Lý thuyết

### ****1. Mô hình với Focal Loss****

**Mục tiêu của Focal Loss:**  
Focal Loss được thiết kế nhằm giải quyết bài toán mất cân bằng lớp (class imbalance) trong quá trình huấn luyện. Nó làm giảm trọng số của các mẫu dễ phân loại và tập trung mạnh hơn vào các mẫu khó, từ đó cải thiện khả năng học của mô hình đối với các mẫu hiếm hoặc khó nhận diện.

**Cách hoạt động:**

Công thức Focal Loss thêm một hệ số (1−pt)γ(1 - p\_t)^\gamma(1−pt​)γ vào hàm binary crossentropy, trong đó ptp\_tpt​ là xác suất dự đoán đúng của mẫu.

Khi mô hình đã dự đoán tốt (tức ptp\_tpt​ cao), hệ số này trở nên nhỏ, giảm tầm quan trọng của mẫu đó. Ngược lại, nếu ptp\_tpt​ thấp (mẫu khó), hệ số sẽ lớn, giúp mô hình chú ý nhiều hơn đến mẫu đó.

Các tham số như α\alphaα (trọng số cho các lớp) và γ\gammaγ (điều chỉnh mức độ tập trung) có thể được điều chỉnh để phù hợp với tính chất của dữ liệu.

**Ưu điểm:**

Giúp cải thiện hiệu suất trên tập dữ liệu mất cân bằng.

Giảm thiểu tác động của các mẫu dễ và cho mô hình tập trung vào học các mẫu khó.

**Nhược điểm:**

Có thể cần thử nghiệm với các giá trị của α\alphaα và γ\gammaγ để tìm được cấu hình tốt nhất cho dữ liệu cụ thể.

Đôi khi có thể làm cho quá trình huấn luyện chậm lại nếu hệ số focal quá mạnh.

### ****2. Mô hình với Dice Loss****

**Mục tiêu của Dice Loss:**  
Dice Loss ban đầu được phát triển cho các bài toán phân đoạn hình ảnh, nhằm đo lường sự trùng khớp giữa vùng dự đoán và vùng thật sự (ground truth). Tuy nhiên, nó cũng được áp dụng cho các bài toán phân loại nhị phân khi dữ liệu bị mất cân bằng.

**Cách hoạt động:**

Dice Loss dựa trên chỉ số Dice, một chỉ số đo lường độ tương đồng giữa hai tập hợp.

Công thức Dice Loss giúp tối đa hóa giao nhau giữa dự đoán và nhãn thực (true label) đồng thời giảm thiểu tổng của các giá trị dự đoán và giá trị thực.

Hàm loss này có tính chất “nhạy” với sự mất cân bằng giữa các lớp vì nó đo lường trực tiếp tỷ lệ trùng khớp giữa các mẫu.

**Ưu điểm:**

Tốt cho các bài toán có dữ liệu mất cân bằng, bởi vì nó không chỉ dựa vào từng mẫu riêng lẻ mà còn quan tâm đến tổng thể vùng dự đoán.

Giúp cải thiện độ chính xác trên các bài toán mà mẫu dương (positive) chiếm tỉ lệ nhỏ.

**Nhược điểm:**

Dice Loss thường được sử dụng trong bài toán phân đoạn nên đôi khi việc áp dụng trực tiếp vào bài toán phân loại có thể cần hiệu chỉnh thêm.

Có thể nhạy cảm với các giá trị nhỏ, do đó cần sử dụng một giá trị “smooth” để tránh chia cho 0.

**Công năng**

### ****Focal Loss****

**Tập trung vào mẫu khó:**

Tự động giảm trọng số cho các mẫu dễ phân loại (đã dự đoán đúng với độ tin cậy cao) và tăng trọng số cho các mẫu khó.

Giúp mô hình học tập trung hơn vào các trường hợp sai lệch hoặc hiếm gặp, từ đó cải thiện khả năng phân biệt.

**Ứng dụng trong nhiều bài toán mất cân bằng:**

Thích hợp cho các bài toán như nhận dạng vật thể (object detection), phân loại đa lớp với số lượng mẫu khác nhau giữa các lớp.

Có thể mở rộng cho các bài toán đa lớp (multi-class) với các biến thể như "Multi-class Focal Loss".

**Khả năng thích nghi cao:**

Tham số γ\gammaγ giúp điều chỉnh mức độ “tập trung” vào các mẫu khó.

Tham số α\alphaα cho phép cân bằng giữa các lớp nếu tỷ lệ mẫu không đồng đều, tạo ra sự linh hoạt khi làm việc với nhiều loại dữ liệu khác nhau.

**Giảm ảnh hưởng của noise:**

Vì các mẫu có dự đoán chính xác (mà có thể do noise hoặc dễ dự đoán) có trọng số nhỏ, nên noise trong dữ liệu ít tác động đến quá trình huấn luyện.

### ****Dice Loss****

**Tối ưu trực tiếp theo chỉ số Dice:**

Dice Loss tối ưu hóa trực tiếp tỷ lệ giao nhau giữa dự đoán và nhãn thực (Dice coefficient), một chỉ số thường được sử dụng để đánh giá chất lượng phân đoạn hoặc xác định vùng có vấn đề.

Điều này giúp mô hình tập trung cải thiện chỉ số chính được sử dụng trong đánh giá (ví dụ như trong segmentation).

**Hiệu quả với các vùng dương nhỏ:**

Trong các bài toán nơi mẫu dương (ví dụ: vùng ung thư, vết tổn thương) chiếm tỉ lệ rất nhỏ so với nền, Dice Loss có khả năng “nhạy” hơn so với các hàm loss truyền thống như binary crossentropy.

Nó giúp mô hình nhận diện tốt hơn những chi tiết nhỏ, vốn có thể bị lãng quên khi sử dụng các loss khác.

**Ổn định trong trường hợp có biến động dữ liệu:**

Thông qua việc sử dụng giá trị “smooth” (một hằng số nhỏ để tránh chia cho 0), Dice Loss đảm bảo quá trình tính toán luôn ổn định, ngay cả khi số lượng pixel dương rất thấp.

Điều này giúp tránh các biến động lớn trong giá trị loss trong quá trình huấn luyện.

**Phù hợp với các bài toán phân đoạn:**

Ngoài ứng dụng trong phân loại nhị phân, Dice Loss thường được sử dụng trong bài toán phân đoạn hình ảnh (ví dụ: segmentation của tế bào, mô hình y tế) để tối ưu hóa kết quả phân đoạn, giúp cải thiện độ chính xác của vùng được dự đoán.