บทที่ 2
ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

บทนี้กล่าวถึงการทบทวนวรรณกรรมที่เกี่ยวข้องกับงานวิจัยนี้ โดยมีรายละเอียดเกี่ยวกับการเตรียมข้อมูลในรูปแบบเชิงตัวเลขเพื่อนำไปใช้สำหรับการวิเคราะห์และประมวลผล รวมถึงการใช้เทคนิคการสุ่มตัวอย่างรูปแบบผสมผสาน ซึ่งประกอบไปด้วยการสุ่มลดจำนวนตัวอย่างในกลุ่มข้อมูลส่วนใหญ่และการสุ่มเพิ่มจำนวนกลุ่มตัวอย่างในกลุ่มข้อมูลส่วนน้อย เพื่อปรับให้กลุ่มข้อมูลมีความสมดุลกัน นอกจากนี้ยังมีการกล่าวถึงกระบวนการที่ใช้ในการจัดการข้อมูล อัลกอริทึมที่ใช้ในการจำแนกประเภท ตลอดจนวิธีการวัดผลและประเมินประสิทธิภาพของแบบจำลอง โดยงานวิจัยนี้มีการอ้างอิงถึงงานวิจัยที่เกี่ยวข้องดังต่อไปนี้

2.1 การจัดการเตรียมข้อมูล

ในปัจจุบันความแม่นยำและความถูกต้องในการจำแนกกลุ่มข้อมูลชนกลุ่มน้อยในงานด้านการเรียนรู้ของเครื่อง เป็นประเด็นที่ได้รับความสนใจ เนื่องจากการกระจายตัวของข้อมูลที่ไม่สมดุล ทำให้เกิดความท้าทายในการแก้ปัญหานี้ (Fotouhi et al., 2019) การกระจายตัวที่ไม่สมดุลนี้เป็นอุปสรรคสำคัญในงานด้านการเรียนรู้ของเครื่องอย่างมาก เนื่องจากแบบจำลองมักเรียนรู้จากข้อมูลส่วนใหญ่ที่มีจำนวนมากกว่า ส่งผลให้ขาดความสามารถเพียงพอในการจำแนกข้อมูลกลุ่มน้อย ทั้งนี้ โมเดลจึงมักทำนายผลลัพธ์ไปในทิศทางเดียวกับข้อมูลกลุ่มใหญ่ (Li et al., 2014)

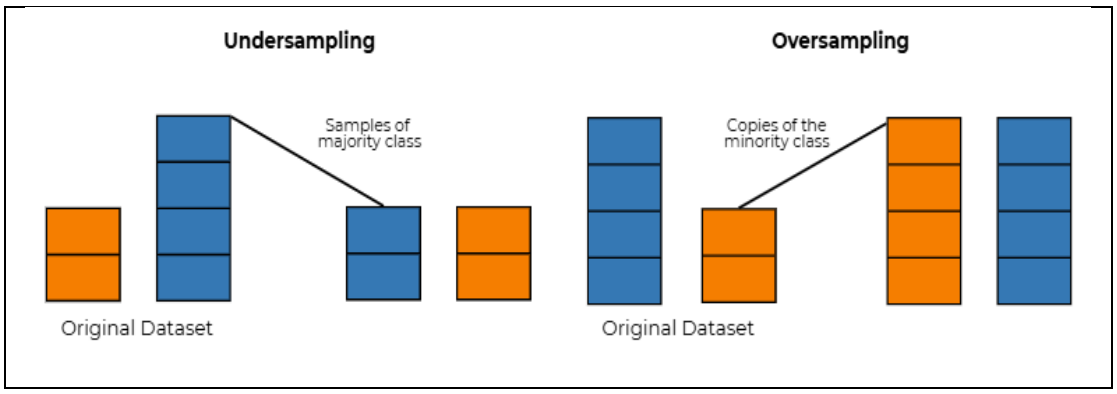
จากการศึกษาพบว่าเพื่อให้สามารถใช้วิธีการสุ่มตัวอย่างได้อย่างมีประสิทธิภาพ จำเป็นต้องแปลงข้อมูลให้เป็นข้อมูลเชิงตัวเลข (Khushi et al., 2021) ก่อนที่จะดำเนินการแก้ปัญหาข้อมูลที่ไม่สมดุล ในงานวิจัยนี้จะใช้วิธีการจัดการข้อมูลในรูปแบบผสมผสาน รูปแบบนี้คือการนำวิธีการจัดการข้อมูลที่ไม่สมดุลมาประยุกต์ใช้ โดยเลือกใช้วิธีการสุ่มลดตัวอย่างกลุ่มข้อมูลร่วมกับการสุ่มเพิ่มตัวอย่างข้อมูล ปกติแล้ววิธีการสุ่มตัวอย่างจะใช้กับชุดข้อมูลที่มีเพียงสองกลุ่ม แต่ชุดข้อมูลในงานวิจัยนี้ประกอบด้วยสามกลุ่ม คือ กลุ่มผู้ป่วยปกติ จำนวน 9,341 ตัวอย่าง จากข้อมูลทั้งหมด กลุ่มผู้ป่วยมีความเสี่ยง จำนวน 123 ตัวอย่าง จากข้อมูลทั้งหมด และกลุ่มผู้ป่วยที่มีภาวะครรภ์เป็นพิษ จำนวน 9 ตัวอย่าง จากข้อมูลทั้งหมด

ดังนั้นจึงได้นำวิธีการสุ่มเพื่อลดจำนวนกลุ่มตัวอย่างมาใช้กับกลุ่มผู้ป่วยปกติ เพื่อให้จำนวนข้อมูลสมดุลกับกลุ่มผู้ป่วยที่มีความเสี่ยง ต่อมาจึงนำวิธีการสุ่มเพื่อเพิ่มจำนวนตัวอย่างกับกลุ่มข้อมูลผู้ป่วยที่มีภาวะครรภ์เป็นพิษ เพื่อให้จำนวนข้อมูลสมดุลกับกลุ่มผู้ป่วยที่มีความเสี่ยง วิธีการนี้ส่งผลให้ทั้งสามกลุ่มข้อมูลมีปริมาณข้อมูลที่สมดุลกันมากขึ้น

ก่อนที่จะทำการสุ่มตัวอย่างข้อมูล จะมีการประเมินความไม่สมดุลของข้อมูลผ่านการคำนวณอัตราส่วนระหว่างกลุ่มข้อมูลที่มีส่วนมากและกลุ่มที่มีส่วนน้อย ซึ่งชุดข้อมูลนี้มีอัตราส่วนความไม่สมดุลอย่างชัดเจน ดังนั้นจึงจำเป็นต้องใช้วิธีการสุ่มตัวอย่างแบบผสมผสานเพื่อลดจำนวนข้อมูลในกลุ่มปกติและเพิ่มข้อมูลในกลุ่มครรภ์เป็นพิษ เพื่อให้ค่าสัดส่วนของชุดข้อมูลเข้าใกล้ 1 ซึ่งแสดงถึงความสมดุลที่ดีขึ้น (Noorhalim et al., 2019) การเลือกใช้วิธีการเหล่านี้ยังอ้างอิงจากงานวิจัยที่ได้นำเทคนิคต่าง ๆ มาปรับใช้ในการจำแนกประเภทข้อมูล (Khushi et al., 2021)

ในด้านวิทยาศาสตร์การแพทย์ เครื่องมือที่ใช้ในการจัดการกับความไม่สมดุลของกลุ่มตัวอย่างข้อมูลได้รับการนำมาใช้ในงานวิจัยหลายชิ้น และมีการจัดการข้อมูลที่ไม่สมดุลในหลากหลายบริบทเพื่อการวินิจฉัย เช่น การวินิจฉัยโรคมะเร็งปอด (Khushi et al., 2021), ผลข้างเคียงทางการแพทย์ (Santiso et al., 2019), และการระบุสารพันธุกรรม (Herndon et al., 2016)

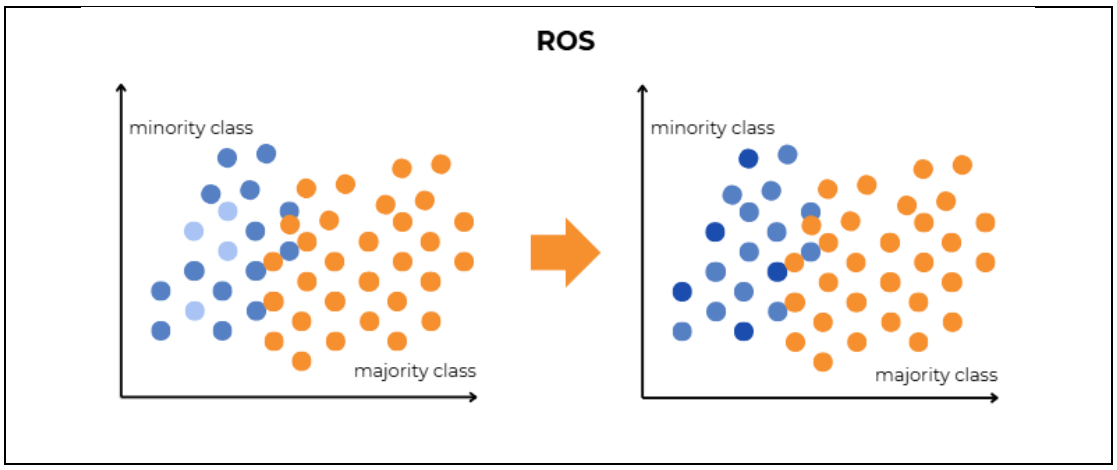
วิธีการระดับข้อมูลในปัจจุบันจะถูกใช้ในขั้นตอนการเตรียมข้อมูล โดยการสุ่มตัวอย่างข้อมูลและจัดการการกระจายตัวของกลุ่มประชากรในข้อมูลใหม่ เพื่อแก้ไขปัญหาความไม่สมดุลของข้อมูล (Longadge and Dongre, 2013) โดยการเพิ่มจำนวนข้อมูลในกลุ่มข้อมูลชนกลุ่มน้อย และลดจำนวนข้อมูลในกลุ่มข้อมูลชนกลุ่มมาก ในวิธีการระดับข้อมูลส่วนถัดไปจะอธิบายวิธีการต่างๆ แบบกระชับสั้น ๆ เกี่ยวกับเทคนิคในระดับข้อมูล



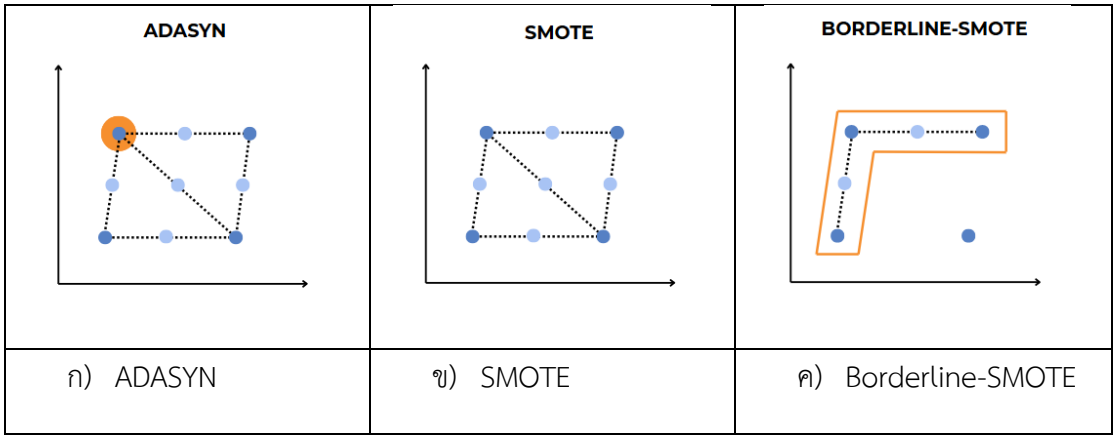
รูปที่ 2.1 หลักการทำงานการสุ่มเพิ่มและลดตัวอย่างข้อมูล

โดยจากการศึกษาการสุ่มเพิ่มตัวอย่างของข้อมูลนั้นจะเป็นการเพิ่มกลุ่มข้อมูลกลุ่มน้อยขึ้นมาและการสุ่มตัวอย่างลดข้อมูลนั้นจะเป็นการลดกลุ่มข้อมูลส่วนใหญ่ลงมานั่นเอง เป็นไปดังรูปที่ 2.1

วิธีการสุ่มเพิ่มตัวอย่างมีดังนี้: Random Oversampling (ROS), Adaptive Synthetic Sampling (ADASYN), Synthetic Minority Over-sampling Technique (SMOTE), Borderline-SMOTE



รูปที่ 2.2 หลักการทำงานการสุ่มเพิ่มข้อมูลด้วยวิธี ROS



ก) ADASYN	ข) SMOTE	ค) Borderline-SMOTE

รูปที่ 2.3 หลักการทำงานการสุ่มเพิ่มข้อมูลด้วยวิธี ADASYN (ก) SMOTE (ข) และ Borderline-SMOTE (ค)

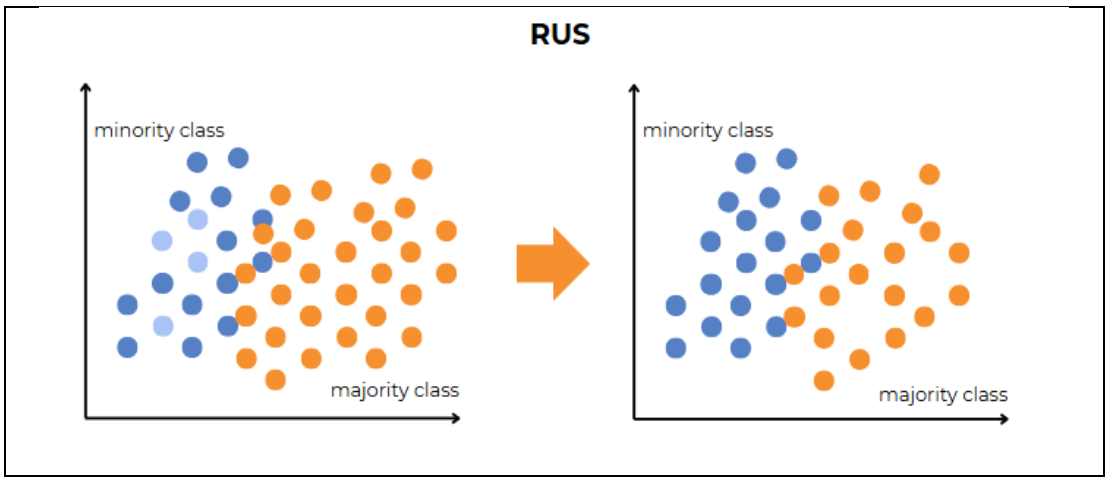
1) ROS เป็นเทคนิคการวิเคราะห์ข้อมูลในขั้นต้นที่ถูกพัฒนาขึ้นเพื่อใช้ในการสุ่มเพิ่มข้อมูลตัวอย่างในคลาสชนกลุ่มน้อย (Batista et al., 2004) ดังรูปที่ 2.2

2) ADASYN ใช้การกระจายตัวถ่วงน้ำหนักในกลุ่มตัวอย่างของข้อมูลคลาสชนกลุ่มน้อยตามระดับความยากในการเรียนรู้ โดยจะสร้างกลุ่มตัวอย่างมากขึ้นในส่วนที่มีความยากในการเรียนรู้สูงกว่า (He et al., 2008) ดังรูปที่ 2.3

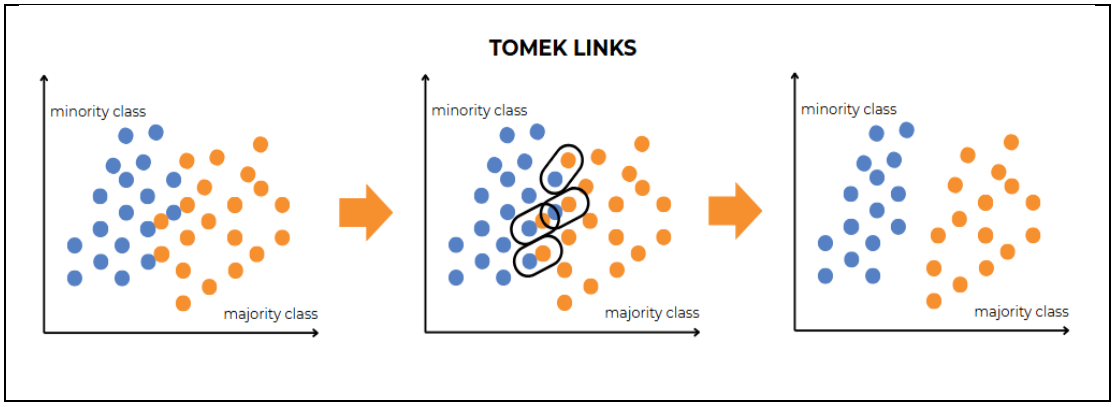
3) SMOTE ใช้ k-Nearest Neighbors (k-NN) เพื่อประมาณค่า ก่อนที่จะทำการสังเคราะห์ข้อมูลใหม่ในคลาสชนกลุ่มน้อยระหว่างข้อมูลเดิม (Chawla et al., 2002) ดังรูปที่ 2.3

4) Borderline-SMOTE เป็นการใช้ SMOTE กับข้อมูลที่อยู่ในเขตชายขอบ ซึ่งมักจะถูกจำแนกผิดประเภท และทำการสังเคราะห์ข้อมูลบริเวณนั้นเพิ่ม (Han et al., 2005) ดังรูปที่ 2.3

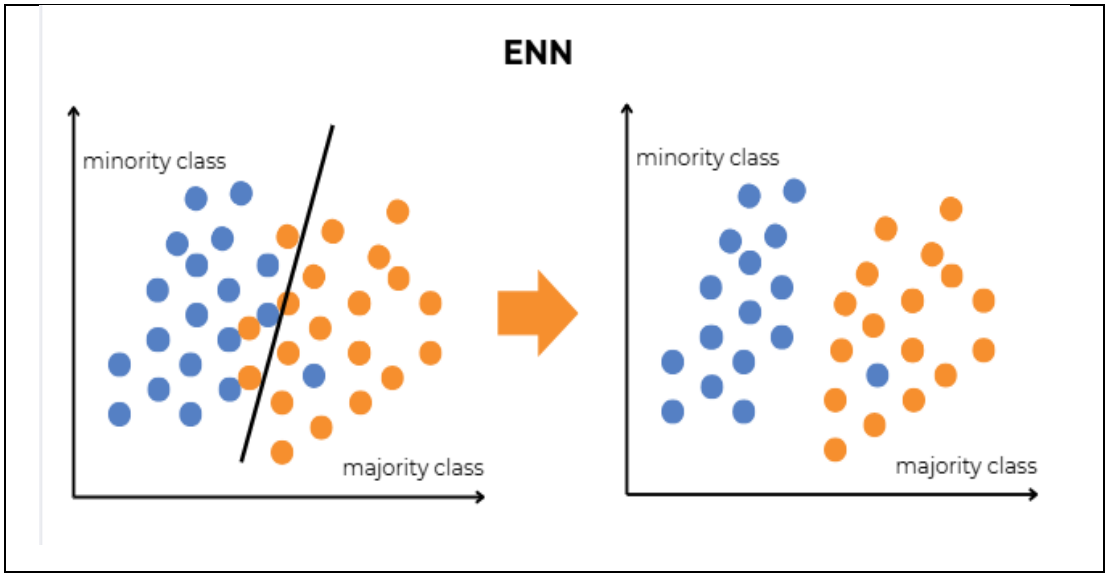
วิธีการสุ่มลดตัวอย่างมีดังนี้: Random Undersampling (RUS), Tomek Links, Edited Nearest Neighbors (ENN), Repeated Edited Nearest Neighbors (RENN)



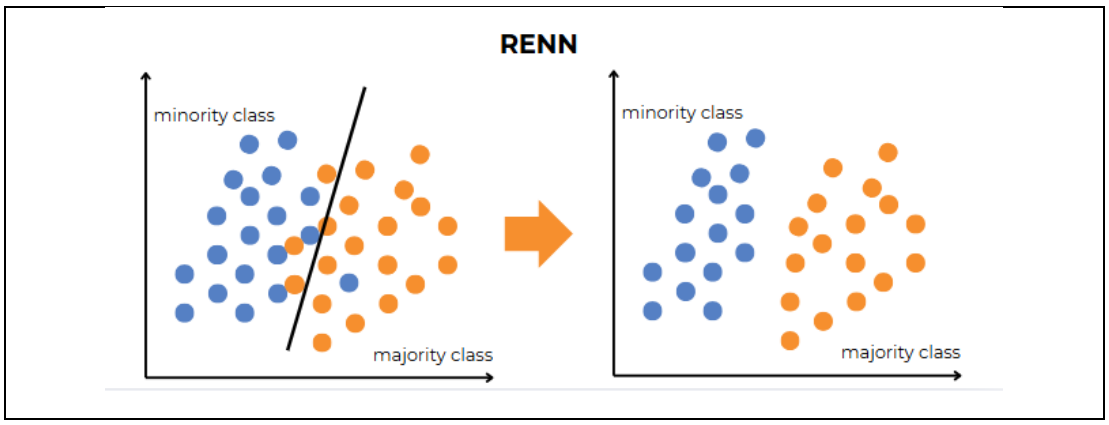
รูปที่ 2.4 หลักการทำงานการสุ่มลดข้อมูลด้วยวิธี RUS



รูปที่ 2.5 หลักการทำงานการสุ่มลดข้อมูลด้วยวิธี Tomek Links



รูปที่ 2.6 หลักการทำงานการสุ่มลดข้อมูลด้วยวิธี ENN



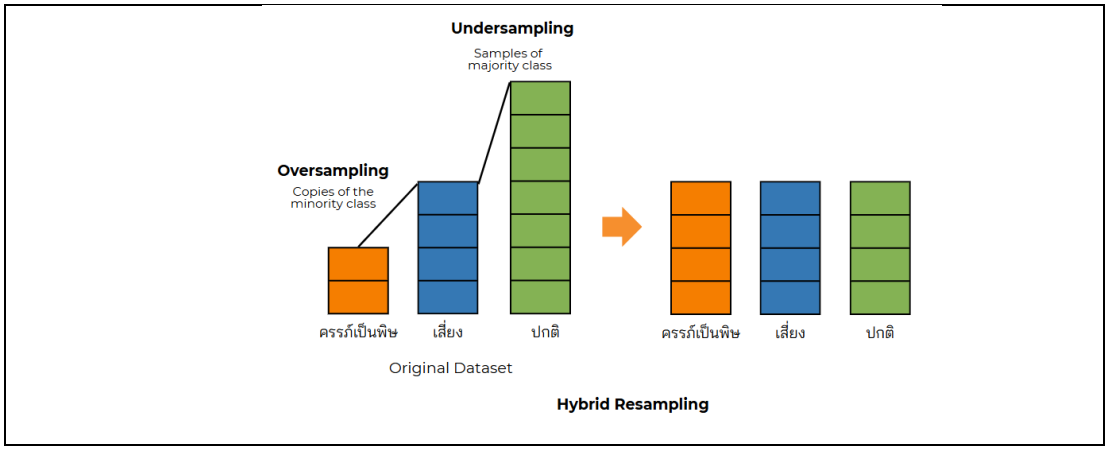
รูปที่ 2.7 หลักการทำงานการสุ่มลดข้อมูลด้วยวิธี RENN

1) RUS เป็นเทคนิคการสุ่มตัวอย่างที่เก่าแก่และได้รับการพัฒนามาตั้งแต่แรก โดยจะลดจำนวนตัวอย่างจากกลุ่มข้อมูลส่วนมาก (Prusa et al., 2015) ดังรูปที่ 2.4

2) Tomek Links สองตัวอย่างข้อมูลจะถูกพิจารณาว่าเป็น Tomek Links หากทั้งสองตัวอย่างอยู่ในกลุ่มข้อมูลที่แตกต่างกัน และเป็นเพื่อนบ้านที่ใกล้ที่สุดในกลุ่มข้อมูล ทั้งนี้สามารถหมายถึงตัวอย่างที่มีข้อผิดพลาดที่อาจจะรบกวน และเป็นตัวอย่างจากกลุ่มข้อมูลส่วนมากจะถูกลบออก (Tomek, 1976) ดังรูปที่ 2.5

3) ENN ในวิธีนี้ กลุ่มตัวอย่างแต่ละกลุ่มจะได้รับการทดสอบโดยใช้ k Nearest Neighbors (k-NN) กับกลุ่มข้อมูลที่เหลือ ตัวอย่างที่ถูกจำแนกไม่ถูกต้องจะถูกตัดทิ้ง และตัวอย่างที่เหลือจะกลายเป็นชุดข้อมูลที่ได้รับการแก้ไข (Broder et al., 1985) ดังรูปที่ 2.6

4) RENN เป็นการทำ ENN ซ้ำ จนกว่าจะสามารถแก้ไขชุดข้อมูลที่ไม่สามารถใช้ในการฝึกสอนได้ โดยจะไม่ถูกกระทบจากการกำจัดตัวอย่างอีกต่อไป (Wilson, 1972) ดังรูปที่ 2.7



รูปที่ 2.8 หลักการทำงานการสุ่มเพิ่มลดตัวอย่างข้อมูลแบบผสมผสาน

วิธีการสุ่มตัวอย่างแบบผสมผสานในงานวิจัยนี้ได้มีการเพิ่มจำนวนประชากรข้อมูลกลุ่มส่วนน้อย และลดจำนวนประชากรของข้อมูลส่วนมาก ที่ผสมผสานจับคู่ระหว่าง การสุ่มลดตัวอย่างกลุ่มข้อมูลชนกลุ่มมาก และ การสุ่มเพิ่มตัวอย่างข้อมูลชนกลุ่มน้อย ดังรูปที่ 2.8

2.2 การออกแบบแบบจำลอง

ในงานวิจัยนี้มีการใช้จะใช้ 6 อัลกอริทึมการจำแนกประเภทด้านการเรียนรู้ของเครื่อง คือ Decision Tree, Extreme Gradient Boosting (XGBoost), Random Forest, Naive Bayes, Neural Network, และ K-Nearest Neighbors (Kaur et al., 2019; Papacharalampous et al., 2021) ซึ่งเป็นตัวที่ใช้ในงานทางการสร้างแบบจำลองเพื่อทำนายข้อมูลที่ไม่สมดุล ปัญหาด้านข้อมูลทางการแพทย์

ตารางที่ 2.1 รายการสรุปการออกแบบแบบจำลอง

อัลกอริทึม	การใช้งาน	การอ้างอิง
Decision Tree	ใช้จำแนกประเภทข้อมูลที่ไม่สมดุล โดยปรับปรุงความแม่นยำผ่านการเลือกอัลกอริทึมที่เหมาะสมและการทดลองในชุดข้อมูลหลากหลาย	(Krawczyk et al., 2014), (Sanz et al., 2015), (Park and Ghosh, 2014)
XGBoost	เปรียบเทียบประสิทธิภาพกับอัลกอริทึมอื่นในงานวิเคราะห์ข้อมูล เช่น ปริมาณน้ำฝนและการวินิจฉัยโรคไตเรื้อรัง พบว่าให้ผลลัพธ์ที่ดีที่สุด	(Papacharalampous et al., 2021), (Ogunleye et al., 2020)
Random Forest	ใช้ในการจำแนกข้อมูลไม่สมดุล เช่น ข้อมูลข้อความและโรคมะเร็งปอด โดยให้ผลลัพธ์ดีที่สุดเมื่อเทียบกับอัลกอริทึมอื่น	(Wu et al., 2014), (Khushi et al., 2021), (Khabsa et al., 2016)
Naive Bayes	ใช้จัดการข้อมูลไม่สมดุล เช่น การจัดกลุ่มนักเรียนและในด้านวิทยาศาสตร์การแพทย์ โดยเน้นการจัดการข้อมูลที่มีมิติสูง	(Yap et al., 2014), (Bhandari and Patel, 2015), (Kaur et al., 2019)
Neural Network	ใช้ในหลากหลายงาน เช่น การประมวลผลข้อมูลลิ้นอิเล็กทรอนิกส์ การตรวจจับการทุจริตบัตรเครดิต และการปรับปรุงการกระจายข้อมูลในชุดข้อมูลที่ไม่สมดุล	(Liu et al., 2013), (Fu et al., 2016), (Jeatrakul et al., 2010)
K-Nearest Neighbors (KNN)	ใช้จำแนกประเภทข้อมูลไม่สมดุล เช่น การพยากรณ์โรคมะเร็ง โดยมีการเพิ่มตัวอย่างในคลาสที่มีจำนวนน้อย และเสนอ Hybrid Weighted Nearest Neighbors เพื่อเพิ่มความแม่นยำในข้อมูลไม่สมดุล	(Patel and Thakur, 2016), (Majid et al., 2014)

2.2.1 Decision Tree เป็นเครื่องมือทางอัลกอริทึมที่สามารถใช้ในการจำแนกประเภทได้ (Krawczyk et al., 2014) โดยการใช้ Decision Tree ในการจำแนกประเภทข้อมูลที่มีลักษณะไม่สมดุล ซึ่งเป็นลักษณะเด่นของชุดข้อมูลจริง ที่มักพบว่ามีกลุ่มข้อมูลที่ไม่สมดุล การเลือกใช้อัลกอริทึมในการจำแนกประเภทจึงมีความสำคัญ นอกจากนี้ยังได้มีการประยุกต์ใช้ในการคำนวณที่มีความแม่นยำดี โดยการนำไปใช้กับชุดข้อมูล 11 ชุด (Sanz et al., 2015) และงานวิจัยนี้ยังได้ใช้ในการแก้ปัญหาข้อมูลที่ไม่สมดุล ซึ่งมีประเภทของกลุ่มข้อมูลหลากหลาย และให้ผลลัพธ์ที่มีประสิทธิภาพในผลการทดลอง (Park and Ghosh, 2014)

2.2.2 XGBoost XGBoost นำไปใช้ในการเปรียบเทียบประสิทธิภาพกับอัลกอริทึมต่าง ๆ ในชุดข้อมูลปริมาณน้ำฝน ซึ่งผลลัพธ์แสดงให้เห็นว่า XGBoost มีประสิทธิภาพดีที่สุดในงานนี้ (Papacharalampous et al., 2021) นอกจากนี้ในงานวิจัยของ (Ogunleye et al., 2020) ได้เสนอให้ใช้ XGBoost ในการวินิจฉัยโรคไตเรื้อรัง

2.2.3 Random ถูกนำเสนอโดย (Wu et al., 2014) เป็นเครื่องมืออัลกอริทึมการเรียนรู้ของเครื่อง ซึ่งนำไปใช้ในการจำแนกประเภทข้อความที่มีข้อมูลไม่สมดุล ในงานการวินิจฉัยโรคมะเร็งปอดวิธีการนี้ได้รับการพิสูจน์ว่าให้ผลลัพธ์ดีที่สุดเมื่อเปรียบเทียบกับวิธีการอื่นๆ (Khushi et al., 2021) และในงานของ (Khabsa et al., 2016) ได้มีการนำเสนอการใช้ Random Forest เป็นกลยุทธ์ที่มีความแม่นยำในการศึกษาที่เกี่ยวข้อง โดยใช้ในการทบทวนระบบที่จำลองมาจากปัญหาการจำแนกประเภทที่ใช้กับข้อมูลไม่สมดุล

2.2.4 Naive Bayes ในงานที่มีการเปรียบเทียบวิธีการที่เหมาะสมในการจัดการกับข้อมูลที่ไม่สมดุล มีการนำเสนอ Naive Bayes เพื่อใช้ในการจำแนกข้อมูลที่มีความไม่สมดุล (Yap et al., 2014) ในงานของ (Bhandari and Patel, 2015) ได้นำเสนอการใช้ Naive Bayes กับชุดข้อมูลที่มีมิติสูงและข้อมูลที่ไม่สมดุลในการจัดกลุ่มนักเรียน และในการทบทวนระบบเกี่ยวกับความท้าทายในการจัดการข้อมูลที่ไม่สมดุลในการเรียนรู้ของเครื่องมือ ยังมีการเสนอวิธีนี้ในงานวิจัยด้านวิทยาศาสตร์การแพทย์ (Kaur et al., 2019)

2.2.5 Neural Network โครงข่ายประสาทเทียมมีการใช้งานที่หลากหลายในอัลกอริทึมการเรียนรู้ของเครื่อง (machine learning) ซึ่ง (Liu et al., 2013) ใช้โครงข่ายประสาทเทียมในการประมวลผลข้อมูลลิ้นอิเล็กทรอนิกส์ เพื่อตรวจจับตัวอย่างเครื่องดื่มส้มและน้ำส้มสายชูจีน และเปรียบเทียบกับอัลกอริทึมอื่นๆ โดยพบว่าโครงข่ายประสาทเทียมให้ผลลัพธ์ที่แม่นยำกว่า (Fu et al., 2016) ใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) ในการ

ตรวจจับการทุจริตบัตรเครดิตจากธุรกรรมของธนาคารพาณิชย์ โดยผลลัพธ์แสดงว่า ใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) มีประสิทธิภาพดีกว่าวิธีดั้งเดิม ในการตรวจจับการทุจริต (Jeatrakul et al., 2010) เสนอการใช้เทคนิคการสุ่มตัวอย่างแบบลดตัวอย่าง และเพิ่มตัวอย่าง โดยใช้โครงข่ายประสาทเทียมที่เสริมสมรรถนะและเทคนิค SMOTE เพื่อแก้ปัญหาความไม่สมดุลของข้อมูล ซึ่งช่วยปรับปรุงการกระจายข้อมูลในชุดข้อมูลและเพิ่มประสิทธิภาพในการฝึกโมเดล

2.2.6 K-Nearest Neighbors เป็นอัลกอริทึมการจำแนกประเภทที่ได้รับความนิยมในด้านการเรียนรู้ของเครื่อง เนื่องจากใช้งานง่ายและไม่ซับซ้อน ในการจัดการข้อมูลที่ไม่สมดุล (Patel and Thakur, 2016) ได้เสนออัลกอริทึมแบบไฮบริดที่เรียกว่า Hybrid Weighted Nearest Neighbors ซึ่งกำหนดน้ำหนักและจำนวน k ตามขนาดของแต่ละคลาส เพื่อเพิ่มความแม่นยำในการจำแนกประเภท (Majid et al., 2014) ศึกษาการพยากรณ์โรคมะเร็งเต้านมและมะเร็งลำไส้ใหญ่ โดยปรับสมดุลข้อมูลผ่านการเพิ่มตัวอย่างในคลาสที่มีจำนวนน้อย จากนั้นเปรียบเทียบผลระหว่าง SVM และ KNN ซึ่งพบว่า SVM ให้ผลลัพธ์ที่ดีกว่า KNN

2.3 งานวิจัยที่เกี่ยวข้องกับภาวะครรภ์เป็นพิษ

ภาวะครรภ์เป็นพิษเป็นภาวะสุขภาพที่ส่งผลต่อหญิงตั้งครรภ์หลังอายุครรภ์ 20 สัปดาห์ โดยมีลักษณะสำคัญคือความดันโลหิตสูงร่วมกับการพบโปรตีนในปัสสาวะ ภาวะนี้สามารถกระทบต่อการทำงานของหลายระบบในร่างกาย และเพิ่มความเสี่ยงต่อภาวะแทรกซ้อนที่รุนแรงทั้งต่อมารดาและทารก เช่น การคลอดก่อนกำหนด หรือความล้มเหลวของไตและตับ การรักษาที่มีประสิทธิภาพที่สุดในปัจจุบันคือการให้คลอดเมื่อถึงระยะเวลาที่เหมาะสม (Vata et al., 2015; Briceño-Pérez et al., 2009)

ตลอดช่วงหลายปีที่ผ่านมา ได้มีความพยายามจากหลากหลายงานวิจัยในการพัฒนาวิธีการคัดกรองหรือทำนายความเสี่ยงของภาวะครรภ์เป็นพิษให้ได้ตั้งแต่ระยะเริ่มต้น โดยใช้ข้อมูลจากประวัติสุขภาพของหญิงตั้งครรภ์ เช่น ดัชนีมวลกาย ระดับความดันโลหิต สารชีวภาพในเลือด หรือการตรวจด้วยอัลตราซาวด์ อย่างไรก็ตาม แนวทางและตัวแปรที่ใช้ในแต่ละงานวิจัยยังแตกต่างกัน ทำให้ยังไม่สามารถพัฒนาแบบจำลองที่ใช้ได้อย่างแพร่หลายในสถานพยาบาลทั่วไปได้ (De Kat et al., 2019)

นอกจากนี้ อัตราการเกิดภาวะครรภ์เป็นพิษยังแตกต่างกันในแต่ละภูมิภาค โดยเฉพาะในประเทศที่มีทรัพยากรจำกัด เช่น ประเทศในแถบแอฟริกา ซึ่งพบอัตราการเกิดสูงกว่าค่าเฉลี่ยโลกอย่างชัดเจน ปัจจัยเสี่ยงที่เกี่ยวข้องอาจรวมถึงภาวะโภชนาการ ความรู้เกี่ยวกับโรค และการเข้าถึงบริการฝากครรภ์ที่เหมาะสม ในบางพื้นที่พบว่าหญิงตั้งครรภ์กว่า 70% ไม่เคยได้รับข้อมูลพื้นฐานเกี่ยวกับภาวะนี้เลย (Vata et al., 2015)

แม้จะมีความพยายามในการพัฒนาแนวทางการป้องกันและการทำนายความเสี่ยงของภาวะครรภ์เป็นพิษ แต่ปัจจุบันองค์ความรู้เกี่ยวกับสาเหตุที่แท้จริงของโรคยังมีจำกัด อีกทั้งยังต้องการงานวิจัยเพิ่มเติม โดยเฉพาะในบริบทของประเทศที่มีข้อจำกัดด้านระบบสาธารณสุข เช่น เอธิโอเปีย ซึ่งจำเป็นต้องมีแนวทางที่เหมาะสมกับบริบทจริง เพื่อช่วยลดความเสี่ยงและผลกระทบที่อาจเกิดขึ้น