ชื่อโครงงาน การใช้การเรียนรู้ของเครื่องสำหรับการคัดแยกโรคผิวหนัง

ชื่อผู้ทำโครงงาน นายวรันธร จันทร์สว่าง

ชั้น มัธยมศึกษาปีที่ 5 สถาบันการศึกษา โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย ปทุมธานี

ชื่อนักวิทยาศาสตร์พี่เลี้ยง ผู้ช่วยศาสตราจารย์ ดร. วสิศ ลิ้มประเสริฐ และศาสตราจารย์ พญ.อรพรรณ โพชนุกูล

สถาบันนักวิทยาศาสตร์พี่เลี้ยง คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต และคณะแพทยศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต

โครงการพัฒนาอัจฉริยภาพทางวิทยาศาสตร์สำหรับเด็กและเยาวชน

ระดับมัธยมศึกษาตอนปลายและปริญญาตรี รุ่นที่ 22 ประจำปี 2562

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

พ.ศ.2562

**สารบัญ**

หน้า

1. บทคัดย่อ 4
2. บทนำ 5  
   ความเป็นมาและความสำคัญของปัญหา 5  
   วัตถุประสงค์ 5  
   ผลที่คาดว่าจะได้รับ 6
3. เนื้อเรื่อง 7  
   ขั้นตอน 7  
   วิธีการดำเนินงาน 7  
   ผลการดำเนินงาน 22
4. บทสรุป 30  
   สรุปผลการดำเนินงาน 30  
   ปัญหา 32  
   ข้อเสนอแนะ 32  
   แนวทางการดำเนินงานต่อเนื่อง 32
5. เอกสารอ้างอิงและภาคผนวก 33

ชื่อโครงงาน การใช้การเรียนรู้ของเครื่องสำหรับการคัดแยกโรคผิวหนัง

สาขา คอมพิวเตอร์

ผู้ปฏิบัติงาน นายวรันธร จันทร์สว่าง

สถาบันการศึกษา โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย ปทุมธานี

นักวิทยาศาสตร์พี่เลี้ยง ผู้ช่วยศาสตราจารย์ ดร. วสิศ ลิ้มประเสริฐ และศาสตราจารย์ พญ.อรพรรณ โพชนุกูล

สถาบันนักวิทยาศาสตร์พี่เลี้ยง คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต และคณะแพทยศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต

ระยะเวลาการปฏิบัติงาน เมษายน 2562 - มิถุนายน 2563

หน่วยงานสนับสนุน สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

**บทคัดย่อ**

โรคผิวหนังมีทั้งที่เป็นอันตรายและไม่อันตราย บางประเภทไม่จำเป็นต้องไปพบแพทย์เพราะสามารถหายเองได้ ในขณะที่โรคผิวหนังบางชนิดอาจส่งผลอันตรายต่อชีวิตได้ เมื่อพาไปพบแพทย์ ซึ่งการที่ไม่มีวิธีการในการคัดกรองส่งผลให้เกิดความแออัดในโรงพยาบาล รวมถึงทำให้แพทย์และพยาบาลมีปริมาณภาระงานที่มากเกินในแต่ละวัน ดังนั้นการมีซอฟต์แวร์ที่สามารถช่วยในการคัดแยกโรคผิวหนังอาจส่งผลดีในหลาย ๆ กรณี จากข้อมูลเบื้องต้นจึงทำให้สนใจในการพัฒนาโมเดลสำหรับการจำแนกประเภทของโรคผิวหนัง โดยใช้การเรียนรู้ของเครื่อง ซึ่งทางผู้ปฏิบัติงานทำได้ดำเนินงานไป 7 ขั้นตอน จากนั้นนำความรู้ทั้งหมดมาใช้ในการดำเนินงานครั้งที่ 8 ซึ่งคือการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้ Convolutional Neural Network และมีการแบ่งการทดลองออกเป็น 3 อย่าง ได้แก่ Oversampling, Class weights และ Focal loss พบว่าได้โมเดลที่ได้ค่า F1-score สูงสุด 3 อันดับแรก ได้แก่ DenseNet121-Class weights, DenseNet121-Focal loss และ ResNet50-Focal loss ซึ่งมีค่า F1-score 0.84, 0.83 และ 0.81 ตามลำดับ จากนั้นนำ Weights ไป Transfer Learning ให้กับการพัฒนาโมเดลสำหรับคัดการแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 พบว่าได้โมเดลที่ได้ค่า F1-score สูงสุด 3 อันดับแรก ได้แก่ DenseNet121-1, DenseNet121-2 และ ResNet50-1 ซึ่งมีค่า F1-score 0.92, 0.86 และ 0.86 ตามลำดับ จากนั้นนำมา Ensemble แบบใช้การเฉลี่ยเลขคณิต ได้ค่า F1-score 0.89

***คำสำคัญ:*** *โรคผิวหนัง, F1-score, Convolutional Neural Network,* Oversampling, Class weights, Focal loss, *Ensemble*

**บทนำ**

**ความเป็นมาและความสำคัญ**

โรคภูมิแพ้เป็นโรคที่พบได้มากในประเทศไทย โดยพบในเด็กสูงถึงร้อยละ 38 และพบในผู้ใหญ่ประมาณร้อยละ 20 จากการศึกษาข้อมูลเกี่ยวกับโรคภูมิแพ้ในปีพ.ศ. 2559 มีผู้ป่วยโรคภูมิแพ้มากขึ้นถึง 3 - 4 เท่าเมื่อเทียบกับ 10 ปีที่ผ่านมา และมีแนวโน้มที่จะสูงขึ้นเรื่อย ๆ เนื่องจากผู้คนมีรูปแบบการใช้ชีวิตที่เปลี่ยนไป มีการออกกำลังกายน้อยลงส่งผลทำให้ร่างกายอ่อนแอ เกิดการติดเชื้อได้ง่าย (สำนักงานกองทุนสนับสนุนการสร้างเสริมสุขภาพ, 2559) โดยโรคภูมิแพ้อาจแสดงออกในรูปของผื่นบนร่างกาย โดยเฉพาะผื่นภูมิแพ้ผิวหนังที่จัดเป็นผื่นผิวอักเสบที่พบได้ประมาณร้อยละ 10 - 20 ในเด็ก และร้อยละ 1 - 3 ในผู้ใหญ่ (ทีมแพทย์และเภสัชกร HonestDocs, 2562) นับว่าเป็นจำนวนที่มากและไม่ควรละเลย)

โดยทั่วไปพบว่าเมื่อเกิดภาวะผื่นภูมิแพ้ผิวหนังที่เกิดขึ้นบนร่างกาย ก็มักจะรีบไปพบแพทย์ที่โรงพยาบาล นำมาซึ่งปัญหาที่ตามมาคือ ภาวะผู้ป่วยมากเกินความสามารถในการปฏิบัติงานของแพทย์ในแต่ละวัน ส่งผลให้แพทย์และพยาบาลมีภาระงานในปริมาณที่สูงในแต่ละวันและอาจส่งผลต่อการปฏิบัติงานของแพทย์และพยาบาลจากอาการเหนื่อยล้า และจากการศึกษาข้อมูลเกี่ยวกับผื่นภูมิแพ้ในเบื้องต้น พบว่าไม่เป็นอันตราย ดังนั้นหากสามารถคัดกรองผู้ป่วยที่มีผื่นภูมิแพ้ที่ไม่เป็รอันตรายออกไปได้จะสามารถช่วยลดปริมาณของผู้ป่วยในโรงพยาบาลได้ นำมาซึ่งการลดภาระงานของแพทย์และพยาบาลในโรงพยาบาลได้ นอกจากนี้แล้วหากสามารถตรวจเช็คได้ว่าโรคผิวหนังเป็นโรคประเภทที่อันตราย หรือเสี่ยงต่ออันตรายจะช่วยให้ผู้ป่วยสามารถถึงมือแพทย์ได้เร็วยิ่งขึ้น ทำให้ลดภาวะเสี่ยงอันตรายต่อชีวิตของผู้ป่วยได้เป็นอย่างดี

การนำการเรียนรู้ของเครื่องมาใช้กับการวิเคราะห์ลักษณะและจำแนกประเภทของโรคผิวหนังเบื้องต้น เพื่อที่จะทราบถึงผื่นที่เกิดขึ้นบนร่างกาย จะทำให้มีความแม่นยำมากยิ่งขึ้น ทำให้สามารถดูแลตนเองเบื้องต้น จนไปถึงการรักษาต่อไปเป็นไปได้อย่างถูกต้อง ผู้ปฏิบัติงานจะเริ่มจากการพัฒนาเป็นโมเดลเบื้องต้นที่สามารถคัดแยกประเภทของโรคผิวหนังก่อน และเมื่อพัฒนาเป็นโมเดลที่สามารถจำแนกประเภทของโรคผิวหนังเบื้องต้นได้แล้ว สามารถนำโมเดลมาพัฒนาต่อเป็นเว็บแอปพลิเคชันที่สามารถจำแนกประเภทของโรคผิวหนังเบื้องต้นและบอกวิธีการดูแลตนเองเบื้องต้น หรือแนะนำว่าควรไปพบแพทย์ภายใน 24 ชั่วโมง ถ้าโรคผิวหนังที่นำมาวิเคราะห์นั้นไม่ใช่ผื่นภูมิแพ้ผิวหนัง เพื่อให้ผู้คนนั้นสามารถใช้งานอย่างสะดวก และเข้าถึงได้ง่าย

**วัตถุประสงค์**

1. เพื่อศึกษาการใช้การเรียนรู้ของเครื่องในการพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนัง
2. เพื่อศึกษาผลของการตรวจวิเคราะห์ลักษณะและจำแนกประเภทของโรคผิวหนังที่ต่างประเภทกัน
3. เพื่อศึกษาผลของการเปรียบเทียบการใช้โมเดลของการเรียนรู้ของเครื่องที่แตกต่างกันในการคัดแยกประเภทของโรคผิวหนัง

**ผลที่คาดว่าจะได้รับ**

1. ได้เรียนรู้การพัฒนาการโมเดลโดยการใช้การเรียนรู้ของเครื่อง
2. สามาถนำไปเปรียบเทียบกับงานวิจัยด้านการเรียนรู้ของเครื่อง
3. สามารถนำโมเดลที่ได้จากการพัฒนาในโครงงานนี้ไปต่อยอดเป็นเว็บแอปพลิเคชันสำหรับคัดโรคผิวหนังต่อไปเพื่อแก้ไขปัญหาต่าง ๆ ได้แก่ ลดความแออัดในโรงพยาบาล ลดมลพิษจากยานพาหนะในการเดินทางไปโรงพยาบาล และลดภาระของแพทย์และพยาบาล

**เนื้อเรื่อง**

**ขั้นตอนการดำเนินงาน**

1. การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทตัวเลขในชุดข้อมูล MNIST โดยการใช้ Convolutional Neural Network
2. การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูลโรคผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัส โดยการใช้ Convolutional Neural Network ฝึกสร้าง Convolutional Neural Network ตั้งแต่เริ่มต้น โดยใช้ชุดข้อมูล Skin Cancer MNIST: HAM10000
3. การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้ Convolutional Neural Network
4. การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการแบ่งส่วนอวัยวะ
5. การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการแบ่งเป็น 2 โมเดล, การปรับปรุงวิธีการทำ Data Augmentation และใช้การปรับค่า Class weights ของแต่ละประเภท
6. การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการปรับปรุงวิธีการทำ Data Augmentation และใช้การปรับค่า Class weights ของแต่ละประเภท
7. การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทภาพในชุดข้อมูล Cifar-10 โดยการใช้ Convolutional Neural Network
8. การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้ Convolutional Neural Network เพื่อนำ Weights ไป Transfer Learning ให้กับการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403
9. การศึกษาการ Ensemble โมเดล โดยการหาค่าเฉลี่ยเลขคณิตของผลการทำนายโมเดลในแต่ละประเภท

**วิธีการดำเนินงาน**

**การดำเนินงานครั้งที่ 1 การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทตัวเลขในชุดข้อมูล MNIST โดยการใช้ Convolutional Neural Network**

**การศึกษาวิธีการเขียนโปรแกรมภาษาไพธอน**

ฝึกเขียนโปรแกรมภาษาไพธอนใน SOLOLEARN เพื่อให้เข้าใจไวยากรณ์ของภาษา

**การศึกษาวิธีการพัฒนา Convolutional Neural Network**

ศึกษาวิธีการเขียนโปรแกรมสำหรับการพัฒนา Convolutional Neural Network

**การพัฒนาโมเดลสำหรับการคัดแยกประเภทตัวเลขในชุดข้อมูล MNIST โดยการใช้** **Convolutional Neural Network**

เขียนโปรแกรมพัฒนาโมเดล Convolutional Neural Network สำหรับการคัดแยกประเภทตัวเลขในชุดข้อมูล MNIST

**การดำเนินงานครั้งที่ 2 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูลโรคผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัส โดยการใช้** **Convolutional Neural Network**

**การศึกษาลักษณะของผื่น, วิธีการดูแลเบื้องต้นและรวบรวมชุดข้อมูลของโรคผื่นภูมิแพ้ผิวหนังผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัส**

ศึกษาลักษณะและวิธีการดูแลเบื้องต้นของโรคผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัสผู้ที่มีผื่นเกิดขึ้นบนร่างกาย โดยเข้าทำการศึกษาจากแพทย์ผู้เชี่ยวชาญในด้านผื่นภูมิแพ้ (ศ.พญ.อรพรรณ โพชนุกูล คณะแพทยศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต) จากนั้นศึกษาข้อมูลจากสมาคมโรคภูมิแพ้ โรคหืด และวิทยาภูมิคุ้มกันแห่งประเทศ และทำการรวบรวมชุดข้อมูลโรคผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัส

**การแบ่งชุดข้อมูล**

แบ่งชุดข้อมูลโรคผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัสเป็น

* Train set ร้อยละ 80 ของทั้งหมด
* Validation set ร้อยละ 10 ของทั้งหมด
* Test set ร้อยละ 10 ของทั้งหมด

**การศึกษาการทำ Data Augmentation**

ศึกษาวิธีการและวิธีการเขียนโปรแกรมเพื่อเพิ่มความหลากหลายให้กับชุดข้อมูล เช่น การพลิกแนวนอน, การพลิกแนวตั้ง, การหมุนภาพตามองศาต่าง ๆ และการซูมเข้า-ออก

**การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูลโรคผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัส โดยการใช้** **Convolutional Neural Network**

พัฒนา Convolutional Neural Network โดยปรับเปลี่ยนสถาปัตยกรรม Convolutional Neural Network ของการทดลองครั้งที่ 1

**การดำเนินงานครั้งที่ 3 การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้** **Convolutional Neural Network**

**การศึกษาชุดข้อมูล Skin Cancer MNIST: HAM10000**

ศึกษาและทำความเข้าใจกับชุดข้อมูล จากนั้นทำการคัดแยกชุดข้อมูลเป็นโรคแต่ละประเภท ได้แก่ Actinic keratoses and intraepithelial carcinoma / Bowen's disease, Basal cell carcinoma, Benign keratosis-like lesions, Dermatofibroma, Melanoma, Melanocytic nevi และ Vascular lesions  
**การแบ่งชุดข้อมูล**

แบ่งชุดข้อมูล Skin Cancer MNIST: HAM10000 เป็น

* + Train set ร้อยละ 80 ของทั้งหมด
  + Validation set ร้อยละ 10 ของทั้งหมด
  + Test set ร้อยละ 10 ของทั้งหมด

**การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000โดยการใช้** **Convolutional Neural Network**

พัฒนา Convolutional Neural Network โดยปรับเปลี่ยนสถาปัตยกรรม Convolutional Neural Network ของการทดลองครั้งที่ 2

**การศึกษาการใช้ Confusion Matrix ในการประเมินผลโมเดล**

ศึกษาความหมายและวิธีการเขียนโปรแกรมสร้าง Confusion Matrix เพื่อคำนวณค่า Metrics ได้แก่ Accuracy, Precision, Recall และ F1 score

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Model** | AccuracyModel | Precision­Model | RecallModel | F1-scoreModel |

**การดำเนินงานครั้งที่ 4 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการแบ่งส่วนอวัยวะ**

**การศึกษาลักษณะวิธีการดูแลเบื้องต้น, และรวบรวมชุดข้อมูลของโรคผิวหนัง 10 อันดับแรกที่พบมากที่สุดในประเทศไทย**

ศึกษาโรคผิวหนัง 10 อันดับแรกที่พบมากที่สุดในประเทศไทย โดยอ้างอิงจากสถิติของสถาบันโรคผิวหนัง ในปี พ.ศ. 2561 และรวบรวมชุดข้อมูลโดยการสร้างเว็บไซต์โดยใช้ WIX.com พร้อมกับ Google Form สำหรับให้ผู้คนทั่วไปอัปโหลดรูปโรคผิวหนังที่ตัวเองเป็นเพื่อที่จะได้นำข้อมูลไปใช้ในการพัฒนาโมเดลต่อไป และรวบรวมชุดข้อมูลภาพจากอินเทอร์เน็ตที่เป็นข้อมูลเปิด จากนั้นตั้งชื่อชุดข้อมูลที่ได้รวบรวมมานั้นชื่อ PJ61403

**การเตรียมชุดข้อมูล**

ตอนที่ 1 รวบรวมภาพแต่ละประเภทและนำมาคัดแยก ดังนี้  
โรคผิวหนังชนิดที่ไม่เป็นอันตราย ได้แก่ โรคผื่นภูมิแพ้ผิวหนัง

โรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย ได้แก่ โรคสะเก็ดเงิน, โรคต่อมไขมันอักเสบ, โรคด่างขาว และโรคกระเนื้อ

ตอนที่ 2 คัดแยกภาพของโรคผิวหนังแต่ละประเภทเป็นแต่ละอวัยวะ ได้แก่ แขน, ลำตัว, มือ และขา โดยที่แต่ละอวัยวะจะแบ่งเป็นโรคผื่นภูมิแพ้ผิวหนัง และโรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย

**A screenshot of a cell phone

Description automatically generated**

**การแบ่งชุดข้อมูล**

แบ่งชุดข้อมูล PJ61403 ในแต่ละอวัยวะเป็น

* Train set ร้อยละ 80 ของทั้งหมด
* Validation set ร้อยละ 10 ของทั้งหมด
* Test set ร้อยละ 10 ของทั้งหมด

**การทำ Oversampling**

* + การพลิกภาพ ในแนวตั้งฉาก แนวระดับ
  + การหมุนภาพ 90 องศา และ 270 องศา
  + การครอบตัดรูปภาพพื้นที่ 50%

**การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในแต่ละส่วนของอวัยวะในชุดข้อมูล PJ61403**

ตอนที่ 1 พัฒนาโมเดลโดยการใช้ Support Vector Machine

พัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในแต่ละส่วนของอวัยวะในชุดข้อมูล PJ61403 โดยการ Flatten ภาพแล้วใช้ Support Vector Machine ตามโมเดลต่าง ๆ จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดลในแต่ละอวัยวะ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **อวัยวะ** | **F1-scoreLinear** | **F1-scoreSigmoid** | **F1-scorePolynomial** | **F1-scoreRBF** |
| **แขน** | F1-scoreLinear-แขน | F1-scoreSigmoid-แขน | F1-scorePolynomial-แขน | F1-scoreRBF-แขน |
| **ลำตัว** | F1-scoreLinear-ลำตัว | F1-scoreSigmoid-ลำตัว | F1-scorePolynomial-ลำตัว | F1-scoreRBF-ลำตัว |
| **มือ** | F1-score Linear-มือ | F1-scoreSigmoid-มือ | F1-scorePolynomial-มือ | F1-scoreRBF-มือ |
| **ขา** | F1-score Linear-ขา | F1-scoreSigmoid-ขา | F1-scorePolynomial-ขา | F1-scoreRBF-ขา |

ตอนที่ 2 พัฒนาโมเดลโดยการใช้ Convolutional Neural Network

พัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในแต่ละส่วนของอวัยวะในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดลในแต่ละอวัยวะ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **อวัยวะ** | **F1-scoreVGG16** | **F1-scoreResNet50** | **F1-scoreMobileNet** | **F1-scoreDenseNet121** | **F1-scoreNASNetMobile** |
| **แขน** | F1-scoreVGG16-แขน | F1-scoreResNet50-แขน | F1-scoreMobileNet-แขน | F1-scoreDenseNet121-แขน | F1-scoreNASNetMobile-แขน |
| **ลำตัว** | F1-scoreVGG16-ลำตัว | F1-scoreResNet50-ลำตัว | F1-scoreMobileNet-ลำตัว | F1-scoreDenseNet121-ลำตัว | F1-scoreNASNetMobile-ลำตัว |
| **มือ** | F1-scoreVGG16-มือ | F1-scoreResNet50-มือ | F1-scoreMobileNet-มือ | F1-scoreDenseNet121-มือ | F1-scoreNASNetMobile-มือ |
| **ขา** | F1-scoreVGG16-ขา | F1-scoreResNet50-ขา | F1-scoreMobileNet-ขา | F1-scoreDenseNet121-ขา | F1-scoreNASNetMobile-ขา |

**การดำเนินงานครั้งที่ 5 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการแบ่งเป็น 2 โมเดล, การปรับปรุงวิธีการทำ Data Augmentation และใช้การปรับค่า Class weights ของแต่ละประเภท**

**การเตรียมชุดข้อมูล**

ครอบภาพให้เหลือเฉพาะแค่บริเวณที่เป็นโรค จากนั้นแบ่งภาพ ดังนี้

ตอนที่ 1 สำหรับการพัฒนาโมเดลที่ 1 เพื่อแยกโรคผิวหนังชนิดที่ไม่เป็นอันตรายกับโรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย ดังนี้

โรคผิวหนังชนิดที่ไม่เป็นอันตราย ได้แก่ โรคผื่นภูมิแพ้ผิวหนัง

โรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย ได้แก่ โรคสะเก็ดเงิน, โรคต่อมไขมันอักเสบ, โรคด่างขาว, โรคกระเนื้อ

ตอนที่ 2 สำหรับการพัฒนาโมเดลที่ 2 เพื่อแยกโรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย ได้แก่ โรคสะเก็ดเงิน, โรคต่อมไขมันอักเสบ, โรคด่างขาว และโรคกระเนื้อ

**การแบ่งชุดข้อมูล**

แบ่งชุดข้อมูล PJ61403 ทั้ง 2 ตอนเป็น

* Train set ร้อยละ 80 ของทั้งหมด
* Validation set ร้อยละ 10 ของทั้งหมด
* Test set ร้อยละ 10 ของทั้งหมด

**การทำ Data Augmentation**

* + การพลิกภาพ ในแนวตั้งฉาก และแนวระดับ
  + การซูมภาพ 70%, 80%, 90%, 110%, 120% และ 130%
  + การหมุนภาพแบบสุ่มระหว่าง 0 - 360 องศา
  + การเลื่อนพิกเซลความสูงและความกว้าง 30%

โดยที่ทุกขั้นตอนจะใช้การสะท้อนพิกเซลในการเติมพิกเซลที่มีค่าเป็น 0 ในภาพที่ได้จากการทำ Data Augmentation แบบนี้ abcddcba|abcd|dcbaabcd

**การศึกษาการปรับ Class weights**

ศึกษาการคำนวณ Class weights ของแต่ละประเภทและปรับเพื่อให้เกิดความสมดุลของจำนวนภาพในแต่ละประเภท โดยใช้สมการนี้

โดย คือ Class weights ของภาพแต่ละประเภท

คือ จำนวนภาพทั้งหมด

คือ จำนวนประเภท

คือ จำนวนภาพของประเภท

**การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนัง**

ตอนที่ 1 พัฒนาโมเดลที่ 1 สำหรับการคัดแยกโรคผิวหนังชนิดที่ไม่เป็นอันตรายกับโรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย

1.1 พัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 ในตอนที่ 1 โดยการ Flatten ภาพแล้วใช้ ใช้ Support Vector Machine ตามโมเดลต่าง ๆ และการปรับ Class weights ของแต่ละประเภท จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Linear-1** | AccuracyLinear-1 | Precision­Linear-1 | RecallLinear-1 | F1-scoreLinear-1 |
| **Polynomial-1** | AccuracyPoly-1 | PrecisionPoly-1 | RecallPoly-1 | F1-scorePoly-1 |
| **Radial Basis Function-1** | AccuracyRBF-1 | PrecisionRBF-1 | RecallRBF-1 | F1-scoreRBF-1 |
| **Sigmoid-1** | AccuracySigmoid-1 | Precision­Sigmoid-1 | RecallSigmoid-1 | F1-scoreSigmoid-1 |

1.2 พัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 ในตอนที่ 1 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด และการปรับ Class weights ของแต่ละประเภท จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16-1** | AccuracyVGG16-1 | Precision­VGG16-1 | RecallVGG16-1 | F1-scoreVGG16-1 |
| **ResNet50-1** | AccuracyResNet50-1 | Precision­ResNet50-1 | RecallResNet50-1 | F1-scoreResNet50-1 |
| **MobileNet-1** | AccuracyMobileNet-1 | Precision­MobileNet-1 | RecallMobileNet-1 | F1-scoreMobileNet-1 |
| **DenseNet121-1** | AccuracyDenseNet121-1 | Precision­DenseNet121-1 | RecallDenseNet121-1 | F1-scoreDenseNet121-1 |
| **NASNetMobile-1** | AccuracyNASNetMobile-1 | Precision­NASNetMobile-1 | RecallNASNetMobile-1 | F1-scoreNASNetMoible-1 |

ตอนที่ 2 พัฒนาโมเดลที่ 2 สำหรับคัดแยกโรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย

2.1 พัฒนาโมเดลโดยการใช้ Support Vector Machine

พัฒนาโมเดลสำหรับคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 ในตอนที่ 2 โดยการ Flatten ภาพแล้วใช้ Support Vector Machine ตามโมเดลต่าง ๆ และการปรับ Class weights ของแต่ละประเภทจากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Linear-2** | AccuracyLinear-2 | Precision­Linear-2 | RecallLinear-2 | F1-scoreLinear-2 |
| **Polynomial-2** | AccuracyPoly-2 | PrecisionPoly-2 | RecallPoly-2 | F1-scorePoly-2 |
| **Radial Basis Function-2** | AccuracyRBF-2 | PrecisionRBF-2 | RecallRBF-2 | F1-scoreRBF-2 |
| **Sigmoid-2** | AccuracySigmoid-2 | Precision­Sigmoid-2 | RecallSigmoid-2 | F1-scoreSigmoid-2 |

2.2 พัฒนาโมเดลโดยการใช้ Convolutional Neural Network

พัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 ในตอนที่ 2 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด และการปรับ Class weights ของแต่ละประเภท จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16-2** | AccuracyVGG16-2 | Precision­VGG16-2 | RecallVGG16-2 | F1-scoreVGG16-2 |
| **ResNet50-2** | AccuracyResNet50-2 | Precision­ResNet50-2 | RecallResNet50-2 | F1-scoreResNet50-2 |
| **MobileNet-2** | AccuracyMobileNet-2 | Precision­MobileNet-2 | RecallMobileNet-2 | F1-scoreMobileNet-2 |
| **DenseNet121-2** | AccuracyDenseNet121-2 | Precision­DenseNet121-2 | RecallDenseNet121-2 | F1-scoreDenseNet121-2 |
| **NASNetMobile-2** | AccuracyNASNetMobile-2 | Precision­NASNetMobile-2 | RecallNASNetMobile-2 | F1-scoreNASNetMoible-2 |

**การดำเนินงานครั้งที่ 6 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการปรับปรุงวิธีการทำ Data Augmentation และใช้การปรับค่า Class weights ของแต่ละประเภท**

**การเตรียมชุดข้อมูล**

แบ่งภาพชุดข้อมูล PJ61403 เป็นโฟลเดอร์ ได้แก่ โรคผื่นภูมิแพ้ผิวหนัง, โรคสะเก็ดเงิน, โรคต่อมไขมัน

อักเสบ, โรคด่างขาว และโรคกระเนื้อ

**A picture containing screenshot

Description automatically generated**

**การแบ่งชุดข้อมูล**

แบ่งชุดข้อมูล PJ61403 เป็น

* Train set ร้อยละ 80 ของทั้งหมด
* Validation set ร้อยละ 10 ของทั้งหมด
* Test set ร้อยละ 10 ของทั้งหมด

**การทำ Data Augmentation**

* + การพลิกภาพ ในแนวตั้งฉาก และแนวระดับ
  + การซูมภาพ 60%, 70%, 80%, 90%, 110%, 120% และ 130%
  + การหมุนภาพแบบสุ่มระหว่าง 0 - 360 องศา
  + การเลื่อนพิกเซลความสูงและความกว้าง 30%

โดยที่ทุกขั้นตอนจะใช้การสะท้อนพิกเซลในการเติมพิกเซลที่มีค่าเป็น 0 ในภาพที่ได้จากการทำ Data Augmentation แบบนี้ abcddcba|abcd|dcbaabcd

**การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนัง**

ตอนที่ 1 พัฒนาโมเดลโดยการใช้ Support Vector Machine

พัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการ Flatten ภาพแล้วใช้ Support Vector Machine ตามโมเดลต่าง ๆ และการปรับ Class weights ของแต่ละประเภท จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Linear** | AccuracyLinear | Precision­Linear | RecallLinear | F1-scoreLinear |
| **Polynomial** | AccuracyPoly | PrecisionPoly | RecallPoly | F1-scorePoly |
| **Radial Basis Function** | AccuracyRBF | PrecisionRBF | RecallRBF | F1-scoreRBF |
| **Sigmoid** | AccuracySigmoid | Precision­Sigmoid | RecallSigmoid | F1-scoreSigmoid |

ตอนที่ 2 พัฒนาโมเดลโดยการใช้ Convolutional Neural Network

พัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด และการปรับ Class weights ของแต่ละประเภท จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **MobileNet** | AccuracyMobileNet | Precision­MobileNet | RecallMobileNet | F1-scoreMobileNet |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |
| **NASNetMobile** | AccuracyNASNetMobile | Precision­NASNetMobile | RecallNASNetMobile | F1-scoreNASNetMoible |

**การดำเนินงานครั้งที่ 7 การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทภาพในชุดข้อมูล Cifar-10** **โดยการใช้ Convolutional Neural Network**

**การแบ่งชุดข้อมูล**

แบ่งชุดข้อมูล Cifar-10 เป็น

* Train set ร้อยละ 70 ของทั้งหมด
* Validation set ร้อยละ 20 ของทั้งหมด
* Test set ร้อยละ 10 ของทั้งหมด

**การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทภาพในชุดข้อมูล Cifar-10 โดยการใช้** **Convolutional Neural Network**

ฝึกการพัฒนา Convolutional Neural Network โดยการใช้ Transfer Learning และการ Fine Tuning ได้แก่ Hyperparameter tuning, Regularization และ Optimization

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Model** | AccuracyModel | Precision­Model | RecallModel | F1-scoreModel |

**การดำเนินงานครั้งที่ 8 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้** **Convolutional Neural Network เพื่อนำ Weights ไป Transfer Learning ให้กับการพัฒนาโมเดลสำหรับคัดการแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403**

**การเตรียมชุดข้อมูล**

มีการเปลี่ยนแปลงประเภทในชุดข้อมูล PJ61403 โดยการเพิ่มประเภทผิวหนังปกติ และตัดโรคด่างขาวออกไปเนื่องจากมีจำนวนที่น้อย จากนั้นครอปรูปให้เหลือบริเวณที่เป็นโรค

![A screenshot of a cell phone

Description automatically generated]()

**การแบ่งชุดข้อมูล**

แบ่งชุดข้อมูล Skin Cancer MNIST: HAM10000 เป็น

* Train set ร้อยละ 63 ของทั้งหมด
* Validation set ร้อยละ 27 ของทั้งหมด
* Test set ร้อยละ 10 ของทั้งหมด

**การทำ Data Augmentation**

* + การพลิกภาพ ในแนวตั้งฉาก และแนวระดับ
  + การซูมภาพ 90% และ 110%
  + การหมุนภาพแบบสุ่มระหว่าง 0 - 360 องศา
  + การเลื่อนพิกเซลความสูงและความกว้าง 10%

โดยที่ทุกขั้นตอนจะใช้การสะท้อนพิกเซลในการเติมพิกเซลที่มีค่าเป็น 0 ในภาพที่ได้จากการทำ Data Augmentation แบบนี้ abcddcba|abcd|dcbaabcd

**การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนัง**

ตอนที่ 1 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการเทรนใหม่ทั้งหมด จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

1.1 ใช้การ Oversampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

1.2 ใช้การปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

1.3 ใช้ Focal loss

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

ตอนที่ 2 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการ Transfer Learning จาก Pre-Trained Weights ของ ImageNet และการ Fine Tuning จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

2.1 ใช้การ Oversampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

2.2 ใช้การปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

2.3 ใช้ Focal loss

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

ตอนที่ 2.5 ทำการตัดโรคผิวหนังประเภท Melanocytic nevi ในชุดข้อมูล Skin Cancer MNIST: HAM10000 ออก ทำให้เหลือโรคผิวหนัง 6 ประเภท เพื่อช่วยแก้ปัญหาในเรื่องความไม่สมดุลของชุดข้อมูล จากนั้นพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการเทรนใหม่ทั้งหมด จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดล

2.5.1 ใช้การ Oversampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

2.5.2 ใช้การปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

2.5.3 ใช้ Focal loss

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | AccuracyVGG16 | Precision­VGG16 | RecallVGG16 | F1-scoreVGG16 |
| **ResNet50** | AccuracyResNet50 | Precision­ResNet50 | RecallResNet50 | F1-scoreResNet50 |
| **DenseNet121** | AccuracyDenseNet121 | Precision­DenseNet121 | RecallDenseNet121 | F1-scoreDenseNet121 |

จากนั้นทำการเลือกโมเดลที่มีค่า F1-score สูงที่สุดมา 3 โมเดลแรกในแต่ละตอน แล้วพิจารณาว่าโมเดลนั้นมีการเทคำตอบหรือไม่ โดยการคำนวณจากการใช้ส่วนเบี่ยงเบนเฉลี่ยของโมเดลตามสูตรนี้

โดย คือ ส่วนเบี่ยงเบนเฉลี่ย

คือ ค่าเฉลี่ยเลขคณิต

คือ คะแนนของข้อมูลแต่ละตัว

คือ จำนวนข้อมูล

*นำค่า* F1-score *ของแต่โมเดลมาลบและบวกด้วยค่า*ส่วนเบี่ยงเบนเฉลี่ย จะได้

ค่าต่ำสุด = F1-score – M.D.

ค่าสูงสุด = F1-score + M.D.

พิจารณาว่าในแต่ละโมเดลจำนวนประเภทที่มีค่า F1-score เกินค่าต่ำสุดกับค่าสูงสุดเป็นจำนวนเท่าไหร่ แล้วทำการเลือกโมเดลที่มีการเทคำตอบของแต่ละประเภทน้อยที่สุดเพื่อนำไป Transfer Learning ต่อในตอนที่ 4

|  |  |  |
| --- | --- | --- |
| **Model** | **F1-score** | **จำนวนประเภทที่มีการเทคำตอบ** |
| Best-Model-1 | F1-scoreBest-Model-1 | Best-Model-1n |
| Best-Model-2 | F1-scoreBest-Model-2 | Best-Model-2n |
| Best-Model-3 | F1-scoreBest-Model-3 | Best-Model-3n |

แบ่งชุดข้อมูล PJ61403 เป็น

* Train set ร้อยละ 50 ของทั้งหมด
* Validation set ร้อยละ 25 ของทั้งหมด
* Test set ร้อยละ 25 ของทั้งหมด

การทำ Data Augmentation

* + การพลิกภาพ ในแนวตั้งฉาก และแนวระดับ
  + การซูมภาพ 90% และ 110%
  + การหมุนภาพแบบสุ่มระหว่าง 0 - 360 องศา
  + การเลื่อนพิกเซลความสูงและความกว้าง 10%

โดยที่ทุกขั้นตอนจะใช้การสะท้อนพิกเซลในการเติมพิกเซลที่มีค่าเป็น 0 ในภาพที่ได้จากการทำ Data Augmentation แบบนี้ abcddcba|abcd|dcbaabcd

ตอนที่ 4 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการ Transfer Learning Weights จากการทดลองตอนที่ 3 จากนั้นทำการ Fine Tuning และเปรียบเทียบค่า F1-score ของแต่ละโมเดล

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Model-X1** | AccuracyModel-X1 | Precision­Model-X1 | RecallModel-X1 | F1-scoreModel-X1 |
| **Model-X2** | AccuracyModel-X2 | PrecisionModel-X2 | RecallModel-X2 | F1-scoreModel-X2 |
| **Model-X3** | AccuracyModel-X3 | Precision­Model-X3 | RecallModel-X3 | F1-scoresModel-X3 |

**การดำเนินงานครั้งที่ 9** **การศึกษาการ Ensemble โมเดล โดยการหาค่าเฉลี่ยเลขคณิตของผลการทำนายโมเดลในแต่ละประเภท**

**การศึกษาการ Ensemble โมเดล**

ศึกษาหลักการทำงานและวิธีการเขียนโปรแกรมการ Ensemble โมเดล โดยการหาค่าเฉลี่ยเลขคณิตของผลการทำนายโมเดลในแต่ละประเภทที่จะได้เป็นความน่าจะเป็น จากนั้นนำโมเดล Model-X1, Model-X2, และ Model-X3 มาทำการ Ensembel ตามที่ได้ศึกษาไว้

โดย = ค่าเฉลี่ยเลขคณิตของผลทำนาย

= ผลการทำนายของโมเดล

= จำนวนโมเดล

**ผลการดำเนินงาน**

**ผลการดำเนินงานครั้งที่ 1 การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทตัวเลขในชุดข้อมูล MNIST โดยการใช้ Convolutional Neural Network**

ทดสอบด้วย Test set ได้ค่า accuracy 0.99

**ผลการดำเนินงานครั้งที่ 2 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูลโรคผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัส โดยการใช้** **Convolutional Neural Network**

ทดสอบด้วย Test set ได้ค่า accuracy 0.50

**ผลการดำเนินงานครั้งที่ 3 การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้** **Convolutional Neural Network**

กราฟแสดงจำนวนภาพโรคผิวหนังแต่ละประเภทในชุดข้อมูล Skin Cancer MNIST: HAM10000

**A screenshot of a cell phone

Description automatically generated**

*3.1 กราฟแสดงชุดข้อมูลแต่ละโรคผิวหนัง*

จากกราฟ 3.1 พบว่าจำนวนชุดข้อมูลโรคผิวหนังไม่มีความสมดุลและมีจำนวนที่ไม่มาก

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Model** | 0.72 | 0.61 | 0.32 | 0.36 |

**ผลการดำเนินงานครั้งที่ 4 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการแบ่งส่วนอวัยวะ**

ตอนที่ 1 ตารางการเปรียบเทียบค่า F1-score ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในแต่ละส่วนของอวัยวะในชุดข้อมูล PJ61403 ตามโมเดลต่าง ๆ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **อวัยวะ** | **F1-scoreLinear** | **F1-scoreSigmoid** | **F1-scorePolynomial** | **F1-scoreRBF** |
| **แขน** | 0.65 | 0.47 | 0.66 | 0.65 |
| **ลำตัว** | 0.48 | 0.26 | 0.26 | 0.29 |
| **มือ** | 0.65 | 0.26 | 0.38 | 0.45 |
| **ขา** | 0.58 | 0.26 | 0.26 | 0.26 |

ตอนที่ 2 ตารางการเปรียบเทียบค่า F1-score ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในแต่ละส่วนของอวัยวะในชุดข้อมูล PJ61403 โดยใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **อวัยวะ** | **F1-scoreVGG16** | **F1-scoreResNet50** | **F1-scoreMobileNet** | **F1-scoreDenseNet121** | **F1-scoreNASNetMobile** |
| **แขน** | 0.33 | 0.54 | 0.33 | 0.84 | 0.62 |
| **ลำตัว** | 0.33 | 0.42 | 0.33 | 0.77 | 0.83 |
| **มือ** | 0.33 | 0.75 | 0.33 | 0.94 | 0.90 |
| **ขา** | 0.33 | 0.53 | 0.33 | 0.83 | 0.69 |

**ผลการดำเนินงานครั้งที่ 5 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการแบ่งเป็น 2 โมเดล, การปรับปรุงวิธีการทำ Data Augmentation ใช้การปรับค่า Class weights** **ของแต่ละประเภท**

ตอนที่ 1 พัฒนาโมเดลที่ 1 สำหรับการคัดแยกโรคผิวหนังชนิดที่ไม่เป็นอันตรายกับโรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย

1.1 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Support Vector Machine ตามโมเดลต่าง ๆ และการปรับ Class weights ของแต่ละประเภท

|  |  |  |  |
| --- | --- | --- | --- |
| **Kernel** | **Precision** | **Recall** | **F1-score** |
| **Linear** | 0.65 | 0.67 | 0.65 |
| **Sigmoid** | 0.47 | 0.69 | 0.56 |
| **Polynomial** | 0.47 | 0.69 | 0.56 |
| **RBF** | 0.47 | 0.69 | 0.56 |

1.2 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด และการปรับ Class weights ของแต่ละประเภท

|  |  |  |  |
| --- | --- | --- | --- |
| **Architecture** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.09 | 0.29 | 0.13 |
| **ResNet50** | 0.50 | 0.50 | 0.50 |
| **MobileNet** | 0.81 | 0.75 | 0.76 |
| **DenseNet121** | 0.66 | 0.71 | 0.65 |
| **NASNetMobile** | 0.88 | 0.88 | 0.88 |

ตอนที่ 2 พัฒนาโมเดลที่ 2 เพื่อแยกโรคผิวหนังชนิดอื่น ๆ ที่อาจเป็นอันตราย

2.1 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Support Vector Machine ตามโมเดลต่าง ๆ และการปรับ Class weights ของแต่ละประเภท

|  |  |  |  |
| --- | --- | --- | --- |
| **Kernel** | **Precision** | **Recall** | **F1-score** |
| **Linear** | 0.62 | 0.69 | 0.65 |
| **Sigmoid** | 0.19 | 0.44 | 0.27 |
| **Polynomial** | 0.19 | 0.44 | 0.27 |
| **RBF** | 0.19 | 0.44 | 0.27 |

2.2 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด และการปรับ Class weights ของแต่ละประเภท

|  |  |  |  |
| --- | --- | --- | --- |
| **Architecture** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.22 | 0.47 | 0.30 |
| **ResNet50** | 0.63 | 0.42 | 0.31 |
| **MobileNet** | 0.67 | 0.74 | 0.68 |
| **DenseNet121** | 0.78 | 0.63 | 0.60 |
| **NASNetMobile** | 0.91 | 0.89 | 0.87 |

**ผลการดำเนินงานครั้งที่ 6 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการปรับปรุงวิธีการทำ Data Augmentation และใช้การปรับค่า Class weights ของแต่ละประเภท**

ตอนที่ 1 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการ Flatten ภาพแล้วใช้ Support Vector Machine ตามโมเดลต่าง ๆ และการปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| Linear | 0.60 | 0.60 | 0.57 | 0.56 |
| Polynomial | 0.60 | 0.58 | 0.54 | 0.54 |
| RBF | 0.47 | 0.44 | 0.39 | 0.38 |
| Sigmoid | 0.38 | 0.09 | 0.25 | 0.14 |

ตอนที่ 2 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด และการปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| VGG16 | 0.27 | 0.07 | 0.25 | 0.11 |
| ResNet50 | 0.27 | 0.07 | 0.25 | 0.11 |
| MobileNet | 0.38 | 0.34 | 0.29 | 0.19 |
| DenseNet121 | 0.38 | 0.70 | 0.40 | 0.35 |
| NASNetMobile | 0.73 | 0.83 | 0.71 | 0.74 |

**ผลการดำเนินงานครั้งที่ 7 การศึกษาการพัฒนาโมเดลสำหรับการคัดแยกประเภทภาพในชุดข้อมูล Cifar-10** **โดยการใช้ Convolutional Neural Network**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Model** | 0.86 | 0.86 | 0.86 | 0.86 |

**ผลการดำเนินงานครั้งที่ 8 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้** **Convolutional Neural Network เพื่อนำ Weights ไป Transfer Learning ให้กับการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403**

**การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนัง**

ตอนที่ 1 ตารางการเปรียบเทียบค่า Metrics ของการการพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการเทรนใหม่ทั้งหมด

1.1 ใช้การ Oversampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.82 | 0.70 | 0.70 | 0.69 |
| **ResNet50** | 0.79 | 0.70 | 0.73 | 0.69 |
| **DenseNet121** | 0.85 | 0.80 | 0.74 | 0.75 |

1.2 ใช้การปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.76 | 0.62 | 0.71 | 0.65 |
| **ResNet50** | 0.81 | 0.67 | 0.74 | 0.70 |
| **DenseNet121** | 0.82 | 0.66 | 0.75 | 0.69 |

1.3 ใช้ Focal loss

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.83 | 0.75 | 0.68 | 0.71 |
| **ResNet50** | 0.85 | 0.75 | 0.73 | 0.73 |
| **DenseNet121** | 0.87 | 0.82 | 0.74 | 0.77 |

ตารางการเปรียบเทียบจำนวนที่มีการเทคำตอบของโมเดลที่มีค่า F1-score สูงสุด 3 อันดับแรกของตอนที่ 1

|  |  |  |
| --- | --- | --- |
| **Model** | **F1-score** | **จำนวนประเภทที่มีการเทคำตอบ** |
| DenseNet121-Focal loss | 0.77 | 5 |
| DenseNet121-Oversampling | 0.75 | 3 |
| ResNet50-Focal loss | 0.73 | 3 |

ตอนที่ 2 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการ Transfer Learning จาก Pre-Trained Weights ของ ImageNet และการ Fine Tuning

2.1 ใช้การ Oversampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.81 | 0.74 | 0.67 | 0.69 |
| **ResNet50** | 0.79 | 0.65 | 0.69 | 0.66 |
| **DenseNet121** | 0.81 | 0.72 | 0.75 | 0.72 |

2.2 ใช้การปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.52 | 0.38 | 0.62 | 0.41 |
| **ResNet50** | 0.70 | 0.55 | 0.71 | 0.60 |
| **DenseNet121** | 0.77 | 0.61 | 0.72 | 0.64 |

2.3 ใช้ Focal loss

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.80 | 0.74 | 0.61 | 0.65 |
| **ResNet50** | 0.79 | 0.75 | 0.54 | 0.59 |
| **DenseNet121** | 0.84 | 0.78 | 0.67 | 0.72 |

ตอนที่ 2.5 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการเทรนใหม่ทั้งหมด โดยมีการตัดโรคผิวหนังประเภท Melanocytic nevi ในชุดข้อมูล Skin Cancer MNIST: HAM10000 ออกทำให้เหลือโรคผิวหนัง 6 ประเภท เพื่อช่วยแก้ปัญหาในเรื่องความไม่สมดุลของชุดข้อมูล

2.5.1 ใช้การ Oversampling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.77 | 0.81 | 0.75 | 0.78 |
| **ResNet50** | 0.77 | 0.78 | 0.79 | 0.78 |
| **DenseNet121** | 0.83 | 0.81 | 0.80 | 0.80 |

2.5.2 ใช้การปรับ Class weights ของแต่ละประเภท

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.77 | 0.76 | 0.79 | 0.77 |
| **ResNet50** | 0.81 | 0.82 | 0.81 | 0.80 |
| **DenseNet121** | 0.84 | 0.85 | 0.83 | 0.84 |

2.5.3 ใช้ Focal loss

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.77 | 0.77 | 0.71 | 0.73 |
| **ResNet50** | 0.83 | 0.82 | 0.80 | 0.81 |
| **DenseNet121** | 0.83 | 0.86 | 0.81 | 0.83 |

ตารางการเปรียบเทียบจำนวนที่มีการเทคำตอบของโมเดลที่มีค่า F1-score สูงสุด 3 อันดับแรกของตอนที่ 2.5

|  |  |  |
| --- | --- | --- |
| **Model** | **F1-score** | **จำนวนประเภทที่มีการเทคำตอบ** |
| **DenseNet121-Class weights** | 0.84 | 2 |
| **DenseNet121-Focal loss** | 0.83 | 2 |
| **ResNet50-Focal loss** | 0.81 | 3 |

ตอนที่ 3 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels และใช้เทคนิคต่าง ๆ ด้วยการ Transfer Learning Weights จากการทดลองตอนที่ 2.5 จากนั้นทำการ Fine Tuning

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **DenseNet121-1** | 0.92 | 0.93 | 0.92 | 0.92 |
| **DenseNet121-2** | 0.86 | 0.88 | 0.86 | 0.86 |
| **ResNet50-1** | 0.84 | 0.85 | 0.85 | 0.84 |

**ผลการดำเนินงานครั้งที่ 9** การ Ensemble โมเดล โดยการหาค่าเฉลี่ยเลขคณิต

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Ensemble** | 0.90 | 0.90 | 0.89 | 0.89 |

**บทสรุป**

**สรุปผลการปฏิบัติงาน**

ก่อนปฏิบัติงานทางผู้ปฏิบัติงานไม่มีพื้นฐานด้านการเรียนรู้ของเครื่องรวมถึงมีพื้นฐานการเขียนโปรแกรมเพียงเล็กน้อยจึงนำไปสู่การดำเนินงานครั้งที่ 1 คือเริ่มจากการฝึกเขียนโปรแกรมภาษาไพธอนจากนั้นก็เรียนรู้วิธีการเขียนโปรแกรมด้านการเรียนรู้ของเครื่องโดยใช้ Convolutional Neural Network เพื่อคัดแยกตัวเลขในชุดข้อมูล MNIST แล้วจึงค่อยเริ่มลงมือปฏิบัติงานในการดำเนินครั้งที่ 2 ซึ่งก็คือเก็บชุดข้อมูลโรคผื่นภูมิแพ้ผิวหนังผื่นภูมิแพ้ผิวหนังและโรคผื่นแพ้ผิวสัมผัส ซึ่งมีจำนวนเพียงแค่ 90 ภาพ เพื่อนำมาพัฒนาโมเดล เมื่อทดสอบโมเดลได้ค่า accuracy 0.59 จึงนำไปสู่การดำเนินครั้งที่ 3 เพื่อที่ทดสอบสมมติฐาน ว่าถ้ามีชุดข้อมูลที่เกี่ยวกับโรคผิวหนังมากขึ้นจะทำให้มีค่า accuracy สูงขึ้นหรือไม่ แต่เนื่องจากชุดข้อมูลนี้มีความไม่สมดุล ทำให้พบว่าข้อมูลที่มีความไม่สมดุลจะส่งผลต่อการทำนายโมเดล โดยโมเดลจะเทคำตอบไปประเภทที่มีจำนวนมาก ซึ่งจะทราบได้จากการใช้ Confusion Matrix แล้วคำนวณออกมาเป็นค่า Metrics ต่าง ๆ ได้แก่ Preicision, Recall และ F1-score การดำเนินครั้งที่ 4 เนื่องจากรูปของแต่ละอวัยวะมึความแตกต่างกัน เช่น แขนจะมีลักษณะเป็นท่อนยาว, ลำตัวจะมีลักษณะเรียบ และมือจะมีนิ้วแยกออกมา แต่เมื่อทำการแบ่งส่วนเป็นอวัยวะจะทำให้ชุดข้อมูลนั้นแบ่งเป็นหลายส่วนที่มีจำนวนในแต่ละส่วนน้อยลง ผู้ปฏิบัติงานจึงใช้ Support Vector Machine ในการคัดแยกเนื่องจากเป็นวิธีที่เหมาะกับชุดข้อมูลที่ไม่มาก จากนั้นจึงใช้ Convolutional Neural Network เพื่อเปรียบเทียบว่าวิธีการใดให้ผลลัพธ์ที่ดีกว่ากัน ซึ่งผลลัพธ์ที่ได้คือจาก Convolutional Neural Network ดีกว่า จากนั้นทางผู้ปฏิบัติงานมีความคิดว่า จะไม่แบ่งเป็นอวัยวะ แต่จะใช้วิธีครอปรูปให้เหลือเพียงบริเวณที่มีโรคเท่านั้นเพื่อตัดส่วนที่ไม่จำเป็นของภาพออกไป จึงนำไปสู่การดำเนินงานครั้งที่ 5 โดยแบ่งเป็น 2 โมเดล โมเดลแรกเพื่อแยกระหว่างโรคกับโรคผิวหนังอื่นๆ และโมเดลที่ 2 ซึ่งในการดำเนินการครั้งที่ 5 นี้ ได้เปลี่ยนวิธีการ Data Augmentation จากการเติมค่าพิกเซลที่มีค่า 0 แบบ abcddcba|abcd|dcbaabcd และใช้วิธีการปรับ Class weights เพื่อแก้ปัญหาชุดข้อมูลไม่สมดุล จากนั้นการดำเนินการต่อครั้งที่ 6 โดยได้มีการคัดแยกภาพในชุดข้อมูลบางส่วนที่ไม่มีความชัดเจนออก ไม่มีการแยกโมเดล และใช้วิธีการปรับ Class weights เพื่อแก้ปัญหาชุดข้อมูลไม่สมดุล จากนั้นพัฒนาโมเดล ผลที่ได้คือ Convolutional Neural Network ให้ผลที่ดีกว่า Support Vector Machine แต่เมื่อตรวจสอบ Grad-CAM พบว่าโมเดลมองจุดที่ไม่ใช่โรค แต่เป็นอย่างอื่นแทน เช่น มุมภาพ และขอบภาพ จากกดำเนินครั้งที่ 4 ถึง 6 จะสังเกตได้ว่าค่า F1-score ที่ได้จาก Convolutional Neural Network ให้ผลที่ดีกว่า Support Vector Machine เนื่องจากการใช้ Support Vector Machine ในโครงงานนี้เป็นเพียงแค่การนำภาพมา Flatten ไม่ได้เป็นการทำ Feature Extraction แต่ Convolutional Neural Network สามารถทำ Feature Extraction เองได้จึงได้ค่า F1-score ที่ดีกว่า จากนั้นผู้ปฏิบัติงานจึงพักการดำเนินครั้งที่ 7 เพื่อหาเทคนิคที่จะเพิ่มเติม ได้แก่ Hyperparameter tuning, Regularization และ Optimization จากนั้นนำความรู้ที่สะสมมาตั้งแต่การดำเนินงานครั้งที่ 1 จนถึงการดำเนินงานครั้งที่ 7 เพื่อนำมาใช้ในการดำเนินครั้งที่ 8 การ ซึ่งในตอนที่ 1 ที่เป็นการเทรนใหม่หมดพบว่าโมเดลที่ได้ค่า F1-score สูงสุด 3 อันดับแรก ได้แก่ และ ซึ่งเมื่อทำการตรวจสอบว่าโมเดลมีการเทคำตอบหรือไม่โดยการใช้ส่วนเบี่ยงเบนเฉลี่ย พบว่าทั้ง 3 โมเดล มีการเทคำตอบไปโมเดลละ 4, 3 และ 3 ประเภทตามลำดับ เนื่องจากเมื่อเราตั้งค่าขนาด Batch เพียงแค่ 32 ซึ่งถ้าตอนสุ่มภาพเข้าเวลาเทรน ประเภทที่มีจำนวนมากกว่าก็จะมีโอกาสที่จะอยู่ใน Batch นั้นมาก และในบางครั้งอาจจะมีแค่ประเภทเดียวหรืออาจจะมีประเภทอื่นอยู่ด้วยแต่มีเพียงจำนวนน้อยที่อยู่ใน Batch ทำให้เน็ตเวิร์คตอบเพียงแค่ประเภทที่มีมากเท่านั้น ทางผู้ปฏิบัติงานจึง นำการดำเนินงานตอนที่ 1 มาปรับปรุงเป็นตอนที่ 2 โดยการ Fine Tuning และเมื่อเปรียบเทียบค่า F1-score ในตอนที่ 1 และตอนที่ 2 จะพบว่าตอนที่ 1 จะมีค่า F1-score ที่สูงกว่าและสาเหตุคือการที่ชุดข้อมูลไม่เหมาะกับการ Transfer Learning จาก Pre-Trained Weights ของ ImageNet เนื่องจาก ImageNet เป็นชุดข้อมูลที่เกี่ยวกับสิ่งของทั่วไป เช่น สุนัข รถยนต์ และเครื่องบิน แต่ชุดข้อมูล Skin Cancer MNIST: HAM10000 เป็นชุดข้อมูลโรคผิวหนัง ทำให้ features จากการเปรียบเทียบของตอนที่ 1 และตอนที่ 2 ทำให้เกิดตอนที่ 3 คือการเทรนใหม่ทั้งหมด แต่มีการตัดโรคผิวหนังประเภท Melanocytic nevi ในชุดข้อมูล Skin Cancer MNIST: HAM10000 ออก ทำให้เหลือโรคผิวหนัง 6 ประเภท เพื่อช่วยแก้ปัญหาในเรื่องความไม่สมดุลของชุดข้อมูล ได้โมเดลที่มีค่า F1-score สูงสุด 3 อันดับแรก ได้แก่ DenseNet121-Class weights, DenseNet121-Focal loss และ ResNet50-Focal loss ซึ่งจะสังเกตได้ว่าจะเป็นโมเดลที่ได้จากการปรับ Class weights และใช้ Focal loss เพราะการปรับ Class weights จะเพิ่มค่า ของ Cross Entropy ซึ่งจะได้เป็น เป็นการถ่วงน้ำหนักให้ของประเภทที่มีจำนวนข้อมูลที่น้อยจะมีค่า ที่มาก และประเภทที่มีจำนวนข้อมูลที่มากจะมี น้อย จึงทำให้เน็ตเวิร์กพยายามตอบจำนวนข้อมูลที่น้อยให้ถูก เพราะถ้าตอบผิดค่า loss จะเพิ่มขึ้นส่วนด้วย การใช้ Focal loss ซึ่งมีสมการคือ ค่าน้ำหนักของค่า loss ในแต่ละตัวจะขึ้นอยู่กับผลการทำนายซึ่งคือความน่าจะเป็นของประเภทนั้นๆ หมายความว่าถ้าเน็ตเวิร์กมีความมั่นใจมาก ซึ่งหมายถึงมีความน่าจะเป็นมาก ทำให้ค่า loss ต่ำลง แต่ถ้าความน่าจะเป็นน้อย ค่า loss ก็จะเพิ่มขึ้น หมายความว่าถ้าค่า loss สูงคือชุดข้อมูลมีความยาก ทำให้เน็ตเวิร์คสนใจกับข้อมูลที่ยากมากยิ่งขึ้น แต่การใช้การ Oversampling คือการสุ่มสร้างข้อมูลเพิ่มซึ่งข้อมูล 1 ตัวอย่าง อาจสร้างเพิ่มเป็นจำนวนที่มากและมีภาพที่ซ้ำกันทำให้ Train accuracy มีค่าสูง แต่ Validation accuracy ไม่มีการพัฒนาเนื่องจาก Oversampling แค่ Train set และเมื่อทำการทดสอบว่าโมเดลมีการเทคำตอบหรือไม่ ผลปรากฏว่ามีการเทคำตอบแค่ 2, 2 และ 3 ประเภท ซึ่งถือว่าโมเดลที่ได้จากการดำเนินงานครั้งที่ 8 ตอนที่ 3 มีคุณภาพดีกว่า และเมื่อทำ Grad-CAM พบกว่าโมเดลมองภาพโรคผิวหนังได้ถูกจุด จากนั้นนำโมเดลที่ได้จากการดำเนินงานครั้งที่ 8 ตอนที่ 3 ไปทำการ Transfer Learning รวมถึง Fine Tuning ต่อให้ได้โมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 ทำให้ได้โมเดลที่มีค่า F1-score 0.84, 0.92 และ 0.86 เมื่อทำ Grad-CAM พบกว่าโมเดลมองภาพโรคผิวหนังได้ถูกจุด จากนั้นเมื่อได้ 3 โมเดลจากการดำเนินครั้งที่ 8 ตอนที่ 3 มาก็นำไปสู่การดำเนินงานครั้งที่ 9 ซึ่งก็คือการนำผลการทำนายมาเฉลี่ยแบบเลขคณิตกันได้ค่า F1-score 0.89

**ปัญหา**

* ผู้ปฏิบัติงานมีประสบการณ์ด้านคอมพิวเตอร์เพียงเล็กน้อยทำให้ต้องใช้เวลาในการเรียนรู้
* ชุดข้อมูลมีจำนวนน้อยและไม่มีคุณภาพ เนื่องจากเป็นข้อมูลทางด้านการแพทย์ ซึ่งเป็นข้อมูลที่ต้องทำการทำเรื่องขออนุญาติจากทางโรงพยาบาลและคนไข้ก่อน แต่การติดต่อนั้นเป็นไปด้วยความลำบาก

**ข้อเสนอแนะ**

* สามาถนำไปเปรียบเทียบกับงานวิจัยด้านการเรียนรู้ของเครื่อง
* สามารถนำโมเดลที่ได้จากการพัฒนาในโครงงานนี้ไปต่อยอดเป็นเว็บแอปพลิเคชันสำหรับคัดโรคผิวหนังต่อไปเพื่อแก้ไขปัญหาต่าง ๆ ได้แก่ ลดความแออัดมนโรงพยาบาล ลดมลพิษจากยานพาหนะในการเดินทางไปโรงพยาบาล และลดภาระของแพทย์และพยาบาล

**แนวทางการดำเนินงานต่อเนื่อง**

* ใช้การปรับ Class weight ร่วมกับ Focal loss
* Preprocess ภาพก่อนการพัฒนาโมเดล เช่น การทำ Histogram Equalization
* ต่อยอดการทำ Data Augmentation โดยใช้ PCA Color Augmentation เป็นพื้นฐาน เพื่อให้ชุดข้อมูลมีความหลากหลายทางสีผิวมากขึ้น
* ใช้เทคนิค One Shot Learning เนื่องจากเหมาะกับชุดข้อมูลที่มีจำนวนน้อย
* เปลี่ยนจากการเทรนโมเดลด้วยภาพ RGB เป็น HSV
* ขอ Dataset จากสถาบันโรคผิวหนังแห่งชาติ เก็บชุดข้อมูลโดยเก็บเป็น VDO เพื่อมาแบ่งเป็นทีละภาพ
* เพิ่มการคัดแยกโรคผิวหนังด้วยการใช้การตอบคำถามร่วมกับการใช้ Classifier

**เอกสารอ้างอิง**

ธนาวุฒิ ประกอบผล. (2552). **โครงข่ายประสาทเทียม**. *มฉก.วิชาการ*, 12(24), 73-87.

ธีรศักดิ์ คำแก้ว. (2557). **การศึกษาความพึงพอใจของพนักงานที่ใช้แอพพลิเคชั่นสั่งสินค้าผ่านทางโทรศัพท์มือถือ: กรณีศึกษาร้านโซล จังหวัดชลบุรี**. (ออนไลน์). สืบค้นได้จาก: http://dspace.bu.ac.th/bitstream/123456789/1951/1/theerasak\_kamk.pdf. [15 กันยายน 2562].

สมาคมแพทย์ผิวหนังแห่งประเทศไทย สมาคมโรคภูมิแพ้และวิทยาภูมิคุ้มกันแห่งประเทศไทย ชมรมแพทย์ผิวหนังเด็กแห่งประเทศไทย. (ม.ป.ป.). **แนวทางการดูแลรักษาโรคผื่นภูมิแพ้ผิวหนัง**. (ออนไลน์). สืบค้นได้จาก: http://www.dst.or.th/files\_news/Atopic\_Dermatitis\_2010.pdf. [17 พฤษภาคม 2562].

อุดมศักดิ์ เลิศสุชาตวนิช และพิเชษฐ์ สืบสายพรหม. (2556). **การประยุกต์ใช้เทคโนโลยีการประมวลภาพสําหรับประเมินระดับความรุนแรงของโรคข้าว**. (ออนไลน์). สืบค้นได้จาก: http://www.lib.ku.ac.th/KUCONF/2556/KC5101046.pdf. [3 ธันวาคม 2561].

Agnieszka Mikołajczyk and Michał Grochowski. (2018). **Data augmentation for improving deep learning in image classification problem.** International Interdisciplinary PhD Workshop, 9-12 May 2018

Andrew et al. (2017). **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications**. (Online). from https://arxiv.org/pdf/1704.04861.pdf. Retrieved 11 September 2019.

Barret et al. (2018). **Learning Transferable Architectures for Scalable Image Recognition**.(Online). from https://arxiv.org/pdf/1707.07012.pdf. Retrieved 11 September 2019.

Christian et al. (2014). **Going deeper with convolutions**.(Online). fromhttps://arxiv.org/pdf/1409.4842.pdf. Retrieved 11 September 2019.

Gao et al. (2017). **Densely Connected Convolutional Networks**. (Online). from https://arxiv.org/pdf/1608.06993.pdf. Retrieved 11 September 2019.

ISAAC. (2003). **ISAAC MANUAL**. (Online). from: http://isaac.auckland.ac.nz/phases/ phaseone/phaseonemanual.pdf. Retrieved 18 September 2019.

Kaiming et al. (2015). **Deep Residual Learning for Image Recognition**. (Online). from https://arxiv.org/pdf/1512.03385.pdf. Retrieved 15 September 2019.

Karen Simonyan and Andrew Zisserman. (2015). **VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION**. (Online). from

https://arxiv.org/pdf/1409.1556.pdf. Retrieved 8 September 2019.

Luke Taylor and Geoff Nitschke. (2018). **Improving Deep Learning with Generic Data Augmentation**. IEEE Symposium Symposium Series on Computational Intelligence SSCI 2018, 18-21 November 2018.

Ovidiu M. Şerban. (2013). **Detection and integration of affective feedback into distributed**

**interactive systems**. (Online). from https://www.researchgate.net/publication/281015715\_Detection\_and\_integration\_of\_ affective\_feedback\_into\_distributed\_interactive\_systems. Retrieved 19 September 2019.

Philipp Tschandl et al. (2018). **The HAM10000 dataset, a large collection of multi-source**

**dermatoscopic images of common pigmented skin lesions**. (Online). from

https://arxiv.org/abs/1803.10417.pdf. Retrieved 12 September 2019.  
Tsung-Yi Lin et al. (2017). **Focal Loss for Dense Object Detection.** (Online). from

https://arxiv.org/abs/1708.02002.pdf. Retrieved 15 May 2020.

**ภาคผนวก**

**ผลการดำเนินงานครั้งที่ 4 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403 โดยการแบ่งส่วนอวัยวะ (เพิ่มเติม)**

ตอนที่ 1 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในแต่ละส่วนของอวัยวะในชุดข้อมูล PJ61403 โดยการ Flatten ภาพแล้วใช้ Support Vector Machine ตามโมเดลต่าง ๆ จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดลในแต่ละอวัยวะ

**แขน**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Linear** | 0.65 | 0.67 | 0.68 | 0.65 |
| **Sigmoid** | 0.48 | 0.58 | 0.56 | 0.47 |
| **Polynomial** | 0.70 | 0.67 | 0.65 | 0.66 |
| **RBF** | 0.65 | 0.67 | 0.68 | 0.65 |

**ลำตัว**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Linear** | 0.50 | 0.48 | 0.48 | 0.48 |
| **Sigmoid** | 0.35 | 0.17 | 0.50 | 0.26 |
| **Polynomial** | 0.35 | 0.17 | 0.50 | 0.26 |
| **RBF** | 0.35 | 0.42 | 0.48 | 0.29 |

**มือ**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Linear** | 0.65 | 0.67 | 0.68 | 0.65 |
| **Sigmoid** | 0.35 | 0.17 | 0.50 | 0.26 |
| **Polynomial** | 0.42 | 0.69 | 0.56 | 0.38 |
| **RBF** | 0.48 | 0.70 | 0.60 | 0.45 |

**ขา**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Kernel** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **Linear** | 0.60 | 0.59 | 0.59 | 0.58 |
| **Sigmoid** | 0.35 | 0.17 | 0.50 | 0.26 |
| **Polynomial** | 0.35 | 0.17 | 0.50 | 0.26 |
| **RBF** | 0.35 | 0.17 | 0.50 | 0.26 |

ตอนที่ 2 ตารางการเปรียบเทียบค่า Metrics ของการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในแต่ละส่วนของอวัยวะในชุดข้อมูล PJ61403 โดยการใช้ Convolutional Neural Network จากการใช้สถาปัตยกรรมที่ใช้ในการแข่งขัน ILSVRC ที่มีการใช้ขนาดรูป 224\*224 pixels ด้วยการเทรนใหม่ทั้งหมด จากนั้นทำการเปรียบเทียบค่า F1-score ของแต่ละโมเดลในแต่ละอวัยวะ

**แขน**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **โมเดล** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.50 | 0.25 | 0.50 | 0.33 |
| **ResNet50** | 0.61 | 0.78 | 0.61 | 0.54 |
| **MobileNet** | 0.50 | 0.25 | 0.50 | 0.33 |
| **DenseNet121** | 0.84 | 0.85 | 0.84 | 0.84 |
| **NASNetMobile** | 0.62 | 0.62 | 0.62 | 0.62 |

**ลำตัว**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.50 | 0.25 | 0.50 | 0.33 |
| **ResNet50** | 0.52 | 0.57 | 0.52 | 0.52 |
| **MobileNet** | 0.50 | 0.25 | 0.50 | 0.33 |
| **DenseNet121** | 0.77 | 0.77 | 0.77 | 0.77 |
| **NASNetMobile** | 0.83 | 0.87 | 0.83 | 0.83 |

**มือ**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.50 | 0.25 | 0.50 | 0.33 |
| **ResNet50** | 0.76 | 0.84 | 0.76 | 0.75 |
| **MobileNet** | 0.50 | 0.25 | 0.50 | 0.33 |
| **DenseNet121** | 0.94 | 0.95 | 0.94 | 0.94 |
| **NASNetMobile** | 0.91 | 0.90 | 0.90 | 0.90 |

**ขา**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architecture** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| **VGG16** | 0.50 | 0.25 | 0.50 | 0.33 |
| **ResNet50** | 0.54 | 0.54 | 0.54 | 0.53 |
| **MobileNet** | 0.50 | 0.25 | 0.5 | 0.33 |
| **DenseNet121** | 0.83 | 0.83 | 0.83 | 0.83 |
| **NASNetMobile** | 0.69 | 0.70 | 0.69 | 0.69 |

**ภาพตัวอย่าง Grad-CAM จากการดำเนินงานครั้งที่ 6 การพัฒนาโมเดลสำหรับการคัดแยกโรคผิวหนังในชุดข้อมูล PJ61403 โดยการปรับปรุงวิธีการทำ Data Augmentation และใช้การปรับค่า Class weights ของแต่ละประเภท  
A picture containing light

Description automatically generated** **A picture containing blurry, white, man, standing

Description automatically generated** A star in the background

Description automatically generated A close up of a colorful background

Description automatically generated

**ภาพตัวอย่าง Grad-CAM จากการดำเนินงานครั้งที่ 8 การพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล Skin Cancer MNIST: HAM10000 โดยการใช้** **Convolutional Neural Network เพื่อนำ Weights ไป Transfer Learning ให้กับการพัฒนาโมเดลสำหรับการคัดแยกประเภทโรคผิวหนังในชุดข้อมูล PJ61403**

ตอนที่ 2.5 โมเดล DenseNet121-Class weights

A star filled sky

Description automatically generatedA picture containing star, sitting, blue, blurry

Description automatically generatedA picture containing star

Description automatically generated

A picture containing star

Description automatically generatedA star filled sky

Description automatically generatedA star in the background

Description automatically generated

ภาพตัวอย่างการใช้โมเดล โมเดล DenseNet121-Class weights ที่ได้จากการทดลองตอนที่ 2.5

![A screenshot of a cell phone

Description automatically generated]()

ตอนที่ 3 โมเดล DenseNet121-1

**A star in the background

Description automatically generatedA picture containing star, clock, light

Description automatically generatedA star filled sky

Description automatically generatedA star in the background

Description automatically generated**