**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

----------🙞🙜🕮🙞🙜----------

**Môn: MẠNG XÃ HỘI**

***Giảng viên hướng dẫn:***

**ThS. Nguyễn Thị Anh Thư**

***Nhóm sinh viên thực hiện:***

**Trần Văn Điệp 18520601**

**Đỗ Hùng Dũng 18520629**

**Dương Thị Nguyệt Minh 18521095**

**Hoàng Lê Nam 18521120**

***Lớp:*** **IS353.M21**

🙡🙢 Tp. Hồ Chí Minh, 3 – 2022 🙠🙣

# LỜI CẢM ƠN

Trong thời gian tìm hiểu đề tài, nhóm xin đặc biệt gửi lời cảm ơn chân thành đến Cô Th.s Nguyễn Thị Anh Thư (Giảng viên dạy môn ***Mạng xã hội***). Cô đã cung cấp kiến thức, tài liệu cùng các bài giảng trên lớp cực kỳ dễ hiểu. Đặc biệt hơn là những tài liệu kèm theo của môn học được cô cung cấp thêm rất chi tiết, phục vụ xuyên suốt quá trình học tập và thực hành của cả lớp.

Xuất phát từ mục đích học tập tìm hiểu về môn học ***Mạng xã hội*** cùng với sự tò mò về cách xử lý các bài toán sử dụng các loại mạng kết hợp với ứng dụng công nghệ Blockchain. Nhóm đã đưa ra quyết định lựa chọn đề tài: “***Blockchain technology: A DNN token-based approach in healthcare and COVID-19 to generate extracted data”.***

Dưới đây là những gì nhóm đã cố gắng tìm hiểu trong suốt quá trình học tập. Dù biết sẽ còn rất nhiều sự thiếu xót trong bài báo cáo này nhưng dù sao đó là nỗ lực của mỗi thành viên trong nhóm. Nhóm rất mong nhận sự góp ý từ phía cô nhằm rút ra những kinh nghiệm quý báu và hoàn thiện vốn kiến thức để nhóm có thể tiếp tục hoàn thành những chủ đề khác trong thời gian tương lai.

Nhóm xin chân thành cảm ơn Cô!

# NHẬN XÉT CỦA GIẢNG VIÊN

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

# MỤC LỤC

[LỜI CẢM ƠN 2](#_Toc102591097)

[NHẬN XÉT CỦA GIẢNG VIÊN 3](#_Toc102591098)

[MỤC LỤC 4](#_Toc102591099)

[Phần 1: TỔNG QUAN 6](#_Toc102591100)

[1. Thông tin bài báo khoa học 6](#_Toc102591101)

[1.1. Thông tin 6](#_Toc102591102)

[1.2. Phát biểu bài toán 8](#_Toc102591103)

[1.3. Thách thức của bài toán 9](#_Toc102591104)

[1.4. Đối tượng và phạm vi 9](#_Toc102591105)

[1.5. Mục tiêu 10](#_Toc102591106)

[1.6. Phương pháp đề xuất 11](#_Toc102591107)

[Phần 2: CÁC NGUYÊN CỨU LIÊN QUAN 13](#_Toc102591108)

[2. Các nghiên cứu sử dụng A DNN token-based 13](#_Toc102591109)

[Phần 3: MÔ HÌNH GIẢI BÀI TOÁN(FRAMEWORK) 15](#_Toc102591110)

[3. Tổng quan về framework 15](#_Toc102591111)

[3.1. Các bước chuẩn bị dữ liệu 15](#_Toc102591112)

[3.2. Các mô hình sử dụng 21](#_Toc102591113)

[3.3. Dữ liệu trích xuất đặc trưng 29](#_Toc102591119)

[Phần 4: THỰC NGHIỆM 31](#_Toc102591120)

[4. Tổng quan bộ dữ liệu và cách tổ chức thực nghiệm 31](#_Toc102591121)

[4.1. Chi tiết bộ dữ liệu 31](#_Toc102591122)

[4.2. Thực nghiệm 34](#_Toc102591123)

# Phần 1: TỔNG QUAN

## Thông tin bài báo khoa học

### Thông tin

* Tên bài báo khoa học: ***Blockchain technology: A DNN token-based approach in healthcare and COVID-19 to generate extracted data.***
* Thời gian xuất bản lần đầu: 23 – 07 – 2021.
* Tác giả: ***Basetty Mallikarjuna***, ***Gulshan Shrivastava***, ***Meenakshi Sharma***
* ***Basetty Mallikarjuna***: Ông nhận bằng tiến sỹ Khoa học máy tính tại trường School of Computing Science and Engineering, Đại học Bharathiar Coimbatore và M.Tech. Hiện tại là trợ lý giáo sư (Hạng III) tại trường School of Computing Science and Engineering tại Đại học Galgotias, thuộc quận Gautam Buddh Nagar, thành phố Greater Noida.
* Kinh Nghiệm quá khứ:
* Từng làm trợ lý giáo sư tại Đại học GITAM.
* Từng là trợ lý giáo sư tại Nedurumalli Balakrishna Reddy Institute of Science and Technology (NBKRIST)
* Có 11 năm kinh nghiệm giảng dạy sau khi tốt nghiệp, và làm trợ lý giáo sư tại PEC Engineering College Hyderabad.
* Link thông tin tác giả: <https://www.galgotiasuniversity.edu.in/school-engineering-technology-socse-cse-faculty-basetty-mallik-arjuna.asp>
* ***Gulshan Shrivastava***: Ông nhận bằng tiến sĩ (Khoa học và Kỹ thuật máy tính) tại trường National Institute of Technology Patna (NIT Patna), India, và M.Tech. Hiện tại, ông đang làm trợ lý giáo sư tại khoa CSE thuộc Đại học Sharda thành phố Greater Noida, U.P.
* Kinh Nghiệm quá khứ:
* Trước khi đảm nhận vai trò hiện tại, ông đã liên kết với Đại học Galgotias và Dronacharya Group of Institutions, Greater Noida, U.P., Ấn Độ.
* Ông đã đến thăm công ty Datec Group, tại Papua New Guinea (PNG) với tư cách là một nhà nghiên cứu và huấn luyện kỹ thuật tự do.
* Ông đã nhận được nhiều chứng chỉ quốc tế từ Coursera, NPTEL, Sun Microsystem, v.v…
* Ông có 5 bằng bằng sáng chế (1 đã được cấp, 4 đã xuất bản) và đã xuất bản 55 bài báo, sách và bài xã luận trên tạp chí International Journals và Hội nghị quốc tế có uy tín cao như: IEEE, Elsevier, Wiley, ACM, Springer, v.v….
* Link thông tin tác giả: <https://www.sharda.ac.in/faculty/details/dr-gulshan-shrivastava>
* ***Meenakshi Sharma***: Bà là hiệu trưởng của trường Đại học University Center of Research & Development (UCRD), và là phó giáo sư tại trường School of Computer Science & Engineering, Đại học Galgotias, thành phố Greater Noida, Ấn Độ.
* Kinh Nghiệm quá khứ:
* Bà đã có hơn 16 năm kinh nghiệm trong việc giảng dạy và nghiên cứu.
* Bà là một chuyên gia có trình độ cao với bằng tiến sĩ Computer Science and M.Tech và Computer Science & Engineering (cả hai đều từ Đại học Kurukshetra).
* Bà ấy đã xuất bản hơn 60 bài báo nghiên cứu trên các tạp chí uy tín cao như: IEEE Transaction, Elsevier, Wiley, v.v…
* Bà đã được cấp 4 bằng sáng chế quốc tế và có 7 bằng sáng chế cấp quốc gia được công bố.
* Bà đã được trao giải thưởng dành cho giáo viên và là nhà nghiên cứu xuất sắc nhất năm 2017-2018.
* Bà là một thành viên có giá trị của các hiệp hội kỹ thuật khác nhau, bao gồm thành viên cấp cao trong IEEE, ISTE, ACM, InSc, ISDS Society, Japan, IEAE, và nhiều tổ chức khác.
* Link thông tin tác giả: <https://www.cgu.edu/people/meenakshi-sharma/>
* Được xuất bản phần 39 tập 3 trong tạp chí: ***Expert Systems: The Journal of Knowledge Engineering.***
* Link bài báo: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8420355/>
* Link dataset: <https://www.kaggle.com/sudalairajkumar/covid19-in-india>

### Phát biểu bài toán

* Phát biểu: Sử dụng mạng thần kinh sâu (DNN), công nghệ blockchain trong chăm sóc sức khỏe và đại dịch COVID ‐ 19, đưa ra quy trình hợp đồng thông minh, để xác định dữ liệu được trích xuất tính năng (FED) từ dữ liệu hiện có. Đồng thời, dữ liệu mới sẽ hữu ích để phân tích các bệnh trong tương lai. Từ đó, qua phân tích để đưa ra phương pháp điều trị thích hợp.
* ***Input:*** Dữ liệu hồ sơ bệnh của các bệnh nhân ở Ấn Độ (bệnh tim, bệnh ung thư máu, bệnh tiêu hóa, bệnh thương hàn, …).
* ***Output:*** Đưa ra tập các trích xuất đặc trưng của bệnh nhân dựa trên hợp đồng sức khoẻ điện tử thông minh. Nhằm phục vụ cho quá trình chuẩn đoán bệnh (Covid 19) cho bệnh nhân.

### Thách thức của bài toán

* Trong lĩnh vực khoa học kỹ thuật, những nghiên cứu về chăm sóc sức khoẻ, dự đoán mầm bệnh, chuẩn đoán truyền nhiễm luôn là một trong những chủ đề nghiên cứu được nhiều người chú ý đến. Nhưng để thực hiện nó những nhà nghiên cứu phải đảm bảo kiến thức về cả y học lẫn công nghệ nằm áp dụng chúng một cách chính xác nhất, mang lại hiệu quả cao nhất cho con người và xã hội. Tuy nhiên, để có thể đạt được một kết quả như mong muốn cần huấn luyện một lượng dữ liệu lưu trữ đủ lâu để có được sự chính xác cao và dĩ nhiên xử lý một lượng lớn dữ liệu liên quan đến lĩnh vực y tế sức khoẻ đó chính là một vấn đề lớn.
* Ngoài ra, thời gian cũng là một vấn đề khiến nhà nghiên cứu quan tâm rất lớn khi sự bùng nổ của dịch bệnh ngày một trở nên nhanh chóng và rộng khắp trên thế giới.

### Đối tượng và phạm vi

* Các công nghệ chăm sóc sức khỏe trong đại dịch Covid-19 đã phát triển vô cùng trong các lĩnh vực khác nhau. Công nghệ Blockchain là một công nghệ chìa khóa trao tay như vậy, đang chuyển đổi dữ liệu một cách an toàn; Để lưu trữ hồ sơ sức khỏe điện tử (EHRs), phát triển các thuật toán học tập sâu, truy cập dữ liệu, xử lý dữ liệu giữa các bác sĩ và bệnh nhân để truy cập EHRs ở dạng seedger phân tán. Công nghệ Blockchain cũng được thực hiện để cung cấp dữ liệu trong đám mây và liên hệ với số lượng lớn dữ liệu chăm sóc sức khỏe, rất khó khăn và phức tạp để xử lý. Vì sự phức tạp trong phân tích dữ liệu đang tăng lên từng ngày, nó đã trở nên cần thiết để giảm thiểu rủi ro phức tạp dữ liệu.
* Trong thời hiện đại, mỗi người cần duy trì hồ sơ sức khỏe điện tử (EHRs) cho toàn bộ cuộc đời của họ bằng cách sử dụng blockchain và ngành chăm sóc sức khỏe hiện tại sẽ chuyển từ bảo trì hồ sơ bệnh nhân kỹ thuật số EHRs đến hệ thống quản lý phi tập trung trong khi sử dụng Công nghệ Blockchain. Mặc dù sử dụng các bác sĩ EHRs có thể phân tích và cung cấp điều trị, nó rất hữu ích để cung cấp tính năng trích xuất dữ liệu, sẽ hữu ích cho các bác sĩ và y tá kê đơn thuốc và duy trì hồ sơ kỹ thuật số của bệnh nhân để cung cấp điều trị hiệu quả.

### Mục tiêu

* Bài báo này hỗ trợ phân tích mạng thần kinh sâu (DNN) trong đại dịch chăm sóc sức khỏe và Covid-19 và cung cấp quy trình hợp đồng thông minh, để xác định tính năng trích xuất dữ liệu (FED) từ dữ liệu hiện có. Đồng thời, sự đổi mới sẽ hữu ích để phân tích các bệnh trong tương lai. Phương pháp đề xuất cũng phân tích các bệnh hiện có đã được báo cáo và cực kỳ hữu ích để hướng dẫn các bác sĩ trong việc cung cấp điều trị phù hợp và cứu mạng. Để đạt được điều này, dữ liệu lớn được tích hợp bằng ngôn ngữ kịch bản Python trong nhiều thư viện khác nhau để thực hiện nhiều chức năng y tế và chăm sóc sức khỏe để suy ra kiến thức hỗ trợ trong chẩn đoán các bệnh chính như bệnh tim, ung thư máu, dạ dày và covid-19.
* Thông tin sức khỏe về định dạng kỹ thuật số bằng blockchain giúp quản lý hồ sơ sức khỏe một cách an toàn và rất hữu ích để ngăn ngừa các bệnh cho các địa điểm khác nhau. Kể từ khi, vị trí cụ thể mà những người mắc bệnh tương tự, để xác minh và thực thi dữ liệu với hợp đồng thông minh, các bệnh, nhiễm trùng và phòng ngừa để cung cấp các biện pháp phòng ngừa có thể với dữ liệu được trích xuất (FED) trở nên khả thi. Bài báo này đề xuất sự khai thác tính năng tăng cường để giảm dữ liệu chăm sóc sức khỏe bằng cách xử lý chính xác cho người dùng cuối dựa trên công nghệ blockchain. Mục tiêu là để giảm dữ liệu và xử lý các loại khác nhau cho các nhóm có thể quản lý.
* Để tính năng trích xuất làm giảm số lượng tài nguyên trong ngành chăm sóc sức khỏe và duy trì tính toàn vẹn của dữ liệu. Trong trích xuất tính năng, các biến hoạt động như các tính năng làm giảm lượng dữ liệu cần xử lý trong chăm sóc sức khỏe trong khi đưa chính xác bộ dữ liệu gốc.
* Việc truyền dữ liệu bằng cách sử dụng trích xuất tính năng mà không mất dữ liệu quan trọng hoặc có liên quan.
* Việc trích xuất tính năng làm giảm lượng dữ liệu dư thừa trong chăm sóc sức khỏe.
* Dữ liệu dự phòng rất hữu ích cho cách tiếp cận học máy để xây dựng các biến để hoạt động như các tính năng.

### Phương pháp đề xuất

* Các cảm biến thu thập dữ liệu, nó phân loại các mã thông báo cung cấp cho bệnh nhân trong một nhật ký tuyệt đối trong lịch sử y tế của họ và nó cũng tạo ra FED (Feature Extracted Data). Các bệnh nhân truy cập dữ liệu thông qua công nghệ blockchain.
* Thông tin bệnh nhân nhận được từ các cảm biến được truy cập thông qua API và xử lý Ethereum Smart Contract (SC), các kiến ​​trúc được quan sát thấy có hỗ trợ dựa trên đám mây.
* Dữ liệu được thu thập từ các cảm biến và có thể được đánh giá bằng các thuật toán học máy, từ đó tạo ra tính năng được trích xuất từ ​​bộ dữ liệu đã cho. Các cảm biến có được dữ liệu theo ứng dụng bệnh nhân, nó tạo ra dữ liệu và thông tin đáng tin cậy cho hợp đồng thông minh (SC), Registrar Contract (RC) sử dụng khóa công khai mật mã để khám phá bản ghi bệnh nhân trong Patient-Provider Relationship Contract(PPR), máy chủ cơ sở dữ liệu cục bộ cung cấp xác thực, mã hóa và xử lý với dữ liệu thông qua blockchain. Bảo mật và chia sẻ dữ liệu đã được thực hiện tại phía Provider. Việc giải mã, nhận khóa, lưu trữ khóa, phát hành chính cũng được thực hiện tại đây.
* Cơ sở dữ liệu Provider thực hiện truy vấn máy chủ và cung cấp truy vấn kết quả với mạng. Provider kiểm tra, xác minh các hợp đồng blockchain, được phép truy vấn nếu địa chỉ không hợp lệ. Trình quản lý Electronic Health Records (EHR) trung tâm được lưu trữ trong Cloud Data Server, nó sẽ duy trì phần mềm Cloud DB (Ethereum duy trì phần mềm cơ sở dữ liệu SQLite), phần mềm cơ sở dữ liệu đám mây linh hoạt với dịch vụ giao diện web và được xây dựng trên khung Python, dịch vụ cơ sở dữ liệu đám mây tương thích với các thiết bị di động, người dùng cuối, hồ sơ.
* Thông tin bệnh nhân nhận được từ các cảm biến và truyền vào lưu trữ đám mây. EHRs được thực hiện với hỗ trợ dựa trên đám mây, công nghệ blockchain.
* **Dữ liệu đám mây** là một dịch vụ được cung cấp bởi một công ty hoặc doanh nghiệp nào đó, có thể là các doanh nghiệp trong nước và nước ngoài cho phép người dùng có thể lưu trữ, quản lý, chia sẻ, sao lưu các dữ liệu như hình ảnh, tập tin,…
* **Smart contract(SC):** Smart Contract (hay Hợp đồng thông minh) là các chương trình chạy trên blockchain. Hợp đồng thông minh cũng giống như một hợp đồng kỹ thuật số bị bắt buộc thực hiện bởi một bộ quy tắc cụ thể. Các quy tắc này do bộ mã máy tính xác định trước mà tất cả các nút (node) trong mạng đều phải sao chép và thực thi các quy tắc đó.
* Về bản chất, Smart Contract chỉ là một đoạn mã chạy trên một hệ thống phân tán (blockchain), cho phép tạo ra các giao thức Permissionless (tức là không cần trao quyền). Điều đó có nghĩa là:
* Hai bên trong hợp đồng có thể đưa ra các cam kết thông qua blockchain mà không cần phải biết về danh tính hay tin tưởng lẫn nhau.
* Họ có thể đảm bảo rằng nếu các điều kiện của hợp đồng không được thỏa mãn, hợp đồng sẽ không được thực thi.
* Ngoài ra, việc sử dụng hợp đồng thông minh loại bỏ nhu cầu đối với các bên trung gian, giúp giảm đáng kể chi phí hoạt động.

# Phần 2: CÁC NGUYÊN CỨU LIÊN QUAN

## Các nghiên cứu sử dụng A DNN token-based

* DNN chắc chắn hỗ trợ phân tích chăm sóc sức khỏe để tạo ra dữ liệu được trích xuất. Dữ liệu được trích xuất nổi bật đến từ dữ liệu được đào tạo mang lại sự tự giám sát và thực hiện các biện pháp phòng ngừa cho kế hoạch đặc trưng, ​​mọi người đang bị các bệnh lớn như bệnh tim, ung thư máu để phân loại dữ liệu và dự đoán dữ liệu chăm sóc sức khỏe. Các blockchain được liên kết với ứng dụng web chăm sóc sức khoẻ và cung cấp một nền tảng bảo mật phi tập trung cung cấp trao đổi bitcoin và cách sử dụng dữ liệu y tế. Sau đó, 'Medtokens' sau đây thay đổi năng suất trong mọi lĩnh vực chăm sóc sức khỏe (Mamoshina et al., 2016). MedtoEKNS hữu ích nhất cho Blockchain riêng tư và rất dễ sử dụng, mua dịch vụ lưu trữ, hầu hết các bác sĩ, các công ty dược phẩm đều nhận được mã thông báo và bác sĩ cung cấp cho TelemeDicine với việc sử dụng mã thông báo y tế.
* Xác định các bệnh chính trong khu vực và cung cấp các biện pháp phòng ngừa với các cơ sở y tế tốt hơn bằng cách sử dụng dữ liệu được trích xuất. Các bác sĩ và thuốc có thể có sẵn theo các bệnh được báo cáo trong khu vực, mỗi EHRs có thể phân tích để biết lý do tại sao một khu vực cụ thể bị ảnh hưởng bởi một căn bệnh cụ thể. Bằng cách phân tích số lượng bệnh, chúng ta có thể phân tích bệnh nào xảy ra nhiều hơn về số lượng. Từ đó, các loại vắc-xin và các biện pháp khác có thể được thực hiện. Phân tích cũng giúp phân tách dữ liệu theo một vị trí cụ thể để biết vị trí nào bị ảnh hưởng nghiêm trọng bởi những bệnh nào và những biện pháp nào có thể được thực hiện để ngăn ngừa các bệnh hoặc giảm số lượng của nó. Nó giúp bệnh viện của khu vực duy trì cơ sở y tế cho những bệnh đó mà khu vực của họ bị ảnh hưởng để giảm vùng tử thần.

# Phần 3: MÔ HÌNH GIẢI BÀI TOÁN(FRAMEWORK)

## Tổng quan về framework

* Xuyên suốt bài báo khoa học tác giả sử dụng mạng DNN và tạo trích xuất đặc trưng(FED).

### Các bước chuẩn bị dữ liệu

* PPR: mối quan hệ giữa nhà cung cấp và bệnh nhân
* SC: hợp đồng thông minh
* RC: hợp đồng đăng ký
* **PoS: bằng chứng cổ phần**
* Proof of Stake (POS) là một thuật toán đồng thuận của blockchain. Có thể hiểu, đây là Bằng chứng ký gửi hay Bằng chứng cổ phần. Trong đó, các Node phải Stake (đặt chỗ -ký gửi) coin để tham gia xác nhận các giao dịch trên block. Nói một cách đơn giản, Node phải gửi coin để xác minh danh tính của nó.
* **Ethereum Technology:**
* Ethereum (ETH) là một nền tảng điện toán có tính chất phân tán, công cộng, mã nguồn mở dựa trên công nghệ Blockchain. Nó có tính năng hợp đồng thông minh (kịch bản), tạo thuận lợi cho các thỏa thuận hợp đồng trực tuyến. Nền tảng này bao gồm một máy ảo hoàn toàn Turing - Ethereum Virtual Machine (EVM), có thể thực thi các kịch bản bằng cách sử dụng một mạng lưới máy tính Ethereum. Ethereum cũng cung cấp một loại tiền mã hóa gọi là "Ether", có thể được chuyển giữa các tài khoản và được sử dụng để trả công cho các thợ đào giúp thực hiện việc tính toán. "Gas" là một cơ chế giá giao dịch nội bộ, được sử dụng để giảm thiểu giao dịch rác (spam) và phân bổ các nguồn lực trên mạng lưới.
* Các điểm khác biệt cơ bản giữa Ethereum và Bitcoin:

|  |  |
| --- | --- |
| **Ethereum** | **Bitcoin** |
| Đồng tiền Ether của mạng lưới Ethereum không được thiết kế như một giải pháp thanh toán thay thế, mà là để thúc đẩy các lập trình viên và các tổ chức sáng tạo và vận hành các ứng dụng phi tập trung trong mạng Ethereum | Bitcoin có thể sử dụng để thanh toán hàng hóa và dịch vụ tại bất cứ nơi nào đồng tiền này được chấp nhận |
| Thời gian tạo khối Ethereum mới là 14 tới 15 giây | Thời gian tạo khối bitcoin mới là 10 phút |
| sử dụng giao thức GHOST giúp giao dịch Ether nhanh hơn Bitcoin. | Chậm hơn Ethereum |
| Ethereum không giới hạn số lượng ether. | Số lượng Bitcoin bị giới hạn ở mức 21 triệu với phần thưởng giảm còn một nửa sau mỗi 4 năm |
| Phí giao dịch của Ethereum được trả bằng Gas (quy đổi được ra ether), được tính dựa trên khối lượng tính toán, băng thông, lưu trữ | phí giao dịch Bitcoin bị cạnh tranh trực tiếp với nhau để vào được khối của Bitcoin mà bị giới hạn |
| Ethereum cho phép chạy mã Turing-complete, cho phép mọi tính toán được thực thi nếu có đủ khả năng tính toán và thời gian. Điều này mang lại nhiều rủi ro bị tấn công hơn cho Ethereum | Cấu trúc của Bitcoin đơn giản hơn nên ít bị tấn công hơn so với Ethereum |
| Có 13% số ether được bán cho lượng người đã tài trợ dự án ban đầu | những người đầu tiên đào Bitcoin nắm giữ số lượng lớn lượng Bitcoin đang phát hành |
| Ethereum chống lại việc sử dụng ASIC như Bitcoin. Người đào Ethereum phải sử dụng card đồ họa vì hàm băm của Ethereum yêu cầu sử dụng bộ nhớ. | Bitcoin sử dụng ASIC |
| Ethereum chống lại việc đào mỏ tập trung bằng cách sử dụng giao thức Ghost. |  |
| Ethereum đã phải chia nhánh sau khi DAO bị tấn công | Bitcoin đã có một lịch sử chưa bao giờ can thiệp vào dữ liệu trên sổ cái |

#### Tạo database

* Các EHR được chia thành 2 lớp gồm: thông tin nhà cung cấp (Provide Information) và thông tin bệnh nhân (Patient Information)
* Bác sĩ sẽ được truy vấn cho nhà cung cấp và nhà cung cấp trả về các câu trả lời của truy vấn đó.
* Khi một hồ sơ mới được xác nhận, nhà cung cấp có thể thêm một trình quản lý EHR, dữ liệu bệnh nhân sẽ được lưu trữ trong máy chủ dữ liệu cục bộ với Mạng đã được kết nối (network connectivity)
* Trình quản lý EHR là một giao diện người dùng API, nó tương tác với hệ thống cơ sở dữ liệu SQLite
* Nền tảng Ethereum có thể tạo ra cơ sở hạ tần để thực thi các hợp đồng thông minh, nó chứ các siêu dữ liệu (Metadata) quyền sở hữu và xem hồ sơ.
* Khi một hồ sơ mới được thêm vào hợp đồng đăng ký và được tạo một bản đăng ký mới (như ảnh trên), nó ánh xạ Mối quan hệ giữa nhà cung cấp và bệnh nhân (patient‐provider relationship contract - PPR) với các kiên kết Ethereum và địa chỉ của các hợp đồng thông minh và lưu trữ trong blockchain.
* Ethereum‐client cung cấp mạng ngang hàng được mã hóa bằng giao thức python, nó ánh xạ địa chỉ của hợp đông thông minh thông qua địa chỉ của hợp đồng đăng ký, và cập nhật hợp đồng thông minh cũng như trạng thái của PPR.
* Ethereum‐client tương tác với trình quản lý EHR và cung cấp thông báo cho khách hàng.

#### Sử dụng Ethereum Blockchain

Diagram

Description automatically generated

* Công nghệ Ethereum sử dụng phương thức chứng thực Bằng chứng cổ phần (Proof of Stake - POS) giao thức đồng thuận hữu ích cho tính toàn vẹn, tính chính xác và sự hợp tác giữa các bên liên quan để phân phối dữ liệu một cách phi tập trung và bảo mật.
* POS xác nhận kích thước của các khối và các chuỗi, nó tự động thực thi các chức năng mà được biết đến như là hợp đồng thông minh.
* Hợp đồng thông minh bổ sung độ chính xác cho dữ liệu và độ thích ứng với chuỗi Ethereum, công nghệ blockchain duy trì các EHR với việc sử dụng hợp đồng thông minh mà có thể được thực hiện bởi Ethereum.
* Nó là một nền tảng phi tập trung chạy trên hợp đồng thông minh.
* Hợp đồng đăng ký liên kết mỗi cái tên với địa chỉ Ethereum và số của hợp đồng thông minh và ánh xạ danh tính được khám phá này với hợp đồng thông minh.
* Hợp đồng thông minh duy trì một bản tóm tắt của PPR, nó quản lý tổng thể các EHR được liên kết với địa chỉ của PPR và trạng thái của mỗi hồ sơ.
* PPR xác định các hồ sơ và cung cấp công cụ để truy cấp các hồ sơ đã được nhận dạng này.
* PPR được liên kết với vài con trỏ, một con trỏ có thể được sử dụng để cung cấp thông tin từ cơ sở dữ liệu.
* Ngoài ra PPR còn có tính năng đặt biệt đó là liên kết với tên mạng máy chủ, số hiệu của cổng và cấu trúc mạng. Nó lưu trữ dữ liệu, cung cấp quyền riêng tư, chất lượng của truy cập và độ chính xác của dữ liệu.

#### Sử dụng điện toán đám mây

Diagram

Description automatically generated

* Ethereum client liên tục giám sát tất các các thực thể bên phía nhà giám sát, nó nhận dạng tất các các hoạt động của người dùng (RC, SC và PPR).
* Điện toán đám mây thu thập, truy xuất dữ liệu và cung cấp thông tin đến cơ sở dữ liệu vì nó có thể chuyển giao dữ liệu đến các thiết bị đeo được (ví dụ như đồng hồ thông minh, thời trang thông minh,…)
* Thư viện backend tương tác với tất cả các thành phần của nhà cung cấp bằng cách phân tích cú pháp giao thức Ethereum và cung cấp đầy đủ dữ liệu truyền tải đến hệ quản lý đám mây của Ethereum.
* Điện toán đám mây hỗ trợ cung cấp khả năng truy cập, quyền sở hữu, sự bảo vệ, sự riêng tư và phương pháp với sự phân cấp của phân tán được xử lý bởi sự đồng thuận của các thực thể đáng tin cậy.
* Blockchain riêng tư cho phép SC tiếp cận dựa trên token‐based; nó hoạt động ở phía nhà cung cấp, thư viện backend tương tác với công nghệ Blockchain và làm cho nó có thể chạy SC, nó liên quan đến nén, mã hóa, giải mã và khóa hệ thống quản lý (khóa giải phóng - key releasing, khóa nhận - key receiving, khóa kho - key storage).
* Điện toán đám mây hỗ trợ cung cấp dữ liệu minh bạch, mối quan hệ đáng tin cậy giữa các thực tể.
* Nhà cung cấp lưu trữ quy trinhg dữ liệu mà không cần sự tương tác của bên thức ba.

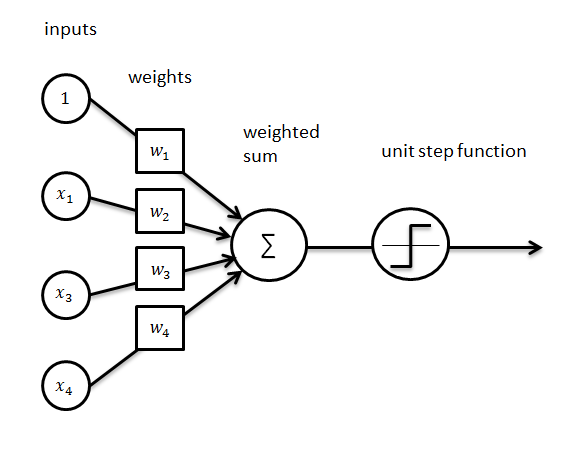
### Các mô hình sử dụng

#### Mạng nơ-ron nhân tạo

* Mạng nơ-ron nhân tạo (Artificial Neural Network – ANN) là một hệ thống tính toán lấy ý tưởng từ (nhưng không giống hoàn toàn) mạng lưới thần kinh sinh học của loài vật. Nó có khả năng "học" để thực hiện một số tác vụ nhất định dựa trên những ví dụ mà nó học được, thường là không cần lập trình trước hay đưa ra bất kì quy luật cụ thể nào cả. Ví dụ, trong bài toán phân loại ảnh, hệ thống này có khả năng xác định trong ảnh có con người hay không bằng cách phân tích những dữ liệu hình ảnh mà được gán nhãn là "có người", hay "không có người". ANN có thể tự học mà không cần có kiến thức gì về những đặc tính của một con người như gương mặt, vóc dáng, tay chân, hay mắt, mũi, miệng, v.v…



* Tương tự như một mạng thần kinh sinh học, một ANN là một mạng lưới nhiều nơ-ron nhân tạo. Mỗi một nơ-ron sẽ nhận đầu vào, thay đổi trạng thái bên trong, sau đó xuất đầu ra dựa vào giá trị đầu vào và hàm kích hoạt.
* Perceptron, một loại nơ-ron nhân tạo, là đơn vị nhỏ nhất thường được sử dụng trong các mô hình phân loại. Một perceptron nhận đầu vào. Mỗi đầu vào sẽ có một trọng số tương ứng. Đầu ra của perceptron được tính bằng tổng các tích đầu vào với trọng số. Sau đó kết quả sẽ được đưa qua một hàm kích hoạt, chẳng hạn như Sigmoid hay ReLU, thường là để chuẩn hóa số liệu và lọc bỏ những kết quả âm.



* Một số kiểu mạng nơ-ron:
* Có hai kiểu mạng nơ ron chính: mạng nơ ron truyền thằng (feedforward neural network) và mạng nơ ron hồi quy (recurrent neural network).
* Mạng truyền thẳng và hồi quy được minh họa như sau:

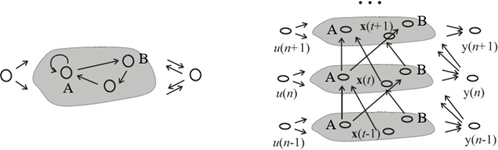
A picture containing text, watch

Description automatically generated

* Dễ thấy, ở mạng nơ ron truyền thẳng, các nơ ron trong tầng ẩn ***n+1*** đều được kết nối với các nơ ron trong tầng ***n***. Do có nhiều tầng ẩn nên chúng ta có thể thấy rằng mạng truyền thẳng kéo dài trong không gian, và là không có bất kỳ đường tuần hoàn (cyclic path) nào nằm trong mạng. Mạng nơ ron truyền thẳng rất phổ biến hiện nay.



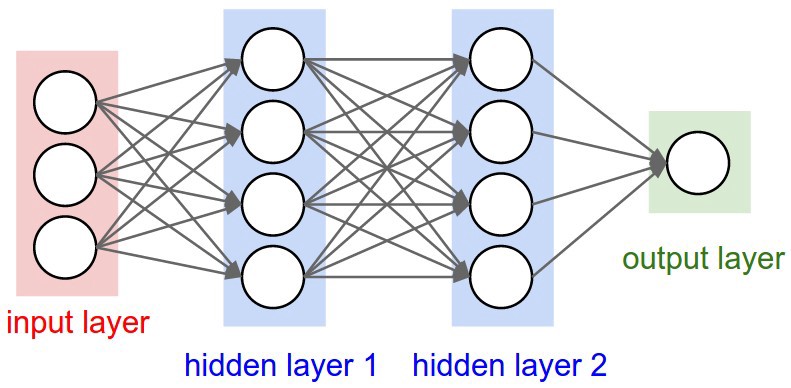
* Một loại khác là mạng nơ ron hồi quy. Không giống như mạng nơ ron truyền thẳng, mạng nơ-ron hồi quy có ít nhất một đường dẫn tuần hoàn. Chúng ta có thể thấy nó ở hình minh họa phía trên. Vì có một đường dẫn tuần hoàn, nên mạng nơ-ron hồi quy có thể gây ra vòng lặp vô cực. Tuy nhiên, mạng nơ ron tuần hoàn có một ứng dụng quan trọng là chúng có thể nhận diện cho các giai đoạn thời gian khác nhau, như hình minh họa sau:



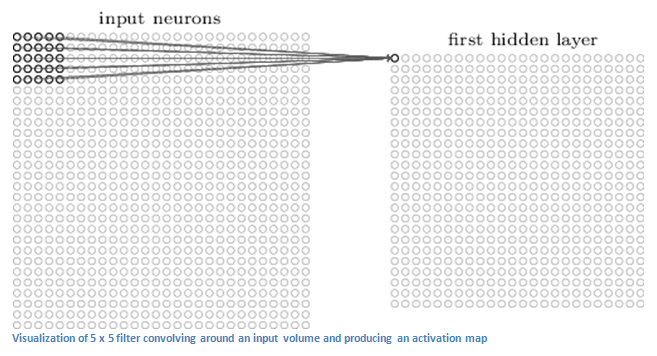
* Như ví dụ trên, có một nút A kết nối với nút B và một chu kỳ đến chính nút A. Mạng nơ ron hồi quy không xử lý đường dẫn tuần hoàn và các kết nối cùng một lúc. Mạng nơ ron hồi quy giả sử rằng đầu ra của nút A trong thời gian n là đầu vào của nút B và nút A trong thời gian n + 1. Vì vậy, ngoài tính chất kéo dài trong không gian khi kết nối với các tầng nơ ron tiếp theo, mạng nơ ron hồi quy cũng nằm sâu trong thời gian. Vì vậy, các mạng nơ ron hồi quy có thể mô hình hóa các hệ thống thay đổi theo bối cảnh. Ví dụ: mạng nơ ron hồi quy thường được sử dụng trong xử lý ngôn ngữ theo ngữ cảnh. Mạng nơ ron hồi quy có thể xử lý các phụ thuộc xa (long-term dependencies) theo mốc thời gian, ví dụ như mạng bộ nhớ dài-ngắn (Long Short Term Memory networks).

#### Tổng quan DNN

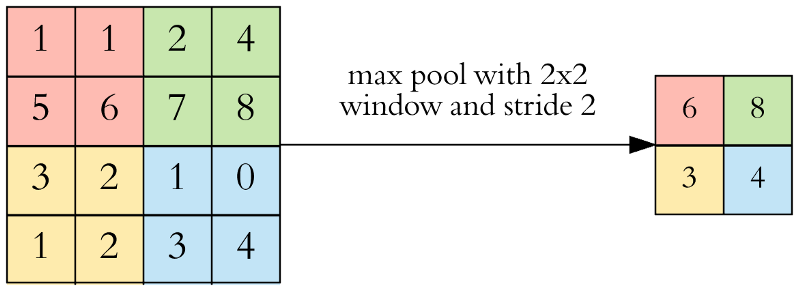
* Mạng học sâu (Deep Learning – DL) là một nhánh con của mạng nơ-ron nhân tạo. DL sử dụng nhiều lớp để lần lượt rút trích các đặc trưng từ thấp đến cao của đầu vào. Ví dụ, trong bài toán xử lí ảnh, những lớp thấp sẽ tập trung vào việc phát hiện góc cạnh của ảnh, trong khi những lớp cao hơn sẽ nhận diện được các đặc trưng thân thuộc với con người hơn như mắt, mũi, chữ cái, v.v…
* Các mô hình DL thường có 3 lớp:
* Lớp vào (input layer): là lớp chứa các đơn vị tương ứng với đầu vào của mạng.
* Lớp ra (output layer): là lớp chứa các đơn vị tương ứng với đầu ra của mạng.
* Lớp ẩn (hidden layer): là lớp nằm ở giữa lớp vào và lớp ra, chứa các đơn vị thể hiện cấu trúc bên trong cũng như cách tính toán của mạng. Lớp ẩn có thể có nhiều hơn một lớp.



* Thông thường, giữa các lớp sẽ được nối đầy đủ các trọng số với nhau. Tuy nhiên, ở một số biến thể tùy vào bài toán, không nhất thiết phải nối đầy đủ giữa tất cả các lớp. Ngoài ra, thực nghiệm cho thấy, việc tăng số lượng các lớp ẩn sẽ giúp cho mô hình có khả năng rút trích nhiều đặc trưng cao cấp hơn, từ đó dễ dàng "học" bộ dữ liệu. Ngược lại, tăng số lượng các perceptron của một lớp sẽ làm cho mô hình có khả năng phân biệt nhiều biến thể của một loại đặc trưng hơn (chẳng hạn như mắt đen, mắt xanh, mắt tròn, v.v…). Tuy nhiên, ở cả hai trường hợp đều có khả năng dẫn đến hiện tượng overfitting (hiện tượng mà khi mô hình "học quá nhiều" một bộ dữ liệu), khiến cho mô hình chỉ hoạt động tốt đúng trên một bộ dữ liệu chứ không thể hoạt động tốt trên bộ dữ liệu khác. Việc tinh chỉnh sao cho số lớp ẩn và số perceptron thích hợp vẫn còn là một câu hỏi lớn.
* Các lớp cơ bản trong một mạng nơ-ron:
* Lớp kết nối đầy đủ: là một dạng lớp trong kiến trúc mạng DL. Trong đó, từng đơn vị của nó được nối (trọng số) với từng đơn vị của lớp trước nó và sau nó (nếu có). Việc nối tiếp tất cả các đơn vị sẽ cho mô hình mạng có thêm tham số để "học" một bộ dữ liệu, từ đó tăng khả năng dự đoán cũng như độ chính xác. Tuy nhiên, điều này có thể dẫn tới hiện tượng overfitting nhiều hơn và tốn nhiều tài nguyên hơn về cả phần cứng lẫn thời gian tính toán. Thông thường, đối với bài toán phân loại, FC Layer nằm ở cuối mạng DL và được dùng để phân loại kết quả.
* Lớp tích chập: là một dạng lớp trong kiến trúc mạng DL. Trong đó, thay vì từng đơn vị được nối đầy đủ với các đơn vị ở lớp trước đó như FC Layer, nó sử dụng một bộ lọc (filter, hay còn gọi là kernel). Một filter sẽ định ra số lượng nơ-ron mà nó sẽ nối, thường là với số lượng đơn vị nhỏ hơn lớp input trước đó. Filter này vẫn có trọng số như một perceptron bình thường và output của nó là một đơn vị tương ứng.



* Lớp Max-Pooling: là một dạng lớp trong kiến trúc mạng DL. Ý tưởng của max-pool layer rất đơn giản, nó làm giảm số lượng tham số mà mạng DL phải học bằng cách sử dụng giá trị lớn nhất tại một vùng nơ-ron đang xét làm đại diện cho vùng đó, và bỏ qua các giá trị nhỏ hơn.



* Max-pool Layer cũng hoạt động giống với Conv Layer, cũng có cửa sổ trượt (kích thước) và stride (số nơ-ron nhảy mỗi lần trượt). Thông thường, Max-Pool Layer sẽ đi kèm phía sau một Conv Layer. Sở dĩ như vậy là bởi vì, thông thường các feature map chỉ có giá trị lớn tại một số vùng nhất định, trong khi đa số các vùng còn lại có giá trị rất nhỏ và không đáng kể. Việc tiếp tục sử dụng feature map đó làm đầu vào cho các lớp phía sau khá là dư thừa. Mặt khác, ta còn có những rào cản về mặt tài nguyên. Do đó, việc áp dụng max-pool layer vào nhằm tập trung vào các con số có giá trị lớn và giảm số lượng tham số cần huấn luyện là quan trọng và có ý nghĩa.

#### Áp dụng DNN trong lĩnh vực chăm sóc sức khoẻ

* Áp dụng DNN trong lĩnh vực chăm sóc sức khoẻ:
* Trong DNN, dữ liệu sẽ được sắp xếp thành đưa về ma trận, mỗi đặc trưng là một cột biểu diễn thành:

Diagram, schematic

Description automatically generated

* Ma trận đầu vào:

Diagram

Description automatically generated

* Độ rộng của ma trận:

Diagram

Description automatically generated with medium confidence

* Ma trận Bias được tạo trong mỗi lớp:

A picture containing text, device

Description automatically generated

* Ma trận tính toán sau mỗi lớp:

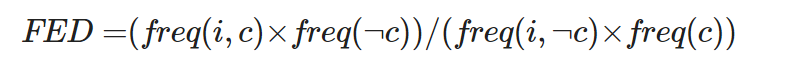


* Những ma trận kết quả của của mỗi layer sẽ là đầu vào của các layer tiếp theo.
* DNN hỗ trợ phân tích chăm sóc sức khỏe để tạo ra dữ liệu được trích xuất. Dữ liệu được trích xuất nổi bật đến từ dữ liệu được training mang lại sự tự giám sát và thực hiện các biện pháp phòng ngừa cho kế hoạch đặc trưng, ​​mọi người đang bị các bệnh lớn như bệnh tim, ung thư máu để phân loại dữ liệu và dự đoán dữ liệu chăm sóc sức khỏe.



### Dữ liệu trích xuất đặc trưng

* Học sâu cung cấp học thứ bậc để tạo FED, dữ liệu được lọc từ một số layer như Hidden Layer, mỗi đầu ra lớp này là đầu vào của lớp kia. Kết quả của mô hình DNN có thể chính xác hơn. Mô hình học cơ bản có thể được bắt đầu từ dữ liệu được thu thập, dữ liệu có thể được phân loại là mã thông báo, tập hợp các mã thông báo có thể được coi là **T = {, , , …., }** trong đó mỗi mã thông báo có thể liên kết với dữ liệusau đó nó tương ứng với = + mỗi kết quả lớp liên tiếp được xác định lại và gửi đến lớp liên tiếp để cung cấp dữ liệu chính xác và tạo mối tương quan. Nó tạo ra một tập hợp các mã thông báo, chính xác từ một lớp sang số lớp khác của các phiên bản **i =**  và mỗi lớp bao gồm một tập hợp các token **C =** , số lượng mã thông báo xảy ra ở 1 lớp cụ thể là là ***freq***(***i, c)***, số lượng mã thông báo không xảy ra ở lớp đó ***freq(i, ¬c)***, không có mã thông báo nào xảy ra là ***freq (¬c)***, chỉ chứa mã thông báo tại một lớp đó ***freq(c)****.*



* Trong DNN, mỗi lớp có thể được gán ***fred*** của lớp đó và dữ liệu có thể được chuyển qua nhiều lần để đào tạo dữ liệu và tinh chỉnh. Nhiều lớp có thể phục vụ các kết quả tinh tế cho lớp tiếp theo. Việc tính toán của FED như sau:

Text, letter

Description automatically generated

* DNN bao gồm một chiến lược đa lớp để cung cấp đầu ra hiệu quả và mong muốn để hoàn thành việc phân loại các nhiệm vụ ở mỗi lớp. Nếu bất kỳ lớp nào xác định các bất thường trong dữ liệu y tế, hình ảnh y tế, báo cáo y tế và nó áp dụng xác suất để tạo Fed theo phương trình sau:

Diagram

Description automatically generated with low confidence

# Phần 4: THỰC NGHIỆM

## Tổng quan bộ dữ liệu và cách tổ chức thực nghiệm

* Tên bộ dữ liệu: COVID-19 in India
* Nguồn: <https://www.kaggle.com/datasets/sudalairajkumar/covid19-in-india>
* Tác giả: SRK và Devakumar K. P.
* Bộ dữ liệu gồm 3 bộ dataset: StatewiseTestingDetails, Covid\_19\_india và covid\_vaccine\_statewise.
* Bộ dữ liệu về số ca nhiễm covid, dữ liệu thử nghiệm và dữ liệu tiêm chủng ở các khu vực của Ấn Độ. Dữ liệu cấp tiểu Bang được lấy từ Ministry of Health & Family Welfare, dữ liệu thử nghiệm và tiêm chủng được thu thập từ covid19india.org

### Chi tiết bộ dữ liệu

* Trong đó:
* Bộ dữ liệu covid\_19\_india thể hiện số ca nhiễm covid theo cấp độ hàng ngày Ấn Độ, được thu thập từ 30/1/2020 đến 11/8/2021. Bộ dữ liệu gồm 18110 hàng và 9 cột với các thuộc tính:
* Sno: mã số series
* Date: ngày quan sát
* Time: thời gian quan sát
* State/UnionTerritory: lãnh thỗ/ Ban được quan sát
* ConfirmedIndianNational: tích lũy số lượng người dân Ấn Độ được xác nhận
* ConfirmedForeignNational: tích lũy số lượng công dân nước ngoài được xác nhận
* Cured: tích lũy số trường hợp được chữa khỏi
* Deaths: tích lũy số trường hợp đã tử vong
* Confirmed: số trường hợp được xác nhận
* Bộ dữ liệu StatewiseTestingDetails cho thấy các chi tiết về các mẫu thử nghiệm covid ở các tiểu bang, được thu thập từ 17/4/2020 đến 10/8/2021. Bộ dữ liệu gồm 16336 hàng và 5 cột với các thuộc tính:
* Date: Ngày quan sát
* State: Tiểu Bang quan sát
* TotalSamples: Tích lũy số lượng mẫu thử nghiệm cho đến ngày được quan sát
* Negative: tích lũy số lượng mẫu âm tính cho đến ngày được quan sát
* Positive: tích lũy số lượng mẫu dương tính cho đến ngày được quan sát
* Bộ dữ liệu covid\_vaccine\_statewise tổng hợp về tình trạng vắc xin ở Ấn Độ, được thu thập từ 16/1/2020 đến 15/8/2021. Bộ dữ liệu gồm 7845 hàng và 24 cột với các thuộc tính:
* Updated On: ngày cập nhật
* State: Tiểu Bang
* Total Doses Administered: tổng số liều tiềm tàng
* Sessions: Phiên làm việc
* Sites: địa điểm
* First Dose Administered: lượng tiêm đầu tiên (Liều quy định)
* Second Dose Administered: Lượng tiêm thứ hai (Liều quy định)
* Male (Doses Administered): Số lượng đàn ông (Liều quy định)
* Female (Doses Administered): Số lượng phụ nữ (Liều quy định)
* Transgender (Doses Administered): Số lượng người chuyển giới (Liều quy định)
* Covaxin (Doses Administered): lượng vắc xin Covaxin (Liều quy định)
* CoviShield (Doses Administered): lượng vác xin CoviShield (Liều quy định)
* Sputnik V (Doses Administered): Lượng vắc xin Sputnik V (liều quy định)
* AEFI: Số lượng ca gặp bất lợi sau khi tiêm chủng
* 18-44 Years (Doses Administered): Số lượng từ 18-44 tuổi (Liều quy định)
* 45-60 Years (Doses Administered): Số lượng từ 45-60 tuổi (Liều quy định)
* 60+ Years (Doses Administered): Số lượng trên 60 tuổi (Liều quy định)
* 18-44 Years (Individuals Vaccinated): Số lượng từ 18-44 tuổi đã tiêm chủng
* 45-60 Years (Individuals Vaccinated): Số lượng từ 45-60 tuổi đã tiêm chủng
* 60+ Years (Individuals Vaccinated): Số lượng trên 60 tuổi đã tiêm chủng
* Male (Individuals Vaccinated): Số lượng đàn ông đã tiêm chủng
* Female (Individuals Vaccinated): Số lượng phụ nữ đã tiêm chủng
* Transgender (Individuals Vaccinated): Số lượng người chuyển giới đã tiêm chủng
* Total Individuals Vaccinated: Tổng số lượng ca đã tiêm chủng
* Bộ dữ liệu về các bệnh như bệnh tim, bệnh ung thư máu, bệnh tiêu hóa, bệnh thương hàn cũng tương tự như bộ dữ liệu Covid 19.
* Bệnh về tim:
* Tên bộ dữ liệu: Heart Diseases
* Link: <https://www.kaggle.com/datasets/mdshamimrahman/heart-diseases>
* Tác giả: MD. SHAMIM RAHMAN
* Bộ dữ liệu thống kê về bệnh tim và các chỉ số liên quan.
* Bộ dữ liệu có 16 cột và 4240 hàng với các thuộc tính: male, age, education, currentSmoker, cigsPerDay, BPMeds, prevalentStroke, prevalentHyp, diabetes, totChol, sysBP, diaBP, BMI, heartRate, glucose, TenYearCHD.
* Bộ dữ liệu thống kê chi tiết về bệnh sốt thương hàn trên từng khu vực và thời gian.
* Tên bộ dữ liệu: US.387754006
* Link: <https://www.tycho.pitt.edu/dataset/US.387754006/>
* Bộ dữ liệu gồm 11709 hàng và 21 cột với các thuộc tính: ConditionName, ConditionSNOMED, PathogenName, PathogenTaxonID, Fatalities, CountryName, CountryISO, Admin1Name, Admin1ISO, Admin2Name, CityName, PeriodStartDate, PeriodEndDate, PartOfCumulativeCountSeries, AgeRange, Subpopulation, PlaceOfAcquisition, DiagnosisCertainty, SourceName, CountValue

### Thực nghiệm

#### Cách tổ chức thực nghiệm

* Để phát triển mô hình đề xuất, mô hình này đã chuẩn bị một bảng câu hỏi với sự trợ giúp của dữ liệu được thu thập từ những người dùng khác nhau trong khu vực NCR Delhi (Ấn Độ), sau đó nhóm tác giả chuyển đổi tệp giá trị về dạng CSV theo quy trình:
* B1: Thu thập dữ liệu từ Google forms
* B2: Cài đặt các thư viện bằng Python
* B3: Sử dụng các thuật toán đề xuất khuyến nghị và truy vấn cơ sở dữ liệu
* B4: Trích xuất đặc trưng (Feature extraction) của dữ liệu thông qua Python và MATALB R2016b
* Huấn luyện: Mạng DNN được huấn luyện dựa trên token‐based learning rate, với từng loại bệnh sẽ có các số lượng tokens khác nhau
* Bệnh tim mạch: Có 6 tokens được sử dụng, với 4 lớp:
* Lớp 0 cho người bình thường
* Lớp 1 cho lần đột quỵ đầu tiên
* Lớp 2 cho lần đột quỵ thứ hai
* Lớp 3 cho ca tử vong

=> Cho ra 24 trường hợp.

* Output layer tạo ra dữ liệu huấn luyện để đạt được trạng thái mục tiêu (được cố định với FED là 1). Input layer bắt đầu với 12% số người bị ảnh hưởng bởi bệnh tim, hiệu suất để tạo FED được hiển thị trong Hình 11:

Chart, line chart

Description automatically generated

* Bệnh ung thư máu: Có 8 tokens được sử dụng, với 4 lớp:
* Lớp 0 cho người bình thường
* Lớp 1 cho giai đoạn đầu
* Lớp 2 cho giai đoạn giữa
* Lớp 3 cho giai đoạn cuối

=> Cho ra 32 trường hợp.

* Output layer tạo ra dữ liệu huấn luyện để đạt được trạng thái mục tiêu (được cố định với FED là 1.2). Input layer bắt đầu với 12% số người bị ảnh hưởng bởi bệnh ung thư máu, hiệu suất để tạo FED được hiển thị trong Hình 12:

Chart, line chart

Description automatically generated

* Bệnh đường tiêu hóa: Có 4 tokens được sử dụng, với 4 lớp:
* Lớp 0 cho trẻ em
* Lớp 1 cho thanh thiếu niên
* Lớp 2 cho người lớn
* Lớp 3 cho người già

=> Cho ra 16 trường hợp.

* Output layer tạo ra dữ liệu huấn luyện để đạt được trạng thái mục tiêu (được cố định với FED là 1.2). Input layer bắt đầu với 10% số người bị ảnh hưởng bởi bệnh đường tiêu hóa, hiệu suất để tạo FED được hiển thị trong Hình 13:

Chart, line chart

Description automatically generated

* Bệnh thương hàn: Có 4 tokens được sử dụng, với 4 lớp:
* Lớp 0 cho giai đoạn 1
* Lớp 1 cho giai đoạn 2
* Lớp 2 cho giai đoạn 3
* Lớp 3 cho giai đoạn 4

=> Cho ra 16 trường hợp.

* Output layer tạo ra dữ liệu huấn luyện để đạt được trạng thái mục tiêu (được cố định với FED là 0.978). Input layer bắt đầu với 10% số người bị ảnh hưởng bởi bệnh thương hàn, hiệu suất để tạo FED được hiển thị trong Hình 14:

Chart, line chart

Description automatically generated

* Bệnh Covid-19: Có 1 tokens được sử dụng, với 4 lớp:
* Lớp 0 cho giai đoạn 1
* Lớp 1 cho giai đoạn 2
* Lớp 2 cho giai đoạn 3
* Lớp 3 cho giai đoạn 4

=> Cho ra 4 trường hợp, đó là số ca dương tính, số ca tiêu cực, số người chết và số người trong diện cách ly.

* Output layer tạo ra dữ liệu huấn luyện để đạt được trạng thái mục tiêu (được cố định với FED là 0.978). Input layer bắt đầu với 13% số người bị ảnh hưởng bởi bệnh Covid-19, hiệu suất để tạo FED được hiển thị trong Hình 15:

Chart, line chart

Description automatically generated

#### Phương pháp thực nghiệm

* Thuật toán 1: Dữ liệu được huấn luyện bằng mô hình DNN dưới sự đánh giá của blockchain
* Input: Tập dữ liệu token: trong không gian trạng thái.
* Output: Bộ phân loại CT, Bộ dữ liệu Blockchain B.
* Bước 1: Khởi tạo tập token CD = ϕ∀D∈ {t1, t2, t3, …… tk}.
* Bước 2: Vòng lặp while (1) do
* Bước 3: CT ← Train (B)
* Bước 4: khởi tạo tập huấn luyện hidden layer: HD = ϕ∀D∈ {t1, t2, t3, …… tk}.
* Bước 5: P (i, c) ← ∀D∈ {t1, t2, t3, …… tk} dựa vào xác suất để phân loại các trường hợp cho lớp C.
* Bước 6: Với mỗi CT ∈ B tiến hành (do):
* Bước 7: Ar ← độ chính xác (P, (i, c))
* Bước 8: If P (i, c)> CT then:
* Bước 9: CD = ϕ∀D∈ {t1, t2, t3, …… tk}
* Bước 10: Else
* Bước 11: HD = ϕ∀D∈ {t1, t2, t3,, …… tk}
* Bước 12: chạy vòng lặp for CD = ϕ∀D∈ {t1, t2, t3, …… tk} CD = ϕ∀D∈ {t1, t2, t3, …… tk}
* Bước 13: CD = ϕ∀D∈ { t1, t2, t3, …… tk }
* Bước 14: Trả về B, CD.
* Thuật toán 2: Tạo dữ liệu trích xuất đặc trưng FED
* Input: Mỗi cá thể mã thông báo tt∈T, tập các lớp .
* Output: FED.
* Bước 1: Khởi tạo.
* Bước 2: for i = 0 to n do

Bươc 3: Trả về:

#### Kết quả thực nghiệm

* Để thu thập các mẫu dữ liệu đã thử nghiệm khác nhau, dữ liệu thử nghiệm đưa ra được tính toán dưới dạng mạng và được huấn luyện khác nhau. Kết quả đầu ra của mạng đã được tính toán theo trọng số đã điều chỉnh. Kết quả thu được được lập bảng như sau:

Table

Description automatically generated