**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

**---🙠**🕮**🙢---**



**SEMINAR 1**

**Đề tài: Xây dựng phần mềm dự báo bức xạ mặt trời sử dụng mô hình LSTM dựa trên phương pháp deep learning**

|  |  |  |
| --- | --- | --- |
| Giảng viên hướng dẫn | : | TS. Nguyễn Đức Tuyên |
| Họ và tên sinh viên | : | Đỗ Huy Khang |
| MSSV | : | 20212027M |
| Khoa/Viện | : | Kĩ thuật điện |

**Hà Nội, 2022**

**LỜI MỞ ĐẦU**

Ngày nay, sự phát triển của nền kinh tế và dân số thế giới kéo theo sự gia tăng nhanh chóng về nhu cầu năng lượng cho sự phát triển kinh tế và xã hội. Do đó, năng lượng đang trở thành một vấn đề cấp thiết. Nguồn năng lượng hóa thạch ngày càng cạn kiệt, ô nhiễm môi trường ngày càng nghiêm trọng. Nhu cầu tìm kiếm năng lượng mới, xanh và tái tạo thay thế các nguồn năng lượng hóa thạch truyền thống, đã là một vấn đề lâu dài đối với cả các nước phát triển và đang phát triển. Việt Nam cũng cần vượt qua khó khăn đó khi chúng ta thiếu năng lượng trong khi các nguồn năng lượng truyền thống dần dần không đủ để sản xuất. Mặt khác, Việt Nam là nước nhiệt đới, có lợi thế về vị trí địa lý, tiềm năng khai thác năng lượng mặt trời rất lớn. Tuy nhiên hiện tại, đa phần các hệ thống pin quang điện (PV), nguồn sản suất điện mặt trời lớn nhất đang tồn tại nhiều nhược điểm, nhất là sự ổn định trong công suất phát. Vì vậy việc dự báo về công suất phát của hệ thống PV cũng là một vấn đề cần giải quyết để đảm bảo hiệu suất của hệ thống điện, giúp hệ thống trở nên tối ưu hơn, từ đó nâng cao hiệu quả về kinh tế. Xuất phát từ yêu cầu thực tế và được sự đồng ý của giáo viên hướng dẫn, em làm seminar này để nghiên cứu về mô hình dự báo bức xạ mặt trời dựa trên phương pháp deep learning.

**MỤC LỤC**

[I. Đặt vấn đề 5](#_Toc105623976)

[1. Tổng quan về năng lương mặt trời (NLMT) 5](#_Toc105623977)

[2. Lí do chọn đề tài 8](#_Toc105623978)

[II. Phương pháp luận 11](#_Toc105623979)

[1. Mô hình Long Short-term Memory (LSTM) 11](#_Toc105623980)

[2. Mạng nơ ron hồi quy (RNN) 13](#_Toc105623981)

[3. Mạng nơ-ron nhân tạo (ANN) 14](#_Toc105623982)

[4. Các sai số đánh giá 15](#_Toc105623983)

[III. Đánh giá kết quả và thảo luận 16](#_Toc105623984)

[1. Thu thập và tiền xử lý dữ liệu 16](#_Toc105623985)

[2. Kết quả và thảo luận 17](#_Toc105623986)

[IV. Xây dựng giao diện phần mềm trên máy tính 20](#_Toc105623987)

[V. Kết luận 26](#_Toc105623988)

**Mục lục hình ảnh**

[Hình 1. Tổng số giờ nắng tại một số trạm quan trắc 6](#_Toc105624018)

[Hình 2. Bản đồ cường độ bức xạ mặt trời trung bình tại Việt Nam 2007-2018 (GHI) 7](#_Toc105624019)

[Hình 3. Tỷ lệ phần trăm nguồn năng lượng trong cơ cấu năng lượng Việt Nam năm 2019 8](#_Toc105624020)

[Hình 4. Cấu trúc của một tế bào LSTM 12](#_Toc105624021)

[Hình 5. Cấu trúc của mạng LSTM 13](#_Toc105624022)

[Hình 6. Cấu trúc mạng nơ-ron hồi quy (RNN) rút gọn (trái) và đầy đủ (phải) 13](#_Toc105624023)

[Hình 7. Mô hình mạng nơ ron đơn giản tương đương một mô hình hồi quy tuyến tính 14](#_Toc105624024)

[Hình 8. Mô hình mạng nơ ron với 4 biến đầu vào và 3 nơ ron trong lớp ẩn 14](#_Toc105624025)

[Hình 9. Tập dữ liệu sau khi được làm sạch 16](#_Toc105624026)

[Hình 10. Kết quả dự báo của các ngày trong tháng 6 năm 2019 17](#_Toc105624027)

[Hình 11. Kết quả dự báo giữa các mô hình đề xuất so với giá trị thực tế vào 1 ngày của tháng 6 18](#_Toc105624028)

[Hình 12. Kết quả dự báo của các ngày trong tháng 12 năm 2019 19](#_Toc105624029)

[Hình 13. Kết quả dự báo giữa các mô hình đề xuất so với giá trị thực tế vào 1 ngày của tháng 12 20](#_Toc105624030)

**Mục lục bảng**

[Bảng 1. Loại dự báo cường độ bức xạ mặt trời và mục đích 10](#_Toc105624055)

# Đặt vấn đề

## Tổng quan về năng lương mặt trời (NLMT)

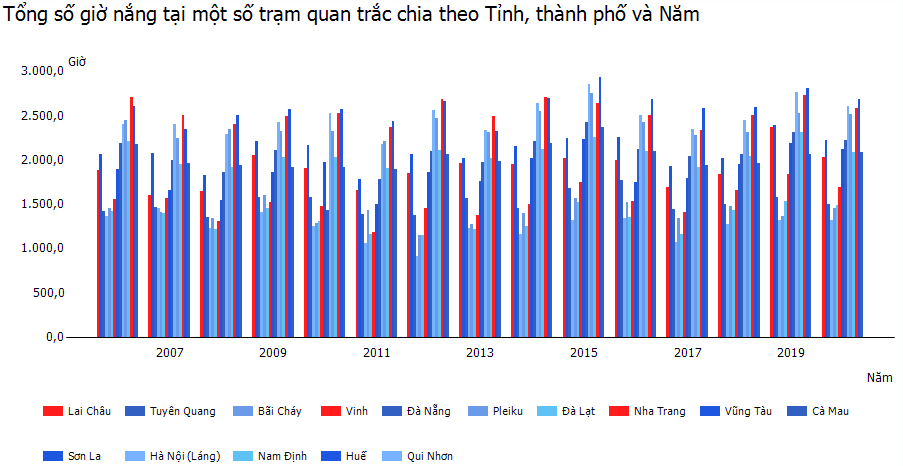
Năng lượng mặt trời là nguồn năng lượng sạch, không gây ô nhiễm môi trường và có trữ lượng vô cùng lớn do có khả năng tái tạo cao. Đồng thời, phát triển công nghiệp sản xuất pin mặt trời sẽ góp phần thay thế các nguồn năng lượng hóa thạch, giảm phát thải khí nhà kính, bảo vệ môi trường. Vì vậy, đây được coi là nguồn năng lượng quý giá, có thể thay thế các dạng năng lượng cũ đang ngày càng cạn kiệt.

Các ưu điểm nổi bật của năng lượng mặt trời:

* Điện mặt trời không gây ô nhiễm và không thải ra khí nhà kính.
* Giảm sự phụ thuộc vào các nguyên liệu hóa thạch.
* Là nguồn điện tái tạo được cung cấp đều đặn hàng ngày trong năm, kể cả những ngày nhiều mây.
* Hệ thống rất it hoặc hầu như không cần bảo trì vì các tấm pin mặt trời có tuổi thọ hơn khá cao.
* Lượng điện dư thừa có thể bán lại cho Điện lực nếu có kết nối với lưới.
* Hệ thống được xây dựng nhanh và đơn giản hơn nhiều so với các loại điện khác.
* Vận hành an toàn hơn so với các nhà máy thủy điện, nhiệt điện hay nhà máy hạt nhân.
* Hiệu suất được cải thiện qua thời gian, không ngừng được nâng cao nhờ các tiến bộ khoa học kĩ thuật.

Tiềm năng phát triển năng lượng mặt trời của Việt Nam là vô cùng lớn, đặc biệt là khu vực miền Trung và miền Nam với cường độ bức xạ mặt trời trung bình khoảng 5 kWh / m2, trong khi ở miền Bắc xấp xỉ 4 kWh / m2. Bức xạ mặt trời trung bình 230-250 kcal / cm2 theo hướng tăng dần về phía Nam chiếm khoảng 2.000 - 3.000 giờ mỗi năm, với tiềm năng lý thuyết ước tính khoảng 43,9 tỷ TOE. Năng lượng mặt trời ở Việt Nam có hầu như quanh năm, khá ổn định và phân bố rộng khắp các vùng miền trên cả nước. Đặc biệt, số ngày nắng trung bình của các tỉnh miền Trung và miền Nam khoảng 300 ngày / năm.

Việt Nam có tiềm năng rất lớn trong việc phát triển điện mặt trời với số giờ nắng trung bình trong năm khá cao được thể hiện trong Hình 1.

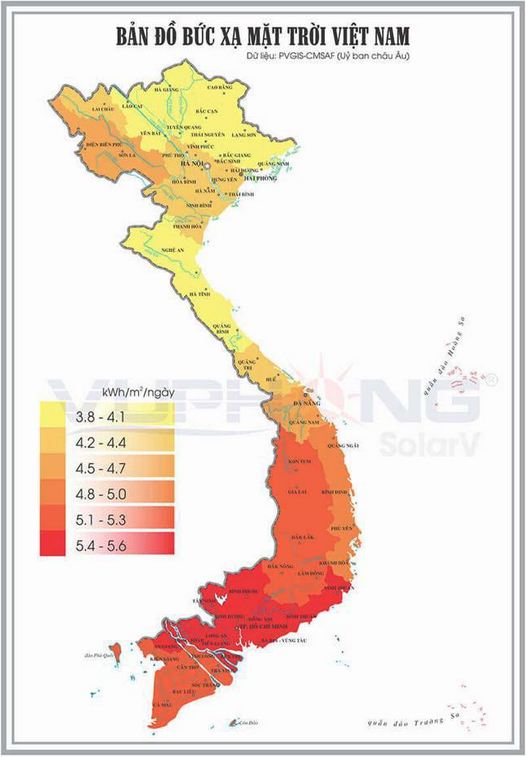


Hình 1. Tổng số giờ nắng tại một số trạm quan trắc

(Nguồn: [https://www.gso.gov.vn](https://www.gso.gov.vn/))

Dựa theo dữ liệu của Tổng cục Thống kê về số giờ nắng trung bình trong 13 năm liên tiếp từ 2016 đến 2019, có thể thấy tiềm năng [điện mặt trời](https://dienmattroihoaluoi.vuphong.vn/dien-nang-luong-mat-troi-hoa-luoi.html) là rất lớn từ Thừa Thiên Huế trở vào Nam và vùng Tây Bắc. Vùng Tây Bắc và vùng Bắc Trung có lượng bức xạ mặt trời khá lớn. Số giờ nắng trung bình cả năm nằm trong khoảng 1800 - 2100 giờ. Từ dưới vĩ tuyến 17, cường độ bức xạ mặt trời vừa cao vừa ổn định trong suốt cả năm, vào mùa khô cao hơn mùa mưa khoảng 20%. Trừ những ngày có mưa rào thì trên 90% số ngày trong năm đều có thể sử dụng năng lượng mặt trời cho sản xuất điện. Số giờ nắng trung bình cả năm nằm trong khoảng 2000 - 2600 giờ. Đây là khu vực ứng dụng năng lượng mặt trời rất hiệu quả.

Bên cạnh số giờ nắng trung bình năm, Việt Nam cũng có thông số về cường độ bức xạ trung bình năm rất phù hợp với định hướng phát triển điện mặt trời. Cường độ bức xạ trung bình trong năm ở các tỉnh và thành phố phía Bắc là 3,8-5,0 kWh/m2 và phía Nam là 4,8-5,6 kWh/m2 được thể hiện như trong Hình 2.



Hình 2. Bản đồ cường độ bức xạ mặt trời trung bình tại Việt Nam 2020 (GHI)

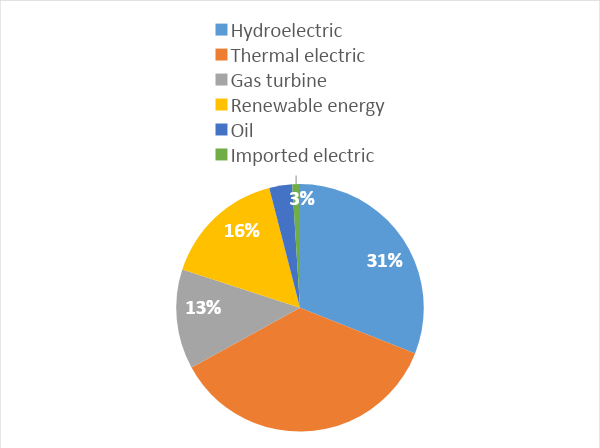
(Nguồn: <https://globalsolaratlas.info/map>)

Trước năm 2017, mặc dù có tiềm năng to lớn nhưng tình hình phát triển điện mặt trời nối lưới được thực hiện ở Việt Nam vẫn còn thấp hơn mong đợi. Tuy nhiên, kể từ khi Chính phủ ban hành Quyết định số 11/2017/QĐ-TTg ngày 11/4/2017 và gần đây nhất quyết định số 13/2020/QĐ-TTg ngày 6/4/2020 về cơ chế khuyến khích phát triển các dự án điện mặt trời tại Việt Nam cũng như Thông tư số 16/2017/TT-BCT quy định về phát triển dự án và hợp đồng mua bán điện mẫu áp dụng cho các dự án điện mặt trời đã khiến loại hình năng lượng tái tạo này trong vòng 3 năm đã phát triển một cách bùng nổ tại Việt Nam.

Theo QHĐ 8, Đến hết năm 2020, nguồn điện mặt trời nối lưới đã được đưa vào vận hành lên tới khoảng 9000 MW (trong đó tỉnh Ninh Thuận và Bình Thuận gần 3,5GW). Quy mô công suất của các dự án điện mặt trời đã được bổ sung quy hoạch là trên 13GW, tổng quy mô đăng ký xây dựng nhưng chưa được bổ sung khoảng 50GW, các con số này đã cho thấy bước phát triển nhảy vọt của điện mặt trời tại Việt Nam.

Bên cạnh các dự án điện mặt trời dạng trang trại (lắp đặt trên mặt đất, mặt nước), các dự án điện mặt trời mái nhà cũng phát triển với tốc độ rất nhanh. Đến hết năm 2019 công suất lắp đặt ĐMT áp mái toàn quốc mới đạt 340MWp (272MW), nhưng đến hết năm 2020 tổng công suất lắp đặt đạt tới 7780MW.

Tuy nhiên, theo Báo cáo Năng lượng tái tạo Việt Nam 2019, lượng năng lượng tái tạo được khai thác chỉ đạt 16% tổng nguồn điện, trong đó năng lượng sản xuất từ mặt trời chưa tương xứng với tiềm năng, chỉ chiếm tỷ lệ nhỏ 8,7%. Việt Nam đang có kế hoạch phấn đấu đạt 20% tổng sản lượng điện khai thác vào năm 2030. Về trung và dài hạn, nguồn năng lượng từ mặt trời sẽ được quan tâm nhiều nhất do ưu điểm ổn định, dễ vận hành và cấu trúc đơn giản của hệ thống.



Hình 3. Tỷ lệ phần trăm nguồn năng lượng trong cơ cấu năng lượng Việt Nam năm 2019

Cuối cùng có thể kết luận rằng, điện mặt trời ở Việt Nam đã, đang và sẽ đóng vai trò quan trọng trong cơ cấu nguồn điện của nước ta. Vai trò và tầm quan trọng của việc phát triển điện mặt trời yêu cầu các đơn vị và tổ chức nghiên cứu cần phải thích ứng, bắt kịp và tập trung nghiên cứu các vấn đề liên quan đề việc phát triển nguồn điện mặt trời.

## Lí do chọn đề tài

Trong nhiều lĩnh vực nghiên cứu khác nhau, dự báo là luôn là một bài toán quan trọng và phổ biến. Cụ thể trong lĩnh vực kỹ thuật, việc dự báo trước được một số dữ liệu thống kê trong tương lai là cơ sở để hoạch định kế hoạch cho các đơn vị quản lý: tổ chức nhân sự, đưa ra giải pháp thi công, vận hành cũng như bảo trì, bảo dưỡng,...trong thời gian ngắn hạn, trung hạn hay dài hạn tùy thuộc tính chất và quy mô của dự án [1]. Chính vì những lý do trên, việc nghiên cứu và phát triển các hệ thống dự báo là một nhiệm vụ vô cùng cần thiết. Dù ở bất kỳ quy mô nào, các mô hình này vẫn cần phải tuân theo một quy trình cơ bản:

* Xác định đối tượng cần dự báo
* Nghiên cứu các phương pháp dự báo
* Lựa chọn phương pháp phù hợp đối với từng vấn đề cụ thể
* Đánh giá và cải thiện kết quả dự báo.

Đối với mỗi một bài toán cụ thể, ta sẽ có một đối tượng dự báo khác nhau. Thêm vào đó, ta cần xác định được những yếu tố có ảnh hưởng đến độ chính xác của kết quả dự báo. Tiếp đó, quy mô dự báo cũng cần được xác định rõ ràng dựa vào mục đích của việc dự báo. Điều này có ý nghĩa quan trọng trọng và ảnh hưởng trực tiếp đến việc chọn một mô hình dự báo phù hợp. Tùy vào quy mô và tính chất của dự án, ta cũng cần chọn được tần suất khảo sát phù hợp. Sau khi xác định được những vấn đề trên, việc tiếp theo cần làm là thu thập dữ liệu dự báo và hiệu chỉnh dữ liệu ấy sao cho phù hợp với yêu cầu dự báo. Dữ liệu dự báo có thể được cung cấp từ các đơn vị, tổ chức thu thập dữ liệu. Đồng thời, với điều kiện công nghệ và các phương pháp đo lường ngày càng tối ưu và chính xác, việc thu thập dữ liệu dự báo ngày càng trở nên thuận lợi.

Cụ thể trong lĩnh vực hệ thống điện, bài toán dự báo có một vài trò hết sức quan trọng trong việc hoạch định các kế hoạch thi công, vận hành, bảo dưỡng hệ thống trong những khoảng thời gian khác nhau, nhằm đạt được mục tiêu kinh tế và kỹ thuật ở mức tối ưu trong phạm vi cho phép. Ví dụ, dựa vào các dữ liệu thống kê về công suất phụ tải ở các khung giờ khác nhau trong nhiều năm liên tiếp, các kỹ sư điều độ hệ thống điện có thể sử dụng các mô hình dự báo để dự đoán về công suất phụ tải ở các khung giờ khác nhau trong những ngày tiếp theo, từ đó đưa ra kế hoạch điều chỉnh công suất phát phù hợp nhằm hạn chế tối đa sự cố trong hệ thống cũng như giảm thiểu lãng phí công suất, từ đó giảm thiểu thiệt hại về kinh tế.

Theo Quy hoạch điện VIII của Bộ Công Thương, một trong những quan điểm phát triển Điện lực quốc gia trong giai đoạn này là: Ưu tiên phát triển nguồn điện sử dụng năng lượng tái tạo (chủ yếu là điện gió trên đất liền, điện gió trên biển; điện mặt trời, thủy điện nhỏ), tạo đột phá trong việc đảm bảo an ninh năng lượng quốc gia, góp phần bảo tồn tài nguyên năng lượng, giảm thiểu tác động tiêu cực tới môi trường trong sản xuất điện; chú trọng phát triển các nguồn điện nhỏ sử dụng năng lượng tái tạo đấu nối với lưới điện phân phối, góp phần giảm tổn thất điện năng [2]. Theo như nội dung Quy hoạch, các nguồn năng lượng tái tạo trong đó có điện mặt trời sẽ ngày càng phát triển và chiểm một tỷ trọng lớn trong cơ cấu nguồn điện Quốc gia trong những năm sắp tới. Điều này đem lại nhiều cơ hội và triển vọng phát triển lớn hơn bao giờ hết cho hệ thống điện Việt Nam, đồng thời đặt ra những thách thức không nhỏ dành cho những người vận hành hệ thống điện nói riêng và nhân lực ngành điện nói chung. Cụ thể, hệ thống điện mặt trời (PV system) bao gồm các tấm pin năng lượng mặt trời (solar panels), có tác dụng chuyển hóa trực tiếp năng lượng ánh sáng mặt trời (quang năng) thành điện năng dựa trên hiện tượng quang điện. Để đảm bảo cho mục tiêu kinh tế và kĩ thuật được tối ưu nhất, hệ thống điện mặt trời cần phải phát huy hết công suất, cung cấp công suất thực cho hệ thống điện. Tuy nhiên, hệ thống này luôn phải chịu những tác động khách quan từ môi trường xung quanh, đặc biệt là những yếu tố thời tiết. Ví dụ, vào những ngày thời tiết nhiều mây hoặc mưa, cường độ bức xạ mặt trời sẽ bị giảm sút. Ngược lại, vào những ngày không mây, nắng gắt, cường độ bức xạ mặt trời lại dồi dào hơn so với bình thường. Điều này dẫn đến sự mất ổn định trong công suất phát, gây ảnh hưởng tiêu cực tới hệ thống hay thậm chí là cả lưới điện nếu không có giải pháp vận hành và điều tiết phù hợp. Những lý do trên đã đặt ra yêu cầu cấp thiết về việc phải dự báo được công suất của hệ thống điện mặt trời dựa trên mối quan hệ với các yếu tố khí tượng. Nếu quá trình dự báo cho ra kết quả với sai số càng thấp thì các giải pháp vận hành và điều độ sẽ được thực hiện với mức độ hiệu quả càng cao, từ đó đạt được mục tiêu toàn diện về cả phương diện kinh tế và kỹ thuật.

Theo như nội dung trên, hệ thống điện mặt trời có công suất phát của phụ thuộc nhiều vào các yếu tố thời tiết trong đó yếu tố có ảnh hưởng trực tiếp đến công suất đầu ra của hệ thống chính là cường độ bức xạ của mặt trời [3]. Tuy nhiên, việc đo đạc và xử lý dữ liệu về công suất phát trong quá khứ thường gặp nhiều khó khăn do phải xây dựng một hệ thống đo lường phức tạp, cồng kềnh và tốn kém, lại càng bất khả thi hơn khi thực hiện quy hoạch hệ thống điện mặt trời ở những khu vực mới, nơi chưa có bất kỳ dữ liệu nào về công suất phát trong quá khứ. Bên cạnh đó, các dữ liệu khí tượng và thời tiết thường được các đơn vị thực hiện dự báo thời tiết đo đạc từ trước bằng những phương pháp hiện đại và tối ưu, khiến cho việc thu thập dữ liệu về cường độ bức xạ mặt trời dễ dàng hơn rất nhiều so với dữ liệu về công suất phát. Từ những dữ liệu ấy, dựa vào các mô hình xử lý phù hợp, ta không chỉ đưa ra được những dự báo được công suất phát trong tương lai mà còn làm tiền đề cho nhiều bài toán quan trọng khác: tìm công suất phát tối ưu, đánh giá phản ứng của hệ thống trong từng trường hợp cụ thể, tái cấu trúc hệ thống tấm pin, đánh giá tuổi thọ tấm pin, tìm ra sự phụ thuộc giữa dòng điện và điện áp của tấm pin vào cường độ bức xạ [4]–[7].

Trong công tác dự báo cường độ bức xạ mặt trời, việc xác định được khoảng thời gian phù hợp là một điều kiện quan trọng ảnh hưởng trực tiếp đến độ chính xác của dự báo. Khoảng thời gian dự báo được chia thành 4 loại: cực ngắn hạn, ngắn hạn, trung hạn và dài hạn. Dự báo cực ngắn hạn là loại dự báo trước một giây đến một giờ, dùng để điều tiết hệ thống trong thời gian thực, nhằm đảm bảo công suất phát được duy trì ở mức ổn định và không bị thay đổi đột ngột. Dự báo ngắn hạn là loại dự báo trước một giờ đến một ngày, giúp cho các đơn vị điều hành nắm bắt thông tin để thực hiện bài toán huy động nguồn, vận hành thị trường điện một cách hiệu quả. Dự báo trung hạn là loại dự báo trong phạm vi một tuần đến một tháng, được sử dụng để lập kế hoạch bảo trì, bảo dưỡng hệ thống. Cuối cùng, dự báo dài hạn là loại dự báo trong phạm vi một tháng tới một năm, giúp các nhà sản xuất và đầu tư nghiên cứu, phát triển các hệ thống điện mặt trời mới. Bảng 1 tổng hợp các loại dự báo và mục đích của từng loại.

Bảng 1. Loại dự báo cường độ bức xạ mặt trời và mục đích

|  |  |
| --- | --- |
| **Loại dự báo cường độ bức xạ mặt trời** | **Mục đích** |
| Cực ngắn hạn (1 giây tới 1 giờ) | Điều phối theo thời gian thực, đảm bảo quá trình sản xuất điện diễn ra ổn định |
| Ngắn hạn (1 giờ tới một ngày) | Tối ưu tổ máy phát, vận hành thị trường điện hiệu quả |
| Trung hạn (1 tuần tới 1 tháng) | Lập kế hoạch bảo trì, bảo dưỡng |
| Dài hạn (1 tháng tới 1 năm) | Lập kế hoạch quy hoạch hệ thống mới |

Bức xạ mặt trời có thể được dự báo bằng rất nhiều các phương pháp, chia thành ba mô hình cơ bản: mô hình vật lý (Physical Model), mô hình thống kê (Statistical Model) và mô hình máy học (Machine Learning Model). Nghiên cứu này sử dụng mô hình Deep Learning (Một nhánh của Machine Learning), cụ thể là mô hình Long Short-term Memory (LSTM) để thực hiện bài toán dự báo. Kết quả dự báo của mô hình trên được so sánh với hai mô hình Deep Learning khác bao gồm mạng nơ-ron hồi quy (RNN) và mạng nơ-ron nhân tạo (ANN). Bên cạnh đó, em xây dựng một giao diện web nhằm dự báo bức xạ mặt trời dựa trên dữ liệu được cung cấp bởi người dùng theo một tiêu chuẩn nhất định. Sau khi dữ liệu được tải lên, phần mềm sẽ bắt đầu tiến hành dự báo bức xạ với quy mô ngắn hạn, sau đó cung cấp các sai số cùng các biểu đồ dự báo để đánh giá độ chính xác các các mô hình.

Nghiên cứu này đưa ra kết quả dự báo bức xạ mặt trời trong khoảng thời gian ngắn hạn( 1 giờ tới 1 ngày) dựa trên mô hình Deep Learning Long-Short term memory (LSTM). Để cho bài nghiên cứu được đảm bảo tuân theo nguyên lý về mối liên hệ phổ biến và nguyên lý về sự phát triển [8], mô hình trên được so sánh với những mô hình tham chiếu đã được nghiên cứu trước đó là mạng nơron nhân tạo (ANN) và mạng nơron hồi quy (RNN). Dựa vào kết quả dự báo thông qua các sai số sẽ được trình bày ở nội dung phía dưới, có thể đánh giá rằng mô hình LSTM được đề xuất thường có độ chính xác cao hơn.

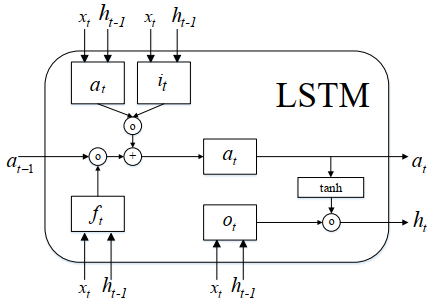
Mô hình nghiên cứu đã được em phát triển trên 1 giao diện web nhờ công cụ hỗ trợ và được lập trình bằng ngôn ngữ Python. Người dùng có thể tương tác thông qua UI (giao diện người dùng) được thiết kế hài hòa, trực quan và dễ thao tác. Dữ liệu đầu vào sẽ được người dùng trực tiếp tải lên. Người dùng có thể điều chỉnh các tham số đầu vào phù hợp với mục đích và nhu cầu. Sau thời gian xử lý dữ liệu đầu vào, dữ liệu đầu ra sẽ được hiển thị dưới dạng đồ thị và bảng số liệu trực quan trên giao diện ứng dụng.

# Phương pháp luận

Dữ liệu lịch sử của các biến khí tượng được thu thập dưới dạng chuỗi thời gian (time series) đồng thời bức xạ mặt trời có độ phụ thuộc lớn vào các biến khí tượng khác. Vì vậy, bên cạnh việc sử dụng các mô hình thống kê như ARIMA thì việc sử dụng mô hình Long Short- term Memory (LSTM) là hoàn toàn phù hợp.

## Mô hình Long Short-term Memory (LSTM)

Long Short-term Memory (LSTM) được đề xuất bởi Hochreiter và Schmidhuber vào  
năm 1997 để tránh việc phụ thuộc dài hạn thông qua thiết kế mục tiêu [9]. Sự cải tiến của mô  
hình LSTM khi so sánh với mô hình RNN gồm một lớp ẩn là mô hình LSTM lưu trữ các  
thông tin trong một đơn vị điều khiển bên ngoài dòng chảy đơn thuần của mô hình RNN, từ  
đó hình thành một đơn vị trạng thái mới 𝑎t (cell state). Hình 1 mô tả cấu trúc của một tế bào  
LSTM. Trong hình này, tại mỗi thời điểm t : it, ft, ot, ãt là input gate, forget gate, output  
gate, và candidate value, có thể được xây dựng tương ứng bởi các công thức sau:



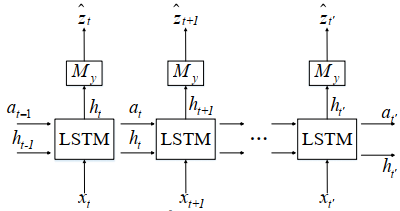
Hình 4. Cấu trúc của một tế bào LSTM

Trong đó *Wi,x, Wi,h, Wf,x, Wf,h, Wã,x, Wã,h*là các ma trận trọng số, *bi, bf, bo, bã* là các véc-tơ  
bias, σ là hàm Sigmoid, *xt* là giá trị đầu vào tại thời điểm *t*, *ht-1* là output của trạng thái tại  
thời điểm *t-1*. Forget gate quyết định xem cần lấy bao nhiêu thông tin từ đơn vị trạng thái  
trước và input gate sẽ quyết định lấy bao nhiêu từ input của trạng thái và lớp ẩn của lớp trước  
[10]. Đơn vị trạng thái tại thời điểm *t* được tính như sau:

Trong đó biểu diễn Hadamad product [11]. Đầu ra *ht* của mạng LSTM tại thời điểm t được tính toán như sau :

Cuối cùng, kết quả dự đoán hay đầu ra của mạng LSTM được tính bằng cách sử dụng *ht* :

Trong đó *My* là ma trận chiếu (project matrix) để giảm số chiều của *ht* . Cấu trúc của  
mạng LSTM được trình bày trong Hình 2. Có thể thấy từ cấu trúc này, một vector đặc trưng  
đầu vào được đưa vào mạng tại thời điểm *t*. Tế bào LSTM trước đó cung cấp phản hồi cho tế  
bào LSTM hiện tại để chứng minh sự phụ thuộc thời gian của mạng này. Quá trình huấn  
luyện mạng có mục đích tối thiểu hóa hàm *f* dựa trên các giá trị mục tiêu thực tế *y* bằng cách sử dụng sự lan truyền ngược (back pro-pagation) với gradient descent:

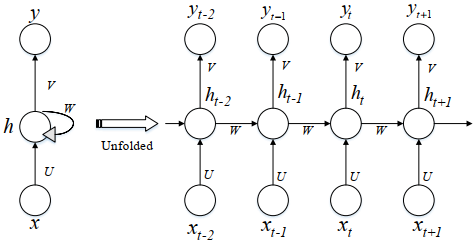


Hình 5. Cấu trúc của mạng LSTM

Để đánh giá hiệu suất cũng như khả năng dự báo của mô hình LSTM, hai mô hình  
tham chiếu được sử dụng lần lượt là mạng nơ-ron hồi quy (RNN) và mạng nơ-ron nhân tạo  
(ANN) được trình bày ở phần dưới.

## Mạng nơ ron hồi quy (RNN)

Mạng nơ-ron hồi quy thường được sử dụng trong mô hình hóa và dự đoán dữ liệu  
chuỗi (sequential data) trong đó đầu ra (output) phụ thuộc vào đầu vào (input). RNN có khả  
năng dự đoán một chuỗi đầu vào ngẫu nhiên dựa trên bộ nhớ trong của mạng này. Bộ nhớ  
trong có thể lưu trữ thông tin về phép tính trước đó. Hình 3 biểu diễn mạng RNN rút gọn và  
mạng RNN đầy đủ. Trong mạng nơ-ron cơ bản, nơ-ron ẩn *h* nhận phản hồi từ các nơ-ron khác  
tại thời điểm trước đó nhân với ma trận trọng số *W*. Khi mạng RNN được biểu diễn một cách  
đầy đủ, ta có thể thấy đầu vào của nơ-ron ẩn là các nơ-ron tại bước thời gian (time step) trước  
đó [12].



Hình 6. Cấu trúc mạng nơ-ron hồi quy (RNN) rút gọn (trái) và đầy đủ (phải)

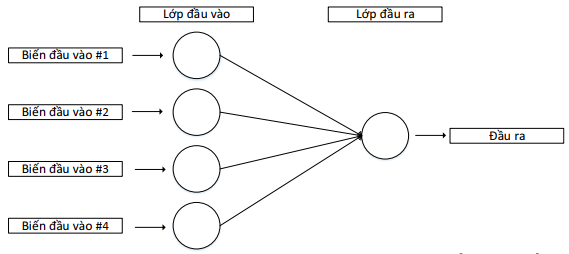
Giá trị đầu vào *𝑥t* tại thời điểm t được nhân với véc-tơ trọng số để đạt được đầu vào của nơ-ron ẩn đầu tiên. Tiếp đó, nơ-ron ẩn tiếp theo *ht* với đầu vào bao gồm *𝑥t-1* và nơ-ron ẩn  
ngay trước đó ht nhân với trọng số W. Nơ-ron đầu ra nhận đầu vào duy nhất là nơ-ron lớp ẩn nhân với ma trận trọng số V. Quá trình hoạt động của mạng RNN được biểu diễn như sau:

Trong đó *f* là hàm hoạt động, ví dụ như Tanh, sigmoid, RELU. Quá trình huấn luyện RNN sử dụng phương pháp Back-propagation through time (BPTT).

## Mạng nơ-ron nhân tạo (ANN)

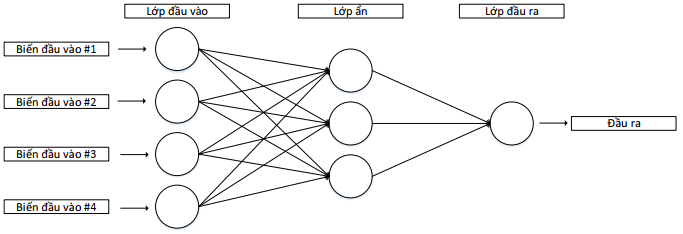
Một mạng nơ ron có thể được tưởng tượng giống như một mạng của các “nơ ron” được  
xếp thành từng lớp. Các yếu tố dự đoán hay các biến đầu vào của mạng nơ ron tạo thành một  
lớp dưới cùng hay còn gọi là lớp đầu vào. Các biến được dự báo hay đầu ra của mạng nơ ron  
tạo thành lớp trên cùng hay lớp đầu ra. Ngoài ra, một mạng nơ ron có thể chứa các lớp trung  
gian (lớp ẩn) bao gồm các “nơ ron ẩn”.

Một Mạng nơ ron đơn giản nhất không chứa các lớp ẩn và chúng tương đương với mô hình hồi quy tuyến tính. Hình 4 biểu diễn phiên bản mạng nơ ron của một mô hình hồi quy tuyến tính với bốn biến đầu vào hay bốn yếu tố dự đoán. Các chỉ số gắn liền với mỗi yếu tố dự đoán được gọi là “weights”. Sự dự báo được thực hiện bởi sự kết hợp tuyến tính của các đầu vào. Các “weight” được lựa chọn và điều chỉnh trong một cấu trúc mạng nơ ron bằng việc sử dụng một “thuật toán học” nhằm mục đích tối thiểu hóa một hàm “cost function”, ví dụ như hàm sai số toàn phương trung bình (MSE). Trong ví dụ đơn giản trên, chúng ta có thể sử dụng hồi quy tuyến tính để đơn giản hóa và hiệu quả tốt hơn thay vì sử dụng mô hình mạng nơ ron để huấn luyện mô hình.



Hình 7. Mô hình mạng nơ ron đơn giản tương đương một mô hình hồi quy tuyến tính

Một khi mạng nơ ron được thêm vào các lớp trung gian gồm các nơ ron ẩn, mạng nơ  
ron đó sẽ trở thành mạng nơ ron phi tuyến. Một ví dụ đơn giản cho trường hợp này được thể  
hiện ở hình 5.



Hình 8. Mô hình mạng nơ ron với 4 biến đầu vào và 3 nơ ron trong lớp ẩn

Mạng nơ ron trên còn được biết với tên khác là mạng nơ ron truyền thẳng nhiều lớp.  
Đặc điểm dễ thấy của mạng nơ ron trên là mỗi nơ ron (nút mạng) của từng lớp nhận các đầu  
vào từ các lớp trước của nó. Đầu ra của mỗi nút mạng trong một lớp là đầu vào của lớp kế  
tiếp. Các biến đầu vào của mỗi nút mạng được hợp nhất bằng việc sử dụng sự kết hợp tuyết  
tính có xét đến các trọng số. Kết quả sau đó được chỉnh sửa bởi một hàm số phi tuyến trước  
khi trở thành đầu ra của mạng nơ ron. Ví dụ, các biến đầu vào của nơ ron ẩn thứ j trong hình  
2 được liên kết một cách tuyến tính như phương trình (11):

Trong lớp ẩn, một hàm số phi tuyến (Ví dụ: Sigmoid) với phương trình (12) sau đây  
được đưa ra để chỉnh sửa lại kết quả trên :

Từ đó đưa ra đầu vào của lớp tiếp theo. Điều này có xu hướng làm giảm sự ảnh hưởng các biến đầu vào có giá trị lớn, từ đó làm cho mô hình mạng có phần mạnh mẽ để trở nên vượt trội hơn.

Các thông số 𝑏1, 𝑏2, 𝑏3 và *W1,1, ..., W4,3* được “học” từ các dữ liệu. Các giá trị của  
weights thường được hạn chế để tránh việc chúng quá lớn. Các chỉ số hạn chế các weights  
này được gọi là “decay parameter” và thường được đặt bằng 0,1.

Các weight nhận các giá trị ngẫu nhiên để bắt đầu, và chúng sau đó được cập nhật với  
việc sử dụng các dữ liệu được xem xét. Hệ quả là một thành phần của sự ngẫu nhiên đó được  
cung cấp bởi mạng nơ ron in việc thực hiện dự báo. Vì vậy, mạng nơ ron thường xuyên được  
huấn luyện một vài lần sử dụng giá trị bắt đầu ngẫu nhiên khác nhau, và kết quả cuối cùng sẽ  
được lấy trung bình của các lần tập huấn đó.

Số lượng của các lớp ẩn hay các nơ ron trong mỗi lớp ẩn phải được chỉ định trước bởi  
vì việc số lượng càng nhiều thì dẫn tới mạng nơ ron trở nên phức tạp hơn và thời gian tính  
toán lâu hơn nhưng kết quả mang lại có thể chính xác hơn.

## Các sai số đánh giá

Ba phép đo sai số được dùng để đánh giá độ chính xác của mô hình dự báo lần lượt là  
sai số trung bình bình phương (RMSE), sai số trung bình tuyệt đối (MAE) và sai số trung  
bình tuyệt đối phần tram (MAPE):

Sai số trung bình bình phương (RMSE): tính sự khác nhau giữa các giá trị thực tế và  
các giá trị dự báo. RMSE càng thấp cho thấy kết quả dự báo càng tốt.

Trong đó *n* là số lượng các giá trị quan sát; *G* là giá trị thực tế; là giá trị dự báo.

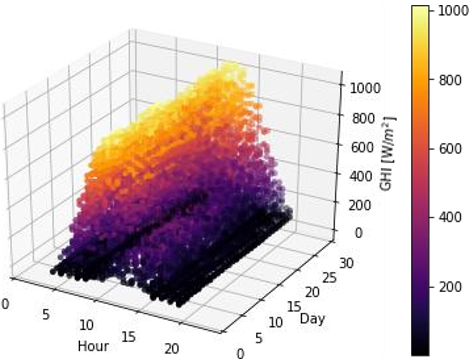
Sai số trung bình tuyệt đối phần trăm (MAPE): là phép đo độ chính xác dự báo của một  
phương pháp dưới dạng sai số phần tram, được biểu diễn như sau:

# Đánh giá kết quả và thảo luận

## Thu thập và tiền xử lý dữ liệu

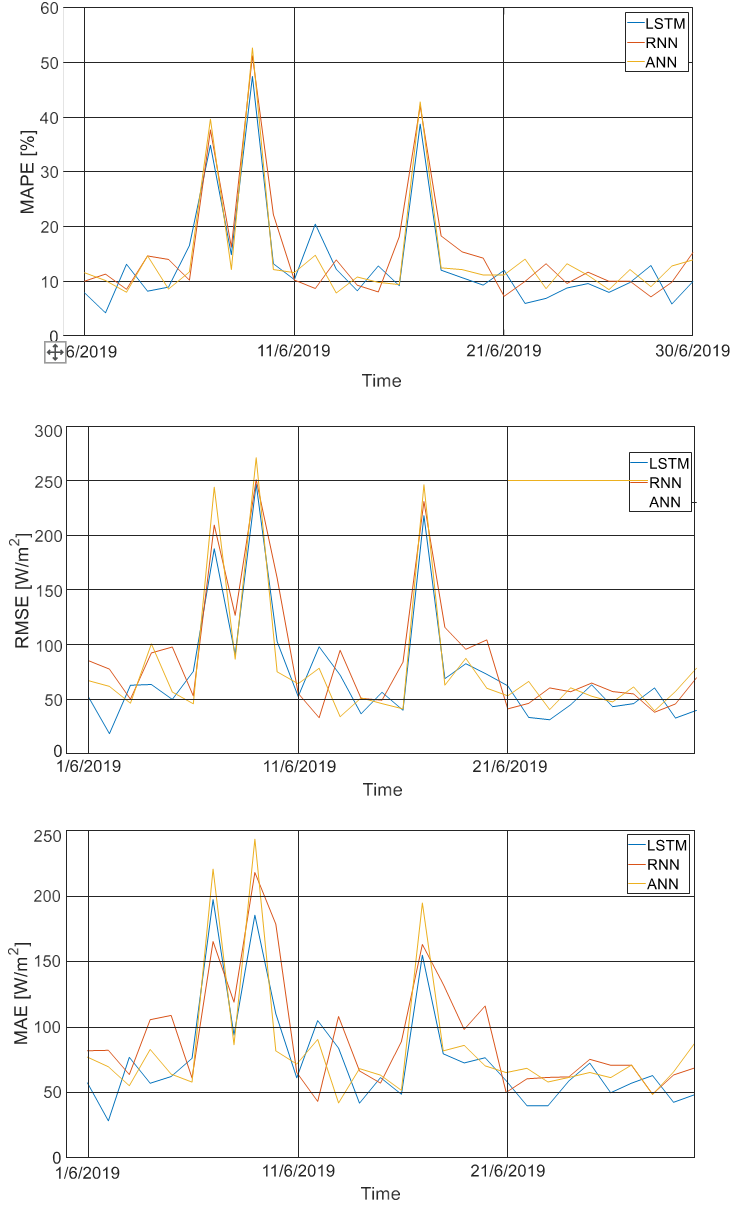
Tập dữ liệu được thu thập từ trang web https://solargis.com với các thông số dữ liệu bức xạ mặt trời và các biến khí tượng được cung cấp từ Plataforma Solar de Almeria, Tây Ban Nha – trung tâm nghiên cứu, phát triển và thử nghiệm công nghệ năng lượng mặt trời tập trung lớn nhất ở châu Âu được thành lập vào năm 1981 [13]. Dữ liệu với độ phân giải 1 giờ được tổng hợp lại dưới dạng file excel với khoảng thời gian từ năm 2016 đến hết năm 2019. Nội dung trong các tệp dữ liệu bao gồm các thông số khí tượng thuỷ văn như nhiệt độ, độ ẩm, tốc độ gió,.. và các thành phần bức xạ mặt trời bao gồm bức xạ chiếu ngang toàn cầu (Global Horizon Irradiance), bức xạ chiếu trực tiếp (Direct Normal Irradiance), bức xạ chiếu ngang khuếch tán (Diffuse Horizon Irradiance). Trong nghiên cứu này, các biến đầu vào bao gồm: bức xạ chiếu ngang toàn cầu (GHI), bức xạ chiếu trực tiếp (DNI), góc thiên đỉnh mặt trời, độ ẩm tương đối, nhiệt độ và bức xạ chiếu ngang khuếch tán (DHI) trong khoảng thời gian từ đầu năm 2016 đến hết năm 2018 được sử dụng làm tập huấn luyện (training set) cho các mô hình dự báo. GHI tại thời điểm tháng 6 và tháng 12 năm 2019 được sử dụng làm tập kiểm tra (testing set) để đánh giá hiệu suất cũng như độ chính xác của mô hình dự báo đã được huấn luyện trước đó.

Sau khi kết thúc quá trình thu thập dữ liệu, tập dữ liệu cần trải qua một bước vô cùng quan trọng là tiền xử lý dữ liệu. Có rất nhiều quy trình để tiền xử lý một bộ dữ liệu. Tuy nhiên, nghiên cứu này chỉ sử dụng hai phương pháp là: làm sạch dữ liệu (Data Cleaning) và chuẩn hóa dữ liệu (Data Normalization). Trong bước làm sạch dữ liệu, các dữ liệu tại thời điểm mà bức xạ có cường độ nhỏ hoặc bằng không (ban đêm) được loại bỏ để giảm thiểu số lượng điểm dữ liệu, từ đó thời gian tính toán của mô hình dự báo được giảm đi đồng thời độ chính xác sẽ được tăng lên. Hình 6 biểu diễn dữ liệu GHI vào tháng 6 sau khi đã được loại bỏ những dữ liệu không cần thiết. Bên cạnh đó, các biến khí tượng được thu thập trên có phạm vi thay đổi giá trị tương đối khác nhau, ví dụ như bức xạ mặt trời thay đổi trong khoảng từ 0 đến 1000 W/m2, nhiệt dộ dao động từ 0°C đến 45°C, v.v. Việc này phần nào gây ra sự khó khăn trong việc tính toán, vì vậy việc chuẩn hóa các giá trị của điểm dữ liệu để thay đổi trong phạm vi từ 0 đến 1 là hết sức cần thiết. Công thức chuẩn hóa được biểu diễn như biểu thức :



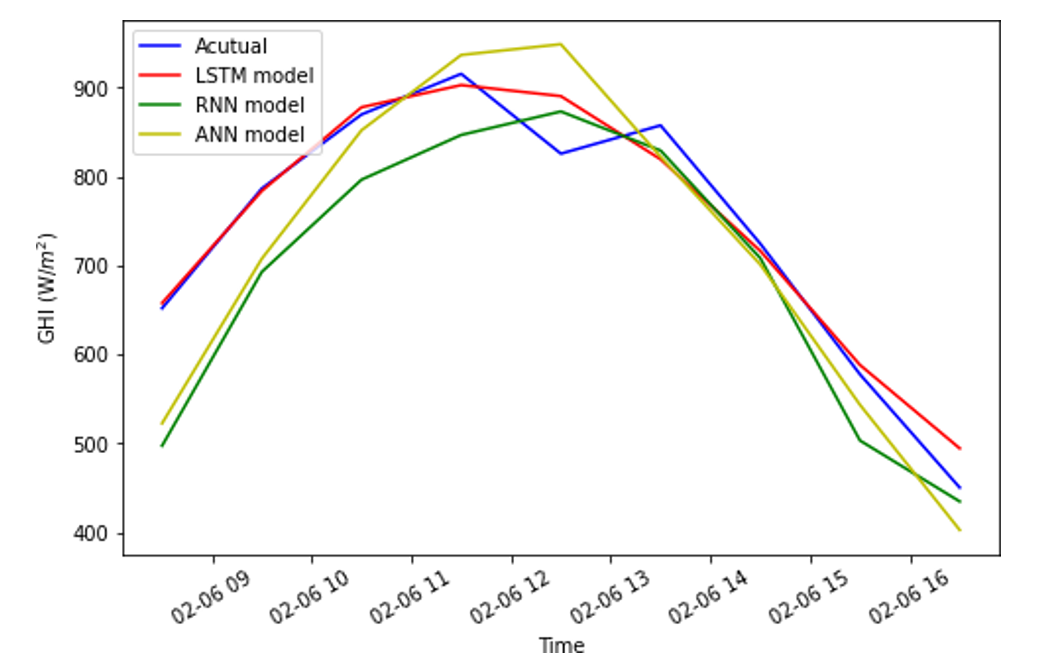
Hình 9. Tập dữ liệu sau khi được làm sạch

## Kết quả và thảo luận

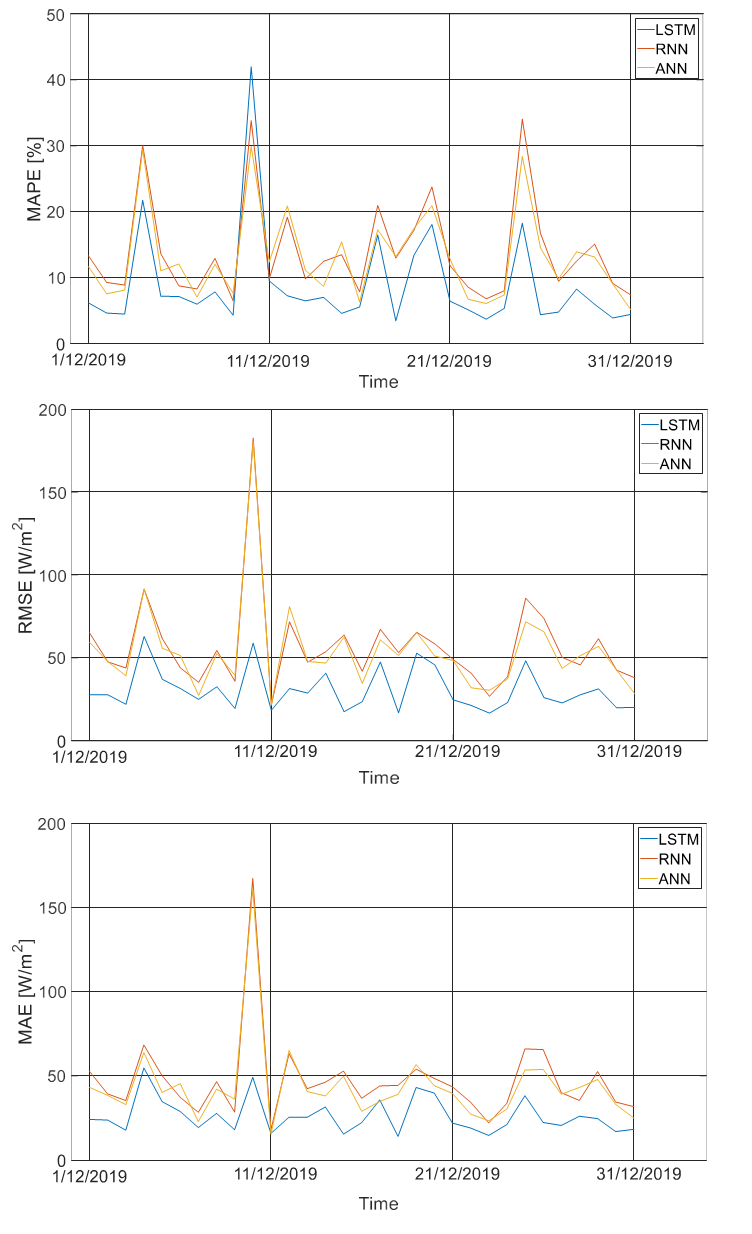


Hình 10. Kết quả dự báo của các ngày trong tháng 6 năm 2019

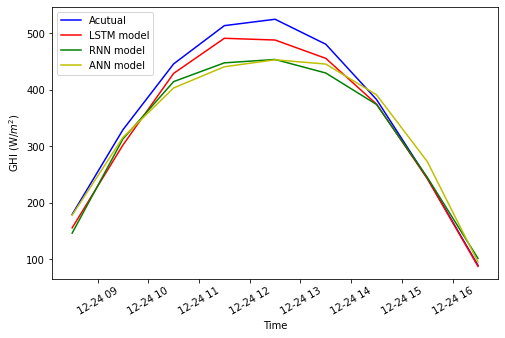
Trong trường hợp nghiên cứu 1, các giá trị GHI theo ngày trong tháng 6 được dùng làm tập kiểm tra nhằm đánh giá độ chính xác khi dự báo ngắn hạn (một ngày) của mô hình LSTM cùng hai mô hình tham chiếu là RNN và ANN. Hình 7 cho thấy kết quả dự báo của các ngày trong tháng 6 năm 2019 dưới dạng MAPE, RMSE, MAE. Từ hình 7a, có thể thấy trong hầu hết các trường hợp, MAPE của mô hình LSTM thấp nhất khi so với hai mô hình còn lại, cụ thể MAPE của mô hình LSTM dao động trong phạm vi từ khoảng 5% đến hơn 30%. Tuy nhiên, từ hình trên, ta thấy sai số của ba mô hình dự báo trong ba ngày lần lượt là 7/6/2019, 9/6/2019 và 17/6/2019 là rất lớn. Điều này chứng tỏ ba ngày kể trên là những ngày có lượng mây lớn hoặc mưa lớn, vì vậy ảnh hưởng rất lớn đến giá trị bức xạ mặt trời trong ngày và việc dự báo trở nên vô cùng khó khăn. Hình 7b và hình 7c lần lượt đưa ra sai số RMSE và MAE của các ngày trong tháng 6. Tương tự như MAPE, mô hình LSTM có độ dự báo chính xác cao nhất trong hầu hết các trường hợp. Cụ thể, sai số RMSE của mô hình LSTM thay đổi từ khoảng 45W/m2 đến hơn 250W/m2 trong khi MAE thay đổi trong khoảng 40W/m2 đến hơn 200W/m2. Hình 8 biểu diễn kết quả dự báo giữa mô hình LSTM cùng hai mô hình tham chiếu RNN, ANN với giá trị GHI thực tế trong ngày 2/6/2019. Có thể thấy từ hình vẽ, giá trị GHI dự báo từ mô hình LSTM có sự khác biệt nhỏ nhất trong ba mô hình khi so sánh với các giá trị GHI thực tế trong ngày. Sự khác nhau lớn nhất xuất hiện khi giá trị GHI thực tế thay đổi đột ngột từ thời điểm 11 giờ sang đến 13 giờ sáng. Điều này làm cho giá trị dự báo khó chính xác. Tuy nhiên, thời gian huấn luyện của mô hình LSTM là lớn nhất trong ba mô hình vì số lượng tham số của mô hình LSTM là rất nhiều. Đây cũng là nhược điểm rất lớn của mô hình LSTM.



Hình 11. Kết quả dự báo giữa các mô hình đề xuất so với giá trị thực tế vào 1 ngày của tháng 6



Hình 12. Kết quả dự báo của các ngày trong tháng 12 năm 2019

******

Hình 13. Kết quả dự báo giữa các mô hình đề xuất so với giá trị thực tế vào 1 ngày của tháng 12

Trong trường hợp nghiên cứu 2, các giá trị GHI theo ngày trong tháng 12 được dùng làm tập kiểm tra nhằm đánh giá độ chính xác trong việc dự báo ngắn hạn (trước 1 ngày) của mô hình LSTM cùng hai mô hình tham chiếu là RNN và ANN. Hình 9 cho thấy kết quả dự báo của các ngày trong tháng 12 năm 2019 dưới dạng MAPE, RMSE, MAE. Từ hình 9a, có thể thấy trong hầu hết các trường hợp, MAPE của mô hình LSTM thấp nhất khi so với hai mô hình còn lại, cụ thể MAPE của mô hình LSTM dao động trong phạm vi từ khoảng 4% đến hơn 40%. Do đặc tính thời tiết cũng như tính mùa vụ nên số lượng ngày có mây hoặc mưa vào tháng 12 lớn hơn so với tháng 6, điều này khiến cho việc dự báo trở nên khó khăn hơn.  
Tuy nhiên phạm vi thay đổi của bức xạ mặt trời là nhỏ hơn so với giá trị bức xạ của tháng 6, do đó đối với những ngày nắng hay ngày trong thì kết quả dự báo tương đối chính xác. Tương tự với sự thay đổi của sai số MAPE, hình 7b và hình 7c cho thấy đồ thị sai số RMSE và MAE của ba mô hình dự báo trong các ngày của tháng 12. Có thể thấy, mô hình LSTM trong trường hợp nghiên cứu thứ 2 cũng có sai số RMSE và MAE nhỏ nhất trong hầu hết các ngày. Sai số RMSE và MAE nhỏ nhất của mô hình LSTM lần lượt là khoảng 25W/m2 và 20W/m2. Đối với hai mô hình tham chiếu còn lại, sai số RMSE của mô hình RNN thay đổi trong khoảng 30W/m2 đến 80W/m2 trong khi sai số này của mô hình ANN thay đổi trong khoảng hơn 40W/m2 đến hơn 170W/m2. Hình 10 cho ta thấy kết quả dự báo giữa các mô hình đề xuất so với giá trị thực tế trong ngày 24/12/2019. Có thể thấy, sự khác nhau của mô hình LSTM với giá trị thực tế là nhỏ nhất so với hai mô hình tham chiếu còn lại là RNN và ANN. Điều này cho thấy độ chính xác của mô hình LSTM là tốt nhất trong ba mô hình. Tuy nhiên, như đã đề cập ở phía trên, mô hình LSTM có thời gian tính toán rất lớn.

# Xây dựng giao diện phần mềm trên máy tính

Phần mềm được xây dựng trên ngôn ngữ lập trình Python. Đây là một ngôn ngữ lập trình bậc cao cho các mục đích lập trình đa năng, do Guido van Rossum tạo ra và lần đầu ra mắt vào năm 1991 và được phát triển trong một dự án mã mở, do tổ chức phi lợi nhuận Python Software Foundation quản lý. Python được thiết kế với ưu điểm mạnh là dễ đọc, dễ học và dễ nhớ. Đồng thời, nó cũng là ngôn ngữ có hình thức rất sáng sủa, cấu trúc rõ ràng, thuận tiện cho người mới học lập trình.

Streamlit là một open-source Python lib, nó giúp ta dễ dàng tạo một web app cho MachineLearning và Data Science. Đây là công cụ rất phù hợp để demo các sản phẩm AI với các lập trình viên không có nhiều kiến thức về web.

Phần code được viết bằng python và phần UI được xử lí bằng Streamlit

Reference: https://docs.streamlit.io/library/api-reference

1. Phần code :

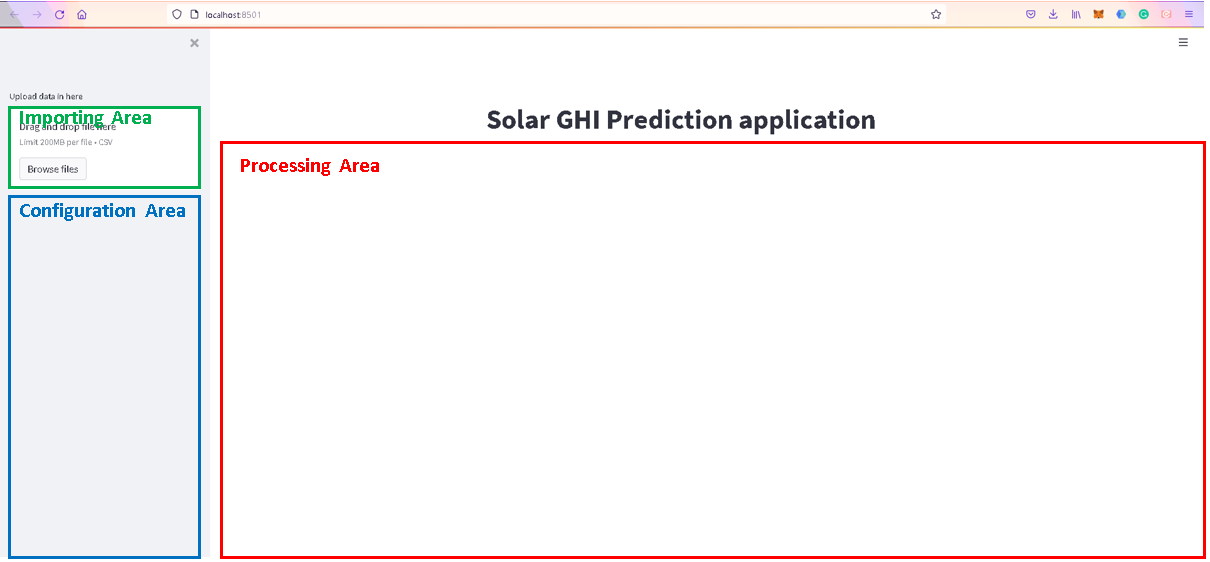
https://colab.research.google.com/drive/1N-QAUcLlZgrcEsyauT178oUu0n5lh8f7#

2. Phần UI Streamlit được thiết kế gồm 3 khu vực

a. Import Area

b. Configuration Area

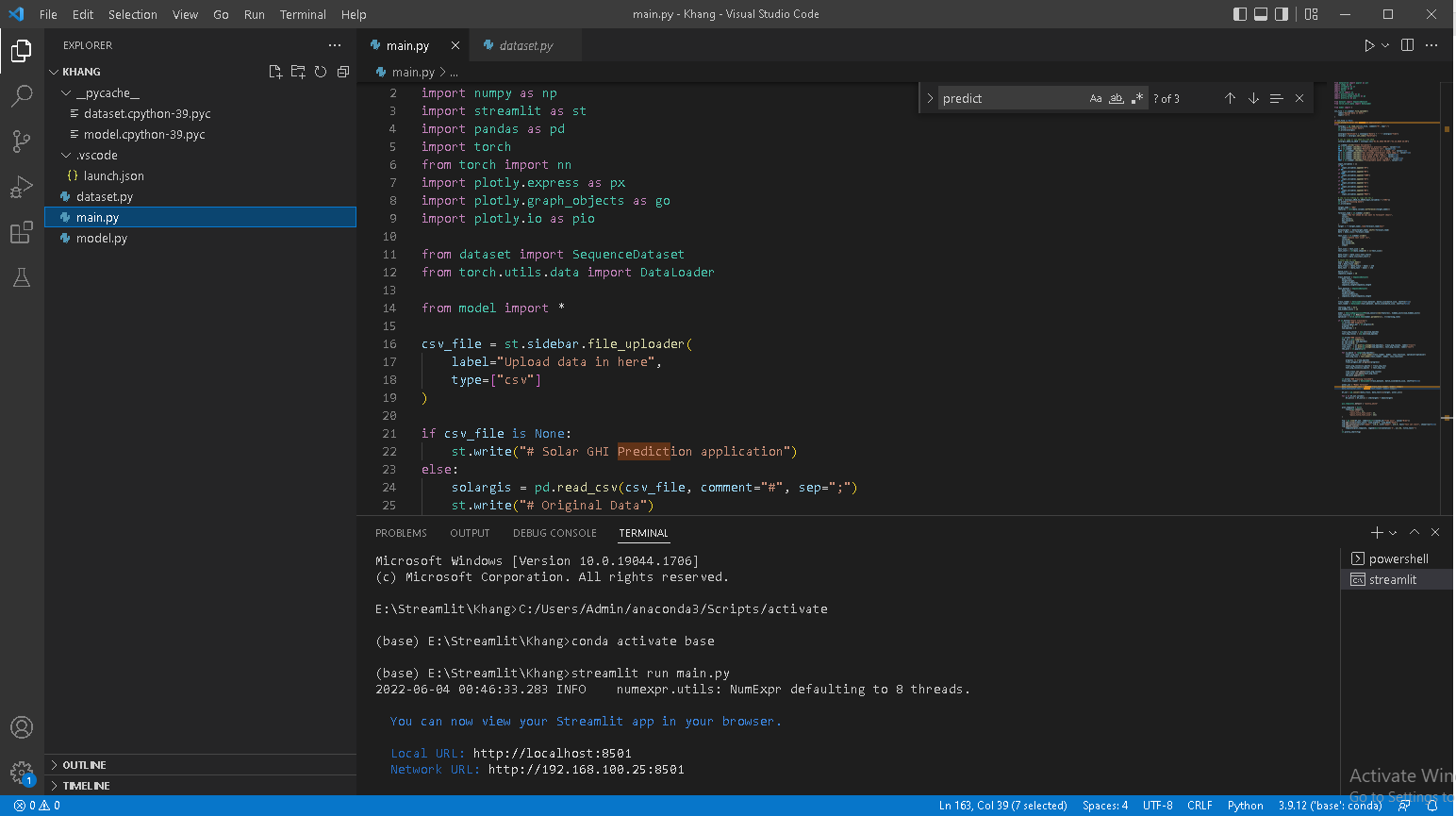
c. Processing Area



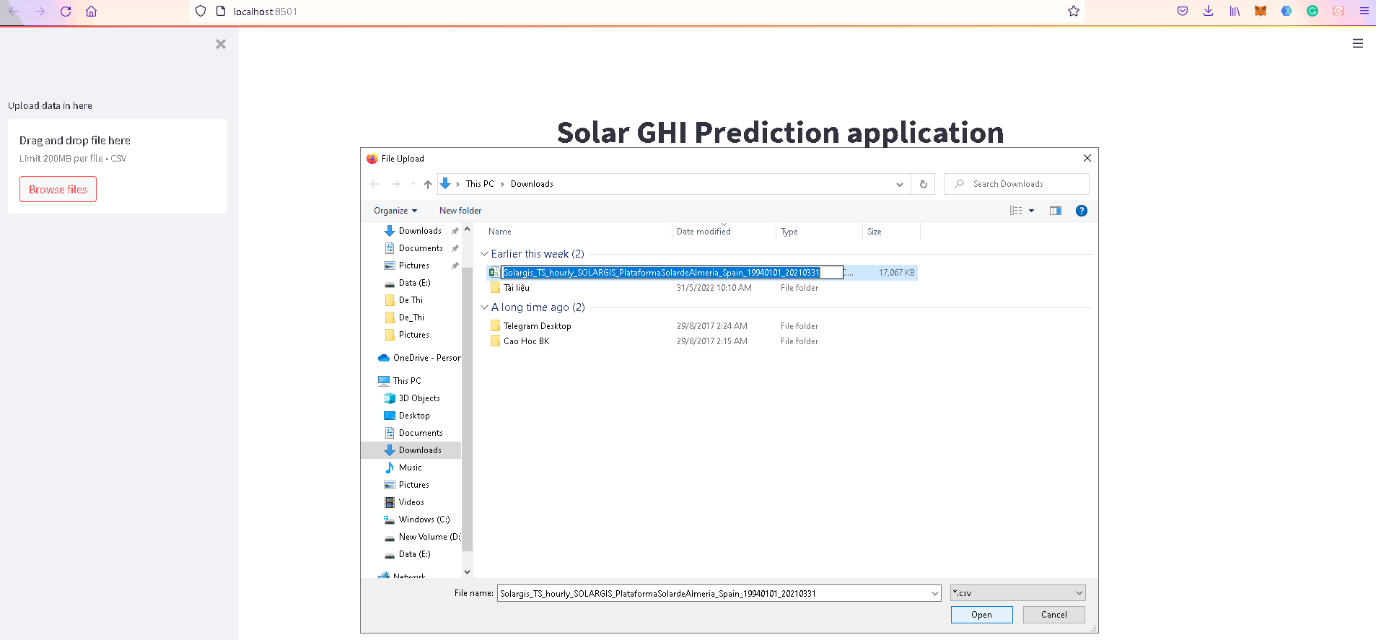
Các bước chính để sử dụng gồm:

* Bước 1: Sử dụng 1 code editor (đã cài các thư viện cần thiết cho streamlit) để mở code.

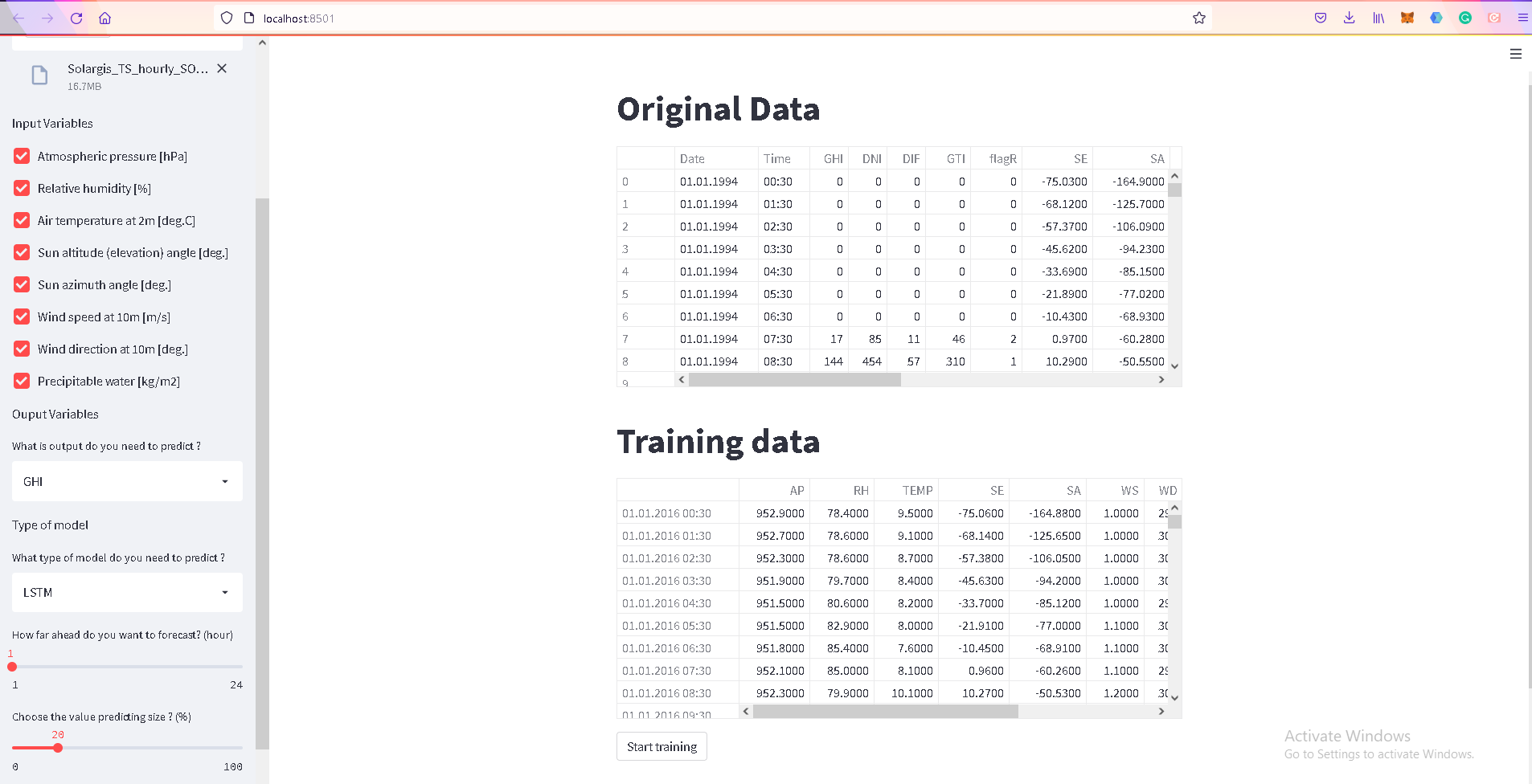
Sử dụng câu lệnh “streamlit run main.py” để mở giao diện Web của Streamlit



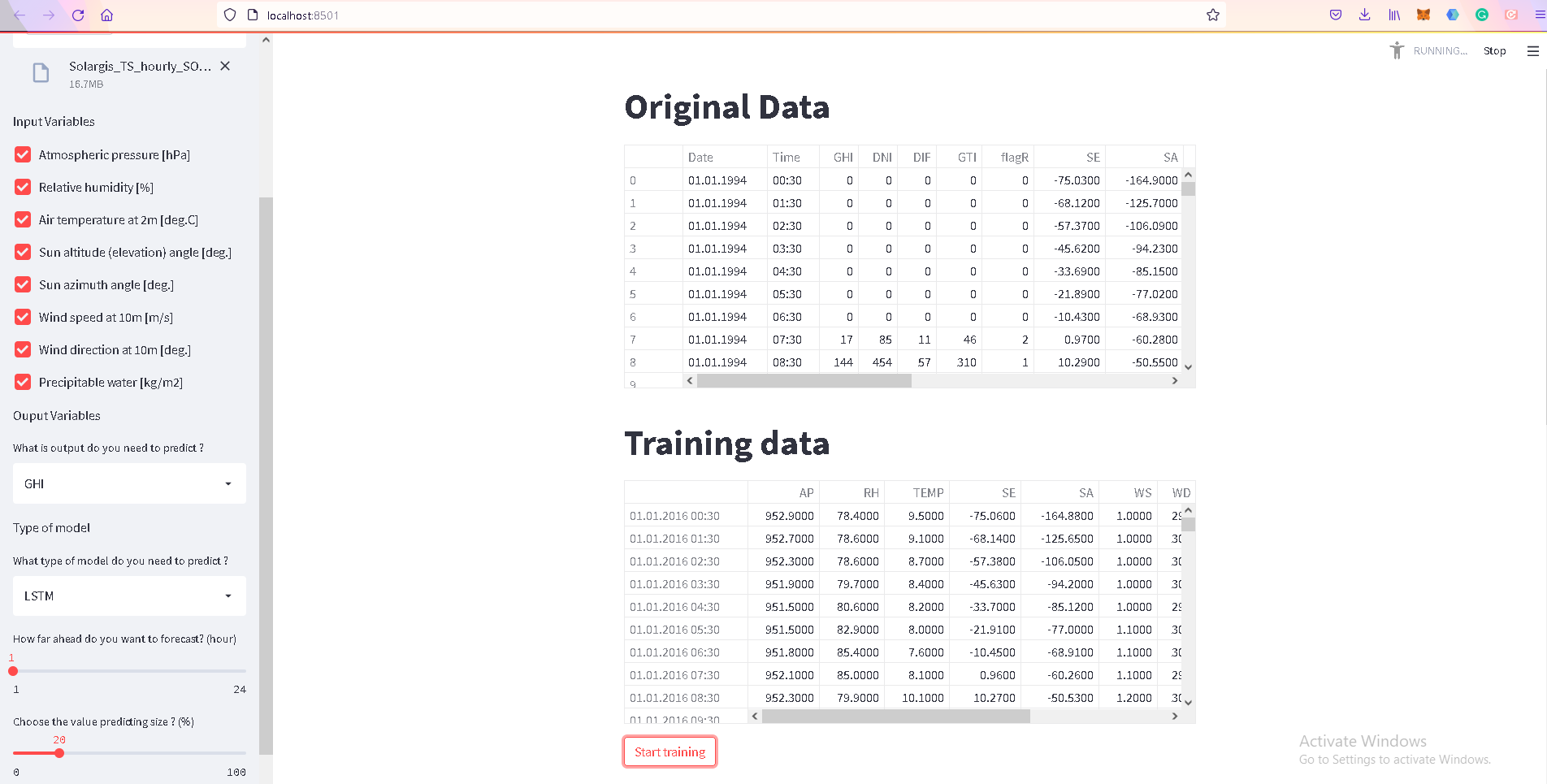
* Bước 2: Click vào nút “Browser Files” để chọn tập dữ liệu muốn sử dụng ở Import Area. Sau khi chọn tập tin, dữ liệu từ tệp sẽ được thể hiện ở Processing Area với định dạng một bảng “Original Data”.



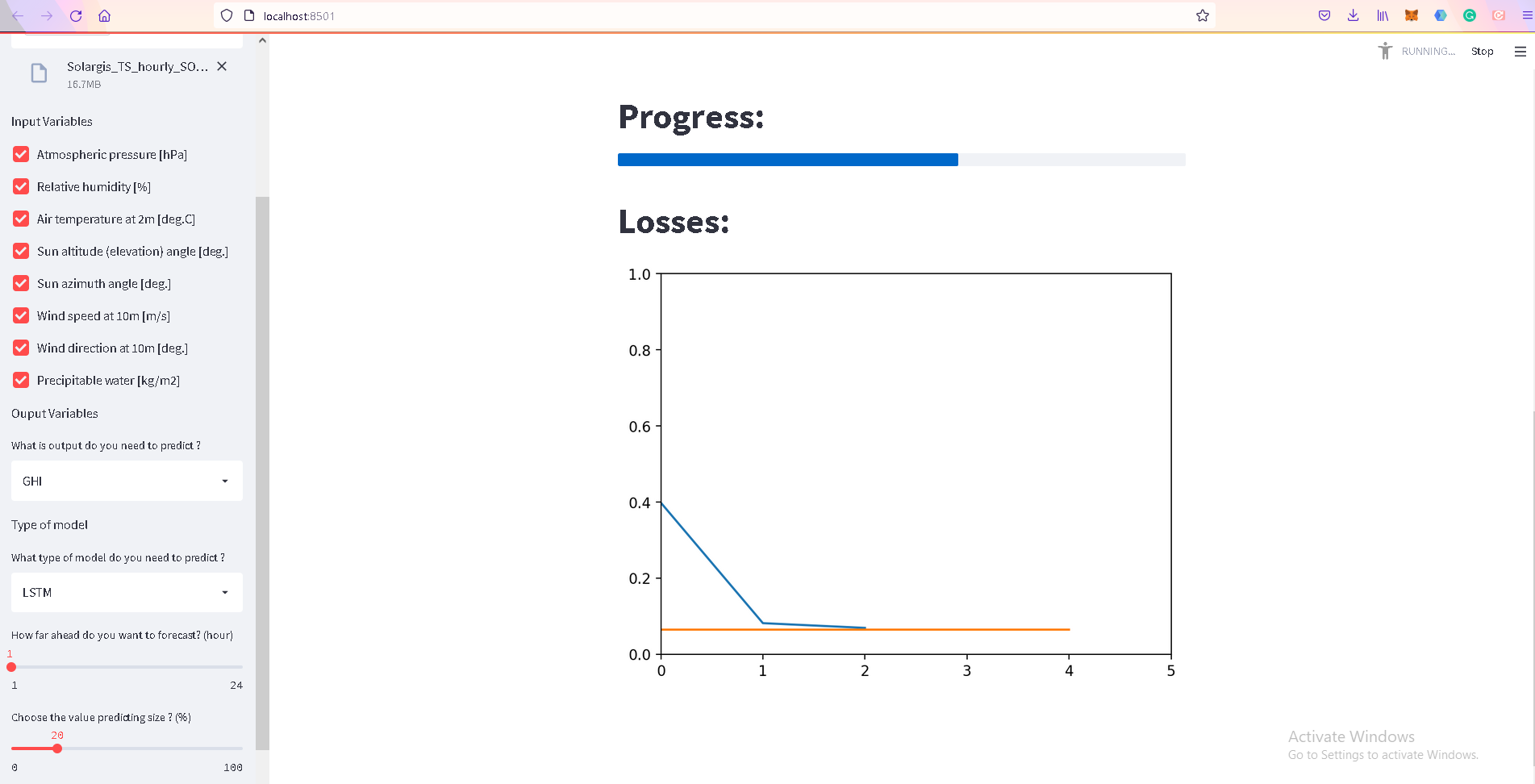
* Bước 3: Người dùng có thể sử dụng các button và slider để chọn các thông số dùng cho dự báo ở Configuration Data
* Chọn các trường dữ liệu đầu vào muốn sử dụng bằng checkbox (Gồm 8 trường : Atmospheric pressure [AP], Relative humidity [RH], Air temperature [TEMP], Sun altitude angle [SE], Sun azimuth angle [SA], Wind speed [WS], Wind direction [WD], Precipitable water [PWAT]).
* Chọn thông số dự báo: Global horizontal irradiation [GHI], Direct normal irradiation [DNI], Diffuse horizontal irradiation [DIF], Global tilted irradiation [GTI].
* Chọn model sử dụng dự báo: LTSM, RNN, ANN
* Chọn thời gian dự báo trước ( 1 - 24h ) bằng slider.
* Chọn tỉ lệ dữ liệu dự báo/ tổng số dữ liệu ( 0 – 100% ) bằng slider.



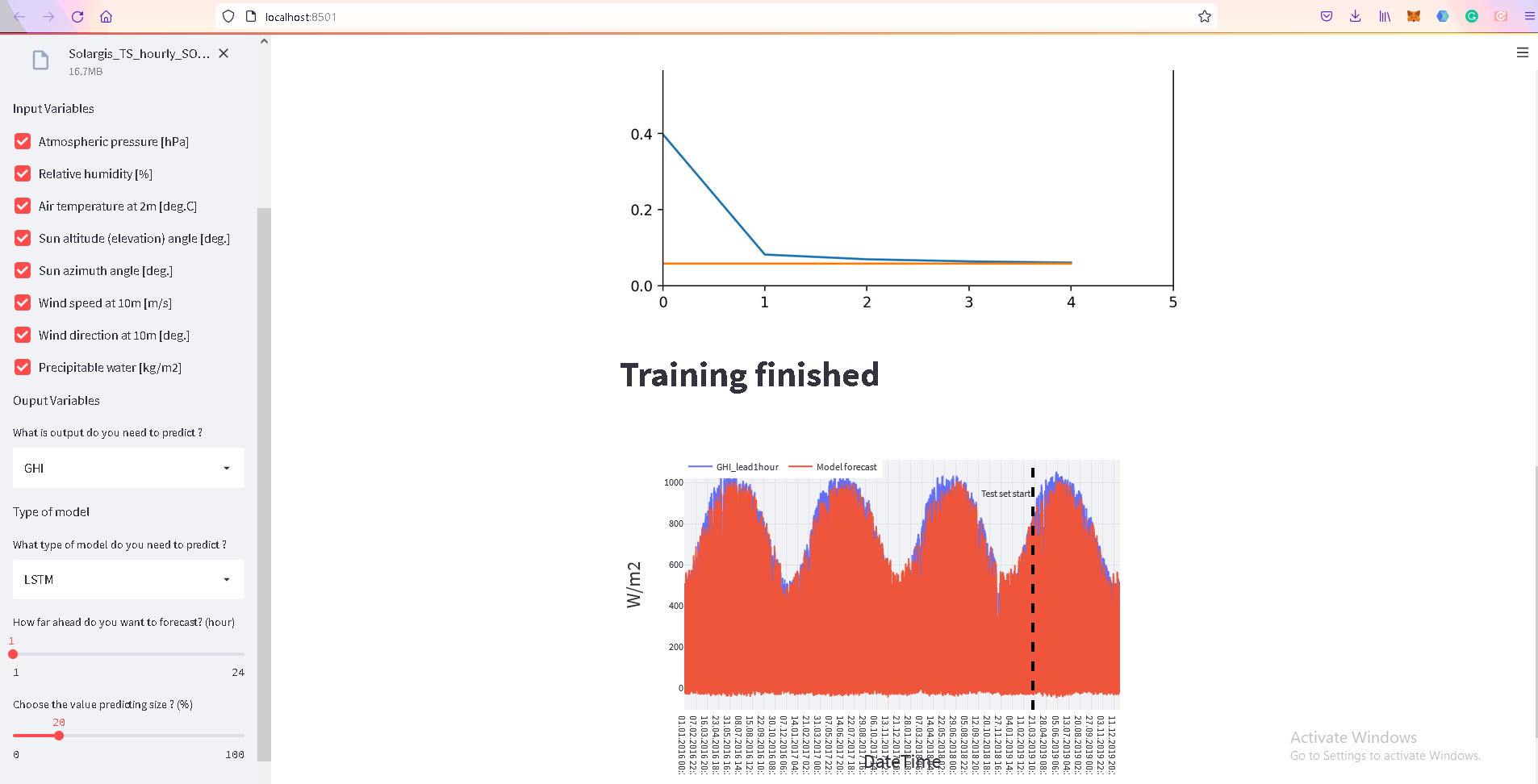
* Bước 4: Sau khi chọn các thông số sử dụng, một bảng giá trị dùng để huấn luyện “Training Data” sẽ được hiển thị ở Processing Area. Bấm “Start training” để bắt đầu



* Bước 5: Chờ đợi quá trình training, có thể quan sát tiến độ ở thanh Progress và Losses ỏ biểu đồ.



* Bước 6: Kết thúc training, sẽ có một đồ thị để dự báo…



# Kết luận

Qua thời gian nghiên cứu thực hiện đề tài, em đã nắm được cách đưa dữ liệu vào một mô hình dự báo, quy trình hoạt động của một số mô hình như mô hình Deep learning Long Short-term memory, mô hình mạng nơ-ron hồi quy và kiểm định độ chính xác qua một số mô hình tham chiếu khác. Bên cạnh đó, em đã biết cách xây dựng một giao diện phần mềm đơn giản nhưng tích hợp được nhiều chức năng thông qua ngôn ngữ lập trình thống kê Python. Một số kết quả chính có thể tổng kết như sau:

* Ứng dụng ba mô hình Long Short-term Memory (LSTM), mạng nơ-ron hồi quy (RNN) và mạng nơ-ron nhân tạo (ANN) vào việc dự báo bức xạ mặt trời trong trường hợp ngắn hạn (cụ thể trước một ngày) trên nền tảng một ứng dụng phần mềm dễ dàng truy cập. Kết quả dự báo tương đối chính xác trong ngày nắng với sai số MAPE khoảng 5%, tuy nhiên đối với các ngày mây thì mô hình có sai số tương đối lớn.
* Xây dựng phần mềm phục vụ nhu cầu dự báo của người tiêu dùng. Người dùng có thể chọn các thông số muốn sử dụng cho việc đánh giá.

Bên cạnh những kết quả đã đạt được, đề tài này vẫn tồn tại những vấn đề chưa giải quyết được. Đây cũng là những hướng phát triển được em dự định thực hiện trong tương lai:

* Áp dụng các thuật toán lựa chọn thông số cho mô hình đi kèm vào phần mềm nhằm nâng cao độ chính xác
* Quy mô dự báo hiện tại của phần mềm chỉ có dự báo trước 1giờ tới 1 ngày, điều đó làm giảm sự đa dạng của ứng dụng. Vì vậy phần mềm sẽ được xây dựng với quy mô dự báo đa dạng hơn để phù hợp với mục đích của người dùng.

Trong tương lai, các mô hình Deep Learning hiện đại với độ chính xác cao hơn sẽ được áp dụng trong ứng dụng.