Scripts

Pretrained ใน nlp อย่างพวก BERT Robert XLNET มีการออกแบบสถาปัตยากรรมที่แตกต่างกัน แต่มีแนวคิดเดียวกันคือใช้ประโยชน์การข้อความที่ไม่มี label กำกับ เพื่อสร้างโมเดล ที่มีความเข้าใจภาษา ก่อนที่จะปรับแต่งอีกรอบและนำไปใช้แบบเฉพาะเจาะจง เช่น qa sentiment ตามรูป 1 จะ pretraining 2 คือ finetune ในการ pretrained มี2 ประเภทคือ โมเดล ภาษา Language model คือการประมวลผลจากซ้ายไปขวา next sentence predict และ MLM คือการทำนายคำที่ถูกปิดบังโดยการเรียนรู้จะเห็น input ทั้งซ้ายไปขวาและขวาไปซ้าย bidirectional การ pretrained โมเดลด้านภาษาใน Bert ถูกทำด้วยการทำนาย 15% ของ token ใน input ซึ่งถูกสุ่มขึ้นมา tokens เหล่านี้ถูกทำ pre processed ทำให้ลดการเรียนรู้ของโมเดลไป นักพัฒนาของ google เลยพัฒนา Electra ขึ้นมาโดยออกแบบวิธีการทำ MLM ของ Bert ดั้งเดิมให้ดีขึ้น และเมื่อนำไป fine tune จะให้ประสิทธิภาพดีกว่าและใช้ทรัพยากรน้อยลง โดยบางขั้นตอนจะผสมระหว่าง RoBERTa และ XLNET

การทำงานของ Electra ใช้วิธีการ pretrained ที่เรียกว่า Replace Token Detection การตรวจหา token มาแทนที่ เป็นโมเดลแบบสองทิศทาง โดยใช้วิธีทำงานคล้ายกับ GANs คือมี ai 2 ตัว โดยมี generator เป็นตัวสร้าง data ปลอมที่ใกล้เคียงกับ data จริง อีกตัวคือ discriminator เป็นตัวแยก data ว่า data ที่สร้างขึ้นมาจาก generator เป็น data จริงหรือ data ปลอม if 1 รูปจริง 0 คือรูปที่มาจาก G ในการทำงาน D จะบอกคำตอบที่ทำนายให้ G หลังจากนี้ ai 2 ตัวก็จะปรับปรุงการทำงานของตัวเองต่อไป คือ g พยายามหลอกให้ได้ d ก็จะพยายามทำให้แม่นยำขึ้นเรื่อย ข้อดีคือ 2 ai ช่วยกันทำงานแล้วแก้ข้อผิดพลาดของตัวเอง โดยไม่ต้องพึ่งมนุษย์ electra จำเป็นต้องมีตัวอย่างให้ก่อนเพื่อให้ได้ประสิทธิภาพเท่าเทียม แต่การแทนโทเค็นส่งผลให้เกิดการเรียนรู้การแทนค่าที่มีประสิทธิภาพ เนื่องจากแบบจำลองต้องเรียนรู้การกระจายข้อมูลที่แม่นยำเพื่อแก้ปัญหา

ผลลัพธ์ ในการเปรียบเทียบ ELECTRA กับแบบจำลอง NLP อื่นๆ พบว่ามีการปรับปรุงอย่างมากในเรื่องงบประมาณในการประมวลผลที่เท่ากัน สามารถมีประสิทธิภาพเทียบเท่ากับRoBERTaและXLNETในขณะที่ใช้การประมวลผลน้อยกว่า 25% เมื่อเทียบกับ ELECTRA ขนาดเล็กให้ความแม่นยำที่ดีบน GPU ตัวเดียวในระยะเวลา 4 วันและมีประสิทธิภาพเหนือกว่า GPT และ ELECTRA ขนาดใหญ่ทดลองในชุดคำถาม SQuAD 2.0 มีประสิทธิภาพสูงกว่า RoBERTa, XLNeT และ ALBERT