TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP

KHOA ĐIỆN TỬ



**ĐỒ ÁN TỐT NGHIỆP**

**NGÀNH : KỸ THUẬT MÁY TÍNH**

**HỆ : ĐẠI HỌC CHÍNH QUY**

**ĐỀ TÀI : XÂY DỰNG MÔ HÌNH DỰ BÁO MỰC NƯỚC SỬ DỤNG MÔ HÌNH HỌC MÁY VỚI CƠ CHẾ ATTENTION**

**THÁI NGUYÊN – 2024**

TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP

**KHOA ĐIỆN TỬ**

**ĐỒ ÁN TỐT NGHIỆP**

**NGÀNH : KỸ THUẬT MÁY TÍNH**

**Đề tài : Xây dựng mô hình dự báo mực nước sử dụng mô hình học máy với cơ chế Attention**

**SINH VIÊN : NGUYỄN DUY CAO**

**LỚP : 56KMT**

**GIÁO VIÊN HƯỚNG DẪN : TS NGHIÊM VĂN TÍNH**

#### THÁI NGUYÊN – 2024

|  |  |
| --- | --- |
| **TRƯỜNG ĐHKTCN** | **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM** |
| **KHOA ĐIỆN TỬ** | ***Độc lập - Tự do - Hạnh phúc*** |

**NHIỆM VỤ THIẾT KẾ TỐT NGHIỆP**

Sinh viên: Nguyễn Duy Cao MSSV: K205480106005

Lớp: K56KMT Ngành: Kỹ thuật máy tính

Giáo viên hướng dẫn: TS Nghiêm Văn Tính

Ngày giao đề tài: 04/10/2024 Ngày hoàn thành: 22/11/2024

1. Tên đề tài: **Xây dựng mô hình dự báo mực nước sử dụng mô hình học máy với cơ chế Attention**

2. Nội dung thực hiện

* Tìm hiểu về tổng quan học máy
* Tìm hiểu về thuật toán Long Sort-Term Memory(LSTM) và cơ chế Attention
* Xây dựng mô hình dự báo

3. Các sản phẩm, kết quả:

- Phần mềm đáp ứng các yêu cầu đặt ra.

- Thuyết minh đồ án theo mẫu chung của khoa Điện tử.

4. Ngày giao nhiệm vụ: 04/10/2024

5. Ngày hoàn thành nhiệm vụ: 22/11/2024

|  |  |  |
| --- | --- | --- |
| BCN KHOA | **TRƯỞNG BỘ MÔN** | **GIÁO VIÊN HƯỚNG DẪN** |
| *(Ký và ghi rõ họ tên)* | *(Ký và ghi rõ họ tên)* | *(Ký và ghi rõ họ tên)* |

**NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN**

*Thái Nguyên, ngày...tháng…năm…*

**GIÁO VIÊN HƯỚNG DẪN**

*(Ký ghi rõ họ tên)*

**NHẬN XÉT CỦA GIÁO VIÊN CHẤM**

*Thái Nguyên, ngày…tháng…năm…*

**GIÁO VIÊN HƯỚNG DẪN**

*(Ký ghi rõ họ tên)*

# LỜI MỞ ĐẦU

Trong những năm gần đây, biến đổi khí hậu đang diễn ra ngày càng rõ rệt, gây ra nhiều hiện tượng thời tiết cực đoan như mưa lớn, bão lũ và hạn hán. Điều này đã làm tăng cao nguy cơ lũ lụt ở nhiều khu vực, ảnh hưởng đến đời sống kinh tế - xã hội và gây tổn thất nghiêm trọng về con người và tài sản. Vì vậy, việc dự báo mực nước sông, hồ và các vùng nước tự nhiên khác đã trở thành một nhu cầu cấp thiết đối với nhiều quốc gia, trong đó có Việt Nam. Khả năng dự báo chính xác mực nước không chỉ góp phần quan trọng trong việc xây dựng các kế hoạch ứng phó với thiên tai mà còn hỗ trợ quản lý nguồn nước, phục vụ nông nghiệp, công nghiệp và các hoạt động khác của đời sống xã hội.

Trước đây, các mô hình dự báo mực nước chủ yếu dựa trên phương pháp thống kê hoặc các mô hình toán học truyền thống. Mặc dù các phương pháp này có hiệu quả nhất định trong những trường hợp đơn giản, nhưng chúng thường gặp nhiều hạn chế khi phải xử lý những dữ liệu lớn, phức tạp hoặc trong những điều kiện biến đổi thất thường của khí hậu. Với sự phát triển của công nghệ dữ liệu lớn (Big Data) và học máy (Machine Learning), các phương pháp dự báo truyền thống đang dần được thay thế bởi những mô hình dự báo hiện đại hơn, có khả năng học hỏi và rút ra các quy luật từ dữ liệu để đưa ra dự đoán chính xác hơn.

Chính vì thấy được lợi ích của việc dự báo trong nhiều lĩnh vực nên em chọn đề tài: “**Xây dựng mô hình dự báo mực nước sử dụng mô hình học máy với cơ chế Attention**”. Việc ứng dụng cơ chế Attention hy vọng sẽ giúp mô hình dự báo chú trọng vào các yếu tố quan trọng nhất trong dữ liệu, từ đó nâng cao độ chính xác và góp phần cung cấp cơ sở cho những quyết định chính sách đúng đắn trong thời gian tới.

# LỜI CẢM ƠN

Em xin gửi lời cảm ơn chân thành và sự tri ân sâu sắc đối với thầy **Nghiêm Văn Tính**, giảng viên của trường Đại Học Kỹ Thuật Công Nghiệp Thái Nguyên đã nhiệt tình hướng dẫn hướng dẫn em hoàn thành tốt báo cáo này.

Em xin gửi lời cảm ơn chân thành tới các thầy, các cô là giảng viên khoa điện tử cùng các thầy, cô bộ môn Công Nghệ Thông Tin trường Đại Học Kỹ Thuật Công Nghiệp Thái Nguyên đã tận tình truyền đạt kiến thức và hướng dẫn cho em trong suốt quá trình học tập tại trường.

Trong quá trình làm báo cáo khó tránh khỏi sai sót, rất mong các thầy, cô bỏ qua, em rất mong nhận được ý kiến góp ý của thầy cô bộ môn để em rút kinh nghiệm cho tương lai.

**Em xin chân thành cảm ơn!**

**Sinh viên thực hiện**

Cao

Nguyễn Duy Cao

# MỤC LỤC

[LỜI MỞ ĐẦU 5](#_Toc182296759)

[LỜI CẢM ƠN 6](#_Toc182296760)

[MỤC LỤC 7](#_Toc182296761)

[CHƯƠNG 1: GIỚI THIỆU VỀ ĐỀ TÀI 9](#_Toc182296762)

[1.1 Lý do chọn đề tài 9](#_Toc182296763)

[1.2. Mục tiêu đề tài 9](#_Toc182296764)

[1.3. Đối tượng nghiên cứu 10](#_Toc182296765)

[1.4. Bài toán nghiên cứu 10](#_Toc182296766)

[CHƯƠNG 2: TỒNG QUAN VỀ HỌC MÁY 11](#_Toc182296767)

[2.1. Định nghĩa về học máy 11](#_Toc182296768)

[2.2.Cách hoạt động của học máy 11](#_Toc182296769)

[2.3.Các loại học máy 12](#_Toc182296770)

[2.3.1.Học máy có giám sát(Supervised Machine Learning) 12](#_Toc182296771)

[2.3.2.Học máy không giám sát(Unsupervised Learning) 13](#_Toc182296772)

[2.3.3.Học máy bán giám sát(Unsupervised Learning) 13](#_Toc182296773)

[2.4.Ưu nhược điểm của học máy 14](#_Toc182296774)

[2.4.1.Ưu điểm 14](#_Toc182296775)

[2.4.2.Nhược điểm 14](#_Toc182296776)

[2.5.Ứng dụng 14](#_Toc182296777)

[CHƯƠNG 3: PHƯƠNG PHÁP XÂY DỰNG MÔ HÌNH DỰ BÁO 16](#_Toc182296778)

[3.1. Long Short-Term Memory (LSTM) 16](#_Toc182296779)

[3.1.1. Tổng quan 16](#_Toc182296780)

[3.1.2. Công thức cập nhật của LSTM 16](#_Toc182296781)

[3.1.3. Ưu điểm của LSTM 17](#_Toc182296782)

[3.1.4. Nhược điểm của LSTM 17](#_Toc182296783)

[3.1.5. Ứng dụng của LSTM 17](#_Toc182296784)

[3.2.Cơ chế Attention 17](#_Toc182296785)

[3.2.1. Scaled Dot-Product Attention 17](#_Toc182296786)

[3.2.2. Seft-Attention 18](#_Toc182296787)

[3.2.3. Global Self-Attention 19](#_Toc182296788)

[3.2.4. DropAttention 20](#_Toc182296789)

[3.2.5. Multi-Head Attention 20](#_Toc182296790)

[3.3. Độ đo đánh giá chất lượng mô hình 21](#_Toc182296791)

[Kết Luận 23](#_Toc182296792)

[3.4.Phương pháp xây dựng mô hình dự báo mực nước ở các trạm 24](#_Toc182296793)

[CHƯƠNG 4: XÂY DỰNG MÔ HÌNH DỰ ĐOÁN MỰC NƯỚC CÁC TRẠM 25](#_Toc182296794)

[1. Mô hình dự đoán giá mực nước 25](#_Toc182296795)

[1.1 Đọc Dữ Liệu 25](#_Toc182296796)

[1.2 Tiền Xử Lý Dữ Liệu 25](#_Toc182296797)

[1.3 Tạo Tập Dữ Liệu Huấn Luyện 26](#_Toc182296798)

[1.4 Tạo Mô Hình Dự Đoán 27](#_Toc182296799)

[1.5 Huấn Luyện Mô Hình và Dự Báo 30](#_Toc182296800)

[1.6 Chuẩn Bị Dữ Liệu Để Vẽ Biểu Đồ 30](#_Toc182296801)

[1.7 Kết quả dự đoán 31](#_Toc182296802)

[2.Đánh giá chất lượng mô hình 32](#_Toc182296803)

[KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 35](#_Toc182296804)

[TÀI LIỆU THAM KHẢO 36](#_Toc182296805)

# CHƯƠNG 1: GIỚI THIỆU VỀ ĐỀ TÀI

## 1.1 Lý do chọn đề tài

Thiên tai như sạt lở và lũ quét không chỉ gây thiệt hại nghiêm trọng về người và tài sản tại Việt Nam mà còn ảnh hưởng lớn đến nhiều quốc gia trên thế giới. Việc biến đổi khí hậu đã khiến các hiện tượng thời tiết cực đoan như mưa lớn và lũ lụt xảy ra ngày càng thường xuyên, gây áp lực lớn lên các cơ quan quản lý và cộng đồng dân cư trong công tác phòng chống thiên tai. Việc dự báo sớm mực nước sông đóng vai trò quan trọng trong việc dự đoán cấp độ lũ, từ đó giúp xây dựng các phương án ứng phó kịp thời và giảm thiểu thiệt hại. Điều này không chỉ bảo vệ tài sản và tính mạng người dân mà còn góp phần vào sự phát triển bền vững của xã hội.

Đề tài này tập trung vào việc xây dựng mô hình dự báo mực nước sông dựa trên Long Short-Term Memory (LSTM) kết hợp với cơ chế Attention. LSTM là một trong những mô hình học sâu nổi bật về khả năng xử lý và dự báo chuỗi thời gian, phù hợp với đặc tính của dữ liệu mực nước. Kết hợp với cơ chế Attention, mô hình có khả năng tập trung vào những yếu tố quan trọng trong dữ liệu đầu vào, từ đó cải thiện độ chính xác và độ tin cậy của dự báo. Việc nghiên cứu và phát triển mô hình này hy vọng sẽ góp phần tạo ra một công cụ dự báo hiệu quả, hỗ trợ công tác phòng chống thiên tai và quản lý nguồn nước một cách bền vững.

## 1.2. Mục tiêu đề tài

* Tìm hiểu và nghiên cứu mô hình Long Short-Term Memory (LSTM) và cơ chế Attention.
* Xây dựng mô hình dự báo mực nước của Hà Nội dựa vào lượng nước vào và xả ra từ các đập thủy điện.
* Làm thực nghiệm để đánh giá chất lượng của mô hình dự báo.

## 1.3. Đối tượng nghiên cứu

* Đối tượng nghiên cứu: Phương pháp Long Short-Term Menory (LSTM) và cơ chế Attention, lượng nước vào, lượng nước xả của các trạm thủy điện Tuyên Quang, Thác Bà, Hòa Bình và mực nước tại trạm Hà Nội.
* Phạm vi nghiên cứu: lượng nước vào, lượng nước xả của các trạm thủy điện Tuyên Quang, Thác Bà, Hòa Bình và mực nước tạo trạm Hà Nội từ ngày 01/01/2008 đến ngày 31/12/2016.

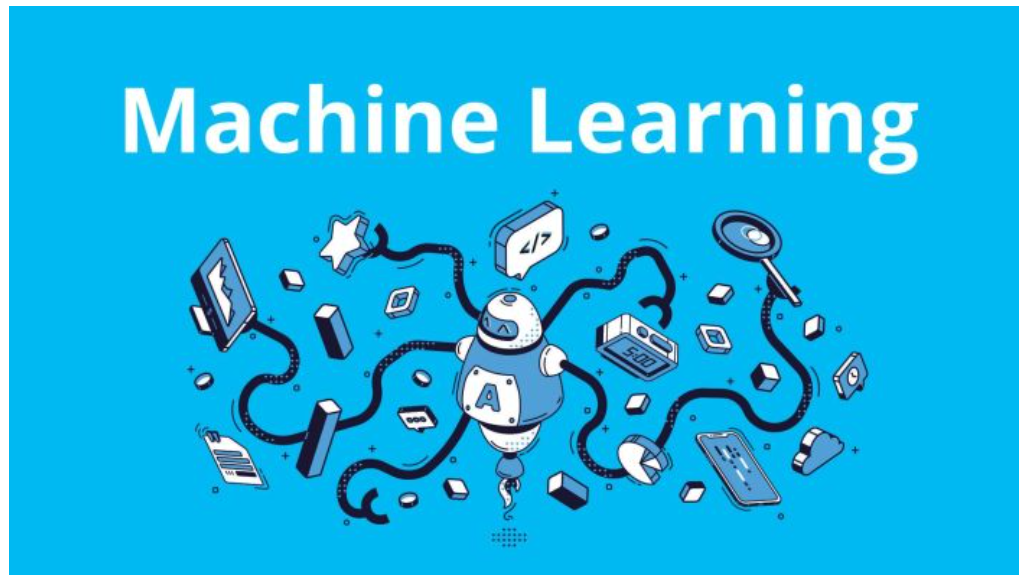
## 1.4. Bài toán nghiên cứu

# CHƯƠNG 2: TỒNG QUAN VỀ HỌC MÁY

## 2.1. Định nghĩa về học máy

**Học máy** (Machine Learning) là một lĩnh vực thuộc trí tuệ nhân tạo, tập trung vào việc nghiên cứu và phát triển các kỹ thuật cho phép hệ thống “học” tự động từ dữ liệu để giải quyết các vấn đề cụ thể.

Hiểu đơn giản, thuật ngữ này nói tới việc con người dạy máy tính nâng cao khả năng thực hiện các tác vụ cụ thể. Cụ thể là cung cấp các dữ liệu và thuật toán có sẵn để máy tính đưa ra dự đoán hoặc tự ra các quyết định. Thông thường, con người chỉ cần lập trình phần mềm với các dòng lệnh cụ thể để máy tính hiểu và thực hiện. Với Machine Learning, máy tính sẽ tự “học” cách giải quyết công việc thông qua những dữ liệu đã được thu thập và cung cấp.



**2.2.Cách hoạt động của học máy**

Quy trình triển khai thuật toán **học máy**thường bao gồm 6 bước như sau:

**Bước 1**: Thu thập dữ liệu (Gathering data/Data collection)

**Bước 2**: Tiền xử lý dữ liệu (Data preprocessing):

* Trích xuất dữ liệu – Data extraction
* Làm sạch dữ liệu – Data cleaning
* Chuyển đổi dữ liệu – Data transformation
* Chuẩn hóa dữ liệu – Data normalization
* Trích xuất đặc trưng – Feature extraction

**Bước 3**: Phân tích dữ liệu (Data analysis)

**Bước 4**: Xây dựng mô hình máy học (Model building)

**Bước 5**: Huấn luyện mô hình (Model training)

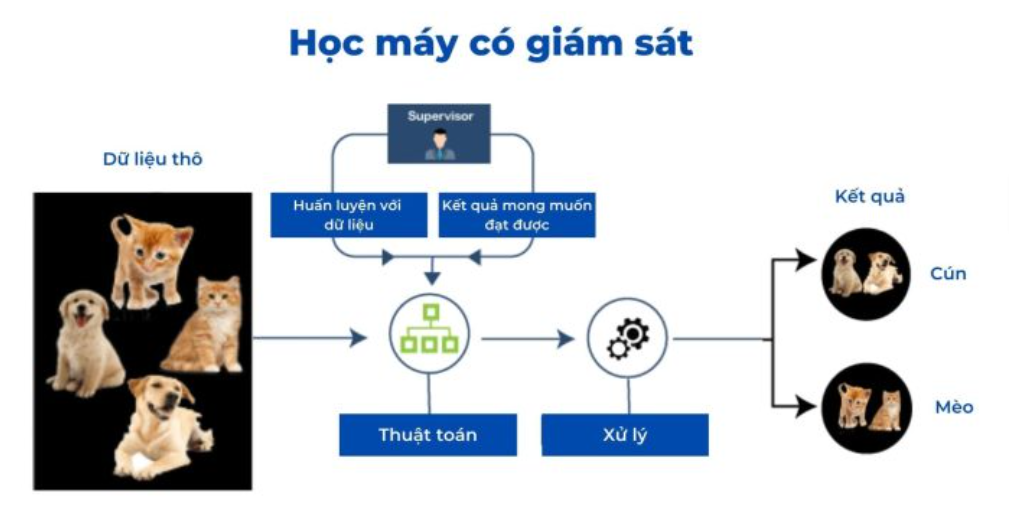
**Bước 6**: Đánh giá mô hình (Model evaluation)

**2.3.Các loại học máy**

**2.3.1.Học máy có giám sát(Supervised Machine Learning)**

**Học máy** có giám sát (Supervised Machine Learning) là phương pháp mà trong đó máy tính được học từ dữ liệu đã được đánh dấu trước, để phát triển các thuật toán có khả năng phân loại hoặc dự đoán kết quả một cách chính xác. Kỹ thuật này thường được áp dụng cho các bài toán phân lớp (Classification).

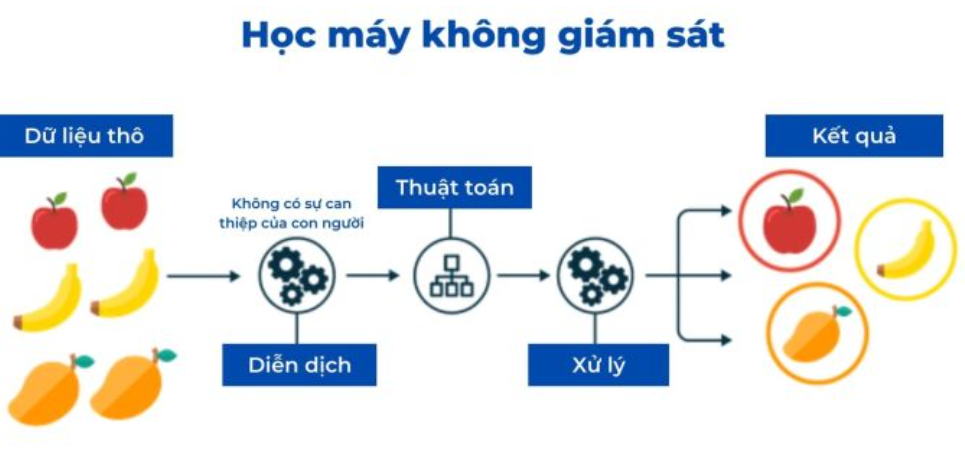
**Ví dụ**: Chúng ta có một tập dữ liệu gồm các hình ảnh của chó và mèo. Trong giai đoạn đầu, máy tính được huấn luyện để nhận biết các hình ảnh này. Khi quá trình huấn luyện hoàn tất, máy tính có thể nhận diện và dự đoán chính xác loại động vật trong hình ảnh mới được cung cấp, dựa trên việc phân tích các đặc điểm như hình dạng và màu sắc. Quá trình này chính là nhận dạng đối tượng trong máy học có giám sát.

****

**2.3.2.Học máy không giám sát(Unsupervised Learning)**

Trái ngược với **học máy** có giám sát, máy học không giám sát (Unsupervised Learning) sử dụng thuật toán để phân tích và phân cụm các dữ liệu không có nhãn. Phương pháp này tự động tìm kiếm các mô hình và cấu trúc ẩn trong dữ liệu mà không yêu cầu sự hỗ trợ từ con người.

**Ví dụ**: Máy học không giám sát nổi bật với khả năng nhận diện các điểm chung và khác biệt trong dữ liệu, làm cho nó trở thành công cụ hữu ích trong việc phân tích dữ liệu khám phá, phân loại khách hàng, phát triển chiến lược bán chéo (Cross-sell), nhận dạng hình ảnh, và nhiều ứng dụng khác.



**2.3.3.Học máy bán giám sát(Unsupervised Learning)**

Trái ngược với **học máy** có giám sát, máy học không giám sát (Unsupervised Learning) sử dụng thuật toán để phân tích và phân cụm các dữ liệu không có nhãn. Phương pháp này tự động tìm kiếm các mô hình và cấu trúc ẩn trong dữ liệu mà không yêu cầu sự hỗ trợ từ con người.

Học bán giám sát kết hợp một lượng nhỏ dữ liệu có gán nhãn và một lượng lớn dữ liệu không gán nhãn để huấn luyện mô hình. Phương pháp này hữu ích khi việc gán nhãn dữ liệu tốn kém hoặc khó khăn.

## 2.4.Ưu nhược điểm của học máy

### **2.4.1.Ưu điểm**

* ***Xác định xu hướng dữ liệu dễ dàng* :** Máy học có thể xử lý lượng lớn dữ liệu và tự động phát hiện các xu hướng mà con người khó có thể nhận ra.
* ***Khả năng tự động hóa cao* :** Sau khi thiết lập, mô hình máy học có thể tự động học hỏi, đưa ra dự đoán và cải thiện hiệu suất mà không cần quá nhiều sự can thiệp từ con người.
* ***Xử lý đa dạng dữ liệu* :** Có thể xử lý nhiều định dạng dữ liệu trong môi trường dữ liệu linh hoạt, khối lượng lớn và phức tạp.

### **2.4.2.Nhược điểm**

**Phụ thuộc vào việc thu thập dữ liệu**

* Cần có tập dữ liệu lớn, đáng tin cậy, không thiên vị và chất lượng tốt để đào tạo mô hình.
* Phải đợi dữ liệu mới được xử lý để đưa vào huấn luyện mô hình học máy.

**Vẫn có khả năng xảy ra lỗi**

* Kết quả máy học có thể không chính xác hoàn toàn do lỗi trong quá trình đào tạo và kiểm tra dữ liệu.
* Việc gỡ lỗi các mô hình máy học phức tạp và tốn thời gian.

## 2.5.Ứng dụng

Các doanh nghiệp/tổ chức có thể ứng dụng Machine Learning trong các lĩnh vực cụ thể sau:

**Sản xuất**

* Hỗ trợ bảo trì dự đoán
* Kiểm soát chất lượng
* Nghiên cứu đổi mới
* Cải thiện các giải pháp hậu cần như quản lý tài sản, chuỗi cung ứng và kho hang

**Chăm sóc sức khỏe và khoa học đời sống**

* Phân tích dữ liệu y tế để hỗ trợ chẩn đoán và điều trị bệnh.
* Phát triển giải pháp phát hiện khối u ung thư và chẩn đoán bệnh về mắt.
* Tác động tới kết quả chăm sóc sức khỏe cho bệnh nhân.

**Dịch vụ tài chính**

* Phân tích rủi ro và quy định hợp lý..
* Phân tích sự biến động của các sàn giao dịch chứng khoán.
* Đánh giá các quỹ phòng hộ và hiệu chỉnh các danh mục tài chính.
* Chống hành vi gian lận và vay nợ có rủi ro cao

**Bán lẻ**

* Chất lượng dịch vụ không ngừng cải thiện theo từng ngày.
* Lượng hàng tồn kho được quản lý hiệu quả
* Tăng cường doanh số bán hàng và thực hiện chiến lược tiếp thị qua nhiều kênh.

**Truyền thông và giải trí**

* Tăng cường hiểu biết về khách hàng mục tiêu
* Cung cấp nội dung chân thực, cá nhân hóa và theo nhu cầu khách hàng.
* Hỗ trợ thiết kế trailer, quảng cáo và đề xuất nội dung phù hợp với từng người dùng
* Tối ưu hóa quy trình sản xuất.

# CHƯƠNG 3: PHƯƠNG PHÁP XÂY DỰNG MÔ HÌNH DỰ BÁO

## 3.1. Long Short-Term Memory (LSTM)

### **3.1.1. Tổng quan**

LSTM có cấu trúc phức tạp hơn so với RNN truyền thống, bao gồm ba cổng chính:

Cổng đầu vào (Input Gate): Quyết định thông tin nào sẽ được thêm vào trạng thái tế bào. Cổng quên (Forget Gate): Quyết định thông tin nào sẽ bị loại bỏ khỏi trạng thái tế bào.

Cổng đầu ra (Output Gate): Quyết định thông tin nào sẽ được xuất ra từ trạng thái tế bào.

### **3.1.2. Công thức cập nhật của LSTM**

Một cell LSTM có ba thành phần chính: cổng quên (Forget gate), cổng đầu vào (Input gate), và cổng đầu ra (Output gate).

* Cổng quên (Forget gate) là cổng sẽ quyết định xem thông tin nào sẽ bị bỏ qua và có công thức như sau:

𝑓𝑡 = 𝜎(𝑊𝑓 . [ℎ𝑡−1, 𝑥𝑡] + 𝑏𝑓)

* Cổng đầu vào (Input gate) là cổng sẽ quyết định bao nhiêu thông tin mới được lưu trữ và nhập vào. Và có công thức như sau:

𝑖𝑡 = 𝜎(𝑊𝑖 . [ℎ𝑡−1,] + 𝑏𝑖)

* Cập nhật trạng thái của tế bào:

𝐶𝑡 = 𝑓𝑡.𝐶𝑡−1 + 𝑖𝑡. 𝐶̃𝑡

* Cổng đầu ra: Quyết định phần nào của trạng thái cell sẽ được xuất ra làm trạng thái ẩn tiếp theo

𝑜𝑡 = 𝜎(𝑊𝑜 . [ℎ𝑡−1, 𝑥𝑡] + 𝑏𝑜)

* Có trạng thái ẩn như sau

ℎ𝑡 = 𝑜𝑡 . tanh(𝐶𝑡)

### **3.1.3. Ưu điểm của LSTM**

Khả năng ghi nhớ lâu dài: LSTM có khả năng lưu trữ và sử dụng thông tin từ xa trong chuỗi dữ liệu, rất hữu ích cho các bài toán cần nhận diện các mẫu lâu dài.

Giảm thiểu vanishing gradient: Cấu trúc cổng giúp kiểm soát thông tin được truyền qua lại, hạn chế vấn đề vanishing gradient gặp phải ở RNN truyền thống.

### **3.1.4. Nhược điểm của LSTM**

Độ phức tạp: Cấu trúc phức tạp của LSTM dẫn đến số lượng tham số lớn hơn, làm cho quá trình huấn luyện lâu hơn và tốn kém tài nguyên hơn.

Khó khăn trong việc điều chỉnh siêu tham số: Việc lựa chọn và điều chỉnh các siêu tham số có thể phức tạp và cần thử nghiệm nhiều lần để có thể chọn ra.

### **3.1.5. Ứng dụng của LSTM**

Dự đoán chuỗi thời gian: Dự đoán giá cổ phiếu, nhu cầu sản phẩm, và dự đoán thời tiết.

Xử lý ngôn ngữ tự nhiên: Phân loại văn bản, dịch máy, và phân tích cảm xúc.

Nhận dạng giọng nói: Chuyển đổi giọng nói thành văn bản và nhận diện mẫu âm thanh.

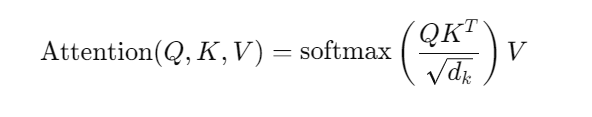
## 3.2.Cơ chế Attention

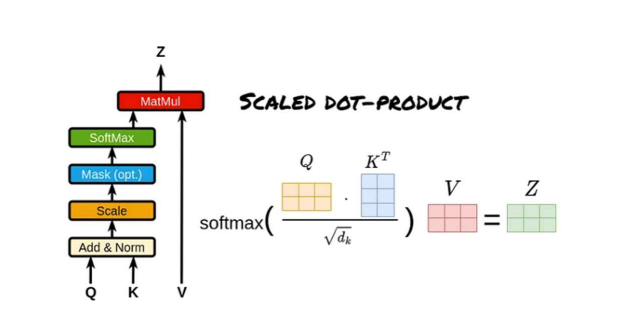
Cơ chế chú ý có thể mô tả như là ánh xạ một truy vấn và tập hợp các Key –

Value đến một đầu ra , trong đó Query, Key, Value và đầu ra đều là các vector [13]. Đầu ra được tính toán dưới dạng tổng có trọng số của Value, trong đó trọng số được gán cho mỗi giá trị được tính toán bằng một hàm tương thích của truy vấn với khóa tương ứng.

### **3.2.1. Scaled Dot-Product Attention**

Trong mô hình **Transformer** đã loại bỏ đi hoàn toàn khái niệm các vector hidden, memori và thay thế chúng bằng ba vector **query, keys, values**. Kết quả đầu ra bây giờ được tính bằng tổng các giá trị **values** đã được đánh trọng số. Trọng số này chính là hàm softmax tính dựa trên **query** và **key** tương ứng. Công thức tính attention weight có tên là **Scaled Dot-Product Attention**.

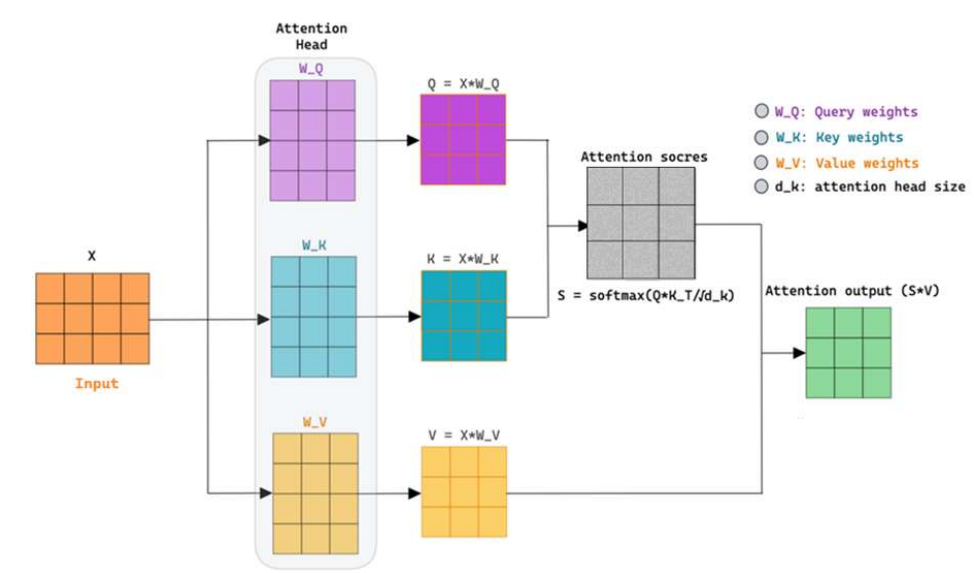




### **3.2.2. Seft-Attention**

**Self-Attention** là một cơ chế trong đó mỗi phần tử trong chuỗi đầu vào "chú ý" đến tất cả các phần tử khác trong chuỗi, bao gồm chính nó. Nó tính toán trọng số attention cho từng phần tử đầu vào dựa trên tất cả các phần tử khác trong chuỗi, giúp nắm bắt mối quan hệ dài hạn trong chuỗi dữ liệu.

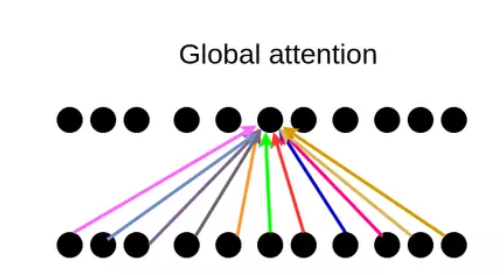
* **Ứng dụng chính**: Trong mô hình Transformer, Self-Attention là cốt lõi giúp nắm bắt mối quan hệ giữa các từ trong câu ở cả ngữ cảnh gần và xa.
* **Ưu điểm**: Tăng khả năng mô hình hiểu được các phụ thuộc dài hạn, đặc biệt hiệu quả trong các chuỗi dữ liệu như văn bản



### **3.2.3. Global Self-Attention**

**Global Self-Attention** là một biến thể của Self-Attention nhưng mở rộng phạm vi "chú ý" ra toàn bộ chuỗi đầu vào. Tức là, nó cho phép mỗi phần tử trong chuỗi chú ý không chỉ các phần tử kế cận mà cả các phần tử ở bất kỳ khoảng cách nào. Nó đặc biệt hữu ích trong các trường hợp mà các mối quan hệ dài hạn trong chuỗi có ảnh hưởng mạnh đến kết quả.

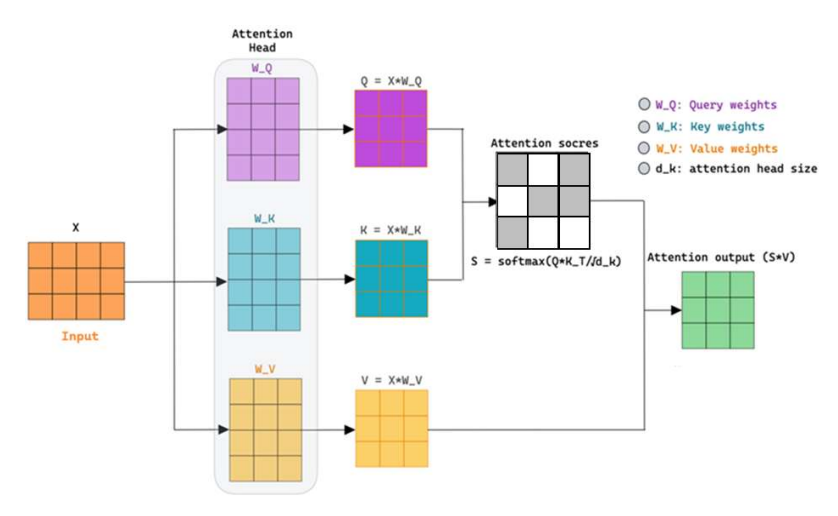
* **Đặc điểm**: Không bị giới hạn trong một phạm vi cục bộ, do đó có thể nắm bắt mối quan hệ toàn cục.
* **Ứng dụng**: Rất hữu ích cho các mô hình dự đoán với dữ liệu có quan hệ dài hạn như thời gian, chuỗi sự kiện.



### **3.2.4. DropAttention**

**DropAttention** là một biến thể của cơ chế attention, giúp giảm bớt các thành phần chú ý không cần thiết để tránh quá tải tính toán. Cơ chế này chỉ giữ lại một phần của các trọng số attention và bỏ qua các trọng số có mức độ chú ý thấp hơn, tương tự cách Dropout hoạt động.

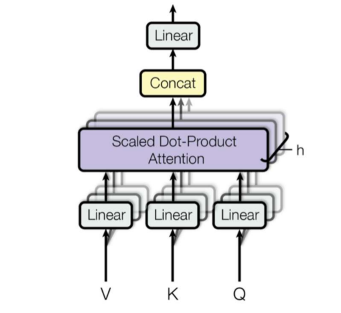
* **Mục tiêu**: Giảm chi phí tính toán và tránh overfitting bằng cách loại bỏ các giá trị chú ý có trọng số thấp.
* **Cách hoạt động**: Thông thường, DropAttention sẽ đặt các trọng số chú ý có giá trị thấp về 0, chỉ giữ lại các trọng số có giá trị lớn hơn ngưỡng nhất định.



### **3.2.5. Multi-Head Attention**

**Multi-Head Attention** là một kỹ thuật mở rộng của Self-Attention. Thay vì tính toán attention một lần duy nhất, Multi-Head Attention thực hiện nhiều phép tính attention song song trên các "đầu attention" (heads) khác nhau, sau đó kết hợp chúng lại để tăng cường khả năng nắm bắt các đặc trưng phức tạp của dữ liệu.

* **Ưu điểm**: Giúp mô hình có thể "nhìn" vào dữ liệu từ nhiều góc độ khác nhau cùng lúc, nắm bắt được các đặc trưng ở các mức độ khác nhau.
* **Ứng dụng**: Là thành phần chủ chốt trong mô hình Transformer, cho phép mô hình xử lý các quan hệ phức tạp giữa các phần tử trong chuỗi.



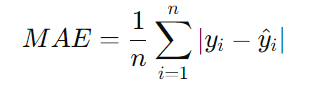
## 3.3. Độ đo đánh giá chất lượng mô hình

Mô hình được đánh giá bằng cách sử dụng các chỉ số hiệu suất như MAE (Mean Absolute Error), Nash–Sutcliffe efficiency (NSE),RMSE (Root Mean Squared Error) và Coefficient of determination () trên tập kiểm tra. Các dự đoán được chuyển đổi về phạm vi giá trị gốc để đánh giá chính xác hiệu suất của mô hình. So sánh các mô hình giúp xác định mô hình nào hoạt động tốt nhất và cần điều chỉnh hay cải thiện gì.

Để đánh giá hiệu suất của một mô hình dự đoán, chúng ta thường sử dụng các chỉ số như MAE, MSE và RMSE. Dưới đây là các công thức và ý nghĩa của từng chỉ số:

**a. Mean Absolute Error (MAE)**

* **Công thức:**

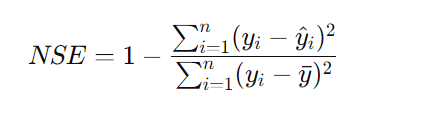


**Ý nghĩa:**

* MAE đo lường sai số trung bình của các dự đoán so với giá trị thực tế mà không tính đến chiều của sai số (sai số âm hoặc dương đều tính như nhau).
* MAE là một số không âm, giá trị càng nhỏ thì mô hình càng chính xác.

**b. Nash–Sutcliffe efficiency (NSE)**

* **Công thức:**

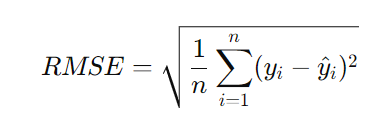


**Ý nghĩa:**

* NSE đánh giá mức độ chính xác của dự đoán so với giá trị trung bình của dữ liệu thực tế. Chỉ số này thường được sử dụng trong các mô hình dự báo chuỗi thời gian.
* NSE dao động từ -∞ đến 1. Giá trị NSE càng gần 1 thì mô hình càng tốt. NSE < 0 nghĩa là mô hình dự đoán kém hơn dự đoán dựa trên trung bình của dữ liệu thực.

**c. Root Mean Squared Error (RMSE)**

* **Công thức:**

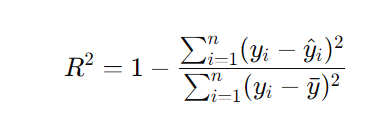


**Ý nghĩa:**

* RMSE là độ đo lỗi phổ biến, nhấn mạnh vào các sai số lớn vì nó bình phương các sai số trước khi tính trung bình. RMSE lớn hơn khi có những sai số lớn, cho thấy mức độ nhạy cảm với các điểm dữ liệu bất thường.
* RMSE là một số không âm, giá trị càng nhỏ thì mô hình càng chính xác.

**d. Coefficient of determination ()**

**Công thức :**



**Ý nghĩa :**

* R² cho biết tỷ lệ phương sai của biến đầu ra (thực tế) được giải thích bởi biến đầu vào trong mô hình.
* Giá trị R² nằm trong khoảng từ 0 đến 1 (đôi khi có thể nhỏ hơn 0 nếu mô hình kém). Giá trị càng gần 1 thì mô hình càng tốt vì nó giải thích được nhiều phương sai của dữ liệu thực tế.

**So sánh và lựa chọn mô hình:**

* Sau khi tính toán các chỉ số MAE, NSE, RMSE và chúng ta sẽ so sánh chúng để đánh giá mô hình nào hoạt động tốt nhất.
* Một mô hình tốt là mô hình có các chỉ số MAE, RMSE càng nhỏ càng tốt cho thấy sai số thấp và độ chính xác cao . Đồng thời, chỉ số **R²** càng gần 1 và **NSE** càng lớn thì mô hình càng có khả năng giải thích biến thiên của dữ liệu một cách chính xác.
* Việc chuyển đổi các dự đoán về phạm vi giá trị gốc (nếu cần thiết) trước khi tính toán các chỉ số này là rất quan trọng để đảm bảo tính chính xác trong đánh giá hiệu suất mô hình.

### **Kết Luận**

Dựa trên kết quả đánh giá, đưa ra kết luận về hiệu suất của các mô hình dự đoán. Đề xuất các bước tiếp theo để cải thiện mô hình nếu cần, chẳng hạn như thu thập thêm dữ liệu, điều chỉnh cấu trúc mô hình hoặc thử nghiệm các kỹ thuật khác. Sử dụng các công cụ trực quan để hiển thị biểu đồ và thông tin liên quan, giúp đánh giá kết quả dự đoán một cách trực quan và rõ ràng.

## 3.4.Phương pháp xây dựng mô hình dự báo mực nước ở các trạm

Mô hình dự báo mực nước được xây dựng qua các bước sau :

Bước 1 : Thu thập và tiền xử lý dữ liệu

* Đọc tệp dữ liệu huấn luyện(train data)
* Loại bỏ những dữ liệu không cần thiết và sau đó bắt đầu gán dữ liệu sao cho lấy những dữ liệu từ khoảng thời gian để dự đoán cho ngay tiếp theo
* Ta chia thành tập dữ liệu X\_train,y\_train và X\_test,y\_test để dùng cho dữ liệu huấn luyện và để đánh giá chất lượng mô hình

Bước 2 : Huấn luyện mô hình

* Thiết kế mô hình huấn luyện : kiến trúc mạng(LSTM,Attention)
* Sử dụng X\_train,y\_train để huấn luyện mô hình

Bước 3 : Đánh giá mô hình

* Sử dụng mô hình ta đã huấn luyện ở bước 2 và X\_test và dùng các độ đo và công thức để đánh giá chất lượng mô hình này

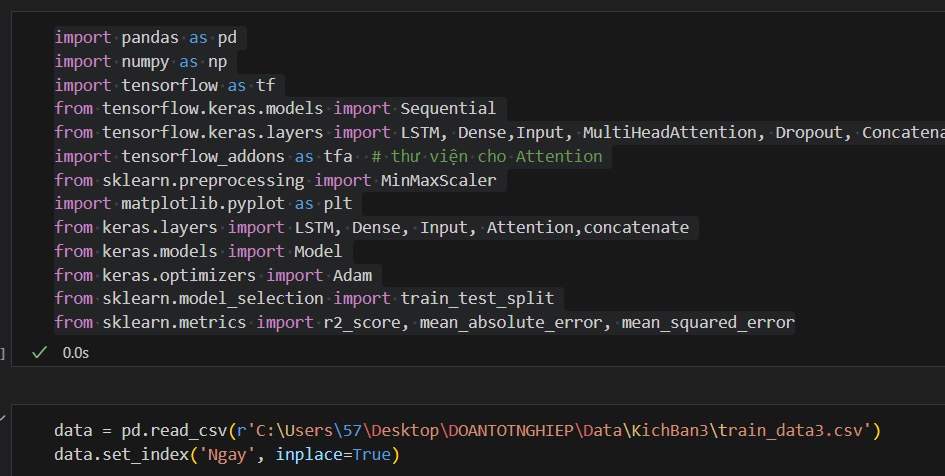
Bước 4 : Sử dụng mô hình

* Dùng mô hình với tệp dữ liệu test\_data để dữ đoán mực nước của từng vùng trong dữ liệu test\_data

# CHƯƠNG 4: XÂY DỰNG MÔ HÌNH DỰ ĐOÁN MỰC NƯỚC CÁC TRẠM

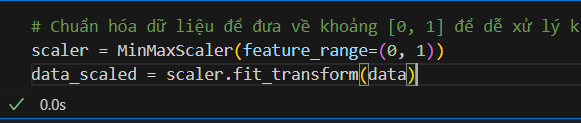
## Mô hình dự đoán giá mực nước

### 1.1 Đọc Dữ Liệu



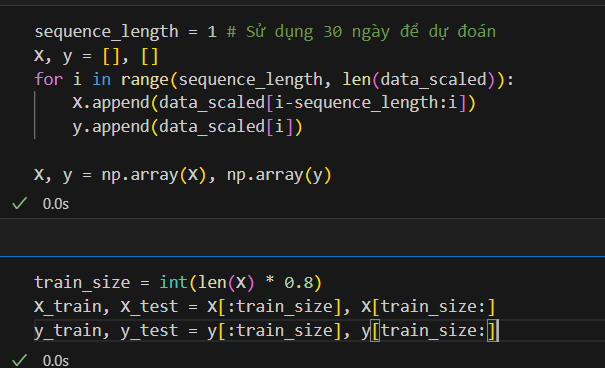
**Giải Thích:** Dữ liệu được đọc từ tệp CSV train\_data3.csv, chứa thông tin về mực nước ở các trạm qua các năm.

### 1.2 Tiền Xử Lý Dữ Liệu



**Giải Thích:** Giá trị mực nước được chuẩn hóa vào khoảng giá trị từ 0 đến 1 bằng MinMaxScaler để dễ dàng xử lý trong mô hình dự đoán.

### 1.3 Tạo Tập Dữ Liệu Huấn Luyện



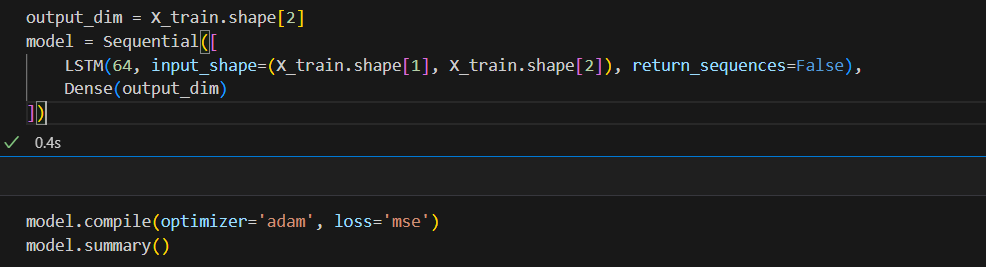
**Giải Thích:**

Trong đoạn mã trên, dữ liệu được chuẩn bị để dự đoán giá trị của một ngày tiếp theo dựa trên một chuỗi ngày trước đó.

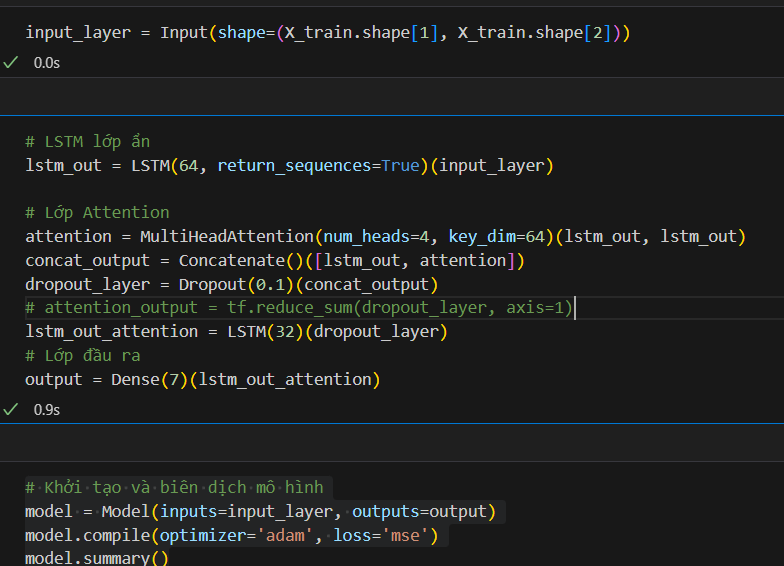
* **sequence\_length = 1**: Đây là độ dài chuỗi ngày dùng để dự đoán.
* **X và y**:
  + **X** đại diện cho các chuỗi con của dữ liệu với độ dài xác định bởi sequence\_length. Mỗi phần tử trong X là dữ liệu của **số ngày trước** dùng để dự đoán.
  + **y** chứa giá trị thực của ngày tiếp theo mà mô hình sẽ dự đoán, tương ứng với từng chuỗi trong X.
* **Chia tập dữ liệu thành tập train và test**: Sau khi xây dựng X và y, tập dữ liệu được chia thành 80% cho tập train (để mô hình học) và 20% cho tập test (để đánh giá hiệu suất mô hình).

### 1.4 Tạo Mô Hình Dự Đoán

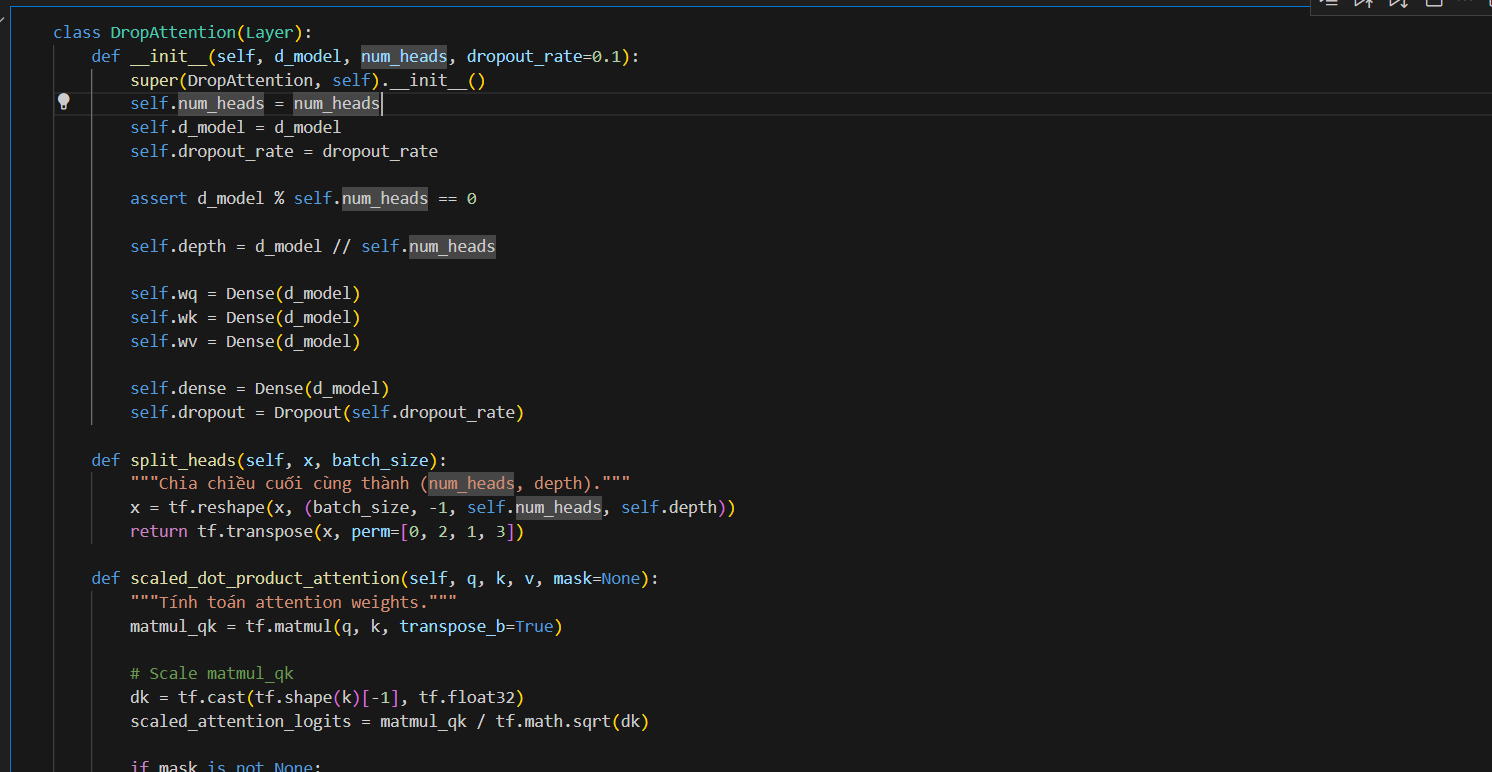
**Mô hình LSTM**

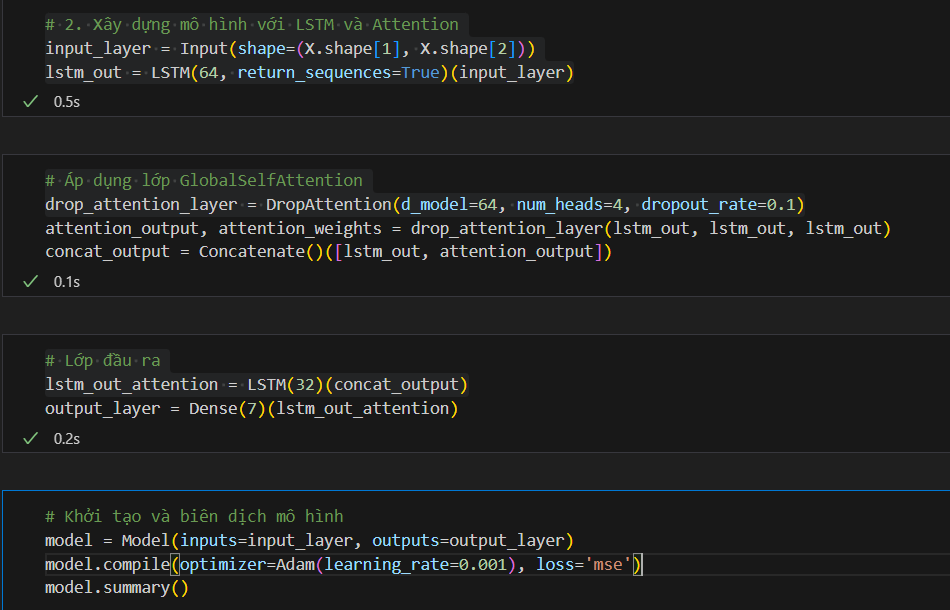


**Mô hình LSTM + MultiHeadAttention**

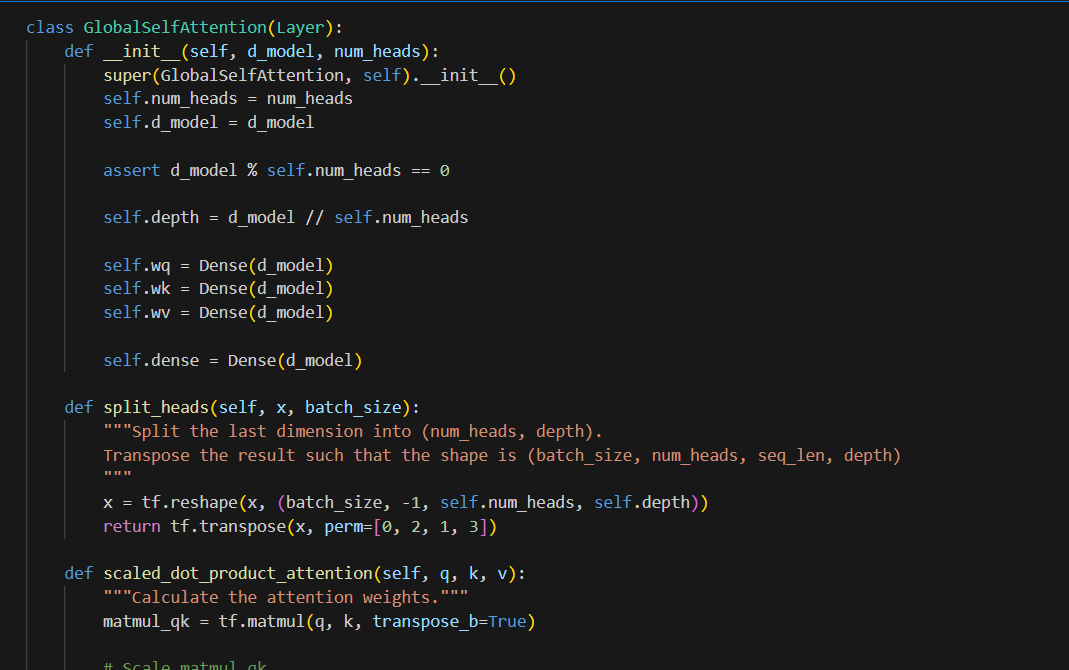


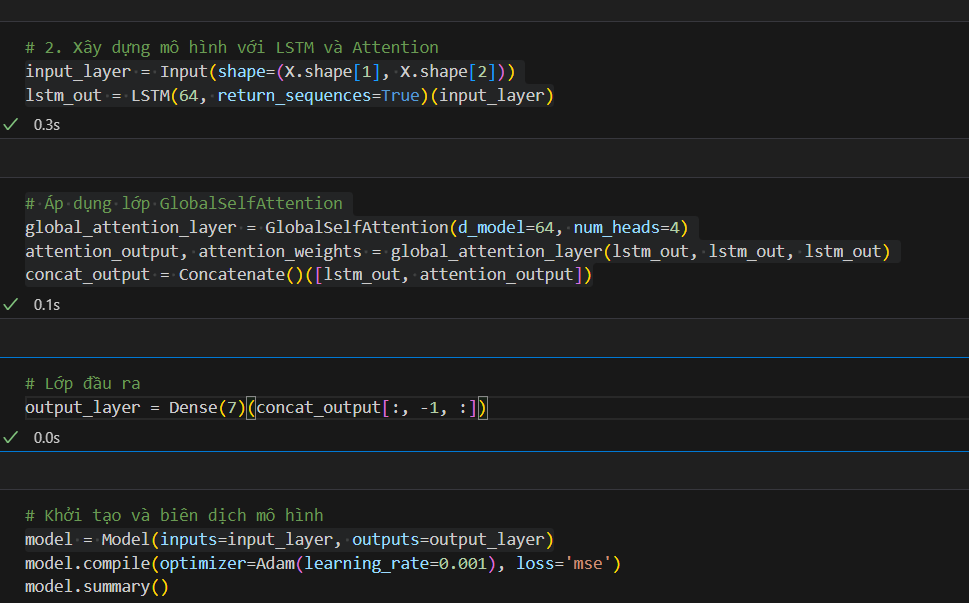
**Mô hình LSTM + DropAttention**



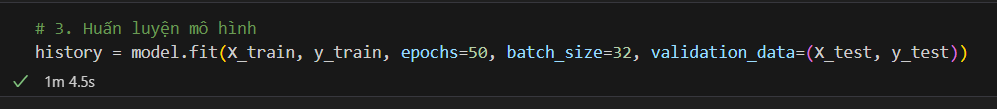


**Mô hình LSTM + Global Self-Attention**



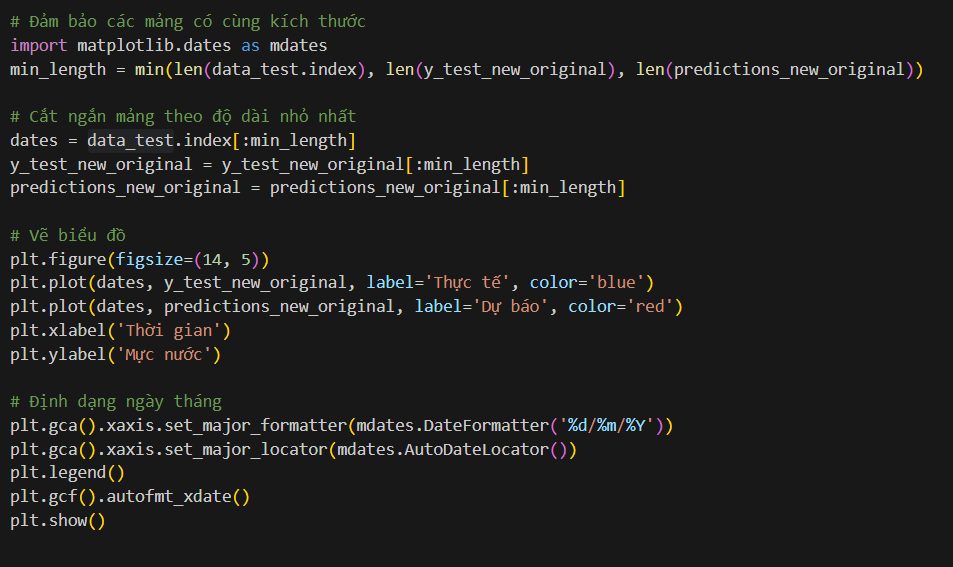
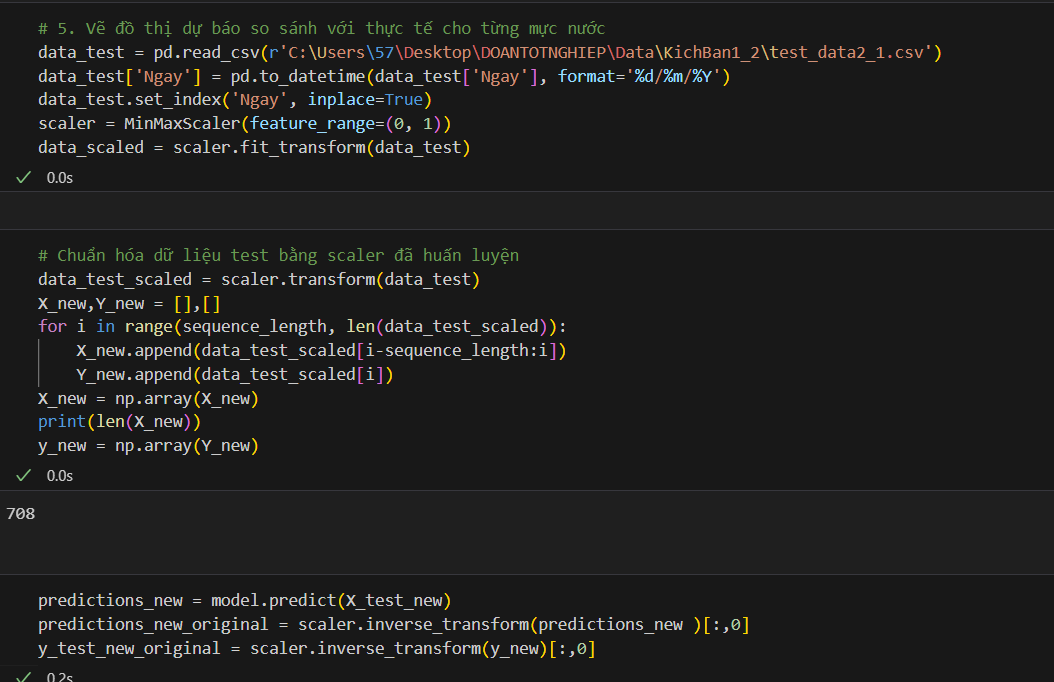


### 1.5 Huấn Luyện Mô Hình và Dự Báo

****

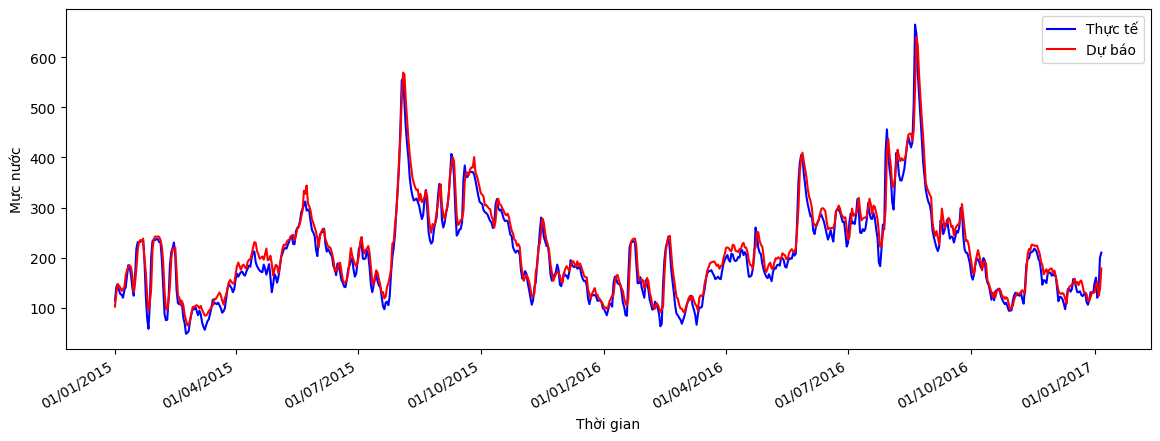
**Giải Thích:** Sau khi tạo mô hình dự đoán cho từng trường hợp ta huấn luyện mô hình cho nó sau 50 vòng lặp liên tục

### 1.6 Chuẩn Bị Dữ Liệu Để Vẽ Biểu Đồ

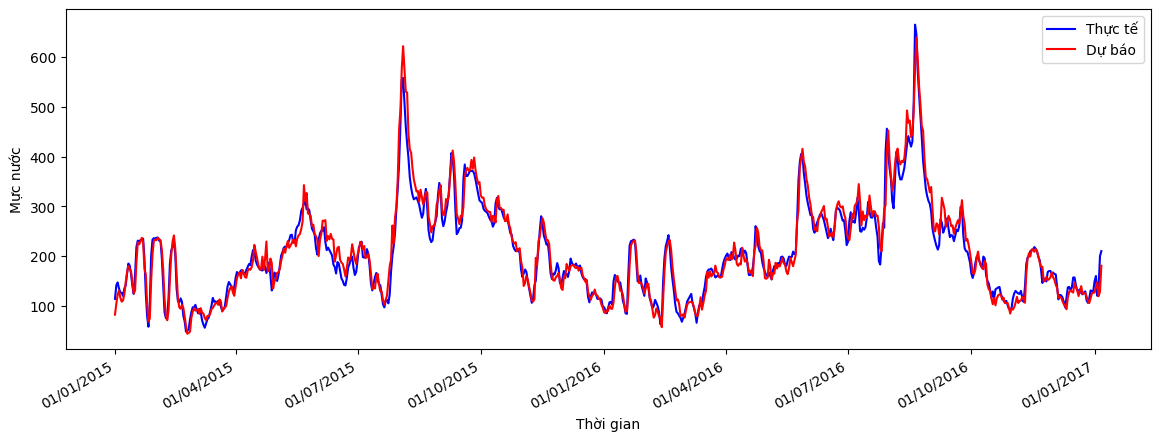
****

### 1.7 Kết quả dự đoán

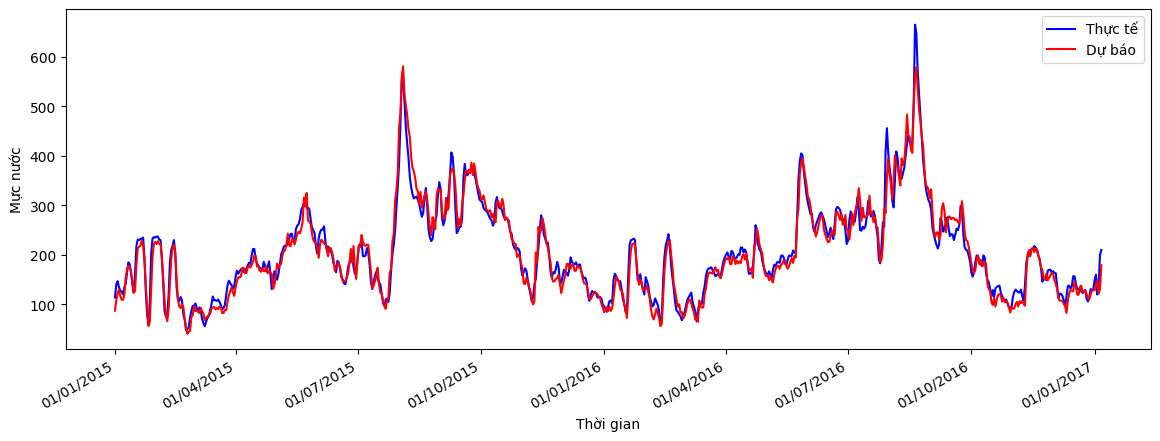
Mô hình LSTM



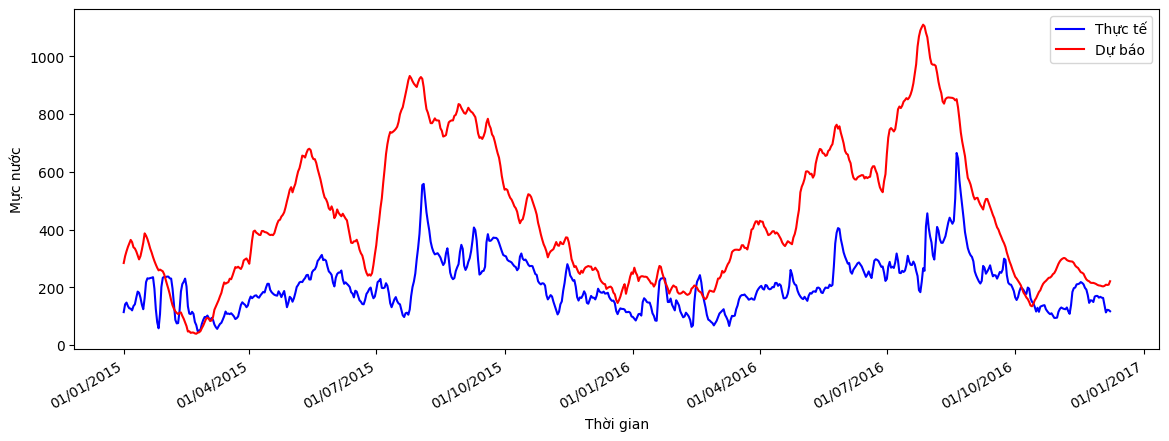
Mô hình LSTM + DropAttention



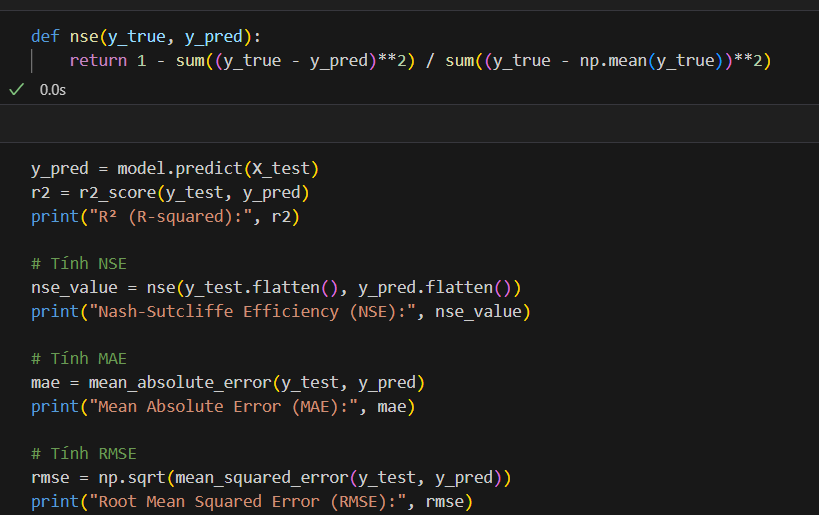
Mô hình LSTM + MultiHeadAttention



Mô hình LSTM + Global Self-Attention



## 2.Đánh giá chất lượng mô hình



**Giải thích:**

* **MAE (Mean Absolute Error)**: Sai số tuyệt đối trung bình.
* **NSE (Nash–Sutcliffe efficiency)**: mức độ chính xác của dự đoán so với giá trị trung bình của dữ liệu thực tế.
* **RMSE (Root Mean Squared Error)**: Căn bậc hai của sai số bình phương trung bình.
* **Coefficient of determination () :** Tỷ lệ phương sai của biến đầu ra

Sau quá trình huấn luyện của tất cả mô hình ta sử dụng đoán code trên để tính toán những tham số dùng đánh giá chất lượng mô hình như sau :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **MAE** | **NSE** | **RMSE** |  |
| LSTM | 0.04217 | 0.8652 | 0.0748 | 0.7928 |
| LSTM+DropAttention | 0.04277 | 0.8650 | 0.07492 | 0.7872 |
| LSTM+MultiHeadAttention | 0.04238 | 0.8654 | 0.0747 | 0.7868 |
| LSTM+GlobalSelfAttention | 0.0445 | 0.8614 | 0.0759 | 0.7723 |

Nhận xét :

* **MAE** : Thấy được LSTM là thấp nhất sau đó tới LSTM+MultiHeadAttention.
* **NSE** : LSTM+MultiHeadAttention đạt NSE cao nhất, chứng tỏ mô hình này có độ chính xác tốt nhất theo chỉ số này.
* **RMSE** **: LSTM+MultiHeadAttention** có RMSE thấp nhất, cho thấy sai số dự đoán trung bình thấp hơn so với các mô hình khác.
* : **LSTM** có R² cao nhất, cho thấy mô hình này giải thích biến thiên của dữ liệu tốt hơn các mô hình khác.

Dựa trên những các chỉ số trên :

* **LSTM** đạt kết quả tốt nhất về MAE và R², cho thấy khả năng dự đoán sát với thực tế và giải thích biến thiên dữ liệu rất tốt.
* **LSTM+MultiHeadAttention** có lợi thế về NSE và RMSE, và kết quả cũng gần với LSTM về các chỉ số khác.

**Kết luận**

* **LSTM** là mô hình phù hợp nhất cho bài toán dự đoán này, vì nó có độ chính xác cao (thể hiện qua MAE và R² tốt nhất) và khả năng giải thích biến thiên trong dữ liệu ổn định. Đây là lựa chọn ưu tiên cho các bài toán đòi hỏi độ chính xác cao.
* **LSTM+MultiHeadAttention** là lựa chọn tốt thứ hai, với kết quả gần như tương đương LSTM ở mọi chỉ số. Đặc biệt, mô hình này đạt điểm cao nhất ở NSE và RMSE, cho thấy tính nhất quán trong việc giảm sai số. **LSTM+MultiHeadAttention** có thể là lựa chọn thay thế hợp lý nếu cần mô hình linh hoạt, có khả năng nắm bắt các mẫu phức tạp hơn với ít tham số hơn.
* **LSTM+DropAttention** cho thấy hiệu suất trung bình, với các chỉ số xấp xỉ nhưng thấp hơn một chút so với LSTM và LSTM+MultiHeadAttention. Mặc dù mô hình này có thể không đạt độ chính xác cao nhất, nhưng vẫn là một lựa chọn cân bằng, đặc biệt trong các bài toán không đòi hỏi độ chính xác quá cao nhưng vẫn cần sự ổn định.
* **LSTM+GlobalSelfAttention** có hiệu suất thấp hơn so với các mô hình khác, đặc biệt thể hiện qua các chỉ số R² và NSE thấp nhất. Điều này cho thấy mô hình gặp khó khăn trong việc nắm bắt và giải thích biến thiên dữ liệu một cách chính xác. Do đó, **LSTM+GlobalSelfAttention** có thể không phải là lựa chọn ưu tiên cho bài toán này, nhưng vẫn có thể áp dụng trong các bài toán yêu cầu ít tài nguyên và độ chính xác dự báo trung bình.

## KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

1. **Kết luận**

Trong nghiên cứu này, em đã triển khai và so sánh các mô hình LSTM và các biến thể kết hợp với các cơ chế Attention khác nhau, bao gồm MultiHeadAttention, DropAttention và GlobalSelfAttention, nhằm dự báo mực nước. Kết quả cho thấy mô hình LSTM đạt hiệu suất tốt nhất, với các chỉ số như MAE, NSE, RMSE và R² đều cao hơn so với các mô hình khác, cho thấy khả năng dự báo chính xác và giải thích tốt các biến động trong dữ liệu.

LSTM+MultiHeadAttention cũng đạt hiệu suất gần tương đương với LSTM, cho thấy đây là một lựa chọn thay thế khả thi, đặc biệt trong các trường hợp yêu cầu mô hình gọn nhẹ, linh hoạt và có khả năng nắm bắt các mối quan hệ phức tạp giữa các biến. Trong khi đó, LSTM+DropAttention và LSTM+GlobalSelfAttention không đạt kết quả cao nhất nhưng vẫn có thể hữu ích cho các bài toán yêu cầu tài nguyên thấp hơn hoặc không cần độ chính xác dự báo tối đa

1. **Hướng phát triển:**
   * **Sử dụng thêm các mô hình học sâu hiện đại**: Các mô hình như Transformer hoặc các phiên bản nâng cao khác của Attention có thể được thử nghiệm nhằm nâng cao hiệu suất, đặc biệt khi xử lý các chuỗi dữ liệu dài và phức tạp hơn.
   * **Tích hợp thêm các yếu tố đầu vào**: Dữ liệu mực nước có thể được kết hợp với các yếu tố khác như lượng mưa, nhiệt độ, độ ẩm và các yếu tố thời tiết khác nhằm tạo nên mô hình dự báo toàn diện và chính xác hơn.
   * **Áp dụng các kỹ thuật giảm nhiễu và xử lý dữ liệu tiền xử lý**: Việc áp dụng các kỹ thuật lọc nhiễu hoặc chọn lọc dữ liệu đầu vào có thể giúp cải thiện chất lượng dự báo, đặc biệt trong các hệ thống phức tạp và biến đổi.

Tổng kết, mặc dù các mô hình hiện tại đã cho thấy hiệu quả trong nhiều tình huống, việc cải thiện chúng bằng cách tích hợp dữ liệu bổ sung và áp dụng các kỹ thuật mô hình hóa tiên tiến sẽ giúp nâng cao độ chính xác và khả năng dự đoán của các mô hình.

## TÀI LIỆU THAM KHẢO

[1]Mekanik, F., Imteaz, M. A., Gato-Trinidad, S., & Elmahdi, A. (2013). Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. Journal of Hydrology, 503, 11–21. https://doi.org/10.1016/j.jhydrol.2013.08.035

[2] Asaad Y. Shamseldin. Artificial neural network model for river flow forecasting in a developing country. Journal of Hydroinformatics, 12(1). 2010.

[3] Jeng-Fung Chen, Ho-Nien Hsieh and Quang Hung Do, Forecasting Hoabinh Reservoir’s Incoming Flow: An Application of Neural Networks with the Cuckoo Search Algorithm. Information, 5(4), p.570-586. 2014. <https://doi.org/10.3390/info5040570>

[4] Ji Youn Sung, Jeongwoo Lee, Il-Moon Chung and Jun-Haeng Heo. Hourly Water Level Forecasting at Tributary Affected by Main River Condition. Water, 9(9):644. 2017. <https://doi.org/10.3390/w9090644>

[5] Sella Nevo, Efrat Morin, Adi Gerzi Rosenthal, Asher Metzger, Chen Barshai, Dana Weitzner, Dafi Voloshin, Frederik Kratzert, Gal Elidan2, Gideon Dror, Gregory Begelman, Grey Nearing, Guy Shalev, Hila Noga, Ira Shavitt, Liora Yuklea, Moriah Royz, Niv Giladi, Nofar Peled Levi, Ofir Reich, Oren Gilon, Ronnie Maor, Shahar Timnat, Tal Shechter, Vladimir Anisimov, Yotam Gigi, Yuval Levin, Zach Moshe, Zvika Ben-Haim, Avinatan Hassidim, Yossi Matias. Flood forecasting with machine learning models in an operational framework. CoRR abs/2111.02780. 2021. Available: <https://doi.org/10.5194/hess-2021-554>.

[6] Guo W-D, Chen W-B, Yeh S-H, Chang C-H, Chen H. Prediction of River Stage Using Multistep-Ahead Machine Learning Techniques for a Tidal River of Taiwan. Water. 2021; 13(7):920. <https://doi.org/10.3390/w13070920>

[7] Ali Najah Ahmed, Ayman Yafouz, Ahmed H. Birima, Ozgur Kisi, Yuk Feng Huang, Mohsen Sherif, Ahmed Sefelnasr & Ahmed El-Shafie. Water level prediction using various machine learning algorithms: a case study of Durian Tunggal river. Malaysia. Engineering Applications of Computational Fluid Mechanics. 2022; 16:1, p.422-440.

[8] Kusudo T, Yamamoto A, Kimura M, Matsuno Y. Development and Assessment of Water-Level Prediction Models for Small Reservoirs Using a Deep Learning Algorithm. Water. 2022; 14(1):55. 54

[9] Hồ Việt Tuấn, Hồ Việt Hùng. Sử dụng mạng nơ ron nhân tạo dự báo mực nước sông chịu ảnh hưởng của thủy triều. Tạp chí Khoa học và Công nghệ thủy lợi. 2019; 52, p.1-9.

[10] Đỗ Văn Đỉnh, Nguyễn Trọng Quỳnh, Vũ Văn Cảnh và Phạm Văn Nam. Dự báo mực nước sông cao nhất, thấp nhất trong ngày sử dụng mô hình hỗn hợp. Tạp chí Nghiên cứu khoa học, Trường Đại học Sao Đỏ, 1(72), p.5-12.

[11] Đinh Nhật Quang, Tạ Quang Chiểu, Đào Thị Huệ, Nguyễn Thị Kim Ngân. Dự báo mực nước trên sông Kiên Giang sử dụng phương pháp hồi quy. Tạp chí Khoa học Kỹ thuật Thủy lợi và Môi Trường.2022; 80, p.71-80.