A screenshot of a computer program

Description automatically generatedA graph of a line graph

Description automatically generated with medium confidence

Nhận xét:

* Khi tốc độ học nhỏ (0.001), Gradient Descent tiến rất chậm và cần nhiều bước lặp để hội tụ.
* Với tốc độ học vừa phải (0.1), Gradient Descent hội tụ nhanh hơn.
* Với tốc độ học quá lớn (1.0), Gradient Descent có thể dao động và khó hội tụ, hoặc thậm chí không hội tụ.