**3.3.1. Giai đoạn 1: Xác định quỹ đạo bay ưa thích (user preferred route - UPR) dùng trí tuệ nhân tạo**

Trên hình [Hình. 2] trình bày quy trình dự doán quỹ đạo bay ưa thích (quỹ đạo AI) theo nhu cầu khai thác thực tiễn của người sử dụng vùng trời, chúng tôi đã xây dựng mô hình ~~học sâu~~ sử dụng mạng nơ-ron thần kinh nhiều lớp (DNN) để dự đoán quỹ đạo bay từ bộ dữ liệu tổng hợp (gồm dữ liệu ADS-B của các chuyến bay đã thực hiện thành công và bộ dữ liệu thời tiết đồng bộ về thời gian trong từng chuyến bay). Các dữ liệu ADS-B được truyền nhận giữa các tàu bay và các trạm mặt đất thông qua hệ thống thiết bị dẫn đường vệ tinh GPS/GNSS/GLONASS, chứa đựng thông thin về vị trí, độ cao, tốc độ, hướng bay, số hiệu tàu bay, mã hiệu chuyến bay, tình hình giao thông, thời tiết, bản tin NOTAM, v.v… Từ bộ dữ liệu ADS-B, mô hình mạng thần kinh đề xuất ~~học sâu~~ có thể dự đoán được đặc tính kỹ thuật vận hành của tàu bay trong quá trình thiết kế đường bay tối ưu đồng thời đảm bảo độ chính xác của quỹ đạo bay trong tầm phủ của hệ thống dẫn đường vệ tinh.

Mô hình

Quỹ đạo bay AI (UPR)

Dữ liệu ADS-B và dữ liệu thời tiết đã xử lý

Các biến đầu vào và đầu ra của mô hình

Dữ liệu huấn luyện

Dữ liệu đánh giá

Dữ liệu kiểm nghiệm

Huấn luyện

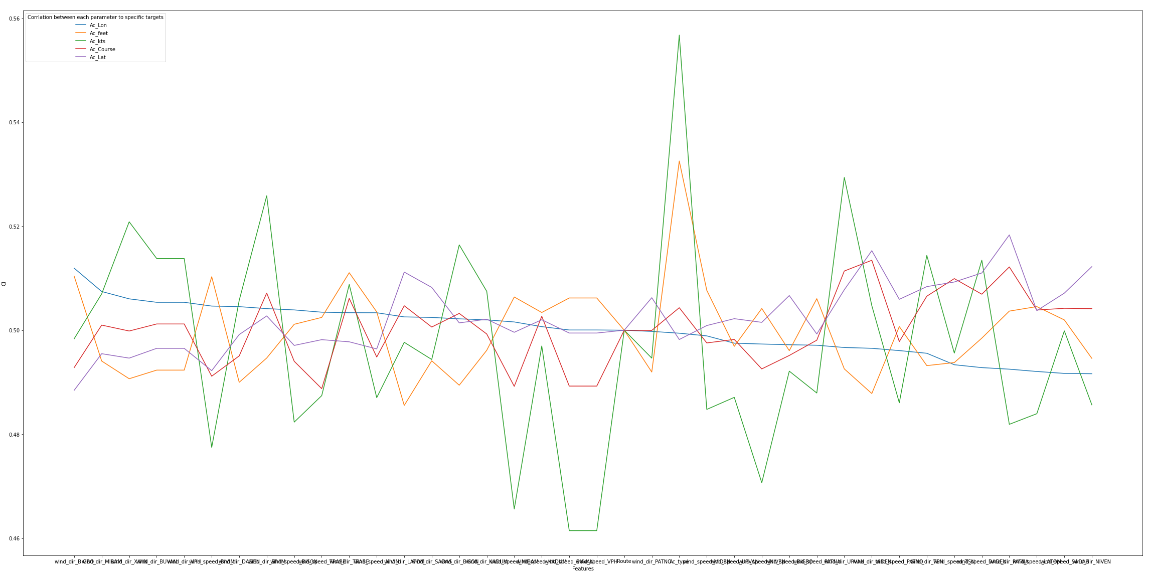
Đánh giá

Hình. 2. Thuật toán học máy

1. **Quản lý dữ liệu**

Quy trình quản lý dữ liệu bao gồm: Thu thập dữ liệu (có thể từ nguồn nội bộ hoặc bên ngoài); Chuẩn bị dữ liệu (tiền xử lý, chọn lọc tham số, chuẩn hóa dữ liệu); Phân chia dữ liệu cho quá trình huấn luyện (thường là các bộ dữ liệu huấn luyện, đánh giá và thử nghiệm); Đánh giá và xác minh bộ dữ liệu (bao gồm độ chính xác, tính đầy đủ, tính đại diện và tính độc lập); Xác định và loại bỏ độ chệch không mong muốn trong các tập dữ liệu.

* **Thu thập dữ liệu**
* *Dữ liệu ADS-B*: dữ liệu các chuyến bay đã thực hiện được thu thập từ trang web <https://flightaware.com/live/findflight>, cung cấp thông tin về mã nhận dạng tàu bay (id), loại tàu bay (at), sân bay đi – đến (od), vĩ độ (), kinh độ (), độ cao (), thời gian (), vận tốc (*v)*, hướng bay ().
* *Dữ liệu thời tiết:* Dữ liệu thời tiết thực được thu thập từ trang web <https://api.meteomatics.com/> và đồng bộ thời gian với từng điểm trên từng quỹ đạo bay, cung cấp thông tin về vận tốc gió, hướng gió, nhiệt độ, áp suất và tình trạng thời tiết (mưa, sương mù, mưa phùn, giông bão, bão cát, tuyết,…)
* **Chuẩn bị dữ liệu:**
* *Chọn lọc dữ liệu:* để phục vụ cho việc sử dụng kỹ thuật huấn luyện mạng neural thần kinh nhiều lớp ~~học sâu~~ có giám sát, bộ dữ liệu ADS-B và bộ dữ liệu thời tiết được kết hợp với nhau thành bộ dữ liệu tổng hợp theo mốc thời gian tương ứng.
* *Tiền xử lý dữ liệu:* các dữ liệu bị thiếu hoặc bị trùng lặp được loại bỏ khỏi các bộ dữ liệu.
* *Phân loại dữ liệu:* theo quy định về mực bay chuyển tiếp giữa phương thức bay đường dài và phương thức bay tiếp cận được công bố trong AIP Việt Nam, bộ dữ liệu tổng hợp trong nghiên cứu này được chia làm hai tập dữ liệu con lần lượt là tập dữ liệu đường dài (có độ cao trên 13,500 feet) và tập dữ liệu tiếp cận (có độ cao từ 13,500 feet trở xuống).
* *Đánh giá mức độ tác động của các biến đầu vào với biến ngõ ra*: sử dụng chỉ số phù hợp concordance index (c-index) để đánh giá mức độ mức ảnh hưởng của từng biến ngõ vào đối với mỗi biến ngõ ra [**Hình 3**]. Từ đó lựa loại bỏ bớt những biến ngõ vào ít ảnh hưởng đến biến ngõ ra để giảm nhiễu và tăng hiệu quả cho quá trình huấn luyện mô hình học sâu.
* *0=<c-index<0.5:*
* *0.5<c-index<=1:*
* *C-index=0.5:*



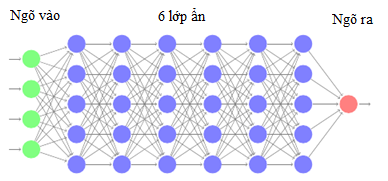
Hình 3. Mức độ tương quan giữa các thông số ngõ vào và thông số ngõ ra

1. **Lựa chọn mô hình và huấn luyện**

Kết quả đánh giá hiệu quả huấn luyện của một số mô hình học máy LR, RF, XG-BOOST, DNN được chúng tôi trình bày trong bài báo [14], cho thấy mô hình DNN là phù hợp nhất và có hiệu quả huấn luyện tốt nhất trong dự đoán quỹ đạo bay.

Vì thế, trong nghiên cứu này, chúng tôi chọn mô hình DNN và điều chỉnh các tham số, tối ưu hóa cấu trúc của mô hình, sử dụng kỹ thuật multitask learning để gia tăng hiệu quả huấn luyện và tăng độ chính xác của mô hình DNN trong dự đoán quỹ đạo bay.

Mô hình DNN học các đặc trưng từ bộ dữ liệu ADS-B và dữ liệu thời tiết tương ứng, có cấu trúc với 1 ngõ vào, 6 lớp ẩn, 1 ngõ ra [**Hình 4**], sử dụng thuật toán tối ưu Adam, hàm kích hoạt ReLU (đơn vị đo tuyến tính được hiệu chỉnh) ở các lớp ẩn, hàm kích hoạt tuyến tính ở ngõ ra. Các tham số huấn luyện: learning rate = 0.001, batch\_size= 3200, epoch = 600 à validation\_split = 0.3.



Hình 4. Mô hình mạng nơ-ron sâu (DNN)

Mô hình DNN được xây dựng trên nền tảng Tensorflow phiên bản 2.1.0 là một framework rất phổ biến dùng cho các nghiên cứu về học sâu (deep learning), học máy (machine learning) nổi tiếng và được sử dụng rộng rãi hiện nay.

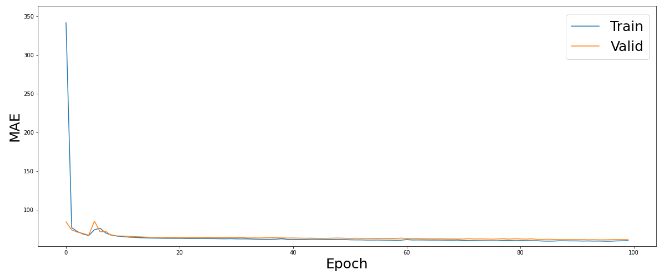
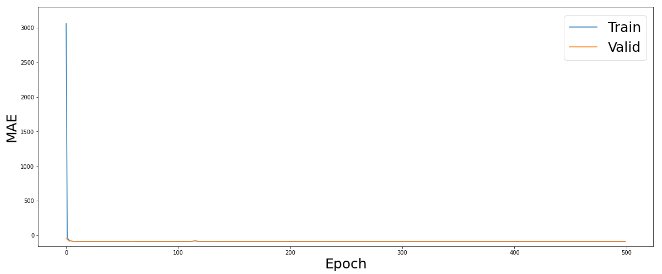
Quá trình phân tích, xử lý dữ liệu, huấn luyện và đánh giá hiệu quả mô hình DNN được thực hiện bằng ngôn ngữ lập trình Python và trên máy vi tính sử dụng bộ xử lý các tác vụ có liên quan tới đồ họa GTX 3090, RAM 48GB, vi xử lý i7.

Thời gian huấn luyện:

Thời gian dự đoán cho một mẫu:

1. **Kết quả huấn luyện mô hình mạng nơ-ron sâu DNN**

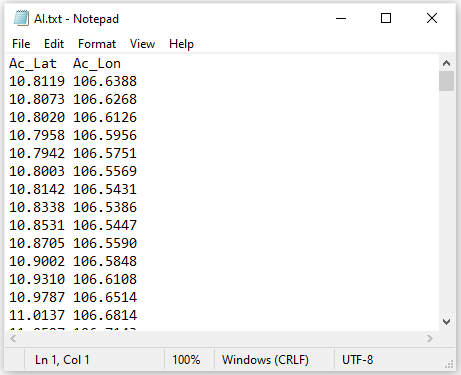
Các mô hình học sâu được sử dụng trong đề tài này cho thấy khả năng học tốt từ các bộ dữ liệu nêu trên với các chỉ số xác định lớn R2= 0.9425 (dự đoán vĩ độ - latitude) và R2= 0.8110 (dự đoán kinh độ - longitude) và giá trị sai số tuyệt đối trung bình (MAE) bé [**Hình 5**].



*а) khi dự đoán vĩ độ - latitude б) khi dự đoán kinh độ - longitude*

Hình 5. Sai số tuyệt đối trung bình (MAE)

Tọa độ của quỹ đạo bay UPR nhận được từ quá trình huấn luyện mô hình mạng nơ-ron sâu (DNN) được lưu dưới dạng file văn bản .txt [**Hình 6**]. Và trong giai đoạn 2, quỹ đạo UPR này sẽ được sử dụng làm tham chiếu cho thuật toán A\* để giải bài toán tối ưu hóa quỹ đạo bay.



Hình 6. Tọa độ địa lý của quỹ đạo bay UPR nhận được từ kết quả học máy