

การออกแบบระบบตรวจให้คะแนนแบบสอบอัตนัย โดยการให้ข้อมูลป้อนกลับอัตโนมัติ เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของเครื่อง Designing Automated Scoring System of Open-Ended Test by Providing Automatic Feedback to Diagnose Mathematical Proficiency Levels through Machine Learning

ศราวรรณ สุขวิชัย^{1*} พัทธี จันทร์เพ็ง² ชัยวัฒน์ ทะวะรุ่งเรือง³ และ ธนพงศ์ อินทะ⁴
Sarawan Sukwichai^{1*}, Putcharee Junpeng², Chaiwat Tawarungruang³ and Thanapong Intharah⁴

(Received: March 31, 2022; Revised: May 23, 2022; Accepted: May 27, 2022)

บทคัดย่อ

การวิจัยครั้งนี้มีความมุ่งหมาย (1) เพื่อวิเคราะห์ผลการตอบของผู้เรียนสู่การออกแบบนวัตกรรมต้นแบบ ประกอบด้วย มโนทัศน์ที่คลาดเคลื่อนของผู้เรียน การกำหนดจุดตัดระดับความสามารถทางคณิตศาสตร์ การให้ข้อมูลป้อนกลับตามมโนทัศน์ที่คลาดเคลื่อนและตัวแปรที่ใช้ในการทำนายผ่านการเรียนรู้ของเครื่อง (machine learning) และ (2) เพื่อออกแบบระบบตรวจให้คะแนนแบบสอบอัตนัย ของผู้เรียนผ่านการเรียนรู้ของเครื่อง (machine learning) โดยใช้ระเบียบวิธีวิจัยการออกแบบ กลุ่มผู้สอบ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 จำนวนนักเรียนทั้งหมด 495 คน เครื่องมือที่ใช้ คือ แบบทดสอบอัตนัย สารสถิติและความน่าจะเป็น ผ่านชุดเครื่องมือวินิจฉัยในระบบการทดสอบออนไลน์ “eMAT-Testing” วิเคราะห์ข้อมูลโดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ MRCML ผลการวิจัยปรากฏดังนี้

¹ นิสิตระดับปริญญาโท หลักสูตรวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

² รองศาสตราจารย์ หลักสูตรการวัดและประเมินผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

³ นักวิจัย ศูนย์จัดการข้อมูลและวิเคราะห์ทางสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

⁴ อาจารย์ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

งานวิจัยนี้ได้รับทุนอำนวยการวิทยานิพนธ์ทุนกลาง ประจำปี 2565 จากสำนักงานการวิจัยแห่งชาติ (วช.) รหัสโครงการวิจัย N42A650236

¹ Master's degree student, Educational Measurement and Evaluation Program, Faculty of Education, Khon Kaen University

² Associate Professor Department of Educational Measurement and Evaluation Faculty of Education Khon Kaen University

³ Researcher, Data Management and Statistical Analysis Center, Faculty of Public Health, Khon Kaen University

⁴ Lecturer, Department of Statistics, Faculty of Science, Khon Kaen University

The research received a research grant for Talented Mid-Career Researchers 2022 from the National Research Council of Thailand (NRCT) Project No. N42A650236

* Corresponding Author E-mail: sarawansukwichai@kkumail.com

1. ผลการวิเคราะห์ผลการตอบของผู้เรียนจากข้อมูลหัตถ์นิยม พบว่า (1) ผู้เรียนมีมีโนทัศน์ทั้งหมด 4 รูปแบบ คือ 1) การใช้ข้อมูล 2) การบิดเบือนทฤษฎีบท กฎ สูตร บทนิยาม และสมบัติ 3) การคำนวณและขาดการตรวจสอบในระหว่างการแก้ปัญหา และ 4) ความผิดพลาดในการตีความด้านภาษาและสัญลักษณ์ทางคณิตศาสตร์ โดยทั้งในมิติกระบวนการทางคณิตศาสตร์ (MAP) และ มิติโครงสร้างความคิดรวบยอด (SLO) นักเรียนมีมีโนทัศน์ที่คลาดเคลื่อนมากที่สุดคือ ด้านการใช้ข้อมูล (2) ผลการกำหนดจุดตัด โดยการกำหนดเกณฑ์พื้นที่บน Wright Map พบว่า มิติกระบวนการทางคณิตศาสตร์ แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุด ที่ -0.99 0.26 0.44 และ 0.61 ตามลำดับ และมิติโครงสร้างความคิดรวบยอด แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุด ที่ -0.70 0.15 0.76 และ 1.39 ตามลำดับ ซึ่งจุดตัดดังกล่าวสามารถนำไปสู่การกำหนดช่วงระดับความสามารถ คะแนนสเกล และคะแนนดิบเพื่อนำไปใช้เป็นเกณฑ์การประเมินระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ และ (3) ตัวแปรจากข้อมูลหัตถ์นิยม ที่นำไปใช้ในการสร้างสมการทำนายระดับความสามารถทางคณิตศาสตร์ผ่านการเรียนรู้ของเครื่อง (machine learning) พบว่า ตัวแปรต้นที่จะส่งผลต่อการสร้างสมการทำนาย คือ เกรดคณิตเทอมที่ผ่านมา เกรดเฉลี่ยรวมเทอมที่ผ่านมา ชั่วโมงเรียนพิเศษและชั่วโมงศึกษาคณิตศาสตร์ด้วยตนเอง และตัวแปรตามประกอบด้วย คะแนนดิบมิติกระบวนการทางคณิตศาสตร์ คะแนนดิบมิติโครงสร้างความคิดรวบยอด และคะแนนรวม 2 มิติ

2. ผลการออกแบบการระบบตรวจให้คะแนนผ่าน machine learning สามารถแบ่งออกเป็น 5 ส่วน ประกอบด้วย (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนแสดงผล (4) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ และ (5) ส่วนรายงานผลการประเมิน โดยผลการประเมินคุณภาพของการออกแบบระบบได้ผ่านผู้เชี่ยวชาญจำนวน 5 คน ด้วยวิธีการสนทนากลุ่ม ครอบคลุมทั้ง 5 มาตรฐาน ประกอบด้วย ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมขอค่าธรรมเนียม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน

คำสำคัญ: การตรวจให้คะแนนแบบสอบอัตโนมัติ การให้ข้อมูลป้อนกลับอัตโนมัติ วิทยาการออกแบบ

โมเดลการตอบสนองข้อสอบพหุมิติ การเรียนรู้ของเครื่อง

Abstract

The study aimed to (1) analyze students' response patterns for the design of the prototypical innovation, including students' misconceptions, determining cut scores for mathematical proficiency levels, and providing the automated feedback based on misconceptions and variables used to make predictions through machine learning; and (2) design the scoring system for the open-ended question test through machine learning. Design research was applied in this study. The sample comprised 495 grade 7 students. The research instrument was an open-ended test on the topic of Statistics and Probability through diagnostic tools in an online testing system, "eMAT-Testing." The collected data were

analyzed using a multidimensional random coefficients multinomial model (MRCMLM). The results are reported below.

1. Regarding the analysis of the students' response patterns based on the secondary data, it was found that (1) the students had 4 types of misconceptions: 1) misused data; 2) distorted theorems, rules, formulas, definitions and properties; 3) unverified solutions; and 4) misinterpreted language and mathematical symbols. In both the mathematical process dimension (MAP) and the conceptual structures dimension (SLO), the students' most common misconception among the 4 types was the misused data. (2) On determining cut scores with defined criterion zones on the Wright map for diagnosing mathematical proficiency levels through Machine learning, it was found that the mathematic proficiency of MAP was divided into five levels based on four cut scores, ranging from the lowest to the highest as follows: -0.99, 0.26, 0.44 and 0.61 respectively. In the same manner, that of SLO was distinguished into five levels according to four cut scores, including -0.70, 0.15, 0.76 and 1.39 respectively. Such cut scores can be employed to determine proficiency level ranges, scale scores, and raw scores as criteria for assessment of mathematical proficiency levels in each dimension. (3) And regarding the variables from secondary data used to generate the equation to predict mathematical proficiency levels through machine learning, the results showed that the independent variables that would affect the generation of the predictive equation were the mathematics GPA in the previous semester, the total GPA of the previous semester, the number of hours of tuition, and that of hours of self-study on mathematics. In addition, the dependent variables included raw scores of MAP, raw scores of SLO, and the total score of both dimensions.

2. The results also showed that the design of the scoring system for the open-ended question test through machine learning featured 5 parts, namely input, processing, output, automated feedback, and assessment reporting. Based on the qualitative assessment of the system design conducted by 5 experts through a focus group discussion, it covered 5 standards, including utility, feasibility, propriety, accuracy, and evaluation accountability.

Keywords: open-ended question scoring system, automated feedback, design research, multidimensional item response model, machine learning

บทนำ

ความสามารถทางคณิตศาสตร์เป็นสิ่งที่สะท้อนให้เห็นว่าผู้เรียนมีความรู้และทักษะที่เพียงพอหรือไม่ต่อการดำรงชีวิตในศตวรรษที่ 21 ดังนั้นการเตรียมผู้เรียนให้มีความสามารถทางคณิตศาสตร์ จึงเป็นสิ่งสำคัญ ซึ่งส่งผลให้ผู้เรียนรู้เท่าทันการเปลี่ยนแปลงของระบบเศรษฐกิจ สังคม วัฒนธรรม และสภาพแวดล้อม สามารถอยู่ร่วมกับประชาคมโลกได้ (Ministry of Education, 2017) และการประเมินความสามารถทางคณิตศาสตร์ถือเป็นสิ่งสำคัญในการขับเคลื่อนการศึกษา โดยเฉพาะอย่างยิ่ง การประเมินความสามารถทางคณิตศาสตร์ที่ไม่ได้มุ่งเน้นเพียงคำตอบที่ถูกต้องเท่านั้น แต่ยังมุ่งเน้นไปที่กระบวนการคิดทางคณิตศาสตร์ด้วยเช่นกัน (Chinjunthuk & Junpeng, 2020)

จากผลการประเมินการเรียนรู้โครงการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ของนักเรียนไทยเทียบกับนานาชาติ (TIMSS) โดยผลการประเมินของประเทศไทยด้านคณิตศาสตร์ในระดับชั้นมัธยมศึกษาปีที่ 2 พบว่า วิชาคณิตศาสตร์ ในปี ค.ศ. 1999 2007 2011 และ 2015 มีคะแนนเฉลี่ยแนวโน้มน้อยลงอย่างต่อเนื่อง (Chinjunthuk & Junpeng, 2020) ประกอบกับผลการทดสอบทางการศึกษาระดับชาตินักเรียนขั้นพื้นฐาน (O-NET) บ่งชี้เช่นเดียวกันว่าผู้เรียนมีคะแนนเฉลี่ยของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ต่ำกว่าร้อยละ 50 ซึ่งเป็นมาตรฐานขั้นต่ำ โดยพบว่า ผลการสอบ O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2563 สารที่ได้คะแนนน้อยที่สุด คือ สารที่ 3 สถิติและความน่าจะเป็น มีคะแนนเฉลี่ย 23.85 คะแนน (The National Institute of Educational Testing Service, 2021) ซึ่งการประเมินทั้งสองรูปแบบนี้ต่างมุ่งเน้นการประเมินความสามารถทางคณิตศาสตร์ของผู้เรียน จากผลการประเมินทั้งในระดับนานาชาติและระดับชาติข้างต้น สะท้อนให้เห็นว่านักเรียนไทยมีปัญหาในการเรียนรู้คณิตศาสตร์

เครื่องมือที่นับว่าสำคัญ และมีประโยชน์ในการค้นหาข้อบกพร่อง ตลอดจนปัญหาของผู้เรียนแต่ละคน คือ แบบสอบวินิจฉัย (diagnostic test) ผลการตอบสามารถระบุได้ว่านักเรียนบกพร่องในทักษะใด (Adulpongpaisal, 2009) จากการศึกษาที่ผ่านมาพบว่า มีงานวิจัยที่พัฒนานวัตกรรมระบบการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนทั้งในบริบทของต่างประเทศและในประเทศ เช่น งานวิจัยของ Junpeng et al. (2020) ได้พยายามพัฒนาระบบการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ผ่านเทคโนโลยีดิจิทัล ในสาระสถิติและความน่าจะเป็น ของนักเรียนระดับมัธยมศึกษาปีที่ 1 อย่างไรก็ตามนวัตกรรมการประเมินดังกล่าวยังมีข้อจำกัด คือ ไม่สามารถตรวจให้คะแนนแบบสอบอัตโนมัติในการที่มีกรให้คะแนนหลายค่าและไม่สามารถให้ข้อมูลย้อนกลับแบบอัตโนมัติเพื่อวินิจฉัยความสามารถผู้เรียนได้ ในกรณีที่มุ่งเน้นไปที่กระบวนการแก้ปัญหาของผู้เรียน โดยสามารถตรวจอัตโนมัติได้เฉพาะในกรณีที่มุ่งเน้นไปที่คำตอบที่ถูกของผู้เรียน และนำผลการตอบที่ได้ไปเทียบกับเกณฑ์ให้คะแนนที่สอดคล้องกับแผนที่เชิงโครงสร้างในแต่ละระดับ จึงสามารถให้คำแนะนำหรือแหล่งในการเรียนรู้เพิ่มเติม เพื่อให้ผู้เรียนปรับปรุงได้ในทันที หากต้องการพิจารณาถึงกระบวนการแสดงวิธีทำของผู้เรียนเพื่อสะท้อนให้เห็นถึงกระบวนการแก้ปัญหา จำเป็นต้องอาศัยระยะเวลาในการให้ครูหรือผู้ประเมิน ตรวจให้คะแนน แล้วจึงนำผลการตรวจให้คะแนนที่ได้ไปสู่การรายงานผลการวินิจฉัยเป็นรายบุคคล และภาพรวมในลำดับต่อไป

จากที่กล่าวมาจะเห็นได้ว่า ระบบดังกล่าวจำเป็นต้องอาศัยระยะเวลาให้ผู้เรียนดำเนินการทดสอบแล้วเสร็จ จึงสามารถประมวลผลและวินิจฉัยระดับความสามารถทางคณิตศาสตร์ได้ จึงถือเป็นจุดอ่อนที่สำคัญ ซึ่งในการวิจัยครั้งนี้ คณะผู้วิจัยจึงสนใจนำการเรียนรู้ของเครื่อง (machine learning) มาเป็นเครื่องมือในการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียน เนื่องจากเป็นเครื่องมือที่กำลังแพร่หลาย สำหรับการวิเคราะห์ข้อมูลจำนวนมาก สามารถเรียนรู้ได้จากข้อมูลที่ซับซ้อนเพื่อสร้างแบบจำลอง (Artificial Intelligence: AI) เป็นกระบวนการที่ทำให้คอมพิวเตอร์มีความสามารถในการเรียนรู้ หลักการทำงานด้วยตนเอง สามารถเลียนแบบความฉลาดและความสามารถของมนุษย์ ผ่านทางชุดคำสั่งในภาษาคอมพิวเตอร์ หรือ อัลกอริทึม สามารถเรียนรู้และปรับตัวตามข้อมูลที่ได้รับ โดยอาศัยแบบจำลอง (model) ที่สร้างจากชุดข้อมูลที่มีความสัมพันธ์กัน และสามารถนำไปทำนาย หรือใช้ในการตัดสินใจภายหลัง (Thamcharoen, 2019; Shalev-Shwartz & Ben-David, 2014) จัดเป็นเทคนิคและเครื่องมือที่ช่วยในการวินิจฉัย และพยากรณ์ได้ว่า ผู้เรียนเป็นอย่างไร มีความสามารถทางคณิตศาสตร์ระดับใด ควรได้รับการปรับปรุงพัฒนาอย่างไร โดยผู้เรียนจะทราบผลการวินิจฉัยได้ ต้องมีการให้ข้อมูลป้อนกลับ การให้ข้อมูลของครูที่เกี่ยวข้องโดยตรงกับเป้าหมายของการเรียนรู้ ช่วยให้ผู้เรียนได้รับข้อมูลสารสนเทศเกี่ยวกับความคืบหน้าหรือจุดบกพร่องของผู้เรียนต่อเป้าหมายดังกล่าว และแนะนำวิธีการที่จะแก้ไขข้อบกพร่องนั้นเพื่อให้ผู้เรียนสามารถทำได้ใกล้เคียงกับเป้าหมายที่คาดหวังไว้ (Brookhart, 2008) โดยในการวิจัยครั้งนี้ให้ข้อมูลป้อนกลับ 3 ลักษณะที่สำคัญ (Hattie & Timperley, 2007) คือ (1) การให้ข้อมูลกระตุ้นการเรียนรู้ (feed up) (2) การให้ข้อมูลป้อนกลับ (feedback) และ (3) การให้ข้อมูลเพื่อการเรียนรู้ต่อยอด (feed forward)

ดังนั้นคณะผู้วิจัย มีความสนใจในการนำ machine learning มาช่วยในการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 1 สาระที่ 3 สถิติและความน่าจะเป็น โดยต่อยอดจากงานวิจัยของ Junpeng et al. (2020) ซึ่งจะทำการวินิจฉัยระดับความสามารถใน 2 มิติ คือ มิติกระบวนการทางคณิตศาสตร์ และมิติโครงสร้างความคิดรวบยอด โดยใช้เทคนิคการวิเคราะห์ข้อมูลตามโมเดลการตอบสนองข้อสอบแบบพหุมิติ (MRCML) (Adams et al., 1997) และโมเดลเชิงโครงสร้าง (Wilson, 2005) ที่เน้นการตีความอย่างมีความหมาย เพื่อนำไปสู่การวินิจฉัยและปรับปรุงพัฒนานักเรียนได้อย่างชัดเจนมากขึ้น (Junpeng, 2018) และสามารถทำนายความสามารถทางคณิตศาสตร์ของผู้เรียนโดยผ่านการเรียนรู้ของเครื่อง (machine learning) สำหรับตัวแปรที่ใช้ในการทำนาย ตัวแปรอิสระ คือ เกรดเฉลี่ยรวมเทอมที่ผ่านมา เกรดคณิตศาสตร์เทอมที่ผ่านมา จำนวนชั่วโมงศึกษาคณิตศาสตร์ และชั่วโมงการเรียนรู้พิเศษ ผลการทำนายนำไปสู่การพัฒนาวัตกรรมการวัดแบบในการวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่าน machine learning สร้างองค์ความรู้ด้านวัดและประเมินผลผู้เรียนโดยใช้เทคโนโลยี นำไปสู่การพัฒนาความสามารถทางคณิตศาสตร์ของผู้เรียนจากระดับหนึ่งไปยังอีกระดับที่สูงขึ้น จะเป็นประโยชน์อย่างยิ่งต่อผู้เรียนหากได้ระบบการตรวจแบบสอบอัตโนมัติ โดยให้ข้อมูลป้อนกลับอัตโนมัติ ที่ผู้เรียนสามารถเข้าถึงการทดสอบและวินิจฉัยระดับความสามารถของผู้เรียนได้ตลอดเวลา เป็นประโยชน์ต่อครูผู้สอน ผู้ปกครอง และผู้ที่เกี่ยวข้อง ในการนำผลการวินิจฉัยความสามารถทางคณิตศาสตร์ ไปสู่การปรับปรุง พัฒนาผู้เรียนให้อยู่ในระดับที่สูงขึ้น นอกจากนี้ยังสามารถวินิจฉัยและตรวจให้คะแนนแบบสอบได้จำนวนหลายคนพร้อมกัน

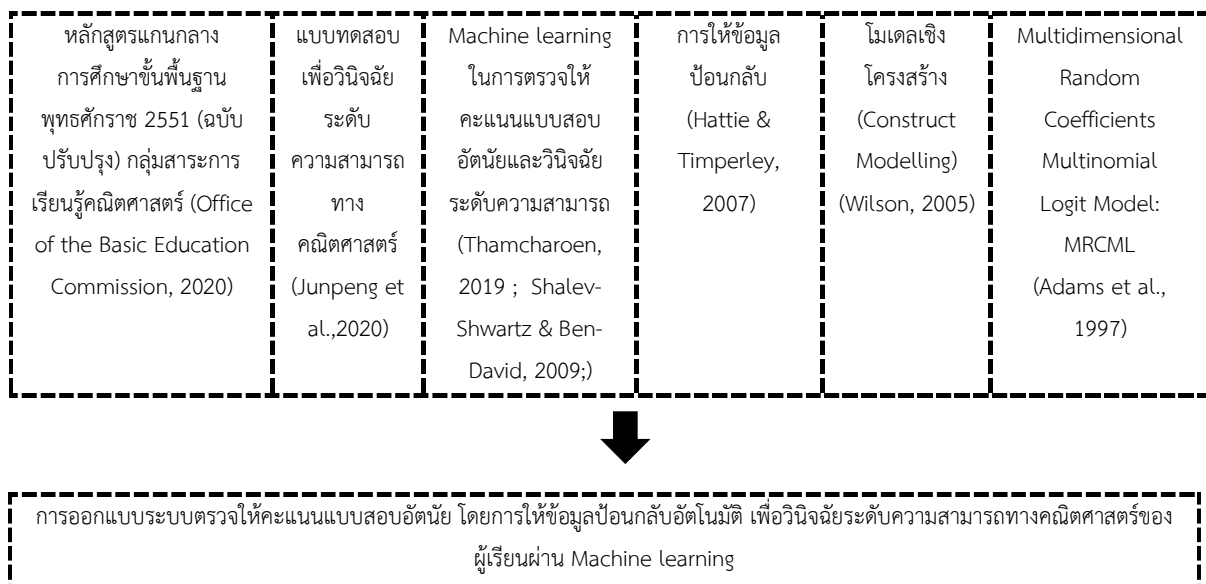
ช่วยลดทรัพยากรบุคคลในการตรวจแบบสอบ และช่วยลดระยะเวลา ขั้นตอน งบประมาณและข้อจำกัดต่าง ๆ ในการตรวจ นำไปใช้ได้ ในสถานศึกษาหลายแห่ง โดยใช้แบบสอบที่เป็นมาตรฐานการตรวจเดียวกัน ได้อย่างแม่นยำและมีประสิทธิภาพต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์ผลการตอบของผู้เรียนสู่การออกแบบนวัตกรรมต้นแบบ ประกอบด้วย มโนทัศน์ที่คลาดเคลื่อนของผู้เรียน การกำหนดจุดตัดระดับความสามารถทางคณิตศาสตร์ การให้ข้อมูลป้อนกลับตามมโนทัศน์ที่คลาดเคลื่อนและตัวแปรที่ใช้ในการทำนาย machine learning
2. เพื่อออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ ของผู้เรียนผ่านการเรียนรู้ของ machine learning

กรอบแนวคิดการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง สามารถกำหนดกรอบแนวคิดการวิจัยได้ดังภาพ 1



ภาพ 1 กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ใช้ระเบียบวิธีวิจัยการออกแบบ (design research) (Wongwanit, 2020) โดยแบ่งเป็น 4 ระยะ คือ (1) วิเคราะห์ผลการตอบของผู้เรียนสู่การออกแบบนวัตกรรมต้นแบบ (2) ออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ ของผู้เรียนผ่านการเรียนรู้ของ Machine learning (3) นำนวัตกรรมต้นแบบไปทดลองใช้เพื่อนำข้อมูลที่ได้มาวิเคราะห์ ปรับปรุงพัฒนาระบบให้สมบูรณ์ และรายงานคุณภาพของนวัตกรรม และ (4) เสนอหลักการออกแบบใหม่ในการพัฒนานวัตกรรมต้นแบบ โดยการนำเสนอในบทความครั้งนี้เป็นการศึกษานำร่องซึ่งจะนำเสนอสองระยะแรก มีรายละเอียดทั้งสองระยะดังนี้

ระยะที่ 1 วิเคราะห์ผลการตอบของผู้เรียนเพื่อนำไปสู่การออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ

คณะผู้วิจัยร่วมกับครุคณิตศาสตร์ วิศวกรรมคอมพิวเตอร์ ผู้เชี่ยวชาญในด้านการวัดและประเมินผล และผู้เชี่ยวชาญในด้านการใช้เทคโนโลยี จำนวน 5 คน เพื่อร่วมกันวิเคราะห์คำตอบทั้งหมดที่ผู้เรียนตอบ ข้อคิดเห็นและวิธีคิดต่าง ๆ ของผู้เรียนที่มีต่อเรื่อง สถิติและความน่าจะเป็น

1. กลุ่มผู้สอบ

กลุ่มผู้สอบที่ใช้วิเคราะห์ผลการตอบ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2562 จำนวนนักเรียนทั้งหมด 495 คน จากข้อมูลทุติยภูมิในโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ของโรงเรียนใน 4 ภูมิภาค คือ ภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคกลางและภาคใต้ จากทั้งหมด 23 โรงเรียนโดยให้ครอบคลุมทุกระดับความสามารถของผู้เรียนทั้งเก่ง ปานกลาง และอ่อน (Junpeng et al., 2020) เนื่องจากการศึกษาในระยะนี้ใช้โมเดล MRCML (Adams et al., 1997) ในการวิเคราะห์ข้อมูลและการกำหนดจุดตัด จึงให้ความสำคัญกับการกำหนดขนาดตัวอย่างและการได้มาของตัวอย่างที่ครอบคลุมเพียงพอในทุกระดับความสามารถในแต่ละมิติที่ทำการศึกษา

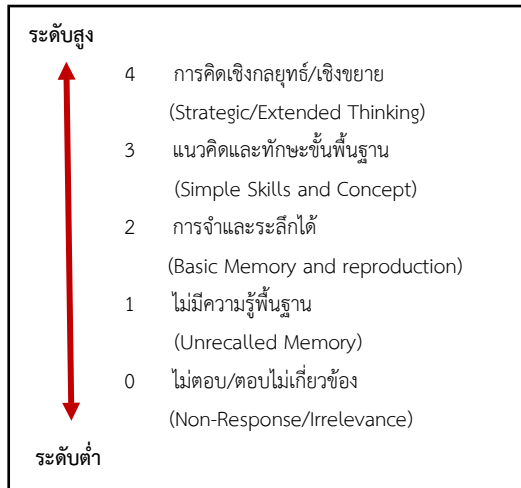
2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นเครื่องมือทุติยภูมิ มาจากโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ ผ่านชุดเครื่องมือวินิจฉัยเรียกว่า “eMAT-Testing” (Junpeng et al., 2020) ซึ่ง “eMAT-Testing” เป็นชุดเครื่องมือในระบบออนไลน์แบบเรียลไทม์ผ่าน Web Application แบบตอบสนองโต้ตอบได้โดยตรง (interactive) รองรับระบบ iOS Android และ Window ซึ่งเป็นแบบทดสอบอัตโนมัติ ตามมาตรฐานและตัวชี้วัดหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 (ฉบับปรับปรุง 2560) กลุ่มสาระการเรียนรู้คณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ในสาระที่ 3 สถิติและความน่าจะเป็น ประกอบด้วย 6 สถานการณ์ 10 ข้อย่อย แบ่งเป็นมิติกระบวนการทางคณิตศาสตร์ จำนวน 5 ข้อ และมิติโครงสร้างความคิดรวบยอด จำนวน 5 ข้อ โดยมีค่าดัชนีความตรงเชิงเนื้อหา รายข้อ (I-CVI) อยู่ระหว่าง 0.50 – 1.00 และ ค่าสัมประสิทธิ์ความเที่ยงแบบ EAP/PV ในมิติที่ 1 และ 2 มีค่าเท่ากับ 0.85 และ 0.83 ตามลำดับ และค่า INFIT MNSQ อยู่ระหว่าง 0.84 - 1.27

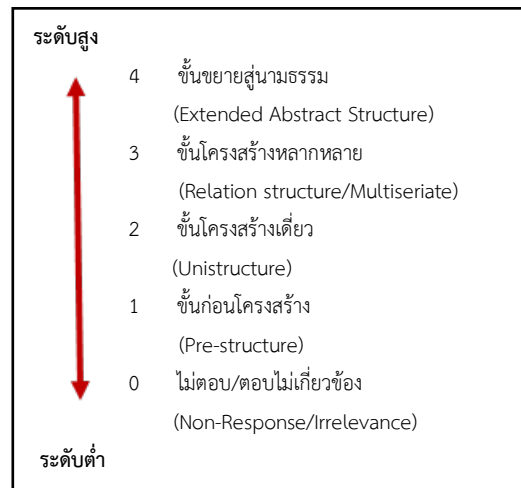
3. การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล

1. ศึกษาแผนที่โครงสร้าง ระดับความสามารถทางคณิตศาสตร์ของโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ โดยทำการสร้างแผนที่โครงสร้างของระดับผลลัพธ์การเรียนรู้เพื่อใช้ในการประเมินระดับความสามารถทางคณิตศาสตร์ 2 มิติ (Junpeng et al., 2020) คือ (1) ด้านกระบวนการทางคณิตศาสตร์ และ (2) ด้านโครงสร้างความคิดรวบยอด เพื่อใช้เป็นแนวทางในการออกแบบระบบตรวจให้คะแนนอัตโนมัติ ดังภาพ 2

มิติที่ 1 ด้านกระบวนการทางคณิตศาสตร์



มิติที่ 2 ด้านโครงสร้างความคิดรวบยอด



ภาพ 2 construct map ในแต่ละด้านเพื่อประเมินความสามารถทางคณิตศาสตร์

(Junpeng et al., 2020)

2. กำหนดจุดตัดเพื่อนำไปสู่การวินิจฉัยระดับความสามารถทางคณิตศาสตร์ในระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการพิจารณากำหนดพื้นที่บน Wright map แต่ละระดับความสามารถ โดยวิเคราะห์ผลการตอบจากข้อมูลหุติภูมิ โดยใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติ MRCML ด้วยโปรแกรม ACER ConQuest Version 2.0 (Wu et al., 2007) ในการประมาณค่าพารามิเตอร์ของข้อสอบและผู้สอบด้วยวิธี MLE พร้อมนำเสนอเป็น Wright map โดย Wright map ทางซ้ายมือนำเสนอค่าความสามารถของผู้เรียนแต่ละข้อ ส่วน Wright map ทางขวามือนำเสนอค่าระดับความสามารถทางคณิตศาสตร์แต่ละขั้นของการตอบ

3. คำนวณจุดตัดร่วมกับการพิจารณาการกำหนดเกณฑ์พื้นที่บน Wright map แต่ละระดับความสามารถทางคณิตศาสตร์ ซึ่งพิจารณาจากค่าเฉลี่ยของความยากหรือระดับ threshold เดียวกัน หาดด้วยจำนวนข้อสอบที่นำมาใช้คำนวณ แยกตามมิติที่ทำการศึกษา และนำคะแนนจุดตัดที่ได้เทียบกับพื้นที่บน Wright map แบ่งเป็นช่วงระดับความสามารถตามที่กำหนดไว้ในโมเดลเชิงโครงสร้างในแต่ละมิติ

4. นำค่าจุดตัดซึ่งเป็นพารามิเตอร์ความสามารถที่เหมาะสมที่ปรากฏใน Wright map และแปลงเป็นคะแนนสเกล (θ_d) ที่เป็นคะแนนมาตรฐานในการวินิจฉัยโดยใช้สูตรดั่งสมการคะแนนสเกล $= 50 + 10(\theta_d)$

5. วิเคราะห์ผลการตอบของผู้เรียนทั้งหมด เพื่อหาแนวโน้มที่คลาดเคลื่อนรูปแบบต่าง ๆ และนำไปสู่การออกแบบรูปแบบการให้ข้อมูลป้อนกลับตามแนวโน้มที่คลาดเคลื่อน และวิเคราะห์ผลการตอบของผู้เรียน เพื่อนำรูปแบบคำตอบที่เป็นไปได้ทั้งหมดไปเป็นข้อมูลต้นแบบในการพัฒนาคำถามและแสดงวิธีหาคำตอบในระบบตรวจให้คะแนนอัตโนมัติ ที่สอดคล้องกับระดับความสามารถตามแผนที่เชิงโครงสร้าง และคำอธิบายในแต่ละระดับความสามารถ

6. วิเคราะห์รูปแบบการให้ข้อมูลป้อนกลับเพื่อพัฒนาผู้เรียนตามแนวโน้มที่คลาดเคลื่อน ในการพัฒนาระดับความสามารถทางคณิตศาสตร์ของผู้เรียน

7. พัฒนาคำถามตามแผนที่โครงสร้าง เป็นการออกแบบข้อคำถามแบบอัตนัยเพื่อให้ได้ข้อมูลตามที่บรรยายไว้ในแผนที่โครงสร้าง โดยได้ออกแบบจากการวิเคราะห์ผลการตอบของผู้เรียน และข้อสอบแต่ละข้อเป็นการตรวจให้คะแนนแบบหลายค่า (polytomously scoring)

8. จัดเตรียมข้อมูลพื้นฐานของผู้เรียนเพื่อเข้าสู่การทำนายระดับความสามารถทางคณิตศาสตร์ของผู้เรียน โดย machine learning

ระยะที่ 2 ออกแบบนวัตกรรมต้นแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติผ่าน Machine learning

นำข้อมูลผลการตอบ ที่ได้จากการวิเคราะห์ด้วยการสนทนากลุ่ม (focus group) ระหว่างผู้วิจัยผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา วิศวกรคอมพิวเตอร์ และผู้เชี่ยวชาญด้านเทคโนโลยี ฯ จำนวน 5 คน จากข้อมูลทฤษฎีในโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ (Junpeng et al., 2020) มาออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ โดยใช้แบบสอบอัตโนมัติจากข้อมูลทฤษฎีจำนวน 10 ข้อ แบ่งเป็นมิติกระบวนการทางคณิตศาสตร์ 5 ข้อ และมิติโครงสร้างความคิดรวบยอด 5 ข้อ ก่อนการนำไปเก็บข้อมูลจริง

1. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือคือแบบประเมินการกำหนดมาตรฐานด้วยแบบอิงมาตรฐาน โดยคุณภาพเครื่องมือตามมาตรฐานการประเมินคุณภาพของนวัตกรรมต้นแบบตาม Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2010; Varasunun, 2011) ประกอบด้วย 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมขอธรรม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน

2. การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูล

1. รวบรวมข้อมูลประกอบด้วย ผลการกำหนดจุดตัดจากระยะที่ 1 ข้อคำถามเพื่อวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนของผู้เรียนรายข้อและภาพรวม การให้ข้อมูลป้อนกลับในแต่ละรูปแบบตามมโนทัศน์ที่คลาดเคลื่อน และชุดข้อมูลปัจจัยนำเข้าในสมการทำนาย

2. ร่วมกันออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ และการให้ข้อมูลป้อนกลับอัตโนมัติตามมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน

3. กำหนดองค์ประกอบของระบบการให้ข้อมูลป้อนกลับอัตโนมัติ แผนผังการทำงานของนวัตกรรมผ่าน machine learning (work flow) พร้อมทั้งหน้าที่ในแต่ละองค์ประกอบ

4. พิจารณาคุณภาพของนวัตกรรมต้นแบบที่พัฒนาขึ้นก่อนนำไปทดลองใช้ โดยผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา ด้านการใช้เทคโนโลยีและครุคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 1 เพื่อพิจารณาความถูกต้อง ความเหมาะสม และความเป็นไปได้ในการนำไปใช้ โดยผลการตรวจสอบคุณภาพของระบบใน 2 ด้าน คือ (1) การประเมินระบบด้วยแบบฮิวริสติก และ (2) การประเมินแบบอิงมาตรฐาน

5. นำผลการประเมินที่ได้มาร่วมกันสะท้อนคิดและปรับปรุงแก้ไขเป็นนวัตกรรมการประเมินต้นแบบก่อนนำไปทดลองใช้ในลำดับถัดไป

ผลการวิจัย

การวิจัยเรื่อง การออกแบบระบบตรวจให้คะแนนแบบสออัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของเครื่อง สรุปผลการวิจัยดังนี้

1. ผลการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนจากผลการตอบของผู้เรียน จุดตัดระดับความสามารถทางคณิตศาสตร์ รูปแบบการให้ข้อมูลป้อนกลับตามมโนทัศน์ที่คลาดเคลื่อนและตัวแปรที่ใช้ในการทำนาย Machine learning

1.1 ผลการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อน ในการแก้โจทย์ปัญหาในสถานการณ์เกี่ยวกับเรื่องสถิติ และความน่าจะเป็น ผู้วิจัยได้ทำการวิเคราะห์ข้อสอบรายข้อ เพื่อหารูปแบบมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน ซึ่งเป็นเครื่องมือจากโครงการการสร้างเครื่องมือวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียน โดยใช้เทคโนโลยีสารสนเทศในการรายงานผลการประเมินเพื่อการเรียนรู้ (Junpeng et al., 2020) พบว่า

1) มิติกระบวนการทางคณิตศาสตร์ (MAP) นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนมากที่สุดคือ ประเภท ด้านการใช้ข้อมูล ละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา ทำผิดคำสั่ง หาคำตอบในสิ่งที่ไม่ต้องการ ร้อยละ 25.45 และมโนทัศน์ที่คลาดเคลื่อนน้อยที่สุดคือ ด้านการคำนวณและขาดการตรวจสอบในระหว่างการแก้ปัญหา ร้อยละ 5.58 สามารถแสดงมโนทัศน์ของผู้เรียนรายข้อ ดังแสดงในตาราง 1

ตาราง 1 ผลการวิเคราะห์มโนทัศน์ที่คลาดเคลื่อนมิติกระบวนการทางคณิตศาสตร์

ข้อที่	สถานการณ์	ด้านการใช้ข้อมูล (ร้อยละ)	ด้านการ บิตเบือนทฤษฎี บท กฎ สูตร บทนิยาม และ สมบัติ (ร้อยละ)	ด้านการคำนวณ และขาดการ ตรวจสอบใน ระหว่างการ แก้ปัญหา (ร้อยละ)	ผิดพลาดในการตีความ ด้านภาษาและ สัญลักษณ์ทาง คณิตศาสตร์ (ร้อยละ)	ไม่มี มโนทัศน์ที่ คลาดเคลื่อน (ร้อยละ)	ไม่ตอบ (ร้อยละ)
1	1.1	33.33	1.62	3.03	29.70	16.16	21.41
2	1.2	21.01	1.21	1.82	14.55	42.22	23.84
4	3	35.35	9.90	13.54	26.26	1.41	23.84
5	4.1	16.36	12.73	5.86	7.07	28.08	34.14
6	4.2	21.21	13.54	3.64	8.69	19.8	37.98
รวม		25.45	7.80	5.58	17.25	21.54	21.53

2) มิติโครงสร้างความคิดรวบยอด (SLO) นักเรียนมีมโนทัศน์ที่คลาดเคลื่อนมากที่สุด คือ ด้าน การใช้ข้อมูล ละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา ทำผิดคำสั่ง หาคำตอบในสิ่งที่ไม่ต้องการ ร้อยละ 9.05 และมโนทัศน์ที่คลาดเคลื่อนน้อยที่สุด คือ ด้านผิดพลาดในการตีความด้านภาษาและสัญลักษณ์ทางคณิตศาสตร์ร้อยละ 4.40 สามารถแสดงมโนทัศน์ของผู้เรียนรายข้อ ดังแสดงในตาราง 2

ตาราง 2 ผลการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อนมิติโครงสร้างความคิดรวบยอด

ข้อที่	สถานการณ์	ด้านการใช้ข้อมูล (ร้อยละ)	ด้าน การบิดเบือน ทฤษฎีบท กฎ สูตร บทนิยาม และสมบัติ (ร้อยละ)	ด้านการคำนวณ และขาด การตรวจสอบ ในระหว่างการ แก้ปัญหา (ร้อยละ)	ผิดพลาดใน การตีความด้าน ภาษาและ สัญลักษณ์ทาง คณิตศาสตร์ (ร้อยละ)	ไม่มี หมโนทัศน์ที่ คลาดเคลื่อน (ร้อยละ)	ไม่ตอบ (ร้อยละ)
3	2	10.71	12.32	3.84	4.65	54.34	17.17
7	5.1	3.64	-	7.68	3.43	73.74	14.95
8	5.2	3.84	-	6.06	1.82	73.54	18.18
9	5.3	7.88	14.55	8.89	8.08	23.43	37.98
10	6	19.19	13.94	5.05	4.04	24.65	34.34
รวม		9.05	8.16	6.30	4.40	49.94	49.93

1.2 การพัฒนาความสามารถทางคณิตศาสตร์ของผู้เรียน ในการออกแบบระบบการให้ข้อมูล ป้อนกลับแบบอัตโนมัติ ตามหมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน พบว่า สามารถกำหนดขอบข่ายของการให้ ข้อมูลป้อนกลับอัตโนมัติ โดยเน้นการให้ข้อมูลป้อนกลับใน 4 รูปแบบ คือ (1) แบบชี้แนะบางส่วน (2) แบบยกตัวอย่าง (3) แบบอธิบายรายละเอียด และ (4) แบบอธิบายรายละเอียดและบอกคำตอบ หลักใน การเลือกใช้แต่ละรูปแบบ ในแต่ละข้ออาจใช้แบบใดแบบหนึ่งหรือหลายแบบบูรณาการรวมกันได้ ขึ้นอยู่กับ หมโนทัศน์ที่คลาดเคลื่อนของผู้เรียน ซึ่งได้จากการวิเคราะห์ผลการตอบผู้เรียน เพื่อนำไปสู่การกำหนดข้อความ ป้อนกลับอัตโนมัติในระบบ ตัวอย่างดังปรากฏในตาราง 3

ตาราง 3 ตัวอย่างการกำหนดขอบข่ายของการให้ข้อมูลป้อนกลับเพื่อนำไปสร้างข้อความอัตโนมัติในระบบ

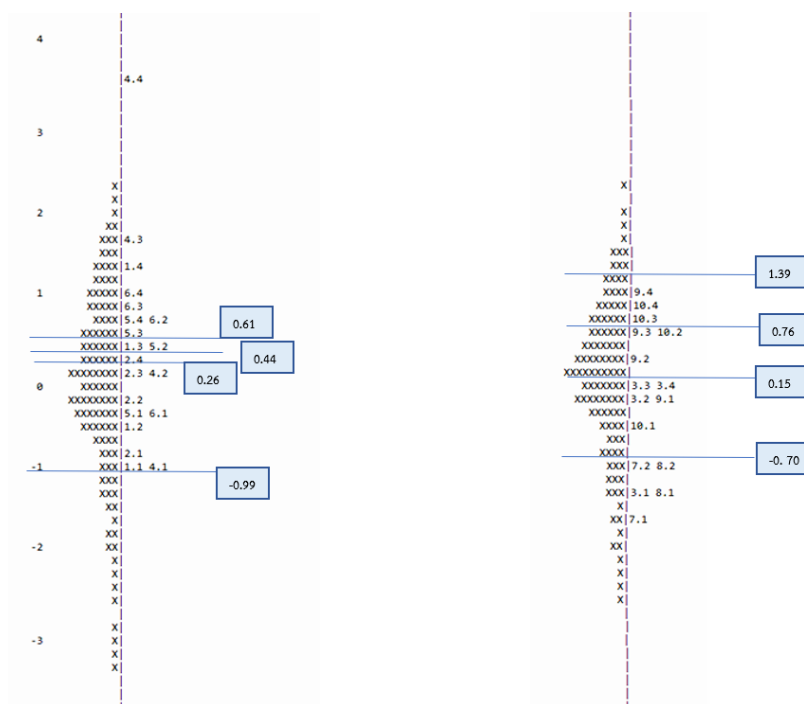
รูปแบบหมโนทัศน์ที่ คลาดเคลื่อน	รูปแบบการให้ ข้อมูลป้อนกลับ	ข้อความป้อนกลับที่ปรากฏในระบบ
ตัวอย่างแบบสอบอัตรย์ สถานการณ์ที่ 6 : ครูนำตัวอักษรจากคำว่า FATHER มาทำเป็นบัตรคำให้นักเรียนสุ่มหยิบขึ้นมา 1 ครั้ง นักเรียนจะมีโอกาสหยิบได้บัตรที่เป็นสระของภาษาอังกฤษ คิดเป็นเท่าไร?		
ด้านการใช้ข้อมูล	แบบชี้แนะบางส่วน	แนวคิด สระในภาษาอังกฤษ คือ A E I O U ดังนั้นคำว่า FATHER มีสระ _ ตัว
ผิดพลาดใน การตีความด้านภาษา และสัญลักษณ์ทาง คณิตศาสตร์	แบบอธิบาย รายละเอียด	แนวคิด จากคำว่า FATHER มีอักษรทั้งหมด _ ตัว และมีสระทั้งหมด _ ตัว ดังนั้น โอกาสหรือความน่าจะเป็น เท่ากับ อัตราส่วนของจำนวนเหตุการณ์ที่เราสนใจ ต่อ จำนวนผลลัพธ์ทั้งหมดที่อาจเกิดขึ้นได้ $P(E) = \frac{n}{N}$
ด้านการคำนวณและ ขาดการตรวจสอบใน ระหว่างการแก้ปัญหา	ยกตัวอย่าง	แนวคิด ยกตัวอย่าง คำว่า MONDAY มีอักษรที่เป็นสระ 2 ตัว คือ O และ A จากสระทั้งหมด 6 ตัว ดังนั้น โอกาสหยิบได้บัตรที่เป็นสระของภาษาอังกฤษ คิดเป็น $\frac{2}{6}$
ด้านการบิดเบือน ทฤษฎีบท กฎ สูตร บทนิยาม และสมบัติ	แบบอธิบาย รายละเอียดและ ยกตัวอย่าง	แนวคิด จากคำว่า FATHER มีอักษรทั้งหมด _ ตัว และมีสระทั้งหมด _ ตัว โอกาสหรือความน่าจะเป็น เท่ากับ อัตราส่วนของจำนวนเหตุการณ์ที่เราสนใจ(n) ต่อ จำนวน ผลลัพธ์ทั้งหมดที่อาจเกิดขึ้นได้(N) ซึ่งมีสูตรในการคิดคำนวณดังนี้ $P(E) = \frac{n}{N}$ ยกตัวอย่าง คำว่า MONDAY มีอักษร 6 ตัว มีสระ 2 ตัว คือ O และ A ดังนั้น โอกาสหยิบได้บัตรที่เป็นสระของภาษาอังกฤษ คิดเป็น $\frac{2}{6}$

1.3 ผลการคำนวณจุดตัดร่วมกับการพิจารณาการกำหนดเกณฑ์พื้นที่บน Wright map แต่ละระดับความสามารถทางคณิตศาสตร์ ซึ่งพิจารณาจากค่าเฉลี่ยของความยากหรือระดับ threshold เดียวกัน หาดด้วยจำนวนข้อสอบที่นำมาใช้คำนวณ แยกตามมิติที่ทำการศึกษา ดังตาราง 4

ตาราง 4 ค่าความยากและค่า threshold ของข้อคำถาม ในแต่ละมิติเพื่อนำไปสู่การคำนวณจุดตัด

มิติกระบวนการทางคณิตศาสตร์						มิติโครงสร้างความคิดรวบยอด							
ข้อที่	ค่า	ระดับชั้นความยาก (threshold)				ข้อที่	ค่า	ระดับชั้นความยาก (threshold)					
		ความ ยาก	1	2	3			4	ความ ยาก	1	2	3	4
3	-0.497	-1.34	-0.27	-0.07	-0.07	1	0.133	-0.97	-0.46	0.48	1.44		
7	-1.281	-1.55	-1.02			2	-0.176	-0.80	-0.12	0.11	0.26		
8	-1.127	-1.30	-0.95			4	1.096	-0.99	0.10	1.69	3.58		
9	0.397	-0.27	0.29	0.62	0.99	5	0.293	-0.43	0.49	0.64	0.69		
10	0.425	-0.50	0.66	0.78	0.91	6	0.533	-0.32	0.75	0.89	0.99		
Mean		-0.99	0.26	0.44	0.61	Mean		-0.70	0.15	0.76	1.39		

ผลการกำหนดจุดตัดเพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ร่วมกับการกำหนดเกณฑ์พื้นที่บน Wright map พบว่า มิติกระบวนการทางคณิตศาสตร์ แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุด ที่ -0.99 0.26 0.44 และ 0.61 ตามลำดับ และมิติโครงสร้างความคิดรวบยอด แบ่งได้ 5 ระดับ 4 จุดตัด จากระดับต่ำสุดไปสูงสุดที่ -0.70 0.15 0.76 และ 1.39 ตามลำดับ ปรากฏดังภาพ 3



ภาพ 3 การกำหนดจุดตัดในมิติด้านกระบวนการทางคณิตศาสตร์(ซ้าย) และมิติด้านโครงสร้างความคิด รวบยอด (ขวา)

จาก 4 จุดตัดในแต่ละมิติดังกล่าวข้างต้น นำไปสู่การกำหนดช่วงการประเมินระดับความสามารถของผู้เรียนในมิติกระบวนการทางคณิตศาสตร์ และ มิติโครงสร้างทางความคิดรวบยอด เป็น 5 ช่วง ตามแผนที่โครงสร้าง โดยมีคะแนนจุดตัดแบ่งระดับความสามารถของผู้เรียน ออกเป็น 5 ระดับ ที่สามารถแปลงเป็นช่วงคะแนนสเกล และช่วงคะแนนดิบ เพื่อนำไปใช้ในการวินิจฉัยผู้เรียนเป็นรายบุคคล โดยแสดงรายละเอียดดังตาราง 5 และตาราง 6

ตาราง 5 เกณฑ์จุดตัดระดับความสามารถทางคณิตศาสตร์ในมิติกระบวนการทางคณิตศาสตร์ (MAP)

จุดตัด	มิติกระบวนการทางคณิตศาสตร์	ช่วงระดับความสามารถ (θ)	คะแนนสเกล	ช่วงคะแนนดิบ (คะแนนเต็ม = 17 คะแนน)
$\theta = 0.61$	ระดับ 5	$\theta = 0.61$ ขึ้นไป	56.10 ขึ้นไป	14 ขึ้นไป
$\theta = 0.44$	ระดับ 4	$\theta = 0.44$ ถึง 0.60	54.4 – 56	9-13
$\theta = 0.26$	ระดับ 3	$\theta = 0.26$ ถึง 0.43	52.60 – 54.30	4-8
$\theta = -0.99$	ระดับ 2	$\theta = -0.99$ ถึง 0.25	40.10 – 52.50	2-3
	ระดับ 1	ต่ำกว่า $\theta = -0.99$	ต่ำกว่า 40.10	0-1

จากตาราง 5 มิติกระบวนการทางคณิตศาสตร์ (MAP) พบว่า มีจุดตัด 4 จุดตัด นำไปสู่การกำหนดช่วงการประเมินระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ เป็น 5 ช่วง ที่สามารถแปลงเป็นช่วงคะแนนสเกล และช่วงคะแนนดิบ เพื่อนำไปใช้ในการวินิจฉัยผู้เรียนเป็นรายบุคคล เช่น ผู้สอบที่ประมาณค่าพารามิเตอร์ระดับความสามารถด้านกระบวนการทางคณิตศาสตร์ตั้งแต่ 0.61 (θ) ≥ 0.61 แสดงว่า ผู้สอบจะได้รับการวินิจฉัยระดับความสามารถทางคณิตศาสตร์อยู่ในระดับที่ 5 โดยมีคะแนนสเกลเท่ากับ 56.10 ขึ้นไป และเมื่อนำไปเทียบกับคะแนนดิบจะมีค่าเท่ากับ 14 คะแนนขึ้นไป จากคะแนนเต็ม 17 คะแนน

ตาราง 6 เกณฑ์จุดตัดระดับความสามารถทางคณิตศาสตร์ในมิติโครงสร้างความคิดรวบยอด (SLO)

จุดตัด	มิติโครงสร้างความคิดรวบยอด	ช่วงระดับความสามารถ (θ)	คะแนนสเกล	ช่วงคะแนนดิบ (คะแนนเต็ม = 21 คะแนน)
$\theta = 1.39$	ระดับ 5	$\theta = 1.39$ ขึ้นไป	63.90 ขึ้นไป	14 ขึ้นไป
$\theta = 0.76$	ระดับ 4	$\theta = 0.76$ ถึง 1.38	57.6 – 63.80	10-13
$\theta = 0.15$	ระดับ 3	$\theta = 0.15$ ถึง 0.75	51.50 – 57.50	4-9
$\theta = -0.70$	ระดับ 2	$\theta = -0.70$ ถึง 0.14	43 – 51.40	2-3
	ระดับ 1	ต่ำกว่า $\theta = -0.70$	ต่ำกว่า 43	0-1

จากตาราง 6 พบว่า มิติโครงสร้างความคิดรวบยอด มีจุดตัด 4 จุดตัด นำไปสู่การกำหนดช่วงการประเมินระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ เป็น 5 ช่วง ที่สามารถแปลงเป็นช่วงคะแนนสเกล และช่วงคะแนนดิบ เพื่อนำไปใช้ในการวินิจฉัยผู้เรียนเป็นรายบุคคล เช่น ผู้สอบที่ประมาณค่าพารามิเตอร์ระดับความสามารถด้านกระบวนการทางคณิตศาสตร์ตั้งแต่ 1.39 (θ) ≥ 1.39 แสดงว่า ผู้สอบจะได้รับการ

การวินิจฉัยระดับความสามารถทางคณิตศาสตร์อยู่ในระดับที่ 5 โดยมีคะแนนสเกลเท่ากับ 63.90 ขึ้นไป และเมื่อนำไปเทียบกับคะแนนดิบจะมีค่าเท่ากับ 14 คะแนนขึ้นไป จากคะแนนเต็ม 21 คะแนน

1.4 ผลการกำหนดแนวทางการให้คะแนนแบบอัตโนมัติในระบบการตรวจ โดยใช้แผนที่โครงสร้างของระดับผลลัพธ์การเรียนรู้ เป็นเกณฑ์ในการตรวจให้คะแนนผู้เรียนในระบบรายข้อ และนำเสนอภาพรวมประกอบด้วย 2 มิติ คือ (1) กระบวนการทางคณิตศาสตร์ และ (2) โครงสร้างความคิดรวบยอด ในแต่ละมิติจะแบ่งระดับคุณลักษณะแบ่งใน 5 ระดับ และมีความเข้มข้นของการคิดเชิงคณิตศาสตร์ลดหลั่นกันจากน้อยไปมากตั้งแต่ 0 - 4 ดังตาราง 7

ตาราง 7 ตัวอย่างการตรวจให้คะแนนแบบหลายค่าตามแผนที่โครงสร้าง

ระดับ	คะแนน	ชื่อ	คำอธิบายระดับความสามารถ
เกณฑ์การให้คะแนนและตัวอย่างคำตอบของข้อคำถามกระบวนการทางคณิตศาสตร์			
ดีมาก	4	การคิดเชิงกลยุทธ์/เชิงขยาย (Strategic/Extended Thinking)	สามารถแสดงการแก้ปัญหาได้อย่างหลากหลายและสามารถสรุปเป็นวิธีการที่สามารถใช้ในระดับที่สูงขึ้นได้ พร้อมแสดงที่มาของคำตอบได้ชัดเจนโดยมีกระบวนการคิดที่เห็นเชิงประจักษ์ เป็นขั้นตอนที่ครบถ้วน
ดี	3	แนวคิดและทักษะขั้นพื้นฐาน (Simple Skills and Concept)	สามารถใช้ความเข้าใจในกระบวนการ หลักการ ทฤษฎี ทางคณิตศาสตร์มาแก้ปัญหาได้อย่างหลากหลาย สามารถแก้ไขปัญหาที่ซับซ้อนได้อย่างมาก 1 ขั้นตอนได้ พร้อมแสดงที่มาของคำตอบได้ชัดเจนโดยมีกระบวนการคิดที่เห็นเชิงประจักษ์ เป็นขั้นตอนบางส่วน
พอใช้	2	การจำและระลึกได้ (Basic Memory and Reproduction)	สามารถใช้ความเข้าใจในกระบวนการพื้นฐานทางคณิตศาสตร์ อย่างง่ายมาประกอบการหาคำตอบ ซึ่งทำให้ได้คำตอบที่ยังคงไม่สมบูรณ์ ขาดการประยุกต์ใช้
ปรับปรุง	1	ไม่มีความรู้พื้นฐาน (Unrecalled Memory)	การตอบคำถามเป็นการนำความรู้บางส่วนมาตอบแต่ผลการตอบสะท้อนให้เห็นว่าผู้เรียนไม่สามารถจดจำเนื้อหาที่สำคัญหรือจำเป็นมาเป็นพื้นฐานในการแก้ปัญหาได้
ปรับปรุง เร่งด่วน	0	ไม่ตอบ (Non- Response/ Irrelevance)	ไม่ตอบหรือตอบไม่เกี่ยวข้องกับคำถาม
เกณฑ์การให้คะแนนและตัวอย่างคำตอบของข้อคำถามทักษะด้านความคิดรวบยอด			
ดีมาก	4	ขั้นขยายสู่นามธรรม (Extended abstract structure)	- นำความคิดรวบยอดเชื่อมโยงและสรุปอ้างอิง - เป็นองค์ความรู้เชิงนามธรรมได้ - สร้างข้อสรุปจากความคิดรวบยอดได้
ดี	3	ขั้นความสัมพันธ์เชิงโครงสร้าง ความคิดรวบยอด (Relation structure)	- เชื่อมโยงความคิดรวบยอดที่ซับซ้อนได้ - จำแนก/ อธิบาย/ บรรยายแต่ละส่วนเข้าด้วยกันได้ - แต่ไม่สามารถสรุปเป็นองค์ความรู้เชิงนามธรรมได้
พอใช้	2	ขั้นโครงสร้างเดี่ยว (Unistructure)	- เชื่อมโยงความคิดรวบยอดอย่างง่ายได้ คิดรวบยอดซับซ้อนไม่ได้ - ระบุสิ่งที่เรียนรู้ในลักษณะของความจำ
ปรับปรุง	1	ขั้นก่อนโครงสร้าง (pre-structure)	- ไม่สามารถสร้างความคิดรวบยอด - ไม่สามารถเชื่อมโยงความคิดรวบยอด - มีความเข้าใจคลาดเคลื่อนในการคิดรวบยอดในเรื่องที่เรียน
ปรับปรุง เร่งด่วน	0	ไม่ตอบ (Non- Response/ Irrelevance)	ไม่ตอบหรือตอบไม่เกี่ยวข้องกับคำถาม

1.5 ผลการวิเคราะห์ตัวแปรที่จะใช้ในการสร้างสมการทำนายจากข้อมูลทฤษฎี

ข้อมูลพื้นฐานของผู้เรียนเพื่อใช้ในการทำนายความสามารถทางคณิตศาสตร์ของผู้เรียน โดย machine learning ได้จากข้อมูลทฤษฎี (Junpeng et al., 2020) นำมาพิจารณาว่าตัวแปรใดบ้างที่ส่งผลต่อระดับความสามารถทางคณิตศาสตร์ของผู้เรียนในแต่ละมิติ ดังแสดงในตาราง 8

ตาราง 8 ตารางแสดงตัวแปรอิสระและตัวแปรตามเพื่อนำไปใช้สร้างสมการทำนาย

ประเภทตัวแปร	ข้อมูลทฤษฎี	ตัวแปร
ตัวแปรอิสระ	เกรดคณิตเทอมที่ผ่านมา	X_1
	เกรดเฉลี่ยรวม	X_2
	ชั่วโมงเรียนพิเศษ	X_3
	ชั่วโมงศึกษาคณิตศาสตร์ด้วยตนเอง	X_4
ตัวแปรตาม	รวมคะแนนดิบมิติกระบวนการคณิตศาสตร์	Y_1
	รวมคะแนนดิบมิติโครงสร้างความคิดรวบยอด	Y_2
	คะแนนรวม 2 มิติ	Y_3

จากตาราง 8 พบว่าตัวแปรต้นที่จะส่งผลต่อการสร้างสมการทำนาย ได้แก่ เกรดคณิตเทอมที่ผ่านมา เกรดเฉลี่ยรวมเทอมที่ผ่านมา ชั่วโมงเรียนพิเศษและชั่วโมงศึกษาคณิตศาสตร์ด้วยตนเอง และตัวแปรตามประกอบด้วยคะแนนดิบมิติกระบวนการคณิตศาสตร์ คะแนนดิบมิติโครงสร้างความคิดรวบยอด และคะแนนรวม 2 มิติ

ผลการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อน การออกแบบรูปแบบการให้ข้อมูลป้อนกลับตามหมโนทัศน์ที่คลาดเคลื่อน ผลการกำหนดจุดตัด และตัวแปรที่ใช้ในการทำนาย ข้อมูลทั้งหมดนี้เป็นข้อมูลป้อนเข้าในการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติผ่าน machine learning

2. ผลการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติเพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของ Machine learning

จากการวิเคราะห์ผลการตอบ รูปแบบการให้ข้อมูลป้อนกลับและผลการกำหนดจุดตัดในระยะแรก นำมาสู่การออกแบบนวัตกรรมต้นแบบผ่าน Machine learning เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์และหมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนทั้ง 2 มิติ โดยระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ ประกอบด้วย 5 ส่วน คือ (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนแสดงผล (4) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ และ (5) ส่วนรายงานผลการประเมิน ดังภาพ 4



ภาพ 4 แสดงการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติผ่าน Machine learning

รายละเอียดขั้นตอนของการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ ผ่าน machine learning ดังนี้

(1) ส่วนรับข้อมูล เป็นการนำเข้าสู่ข้อมูล เพื่อใช้วินิจฉัยโมเดลที่คลาดเคลื่อนทางคณิตศาสตร์ และรายงานระดับความสามารถทางคณิตศาสตร์ในแต่ละมิติ ประกอบด้วย ผลการกำหนดจุดตัด คะแนนสเกลและคะแนนดิบ ข้อคำถามเพื่อใช้ในการวิเคราะห์โมเดลที่คลาดเคลื่อนของผู้เรียน รูปแบบคำตอบที่ถูกต้องที่เป็นไปได้ทั้งหมดของแบบสอบอัตโนมัติ ข้อมูลป้อนกลับทั้ง 4 รูปแบบในรายข้อกรณีตอบไม่ถูกต้อง คำอธิบายระดับความสามารถในแต่ละมิติ และตัวแปรเพื่อใช้ในการทำนายความสามารถทางคณิตศาสตร์ผ่าน machine learning เช่น เกรดเฉลี่ยวิชาคณิตศาสตร์ การเรียนพิเศษ ชั่วโมงการทบทวนความรู้ด้วยตนเอง และต้องอาศัยข้อมูลของโรงเรียนที่ใช้ทำการทดสอบ ประกอบด้วย รหัสโรงเรียน รายชื่อโรงเรียน รหัสประจำตัว เพื่อให้ผู้เรียน login username และ password ก่อนดำเนินการสอบ และเมื่อผู้เรียนดำเนินการสอบถือเป็นขั้นตอนสำคัญในการแสดงวิธีหาคำตอบ ลงในสถานการณ์ปลายเปิดที่เป็นแบบทดสอบอัตโนมัติที่กำหนดไว้

(2) ส่วนประมวลผล เป็นการตรวจให้คะแนนผู้เรียนรายข้อผ่าน machine learning หากผู้เรียนตอบผิดหรือไม่ตอบ จะมีการให้ข้อมูลป้อนกลับในแต่ละส่วน ตามโมเดลที่คลาดเคลื่อนที่กำหนดไว้ในข้อนั้น ๆ เมื่อผู้เรียนทำแบบทดสอบครบ 10 ข้อ ระบบจะตรวจให้คะแนนประมวลผลความสามารถทางคณิตศาสตร์ของผู้เรียนใน 2 มิติ คือมิติกระบวนการทางคณิตศาสตร์ และมิติโครงสร้างความคิดรวบยอด และรายงานความสามารถในแต่ละมิติ โดยแบ่งระดับความสามารถในแต่ละมิติจากจุดตัดที่กำหนดไว้ และระบบสามารถประมวลผลและทำนายระดับความสามารถทางคณิตศาสตร์ของผู้เรียนก่อนเริ่มทำการทดสอบ โดยใช้ machine learning จากข้อมูลส่วนตัว เช่น เพศ อายุ เกรดเฉลี่ย การเรียนพิเศษ ชั่วโมงการทบทวนความรู้ และอื่นๆ จากสมการทำนาย

(3) ส่วนแสดงผล เป็นการให้ข้อมูลที่ได้รับจากส่วนนำเข้ามาสู่กระบวนการประมวลผล โดยเน้นไปที่กระบวนการงานของโปรแกรม (algorithm) เช่น การแสดงผลการตรวจคำตอบแต่ละข้อ และแสดงข้อมูลป้อนกลับแก่ผู้เรียนกรณีตอบผิด และในขั้นตอนสุดท้ายจะแสดงผลระดับความสามารถทางคณิตศาสตร์ของผู้เรียนโดยระบบจะทำการรวมคะแนนดิบของข้อสอบทุกข้อในแต่ละมิติ ไปเทียบกับคะแนนดิบที่กำหนดไว้ จากนั้นระบบจะนำไปประมวลผลเทียบกับคะแนนสเกล และนำไปสู่การวินิจฉัยระดับความสามารถทางคณิตศาสตร์ ในมิตินั้น ๆ ว่าอยู่ในระดับใด โดยจะแสดงคำอธิบายระดับความสามารถที่กำหนดไว้ในแผนที่โครงสร้าง ซึ่งถือเป็นสิ่งที่ผู้เรียนทำได้ สิ่งที่คุณเรียนควรพัฒนาต่อ

(4) ส่วนให้ข้อมูลป้อนกลับอัตโนมัติ ระบบจะให้ข้อมูลป้อนกลับรายข้อ โดยระบบจะแสดงข้อความคำอธิบายเพิ่มเติมในแต่ละข้อเมื่อผู้เรียนไม่ตอบหรือตอบผิด ตามโมเดลที่คลาดเคลื่อนแต่ละรูปแบบ เพื่อเป็นแนวทางในการตอบคำถามเพื่อพัฒนาความสามารถทางคณิตศาสตร์

(5) ส่วนรายงานผลการประเมิน เป็นการรายงานผลการทดสอบทั้งหมดเพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ เป็นการนำผลจากการประมวลผลมารายงานให้ผู้ใช้งานเข้าใจได้ง่ายมากยิ่งขึ้น เป็นการรายงานผลรายบุคคล (individual report) การรายงานในส่วนนี้จะเป็นการรายงานผลจากการประเมินและจากการวิเคราะห์ของระบบ โดยจะแสดงผลเป็นรายบุคคล สำหรับผู้เรียนในการรายงาน

จะประกอบด้วย 1) ข้อมูลส่วนบุคคล 2) ข้อมูลสารสนเทศและสถิติของการประเมิน 3) ข้อมูลจากการประเมินผลระดับความสามารถทางคณิตศาสตร์

การประเมินคุณภาพของผลการออกแบบระบบในครั้งนี้ได้ผ่านการประเมินโดยผู้เชี่ยวชาญ จำนวน 5 คน ด้วยวิธีการสนทนากลุ่ม (focus group) ครอบคลุม 5 มาตรฐาน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมขอรรถธรรม ด้านความถูกต้อง และด้านความรับผิดชอบการประเมิน โดยพบว่ากระบวนการออกแบบสะดวกต่อการใช้งาน มีความเหมาะสม และง่ายต่อการเรียนรู้ มีความเป็นขั้นตอนในการนำเข้า ทำให้เข้าใจง่าย รวมถึงสามารถนำไปใช้ได้จริง ซึ่งจะเป็นประโยชน์อย่างมากต่อผู้เรียน หากมีการพัฒนาให้ครบวงจรได้ โดยไม่มีกระบวนการเตรียมข้อมูล ระบบจะมีความสมบูรณ์มากยิ่งขึ้น

อภิปรายผล

1. จากผลการกำหนดจุดตัดพบว่า มีความเหมาะสมที่สามารถวินิจฉัยความสามารถทางคณิตศาสตร์ได้ ดังจะเห็นได้จากความสามารถของผู้เรียนในแต่ละมิติ มีการกระจายความสามารถตั้งแต่ -4 ถึง 4 ซึ่งสอดคล้องกับแผนที่โครงสร้างที่ได้กำหนดไว้ เนื่องจากผลการตอบที่ใช้ในการวิเคราะห์ข้อมูล เป็นกลุ่มตัวอย่างที่ครอบคลุมทุกระดับความสามารถ จาก 4 ภูมิภาคของประเทศไทย สอดคล้องกับแนวคิดของ Demars (2010) ที่ให้ความสำคัญกับการกำหนดตัวอย่างและการได้มาของตัวอย่างในการกำหนดจุดตัดที่มีคุณลักษณะของระดับความสามารถที่ครอบคลุม เพียงพอในทุกระดับความสามารถในแต่ละมิติที่ทำการศึกษา มากกว่ามุ่งเน้นการสุ่มตัวอย่างให้มีความเป็นตัวแทนของประชากรเหมือนการออกแบบการวิจัยที่มุ่งเน้นการสรุปอ้างอิงนัยทั่วไปจากกลุ่มผู้สอบไปยังประชากร (Demars, 2010) แต่เมื่อพิจารณาในมิติกระบวนการทางคณิตศาสตร์ มีจุดตัดที่ใกล้เคียงกันอย่างเห็นได้ชัด โดยมีจุดตัดจากระดับที่ 2 ไประดับที่ 3 ระดับที่ 3 ไประดับที่ 4 ซึ่งจุดตัดที่ใกล้กันจะส่งผลต่อเกณฑ์การให้คะแนน โดยการกำหนดค่าอธิบายระดับความสามารถให้มีความแตกต่างและชัดเจนมากขึ้นกว่าเดิม เพื่อให้ข้อสอบที่พัฒนาขึ้นแสดงให้เห็นถึงมวลเนื้อหาหรือสิ่งที่มุ่งวัดในแต่ละชุดครอบคลุมผู้สอบหรือไม่ หากครอบคลุมจะสะท้อนให้เห็นถึงหลักฐานความตรงเชิงเนื้อหาของรายการประเมิน (Wilson, 2005)

2. ผลการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อน ในการแก้โจทย์ปัญหาในสถานการณ์เกี่ยวกับเรื่องสถิติ และความน่าจะเป็น พบว่า ทั้งมิติกระบวนการทางคณิตศาสตร์ (MAP) และมิติโครงสร้างความคิดรวบยอดนักเรียนมีหมโนทัศน์ที่คลาดเคลื่อนมากที่สุด คือ ด้านการใช้ข้อมูล ละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา ทำผิดพลาด หาคำตอบในสิ่งที่ไม่ต้องการ อาจเนื่องมาจากผู้เรียนเลือกใช้หน่วยของตัวแปรผิด หรือการลอกรายละเอียดเกี่ยวกับโจทย์ผิด โดยข้อผิดพลาดอาจเกิดขึ้นตั้งแต่เริ่มต้น หรืออาจเกิดขึ้นในช่วงระหว่าง การดำเนินการกับข้อมูล สอดคล้องกับผลการวิจัยของ Sukkrom (2015) ที่พบว่า ข้อผิดพลาดนักเรียนที่ได้จากการทำแบบวัดหมโนทัศน์และข้อผิดพลาดทางคณิตศาสตร์ แบบอัตนัย นักเรียนมีข้อผิดพลาดมากที่สุด คือ การใช้ข้อมูลผิด ไม่ทำตามที่โจทย์ระบุอย่างชัดเจน ใช้ข้อมูลอื่นที่ไม่สอดคล้องต่อการใช้แก้ปัญหา ซึ่งผลจากการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อน จะใช้เป็นข้อมูลในการสะท้อนหมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนในระบบการตรวจให้คะแนน โดยพิจารณาจากการตอบคำถาม ในข้อคำถามที่ผู้วิจัยได้ออกแบบเพื่อวินิจฉัยหมโนทัศน์

ที่คลาดเคลื่อนไปแล้ว ซึ่งการละเลยการใช้ข้อมูลที่จำเป็นในการแก้ปัญหา จะส่งผลกระทบต่อคะแนนเป็นอย่างมาก เพราะถ้าหากใช้ข้อมูลผิด จะส่งผลต่อการหาคำตอบที่ผิดในขั้นต่อไป

3. การให้ข้อมูลย้อนกลับ ถือเป็นรูปแบบกระบวนการหรือวิธีการตรวจสอบผู้เรียนโดยอาศัยทั้งข้อมูลและสารสนเทศ เพื่อทำการประเมินการเรียนรู้ผู้เรียน ทั้งที่เกิดขึ้นทั้งภายในและภายนอกเป็นรายบุคคล เพื่อที่จะนำไปสู่การปรับเปลี่ยน วางแผนกระบวนการจัดการเรียนรู้ที่เหมาะสมกับผู้เรียนแต่ละบุคคล (Fisher & Frey, 2011) ซึ่งในระบบตรวจให้คะแนนแบบสอบอัตนัยครั้งนี้ เน้นการให้ข้อมูลย้อนกลับใน 4 รูปแบบ คือ (1) แบบชี้แนะบางส่วน (2) ยกตัวอย่าง (3) แบบอธิบายรายละเอียด และ (4) แบบอธิบายรายละเอียดและยกตัวอย่าง ในแต่ละข้ออาจใช้แบบใดแบบหนึ่งหรือหลายแบบบูรณาการรวมกันได้ขึ้นอยู่กับมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนในแต่ละข้อ ซึ่งแต่ละรูปแบบได้ให้ข้อมูลที่จำเป็น บอกรายละเอียดในแก้โจทย์ปัญหา พร้อมยกตัวอย่างประกอบใกล้เคียงกับโจทย์ที่ผู้เรียนกำลังเผชิญอยู่ เพื่อกระตุ้นความคิดของผู้เรียน มากกว่าการบอกว่าถูกหรือผิดเพียงอย่างเดียว จึงเป็นประโยชน์ต่อการนำไปใช้พัฒนาระดับความสามารถทางคณิตศาสตร์ให้สูงขึ้น ซึ่งสอดคล้องกับผลการวิเคราะห์ห่อถักของ Van Der Kleij et al. (2015) ที่พบว่า ข้อมูลย้อนกลับแบบอธิบายรายละเอียดมีขนาดอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนมากกว่าข้อมูลย้อนกลับแบบบอกความถูกต้อง และข้อมูลย้อนกลับแบบบอกผลการกระทำ ซึ่งการให้ข้อมูลย้อนกลับช่วยกระตุ้นให้นักเรียนได้คิดแก้ปัญหา สามารถเชื่อมโยงความรู้เดิมเข้ามาปรับใช้กับสิ่งที่ตนเองยังไม่ทราบหรือยังมีมโนทัศน์ที่ไม่แม่นยำ

4. ผลการออกแบบระบบตรวจให้คะแนนแบบสอบอัตนัย โดยการให้ข้อมูลย้อนกลับอัตโนมัติ เพื่อวินิจฉัยระดับความสามารถทางคณิตศาสตร์ของผู้เรียนผ่านการเรียนรู้ของ machine learning ประกอบด้วย 5 ส่วน คือ (1) ส่วนรับข้อมูล (2) ส่วนประมวลผล (3) ส่วนแสดงผล (4) ส่วนให้ข้อมูลย้อนกลับอัตโนมัติ และ (5) ส่วนรายงานผลการประเมิน ซึ่งทุกส่วนล้วนมีความสำคัญในการตรวจให้คะแนนผู้เรียนให้ถูกต้องและแม่นยำ แต่ควรมุ่งเน้นไปที่ส่วนการรับข้อมูลซึ่งเป็นส่วนแรกของการทำงานของระบบ เช่น หากข้อมูลในส่วนจุดตัดไม่ถูกต้อง ก็จะส่งผลต่อการอธิบายความสามารถของผู้เรียนในแต่ละมิติได้ไม่ถูกต้อง หากคำตอบที่นำเข้ามาไม่ครอบคลุมคำตอบที่เป็นไปได้ทั้งหมดที่ผู้เรียนตอบ ก็จะส่งผลต่อการตรวจคำตอบของระบบและการให้คะแนนไม่ถูกต้อง ทั้งนี้เพื่อให้ได้ระบบการตรวจให้คะแนนอัตโนมัติที่มีมาตรฐาน ควรมีผู้เชี่ยวชาญตรวจสอบผลการออกแบบทุกขั้นตอน ซึ่งสอดคล้องกับงานวิจัยของ (Yadav et al., 2017) ที่พบว่าการปรับปรุงสมรรถนะของระบบหรือแพลตฟอร์มดิจิทัลใดๆ ต้องใช้กระบวนการมาตรฐานและควรมีผู้เชี่ยวชาญตรวจสอบผลการออกแบบและพัฒนารวมถึงการประเมินผลในทุกขั้นตอน โดยจะสามารถสร้างแพลตฟอร์มที่สามารถตอบสนองต่อความต้องการของผู้ใช้และได้มาตรฐานในระดับสากล

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 จากผลการวิเคราะห์จะพบว่าในมิติกระบวนการทางคณิตศาสตร์ มีจุดตัดที่ใกล้เคียงกัน อย่างเห็นได้ชัดจากระดับที่ 2 ไประดับที่ 3 ระดับที่ 3 ไประดับที่ 4 ดังนั้นในการกำหนดคำอธิบายความสามารถ

ในมิติกระบวนการทางคณิตศาสตร์ ควรมีการเขียนเขียนคำอธิบายให้แตกต่างและชัดเจนเพื่อนำไปสู่การตรวจให้คะแนนในระบบได้อย่างถูกต้อง

1.2 ผลการวิเคราะห์หมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนใน 2 มิติพบว่า ทั้ง 2 มิติ ผู้เรียนมีหมโนทัศน์ที่คลาดเคลื่อนมากที่สุดคือ ด้านการใช้ข้อมูล ดังนั้นในการออกแบบระบบการให้ข้อมูลป้อนกลับ ควรอธิบายรายละเอียดข้อมูลที่โจทย์กำหนด เพื่อให้ผู้เรียนใช้ข้อมูลได้อย่างถูกต้อง

1.3 การออกแบบระบบควรมุ่งเน้นไปที่ส่วนการรับข้อมูล ซึ่งเป็นส่วนแรกของการทำงานของระบบ เช่น หากข้อมูลในส่วนจุดตัดไม่ถูกต้อง ก็จะส่งผลต่อการอธิบายความสามารถของผู้เรียนในแต่ละมิติได้ไม่ถูกต้อง หากคำตอบที่นำเข้าไปไม่ครอบคลุมคำตอบที่เป็นไปได้ทั้งหมดที่ผู้เรียนตอบ ก็จะส่งผลต่อการตรวจคำตอบของระบบและการให้คะแนนไม่ถูกต้อง ทั้งนี้เพื่อให้ได้ระบบการตรวจให้คะแนนอัตโนมัติที่มีมาตรฐาน ควรมีผู้เชี่ยวชาญตรวจสอบผลการออกแบบทุกขั้นตอน

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 เนื่องจากการออกแบบระบบตรวจให้คะแนนแบบสอบอัตโนมัติโดยการให้ข้อมูลป้อนกลับผ่าน machine learning มีความจำเป็นต้องสร้างผลการตอบให้สอดคล้องกับการทำงานของโมเดล machine learning ดังนั้นในการวิจัยครั้งต่อไป ควรมีการเปรียบเทียบโมเดลที่เหมาะสมในการตรวจให้คะแนนแบบสอบอัตโนมัติ

2.2 การวิจัยครั้งนี้สามารถตรวจให้คะแนนแบบสอบอัตโนมัติ โดยกำหนดรูปแบบในการตอบตามกรอบที่ผู้วิจัยกำหนดไว้ เพื่อให้วินิจฉัยหมโนทัศน์ที่คลาดเคลื่อนของผู้เรียนได้อย่างถูกต้อง ในการวิจัยครั้งต่อไป ควรพัฒนาการตรวจให้คะแนนแบบสอบอัตโนมัติ โดยการให้ข้อมูลป้อนกลับอัตโนมัติ ผ่าน machine learning ที่ผู้เรียนสามารถเขียนแสดงกระบวนการแก้ปัญหาด้วยตนเองทั้งหมด โดยไม่มีกรอบในการตอบ เพื่อให้ผู้เรียนสามารถแสดงแนวคิดและการแก้ปัญหาได้อย่างหลากหลาย

References

- Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1-23.
- Brookhart, S. M., Moss, C. M., & Long, B. A. (2008). *Teacher inquire into formative assessment practices among Title I reading teachers and literacy coaches*.
http://www.castl.duq.edu/Castl_TechReports.htm
- Demars, C. (2010). *Item Response Theory: Understanding Statistics Measurement*. Oxford University Press.
- Finnegan, D., Kauppinen, A., & Wärensby, A. (2015). *Automated Feedback in a Blended Learning Environment: Student Experience and Development*. In G. Rijlaarsdam (Eds.), M. Deane & T. Guasch. (Vol. Eds.), *Studies in Writing: Vol. x, Learning and teaching writing online: Strategies for success* (pp. 31-45). Brill.

- Fisher, D., Grant, M., Frey, N., & Johnson, C. (2007). Taking formative assessments schoolwide. *Educational Leadership*, 65(4), 64–68.
- Fisher, D., & Frey, N. (2011). *The formative assessment action plan*. ASCD.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81-112.
- Junpeng, P., Krotha, J., Chanayota, K., Tang, K. N., & Wilson, M. (2019). Constructing Progress Maps of Digital Technology for Diagnosing Mathematical Proficiency. *Journal of Education and Learning*, 8(6), 90-102
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine learning: From theory to algorithms*. Cambridge university press
- Stiggins, R. J., & Chappuis, J. (2011). *An Introduction to Student-Involved Assessment FOR Learning* (6th Ed.). Pearson.
- Van Der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, 85(4), 475-511.
- Varasunun, P. (2011). The Program Evaluation Standards (3rd Ed.) in 2010. *Journal of Research Methodology*, 24(2), 273-278.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Routledge.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (2007). *ACER Conquest: Generalized item response modelling software*. ACER Press.
- Yadav, R., Tiruwa, A., & Suri, P. K. 2017. Internet Based Learning (IBL) in Higher Education: A Literature Review. *Journal of International Education in Business*, 10(2), 102-129.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program Evaluation standards: A guide for evaluators and evaluation users* (3rd Ed.). Sage.

Translate Thai References

- Adulpongpaial, S. (2009). *Construction of Mathematics Diagnostic Test on Relations and Functions for MathayomSuksa 5 Students* [Master's thesis]. Chiang Mai University. (in Thai)
- Chinjunthuk, S., & Junpeng, P. (2020). Assessment Guidelines for Student's Personalized Mathematical Proficiency Development. *Journal of Educational Measurement, Mahasarakham University*, 26(1), 47- 64. (in Thai)

- Junpeng, P., Marwiang, M., Chinjunthuk, S., Suwannatrai, P., Krotha, J., Chanayota, K., Tawarungruang, C., Thuanman, J., Tang K. N., & Wilson M. (2020). *Developing Students' Mathematical Proficiency Level Diagnostic Tools through Information Technology in Assessment for Learning Report*. The Thailand Research Fund and Khon Kaen University. (in Thai)
- Junpeng, P. (2018). *Applications of the multidimensional item response theory of research* (4th Ed.). Khon Kaen University Printing. (in Thai)
- Ministry of Education. (2017). *Indicators and subjects of learning the core Learning group Mathematics according to the basic education core curriculum (Revised version 2017)*. Agricultural Co-operative Federation of Thailand Limited. (in Thai)
- Office of the Basic Education Commission. (2020). *Instruction manual learning mathematics (revised version 2020) According to the core curriculum of basic education version 2017*. Agricultural Co-operative Federation of Thailand Limited. (in Thai)
- Sukkrom, P. & Mekanong, A. (2015). A Study of Misconceptions and Mistakes in Mathematics of Eleventh Grade Students in Schools Under the Secondary Educational Service Area Office 1 and 2. *OJED*, 10(4), 599-611. (in Thai)
- Thamcharoen, T. (2019). *Fault Detection and Diagnosis of Machine Downtime in Slider Attachment Process using Machine learning Technique* [Master's thesis]. Suranaree University of Technology. (in Thai)
- The National Institute of Educational Testing Service. (2021). *The result of the national basic educational testing (O-NET) Mathayom 3 Year 2020*. <http://www.niet.or.th> (in Thai)
- Wongwanit, S. (2020). *Design Research in Education*. Chulalongkorn University. (in Thai)