khung1khung3khung4

**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TPHCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

 **KỸ THUẬT DỮ LIỆU**

**BÁO CÁO TIỂU LUẬN CHUYÊN NGÀNH**

**Tìm hiểu DEEP LEARNING**

**VÀ BÀI TOÁN pHÂN LOẠI VỚI TENSORFLOW**

**Giảng viên hướng dẫn:** **ThS. Lê Thị Minh Châu**

**Nhóm thực hiện: Tên thành viên MSSV**

**Nguyễn Tiến Duy 17133008**

**Nguyễn Hiền Nhung 17133044**

**Phan Tấn Thịnh 17133060**

**KHÓA 2017-2021**

khung2

**ĐH SƯ PHẠM KỸ THUẬT TP.HCM**

**KHOA CNTT**

**\*\*\*\*\*\*\***

**CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**Độc Độc lập – Tự do – Hạnh Phúc**

**\*\*\*\*\*\*\***

**PHIẾU NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN**

Họ và tên Sinh viên 1: Nguyễn Tiến Duy MSSV: 17133008

Họ và tên Sinh viên 2: Nguyễn Hiền Nhung MSSV: 17133044

Họ và tên Sinh viên 3: Phan Tấn Thịnh MSSV: 17133060

Tên đề tài: TÌM HIỂU DEEP LEARNING VA BÀI TOÁN PHÂN LOẠI VỚI TENSORFLOW

Họ và tên Giáo viên hướng dẫn: ThS. Nguyễn Thị Minh Châu

**NHẬN XÉT**

1. Về nội dung đề tài khối lượng thực hiện: .................................................................................................................................................................. .................................................................................................................................................................. .................................................................................................................................................................. ..................................................................................................................................................................

2. Ưu điểm: .................................................................................................................................................................. .................................................................................................................................................................. .................................................................................................................................................................. ..................................................................................................................................................................

3. Khuyết điểm .................................................................................................................................................................. .................................................................................................................................................................. .................................................................................................................................................................. ..................................................................................................................................................................

4. Đề nghị cho bảo vệ hay không?

5. Đánh giá loại:

6. Điểm:

Tp. Hồ Chí Minh, ngày tháng năm 2020

Giáo viên hướng dẫn

(Ký & ghi rõ họ tên)

**ĐH SƯ PHẠM KỸ THUẬT TP.HCM**

**KHOA CNTT**

**\*\*\*\*\*\*\***

**CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**Độc Độc lập – Tự do – Hạnh Phúc**

**\*\*\*\*\*\*\***

**PHIẾU NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN**

Họ và tên Sinh viên 1: Nguyễn Tiến Duy MSSV: 17133008

Họ và tên Sinh viên 2: Nguyễn Hiền Nhung MSSV: 17133044

Họ và tên Sinh viên 3: Phan Tấn Thịnh MSSV: 17133060

Tên đề tài: TÌM HIỂU DEEP LEARNING VA BÀI TOÁN PHÂN LOẠI VỚI TENSORFLOW

Họ và tên Giáo viên phản biện: TS. Trần Nhật Quang

**NHẬN XÉT**

1. Về nội dung đề tài khối lượng thực hiện: .................................................................................................................................................................. .................................................................................................................................................................. .................................................................................................................................................................. ..................................................................................................................................................................

2. Ưu điểm: .................................................................................................................................................................. .................................................................................................................................................................. .................................................................................................................................................................. ..................................................................................................................................................................

3. Khuyết điểm .................................................................................................................................................................. .................................................................................................................................................................. .................................................................................................................................................................. ..................................................................................................................................................................

4. Đề nghị cho bảo vệ hay không?

5. Đánh giá loại:

6. Điểm:

Tp. Hồ Chí Minh, ngày tháng năm 2020

Giáo viên phản biện

(Ký & ghi rõ họ tên)

**LỜI CẢM ƠN**

Chúng tôi xin được gửi lời cảm ơn chân thành đến Cô Lê Thị Minh Châu. Cô đã cung cấp cho chúng tôi tài liệu và hướng dẫn tận tình cho chúng tôi trong suốt quá trình thực hiện tiểu luận chuyên ngành. Cô luôn theo dõi tiến độ và giải đáp, giúp chúng tôi vượt qua những khó khăn. Chúng tôi rất trân quý sự tâm huyết và trách nhiệm của Cô trong công việc giảng dạy và truyền thụ kiến thức.

Chúng tôi cũng xin gửi lời cảm ơn sâu sắc đến Thầy Cô khoa Công nghệ thông tin- Đại học Sư Phạm Kỹ Thuật thành phố Hồ Chí Minh đã truyền dạy kiến thức và hỗ trợ chúng tôi trong suốt quá trình học tập và thực hiện tiểu luận chuyên ngành. Chúng tôi xin cảm ơn mái trường Sư phạm kỹ thuật đã tạo nhiều điều kiện thuận lợi cho hoạt động học tập của sinh viên chúng tôi, đặc biệt là thư viện với nguồn tri thức vô tận.

Tuy nhiên, do chưa có nhiều kinh nghiệm làm để tài cũng như những hạn chế về kiến thức, nên bài tiểu luận chắc chắn sẽ không tránh khỏi những thiếu sót. Rất mong nhận được sự nhận xét, ý kiến đóng góp, phê bình từ phía thầy/cô để bài tiểu luận được hoàn thiện hơn.

Lời cuối cùng, chúng tôi xin kính chúc thầy/cô nhiều sức khỏe, thành công và hạnh phúc.

Nguyễn TIến Duy

Nguyễn Hiền Nhung

Phan Tấn Thịnh

**MỤC LỤC**

[**LÍ DO CHỌN ĐỀ TÀI** 7](#_Toc61794783)

[**CHƯƠNG 1: TỔNG QUAN NGHIÊN CỨU** 9](#_Toc61794784)

[**1.** **Deep learning** 9](#_Toc61794785)

[**.1.1.** **Tổng quan Deep learning** 9](#_Toc61794786)

[**.1.2.** **Hoạt động của Deep Learning** 10](#_Toc61794787)

[**.1.3.** **thuật toán Deep Learning** 11](#_Toc61794788)

[**.1.4.** **Khác nhau giữa Machine Learning và Deep Learning** 11](#_Toc61794789)

[**.1.5.** **Một số ứng dụng của Deep Learning** 13](#_Toc61794790)

[**2.** **Thuật toán Convolutional Neural Network (CNN)** 18](#_Toc61794800)

[**.2.1.** **Neural network (NN-Mạng nơ-ron)** 18](#_Toc61794801)

[**.2.2.** **Convolutional Neural Network là gì?** 19](#_Toc61794802)

[**.2.3.** **Cấu trúc CNN** 20](#_Toc61794803)

[**.2.4.** **Hoạt động CNN** 21](#_Toc61794804)

[**3.** **Thư viện Tensorflow** 14](#_Toc61794791)

[**.3.1.** **Tổng quan về Tensorflow** 14](#_Toc61794792)

[**.3.2.** **Một số khái niệm quan trọng** 15](#_Toc61794793)

[**.3.2.1.** **Tensor** 15](#_Toc61794794)

[**.3.2.2.** **Graph** 15](#_Toc61794795)

[**.3.3.** **Cách Tensorflow hoạt động** 17](#_Toc61794796)

[**.3.4.** **Ưu điểm của Tensorflow** 17](#_Toc61794797)

[**.3.5.** **Hạn chế của Tensorflow** 18](#_Toc61794798)

[**.3.6.** **Các thuật toán nổi bật được hỗ trợ bởi Tensorflow** 18](#_Toc61794799)

[**4.** **Django** 22](#_Toc61794805)

[**.4.1.** **Tổng quan về Django** 22](#_Toc61794806)

[**.4.2.** **Kiến trúc Django** 22](#_Toc61794807)

[**.4.3.** **Ưu điểm của Django** 23](#_Toc61794808)

[**.4.4.** **Hạn chế của Django** 23](#_Toc61794809)

[**CHƯƠNG 2: MÔ TẢ BÀI TOÁN** 24](#_Toc61794810)

[**1.** **Mô tả tập dữ liệu** 25](#_Toc61794811)

[**2.** **Sơ lược về những thuật toán có sử dụng trong bài toán** 26](#_Toc61794812)

[**.2.1.** **Thuật toán KNN (K-nearest neighbors)** 26](#_Toc61794813)

[**.2.2.** **Pre-trained CNN Model** 27](#_Toc61794814)

[**3.** **Các bước xử lí** 28](#_Toc61794815)

[**4.** **So sánh độ đo giữa các thuật toán** 28](#_Toc61794816)

[**CHƯƠNG 3: KẾT LUẬN** 30](#_Toc61794817)

[**1.** **KẾT QUẢ ĐẠT ĐƯỢC** 30](#_Toc61794818)

[**.1.1.** **Ý nghĩa khoa học** 30](#_Toc61794819)

[**.1.2.** **Ý nghĩa thực tiễn** 30](#_Toc61794820)

[**2.** **HẠN CHẾ** 31](#_Toc61794821)

[**3.** **HƯỚNG PHÁT TRIỂN** 31](#_Toc61794822)

[**TÀI LIỆU THAM KHẢO** 32](#_Toc61794823)

# **LÍ DO CHỌN ĐỀ TÀI**

**Trong những năm qua, đối với thế giới nói chung và Việt Nam nói riêng đã và đang được chứng kiến những bước tiến khổng lồ về chất lượng cùng “độ thần thánh” của các sản phẩm công nghệ chúng ta vẫn sử dụng hàng ngày. Đầu tiên phải nói đến những công nghệ hàng đầu: AI, Machine Learning, Deep Learning, Cognitive Computing.**

Những công nghệ nhận diện giọng nói được nâng cấp rất nhiều so với trước đây. Nhờ có nó mà người dùng hiện nay có thể sử dụng khẩu lệnh để tương tác nhiều hơn với các thiết bị thông minh.Những trợ lý ảo giọng nói như Alexa của Amazon, Siri của Apple, Cortana của Microsoft cùng các hệ thống nhận diện giọng nói có mặt trên hầu như mọi sản phẩm của Google đang đồng loạt nở rộ trên nhiều nền tảng, giúp người dùng thực hiện nhiều tác vụ khác nhau. Ở Trung Quốc, Baidu cũng tỏ ra không hề kém cạnh với thống kê cho thấy người dùng các sản phẩm của công ty đã sử dụng nhận diện giọng nói nhiều gấp ba lần chỉ trong vòng 18 tháng qua.

Và rồi chúng ta lại có công nghệ nhận diện hình ảnh – tính năng đã có mặt rộng rãi trên các sản phẩm của 4 gã khổng lồ nêu trên. Bạn có thể tìm kiếm và sắp xếp những bức ảnh của mình mà không cần phải gắn thẻ chúng, mà chỉ dựa cần trên những gì hiện diện trong hình, từ cụ thể như một chú chó, cảnh tuyết rơi cho đến trừu tượng như những cái ôm. Nhiều sản phẩm trong số này thậm chí còn có thể đọc mô tả lại các yếu tố trên bức ảnh cho người dùng khiếm thị

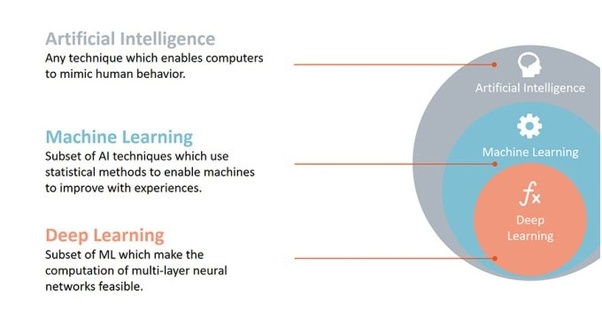
Những công nghệ này đã phát triển rất sôi động trong rất nhiều lĩnh vực trên toàn thế giới, trong đó có Việt Nam được xem là nước đang phát triển, hầu hết các hoạt động hàng ngày của chúng ta đều liên quan đến công nghệ thông tin. Trong thực tế ngày nay, ứng dụng công nghệ thông tin đã được đưa vào các lĩnh vực kinh tế, chính trị, xã hội, quân sự…, đây là cơ hội phát triển và đồng thời cũng là thách thức to lớn đối với việc phát triển xã hội nói chung. Cho đến nay, vấn đề này đã được nhiều nghiên cứu trên thế giới quan tâm. Nhưng những nghiên cứu này chưa thực sự phổ biến mạnh tại Việt. Trong khi đó, có rất nhiều vấn đề thực tiễn cần đến sự có mặt của công nghệ thông tin nói chung cũng như là Deep Learning nói riêng để nhằm đẩy mạnh phát triển đất nước.

Vì vậy, Trong tiểu luận chuyên ngành, nhóm tôi chọn đề tài *“Tìm hiểu Deep Learning và bài toán phân loại với Tensorflow”* để nghiên cứu với mục đích tìm hiểu được nền tảng công nghệ, các lý thuyết cơ bản trong bài toán xử lý ảnh, xử lí giọng nói, học máy và áp dụng vào các lĩnh vực trong cuộc sống.

**CHƯƠNG 1: TỔNG QUAN NGHIÊN CỨU**

* 1. **Deep learning**
     1. **Tổng quan Deep learning**

Deep Learning là một chủ đề Trí truệ nhân tạo (AI) và là một nhánh nhỏ của Machine Learning với đặc thù mang độ phức tạp cao hơn. Deep Learning có khả năng khác biệt ở một số khía cạnh quan trọng so với Machine Learning nông truyền thống, cho phép máy tính giải quyết một loạt các vấn đề phức tạp không thể giải quyết được.



*Hình 1: Quan hệ giữ AI, Machine Learning và Deep Learning.**Nguồn ảnh:KDnuggetsTM.vn*

Trong thực tế, Deep Learning bao gồm nhiều lớp ẩn trong một mạng lưới thần kinh và thuộc lớp sau cùng. Việc đi qua nhiều số lượng lớp và mạng phức tạp được cho là độ sâu. Ngày nay, sự thay đổi lớn nhất trong Deep Learning là độ sâu của mạng lưới thần kinh đã phát triển từ một vài lớp đến hàng trăm trong số chúng. Độ sâu hơn có nghĩa là khả năng nhận dạng các mẫu lớn hơn, với nguồn thông tin lớn hơn giúp tăng khả năng tiếp nhận các đối tượng trở nên rộng hơn, chi tiết hơn.

Deep Learning đang trở thành một trong những lĩnh vực quan trọng trong khoa học máy tính. Chỉ trong vài năm, Deep Learning đã thúc đẩy tiến bộ trong đa dạng các lĩnh vực như nhận thức sự vật (object perception), dịch tự động (machine translation), nhận diện giọng nói - đó là những vấn đề từng rất khó khăn với các nhà nghiên cứu trí tuệ nhân tạo.

* + 1. **Hoạt động của Deep Learning**

Deep learning cho phép chúng ta huấn luyện một AI có thể dự đoán các đầu ra đựa vào một tập các đầu vào. Cả hai phương pháp có giám sát và không có giám sát đều có thể sử dụng để huấn luyện.

Cách thức hoạt động của thuật toán Deep Learning diễn ra như sau : các dòng thông tin sẽ được truyền vào mạng thần kinh nhân tạo bao gồm nhiều lớp và được sắp xếp phân cấp . Mạng sẽ học một cái gì đó đơn giản ở cấp ban đầu trong hệ thống phân cấp và sau đó gửi thông tin này đến cấp tiếp theo. Cấp tiếp theo lấy thông tin đơn giản này, kết hợp với cái nó học được thành một thứ gì đó phức tạp hơn một chút và chuyển nó tới cấp thứ ba. Quá trình này tiếp tục khi mỗi cấp trong hệ thống phân cấp xây dựng một thứ gì đó phức tạp hơn từ đầu vào mà nó nhận được từ cấp trước.

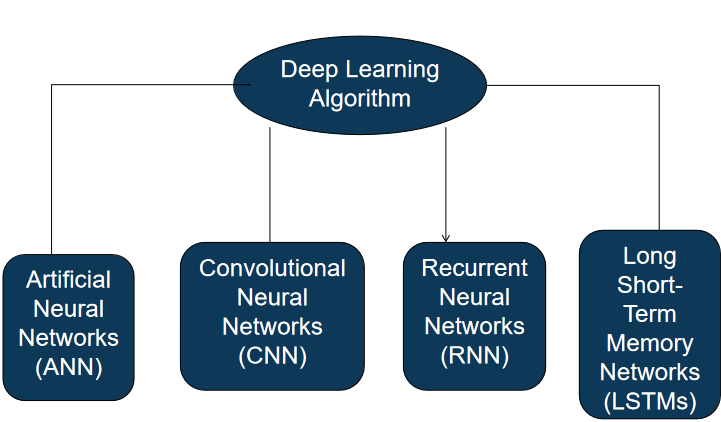
Ví dụ:

- Nhận diện con chó: cấp độ ban đầu của mạng có thể sử dụng sự khác biệt về vùng sáng và vùng tối của hình ảnh để tìm hiểu vị trí của các cạnh hoặc đường trong hình ảnh một con chó. Cấp ban đầu chuyển thông tin này đến cấp thứ hai, kết hợp các cạnh thành các hình dạng đơn giản như một đường chéo hoặc một góc vuông. Cấp độ thứ ba kết hợp các hình dạng đơn giản thành các đối tượng phức tạp hơn như hình bầu dục hoặc hình chữ nhật. Cấp độ tiếp theo có thể kết hợp hình bầu dục và hình chữ nhật thành râu, bàn chân và đuôi thô sơ. Quá trình tiếp tục cho đến khi nó đạt đến cấp cao nhất trong hệ thống phân cấp .

* Ứng dụng trợ lý giọng nói như Alexa hoặc Siri sử dụng Deep learning cho các trải nghiệm trò chuyện tự nhiên. Trong các cấp độ ban đầu của mạng thần kinh (neural network), khi trợ lý giọng nói được cung cấp dữ liệu, nó sẽ cố gắng xác định tình trạng giọng nói, ngữ điệu và hơn thế nữa. Đối với các cấp độ cao hơn, nó sẽ thu thập thông tin về từ vựng và thêm các phát hiện của các cấp độ trước đó. Trong các cấp độ sau, nó sẽ phân tích các gợi ý và kết hợp tất cả các kết luận của nó. Đối với cấp cao nhất của cấu trúc phân cấp, trợ lý giọng nói sẽ cố gắng học đủ để có thể phân tích một đoạn hội thoại và dựa vào các đầu vào đó, nó sẽ đưa ra một hành động tương ứng.
  + 1. **thuật toán Deep Learning**

Các thuật toán Deep Learning phổ biến:

* Convolutional Neural Network (CNN)
* Recurrent Neural Networks (RNNs)
* Long Short-Term Memory Networks (LSTMs)
* Stacked Auto-Encoders.
* Deep Boltzmann Machine (DBM)
* Deep Belief Networks (DBN)



*Hình 2. Thuật toán phổ biến của Deep Learning*

* + 1. **Khác nhau giữa Machine Learning và Deep Learning**

Machine Learning trích xuất những thuộc tính có sẵn trong khi Deep Learning trích xuất thuộc tính từ dữ liệu thô

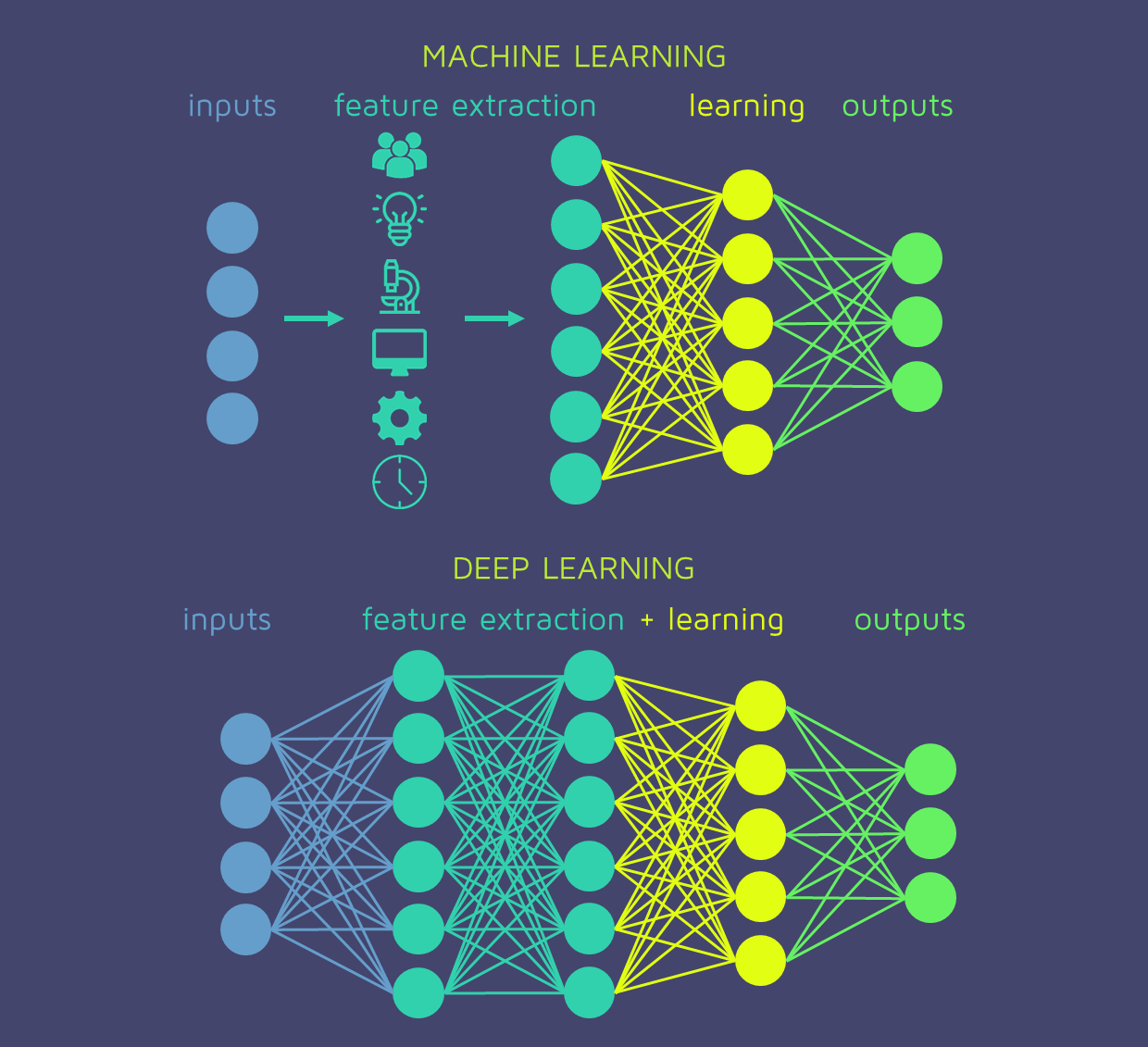
Với Machine Learning, cần ít dữ liệu hơn để đào tạo thuật toán so với Deep Learning. Deep Learning đòi hỏi một bộ dữ liệu rộng lớn và đa dạng để xác định cấu trúc cơ bản. Vì nếu dữ liệu quá nhỏ thì sẽ dễ gây ra hiện tương overfitting.

Bên cạnh đó, Machine Learning cung cấp một mô hình được đào tạo nhanh hơn. Hầu hết các kiến trúc Deep Learning tiên tiến có thể mất vài ngày đến một tuần để đào tạo. Tuy nhiên quy mô mạng lưới của Deep learning càng lớn và nạp càng nhiều dữ liệu thì hệ thống vận hành càng nhanh.

Việc điều chỉnh siêu tham số sẽ dễ dàng và linh hoạt theo nhiều cách khác nhau khi sử dụng Deep Learning.Còn đối với Machine Learning việc điều chỉnh có phần hạn chế hơn.

Ưu điểm của Deep Learning so với Machine Learning là nó có độ chính xác cao. Ta không cần phải hiểu những tính năng nào là đại diện tốt nhất cho dữ liệu; mạng nơ-ron đã học cách chọn các tính năng quan trọng. Trong Machine Learning, ta cần chọn cho mình những tính năng cần có trong mô hình.

Ngoài ra Deep Learning còn là một thuật toán có khả năng theo kịp sự phát triển của công nghệ và dữ liệu lớn. Và nếu bạn cung cấp cho hệ thống hàng tấn thông tin , nó sẽ bắt đầu tìm hiểu và phản hồi theo nhiều cách hữu ích.



*Hình 3. Sơ đồ khác biệt giữa Machine Learning và Deep Learning.* *Nguồn: Codecamp.vn*

* + 1. **Một số ứng dụng của Deep Learning**

Cá nhân hóa: Các doanh nghiệp của họ phát triển mạnh trong việc đưa ra các tùy chọn trong tiềm thức của người sử dụng dựa trên sở thích đó, các mục được truy cập gần đây, mối quan hệ với các thông tin và lịch sử duyệt web tổng thể trên nền tảng của họ.

Nhận diện hình ảnh, phân tích video, ảnh MRI: Mạng thần kinh chuyển đổi cho phép xử lý hình ảnh kỹ thuật số có thể được phân tách thành nhận dạng khuôn mặt, nhận dạng đối tượng, phân tích chữ viết tay,…

Xử lý ngôn ngữ tự nhiên: tóm tắt tài liệu, trả lời câu hỏi, mô hình hóa ngôn ngữ, phân loại văn bản, phân tích tình cảm là một số ứng dụng phổ biến đã bắt kịp đà phát triển.

Trợ lý ảo: Amazon Echo, Google Assistant, Alexa và Siri đều đang khai thác các khả năng học sâu để xây dựng trải nghiệm người dùng tùy chỉnh cho ta.



*Hình 4. Ứng dụng của Deep Learning. Nguồn: Tepco.com*

* 1. **Thuật toán Convolutional Neural Network (CNN)**
     1. **Neural network (NN-Mạng nơ-ron)**

Neural là tính từ của neuron (nơ-ron), network chỉ cấu trúc đồ thị nên neural network (NN) là một hệ thống tính toán lấy cảm hứng từ sự hoạt động của các

nơ-ron trong hệ thần kinh.

Tuy nhiên NN chỉ là lấy cảm hứng từ não bộ và cách nó hoạt động, chứ không phải bắt chước toàn bộ các chức năng của nó. Việc chính của chúng ta là dùng mô hình đấy đi giải quyết các bài toán chúng ta cần.

Các nơ-ron được nhóm và 3 loại layer khác nhau:

* Input layer
* Các hidden layer
* Output layer

Input layer: nhận các dữ liệu đầu vào(các label). Input layer sẽ đưa các đầu vào này vào hidden layer thứ nhất.

Các hidden layer thực hiện các phép tính toán cho các đầu vào. Thử thách lớn nhất trong việc tạo mạng nơ ron là quyết định số lượng các hidden layer này, cũng như số các nơ ron cho mỗi layer.

Từ “Deep” trong Deep Learning chỉ đến việc có nhiều hơn một hidden layer.

Output layer trả về dữ liệu đầu ra, trường hợp của ta sẽ là đưa ra dự đoán về giá vé.

Mối liên kết giữa nơ ron được kết hợp với một trọng số, nó chỉ ra được tầm quan trọng của giá trị đầu vào.

Việc lặp lại tập dữ liệu và so sánh các đầu ra sẽ sinh ra cost function giúp chỉ ra sai sót của AI so với đầu ra thực tế. Và việc của chúng ta là tính toán để làm giảm giá trị này. Khi Cost Function của ta sẽ trả về 0, khi đó đầu ra của AI cũng sẽ giống với đầu ra thực tế từ tập dữ liệu.

* + 1. **Convolutional Neural Network là gì?**

Convolutional Neural Network (CNN – Mạng nơ-ron tích chập) là một trong những mô hình Deep Learning tiên tiến. Nó giúp cho chúng ta xây dựng được những hệ thống thông minh với độ chính xác cao như hiện nay.

CNN thu nhận các thuộc tính đã học với dữ liệu đầu vào, và sử dụng các lớp convolution 2 chiều, làm cho cấu trúc này phù hợp với dữ liệu 2 chiều, chẳng hạn như hình ảnh.

CNN trích xuất trực tiếp các thuộc tính từ hình ảnh thô, những thuộc tính chưa được đào tạo trước nó sẽ học sau khi mạng được huấn luyện từ một nhóm các hình ảnh.

CNN được sử dụng nhiều trong các bài toán nhận dạng các object trong ảnh. Để tìm hiểu tại sao thuật toán này được sử dụng rộng rãi cho việc nhận dạng (detection)

* + 1. **Cấu trúc CNN**

CNN bao gồm tập hợp các lớp cơ bản bao gồm:

* Convolution layer + Nonlinear layer:
  + Convolution layer là lớp quan trọng nhất và cũng là lớp đầu tiên của của mô hình CNN. Lớp này có chức năng chính là phát hiện các đặc trưng có tính không gian hiệu quả. Trong tầng này có 4 đối tượng chính là: ma trận đầu vào, bộ filters, và receptive field, feature map.
  + ReLU (Rectified Linear Units, f = max (0, x)) là hàm kích hoạt phổ biến nhất cho CNN. Hàm ReLU được ưa chuộng vì tính toán đơn giản, giúp hạn chế tình trạng vanishing gradient, và cũng cho kết quả tốt hơn.
* Pooling layer: một số loại pooling layer phổ biến như là max-pooling, average pooling, với chức năng chính là giảm chiều của tầng trước đó
* Fully connected layer (là mỗi hidden layer): có chức năng chuyển ma trận đặc trưng ở tầng trước thành vector chứa xác suất của các đối tượng cần được dự đoán

Các lớp này liên kết với nhau theo một thứ tự nhất định. Thông thường, một ảnh sẽ được lan truyền qua tầng convolution layer + nonlinear layer đầu tiên, sau đó các giá trị tính toán được sẽ lan truyền qua pooling layer, bộ ba convolution layer + nonlinear layer + pooling layer có thể được lặp lại nhiều lần trong network. Và sau đó được lan truyền qua tầng fully connected layer và softmax để tính xác xuất ảnh đó chứa vật thế gì.



*Hình 8. Cấu trúc CNN trong nhận diện xe. Nguồn: pbcquoc.github.io*

* + 1. **Hoạt động CNN**
* Đầu vào của lớp tích chập là hình ảnh
* Chọn đối số, áp dụng các bộ lọc với các bước nhảy, padding nếu cần. Thực hiện tích chập cho hình ảnh và áp dụng hàm kích hoạt ReLU cho ma trận hình ảnh.
* Thực hiện Pooling để giảm kích thước cho hình ảnh.
* Thêm nhiều lớp tích chập sao cho phù hợp.
* Xây dựng đầu ra và dữ liệu đầu vào thành 1 lớp được kết nối đầy đủ (Full Connected).
  1. **Thư viện Tensorflow**
     1. **Tổng quan về Tensorflow**

Thư viện Tensorflow là thư viện mã nguồn mở và sử dụng phổ biến cho Machine Learning và Deep Learning đã giúp việc tiếp cận các bài toán trở nên đơn giản, nhanh chóng và tiện lợi hơn nhiều.Việc xử lí tính toán số học dựa trên biểu đồ mô tả sự thay đổi của dữ liệu.

Kiến trúc linh hoạt cho phép chúng ta triển khai tính toán trên một hoặc nhiều CPU hoặc GPU trong một máy tính để bàn, máy chủ, hoặc thiết bị di động với một API đơn.

Kiến trúc Tensorflow hoạt động được chia thành 3 phần:

* Tiền xử lý dữ liệu
* Dựng model
* Train và ước tính model

TensorFlow có các API với một số ngôn ngữ lập trình cho cả xây dựng và thực thi một đồ thị TensorFlow. Python API là hiện tại hoàn thiện nhất và dễ sử dụng nhất, nhưng API C++ có một vài ưu điểm về hiệu năng trong việc thực thi đồ thị, và hỗ trợ triển khai các thiết bị nhỏ như Android.

* + 1. **Một số khái niệm quan trọng**
       1. **Tensor**

Đơn vị dữ liệu chính của Tensorflow là Tensor hay Tensorflow là dòng chảy của những Tensor.Tensor là một kiểu dữ liệu dạng mảng có nhiều chiều. Hiểu một cách đơn giản là trong thế giới của Tensorflow, mọi kiểu dữ liệu đều được quy về một loại dữ liệu dạng mảng đó là Tensor.

Một tensor có thể là một **vector** hay **ma trận** của n-chiều không gian đại diện cho tất cả loại dữ liệu.

Rank của một Tensor là số chiều hay là bậc của nó. Việc phân Rank này khá quan trọng vì nó giúp phân loại dữ liệu của Tensor. Tùy thuộc vào các trường hợp của Rank mà Tensor có nhưng tên gọi khác nhau.

Shape là một bộ các biểu diễn số lượng phần tử có trong mỗi chiều.

Để hiểu hơn ta sẽ xem qua ví dụ sau:

* 3.0: Tensor với Rank là 0
* [1., 2., 3.]: Tensor với Rank là 1, là một vector có shape là [3]
* [ [1., 2., 3.], [4., 5., 6.] ]: Tensor với Rank là 2, là một ma trận có shape là [ 2, 3]
* [ [ [1., 2., 3. ] ], [ [7., 8., 9.] ] ]: Tensor có Rank là 3, là một ma trận có shape là [ 2, 1, 3]
  + - 1. **Graph**

Graph sử dụng framework dạng biểu đồ, là một tập hợp và mô tả các phép toán được thực hiện trong quá trình training. Một graph được cấu thành bởi 2 loại đối tượng chính:

* tf.Operation: Là các node của Graph. Operation mô tả sự tính toán để tạo ra các tensor.
* tf.Tensor: Là các cạnh của Graph. Chúng biểu diễn các giá trị dữ liệu xuyên suốt đồ thị. Hầu hết các hàm của Tensorflow đều trả lại tf.Tensors

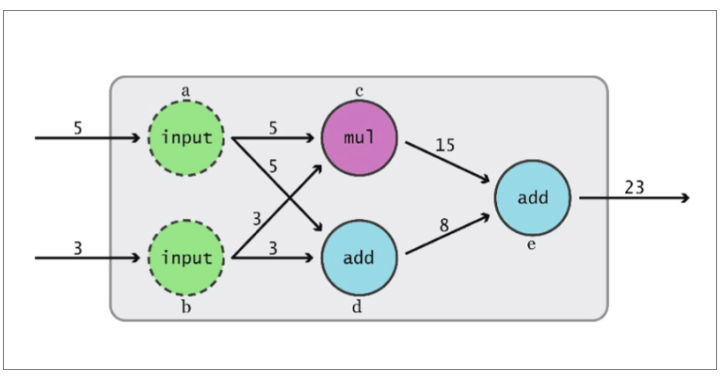


*Hình 5: Dòng chạy của Tensor. Nguồn: kipalog.com*

Tất cả các tính toán trong graph được thực hiện bằng cách kết nối các Tensor lại với nhau. Một Tensor có một node và một edge. Node mang operation toán học và sản xuất các output ở đầu cuối. Các edge giải thích mối quan hệ input/output giữa các node.

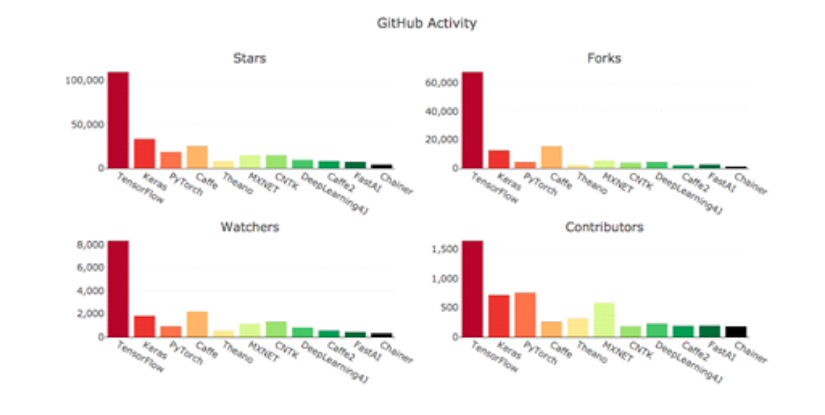
* + 1. **Cách Tensorflow hoạt động**

TensorFlow hoạt động theo cơ chế đồ thị tính toán tĩnh (static computational graph), nghĩa là ta phải xây dựng lên một đồ thị tính toán trước, sau đó mới thực hiện các tính toán trên đó. Khi đã có một graph, ta tạo ra một Session để thực hiện những tính toán đã được định nghĩa trên đồ thị



*Hình 6: Mô phỏng đồ thị được xậy dựng Tensorflow. Nguồn: idz.vn*

* + 1. **Ưu điểm của Tensorflow**
* TensorFlow có cấu trúc đáp ứng khi bạn có thể dễ dàng hình dung từng phần của biểu đồ.
* Tính linh hoạt nền tảng, có nghĩa là nó là mô-đun và một số phần của nó có thể độc lập trong khi các phần khác kết hợp lại.
* Dễ dàng huấn luyện trên CPU cũng như GPU để tính toán phân tán.
* Khả năng phân biệt tự động, có lợi cho các thuật toán học máy dựa trên độ dốc, nghĩa là bạn có thể tính toán các đạo hàm của các giá trị đối với các giá trị khác, dẫn đến mở rộng biểu đồ.
* Hỗ trợ nâng cao cho các luồng, tính toán không đồng bộ và hàng đợi.
* Tùy biến và nguồn mở.



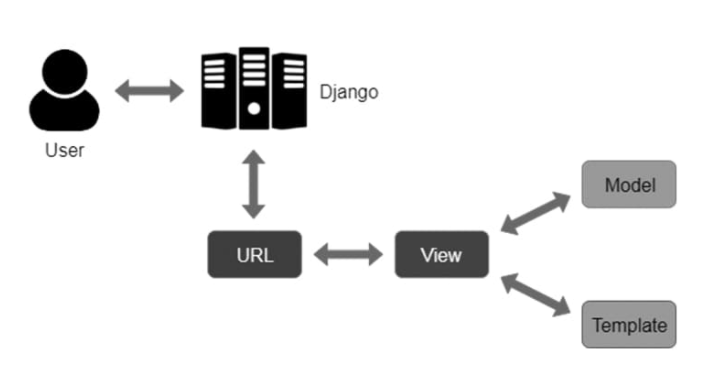
*Hình 7: Thống kê độ phổ biến framework TensorFlow. Nguồn: onetech.vn*

* + 1. **Hạn chế của Tensorflow**
* Có bộ nhớ GPU xung đột với Theano nếu được nhập trong cùng phạm vi.
* Không hỗ trợ OpenCL.
* Yêu cầu kiến thức chuyên sâu.
  + 1. **Các thuật toán nổi bật được hỗ trợ bởi Tensorflow**
* Linear regression.
* Classification.
* Deep learning classification.
* Deep learning wipe and deep
* Booster tree regression
* Booster tree classification
  1. **Django**
     1. **Tổng quan về Django**

Django là một Framework lập trình wed bậc cao, mã nguồn mở được viết bằng Ngôn ngữ lập trình Python.

Django giúp lập trình web trong thời gian ngắn. Hơn nữa, Django có tài liệu rất tốt và tuân theo nguyên tắc DRY (Đừng lặp lại chính mình) trong khi những Framework khác không quan tâm nhiều về điều này. Django cũng hỗ trợ ORM (Object Relistic Mapping).

* + 1. **Kiến trúc Django**



*Hình 9. Mô hình MVT của Django. Nguồn: niithnoi.edu.vn*

Django sử dụng mô hình MVT (Model-View-Temmplate), mô hình MVT được sử dụng trong khi tạo một ứng dụng với Tương tác người dung.

Mô hình này thì bao gồm code HTML với Django Templage Language (DTL)

Controller là mã được viết để kiểm soát sự tương tác giữa Model và View và Django dễ dàng chăm sóc nó.

Bất cứ khi nào người dùng người request, nó xử lý request của người dùng đó bằng Model, View và Template.

Nó hoạt động như một Controller để kiểm tra xem nó có khả dụng hay không bằng cách ánh xạ URL và nếu URL ánh xạ thành công thì View sẽ bắt đầu tương tác với Model và gửi lại Template cho người dùng dưới dạng respone.

* + 1. **Ưu điểm của Django**

Ngôn ngữ Python: Django được viết bằng python. Và nó mang đầy đủ tính chất của Python. Tính đơn giản, dễ học và dễ thực hiện, code ngắn hơn, thư viện lớn hơn.

Đảm bảo về tính bảo mật: Django rất coi trọng vấn đề bảo mật và giúp các nhà phát triển tránh được nhiều lỗi bảo mật phổ biến. Hệ thống xác thực người dùng của nó cung cấp một cách an toàn để quản lý tài khoản và mật khẩu người dùng.

Thư viện hỗ trợ đầy đủ: Django cung cấp các thư viện bao gồm các thao tác với chuỗi, web services, giao diện hệ điều hành và giao thức chuẩn.

Khả năng mở rộng: Django được sử dụng để đáp ứng nhu cầu lưu lượng truy cập lớn và cực lớn.

* + 1. **Hạn chế của Django**

Django không có khả năng quản lý các yêu cầu khác nhau cùng một lúc vì nó không trao quyền cho các thủ tục riêng lẻ để xử lý nhiều yêu cầu cùng một lúc.

Hệ thống ORM (Object Relation Mapping) cho phép các lập trình viên làm việc trên các cơ sở dữ liệu khác nhau cùng một lúc nhưng nó thiếu một số tính năng chủ yếu được cung cấp bởi các hệ thống ORM khác. Nó được thiết kế theo cách mà không cho phép các lập trình viên sử dụng SQLAlchemy hoàn toàn.

# **CHƯƠNG 2: MÔ TẢ BÀI TOÁN**

**I – MÔ TẢ BÀI TOÁN**

Việt Nam là một nước có đa dạng các loại hình phương tiện di chuyển và cá thể loại xe như xe máy, ô tô, xe ba gác, xe, xe cấp cứu, xe cứu hỏa, xe bộ….Từ việc có nhiều loại hình xe và các thể loại xe khác nhau dẫn đến việc điều tiết, phân luồng và quản lí các loại xe sẽ trở nên phức tạp hơn, đồng nghĩa là sẽ có nhiều loại biển báo giao thông hơn.

Điều kiện về thời tiết, địa hình hay về con người cũng là một yếu tố quan trọng trong việc điều khiển phương tiện giao thông, có thể điều khiển xe trong ban đêm, ánh sáng của xe quá yếu để có thể nhìn được biển báo bằng mắt thường hoặc khoảng cách xa biển báo để có thể thấy biển báo.



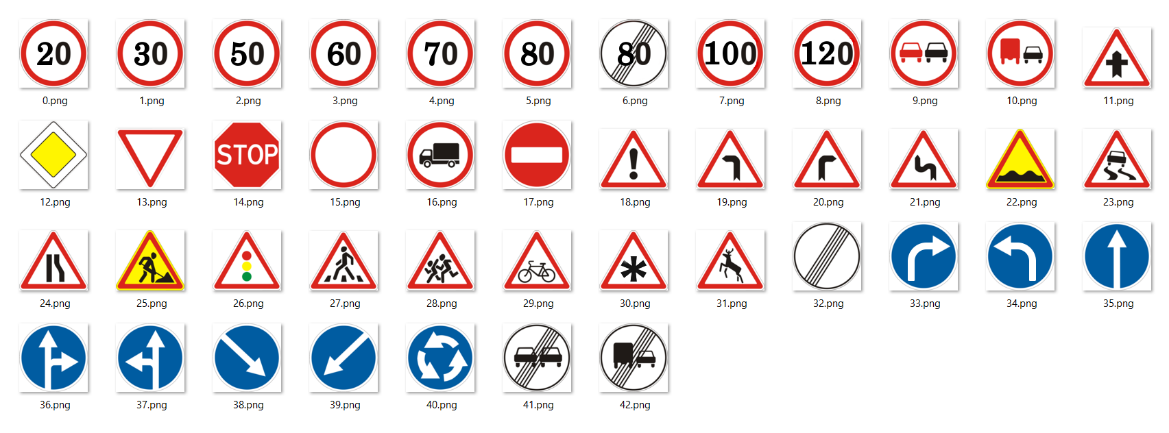
*Hình 10: Các trường hợp biển báo có thể gặp khi di chuyển.*

Để việc quản lí, phân luồng được nhanh chóng, tiết kiệm thời gian cũng như giải quyết các vấn đề liên quan đến các yếu tố thời tiết, địa hình. Nhóm đã chọn ra đề tài “Phân Loại Biển Báo Giao Thông”, từ những hình ảnh đầu vào, ta sẽ xác định hình ảnh đó thuộc dạng biển nào và đưa ra trả lời trong thời gian nhanh nhất có thể, vì trên thực tế có rất nhiều các loại biển báo khác nhau. Hiện nay thì số lượng biển báo rất là nhiều, có đến 6 nhóm các loại biển báo, trong các nhóm sẽ có nhiều kiểu biển báo khác nhau, nên việc để ghi nhớ được những loại biển đó cũng rất mất thời gian, khi đó ta cần một bên thứ ba để cải thiện cho việc phân loại và nhận diện các loại biển một cách nhanh chóng.

* 1. **Mô tả tập dữ liệu**

Tập dữ liệu này được sử dụng từ tập dữ liệu German Traffic Sign Recognition Benchmark (GTSRB) được cung cấp bởi Kanggle. Dữ liệu ban đầu khoảng 40.000 hình cho tập training, tất cả đều có dạng (.png) với các kích thước và độ rõ nét của hình khác nhau. Hình ảnh trong tập này được chia thành 42 tệp tương ứng với mỗi tệp là một loại biển báo khác nhau. Tập dữ liệu trani này được dùng để huấn luyện cho máy.

Ngoài ra, ta còn tập Test với khoảng gần 13.000 hình ảnh ở dạng .png và cũng có các kích thước và độ rõ nét trên từng hình là khác nhau. Điều khác là trong tập này các hình ảnh không được dán nhãn, các hình ảnh ở các dạng biển báo khác nhau sẽ nằm chung với nhau. Mục đích của bài toán là tìm ra được hình ảnh trong tập test đó là thuộc loại biển báo nào.

**

*Hình 11: Các loại biển báo trong tập dữ liêu GTSRB.*

Input sẽ là hình ảnh nằm trong tập chưa được train, output là ảnh đó là ảnh gì, hay cụ thể hơn là biển báo gì.



*Hình 12: Giao diện website sử dụng framework Django.*

Kết quả sau cùng sẽ là thông tin về ảnh đó kèm theo hình ảnh vừa được chọn để nhận biết.

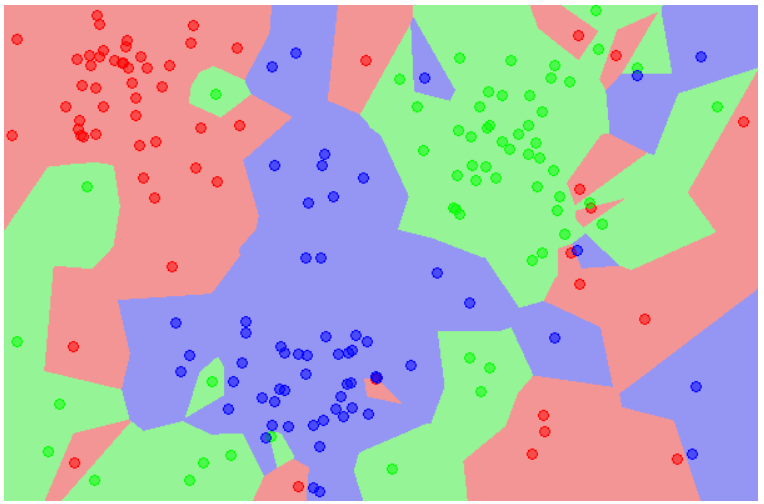
* 1. **Sơ lược về những thuật toán có sử dụng trong bài toán**
     1. **Thuật toán KNN (K-nearest neighbors)**

K-nearest neighbor là một trong những thuật toán supervised-learning đơn giản nhất (mà hiệu quả trong một vài trường hợp) trong Machine Learning.

Khi training, thuật toán này không học một điều gì từ dữ liệu training, đây cũng là lý do thuật toán này được xếp vào loại lazy learning, mọi tính toán được thực hiện khi nó cần dự đoán kết quả của dữ liệu mới

Với KNN, trong bài toán Classification, label của một điểm dữ liệu mới được suy ra trực tiếp từ K điểm dữ liệu gần nhất trong training set. Label của một test data có thể được quyết định bằng major voting (bầu chọn theo số phiếu) giữa các điểm gần nhất, hoặc nó có thể được suy ra bằng cách đánh trọng số khác nhau cho mỗi trong các điểm gần nhất đó rồi suy ra label.

Một cách ngắn gọn, KNN là thuật toán đi tìm đầu ra của một điểm dữ liệu mới bằng cách chỉ dựa trên thông tin của K điểm dữ liệu trong training set gần nó nhất (K-lân cận), không quan tâm đến việc có một vài điểm dữ liệu trong những điểm gần nhất này là nhiễu.



*Hình 13: Mô phỏng thuật toán KNN. Nguồn: machinelearningcoban.com*

Ví dụ trên đây là bài toán Classification với 3 classes: Đỏ, Lam, Lục. Mỗi điểm dữ liệu mới (test data point) sẽ được gán label theo màu của điểm mà nó thuộc về. Trong hình này, có một vài vùng nhỏ xem lẫn vào các vùng lớn hơn khác màu. Ví dụ có một điểm màu Lục ở gần góc 11 giờ nằm giữa hai vùng lớn với nhiều dữ liệu màu Đỏ và Lam. Điểm này rất có thể là nhiễu. Dẫn đến nếu dữ liệu test rơi vào vùng này sẽ có nhiều khả năng cho kết quả không chính xác.

* + 1. **Pre-trained CNN Model**

Pre-trained CNN Model là mô hình được tạo ra bởi một số người khác để giải quyết một vấn đề tương tự. Thay vì xây dựng mô hình từ đầu để giải quyết một vấn đề tương tự, bạn sử dụng mô hình được đào tạo về vấn đề khác làm điểm khởi đầu.

Một mô hình được đào tạo trước có thể không chính xác 100% trong ứng dụng của bạn, nhưng nó giúp tiết kiệm những nỗ lực lớn cần thiết để phát minh lại bánh xe. Hãy để tôi cho bạn thấy điều này với một ví dụ gần đây.

Một số Pre-trained Model nổi bật như là: VGG-16, Inceptionnet, ResNet, InceptionResNet, MobieNet, …

Ta nên cẩn thận trong khi chọn Pre-trained Model mà ta nên sử dụng trong trường hợp của mình. Nếu việc chọn mô hình không phù hợp thì sẽ dẫn đến dự đoán mà chúng ta nhận được sẽ rất không chính xác. Ví dụ: một mô hình được đào tạo trước đây để nhận dạng giọng nói sẽ không đạt kết quả tốt nếu chúng ta sử dụng nó để xác định hình ảnh.

Ở bài toán phân loại biển báo giao thông thì sử dụng sử dụng mạng MobieNet.

* 1. **Các bước xử lí**
* Cài đặt và import các thư viện cần sử dụng
* Load tập dữ liệu lên.
* Đọc và chuyển đổi nhãn từ dạng số thành nhãn dạng tên biển báo
* Hình ảnh không có cùng kích thước sẽ biến đổi chúng để có kích thước 224x224 bằng nhau.
* Xây dựng model
* Train model
* Lưu model vừa được train lại.
* Đánh giá độ chính xác trên model vừa được train
* Load dữ liệu từ tập test.
* Kiểm tra độ chính xác của dữ liệu.
* Thử nghiệm với các trường hợp đặc biệt trên hình ảnh tự chụp.
  1. **So sánh độ đo giữa các thuật toán**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Thuật toán | CNN | VGG-16 | MobiNet | KNN |
| Độ chính xác | 98% | 96% | 63% | 38% |
| Thời gian | >1h | 45p – 1h | 25p | 30p |
| Thời gian cải thiện | 5p | 3-5p | 3-5p | 3-5p |
| h: giờ, p: phút |  |  |  |  |

**Bảng: So sánh giữa các thuật toán**

Lí do sự chênh lệch độ chính xác mà nhóm nhận thấy là:

* Bản chất của thuật toán KNN là tìm dữ liệu đầu ra của dữ liệu bằng cách dựa vào K điểm dữ liệu train set gần nó mà không quan tâm dữ liệu có bị nhiễu hay không, dù cho K điểm gần nó có dự đoán sai thì kết dữ liệu ra sẽ vẫn được chọn là điểm đã dự đoán sai, nên độ chính xác không đạt được ở mức cao.
* Mạng MobiNet là một Pre-trained Model, vì là được đào tạo sẵn nên có thể mục đích đào tạo nên mạng sẽ khác hoặc là đào tạo ít hơn, dẫn đến sẽ không phù hợp trong một số trường hợp. Trong trường hợp phân loại biển báo thì hiệu quả kém. Thứ hai, điểm cải tiến của mô hình này là sử dụng tích chập Depthwise Separable Convolution, mục đích là dùng để giảm kích thước và giảm độ phức tạp khi tính toán. Do đó nên mạng này sẽ phù hợp hơn nếu sử dụng trên các thiết bị điện thoại di động hoặc các hệ thống nhúng.
* Mạng VGG-16 cũng là một Pre-trained Model nhưng do kiến trúc sâu hơn, bao gồm 12 lớp tích chập 2 chiều và ta cũng có thể cấu hình trên GPU/TPU nên độ chính xác so với mạng CNN thì cũng xấp sỉ.

**CHƯƠNG 3: KẾT LUẬN**

* 1. **Kết quả đạt được**
     1. **Ý nghĩa khoa học**

Báo cáo đã trình bày cơ sở lý thuyết Deep Learning, Nerual Network, mô hình CNN, thư viện Tensorflow, Django. Nội dung chính của đề tài là trình bày Deep Learning, thư viện Tensorflow và thuật toán CNN và áp dụng để xậy dựng bài toán phân loại và nhận diện biển báo giao thông. Sử dụng Django để xây dựng website. Thông qua đề tài, chúng tôi biết được quá trình xây dựng và giải quyết bài toán phân loại đối tượng trong Deep Learning và biểu diễn kết quả của bài toán thông qua mô hình và để gần gũi với người dùng thì được thể hiện dưới dạng ứng dụng trang web, nâng cao hiểu biết và kỹ năng sử dụng Python cùng các thư viện hỗ trợ để phân tích dữ liệu. Bên cạnh đó, chúng tôi còn nâng cao khả năng đọc hiểu tài liệu, khả năng làm việc nhóm và khả năng trình bày báo cáo khoa học.

* + 1. **Ý nghĩa thực tiễn**

Chúng tôi biết được nhiều hơn các thuật toán trong lĩnh vực Deep Learning, thư viện Tensorflow đặc biệt là những bài toán phân loại (Classification) và nhận diện (identification) sử dụng CNN. Áp dụng được kiến thức đã tìm hiểu và được giảng dạy để thu thập và phân tích bài toán phân loại đối tượng một cách tốt nhất.

Thông qua đề tài, chúng tôi đạt được các kết quả:

* Nắm rõ cơ sở lý thuyết về Deep Learning, thư viện Tensorflow
* Hiểu được thuật toán CNN.
* Vận dụng thuật toán Cnn vào bài toán.
* Xây dụng được mô hình bài toán.
* Sử dụng Django làm website.

Cụ thể chúng tôi có đuợc tập dữ liệu về biển báo giao thông, và xây dựng được website phân loại và nhận diện biển báo giao thông. Với trang website này giúp cho các tài xế khi tham gia giao thông nhận diện được biển báo giao thông làm giảm được tai nạn giao thông.

* 1. **Hạn chế**

Do sự giới hạn về thời gian cũng như nguồn lực, chúng tôi chỉ mới phân loại và nhận diện từ những hình ảnh có sẵn chưa từ nguồn dữ liệu tự thu thập nên tỷ lệ khi thực tế sẽ có sai sót.

* 1. **Hướng phát triển**

Hướng phát triển tiếp theo của đề tài này là xậy dựng thành một ứng dựng và phát triển các thiết bị giao thông đường bộ để mọi người có thể sử dụng rỗng rãi và tiện lợi. Ứng dụng sẽ phát triển hơn khi có thể phát hiện, phân loại và nhận diện từ nguồn là video được quay trực tiếp từ camera của phương tiện. Nhóm có thể nghiên cứu thêm những thuật toán khác để cải thiện độ chính xác cũng như tốc độ nhận diện của ứng dụng.

# **TÀI LIỆU THAM KHẢO**

[1] GTSRB - German Traffic Sign Recognition Benchmark. (2017). Mykola <https://www.kaggle.com/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign>

[2] Traffic Sign Classification with Keras and Deep Learning. (2019). Adrian Róebrock

<https://www.pyimagesearch.com/2019/11/04/traffic-sign-classification-with-keras-and-deep-learning/>

[3] 23 Amazing Deep Learning Project Ideas. (2020). Team Data Flair

<https://data-flair.training/blogs/deep-learning-project-ideas/>

[4] Basic classification: Classify images of clothing. (2020). Team Tensorflow Core

<https://www.tensorflow.org/tutorials/keras/classification?hl=vi&fbclid=IwAR2qmQLjLrrfqZh_oNTnbSzwAO1YoAl4FAfyKjyVfbvCi27YXFCyQ-QCvpM>

[5] Deep learning vs machine learning: a simple way to understand the difference. (2020). Brett Grossfeld

<https://www.zendesk.com/blog/machine-learning-and-deep-learning/#:~:text=To%20recap%20the%20differences%20between,intelligent%20decisions%20on%20its%20own>

[6] Creating a Machine Learning Based Web Application Using Django. (2020). Akash Chauhan

<https://towardsdatascience.com/creating-a-machine-learning-based-web-application-using-django-5444e0053a09>

[7] Full-stack Django & ReactJS || Home Loan Classification Problem || Machine Learning Application. (2020). Technology for Noobs

<https://www.youtube.com/watch?v=Tnto7gi6Zzo>

[8] K-nearest neighbors (2017). Vũ Hưu Long

<https://machinelearningcoban.com/2017/01/08/knn/>

[9] A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. (12-2018). Saha.

<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

[10] Sách deep learning cơ bản v2. (8-2020). Nguyễn Thanh Tuấn. <https://drive.google.com/file/d/1lNjzISABdoc7SRq8tg-xkCRRZRABPCKi/view>