

- การเพิ่มจำนวน Layer สามารถช่วยให้โมเดลสามารถเรียนรู้ฟีเจอร์ที่ซับซ้อนขึ้นได้ แต่ก็มักทำให้การฝึกฝนใช้เวลานานขึ้น และอาจทำให้โมเดลเกิด overfitting
- การเพิ่มจำนวน node ในแต่ละ Layer สามารถช่วยให้โมเดลเรียนรู้ข้อมูลได้ละเอียดขึ้น แต่จะเพิ่มความซับซ้อนและอาจทำให้การฝึกฝนใช้เวลานานขึ้น
- Learning Rate ที่สูงเกินไปอาจทำให้โมเดลไม่สามารถหาค่าพารามิเตอร์ที่เหมาะสมได้ (overshooting) ขณะที่ Learning Rate ที่ต่ำเกินไปอาจทำให้การฝึกฝนช้าและต้องใช้จำนวน epoch มากขึ้น
- Activation Function ที่แตกต่างกันสามารถมีผลต่อการเรียนรู้และประสิทธิภาพของโมเดล เช่น ReLU มักใช้เนื่องจากมันช่วยลดปัญหา vanishing gradients, แต่บางครั้งการใช้ Sigmoid หรือ Tanh อาจดีกว่าในบางกรณี

layers': 1, 'neurons': 128, 'learning\_rate': 0.01,  
'activation': 'relu'

Accuracy: 0.97500000238418579

Training Time: 1.8493187427520752 seconds