**Chương 12. Học sâu với mạng Neural**

**Outline**

[**I. CÁC CÂU HỎI CHƯƠNG 12 CÓ TRONG CÁC ĐỀ THI 2**](#_43e9l3bvebs3)

[1. Fin\_201.pdf\_HKI\_(2020-2021) 2](#_ghd6d3hz7wos)

[2. Fin\_202.pdf\_HKII\_(2020-2021) 2](#_3tl6br8wrpcf)

[3. Fin\_12\_2021.pdf\_HKI\_(2021-2022) 2](#_j9gq0wo6z7xt)

[4. Fin\_12\_2021.pdf\_HKII\_(2021-2022) 2](#_lxbo7ly6sudw)

[**II. CÂU HỎI CHI TIẾT TRONG ĐỀ 2**](#_redpzuq8njsd)

[1. Fin\_201.pdf\_HKI\_(2020-2021) 2](#_3xx3b76v0id3)

[2. Fin\_202.pdf\_HKII\_(2020-2021) 3](#_7ojfifuqfd31)

[3. Fin\_12\_2021.pdf\_HKI\_(2021-2022) 3](#_s1im3gd9yoi)

[4. Fin\_12\_2021.pdf\_HKII\_(2021-2022) 3](#_xr5cagvendmm)

[**II. GIẢI ĐỀ …To do … 4**](#_5yj3rsby3yc0)

[1. Fin\_201.pdf\_HKI\_(2020-2021) 4](#_kh31ffyhop8c)

[**Câu 1. Phần 1.2 & 1.4 4**](#_s99pm437lase)

[Đáp án Phần 1.2: 4](#_gppk1yo3rmgn)

[Đáp án Phần 1.4: 4](#_y8fxyf1v9p8z)

[**Câu 8. (2 điểm) → Chương 12 4**](#_a9kmkjjx60fc)

[Đáp án câu a: 4](#_ebcnsmbt1hot)

[Đáp án câu b: 4](#_kkr1bpfpn792)

[Đáp án câu c: 5](#_om9tbozgwnhy)

[Đáp án câu d: 5](#_zb9hap2ip2iz)

[Đáp án câu e: 5](#_f18qh3ysk7fg)

[**Câu 9 (0.75 điểm) 6**](#_lb1gxff52pbg)

[Đáp án: 6](#_v3iuncs1qggt)

[2. Fin\_202.pdf\_HKII\_(2020-2021) 6](#_u00vl3s0zkwo)

[**Câu 2 (0.75 điểm) 6**](#_4xyjtssco6ho)

[Đáp án: 6](#_u8hdcipdtx78)

[3. Fin\_12\_2021.pdf\_HKI\_(2021-2022) 6](#_rsqw0qbvm2l2)

[Câu 10 a.b.c.d (Trùng với câu 8 đề Fin\_201.pdf\_HKI\_(2020-2021)) 6](#_orw2lbgqgunu)

[Đáp án câu d 7](#_2gupovgqf1ux)

[4. Fin\_12\_2021.pdf\_HKII\_(2021-2022) 9](#_gx1cpljwy3d1)

[Câu 1 9](#_tnby6qvbmz18)

[Đáp án câu 1.2 e 9](#_9947p4k8xmhw)

## **I. CÁC CÂU HỎI CHƯƠNG 12 CÓ TRONG CÁC ĐỀ THI**

##### [Fin\_201.pdf\_HKI\_(2020-2021)](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md)

* + Câu 1. Phần 1.2, Phần 1.4 - **Đạt - Trâm**
  + Câu 8 - **Đạt**
  + Câu 9 - **Đạt**

##### [Fin\_202.pdf\_HKII\_(2020-2021)](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)

* + Câu 2 - **Đạt**
  + Câu 9  **(Trùng với câu 8** [**Fin\_201.pdf\_HKI\_(2020-2021)**](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md)

##### [Fin\_12\_2021.pdf\_HKI\_(2021-2022)](https://docs.google.com/document/d/1FBqWzRhwBudjJDBuK2_z72pM98EndxA2/edit#heading=h.t37itj2wog4b)

* + Câu 1  **(Trùng với câu 1 đề** [**Fin\_202.pdf\_HKII\_(2020-2021)**](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)**)**
  + Câu 10
    1. **a.b.c.d (Trùng với câu 8 đề** [**Fin\_201.pdf\_HKI\_(2020-2021)**](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md)**)**
    2. e **- Trâm**

##### [Fin\_12\_2021.pdf\_HKII\_(2021-2022)](https://docs.google.com/document/d/1BLSqsCee6LnqI5rh2J_ZeN3whI-TOJl-VinGDjto97A/edit#)

* + Câu 1. Phần 1.2 **- Trâm**
  + Câu 9  ***(*Trùng *với câu 9.*** [***Fin\_202.pdf\_HKII\_(2020-2021)***](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)***,***

## **II. CÂU HỎI CHI TIẾT TRONG ĐỀ**

#### **1.** [**Fin\_201.pdf\_HKI\_(2020-2021)**](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md)

**Câu 1.**

**1.2** Thư viện phần mềm học sâu nào sau đây có thể làm việc với môi trường mobile:

A. Keras B. TensorFlow → Tensorflow Lite

C. Torch → TorchVision D. Pytorch → Caffe2, PyTorch Mobile  
 **→ Chương 12**

**1.4 (Đúng/Sai)** Trong mạng nơ ron tích chập (CNN) các tầng gộp (pooling layer) nhằm thu giảm độ phân giải không gian của hình ảnh khi phân lớp hình ảnh.  
 **→ Chương 12**

**Câu 8. (1.5 điểm)(2 điểm) → Chương 12**

**a)** Nêu công dụng chính của kiến trúc nhiều tầng trong mạng nơ ron học sâu.

(*0.25 điểm*)

**b)** Trong lãnh vực mạng nơ ron học sâu, *mô hình sinh* (generative model) và *mô hình phân biệt* (discriminative model) khác biệt nhau như thế nào? (*0.5 điểm*)

**c)** Nêu tên hai giải thuật dùng trong hai giai đoạn của quá trình huấn luyện mạng Deep Belief Network; một giai đoạn là học không giám sát và một giai đoạn là học có giám sát. (*0.5 điểm*)

**d)** Giải thích công dụng của tầng tích chập (convolution) và tầng gộp (pooling) trong mạng nơ ron tích chập. (*0.5 điểm*)

**e)** Hãy nêu tên giải thuật dùng để huấn luyện mạng nơ ron tích chập (CNN) (*0.25 điểm*)

**Câu 9 (0.75 điểm)** Ứng với các phần mềm mã nguồn mở sau đây để hỗ trợ cho lãnh vực học sâu, hãy nêu tên ngôn ngữ lập trình được dùng để hiện thực mỗi phần mềm đó. TensorFlow

Keras

PyTorch  
**→ Chương 12**

#### **2.** [**Fin\_202.pdf\_HKII\_(2020-2021)**](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)

**Câu 2 (0.75 điểm)** Khi xác định cấu hình của mạng nơ ron cho một bài toán phân lớp, căn cứ vào đâu để xác định số nút của tầng nhập và số nút của tầng xuất? Và bằng cách nào xác định số nút thích hợp cho tầng ẩn ?  
 **→ Chương 12**

**Câu 9. (Trùng với câu 8** [**Fin\_201.pdf\_HKI\_(2020-2021)**](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md)

#### **3.** [**Fin\_12\_2021.pdf\_HKI\_(2021-2022)**](https://docs.google.com/document/d/1FBqWzRhwBudjJDBuK2_z72pM98EndxA2/edit#heading=h.t37itj2wog4b)

**Câu 1 (Trùng với câu 1 đề** [**Fin\_202.pdf\_HKII\_(2020-2021)**](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)**)**

**Câu 10 a.b.c.d (Trùng với câu 8 đề** [**Fin\_201.pdf\_HKI\_(2020-2021)**](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md)**)**

1. Nêu công dụng chính của kiến trúc nhiều tầng trong mạng nơ ron học sâu.

(*0.25 điểm*)

1. Trong lãnh vực mạng nơ ron học sâu, *mô hình sinh* (generative model) và *mô hình phân biệt* (discriminative model) khác biệt nhau như thế nào? (*0.5 điểm*)
2. Nêu tên hai giải thuật dùng trong hai giai đoạn của quá trình huấn luyện mạng Deep Belief Network; một giai đoạn là học không giám sát và một giai đoạn là học có giám sát. (*0.5 điểm*)
3. Giải thích công dụng của tầng tích chập (convolution) và tầng gộp (pooling) trong mạng nơ ron tích chập. (*0.5 điểm*)
4. **Mô tả cách huấn luyện mạng nơ ron LSTM để dự báo dữ liệu chuỗi thời gian. (*0.5 điểm*)**

#### **4.** [**Fin\_12\_2021.pdf\_HKII\_(2021-2022)**](https://docs.google.com/document/d/1BLSqsCee6LnqI5rh2J_ZeN3whI-TOJl-VinGDjto97A/edit#)

**Câu 1**

**1.2** Phát biểu nào sau đây là không đúng  
A. Bài toán *quy hoạch toàn phương* (Quadratic Programming) liên quan đến máy vector hỗ trợ thuộc loại tối ưu hóa có ràng buộc (constrained optimization). *→ Chương 8*  
B. Trong giải thuật ADABOOST, việc xây dựng các bộ phận lớp thành phần là độc lập với nhau.*→ Chương 9*  
C. Bagging và Rừng Ngẫu nhiên sử dụng cùng chiến lược phiếu bầu: đa số phiếu (majority voting).*→ Chương 9*  
D. Quá trình huấn luyện mạng nơ ron truyền thẳng thường nhanh hơn quá trình huấn luyện mạng nơ ron RBF.*→ Chương 7*  
E. Gibb Sampling là một kỹ thuật toán học được dùng để huấn luyện mạng Deep Belief network mà dựa vào máy Boltzmann giới hạn. *→ Chương 12*  
F. Học không có giám sát không cần đến thông tin nhân lớp của các mẫu trong tập dữ liệu huấn luyện. *→ Chương 6*  (0.5 điểm)

**→ Chương 6 + Chương 7 + Chương 8 + Chương 9 + Chương 12**

**Câu 9 *(*Trùng *với câu 9.*** [***Fin\_202.pdf\_HKII\_(2020-2021)***](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)**)**

## **II. GIẢI ĐỀ** …To do …

#### **1.** [**Fin\_201.pdf\_HKI\_(2020-2021)**](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md)

##### **Câu 1. *Phần 1.2 & 1.4***

**1.2** Thư viện phần mềm học sâu nào sau đây có thể làm việc với môi trường mobile:

A. Keras B. TensorFlow

C. Torch D. Pytorch

###### **Đáp án Phần 1.2:**

**[Đạt]** B và D

=> TensorFlow có Tensorflow lite, Pytorch có Pytorch mobile dành cho môi trường mobile (android, IOS, thiết bị nhúng linux)

**[Trâm]** A, B, D

A: Keras cũng có triển khai mô hình Keras trên môi trường di động. Keras có thể làm việc với môi trường mobile thông qua tích hợp với các công cụ phát triển ứng dụng di động như TensorFlow Lite hoặc Core ML <https://keras.io/api/applications/>

**[Lâm]** B và D

Theo em tìm hiểu thì chỉ có TensorFlow với Tensorflow Lite, và Pytorch với Pytorch mobile, là có hỗ trợ official cho mobile.

**1.4 (Đúng/Sai)** Trong mạng nơ ron tích chập (CNN) các tầng gộp (pooling layer) nhằm thu giảm độ phân giải không gian của hình ảnh khi phân lớp hình ảnh.

###### **Đ*áp án Phần 1.*4*:***

**[Đạt]** Đúng. Việc này giúp giảm số lượng tham số và tính toán trong mạng nơ ron, tăng tốc độ tính toán và giảm khả năng overfitting.

**[Trâm]** Đúng

**[Lâm]** Đúng

##### **Câu 8. (2 điểm) → Chương 12**

**a)** Nêu công dụng chính của kiến trúc nhiều tầng trong mạng nơ ron học sâu.

###### **Đáp án câu a:**

**[Đạt]** Công dụng chính của kiến trúc nhiều tầng là khả năng học và trích xuất các đặc trưng phức tạp của dữ liệu, thông qua việc kết hợp các tầng để tạo ra các đặc trưng phức tạp hơn. Ngoài ra, kiến trúc nhiều tầng còn có tính tự động hóa và khả diễn, giúp tăng tính khả chuyển đổi và linh hoạt của mô hình, từ đó giúp giảm thiểu thời gian và chi phí đào tạo mô hình cho các bài toán khác nhau.

**b)** Trong lãnh vực mạng nơ ron học sâu, *mô hình sinh* (generative model) và *mô hình phân biệt* (discriminative model) khác biệt nhau như thế nào?

###### **Đáp án câu b:**

**[Đạt]**

* ***Mô hình sinh (Generative model)*** là mô hình học sâu có khả năng tạo ra các mẫu dữ liệu mới từ phân phối xác suất của dữ liệu đầu vào. Mô hình sinh học được phân phối xác suất của dữ liệu đầu vào và sau đó sử dụng phân phối xác suất này để tạo ra các mẫu dữ liệu mới. Một số ví dụ Restricted Boltzmann Machines (RBM), Variational Autoencoder (VAE), và Generative Adversarial Networks (GAN).
* ***Mô hình phân biệt (Discriminative model):*** là mô hình học sâu có khả năng phân loại dữ liệu vào các lớp khác nhau, tập trung vào học cách xác định mối quan hệ giữa đầu vào và đầu ra, học cách phân loại dữ liệu vào các lớp khác nhau dựa trên các đặc trưng của dữ liệu đầu vào. Một số ví dụ Logistic Regression, Support Vector Machines (SVM), Convolutional Neural Networks (CNN) và Recurrent Neural Networks (RNN).

**c)** Nêu tên hai giải thuật dùng trong hai giai đoạn của quá trình huấn luyện mạng Deep Belief Network; một giai đoạn là học không giám sát và một giai đoạn là học có giám sát.

###### **Đáp án câu c**:

**[Đạt]**

* Giai đoạn học không giám sát: đầu tiên, DBN sẽ được huấn luyện bằng giải thuật Contrastive Divergence
* để tối ưu hóa hàm mục tiêu. Giải thuật sử dụng để huấn luyện từng tầng của DBN một cách độc lập.
* Giai đoạn học có giám sát: Sau khi huấn luyện xong tầng cuối cùng của mạng trong giai đoạn không giám sát, DBN được chuyển sang giai đoạn huấn luyện có giám sát bằng giải thuật lan truyền ngược (Backpropagation). Giải thuật này được sử dụng để tính toán đạo hàm của hàm mất mát (loss function) và cập nhật các trọng số (weights) của mạng. Trong giai đoạn này, DBN sẽ được huấn luyện để phân loại các dữ liệu đầu vào thành các nhãn (labels), tùy thuộc vào mục đích của bài toán cụ thể.

**d)** Giải thích công dụng của tầng tích chập (convolution) và tầng gộp (pooling) trong mạng nơ ron tích chập.

###### **Đáp án câu d:**

**[Đạt]**

* Tầng tích chập (convolution) giúp mạng học được các tính năng của dữ liệu đầu vào dựa trên việc áp dụng một bộ lọc (filter).
* Tầng gộp (pooling) lấy mẫu và giúp giảm kích thước của feature map, tăng tốc độ tính toán. Ngoài ra tầng gộp còn giảm thiểu overfitting bằng cách loại bỏ các thông tin không quan trọng trong feature map và ngăn chặn mạng học những tính năng quá cụ thể cho tập huấn luyện.

**e)** Hãy nêu tên giải thuật dùng để huấn luyện mạng nơ ron tích chập (CNN)

###### **Đáp án câu e**:

**[Đạt]** Giải thuật dùng để huấn luyện CNN: Stochastic Gradient Descent (SGD) và các biến thể của nó như Momentum, Nesterov Momentum và Adaptive learning rate methods như Adam và RMSProp.

##### **Câu 9 (0.75 điểm)**

Ứng với các phần mềm mã nguồn mở sau đây để hỗ trợ cho lãnh vực học sâu, hãy nêu tên ngôn ngữ lập trình được dùng để hiện thực mỗi phần mềm đó.

###### **Đáp án:**

**[Đạt]**

+ TensorFlow: Python, ngoài ra cũng có API cho C++, Java, và Go

+ Keras: Python

+ PyTorch : Python, ngoài ra có API cho C++ và Java.

#### **2.** [**Fin\_202.pdf\_HKII\_(2020-2021)**](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)

##### **Câu 2 (0.75 điểm)**

Khi xác định cấu hình của mạng nơ ron cho một bài toán phân lớp, căn cứ vào đâu để xác định số nút của tầng nhập và số nút của tầng xuất? Và bằng cách nào xác định số nút thích hợp cho tầng ẩn ?

###### **Đáp án**:

**[Đạt]** chúng ta có thể căn cứ vào các yếu tố sau để xác định số nút của tầng nhập và số nút của tầng xuất:

* Tầng nhập có số nút bằng với số lượng đặc trưng của dữ liệu đầu vào.
* Tầng xuất có số nút bằng với số lượng lớp đầu ra của bài toán phân lớp.

Để xác định số nút thích hợp cho các tầng ẩn, có thể sử dụng các phương pháp sau:

1. Tìm kiếm tốt nhất (Grid Search): huấn luyện nhiều mô hình với các giá trị khác nhau về số lượng nút trong tầng ẩn và chọn mô hình tốt nhất dựa trên kết quả đánh giá hiệu suất trên tập kiểm tra.
2. Các kỹ thuật tối ưu như các thuật toán tối ưu hóa ngẫu nhiên (random search) hoặc thuật toán tối ưu hóa gradient (gradient-based optimization)

=> Đây là một quá trình thử và sai, phụ thuộc vào nhiều yếu tố như độ phức tạp của bài toán, kích thước và đa dạng của dữ liệu đầu vào,... nên không có phương án tối ưu, phải đảm bảo cân bằng giữa độ phức tạp của mô hình và khả năng học tập của mô hình

#### **3.** [**Fin\_12\_2021.pdf\_HKI\_(2021-2022)**](https://docs.google.com/document/d/1FBqWzRhwBudjJDBuK2_z72pM98EndxA2/edit#heading=h.t37itj2wog4b)

**Câu 1 (Trùng với câu 1 đề** [**Fin\_202.pdf\_HKII\_(2020-2021)**](https://docs.google.com/document/d/1DkbvgkUJOOQcx9eNk-DsMT5MwyhNMvLE/edit#)**)**

##### **Câu 10** a.b.c.d (Trùng với câu 8 đề [Fin\_201.pdf\_HKI\_(2020-2021)](https://docs.google.com/document/d/1O30YVMcYqIh3FvXyK30TFXqQB4cgPsMO/edit#heading=h.k859l5ku3md))

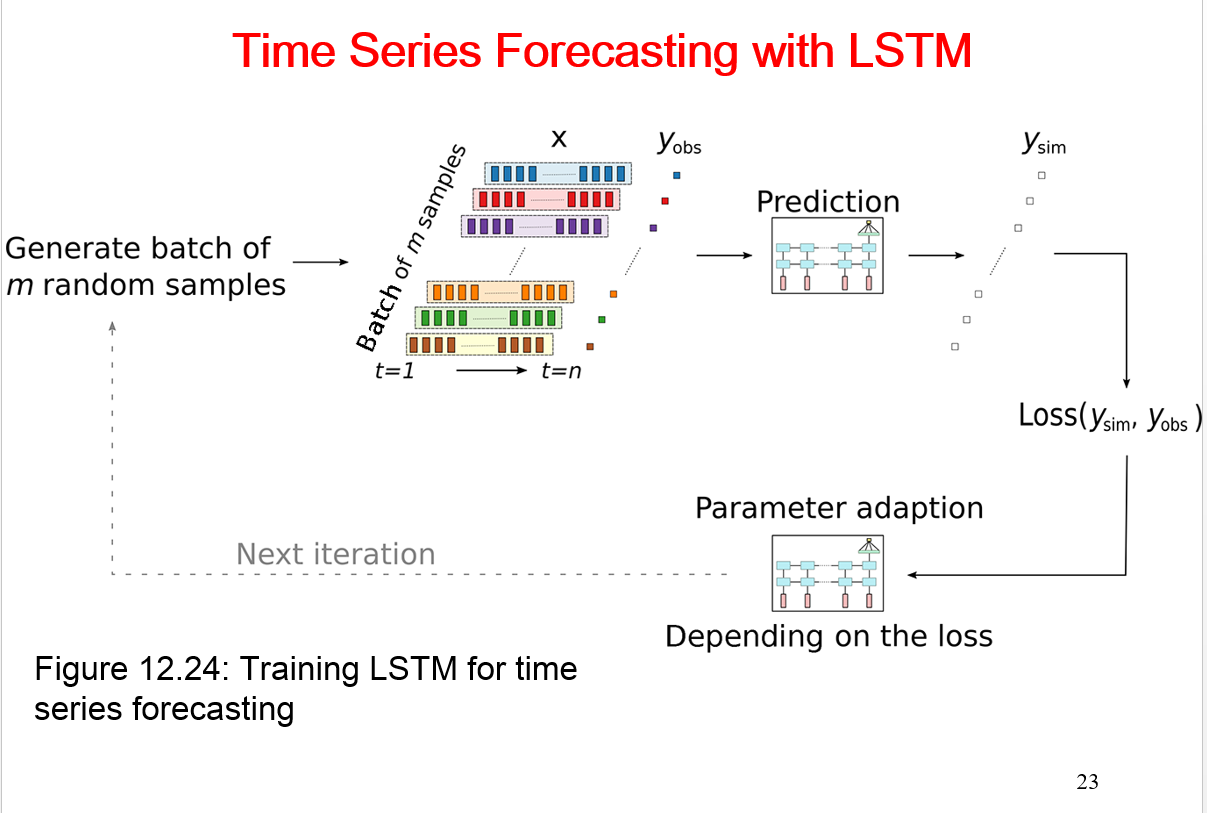
1. Nêu công dụng chính của kiến trúc nhiều tầng trong mạng nơ ron học sâu.

(*0.25 điểm*)

1. Trong lãnh vực mạng nơ ron học sâu, *mô hình sinh* (generative model) và *mô hình phân biệt* (discriminative model) khác biệt nhau như thế nào? (*0.5 điểm*)
2. Nêu tên hai giải thuật dùng trong hai giai đoạn của quá trình huấn luyện mạng Deep Belief Network; một giai đoạn là học không giám sát và một giai đoạn là học có giám sát. (*0.5 điểm*)
3. Giải thích công dụng của tầng tích chập (convolution) và tầng gộp (pooling) trong mạng nơ ron tích chập. (*0.5 điểm*)
4. **Mô tả cách huấn luyện mạng nơ ron LSTM để dự báo dữ liệu chuỗi thời gian. (*0.5 điểm*)**

###### **Đáp án câu d**

**[Trâm]**



Hình 12.24 minh họa một bước lặp trong quy trình huấn luyện của LSTM. Một lô dữ liệu đầu vào ngẫu nhiên x bao gồm m mẫu huấn luyện độc lập (được mô tả bằng màu sắc) được sử dụng trong mỗi bước. Mỗi mẫu đào tạo bao gồm n điểm dữ liệu và một giá trị mục tiêu (yobs) để dự đoán. Tổn thất được tính toán từ giá trị mục tiêu và ysim dự đoán của mạng và được sử dụng để cập nhật các tham số mạng.

Để huấn luyện mạng nơ ron LSTM (Long Short-Term Memory) để dự báo dữ liệu chuỗi thời gian bao gồm một số bước.

**1. Chuẩn bị dữ liệu:**

* Thu thập dữ liệu chuỗi thời gian mà bạn muốn dự đoán, bao gồm các giá trị được quan sát và dấu thời gian tương ứng của chúng.
* Chia dữ liệu thành các tập huấn luyện và kiểm tra. Thông thường, bạn sẽ giữ một phần dữ liệu ở phần cuối dưới dạng bộ kiểm tra để đánh giá hiệu suất của mô hình được đào tạo.

**2. Chuẩn hóa dữ liệu:**

* Dữ liệu chuỗi thời gian thường có quy mô và phân phối khác nhau. Điều quan trọng là chuẩn hóa dữ liệu để đưa chúng về cùng một phạm vi hoặc phân phối. Một phương pháp phổ biến là chuẩn hóa Min-Max, trong đó dữ liệu được chia tỷ lệ thành phạm vi [0, 1].

**3. Chuẩn bị dữ liệu cho LSTM:**

* Dữ liệu chuỗi thời gian cần được chuyển đổi thành đầu vào và đầu ra thích hợp cho mạng LSTM.
* Một cách tiếp cận phổ biến là sử dụng cửa sổ trượt để tạo các cặp đầu vào-đầu ra. Với độ dài cửa sổ nhất định, bạn trượt nó qua chuỗi thời gian và tạo các cặp. Đầu vào của mỗi cặp là một chuỗi các giá trị được quan sát gần đây và đầu ra là giá trị tiếp theo trong chuỗi.

**4. Xây dựng mạng LSTM:**

* Mạng LSTM bao gồm một hoặc nhiều lớp LSTM kết hợp với các lớp khác như lớp Dropout và lớp kết nối đầy đủ.
* Bạn có thể sử dụng các thư viện máy học như TensorFlow hoặc PyTorch để xây dựng mô hình LSTM. Định cấu hình các lớp LSTM với các tham số như số lượng đơn vị LSTM, tỷ lệ bỏ học và chức năng kích hoạt.

**5. Huấn luyện mô hình:**

* Sử dụng tập huấn luyện đã chuẩn bị để huấn luyện mạng LSTM.
* Định nghĩa hàm mất mát (loss function) phù hợp cho bài toán dự báo chuỗi thời gian, ví dụ như hồi quy (regression)
* Định nghĩa hàm mất mát (loss function) phù hợp cho bài toán dự báo chuỗi thời gian, ví dụ như hồi quy (regression) hoặc hồi quy dạng phân loại (classification).
* Chọn một thuật toán tối ưu hóa như stochastic gradient descent (SGD) hoặc Adam để điều chỉnh các tham số mạng nơ ron.
* Đặt các siêu tham số như tỷ lệ học (learning rate), số lần lặp (epochs), kích thước batch (batch size).
* Huấn luyện mạng LSTM bằng cách truyền các đầu vào qua mạng, tính toán hàm mất mát và điều chỉnh các trọng số dựa trên gradient. Lặp lại quá trình này cho tất cả các mẫu huấn luyện trong nhiều epochs.

**6. Đánh giá mô hình:**

* Sử dụng tập kiểm tra để đánh giá hiệu suất dự báo của mô hình đã được huấn luyện.
* Đo lường các độ đo đánh giá như sai số trung bình (mean squared error - MSE), hệ số tương quan (correlation coefficient), hay bất kỳ độ đo nào phù hợp với bài toán của bạn.

**7. Dự báo dữ liệu:**

* Sử dụng mô hình đã huấn luyện để dự báo giá trị chuỗi thời gian trong tương lai.
* Cung cấp đầu vào chuỗi quan sát gần đây, mô hình sẽ sinh ra dự báo cho các giá trị tiếp theo.

**8. Điều chỉnh và cải thiện mô hình:**

* Kiểm tra hiệu suất dự báo và cân nhắc điều chỉnh các siêu tham số, kiến trúc mạng hoặc phương pháp tiền xử lý dữ liệu để cải thiện kết quả.
* Lặp lại quá trình huấn luyện và đánh giá để tiếp tục cải thiện mô hình.

<https://books.google.com.vn/books?hl=vi&lr=&id=o5qnDwAAQBAJ&oi=fnd&pg=PP1&dq=%22Deep+Learning+for+Time+Series+Forecasting%22+(Jason+Brownlee)&ots=yH76sQpc0b&sig=6fZF5lOCMUcYFDXFbzxzUiUNYYg&redir_esc=y#v=onepage&q=%22Deep%20Learning%20for%20Time%20Series%20Forecasting%22%20(Jason%20Brownlee)&f=false>

#### **4.** [**Fin\_12\_2021.pdf\_HKII\_(2021-2022)**](https://docs.google.com/document/d/1BLSqsCee6LnqI5rh2J_ZeN3whI-TOJl-VinGDjto97A/edit#)

##### **Câu 1**

**1.2** Phát biểu nào sau đây là không đúng  
A. Bài toán *quy hoạch toàn phương* (Quadratic Programming) liên quan đến máy vector hỗ trợ thuộc loại tối ưu hóa có ràng buộc (constrained optimization). *→ Chương 8*  
B. Trong giải thuật ADABOOST, việc xây dựng các bộ phận lớp thành phần là độc lập với nhau.*→ Chương 9*  
C. Bagging và Rừng Ngẫu nhiên sử dụng cùng chiến lược phiếu bầu: đa số phiếu (majority voting).*→ Chương 9*  
D. Quá trình huấn luyện mạng nơ ron truyền thẳng thường nhanh hơn quá trình huấn luyện mạng nơ ron RBF.*→ Chương 7*  
E. Gibb Sampling là một kỹ thuật toán học được dùng để huấn luyện mạng Deep Belief network mà dựa vào máy Boltzmann giới hạn. *→ Chương 12*  
F. Học không có giám sát không cần đến thông tin nhân lớp của các mẫu trong tập dữ liệu huấn luyện. *→ Chương 6*  (0.5 điểm)

**→ Chương 6 + Chương 7 + Chương 8 + Chương 9 + Chương 12**

###### **Đáp án câu 1.2 e**

E. Gibb Sampling là một kỹ thuật toán học được dùng để huấn luyện mạng Deep Belief network mà dựa vào máy Boltzmann giới hạn. *→ Chương 12*

**[Trâm]**

Không đúng. Giải thuật Gibbs Sampling là một phương pháp mẫu ngẫu nhiên (Monte Carlo) dùng để xấp xỉ phân bố xác suất của biến ngẫu nhiên đa biến trong một mô hình thống kê. Phương pháp này không liên quan trực tiếp đến việc huấn luyện mạng Deep Belief Network (DBN) hoặc máy Boltzmann giới hạn (RBM).

Slide [Chapter\_12\_PartI.ppt] page 22, 23, 24, 25