**ชื่อเรื่อง** การศึกษาและเปรียบเทียบประสิทธิภาพของอัลกอรึทึมของ Machine learning ในการตรวจจับภาพขนาดเล็กภายใต้กล้องจุลทรรศน์ โดยข้อมูลจำนวนจำกัด

A Comparative Study of Machine Learning Algorithms for Complex Image Classification under Microscope with Limited Data

**ชื่อผู้จัดทำ** นาย ศุภวิชญ์ มารยาท

นาย อชิรวิทย์ ประสม

**ครูที่ปรึกษา** ว่าที่ ร้อยตรี สําธิต ธรรมขันทํา

นาง มนัสชนก ตามวงค์

**ที่ปรึกษาร่วม** -

**สถานศึกษา** โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย เชียงราย

**ปีการศึกษา** 2565

**บทคัดย่อ**

โครงงาน การศึกษาและเปรียบเทียบประสิทธิภาพของอัลกอรึทึมของ Machine learning ในการตรวจจับภาพขนาดเล็กภายใต้กล้องจุลทรรศน์ โดยข้อมูลจำนวนจำกัด ในครั้งนี้มีวัตถุประสงค์เพื่อ ศึกษาอัลกอรึทึม ที่ใช้ใน Machine learning สำหรับการทำ Image Classification โดยมีวิธีดำเนินงาน 5 ขั้นตอน ดังนี้ ขั้นตอนที่ 1 Data preparation ทำการเตรียมพร้อมข้อมูลจาก Data set อย่าง BCCD Data set ขั้นตอนที่ 2 Data augmentation โดยทำการ Data augmentation ประกอบไปด้วย Geometric และ Color space transformations ขั้นตอนที่ 3 Training การทำการดึงข้อมูลออกมาเป็นรูปแบบเวกเตอร์ ทำการแบ่งข้อมูลเป็น Test set และ Train set ใน อัตราส่วน 70 ต่อ 30 โดยการทำ Hyperparameter tunning ร่วมกับอัลกอรึทึมอย่าง Support Vector Machines (SVMs) โดยใช้ร่วมกับ Kernel ต่างๆประกอบไปด้วย Linear, Radial Basis Function (RBF) และ Polynomial (Poly), Naive Bayes โดยใช้ Event Model ที่ประกอบไปด้วย Gaussian naive Baye, Multinomial naive Bayes และ Bernoulli naive Bayes , Decision tree, Random Forest, Ada Boosting, Gradient Boosting, Hist Gradient Boosting, eXtreme Gradient Boosting, light gradient-boosting และ Cat boost และ ขั้นตอนที่ 5 evaluate ทำการ predict ข้อมูล และเทียบข้อมูลกับ metric scoring ผลการศึกษา พบว่า ในขั้นตอน Data augmentation ทำให้เห็นว่าสามารถเพิ่มปริมาณข้อมูลจากข้อมูลจำนวน 352 เป็น 10,028 ภาพ และในขั้นตอนการ Training และ Evaluate พบว่าอัลกอรึทึมที่มีประสิทธิภาพในการทำ Image classification มากที่สุด ตามมาตรา ROC AUC คือ eXtreme gradient boosting ที่ 0.93 ตามมาด้วย light gradient-boosting machine และ gradient boosting ที่ 0.91 และ Cat Boost ที่ 0.87

**คำสำคัญ** : Machine learning, Data augmentation, Hyperparameter tunning