แบบสอบกลางภาค

รายวิชา 2758623 หลักการเรียนรู้ของเครื่องและการประยุกต์

ข้อ 1 วิพากษ์จุดเด่นและข้อสังเกตในงานวิจัย

การกำหนดชื่อภูมินิพนธ์

จุดเด่น

ชื่อเรื่องระบุตัวแปรและองค์ประกอบสำคัญของการวิจัยได้ครบถ้วน ทั้งเทคนิคที่ใช้ (เทคนิคเหมืองข้อมูล) เป้าหมายหลัก (พยากรณ์ผลการเรียน) ตัวแปรต้น (พฤติกรรมการเรียนรู้) และบริบท (ระบบ STOU eLearning ของนักศึกษาระดับปริญญาโท)

ข้อสังเกต

ชื่อเรื่องใช้คำว่า เทคนิคเหมืองข้อมูล แต่การศึกษาจริงเลือกใช้เฉพาะ วิธีต้นไม้ตัดสินใจ ดังนั้นชื่ออาจยังไม่เฉพาะเจาะจงพอเมื่อเทียบกับวิธีวิจัยจริง

ชื่อเรื่องระบุขอบเขตกว้างเหมือนครอบคลุมนักศึกษาระดับปริญญาโทของมหาวิทยาลัยทั้งหมด แต่ข้อมูลจริงมาจากนักศึกษาปริญญาโท สาขาพยาบาลศาสตร์เพียง 1 ชุดวิชา

ข้อเสนอแนะเพื่อการพัฒนา

ควรปรับขอบเขตของชื่อเรื่องให้รัดกุมและสะท้อนบริบทของการศึกษาที่แท้จริง เพื่อลดความเข้าใจผิดของผู้อ่าน เช่น การพยากรณ์ผลการเรียนจากพฤติกรรมการเรียนรู้บนระบบ STOU eLearning โดยใช้เทคนิคต้นไม้ตัดสินใจ กรณีศึกษา: นักศึกษาระดับปริญญาโท สาขาพยาบาลศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช

วัตถุประสงค์ของการวิจัย

จุดเด่น

วัตถุประสงค์ที่เขียนไว้มีความสอดคล้องกับชื่อเรื่อง กรอบแนวคิด และวิธีดำเนินการวิจัย โดยชี้ชัดว่าต้องการใช้เทคนิคเหมืองข้อมูลด้วย วิธีต้นไม้ตัดสินใจเพื่อพยากรณ์ผลการเรียนจากพฤติกรรมการเรียนรู้บน STOU eLearning

ข้อสังเกต

ผู้วิจัยเขียนวัตถุประสงค์ไว้เพียงข้อเดียวแบบรวมๆ แต่เมื่อไปพิจารณาในส่วนของกระบวนการและผลการวิจัยจะพบว่า ผู้วิจัยได้ดำเนินการและรายงานผลลัพธ์หลัก ๆ ออกเป็น 2 ส่วนที่แยกจากกัน

ข้อเสนอแนะเพื่อการพัฒนา

ควรแตกย่อยวัตถุประสงค์หลักออกเป็นข้อย่อย เพื่อให้สอดคล้องกับผลการดำเนินงานจริง

การทบทวนวรรณกรรม

จุดเด่น

ผู้วิจัยมีการทบทวนวรรณกรรมที่ครอบคลุมแนวคิดหลักของงานวิจัย โดยเริ่มจากการอธิบายความหมายและเทคนิคต่าง ๆ ของเหมืองข้อมูล ก่อนจะเจาะจงลงไปที่เทคนิคต้นไม้ตัดสินใจ และเชื่อมโยงไปสู่บริบทของมหาวิทยาลัยสุโขทัยธรรมาธิราช ซึ่งเป็นการศึกษาทางไกล และรองรับนโยบายมหาวิทยาลัย

ข้อสังเกต

ลักษณะการเขียนยังออกไปทางการรวบรวมและบรรยาย มากกว่า การสังเคราะห์และวิพากษ์ และยังให้เพียงหัวข้อของเทคนิคต่าง ๆ แต่ไม่ลงลึก และพบลักษณะการอ้างอิงบางส่วนเป็นแบบอ้างถึงใน

ข้อเสนอแนะเพื่อการพัฒนา

ควรปรับการทบทวนวรรณกรรมจากแบบเล่า ไปเป็นแบบสังเคราะห์มากขึ้น ควรย่อหรือตัดส่วนที่อธิบายเทคนิคเหมืองข้อมูลแบบอื่นๆ ออก แล้วไปอธิบายต้นไม้ตัดสินใจให้มากขึ้น และควรเขียนช่องว่างการวิจัย (research gap) ให้ชัดขึ้น พร้อมทั้งเข้าถึงต้นฉบับแทนการอ้างถึงใน

กรอบแนวคิดของการวิจัย

จุดเด่น

กรอบแนวคิดมีความ เรียบง่าย ชัดเจน และอ่านเข้าใจง่าย และมีความสอดคล้องกับวัตถุประสงค์และวิธีดำเนินการวิจัย

ข้อสังเกต

ยังมีลักษณะเป็นกรอบการดำเนินงาน มากกว่ากรอบแนวคิดเชิงวิชาการ และขาดการแสดงความสัมพันธ์ของตัวแปร

ข้อเสนอแนะเพื่อการพัฒนา

ควรปรับให้แสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามอย่างชัดเจน โดยกำหนดให้พฤติกรรมการเรียนรู้และลักษณะพื้นฐานของผู้เรียนเป็นตัวแปรพยากรณ์ และผลการเรียนเป็นตัวแปรตาม พร้อมทั้งแยกส่วนที่เป็นขั้นตอนการดำเนินการวิจัยออกจากกรอบแนวคิด

วิธีดำเนินการวิจัย

จุดเด่น

ผู้วิจัยใช้ ข้อมูลจริงจาก log files ของระบบการเรียนออนไลน์ ไม่ได้อาศัยเพียงแบบสอบถามหรือการประเมินตนเองของผู้เรียน จึงทำให้ข้อมูลพฤติกรรมการเรียนมีลักษณะเป็นข้อมูลเชิงประจักษ์ สะท้อนการใช้งานจริงของผู้เรียนในระบบ และผู้วิจัยมีการอธิบายกระบวนการจัดการกับข้อมูลก่อนนำไปวิเคราะห์ได้ดี

ข้อสังเกต

ขนาดของกลุ่มตัวอย่างเล็กเกินไปสำหรับงานเหมืองข้อมูล ส่งผลให้เกิดปัญหา Overfitting

ในหัวข้อวิธีดำเนินการวิจัย ผู้วิจัยเกริ่นไว้ว่าแบ่งเป็น 7 ขั้นตอน แต่เมื่อเขียนบรรยายจริงกลับมีเพียง 5 ขั้นตอน นอกจากนี้ในข้อ 1) ยังมีการเรียงลำดับหัวข้อย่อยผิดพลาดเป็น (1), (3), (4), (5) โดยหัวข้อที่ (2) หายไป

ข้อมูลที่นำมาวิเคราะห์มีนักศึกษาสอบผ่านถึง 31 คน และสอบไม่ผ่านเพียง 8 คน แสดงถึงปัญหาความไม่สมดุลของข้อมูล (Class Imbalance) และขั้นตอนการเตรียมข้อมูลและการวิเคราะห์ ผู้วิจัยไม่ได้ใช้วิธีการจัดการกับข้อมูลที่ไม่สมดุลนี้

ข้อเสนอแนะเพื่อการพัฒนา

ควรเพิ่มขนาดกลุ่มตัวอย่างและช่วงเวลาในการเก็บข้อมูล เพื่อให้ได้จำนวน Records ที่เพียงพอกับเทคนิคเหมืองข้อมูล

เพิ่มขั้นตอนการจัดการข้อมูลที่ขาดความสมดุล เพื่อปรับสัดส่วนกลุ่มคนที่สอบผ่านและสอบไม่ผ่านให้ใกล้เคียงกัน ซึ่งจะช่วยเพิ่มค่าความแม่นยำในการตรวจจับ นักศึกษาที่เสี่ยงสอบไม่ผ่าน ได้ดียิ่งขึ้น

ควรตรวจทานการรันหมายเลขหัวข้อและจำนวนขั้นตอนการวิจัยในบทความให้สอดคล้องกันทั้งเล่ม เพื่อความน่าเชื่อถือทางวิชาการ

ผลการวิจัย

จุดเด่น

ผู้วิจัยไม่ได้นำตัวแปรทั้งหมดเข้าสู่โมเดลทันที แต่มีการใช้ Operator Correlation Matrix เพื่อหาค่าสหสัมพันธ์ระหว่างตัวแปรต้น 10 ตัวแปร กับผลการเรียนก่อน ซึ่งช่วยลดความซับซ้อนและเพิ่มประสิทธิภาพของโมเดล

ตัวแบบที่ได้มีความเรียบง่ายและตีความได้ง่าย เพราะต้นไม้ตัดสินใจสรุปออกมาเหลือเพียง 2 ตัวแปรสำคัญ คือ จำนวนครั้งที่เข้าสู่ระบบ (Learn) และ การทำแบบทดสอบก่อน–หลังเรียน (Pre-Post Test) แล้วจำแนกออกเป็น 3 เส้นทางอย่างชัดเจน เช่น หากเข้าเรียนมากกว่า 376 ครั้ง และไม่ทำแบบทดสอบ จะพยากรณ์ว่า สอบไม่ผ่าน ทำให้ผู้สอนหรือผู้ดูแลระบบสามารถนำไปตั้งเงื่อนไขคัดกรองนักศึกษากลุ่มเสี่ยงในระบบได้ทันที

ข้อสังเกต

แม้ความถูกต้องรวมจะดูสูง (ร้อยละ 84.62) แต่ความถูกต้องในการพยากรณ์นักศึกษากลุ่มที่ สอบไม่ผ่านกลับมีเพียงร้อยละ 50.00 ซึ่งในทางปฏิบัติ การตรวจจับนักศึกษาที่จะสอบตกหรือออกกลางคันคือเป้าหมายสำคัญที่สุดของระบบแจ้งเตือนล่วงหน้า

มีความไม่สอดคล้องภายในรายงานผลเอง ในข้อความระบุว่า นักศึกษาทั้งหมดมี 39 คน แต่โปรแกรมพยากรณ์สอบผ่านและไม่ผ่านรวมกันได้เพียง 33 คน ข้อความอธิบายผลส่วนนี้มีแนวโน้ม คำนวณหรืออธิบายผิด และควรแก้ไข

ข้อเสนอแนะเพื่อการพัฒนา

เนื่องจากตัวแบบทำนายกลุ่มที่สอบผ่านได้ดีมาก แต่ทำนายกลุ่มที่ตกได้แย่ สาเหตุหลักมาจากข้อมูลเบื้องต้นมีคนผ่าน 31 คน และตกเพียง 8 คน ในการวิเคราะห์รอบหน้าควรเพิ่มเทคนิคการปรับสมดุลข้อมูล

ควรเพิ่มการรายงานตัวชี้วัดอื่นนอกเหนือจาก accuracy เช่น precision, recall, F1-score หรือ sensitivity สำหรับกลุ่มไม่ผ่าน เนื่องจากผลวิจัยชี้อยู่แล้วว่าโมเดลทำนายกลุ่มสอบไม่ผ่านได้ไม่สูงนัก การมีตัวชี้วัดหลายตัวจะช่วยให้เห็นจุดแข็งและจุดอ่อนของโมเดลได้ชัดกว่าการดู accuracy รวมเพียงค่าเดียว

ข้อเสนอแนะของการวิจัย

จุดเด่น

ผู้วิจัยมีการแบ่งข้อเสนอแนะออกเป็น 2 ส่วนหลักอย่างชัดเจน ได้แก่ ข้อเสนอแนะในการนำผลการวิจัยไปใช้ประโยชน์ และข้อเสนอแนะในการทำวิจัยครั้งต่อไป ทำให้ผู้อ่านสามารถนำไปประยุกต์ใช้ได้ง่าย

ข้อสังเกต

มีลักษณะค่อนข้าง กว้าง และยังไม่ลงสู่ระดับ แนวปฏิบัติที่ชัดเจน มากพอ ยังไม่ตอบชัดว่า ควรปรับอะไรอย่างไร

งานวิจัยนี้เริ่มต้นด้วยตัวแปรพฤติกรรมถึง 8 ตัวแปร แต่เมื่อโมเดลคัดเหลือเพียง 2 ตัวแปร ผู้วิจัยกลับไม่มีข้อเสนอแนะถึงตัวแปรที่เหลือ

ข้อเสนอแนะเพื่อการพัฒนา

ในการเสนอแนะงานวิจัยครั้งต่อไป แทนที่จะกล่าวเพียงให้ ใช้วิธีอื่นๆ ควรระบุเจาะจงไปเลยว่า ควรลองใช้เทคนิคที่จัดการกับ ข้อมูลที่ไม่สมดุล (Imbalanced Data) ได้ดี เช่น Random Forest หรือ Support Vector Machine (SVM)

ควรทำให้ข้อเสนอแนะ โยงตรงกับผลวิจัยหลัก ให้ชัด เช่น งานนี้พบว่าเพียง เข้าใช้งานจำนวนมาก ไม่ได้แปลว่าจะผ่านเสมอไป แต่การเข้าใช้งานร่วมกับการทำ pre-post test จึงสำคัญกว่า