

# BÁO CÁO MÔN Điện Toán Đám Mây

Đề tài: " Triển khai mô hình học sâu dự đoán giá sim điện trên nền tảng Amazon Web Services"

Nhóm thực hiện: Nhóm 5

Gồm : Nguyễn Thái Học, Nông Minh Đức, Lý Quốc Huy

# Nội dung

- 1 Giới thiệu về bài toán định giá sim
- 2 Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

# Giới thiệu về bài toán định giá

## 1. Giới thiệu chung

- Sim là từ viết tắt của "Subscriber Identity Module" (mô-đun nhận dạng thuê bao) và là một thiết bị nhỏ có chứa thông tin nhận dạng cá nhân của một thuê bao di động, cho phép người dùng thực hiện các cuộc gọi, gửi tin nhắn và truy cập các dịch vụ di động khác.
- Để định giá một sim điện thoại, các yếu tố sau thường được đưa vào để đánh giá:
  - Số lượng
  - Phân loại
  - Tính độc đáo
  - Thị trường

# Gới thiệu về bài toán định giá sim

## 2. Phương pháp triển khai bài toán

- Ưu điểm
  - Học máy có thể được áp dụng để dự đoán giá trị của một số điện thoại trong lý thuyết. Tuy nhiên, giá trị của sim điện thoại không phải là một giá trị liên tục, điều này có thể làm cho mô hình học máy bị sai lệch và không hoạt động hiệu quả, vì các thuật toán học máy thường hoạt động tốt trên các giá trị biến liên tục. Số sim điện thoại bao gồm các con số từ 0 đến 9 và giá trị của sim phụ thuộc vào sự kết hợp và vị trí của các con số với nhau. Học sâu, hay deep learning, được áp dụng để giải quyết các vấn đề này bởi vì nó có khả năng nhận biết sự đặc biệt và học được các quy luật giữa các con số.

# Giới thiệu về bài toán định giá sim

## 2. Phương pháp triển khai bài toán

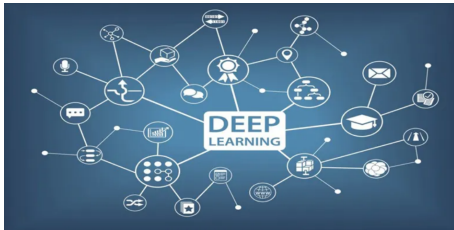
- Hạn chế

- Giá sim điện thoại có rất nhiều khoảng giá và luôn luôn thay đổi theo thời gian và thị trường. Để một mô hình có hiệu suất, hiệu quả cao thì cần có một tập dữ liệu đủ lớn và luôn luôn được cập nhật. Việc giá sim điện thoại có rất nhiều khoảng giá đồng nghĩa với việc mô hình sẽ có rất nhiều đầu ra do mô hình phải xác định xem số điện thoại này có giá trị là bao nhiêu. Đối với dữ liệu có quá nhiều đầu ra, mô hình sẽ trở nên phức tạp, học hỏi quá chi tiết từ dữ liệu cũng sẽ làm giảm hiệu suất, hiệu quả của mô hình.

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

## 1. Tổng quan về Deep Learning

- Học sâu (Deep Learning) là một kỹ thuật học máy tiên tiến dựa trên việc học đại diện. Cách tiếp cận mạnh mẽ này cho phép máy móc tự động học các biểu diễn tính năng cao cấp từ dữ liệu.



Hình: 1. Deep Learning

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

## 2. Cách hoạt động của học sâu

- Mạng nơ-ron học sâu cố gắng bắt chước bộ não con người thông qua sự kết hợp của dữ liệu đầu vào, trọng số và bias (tham số trong quá trình huấn luyện). Các yếu tố này hoạt động cùng nhau để nhận dạng, phân loại và mô tả chính xác các đối tượng trong dữ liệu.
- Mạng nơ-ron học sâu bao gồm nhiều lớp được kết nối với nhau, mỗi lớp được xây dựng dựa trên lớp trước đó để tinh chỉnh và tối ưu hoá dự đoán hoặc phân loại. Sự tiến triển của các tính toán thông qua mạng được gọi là lan truyền chuyển tiếp. Các lớp đầu vào và đầu ra là nơi mô hình Deep Learning thu nạp dữ liệu để xử lý và lớp đầu ra là nơi đưa ra dự đoán hoặc phân loại cuối cùng.

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

## 3. Lĩnh vực và ứng dụng trong học sâu

- Về lĩnh vực:

- NLP
- Thị giác máy tính
- Học tăng cường

- Về ứng dụng:

- Phân loại hình ảnh
- Nhận dạng giọng nói
- Phát hiện gian lận
- Trong y tế



# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

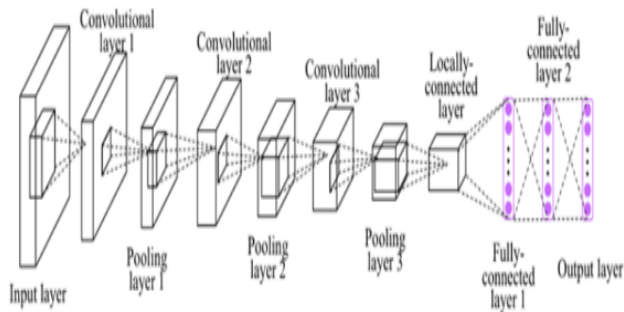
## 4. Một số mạng tiêu biểu trong học sâu

**4.1 CNN** - Mạng nơ-ron tích chập (Convolutional Neural Network - CNN) là một kiến trúc mạng nơ-ron được thiết kế đặc biệt cho việc xử lý và phân tích dữ liệu không gian như hình ảnh và video.

- Tổng quan về CNN:

- Convolutional layer ( lớp tích chập )
- Pooling layer(Lớp gộp)
- Activation function(hàm kích hoạt)
- Fully connected layer(lớp hoàn toàn kết nối)
- Huấn luyện

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

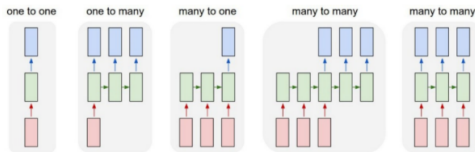


Hình: 2. Cấu trúc mạng mô hình nơ ron tích chập

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

## 4.2 RNN

### - Phân loại bài toán RNN

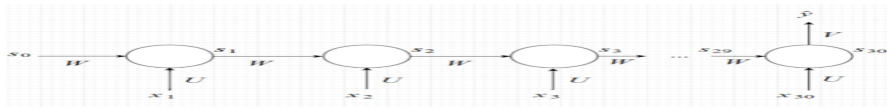


Hình: 3. Các dạng bài toán RNN

- one to one: mẫu bài 1 input và 1 output
- one to many : bài toán 1 input nhưng nhiều output
- many to one : bài toán nhiều input nhưng chỉ 1 output
- many to many : bài toán nhiều input và nhiều output

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

## - Mô hình RNN



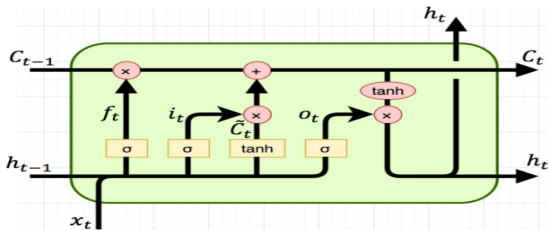
Hình: 4. Mô hình RNN

- Mỗi hình tròn được gọi là 1 state (trạng thái), state  $t$  có input là  $x_t$  và  $s_{t-1}$  (output của state trước), output là  $s_t = f(U x_t + W s_{t-1})$ .
- Dễ dàng nhận thấy  $S_t$  chính là bộ nhớ, nhớ các đặc điểm của các input từ  $x_1$  đến  $x_t$ .
- Đầu ra sẽ được đặt ở state cuối cùng, khi đó state cuối sẽ học được thông tin từ tất cả các input trước đó và đưa ra dự đoán.

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

## 4.3 Long Short Term Memory

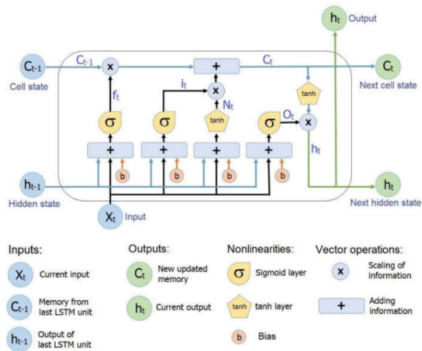
- Giới thiệu : Long short term memory hay còn gọi là mạng trí nhớ ngắn hạn định hướng dài hạn viết tắt là LSTM. LSTM chính là phiên bản mới của RNN nhằm khắc phục những vấn đề mà RNN đang gặp phải.



Hình: 5. LSTM

# Cơ sở lý thuyết về học sâu và mô hình mạng LSTM

- Cấu trúc mạng LSTM:



Hình: 6. Cấu trúc mạng LSTM