COS4312 Deep Learning

นายพลวัฒน์ ช้างโต รหัสนักศึกษา 6601001123

---

1. หลัง vectorize แล้ว x_train.shape มีค่าเป็นอะไร

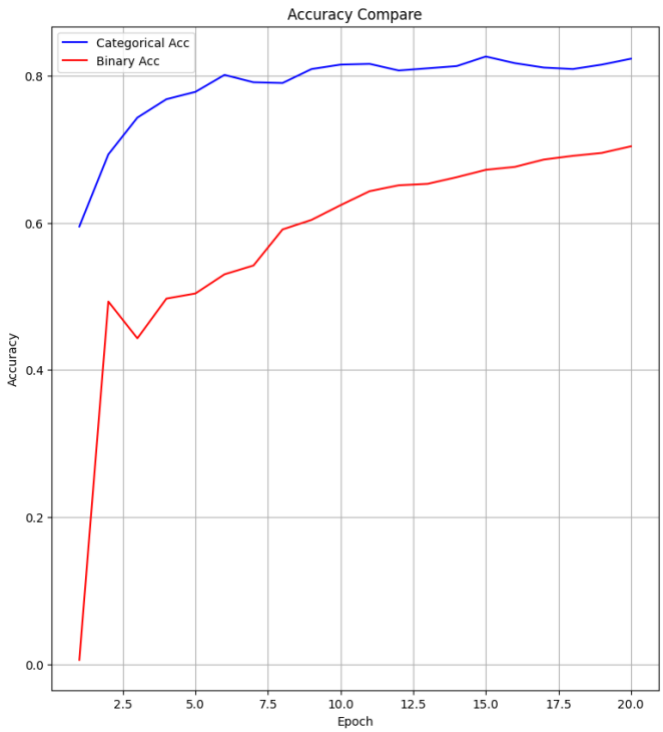
<u>ตอบ</u> x_train ก่อน vectorize_sequences(…) จะเป็น (8982,) และหลัง vectorize_sequences(x_train) แล้ว x_train.shape จะได้ (8982,10000)

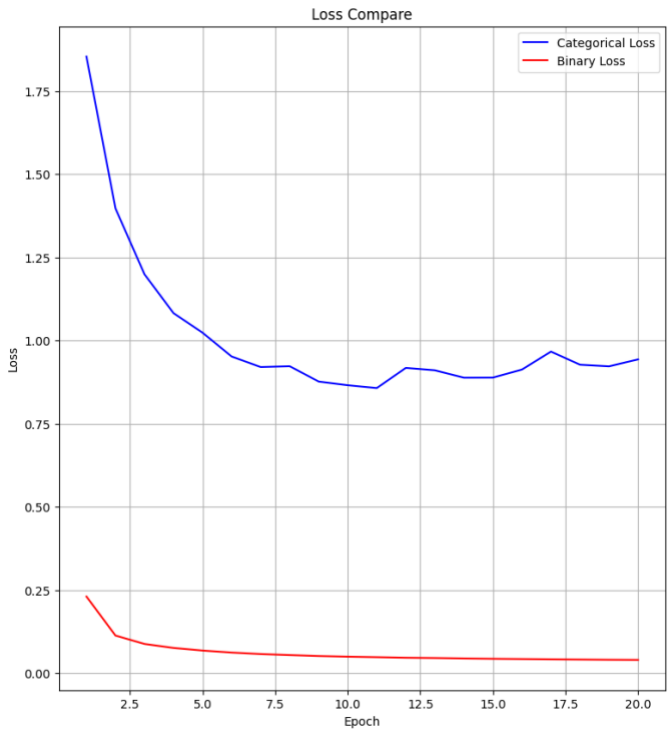
2. categorical_crossentropy เหมาะกับปัญหาประเภทใด

<u>ตอบ</u> categorical_crossentropy เหมาะกับปัญหาที่มีความเป็นไปได้ของคำตอบมากกว่า 2 รูปแบบ

3. หากใช้ binary_crossentropy จะเกิดปัญหาอะไร (อธิบายเชิงแนวคิดและรันแสดงผล)

<u>ตอบ</u>

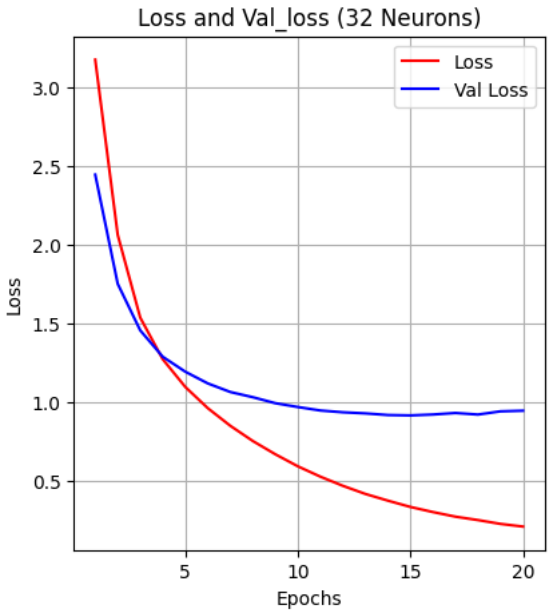


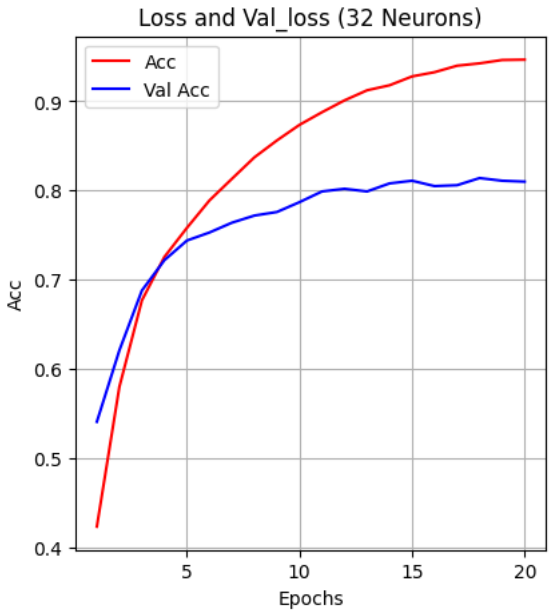


สาเหตุที่เมื่อใช้ binary_crossentropy (BCE) แทน categorical_crossentropy แล้วทำให้มีค่า loss ที่น้อยลงนั้น เพราะว่า ณ Output Layer มีจำนวน Neuron อยู่ 46 Neuron เพราะฉะนั้น BCE จะทำการคำนวณค่า Loss แยกแต่ละ Neuron ทำให้ Sample ที่มีค่าเป็น 0 ตรงกับคำเฉลย ซึ่งอาจมีอยู่มากถึง 45 ตัวที่ตรงกับคำเฉลยนั้น ส่งผลให้ค่า Loss ต่ำลง และต้องนำไปเฉลี่ยกับจุดที่ทำนายผิด ทำให้ค่า Loss ที่ได้รับนั้นต่ำมาก ๆ

4. ลองลด neuron จาก 64 ⟶ 32 ผลเปลี่ยนอย่างไร

<u>ตอบ</u>





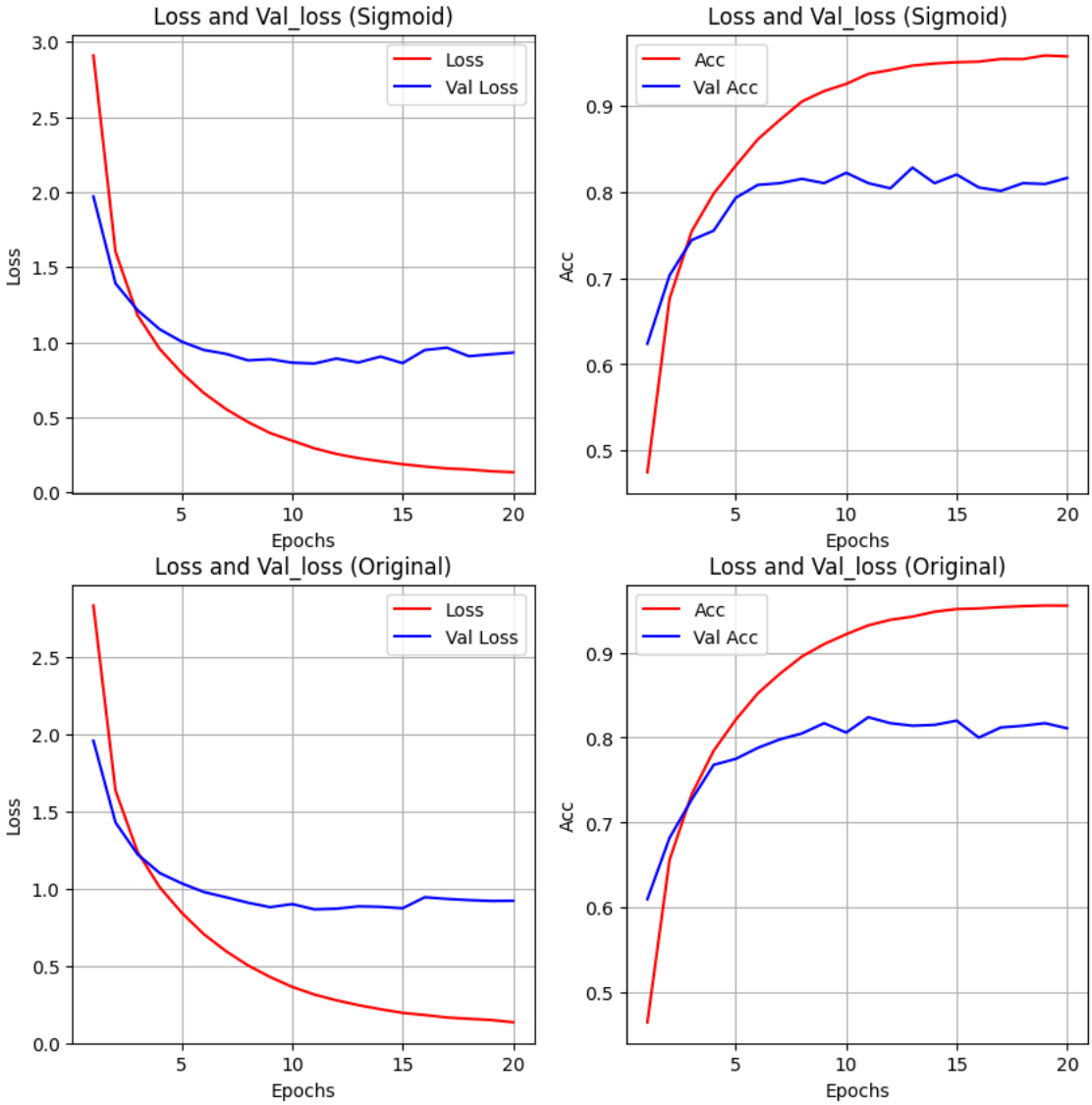
เมื่อลดจำนวน Neuron ใน Hidden Layers ทั้ง 2 ชั้น ปรากฏว่า อัตราการเรียนรู้ และอัตราการเกิด Overfitting ของ Model มีแนวโน้มลดลง สังเกตจากกราฟเส้นสีแดง ทียังคงมีความโค้งสูงอยู่ แสดงถึงความชันที่ยังคงมีอยู่ในระดับหนึ่ง ซึ่งต่างกับโมเดลที่มี Neuron อยู่ 64 ตัว ที่จะส่งผลให้โมเดลสามารถจดจำรายละเอียดของข้อมูลที่นำเข้าไปเทรนได้อย่างรวดเร็ว ส่งผลให้อัตราการเรียนรู้ และอัตราการเกิด Overfitting สูงตามไปด้วย

5. ลองลด num_words จาก 10000 ⟶ 5000 ผลเปลี่ยนอย่างไร

<u>ตอบ</u> num_words คือคำที่พบบ่อยสุดจากข่าว เมื่อลดจำนวนข้อมูลลง แต่ Neuron เท่าเดิม อาจทำให้เกิดการ Overfitting มากขึ้นได้

6. ถ้าใช้ sigmoid ร่วมกับ categorical_crossentropy ในชั้น output layer จะเกิดอะไรขึ้น

<u>ตอบ</u>



กราฟ Loss จากโมเดลที่ใช้ Sigmoid มีการแกว่งมากขึ้นเล็กน้อย