

การจำแนกเว็บไซต์พนันออนไลน์จากหน้าจอด้วย MobileNetV2

Screenshot-Based Classification of Online Gambling Websites with MobileNetV2

อมรเทพ สืบสาย และ ชัยพร เขมะภาคะพันธ์*

สาขาวิชาวิศวกรรมคอมพิวเตอร์ วิทยาลัยวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิต

ถนนประชาชื่น พุ่งสองห้อง เขตหลักสี่ กรุงเทพฯ 10210

*Corresponding Author: chaiyaporn@dpu.ac.th

Received 23 February 2026; Received in revised form 7 April 2026; Accepted 24 April 2026

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินประสิทธิภาพของระบบอัตโนมัติสำหรับการจำแนกเว็บไซต์พนันออนไลน์จากหน้าจอเว็บไซต์ โดยประยุกต์ใช้เทคนิคการประมวลผลภาพร่วมกับการเรียนรู้เชิงลึกผ่านสถาปัตยกรรมโมบายเน็ตวีทู ซึ่งเป็นโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ผ่านการฝึกสอนล่วงหน้าบนชุดข้อมูลอิมเมจเน็ต และนำมาปรับใช้ด้วยเทคนิคการเรียนรู้แบบถ่ายโอนเพื่อเพิ่มประสิทธิภาพกับชุดข้อมูลเฉพาะทาง ชุดข้อมูลที่ใช้ทดลองประกอบด้วยหน้าจอเว็บไซต์ 1,200 ภาพ แบ่งเป็นเว็บไซต์การพนันและเว็บไซต์ทั่วไปในสัดส่วนที่เท่ากัน กระบวนการเตรียมข้อมูลรวมถึงการปรับขนาดภาพมาตรฐานและการทำอนอร์มัลไลเซชัน โดยแบ่งชุดข้อมูลเพื่อฝึกสอน ตรวจสอบ และทดสอบ ในสัดส่วนร้อยละ 70-20-10 และร้อยละ 80-10-10 เพื่อเปรียบเทียบผลลัพธ์เชิงประจักษ์ นอกจากนี้ได้ปรับค่าพารามิเตอร์หลัก ได้แก่ ขนาดแบทช์ 8 และ 16 ร่วมกับจำนวนรอบการฝึกสอนระหว่าง 10-50 รอบ พร้อมประยุกต์ใช้เทคนิคการหยุดก่อนกำหนดเพื่อป้องกันปัญหาการเรียนรู้เกินและเพิ่มเสถียรภาพของแบบจำลอง ทั้งนี้ได้ประเมินประสิทธิภาพผ่านตัวชี้วัดความแม่นยำ ค่าความระลึก ค่าเอฟวัน และค่าพื้นที่ใต้เส้นโค้งลักษณะเฉพาะ เพื่อสะท้อนความสามารถในการจำแนกข้อมูลอย่างครอบคลุม ผลการทดลองพบว่า การแบ่งข้อมูลในสัดส่วน 80-10-10 ร่วมกับ Batch Size เท่ากับ 8 และ Epoch เท่ากับ 50 ให้ประสิทธิภาพสูงสุด โดยมีค่า Accuracy เท่ากับ 0.97, Precision เท่ากับ 0.98, Recall เท่ากับ 0.95, F1-Score เท่ากับ 0.95 และ ROC AUC เท่ากับ 0.9969 ซึ่งอยู่ในระดับยอดเยี่ยม ผลการวิจัยพบว่าโมบายเน็ตวีทูมีประสิทธิภาพสูงในการจำแนกเว็บไซต์การพนันออนไลน์อย่างแม่นยำและมีเสถียรภาพ ด้วยจุดเด่นด้านขนาดแบบจำลองที่เล็กและใช้ทรัพยากรต่ำ จึงเหมาะสมต่อการประยุกต์ใช้ในระบบเฝ้าระวังและสนับสนุนการบังคับใช้กฎหมาย อย่างไรก็ตาม งานวิจัยยังมีข้อจำกัดด้านขนาดชุดข้อมูลและการจำแนกเพียงสองกลุ่ม งานวิจัยในอนาคตจึงควรเพิ่มความหลากหลายของข้อมูล ทดลองเปรียบเทียบกับสถาปัตยกรรมอื่น

คำสำคัญ: โมบายเน็ตV2, การจำแนกหน้าจอเว็บไซต์, เว็บไซต์พนัน, การเรียนรู้เชิงลึก, การถ่ายทอดการเรียนรู้

Abstract

This study aimed to develop and evaluate an automated system for classifying online gambling websites from screenshots by integrating image processing with deep learning via the MobileNetV2 architecture. This convolutional neural network, pre-trained on the ImageNet dataset, was implemented using transfer learning to enhance performance on domain-specific data. The experimental dataset consisted of 1,200 screenshots, equally divided between



gambling and general websites. Data preparation involved standard resizing and normalization, with the dataset partitioned into train-validation-test sets at ratios of 70:20:10 and 80:10:10 for empirical comparison. Key parameters, including batch sizes of 8 and 16 and training epochs ranging from 10 to 50, were adjusted alongside the application of early stopping to prevent overfitting and ensure model stability. Performance was evaluated using accuracy, precision, recall, F1-score, and ROC AUC to reflect comprehensive classification capabilities. The results demonstrated that an 80:10:10 data split combined with a batch size of 8 and 50 epochs yielded the highest performance, achieving an accuracy of 0.97, precision of 0.98, recall of 0.95, F1-score of 0.95, and an excellent ROC AUC of 0.9969. The findings indicated that MobileNetV2 was highly effective in accurately and stably identifying gambling websites. Due to its compact size and low computational requirements, the model is well-suited for real-world monitoring systems and law enforcement support. However, the study was limited by its dataset size and binary classification scope. Future research should increase data diversity, compare alternative architectures, and develop multi-class classification to expand applicability in more complex contexts.

Keywords: MobileNetV2, Website Screenshot Classification, Gambling Websites, Deep Learning, Transfer Learning

บทนำ

การแพร่กระจายของเว็บไซต์พนันออนไลน์ก่อให้เกิดผลกระทบต่อสังคม เศรษฐกิจ และความมั่นคงปลอดภัยไซเบอร์ในวงกว้าง แม้ว่าประเทศไทยจะมีกรอบกฎหมาย เช่น พระราชบัญญัติว่าด้วยการกระทำความผิดเกี่ยวกับคอมพิวเตอร์ พ.ศ. 2550 และพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 ที่เปิดช่องให้สามารถปิดกั้นเว็บไซต์ที่ผิดกฎหมายได้ แต่กระบวนการตรวจสอบแบบใช้คำสำคัญหรือการทำงานด้วยเจ้าหน้าที่เพียงอย่างเดียวยังคงมีข้อจำกัดด้านความเร็วและความแม่นยำ

การเข้าถึงอินเทอร์เน็ตได้อย่างกว้างขวางและสะดวกสบายทำให้พฤติกรรมของผู้บริโภค ในหลายด้านเปลี่ยนแปลงอย่างมีนัยสำคัญหนึ่งในนั้นคือการพนันออนไลน์ (Online Gambling) ซึ่งแพร่ขยายอย่างรวดเร็ว เนื่องจากผู้ใช้งานสามารถเข้าถึงบริการได้ทุกที่ทุกเวลา โดยไม่จำเป็นต้องเดินทางไปยังสถานประกอบการจริง ลักษณะดังกล่าวทำให้การพนันออนไลน์มีความเข้าถึงง่าย (Accessibility) สะดวก (Convenience) และไม่เปิดเผยตัวตน (Anonymity)

ซึ่งดึงดูดกลุ่มผู้ใช้งานได้หลากหลายตั้งแต่วัยรุ่นจนถึงผู้ใหญ่

อย่างไรก็ตามการพนันออนไลน์ก่อให้เกิดผลกระทบเชิงลบต่อหลายมิติ เช่น เศรษฐกิจ: การสูญเสียรายได้ครัวเรือน การหมุนเวียนเงินในระบบใต้ดิน และการสูญเสียรายได้ภาษีของรัฐ และสังคม: การเสพติดการพนัน การเพิ่มขึ้นของหนี้สินและอาชญากรรมทางเศรษฐกิจ สถานการณ์ดังกล่าวยังสะท้อนให้เห็นถึงความสำคัญของ เครื่องมือในการตรวจสอบและเฝ้าระวังเว็บไซต์พนันออนไลน์ ที่ต้องมีประสิทธิภาพมากขึ้น ซึ่งมีกรอบกฎหมายที่เกี่ยวข้องได้แก่ พระราชบัญญัติว่าด้วยการกระทำความผิดเกี่ยวกับคอมพิวเตอร์ พ.ศ. 2550 [1] พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 [2] และพระราชบัญญัติการพนัน พ.ศ. 2478 มาตรา 12 [3] โดยในส่วนของการดำเนินการระงับการแพร่หลายซึ่งข้อมูลความผิดที่ผิดกฎหมาย เช่น การระงับข้อมูลที่ผิดกฎหมายตาม Domain Name/URL นั้น องค์ประกอบของภาพหลักฐานในการประกอบการดำเนินการ เพื่อระงับการแพร่หลายซึ่งข้อมูลดังกล่าว

หลักฐานที่เหมาะสมต้องมีองค์ประกอบดังนี้

1. ต้องแสดงให้เห็น URL (URL bar) อย่างชัดเจน
2. ต้องมีคำหรือรูปภาพที่เข้าข่ายผิดกฎหมาย (เช่น keyword การพนัน)
3. ต้องมี Taskbar หรือ System Clock แสดงวันเดือนปีและเวลา ณ ขณะจับภาพหน้าจอ
4. หน้าจอแสดงผลต้องไม่ถูกปิดเป็นรูป crop จนไม่สามารถยืนยันแหล่งที่มาของข้อมูลได้

ในประเทศไทย แม้ว่าการพนันออนไลน์ยังคงเป็นสิ่งผิดกฎหมายตามพระราชบัญญัติการพนัน พ.ศ. 2478

อย่างไรก็ตาม แม้มีกฎหมายรองรับแต่การปฏิบัติงานในปัจจุบันยังคงเผชิญข้อจำกัดสำคัญจากการพึ่งพาการตรวจสอบด้วยมนุษย์ และการกรองด้วยคำสำคัญ ซึ่งไม่สามารถตรวจจับเว็บไซต์ที่เปลี่ยนชื่อโดเมนได้ทันที และไม่ครอบคลุมการใช้คำค้นหาที่หลากหลายในรูปแบบภาษาไทยและอังกฤษที่มีการปรับเปลี่ยนตลอดเวลา และแม้การวิจัยในอดีตจะนำเทคนิคการประมวลผลภาษามาใช้จำแนกเว็บไซต์จากรหัสต้นฉบับ แต่ยังพบช่องว่างในการตรวจจับอัตลักษณ์ทางสายตาของเว็บไซต์พนันสมัยใหม่ที่เน้นการใช้กราฟิกและแบนเนอร์เป็นหลัก [4] งานวิจัยนี้จึงมุ่งปิดช่องว่างดังกล่าวด้วยการประยุกต์ใช้สถาปัตยกรรมโมบายเน็ตวีทู (MobileNetV2) ซึ่งเป็นโครงข่ายประสาทเทียมแบบคอนโวลูชันที่มีประสิทธิภาพสูงในการสกัดคุณลักษณะเชิงทัศน์ มาจำแนกประเภทเว็บไซต์จากภาพหน้าจอโดยตรง [5] แนวทางนี้ไม่เพียงแต่ช่วยเพิ่มความแม่นยำในการระบุเว็บไซต์ผิดกฎหมายที่ซ่อนเร้นเนื้อหาเชิงข้อความ แต่ยังสนับสนุนการรวบรวมพยานหลักฐานดิจิทัลให้สอดคล้องกับพระราชบัญญัติว่าด้วยการกระทำความผิดเกี่ยวกับคอมพิวเตอร์ พ.ศ. 2550 เพื่อยกระดับการเฝ้าระวังและบังคับใช้กฎหมายได้อย่างมีประสิทธิภาพและยั่งยืน [6]

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาโมเดลการเรียนรู้เชิงลึก (Deep Learning) สำหรับจำแนกเว็บไซต์พนันออนไลน์จากภาพหน้าจอเว็บไซต์

2. เพื่อประเมินและเปรียบเทียบประสิทธิภาพของโมเดล MobileNetV2 ภายใต้พารามิเตอร์การฝึกสอนที่ต่างกัน

วิธีดำเนินงาน

1. การเก็บและจัดเตรียมข้อมูล

ชุดข้อมูล (Dataset) ในการวิจัยครั้งนี้ ผู้วิจัยได้กำหนดกรอบการสร้างชุดข้อมูลอย่างเป็นระบบ โดยเริ่มจากการกำหนด เกณฑ์การติดฉลาก (Labeling Criteria) ล่วงหน้า เพื่อให้การจำแนกประเภทข้อมูลมีความสอดคล้อง (Consistency) และลดความคลาดเคลื่อนระหว่างผู้ประเมิน (Inter-rater Variability) ทั้งนี้ ชุดข้อมูลประกอบด้วยภาพหน้าจอเว็บไซต์จำนวนทั้งสิ้น 1,200 ภาพ แบ่งออกเป็น 2 กลุ่มหลัก ได้แก่ เว็บไซต์พนันออนไลน์ และเว็บไซต์ทั่วไปที่ไม่เกี่ยวข้องกับการพนัน

1.1 กลุ่มภาพหน้าจอเว็บไซต์พนันออนไลน์

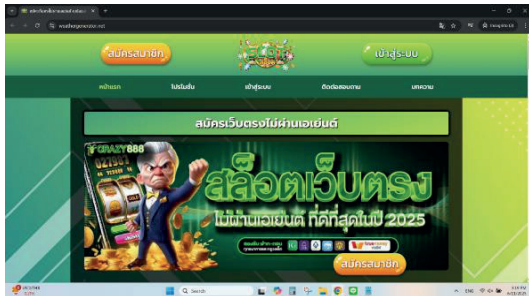
ในกลุ่มข้อมูลนี้ประกอบด้วยภาพหน้าจอเว็บไซต์พนันออนไลน์ จำนวน 600 ภาพ ที่มีวัตถุประสงค์หลัก ในการนำเสนอหรืออำนวยความสะดวกแก่กิจกรรมการพนัน โดยผู้วิจัยกำหนดเกณฑ์การพิจารณาจากองค์ประกอบเชิงโครงสร้างของหน้าเว็บ (Structural Features) และองค์ประกอบเชิงเนื้อหา (Content Features) ดังต่อไปนี้

- 1) การปรากฏของปุ่ม “สมัครสมาชิก”, “เข้าสู่ระบบ”, “เติมเงิน”, “ฝาก – ถอน”
- 2) การแสดงแบนเนอร์ส่งเสริมการเล่นเกมหรือโบนัส
- 3) การแสดงตารางอัตราต่อรอง (Odds Table)
- 4) อินเทอร์เฟซเกมสล็อต หรือห้องคาสิโนสด
- 5) สัญลักษณ์หรือข้อความที่สื่อถึงกิจกรรมการพนันโดยตรง

ภาพหน้าจอเว็บไซต์ที่อยู่ในกลุ่มภาพหน้าจอเว็บไซต์พนันออนไลน์ ซึ่งสะท้อนลักษณะร่วม (Common Visual Patterns) ได้แก่ การใช้สีสดที่ดึงดูดความสนใจ ปุ่มกระตุ้นการสมัครใช้งานขนาดใหญ่ การจัดวางองค์ประกอบที่เน้นกิจกรรมการเติมเงิน และการแสดงภาพกราฟิกของเกมหรือเหรียญรางวัล ดังแสดงในรูปที่ 1 และรูปที่ 2



รูปที่ 1 ตัวอย่างภาพหน้าจอของเว็บไซต์พนัน



รูปที่ 2 ตัวอย่างภาพหน้าจอของเว็บไซต์พนัน

จากรูปที่ 1 และรูปที่ 2 จะเห็นได้ว่าเว็บไซต์มีการออกแบบอินเตอร์เฟซที่เน้นการกระตุ้นการมีส่วนร่วมของผู้ใช้งานผ่านองค์ประกอบภาพ (Visual Stimuli) และข้อความเชิญชวน ซึ่งถือเป็นคุณลักษณะสำคัญที่โมเดลการเรียนรู้เชิงลึกสามารถเรียนรู้ได้ผ่านกระบวนการสกัดคุณลักษณะ (Feature Extraction) ดังนั้น การคัดเลือกภาพตัวอย่างที่มีความหลากหลายในด้าน เช่น โทนสี โครงสร้างหน้าเว็บ รูปแบบการจัดวางปุ่ม ประเภทเกมที่แสดงผล

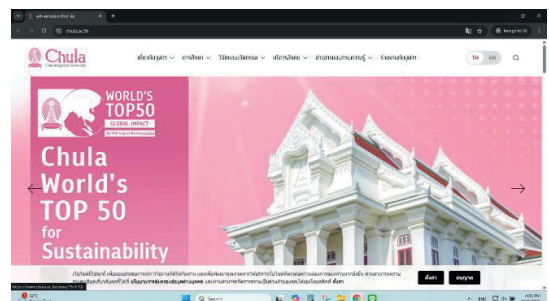
1.2 กลุ่มภาพหน้าจอเว็บไซต์ทั่วไปที่ไม่เกี่ยวข้องกับการพนัน นอกจากกลุ่มเว็บไซต์พนันออนไลน์แล้ว ผู้วิจัยได้จัดเตรียมชุดข้อมูลในกลุ่มเว็บไซต์ทั่วไปที่ไม่เกี่ยวข้องกับการพนัน จำนวน 600 ภาพ เพื่อใช้เป็นกลุ่มเปรียบเทียบ (Negative Class) สำหรับการฝึกและประเมินประสิทธิภาพของโมเดลการจำแนกประเภทภาพ โดยการกำหนดเกณฑ์การคัดเลือกเว็บไซต์ในกลุ่มนี้พิจารณาจากเงื่อนไขหลัก ได้แก่

- 1) ไม่มีเนื้อหาเกี่ยวข้องกับการพนันโดยตรงหรือโดยอ้อม
- 2) ไม่มีปุ่มหรือองค์ประกอบที่สื่อถึงการสมัครสมาชิกเพื่อเติมเงิน

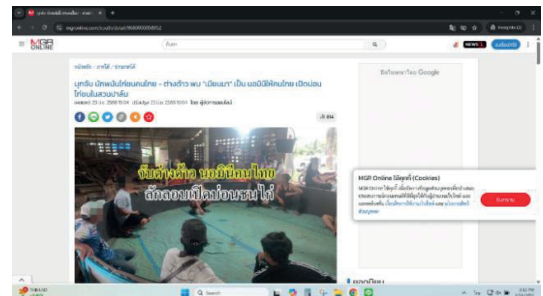
3) ไม่มีตารางอัตราต่อรอง ระบบฝาก-ถอน หรืออินเตอร์เฟซเกมคาสิโน

4) ไม่ปรากฏข้อความส่งเสริมการเล่นเกมที่มิลักษณะเป็นกิจกรรมเสี่ยงโชค

เว็บไซต์ในกลุ่มนี้ครอบคลุมหลายหมวดหมู่ เช่น เว็บไซต์ข่าวสาร เว็บไซต์องค์กรภาครัฐและเอกชน เว็บไซต์พาณิชย์อิเล็กทรอนิกส์ เว็บไซต์การศึกษา และเว็บไซต์บล็อกหรือบทความทั่วไป ทั้งนี้เพื่อให้ข้อมูลมีความหลากหลายเชิงโครงสร้าง (Structural Diversity) และสะท้อนสภาพแวดล้อมจริงของโลกออนไลน์ ดังแสดงในรูปที่ 3 และรูปที่ 4



รูปที่ 3 ตัวอย่างภาพหน้าจอของเว็บไซต์ทั่วไป



รูปที่ 4 ตัวอย่างภาพหน้าจอของเว็บไซต์ทั่วไป

จากภาพตัวอย่างในกลุ่มเว็บไซต์ทั่วไป จะเห็นลักษณะการออกแบบที่แตกต่างจากภาพในกลุ่มเว็บไซต์พนันอย่างชัดเจน กล่าวคือ โครงสร้างหน้าเว็บเน้นการนำเสนอข้อมูลหรือเนื้อหาบทความ การจัดวางองค์ประกอบเป็นไปตามลำดับข้อมูล (Information Hierarchy) ไม่มีองค์ประกอบภาพที่เน้นการกระตุ้นพฤติกรรมเสี่ยงหรือการทำธุรกรรมเดิมพัน และการใช้สีและกราฟิกมีลักษณะเป็นกลางหรือสอดคล้องกับอัตลักษณ์องค์กร ความแตกต่างดังกล่าวเป็นปัจจัยสำคัญที่ช่วยให้โมเดลสามารถเรียนรู้คุณลักษณะเชิงภาพ (Visual Features) ที่



จำแนกความแตกต่างระหว่างเว็บไซต์พนักกับเว็บไซต์ทั่วไปได้อย่างมีประสิทธิภาพ โดยเฉพาะในมิติของรูปแบบปุ่ม (Button Patterns) การจัดวางแบนเนอร์ (Banner Layout) และองค์ประกอบกราฟิกที่สื่อถึงกิจกรรมเกมเติมฟัน

2. ขอบเขตข้อมูล (Data Scope)

ข้อมูลที่ใช้ในงานวิจัยนี้ถูกเก็บรวบรวมจากเว็บไซต์ที่สามารถเข้าถึงได้โดยสาธารณะ (Publicly Accessible Websites) โดยไม่มีการเข้าถึงข้อมูลที่ต้องใช้บัญชีผู้ใช้หรือสิทธิ์เฉพาะ ทั้งนี้เพื่อให้สอดคล้องกับหลักจริยธรรมการวิจัยด้านข้อมูลดิจิทัล และหลีกเลี่ยงการเข้าถึงข้อมูลส่วนบุคคล โดยมีรายละเอียด ดังนี้

2.1 ขอบเขตเชิงโครงสร้างหน้าเว็บ (Page-Level Scope) ผู้วิจัยกำหนดให้เก็บข้อมูลเฉพาะหน้าเว็บหลัก (Homepage หรือ Landing Page) ของแต่ละเว็บไซต์เท่านั้น โดยไม่รวมถึงหน้าลงทะเบียน (Registration Page) หน้าระบบสมาชิก (Member Area) หรือหน้าที่ต้องเข้าสู่ระบบก่อนเข้าถึงเว็บไซต์ ซึ่งเหตุผลในการจำกัดเฉพาะหน้าเว็บหลัก เนื่องจากหน้าแรกเป็นหน้าที่สะท้อนอัตลักษณ์และวัตถุประสงค์หลักของเว็บไซต์ และเป็นหน้าที่ผู้ใช้งานทั่วไปสามารถเข้าถึงได้

2.2 ขอบเขตเชิงเวลา (Temporal Scope) ข้อมูลถูกเก็บในช่วงเวลาปัจจุบันเนื่องจากหน้าเว็บหลักมีการเปลี่ยนแปลงตลอดเวลา ทั้งด้านเนื้อหา โปรโมชัน และการออกแบบอินเทอร์เฟซ ดังนั้นชุดข้อมูลจึงสะท้อนสภาพเว็บไซต์ ณ ช่วงเวลาที่ทำการเก็บข้อมูลเท่านั้น

2.3 ขอบเขตเชิงเนื้อหา (Content Scope) การวิเคราะห์ในงานวิจัยนี้พิจารณาเฉพาะ ลักษณะเชิงภาพ (Visual Features) ของหน้าเว็บไซต์ เช่น การจัดวางองค์ประกอบ (Layout Structure) รูปแบบปุ่มและแบนเนอร์ โทสนีและองค์ประกอบกราฟิก

2.4 ข้อจำกัดของขอบเขตข้อมูล (Scope Limitations) การจำกัดข้อมูลเฉพาะหน้าเว็บหลักอาจส่งผลให้ไม่ครอบคลุมรูปแบบเว็บไซต์ที่ซ่อนเนื้อหาหลักไว้หลังระบบสมัครสมาชิก และไม่สามารถ

ตรวจจับเว็บไซต์ที่ปรับเปลี่ยนหน้าแรกให้มีลักษณะเป็นกลางเพื่อหลีกเลี่ยงการตรวจสอบได้

3. การเตรียมข้อมูล (Data Preprocessing)

ก่อนนำภาพเข้าสู่กระบวนการฝึกสอนโมเดล ผู้วิจัยได้ดำเนินการเตรียมข้อมูลให้สอดคล้องกับข้อกำหนดของสถาปัตยกรรม MobileNetV2 โดยมีขั้นตอน ดังต่อไปนี้

3.1 การปรับขนาดภาพ (Image Resizing) ภาพหน้าจอทั้งหมดถูกปรับขนาดเป็น 224×224 พิกเซล ซึ่งเป็นขนาดอินพุตมาตรฐานของ MobileNetV2 ตามที่เสนอโดย Sandler และคณะ [5] การกำหนดขนาดภาพให้คงที่มีความสำคัญในเชิงโครงสร้าง เนื่องจากโครงข่าย Convolutional Neural Network (CNN) ต้องการมิติข้อมูลที่สม่ำเสมอเพื่อให้สามารถคำนวณ Convolution Operation ได้อย่างถูกต้อง นอกจากนี้ การลดขนาดภาพยังช่วยลดจำนวนพารามิเตอร์ที่ต้องประมวลผล ส่งผลให้ลดภาระด้านหน่วยความจำและเวลาในการฝึกสอน ซึ่งสอดคล้องกับแนวคิดของ MobileNetV2 ที่มุ่งเน้นความมีประสิทธิภาพเชิงคำนวณ (Computational Efficiency) และความเบาของโมเดล

3.2 การปรับค่าความเข้มของพิกเซล (Pixel Normalization) ค่าความเข้มของพิกเซล (0-255) ถูกปรับให้อยู่ในช่วง 0-1 โดยการหารด้วย 255 ทั้งนี้เพื่อให้ข้อมูลมีสเกลที่เหมาะสมต่อการเรียนรู้ของโครงข่ายประสาทเทียมตามแนวคิดของ Goodfellow et al. [7] การทำ Normalization มีบทบาทสำคัญในการ

- 1) ลดความแปรปรวนของข้อมูลนำเข้า
- 2) ป้องกันปัญหา Gradient ที่มีค่าสูงเกินไป (Exploding Gradients)
- 3) ช่วยให้กระบวนการ Gradient Descent มีเสถียรภาพมากขึ้น
- 4) เพิ่มความเร็วในการลู่เข้าสู่ค่าที่เหมาะสม (Faster Convergence)

ดังนั้น การปรับสเกลข้อมูลจึงเป็นขั้นตอนมาตรฐานในงาน Deep Learning และมีผลโดยตรงต่อประสิทธิภาพของโมเดล

3.3 การจัดรูปแบบข้อมูล (Tensor Formatting) ภาพที่ผ่านการปรับขนาดและ Normalization ถูกจัดให้อยู่ในรูปแบบ Tensor ขนาด $224 \times 224 \times 3$ (RGB Channels) เพื่อให้สอดคล้องกับโครงสร้างอินพุตของ MobileNetV2 ซึ่งออกแบบให้รองรับภาพสีแบบสามช่องสัญญาณ

3.4 เหตุผลเชิงสถาปัตยกรรม (Architectural Consistency) MobileNetV2 ใช้โครงสร้าง Inverted Residual Blocks และ Linear Bottlenecks ซึ่งออกแบบมาเพื่อประมวลผลข้อมูลภาพที่ผ่านการปรับสเกลอย่างเหมาะสม หากข้อมูลมีสเกลที่ไม่สมดุล อาจส่งผลให้การเรียนรู้ภายในชั้น Convolution และ Batch Normalization ไม่มีประสิทธิภาพในเชิงทฤษฎี การเตรียมข้อมูลให้มีสเกลที่เหมาะสมจึงสอดคล้องกับหลักการของการเรียนรู้เชิงลึกที่อธิบายโดย Goodfellow et al. [7] ว่าการปรับรูปแบบข้อมูลก่อนการฝึกสอนเป็นองค์ประกอบสำคัญของ Pipeline การเรียนรู้

3.5 ข้อจำกัดของโมเดล MobileNetV2 เนื่องจากข้อจำกัดสถาปัตยกรรมของโมเดล MobileNetV2 ถูกออกแบบให้รับข้อมูลนำเข้าขนาดมาตรฐานรูปภาพขนาด 224×224 พิกเซล ในการประมวลผล ดังนั้นในการวิจัยนี้จึงจำเป็นต้องมีการปรับความละเอียดของรูปภาพ ดังแสดงในรูปที่ 5 ซึ่งทำการปรับลดขนาดรูปภาพต้นฉบับ (Full HD) ตามสัดส่วนก่อนทำการครอบ (Crop) ภาพให้เหลือขนาด 224×224 พิกเซล ในขั้นสุดท้ายก่อนนำเข้าสู่โมเดล



รูปที่ 5 แสดงผลการปรับขนาดภาพของโมเดล MobileNetV2

อย่างไรก็ตาม งานวิจัยนี้มุ่งเน้นการจำแนกประเภทเว็บไซต์จากภาพหน้าจอโดยพิจารณารูปแบบโดยรวมของหน้าเว็บ มากกว่าการวิเคราะห์ข้อความเชิงลึก การปรับขนาดภาพเป็น 224×224 พิกเซล

จึงยังคงรักษาคุณลักษณะที่เพียงพอต่อการเรียนรู้ของโมเดล

4. การแบ่งชุดข้อมูล (Dataset Splitting)

เพื่อให้การประเมินประสิทธิภาพของแบบจำลองมีความน่าเชื่อถือและสะท้อนความสามารถในการจำแนกข้อมูลใหม่ จึงเลือกใช้กระบวนการแบ่งชุดข้อมูลแบบโฮลด์เอาต์ โดยแบ่งข้อมูลออกเป็น 3 ส่วนที่เป็นอิสระต่อกัน ได้แก่ ชุดข้อมูลฝึกสอน (Training Set) ชุดข้อมูลตรวจสอบ (Validation Set) และชุดข้อมูลทดสอบ (Test Set) โดยมีรายละเอียดดังนี้

4.1 การวิเคราะห์ความไวต่อปริมาณข้อมูล โดยกำหนดสัดส่วนการทดลองแบ่งข้อมูลเป็น กรณี 2 โดยมี 10-10-80 และ 10-20-70 ได้แก่ วัตถุประสงค์เพื่อวิเคราะห์ความไวของสถาปัตยกรรม MobileNetV2 ต่อการเปลี่ยนแปลงปริมาณข้อมูลฝึกสอนภายใต้ทรัพยากรข้อมูลที่จำกัด การเปรียบเทียบนี้จะช่วยหาจุดสมดุลระหว่างปริมาณข้อมูลที่ใช้ในการเรียนรู้พารามิเตอร์ และปริมาณข้อมูลที่ใช้ในการตรวจสอบความถูกต้องเพื่อหาค่าพารามิเตอร์ที่เหมาะสมที่สุด

4.2 หน้าที่ของชุดข้อมูลย่อย

1. ชุดข้อมูลฝึกสอน ใช้สำหรับให้แบบจำลองเรียนรู้และสกัดคุณลักษณะเชิงทัศน์ (Visual Features) จากภาพหน้าจอเว็บไซต์
2. ชุดข้อมูลตรวจสอบ ทำหน้าที่เป็นข้อมูลชุดยืนยันในระหว่างการฝึกสอน เพื่อใช้ปรับจูนค่า Batch Size, Epoch และใช้ร่วมกับเทคนิค Early Stopping เพื่อป้องกันปัญหาการเรียนรู้เกินอย่างเป็นระบบ
3. ชุดข้อมูลทดสอบ ทำหน้าที่เป็นข้อมูลที่ไม่เคยผ่านตาแบบจำลองมาก่อน โดยถูกแยกไว้ประเมินประสิทธิภาพเชิงประจักษ์ในขั้นตอนสุดท้ายเพียงครั้งเดียว เพื่อให้ได้ตัวชี้วัดที่สะท้อนการใช้งานจริงอย่างเป็นกลาง

4.3 วิธีการสุ่มแบ่งข้อมูล

เพื่อให้การแบ่งข้อมูลเป็นไปตามหลักการสถิติ ผู้วิจัยได้ประยุกต์ใช้เทคนิคการสุ่มแบบแบ่งชั้นภูมิ (Stratified Sampling) เพื่อรักษาสัดส่วนของเว็บไซต์การพนันและเว็บไซต์ทั่วไปให้มีความสมดุลเท่ากันใน

ทุกชุดข้อมูลย่อย นอกจากนี้ยังกำหนดเงื่อนไขให้ภาพที่มาจากโดเมนเว็บไซต์เดียวกันต้องถูกจัดอยู่ในชุดข้อมูลกลุ่มเดียวกันเท่านั้น เพื่อป้องกันปัญหาข้อมูลรั่วไหล ซึ่งอาจส่งผลให้แบบจำลองจดจำรูปแบบการออกแบบเฉพาะของโดเมนแทนการเรียนรู้คุณลักษณะทั่วไปของคลาส ดังตารางที่ 1

ตารางที่ 1 การแบ่งจำนวนภาพตามคลาสและสัดส่วนการฝึกสอน-การทดสอบประสิทธิภาพ-การทดสอบ (Train-Validation-Test Split)

Split	Class	Train	Validation	Test	Total
70 – 20–10	เว็บไซต์พนัน	420	120	60	600
	เว็บไซต์ทั่วไป	420	120	60	600
	รวม	840	240	120	1,200
	เว็บไซต์พนัน	480	60	60	600
80 – 10–10	เว็บไซต์ทั่วไป	480	60	60	600
	รวม	960	120	120	1,200

5. การฝึกสอนโมเดล (Model Training)

ในการวิจัยนี้ได้ทำการทดลองฝึกสอนโมเดลภายใต้จำนวนรอบการเรียนรู้ (Epochs) ที่แตกต่างกัน ตั้งแต่ 10 รอบ ถึง 50 รอบ โดยเพิ่มครั้งละ 10 รอบ (คือ 10, 20, 30, 40 และ 50 รอบ) เพื่อวิเคราะห์ผลกระทบต่อประสิทธิภาพของโมเดล นอกจากนี้ยังได้ทดสอบขนาดแบตช์ (Batch Size) ที่ 2 ค่า คือ 8 และ 16 เพื่อเปรียบเทียบสมรรถนะในการเรียนรู้ของโมเดล โดยมีการใช้ TensorBoard เพื่อติดตามค่าตัวชี้วัดสำคัญ ได้แก่ Accuracy, Loss และ ROC AUC ตลอดกระบวนการฝึกสอนโมเดล ซึ่งดำเนินการบนชุดข้อมูล Training Set และตรวจสอบความถูกต้องด้วย Validation Set โดยในงานวิจัยนี้กำหนดจำนวนรอบการเรียนรู้สูงสุด (Epochs) ที่ 50 รอบ และใช้กลไก Early Stopping เพื่อป้องกันการเกิด Overfitting โดยเฝ้าติดตามค่า Validation ROC AUC (val_auc)

หากค่าไม่ปรับตัวดีขึ้นภายในจำนวนรอบที่กำหนด โมเดลจะหยุดการเรียนรู้โดยอัตโนมัติและบันทึกพารามิเตอร์ที่ดีที่สุดไว้

5.1 การปรับพารามิเตอร์และกระบวนการเรียนรู้

โมเดลถูกปรับพารามิเตอร์ด้วย Adam Optimizer โดยกำหนดอัตราการเรียนรู้ (learning rate) ที่ 0.001 ซึ่งเป็นค่ามาตรฐานที่เหมาะสมสำหรับการฝึกโมเดล Deep Learning ในงานจำแนกภาพ เนื่องจากสามารถปรับอัตราการเรียนรู้แบบ adaptive ตามค่าเกราดิเอนต์ของแต่ละพารามิเตอร์ได้ และใช้ฟังก์ชันการสูญเสีย (loss function) แบบ Categorical Cross-Entropy ซึ่งแม้งานวิจัยนี้จะมีเพียงสองคลาสแต่ยังถือว่าเหมาะสมสำหรับการจำแนกประเภท (multi-class classification)

5.2 การศึกษาผลกระทบของ Batch Size

เพื่อศึกษาผลกระทบของขนาด batch (Batch Size) ได้เปรียบเทียบประสิทธิภาพของโมเดลเมื่อใช้ batch size ที่แตกต่างกันสองค่า คือ 8 และ 16 โดย batch ขนาดเล็กมีแนวโน้มทำให้การเรียนรู้มีความละเอียดมากขึ้น แต่ต้องใช้เวลาประมวลผลนานกว่า ในขณะที่ batch ขนาดใหญ่ช่วยให้การฝึกเร็วขึ้นแต่มีโอกาสลดทอนความสามารถในการ generalize โมเดล

6. การประเมินผลโมเดล (Model Evaluation)

หลังจากเสร็จสิ้นกระบวนการฝึกสอน โมเดลที่ได้จะถูกนำมาประเมินประสิทธิภาพด้วยชุดข้อมูล Test Set ซึ่งไม่เคยถูกใช้ในการฝึกหรือปรับแต่งพารามิเตอร์มาก่อน เพื่อให้การประเมินสะท้อนความสามารถในการทำนายข้อมูลใหม่ ได้อย่างเป็นกลาง

ในการวิจัยนี้ใช้การประเมินผลเชิงปริมาณ (Quantitative Evaluation) เป็นหลัก โดยพิจารณาตัวชี้วัดหลายรายการ เพื่อให้ครอบคลุมมิติของประสิทธิภาพที่ต่างกันของโมเดลจำแนกประเภทแบบสองคลาส (Binary Classification) โดยมีตัวชี้วัดดังนี้

6.1 Confusion Matrix เป็นตารางสรุปผลการทำนายของโมเดลแบ่งออกเป็น 4 กรณี ได้แก่



True Positive (TP): ทำนายว่าเป็นเว็บไซต์พนัน และเป็นเว็บไซต์พนันจริง

True Negative (TN): ทำนายว่าเป็นเว็บไซต์ทั่วไป และเป็นเว็บไซต์ทั่วไปจริง

False Positive (FP): ทำนายว่าเป็นเว็บไซต์พนัน แต่เป็นเว็บไซต์ทั่วไปจริง

False Negative (FN): ทำนายว่าเป็นเว็บไซต์ทั่วไป แต่เป็นเว็บไซต์พนันจริง

6.2 Accuracy (ความแม่นยำรวม) แสดงสัดส่วนของจำนวนตัวอย่างที่โมเดลทำนายถูกต้องต่อจำนวนตัวอย่างทั้งหมด ดังสมการ

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

6.3 Precision (ความแม่นยำเชิงบวก) วัดความถูกต้องของตัวอย่างที่โมเดลทำนายว่าเป็นเว็บไซต์พนัน ดังสมการ

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

6.4 Recall (การครอบคลุม หรือ Sensitivity) วัดความสามารถของโมเดลในการตรวจจับเว็บไซต์พนันทั้งหมดที่มีอยู่จริง ดังสมการ

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

6.5 F1-Score เป็นค่าเฉลี่ยเชิงฮาร์โมนิกของ Precision และ Recall ซึ่งสะท้อนความสมดุลระหว่างการทำนายถูกและการตรวจจับครอบคลุม

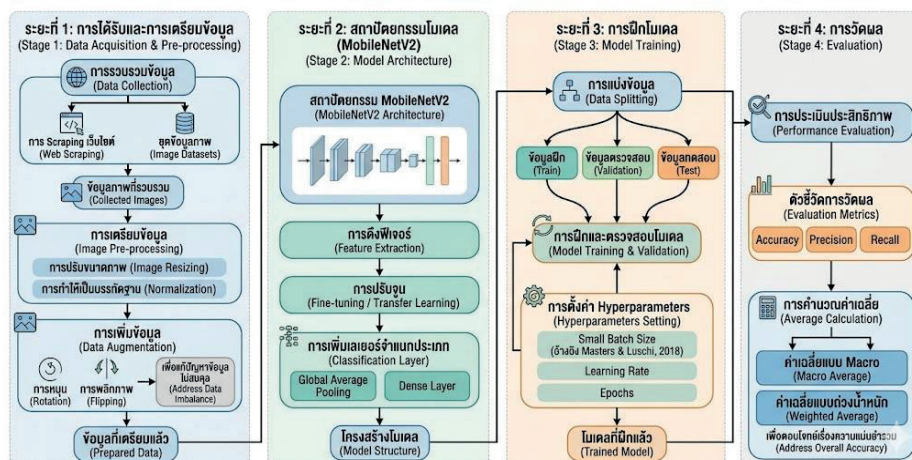
$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

6.6 ROC AUC (Receiver Operating Characteristic – Area Under the Curve) ตัวชี้วัดสมรรถนะของโมเดลที่ใช้ประเมินความสามารถในการจำแนกเว็บไซต์พนันและเว็บไซต์ทั่วไป โดยพิจารณาจากพื้นที่ใต้กราฟ ROC ซึ่งแสดงความสัมพันธ์ระหว่างค่าอัตราการทำนายบวกจริง (True Positive Rate: TPR) และอัตราการทำนายบวกเทียม (False Positive Rate: FPR) ดังสมการ

$$\text{TPR} = \frac{TP}{TP + FN}, \quad \text{FPR} = \frac{FP}{FP + TN} \quad (5)$$

ค่า ROC AUC ที่ใกล้เคียง 1.0 หมายถึงโมเดลสามารถจำแนกได้ดีมาก ขณะที่ค่าที่ใกล้ 0.5 แสดงว่าโมเดลไม่สามารถจำแนกได้ดีกว่าการสุ่ม (random guess) นอกจากนี้ได้สร้าง Confusion Matrix เพื่อแสดงจำนวนการทำนายถูกและผิดในแต่ละคลาสโดยตรง และสร้างรายงานผลในรูปแบบ Classification Report ที่สรุปค่า Precision, Recall และ F1-Score ของแต่ละคลาส พร้อมค่าเฉลี่ย Macro และ Weighted Average [8]

จากวิธีการดำเนินงานที่กล่าวมา สามารถสรุปเป็นแผนภาพขั้นตอนการทำงาน (Research Workflow) ได้ดังรูปที่ 6



รูปที่ 6 แผนภาพขั้นตอนการดำเนินงานวิจัย (Research Workflow)



ผลการทดลองและการอภิปรายผล

บทนี้นำเสนอผลการทดลองของการประยุกต์ใช้โมเดล MobileNetV2 ในการจำแนกเว็บไซต์พนันออนไลน์จากภาพหน้าจอบเว็บไซต์ โดยการทดลองดำเนินการภายใต้ขั้นตอนและการตั้งค่าที่กำหนดไว้ การประเมินประสิทธิภาพของโมเดลครอบคลุมตัวชี้วัดมาตรฐาน ได้แก่ Accuracy, Precision, Recall, F1-score และ ROC AUC เพื่อสะท้อนถึงความถูกต้องโดยรวมและความสามารถในการจำแนกเชิงลึกในระดับคลาส

1. ภาพรวมประสิทธิภาพของโมเดล

ผลการทดลองพบว่าโมเดลสามารถจำแนกเว็บไซต์พนันออนไลน์ได้อย่างมีประสิทธิภาพโดยมีค่า

1.1 Accuracy = 97.0%

1.2 F1-score = 0.97

1.3 Precision และ Recall อยู่ในช่วง 0.95–0.98

1.4 ROC AUC มีค่าใกล้เคียง 0.99
ค่าดังกล่าวสะท้อนว่าโมเดลมีความสามารถในการแยกแยะลักษณะเชิงภาพ (Visual Features) ของเว็บไซต์พนันกับเว็บไซต์ทั่วไปได้อย่างชัดเจน ทั้งในเชิงความถูกต้องโดยรวม (Accuracy) และความสมดุลระหว่าง Precision กับ Recall (F1-score) ค่า F1-score ที่สูงแสดงให้เห็นว่าโมเดลไม่เพียงแต่ทำนายถูกต้องจำนวนมาก แต่ยังสามารถลดทั้ง False Positive และ False Negative ได้อย่างมี

ประสิทธิภาพ ซึ่งมีความสำคัญอย่างยิ่งในบริบทของการตรวจจับเว็บไซต์พนันออนไลน์

2. การเปรียบเทียบกับงานวิจัย

เมื่อเปรียบเทียบกับงานวิจัยที่ใช้ MobileNetV2 ในโดเมนอื่น ๆ ดังแสดงในตารางที่ 2 พบว่าประสิทธิภาพของงานวิจัยนี้อยู่ในระดับใกล้เคียงกับงานก่อนหน้า แม้บริบทข้อมูลจะแตกต่างกันอย่างมีนัยสำคัญ ผลลัพธ์ที่ได้ (97%) อยู่ในระดับใกล้เคียงกับงานด้านภาพธรรมชาติ เช่น งานของ Nair et al. [4] และ Aufar and Kaloka [10] ทั้งที่ข้อมูลเว็บไซต์มีลักษณะซับซ้อนกว่า (มีข้อความ ไอคอน โฆษณา และ

องค์ประกอบกราฟิกหลากหลาย) จึงสามารถสรุปได้ว่าโครงสร้างของ MobileNetV2 ซึ่งใช้แนวคิด Inverted Residuals และ Depthwise Separable Convolution มีความสามารถในการดึงคุณลักษณะเชิงภาพที่มีโครงสร้างซับซ้อนได้ดี แม้โดเมนจะไม่ใช่วัตถุทั่วไป

ตารางที่ 2 เปรียบเทียบผลลัพธ์ MobileNetV2

งานวิจัย	โดเมนข้อมูล	Accuracy	F1-score
Patel et al. [9]	การแพทย์	91.2%	0.91
Aufar and Kaloka [10]	เกษตร	99.9%	0.99
Nair B. J. et al. [4]	ดอกไม้ สมุนไพร	98.3%	0.98
งานวิจัยนี้	เว็บไซต์พนัน	97.0%	0.97

3. ผลกระทบของพารามิเตอร์การฝึกสอน

ในการทดลอง ผู้วิจัยได้ทำการปรับพารามิเตอร์หลักดังนี้

3.1 Epochs: 10, 20, 30, 40, 50

3.2 Batch Size: 8 และ 16

3.3 Data Split: 70–20–10 และ 80–10–10

ผลการทดลองพบว่า

3.4 เมื่อเพิ่มจำนวน Epoch จาก 10 เป็น 30 ค่า Accuracy เพิ่มขึ้นอย่างต่อเนื่อง

3.5 หลัง 40 - 50 Epoch ค่าประสิทธิภาพเริ่มคงที่ (Plateau)

3.6 Batch Size = 8 ให้ค่า Validation Accuracy สูงกว่าเล็กน้อยในบางกรณี โดยจากการทดลองพบว่า Batch Size ขนาดเล็ก (8) ให้ค่าความแม่นยำสูงกว่าขนาดใหญ่ (16) เนื่องจากในทางทฤษฎีการเรียนรู้เชิงลึก Batch Size ขนาดเล็กช่วยให้กระบวนการ Gradient Descent มีความถี่ในการปรับปรุงค่าน้ำหนัก (Weights) บ่อยขึ้นในหนึ่งรอบ



การเรียนรู้ ส่งผลให้โมเดลสามารถเรียนรู้ลักษณะเฉพาะที่ละเอียดอ่อน (Fine-grained features) ของหน้าเว็บไซต์นั้น เช่น รูปแบบของปุ่มขนาดเล็กหรือสัญลักษณ์เฉพาะได้ดีกว่า และยังช่วยให้โมเดลหลีกเลี่ยงการติดอยู่ในจุด Local Minima ที่ไม่เหมาะสม (Poor Local Minima) ได้ดีกว่า Batch Size ขนาดใหญ่ [11]

3.7 Data split แบบ 80-10-10 ให้ผลลัพธ์ที่เสถียรมากกว่าเล็กน้อย เนื่องจากสถาปัตยกรรม MobileNetV2 แม้จะเป็นแบบจำลองขนาดเล็ก แต่ยังคงต้องการปริมาณข้อมูลฝึกสอนที่เพียงพอเพื่อปรับจูนพารามิเตอร์ให้ครอบคลุมความหลากหลายของข้อมูล (Data Diversity) การเพิ่มสัดส่วนข้อมูลฝึกสอนเป็นร้อยละ 80 ช่วยให้โมเดลได้เห็นรูปแบบของหน้าเว็บไซต์ทั่วไปและเว็บไซต์พิเศษที่มีความซับซ้อนของกราฟิกสูงได้หลากหลายขึ้น ส่งผลให้โมเดลมีโอกาสเกิดการเรียนรู้ที่ครอบคลุม (Better

Generalization) และลดความแปรปรวนเมื่อนำไปทดสอบกับข้อมูลที่ไม่เคยเห็น ส่งผลให้ได้ค่าความแม่นยำสูงถึง 0.97 ซึ่งสะท้อนถึงประสิทธิภาพการเรียนรู้ที่เหมาะสมภายใต้ทรัพยากรข้อมูลที่มีอยู่อย่างจำกัด

4. การเปรียบเทียบผลลัพธ์ของโมเดลบนชุด Validation จากผลการทดลองของโมเดล MobileNetV2 ภายใต้เงื่อนไขการแบ่งข้อมูล (Data Splitting) และค่าพารามิเตอร์การฝึกสอนที่แตกต่างกัน พบว่าโมเดลมีประสิทธิภาพอยู่ในระดับสูงอย่างสม่ำเสมอ โดยมีค่า Accuracy อยู่ในช่วง 0.93 – 0.97 ขณะที่ค่า Precision, Recall และ F1-Score อยู่ระหว่าง 0.92 – 0.98 ซึ่งสะท้อนให้เห็นถึงความเสถียรของโมเดลต่อการเปลี่ยนแปลงพารามิเตอร์หลัก ได้แก่ จำนวน Epoch และขนาด Batch Size ดังแสดงในตารางที่ 3

ตารางที่ 3 ผลลัพธ์ของโมเดลบนชุด Validation ตามการแบ่งชุดข้อมูลและขนาดแบตช์

Split	Epoch	Batch Size	Accuracy	Precision	Recall	F1-Score	ROC AUC
70 -20-10	10	8	0.94	0.93	0.95	0.94	0.9689
		16	0.95	0.95	0.95	0.95	0.9706
	20	8	0.96	0.94	0.98	0.96	0.9797
		16	0.95	0.94	0.97	0.95	0.9822
	30	8	0.96	0.95	0.97	0.96	0.9794
		16	0.96	0.94	0.98	0.96	0.9806
	40	8	0.96	0.94	0.98	0.96	0.9811
		16	0.96	0.94	0.98	0.96	0.9831
	50	8	0.96	0.94	0.98	0.96	0.9781
		16	0.96	0.94	0.98	0.96	0.9825
80 -10-10	10	8	0.95	0.98	0.92	0.95	0.9892
		16	0.95	0.98	0.92	0.95	0.9900
	20	8	0.96	0.98	0.93	0.96	0.9958
		16	0.95	0.98	0.92	0.95	0.9961
	30	8	0.96	0.98	0.93	0.96	0.9964
		16	0.95	0.97	0.93	0.95	0.9967
	40	8	0.96	0.95	0.97	0.96	0.9928
		16	0.96	0.98	0.93	0.96	0.9961
	50	8	0.97	0.98	0.95	0.97	0.9969
		16	0.93	0.95	0.92	0.93	0.9942



สรุปผลการทดลอง

จากผลการศึกษาค้นคว้านี้สามารถสรุปข้อเสนอแนะได้สองประการสำคัญ ดังนี้

1. เชิงการประยุกต์ใช้งานจริง

1.1 โมเดลนี้สามารถพัฒนาเป็น เครื่องมือสนับสนุนเจ้าหน้าที่ เพื่อใช้ตรวจสอบเว็บไซต์ต้องสงสัยได้อย่างรวดเร็วและอัตโนมัติ

1.2 ระบบสามารถบูรณาการเข้ากับ หน่วยงานรัฐหรือเอกชน ที่ทำหน้าที่กำกับดูแลการใช้งานอินเทอร์เน็ต เพื่อเพิ่มประสิทธิภาพในการบังคับใช้กฎหมาย

2. ข้อจำกัดของงานวิจัย

2.1 งานวิจัยนี้ มีข้อจำกัดในเรื่องของขนาดชุดข้อมูล โดยชุดข้อมูล 1,200 ภาพ แม้จะเพียงพอต่อการเรียนรู้แบบถ่ายโอน แต่ยังไม่ครอบคลุมความหลากหลายของรูปแบบเว็บไซต์พนันที่มีการปรับเปลี่ยนหน้าตา (UI/UX) อยู่ตลอดเวลาเพื่อหลีกเลี่ยงการตรวจจับ

2.2 งานวิจัยนี้มุ่งเน้นเพียงการแยกเว็บไซต์พนันออกจากเว็บไซต์ทั่วไปเท่านั้น จึงยังไม่สามารถระบุประเภทเฉพาะของกิจกรรมพนัน เช่น คาสโนสดพนันกีฬา หรือสล็อต ซึ่งมีลักษณะกราฟิกที่แตกต่างกันได้

2.3 การพิจารณาเพียงภาพหน้าจอหลัก (Landing Page) ดังนั้นการเก็บข้อมูลจึงจำกัดเฉพาะหน้าแรกของเว็บไซต์ อาจทำให้ไม่สามารถตรวจจับเว็บไซต์ที่ออกแบบหน้าแรกให้ดูเป็นกลาง เพื่อปกปิดเนื้อหาการพนันที่ซ่อนอยู่ภายในระบบสมาชิกได้

3. เชิงการวิจัยในอนาคต

3.1 ขยายชุดข้อมูลให้ใหญ่ และหลากหลายขึ้น ทั้งในด้านภาษา (ไทย/อังกฤษ), การออกแบบเว็บไซต์ และลักษณะกราฟิก

3.2 ทดลองใช้สถาปัตยกรรมใหม่ เช่น EfficientNet, ConvNeXt หรือ Vision Transformer (ViT) เพื่อเปรียบเทียบเชิงประสิทธิภาพกับ MobileNetV2

3.3 พัฒนาโมเดลสำหรับ Multi-class Classification เพื่อตรวจจับเว็บไซต์ผิดกฎหมายประเภทอื่น

3.4 เพิ่มการตรวจสอบชื่อเว็บไซต์ที่เข้าข่ายเว็บไซต์พนันออนไลน์

เอกสารอ้างอิง

[1] พระราชบัญญัติการพนัน พุทธศักราช 2478. (2478, 5 กุมภาพันธ์). ราชกิจจานุเบกษา. เล่ม 52, ตอนที่ 0 ก. หน้า 1980-2003.

[2] พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562. (2562, 27 พฤษภาคม). ราชกิจจานุเบกษา. เล่ม 136, ตอนที่ 69 ก. หน้า 52-95.

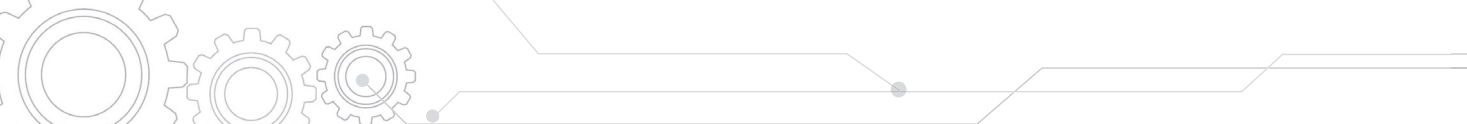
[3] พระราชบัญญัติว่าด้วยการกระทำความผิดเกี่ยวกับคอมพิวเตอร์ พ.ศ. 2550. (2550, 18 มิถุนายน). ราชกิจจานุเบกษา. เล่ม 124, ตอนที่ 27 ก. หน้า 4-13.

[4] Nair, B. J. B., Arjun, B., Abhishek, S., Abhinav, N. M., and Madhavan, V. (2024). Classification of Indian medicinal flowers using MobileNetV2. 2024 11th International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1512-1518.

[5] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linearized bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510-4520.

[6] พระราชบัญญัติการบริหารงานและการให้บริการภาครัฐผ่านระบบดิจิทัล พ.ศ. 2562. (2562, 22 พฤษภาคม). ราชกิจจานุเบกษา. เล่ม 136, ตอนที่ 67 ก. หน้า 40-52.

[7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.



[8] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830.

[9] Patel, S., Kumar, A., & Singh, R. (2020). AI for inclusive education: A case study. *International Journal of Special Education*, vol. 35(3), pp. 10-23.

[10] Aifar, Y., & Kaloka, T. P. (2021). Robusta coffee leaf detection based on YOLOv3-MobileNetV2 model. *International Journal of Electrical and Computer Engineering (IJECE)*. vol. 12(6), pp. 6675-6683.

[11] Masters, D., & Luschi, C. (2018). Revisiting small batch training for deep neural networks. arXiv:1804.07612. pp. 1-18.