**Data Augmentation** là một kỹ thuật mở rộng tập dữ liệu bằng cách tạo ra các phiên bản mới từ dữ liệu hiện có. Mục tiêu của việc này là tăng cường độ đa dạng của tập dữ liệu, giúp mô hình học máy học được các biến động và biến đổi trong dữ liệu, từ đó cải thiện khả năng tổng quát hóa của mô hình. Data Augmentation thường được sử dụng trong các bài toán học máy, đặc biệt là khi dữ liệu huấn luyện có lượng dữ liệu hạn chế.

Dưới đây là một số phương pháp Data Augmentation và ví dụ đi kèm:

**1. Rotation (Xoay ảnh):**

* **Mô tả:** Xoay ảnh một góc ngẫu nhiên để tạo ra các phiên bản mới.
* **Ví dụ:**
  + Original Image:
  + Augmented Image:

**2. Flip (Lật ảnh):**

* **Mô tả:** Lật ảnh theo chiều ngang hoặc chiều dọc để tạo ra các phiên bản mới.
* **Ví dụ:**
  + Original Image:
  + Augmented Image:

**3. Zoom (Phóng to/Thu nhỏ):**

* **Mô tả:** Phóng to hoặc thu nhỏ một phần của ảnh để tạo ra sự đa dạng.
* **Ví dụ:**
  + Original Image:
  + Augmented Image:

**4. Brightness and Contrast (Độ sáng và độ tương phản):**

* **Mô tả:** Điều chỉnh độ sáng và độ tương phản của ảnh.
* **Ví dụ:**
  + Original Image:
  + Augmented Image:

**5. Crop (Cắt ảnh):**

* **Mô tả:** Cắt một phần của ảnh để tạo ra các phần ảnh mới.
* **Ví dụ:**
  + Original Image:
  + Augmented Image:

**6. Color Jittering (Làm mờ màu):**

* **Mô tả:** Thay đổi một chút giá trị màu sắc của pixel để tạo ra sự đa dạng.
* **Ví dụ:**
  + Original Image:
  + Augmented Image:

**Lưu ý:**

* Data Augmentation không thích hợp cho mọi loại dữ liệu. Đối với một số bài toán như nhận dạng khuôn mặt, việc xoay ảnh có thể làm mất đi ý nghĩa của dữ liệu.
* Khi áp dụng Data Augmentation, bạn cần đảm bảo rằng ý nghĩa của dữ liệu không bị thay đổi quá mức.
* Các thư viện như TensorFlow và Keras cung cấp các công cụ hỗ trợ Data Augmentation trong quá trình xây dựng mô hình.

Top of Form