**BDRPC187TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HCM**

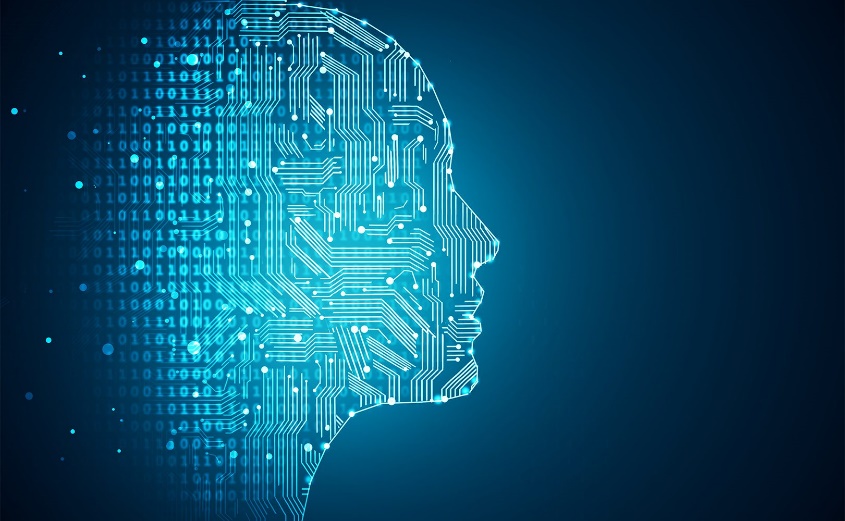
**KHOA ĐÀO TẠO CHẤT LƯỢNG CAO**

**BÁO CÁO**

**MACHINE LEARNING**

*ĐỀ TÀI:*

**ỨNG DỤNG THUẬT TOÁN MÁY HỌC TRONG VIỆC XỬ LÝ ẢNH ĐỂ CHẨN ĐOÁN BỆNH PHỔI TRONG Y KHOA**

****

**Giảng viên: TS. Nguyễn Quang Huy**

* **Phạm Hồng Cang 15110014**
* **Lê Minh Chương 15110020**
* **Nguyễn Thành Luân 15110075**

**Nhận xét của giáo viên hướng dẫn**

1. **Giới thiệu:**
2. Machine Learning là gì ?

Machine Learning là một lĩnh vực nhỏ của Khoa Học Máy Tính, nó có khả năng tự học hỏi dựa trên dữ liệu đưa vào mà không cần phải được lập trình cụ thể. Machine Learning gây nên cơn sốt công nghệ trên toàn thế giới trong vài năm nay. Trong giới học thuật, mỗi năm có hàng ngàn bài báo khoa học về đề tài này. Trong giới công nghiệp, từ các công ty lớn như Google, Facebook, Microsoft đến các công ty khởi nghiệp đều đầu tư vào Machine Learning.

Có khá nhiều ứng dụng của công nghệ học máy đã được áp dụng trong cuộc sống hàng ngày như: Google dịch, xe ô tô tự lái, hệ thống gợi ý mua hàng, hệ thống nhận diện khuôn mặt... Đặc biệt, Machine Learning là 1 xu hướng phát triển nhanh chóng trong ngành chăm sóc sức khỏe, nhờ vào sự ra đời của các thiết bị và máy cảm ứng đeo được sử dụng dữ liệu để đánh giá tình hình sức khỏe của bệnh nhân trong thời gian thực (real-time). Công nghệ Machine Learning còn giúp các chuyên gia y tế xác định những xu hướng hoặc tín hiệu để cải thiện khả năng điều trị, chẩn đoán bệnh.

1. Lí do chọn đề tài:

a, Hiện tại:

* + Ngành y khoa:
    - Nhu cầu hỗ trợ bác sĩ trong y khoa càng lúc càng tăng theo thời gian. Việc đọc, phân tích hình ảnh và liên hệ bệnh lý cần đòi hỏi bác sĩ có kiến thức rộng và kinh nghiệm nhiều. Đọc thủ công không những mất rất nhiều thời gian mà lại có độ tin cậy không cao.
    - Số lượng bệnh nhân càng lúc càng tăng và các loại bệnh cơ bản lặp đi lặp lại nhiều nên cần có công cụ hỗ trợ các bác sĩ chuẩn đoán nhanh các bệnh cơ bản đó.
    - Y tế được biết đến là dịch vụ về chăm sóc sức khỏe con người cũng như là một trong những dịch vụ yêu cầu bắt buộc phải phát triển trên mức kinh tế để có thể đảm bảo cuộc sống cho người dân cũng như là sự tồn vong của nhân loại. Những ứng dụng của khoa học công nghệ luôn chú trọng vào việc có thể ứng dụng vào y tế rất nhiều và được ưu tiên hàng đầu.
    - Thế kỷ XXI là một thế kỷ của sự hội nhập kinh tế. Vì thế mà Việt Nam cần đẩy mạnh về việc có thể giao lưu kinh tế với các quốc gia trong và ngoài khu vực. Để có thể đảm bảo về cơ sở vật chất, dịch vụ thì việc có được những chiến lược, những áp dụng công nghệ, khoa học vào y tế là điều cần thiết. Đây được xem là một trong những lĩnh vực có thể thu hút được đầu tư từ các đơn vị nước ngoài.
    - Trong kỷ nguyên mới ngành y tế càng cần thiết để con người có được sức khỏe tốt nhất cống hiến cho xã hội. Ngành y tế ứng dụng các thiết bị khoa học để phát triển hơn, khi sức khỏe con người trở nên ổn định thì đương nhiên mọi hoạt động kinh tế, xã hội đều được thúc đẩy mạnh mẽ.
  + Bệnh phổi:
    - Vì quá trình công nghiệp hóa hiện đại hóa đang diễn ra với tốc độ nhanh chóng, từ đó kéo theo rất vấn đề như: ô nhiễm môi trường đặc biệt là không khí ,… Dẫn đến các bệnh liên quan về đường hô hấp gia tăng và ngày càng phổ biến. Nên có 1 lượng dữ liệu lớn về bệnh phổi - sau khi chuẩn đoán (bằng phương pháp chụp X - ray).
    - Theo số liệu thống kê, năm 2015 có 1,8 triệu người bị chết do lao phổi, trong số 10,4 triệu người mắc bệnh.
    - Theo dự đoán của WHO, đến năm 2020, bệnh phổi sẽ là nguyên nhân tử vong xếp hàng thứ 3 trên toàn thế giới và là nguyên nhân thứ 5 trong các bệnh gây ra tàn phế.

b, Tương lai:

* + Machine Learning nói riêng và AI nói chung trong tương lai sẽ phát triển rất nhanh chóng và hỗ trợ, phục vụ con người trong nhiều lĩnh vực như: kinh tế, công nghiệp, đời sống, giáo dục,… và đặc biệt là y khoa.
  + Machine Learning giúp phổ cập rộng rãi các hệ thống hồ sơ y tế điện tử, cải thiện tốc độ xử lý chẩn đoán.
  + Khi triệu chứng bệnh xuất hiện, có thể có hàng triệu nghiên cứu và kết quả liên quan đến chúng và có thể tốn rất nhiều thời gian để phân tích một lượng dữ liệu lớn như vậy. Machine Learning sẽ vượt trội hơn so với cách truy vấn dữ liệu truyền thống, để quét, phân tích và tóm tắt các văn bản y tế giúp các bác sĩ tập trung vào chẩn đoán bệnh và có thể đưa ra kết luận trong thời gian sớm nhất.
  + Mô hình sử dụng công nghệ sẽ nhanh hơn và chính xác hơn so với mô hình được hỗ trợ bởi máy tính truyền thống và có thể tự cải thiện khả năng phát hiện và phân loại bệnh của chúng thông qua khả năng tự học hỏi và nhận thức của máy móc. Qua đó, giúp các bác sĩ có thể phản ứng nhanh chóng và điều trị kịp thời cho bệnh nhân.
  + Bằng việc kết hợp khoa học dữ liệu và phân tích dữ liệu, Machine Learning và AI trong tương lai có thể tạo ra bản sao kỹ thuật số của con người dựa trên những dữ liệu chi tiết của họ, qua đó cung cấp hình ảnh toàn diện về bệnh nhân giúp hỗ trợ và giảm rủi ro trong điều trị bệnh.
  + Nghiên cứu cho thấy GDP toàn cầu vào năm 2030 sẽ tăng đáng kể nhờ vào lĩnh vực trí tuệ nhân tạo, biến nó trở thành cơ hội thương mại lớn nhất trong nền kinh tế thay đổi nhanh chóng ngày nay.

1. Ý nghĩa đề tài:

* Áp dụng Machine Learning & AI vào trong y khoa sẽ giúp các bác sĩ dễ dàng và nhanh chóng hơn trong việc chuẩn đoán bệnh để từ đó kê đơn thuốc đúng và cho từng loại bệnh của các bệnh nhân.
* Giảm thiểu sự sai sót trong chuẩn đoán bệnh.
* Hỗ trợ các bác sĩ từ các quốc gia có nên y khoa chưa phát triển từ đó giúp giảm thiểu rủi ro tử vong của bệnh nhân vì chẩn đoán không chính xác .

VD: Mỗi năm, riêng ở Mỹ, ước tính có khoảng 40.000 đến 80.000 người chết do chẩn đoán sai. (theo BCC News: Lỗi chẩn đoán gây tử vong cho bệnh nhân).

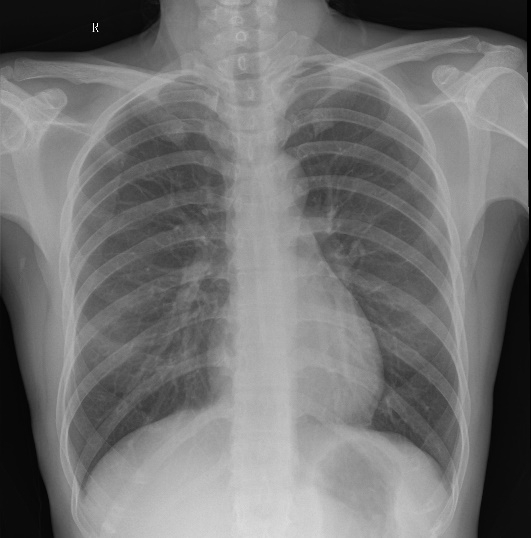
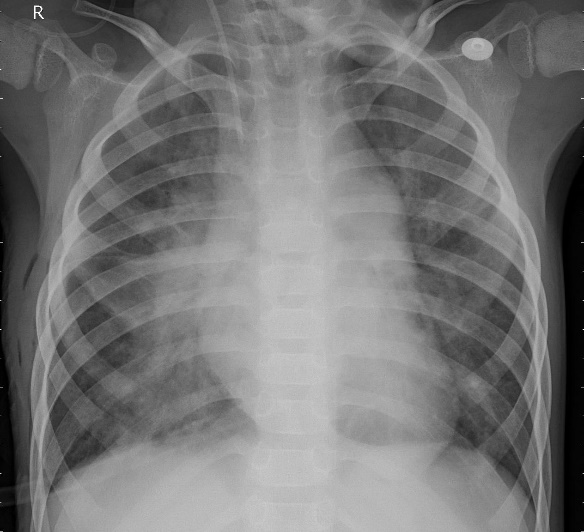
1. **Mục tiêu:**

* Nhóm hoàn thành dự án đúng mong đợi:
* Sau tuần học thứ 9 các thành viên trong nhóm đã có thể nắm rõ các thuật toán cơ bản của Machine Learning và chọn được thuật toán phù hợp cho đề tài.
* Sau tuần học thứ 14, nhóm hoàn thành việc thực hiện training model và test, chọn ra được hướng đi có tỉ lệ chính xác cao nhất.
* Các thành viên trong nhóm nắm bắt tiến trình làm việc, dễ dàng hỗ trợ lẫn nhau.
* Từ dữ liệu có sẵn, training ra được bộ model cho chẩn đoán sau này.
* Với hình ảnh test sẽ dự đoán ra được có bệnh hay không có bệnh.
* Kết quả dự đoán đạt độ chính xác cao (trên 80%).
* Có thể áp dụng dự án vào trong thực tế.
* Tương lai sẽ phát triển dự án theo mô hình có khả năng phát hiện và phân loại giai đoạn phát triển của bệnh phổi.

1. **Quá trình thực hiện:**
2. Đặt vấn đề:

* Các thành viên trong nhóm tìm hiểu các khái niệm cơ bản, tìm ra thực trạng và nhu cầu sử dụng cũng như ý nghĩa của đề tài.
  + Sau khi hoàn thành giai đoạn tìm hiểu, các thành viên trong nhóm đã phân tích các vấn đề thực tế có thể áp dụng cho đề tài. Từ đó, hội ý và đưa ra giải pháp cuối cùng để thực hiện đề tài có kết quả tốt nhất có thể.
  + Với tập dữ liệu mới được phân loại thành 2 nhóm: Pneumonia và Normal, nhóm sẽ thực hiện thuật toán Supervised Learning cho đề tài và đáp ứng các mục tiêu, ý nghĩa đã đề ra.

1. Chuẩn bị dữ liệu:
   * Thông qua Internet, nhóm đã tìm được nguồn dữ liệu hình ảnh X-quang ngực được chọn từ một nhóm nghiên cứu nước ngoài.
   * Tập dữ liệu hình ảnh bao gồm 3 thư mục chính (train, test, val) và chứa 2 thư mục con đã gắn nhãn (Pneumonia, Normal). Tổng cộng có 5863 hình ảnh X-quang (JPEG).
   * Dưới đây là một số hình ảnh trong tập dữ liệu:

Bình thường Bị bệnh

1. Chọn thuật toán:

* Nhóm thực hiện nghiên cứu, tìm hiểu các thuật toán cơ bản và phổ biến trong chương trình học thông qua các trang web, sự hướng dẫn của giảng viên để phân tích ưu nhược điểm của từng thuật toán và đánh giá mức độ phù hợp cho đề tài.
  + Sau khi tìm hiểu nhiều thuật toán thông qua Internet và trong chương trình học ví dụ như: Support Vector Machine, Convolutional Neural Network, K-means, … và sự hướng dẫn của giảng viên, nhóm đã xác định được thuật toán Convolutional Neural Network phù hợp với đề tài.
  + Lý do chọn thuật toán CNN:

1. Training model:
   * Nhóm thực hiện tìm hiểu và sử dụng các thư viện hỗ trợ cho Python như matplotlib, keras, sklearn,…
   * Áp dụng các Layer dùng cho thuật toán CNN:
   * Conv2D: (Convolutional Layers: chứa các layer trong mạng nơ ron tích chập) là convolution layer dùng để lấy feature từ image
   * MaxPool2D: (Pooling Layers : Chứa các layer dùng trong mạng CNN.) dùng để lấy feature nổi bật dùng max) và giúp giảm parameter khi training.
   * Dense: layer này sử dụng như một layer neural network bình thường.
   * Dropout: layer này dùng như regularization cho các layer hạn chế overfiting
   * Softmax: (chọn activation function) dùng trong multi classifier
   * Input: layer này sử dụng input như 1 layer
   * Flatten: dùng để lát phẳng layer để fully connection
   * Activation: dùng để chọn activation.

* Chạy code với các thông số khác nhau về learning rate, tỷ lệ hình trong 2 thư mục (training, test) và cho ra model.

1. Kiểm thử:
   * Với mỗi model đã được training trong phần trên, nhóm thực hiện import model và kiểm tra tỷ lệ chính xác (%), lưu kết quả kiểm tra với thông số tương ứng.
   * Với kết quả kiểm tra, đưa ra được model có độ chính xác cao nhất và chọn đó làm kết quả cuối cùng.
2. **Cơ sở lý thuyết:**
3. Xử lý ảnh:
   1. Đặc trưng của ảnh (Feature):

Đặc trưng ảnh là những chi tiết xuất hiện trong ảnh, từ đơn giản như cạnh, hình khối, chữ viết tới phức tạp như mắt, mặt, chó, mèo, bàn, ghế, xe, đèn giao thông, v.v..

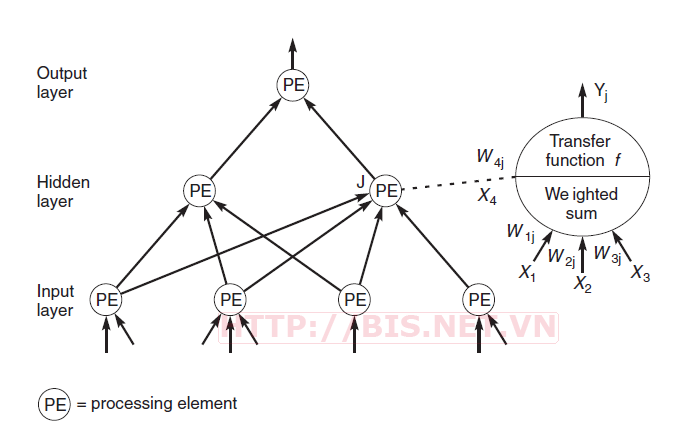
* 1. Bộ lọc ảnh:

Là bộ lọc giúp phát hiện và trích xuất các đặc trừng của ảnh, có thể là bộ lọc góc, cạnh, đường chéo, hình tròn, hình vuông, v.v.

* 1. Trong bài toán của chúng ta thì các bức ảnh là các tấm phim X-ray được chụp của nhiều bệnh nhân và người không bị bệnh. Mỗi hình ảnh có kích thước khác nhau nên muốn xử lý những hình ảnh đó cần trải qua các bước tách các đặt trưng của ảnh để tiện cho việc xử lý. Bước này được gọi là bước tiền xử lý.

1. Neural Networks:
   1. Tổng quan:

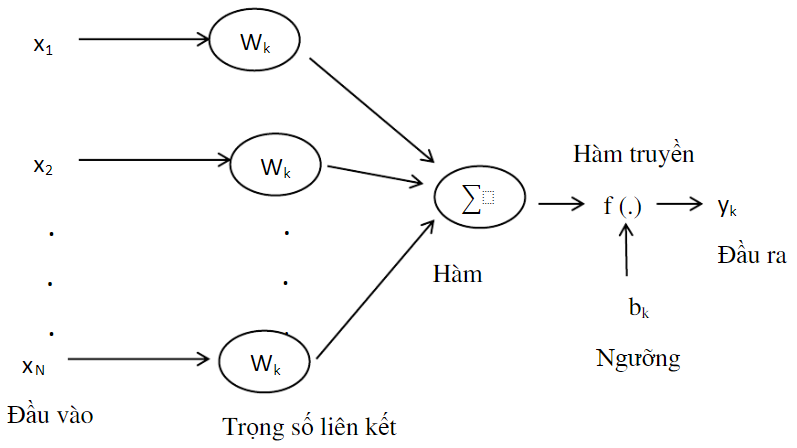
* Mạng nơron nhân tạo, Artificial Neural Network (ANN) là một mô hình xử lý thông tin phỏng theo cách thức xử lý thông tin của các hệ nơron sinh học. Nó được tạo nên từ một số lượng lớn các phần tử (nơron) kết nối với nhau thông qua các liên kết (trọng số liên kết) làm việc như một thể thống nhất để giải quyết một vấn đề cụ thể nào đó. Một mạng nơron nhân tạo được cấu hình cho một ứng dụng cụ thể (nhận dạng mẫu, phân loại dữ liệu,... ) thông qua một quá trình học từ tập các mẫu huấn luyện. Về bản chất học chính là quá trình hiệu chỉnh trọng số liên kết giữa các nơron.
  1. Kiến trúc:



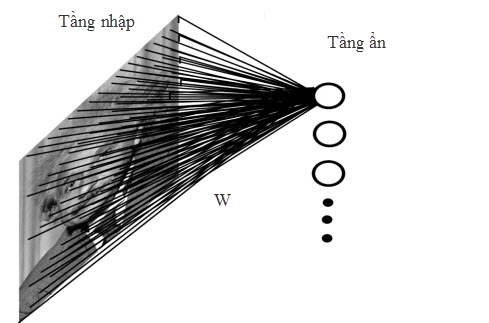
Hình 1: Kiến trúc mạng Neural Networks

* Kiến trúc chung của một NN gồm 3 thành phần đó là Input Layer, Hidden Layer và Output Layer.
  1. Quá trình xử lý thông tin:





* Inputs: Mỗi Input tương ứng với 1 thuộc tính (attribute) của dữ liệu (patterns). Ví dụ như trong ứng dụng của ngân hàng xem xét có chấp nhận cho khách hàng vay tiền hay không thì mỗi Input là một thuộc tính của khách hàng như thu nhập, nghề nghiệp, tuổi, số con,…
* Output: Kết quả của một ANN là một giải pháp cho một vấn đề, ví dụ như với bài toán xem xét chấp nhận cho khách hàng vay tiền hay không thì output là yes (cho vay) hoặc no (không cho vay).
* Connection Weights (Trọng số liên kết) : Đây là thành phần rất quan trọng của một ANN, nó thể hiện mức độ quan trọng (độ mạnh) của dữ liệu đầu vào đối với quá trình xử lý thông tin (quá trình chuyển đổi dữ liệu từ Layer này sang layer khác). Quá trình học (Learning Processing) của ANN thực ra là quá trình điều chỉnh các trọng số (Weight) của các input data để có được kết quả mong muốn.
* Bộ tổng (Summing function): Thường dùng để tính tổng của tích các đầu vào với trọng số liên kết của nó
* Ngưỡng (còn gọi là một độ lệch - bias): Ngưỡng này thường được đưa vào như một thành phần của hàm truyền.
* Hàm truyền (Transfer function): Hàm này được dùng để giới hạn phạm vi đầu ra của mỗi nơron. Nó nhận đầu vào là kết quả của hàm tổng và ngưỡng.
* Tuy nhiên, mạng Nơ-ron truyền thống (Neural Network) hoạt động không thực sự hiệu quả với dữ liệu đầu vào là hình ảnh. Nếu coi mỗi điểm ảnh là một thuộc tính (feature), một ảnh RBG kích thước (64×64) có 12288 (=64×64×3) thuộc tính. Nếu kích thước ảnh tăng lên 1000×10000, chúng ta có 3 triệu (3M) thuộc tính cho mỗi ảnh đầu vào. Nếu sử dụng mạng liên kết đầy đủ (fully connected NN) và giả sử lớp thứ 2 có 1000 thành phần (units/ neurons), ma trận trọng số sẽ có kích thước 1000×3M tương đương với 3 tỷ trọng số cần huấn luyện (learning).



Hình: ANN (Artificial Neural Network) với cấu trúc Full - connected

Điều này yêu cầu khối lượng tính toán cực lớn (expensive computational cost) và thường dẫn đến overfitting do không đủ dữ liệu huấn luyện. Nên dẫn đến Convolutional Neural Networks ra đời để khắc phục các vấn đề của mạng Nơ-ron truyền thống (Neural Network).

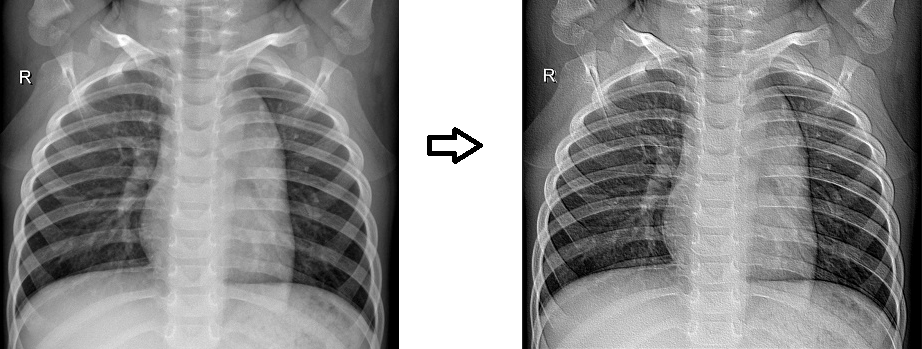
1. Convolutional Neural Networks – CNN: (Mạng nơ-ron tích chập)
   1. Tổng quan:
   * CNN là mạng nơ-ron phổ biến nhất được dùng cho dữ liệu ảnh. Bên cạnh các lớp liên kết đầy đủ (FC layers), CNN còn đi cùng với các lớp ẩn đặc biệc giúp phát hiện và trích xuất những đặc trưng - chi tiết (patterns) xuất hiện trong ảnh gọi là Lớp Tích chập (Convolutional Layers). Chính những lớp tích chập này làm CNN trở nên khác biệt so với mạng nơ-ron truyền thống và hoạt động cực kỳ hiệu quả trong bài toán phân tích ảnh.
   1. Các thành phần cơ bản của mạng CNN:
2. Convolution Layer:
   * + Convolution Layer là lớp quan trọng nhất trong cấu trúc mạng CNN.
     + Conv dựa trên lý thuyết xử lý tín hiệu số, việc lấy tích chập sẽ giúp trích xuất được những thông tin quan trọng từ dữ liệu.



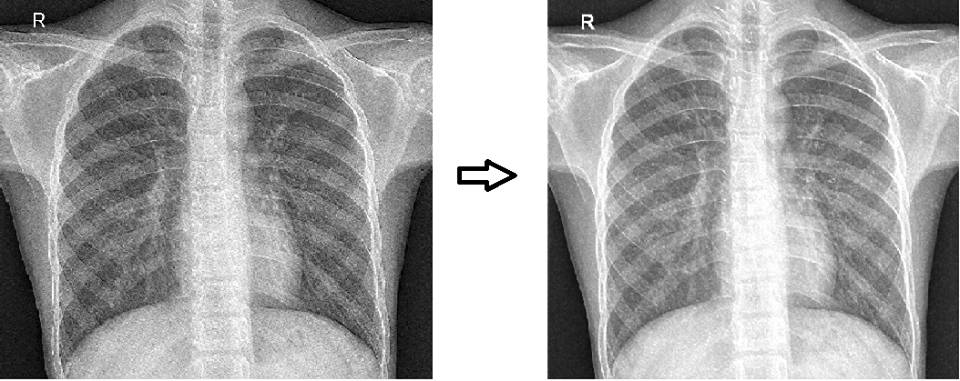
Hình: Mô hình phép toán tích chập

Ví dụ: Một ảnh phổi sau khi được Convolution với

kernel = [-2,-1,0,-1,1,1,0,1,2]



Hình: So sánh trước và sau khi Convolution



Hình: So sánh sau khi Max Pooling từ ảnh đã được Convolution

1. Activation function:

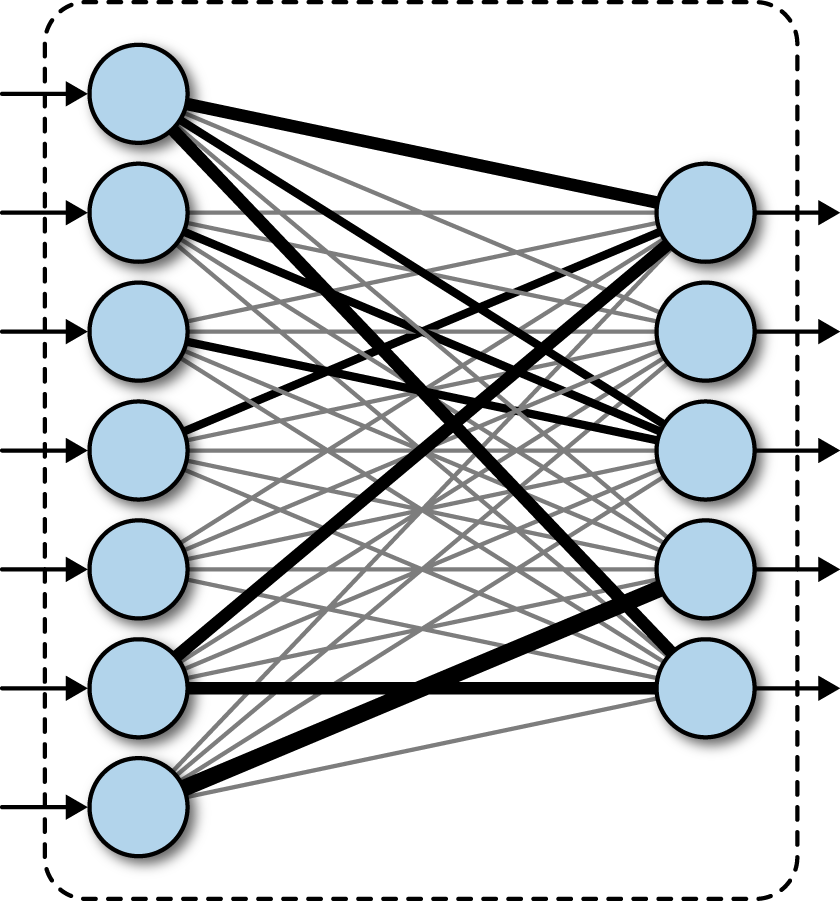
Về cơ bản, Covolution là một phép biển đổi tuyến tính. Nếu tất cả các neural được tổng hợp bởi các phép biến đổi tuyến tính thì một mạng neural đều có thể đưa về dưới dạng một hàm tuyến tính. Khi đó sẽ đưa các bài toán về logistic regression. Do đó tại mỗi neural cần có một hàm truyền dưới dạng phi tuyến. Có nhiều dạng hàm phi tuyến được sử dụng trong quá trình này như tanh, sigmoid function, hay ReLUs.

Tuy nhiên, các nghiên cứu gần đây chứng minh được việc sử dụng hàm ReLu (Rectified Linear Unit) cho kết quả tốt hơn ở các khía cạnh:

* Tính toán đơn giản.
* Tạo ra tính thưa (sparsity) ở các neural ẩn. Ví dụ như sau bước khởi tạo ngẫu nhiên các trọng số, khoảng 50\% các neural ẩn được kích hoạt (có giá trị lớn hơn 0).
* Quá trình huấn luyện nhanh hơn ngay cả khi không phải trải qua bước tiền huấn luyện.
* Trong bài toán này chúng ta sử dụng hàm ReLU để làm hàm Activation.

1. Fully – Connected:

Là cách kết nối các neural ở hai tầng với nhau trong đó tầng sau kết nối đẩy đủ với các neural ở tầng trước nó. Đây cũng là dạng kết nối thường thấy ở ANN (Artificial Neural Network ), trong CNN tầng này thường được sử dụng ở các tầng phí cuối của kiến trúc mạng.



Hình: Cách hoạt động của Fully-Connected

* 1. Cấu trúc mạng CNN và cách hoạt động:

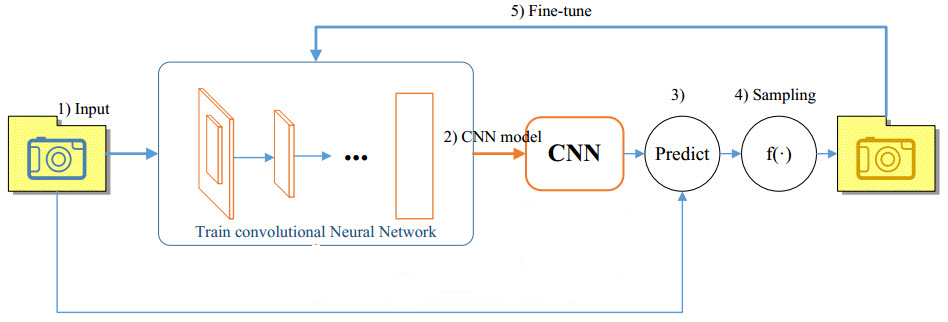
CNN có kiến trúc được hình thành từ các thành phần cơ bản bao gồm Convolution (CONV), Pooling (POOL), ReLU, Fully-connected (FC) về mặt xây dựng kiến trúc tổng quát CNN được mô tả như sau (dấu mũi tên thể hiện thứ tự sắp xếp các tầng từ trước đến sau):



Quá trình hoạt động của mạng CNN:

* + Convolution + ReLU tức là trong kiến trúc này sau khi Convolution thì sẽ kết hợp ReLU, trong CNN kiến trúc lớp này có thể lặp N lần.
  + Pooling là tầng Pooling cho người thiết kế quyết định có thể có hoặc không.
  + Tập hợp các lớp Convolution + ReLU và Pooling được gọi là giai đoạn Feature Learning.
  + Sau đó tập hợp các Convolution Layer chồng lên nhau và sử dụng các nonlinear activation function như ReLU để kích hoạt các trọng số trong các node.
  + Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo.
  + Các layer liên kết được với nhau thông qua cơ chế convolution. Layer tiếp theo là kết quả convolution từ layer trước đó, nhờ vậy mà ta có được các kết nối cục bộ. Nghĩa là mỗi nơ-ron ở layer tiếp theo sinh ra từ filter áp đặt lên một vùng ảnh cục bộ của nơ-ron layer trước đó.
  + Sau khi kết thúc quá trình Feature Learning thì có 2 lớp Fatten và Fully Connected sẽ tạo ra các liên kết đầy đủ như một mạng Neurual Network bình thường. Cuối cùng sẽ cho ra kết quá mỗi hình được xử lí thuộc nhóm nào.
* Trong mô hình CNN có 2 khía cạnh cần quan tâm là tính bất biến (Location Invariance) và tính kết hợp (Compositionality):
* Tính bất biến: Với cùng một đối tượng, nếu đối tượng này được chiếu theo các gốc độ khác nhau (translation, rotation, scaling) thì độ chính xác của thuật toán sẽ bị ảnh hưởng đáng kể. Pooling layer sẽ cho bạn tính bất biến đối với phép dịch chuyển (translation), phép quay (rotation) và phép co giãn (scaling).
* Tính kết hợp: cho ta các cấp độ biểu diễn thông tin từ mức độ thấp đến mức độ cao và trừu tượng hơn thông qua convolution từ các filter.
  + Trong suốt quá trình huấn luyện, CNNs sẽ tự động học được các thông số.

1. **Model:**



Hình: Qúa trình triển khai xử lý bài toán

Quá trình:

Bước 1: Input bộ data gồm bộ data training và bộ data testing

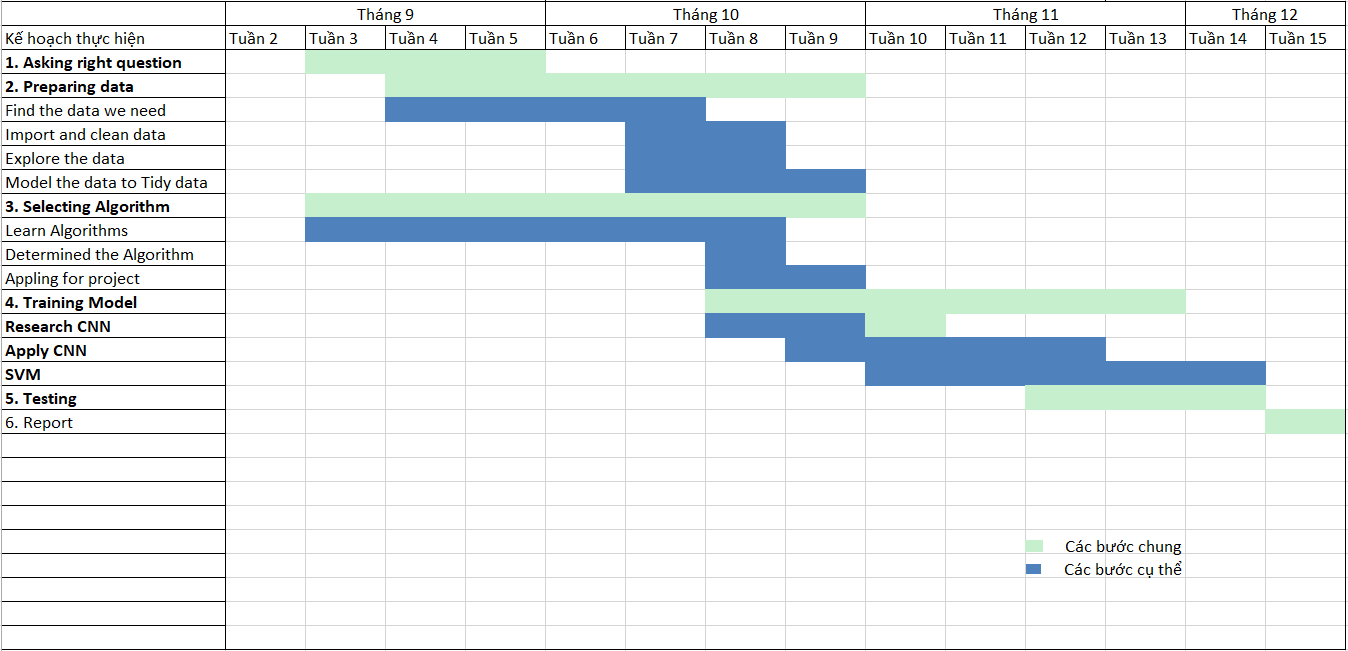
Bước 2: Từ bộ training trải qua quá trình xử lý CNN tạo ra được bộ Model

Bước 3: Sử dụng bộ Model tạo được so sánh với bộ testing

Bước 4: Đánh giá mô hình

Bước 5: Nếu kết quả đánh giá đáp ứng nhu cầu thì dừng lại. Ngược lại thì quay lại bước 2.

1. **Thống kê, kế hoạch, phân công công việc:**



Hình: Bảng phân công việc và kế hoạch.

1. **Kết quả thực tế:**
2. **Phụ lục:**

Tài liệu tham khảo:

1. [https://machinelearningcoban.com](https://machinelearningcoban.com/)
2. <https://www.coursera.org/learn/machine-learning> (Andrew Ng)
3. Machine learning in Python Essential Techniques For Predictive Analysis (Book)
4. <https://www.tensorflow.org/>
5. <https://viblo.asia/p/ung-dung-convolutional-neural-network-trong-bai-toan-phan-loai-anh-4dbZNg8ylYM>
6. <https://app.pluralsight.com/library/courses/python-understanding-machine-learning>