เทคนิคการเรียนรู้ของเครื่องสำหรับการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อยจากแบบประเมินความคล่องทางภาษา

นายสุพพัต เมธารักษ์ชีพ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2564
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

MACHINE LEARNING TECHNIQUES FOR DIAGNOSIS OF MILD COGNITIVE IMPAIRMENT FROM THAI VERBAL FLUENCY ASSESSMENT

Mr. Suppat Metarugcheep

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2021

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	เทคนิคการเรียนรู้ของเครื่องสำหรับการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อยจากแบบประเมินความคล่องทางภาษา
โดย	นายสุพพัต เมธารักษ์ชีพ
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	อาจารย์ ดร.นฤมล ประทานวณิช

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

---------- คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

---------- ประธานกรรมการ
(อาจารย์ ดร.เอกพล ช่วงสุวนิช)

---------- อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ)

---------- อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(อาจารย์ ดร.นฤมล ประทานวณิช)

---------- กรรมการ
(รองศาสตราจารย์ แพทย์หญิงโสพพัทธ์ เหมรัญช์โรจน์)

---------- กรรมการ
(อาจารย์ ดร.นพ.ชัยภัทร ชุณหรัศมิ์)

---------- กรรมการภายนอกมหาวิทยาลัย
(ดร.ธีรวิทย์ วิไลประสิทธิ์พร)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สุพพัต เมธารักษ์ชีพ : เทคนิคการเรียนรู้ของเครื่องสำหรับการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อยจากแบบประเมินความคล่องทางภาษา. (MACHINE LEARNING TECHNIQUES FOR DIAGNOSIS OF MILD COGNITIVE IMPAIRMENT FROM THAI VERBAL FLUENCY ASSESSMENT) อ.ที่ปรึกษาหลัก : รศ. ดร.โปรดปราน บุณยพุกกณะ, อ.ที่ปรึกษาร่วม : อ. ดร.นฤมล ประทานวณิช

แบบทดสอบความคล่องแคล่วทางภาษาของหน่วยเสียงเป็นแบบประเมินสำหรับภาวะความรู้คิดบกพร่อง โดยให้ผู้เข้าร่วมทำการทดสอบพูดคำที่ขึ้นต้นด้วยตัวอักษรที่กำหนดให้ได้มากที่สุดเป็นเวลา 1 นาที มีงานวิจัยจำนวนมากนำกลวิธีการเรียนรู้ของเครื่องซึ่งเป็นเทคโนโลยีที่ทันสมัย และใช้กันอย่างแพร่หลายร่วมกับเทคนิคการสกัดคุณลักษณะในเสียงเพื่อวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อย เช่น คุณลักษณะที่ขึ้นอยู่กับส่วนเงียบ การจัดกลุ่มคำ การเปลี่ยนกลุ่มคำ และคุณลักษณะทางความหมาย คุณลักษณะเหล่านี้จะถูกนำไปเป็นชุดข้อมูลสำหรับฝึกฝนและทดสอบในกระบวนการของการเรียนรู้ของเครื่อง เพื่อคัดแยกภาวะความรู้คิดบกพร่องเล็กน้อยต่อไป ปัจจุบันมีงานวิจัยที่ศึกษาเกี่ยวกับแบบทดสอบความคล่องแคล่วทางภาษาของหน่วยเสียงในหลายภาษาแต่ยังคงไม่พบการศึกษาในภาษาไทย การประยุกต์วิธีการที่ใช้ในปัจจุบันกับแบบทดสอบความคล่องทางภาษาฉบับภาษาไทยให้ผลลัพธ์ไม่เป็นที่น่าพอใจ เนื่องจากภาษาไทยมีลักษณะเฉพาะที่แตกต่างจากภาษาอังกฤษ เป็นผลให้วิธีการที่ใช้ปัจจุบัน เช่น การสกัดคุณลักษณะจะต้องปรับวิธีการคำนวนให้เหมาะสม งานวิจัยของเรามีเป้าหมายที่จะแยกแยะภาวะความรู้คิดบกพร่องเล็กน้อยจากแบบประเมินความคล่องทางภาษาฉบับภาษาไทย งานวิจัยของเรายังนำเสนอเทคนิคใหม่สำหรับการจัดกลุ่มคำสำหรับภาษาไทย และการใช้คุณลักษณะแบบผสมผสานโดยคัดเลือกจากผลลัพธ์ทางสถิติ โดยอาศัยข้อมูลจากโปรแกรมประยุกต์โมคาจำนวน 100 ตัวอย่าง โดยแบ่งเป็น 41 คน ที่มีภาวะความรู้คิดบกพร่อเล็กน้อยและ 59 คน ที่เป็นกลุ่มควบคุม ผลลัพธ์แสดงให้เห็นว่าการใช้คุณลักษณะที่คัดเลือกจากผลลัพธ์ทางสถิติช่วยเพิ่มความแม่นยำในการคัดแยกผู้ป่วยได้ดีอย่างมีนัยสำคัญ และคุณลักษณะแต่ละประเภทมีประสิทธิภาพต่อของตัวจำแนกประเภทไม่เท่ากัน

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่อนิสิต .....
ปีการศึกษา 2564 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

6270302421 : MAJOR COMPUTER SCIENCE

KEYWORD: Mild Cognitive Impairment, Machine Learning, Verbal Fluency, Feature Extraction

Suppat Metarugcheep : MACHINE LEARNING TECHNIQUES FOR DIAGNOSIS OF MILD COGNITIVE IMPAIRMENT FROM THAI VERBAL FLUENCY ASSESSMENT. Advisor: Assoc. Prof. Proadpran Punyabukkana Co-advisor: Ph.D. NARUEMON PRATANWANICH

Phonemic Verbal Fluency task (PVF) is an assessment for cognitive impairment that participants were asked to produce words starting with the given letters, such as the letter “F” in English and “ก” in Thai. Many studies applied the state-of-the-art machine learning method with feature extraction to diagnose mild cognitive impairment (MCI). For example, acoustic features, clustering, semantic feature. These features are measured the value from the information in audio data and then trained in machine learning model for classifying patients. Although, PVF has been studied in many languages but not found in Thai. Using the existing method for Thai PVF gives unpleasant results due to the difference characteristic in Thai compared to English. Thus, the method for Thai PVF needs to be rethought regarding analytical process of feature extraction. Our research focuses on diagnosis of MCI for Thai PVF using the data from The Montreal Cognitive Assessment (MoCA) application. We also propose a new method for clustering for Thai language and the combination of multiple features (silenced-based, similarity-based, and clustering) by statistical selection. The performance is evaluated with K-fold cross validation using the data set of 41 MCI and 59 HC (n=100). The result shows that using the significant feature from the combination of multiple feature groups yields classification accuracy, and each of feature groups has different predictive power for classification in Thai PVF.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Field of Study: Computer Science Student's Signature .....

Academic Year: 2021 Advisor's Signature .....

Co-advisor's Signature .....

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ รองศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้ความช่วยเหลือ คำชี้แนะ คอยผลักดัน และตรวจทานแก้ไขข้อบกพร่องเสมอมา และขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.ฑิตยา หวานวารี และ อาจารย์ ดร.นฤมล ประทานวณิช อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วมที่กรุณาให้ความรู้ ให้คำแนะนำมาโดยตลอด และขอขอบพระคุณ รองศาสตราจารย์ พญ.โสฬพัทธ์ เหมรัญช์โรจน์ และ อาจารย์ ดร.นพ.ชัยภัทร ชุณหรัศมิ์ กรรมการสอบวิทยานิพนธ์ ที่กรุณาให้คำปรึกษา และข้อมูลที่จำเป็นต่องานวิจัยมาโดยตลอดจนวิทยานิพนธ์เล่มนี้สำเร็จสมบูรณ์

ขอขอบพระคุณ อาจารย์ ดร.เอกพล ช่วงสุวนิช ประธานกรรมการสอบวิทยานิพนธ์ และผู้ช่วยศาสตราจารย์ ดร.ธีรวิทย์ วิไลประสิทธิ์พร กรรมการภายนอกมหาวิทยาลัย ที่กรุณาเสียสละเวลาชี้แนะ และให้คำแนะนำจนวิทยานิพนธ์เล่มนี้สำเร็จสมบูรณ์มากยิ่งขึ้น

ขอขอบพระคุณ บุคลากรณ์ทางการแพทย์ และนักจิตวิทยา โรงพยาบาลจุฬาลงกรณ์ ฝ่ายจิตเวชศาสตร์ ที่กรุณาให้ข้อมูลที่สำคัญต่องานวิจัย

สุดท้ายนี้ ขอขอบพระคุณบิดา มารดา และญาติพี่น้องของข้าพเจ้าที่เป็นกำลังใจสำคัญมาโดยตลอด และขอขอบคุณเพื่อนๆ พี่ๆ ที่คอยช่วยเหลือ และให้คำแนะนำเสมอมาจนวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงลงได้ด้วยดี

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สุพพัต เมธารักษ์ชีพ

สารบัญ

หน้า

.....ค

บทคัดย่อภาษาไทย.....ค

.....ง

บทคัดย่อภาษาอังกฤษ.....ง

กิตติกรรมประกาศ.....จ

สารบัญ.....ฉ

สารบัญตาราง.....ฌ

สารบัญภาพ.....ญ

บทที่ 1 บทนำ.....11

1.1 ที่มาและความสำคัญของปัญหา.....11

1.2 วัตถุประสงค์ของการวิจัย.....12

1.3 ขอบเขตการดำเนินงาน.....12

1.4 ประโยชน์ที่คาดว่าจะได้รับ.....13

1.5 ขั้นตอนการดำเนินงาน.....13

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....14

2.1 ทฤฎีที่เกี่ยวข้อง.....14

2.1.1 การพัฒนาโปรแกรมประยุกต์โมคา.....14

2.1.2 แบบประเมินความคล่องแคล่วทางภาษาของหน่วยเสียงในโมคา.....15

2.1.3 ความคล้ายแบบอักขรวิธี (Orthographic Similarity).....16

2.1.4 ระยะทาง Levenshtein.....17

2.1.5 ความคล้ายเชิงความหมาย (Semantic Similarity).....18

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.6 การตรวจจับความเงียบในเสียงบันทึก 19
2.1.7 K-Fold Cross Validation 20
2.1.8 ลักษณะของคำในภาษาไทย 20
2.2 งานวิจัยที่เกี่ยวข้อง 21
2.2.1 งานวิจัยที่เกี่ยวกับคุณลักษณะที่ขึ้นอยู่กับความเงียบ 21
2.2.2 งานวิจัยที่เกี่ยวกับคุณลักษณะที่ขึ้นอยู่กับความคล้าย 21
2.2.3 คุณลักษณะที่มาจากการจัดกลุ่มคำและการสลับกลุ่มคำ 22
บทที่ 3 ระเบียบวิธีการวิจัย 23
3.1 การทำแบบประเมินและการเก็บข้อมูล 23
3.2 การสกัดคุณลักษณะ 24
3.2.1 เทคนิคการจัดกลุ่มคำสำหรับภาษาไทย 26
3.2.2 เทคนิคการเปลี่ยนกลุ่มคำ 27
3.2.3 เทคนิคการใช้ค่า SHAP เพื่อหาความสำคัญของคุณลักษณะ 28
3.3 การจำแนกประเภท 28
3.3.1 การแบ่งข้อมูลด้วยเทคนิค 10-Fold Cross Validation 28
3.3.2 การฝึกฝนและการทดสอบแบบจำลอง 29
3.3.3 การปรับละเอียดแบบจำลองสำหรับคุณลักษณะสำคัญ 30
3.4 การคัดเลือกคุณลักษณะ 32
3.5 การประเมินผลแบบจำลอง 33
บทที่ 4 ผลการทดลอง 35
4.1 ผลการประเมินแบบจำลองเพื่อการคัดเลือกคุณลักษณะ 35
4.1.1 ผลลัพธ์ของคุณลักษณะเดี่ยว 35
4.1.2 ผลลัพธ์ของคุณลักษณะแบบกลุ่ม 39
4.2 ผลการประเมินแบบจำลองสำหรับคุณลักษณะสำคัญ 40

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.3 การพิจารณาความสำคัญของคุณลักษณะด้วยค่า SHAP....41
4.3.1 ค่า SHAP ในตัวจำแนก XGBoost....41
4.3.2 ค่า SHAP ในตัวจำแนก Support Vector Machine....42
4.3.3 ค่า SHAP ในตัวจำแนก Random Forest....43
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ....45
5.1 การสรุปผลลัพธ์ในภาพรวม....45
5.2 ปัจจัยของคุณลักษณะแต่ละชนิดต่อภาวะความรู้คิดบกพร่องเล็กน้อย....45
5.2.1 คุณลักษณะที่ขึ้นอยู่กับความเงียบ....46
5.2.2 คุณลักษณะที่ขึ้นอยู่กับความคล้าย....46
5.2.3 การจัดกลุ่มคำและการเปลี่ยนกลุ่มคำ....47
5.2.4 จำนวนปีการศึกษา....48
5.3 ผลลัพธ์จากการคัดเลือกคุณลักษณะสำคัญ....49
5.4 ข้อจำกัดของงานวิจัย....49
5.5 ข้อเสนอแนะสำหรับการนำงานวิจัยไปศึกษาต่อ....50
บรรณานุกรม....51
ประวัติผู้เขียน....56

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

หน้า

ตารางที่ 1 การแจกแจงข้อมูลประชากรของชุดข้อมูล 24
ตารางที่ 2 แสดงรายการคุณลักษณะทั้งหมดในงานวิจัย 25
ตารางที่ 3 แสดงรายการของกลุ่มคำ 26
ตารางที่ 4 แสดงค่าผลลัพธ์ของคุณลักษณะแบบเดี่ยว 36
ตารางที่ 5 แสดงค่าผลลัพธ์ของคุณลักษณะแบบกลุ่ม 39
ตารางที่ 6 แสดงค่าผลลัพธ์ของคุณลักษณะสำคัญ 40
ตารางที่ 7 ตารางพิสัยการแจกแจงข้อมูลคุณลักษณะที่ขึ้นอยู่กับความเงียบ.....46
ตารางที่ 8 ตารางพิสัยการแจกแจงข้อมูลคุณลักษณะเชิงความหมาย 47
ตารางที่ 9 ตารางพิสัยความถี่ข้อมูลของการจัดกลุ่มคำ 47
ตารางที่ 10 ตารางพิสัยการแจกแจงข้อมูลของการเปลี่ยนกลุ่มคำ.....48
ตารางที่ 11 ตารางพิสัยและการแจกแจงความถี่ของจำนวนปีการศึกษา 48
ตารางที่ 12 คุณลักษณะสำคัญที่มีผลต่อความแม่นยำในแต่ละตัวจำแนก.....49

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญภาพ

หน้า

ภาพที่ 1 โปรแกรมประยุกต์โมคา 14
ภาพที่ 2 แบบประเมินความคล่องทางภาษาในแบบประเมินโมคา 16
ภาพที่ 3 ภาพประกอบการคำนวนความคล้ายแบบอักขรวิธี 17
ภาพที่ 4 ภาพประกอบความคล้ายเชิงความหมายระหว่างคำ 19
ภาพที่ 5 การตรวจจับความเงียบในเสียงบันทึก 19
ภาพที่ 6 โครงสร้าง K-Fold Cross Validation (K=5) 20
ภาพที่ 7 ภาพแสดงขั้นตอนการทำวิจัยโดยรวม 23
ภาพที่ 8 ขั้นตอนวิธีการของการจัดกลุ่มคำ 27
ภาพที่ 9 ภาพแสดงการนับการเปลี่ยกลุ่มคำ 28
ภาพที่ 10 ภาพแสดงการแบ่งข้อมูลด้วย 10-Fold Cross Validation 29
ภาพที่ 11 แสดงตัวอย่างชุดคำสั่งการฝึกฝนและทดสอบแบบจำลอง 30
ภาพที่ 12 แสดงชุดคำสั่งการปรับละเอียดในตัวจำแนก Support Vector Machine 31
ภาพที่ 13 แสดงชุดคำสั่งการปรับละเอียดในตัวจำแนก Random Forest 31
ภาพที่ 14 แสดงชุดคำสั่งการปรับละเอียดในตัวจำแนก XGBoost 32
ภาพที่ 15 แผนภูมิสายงานการคัดเลือกคุณลักษณะ 33
ภาพที่ 16 ภาพแสดงค่าพื้นที่ใต้กราฟ ROC (AUC) 34
ภาพที่ 17 Confusion Matrix ของคุณลักษณะสำคัญ 41
ภาพที่ 18 ค่า SHAP ของตัวจำแนก XGBoost 42
ภาพที่ 19 ค่า SHAP ของตัวจำแนก Support Vector Machine 43
ภาพที่ 20 ค่า SHAP ของตัวจำแนก Random Forest 44

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ประเทศไทยกำลังก้าวเข้าสู่สังคมผู้สูงอายุ จากรายงานในปีพุทธศักราช 2544 พบว่าประชากรที่มีอายุมากกว่า 65 ปี หรือประมาณ 7 เปอร์เซ็น ของประชากรทั้งหมด มีการประมาณการว่าภายในปีพุทธศักราช 2593 ประเทศไทยจะมีประชากรผู้สูงอายุเพิ่มสูงขึ้นถึง 35.8 เปอร์เซ็น หรือคิดเป็นประมาณ 20 ล้านคน จากสถิติพบว่ามีโอกาสที่จะพบภาวะรู้คิดบกพร่องเล็กน้อย (Mild Cognitive Impairment; MCI) ในผู้สูงอายุประมาณ 20 เปอร์เซ็น [1] อัตราการพบภาวะความรู้คิดบกพร่องเล็กน้อยจึงเพิ่มขึ้นตามประชากรสูงอายุที่เพิ่มขึ้นด้วย ด้วยเหตุนี้การตรวจพบภาวะความรู้คิดบกพร่องเล็กน้อยตั้งแต่แรกเริ่มจึงเป็นสิ่งจำเป็น ภาวะความรู้คิดบกพร่องเล็กน้อยมีสาเหตุมาจากการเปลี่ยนแปลงการรู้จำในผู้ที่มีอายุมากกว่า 65 ปีขึ้นไป ภาวะนี้ยังสามารถพัฒนาไปเป็นโรคอัลไซเมอร์ (Alzheimer's disease) ได้อีกด้วย การตรวจพบภาวะรู้คิดบกพร่องเล็กน้อยตั้งแต่เนิ่นๆ นั้นมีความสำคัญซึ่งช่วยให้ครอบครัวของผู้ป่วยสามารถประคับประคองอาการ หรือชะลออาการของโรคอัลไซเมอร์ได้ อย่างไรก็ตามการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อยใช้เวลาค่อนข้างมาก และยังต้องการแพทย์ผู้เชี่ยวชาญสำหรับการวินิจฉัย ด้วยเหตุผลนี้ผู้วิจัยคิดว่าการนำเทคโนโลยีทางคอมพิวเตอร์มาพัฒนาจนเป็นระบบที่สามารถช่วยเหลือในรูปแบบของการเครื่องมือคัดกรองอัตโนมัติจะช่วยลดระยะเวลาและอำนวยความสะดวกต่อบุคลากรทางการแพทย์ได้

แบบประเมินโมคา (The Montreal Cognitive Assessment) เป็นที่นิยมอย่างกว้างขวางสำหรับการคัดกรองและการวินิฉัยความบกพร่องทางสติปัญญา (Cognitive Impairment) [2] การวินิจฉัยภาวะรู้คิดบกพร่องเล็กน้อยสำหรับแบบประเมินโมคาจะทำโดยพิจารณาคะแนนรวม โดยคะแนนรวมที่มากว่าหรือเท่ากับ 25 จะถือเป็นปกติ คะแนนรวมที่ต่ำกว่า 25 จะถือว่ามีภาวะความรู้คิดบกพร่องเล็กน้อย อย่างไรก็ตามแบบประเมินโมคามีข้อด้อย 4 ข้อ ดังนี้ 1.) แบบประเมินโมคาต้องการบุคลากรทางการแพทย์อย่างน้อย 1 คน เพื่อทำการประเมิณจนจบ 2.) การประเมินต้องใช้เวลาอย่างน้อย 15-30 นาที ในการทำประเมินแต่ละครั้ง 3.) แบบประเมินไม่สามารถใช้ได้กับผู้เข้ารับการประเมินที่มีความผิดปกติทางสายตาหรือผู้พิการ 4.) ข้อมูลจากการประเมินส่วนใหญ่จะถูกบันทึกลงบนกระดาษในรูปแบบของการสรุปคะแนนเท่านั้น

ด้วยเหตุผลเหล่านี้แบบประเมินความคล่องทางภาษาจึงถูกนำมาแก้ปัญหา ด้วยความที่มันมีความซับซ้อนในการประเมินน้อยกว่า [3] โดยทั่วไปแล้วแบบประเมินความคล่องทางภาษาจะแบ่งเป็น 2 ประเภท คือ 1.) แบบประเมินความคล่องทางภาษาของความหมาย (Semantic Verbal Fluency) 2.) แบบประเมินความคล่องทางภาษาของหน่วยเสียง (Phonemic Verbal Fluency) โดยจะมี

ขั้นตอนการทำแบบประเมิน คือ ผู้เข้ารับการประเมินจะได้รับคำสั่งให้คิดคำพูดให้มากที่สุดภายในเวลา 1 นาที โดยมีเงื่อนไขว่าคำนั้นจะต้องมีความหมายอยู่ในประเภทของ สัตว์ หรือผลไม้สำหรับแบบประเมินความคล่องทางภาษาของความหมาย และคำที่ขึ้นต้นด้วยตัวอักษร “ก” สำหรับแบบประเมินความคล่องทางภาษาของหน่วยเสียง ในการให้คะแนนสำหรับแบบประเมินความคล่องทางภาษานั้นจะพิจารณาจากจำนวนคำที่ผู้เข้ารับการประเมินพูดถูกต้องตามเงื่อนไขทั้งหมดโดยไม่คิดคำซ้ำและชื่อเฉพาะ เช่น กรุงเทพมหานคร กฤษฎา เป็นต้น โดยมีเงื่อนไข คือ ถ้าหากคะแนนรวมมากกว่าหรือเท่ากับ 11 คำ จะถือว่าเป็นปกติ

ในงานวิจัยที่ผ่านมาการนำเทคโนโลยีการประมวลผลภาษาธรรมชาติ (Natural Language Processing) มาใช้วิเคาะห์ข้อมูลเพื่อสกัดคุณลักษณะเฉพาะในเสียงสำหรับแบบประเมินความคล่องทางภาษา เช่น ความสัมพันธ์ทางความหมายของคำ ความคล้ายของคำ การจัดกลุ่มคำ และคุณลักษณะอื่นๆ โดยมีจุดประสงค์ที่จะแยกแยะความบกพร่องทางสติปัญญา (Cognitive Impairment) [4-6] ถึงแม้ว่าวิธีที่กล่าวถึงนั้นจะมีประสิทธิภาพที่ดีโดยเฉพาะในภาษาอังกฤษ แต่เมื่อนำมาใช้กับภาษาไทยแล้วกลับได้ผลลัพธ์ไม่เป็นที่น่าพอใจ เนื่องจากความแตกต่างระหว่างภาษาไทยกับภาษาอังกฤษ โดยเฉพาะอย่างยิ่งการจัดกลุ่มคำในภาษาไทยที่มีความแตกต่างจากภาษาอังกฤษอย่างชัดเจน จึงต้องมีการจัดกลุ่มใหม่เพื่อให้เหมาะสมกับภาษาไทย เช่น กลุ่มคำควบกล้ำ กลุ่มคำพ้องเสียง กลุ่มคำอาการนาม กลุ่มคำประวิสรรชนีย์ และอื่นๆ ด้วยเหตุนี้วิธีการดั่งเดิมจะต้องปรับวิธีการคำนวนเพื่อสกัดคุณลักษณะภายในเสียงใหม่เพื่อให้เหมาะสมกับการใช้งานในภาษาไทย

สำหรับงานวิจัยนี้เรานำเสนอเทคนิคใหม่สำหรับการจัดกลุ่มคำในภาษาไทย และเทคนิคการเรียนรู้ของเครื่องร่วมกับการใช้คุณลักษณะแบบผสม ด้วยการคัดเลือกคุณลักษณะจากผลลัพธ์ทางสถิติ เพื่อการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อยสำหรับแบบประเมินความคล่องแคล่วทางภาษาแบบหน่วยเสียงในภาษาไทย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาการสกัดคุณลักษณะภายในข้อมูลเสียง
2. เพื่อศึกษาและพัฒนาเทคนิคการเรียนรู้ของเครื่องในการจำแนกภาวะความรู้คิดบกพร่องเล็กน้อย จากข้อมูลเสียงในแบบประเมินความคล่องทางภาษา
3. เพื่อศึกษาปัจจัยทางคุณลักษณะที่มีผลต่อประสิทธิภาพการจำแนกประเภท

1.3 ขอบเขตการดำเนินงาน

1. แหล่งข้อมูลมาจากการบันทึกเสียงผ่านโปรแกรมประยุกต์โมคาระหว่างการทำแบบประเมินของผู้เข้าร่วม

2. ข้อมูลเสียงเป็นเป็นการสนทนาภาษาไทย
3. ทำการศึกษาและพัฒนาแบบจำลองเพื่อใช้สำหรับภาวะความรู้คิดบกพร่องเล็กน้อยเท่านั้น
4. สภาพแวดล้อมของการบันทึกเสียงจะต้องอยู่ในความเหมาะสำหรับการบันทึก โดยปราศจากเสียงรบกวนภายนอก

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ความรู้จากการศึกษาการสกัดคุณลักษณะภายในข้อมูลเสียง
2. ได้ความรู้จากการพัฒนาแบบจำลองเพื่อใช้คัดแยกภาวะความรู้คิดบกพร่องเล็กน้อย
3. ได้ความรู้จากการศึกษาการอธิบายแบบจำลองด้วยวิธีที่ทันสมัยต่างๆ
4. ได้เทคนิคสำหรับการคัดแยกภาวะความรู้คิดบกพร่องเล็กน้อยในโปรแกรมประยุกต์โมคา
5. เพื่อเป็นแนวทางสำหรับการศึกษาต่อเกี่ยวกับโรคสมองเสื่อม (Dementia) ในอนาคต

1.5 ขั้นตอนการดำเนินงาน

1. ศึกษาค้นคว้างานวิจัยและทฤษฎีที่เกี่ยวข้อง
2. รวบรวมข้อมูลที่ต้องใช้สำหรับงานวิจัย
3. ทำความสะอาดข้อมูลเบื้องต้น
4. พัฒนาโปรแกรมคอมพิวเตอร์เพื่อสกัดคุณลักษณะจากข้อมูลเสียง
5. ออกแบบและสร้างแบบจำลองเพื่อจำแนกภาวะความรู้คิดบกพร่องเล็กน้อย
6. หาข้อผิดพลาด และปรับแต่งแบบจำลองให้ดียิ่งขึ้น
7. ดำเนินการวัดผลลัพธ์
8. สรุปผลการดำเนินงาน
9. เรียบเรียงผลลัพธ์และจัดทำวิทยานิพนธ์

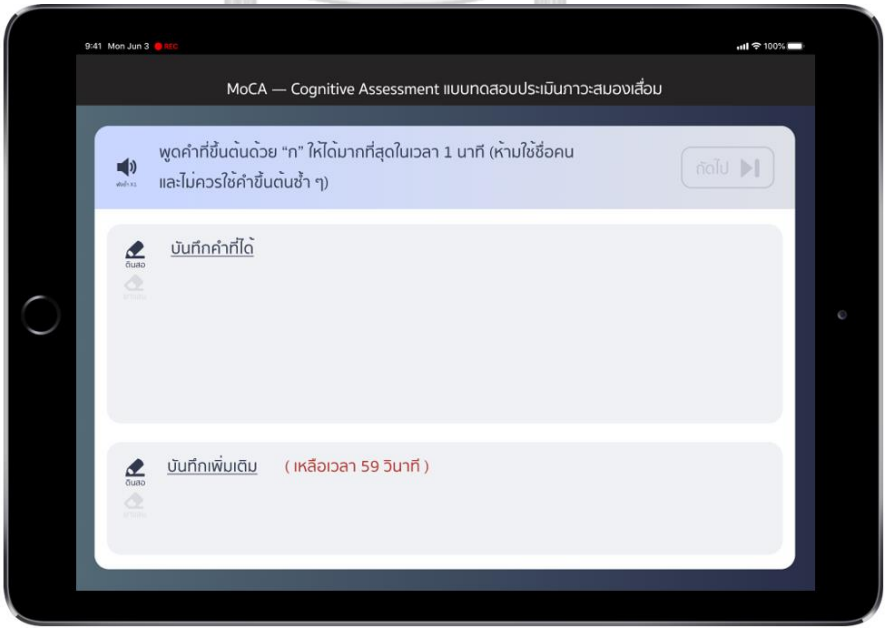
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤฎีที่เกี่ยวข้อง

2.1.1 การพัฒนาโปรแกรมประยุกต์โมคา

ปัจจุบันแบบประเมินโมคาในประเทศไทยมี 2 แบบ คือ 1.) แบบมาตฐานที่บันทึกข้อมูลและคะแนนลงบนกระดาษ 2.) แบบโปรแกรมประยุกต์บนแท็บเล็ตพีซี (Tablet PC) สำหรับแบบโปรแกรมประยุกต์โมคานั้นถูกพัฒนาโดยได้รับความร่วมมือจากฝ่ายจิตเวชศาสตร์ โรงพยาบาลจุฬาลงกรณ์ และภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย [7] การใช้งานโปรแกรมประยุกต์โมคาทำให้เกิดการเปิดโอกาสที่จะนำเทคโนโลยีในสาขาวิชาต่างๆ มาประยุกต์ใช้เพื่อให้เกิดประโยชน์สูงสุด เช่น เทคโนโลยีปัญญาประดิษย์ (Artificial Intelligence) เทคนิคการเรียนรู้ของเครื่อง (Machine Learning Technique) การเรียนรู้เชิงลึก (Deep Learning) ทั้งนี้การพัฒนาโปรแกรมโมคาประยุกต์มีเป้าหมายเบื้องต้น 2 ประการ คือ 1.) การให้คะแนนการประเมินแบบอัตโนมัติ 2.) การตรวจพบโรคที่เกิดจากความรู้คิดบกพร่องแบบอัตโนมัติโดยอาศัยข้อมูลบันทึกได้ ในโปรแกรมโมคาประยุกต์ใช้จะขั้นตอนการประเมิน และหนังสือการขอความยินยอมแบบเดียวกับแบบประเมินโมคาแบบมาตรฐาน ในขณะที่ผู้เขียนกำลังเขียนงานวิจัยนี้ โปรแกรมโมคาประยุกต์ยังคงเป็นรุ่นต้นแบบในการพัฒนา และงานวิจัยนี้เป็นโครงการริเริ่มเพื่อเพื่ออรรถประโยชน์ทางเทคโนโลยีที่ล้ำสมัย



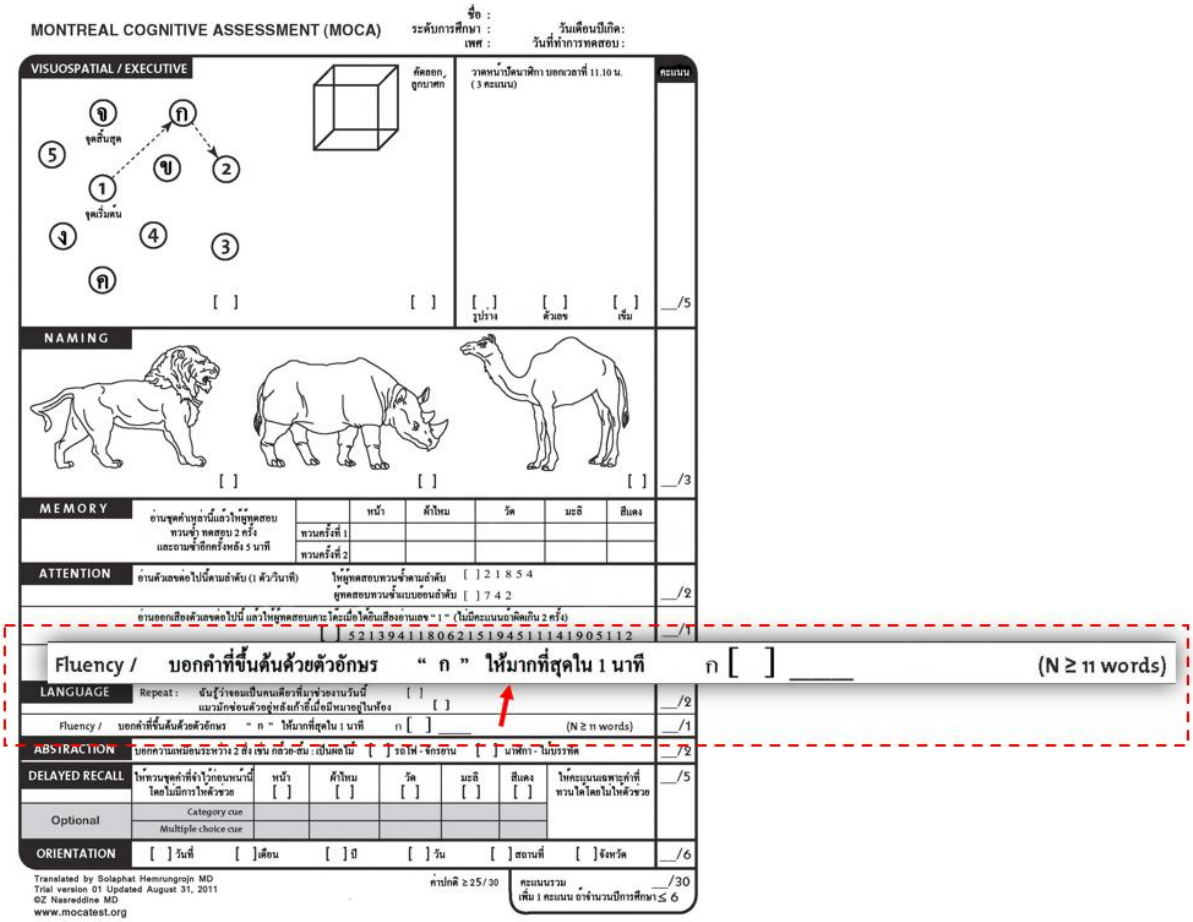
ภาพที่ 1 โปรแกรมประยุกต์โมคา

2.1.2 แบบประเมินความคล่องแคล่วทางภาษาของหน่วยเสียงในโมคา

แบบประเมินโมคามีคะแนนเต็ม 30 คะแนนใช้เวลาทำประเมินประมาณ 15-30 นาที [8] และประกอบด้วย 8 หัวข้อย่อยที่วัดทักษะแตกต่างกันในแต่ละโดเมน แบบประเมินความคล่องแคล่วทางภาษาอยู่ภายใต้ข้อย่อยที่ 8 ดังแสดงในภาพที่ 2 [9] โดยจะวัดทักษะด้านความสามารถทางภาษา (Language) โดยมีคำสั่งให้ผู้รับการประเมินคิดคำพูดที่ขึ้นต้นด้วยตัวอักษร “ก” ให้ได้มากที่สุดภายใน 1 นาที สำหรับการให้คะแนนจะพิจารณาจากจำนวนคำที่ผู้เข้ารับการประเมินพูดถูกต้องตามเงื่อนไขทั้งหมดโดยไม่นับคำซ้ำ คำที่ไม่มีความหมาย และชื่อเฉพาะ เช่น กรุงเทพมหานคร กฤษฎา ถ้าหากคะแนนรวมมากกว่าหรือเท่ากับ 11 คำ จะถูกพิจารณาให้คะแนนสำหรับข้อย่อยนี้ 1 คะแนน

ความคล่องแคล่วทางภาษาของหน่วยเสียงนั้นเกี่ยวข้องกับการทำงานของสมองส่วนหน้า โดยมีงานวิจัยระบุว่าความคล่องแคล่วทางภาษาของหน่วยเสียงต้องการทักษะหลายๆ อย่าง เช่น กลยุทธ์การสืบค้น (Search Strategy) ความจำขณะทำงาน (Working Memory) การเปลี่ยนกลุ่มคำ (Switching) ความจำเชิงความหมาย (Semantic Memory) และการยับยั้งความไม่เกี่ยวข้อง (Suppressing Irrelevant) [10-12] คนไข้ที่มีบาดแผลในส่วนหน้าของสมองมีแนวโน้มที่จะคิดคำพูดได้น้อยกว่ากลุ่มควบคุม มีงานวิจัยที่บ่งชี้ว่าความคล่องแคล่วทางภาษาของหน่วยเสียงกระตุ้นให้เกิดการทำงานที่ซับซ้อนของสมองส่วนหน้าร่วมกับส่วนอื่นๆ โดยใช้การถ่ายภาพสมอง (Neuroimaging) [13, 14] และยังมีงานวิจัยที่บ่งชี้ว่าจำนวนปีการศึกษาของผู้เข้ารับประเมินมีความสัมพันธ์กับจำนวนคำพูดที่ถูกต้องอีกด้วย [15-17] ด้วยการสนับสนุนจากงานวิจัยเหล่านี้อาจสรุปได้ว่าโรคที่มีสาเหตุจากความผิดปกติของการทำงานในสมองส่วนหน้า เช่น ภาวะความรู้คิดบกพร่องเล็กน้อย และโรคอัลไซเมอร์ จะแสดงความผิดปกติในแบบประเมินความคล่องแคล่วทางภาษา [18] ปัจจุบันประเทศไทยมีการนำโปรแกรมประยุกต์โมคาไปใช้แล้วในบางพื้นที่ ข้อมูลของแบบทดสอบย่อยนี้จะเป็นสิ่งที่ผู้วิจัยจะนำมาศึกษาต่อไป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



ภาพที่ 2 แบบประเมินความคล่องทางภาษาในแบบประเมินโมคา

2.1.3 ความคล้ายแบบอักขรวิธี (Orthographic Similarity)

ความคล้ายแบบอักขรวิธี คือ ค่าที่บ่งบอกความคล้ายระหว่าง 2 คำ ซึ่งจะมีค่าตั้งแต่ 0 จนถึง 1 โดยที่ 1 หมายถึงคำที่มีความเหมือนกัน 100 เปอร์เซ็น และ 0 หมายถึงคำที่ไม่มีจุดใดเหมือนกันเลย [19] สำหรับงานวิจัยนี้เรานำขั้นตอนการคำนวนความคล้ายแบบอักขรวิธีแบบฉบับภาษาอังกฤษ มาใช้กับภาษาไทย โดยมีขั้นตอนที่เปลี่ยนแปลงเล็กน้อยเพื่อให้มีความเหมาะสมกับลักษณะของการประสมคำ และสระในภาษาไทยที่สามารถเขียนอยู่ด้านบนหรือด้านล่างของพยัญชนะได้ จากภาพที่ 3 แสดงขั้นตอนการการคำนวนความคล้ายแบบอักขรได้ดังนี้

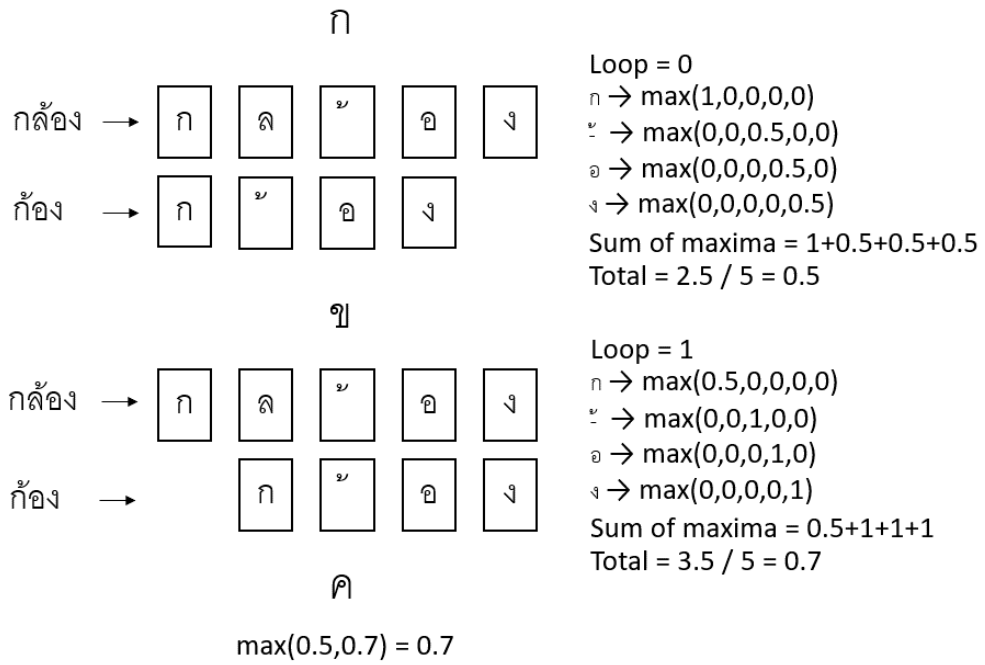
1) ขั้นตอน “ก” คำที่ 1 และคำที่ 2 จะถูกวางไว้ที่ตำแหน่งเดียวกัน โดยนำตัวอักษรที่มีความยาวมากกว่าเป็นตัวตั้งเสมอ และจำนวนรอบให้มีค่าเท่ากับ 0 จากนั้นเปรียบเทียบตัวอักษรตัวแรกของคำที่สั้นกว่ากับตัวอักษรทุกตำแหน่งของคำที่ยาวกว่า จากนั้นให้ผลการเปรียบเทียบเท่ากับ 1 หากตัวอักษรเหมือนกันในตำแหน่งเดียวกัน ให้ผลการเปรียบเทียบเป็น 1/2 หากตัวอักษรเหมือนกันในตำแหน่งถัดไป ให้ผลการเปรียบเทียบ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

เท่ากับ 1/3 หากตัวอักษรเหมือนกันในตำแหน่งที่ 3 จะเห็นว่าใช้ผลหารเป็นอนุกรมเศษส่วนต่อเนื่องเช่นนี้ไปเรื่อยๆ จากนั้นทำซ้ำจนครบทุกตัวอักษรของคำที่สั้นกว่าแล้วให้หาค่าสูงสุดของการเปรียบเทียบในแต่ละตัวอักษรของคำที่สั้นกว่า จากนั้นหารผลรวมด้วยความยาวตัวอักษรของคำที่ยาวกว่า

2) ขั้นตอน “ข” เลื่อนคำที่สั้นกว่าออกไป 1 ตำแหน่ง เปรียบเทียบตัวอักษรของคำที่สั้นกว่ากับทุกตัวอักษรของคำที่ยาวกว่า โดยมีกระบวนการเดียวกับขั้นตอน “ก”

3) ขั้นตอน “ค” หาค่าสูงสุดของผลลัพธ์จากขั้นตอน “ก” และ “ข”



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาพที่ 3 ภาพประกอบการคำนวนความคล้ายแบบอักขรวิธี

2.1.4 ระยะทาง Levenshtein

ระยะทาง Levenshtein เป็นวิธีการวัดความแตกต่างระหว่างคำ 2 คำ ที่นำเสนอโดยงานวิจัย [20] โดยมีวิธีการคำนวนจากการนับจำนวนการดำเนินการทั้งหมดที่จะแปลงคำหนึ่งไปสู่อีกคำหนึ่ง สำหรับการดำเนินการจะประกอบด้วย 3 ลักษณะ คือ การตัดออก (Deletion) การแทรก (Insertion) และการแทนที่ (Substitution) ระยะทาง Levenshtein ถูกใช้อย่างทั่วไปในเทคนิคการสกัดคุณลักษณะ (Feature Extraction) ในการการวิเคราะห์เชิงภาษาเพื่อนำเข้าสู่ขั้นตอนการเรียนของเครื่อง มีงานวิจัยที่นำเสนอการตรวจจับความบกพร่องทางสติปัญญาในผู้ป่วยที่มีภาวะความรู้คิดบกพร่องเล็กน้อยโดยใช้ระยะทาง Levenshtein โดยให้เป็นคุณลักษณะที่มีการท่วงน้ำหนักตามลำดับ

ของหน่วยเสียง [21, 22] ในงานวิจัยนี้จะใช้ค่าสัดส่วนระยะทาง Levenshtein (Lavenshtein Distance Ratio) โดยมีสมการการคำนวนดังนี้

$$Levenshtein\ Distance\ Ratio\ = \frac{Levenshtein\ Distance}{Aligment\ Length} \tag{1}$$

โดยระยะทาง Levenshtein นั้นสามารถเขียนได้ในรูป $lev(a,b)$ ซึ่งจะมีสมการการคำนวนดังนี้

$$lev_{a,b}(i,j) = \begin{cases} max(i,j) & if\ mins(i,j) = 0, \\ min\begin{cases} lev_{a,b}(i-1,j)+1 \\ lev_{a,b}(i,j-1)+1 \\ lev_{a,b}(i-1,j-1)+1_{(a_i \neq b_j)} \end{cases} & otherwise. \end{cases} \tag{2}$$

โดยที่ “a” และ “b” คือ คำที่ 1 และคำที่ 2 ตามลำดับ และสำหรับ “i” และ “j” คือตำแหน่งของตัวอักษรสำหรับคำที่ 1 และคำที่ 2 ตามลำดับ และเงื่อนไข $(a_i \neq b_j)$ คือเงื่อนไขที่จะคืนค่าเป็น 0 เมื่อ $a_i \neq b_j$ เป็นเท็จ

2.1.5 ความคล้ายเชิงความหมาย (Semantic Similarity)

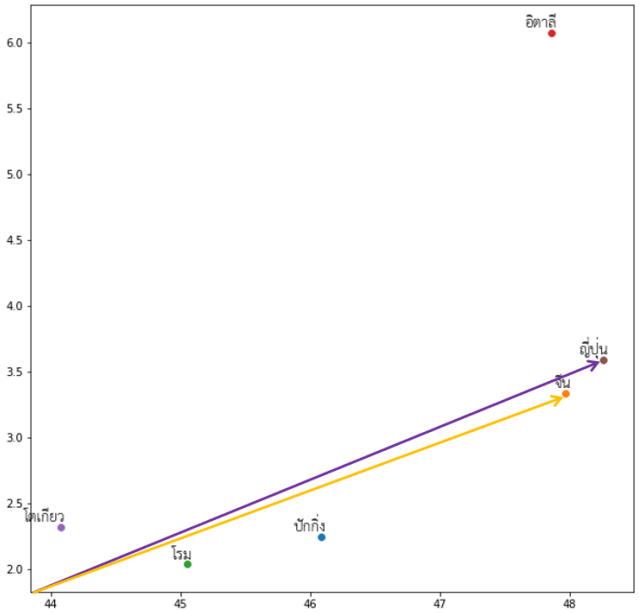
ความคล้ายเชิงความหมาย คือ ระยะห่างระหว่างคำ 2 คำ โดยมีพื้นฐานที่ขึ้นอยู่กับความหมาย ระยะที่สั้นจะแสดงถึงความคล้ายกันมาก และระยะที่ยาวหมายถึงความคล้ายกันน้อย ตัวอย่างเช่น คำว่า “จีน” จะมีระยะทางใกล้กับคำว่า “ญี่ปุ่น” ดังแสดงในภาพที่ 4 เนื่องมาจากมีความหมายเป็นชื่อของประเทศเหมือนกัน โดยทั่วไปแล้วความคล้ายเชิงความหมายสามารถคำนวนได้จากค่าความคล้ายโคไซน์ (Cosine Similarity) โดยมีสมการดังนี้

$$Similarity\ (A,B)\ = \frac{A \cdot B}{\|A\| \times \|B\|} \tag{3}$$

เมื่อ A และ B เป็นองค์ประกอบของเวกเตอร์ A และ B ตามลำดับ

คลังโปรแกรม “PyThaiNLP” เป็นโปรแกรมสำเร็จรูปที่สามารถใช้งานได้ง่าย [23] และยังมีโมดูลการแปลงคำเป็นองค์ประกอบเวกเตอร์ “Thai2Vec” ที่เกิดจากการรวบรวมการซ้อนของคำ (Word Embedding) กว่า 50,000 ชุด และยังมีมิติของคำ (Word Dimension) กว่า 300 มิติ ที่ถูกฝึกสอนด้วย “fast.ai” โดยในงานวิจัยนี้จะใช้คลังโปรแกรม “PyThaiNLP” เพื่อหาค่าความคล้ายเชิงความหมายระหว่างคำ

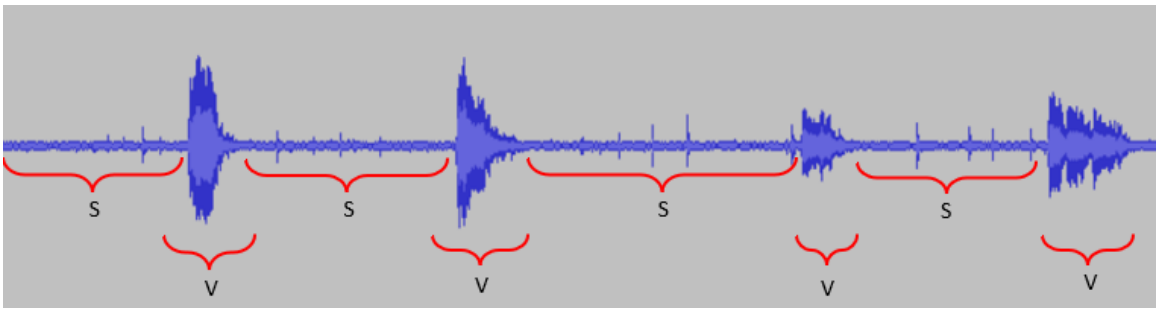
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



ภาพที่ 4 ภาพประกอบความคล้ายเชิงความหมายระหว่างคำ

2.1.6 การตรวจจับความเงียบในเสียงบันทึก

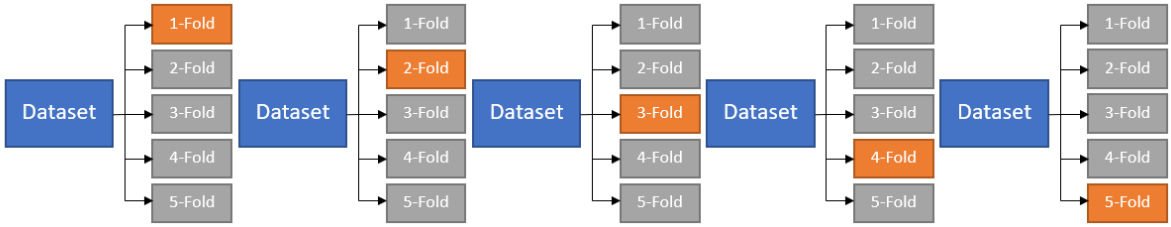
การตรวจจับความเงียบในเสียงบันทึก จะใช้เทคนิคที่เรียกว่าการตรวจจับเสียงพูด (Voice Activity Detection; VAD) โดยที่เสียงในสิ่งบันทึกจะถูกวัดค่าเป็นความดัง มีหน่วยเป็นเดซิเบล (dB) จากแอมพลิจูด (Amplitude) ของคลื่นเสียง มีงานวิจัยที่บ่งชี้ว่าคุณลักษณะที่มีพื้นฐานจากความเงียบ (Silenced-based Feature) เช่น ช่วงของความเงียบทั้งหมด (Total Silence Duration) อัตราการพูด (Word Rate) จำนวนของส่วนเงียบ (Number of Silence Segment) และอื่นๆ เป็นตัวบ่งชี้ของภาวะความรู้คิดบกพร่องเล็กน้อยในแบบประเมินความคล่องทางภาษา [24, 25] ในงานวิจัยนี้คลื่นสียงที่มีแอมพลิจูดต่ำมากจะถูกพิจารณาให้เป็นความเงียบ ดังแสดงในภาพที่ 5



ภาพที่ 5 การตรวจจับความเงียบในเสียงบันทึก

2.1.7 K-Fold Cross Validation

K-Fold Cross Validation เป็นเทคนิคที่ใช้กันอย่างแพร่หลายเพื่อลดความอคติ (Bias) สำหรับแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning Model) โดยทั่วไปแล้วจำนวน K ที่มากขึ้นแสดงถึงความอคติที่ลดลงของแบบจำลองการเรียนรู้ของเครื่องด้วย ชุดข้อมูลจะถูกแบ่งออกเป็น K ส่วนเท่าๆกัน คือ ชุดข้อมูลสำหรับการทดสอบ 1 ส่วน และชุดข้อมูลสำหรับการฝึกสอนจำนวน K-1 ส่วน จากนั้นชุดข้อมูลถูกนำไปฝึกสอน และทดสอบในแบบจำลองเป็นจำนวน K รอบ ดังแสดงในภาพที่ 6 สำหรับงานวิจัยฉบับนี้จะใช้คลังโปรแกรมสำเร็จรูป “sci-kit learn” [26]



ภาพที่ 6 โครงสร้าง K-Fold Cross Validation (K=5)

2.1.8 ลักษณะของคำในภาษาไทย

ในส่วนนี้จะอธิบายลักษณะเฉพาะของคำในภาษาไทยที่เกี่ยวข้องกับการจัดกลุ่มคำที่ใช้ในงานวิจัยนี้ [27] โดยประกอบด้วย 6 ส่วนดังนี้

1. อาการนาม คือ คำนามที่บอกกิริยา หรือการบรรยายการกระทำซึ่งมาจากคำกิริยา หรือคำวิเศษณ์ และมักจะมีรูปของ การ หรือ ความ นำหน้าอยู่ เช่น การบ้าน การยืน การเดินทาง ความรัก ความซื่อสัตย์ เป็นต้น
2. คำประวิสรรชนีย์ คือ การนำพยัญชนะมาผสมสระ อะ และคงรูปสระไว้หลังพยัญชนะ เช่น กระบี่ กระบอง กะทะ กะปิ กระโดด เป็นต้น
3. คำพ้องเสียง คือ คำที่ออกเสียงเหมือนกันแต่มักมีความหมายไม่เหมือนกัน และเขียนต่างกัน เช่น เก้า ก้าว เป็นต้น
4. คำที่เกิดจากสระทั้ง 21 รูป เช่น ไก่ เกิด แก่ กิน กอด เป็นต้น
5. คำที่เกิดจากวรรยุกต์ คือ การนำวรรยุกต์ทั้ง 5 รูป ซึ่งเป็นเครื่องหมายของระดับเสียงในภาษาไทย นำมาเขียนไว้บนพยัญชนะต้นของคำ ทำให้เกิดความหลากหลายและความหมายของเสียง เช่น กู กู่ กู้ กู๊ กู๋
6. คำควบกล้ำ คือ การนำพยัญชนะ 2 ตัว มาเขียนเรียงกันเพื่อสร้างการออกเสียงใหม่ ทำให้เกิดคำและความหมายมากขึ้น เช่น กราบ กลาง กวาง คลาย คลอง เป็นต้น

2.2 งานวิจัยที่เกี่ยวข้อง

จากการทบกวนวรรณกรรมที่เกี่ยวข้องกับการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อยด้วยแบบประเมินความคล่องแคล่วทางภาษา พบว่าคุณลักษณะ (Feature) ที่ใช้กันอย่างแพร่หลายประกอบด้วยพื้นฐาน 3 ส่วน คือ 1.) คุณลักษณะที่ขึ้นอยู่กับความเงียบ (Silenced-Based Feature) 2.) คุณลักษณะที่ขึ้นอยู่กับความคล้าย (Similarity-Based Feature) 3.) คุณลักษณะที่มาจากการจัดกลุ่มคำและการสลับกลุ่มคำ (Clustering & Switching)

2.2.1 งานวิจัยที่เกี่ยวกับคุณลักษณะที่ขึ้นอยู่กับความเงียบ

ความคล่องแคล่วในการพูด (Spontaneous Speech) เป็นค่าที่คำนวนได้จากส่วนเงียบ (Silence Segment) และส่วนเสียง (Voice Segment) ซึ่งเป็นตัวแปรเสริมที่มีประสิทธิภาพและความไวในการตรวจจับความบกพร่องทางสติปัญญาสำหรับแบบประเมินความคล่องทางภาษา Hoffman และคณะ [28] นำเสนอ 4 ตัวแปรเสริมชั่วคราวซึ่งประกอบด้วย 1.) อัตราการออกเสียง (Articulation Rate) 2.) อัตราการลังเล (Hesitation Ratio) 3.) ความเร็วในการพูด (Speech Tempo) 4.) ความผิดพลาดทางไวยากรณ์ (Grammatical Error) ซึ่งผลลัพธ์แสดงให้เห็นว่าตัวแปรเสริมทั้ง 4 ตัว มีประสิทธิภาพในการคัดแยกโรคอัลไซเมอร์ โดยอัตราการลังเลให้ผลลัพธ์ที่ดีที่สุด T´oth และคณะ [29] นำเสนอระบบแยกคุณลักษณะอัตโนมัติด้วยเทคนิคการรู้จำเสียงพูดแบบอัตโนมัติ (Automatic Speech Recognition) โดยให้ผลลัพธ์ความแม่นยำ F1-Score ที่ 0.853 ซึ่งมีความแม่นยำมากพอที่จะใช้เป็นอุปกรณ์เพื่อการคัดกรองสำหรับภาวะความรู้คิดบกพร่องเล็กน้อย T´oth และคณะ [30] ใช้คุณลักษณะที่ขึ้นอยู่กับส่วนเงียบในเสียงบันทึก เช่น จำนวนส่วนเงียบ (Number of Silence Segment) คำแสดงช่วงหยุด (Filler Pause) และความยาวของส่วนเงียบ (Silence Duration) โดยให้ผลลัพธ์ความแม่นยำ F1-Score ที่ 0.788 สำหรับการคัดแยกผู้ป่วยภาวะความรู้คิดบกพร่องเล็กน้อย Edward และคณะ [31] นำเสนอขั้นตอนวิธีที่ขึ้นอยู่กับรูปแบบความเงียบชั่วคราว ซึ่งผลลัพธ์แสดงให้เห็นว่าคุณลักษณะที่ขึ้นอยู่กับความเงียบให้ความแม่นยำดีที่สุด โดยอาศัยฐานข้อมูล “AcceXible” และ “ADReSS” โดยสรุปแล้วงานวิจัยหลายๆชิ้นต่างชี้ไปในทางเดียวกันว่าคุณลักษณะที่ขึ้นอยู่กับความเงียบเป็นตัวบ่งชี้ที่สามารถระบุความบกพร่องทางสติปัญญาสำหรับแบบประเมินความคล่องทางภาษาได้

2.2.2 งานวิจัยที่เกี่ยวกับคุณลักษณะที่ขึ้นอยู่กับความคล้าย

คุณลักษณะที่ขึ้นอยู่กับความคล้ายของคำเป็นอีกเทคนิคหนึ่งได้ผลอย่างดีกับการแยกแยะความบกพร่องทางสติปัญญา Lindsay และคณะ [5] นำเสนอตัววัด (Metric) ความคล้ายซึ่งประกอบด้วยระยะทาง Levenshtein โดยการถ่วงน้ำหนักความคล้ายเชิงหน่วยเสียง (Weighted Phonemic

Similarity) ความคล้ายเชิงความหมาย (Semantic Similarity) และการจัดกลุ่มคำ (Clustering) โดยผลลัพธ์แสดงให้เห็นว่าคุณลักษณะความคล้ายเชิงหน่วยเสียงมีประสิทธิภาพดีที่สุด และคุณลักษณะที่ขึ้นอยู่กับความคล้ายนั้นสามารถเพิ่มความแม่นยำได้ถึง 29 เปอร์เซ็น [6, 21] นอกจากนี้ยังพบงานวิจัยที่เกี่ยวกับภาษาศาสตร์จิตวิทยา (Psycholinguistics) โดยใช้ความคล้ายแบบอักขรวิธี ซึ่งจากผลลัพธ์บ่งชี้ว่าความคล้ายแบบอักขรวิธีมีผลต่อการเข้าถึงศัพท์ (Lexical Access) ในความจำ และในงานวิจัยนี้ยังศึกษาในส่วนของความคล้ายเชิงความหมาย ซึ่งบ่งชี้ว่าความคล้ายชนิดนี้มีความสัมพันธ์กับคะแนนของแบบประเมินความคล่องทางภาษา [19, 32]

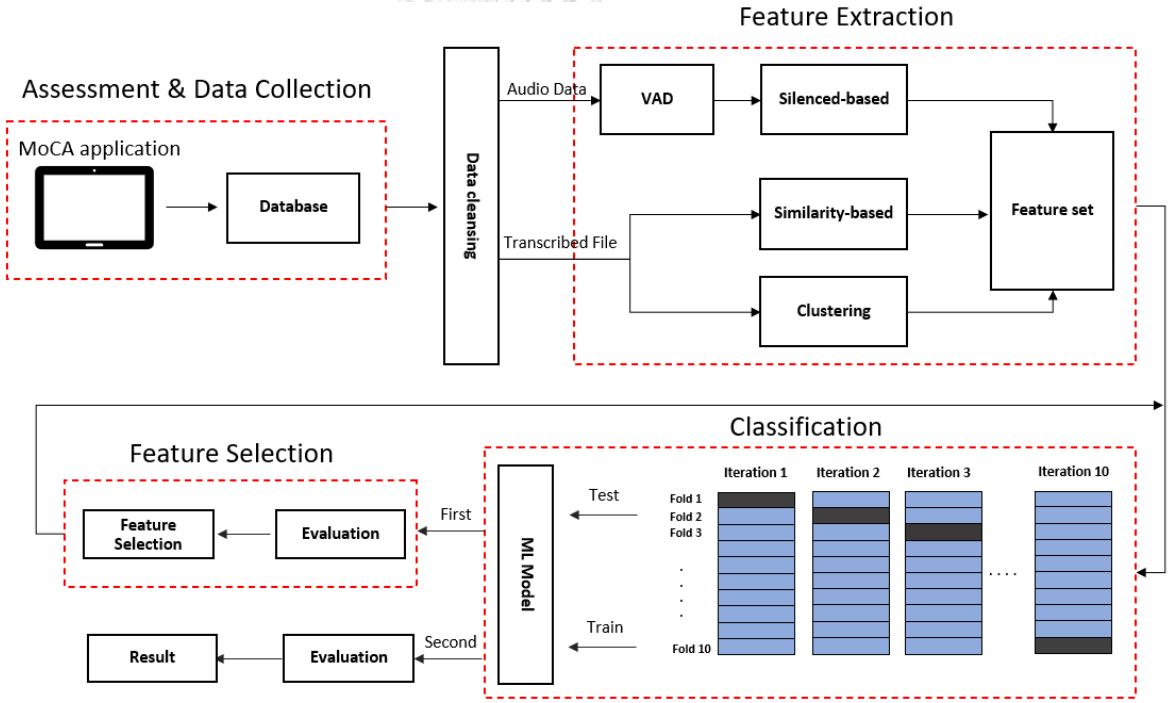
2.2.3 คุณลักษณะที่มาจากการจัดกลุ่มคำและการสลับกลุ่มคำ

Troyer และคณะ [33] นำเสนอคุณลักษณะที่สำคัญสำหรับแบบประเมินความคล่องแคล่วทางภาษา 2 ประเภทได้แก่ การจัดกลุ่มคำ (Clustering) คือ จำนวนของกลุ่มคำที่แบ่งได้จากคำที่ผู้รับการประเมินพูดคำที่ถูกต้องทั้งหมด และการเปลี่ยนกลุ่มคำ (Switching) คือ จำนวนที่ผู้รับการประเมินพูดคำที่ต่างออกไปจากกลุ่มคำก่อนหน้า โดยผลลัพธ์ของงานวิจัยนี้บ่งชี้ว่า การเปลี่ยนกลุ่มคำมีความสัมพันธ์อย่างมีนัยสำคัญกับจำนวนคำพูดที่ถูกต้องในแบบประเมินความคล่องแคล่วทางภาษา Ryan และคณะ [34] เสนองานวิจัยที่สังเกตุภาวะความบกพร่องทางสติปัญญา โดยเปรียบเทียบระหว่างนักกีฬาศิลปะการต่อสู้แบบผสมอาชีพ (Mix Martial Arts) และนักกีฬาสมัครเล่น โดยให้ผู้สมัครทำแบบประเมินความคล่องแคล่วทางภาษา พบว่าจำนวนครั้งของการต่อสู้มีผลต่อทักษะการเปลี่ยนกลุ่มคำอย่างเห็นได้ชัด Muller และคณะ [35] ศึกษาการความสัมพันธ์ของโรคอัลไซเมอร์และภาวะความรู้คิดบกพร่องและสมองเสื่อมเล็กน้อย (Amnestic Mild Cognitive Impairment) ที่มีต่อแบบประเมินความคล่องแคล่วทางภาษา พบว่าผู้ป่วยที่มีภาวะความรู้คิดบกพร่องและสมองเสื่อมเล็กน้อยทำคะแนนได้น้อยกว่าอย่างชัดเจนเมื่อเทียบกับกลุ่มควบคุม นอกจากนี้ยังมีงานวิจัยที่บ่งชี้ว่าผู้เข้ารับการประเมินมีแนวโน้มที่จะพูดคำในกลุ่มเดียวกันมากกว่า และชี้ว่าการเปลี่ยนกลุ่มคำจะสื่อถึงการตัดสินใจอย่างรอบคอบในการเปลี่ยนกลุ่มคำ ซึ่งเหมาะสมที่จะนำมาเป็นตัวชี้วัดมากกว่า **[19]**

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 3
ระเบียบวิธีการวิจัย

งานวิจัยนี้เรานำเสนอเทคนิคใหม่สำหรับการจัดกลุ่มคำในภาษาไทย และเทคนิคการเรียนรู้ของเครื่องร่วมกับการใช้คุณลักษณะแบบผสม ด้วยการคัดเลือกคุณลักษณะจากผลลัพธ์ทางสถิติ เพื่อการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อยสำหรับแบบประเมินความคล่องแคล่วทางภาษาของหน่วยเสียงในภาษาไทย ซึ่งในส่วนนี้จะอธิบายถึงขั้นตอนและวิธีการดำเนินการซึ่งประกอบด้วย 4 ส่วนดังนี้ 1.) การทำแบบประเมินและการเก็บข้อมูล 2.) การสกัดคุณลักษณะ 3.) การจำแนกประเภทและการประเมินผล 4.) การคัดเลือกคุณลักษณะและการประเมินผล ดังแสดงในภาพที่ 7



ภาพที่ 7 ภาพแสดงขั้นตอนการทำวิจัยโดยรวม

3.1 การทำแบบประเมินและการเก็บข้อมูล

ผู้เข้าร่วมจะทำแบบประเมินผ่านโปรแกรมประยุกต์โมคา โดยมีนักจิตวิทยาคอยชี้แจงเกี่ยวกับคำสั่งและให้เวลาเพื่อทำความคุ้นเคยกับอุปกรณ์แท็บเล็ตพีซี เมื่อผู้เข้าร่วมพร้อมแล้วโปรแกรมประยุกต์โมคาจะอ่านคำสั่งจนจบและจะสามารถฟังคำสั่งซ้ำได้ 1 ครั้ง จากนั้นเริ่มเก็บข้อมูลของผู้เข้ารับการประเมินผ่านอุปกรณ์รับสัญญาณอิเล็กทรอนิกส์ ดังนี้ 1.) การเก็บข้อมูลการวาดเส้นด้วยปากกาอิเล็กทรอนิกส์ 2.) การเก็บข้อมูลการแตะด้วยหน้าจอสัมผัส 3.) การเก็บข้อมูลเสียงพูดด้วยไมโครโฟน 4.) การเก็บข้อมูลตัวหนังสือและตัวเลขผ่านการกรอกข้อมูล

งานวิจัยนี้มุ่งเน้นไปที่ข้อมูลประเภทเสียง โดยมีรูปแบบของข้อมูลการเก็บประเภท ".m4a" มีอัตราบิตที่ 44,100 กิโลเฮิร์ต และขนาดประเภท 32 บิต โดยข้อมูลเสียง ข้อมูลส่วนบุคคล และข้อมูลคะแนนจะถูกดาวน์โหลดมาจากฐานข้อมูลบนเซิร์ฟเวอร์ ซึ่งมีจำนวน 100 ตัวอย่าง โดยแบ่งเป็น 2 กลุ่ม คือ 1.) กลุ่มควบคุม คือ ผู้เข้าร่วมที่ได้คะแนนรวมมากว่าหรือเท่ากับ 25 2.) กลุ่มที่มีภาวะความรู้คิดบกพร่องเล็กน้อย คือ ผู้เข้าร่วมที่ได้คะแนนน้อยกว่า 25 ดังแสดงไว้ในตารางที่ 1

ตารางที่ 1 *การแจกแจงข้อมูลประชากรของชุดข้อมูล*

	กลุ่มภาวะความรู้คิดบกพร่องเล็กน้อย	**กลุ่มควบคุม**
เพศชาย	7	10
เพศหญิง	34	49
จำนวนคำที่พูดถูกต้อง (ค่าเฉลี่ย)	3 – 15 (9.61)	2 – 24 (10.1)
พิสัยคะแนนโมคา (ค่าเฉลี่ย)	10 – 24 (21.59)	25 - 29 (27)

3.2 การสกัดคุณลักษณะ

การสกัดคุณลักษณะ คือ การดึงเอาข้อมูลที่มีประโยชน์จากข้อมูลภายในเสียง เพื่อนำไปใช้ในกระบวนสกัดของคุณลักษณะต่างๆ เช่น ความยาวของส่วนเงียบที่เกิดจากการหาผลรวมของช่วงเวลาที่มีเสียงเงียบ เป็นต้น งานวิจัยนี้แบ่งการสกัดคุณลักษณะออกเป็น 2 แนวทาง คือ 1.) การสกัดคุณลักษณะจากข้อมูลเสียง 2.) การสกัดคุณลักษระจากแฟ้มคัดลอกข้อมูล (Transcribed File) ตารางที่ 2 แสดงคุณลักษณะและความหมายของคุณลักษณะทั้งหมดที่ใช้ในงานวิจัยนี้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 2 แสดงรายการคุณลักษณะทั้งหมดในงานวิจัย

Type	Feature	Description	Source
Silenced-based feature	Total silence	Total length of silence during test	Audio File
	Number of Silence segments	Total number of silence segment	
	Average silence between word	Total silence divided by number of silence segment	
	Q1 Silence	Total silence in first 30 seconds of the audio file	
	Q2 Silence	Total silence in last 30 seconds of the audio file	
	Silence before first word	Silence length before participant speaking first word	
	Different silence between Q1 and Q2	Total silence in first 30 second minus last 30 second	
	Total voiced	Total length of voiced during test	
	Number of Voiced segments	Total number of voiced segments	
Similarity-based feature	Mean Orthographic Similarity	Average orthographic similarity value of all words	Transcribed File
	Mean Levenshtein Distance	Average Levenshtein distance ratio of all words	
	Mean Semantic Similarity	Average semantic similarity value of all words	
Clustering & switching	Phonemic Clustering	Group words into phonemic categories	Transcribed File
	Number of Switches	Total number of cluster transition	
Global features	Years of Education	Total year of participant’ s education	Raw data
	Word Count	Total correct word that produced from participant during the task	

จุฬาลงกรณ์มหาวิทยาลัย CHULALONGKORN UNIVERSITY

3.2.1 เทคนิคการจัดกลุ่มคำสำหรับภาษาไทย

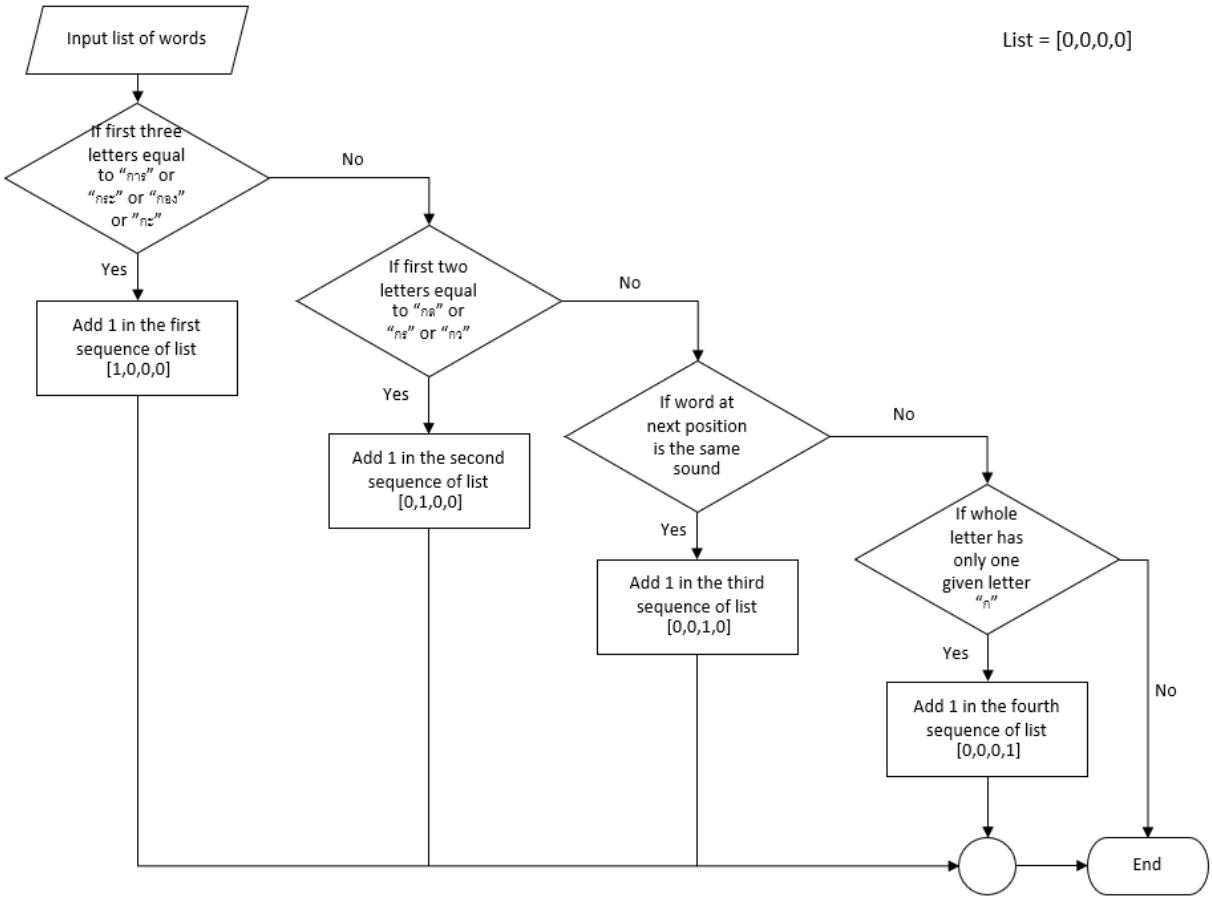
การจัดกลุ่มคำในงานวิจัยนี้มีแนวคิดมาจากความเป็นไปได้ทั้งหมดที่จะเกิดชนิดของคำจากการพูดในขณะที่ผู้เข้าร่วมกำลังทำแบบประเมินความคล่องแคล่วทางภาษาของหน่วยเสียง “ก” โดยเมื่อพิจารณาแล้วจะสรุปชนิคำทั้งหมดแบ่งออกเป็น 4 กลุ่มคำ ดังแสดงไว้ในตารางที่ 3

ตารางที่ 3 แสดงรายการของกลุ่มคำ

กลุ่ม	ชนิดของคำ	ตัวอย่าง
1	กลุ่มอาการนามและคำประวิสรรชนีย์	การบ้าน, การเรียน, กระต่าย, กระเด็น
2	กลุ่มคำควบกล้ำ	กลวง, กราบ, กวาด, กลิ่น
3	กลุ่มคำพ้องเสียง	(ก้าว-เก้า), (กัน-กัณฑ์), (กาน-กาล)
4	กลุ่มคำที่เกิดจากสระ และวรรณยุกต์	ไก่, กู่, เกิด, แก่, เก็บ

ขั้นตอนวิธีสำหรับการจัดกลุ่มคำแสดงไว้ในภาพที่ 8 โดยที่โปรแกรมจะรับรายการคำพูดจากแฟ้มคัดลอกข้อมูล และทำวงวนเพื่อป้อนแต่ละคำพูดให้โปรแกรมเพื่อตรวจสอบคำกับเงื่อนไข เมื่อคำพูดถูกเปรียบเทียบแล้วตรงกับเงื่อนไขใดเงื่อนไขหนึ่ง จะบวกเพิ่มเท่ากับ 1 ให้กับลำดับตัวเลข (Array) ที่ตรงกับเงื่อนไขนั้น ทำขั้นตอนซ้ำจนครบทุกคำพูดในรายการคำพูด สำหรับการให้คะแนนจะพิจารณาจากจำนวนคำในลำดับตัวเลขที่ไม่เท่ากับ 0 เช่น โปรแกรมประมวลผลลำดับตัวเลขได้ผลลัพธ์เป็น [4,5,0,3] จะพิจารณาให้คะแนนเท่ากับ 3 เป็นต้น

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY



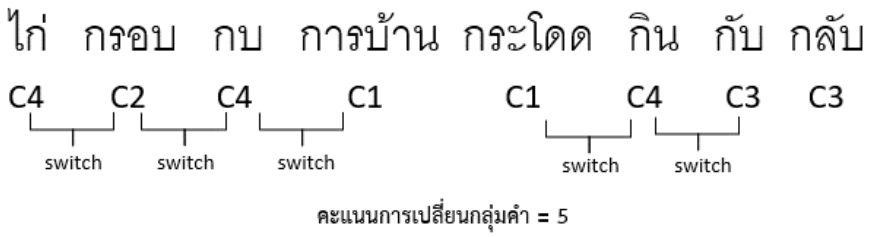
ภาพที่ 8 ขั้นตอนวิธีการของการจัดกลุ่มคำ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.2.2 เทคนิคการเปลี่ยนกลุ่มคำ

เทคนิคนี้เป็นส่วนที่สืบเนื่องมาจากการจัดกลุ่มคำ เมื่อผู้เข้ารับการประเมินพูดคำที่แตกต่างจากกลุ่มคำก่อนหน้าจะพิจารณาให้คะแนนในการเปลี่ยนกลุ่มคำเท่ากับ 1 โดยจะไม่มีเพดานคำแนนเต็มสำหรับการเปลี่ยนกลุ่มคำ ดังแสดงไว้ในภาพที่ 9 โดยอธิบายขั้นตอนวิธีการคำนวนดังนี้

1. พิจารณาให้คำพูดที่ 1 เป็นกลุ่มคำที่ 4
2. พิจารณาให้คำพูดที่ 2 เป็นกลุ่มคำที่ 2 เปรียบเทียบกลุ่มคำระหว่างคำพูดที่ 1 และคำพูดที่ 2 หากเหมือนกัน +0 คะแนน หากไม่เหมือนกัน +1 คะแนน
3. เลื่อนไปพิจารณาระหว่างคำพูดที่ 2 และคำพูดที่ 3 ด้วยกระบวนการเดียวกับข้อ 1 และ 2
4. ทำซ้ำจนครบทุกคำพูด และทำการรวบรวมคะแนน



ภาพที่ 9 ภาพแสดงการนับการเปลี่ยกลุ่มคำ

3.2.3 เทคนิคการใช้ค่า SHAP เพื่อหาความสำคัญของคุณลักษณะ

ค่า SHAP เป็นคลังโปรแกรมประยุกต์ ที่ใช้เพื่อธิบายแบบจำลอง ซึ่งมีมีแนวมโนทัศน์มาจากทฤษฎีเกมส์ (Game Theory) ในงานวิจัยฉบับนี้นำคลังโปรแกรมประยุกต์ SHAP [36] มาปรับใช้ในการอธิบายแบบจำลองที่นำเสนอ เพื่อให้เกิดความเข้าใจในแบบจำลองมากยิ่งขึ้น

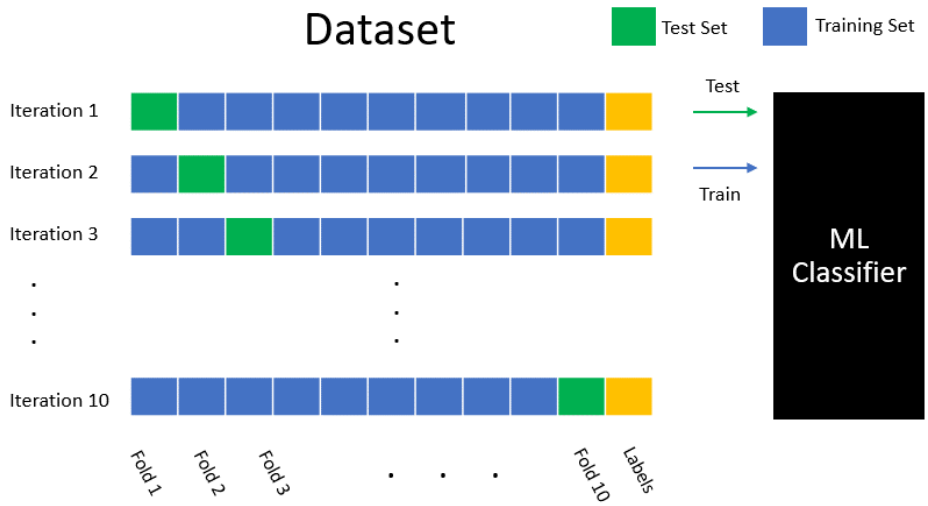
3.3 การจำแนกประเภท

งานวิจัยนี้แบ่งการจำแนกประเภทเป็น 2 ครั้ง โดยครั้งแรก คือ การจำแนกประเภทเพื่อทำการวัดประเมินประสิทธิภาพคุณลักษณะแบบเดี่ยว และครั้งที่สอง คือ การจำแนกประเภทเพื่อวัดประสิทธิภาพของคุณลักษณะสำคัญ ดังแสดงในภาพที่ 7

ซึ่งกระบวนการจำแนกประเภทจะประกอบไปด้วยเทคนิค 3 ส่วน ได้แก่ 1.) การแบ่งข้อมูลด้วย 10-Fold Cross Validation 2.) การฝึกฝนและการทดสอบแบบจำลอง 3.) การปรับละเอียดแบบจำลองสำหรับคุณลักษณะสำคัญ

3.3.1 การแบ่งข้อมูลด้วยเทคนิค 10-Fold Cross Validation

สำหรับการแบ่งข้อมูลด้วย 10-Fold Cross Validation จะกระทำโดยใช้คลังโปรแกรม “scikit-learn” [26] ร่วมกับเทคนิค Stratified K-Folds Cross Validator ซึ่งเป็นคลังโปรแกรมที่นิยมสำหรับภาษา Python เพื่อแบ่งข้อมูลออกเป็น 10 ส่วน โดยที่ 1 ส่วน จะเป็นชุดข้อมูลสำหรับทดสอบ และอีก 9 ส่วน จะเป็นชุดข้อมูลสำหรับฝึกฝนดังภาพที่ 10 และมีตัวอย่างชุดคำสั่งดังภาพที่ 11



ภาพที่ 10 ภาพแสดงการแบ่งข้อมูลด้วย 10-Fold Cross Validation

3.3.2 การฝึกฝนและการทดสอบแบบจำลอง

ในงานวิจัยนี้มีกระบวนการจำแนกประเภทจะดำเนินการทั้งหมด 2 ครั้ง ได้แก่ 1.) ครั้งแรกเป็นการดำเนินการเพื่อค้นหานัยสำคัญของคุณลักษณะที่มีผลต่อความแม่นยำของตัวจำแนก (Classifier) แต่ละชนิด ซึ่งประกอบด้วย Support Vector Machine, XGBoost, Random Forest 2.) ครั้งที่สองเป็นการดำเนินการเพื่อให้ได้ผลลัพธ์ตามสมมติฐานงานวิจัย หลังจากการคัดเลือกคุณลักษณะจากผลลัพธ์การจำแนกประเภทครั้งแรก โดยมีตัวอย่างชุดคำสั่งสำหรับการฝึกฝนและทดสอบแบบจำลองดังภาพที่ 11 ซึ่งในงานวิจัยนี้จะใช้โปรแกรม Google Colab ในการเขียนชุดคำสั่งขึ้นทั้งหมด

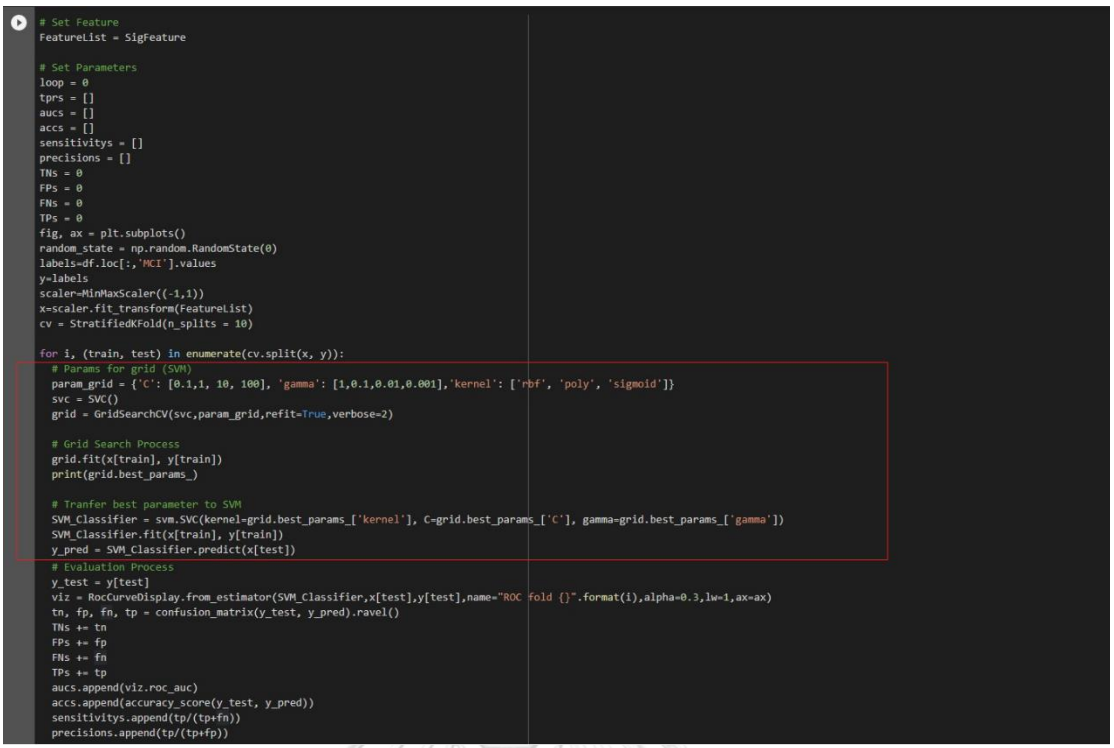
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
# Set Parameters
loop = 0
tprs = []
aucs = []
accs = []
sensitivitys = []
precisions = []
TNs = 0
FPs = 0
FNs = 0
TPs = 0
fig, ax = plt.subplots()
random_state = np.random.RandomState(0)
labels=df.loc[:,'MCI'].values
y=labels
scaler=MinMaxScaler((-1,1))
x=scaler.fit_transform(FeatureList)
SVM_Classifier = svm.SVC(kernel="linear", probability=True, random_state=random_state)
cv = StratifiedKFold(n_splits = 10)

for i, (train, test) in enumerate(cv.split(x, y)):
  SVM_Classifier.fit(x[train], y[train])
  y_pred = SVM_Classifier.predict(x[test])
  y_test = y[test]
  viz = RocCurveDisplay.from_estimator(SVM_Classifier,x[test],y[test],name="ROC fold {}".format(i),alpha=0.3,lw=1,ax=ax)
  tn, fp, fn, tp = confusion_matrix(y_test, y_pred).ravel()
  TNs += tn
  FPs += fp
  FNs += fn
  TPs += tp
  aucs.append(viz.roc_auc)
  accs.append(accuracy_score(y_test, y_pred))
  sensitivitys.append(tp/(tp+fn))
  precisions.append(tp/(tp+fp))
```

ภาพที่ 11 แสดงตัวอย่างชุดคำสั่งการฝึกฝนและทดสอบแบบจำลอง

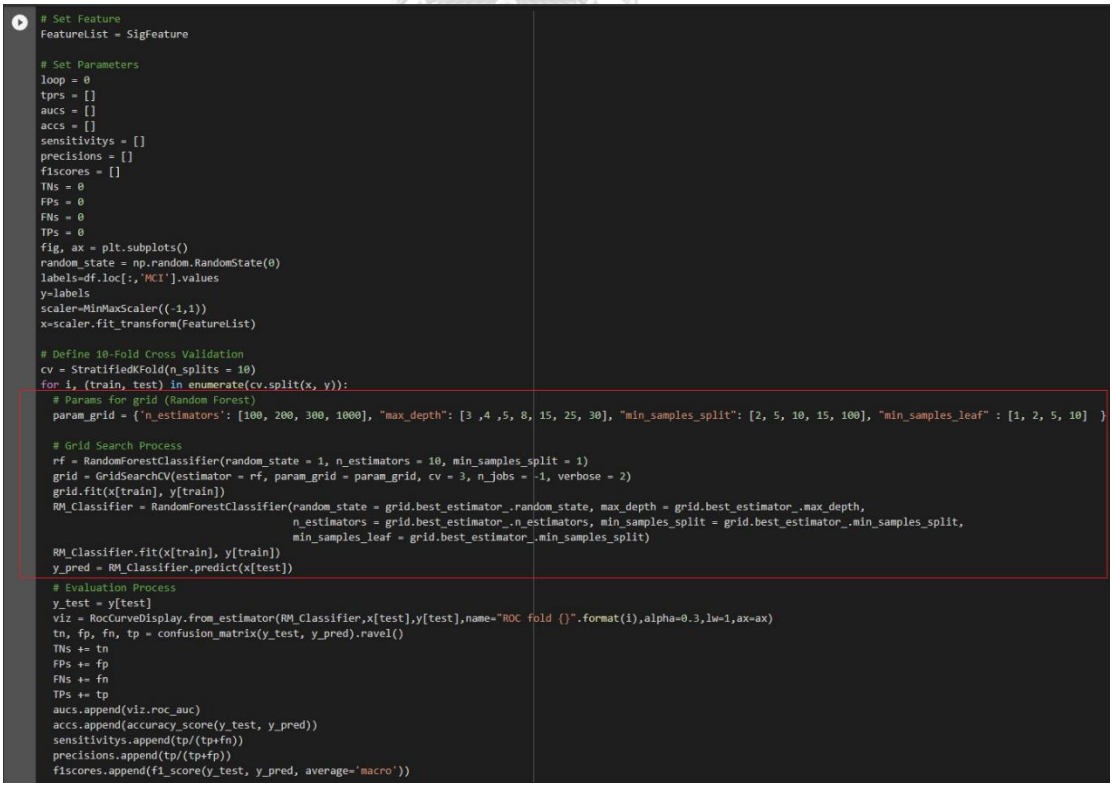
3.3.3 การปรับละเอียดแบบจำลองสำหรับคุณลักษณะสำคัญ

สำหรับการปรับละเอียดแบบจำลอง (Fine-Tuning) ในงานวิจัยนี้มีจุดประสงค์เพื่อค้นหาพารามิเตอร์ที่เหมาะสมที่สุดสำหรับชุดข้อมูลที่ใช้ในงานวิจัย และเพื่อลดความความลำเอียงของพารามิเตอร์จากการกำหนดตั้งค่าแบบตายตัว ในงานวิจัยนี้จะใช้เทคนิค Grid Search และเทคนิค Random Search เพื่อค้นหาพารามิเตอร์ที่เหมาะสมที่สุดในแต่ละการวนรอบของเทคนิค K-Fold Cross Validation จากนั้นโปรแกรมจะส่งผ่านค่าพารามิเตอร์ที่เหมาะสมที่สุดของแต่ละการวนรอบนั้นให้กับตัวจำแนกประเภท เพื่อเข้าสู่กระบวนการจำแนกต่อไป โดยจะมีตัวอย่างชุดคำสั่งสำหรับการปรับละเอียดแบบจำลองในตัวจำแนก Support Vector Machine และ Random Forest ด้วย Grid Search ดังภาพที่ 12, 13 และการปรับละเอียดแบบจำลองในตัวจำแนก XGBoost ด้วย Random Search ดังภาพที่ 14



```
# Set Feature
FeatureList = SigFeature

# Set Parameters
loop = 0
tprs = []
aucs = []
accs = []
sensitivitys = []
precisions = []
TNs = 0
FPs = 0
FNs = 0
TPs = 0
fig, ax = plt.subplots()
random_state = np.random.RandomState(0)
labels=df.loc[:,'MCI'].values
y=labels
scaler=MinMaxScaler((-1,1))
x=scaler.fit_transform(FeatureList)
cv = StratifiedKFold(n_splits = 10)

for i, (train, test) in enumerate(cv.split(x, y)):
  # Params for grid (SVM)
  param_grid = {'C': [0.1,1, 10, 100], 'gamma': [1,0.1,0.01,0.001],'kernel': ['rbf', 'poly', 'sigmoid']}
  svc = SVC()
  grid = GridSearchCV(svc,param_grid,refit=True,verbose=2)

  # Grid Search Process
  grid.fit(x[train], y[train])
  print(grid.best_params_)

  # Tranfer best parameter to SVM
  SVM_Classifier = svm.SVC(kernel=grid.best_params_['kernel'], C=grid.best_params_['C'], gamma=grid.best_params_['gamma'])
  SVM_Classifier.fit(x[train], y[train])
  y_pred = SVM_Classifier.predict(x[test])
  # Evaluation Process
  y_test = y[test]
  viz = RocCurveDisplay.from_estimator(SVM_Classifier,x[test],y[test],name="ROC fold {}".format(i),alpha=0.3,lw=1,ax=ax)
  tn, fp, fn, tp = confusion_matrix(y_test, y_pred).ravel()
  TNs += tn
  FPs += fp
  FNs += fn
  TPs += tp
  aucs.append(viz.roc_auc)
  accs.append(accuracy_score(y_test, y_pred))
  sensitivitys.append(tp/(tp+fn))
  precisions.append(tp/(tp+fp))
```

ภาพที่ 12 แสดงชุดคำสั่งการปรับละเอียดในตัวจำแนก Support Vector Machine



```
# Set Feature
FeatureList = SigFeature

# Set Parameters
loop = 0
tprs = []
aucs = []
accs = []
sensitivitys = []
precisions = []
f1scores = []
TNs = 0
FPs = 0
FNs = 0
TPs = 0
fig, ax = plt.subplots()
random_state = np.random.RandomState(0)
labels=df.loc[:,'MCI'].values
y=labels
scaler=MinMaxScaler((-1,1))
x=scaler.fit_transform(FeatureList)

# Define 10-Fold Cross Validation
cv = StratifiedKFold(n_splits = 10)
for i, (train, test) in enumerate(cv.split(x, y)):
  # Params for grid (Random Forest)
  param_grid = {'n_estimators': [100, 200, 300, 1000], "max_depth": [3 ,4 ,5, 8, 15, 25, 30], "min_samples_split": [2, 5, 10, 15, 100], "min_samples_leaf" : [1, 2, 5, 10]  }

  # Grid Search Process
  rf = RandomForestClassifier(random_state = 1, n_estimators = 10, min_samples_split = 1)
  grid = GridSearchCV(estimator = rf, param_grid = param_grid, cv = 3, n_jobs = -1, verbose = 2)
  grid.fit(x[train], y[train])
  RM_Classifier = RandomForestClassifier(random_state = grid.best_estimator_.random_state, max_depth = grid.best_estimator_.max_depth,
                                         n_estimators = grid.best_estimator_.n_estimators, min_samples_split = grid.best_estimator_.min_samples_split,
                                         min_samples_leaf = grid.best_estimator_.min_samples_split)
  RM_Classifier.fit(x[train], y[train])
  y_pred = RM_Classifier.predict(x[test])
  # Evaluation Process
  y_test = y[test]
  viz = RocCurveDisplay.from_estimator(RM_Classifier,x[test],y[test],name="ROC fold {}".format(i),alpha=0.3,lw=1,ax=ax)
  tn, fp, fn, tp = confusion_matrix(y_test, y_pred).ravel()
  TNs += tn
  FPs += fp
  FNs += fn
  TPs += tp
  aucs.append(viz.roc_auc)
  accs.append(accuracy_score(y_test, y_pred))
  sensitivitys.append(tp/(tp+fn))
  precisions.append(tp/(tp+fp))
  f1scores.append(f1_score(y_test, y_pred, average='macro'))
```

ภาพที่ 13 แสดงชุดคำสั่งการปรับละเอียดในตัวจำแนก Random Forest

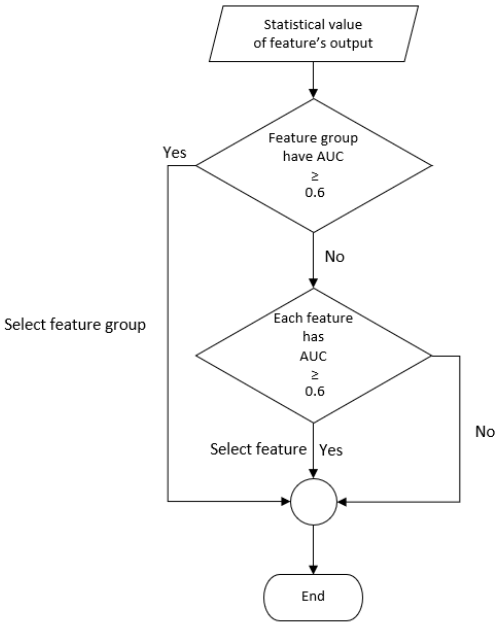
```
# Set Feature
FeatureList = SigFeature

# Set Parameters
loop = 0
tprs = []
aucs = []
accs = []
sensitivitys = []
precisions = []
TNs = 0
FPs = 0
FNs = 0
TPs = 0
fig, ax = plt.subplots()
random_state = np.random.RandomState(0)
labels=df.loc[:,'MCI'].values
y=labels
scaler=MinMaxScaler((-1,1))
x=scaler.fit_transform(FeatureList)

# Define 10-Fold cross validation
cv = StratifiedKFold(n_splits = 10)
for i, (train, test) in enumerate(cv.split(x, y)):
  # Params for Random search (XGBoost)
  params = {'gamma': [0.001, 0.01, 0.1, 1, 10, 100], 'learning_rate': [0.001, 0.001, 0.01, 0.1, 0.2, 0.3], 'max_depth': [1,2,3,4,5,7,9],
            'n_estimators': [50,65,80,100,115,130,150], 'reg_alpha': [0,0.1,0.2,0.4,0.8,1,3,10,100],'reg_lambda': [0,0.1,0.2,0.4,0.8,1,3,10,100],
            'objective': ["binary:logistic","multi:softmax","multi:softprob"]}
  folds = 3
  param_comb = 5

  # Random Search Process
  xgb = XGBClassifier()
  random_search = RandomizedSearchCV(xgb, param_distributions=params, n_iter=param_comb, scoring='roc_auc', n_jobs=4, cv=10, verbose=3, random_state=1001)
  random_search.fit(x[train], y[train])

  # Tranfer best params to XGBoost
  XGB_Classifier = XGBClassifier(gamma= random_search.best_estimator_.gamma, learning_rate= random_search.best_estimator_.learning_rate,
                                 max_depth= random_search.best_estimator_.max_depth, n_estimators= random_search.best_estimator_.n_estimators,
                                 reg_alpha= random_search.best_estimator_.reg_alpha, reg_lambda= random_search.best_estimator_.reg_lambda,
                                 objective= random_search.best_estimator_.objective)
  XGB_Classifier = XGB_Classifier.fit(x[train], y[train])
  y_pred = XGB_Classifier.predict(x[test])

  # Evaluation Process
  y_test = y[test]
  viz = RocCurveDisplay.from_estimator(XGB_Classifier,x[test],y[test],name="ROC fold {}".format(i),alpha=0.3,lw=1,ax=ax)
  tn, fp, fn, tp = confusion_matrix(y_test, y_pred).ravel()
  TNs += tn
  FPs += fp
  FNs += fn
  TPs += tp
  aucs.append(viz.roc_auc)
  accs.append(accuracy_score(y_test, y_pred))
  sensitivitys.append(tp/(tp+fn))
  precisions.append(tp/(tp+fp))
```

ภาพที่ 14 แสดงชุดคำสั่งการปรับละเอียดในตัวจำแนก XGBoost

3.4 การคัดเลือกคุณลักษณะ

การคัดเลือกคุณลักษณะ (Feature Selection) คือ กระบวนการคัดเลือกคุณลักษณะที่มีผลในเชิงบวกต่อแบบจำลองฝึกสอน เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด และลดระยะเวลาในการสกัดคุณลักษณะให้ได้มากที่สุด ในงานวิจัยนี้การคัดเลือกคุณลักษณะจะพิจารณาจากผลลัพธ์จากการจำแนกประเภทครั้งแรก โดยมีขั้นตอนดังที่แสดงดังภาพที่ 15

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



ภาพที่ 15 แผนภูมิสายงานการคัดเลือกคุณลักษณะ

3.5 การประเมินผลแบบจำลอง

การประเมินผลแบบจำลองในงานวิจัยนี้จะคำนวนจาก ค่าความแม่นยำของชุดฝึกสอน และชุดทดสอบนร่วมกัน โดยในแต่ละรอบการวนซ้ำของเทคนิค 10-Fold Cross Validation ค่าความแม่นยำจะถูกนำไปรวมกันจนครบ 10 รอบการวนซ้ำเพื่อหาค่าเฉลี่ยต่อไป สำหรับค่าตัววัดสมรรถนะแบบจำลองจะใช้ค่าพื้นที่ใต้กราฟ ROC (AUC) เป็นหลักสำหรับงานวิจัยนี้ โดยแสดงภาพประกอบในภาพที่ 16 และยังใช้ค่าสภาพจำเพาะ (Specificity) ค่าสภาพไว (Sensitivity) ค่าความแม่นยำ (Accuracy) เพื่อประกอบการพิจารณาในมุมมองอื่นๆ โดยมีสูตรคำนวนดังสมการที่ 6 ไปจนถึงสมการที่ 10 ตามลำดับ

$$AUC(f) = \frac{\sum_{t_0 \in D^0} \sum_{t_1 \in D^1} 1[f(t_0) < f(t_1)]}{|D^0| \cdot |D^1|} \tag{6}$$

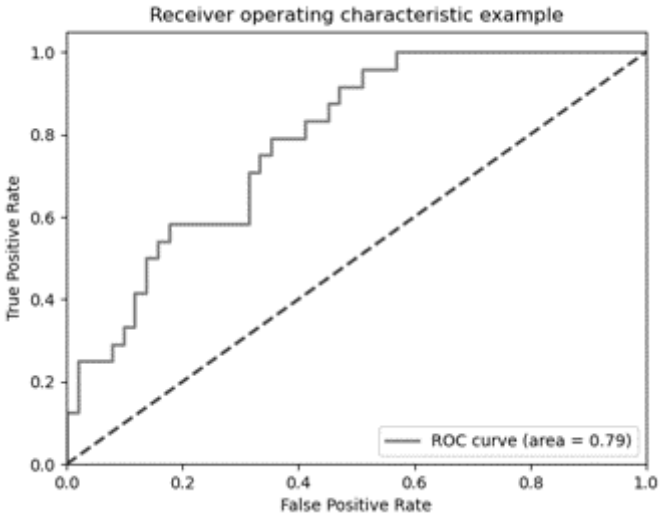
โดยที่ $1[\mathrm{f}(\mathrm{t}_0) < \mathrm{f}(\mathrm{t}_1)]$ หมายถึงฟังก์ชันตัวบ่งชี้ที่คืนค่าเท่ากับ 1 ถ้าเงื่อนไข $f(t_0) < f(t_1)$ เป็นจริง และจะคืนค่า 0 เมื่อเป็นเท็จ และ D^0 คือ เซตของตัวอย่างเชิงลบ และ D^1 คือ เซตของตัวอย่างเชิงบวก

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{7}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{8}$$

$$Acc = \frac{2 \times True\ Positive}{2 \times True\ Positive + False\ Positive + False\ Negative} \tag{9}$$

$$Acc = \frac{True\ Positive\ +\ True\ Negative}{True\ Positive\ + True\ Negative + False\ Positive\ +\ False\ Negative} \tag{10}$$



ภาพที่ 16 ภาพแสดงค่าพื้นที่ใต้กราฟ ROC (AUC)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 4
ผลการทดลอง

ในบทนี้จะกล่าวถึงผลการทดลองซึ่งประกอบด้วย 3 ส่วน คือ ผลการประเมินแบบจำลองเพื่อการคัดเลือกคุณลักษณะ ผลการประเมินแบบจำลองสำหรับคุณลักษณะสำคัญ และการพิจารณาความสำคัญของคุณลักษณะด้วยค่า SHAP

4.1 ผลการประเมินแบบจำลองเพื่อการคัดเลือกคุณลักษณะ

จากข้อมูลทั้งหมด 100 ตัวอย่าง ผ่านการแบ่งข้อมูลด้วยเทคนิค 10-Fold Cross Validation ร่วมกับตัวจำแนกประเภททั้งหมด 3 ชนิด ได้แก่ Support Vector Machine, XGBoost และ Random Forest โดยจะแบ่งการอธิบายผลลัพธ์เป็น 2 กลุ่ม คือ ผลลัพธ์ของคุณลักษณะเดี่ยว และผลลัพธ์ของคุณลักษณะแบบกลุ่ม

4.1.1 ผลลัพธ์ของคุณลักษณะเดี่ยว

ผลลัพธ์จากการวัดประสิทธิภาพของแบบจำลอง โดยใช้คุณลักษณะเดี่ยว ร่วมกับการวนซ้ำด้วยเทคนิค 10-Fold Cross Validation ในตัวจำแนกประเภททั้ง 3 ชนิด เมื่อนำมาหาค่าเฉลี่ยความแม่นยำจากผลรวมของผลลัพธ์สำหรับทุกรอบการวนซ้ำแล้ว แสดงผลลัพธ์ไว้ดังตารางที่ 4

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ **4** แสดงค่าผลลัพธ์ของคุณลักษณะแบบเดี่ยว

Features	XGBoost				
	AUC	Sensitivity	Precision	F1-Score	ACC
Total silence	0.604 ±0.18	0.345 ±0.17	0.483 ±0.26	0.501 ±0.13	0.540 ±0.14
Q1 Silence	0.601 ±0.21	0.445 ±0.22	0.516 ±0.23	0.532 ±0.14	0.560 ±0.15
Q2 Silence	0.595 ±0.24	0.415 ±0.25	0.503 ±0.25	0.557 ±0.16	0.600 ±0.13
Different silence between Q1 and Q2	0.538 ±0.19	0.330 ±0.30	0.319 ±0.26	0.461 ±0.15	0.530 ±0.12
Silence before first word	0.613 ±0.20	0.515 ±0.21	0.569 ±0.25	0.582 ±0.14	0.600 ±0.14
Average silence between word	0.647 ±0.17	0.470 ±0.31	0.414 ±0.22	0.525 ±0.15	0.560 ±0.12
Number of Silence Segment	0.640 ±0.21	0.605 ±0.28	0.521 ±0.24	0.594 ±0.17	0.620 ±0.15
Number of Voiced Segment	0.640 ±0.21	0.605 ±0.28	0.521 ±0.24	0.594 ±0.17	0.620 ±0.15
Total voiced	0.615 ±0.22	0.530 ±0.18	0.585 ±0.20	0.605 ±0.14	0.630 ±0.14
Average Orthographic Similarity	0.494 ±0.11	0.375 ±0.20	0.392 ±0.22	0.481 ±0.16	0.510 ±0.16
Average Levenshtein Distance	0.548 ±0.20	0.435 ±0.20	0.491 ±0.23	0.519 ±0.15	0.550 ±0.16
Average Semantic Similarity	0.495 ±0.17	0.340 ±0.16	0.495 ±0.16	0.533 ±0.13	0.580 ±0.12
Clustering	0.547 ±0.18	0.000 ±0.00	0.000 ±0.00	0.346 ±0.06	0.540 ±0.12
Switching	0.642 ±0.23	0.320 ±0.28	0.318 ±0.24	0.429 ±0.17	0.470 ±0.17
Word Count (Baseline)	0.680 ±0.21	0.530 ±0.21	0.607 ±0.28	0.622 ±0.18	0.640 ±0.17
Years of Education	0.725 ±0.17	0.410 ±0.21	0.783 ±0.32	0.655 ±0.16	0.710 ±0.13

ตารางที่ **4** แสดงค่าผลลัพธ์ของคุณลักษณะแบบเดี่ยว (ต่อ)

Features	Support Vector Machine				
	AUC	Sensitivity	Precision	F1-Score	ACC
Total silence	0.558 ±0.27	0.000 ±0.00	0.000 ±0.00	0.371 ±0.01	0.590 ±0.03
Q1 Silence	0.588 ±0.26	0.200 ±0.24	0.628 ±0.34	0.478 ±0.15	0.600 ±0.12
Q2 Silence	0.494 ±0.11	0.000 ±0.00	0.000 ±0.00	0.371 ±0.01	0.590 ±0.03
Different silence between Q1 and Q2	0.488 ±0.26	0.275 ±0.26	0.738 ±0.35	0.531 ±0.18	0.630 ±0.14
Silence before first word	0.449 ±0.21	0.075 ±0.11	0.625 ±0.41	0.426 ±0.10	0.600 ±0.06
Average silence between word	0.402 ±0.31	0.000 ±0.00	0.000 ±0.00	0.362 ±0.02	0.570 ±0.05
Number of Silence Segment	0.715 ±0.24	0.375 ±0.28	0.507 ±0.34	0.553 ±0.21	0.610 ±0.18
Number of Voiced Segment	0.715 ±0.24	0.375 ±0.28	0.507 ±0.34	0.559 ±0.20	0.620 ±0.16
Total voiced	0.521 ±0.29	0.000 ±0.00	0.000 ±0.00	0.358 ±0.03	0.560 ±0.07
Average Orthographic Similarity	0.597 ±0.24	0.000 ±0.00	0.000 ±0.00	0.371 ±0.01	0.590 ±0.03
Average Levenshtein Distance	0.564 ±0.21	0.000 ±0.00	0.000 ±0.00	0.371 ±0.01	0.590 ±0.03
Average Semantic Similarity	0.427 ±0.20	0.000 ±0.00	0.000 ±0.00	0.356 ±0.04	0.560 ±0.09
Clustering	0.522 ±0.18	0.000 ±0.00	0.000 ±0.00	0.371 ±0.01	0.590 ±0.03
Switching	0.690 ±0.22	0.300 ±0.29	0.454 ±0.22	0.498 ±0.19	0.570 ±0.17
Word Count (Baseline)	0.754 ±0.23	0.580 ±0.23	0.683 ±0.29	0.694 ±0.16	0.720 ±0.14
Years of Education	0.725 ±0.17	0.320 ±0.23	0.759 ±0.32	0.590 ±0.17	0.670 ±0.13

ตารางที่ 4 แสดงค่าผลลัพธ์ของคุณลักษณะแบบเดี่ยว (ต่อ)

Features	Random Forest				
	AUC	Sensitivity	Precision	F1-Score	ACC
Total silence	0.606 ±0.17	0.415 ±0.19	0.501 ±0.27	0.554 ±0.15	0.590 ±0.14
Q1 Silence	0.609 ±0.21	0.355 ±0.21	0.630 ±0.32	0.565 ±0.15	0.620 ±0.14
Q2 Silence	0.622 ±0.22	0.270 ±0.24	0.415 ±0.36	0.488 ±0.17	0.560 ±0.14
Different silence between Q1 and Q2	0.473 ±0.24	0.205 ±0.28	0.511 ±0.41	0.468 ±0.19	0.590 ±0.16
Silence before first word	0.631 ±0.23	0.265 ±0.23	0.518 ±0.34	0.515 ±0.12	0.610 ±0.07
Average silence between word	0.694 ±0.16	0.595 ±0.21	0.582 ±0.18	0.604 ±0.11	0.620 ±0.11
Number of Silence Segment	0.628 ±0.21	0.515 ±0.24	0.497 ±0.24	0.547 ±0.15	0.570 ±0.14
Number of Voiced Segment	0.628 ±0.21	0.515 ±0.24	0.497 ±0.24	0.547 ±0.15	0.570 ±0.14
Total voiced	0.607 ±0.22	0.530 ±0.18	0.567 ±0.22	0.613 ±0.16	0.630 ±0.16
Average Orthographic Similarity	0.549 ±0.14	0.175 ±0.16	0.446 ±0.36	0.445 ±0.10	0.550 ±0.09
Average Levenshtein Distance	0.488 ±0.18	0.245 ±0.22	0.442 ±0.36	0.486 ±0.18	0.560 ±0.17
Average Semantic Similarity	0.487 ±0.19	0.220 ±0.18	0.639 ±0.43	0.508 ±0.17	0.600 ±0.16
Clustering	0.505 ±0.20	0.020 ±0.06	0.250 ±0.43	0.361 ±0.08	0.540 ±0.12
Switching	0.640 ±0.22	0.320 ±0.28	0.368 ±0.31	0.441 ±0.17	0.490 ±0.16
Word Count (Baseline)	0.696 ±0.22	0.485 ±0.21	0.593 ±0.29	0.609 ±0.19	0.630 ±0.19
Years of Education	0.720 ±0.17	0.370 ±0.21	0.783 ±0.32	0.629 ±0.16	0.690 ±0.13

จากตารางที่ 4 พบว่าคุณลักษณะที่มีผลต่อความแม่นยำในการจำแนกประเภทมากที่สุด 5 ลำดับแรก คือ จำนวนปีการศึกษา ค่าเฉลี่ยความเงียบระหว่างคำ จำนวนช่วงของความเงียบ จำนวนช่วงของเสียงพูด และความเงียบก่อนเริ่มพูดคำแรก โดยแบ่งอธิบายตามตัวจำแนกประเภท 3 ชนิด คือ1.) ค่า AUC เท่ากับ 0.725, 0.647, 0.640, 0.640 ในตัวจำแนกประเภท XGBoost ตามลำดับ 2.) ค่า AUC เท่ากับ 0.725, 0.402, 0.715, 0.715 ในตัวจำแนกประเภท Support Vector Machine ตามลำดับ 3.) ค่า AUC เท่ากับ 0.725, 0.694, 0.628, 0.628 ในตัวจำแนกประเภท Random Forest ตามลำดับ

4.1.2 ผลลัพธ์ของคุณลักษณะแบบกลุ่ม

ผลลัพธ์จากการวัดประสิทธิภาพแบบจำลองโดยใช้คุณลักษณะแบบกลุ่ม ได้แก่ กลุ่มคุณลักณะที่ขึ้นอยู่กับความเงียบ กลุ่มคุณลักษณะที่ขึ้นอยู่กับความคล้าย การจัดกลุ่มคำ และการเปลี่ยนกลุ่มคำ ร่วมกับเทคนิค 10-Fold Cross Validation ในตัวจำแนกประเภททั้ง 3 ชนิด เมื่อนำมาหาค่าเฉลี่ย ความแม่นยำสำหรับทุกรอบการวนซ้ำแล้ว จะแสดงผลลัพธ์ดังตารางที่ 5

ตารางที่ 5 แสดงค่าผลลัพธ์ของคุณลักษณะแบบกลุ่ม

Classifier	Feature	AUC	Sensitivity	Precision	F1-Score	Acc
XGBoost	Silenced-based features	0.654 ±0.19	0.410 ±0.24	0.513 ±0.23	0.545 ±0.16	0.590 ±0.14
	Similarity-based features	0.468 ±0.18	0.360 ±0.21	0.484 ±0.31	0.467 ±0.16	0.500 ±0.18
	Clustering & switching	0.626 ±0.23	0.490 ±0.25	0.459 ±0.24	0.520 ±0.19	0.540 ±0.19
	All features	0.629 ±0.23	0.505 ±0.26	0.513 ±0.23	0.600 ±0.16	0.630 ±0.13
Support Vector Machine	Silenced-based features	0.696 ±0.21	0.190 ±0.17	0.687 ±0.43	0.621 ±0.20	0.620 ±0.14
	Similarity-based features	0.457 ±0.17	0.390 ±0.16	0.535 ±0.26	0.352 ±0.04	0.570 ±0.16
	Clustering & switching	0.679 ±0.21	0.350 ±0.28	0.521 ±0.18	0.532 ±0.17	0.600 ±0.15
	All features	0.687 ±0.24	0.350 ±0.32	0.592 ±0.38	0.575 ±0.17	0.650 ±0.17
Random Forest	Silenced-based features	0.659 ±0.20	0.410 ±0.32	0.392 ±0.33	0.510 ±0.22	0.550 ±0.20
	Similarity-based features	0.509 ±0.18	0.270 ±0.14	0.568 ±0.32	0.483 ±0.13	0.550 ±0.14
	Clustering & switching	0.633 ±0.23	0.490 ±0.25	0.415 ±0.36	0.453 ±0.14	0.520 ±0.14
	All features	0.617 ±0.23	0.430 ±0.32	0.435 ±0.32	0.542 ±0.22	0.580 ±0.19

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

จากตารางที่ 5 เมื่อจัดคุณลักษณะแต่ละชนิดรวมเป็นกลุ่มโดยแบ่งแยกแต่ละประเภทพบว่ากลุ่มคุณลักษณะที่ขึ้นอยู่กับความเงียบให้ผลลัพธ์ที่ดีที่สุด ตามมาด้วยการจัดกลุ่มคำและการเปลี่ยนกลุ่มคำ ส่วนกลุ่มคุณลักษณะที่ขึ้นอยู่กับความคล้ายทำผลงานได้ในลำดับสุดท้าย

โดยเมื่อพิจาณาจากผลลัพธ์ที่ได้แล้ว กลุ่มคุณลักษณะที่ขึ้นอยู่กับความเงียบ การจัดกลุ่มคำและการเปลี่ยนกลุ่มคำเป็นกลุ่มที่มีความเหมาะสมที่จะถูกเลือกเป็นกลุ่มคุณลักษณะสำคัญ ซึ่งจะถูกนำไปไปใช้ต่อในเทคนิคการเรียนรู้ของเครื่อง เนื่องจากมีประสิทธิภาพในการจำแนกประเภทที่ค่อนข้างดีโดยมีค่า AUC เท่ากับ 0.655 และ 0.626 ตามลำดับ ในขณะที่กลุ่มคุณลักษณะที่ขึ้นอยู่กับความคล้ายจะถูกคัดออกไป เนื่องจากประสิทธิภาพในการจำแนกประเภทค่อนข้างต่ำที่ค่า AUC เท่ากับ 0.468 เท่านั้น

4.2 ผลการประเมินแบบจำลองสำหรับคุณลักษณะสำคัญ

เมื่อนำคุณลักษณะที่ได้จากการคัดเลือกกลับเข้าสู่กระบวนการฝึกสอนและทดสอบอีกครั้ง โดยมีขั้นตอนการแบ่งชั้นข้อมูลแบบ 10-Fold Cross Validation ในตัวจำแนกประเภททั้ง 3 ชนิด และมีกระบวนการปรับละเอียดในทุกการวนรอบเพื่อให้ได้พารามิเตอร์ที่เหมาะสมที่สุด โดยผลลัพธ์จากการประเมินประสิทธิภาพแบบจำลอง ได้แก่ AUC Sensitivity Precision และ Accuracy จะแสดงเป็นค่าเฉลี่ยในตารางที่ 6 และมี Confusion Matrix ดังภาพที่ 17

ตารางที่ 6 แสดงค่าผลลัพธ์ของคุณลักษณะสำคัญ

Features	AUC	Sensitivity	Precision	F1-Score	Acc
XGBoost					
Word Count (Baseline)	0.680 ±0.21	0.530 ±0.21	0.607 ±0.28	0.622 ±0.18	0.640 ±0.17
Significant features	0.720 ±0.16	0.510 ±0.25	0.680 ±0.22	0.638 ±0.17	0.680 ±0.15
Support Vector Machine					
Word Count (Baseline)	0.754 ±0.23	0.555 ±0.22	0.688 ±0.22	0.694 ±0.16	0.710 ±0.17
Significant features	0.708 ±0.19	0.385 ±0.26	0.717 ±0.39	0.635 ±0.20	0.700 ±0.15
Random Forest					
Word Count (Baseline)	0.696 ±0.22	0.485 ±0.21	0.593 ±0.29	0.609 ±0.19	0.630 ±0.19
Significant features	0.758 ±0.23	0.550 ±0.31	0.610 ±0.29	0.660 ±0.21	0.660 ±0.22

จากตารางที่ 6 ผลลัพธ์ของคุณลักษณะสำคัญที่ผ่านการคัดเลือก พบว่าคุณลักษณะสำคัญมีประสิทธิภาพที่ดีกว่าแบบจำลองพื้นฐานในตัวจำแนกประเภท XGBoost และ Random Forest ที่ค่า AUC เท่ากับ 0.720 และ 0.758 ตามลำดับ แต่เมื่อพิจารณาผลลัพธ์สำหรับตัวจำแนกประเภท Support Vector Machine พบว่าคุณลักษณะสำคัญ ทำประสิทธิได้ด้อยกว่าแบบจำลองพื้นฐานที่ค่า AUC เท่ากับ 0.708

จากผลลัพธ์ที่แสดงให้เห็นนี้นำไปสู่การยืนยันสมมุติฐานของงานวิจัยที่ว่า การใช้ข้อมูลแบบประเมินความคล่องทางภาษามีความเป็นไปได้ในการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อย และการใช้คุณลักษณะแบบผสมจากการคัดเลือกโดยค่าทางสถิติช่วยเพิ่มความแม่นยำในการจำแนกประเภทให้มากขึ้นได้

XGBoost

	MCI	Normal
MCI	21	20
Normal	12	47

SVM

	MCI	Normal
MCI	14	27
Normal	4	55

Random Forest

	MCI	Normal
MCI	23	18
Normal	16	43

ภาพที่ 17 Confusion Matrix ของคุณลักษณะสำคัญ

4.3 การพิจารณาความสำคัญของคุณลักษณะด้วยค่า SHAP

ในขั้นตอนนี้เป็นกระบวนการนำค่า SHAP มาประยุกต์เข้ากับแบบจำลองเพื่ออธิบายขีดความสามารถของคุณลักษณะสำคัญ ว่าคุณลักษณะแต่ละชนิดมีผลต่อการจำแนกประเภทใน 3 ตัวจำแนกประเภทอย่างไร

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

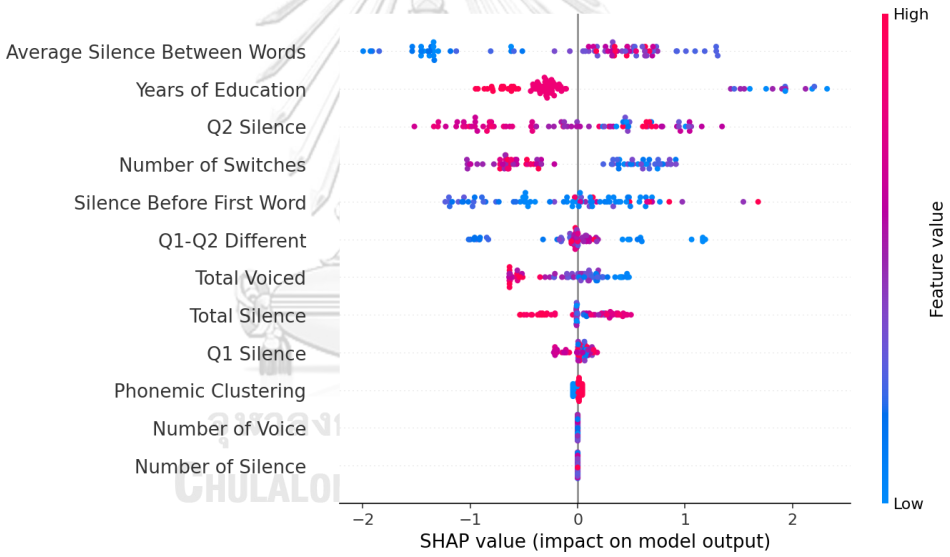
4.3.1 ค่า SHAP ในตัวจำแนก XGBoost

จากภาพที่ 18 พบว่า 5 ลำดับแรกของคุณลักษณะที่มีผลต่อตัวจำแนกสูงที่สุดได้แก่ ค่าเฉลี่ยความเงียบระหว่างคำ จำนวนปีการศึกษา ความเงียบในย่านที่ 2 การเปลี่ยนกลุ่มคำ และความเงียบก่อนเริ่มพูดคำแรก ตามลำดับ เมื่อจารณาอย่างละเอียดลงไปแล้วพบว่าคุณลักษณะที่โดดเด่นที่สุดมีเพียง 3 ชนิด คือ ค่าเฉลี่ยความเงียบระหว่างคำ จำนวนปีการศึกษา และการเปลี่ยนกลุ่มคำ โดยมีคำอธิบายดังนี้

- สำหรับจำนวนปีการศึกษา ค่าที่มาก (จุดสีแดง) จะมีผลในทางลบต่อตัวจำแนก XGBoost ในทางกลับกัน ค่าที่น้อย (จุดสีฟ้า) จะมีผลในทางบวกสูงต่อตัวจำแนก และค่าปานกลาง (จุดสีม่วง) จำนวนหนึ่งมีผลต่อตัวจำแนกในทางบวกสูงๆ สังเกตุได้อีกว่าค่าที่

มาก-น้อย มีการแยกตัวออกจากจุดกึ่งกลางอย่างชัดเจน ทำให้จำนวนปีการศึกษาถือเป็นคุณลักษณะที่มีประสิทธิภาพดีเยี่ยม

- สำหรับค่าเฉลี่ยความเงียบระหว่างคำ ค่าที่น้อย (จุดสีฟ้า) โดยส่วนใหญ่จะส่งผลต่อตัวจำแนก XGBoost ในทางลบ และยังมีค่าที่น้อยบางส่วนที่ส่งผลต่อตัวจำแนกในทางบวกอีกด้วย ค่าที่มาก (จุดสีแดง) ส่งผลต่อตัวจำแนกในทางบวกเช่นกัน ทำให้ค่าเฉลี่ยความเงียบระหว่างคำถือเป็นคุณลักษณะที่มีประสิทธิภาพดีปานกลาง
- สำหรับการเปลี่ยนกลุ่มคำ ค่าที่มาก (จุดสีแดง) จะมีผลในทางลบต่อตัวจำแนก XGBoost ค่าที่น้อย (จุดสีฟ้า) จะมีผลในทางบวกต่อตัวจำแนก ค่าปานกลาง (จุดสีม่วง) จำนวนหนึ่งมีผลต่อตัวจำแนกทั้งด้านบวก และลบคละกันไป สังเกตุว่าค่าที่มาก-น้อยทั้งหมดแสดงการแยกตัวจากจุดกึ่งกลางอย่างชัดเจน ทำให้การเปลี่ยนกลุ่มคำถือเป็นคุณลักษณะที่มีประสิทธิภาพดี

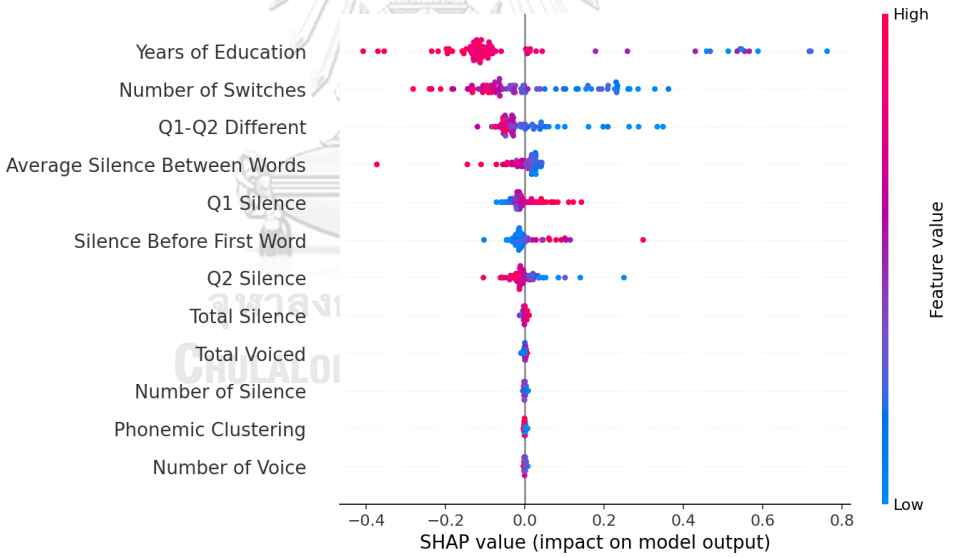


ภาพที่ 18 ค่า SHAP ของตัวจำแนก XGBoost

4.3.2 ค่า SHAP ในตัวจำแนก Support Vector Machine

จากภาพที่ 19 พบว่า 5 ลำดับแรกของคุณลักษณะที่มีผลต่อตัวจำแนกสูงที่สุดได้แก่ จำนวนปีการศึกษา จำนวนการเปลี่ยนกลุ่มคำ ความต่างความเงียบระหว่างย่านที่ 1 และย่านที่ 2 ค่าเฉลี่ยความเงียบระหว่างคำ และความเงียบย่านที่ 1 ตามลำดับ เมื่อพิจารณาอย่างละเอียดแล้วพบว่ามีคุณสมบัติที่โดดเด่นทั้งหมด 3 ชนิดด้วยกัน คือ จำนวนปีการศึกษา จำนวนการเปลี่ยนกลุ่มคำ และความต่างความเงียบระหว่างย่านที่ 1 และย่านที่ 2 โดยมีรายละเอียดดังนี้

- สำหรับจำนวนปีการศึกษา ค่าที่มาก (จุดสีแดง) จะส่งผลในทางลบต่อตัวจำแนก Support Vector Machine และค่าที่น้อย (จุดสีฟ้า) จะส่งผลในทางบวกมากต่อตัวจำแนก ซึ่งสังเกตุได้ว่าค่าน้อย-มากมีการแยกตัวออกจากกันอย่างชัดเจน ทำให้จำนวนปีการศึกษาถือเป็นคุณลักษณะที่มีประสิทธิภาพดีที่สุด
- สำหรับจำนวนการเปลี่ยนกลุ่มคำ ค่าที่มาก (จุดที่แดง) จะส่งผลในทางลบต่อตัวจำแนก Support Vector Machine และค่าที่น้อย (จุดสีฟ้า) จะส่งผลในทางบวกต่อตัวจำแนก และยังมีค่าปานกลาง (จุดสีม่วง) จำนวนหนึ่งส่งผลต่อตัวจำแนกทางด้านลบน้อยๆ ทำให้จำนวนการเปลี่ยนกลุ่มคำถือเป็นคุณลักษณะที่มีประสิทธิภาพดีปานกลาง
- สำหรับความต่างความเงียบระหว่างย่านที่ 1 และย่านที่ 2 ค่าที่มาก (จุดสีแดง) จะส่งผลในทางลบต่อตัวจำแนก Support Vector Machine ค่าที่น้อย (จุดสีฟ้า) จะส่งผลต่อตัวจำแนกในทางบวก และยังมีค่าปานกลาง (จุดสีม่วง) จะส่งผลต่อตัวจำแนกโดยเกาะตัวอยู่ระหว่างจุดกึ่งกลางไปจนถึงด้านลบ ทำให้ความต่างความเงียบระหว่างย่านที่ 1 และย่านที่ 2 ถือเป็นคุณลักษณะมีประสิทธิภาพปานกลาง



ภาพที่ 19 ค่า SHAP ของตัวจำแนก Support Vector Machine

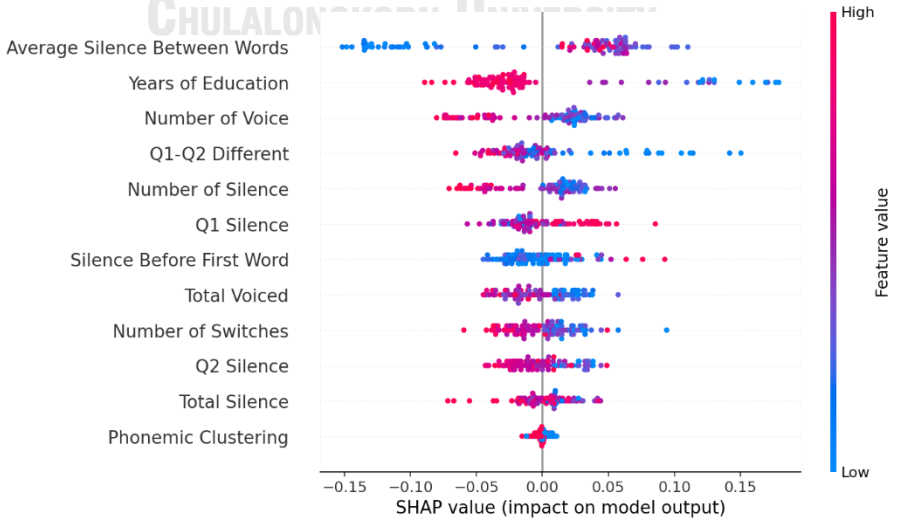
4.3.3 ค่า SHAP ในตัวจำแนก Random Forest

จากภาพที่ 20 พบว่า 5 ลำดับแรกของคุณลักษณะที่มีผลต่อตัวจำแนกสูงที่สุดได้แก่ ค่าเฉลี่ยความเงียบระหว่างคำ จำนวนปีการศึกษา จำนวนของช่วงเสียงพูด ความต่างความเงียบระหว่างย่านที่ 1 กับย่านที่ 2 และจำนวนช่วงเงียบ ตามลำดับ เมื่อจารพิจารณาอย่างละเอียดแล้วพบว่ามีคุณสมบัติที่

โดดเด่นทั้งหมด 4 ชนิดด้วยกัน คือ จำนวนปีการศึกษา ค่าเฉลี่ยความเงียบระหว่างคำ จำนวนของช่วงเสียงพูด จำนวนช่วงเงียบ โดยมีรายละเอียดดังนี้

- สำหรับจำนวนปีการศึกษา ค่าที่มาก (จุดสีแดง) ส่งผลต่อตัวจำแนก Random Forest ทางด้านลบมาก ค่าที่น้อย (จุดสีฟ้า) ส่งผลต่อตัวจำแนกทางด้านบวกมาก และค่าปานกลาง (จุดสีม่วง) ส่งผลต่อตัวจำแนกทางด้านบวกเช่นกัน ทำให้จำนวนปีการศึกษาถือเป็นคุณลักษณะที่มีประสิทธิภาพดีเยี่ยม
- สำหรับค่าเฉลี่ยความเงียบระหว่างคำ ค่าที่น้อย (จุดสีฟ้า) จะส่งผลต่อตัวจำแนกในทางลบมาก ค่าที่มาก (จุดสีแดง) จะส่งผลต่อตัวจำแนกในทางบวก และค่าปานกลาง (จุดสีม่วง) จะส่งผลต่อตัวจำแนกในทางบวกเช่นกัน ทำให้ค่าเฉลี่ยความเงียบระหว่างคำถือเป็นคุณลักษณะที่มีประสิทธิภาพดี
- สำหรับจำนวนของช่วงเสียงพูด ค่าที่มาก (จุดสีแดง) จะส่งผลต่อตัวจำแนกในทางลบ ค่าที่น้อย (จุดสีฟ้า) จะส่งผลต่อตัวจำแนกในทางบวก และค่าปานกลาง (จุดสีม่วง) โดยส่วนมากจะส่งผลต่อตัวจำแนกในทางบวกและมีบางส่วนที่ส่งผลต่อตัวจำแนกในทางปานกลางถึงลบ ทำให้จำนวนช่วงเสียงพูดเป็นคุณลักณะที่มีประสิทธิภาพพอใช้ได้
- สำหรับจำนวนของช่วงเงียบ ค่าที่มาก (จุดสีแดง) ส่งผลต่อตัวจำแนกในด้านลบ ค่าที่น้อย (จุดสีฟ้า) ส่งผลต่อตัวจำแนกในด้านบวก และค่าปานกลาง (จุดสีม่วง) ส่วนมากส่งผลต่อตัวจำแนกในด้านบวกและมีบางส่วนที่ส่งผลในด้านลบด้วย ทำให้จำนวนของช่วงเงียบเป็นคุณลักษณะที่มีประสิทธิภาพพอใช้ได้ และคุณลักษณะนี้มีพฤติกรรมใกล้เคียงกับจำนวนของช่วงเสียงพูด

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



ภาพที่ 20 ค่า SHAP ของตัวจำแนก Random Forest

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

วิทยานิพนธ์ฉบับนี้นำเสนอเทคนิคการเรียนรู้ของเครื่องเพื่อการวินิจฉัยภาวะความรู้คิดบกพร่องเล็กน้อย โดยอาศัยข้อมูลจากแบบทดสอบย่อยความคล่องทางภาษาของหน่วยเสียงฉบับภาษาไทย ภายในโปรแกรมประยุกต์โมคา งานวิจัยฉบับนี้ใช้ข้อมูลที่น่าเชื่อถือจำนวน 100 ตัวอย่าง จากการทำแบบประเมินจริงของผู้เข้ารับการประเมิน โดยมีบุคลากรทางการแพทย์กำกับดูแลการดำเนินการของกระบวนการประเมินในทุกขั้นตอน ณ โรงพยาบาลจุฬาลงกรณ์ ฝ่ายจิตเวชศาสตร์

ในบทนี้จะกล่าวถึงผลการวิจัย โดยสรุปตามการวิเคราะห์จากผลลัพธ์ที่เกิดขึ้น ได้แก่ 1.) การสรุปผลลัพธ์ในภาพรวม 2.) ปัจจัยของคุณลักษณะแต่ละชนิดต่อภาวะความรู้คิดบกพร่องเล็กน้อย 3.) ผลลัพธ์จากการคัดเลือกคุณลักษณะสำคัญ 4.) ข้อจำกัดของงานวิจัย 5.) ข้อเสนอแนะสำหรับการนำงานวิจัยไปศึกษาต่อ

5.1 การสรุปผลลัพธ์ในภาพรวม

ผลลัพธ์จากวิจัยชิ้นนี้แสดงให้เห็นว่า การใช้เทคนิคการเรียนรู้ของเครื่องร่วมกับการผสมผสานคุณลักษณะสำคัญจากกระบวนการคัดเลือกคุณลักษณะให้ผลเป็นที่น่าพึงพอใจ และดีกว่าแบบจำลองพื้นฐานที่เกิดจากการใช้จำนวนคำพูดที่เกิดขึ้นในการทำประเมินเพียงอย่างเดียว โดยมีค่าAUC เท่ากับ 0.720, 0.708, 0.758 จากตัวจำแนกประเภท XGBoost, Support Vector Machine, Random forest ตามลำดับ อีกทั้งผลลัพธ์จากงานวิจัยนี้ยังบ่งชี้ให้เห็นว่า การใช้ข้อมูลจากแบบประเมินความคล่องแคล่วทางภาษาของหน่วยเสียงฉบับภาษาไทย มีความเป็นไปได้ที่จะนำไปใช้เพื่อการแยะแยะภาวะรู้คิดบกพร่องเล็กน้อย เนื่องด้วยอรรถประโยชน์จากความเรียบง่าย และการใช้เวลาน้อยของตัวแบบประเมิน ทำให้ข้อมูลเสียงที่ได้จากแบบประเมินความคล่องทางภาษาเหมาะที่จะนำมาพัฒนาต่อจนสามารถเป็นอุปกรณ์ช่วยคัดกรองภาวะความรู้คิดบกพร่องเล็กน้อยได้ ซึ่งผู้วิจัยคิดว่าจะช่วยลดภาระงานให้กับบุคลากรณ์ทางการแพทย์ และยังเหมาะสมที่จะเป็นพัฒนาเป็นโปรแกรมประยุกต์ออนไลน์ สำหรับการประเมินความเสี่ยงภาวะความรู้คิดบกพร่องเล็กน้อยให้กับผู้ที่เข้าถึงโรงพยาบาลได้ยากอีกด้วย

5.2 ปัจจัยของคุณลักษณะแต่ละชนิดต่อภาวะความรู้คิดบกพร่องเล็กน้อย

ในส่วนนี้จะกล่าวถึงการอภิปรายความสัมพันธ์ระหว่างคุณลักษณะแต่ละชนิดกับภาวะความรู้คิดบกพร่องเล็กน้อย ได้แก่ คุณลักษณะที่ขึ้นอยู่กับความเงียบ คุณลักษณะที่ขึ้นอยู่กับความคล้าย การจัดกลุ่มคำ การเปลี่ยนกลุ่มคำ และจำนวนปีการศึกษา

5.2.1 คุณลักษณะที่ขึ้นอยู่กับความเงียบ

กลุ่มคุณลักษณะที่ขึ้นอยู่กับความเงียบมีประสิทธิภาพมากที่สุด เมื่อเทียบกับกลุ่มอื่นๆ โดยมีพิสัยการแจกแจงข้อมูลดังตารางที่ 7 จะสังเกตุเห็นได้ว่า เมื่อเปรียบเทียบค่าเฉลี่ยของคุณลักษณะเดี่ยวต่างๆพบว่า มีช่องว่างอยู่พอสมควรระหว่างผู้เข้ารับการประเมินที่มีภาวะความรู้คิดบกพร่องเล็กน้อย กับกลุ่มควบคุม ค่า AUC ของกลุ่มคุณลักษณะนี้เท่ากับ 0.654, 0.696, 0.659 จากตัวจำแนกประเภท XGBoost, Support Vector Machine, Random forest ตามลำดับ เมื่อพิจารณาคุณลักษณะแบบเดี่ยวแล้วพบว่า ค่าเฉลี่ยของช่วงเงียบระหว่างคำมีค่า SHAP สูงที่สุดในกลุ่ม จึงนำไปสู่ความเป็นไปได้ว่าค่าเฉลี่ยของช่วงเงียบระหว่างคำมีความเกี่ยวโยงกับทักษะในการทำแบบประเมิน 2 ประการ ได้แก่ 1.) การเข้าถึงคำพูดในความจำของผู้เข้ารับการประเมิน ซึ่งผู้เข้ารับการประเมินที่มีภาวะความรู้คิดบกพร่องเล็กน้อยมีแนวโน้มจะใช้เวลาคิดนานกว่ากลุ่มควบคุม 2) คุณลักษณะที่ขึ้นอยู่กับความเงียบมีความเกี่ยวพันกับจำนวนคำพูดในการทำแบบประเมิน คุณลักษณะที่ขึ้นอยู่กับความเงียบจึงมีความเหมาะสมที่จะนำไปใช้ร่วมกับเทคนิคการเรียนรู้ของเครื่อง

ตารางที่ 7 ตารางพิสัยการแจกแจงข้อมูลคุณลักษณะที่ขึ้นอยู่กับความเงียบ

	MCI (n=41)	HC (n=59)
Total Silence (Mean)	49.6 - 60.2 (56.5)	44.1 - 60.5 (55.0)
Q1Silence (Mean)	21.1 - 28.2 (25.5)	28.9 - 16.2 (23.9)
Q2 Silence (Mean)	26.2 - 35.8 (31.0)	27.1 - 36.3 (31.1)
Different Silence between Q1 and Q2 (Mean)	0.17 - 13.16 (5.6)	2.62 - 13.96 (7.3)
Average Silence Between Words (Mean)	2.82 - 14.94 (5.5)	1.76 - 20.17 (4.3)
Silence Before First Word (Mean)	0.10 - 11.64 (2.0)	0.11 - 8.8 (1.3)
Number of Silence Segment (Mean)	4 - 18 (11.7)	3 - 25 (15.3)
Number of Voiced Segment (Mean)	3 - 17 (10.7)	2 - 24 (14.3)
Total Voice (Mean)	2 - 16.1 (7.7)	1 - 23.1 (10.1)

5.2.2 คุณลักษณะที่ขึ้นอยู่กับความคล้าย

กลุ่มคุณลักษณะที่ขึ้นอยู่กับความคล้ายมีประสิทธิที่ค่อนข้างแย่เมื่อเทียบกับกลุ่มอื่นๆ แต่เมื่อพิจารณาคุณลักษณะแบบเดี่ยวภายในกลุ่มแล้ว กลับพบว่าค่าเฉลี่ยของความคล้ายเชิงความหมายมีประสิทธิภาพโดดเด่นที่สุดเมื่อเทียบกับกลุ่มตัวเอง ความคล้ายเชิงความหมายมีความสัมพันธ์กับทักษะในการพูดคำที่มีความหายหลากหลาย มีความเป็นไปได้ว่าผู้เข้ารับการประเมินที่มีภาวะความรู้คิดบกพร่องเล็กน้อยมีแนวโน้มที่จะพูดคำที่มีความหมายคล้ายกันมากกว่ากลุ่มควบคุม เมื่อพิจารณาพิสัย

การแจกแจงข้อมูลจากตารางที่ 7 พบว่า ช่องว่างของค่าเฉลี่ยความคล้ายทางความหมายมีค่ามากที่สุด ซึ่งอาจเป็นข้อมูลที่สนับสนุนพฤติกรรมข้างต้นได้

ตารางที่ 8 ตารางพิสัยการแจกแจงข้อมูลคุณลักษณะเชิงความหมาย

	MCI (n=41)	HC (n=59)
Average Orthographic Similarity (Mean)	0.178 - 0.390 (0.274)	0.191 - 0.500 (0.270)
Average Levenshtein Distance (Mean)	0.252 - 0.449 (0.359)	0.270 - 0.750 (0.363)
Average Semantic Similarity (Mean)	0.060 - 0.415 (0.186)	-0.006 - 0.397 (0.156)

5.2.3 การจัดกลุ่มคำและการเปลี่ยนกลุ่มคำ

การจัดกลุ่มคำและการเปลี่ยนกลุ่มคำมีประสิทธิภาพที่ค่อนข้างดี โดยมีค่า AUC เท่ากับ 0.626, 0.679, 0.633 จากตัวจำแนกประเภท XGBoost, Support Vector Machine, Random forest ตามลำดับ เมื่อแยกพิจารณาแบบเดี่ยวแล้วพบว่าการเปลี่ยนกลุ่มคำมีประสิทธภาพที่ดีกว่า เนื่องจากการเปลี่ยนกลุ่มคำอาศัยทักษะในการตัดสินใจที่อยางถี่ถ้วนที่จะพูดคำที่แตกต่างออกไป ผู้เข้ารับการประเมินที่มีภาวะความรู้คิดบกพร่องเล็กน้อยมีแนวโน้มที่จะได้แต้มน้อยกว่าในการเปลี่ยนกลุ่มคำเมื่อเทียบกับกลุ่มควบคุม

ตารางที่ 9 ตารางพิสัยความถี่ข้อมูลของการจัดกลุ่มคำ

	MCI (n=41)		HC (n=59)	
การจัดกลุ่มคำ	คะแนน 1	3	คะแนน 1	2
	คะแนน 2	16	คะแนน 2	19
	คะแนน 3	21	คะแนน 3	35
	คะแนน 4	1	คะแนน 4	3

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จากตารางที่ 9 แสดงพิสัยความถี่ข้อมูลของการจัดกลุ่มคำ ซึ่งบ่งชี้ว่าการจัดกลุ่มคำพูดของผู้เข้ารับการประเมินจะแบ่งได้ดังนี้

1) คำพูดแบ่งได้เป็น 3 กลุ่มมากที่สุด ทั้งผู้เข้ารับการประเมินที่มีภาวะความรู้คิดบกพร่องเล็กน้อย และกลุ่มควบคุม
2) คำพูดแบ่งได้เป็น 2 กลุ่มรองลงมา ทั้งผู้เข้ารับการประเมินที่มีภาวะความรู้คิดบกพร่องเล็กน้อย และกลุ่มควบคุม
3) มีความเป็นไปได้น้อยที่ผู้เข้ารับการประเมินจะพูดเพียง 2 หรือ 1 กลุ่มคำ

ตารางที่ 10 ตารางพิสัยการแจกแจงข้อมูลของการเปลี่ยนกลุ่มคำ

	MCI (n=41)	HC (n=59)
การเปลี่ยนกลุ่มคำ (ค่าเฉลี่ย)	1 – 13 (4.66)	1 – 13 (6.54)

จากตารางที่ 10 แสดงพิสัยการแจกแจงข้อมูลของการเปลี่ยนกลุ่มคำ ซึ่งบ่งชี้ว่าผู้เข้ารับการประเมินที่มีภาวะความรู้คิดบกพร่องเล็กน้อยมีค่าเฉลี่ยการเปลี่ยนกลุ่มคำที่ต่ำกว่ากลุ่มควบคุมอย่างชัดเจน

5.2.4 จำนวนปีการศึกษา

จำนวนปีการศึกษาเป็นคุณลักษณะที่มาจากข้อมูลพื้นฐานที่บันทึกจากการกรอกข้อมูลก่อนเริ่มการทำแบบประเมิน จำนวนปีการศึกษามีประสิทธิภาพที่สูงในการคัดแยกภาวะความรู้คิดบกพร่องเล็กน้อยด้วยตัวมันเอง โดยที่มีค่า AUC สูงถึง 0.725 ในทุกตัวจำแนกประเภท ซึ่งผลลัพธ์ที่ได้นี้สอดกับงานวิจัย [27] ที่บ่งชี้ว่า ผู้เข้ารับการประเมินที่มีจำนวนปีการศึกษาสูงมักจะมีความคุ้นเคยกับการทำแบบทดสอบ เป็นผลให้สามารถทำคะแนนรวมในแบบประเมินโมคาได้ดีตามไปด้วย ซึ่งในงานวิจัยนี้แบ่งกลุ่มตัวอย่างโดยใช้คะแนนรวมจากแบบระเมินโมคา จึงเป็นเหตุผลที่ทำให้ค่าความแม่นยำที่ได้สำหรับคุณลักษณะนี้สูงในผู้เข้ารับการประเมินที่มีจำนวนปีการศึกษาสูงตามไปด้วย สำหรับพิสัยการแจกแจงข้อมูลจะแสดงไว้ดังตารางที่ 11

ตารางที่ 11 ตารางพิสัยและการแจกแจงความถี่ของจำนวนปีการศึกษา

	MCI (n=41)	HC (n=59)
Education Range		
Min - Max (Mean)	6 - 18 (13.15)	6 - 21 (16.22)
Year of Education by frequency		
6 or lower	7	1
9 to 11 years	5	1
12 to 15 years	5	3
16 to 17 years	20	36
18 to 20 years	4	16
21 or more	0	2

CHULALONGKORN UNIVERSITY

5.3 ผลลัพธ์จากการคัดเลือกคุณลักษณะสำคัญ

การผสมผสานคุณลักณะกลุ่มต่างๆ จากการคัดเลือกคุณลักษณะด้วยค่าทางสถิติร่วมกับเทคนิคการเรียนรู้ของเครื่อง ให้ผลลัพธ์ที่ดีกว่าแบบจำลองพื้นฐานอันมาจากการใช้จำนวนคำพูดที่ถูกต้องเพียงอย่างเดียว โดยตารางสรุปคุณลักษณะสำคัญแสดงไว้ในตารางที่ 12 จากตารางสรุปนี้พบว่าคุณลักษณะสำคัญในตัวจำแนกแต่ละประเภท มีความทับซ้อนกันอยู่ 2 ชนิด คือ 1.) จำนวนปีการศึกษา 2.) ค่าเฉลี่ยความเงียบระหว่างคำ ซึ่งอาจสรุปได้ว่าคุณลักษณะทั้ง 2 ชนิดนี้อาจเป็นตัวบ่งชี้สำหรับการตรวจจับภาวะความรู้คิดบกพร่องเล็กน้อยได้ดี ทั้งนี้ผลลัพธ์จากการทดลองของเราช่วยยืนยันสมมุติฐานของงานวิจัยได้อย่างดีว่า การใช้คุณลักษณะแบบผสมผสานด้วยการคัดเลือกจากค่าทางสถิติ และการใช้ข้อมูลจากแบบประเมินความคล่องแคล่วทางภาษาแบบหน่วยเสียงฉบับภาษาไทยมีความเป็นได้ในการแยกยะภาวะความรู้คิดบกพร่องเล็กน้อยออกจากกลุ่มควบคุม

ตารางที่ 12 คุณลักษณะสำคัญที่มีผลต่อความแม่นยำในแต่ละตัวจำแนก

Classifier	Top 5 Feature Importance
XGBoost	■ ■ □ ■ ●
SVM	■ ■ ● ■ ■
RF	■ ■ ○ ● □

■ Year of Education
■ Average Silence Between Word
■ Number of Switches
■ Q1 Silence
□ Q2 Silence
□ Number of Silence Segment
○ Number of Voiced Segment
● Silence Before First Word
● Q1-Q2 Difference

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5.4 ข้อจำกัดของงานวิจัย

ในส่วนนี้จะกล่าวถึงข้อจำกัดในงานวิจัยซึ่งเป็นอุปสรรคสำหรับงานสิจัยฉบับนี้ไว้ ดังนี้

1. ในขณะที่ผู้วิจัยกำลังเขียนวิทยานิพนธ์นี้ โปรแกรมประยุกต์โมคายังคงเป็นเวอร์ชั่นต้นแบบทำให้ยังคงมีจุดอ่อนสำหรับการเก็บข้อมูลประเภทเสียง เป็นผลให้ข้อมูลประเภทเสียงที่ได้มานั้นมีเสียงรบกวนที่ค่อนข้างซับซ้อน และการนำเทคโลยีรู้จำเสียงพูด (Automatic Speech Recognition) มาประยุกต์ใช้ ให้ผลลัพธ์ที่ไม่สามารถยอมรับได้ในการแปลงเสียงคำพูดเป็นตัวอักษร ซึ่งด้วยเหตุผลนี้ทำให้ผู้วิจัยจะต้องใช้แฟ้มข้อมูลถอดความแทน (Transcribed File) สำหรับขั้นตอนการแปลงเสียงพูดเป็นตัวอักษร
2. ข้อมูลที่ใช้งานวิจัยยังขาดความหลากหลาย เนื่องจากจากเป็นเพียงข้อมูลที่มาจากแบบประเมินความคล่องแคล่วทางภาษาเพียงประเภทเดียว คือ การพูดคำที่ขึ้นต้นด้วยตัวอักษร

“ก” ให้ได้มากที่สุดเพียงตัวอักษรเดียว ซึ่งสำหรับฉบับภาษาไทยยังมีแบบประเมินความคล่องแคล่วทางภาษาอีกหลากหลายรูปแบบที่ยังรอการศึกษาค้นคว้าอยู่ เช่น ตัวอักษร “ค”, คำที่อยู่ในกลุ่มผลไม้ คำที่อยู่ในกลุ่มสัตว์ และการอ่านบทความสั้นๆ

5.5 ข้อเสนอแนะสำหรับการนำงานวิจัยไปศึกษาต่อ

เนื่องจากปัญหาข้อจำกัดที่กล่าวไปข้างต้น การประยุกต์โดยนำเทคโนโลยีการรู้จำเสียงพูดจึงเป็นก้าวสำคัญที่จะทำให้เทคนิคในงานวิจัยชิ้นนี้สัมฤทธิ์ผลในเชิงการใช้งานจริงแบบอัตโนมัติ อีกทั้งการเพิ่มจำนวนข้อมูลให้มีจำนวนตัวอย่างและความหลากหลายทางแบบทดสอบที่มากขึ้นเพื่อความแข็งแรงของแบบจำลองก็เป็นอีกสิ่งจำเป็นสำหรับการนำไปใช้งานจริง และการเพิ่มเทคนิคใหม่ๆ สำหรับขั้นตอนการสกัดคุณลักษณะในภาษาไทย ก็เป็นอีกความท้าทายหนึ่งในการเพิ่มความแม่นยำสำหรับงานวิจัยในอนาคต

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บรรณานุกรม

1. *Number of Population from Registration by Age Group Province and Region:* 2011 - 2020. [cited 2021 17 Aug]; Available from: http://statbbi.nso.go.th/staticreport/page/sector/th/01.aspx.
2. Nasreddine, Z.S., et al., *The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment.* Journal of the American Geriatrics Society, 2005. 53(4): p. 695-699.
3. Chi, Y.K., et al., *Development of a screening algorithm for Alzheimer's disease using categorical verbal fluency.* PloS one, 2014. 9(1): p. e84111.
4. König, A., et al., *Fully automatic speech-based analysis of the semantic verbal fluency task.* Dementia and geriatric cognitive disorders, 2018. 45(3-4): p. 198-209.
5. Lindsay, H., et al. *Automatic data-driven approaches for evaluating the phonemic verbal fluency task with healthy adults*. in *Proceedings of the 3rd International Conference on Natural Language and Speech Processing*. 2019.
6. Lindsay, H., et al. *What difference does it make? early dementia detection using the semantic and phonemic verbal fluency task*. in *LREC 2020 Language Resources and Evaluation Conference 11-16 May 2020*. 2020.
7. Kantithammakorn, P., *Using Automatic Speech Recognition to Assess Thai Speech Language Fluency in Montreal Cognitive Assessment (MoCA).* 2021.
8. Maust, D., et al., *Psychiatric rating scales.* Handbook of Clinical Neurology, 2012. 106: p. 227-237.
9. Julayanont, P. and Z.S. Nasreddine, *Montreal Cognitive Assessment (MoCA): concept and clinical review*, in *Cognitive screening instruments*. 2017, Springer. p. 139-195.
10. Baldo, J.V. and A.P. Shimamura, *Letter and category fluency in patients with frontal lobe lesions.* Neuropsychology, 1998. 12(2): p. 259.
11. Baldo, J.V., et al., *Verbal and design fluency in patients with frontal lobe lesions.* Journal of the international neuropsychological society, 2001. 7(5): p. 586-596.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

12. Troyer, A.K., et al., *Clustering and switching on verbal fluency: The effects of focal frontal-and temporal-lobe lesions.* Neuropsychologia, 1998. 36(6): p. 499-504.
13. Parks, R.W., et al., *Cerebral metabolic effects of a verbal fluency test: A PET scan study.* Journal of clinical and experimental neuropsychology, 1988. 10(5): p. 565-575.
14. Phelps, E.A., et al., *FMRI of the prefrontal cortex during overt verbal fluency.* Neuroreport, 1997. 8(2): p. 561-565.
15. Ostrosky-Solís, F., A. Ardila, and M. Rosselli, *NEUROPSI: A brief neuropsychological test battery in Spanish with norms by age and educational level.* Journal of the international Neuropsychological Society, 1999. 5(5): p. 413-433.
16. Ostrosky-Solis, F., et al., *Neuropsychological test performance in illiterate subjects.* Archives of Clinical Neuropsychology, 1998. 13(7): p. 645-660.
17. Rosselli, M., A. Ardila, and P. Rosas, *Neuropsychological assessment in illiterates: II. Language and praxic abilities.* Brain and Cognition, 1990. 12(2): p. 281-296.
18. Murphy, K.J., J.B. Rich, and A.K. Troyer, *Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia.* Journal of the International Neuropsychological Society, 2006. 12(4): p. 570-574.
19. Clark, D., et al., *Lexical factors and cerebral regions influencing verbal fluency performance in MCI.* Neuropsychologia, 2014. 54: p. 98-111.
20. Levenshtein, V.I. *Binary codes capable of correcting deletions, insertions, and reversals*. in *Soviet physics doklady*. 1966. Soviet Union.
21. Lindsay, H., et al. *Dissociating Semantic and Phonemic Search Strategies in the Phonemic Verbal Fluency Task in early Dementia*. in *Proceedings of the Seventh Workshop on Computational Linguistics and Clinical Psychology: Improving Access*. 2021.
22. Fontan, L., et al. *Using phonologically weighted Levenshtein distances for the prediction of microscopic intelligibility*. in *Annual conference Interspeech (INTERSPEECH* 2016). 2016.
23. Phatthiyaphaibun, W., et al., *PyThaiNLP: Thai Natural Language Processing in*

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Python, June 2016. URL http://doi. org/10.5281/zenodo. 3519354.

24. Weiner, J., C. Herff, and T. Schultz. *Speech-Based Detection of Alzheimer's Disease in Conversational German*. in *INTERSPEECH*. 2016.
25. Satt, A., et al. *Evaluation of speech-based protocol for detection of early-stage dementia*. in *Interspeech*. 2013.
26. Pedregosa, F., et al., *Scikit-learn: Machine learning in Python.* the Journal of machine Learning research, 2011. 12: p. 2825-2830.
27. Tingsabadh, M.K. and A.S. Abramson, *Thai.* Journal of the International Phonetic Association, 1993. 23(1): p. 24-28.
28. Hoffmann, I., et al., *Temporal parameters of spontaneous speech in Alzheimer's disease.* International journal of speech-language pathology, 2010. 12(1): p. 29-34.
29. Tóth, L., et al., *A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech.* Current Alzheimer Research, 2018. 15(2): p. 130-138.
30. Tóth, L., et al. *Automatic detection of mild cognitive impairment from spontaneous speech using ASR*. in *Sixteenth Annual Conference of the International Speech Communication Association*. 2015.
31. Campbell, E.L., et al., *Paralinguistic and linguistic fluency features for Alzheimer's disease detection.* Computer Speech and Language, 2021. 68.
32. Siew, C.S., *The orthographic similarity structure of English words: Insights from network science.* Applied Network Science, 2018. 3(1): p. 1-18.
33. Troyer, A.K., M. Moscovitch, and G. Winocur, *Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults.* neuropsychology, 1997. 11(1): p. 138.
34. Ryan, J.O., et al. *Computerized analysis of a verbal fluency test*. in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 2013.
35. Mueller, K.D., et al., *Verbal fluency and early memory decline: results from the Wisconsin registry for Alzheimer's prevention.* Archives of Clinical Neuropsychology, 2015. 30(5): p. 448-457.
36. Lundberg, S.M., et al., *From local explanations to global understanding with*

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

explainable AI for trees. Nature Machine Intelligence, 2020. 2(1): p. 56-67.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียน

ชื่อ-สกุล	นาย สุพพัต เมธารักษ์ชีพ
วัน เดือน ปี เกิด	19 มกราคม 2534
สถานที่เกิด	สมุทรสาคร
วุฒิการศึกษา	ปริญญาตรี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ที่อยู่ปัจจุบัน	199/6 ซอยจินดาถวิล ถนนสี่พระยา แขวงมหาพฤฒาราม อำเภอบางรัก กรุงเทพมหานคร 10500

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University