

## Programming Homework: Guideline (คะแนน 20%)

1. การบ้านนี้เป็นงานกลุ่ม โดยกำหนดให้ 1 กลุ่มมีสมาชิก 4 คน (กรณีจำนวนนักศึกษาในคลาสหาร 4 ไม่ลงตัว อาจารย์จะแจ้งวิธีการแบ่งกลุ่มให้ในคลาสดะ)
2. การบ้านครั้งนี้ไม่จำเป็นต้องส่งเป็น Github และไม่จำเป็นต้องส่งรายงานใด ๆ แนะนำเสนอผลการทดลองที่ได้ โดยละเอียดผ่านการ presentation ในคาบและส่งข้อมูลที่อยู่ในไฟล์ presentation ก็พอ ทั้งนี้หากกลุ่มไหน ต้องการงานเก็บไว้ใน Github เพื่อเป็น profile/portfolio อาจารย์มีตัวอย่างงาน Github ที่ทำดี ๆ ของรุ่นพี่ให้ดูดังนี้ค่ะ (ลิงก์ที่ 2 จะหิวข้อใกล้เคียงกับ Programming Homework ของเราค่ะ)
  - [https://github.com/lukplamino/DADS7202\\_HW01\\_MNLP\\_Group](https://github.com/lukplamino/DADS7202_HW01_MNLP_Group)
  - [https://github.com/lukplamino/DADS7202\\_HW02-CNN\\_MNLP\\_Group](https://github.com/lukplamino/DADS7202_HW02-CNN_MNLP_Group) , <https://github.com/dads7202/assignment2>
3. หัวข้อการบ้านคือให้แต่ละกลุ่มเตรียสร้าง Image Classifier ด้วย CNN บน A new image dataset ที่พวกเรารวบรวมมาเอง โดยมีเงื่อนไขดังนี้ค่ะ
  - 1) คลาสที่จะเลือกมาให้โมเดลพยากรณ์ต้องเป็นคลาสที่อยู่นอกเหนือจาก 1,000 ImageNet classes หรือก็คือ หากเอาภาพคลาสเหล่านี้ไปให้ CNN (pre-trained on ImageNet) พยากรณ์เลยโดยไม่เทรนใด ๆ เพิ่มเติม ผลลัพธ์การพยากรณ์จะไม่ถูกต้อง (เช่น จำแนกผิดไปเป็นคนละคลาสเลย) หรือถูกต้องแค่บางส่วน (เช่น จำแนกได้ว่าเป็นรถ แต่จำแนกละเอียดลงไปว่าเป็นรถตึก ๆ ไม่ได้)
    - ตัวอย่างยอดฮิตของรุ่นพี่ เช่น คลาสรถสไต้ไทย ๆ (ตึกๆ กระบะ อีแต่น ฯลฯ) คลาสขนมไทย คลาสผลไม้ไทย ฯลฯ
  - 2) นักศึกษาในกลุ่มต้องเป็นผู้รวบรวม A new image dataset ของกลุ่มขึ้นมาเอง โดยอาจเป็นการรวมภาพจาก public dataset หลาย ๆ อันมาประกอบเข้าด้วยกัน, การทำ web scraping ภาพมาเพิ่มเติม, การถ่ายรูปมาเองบางส่วน, หรือใช้เทคนิคข้างต้นผสมผสานกันเพื่อให้ได้เซตภาพเป้าหมายตามที่ต้องการ แต่จะต้องไม่ใช่การนำ Image dataset ที่มีอยู่แล้วมาใช้ตรง ๆ โดยนักศึกษาไม่ได้มีการรวบรวมข้อมูลเพิ่มเติมเองเพื่อสร้าง A new dataset นี้เลย
  - 3) เพื่อให้ dataset ที่ได้มีความหลากหลาย A new image dataset ของแต่ละกลุ่มจะต้องหิวข้อไม่ซ้ำกันนะค่ะ กลุ่มไหนจงก่อนได้ก่อน พิมพ์จงโดยกด reply ใต้โพสต์นี้ บอกชื่อกลุ่ม ชื่อ class ที่จะทำให้ และใส่ @ tag ชื่อสมาชิกในกลุ่มทุกคนด้วยค่ะ

- 4) ในส่วนของตัวโมเดล Foundation (Pre-trained) CNN ที่จะเลือกมาใช้เป็น backbone (feature extractor) จะต้องเลือกอย่างน้อย 3-5 architectures ที่แตกต่างกันอย่างมีนัยสำคัญ เพื่อทดลองเปรียบเทียบกันว่า architecture ใดจะให้ผลลัพธ์ที่ดีที่สุดในการวิเคราะห์แต่ละด้าน
    - เช่น ถ้าเลือก VGG-16 กับ VGG-19 มา อาจารย์จะนับว่าเป็นแค่การทดลองบน 1 architecture แต่ถ้าเลือก VGG-16, ResNet-50, EfficientNet-B7 อย่างนี้จึงจะนับเป็นการทดลองบน 3 architectures ค่ะ
  - 5) ในส่วนของการเทรนโมเดลนั้น ไม่จำกัดว่าจะใช้ transfer learning (ไม่ต้อง ImageNet weights ของเดิมเลย) หรือ finetuning (มีการอัปเดต ImageNet weights ในบางเลเยอร์ของโมเดล) นักศึกษาแต่ละกลุ่มสามารถทดลองได้เอง เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด
4. การเตรียมสไลด์นำเสนอให้เตรียมเรียงตามหัวข้อนี้เลยคะ เป็น standard ของการรายงานผลการทดลองด้าน Deep Learning (1 กลุ่มมีเวลานำเสนอแค่ประมาณ 5 นาทีนะคะ อาจต้องรีบพูดทำเวลานิดนึง)
- 1) Dataset description
    - ที่มาของข้อมูลว่านำมาจากแหล่งไหนบ้าง ถ้าถ่ายรูปมาเองให้บอก setting ต่าง ๆ ของการถ่ายรูปโดยละเอียด (เพื่อแสดงว่ากระบวนการเก็บภาพไม่มี bias)
    - Exploratory Data Analysis (EDA) ของ dataset เบื้องต้น เช่น จำนวนภาพแต่ละคลาสมีเท่าไร, ภาพแต่ละภาพขนาดเท่าไร, ภาพสีก็ภาพ ภาพ grayscale ก็ภาพ, ฯลฯ
    - Dataset ที่นำมาใช้มีความ imbalanced ในจำนวนภาพระหว่างคลาสหรือไม่? ถ้ามีต้องนำเสนอว่าจะแก้ปัญหา imbalanced dataset นี้ยังไง
  - 2) Data preparation
    - การ preprocess และ/หรือ postprocess ข้อมูลรูปภาพสำหรับนำไปใช้เป็น training sample เช่น การย่อขนาดภาพ, การ crop ภาพ, ฯลฯ
    - วิธีการทำ train / validation / test split ว่ามีวิธีแบ่งยังไง แบ่งแล้วได้เซตละเท่าไร
    - การทำ data augmentation บนรูปภาพที่จะใช้เป็น training sample ว่าเราเลือกใช้ operation ใดบ้าง (เลือกตามความเหมาะสมของ dataset นะคะ) \*\*\*อันนี้ขอบังคับว่าทุกกลุ่มต้องทำ data augmentation นะคะ แต่จะใช้วิธีการ offline/online ก็แล้วแต่คะ\*\*\*
  - 3) Model architecture
    - บอกว่ากลุ่มเราเลือก CNN backbone ตัวไหนมาบ้าง พร้อมเหตุผลสั้น ๆ

- ประสิทธิภาพของโมเดล CNN backbone (ที่ยังไม่ทำการเทรนอะไรเพิ่มเลย) บน A new image dataset ของเรา (เพื่อแสดงว่าถ้าไม่เทรนอะไรเพิ่มเลย โมเดลมันยังไม่รู้จัก Image dataset ของเราจริง ๆ ) เช่น อาจแคปผลลัพธ์ภาพบางคลาสมาให้ดูเป็นตัวอย่าง
- อธิบายว่าเราทำการตัดตอน original backbone ส่วนไหนออกไป หรือเก็บส่วนไหนไว้บ้าง
- ส่วนต่อเติมใหม่ที่เราเพิ่มเข้าไป (เช่น ส่วน classifier) ได้แก่ ส่วนไหนบ้าง, เพิ่มไปกี่เลเยอร์, แต่ละเลเยอร์ต่อกันยังไง, และ รายละเอียดพารามิเตอร์ต่าง ๆ ของเลเยอร์ที่เพิ่มเข้าไปเป็นยังไง

#### 4) Training method

- อธิบายวิธีการเทรน เช่น เรา freeze weights ของเลเยอร์ไหนไว้บ้าง, เลเยอร์ไหนบ้างที่จะถูก update weights ในระหว่างการเทรน, วิธีการอื่น ๆ ถ้าไม่ใช่การเทรนแบบมาตรฐาน end-to-end training, ฯลฯ
- พารามิเตอร์ที่ใช้ในการเทรน เช่น optimizer ที่เลือกใช้, สมการ loss ที่เลือกใช้, initial learning rate, การตั้งค่า hyperparameter อื่น ๆ
- เทคนิคการทำ hyperparameter tuning ของกลุ่มเรา (ในกรณีงานกลุ่ม แนะนำให้ใช้เครื่องมือ WandB ที่เคยสอนไป เพื่อให้รวบรวมผลการทดลองของทุกคนในกลุ่มไว้ใน project เดียวกัน และคัดเลือกผลที่ดีที่สุดขึ้นมาได้ง่ายนะคะ)

#### 5) Evaluation metrics

- อธิบายว่างานนี้จะใช้ Metric อะไรในการประเมินผลลัพธ์บ้าง สำหรับงาน Image classification ก็คือตัวเบสิกค่ะ พวก accuracy, precision, recall, F1 score

#### 6) Experimental results

- ดูกราฟผลลัพธ์การเทรนว่าเกิด overfit / underfit หรือเปล่า
- เปรียบเทียบผลลัพธ์ของโมเดลแต่ละตัว บน Evaluation metrics ทุกตัวที่พูดถึงในหัวข้อ 5
- \*\*\* ตัวเลขแต่ละตัวที่รายงาน จะต้องรายงานเป็นค่า  $\text{mean} \pm \text{SD}$  นะคะ \*\*\* โดยสำหรับโมเดล A ตัวที่ fix ค่า hyperparameter ทุกอย่างเรียบร้อยแล้ว ให้ทำการเทรนโมเดลนั้นซ้ำตั้งแต่ต้นอย่างน้อย 3 – 10 หน (แต่ละหนคือโมเดล A โครงสร้างเหมือนกันเป๊ะ แต่จะมี initial weights ต่างกัน) แล้วเอาโมเดล 3-10 ตัวที่ได้ (A1, A2, A3, ..., A10) นั้นมาวัดประสิทธิภาพบน Evaluation metrics ทุกตัว เพื่อแสดงผลโดยเฉลี่ยของโมเดล A ตัวนี้ในรูป  $\text{mean} \pm \text{SD}$  ค่ะ

## 7) Discussion and conclusions

- วิเคราะห์ว่าผลการทดลองในหัวข้อ 6 สรุปได้ว่าตัวไหนดีที่สุด (แล้วดีที่สุດอย่างมีนัยสำคัญหรือเปล่า? หรือชนะแค่เฉียดๆ) ตัวไหนแย่ที่สุด ตัวไหนดูแล้วเหมาะสมที่สุดที่จะนำไปใช้ในบริบทของงานเรา
  - วิเคราะห์ค่าผลการทดลองว่ามีจุดไหนที่ค่าผิดปกติบ้างไหม ถ้ามีคิดว่าความผิดปกตินั้นน่าจะเกิดจากอะไร อนาคตควรแก้ไขยังไง
  - GradCAM analysis รัน GradCAM เพื่อยืนยันว่าโมเดลของเราพยากรณ์โดยมองภาพถูกบริเวณที่ควรจะเป็นอยู่หรือไม่ \*\*\* อันนี้ขอว่าต้องมีนะสำหรับการบ้านนี้ \*\*\*
  - Eyeball analysis แคปตัวอย่างผลลัพธ์ที่น่าสนใจเป็นภาพแยกมาให้คนดูเป็นตัวอย่างชัด ๆ เช่น ผลลัพธ์เปรียบเทียบระหว่าง โมเดล pre-trained (ที่ยังไม่ได้เทรนเพิ่ม) vs. โมเดล A ของเรา vs. โมเดล B ของเรา
5. สิ่งที่ต้องส่งในวันพรีเซนตการบ้าน (ตัวแทนกลุ่ม 1 คนส่งในโพสต์ที่กำหนด โดยระบุชื่อกลุ่ม และ tag สมาชิกทุกคนในกลุ่มนะคะ)
- 1) ไฟล์ PDF รายชื่อสมาชิกและหน้าที่ความรับผิดชอบของแต่ละคน ตั้งชื่อไฟล์ว่า HW\_ชื่อกลุ่ม\_member.pdf
  - 2) ไฟล์ PDF ของ presentation ที่นำเสนอ ตั้งชื่อไฟล์ว่า HW\_ชื่อกลุ่ม\_presentation.pdf

\*\*\*\*\* THE END \*\*\*\*\*