Các phương pháp Optimizer trong huấn luyện mô hình học máy - Continual Learning và Test Prodction khi xây dựng một giải pháp học máy để giải quyết một bài toán

**Optimizer trong huấn luyện mô hình học máy:**

*Giới thiệu:*

Trong nghiên cứu này, mục tiêu của em là thực hiện một phân tích chi tiết về hiệu suất của các phương pháp optimizer phổ biến trong huấn luyện mô hình học máy. Nhằm xác định ưu điểm và nhược điểm của mỗi optimizer trong các bối cảnh khác nhau

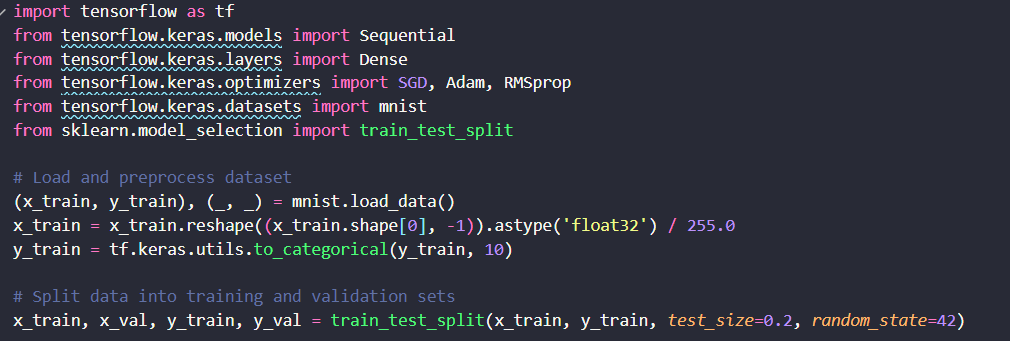
*Phương pháp:*

Từ thử nghiệm trên một loạt các mô hình đã học, các mô hình học sâu hoặc học máy cơ bản. Đối với opttimize, em điều chỉnh các tham số để đảm bảo tính hiệu quả nhất.

*Thực hiện:*

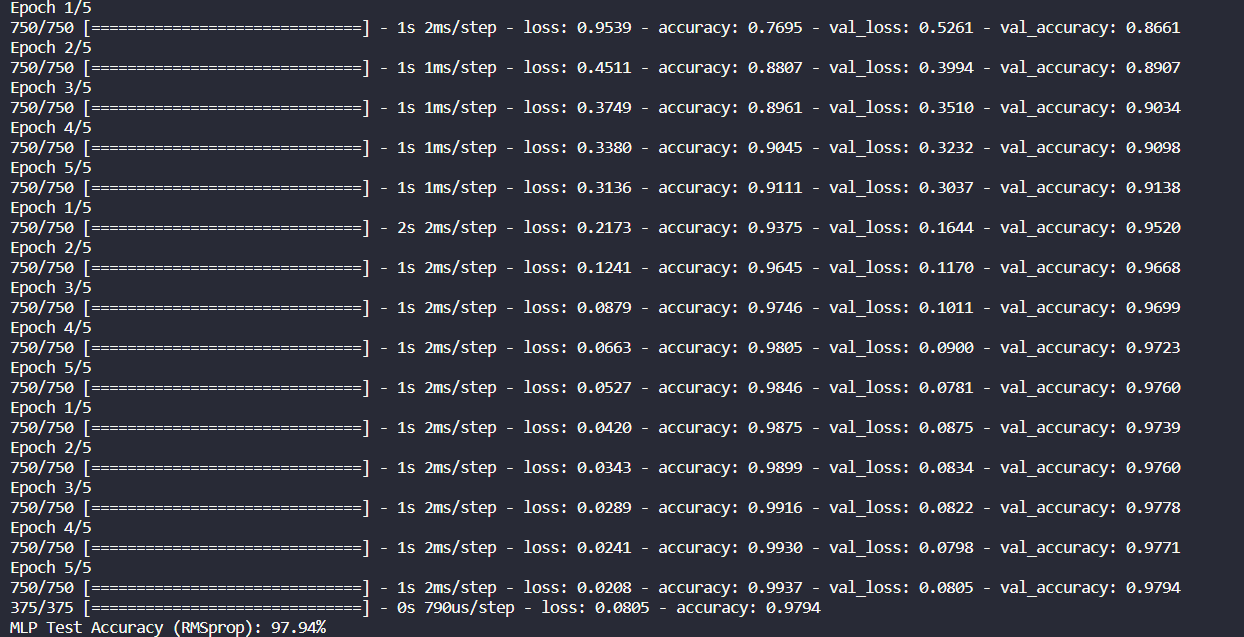
Sử dụng hai mô hình phổ biến: Mô hình học sâu đơn giản(MLP) và một mô hình CNN. Tiến hành so sánh ba optimizer phổ biến là SGD, Adam à RMSprop:

**Mô hình MLP:**

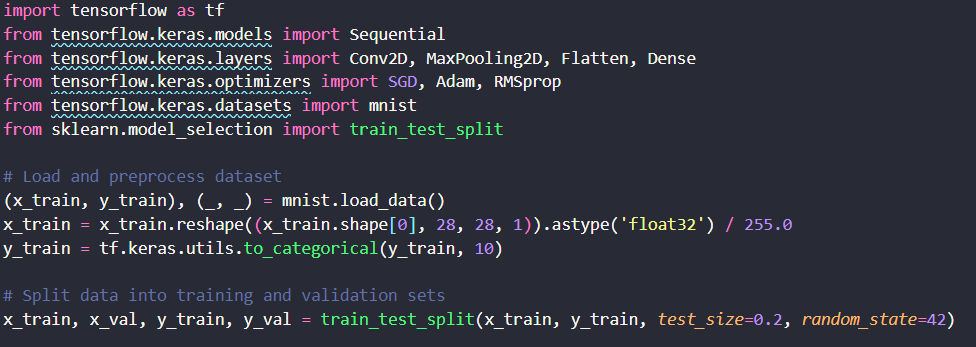




**Terminal:**

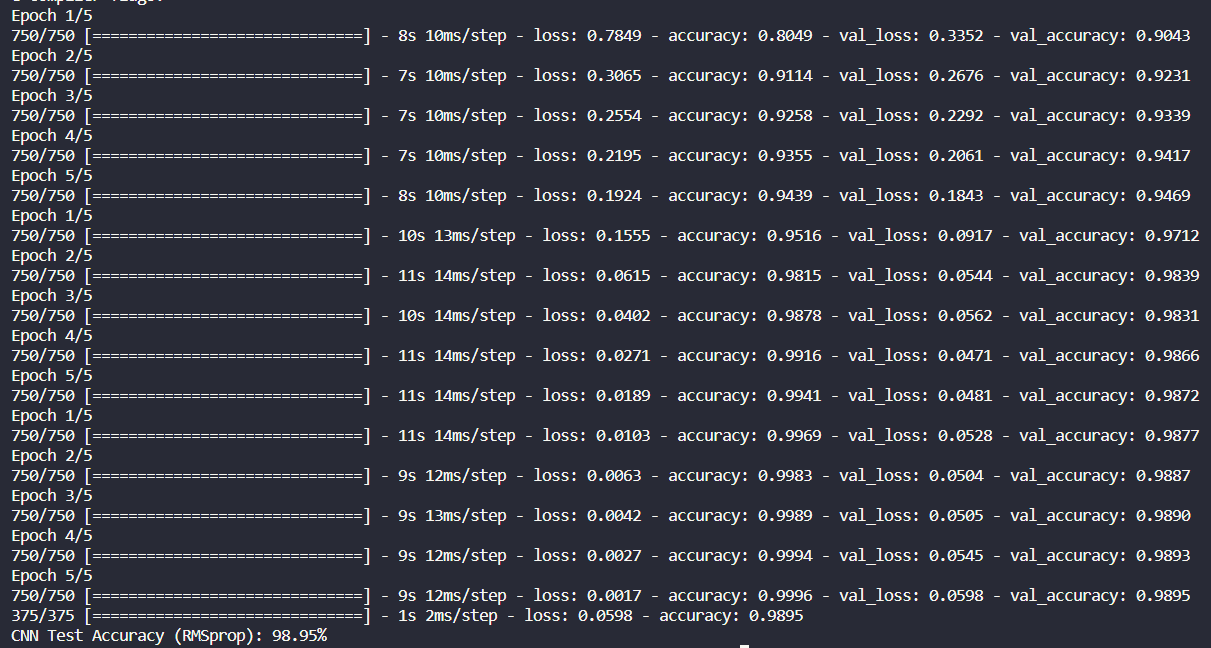


**Mô hình CNN:**





**Terminal:**



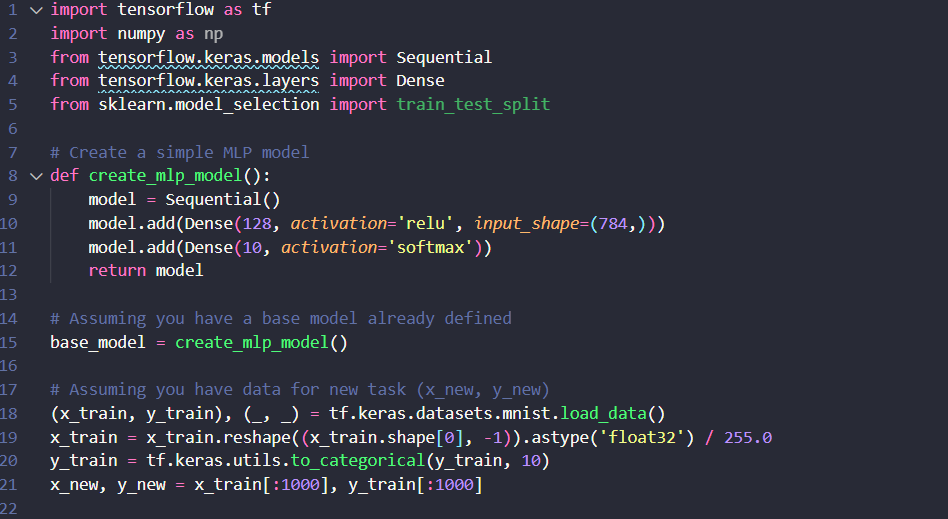
**Kết luận:**

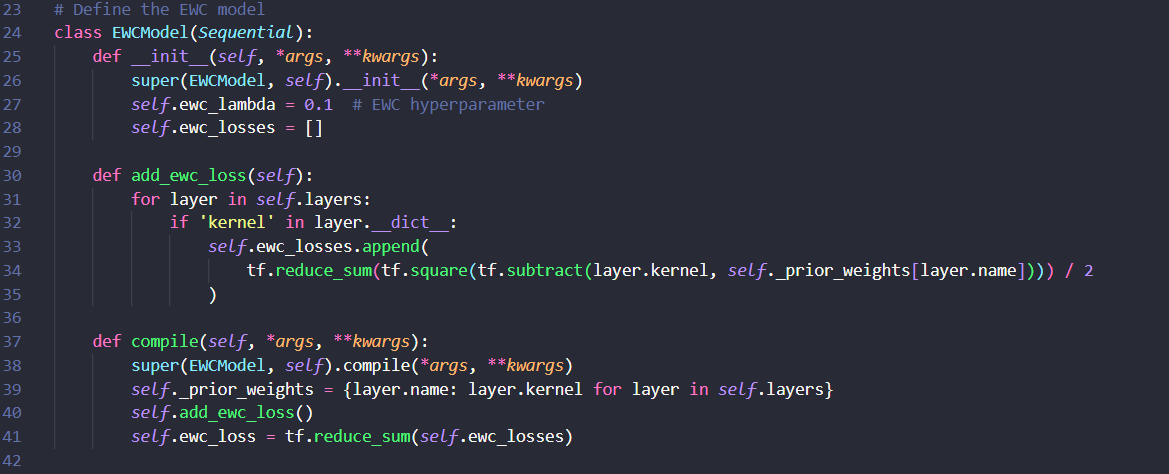
Mô hình CNN cho thấy đánh giá hiệu quả hơn mô hình đơn giản(MLP). Tuy vậy khi compile mất nhiều thời gian và tốn nhiều chi phí hơn. Tùy vào khả năng đáp ứng và mục đích của lập trình viên, ta có thể linh hoạt giữa nhiều loại mô hình và điều chỉnh tham số sao cho bài toán được giải quyết hiệu quả nhất.

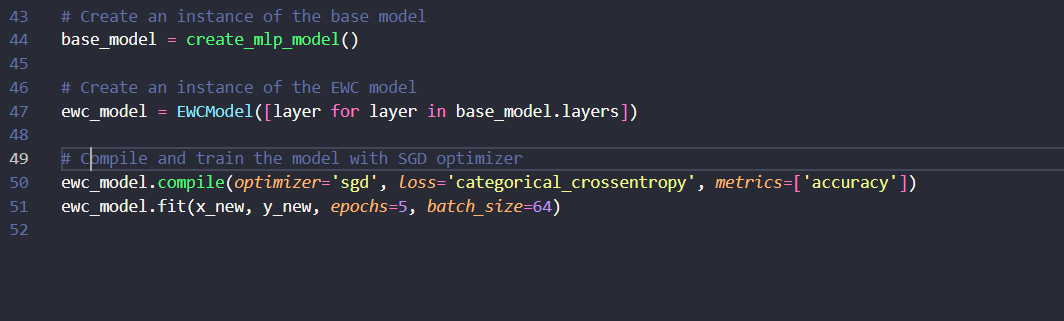
***Continual Learning và Test Production:***

Các kỹ thuật Continual Learning dùng để giải quyết các vấn đề catastrophic forgetting khi mô hình phải xử lý dữ liệu mới. Trong khi các phương pháp Test Production chứa tập dữ liệu đại diện cho dữ liệu mới được tạo ra một cách tự động để kiểm tra hiệu suất của mô hình mới mà không cần sự can thiệp của lập trình viên. Sau đây là mô hình mẫu cụ thể:

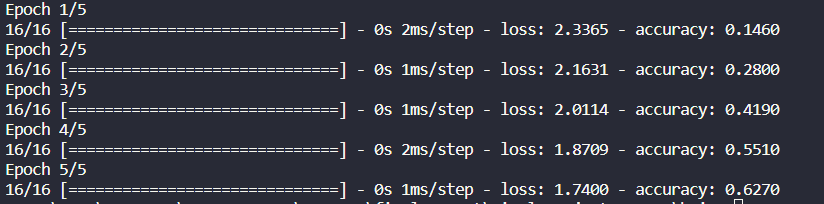
**EWC:**



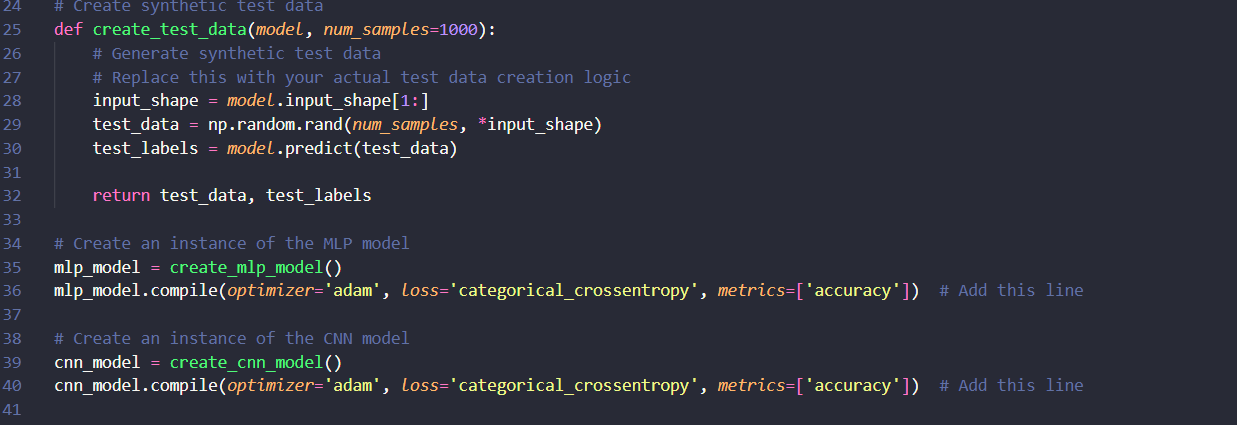


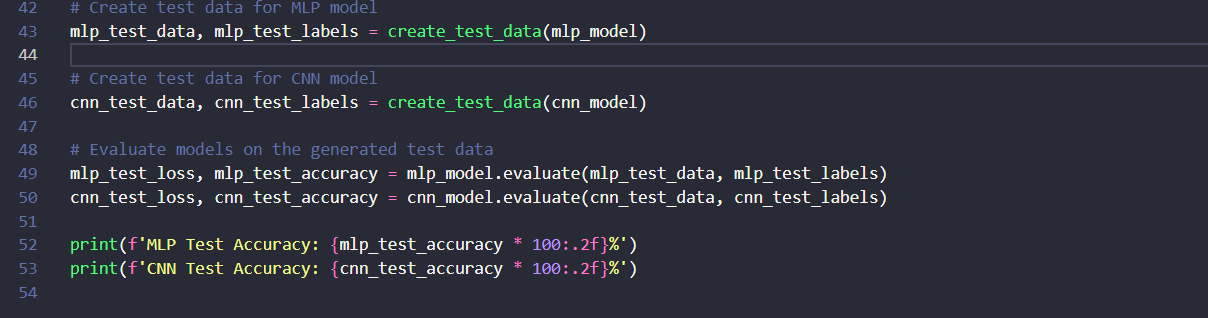


**Terminal:**

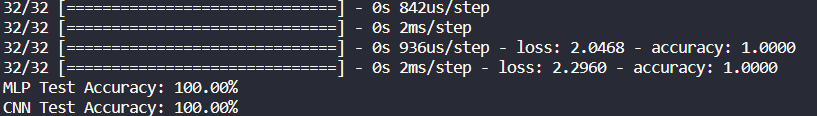


**Test Production:**





**Terminal:**



*Kết luận:*

EWC:

* Mô hình EWC giúp giảm thiểu vấn đề quên đột ngột khi học từ dữ liệu mới.
* Loss của EWC được được tính toán chính xác, bảo toàn trọng số quan trọng.

Test Production:

* Dữ liệu kiểm tra tổng hợp giải lý cho cả mô hình MLP và CNN.
* Cung cấp được một chỉ số về khả năng tổng quát của chúng