และรักษาอำนาจ

Influence ใช้แรงขับ ความพยายามเพื่อสร้างอิทธิพลในหมู่น คน โดยการโน้มน้าว ชักจูง มุ่ง แสดงออก และการได้รับความสนใจ

Steadiness ใช้แรงขับ ความพยายามเพื่อสร้างความมั่นคง สม่่าเสมอ มุ่งเน้นสายสัมพันธ์ โดยเข้าให้การสนับสนุน ช่วยเหลือ

Conscientious ใช้แรงขับ ความพยายามเพื่อความถูกต้อง แน่นอน ปลอดภัย มุ่งสาระ
กระบวนการ อย่างมีตรรกะและขั้นตอน

อย่างไรก็ตาม รายงานผลจากแบบประเมินไม่สามารถใช้ตัดสินพฤติกรรม ซึ่งถูกผิดหรือ
ชี้ชัดอนาคต อีกทั้งไม่สะท้อนระดับการศึกษา ความฉลาดเฉลียว แต่การประเมินจะให้ข้อมูลเชิง
สัมพันธ์กับความเป็นตัวตนของบุคคลนั้น เป็นเครื่องมือในการช่วยสังเกตและวิเคราะห์ฐาน
ความคิด (ซึ่งส่งผลถึงพฤติกรรมของบุคคล) เพื่อใช้ประกอบพิจารณาตีความคุณลักษณะบุคคลไม่ว่า
จะเป็นแบบใดก็ตาม มีทั้งข้อดี และข้อด้อยของตัวเอง ไม่ควรนำมาวัดผลว่าแบบไหนที่ดีหรือด้อย
กว่าอีกแบบประโยชน์ที่ได้รับจากการรู้จักเข้าใจตน โดยทั่วไป บุคคลมักจะให้ความสนใจในการทำ
ความรู้จักเข้าใจเรื่องสิ่งแวดล้อมภายนอกก่อน เช่น เทคโนโลยี เครื่องมือเครื่องใช้ อุปกรณ์การเงิน
รวมถึงบุคคลอื่น มากกว่าทำความเข้าใจกับตัวตนของตัวเอง คนจำนวนมากจึงมักจะรู้สึกสับสน
ในความเป็นตัวตนของเรา จากที่บุคคลภายนอกข้างเคียงให้ข้อมูลป้อนกลับเข้ามาไม่ว่าจะในรูปแบบของ
คำพูด คำวิพากษ์ ความเห็น ความรู้สึกสะท้อน ซึ่งเกิดขึ้นจากโลกทัศน์ของคนเหล่านั้น โดยที่อาจไม่
เป็นไปตามที่ตัวเราเป็น การได้ทำความเข้าใจตัวเองทั้งในแง่บวก และลบ ผ่านเครื่องมือ DISC เป็น
เหมือนการได้ข้อมูลป้อนกลับที่ผ่านทัศนคติของเราเอง สะท้อนแนวโน้มพฤติกรรมความถนัดใน
การแสดงออก ความชอบ ความไม่ชอบ ความปรารถนา สิ่งแวดล้อมที่พึงใจ แรงกระตุ้นจิตใจ สิ่งเร้า
ความตึงเครียด เป็นต้น ทำให้เราสามารถคาดการณ์ได้ถึงผลกระทบที่อาจตามมาจากการแสดง
พฤติกรรม ในแต่ละสถานการณ์ที่แตกต่างกัน อีกทั้งบุคคลที่เข้าใจถึงความแตกต่างระหว่างบุคคล
จะสามารถยอมรับ และปรับตัวเข้ากับบุคคลอื่นและสถานการณ์ได้ดีขึ้น สร้างสัมพันธภาพที่ดีกับ
บุคคลที่แวดล้อม

จึงกล่าวได้ว่า การตระหนักรับรู้ตัวตนเป็นรากฐานสำคัญต่อความก้าวหน้า ซึ่ง
ความสำเร็จทั้งชีวิตส่วนตัว และหน้าที่การงานการรู้จักตัวตนของตัวเอง เป็นพื้นฐานสำคัญที่จะทำ
ให้บุคคลเติบโตขึ้นอย่างมั่นคง รู้ทิศทางที่ชัดเจนคือ ประเมินตนเองได้อย่างชัดเจน ทั้งจุดแข็งและ
จุดอ่อน สร้างสมดุลด้านบวก ด้านลบ ของคุณลักษณะตัวเองได้ดีขึ้น แสดงตนได้อย่างเหมาะสม
เต็มศักยภาพ ทั้งทางด้านการแสดงความคิดเห็น การร่วมกิจกรรมต่างๆ พัฒนาตัวเอง เสริมจุดแข็ง
ปรับจุดอ่อน อย่างมีทิศทาง สื่อสาร สร้างสัมพันธภาพด้วยความเข้าใจกันจากการที่ยอมรับในความ
แตกต่างของแต่ละบุคคล ปรับตัว ยืดหยุ่น ไปด้วยแต่ละสถานการณ์ได้อย่างผ่อนคลายมากขึ้น

การประยุกต์ใช้ DISC สามารถนำไปประยุกต์ใช้อย่างหลากหลายในทุกกิจกรรมที่มี
ความเกี่ยวข้องกับบุคคลในเชิงธุรกิจ เพื่อใช้ส่งเสริมการพัฒนาบุคลากร บริหารทรัพยากรบุคคลใน
องค์กร และสรรหา อาชีพ ผลการประเมินสามารถสะท้อนคุณลักษณะเฉพาะของบุคคล สไตส์การ

สื่อสารปฏิสัมพันธ์ในสังคม สไตล์การตัดสินใจ และเผชิญปัญหา บทบาทที่ถนัดในการแสดงออก และตอบสนองต่อสภาพแวดล้อม ความถนัด และแรงกระตุ้นเชิงใจ ได้อย่างชัดเจนยิ่งขึ้น ข้อมูลจำเพาะเหล่านี้เป็นตัวช่วยที่สำคัญในการบริหารจัดการบุคลากร และทีมงานในองค์กร เพิ่มขีดความสามารถ เกิดประสิทธิภาพและมูลค่าเพิ่มในเนื้องาน สร้างความโดดเด่นให้องค์กรมุ่งสู่ความสำเร็จได้อย่างยั่งยืน

2.3 แนวคิดและทฤษฎีเกี่ยวกับผลการปฏิบัติงานของพนักงาน

ผลการปฏิบัติงานของพนักงาน (Job Performance) หมายถึง การที่พนักงานแต่ละรายสามารถทำงานในหน้าที่ๆ ตนได้รับมอบหมายให้บรรลุเป้าหมาย ตามที่กิจการกำหนดไว้โดยใช้ปัจจัยการทำงานหลายๆ ด้านเป็นตัวชี้วัด ซึ่งมีหลักสำคัญคือ การปฏิบัติงานอย่างหนัก และอยู่ภายใต้กฎระเบียบของบริษัท ซึ่งผลสำเร็จในการปฏิบัติงานอาจมาจากทัศนคติ หรือการรับรู้ของหัวหน้างานที่มีต่อการปฏิบัติงานของพนักงานในสังกัดของตนเอง ซึ่งแสดงออกมาเป็นคะแนนการพิจารณาประเมินผลสำเร็จในการปฏิบัติงานประจำปี ตามที่บริษัทได้กำหนดไว้ให้ผู้บังคับบัญชาประเมินผลการปฏิบัติงานของพนักงานตามหลักเกณฑ์ที่บริษัทเป็นผู้กำหนด ผลการปฏิบัติงานของพนักงานสามารถพิจารณาตัดสินได้ว่าพนักงานปฏิบัติงานได้ดีเพียงใด โดยนำผลไปเปรียบเทียบกับมาตรฐานที่ได้กำหนดไว้และสื่อสารให้พนักงานทราบถึงผลการประเมินดังกล่าว (Mathis & Jackson, 2000)

กล่าวโดยสรุปได้ว่า ผลการปฏิบัติงานของพนักงาน หมายถึง ผลการปฏิบัติงานของพนักงานคนหนึ่งในช่วงเวลาหนึ่งว่าสามารถปฏิบัติงานได้ดีเพียงใด โดยมีเกณฑ์การประเมินในด้านต่างๆ คือ ประสิทธิภาพของงาน คุณภาพของงาน ความสามารถเกี่ยวกับงาน ความสามารถในการใช้อุปกรณ์และเครื่องมือ ทักษะในการสื่อสาร ความเชื่อถือไว้วางใจได้ ความสัมพันธ์กับเพื่อนร่วมงาน ความสัมพันธ์กับผู้บังคับบัญชา การปฏิบัติตามกฎ ระเบียบ ความคิดริเริ่มสร้างสรรค์ ความรับผิดชอบ การปรับตัว การควบคุมอารมณ์ และความตรงต่อเวลา ซึ่งเป็นกระบวนการเกี่ยวกับพฤติกรรม

ทั้งนี้เกณฑ์ที่ใช้วัดผลการปฏิบัติงานของพนักงานที่มีการพิจารณาทั้งในผลผลิตที่ทำได้กับความคุ้มค่า และความสามารถส่วนบุคคลของผู้บริหาร คือ ความมีประสิทธิภาพในการปฏิบัติงาน (Effectiveness) ประสิทธิภาพในการปฏิบัติงาน (Efficiency) และพฤติกรรมในการทำงาน (Behavior) หมายถึง องค์กรต้องสามารถทำกำไรได้มากกว่าคู่แข่ง องค์กรต้องเติบโตเร็วกว่าคู่แข่ง

แข่งขัน การวัดผลงานที่ทำได้เทียบกับเป้าหมาย ถ้าหากสามารถทำให้ตามเป้าหมายที่วางไว้ก็แสดงว่ามี ผลการปฏิบัติงานของพนักงาน

2.3.1 ประสิทธิภาพในการปฏิบัติงาน

ประสิทธิภาพ เป็นปัจจัยสำคัญที่จะเป็นเครื่องตัดสินใจขั้นสุดท้ายว่าการบริหาร และ องค์การประสบความสำเร็จมากน้อยเพียงใด แต่ก็ยังมีความความเข้าใจที่แตกต่างกันของนักวิชาการ ในแต่ละสาขา ดังนี้

Gibson และคณะ (1988) ได้ให้ความหมายของประสิทธิภาพ เอาไว้ว่า หมายถึง การกระทำใดๆ หรือความพยายามใดๆ ที่มีความมุ่งหมายจะได้รับผลอะไรสักอย่างให้เกิดขึ้น และการกระทำจะมีประสิทธิภาพสูงเพียงใดขึ้นอยู่กับว่าผลที่ได้รับนั้นตรง ครบถ้วน ในเชิงปริมาณและคุณภาพ และใช้พลังงานมากน้อยเพียงใด

Robbins และ Coulter (1999) ประสิทธิภาพ หมายถึง การทำงานให้ประสบความสำเร็จ ในแง่การจัดการ คือการปฏิบัติงานบรรลุเป้าหมายขององค์กร ประกอบด้วยคุณลักษณะที่สำคัญ ประสิทธิภาพขององค์กร ต้องวิเคราะห์ต่อไปถึงความสำเร็จของเป้าหมายว่า เป็นเป้าหมายระยะสั้น หรือระยะยาว และเป็นเป้าหมายที่แท้จริงหรือเป็นเพียงเป้าหมายทางการที่ระบุไว้ให้บุคคลภายนอกทราบเป้าหมายต่างๆ ที่กำหนด และทำบรรลุผลลงได้นั้นได้ตกเป็นของกลุ่มผลประโยชน์กลุ่มใดอีก ด้วย เช่น เป็นของผู้ถือหุ้น พนักงานฝ่ายจัดการ หรือกับลูกค้า และบุคคลภายนอกรวมทั้งสังคมโดยรวม

วรัท พุกษากุลนันท์ (2550) กล่าวว่า ประสิทธิภาพ หมายถึง การปฏิบัติงานให้บรรลุวัตถุประสงค์หรือเป้าหมายที่ตั้งไว้ หรือความสามารถในการดำเนินการให้บรรลุวัตถุประสงค์ที่วางไว้ จุดสำคัญของประสิทธิภาพอยู่ที่ความสัมพันธ์ระหว่างผลผลิตที่ถูกคาดหวังตามวัตถุประสงค์ที่วางไว้และผลผลิตจริงที่มีขึ้น

อนันท์ งามสะอาด (2558) กล่าวว่า ประสิทธิภาพ หมายถึงผลสำเร็จของงาน ที่เป็นไปตามความมุ่งหวังที่กำหนดไว้ในวัตถุประสงค์ หรือเป้าหมาย ได้แก่

1. เป้าหมายเชิงปริมาณ จะกำหนดชนิดประเภท และจำนวนของผลผลิตสุดท้ายที่ต้องการได้รับ เมื่อการดำเนินงานเสร็จสิ้นลง
2. เป้าหมายเชิงคุณภาพ จะแสดงถึงคุณค่าของผลผลิตที่ได้รับจากการดำเนินงานนั้นๆ
3. มุ่งเน้นที่จุดสิ้นสุดของกิจกรรมหรือการดำเนินงานว่าได้ผลตามที่ตั้งไว้หรือไม่ และที่สำคัญต้องมีตัวชี้วัดที่ชัดเจน

ซึ่งประสิทธิผลขององค์กร เกิดขึ้นมาจากความสัมพันธ์ของประสิทธิผล 3 ระดับ คือ ระดับบุคคล ระดับกลุ่ม และระดับองค์กร โดยประสิทธิผลของบุคคล (Individual Effectiveness) เป็นผลสำเร็จในการปฏิบัติงานตามหน้าที่งานของพนักงานหรือสมาชิกในองค์กร ประสิทธิผลของกลุ่ม (Group Effectiveness) เป็นผลรวมของการช่วยเหลือสนับสนุนของสมาชิกในกลุ่มทั้งหมด โดยที่ประสิทธิผลของกลุ่มเป็นมากกว่าผลรวมของการช่วยเหลือสนับสนุนของแต่ละบุคคลในงานประเภทระบบสายพาน แต่เป็นในรูปของพลังร่วม (Synergy) ที่แสดงถึงการรวมตัวกันของการร่วมมือร่วมใจช่วยเหลือกันของบุคคลที่มีลักษณะเกินกว่า การรวมตัวกันธรรมดา และประสิทธิผลขององค์กร (Organizational Effectiveness) องค์กรนั้นประกอบไปด้วยบุคคล และกลุ่ม ประสิทธิผลขององค์กรจึงประกอบด้วยประสิทธิผลของบุคคล และกลุ่ม โดยประสิทธิผลขององค์กร ได้รับผลกระทบของพลังร่วม (Synergistic Effects) ที่ทำให้องค์กรมีประสิทธิผลในระดับสูงกว่าผลรวมจากส่วนต่างๆ ธรรมดา ทั้งนี้ ความสัมพันธ์ระหว่างประสิทธิผลระดับบุคคล ระดับกลุ่ม และระดับองค์กร แสดงให้เห็นว่าประสิทธิผลของกลุ่มจะขึ้นอยู่กับประสิทธิผลของบุคคล ในขณะที่ประสิทธิผลขององค์กรขึ้นอยู่กับประสิทธิผลของบุคคลและกลุ่ม ความสัมพันธ์ทั้ง 3 ระดับนี้จะมีพลังร่วม (Synergistic Effects) ซึ่งทำให้ประสิทธิผลของกลุ่มมีมากกว่าผลรวมของประสิทธิผลของบุคคล (Gibson et al., 1997)

Steers (1977) ได้เสนอแนวคิดเกี่ยวกับการวัดประสิทธิผลในการปฏิบัติงานว่าสามารถแบ่งออกได้เป็น 2 ประเภท คือ การวัดประสิทธิผล โดยใช้ตัวบ่งชี้เดียว จะใช้ตัวบ่งชี้ที่เหมาะสมที่สุดเพียงตัวเดียวในการวัดประสิทธิผล ในการปฏิบัติงาน เช่น ความสามารถในการผลิตความสำเร็จในภารกิจ และการบรรลุเป้าหมาย เป็นต้น และการวัดประสิทธิผลโดยใช้ตัวบ่งชี้หลายตัวเป็นการวัดที่ได้รับการนิยามอย่างกว้างขวาง โดยจะใช้ตัวบ่งชี้หลายตัว ที่เหมาะสมในการวัดประสิทธิผล ดังที่ Hoy และ Miskle (1991) ได้เสนอการวัดประสิทธิผลจากหน้าที่หลัก 4 ประการ ดังนี้ ประการแรก การปรับตัว หมายถึง การปรับตัวให้สอดคล้องกับสิ่งแวดล้อมภายนอก โดยการปรับเปลี่ยนการดำเนินงานให้ตอบสนองต่อสภาพการณ์ใหม่ๆ ซึ่งตัวบ่งชี้ที่ใช้วัด อาทิ ความสามารถในการปรับเปลี่ยน นวัตกรรม ความเจริญเติบโต และการพัฒนา ประการที่สองการบรรลุเป้าหมาย หมายถึง การกำหนดวัตถุประสงค์ การจัดหา และการใช้ทรัพยากรต่างๆ เพื่อให้การดำเนินงานบรรลุวัตถุประสงค์ที่วางไว้ ซึ่งตัวบ่งชี้วัด อาทิ ผลสัมฤทธิ์คุณภาพ การแสวงหาทรัพยากร และประสิทธิภาพ ประการต่อมาการบูรณาการ หมายถึง การประสานความสัมพันธ์ของสมาชิกภายในองค์กร เพื่อรวมพลังให้มีความเป็นอันหนึ่งอันเดียวกันในการปฏิบัติงานขององค์กร ซึ่งตัวบ่งชี้วัด อาทิ ความพอใจ บรรยากาศ การติดต่อสื่อสาร และความขัดแย้ง และประการสุดท้ายการคงสภาพ

ความสมบูรณ์ของระบบค่านิยม หมายถึง การดำรงรักษาระบบค่านิยมขององค์กร ซึ่งตัวบ่งชี้วัด อาทิ ความภักดี ศูนย์กลางความสนใจในชีวิต แรงจูงใจ และเอกลักษณ์

2.3.2 ประสิทธิภาพในการปฏิบัติงาน

ประสิทธิภาพ (Efficiency) คือ ความสัมพันธ์ระหว่างสิ่งที่นำเข้า (Input) กับผลที่ได้รับ (Output) กล่าวอีกนัยหนึ่ง คือ ความสัมพันธ์ระหว่างทรัพยากรที่นำเข้ากับผลผลิตที่ได้ออกมา ถ้าผลผลิตที่นำออกมาได้มากกว่าทรัพยากร คน เงิน วัสดุสิ่งของ เครื่องไม้เครื่องมือต่างๆ และวิธีการผลิต ที่นำเข้าหมายความว่าองค์กรได้เพิ่มประสิทธิภาพกระบวนการในการทำงานขึ้นแล้ว (มองในแง่ผลผลิตที่ได้รับ) ในขณะที่เดียวกันหากผลผลิตที่ได้รับออกมาใช้ทรัพยากรที่น้อยหรือประหยัดที่สุด หรือการทำสิ่งต่างๆ ให้ถูกต้องเหมาะสมเมื่อนำมารวมเข้ากับความพึงพอใจของลูกค้าย่อมหมายถึง คุณภาพในตัวผลิตภัณฑ์ที่มีโอกาสซื้อซ้ำหรือบอกต่อของลูกค้าว่าเป็นผลิตภัณฑ์ที่มีคุณภาพ พร้อมด้วยประสิทธิผล คือ การทำงานให้ประสบผลสำเร็จในแง่การจัดการ หมายถึง การบรรลุเป้าหมายขององค์กรนั้นคือการทำถูกต้องตามสิ่งที่ควรทำ (Robbins & Coulter, 1999) ซึ่งบางครั้งการบรรลุเป้าหมายขององค์กรในแง่ผลผลิตอาจไม่คำนึงถึงต้นทุนหรือทรัพยากรที่ใช้ไปว่าจะมากหรือน้อยเพียงใด ในภาพรวมแล้วประสิทธิภาพที่จะเกี่ยวข้องกับวิธีการ ส่วนประสิทธิผลจะเกี่ยวข้องกับจุดมุ่งหมายหรือเป้าหมาย

ชัดเจนว่าประสิทธิภาพสูงย่อมมีส่วนที่สำคัญกับประสิทธิผลสูง การจัดการที่ดีจึงต้องคำนึงถึงหลักที่ว่าการบริหารที่ไร้ประสิทธิภาพและประสิทธิผล หรือมีประสิทธิผลแต่ขาดประสิทธิภาพ ซึ่งมีโอกาสเกิดขึ้นได้ตลอดเวลาหาก ผู้บริหาร ไม่ตระหนักและคอยระมัดระวังอย่างเพียงพอ หลักการนี้ย่อมเป็นข้อเตือนใจที่ผู้บริหารต้องพึงคอยระมัดระวังเป็นอย่างยิ่ง

Harrington Emerson (1982) ได้นำเสนอแนวคิดโดยการเน้นการจัดการทรัพยากรและขจัดความสูญเปล่า โดยการยอมรับการบริหารจัดการแบบวิทยาศาสตร์และให้ความสำคัญที่โครงสร้างและเป้าหมายขององค์กร จากการกำหนดจุดมุ่งหมายที่ชัดเจน ใช้หลักเหตุผลทั่วไป ให้คำแนะนำที่ดี มีวินัย ยุติธรรม มีข้อมูลที่เชื่อถือได้เป็นปัจจุบัน ถูกต้องและแน่นอน มีความจับใจของการจัดส่ง ทำงานได้ มาตรฐานและตารางเวลา รักษาสภาพแวดล้อมให้มีสภาพมาตรฐาน ปฏิบัติการที่มีมาตรฐาน มีคำสั่งการปฏิบัติงานที่มีมาตรฐานระบุไว้ และรู้จักการให้รางวัลที่มีประสิทธิภาพ

2.3.3 การวัดผลการปฏิบัติงานของพนักงาน

ผลสำเร็จในการปฏิบัติงานพนักงานในการวิจัยครั้งนี้หมายถึง ผลการปฏิบัติงานของพนักงานแต่ละคนในช่วงเวลาหนึ่งว่าสามารถปฏิบัติงานได้ดีเพียงใด รวมทั้งสามารถทำงานใน

หน้าที่ที่ตนได้รับมอบหมายให้สำเร็จตามวัตถุประสงค์ และเป้าหมายที่บริษัทกำหนดไว้ โดยใช้ปัจจัยการทำงานหลายๆ ด้านเป็นตัวชี้วัด การวิจัยครั้งนี้มุ่งเน้นผลการปฏิบัติงานของพนักงาน 3 ด้าน คือ ด้านยอดขาย ด้านการให้บริการ และด้านพฤติกรรมการทำงาน

จากการทบทวนวรรณกรรมเกี่ยวกับการวัดผลการปฏิบัติงานของพนักงานพบว่า งานวิจัยเชิงประจักษ์เกี่ยวกับการวัดผลการปฏิบัติงานของพนักงานมีหลายมาตรวัด เช่น งานวิจัยของ Ahire et al.,(1996), งานวิจัยของ Murray, Snider และ Midkiff (1999) และ งานวิจัยของ Goris et al. (2003) เป็นต้น

2.4 งานวิจัยด้านบุคลิกภาพที่ส่งผลต่อการทำงาน

กล้าหาญ ฦ น่าน (2557) ได้ทำการศึกษาอิทธิพลของคุณลักษณะบุคลิกภาพ และความพึงพอใจที่มีต่อการปรับตัวในการทำงานของผู้เข้าสู่ตลาดบัณฑิตใหม่ และพบว่าคุณลักษณะบุคลิกภาพมีอิทธิพลทางตรงต่อความพึงพอใจในงานและการปรับตัวในการทำงาน โดยบัณฑิตที่มีบุคลิกภาพแบบประนีประนอม บุคลิกภาพแบบแสดงตัว บุคลิกภาพแบบมีมโนสำนึก และบุคลิกภาพแบบเปิดรับประสบการณ์ จะมีความพึงพอใจภายในงานและภายนอกงาน รวมถึงการปรับตัวให้เข้ากับงาน สภาพแวดล้อม และเพื่อนร่วมงานได้ดีกว่าแบบอื่น

Kim Yun-Young และคณะ (2012) พบว่า ความพึงพอใจในงาน ความผูกพันต่อองค์กร และประสิทธิภาพในการทำงาน มีสหสัมพันธ์เชิงบวกต่อกัน และยังพบว่ากลุ่ม Dominance มีคะแนนการประเมินตนเองว่ามีประสิทธิภาพในการทำงาน มีความท้าทายในการทำงาน และมีความสุขเมื่อพูดถึงบริษัทที่ทำงาน ที่สูงกว่าคะแนนที่กลุ่ม Steadiness ประเมินตนเองอย่างมีนัยสำคัญทางสถิติ

2.5 วิธีการจำแนกข้อมูลแบบหลายลาเบล (Multi-label Classification)

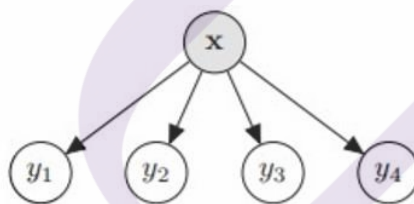
โดยปกติข้อมูลทั่วไปจะใช้คุณลักษณะหนึ่งคุณลักษณะ (Feature) หรือมากกว่า เพื่อทำนาย หรือวิเคราะห์ลาเบล (Label) ของข้อมูล ซึ่งมักจะมีเพียงแค่ตัวเดียว เช่น การที่บริษัท โทรคมนาคมทำวิเคราะห์หาลูกค้าที่มีแนวโน้มจะย้ายค่าย ซึ่งลาเบลคือ การย้ายค่ายหรือไม่ย้าย หรือ การทำนายว่าพนักงานในบริษัทมีแนวโน้มที่จะลาออกหรือไม่ เป็นต้น

แต่ก็มีข้อมูลอีกหลายประเภทที่ข้อมูลลาเบลที่ถูกใช้ในการจำแนกสามารถมีได้หลายลาเบล ซึ่งข้อมูลในลักษณะนี้ถือว่าเป็นข้อมูลที่ลาเบลไม่สามารถแยกออกจากกันได้โดยเด็ดขาด (Not Mutually Exclusive) เช่น ภาพยนตร์เรื่องหนึ่งสามารถมีได้หลากหลายประเภท (ลาเบล) ทั้งตลก แอ็คชั่น และดราม่าภายในเรื่องเดียว หรือจะเป็นประเภทเพลงที่อาจจะเป็นที่ป๊อป (Pop) และร็อก (Rock) เป็นต้น

ซึ่งการจำแนกประเภทข้อมูลในลักษณะนี้จึงจำเป็นต้องใช้การจำแนกหลายลาเบลมาช่วยในการแก้ไขปัญหา โดยในงานวิจัยนี้จะเลือกใช้วิธีการจำแนกหลายลาเบลแบบ Problem Transformation methods เพียงอย่างเดียวเท่านั้น โดยวิธีการที่ใช้จะมีทั้งหมด 3 วิธีการ ได้แก่

2.5.1 Binary Relevance (BR)

เป็นวิธีการที่ง่ายที่สุด คือจะนำคุณลักษณะที่ใช้ทำนาย (X) ไปทำการทำนายลาเบลแต่ละตัว โดยแต่ละลาเบลนั้นจะถูกแยกออกจากกันอย่างอิสระ ดังภาพ



ภาพที่ 2.1 การทำงานของ Binary Relevance

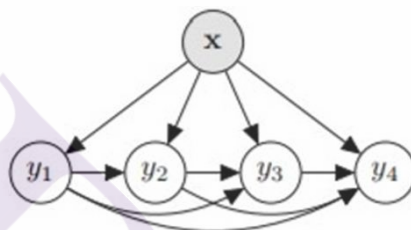
2.5.2 Classifier Chains (CC)

เป็นการรวบรวมตัวจำแนกประเภทแบบไบนารีต่างๆ ให้รวมเป็นหนึ่งเดียว เพื่อที่จะสามารถเทียบความสัมพันธ์กับลาเบลที่ต้องการจำแนกได้ หากในชุดข้อมูลมีข้อมูลทั้งหมด N คลาส

(Class) ก็จะมีการสร้างตัวจำแนกแบบไบนารีทั้งหมด N ตัว โดยจะระบุตัวเลขระหว่าง 0 ถึง $N-1$ โดยตัวเลขเหล่านี้จะหมายถึงลำดับของแบบจำลองที่อยู่ในโซ่

โดยหลังจากที่ตัวจำแนกประเภทแต่ละตัวทำการเรียนรู้จากชุดข้อมูลสอนแล้วก็จะถูกรวมเข้ากับข้อมูลลาเบลที่แท้จริงของแต่ละคลาส หลังจากนั้นผลของข้อมูลที่ทำนายได้จะถูกส่งไปยังแบบจำลองย่อยที่อยู่ในลำดับถัดไปของโซ่เพื่อที่จะใช้เป็นคุณลักษณะในการทำนายต่อไป

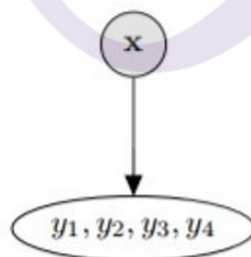
จะเห็นได้ว่าลำดับของโซ่นั้นมีความสำคัญอย่างมาก เพราะสำหรับแบบจำลองที่อยู่ในลำดับแรกนั้นจะไม่มีข้อมูลของลาเบลอื่นๆ ในขณะที่แบบจำลองลำดับสุดท้ายจะมีข้อมูลทั้งหมดจากแบบจำลองในลำดับก่อนหน้า อย่างไรก็ตามก็ได้เป็นการยากที่จะทำการจัดลำดับที่ดีที่สุดของโซ่ ดังนั้นจึงใช้วิธีการสุ่มลำดับเพื่อแก้ไขปัญหานี้แทน



ภาพที่ 2.2 การทำงานของ Classifier Chains

2.5.3 Label Powerset (LP)

จะทำการแปลงปัญหาจากการจำแนกแบบหลายลาเบลไปเป็นการจำแนกแบบหลายคลาส (Multi-class classification) โดยลาเบลแต่ละตัวจะถูกรวมเข้าไว้ด้วยกัน และการรวมกันในแต่ละรูปแบบจะถูกแปลงให้เป็นคลาสใหม่ ซึ่งแตกต่างจากคลาสอื่นๆ ทำให้วิธีการนี้จะมีการสร้างคลาสขึ้นมาจำนวนมาก



ภาพที่ 2.3 การทำงานของ Label Powerset

2.6 การจำแนกประเภท (Classification)

การจำแนกประเภท เป็นกระบวนการสร้างแบบจำลองจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนด ซึ่งถือเป็นการเรียนรู้แบบมีผู้สอน (Supervised Learning) โดยจะนำข้อมูลส่วนหนึ่งจากข้อมูลทั้งหมดที่มีอยู่มาเป็นข้อมูลสอน (Training Data) ให้เครื่องจักรเรียนรู้ ซึ่งข้อมูลสอนจะประกอบไปด้วยข้อมูลต่างๆ (input) และผลลัพธ์ (output) โดยโปรแกรมจะเรียนรู้รูปแบบต่างๆ จากข้อมูลสอน เพื่อนำไปวิเคราะห์หาผลลัพธ์จากข้อมูลอื่นๆ เพื่อหาวิธีการจำแนกข้อมูลออกเป็นกลุ่มตามที่ได้กำหนดไว้ โดยผลลัพธ์ที่ได้จากการเรียนรู้คือ แบบจำลองการจำแนกข้อมูล (Classifier Model) หลังจากนั้นจะนำข้อมูลอีกส่วนหนึ่งมาใช้เป็นข้อมูลสำหรับทดสอบ (Testing Data) และใช้เปรียบเทียบกับกลุ่มที่หามาได้จากแบบจำลองเพื่อทดสอบความถูกต้อง โดยเทคนิคพื้นฐานที่นิยมนำมาใช้สำหรับการจำแนกประเภทข้อมูล ได้แก่

2.6.1 ต้นไม้ตัดสินใจ (Decision Tree)

เป็นแบบจำลองทางคณิตศาสตร์เพื่อการหาทางเลือกที่ดีที่สุด โดยการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบของแผนภาพต้นไม้ โดยสามารถจำแนกได้ทั้งปัญหาแบบไบนารี (Binary classes) และปัญหาแบบหลายคลาส (Multiple classes) รวมทั้งยังสามารถนำไปใช้แสดงโครงสร้างข้อมูลในฐานข้อมูล (Database) เพื่อความรวดเร็วในการสืบค้น หรือการวิเคราะห์การถดถอย (Regression Analysis) ได้อีกด้วย

โดยต้นไม้ตัดสินใจจะแสดงทางเลือกและเหตุการณ์เพื่อแสดงลำดับขั้นตอนทางเลือกที่มีอยู่ทั้งหมด และเหตุการณ์ต่าง ๆ ที่อาจจะเกิดขึ้น รวมไปถึงผลของการเลือกทางเลือกภายใต้เหตุการณ์ต่าง ๆ โดยจะประกอบไปด้วย โหนด (Node) ที่แสดงคุณลักษณะของข้อมูล ซึ่งใช้เป็นจุดตัดสินใจกิ่ง (Branch) จะแสดงถึงผลลัพธ์ที่ได้จากการพิจารณาเงื่อนไขของทางเลือก หรือเหตุการณ์ต่าง ๆ ที่เป็นไปได้จากโหนด ซึ่งแต่ละกิ่งจะนำไปสู่ผลลัพธ์สุดท้ายคือ ใบ (Leaf) ซึ่งแสดงกลุ่มของผลลัพธ์ที่เป็นค่าเป้าหมาย (Target) หรือคลาส (Class) ของปัญหา

สำหรับการสร้าง Decision Tree ของแต่ละอัลกอริทึมนั้นจะมีลักษณะที่คล้ายกันคือ เริ่มต้นทำการคัดเลือกแอตทริบิวต์ที่มีความสัมพันธ์กับคลาสมากที่สุดขึ้นมาเป็นโหนดบนสุดของต้นไม้ (Root Node) หลังจากนั้นจะทำการแตกกิ่งแอตทริบิวต์ออกไปเรื่อยๆ จนสามารถแบ่งข้อมูลออกเป็นคลาสได้ชัดเจน

2.6.2 การสุ่มของป่า (Random Forest)

เป็นอัลกอริทึมที่นำเสนอโดย Brieman (2001) การสุ่มของป่าจะประกอบไปด้วยต้นไม้ตัดสินใจ (Decision Tree) จำนวนหลายต้น แต่ละต้นเป็นอิสระต่อกัน ในการพยากรณ์ผลลัพธ์จะ

เริ่มต้นด้วยการนำเข้าข้อมูลสู่ต้นไม้ตัดสินใจของแต่ละต้น โดยต้นไม้แต่ละต้นจะทำการจำแนกประเภท หรือพยากรณ์ผลลัพธ์ออกมา หลังจากนั้นผลลัพธ์สุดท้ายจะได้จากการโหวต (Majority Vote) ผลลัพธ์ของต้นไม้แต่ละต้น โดยผลลัพธ์ที่ได้ผลโหวตมากที่สุดจะเป็นผลลัพธ์สุดท้าย ซึ่งลักษณะของต้นไม้ที่อยู่ภายในป่าของเทคนิคการสุ่มของป่าจะถูกควบคุมด้วย 3 ปัจจัยได้แก่

1. ต้นไม้แต่ละต้นจะถูกสอน (Train) โดยการใช้เซตย่อยจากข้อมูลตัวอย่าง
2. เมื่อต้นไม้โตขึ้น จะสามารถค้นหา โหนด (Node) แต่ละ โหนดที่อยู่ในกิ่งที่ดีที่สุดของต้นไม้ โดยการใช้การสุ่มเลือกคุณสมบัติจาก N คุณสมบัติ
3. ต้นไม้แต่ละต้นจะ ไม่มีการตัดออก แต่จะปล่อยให้ต้นไม้โตขึ้นไปเรื่อยๆ จนได้ผลลัพธ์ที่ดีที่สุดหลังจากการสร้างป่า แล้วทำการให้คะแนน (Vote) โดยต้นไม้ภายในป่า หากต้นไม้ต้นใดได้คะแนนสูงที่สุดก็จะนำเอาต้นไม้ต้นนั้นออกมาสร้างเป็นแบบจำลอง

2.6.3 การหาเพื่อนบ้านใกล้สุด K ตัว (K-Nearest Neighbor)

เป็นเทคนิคที่ใช้กับปัญหาแบบการจำแนกกลุ่มข้อมูล โดยเทคนิคการหาเพื่อนบ้านที่ใกล้ที่สุดจะตัดสินใจว่าคลาสใดที่จะแทนเงื่อนไขหรือกรณีใหม่ ๆ ได้บ้าง โดยจะมีการระบุค่าตัวเลขจำนวนเพื่อนบ้านหรือค่า K ซึ่งจะเป็ค่าที่บอกถึงจำนวนเงื่อนไขหรือกรณีต่างๆ สำหรับแต่ละคลาส รวมไปถึงการกำหนดเงื่อนไขใหม่ๆ ให้คลาสที่เหมือนกับคลาสที่ใกล้เคียงกันมากที่สุด ซึ่งจะมีความแตกต่างจากเทคนิคอื่นที่ไม่มีการแบ่งข้อมูลชุดสอนในการสร้างแบบจำลองแต่จะใช้ข้อมูลทั้งหมดมาเป็นแบบจำลอง

โดยเทคนิคของขั้นตอนวิธีการหาเพื่อนบ้านใกล้ที่สุด จะเป็นการหาระยะห่างระหว่างแต่ละตัวแปรในข้อมูล จากนั้นก็คำนวณค่าออกมา ซึ่งจะเหมาะสมสำหรับข้อมูลแบบตัวเลข แต่ตัวแปรที่เป็นค่าแบบไม่ต่อเนื่องนั้นก็ยังสามารถทำได้ เพียงแต่ต้องการการจัดการแบบพิเศษเพิ่มขึ้น โดยการใช้งานเทคนิคนี้จำเป็นจะต้องมีการระบุค่าตัวเลขจำนวนเพื่อนบ้านหรือค่า K ซึ่งค่านี้จะเป็นตัวบอกจำนวนของกรณีที่จะต้องค้นหาในการทำนายกรณีใหม่ซึ่งได้แก่ 1-NN , 2-NN , 3-NN , , K-NN โดยที่ K เป็นจำนวนเต็มบวก เช่น 2-NN หมายถึงจะมีการค้นหา 2 กรณีที่มีลักษณะใกล้เคียงกับกรณีใหม่ (Nearest Cases) ในการทำนายกรณีใหม่ซึ่งสมการที่เกี่ยวข้องกับ KNN คือ

$$D = \frac{(TV - OB)}{TV} * 100 \quad (\text{สมการที่ 2.6.3})$$

โดย D คือ ค่าระยะห่างของจุด 2 จุดของแบบ Percentage Disagreement

TV คือ ค่าทางทฤษฎี (Theoretical Value)

OB คือ ค่าที่สังเกตเห็น (Observed Value)

2.6.4 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล โดยอาศัยแบบจำลองทางคณิตศาสตร์ในเรื่องของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกต้องป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นเลือกเส้นที่แบ่งข้อมูลได้ดีที่สุดโดยไม่จำเป็นต้องเป็นเส้นตรงเสมอไป

หลักการของอัลกอริทึมคือการนำข้อมูลที่ใช้สอนมากระจายเป็นเวกเตอร์ในระนาบ หรือสเปซ N มิติ (Feature Space) จากนั้นคำนวณหาระนาบการตัดสินใจที่เรียกว่า ระนาบเกิน หรือไฮเปอร์เพลน (Hyperplane) ซึ่งต้องเป็นเส้นที่มีค่ามาร์จิ้น (Margin) หรือระยะห่างระหว่างจุดของข้อมูลที่อยู่ใกล้กับระนาบเกินทั้งสองด้าน คือ $d+$ และ $d-$ มากที่สุด เพราะถ้าค่ามาร์จิ้นน้อย แสดงว่าอาจมีจุดหนึ่งของระนาบเกินที่อยู่ใกล้ข้อมูลของแต่ละคลาสมากเกินไป ทำให้ข้อมูลใหม่ที่อยู่ห่างออกไปเล็กน้อยเกิดการทำนายผิดพลาดได้ โดยข้อมูลที่อยู่บนขอบของมาร์จิ้นเรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vector)

โดยในการหาเส้นระนาบเกินที่เหมาะสมที่สุด จะทำการหาดำแหน่งของซัพพอร์ตเวกเตอร์เพื่อใช้เป็นตัวแทนของกลุ่มข้อมูลทั้งหมด ในการพิจารณาเกณฑ์แบ่งกลุ่มโดยอาศัยหลักการคือ จะใช้ระนาบเกินที่เป็นระยะห่างที่สุดระหว่างข้อมูล 2 กลุ่ม ที่อยู่ใกล้กันมากที่สุดเพียงระนาบเดียวเท่านั้น จากนั้นจึงหาระนาบที่รักษาระยะห่างจากขอบมากที่สุดและถือว่าระนาบดังกล่าวคือ ระนาบสำหรับการแบ่งกลุ่มที่เหมาะสมที่สุด

จากที่กล่าวไปข้างต้นว่าซัพพอร์ตเวกเตอร์แมชชีนไม่เพียงแต่ทำการจำแนกประเภทข้อมูลได้แค่แบบเส้นตรงเท่านั้น แต่ยังสามารถประยุกต์ใช้ในการแก้ปัญหาในแบบที่ไม่เป็นเชิงเส้น (Nonlinear) ได้อีกด้วย โดยทำการแปลงข้อมูลผ่านฟังก์ชันที่ไม่เป็นเชิงเส้นและสร้างฟังก์ชันวัดระยะห่างที่เรียกว่าเคอร์เนลฟังก์ชัน (Kernel Function) บน Feature Space ซึ่งจะเหมาะสมสำหรับข้อมูลที่มีมิติของข้อมูลสูง (High Dimension) โดยมีวัตถุประสงค์ที่จะพยายามจะทำการลดความผิดพลาดในการทำนายกลุ่มข้อมูล พร้อมกับเพิ่มระยะแยกแยะโดยพยายามสร้างเส้นแบ่งกลางระหว่างกลุ่มให้มีระยะห่างระหว่างขอบเขตทั้งสองกลุ่มมากที่สุด

การคำนวณหาระนาบเกินสามารถทำได้ดังสมการที่ 2.5.4

$$\square^{\square} \square + \square = 0 \quad (\text{สมการที่ 2.6.4})$$

โดย	W	คือ ค่าเวกเตอร์น้ำหนัก
	X	คือ ข้อมูลนำเข้า
	b	คือ ค่าไบแอสหรือระยะเลื่อนขนานจากจุดกำเนิด

2.6.5 ตัวแยกแยะเบย์อย่างง่าย (Naïve Bayes)

เป็นแบบจำลองการจัดกลุ่มที่ใช้การคำนวณหาความน่าจะเป็นในการทำนายผล โดยเป็นเทคนิคในการแก้ปัญหาแบบจำแนกประเภทที่สามารถคาดการณ์ผลลัพธ์ได้ ซึ่งอยู่บนพื้นฐานของ Bayes' Theorem และสมมติฐานที่กำหนดให้การเกิดของเหตุการณ์ต่างๆ ที่ใช้ในการจัดกลุ่มนั้นเป็นอิสระต่อกัน ซึ่งการเรียนรู้จำแนกด้วยกระบวนการของ Naïve Bayes ได้รับความนิยมเนื่องจากมีการทำงานที่ไม่ซับซ้อน แต่ทำงานได้อย่างมีประสิทธิภาพ เช่น งานวิจัยด้านการจำแนกเอกสาร (Text Categorization) ประเภทต่างๆ เป็นต้น

ในการจัดกลุ่ม Naïve Bayes Model อาจจะมีการเกิดขึ้นของเหตุการณ์ต่างๆ ที่ใช้ในการจัดกลุ่มมากกว่า 1 ชนิด และเมื่อนำมาประยุกต์ใช้กับ Bayes' Theorem จะทำให้เกิดการคำนวณที่ซับซ้อนเนื่องจากการขึ้นต่อกันของการเกิดของเหตุการณ์ (dependence) ดังนั้น Naïve Bayes Model นั้นจึงตั้งสมมติฐานให้แต่ละเหตุการณ์ที่ใช้ในการจัดกลุ่มนั้นเป็นอิสระต่อกัน ซึ่งเป็นที่มาของคำว่า Naïve โดยเขียนเป็นสมการได้ดังนี้

$$P(H|E) = \frac{P(H) \prod P(E_i|H)}{\prod P(E_i)} \quad (\text{สมการที่ 2.6.5})$$

โดย

- $P(H|E)$ คือ ค่าความน่าจะเป็นที่จะเกิดเหตุการณ์ H เมื่อเกิดเหตุการณ์ E
- $P(E|H)$ คือ ค่าความน่าจะเป็นที่จะเกิดเหตุการณ์ E เมื่อเกิดเหตุการณ์ H
- $P(H)$ คือ ค่าความน่าจะเป็นก่อนหน้าของสมมติฐาน H
- $P(E)$ คือ ค่าความน่าจะเป็นก่อนหน้าของชุดข้อมูลตัวอย่าง E

2.6.6 การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression)

เป็นเทคนิคที่ใช้ในการหาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ แล้วนำสมการที่ได้จากความสัมพันธ์เหล่านั้นมาสร้างเป็นแบบจำลอง หลังจากนั้นนำแบบจำลองมาใช้ในการพยากรณ์ต่อไป ประเภทของการถดถอยแบบโลจิสติกมี 2 ประเภท คือ

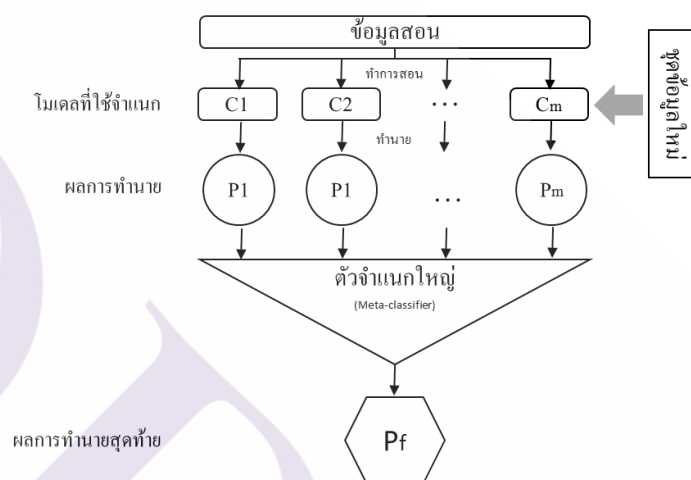
1. Binary Logistic โดยวิธีนี้จะใช้เมื่อตัวแปรตามเป็นตัวแปรเชิงกลุ่มที่มีค่าได้เพียง 2 ค่า
2. Multinomial Logistic โดยวิธีนี้จะใช้เมื่อตัวแปรตามเป็นตัวแปรเชิงกลุ่มที่มีค่ามากกว่า 2 ค่าขึ้นไป

2.6.7 การจำแนกแบบกลุ่มด้วยวิธี Stacking

วิธี Stacking เป็นการจำแนกแบบรวมกลุ่มประเภทหนึ่ง ถูกเสนอโดย David H. Wolpert (1992) เป็นวิธีที่ใช้แบบจำลองหลัก หรือตัวจำแนกใหญ่ (Meta-classifier) ในการรวมผลลัพธ์ของแบบจำลองระดับย่อยที่ใช้จำแนกต่างๆ เพื่อให้ได้ค่าความถูกต้องที่สูงขึ้น

โดยวิธีการสแตกกิ่งจะทำการรวบรวมผลการทำนายของแต่ละแบบจำลองย่อยเพื่อนำมาสร้างชุดข้อมูลใหม่ ซึ่งบ่อยครั้งที่วิธีการนี้จะสร้างผลการทำนายได้ดีกว่าแบบจำลองระดับย่อยที่ใช้ อัลกอริทึมเพียงตัวเดียว เนื่องจากวิธีการสแตกกิ่งจะทำการเลือกจุดเด่นของแบบจำลองที่มีประสิทธิภาพ ขณะเดียวกันก็จะไม่ให้ความสำคัญกับแบบจำลองที่ไม่มีประสิทธิภาพ ดังภาพที่

2.6.7



ภาพที่ 2.4 แผนผังการวิธีการสแตกกิ่ง

ซึ่งจากวิธีการของสแตกกิ่ง หากแบบจำลองระดับย่อยที่ถูกนำมาใช้มีความแตกต่างกันอย่างมีนัยสำคัญก็จะทำให้ตัวแบบจำลองหลักมีประสิทธิภาพที่ดีด้วย

2.7 การเลือกคุณลักษณะ (Feature Selection)

บ่อยครั้งที่การทำกรจำแนกประเภทข้อมูลมักที่จะเจอปัญหาเกี่ยวกับจำนวนคุณลักษณะหรือตัวแปร (feature) ที่มีจำนวนมาก ซึ่งการวิเคราะห์ข้อมูลที่มีคุณลักษณะจำนวนมากนั้น ถือเป็นเรื่องยากและมีความซับซ้อนสูง และขั้นตอนสำหรับการเรียนรู้ที่จะนำมาใช้จัดการกับข้อมูลเหล่านี้จะต้องใช้ทรัพยากรทั้งในด้านการคำนวณและการใช้หน่วยความจำจำนวนมาก ในขณะที่ประสิทธิภาพของกระบวนการเรียนรู้อาจลดลงเนื่องจากอาจมีสัญญาณรบกวน (Noise) ในข้อมูลที่เกิดจากคุณลักษณะที่ไม่เกี่ยวข้องกับการวิเคราะห์ข้อมูลนั้นๆ ดังนั้น เพื่อแก้ปัญหาดังกล่าวคุณลักษณะของข้อมูลเหล่านั้นจะต้องถูกลดจำนวนลงด้วยวิธีการเลือกคุณลักษณะที่สำคัญสำหรับการวิเคราะห์ข้อมูล (Feature Selection) ซึ่งเป็นกระบวนการลดจำนวนคุณลักษณะลง เพื่อให้เหลือ

เพียงแคคุณลักษณะที่มีความเกี่ยวข้องกับการจำแนกประเภท ซึ่งจะเป็นการช่วยปรับปรุงประสิทธิภาพในการจำแนกประเภทให้ดียิ่งขึ้น และยังสามารถทำให้การประมวลผลเป็นไปได้อย่างรวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น โดยการเลือกคุณลักษณะสามารถแบ่งประเภทตามวิธีการในการเลือกคุณลักษณะได้เป็น 3 ประเภท ได้แก่ Filter, Wrapper และ Embedded (Molina et al., 2002)

การเลือกคุณลักษณะแบบ Embedded

เป็นวิธีที่การเลือกคุณลักษณะเป็นส่วนหนึ่งในกระบวนการจำแนกประเภท ซึ่งจะทำการเลือกคุณลักษณะที่เหมาะสมในขั้นตอนของการเรียนรู้ โดยขั้นตอนวิธีเหล่านี้จะมีการเลือกคุณลักษณะที่เหมาะสมสำหรับการสร้างแบบจำลองในการแก้ปัญหาต่างๆ โดยไม่จำเป็นต้องเพิ่มขั้นตอนวิธีในการเลือกคุณลักษณะอื่นๆ เข้ามาช่วย ตัวอย่างของขั้นตอนวิธีสำหรับการเรียนรู้ที่มีการเลือกคุณลักษณะแบบ Embedded ได้แก่ ต้นไม้ตัดสินใจ และ กำหนดการพันธุกรรม

การเลือกคุณลักษณะแบบ Wrapper

เป็นวิธีที่ใช้การจำแนกประเภทเพื่อคัดเลือกเซตย่อยของคุณลักษณะทั้งหมดของข้อมูล โดยจะทำการเลือกเซตย่อยของคุณลักษณะที่มีความแม่นยำ และเหมาะสมกับขั้นตอนวิธีการเรียนรู้แบบใดแบบหนึ่งโดยเฉพาะ ทำให้วิธีการเลือกคุณลักษณะแบบนี้ถือได้ว่าเป็นการเพิ่มประสิทธิภาพของขั้นตอนวิธีการเรียนรู้ได้ดีที่สุด แต่เนื่องจากการที่จะต้องคัดเลือกเซตย่อยของคุณลักษณะทั้งหมดของข้อมูล ทำให้วิธีการนี้จะต้องใช้เวลาในการเรียนรู้และประมวลผลที่ค่อนข้างมาก นอกจากนี้เซตย่อยของคุณลักษณะที่เลือกได้จะเหมาะกับวิธีการเรียนรู้แบบหนึ่ง ซึ่งอาจจะไม่เหมาะกับวิธีการเรียนรู้แบบอื่นๆ

การเลือกคุณลักษณะแบบ Filter

เป็นการคัดเลือกคุณลักษณะโดยประเมินความเกี่ยวข้อง หรือความสำคัญของคุณลักษณะแต่ละตัวว่ามีความเหมาะสมกับการวิเคราะห์ข้อมูลมากน้อยเพียงใด โดยไม่ขึ้นกับขั้นตอนวิธีการเรียนรู้แบบใดแบบหนึ่ง การเลือกคุณลักษณะแบบนี้จะทำการจัดลำดับ (Ranking) ตามความเกี่ยวข้อง หรือความสำคัญของคุณลักษณะแต่ละตัว และเลือกคุณลักษณะที่มีค่าความสำคัญสูงสุด

การเลือกคุณลักษณะแบบ Wrapper และ Embedded ถือว่ามีความซับซ้อนในการคำนวณสูงกว่าแบบ Filter และนอกจากนี้ทั้ง 2 แบบยังมีแนวทางการคัดเลือกเซตย่อยของตัวแปรที่มีความเฉพาะเจาะจงกับวิธีการเรียนรู้ ซึ่งจะทำให้เกิดปัญหา Overfitting ตามมา กล่าวคือสามารถเรียนรู้จากตัวได้เป็นอย่างดี แต่เมื่อนำไปใช้กับข้อมูลแบบอื่นกลับได้ผลที่ไม่ดี นั่นจึงทำให้วิธีการเลือกคุณลักษณะแบบ Filter จึงมักจะถูกนำมาใช้

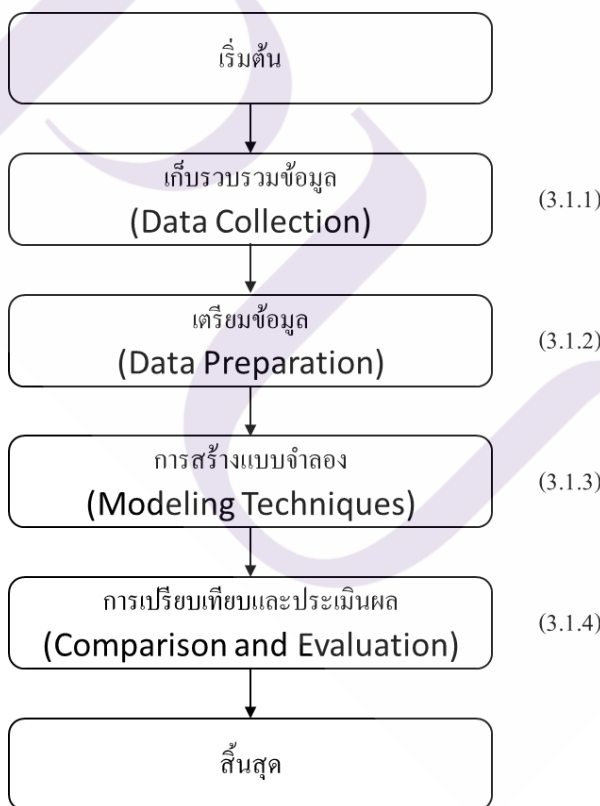
บทที่ 3

ระเบียบวิธีวิจัย

การวิจัยเรื่องการจำแนกบุคคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงานด้วยวิธีการจำแนกแบบหลายเวลาเบลนั้นก็เพื่อมุ่งหวังว่าสามารถช่วยให้องค์กรสามารถสรรหาพนักงานได้ตรงกับประเภทงานแต่ละงานได้อย่างเหมาะสมที่สุด โดยขั้นตอนในการศึกษาวิจัยแบ่งขั้นตอนออกเป็นดังนี้

3.1 แนวทางการวิจัย

แนวทางการวิจัยมีขั้นตอนดังภาพที่ 3.1



ภาพที่ 3.1 แผนผังการดำเนินงานโดยรวมของระบบ

3.1.1 วิธีการเก็บรวบรวมข้อมูล (Data Collection)

1. ทำการเก็บแบบสอบถามจากพนักงานในบริษัทด้านการเงินแห่งหนึ่งของไทย เพื่อวิเคราะห์ว่าพนักงานแต่ละคนมีบุคลิกภาพแบบใด โดยมีผู้ร่วมตอบแบบสอบถามทั้งสิ้น 2,541 คน โดยแบ่งเป็นระดับผู้จัดการขึ้นไป 636 คน และพนักงานในระดับปฏิบัติการ 1,905 คน

2. รวบรวมข้อมูลผลการปฏิบัติงาน 3 ปีย้อนหลังของพนักงานทั้งหมดจากฐานข้อมูลของบริษัท ซึ่งมีทั้งหมด 4 เรื่อง ได้แก่ ยอดขายสินค้า A (ยอดขายเป็นจำนวน), ยอดขายสินค้า B (ยอดขายเป็นจำนวน), คุณภาพการให้บริการ (คะแนนเต็ม 10 คะแนน) และการประเมินผล 360 องศา (คะแนนเต็ม 100 คะแนน)

3.1.2 การเตรียมข้อมูล (Data Preparation)

การรวมข้อมูล

1. ทำการรวมข้อมูล (Join) ระหว่างข้อมูลผลการปฏิบัติงานในด้านต่างๆ ทั้ง 4 ด้าน โดยใช้รหัสพนักงานเป็นตัวเชื่อม (Key)

2. ทำการทำความสะอาดข้อมูล (Cleansing Data) จากแบบสอบถาม โดยจะทำการตัดข้อมูลที่ผิดปกติออกไป เช่น กรอกรหัสพนักงานไม่ครบถ้วน, คำตอบในข้อคำถามแต่ละข้อเหมือนกันหมด (เช่น ตอบ 5 ทั้งหมดทุกข้อ) และรหัสพนักงานซ้ำ เป็นต้น หลังจากทำความสะอาดข้อมูลแล้วพบว่าข้อมูลของผู้ตอบแบบสอบถามเหลืออยู่ทั้งสิ้น 2,317 คน โดยแบ่งเป็นระดับผู้จัดการขึ้นไป 564 คน และพนักงานในระดับปฏิบัติการ 1,753 คน

3. ทำการรวมข้อมูล (Join) ระหว่างข้อมูลของแบบสอบถาม และข้อมูลจากฐานข้อมูลที่ได้ทำการรวบรวมมาแล้วก่อนหน้านี้ โดยใช้รหัสพนักงานเป็นตัวเชื่อม (Key)

4. หลังจากนั้นแปลงค่าผลการปฏิบัติงานให้อยู่ในรูปแบบไบนารี (Binary) โดยคะแนนของผลการปฏิบัติงานที่ต่ำกว่าค่าเฉลี่ยของผลการปฏิบัติงานในแต่ละด้านนั้น จะมีค่าเป็น 0 แต่ถ้าหากมากกว่าค่าเฉลี่ยในแต่ละด้านจะมีค่าเป็น 1 โดยจะทำแบบเดียวกันนี้กับทุกผลการปฏิบัติงานที่รวบรวมได้

การกรองข้อมูล

ทำการคัดเลือกลเฉพาะพนักงานที่มีอายุงานตั้งแต่ 1 ปีขึ้นไป และมีผลการปฏิบัติงานเก็บไว้ในฐานข้อมูลเท่านั้น

การแบ่งข้อมูล

สำหรับการแบ่งข้อมูลในการวิจัยนี้จะใช้วิธี Split Test โดยจะทำการสุ่มข้อมูลที่ใช้สอนในอัตรา 70% ของข้อมูลทั้งหมด และทำการสุ่มข้อมูลที่ใช้ทดสอบในอัตรา 30% ของข้อมูลทั้งหมด โดยทั้งข้อมูลสอนและข้อมูลทดสอบจะมีสัดส่วนของตำแหน่งงานเดียวกับชุดข้อมูลทั้งหมด

3.1.3 การสร้างแบบจำลอง (Modeling Techniques)

สำหรับการสร้างแบบจำลองในงานวิจัยนี้ จะมีการสร้างแบบจำลองทั้งหมด 3 รูปแบบ ได้แก่

3.1.3.1 แบบที่ 1: สร้างการจำแนกแบบหลายลาเบลโดยแยกรายอัลกอริทึม

เมื่อได้ข้อมูลทั้งหมดที่ต้องการมาแล้ว ก็จะเข้าสู่กระบวนการสร้างการจำแนกแบบหลายลาเบลโดยแยกรายอัลกอริทึม โดยในงานวิจัยนี้จะใช้วิธีการจำแนกหลายลาเบล 3 แบบ ได้แก่ Binary Relevance (BR), Label Powerset (LP) และ Classifier Chains(CC) และจะใช้อัลกอริทึมทั้งหมด 6 ตัว เพื่อสร้างแบบจำลองทำนาย ได้แก่

Logistic Regression ไม่มีการกำหนดค่าตัวแปร

Decision Tree มีการปรับค่า Criterion ให้เป็น Entropy และ Max Depth เท่ากับ 1

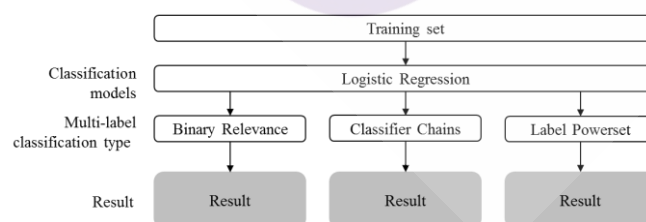
Random Forest มีการกำหนดค่า Random State ให้เท่ากับ 1

K-Nearest Neighbor มีการกำหนดค่า k เท่ากับ 5

Support Vector Machine แบบ Linear

Naïve Bayes ไม่มีการกำหนดค่าตัวแปร

โดยจะเริ่มต้นจากการแบ่งข้อมูลออกเป็น 2 ส่วน โดยส่วนแรกเป็นข้อมูลเพื่อนำมาสอนคิดเป็น 70% และอีกส่วนเป็นข้อมูลสำหรับใช้ทดสอบคิดเป็น 30% หลังจากนั้นก็จะเอาข้อมูลส่วนที่จะนำมาสอนมาเข้าโมเดล ซึ่งจะทดสอบด้วยอัลกอริทึมทั้งหมด 6 อัลกอริทึม โดยแต่ละอัลกอริทึมจะมีผลลัพธ์ออกมา 3 ผลลัพธ์ ซึ่งได้จากวิธีการจำแนกหลายลาเบล 3 แบบที่ได้กล่าวไปข้างต้น ตามภาพที่ 3.2



ภาพที่ 3.2 แผนผังการสร้างชุดข้อมูลใหม่สำหรับการทดลองแบบที่ 1

หลังจากนั้นก็นำข้อมูลใหม่ที่ได้ไปรวมกับผลการปฏิบัติงานที่ได้แปลงข้อมูลให้เป็นไบนารี แล้วจึงไปข้อมูลทั้งหมดไปทำนายผลการปฏิบัติงาน

3.1.3.2 แบบที่ 2: นำวิธีการจากแบบที่ 1 มาประยุกต์ใช้กับวิธีการสแตกกิ่ง

นำวิธีการแบบที่ 1 มาประยุกต์ใช้กับวิธีการจำแนกแบบสแตกกิ่ง โดยจะใช้ตัวจำแนกใหญ่ (Meta-classifier) ด้วยอัลกอริทึมทั้งหมด 6 ตัว เพื่อที่จะสามารถนำมาเปรียบเทียบผลลัพธ์ โดยใช้อัลกอริทึม ดังนี้

Logistic Regression ไม่มีการกำหนดค่าตัวแปร

Decision Tree มีการปรับค่า Criterion ให้เป็น Entropy และ Max Depth เท่ากับ 1

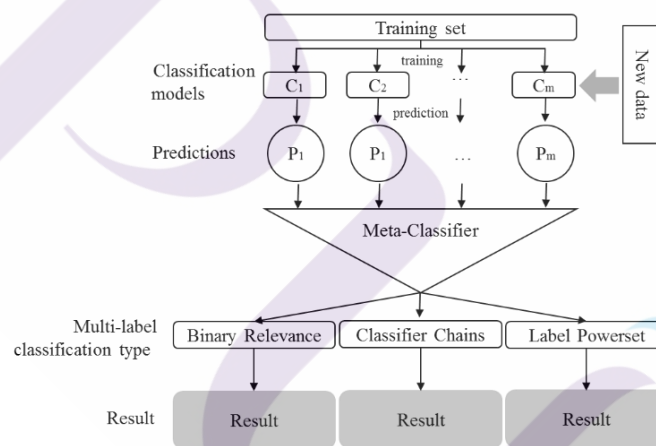
Random Forest มีการกำหนดค่า Random State ให้เท่ากับ 1

K-Nearest Neighbor มีการกำหนดค่า k เท่ากับ 5

Support Vector Machine แบบ Linear

Naïve Bayes ไม่มีการกำหนดค่าตัวแปร

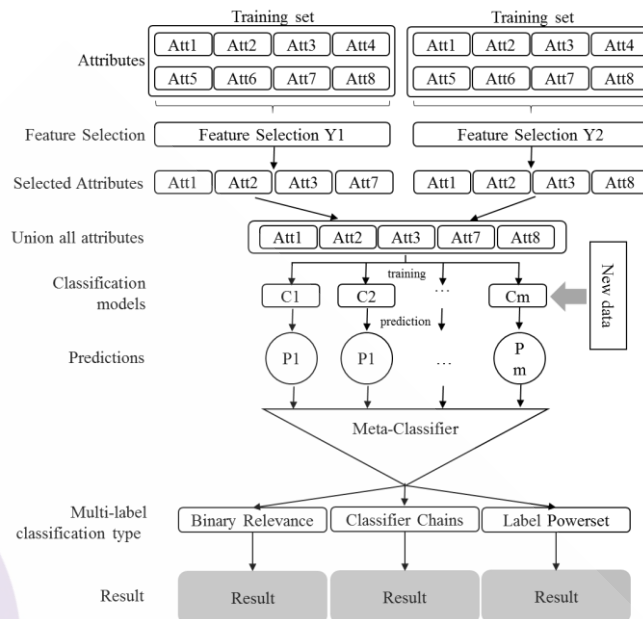
โดยกระบวนการทำงานของการทดลองแบบที่ 2 จะเป็นตามภาพที่ 3.3



ภาพที่ 3.3 แผนผังการสร้างชุดข้อมูลใหม่สำหรับการทดลองแบบที่ 2

3.1.3.3 แบบที่ 3: นำข้อมูลจากแบบที่ 2 มาทำการคัดเลือกคุณลักษณะแบบรวม แล้วใช้วิธีการสแตกกิ่ง

โดยแบบรวม (Union) จะเป็นการหาคุณลักษณะที่สำคัญด้วยวิธีการขจัดออก (Backward Elimination) ที่สามารถทำนายผลการปฏิบัติงานในด้านต่างๆ ได้ หลังจากนั้นก็จะทำการเลือกคุณลักษณะของแต่ละลาเบลมารวมกัน ดังภาพที่ 3.4



ภาพที่ 3.4 แผนผังการสร้างชุดข้อมูลใหม่สำหรับการทดลองแบบที่ 3

หลังจากนั้นก็นำคุณลักษณะทั้งหมดมาหาวิธีการจำแนกแบบเสตคกิ้ง โดยใช้วิธีการและอัลกอริทึมต่างๆ เหมือนกับในแบบที่ 2

3.1.6 การเปรียบเทียบและประเมินผล (Comparison and Evaluation)

การประเมินผลจะประเมินตามค่าการคำนวณทั้งสิ้น 5 ค่า ได้แก่ Hamming Loss, Accuracy, F1 score, Precision และ Recall

โดย Hamming Loss จะเป็นค่าที่คำนวณมาจากลาเบลที่จำแนกผิด ซึ่งค่ายิ่งน้อยยิ่งแสดงให้เห็นถึงประสิทธิภาพของการจำแนกประเภทที่ดียิ่งขึ้น

$$\text{Hamming Loss} = \frac{1}{m} \sum_{i=1}^m \frac{|Y_i \Delta Z_i|}{M}$$

Accuracy เป็นค่าที่คำนวณมาจากจำนวนลาเบลที่จำแนกได้ถูกต้องหารด้วยจำนวนการจำแนกทั้งหมด นั่นหมายความว่ายิ่งค่านี้มีค่าสูงเท่าใดก็หมายความว่าโมเดลนั้นมีการจำแนกประเภทได้อย่างแม่นยำมากขึ้นเท่านั้น

$$\text{Accuracy} = \frac{1}{m} \sum_{i=1}^m \frac{|Y_i \cap Z_i|}{|Y_i \cup Z_i|}$$

Precision เป็นสัดส่วนของจำนวนการจำแนกที่เป็น true positive ต่อจำนวนการจำแนกที่เราสนใจ หรือเรียกอีกอย่างหนึ่งว่าเป็นความแม่นยำที่สนใจแค่ในส่วนของการทำนาย

$$Precision = \frac{1}{m} \sum_{i=1}^m \frac{|Y_i \cap Z_i|}{|Z_i|}$$

Recall เป็นสัดส่วนของจำนวนการจำแนกที่เป็น true positive ต่อข้อมูลที่เกิดขึ้นจริง เพื่อดูว่าเมื่อเทียบกับการจำแนกที่เราสนใจแล้ว มีความแม่นยำมากน้อยแค่ไหนเมื่อเทียบกับข้อมูลจริง

$$Recall = \frac{1}{m} \sum_{i=1}^m \frac{|Y_i \cap Z_i|}{|Y_i|}$$

F-Measure เป็นการผสมระหว่าง Precision และ Recall เข้าด้วยกัน เพื่อใช้วัดความสามารถของโมเดล

$$F - Measure = \frac{1}{m} \sum_{i=1}^m \frac{2|Y_i \cap Z_i|}{|Y_i| + |Z_i|}$$

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 ภาษาไพทอน (Python)

ภาษาไพทอน เป็นภาษาที่ใช้สำหรับการเขียนโปรแกรม เป็นภาษาสคริปต์ที่สร้างโดย กิโด ฟาน รอสซัม (Guido van Rossum) ในพ.ศ. 2533 ซึ่งได้ออกแบบมาเพื่อให้ทำงานได้กับ Web Application ที่ลักษณะคล้ายกับภาษา Perl, PHP, JAVA และ ASP

เนื่องจากภาษาไพทอน เป็นภาษาที่ใหม่จึงมีคุณสมบัติที่ดีดังต่อไปนี้

1. สามารถใช้ได้ทุกแพลตฟอร์ม กล่าวคือ สามารถทำงานได้ทุก ๆ CPU หลาย ๆ ระบบปฏิบัติการ เพียงแต่ผู้เขียน โปรแกรมเขียนจากแพลตฟอร์มใด ๆ แล้วนำโปรแกรมที่ได้ไปให้ทำงานต่างแพลตฟอร์มกันได้
2. ไม่ต้องเสียค่าใช้จ่ายในการจัดซื้อโปรแกรมต้นฉบับ โดยปกติแล้วโปรแกรมภาษาทั่ว ๆ ไป จะต้องจัดซื้อโปรแกรมต้นฉบับเพื่อนำมาติดตั้งในราคาที่แพงมาก แต่โปรแกรมภาษาไพทอนสามารถดาวน์โหลดจาก www.python.org ได้โดยตรง แล้วนำมาติดตั้งและศึกษาการใช้ด้วยตนเอง เพราะเป็นโปรแกรมประเภท Open Source
3. ภาษาไพทอนได้นำเอาข้อดีของโปรแกรมในอดีตเข้ามาไว้ด้วยกัน เช่น ภาษา C, C++, Java และ Perl เป็นต้น

4. มีความปลอดภัยสูง เนื่องจากภาษาไพธอนทำงานอยู่ด้าน Server เป็นหลัก เมื่อมีการร้องขอจากเครื่อง Client จะประมวลผลที่เครื่อง Server ทำให้ผู้ใช้ทั่วไปไม่สามารถเข้าถึงเครื่อง Server ได้โดยตรงจึงมีความปลอดภัยสูงกว่า
5. ใช้ในการพัฒนา Web Service ซึ่งในปัจจุบันการพัฒนาซอฟต์แวร์ได้เน้นที่มีการแลกเปลี่ยนข้อมูลซึ่งกันและกันทั้งในองค์กรเดียวกันหรือแม้แต่ต่างองค์กรกัน ทำให้เกิดความ สะดวกสบาย ไม่ต้องใช้ซอฟต์แวร์อื่น ๆ มาแปลงข้อมูลเพื่อให้เข้ากันได้อีกต่อไปเรียนรู้ได้เร็วกว่าโปรแกรมภาษาอื่น ๆ เพราะมีโครงสร้างภาษาที่ไม่ซับซ้อน ซึ่งโครงสร้างภาษาคล้ายคลึงกับภาษา C ถ้าโปรแกรมเมอร์ที่เคยใช้ภาษา C มาก่อนจะทำให้เรียนรู้ได้เร็วยิ่งขึ้น นอกจากนี้การเขียนโปรแกรมด้วยภาษาไพธอนจะมีความกระชับและสั้นกว่าภาษาซี

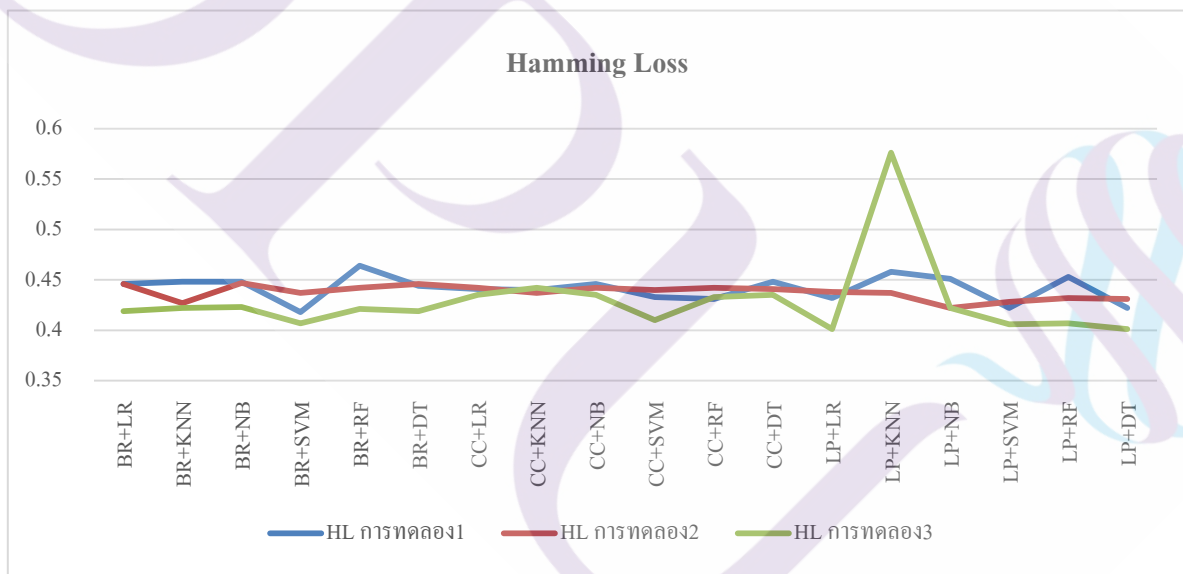


บทที่ 4

ผลการวิจัย

ในบทนี้จะนำเสนอถึงผลการวิจัยของการศึกษาการจำแนกบุคคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงานด้วยวิธีการจำแนกแบบหลายลาเบลในการทดลองแบบต่างๆ โดยผลสรุปจะประกอบไปด้วยค่าการคำนวณทั้งสิ้น 5 ค่า ได้แก่ Hamming Loss, Accuracy, F1 score, Precision และ Recall

4.1 ผลการทดสอบประสิทธิภาพและการอภิปรายผล



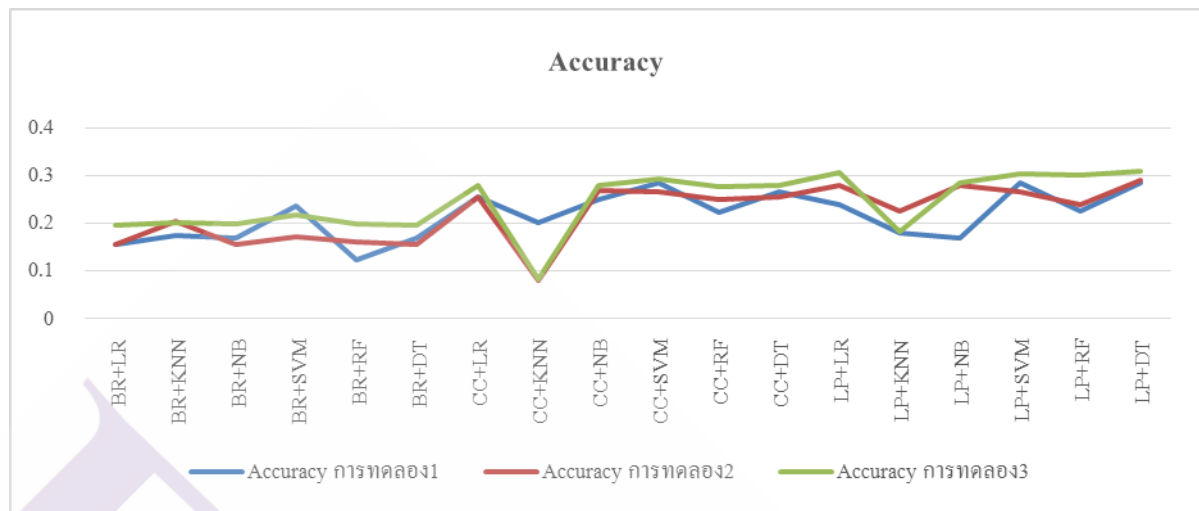
ภาพที่ 4.1 ภาพเปรียบเทียบค่า Hamming Loss ของการทดสอบในแต่ละรูปแบบ

ตารางที่ 4.1 ตารางเปรียบเทียบค่า Hamming Loss ของการทดสอบในแต่ละรูปแบบ

Algorithm	HL การทดลอง1	HL การทดลอง2	HL การทดลอง3
BR+LR	0.446	0.446	0.419
BR+KNN	0.448	0.427	0.422
BR+NB	0.448	0.447	0.423
BR+SVM	0.418	0.437	0.407
BR+RF	0.464	0.442	0.421
BR+DT	0.444	0.446	0.419
CC+LR	0.441	0.442	0.435
CC+KNN	0.44	0.437	0.442
CC+NB	0.446	0.442	0.435
CC+SVM	0.433	0.44	0.41
CC+RF	0.431	0.442	0.433
CC+DT	0.448	0.441	0.435
LP+LR	0.432	0.438	0.401
LP+KNN	0.458	0.437	0.576
LP+NB	0.451	0.422	0.422
LP+SVM	0.422	0.428	0.406
LP+RF	0.453	0.432	0.407
LP+DT	0.422	0.431	0.401

จากตารางที่ 4.1 พบว่า โดยภาพรวมแล้ว การใช้วิธีการแสดงคั้งผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) มีค่า Hamming Loss ที่ดีกว่า (ต่ำกว่า) การทดลองแบบแยกสายอัลกอริทึม หรือการทดลองที่ใช้แสดงคั้งเพียงอย่างเดียว

นอกจากนี้ยังพบว่า วิธีการจำแนกแบบหลายลาเบลแบบ Label Powerset โดยรวมแล้ว (หากไม่นับ LP+KNN) มีค่า Hamming Loss ที่ต่ำกว่าวิธีการจำแนกแบบหลายลาเบลวิธีอื่น โดยเฉพาะ LP+DT ที่มีค่า Hamming Loss ที่ต่ำที่สุดที่ 0.401



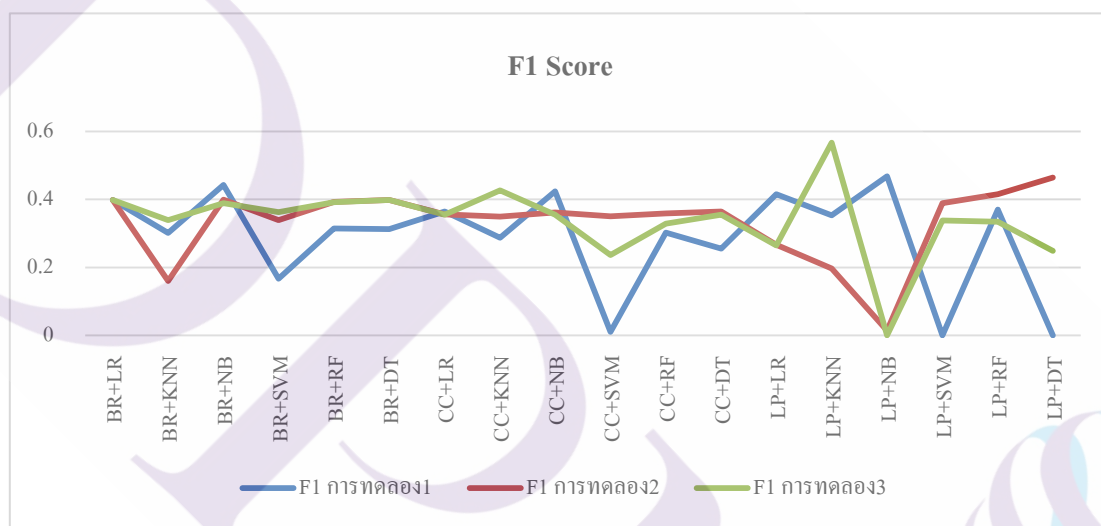
ภาพที่ 4.2 ภาพเปรียบเทียบค่า Accuracy ของการทดสอบในแต่ละรูปแบบ

ตารางที่ 4.2 ตารางเปรียบเทียบค่า Accuracy ของการทดสอบในแต่ละรูปแบบ

Algorithm	Accuracy การทดลอง1	Accuracy การทดลอง2	Accuracy การทดลอง3
BR+LR	0.155	0.155	0.196
BR+KNN	0.173	0.205	0.201
BR+NB	0.169	0.155	0.199
BR+SVM	0.235	0.171	0.217
BR+RF	0.122	0.16	0.198
BR+DT	0.168	0.155	0.196
CC+LR	0.254	0.256	0.28
CC+KNN	0.201	0.081	0.083
CC+NB	0.25	0.268	0.28
CC+SVM	0.284	0.265	0.293
CC+RF	0.222	0.25	0.277
CC+DT	0.266	0.254	0.28
LP+LR	0.238	0.279	0.307
LP+KNN	0.18	0.226	0.183
LP+NB	0.168	0.279	0.284
LP+SVM	0.284	0.265	0.303
LP+RF	0.226	0.238	0.302
LP+DT	0.284	0.291	0.31

จากตารางที่ 4.2 พบว่า การใช้วิธีการแสดงถึงผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) มีค่า Accuracy ที่ดีกว่า (สูงกว่า) การทดลองแบบแยกย่อยอัลกอริทึมหรือการทดลองที่ใช้แสดงถึงเพียงอย่างเดียว โดยมีค่าสูงกว่าในทุกการจำแนกแบบหลายลาเบลและทุกอัลกอริทึม ยกเว้นการประยุกต์ใช้กับ KNN และ BR+SVM

นอกจากนี้ยังพบว่า ค่า Accuracy ที่สูงที่สุดอันดับต้นๆ จากการทดลองทั้งหมดจะได้จากวิธีการจำแนกแบบหลายลาเบลแบบ Label Powerset โดย LP+DT จะมีค่า Accuracy ที่สูงที่สุดที่ 0.31



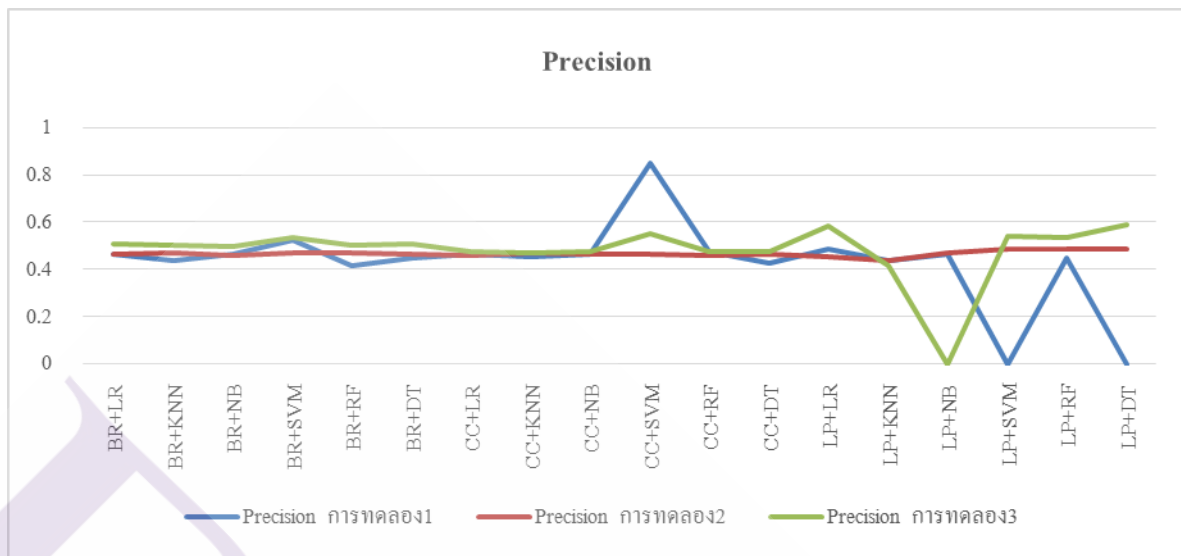
ภาพที่ 4.3 ภาพเปรียบเทียบค่า F1 Score ของการทดสอบในแต่ละรูปแบบ

ตารางที่ 4.3 ตารางเปรียบเทียบค่า F1 Score ของการทดสอบในแต่ละรูปแบบ

Algorithm	F1 การทดลอง1	F1 การทดลอง2	F1 การทดลอง3
BR+LR	0.398	0.398	0.398
BR+KNN	0.301	0.16	0.339
BR+NB	0.443	0.399	0.389
BR+SVM	0.167	0.339	0.363
BR+RF	0.315	0.393	0.392
BR+DT	0.313	0.398	0.398
CC+LR	0.364	0.356	0.355
CC+KNN	0.287	0.349	0.427
CC+NB	0.424	0.362	0.355
CC+SVM	0.01	0.35	0.236
CC+RF	0.302	0.359	0.329
CC+DT	0.255	0.364	0.355
LP+LR	0.415	0.267	0.265
LP+KNN	0.353	0.197	0.567
LP+NB	0.468	0.014	0
LP+SVM	0	0.389	0.338
LP+RF	0.37	0.415	0.334
LP+DT	0	0.464	0.249

จากตารางที่ 4.3 แม้ว่าภาพรวมคะแนนในแต่ละการทดลองจะค่อนข้างมีความผันผวนของค่า F1 Score จากการจำแนกแบบหลายลาเบลเมื่อเปลี่ยนอัลกอริทึม และไม่มีกรทดลองหรือการจำแนกแบบหลายลาเบลใด ที่ชนะค่าอื่นทุกครั้ง แต่จะเห็นว่า 3 อันดับแรกที่มีค่า F1 Score สูงที่สุดอยู่ในการทดลองที่ใช้การจำแนกแบบหลายลาเบลแบบ Label Powerset

โดยอันดับหนึ่งได้แก่ LP+KNN จากการใช้วิธีการแสดงคั้งผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) ที่มีค่า F1 Score อยู่ที่ 0.567 ตามมาด้วย LP+NB จากการทดลองแบบแยกรายอัลกอริทึม (การทดลอง 1) โดยมีค่า F1 Score อยู่ที่ 0.468 และ LP+DT จากการทดลองที่ใช้แสดงคั้งเพียงอย่างเดียว (การทดลอง 2) โดยมีค่า F1 Score อยู่ที่ 0.464



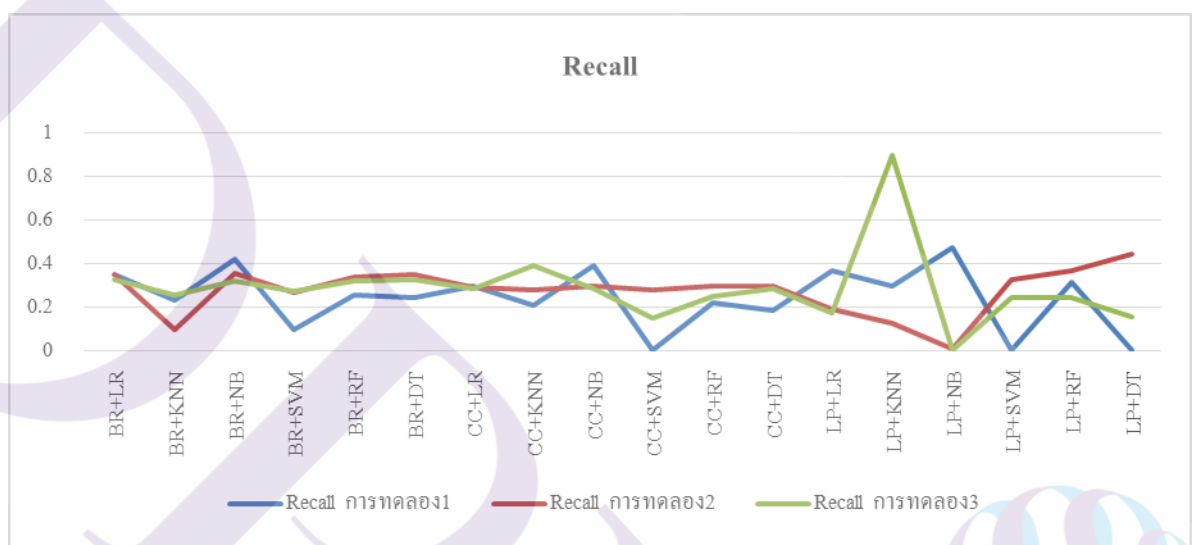
ภาพที่ 4.4 ภาพเปรียบเทียบค่า Precision ของการทดสอบในแต่ละรูปแบบ

ตารางที่ 4.4 ตารางเปรียบเทียบค่า Precision ของการทดสอบในแต่ละรูปแบบ

Algorithm	Precision การทดลอง1	Precision การทดลอง2	Precision การทดลอง3
BR+LR	0.462	0.462	0.505
BR+KNN	0.439	0.469	0.5
BR+NB	0.466	0.46	0.497
BR+SVM	0.522	0.468	0.533
BR+RF	0.417	0.467	0.502
BR+DT	0.449	0.462	0.505
CC+LR	0.464	0.461	0.474
CC+KNN	0.453	0.47	0.47
CC+NB	0.465	0.462	0.474
CC+SVM	0.85	0.464	0.55
CC+RF	0.476	0.461	0.473
CC+DT	0.426	0.464	0.474
LP+LR	0.483	0.453	0.582
LP+KNN	0.437	0.436	0.415
LP+NB	0.465	0.467	0
LP+SVM	0	0.488	0.54
LP+RF	0.447	0.483	0.537
LP+DT	0	0.487	0.59

จากตารางที่ 4.4 โดยส่วนใหญ่ ค่า Precision ของที่ได้จากการจำแนกในแต่ละแบบ ของแต่ละการทดลองจะมีค่าที่เกาะกลุ่มกันอยู่ระหว่าง 0.4 ถึง 0.5 อย่างไรก็ตามค่า Precision ของ CC+SVM จากการทดลองแบบแยกรายอัลกอริทึม (การทดลอง 1) มีค่าที่สูงที่สุด โดยมีค่าอยู่ที่ 0.85 ซึ่งถือว่าสูงกว่าค่าของการจำแนกอื่น และการทดลองอื่นค่อนข้างมาก

และเมื่อสังเกตข้อมูลโดยดูจากวิธีการจำแนกแบบหลายลาเบล ก็พบว่าผลที่ได้จากการใช้วิธีการแบบ Label Powerset มีหลายค่าที่มีค่าน้อยกว่าค่าเฉลี่ยโดยรวม



ภาพที่ 4.5 ภาพเปรียบเทียบค่า Recall ของการทดสอบในแต่ละรูปแบบ

ตารางที่ 4.5 ตารางเปรียบเทียบค่า Recall ของการทดสอบในแต่ละรูปแบบ

Algorithm	Recall การทดลอง1	Recall การทดลอง2	Recall การทดลอง3
BR+LR	0.349	0.349	0.328
BR+KNN	0.229	0.096	0.256
BR+NB	0.422	0.353	0.32
BR+SVM	0.099	0.266	0.275
BR+RF	0.253	0.339	0.322
BR+DT	0.241	0.349	0.328
CC+LR	0.299	0.29	0.283
CC+KNN	0.21	0.277	0.391
CC+NB	0.39	0.297	0.283
CC+SVM	0.005	0.28	0.151
CC+RF	0.221	0.294	0.252
CC+DT	0.182	0.299	0.283
LP+LR	0.364	0.189	0.172
LP+KNN	0.296	0.128	0.893
LP+NB	0.471	0.007	0
LP+SVM	0	0.323	0.246
LP+RF	0.316	0.364	0.243
LP+DT	0	0.442	0.158

แม้ว่าการจำแนกแบบ CC+SVM จากการทดลองแบบแยกรายอัลกอริทึม (การทดลอง 1) จะมีค่า Precision ที่สูงกว่าการจำแนก และการทดลองแบบอื่นเป็นอย่างมาก แต่จากตารางที่ 4.5 พบว่า ค่า Recall ของ CC+SVM ของการทดลอง 1 นั้นมีค่าที่น้อยมากๆ อยู่ที่ 0.005

โดยเมื่อเปรียบเทียบในส่วนของค่า Recall นั้นพบว่า LP+KNN จากการใช้วิธีการแสดงก็่งผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) นั้นมีค่า Recall ที่สูงที่สุด อยู่ที่ 0.893

ตารางที่ 4.6 เปรียบเทียบความสัมพันธ์ของแต่ละลาเบล

	Product1	Product2	Product3	Product4
Product1	1	0.405232	0.383818	0.377559
Product2	0.405232	1	0.210349	0.314735
Product3	0.383818	0.210349	1	0.352494
Product4	0.377559	0.314735	0.352494	1

เมื่อทำการหาความสัมพันธ์ระหว่างลาเบลแต่ละตัว พบว่าส่วนใหญ่ค่าของลาเบล มีสหสัมพันธ์กับในระดับต่ำถึงปานกลาง ซึ่งนั่นหมายถึง ข้อมูลของลาเบลแต่ละตัวไม่ได้รับอิทธิพล ซึ่งกันและกันโดยตรง

บทที่ 5

สรุปผลและข้อเสนอแนะ

ในบทนี้เป็นการสรุปผลที่ได้จากการวิจัยในเรื่องของการจำแนกบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงานด้วยวิธีการจำแนกแบบหลายลาเบล ตลอดจนข้อเสนอแนะสำหรับการพัฒนาหรือต่อยอดการวิจัยที่เกี่ยวข้องอื่นๆ ในอนาคต

5.1 สรุปผลการวิจัย

การวิจัยเรื่องการจำแนกบุคลิกภาพที่ส่งผลต่อผลการปฏิบัติงานของพนักงานด้วยวิธีการจำแนกแบบหลายลาเบล เป็นการวิจัยที่ต้องการศึกษาความแตกต่างของผลการปฏิบัติงานของพนักงานที่มีบุคลิกภาพแตกต่างกันไป ซึ่งเป็นการผสมความรู้จากหลายศาสตร์ ไม่ว่าจะเป็นจิตวิทยาและการใช้การเรียนรู้ของเครื่องจักรมาประยุกต์ใช้

โดยผลจากการวิจัยนี้พบว่า ค่าที่ได้จากการคำนวณจากทั้ง 5 ค่า ไม่ว่าจะเป็น Hamming Loss, Accuracy, F1 score, Precision หรือ Recall ของการจำแนกประเภทแต่ละแบบ มีความแตกต่างกันไป ไม่มีการจับคู่ระหว่างวิธีการจำแนกแบบหลายลาเบล กับอัลกอริทึม และแบบการทดลองใด ที่สมบูรณ์แบบ หรือมีค่าทั้ง 5 ค่าดีกว่ารูปแบบอื่นไปทั้งหมด

โดยการวิจัยพบว่า การจำแนกแบบ LP+DT จากการใช้วิธีการแสดงคั้งผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) มีค่า Hamming Loss ต่ำที่สุด ที่ 0.401 และมีค่า Accuracy ที่สูงที่สุด ที่ 0.31

เมื่อพิจารณาค่า F1 Score พบว่า LP+KNN จากการใช้วิธีการแสดงคั้งผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) ที่มีค่า F1 Score ที่สูงที่สุด อยู่ที่ 0.567

ในขณะที่ CC+SVM จากการทดลองแบบแยกรายอัลกอริทึม (การทดลอง 1) มีค่า Precision ที่สูงที่สุด โดยมีค่าอยู่ที่ 0.85 แต่กลับมีค่า Recall ที่ต่ำมากที่สุดที่ 0.005

ซึ่งการจำแนกที่มีค่า Recall สูงที่สุด คือการจำแนกแบบ LP+KNN จากการใช้วิธีการแสดงคั้งผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) โดยมีค่า Recall อยู่ที่ 0.893

อย่างไรก็ตาม แม้ว่าจะไม่มีการจำแนกรูปแบบใด หรือการทดลองใดที่ดีที่สุด แต่ก็จะพบว่า วิธีการจำแนกแบบหลายลาเบลแบบ Label Powerset โดยส่วนใหญ่จะให้ผลลัพธ์ที่ค่อนข้างดีกว่า ซึ่งจะเห็นได้จาก 4 ใน 5 ของการจำแนกที่ได้ค่าที่ดีที่สุดจากการคำนวณทั้ง 5 คำนับ เป็นวิธีการที่มาจากวิธีการจำแนกแบบหลายลาเบลแบบ Label Powerset

นอกจากนี้ การใช้วิธีการแสดงคั้งผสมกับการคัดเลือกคุณลักษณะแบบรวมคุณลักษณะ (การทดลอง 3) ก็ให้ผลลัพธ์ที่ดีกว่าวิธีการทดลองแบบแยกรายอัลกอริทึม (การทดลอง 1) หรือการทดลองที่ใช้แสดงคั้งเพียงอย่างเดียว (การทดลอง 2) โดยจะเห็นได้จาก 4 ใน 5 ของการจำแนกที่ได้ค่าที่ดีที่สุดจากการคำนวณทั้ง 5 คำนับ มาจากการทดลอง 3

5.2 ข้อจำกัดและแนวทางการแก้ไขของงานวิจัย

ข้อจำกัดและแนวทางแก้ไขของงานวิจัยนี้ สามารถแบ่งออกได้เป็น 3 ข้อดังนี้

5.2.1 แหล่งข้อมูลที่เฉพาะเจาะจงเกินไป โดยการวิจัยนี้ได้ทำการศึกษากลุ่มตัวอย่างจากเพียงบริษัทเดียวเท่านั้น ซึ่งอาจจะส่งผลให้เกิดอคติจากข้อมูลได้ จากทั้งเหตุผลในแง่ของอุตสาหกรรมที่บริษัทนั้นทำอยู่ ซึ่งแม้ว่าจะเป็นอาชีพเดียวกัน แต่ในแต่ละอุตสาหกรรมก็มีความต้องการลักษณะของพนักงานที่แตกต่างกันไป ยกตัวอย่างเช่น พนักงานขายที่อยู่ในธุรกิจเครื่องสำอาง หรือของฟุ่มเฟือย ก็อาจจะมีบุคลิกภาพและมีวิธีการวัดผลการปฏิบัติงานที่แตกต่างไปจากพนักงานขายเครื่องจักรอุตสาหกรรม เป็นต้น

และแม้ว่าจะเป็นบริษัทที่อยู่ในอุตสาหกรรมเดียวกัน ก็อาจจะมีความต้องการพนักงานที่มีลักษณะบุคลิกภาพและมีวิธีการวัดผลการปฏิบัติงานที่แตกต่างกันได้เช่นเดียวกัน ไม่ว่าจะเป็นสาเหตุจากวัฒนธรรมการทำงานของแต่ละบริษัทที่แตกต่างกัน รวมไปถึงการคัดเลือก การฝึกอบรม หรือผลลัพธ์ที่ถูกคาดหวังที่อาจมีความแตกต่างกันไป

5.2.2 วิธีการจำแนก และการทดลองที่ยังไม่ครอบคลุม เนื่องจากการวิจัยนี้มีการใช้อัลกอริทึมทั้งหมด 6 ตัว (ตามที่ได้กล่าวไปแล้วในบทที่ 3) และมีวิธีการทดลองเพียงแค่ 3 แบบเท่านั้น ซึ่งอาจจะยังไม่ครอบคลุมวิธีการทั้งหมด การศึกษาในอนาคตจึงสามารถนำการวิจัยนี้ไปต่อยอดด้วยอัลกอริทึม หรือวิธีการทดลองแบบอื่น เพื่อพัฒนาและหาผลลัพธ์ที่ดีขึ้นได้ต่อไป

5.2.3 การวิจัยนี้ใช้ค่าพารามิเตอร์มาตรฐานที่มีในโปรแกรม Python ซึ่งผลลัพธ์ที่ได้อาจมีความแตกต่างจากค่ามาตรฐานของโปรแกรมอื่น นอกจากนี้การปรับค่าพารามิเตอร์เพิ่มเติมจากวิธีการที่นำเสนอไป ก็อาจจะทำให้ได้ผลลัพธ์ที่ดีขึ้นกว่าเดิมได้



บรรณานุกรม

บรรณานุกรม

ภาษาไทย

กล้าหาญ ณ น่าน (2557). อิทธิพลของคุณลักษณะบุคลิกภาพ และความพึงพอใจที่มีต่อการปรับตัว
ในการทำงานของผู้เข้าสู่ตลาดบัณฑิตใหม่. Journal of Management Sciences Vol.31

No.1 January-June 2014

เต็มศักดิ์ คทวณิช (2547). จิตวิทยาทั่วไป .กรุงเทพฯ : บริษัท ซีเอ็ดดูเคชั่น จำกัด (มหาชน)

วรัท พฤษภากุลนันท์ (2550). ประสิทธิภาพและประสิทธิผล (Efficiency & Effectiveness) , สืบค้น
13 มีนาคม 2562, จาก <http://peenet.blogspot.com/2008/07/efficiency-effectivenessadministrator.html>.

อนันท์ งามสะอาด. (2558). ประสิทธิภาพ (Efficiency) และ ประสิทธิผล (Effective) ต่างกันอย่างไร
, สืบค้น 13 มีนาคม 2562, จาก <http://www.kat.ac.th/index.php/81-2011-12-30-02-36-17/2584-28aaa>.

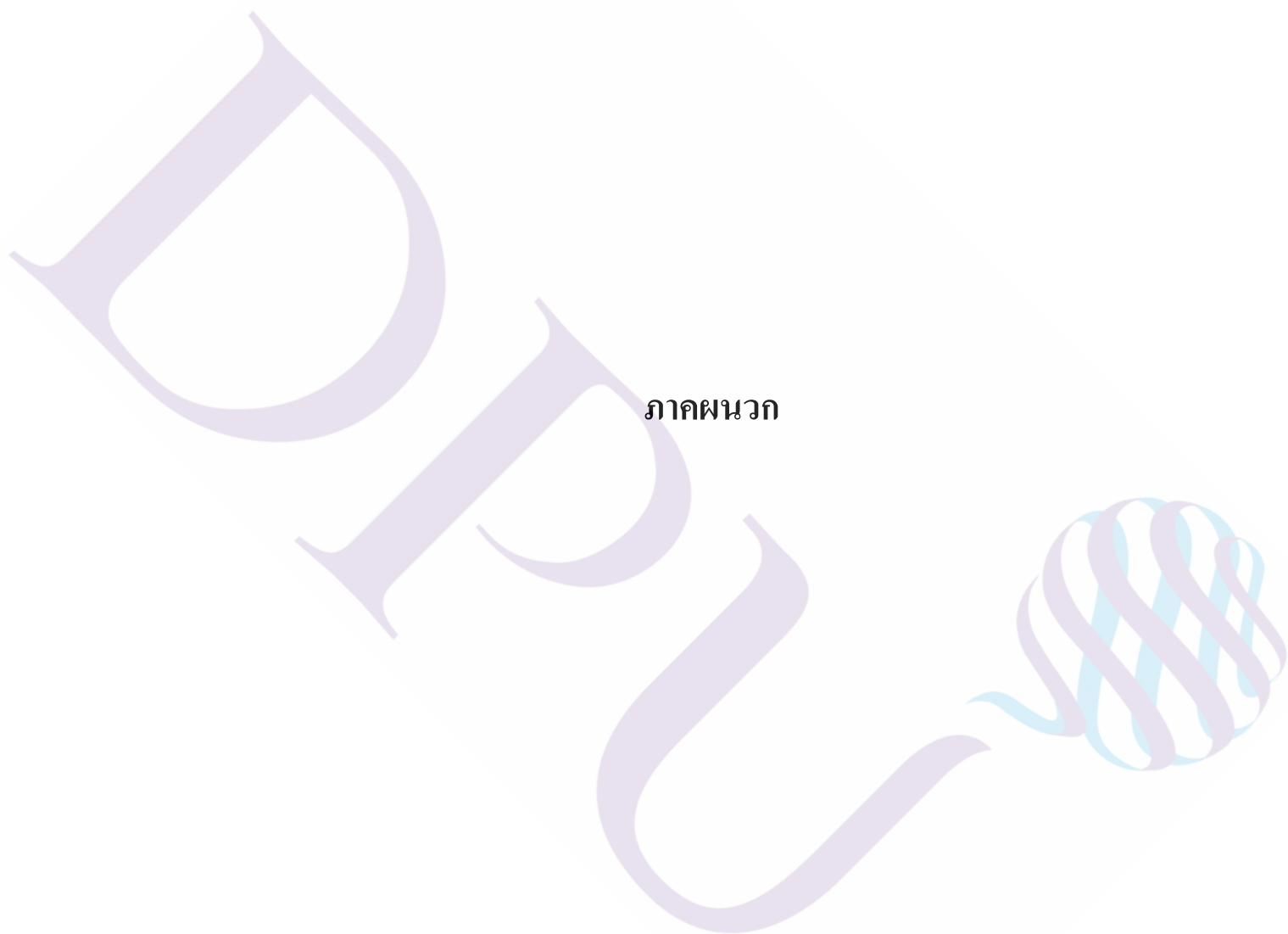
ภาษาต่างประเทศ

- Ahire, S.L., Golhar, D.Y. & Waller, M.A. (1996). *Development and validation of TQM implementation constructs*. *Decision Science*, 27
- Allport, G. W. (2009). *Personality: A psychology interpretation*. New York: Henry Holt.
- Bernard, H. W. (2009). *Mental health in the classroom*. New York: MacGraw-Hill Inc.
- Guilford, J. P. (2010). *Personality*. New York: MacGraw-Hill Inc.
- Gibson, J. L. & Others. (1988). *Organizational :Behavior, Sstructure, Process*. 3rd ed. Dallas; Texas : Business Publications, Inc.
- Gibson, G.R., Saveedra, J.M., MacFarlane, S., & MacFarlane, G.T. (1997). *Probiotics and intestinal infections*. *Applications and Practical Aspects*, pp. 10– 39.
- Goris AHC, Vermeeren MAP, Wouters EFM, Schols AMWJ & Westerterp KR. (2003). *Energy balance in depleted ambulatory patients with chronic obstructive pulmonary disease: the effect of physical activity and oral nutritional supplementation*. *Br J Nutr*, 89, pp. 725–729.
- Harrington, E. (1982). *The Twelve principles of Efficiency*. New York The Engineering Magazine, pp. 49.
- Hilgard, E. R. (2009). *Introduction to education*. New York: Harcourt Brace and World Inc.
- Hoy, W.K., & Miskel, C.G. (1991). *Educational administration: Theory research and Practice (4th ed.)*. New York: Random House.
- Slowikowski, M.K., 2005, *Using the DISC Behavioral Instrument to Guide Leadership and Communication*, *AORN Journal* Volume 82, Issue 5, November 2005
- Mathis, R.L., & Jackson, J.H. (2000). *Human resource management (9th ed.)*. Ohio: South Western College.
- Murray, J. L., Snider, B. R., & R. M. Midkiff, J. (1999). *The effects of training on resident assistant job performance*. *Journal of College Student Development*, 40(6). pp.744-747.
- Robbins, S.R. & Coulter, M. (1999). *Management (6th Ed.)*. New Jersey: Prentice Hall, Upper Saddle River.
- Robbins, S.P. (2003). *Organisational behaviour (10th ed)*. San Diego: Prentice Hall.

Steers, R.M. (1977). *Antecedents and Outcomes of Organizational Commitment*. Administrative Science Quarterly. 22(March 1977)

Yong, K. Y., Hwa, B. Y., Hyun, P. H., Hyang, Y. J., and Su, J. E. (2012). *The Effects of DISC Behavior Styles of Office Workers on Job Satisfaction, Organizational Commitment and Job Performance*. Korean J Occup Health Nurs. 2012 Aug;21(2):98-107





ภาคผนวก

Multi-Label Classification of Employee Job Performance Prediction by DISC Personality

Patompat Kamtar
Big Data Engineering Program,
College of Innovative Technology
and Engineering,
Dhurakij Pundit University
110/1-4 Pracha Chuen Rd,
Khwaeng Thung Song Hong, Khet
Lak Si, Bangkok 10210, Thailand
605162020019@dpu.ac.th

Duangjai Jitkongchuen
Big Data Engineering Program,
College of Innovative Technology
and Engineering,
Dhurakij Pundit University
110/1-4 Pracha Chuen Rd,
Khwaeng Thung Song Hong, Khet
Lak Si, Bangkok 10210, Thailand
duangjai.jit@dpu.ac.th

Eakasit Pacharawongsakda
Big Data Engineering Program,
College of Innovative Technology
and Engineering,
Dhurakij Pundit University
110/1-4 Pracha Chuen Rd,
Khwaeng Thung Song Hong, Khet
Lak Si, Bangkok 10210, Thailand
eakasit.pac@dpu.ac.th

ABSTRACT

The objective of this study was to automate job performance prediction based on DISC personality test. We transformed this problem to Multi-Label Classification (MLC) by using employee's job performances as labels. In this study, three widely used MLC techniques have been employed such as Binary Relevance (BR), Label Powerset (LP) and Classifier Chains (CC) for prediction of job performances. However, these traditional techniques didn't show promising results. Therefore, we proposed another approach by building stacking MLC with model selection. The proposed method has three steps: (1) building MLC model; (2) using process from the first step and applying with a stacking model and (3) utilizing feature selection technique to select the proper models for final prediction. Using the surveys from a big financial company in Thailand, we found that the last proposed approach shows better performance, compared to the traditional MLC.

CCS Concepts

Information systems → Data management systems → Database design and models → Data model extensions
Computing methodologies → Modeling and simulation → Simulation types and techniques → Massively parallel and high-performance simulations

Keywords

Multi-label Classification; Stacking; Personality; DISC personality test; Psychology

1. INTRODUCTION

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

ICCBID 2019, October 18–20, 2019, TAICHUNG, Taiwan

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-7290-9/19/10...\$15.00

<https://doi.org/10.1145/3366650.3366666>

The relationship between personality and job performance has been studied in many industrial-organizational psychology research, the result of many research found that personality is one of the most important factors that affects to the employee's job performance [1]. Job performance generally relates to the positive things that people do to succeed at work, including task performance, discretionary behaviors, future-focused and improvement behaviors [2].

However, not every personality is suited for every job position, so it's important to recognize personality traits and to pair employees with the job that fit their personalities the best. This can lead to increased productivity and job satisfaction, helping business function to be more efficient.

To put the right man on the right job, understanding personality which suited for job position is crucial.

Previous research and study towards personality and job performance were conducted with only single-label of performance, but in fact, each employee can perform various performance at the same time. Single-label study cannot apply to complicated problem [3].

The classification of personality consists of comparing a user's personality against the standard personality tests taken. DISC is one of the most popular and standardized personality test, it shows that the human psychology and personality of normal people can be identified and observed and divides it into four categories. These four factors are Dominance, Influence, Steadiness and Compliance [4].

Section 2 is about the related works to conduct this study. Section 3 talks about the proposed method in this study. Section 4 describes the process flow. Section 5 and 6 talks about the results and conclusions drawn from this study

2. RELATED WORKS

Kim Yun-Yong et. al. studied about the effects of DISC behavior style of office workers on job satisfaction, organizational commitment and job performance. The study compared the difference between group of DISC type and found that Dominance people has a significantly higher recognized as a person who has good job performance than Steadiness people [5].

Fariha Tabasum et. al. had investigated the relationship between salesman's personality traits on sales performance for FMCG. This study shown the positive and significant relationship between salesman personality and consumer perception and sales. The result shows that customer perception and sales of specific product or service can be enhanced with the attractive personality of salesmen [6].

Joy Eberechukwu Agodi, Emmanuel Onyedikachi Ahaiwe and Aniekam Eyo Awah found a strongly positive relationship between personality traits and sales performance. Successful salesman can be described as being empathetic, assertive and ambitious [7].

Anisha Yata et. al. studied about personality prediction of the users by implementing multi-label classifiers on textual data which are the views and opinions of the users posted by them on the social networking [8].

Gauthier Doquire and Michel Verleysen proposed the use of mutual information for feature selection in multi-label classification. The results show the interest of the approach which allows one to sharply reduce the dimension of the problem and to enhance the performance of classifiers [9].

In this study, we combined techniques between Multi-Label Classification, Stacking and Feature Selection to gain a better result.

2.1 Multi-Label Classification

Multi-label Classification is the task of assigning data points to a set of classes or categories which are not mutually exclusive, which means that a point can belong simultaneously to different classes.

Multi-label classification methods can be broadly categorized into two different groups:

- i) Problem Transformation methods
- ii) Algorithm Adaptation methods

The first group contains methods that are algorithm independent approach, since its functioning does not depend directly on the classification method used. This method transforms the original multi-label classification problem into one or more single label classification, regression or ranking tasks.

The second group contains methods that extend specific learning algorithms or develop new algorithms in order to handle multi-label problems directly.

This paper aims at solving multi-label classification by implementing Problem Transformation method. The multi-label classifiers proposed to use in this study are Binary Relevance (BR), Label Powerset (LP) and Classifier Chains(CC) which are all Problem Transformation Methods.

The first two approaches were selected because they are the most basic approaches for multi-label classification tasks. BR considers the prediction of each label as an independent binary classification task and then constructing a binary training set for each particular label, while LP considers each unique set of labels that exists in a multi-label training set as one of the labels of a new single-label

classification task. For CC, it transforms the problem into a hierarchical chain of binary classification problems, where each one is built upon the previous prediction [10] [11] [12].

2.2 STACKING

Stacking is an ensemble learning technique in which a number of base classifiers are combined using one meta-classifier which learns their outputs. The individual classification models are trained based on the complete training set. Then, instead of using the original input attributes, Stacking uses the classifications predicted by the base-classifiers as the input attributes.

The meta-classifier that has been produced combines the different predictions into a final prediction. [13][14].

2.3 FEATURE SELECTION

Many classification problems rely on a large set of features, feature selection methods can effectively reduce data dimensionality by removing irrelevant and redundant features which can be used to get rid of this problem. There are 3 types of feature selection methods which are filter, wrapper, and embedded methods [15].

Filter methods is the simplest and most scalable of the methods, this method rank the features based on relevance measure and select the highest k ranked features according to that measure.

Wrapper method evaluates all possible combinations of the features and selects the combination that yields the best result for a specific machine learning algorithm.

Embedded methods using specific algorithms which incorporate feature selection as part of the training process, this method has the best ability to discriminate among classes.

In this study we applied Backward Elimination which is the wrapper method of feature selection to select relevance attributes.

3. PROPOSED METHOD

We proposed a method of combining of MLC, Stacking and Feature Selection. 3 steps were created as following;

Step 1: building MLC model by applying each algorithm with MLC to the training set.

Because of each employee can perform various performance, so, MLC would be applied to the machine learning process. Figure 1 demonstrates the structure of Step1. In this step, BR, LP and CC were run using 6 classification models (Logistic regression, KNN, Naive Bayes, SVM, Random Forest and Decision Tree) as the base classifier. We used a split test validation (train 70%, test 30%) on the training set. Finally, we would get the result of each base classifier in each MLC algorithm.



Figure 1, Building MLC model

Step 2: using process from the first step and applying with a stacking model.

Nanak Chand et al. (2016) measured and compared performance of SVM and its stacking with 9 other machine learning algorithms, and found that Stacking of SVM was providing a better performance [16].

Nazlia Omar et al.(2013) made a comparative study carried out on the effectiveness of individual supervised classifiers and ensemble methods for Subjectivity and Sentiment Analysis of Arabic Customers' Reviews. The results showed that ensemble of classification algorithms with meta learner ensemble technique performed robustly better than all the other individual classifier [17].

In this study, we applied Stacking model with the previous step. As we discussed in the reviews above, Stacking combines multiple classifiers generated by different machine learning algorithms. So, in this step, we used the top three performance algorithms to be base-classifiers (classification models) and each base classifier will predict the result from training set, then created a new data by transformed each base classifier result to be attributes of new data. 6 classification models (Logistic regression, KNN, Naive Bayes, SVM, Random Forest and Decision Tree) were selected for meta-classifier, and we used it to predict the final result by applied with BR, LP and CC. (Figure 2)

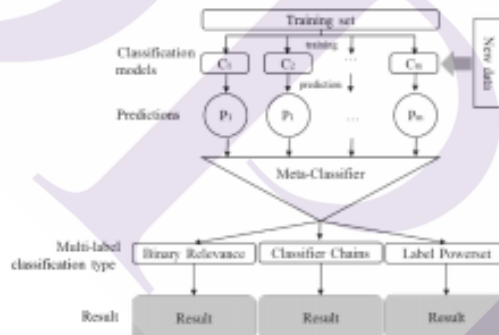


Figure 2, Using the process from the first step to build a stacking model,

Step 3: we added feature selection technique into step 2 process in order to select the proper models for the final prediction.

The number of attributes from questionnaire have 144 attributes, the high number of attributes that use to train the models may have a huge influence on the performance.

Erik M. Schmidt et al. (2010) found that Stacking with feature selection technique provided the best performance across the classification and regression tasks for automated systems to recognize the emotional content of music [18].

In order to find relevance attributes and achieve better accuracy, we applied Backward Elimination method which is one of feature selection method in the first process. However, the selected attributes from feature selection process of each label were difference. We joined (union) all selected attributes of each label together and then we repeated the same process of step2. (Figure 3)



Figure 3, Utilizing feature selection technique to select the proper models for final prediction

4. EXPERIMENT

4.1 Dataset Collection

For data collection process, this study collected data from the DISC personality test from 2,317 respondents which are Manager level (564 respondents) and Operation level (1,753 respondents), and also collected past 3 years of job performance record which consist of 4 job performance types.

4.2 Dataset Preparation

After that move to the data preparation process, some data such as wrong employee ID, duplicated record (1 ID did the test more than 1 times) and employee data who have been working less than 1 year should be eliminated. So, in total of 1,888 respondents are remained (406 managers, 1,482 operations). Joined data from personality test with data from database.

Next step, we transform the data into binary data type. For job performance, 2.5% of each high and low outlier would be eliminated, then transform by using the average value of each performance as a cut point (if it is higher than average, it will be transformed to 1, but if it is lower than average the value will be 0).

4.3 Evaluation

We have performed these experiments on Python version 3.7 by using DISC and job performance dataset. This study used five

measures which are: Hamming loss, Accuracy, Precision, Recall and F-Measure [19][20].

Hamming Loss: is the fraction of labels that are incorrectly predicted. A low value of hamming loss is required to show better classification performance.

$$\text{Hamming Loss} = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_i \Delta Z_i}{M} \right|$$

Accuracy: for each instance is defined as the proportion of the predicted correct labels to the total number (predicted and actual) of labels for that instance.

$$\text{Accuracy} = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_i \cap Z_i}{Y_i \cup Z_i} \right|$$

Precision: Precision can be defined as the percentage of true positive examples from all the examples classified as positive by the classification model.

$$\text{Precision} = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_i \cap Z_i}{Z_i} \right|$$

Recall: is the percentage of examples classified as positive by a classification model that are true positive.

$$\text{Recall} = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_i \cap Z_i}{Y_i} \right|$$

F-Measure: is a combination of Precision and Recall. It is the harmonic average of the two metrics and it is used as an aggregated performance score.

$$F - \text{Measure} = \frac{1}{m} \sum_{i=1}^m \frac{2|Y_i \cap Z_i|}{|Y_i| + |Z_i|}$$

Table 1. Evaluation Measures Comparison of Step 1

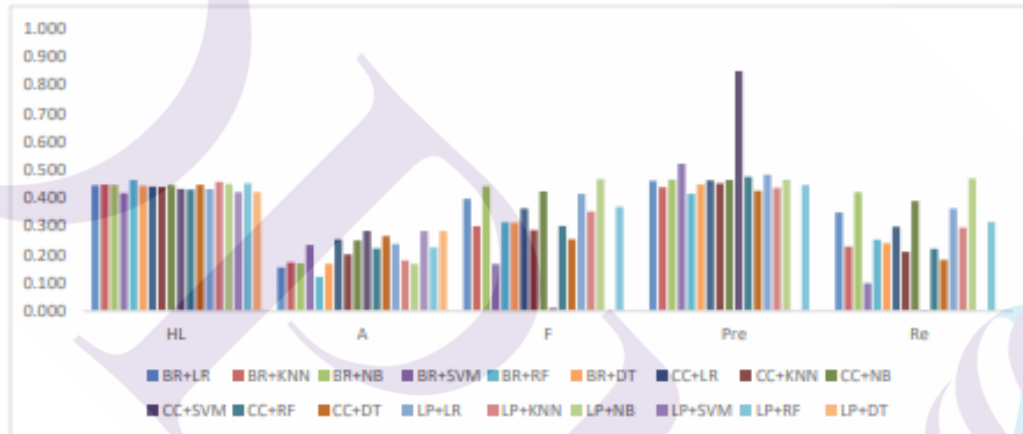


Table 2. Evaluation Measures Comparison of Step 2

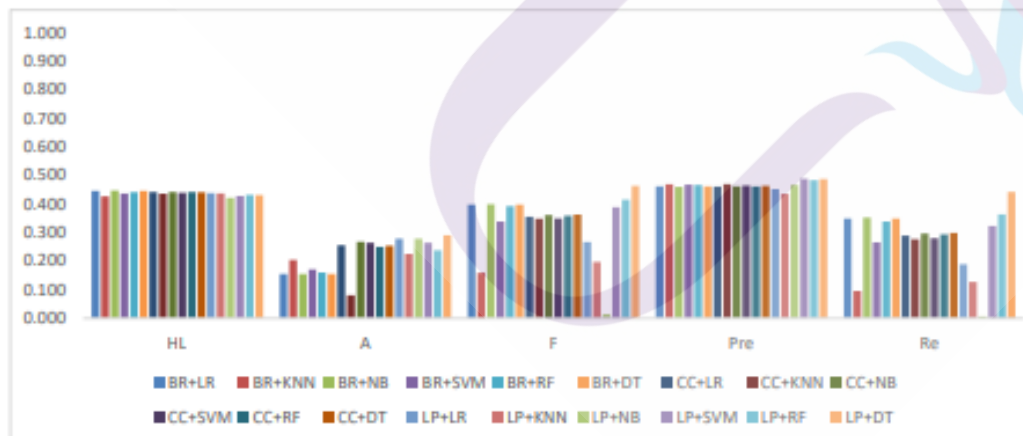
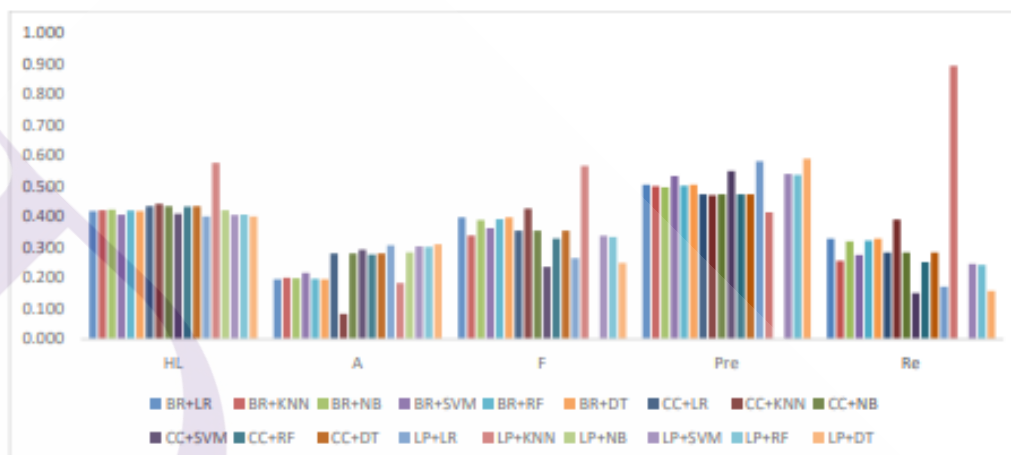


Table 3, Evaluation Measures Comparison of Step 3



5. RESULTS

Table 1-3 shows the predictive performance of all evaluation measures of each pair between base classifiers and MLC algorithms in each step.

Here, HL, A, F1, Pre and Re represents Hamming Loss, Accuracy, F-Measure, Precision and Recall respectively.

From Table 1-3, we can draw the conclusion that Step 3 that utilizing feature selection technique to MLC with Stacking method dominates the other steps in mostly of measures result. Moreover, Decision Tree base classifier gives the best overall performance (lowest Hamming Loss and highest Accuracy) when used with the MLC Label Powerset.

6. CONCLUSIONS

We proposed MLC of employee job performance prediction by disc personality. For improving MLC performance, we focus on step 3 that combined feature selection process for MLC with Stacking. The top three performance algorithms (Logistic regression, SVM, and Decision Tree) were applied to be base classifiers with six Meta-classifier's algorithms, while three MLC algorithms including BR, LP and CC. The results of step 3 revealed that this step able to gain a better performance than traditional MLC or only MLC with Stacking.

7. REFERENCES

- [1] Cascio, F. W. and Aguinis, H. 2008. Research in Industrial and Organizational Psychology From 1963 to 2007: Changes, Choices, and Trends. in *Journal of Applied Psychology* 2008, Vol. 93, No. 5, 1062-1081
- [2] Spitzmuller, M., Dyne, L. V. and Ilie, R. 2008. Organizational Citizenship Behavior. A Review and Extension of its Nomological Network
- [3] Waheed, A., Yang, J. and Webber, J. 2018. THE EFFECT OF PERSONALITY TRAITS ON SALES PERFORMANCE: AN EMPIRICAL INVESTIGATION TO TEST THE FIVE-FACTOR MODEL (FFM) IN PAKISTAN. in *Interdisciplinary Journal of Information, Knowledge, and Management*, Vol. 12
- [4] Weiming, G. 2011. Study on the Application of DISC Behavioral Style in Talent Management in Banking Industry. *Proceedings of the 8th International Conference on Innovation & Management*
- [5] Yong, K. Y., Hwa, B. Y., Hyun, P. H., Hyang, Y. J., and Su, J. E. 2012. The Effects of DISC Behavior Styles of Office Workers on Job Satisfaction, Organizational Commitment and Job Performance. *Korean J Occup Health Nurs*. 2012 Aug;21(2):98-107
- [6] Tabasam, F., Ibrahim, M. Rabbani, M. and Asif, M. 2014. Impact of Salesmen Personality on Customer Perception and Sales. *Global Journal of Management and Business Research: E Marketing*
- [7] Agodi, J. E., Ahaiwe, E. O. and Awah, A. E. 2017. Salesman's Personality Trait and Its Effect on Sales Performance: Study of Fast Moving Consumer Goods (FMCG) in Abia State, Nigeria. in *Journal of Economics and Sustainable Development*, Vol.8, No.24
- [8] Yata, A., Kante, P., Sravani, T. and Malathi, B. 2018. Personality Recognition using Multi-Label Classification. *International Research Journal of Engineering and Technology (IRJET)*, Volume: 05 Issue: 03
- [9] Doquire, G. and Verleysen, M. 2011. Feature Selection for Multi-label Classification Problems. *Universit'e catholique de Louvain*
- [10] Kafrawy, P. E., Mousad, A. and Esmail, H. 2015. Experimental Comparison of Methods for Multi-Label

- Classification in Different Application Domains. in *International Journal of Computer Applications*, Volume 114
- [11] Santos, A. M., Canuto, A. M. P. and Neto, A. F. 2011. A Comparative Analysis of Classification Methods to Multi-label Tasks in Different Application Domains. in *International Journal of Computer Information Systems and Industrial Management Applications*, Volume 3
- [12] Ganda, D. and Buch, R. 2018. A Survey on Multi Label Classification. in *Recent Trends in Programming Languages*, Volume 5, Issue 1
- [13] Menahem, E., Rokach, L. and Elovici, Y. 2009. An improved stacking schema for classification tasks. Department of Information Systems Engineering, Ben-Gurion University and Deutsche Telekom Laboratories at Ben-Gurion University, Be'er Sheva 84105, Israel
- [14] Wolpert, D. H. 1992. Stacked generalization. *Neural networks*, Volume 5, Issue 2, Pages 241-259
- [15] Spola'or, N., Cherman, E. A., Monard, M. C. and Lee, H. D. 2013. A Comparison of Multi-label Feature Selection Methods using the Problem Transformation Approach. in *Electronic Notes in Theoretical Computer Science* 292:135-151
- [16] Nanak, C., Preeti, M., C., Rama, K., Emmanuel, S. P. and Mahesh, C. G. 2016. A Comparative Analysis of SVM and its Stacking with other Classification Algorithm for Intrusion Detection. *International Conference on Advances in Computing, Communication, & Automation (ICACCA)*
- [17] Nazlia, O., Mohammed, A., Adel, Q. A., Tareq, A. 2013. Ensemble of Classification Algorithms for Subjectivity and Sentiment Analysis of Arabic Customers' Reviews. in *International Journal of Advancements in Computing Technology*
- [18] Erik, M. S., Douglas, T. and Youngmoo, E. K. 2010. Feature Selection for Content-Based, Time-Varying Musical Emotion Regression. Published in *Multimedia Information Retrieval*.
- [19] Asim, M. N., Rehman, A. and Shoab, U. 2017. Accuracy Based Feature Ranking Metric for Multi-Label Text Classification. in *International Journal of Advanced Computer Science and Applications*, Vol. 8, No. 10
- [20] Santos, A. M., Canuto, A. M. P. and Neto, A. F. 2011. A Comparative Analysis of Classification Methods to Multi-label Tasks in Different Application Domains. in *International Journal of Computer Information Systems and Industrial Management Applications*, Vol. 3, pp. 218-227

ประวัติผู้เขียน

ชื่อ นามสกุล
ประวัติการศึกษา

ตำแหน่งและสถานที่ทำงานปัจจุบัน

ปฐมภัทร คำตา
พ.ศ. 2554 ปริญญาตรี
คณะจิตวิทยา
จุฬาลงกรณ์มหาวิทยาลัย
พ.ศ. 2560 ปริญญาโท
คณะพาณิชยศาสตร์และการบัญชี ภาควิชาการตลาด
(หลักสูตรนานาชาติ)
มหาวิทยาลัยธรรมศาสตร์
นักวิทยาศาสตร์ข้อมูล
บริษัท เงินติดล้อ จำกัด