

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ
KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG



NIÊN LUẬN CƠ SỞ NGÀNH
NGÀNH KHOA HỌC MÁY TÍNH

Đề tài

**DỰ ĐOÁN BIỂU HIỆN TẾ BÀO B COVID-19/SARS
PHÂN TÍCH ĐẠI DỊCH COVID-19 – SARS-CoV-2**

Sinh viên thực hiện: Nguyễn Hữu Tính
Mã số: B1710355
Khóa: 43

Cần Thơ, 1/2021

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ
KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG



NIÊN LUẬN CƠ SỞ NGÀNH
NGÀNH KHOA HỌC MÁY TÍNH

Đề tài

**DỰ ĐOÁN BIỂU HIỆN TẾ BÀO B COVID-19/SARS
PHÂN TÍCH ĐẠI DỊCH COVID-19 – SARS-CoV-2**

Giáo viên hướng dẫn:
TS. Trần Nguyễn Dương Chi

Sinh viên thực hiện:
Nguyễn Hữu Tính
Mã số: B1710355
Khóa: 43

Cần Thơ, 1/2021

NHẬN XÉT CỦA GIẢNG VIÊN

Cần Thơ, ngày tháng năm

(GVHD ký và ghi rõ họ tên)

LỜI CẢM ƠN

Để có được bài niêm luận này, em xin được bày tỏ lòng biết ơn chân thành và sâu sắc đến Cô Trần Nguyễn Dương Chi – người đã trực tiếp tận tình hướng dẫn, giúp đỡ em. Trong suốt quá trình thực hiện niêm luận, nhờ những sự chỉ bảo và hướng dẫn quý giá đó mà bài niêm luận này được hoàn thành một cách tốt nhất.

Em cũng xin gửi lời cảm ơn chân thành đến các Thầy Cô Giảng viên Đại học Cần Thơ, đặc biệt là các Thầy Cô ở Khoa CNTT & TT, những người đã truyền đạt những kiến thức quý báu trong thời gian qua.

Em cũng xin chân thành cảm ơn bạn bè cùng với gia đình đã luôn động viên, khích lệ và tạo điều kiện giúp đỡ trong suốt quá trình thực hiện để em có thể hoàn thành bài niêm luận một cách tốt nhất.

Tuy có nhiều cố gắng trong quá trình thực hiện niêm luận, nhưng không thể tránh khỏi những sai sót. Em rất mong nhận được sự đóng góp ý kiến quý báu của quý Thầy Cô và các bạn để bài niêm luận hoàn thiện hơn.

Cần Thơ, ngày tháng 1 năm 2021

Người viết

Nguyễn Hữu Tính

MỤC LỤC

MỤC LỤC	1
DANH MỤC HÌNH	3
DANH MỤC BẢNG.....	6
ABSTRACT.....	7
TÓM TẮT.....	8
PHẦN GIỚI THIỆU.....	9
1. Đặt vấn đề	9
2. Lịch sử giải quyết vấn đề	10
3. Mục tiêu đề tài	10
4. Đối tượng và phạm vi nghiên cứu	11
5. Phương pháp nghiên cứu	11
6. Bố cục đề tài	11
PHẦN NỘI DUNG.....	12
CHƯƠNG 1.....	12
MÔ TẢ BÀI TOÁN	12
1. Mô tả chi tiết bài toán	12
2. Vấn đề và giải pháp liên quan đến bài toán	24
CHƯƠNG 2.....	26
THIẾT KẾ VÀ CÀI ĐẶT	26
GIAI ĐOẠN 1: DỰ ĐOÁN BIỂU HIỆN TÉ BÀO B-CELL/SARS	26
A. NHIỆM VỤ 1: DỰ ĐOÁN SARS VỚI DỮ LIỆU B-CELL.....	26
I. Yêu cầu của nhiệm vụ:	26
II. Ý tưởng giải quyết vấn đề:	26
III. Tập dữ liệu:.....	27
1. Sơ lược về tập dữ liệu:.....	27
2. Tiền xử lý tập dữ liệu:	27
IV. Nội dung:	27
B. NHIỆM VỤ 2: DỰ ĐOÁN COVID-19 VỚI DỮ LIỆU B-CELL VÀ SARS	33
I. Yêu cầu của nhiệm vụ:	33
II. Ý tưởng giải quyết vấn đề:	33
III. Sơ lược về các tập dữ liệu:	33
IV. Nội dung:	33

GIAI ĐOẠN 2: PHÂN TÍCH ĐẠI DỊCH COVID-19	34
C. NHIỆM VỤ 3: NHỮNG GÌ ĐÃ BIẾT VỀ SỰ LÂY TRUYỀN, Ủ BỆNH VÀ ÔN ĐỊNH CỦA COVID-19 ĐỐI VỚI MÔI TRƯỜNG	34
I. Yêu cầu của nhiệm vụ:	34
II. Ý tưởng giải quyết vấn đề:	34
III. Sơ lược về các tập dữ liệu:	4
IV. Nội dung:	6
D. NHIỆM VỤ 4: SỐ TRƯỜNG HỢP NHIỄM BỆNH, TỔNG SỐ NGƯỜI CHẾT VÀ ĐẠI DỊCH PHÁT TRIỂN Ở CÁC QUỐC GIA TRÊN THẾ GIỚI	24
I. Yêu cầu của nhiệm vụ:	24
II. Sơ lược về các tập dữ liệu:	24
III. Nội dung:	24
E. NHIỆM VỤ 5: TẠO BẢNG NGHIÊN CỨU LIÊN QUAN ĐẾN COVID-19 VÀ THỐNG KÊ CÂU HỎI LIÊN QUAN ĐẾN BỆNH NHÂN	35
I. Yêu cầu của nhiệm vụ:	35
II. Sơ lược về các tập dữ liệu:	35
III. Nội dung:	36
PHẦN KẾT LUẬN	51
1. Kết quả đạt được:	51
2. Những hạn chế:	51
3. Hướng phát triển:	52
TÀI LIỆU THAM KHẢO	53

DANH MỤC HÌNH

Hình 1. Thuộc tính start_position	13
Hình 2. Thuộc tính end_position	13
Hình 3. Thuộc tính chou_fasman.....	14
Hình 4. Thuộc tính emini	14
Hình 5. Thuộc tính kolaskar_tongaonkar	15
Hình 6. Thuộc tính parker.....	16
Hình 7. Thuộc tính isoelectric_point	17
Hình 8. Thuộc tính aromacity	17
Hình 9. Thuộc tính hydrophobicity.....	18
Hình 10. Thuộc tính stability	18
Hình 11. Giá trị nhăn	19
Hình 12. Tập dữ liệu Covid_19_data.csv	19
Hình 13. Tập dữ liệu PatienInfo.csv	21
Hình 14. Tập dữ liệu Case.csv	22
Hình 15. Tập dữ liệu COVID 19 Containment measures data.csv	22
Hình 16. Tập dữ liệu covid_19_data.csv sau khi tiền xử lý	23
Hình 17. Tập dữ liệu time_series_covid_19_confirmed.csv	24
Hình 18. Tập dữ liệu metadata.csv	24
Hình 19. Hình ảnh tập dữ liệu B-cell	27
Hình 20. Tập dữ liệu B-cell sau khi tiền xử lý.....	27
Hình 21. Biểu đồ y sinh thuộc tính quan trọng.....	28
Hình 22. Biểu đồ các thuộc tính quan trọng	29
Hình 23. Biểu đồ dữ liệu gen.....	29
Hình 24. Biểu đồ phân chia tập dữ liệu train_test	30
Hình 25. Tập dữ liệu Covid.csv	32
Hình 26. Tập dữ liệu time_series_covid_19_confirmed.csv sau tiền xử lý.....	5
Hình 27. Gom dữ liệu theo ngày.....	5
Hình 28. Tập dữ liệu time_series_covid_19_confirmed.csv sau khi	5
Hình 29. Sự lan rộng của covid 19 trên toàn thế giới	6
Hình 30. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước Asian	6
Hình 31. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước	7
Hình 32. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước	7
Hình 33. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước	8
Hình 34. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước Europe	8
Hình 35. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở khu vực Africa.....	9
Hình 36. Tốc độ tử vong, xác nhận và bình phục của các nước	10
Hình 37. Beta, gamma, delta ở các nước	10
Hình 38. Tốc độ truyền, phục hồi và tử vong ở các nước Asian	11
Hình 39. Tốc độ truyền, phục hồi và tử vong ở các nước phía bắc America	11
Hình 40. Tốc độ truyền, phục hồi và tử vong ở các nước Oceania	11
Hình 41. Tốc độ truyền, phục hồi và tử vong ở các nước phía nam America	12
Hình 42. Tốc độ truyền, phục hồi và tử vong ở các nước Europe	12
Hình 43. Tốc độ truyền, phục hồi và tử vong ở các nước Africa	12
Hình 44. Diễn giải kiểm soát dịch bệnh ở lục địa Oceania.....	13
Hình 45. Tốc độ lan truyền, tử vong và phục hồi ở lục địa Oceania	13
Hình 46. Chuyển động của R0 theo thời gian ở Oceania	14

Hình 47. Diễn giải kiểm soát dịch bệnh ở Trung Quốc	14
Hình 48. Tốc độ lan truyền, tử vong và phục hồi ở Trung Quốc.....	15
Hình 49. Chuyển động của R0 theo thời gian ở Trung Quốc	15
Hình 50. Các trường hợp xác nhận nhiều nhất ở Hàn Quốc	16
Hình 51. Các trường hợp xác nhận ít nhất ở Hàn Quốc	16
Hình 52. Các trường hợp nhiễm trùng nhiều nhất ở Hàn Quốc.....	17
Hình 53. Ảnh hưởng của covid đối với giới tính	18
Hình 54. Nhóm tuổi dễ nhiễm bệnh.....	18
Hình 55. Nhóm tuổi dễ tử vong	19
Hình 56. Nhóm tuổi có khả năng dễ phục hồi	19
Hình 57. Tương quan giữa tình trạng bệnh và nhóm tuổi.....	20
Hình 58. Các biện pháp ngăn chặn và giảm thiểu của Trung Quốc	20
Hình 59. Các biện pháp ngăn chặn và giảm thiểu của Trung Quốc sau dịch bệnh	22
Hình 60. Tình hình dịch bệnh ở Trung Quốc giai đoạn đầu	22
Hình 61. Hiệu quả của việc phòng chống đối với dịch bệnh ở Trung Quốc	22
Hình 62. Tỷ lệ tử vong, phục hồi và xác nhận ở Trung Quốc sau khi áp dụng lối ngăn chặn	23
Hình 63. Các trường hợp nhiễm bệnh từ ngày 22/1/2020	25
Hình 64. Các trường hợp nhiễm bệnh ở Ấn Độ.....	25
Hình 65. Top 5 quốc gia có số ca đã xác nhận cao nhất.....	26
Hình 66. Top 5 quốc gia có số ca tử vong cao nhất.....	26
Hình 67. Top 5 quốc gia có số ca phục hồi cao nhất	26
Hình 68. Top 10 quốc gia có số ca được xác nhận trung bình	27
Hình 69. Top 10 quốc gia có số ca phục hồi trung bình	27
Hình 70. Top 10 quốc gia có số ca tử vong trung bình.....	28
Hình 71. Phân phối các trường hợp đã xác nhận	28
Hình 72. Phân phối các triệu chứng của covid 19	29
Hình 73. Tốc độ đại dịch covid phát triển trên thế giới.....	30
Hình 74. Tỷ lệ phục hồi và tử vong của Trung Quốc	30
Hình 75. So sánh với Trung Quốc	31
Hình 76. Sự tương quan giữa thời gian kể từ trường hợp đầu tiên và sự phục hồi.....	31
Hình 77. Trường hợp không đồng đều.....	32
Hình 78. Số trường hợp nhiễm bệnh theo thời gian	33
Hình 79. Xu hướng nhiễm bệnh theo thời gian	33
Hình 80. Tập dữ liệu metadata.csv sau khi tiền xử lý dữ liệu	36
Hình 81. Các chuẩn đoán về virus corona	37
Hình 82. Tìm kiếm thông qua tóm tắt với từ CORONA cho CÔNG CỤ (diagnostic tolos)	38
Hình 83. Tìm kiếm thông tin qua từ kiểm tra (test)	39
Hình 84. Tìm kiếm thông tin qua từ COVID	39
Hình 85. Tìm kiếm thông tin qua từ FDA	40
Hình 86. Tìm kiếm thông tin qua từ Viral Test	41
Hình 87. Các thuật ngữ trong 111 bài viết ABSTRACTS	42
Hình 88. Thời gian virus phát tán sau khi khởi phát.....	42
Hình 89. Thời gian ủ bệnh ở độ tuổi khác nhau	43
Hình 90. Thời gian ủ bệnh ở nhóm tuổi khác nhau	43
Hình 91. Hình E/IV/3: Thời kỳ ủ bệnh của virus	44
Hình 92. Tỷ lệ bệnh nhân không triệu chứng	45

Hình 93. Bệnh nhi không có triệu chứng 1	45
Hình 94. Bệnh nhi không có triệu chứng 2	46
Hình 95. Lây truyền không có triệu chứng trong thời gian ủ bệnh.....	47
Hình 96. Lịch sử tự nhiên của virus từ người nhiễm bệnh	48
Hình 97. Thời gian rụng trung bình của virus	49
Hình 98. Khoảng thời gian dài nhất của virus	50

DANH MỤC BẢNG

Bảng 1. Độ chính xác của các giải thuật.....	30
Bảng 2. Độ chính xác của hai giải thuật sau các lần lặp.....	31
Bảng 3. So sánh và kết luận hóa trị kháng sinh	32
Bảng 4. Kết luận về hóa trị kháng sinh.....	33
Bảng 5. Bảng số trường hợp dương tính.....	9

ABSTRACT

TÓM TẮT

PHẦN GIỚI THIỆU

1. Đặt vấn đề

Do sự lây lan của COVID-19, việc phát triển vắc-xin đang được tiến hành càng sớm càng tốt. Với tầm quan trọng của phân tích dữ liệu trong quá trình phát triển vắc xin, tuy không có nhiều bộ dữ liệu phân tích đơn giản để có thể xử lý. Đề tài đã xuất bản bộ dữ liệu và mã mẫu cho dự đoán Bcell epitope, một trong những chủ đề nghiên cứu quan trọng trong việc phát triển vắc xin. Bộ dữ liệu này được phát triển trong quá trình nghiên cứu và được lấy từ IEDB và UniProt . Đề tài mô tả ngắn gọn các dự đoán của epitope té bào B được đề cập trong tập dữ liệu. Té bào B tạo ra đáp ứng miễn dịch đặc hiệu với kháng nguyên in vivo tạo ra một lượng lớn kháng thể đặc hiệu với kháng nguyên bằng cách nhận biết các tiểu vùng (vùng biểu mô) của protein kháng nguyên. Chúng có thể ức chế hoạt động của chúng bằng cách liên kết các kháng thể với protein kháng nguyên. Dự đoán các vùng epitope có lợi cho việc thiết kế và phát triển vắc-xin nhằm mục đích tạo ra kháng thể đặc hiệu với kháng nguyên. Bộ dữ liệu và mã này cho thấy rằng chúng sẽ hữu ích rộng rãi không chỉ cho COVID-19 mà còn cho việc phân tích dữ liệu y tế trong tương lai.

Bộ dữ liệu COVID-19 đại diện cho bộ sưu tập tài liệu về coronavirus có thể đọc được bằng máy mở rộng nhất hiện có để khai thác dữ liệu cho đến nay. Điều này cho phép cộng đồng nghiên cứu AI trên toàn thế giới có cơ hội áp dụng các phương pháp tiếp cận khai thác dữ liệu và văn bản để tìm câu trả lời cho các câu hỏi bên trong và kết nối thông tin chi tiết về nội dung này để hỗ trợ các nỗ lực ứng phó COVID-19 đang diễn ra trên toàn thế giới. Ngày càng có nhiều cấp bách đối với các phương pháp tiếp cận này vì sự gia tăng nhanh chóng trong các tài liệu về coronavirus, khiến cộng đồng y tế khó theo kịp.

Danh sách các câu hỏi chính ban đầu được tìm thấy trong phần nhiệm vụ của tập dữ liệu này. Những câu hỏi khoa học quan trọng này được rút ra từ các chủ đề nghiên cứu SCIED của NASEM (Viện Hàn lâm Khoa học, Kỹ thuật và Y học Quốc gia về các bệnh truyền nhiễm mới nổi và các mối đe dọa sức khỏe thế kỷ 21) và Kế hoạch chi tiết về R&D của Tổ chức Y tế Thế giới cho COVID-19.

Nhiều câu hỏi trong số này phù hợp để khai thác văn bản và các nhà nghiên cứu phát triển các công cụ khai thác văn bản để cung cấp thông tin chi tiết về những câu hỏi này.

2. Lịch sử giải quyết vấn đề

Để đối phó với đại dịch COVID-19, Nhà Trắng và liên minh các nhóm nghiên cứu hàng đầu đã chuẩn bị Bộ dữ liệu nghiên cứu mở COVID-19 (CORD-19). CORD-19 là một nguồn tài nguyên của hơn 200.000 bài báo học thuật, trong đó có hơn 100.000 bài báo có nội dung đầy đủ, về COVID-19, SARS-CoV-2 và các coronavirus liên quan. Bộ dữ liệu có sẵn miễn phí này được cung cấp cho cộng đồng nghiên cứu toàn cầu để áp dụng những tiến bộ gần đây trong xử lý ngôn ngữ tự nhiên và các kỹ thuật AI khác để tạo ra những hiểu biết mới nhằm hỗ trợ cuộc chiến đang diễn ra chống lại căn bệnh truyền nhiễm này. Ngày càng có nhiều cấp bách đối với các phương pháp tiếp cận này do các tài liệu về coronavirus mới tăng tốc nhanh chóng, khiến cộng đồng nghiên cứu y học khó theo kịp. Đề tài đã và đang kêu gọi các chuyên gia trí tuệ nhân tạo trên thế giới phát triển các công cụ khai thác dữ liệu và văn bản có thể giúp cộng đồng y tế phát triển câu trả lời cho các câu hỏi khoa học có mức độ ưu tiên cao.

3. Mục tiêu đề tài

Đề tài gồm 5 nhiệm vụ với các mục tiêu như sau:

A. Nhiệm vụ 1: Dự đoán SARS với dữ liệu B-cell

→ Phân loại peptit thành hai loại là kháng thể có đặc tính cảm ứng (dương tính) và kháng thể không có đặc tính cảm ứng (âm tính) trên tập dữ liệu tế bào B. Dự đoán các hóa trị kháng thể có giúp ích cho việc tạo kháng sinh không.

B. Nhiệm vụ 2: Dự đoán Covid-19 với dữ liệu B-cell và SARS

→ Phân loại peptit thành hai loại là kháng thể có đặc tính cảm ứng (dương tính) và kháng thể không có đặc tính cảm ứng (âm tính) trên sự kết hợp hai tập dữ liệu tế bào B và SARS-CoV. Dự đoán các hóa trị kháng thể có giúp ích cho việc tạo kháng sinh không.

C. Nhiệm vụ 3: Những gì đã biết về sự lây truyền, ủ bệnh và ổn định của Covid-19 đối với môi trường

→ Tìm ra công thức tính thời gian của sự lây truyền, tốc độ truyền, tốc độ phục hồi, tỷ lệ tử vong trên thế giới. Thể hiện sự lây lan lên bản đồ toàn cầu và cho biết được các độ tuổi nhiễm bệnh cùng các yếu tố liên quan. Các biện pháp ngăn chặn và giảm thiểu sự thiệt hại do Covid-19 gây ra.

D. Nhiệm vụ 4: Số trường hợp nhiễm bệnh, tổng số người chết và đại dịch phát triển ở các quốc gia trên thế giới

→ Giải đáp được quốc gia có số ca tử vong, được xác nhận và phục hồi cao nhất, tạo bảng phân phối các trường hợp và triệu chứng của Covid-19. Diễn giải sự tương quan giữa thời gian kể từ trường hợp đầu tiên và sự phục hồi và dự đoán số người chết theo thời gian trên thế giới.

E. Nhiệm vụ 5: Tạo bảng nghiên cứu liên quan đến Covid-19 và thống kê câu hỏi liên quan đến bệnh nhân

→ Tạo bảng gồm các bài báo nói về vấn đề nghiên cứu Covid-19 cùng các tiến bộ trong việc chuẩn đoán và bảng các câu hỏi về thời gian ủ bệnh, thời gian phát tán của Covid-19, thời kỳ ủ bệnh, sự lây truyền, lịch sử tự nhiên của Covid-19 cũng như tỷ lệ bệnh nhân không có triệu chứng và có triệu chứng, ...

4. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu: Ngôn ngữ Python

Phạm vi nghiên cứu: Các vấn đề về đại dịch Covid-19

5. Phương pháp nghiên cứu

- . Phát triển các mô hình máy học, các phương pháp tự động và phân tích dữ liệu để giải quyết các mục tiêu của đề tài đưa ra

6. Bố cục đề tài

Phần giới thiệu

Giới thiệu tổng quát về đề tài.

Phần nội dung

Chương 1: Mô tả bài toán.

Chương 2: Thiết kế, cài đặt giải thuật, biểu diễn cơ sở dữ liệu, phân tích làm rõ mục tiêu ở từng nhiệm vụ

Chương 3: Đánh giá độ chính xác.

Phần kết luận

Trình bày kết quả đạt được và hướng phát triển hệ thống.

PHẦN NỘI DUNG

CHƯƠNG 1

MÔ TẢ BÀI TOÁN

1. Mô tả chi tiết bài toán

Đề tài gồm năm nhiệm vụ được chia thành hai giai đoạn chính: Dự đoán biểu hiện tế bào B Covid-19/SARS và phân tích đại dịch Covid-19

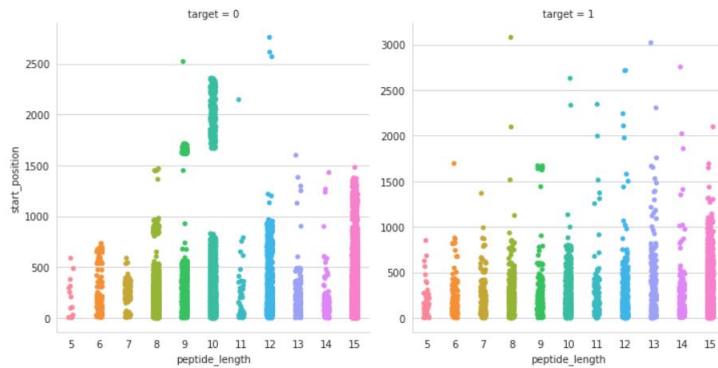
Dự đoán biểu hiện tế bào B Covid-19/SARS:

Bộ dữ liệu gồm 3 tập dữ liệu chính:

- *input_bcell.csv*: đây là dữ liệu đào tạo chính. Số hàng là 14387 cho tất cả các tổ hợp gồm 14362 peptit và 757 protein.
- *input_sars.csv*: đây cũng là dữ liệu đào tạo chính. Số hàng là 520.
- *input_covid.csv*: đây là dữ liệu mục tiêu và không có dữ liệu nhãn trong các cột.
➔ Đây là bộ dữ liệu thuộc học bán dám sát (có hai tập dữ liệu được gán nhãn và một tập dữ liệu không có nhãn). Khi đó nhãn của tập dữ liệu chưa được gán sẽ được tính trọng số bởi các ước tính xác suất của chúng. Nghĩa là, lớp được dự đoán là lớp có ước tính xác suất trung bình cao nhất trên các cây.

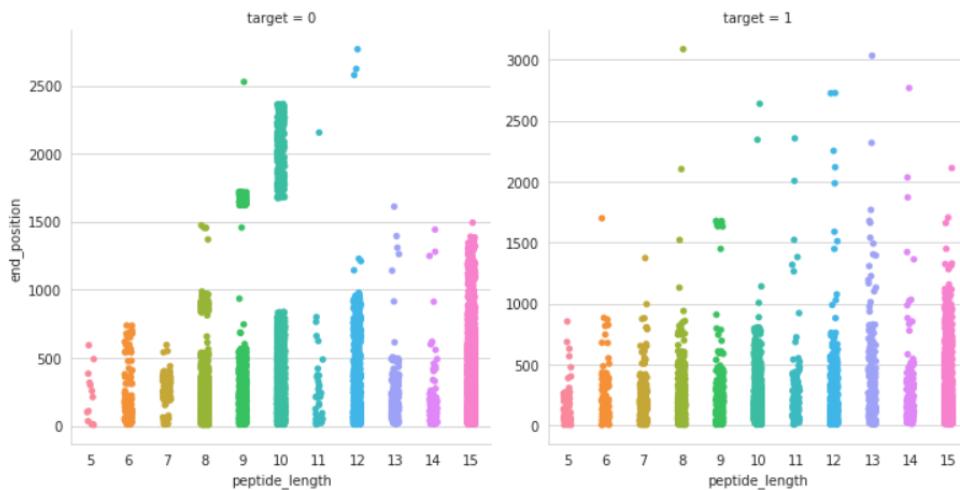
Tất cả ba bộ dữ liệu bao gồm thông tin của protein và peptit: phân tích rõ từng thuộc tính

- parent_protein_id: ID protein mẹ
- protein_seq: trình tự protein mẹ. SARS-CoV-2 có độ dài khoảng 25-32 kilobase
- start_position: [2] vị trí bắt đầu của peptit, vùng 5'UTR. Trên đoạn protein ta sẽ cắt ngẫu nhiên bất kỳ một vị trí để làm vị trí bắt đầu



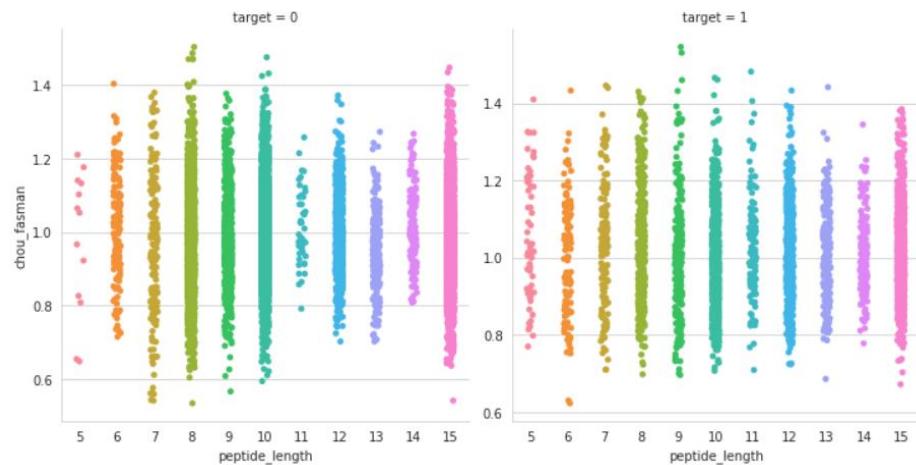
Hình 1. Thuộc tính start_position

- end_position: ⁰vị trí cuối của peptit, vùng 3'UTR và cuối cùng là Đuôi-poly (A). Từ vị trí bắt đầu đã được cắt lấy 4 đơn vị tiếp theo để xác định vị trí kết thúc của peptit. Lưu ý các đoạn peptit phải bằng nhau, như thế sẽ dễ dàng tính toán và đưa ra dữ liệu chính xác.



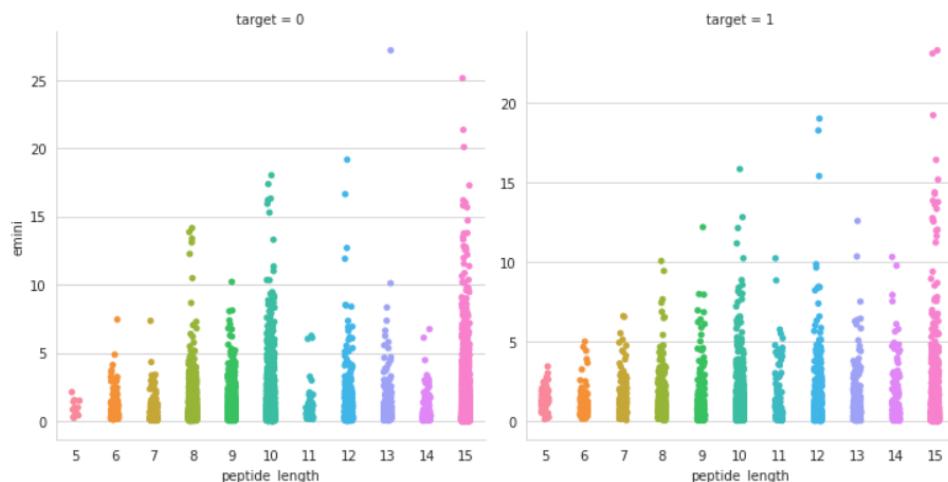
Hình 2. Thuộc tính end_position

- peptide_seq: chuỗi peptit
- chou_fasman: ^[1]tính năng peptit, lần lượt β, gây bệnh ở dơi, chuột, các virus SARS được cho là từ dơi và MERS-CoV được cho là từ dơi và lạc đà truyền sang người, các virus HCoV043 và HCoV-HKU1 gây cảm lạnh ở người cũng thuộc nhóm này, gây ra khoảng 20% số ca nhiễm trùng đường hô hấp. Nói cách khác, khi con người ở gần các loài dơi hay lạc đà có dấu hiệu của virus thì các virus sẽ có khả năng xâm nhập vào cơ thể người và gây ra các dấu hiệu như cảm lạnh bất thường và hô hấp sẽ trở nên khó khăn hơn.



Hình 3. Thuộc tính chou_fasman

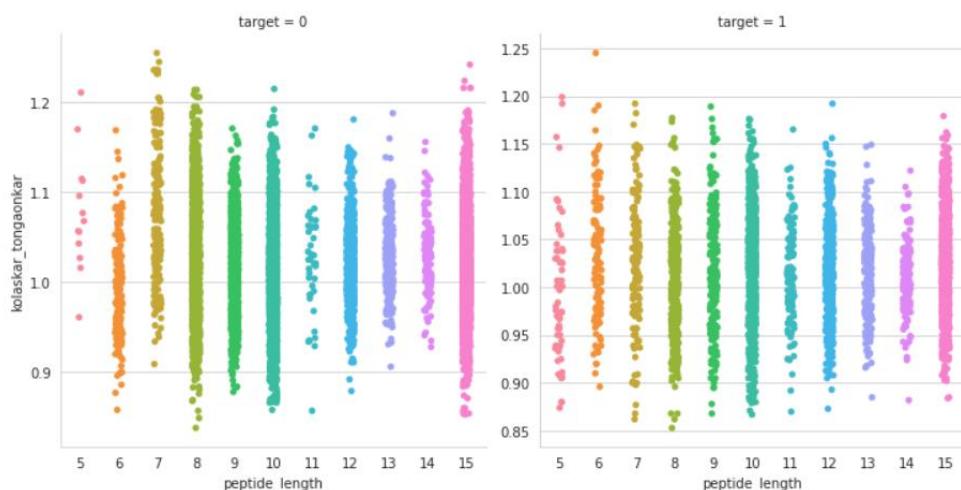
- emini: ^[3] tính năng peptide, khả năng tiếp cận bề mặt tương đối. Khả năng tiếp cận dung môi dư (RSA) trong protein là phạm vi diện tích bề mặt có thể tiếp cận được của một chất cặn nhất định và có liên quan đến sự sắp xếp và đóng gói không gian của cặn. Nó cho thấy trạng thái gấp của protein và được coi là một phép đo định lượng đáng kể cho cấu trúc ba chiều của protein. Khả năng tiếp cận dung môi liên quan chặt chẽ đến việc xác định miền cấu trúc. Giá trị càng cao thì kháng nguyên càng hấp dẫn, dễ cho việc tạo kháng thể. Hay nói một cách dễ hiểu, khả năng tiếp cận bề mặt tương đối này có thể xem như chất kết dính của protein, như các đóng cộc để ngăn chong nước lũ (covid), khi các cộc có sự kết chặt với nhau thì nước lũ sẽ khó có thể phá vỡ được mà tràn vào đất liền giúp cho đất liền không bị xâm ngập trong lũ (khả năng tạo kháng thể mạnh).



Hình 4. Thuộc tính emini

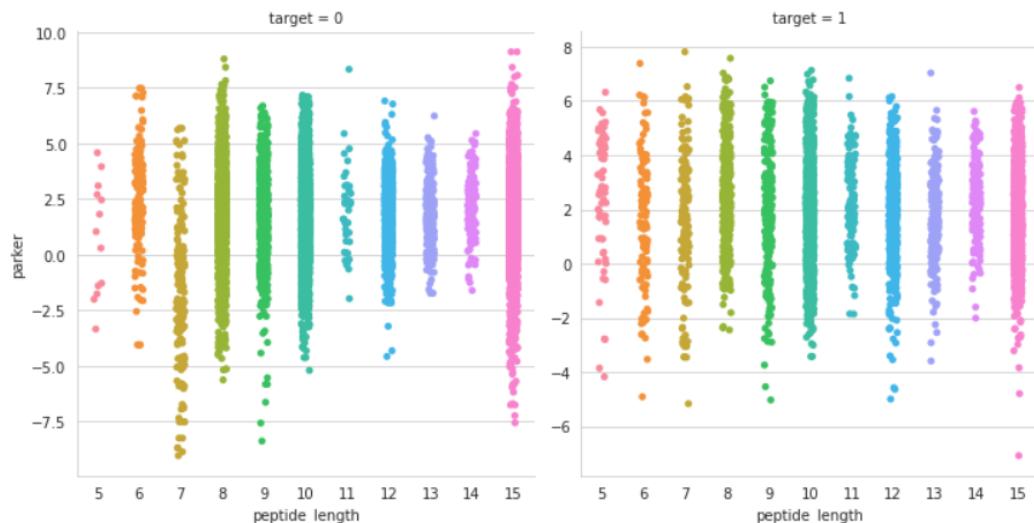
- kolaskar_tongaonkar: ^[4] tính năng peptide, tính kháng nguyên, Peptide là những phân tử lớn hơn, vì vậy chúng thực sự có thể bám vào coronavirus và

úc ché sự xâm nhập vào tế bào, trong khi nếu bạn sử dụng một phân tử nhỏ, rất khó để chặn toàn bộ khu vực mà virus đang sử dụng, các kháng thể cũng có diện tích bề mặt lớn, vì vậy chúng cũng có thể hữu ích. Chúng chỉ mất nhiều thời gian hơn để sản xuất và khám phá. Điểm kháng nguyên nằm trong khoảng từ 0,9567 đến 1,6969. Có thể xem tính kháng nguyên này như hiện tượng kháng sinh giữa nấm mốc và vi khuẩn, khi đó nấm sẽ tiết ra chất gây úc ché sự sinh sản và phát triển mạnh của vi khuẩn, khiến cho vi khuẩn ngày càng yếu dần đi và không còn khả năng hoạt động nữa.



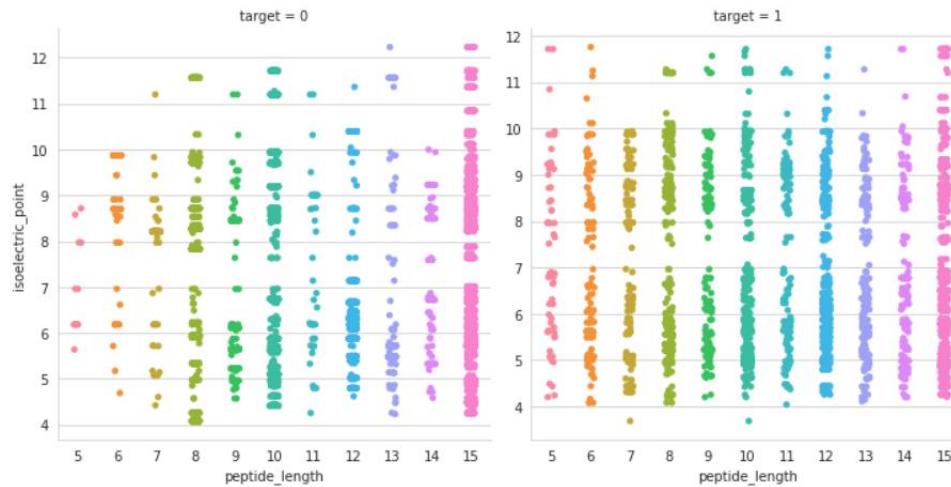
Hình 5. Thuộc tính kolaskar_tongaonkar

- parker: tính năng peptide, tính kỹ nước.^[4] Trong dự đoán tính ưa nước của Parker, điểm tính ưa nước trung bình của protein là 1,480, với tối đa là 4,929 và tối thiểu là -6,843, tất cả các giá trị đều bằng hoặc lớn hơn ngưỡng mặc định -0. 911 có khả năng ưa nước. Như biểu đồ tương quan cho thấy khi protein nằm ở tập trung ở mức ưa nước thì liên kết chặt chẽ hơn nhiều giúp cho việc tạo kháng thể chống lại virus một phần có lợi thế hơn.



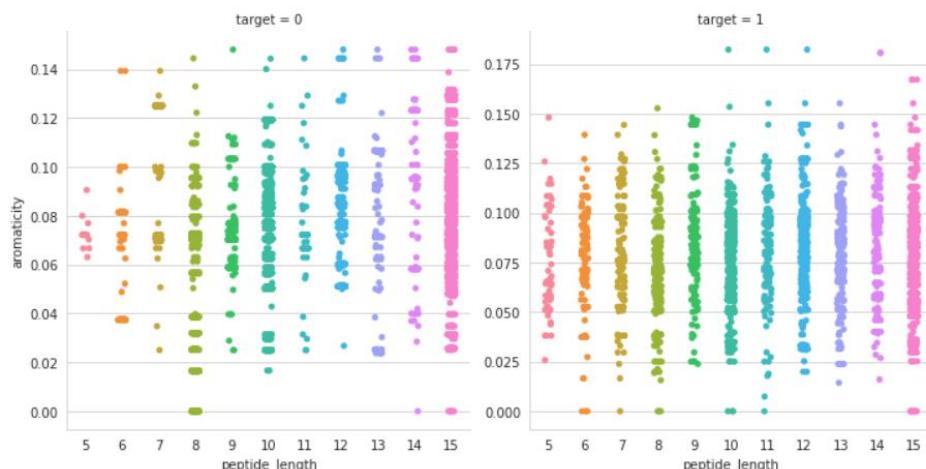
Hình 6. Thuộc tính parker

- isoelectric_point: tính năng protein_ Điểm đẳng điện,^[5] điểm đẳng điện là rất quan trọng bởi vì độ hòa tan và lực đẩy điện là thấp nhất tại p tối. Do đó, xu hướng kết tụ và kết tủa là cao nhất. Trong trường hợp của virus, giá trị do đó cung cấp thông tin về điện tích bề mặt của virus trong một môi trường cụ thể. Các điểm đẳng điện của virus trong đó họ đánh giá 137 p I các phép đo của 104 loại virus. Đánh giá cho thấy điểm đẳng điện (IEP) của vi rút nằm trong khoảng từ 1,9 đến 8,4, nhưng hầu hết có thể được tìm thấy trong khoảng từ 3,5 đến 7, có nghĩa là vi rút có giá trị p I trong phạm vi cơ bản mạnh vẫn chưa được mô tả. Các đề xuất chính là vi rút có thể có nhiều hơn một p I và giá trị p I phụ thuộc vào điều kiện điện phân mà chúng tiếp xúc. Sự phụ thuộc của các trạng thái điện tích, và do đó, các giá trị p I, vào sự có mặt của các ion kim loại thường được quan sát thấy. Hiện tại, không có dữ liệu thực nghiệm nào cho giá trị p I của SARS - CoV - 2 hoặc các protein cấu trúc, phi cấu trúc hoặc phụ kiện của nó. Các giá trị điểm đẳng điện càng nhỏ càng tốt. nói cách khác pI của protein té bào chủ cao hơn so với của protein đột biến, lớn hơn sẽ là ái lực của tăng đột biến protein để gắn vào nó. Một cách dễ hiểu, có thể xem điểm đẳng điện này như các sai số trong quá trình tính toán của việc đầu tư cổ phiếu, cũng như đã nói thì các sai số (điểm đẳng điện) này cực kỳ quan trọng, nó ảnh hưởng đến việc đầu tư vào cổ phiếu nào, khi ta thấy giá cổ phiếu nào có độ sai lệch nguy hiểm cao thì sẽ không chọn và ngược lại để có thể đầu tư vào cổ phiếu tốt nhất. Điểm đẳng điện cũng vậy, đây là thuộc tính có lợi cho covid corona, vì thế giá trị điểm đẳng điện càng nhỏ sẽ càng bất lợi với chúng nhưng sẽ rất có lợi cho việc tạo ra kháng thể.



Hình 7. Thuộc tính isoelectric_point

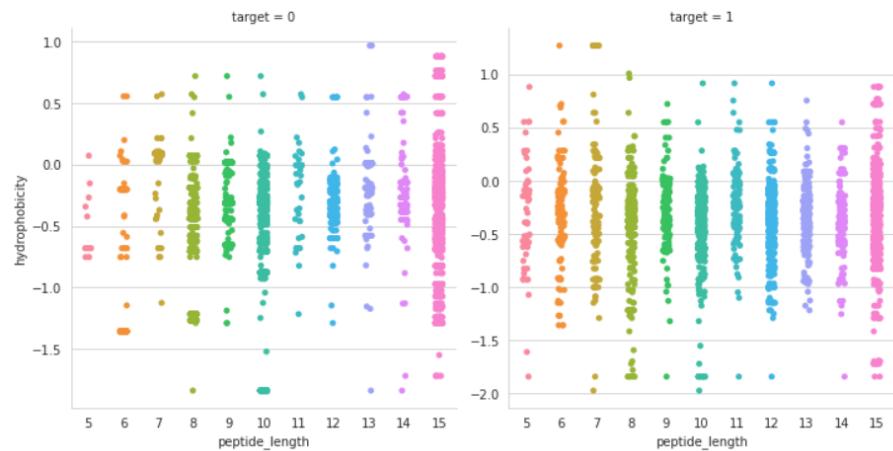
- aromacity: tính năng protein, là tác nhân hợp lý hóa^[6] vai trò dường như vô tội của histidine (His41 of M pro). Nó có liên kết hydro với nhóm hydroxyl và là chất mang proton của thiol của Cys145 với chi phí năng lượng gần như bằng không, tạo điều kiện cho tương tác với chất ức chế hoạt động như chất nhận Michael. Nói cách khác, có thể xem aromacity như yếu tố thời tiết, hydro là giống cây trồng và hydroxyl là nguồn nước. Để có một vụ mùa bội thu cần ngoài nguồn nước và giống tốt cần phải có thời tiết tốt là môi trường trung gian hòa hợp lại, việc tạo ra kháng thể chống lại virus corona cũng vậy.



Hình 8. Thuộc tính aromacity

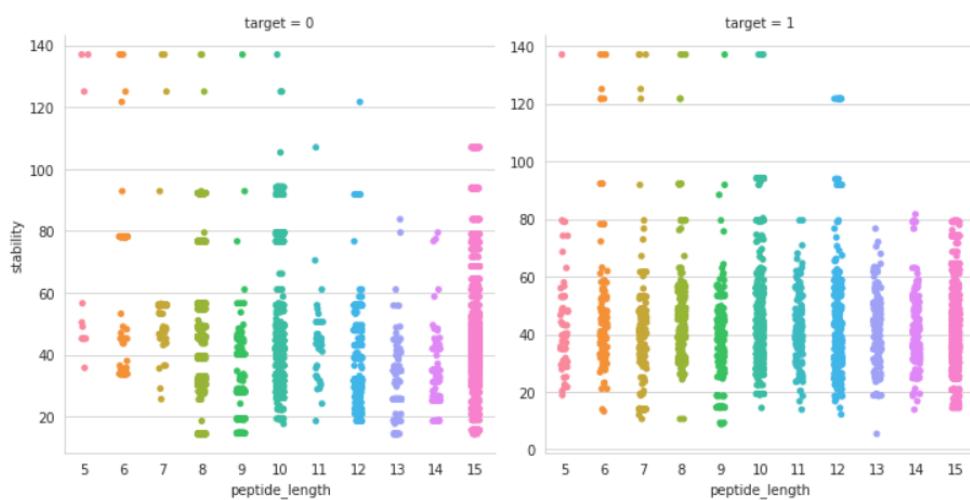
- hydrophobicity: tính năng protein, hàm lượng nước^[7] thấp trong bioaerosol dẫn đến nồng độ H tự do cao+ và do đó pH thấp. Điều này gây ra những thay đổi về cấu trúc đối với glycoprotein của virus và phá hủy khả năng tồn tại của virus. Sự gắn kết của vi rút trên các vật liệu kỵ nước dẫn đến phá hủy cấu trúc của vi rút thông qua các lực như sự sắp xếp lại cấu trúc của súc cảng bề mặt

và ứng suất cắt. Vì SARS-CoV-2 bị vô hiệu hóa trong môi trường khô ráo, nên việc áp dụng các vật liệu kỹ nước có thể làm mất độ ẩm của vi rút có thể được áp dụng như một phương pháp ngăn chặn sự lây lan của nó. Việc này có thể dễ dàng thấy được khi nước ta dễ dàng không chế virus và điều trị cho các bệnh nhân khỏi bệnh hơn các nước bạn như Hàn Quốc,... Do nước ta có khí hậu nóng – khắc tinh của virus corona.



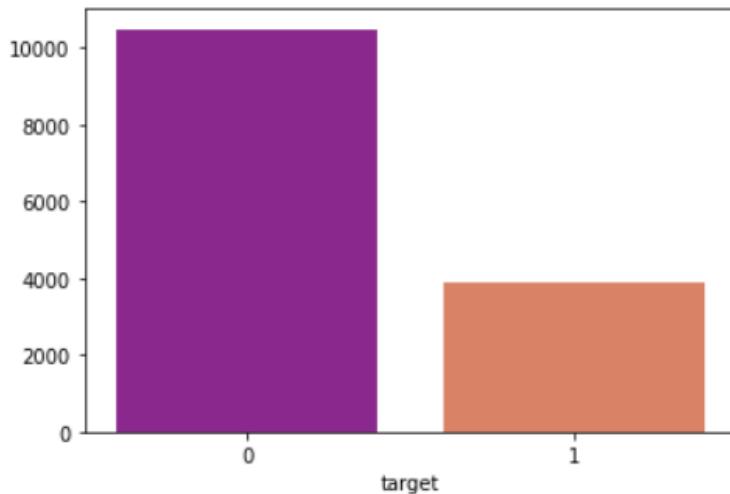
Hình 9. Thuộc tính hydrophobicity

- stability: tính năng protein, CoV tồn tại lâu hơn trong môi trường ở nhiệt độ [8] thấp hơn và độ ẩm tương đối thấp hơn. Nguồn 22, và pH cao hơn bình thường thì tồn tại lâu hơn. Như ta đã thấy ở thực tế, Hàn Quốc và các nước bạn có khí hậu lạnh khó không chế độ lây lan của covid do chúng ưa môi trường khí hậu lạnh, khí hậu lạnh sẽ giúp chúng tồn tại lâu với sức sống mạnh mẽ và ngược lại ở môi trường có khí hậu nóng ẩm.



Hình 10. Thuộc tính stability

- target: hóa trị kháng thể (giá trị đích), trong đó 1 là dương tính và 0 là âm tính. Càng có nhiều giá trị âm tính thì việc tạo kháng thể càng thành công.



Hình 11. Giá trị nhãn

Ở đây giá trị âm tính (0) có số lượng lớn hơn (60%) so với giá trị dương tính (1) có số lượng khoảng 40%. Với tỷ lệ 6:4 này là mất cân bằng dữ liệu, tuy nhiên tỷ lệ này sẽ không ảnh hưởng nhiều đến mô hình dự đoán và độ chính xác của giải thuật.

→ Dữ liệu được cung cấp từ [Cơ sở dữ liệu Epitope miễn dịch \(IEDB\)](#) và [UniProt](#).

Phân tích đại dịch Covid-19:

Các tập dữ liệu được sử dụng trong giai đoạn này gồm: Covid_19_data.csv, COVID 19 Containment measures data.csv, PatientInfo.csv, Case.csv, Time_series_covid_19_confirmed.csv, COVID19_open_line_list.csv, COVID19_line_list_data.csv, Metadata.csv cùng hai tệp giải nén đính kèm.

❖ covid_19_data.csv

SNo	ObservationDate	Province/State	Country/Region	Last Update	Confirmed	Deaths	Recovered	
0	1	01/22/2020	Anhui	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
1	2	01/22/2020	Beijing	Mainland China	1/22/2020 17:00	14.0	0.0	0.0
2	3	01/22/2020	Chongqing	Mainland China	1/22/2020 17:00	6.0	0.0	0.0
3	4	01/22/2020	Fujian	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
4	5	01/22/2020	Gansu	Mainland China	1/22/2020 17:00	0.0	0.0	0.0

Hình 12. Tập dữ liệu Covid_19_data.csv

Trong đó:

- Observation Date: Ngày bắt đầu quan sát tình hình dịch bệnh
- Province/State: Tỉnh hay bang được theo dõi tình hình phát sinh dịch bệnh
- Country/Region: Quốc gia, khu vực được điều tra dịch bệnh
- Last Update: Lần cập nhật cuối cùng, thời gian tính được cập nhật cho tỉnh hoặc quốc gia nhất định.
- Confirmed: Tích lũy số trường hợp được xác nhận cho đến thời điểm quan sát theo dõi.
- Deaths: Số người chết cộng dồn đến thời điểm quan sát.
- Recovered: Số trường hợp đã phục hồi cho đến thời điểm hiện tại quan sát.

❖ *PatientInfo.csv*

patient_id	sex	age	country	province	city	infection_case	infected_by	contact_number	symptom_onset_date	confirmed_date	released_date	deceased_date	state
0	1000000001	male	50s	Korea	Seoul	Gangseo-gu	overseas inflow	NaN	75	2020-01-22	2020-01-23	2020-02-05	NaN released
1	1000000002	male	30s	Korea	Seoul	Jungnang-gu	overseas inflow	NaN	31	NaN	2020-01-30	2020-03-02	NaN released
2	1000000003	male	50s	Korea	Seoul	Jongno-gu	contact with patient	2002000001	17	NaN	2020-01-30	2020-02-19	NaN released
3	1000000004	male	20s	Korea	Seoul	Mapo-gu	overseas inflow	NaN	9	2020-01-26	2020-01-30	2020-02-15	NaN released
4	1000000005	female	20s	Korea	Seoul	Seongbuk-gu	contact with patient	1000000002	2	NaN	2020-01-31	2020-02-24	NaN released

Hình 13. Tập dữ liệu PatientInfo.csv

Trong đó:

- patient_id: id bệnh nhân
- sex: giới tính
- age: độ tuổi
- country: quốc gia
- province: tỉnh
- city: thành phố
- infection_case: tình trạng nhiễm trùng
- infected_by: lý do nhiễm trùng
- contact_number: số liên lạc
- symptom_onset_date: ngày có triệu chứng
- confirmed_date: ngày xác nhận
- released_date: ngày kết thúc
- deceased_date: ngày qua đời
- state: thuộc tiểu bang

❖ Case.csv

	case_id	province	city	group	infection_case	confirmed	latitude	longitude
0	1000001	Seoul	Yongsan-gu	True	Itaewon Clubs	139	37.538621	126.992652
1	1000002	Seoul	Gwanak-gu	True	Richway	119	37.48208	126.901384
2	1000003	Seoul	Guro-gu	True	Guro-gu Call Center	95	37.508163	126.884387
3	1000004	Seoul	Yangcheon-gu	True	Yangcheon Table Tennis Club	43	37.546061	126.874209
4	1000005	Seoul	Dobong-gu	True	Day Care Center	43	37.679422	127.044374

Hình 14. Tập dữ liệu Case.csv

Trong đó:

- Province: tỉnh
- City: thành phố
- Group: nhóm
- Infection_case: trường hợp nhiễm trùng
- Confirmed: trường hợp đã được xác nhận
- Latitude: chiều rộng
- Longitude: chiều dài

❖ COVID 19 Containment measures data.csv

Country	Date Start	Description of measure implemented
1458	South Korea Apr 01, 2020	All arrival passengers from the world (including travel and short term travellers)- must be in 14 days quarantine either in self-isolation (if you... text alert system announced
1695	US:Louisiana Apr 02, 2020	
1690	US: Rhode Island Apr 03, 2020	State beaches and parks closed as of this date, Governor Raimondo cites a lack of social distancing as the reason for the decision
1692	US: Rhode Island Apr 03, 2020	Governor Raimondo announces three field hospitals will be set up in the state to add 1,000 hospital beds in case of hospital overflow
1638	US:Hawaii Apr 04, 2020	Gov. Ige and the Hawaii Department of Health and the Hawaii Emergency Management Agency are formulating a statewide policy on the voluntary use ...

Hình 15. Tập dữ liệu COVID 19 Containment measures data.csv

Trong đó:

- Country: quốc gia được nói đến
- Date Start: ngày bắt đầu biện pháp
- Description of measure implemented: cách thức biện pháp đã thực hiện trong quá trình

❖ covid_19_data.csv

Country/Region	ObservationDate	Confirmed	Deaths	Recovered	CurrentConfirmed
Azerbaijan	02/28/2020	1.0	0.0	0.0	1.0
('St. Martin',)	03/10/2020	2.0	0.0	0.0	2.0
Afghanistan	02/24/2020	1.0	0.0	0.0	1.0
	02/25/2020	1.0	0.0	0.0	1.0
	02/26/2020	1.0	0.0	0.0	1.0

Hình 16. Tập dữ liệu covid_19_data.csv sau khi tiền xử lý

Với:

- Confirmed: tỉ lệ đã xác nhận nghi ngờ
- Deaths: Tỉ lệ tử vong
- Recovered: tỉ lệ phục hồi
- CurrentConfirmed: tỉ lệ trường hợp đã được xác nhận ở hiện tại
- Country/Region: Quốc gia/ khu vực được quan sát
- ObservationDate: ngày bắt đầu quan sát

❖ time_series_covid_19_confirmed.csv

Trong đó:

- Country/Region: Quốc gia/ khu vực được quan sát
- Long: số lượng tới ngày cuối cùng
- Lat: số lượng khi bắt đầu khảo sát
- Các cột còn lại là số lượng người nhiễm trong từng ngày

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	1/28/20	1/29/20	1/30/20	1/31/20	2/1/20	2/2/20	2/3/20	2/4/20
0	Nan	Afghanistan	33.93911	67.709953	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Nan	Albania	41.15330	20.168300	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Nan	Algeria	28.03390	1.659600	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Nan	Andorra	42.50630	1.521800	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Nan	Angola	-11.20270	17.873900	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Hình 17. Tập dữ liệu time_series_covid_19_confirmed.csv

❖ metadata.csv

cord_uid	sha	source_x	title	doi	pmcid	pubmed_id	license	abstract	publish_time	authors	journal	mag_id
0 ug7v899j	d1aafb70c066a2068b02786f8929fd9c900897fb	PMC	Clinical features of culture-proven Mycoplasma...	10.1186/1471-2334-1-6	PMC35282	11472636	no-cc	OBJECTIVE: This retrospective chart review des...	2001-07-04	Madani, Tariq A; Al-Ghamdi, Aisha A	BMC Infect Dis	NaN
1 02lnwd4m	6b0567729c2143a66d737eb0a2f63f2dce2e5a7d	PMC	Nitric oxide: a pro-inflammatory mediator in I...	10.1186/rr14	PMC59543	11667967	no-cc	Inflammatory diseases of the respiratory tract...	2000-08-15	Vliet, Albert van der; Eisnerich, Jason P; Cros...	Respir Res	NaN
2 ejv2xin0	06ced00a5fc04215949aa72528f2eeaae1d58927	PMC	Surfactant protein-D and pulmonary host defense	10.1186/rr19	PMC59549	11667972	no-cc	Surfactant protein-D (SP-D) participates in th...	2000-08-25	Crouch, Erika C	Respir Res	NaN

Hình 18. Tập dữ liệu metadata.csv

Trong đó có các cột cần chú ý:

- Title: tiêu đề bài báo
- Abstract: tóm tắt bài báo
- Authors: tác giả của bài báo

2. Vấn đề và giải pháp liên quan đến bài toán

Do sự lây lan của COVID-19, việc phát triển vắc-xin đang được tiến hành càng sớm càng tốt. Với tầm quan trọng của phân tích dữ liệu trong quá trình phát triển vắc xin, tuy không có nhiều bộ dữ liệu phân tích đơn giản để có thể xử lý. Đề tài đã xuất bản bộ dữ liệu và mã mẫu cho dự đoán Bcell epitope, một trong những chủ đề nghiên cứu quan trọng trong việc phát triển vắc xin.

Vấn đề về việc liệu một peptit axit amin có biểu hiện hoạt tính tạo kháng thể hay không được lấy từ IEDB, đã được sử dụng trong nhiều nghiên cứu trước đây. Theo đó, thông tin này đã được sử dụng làm dữ liệu nhãn. Các tổ chức y tế cũng thu được trình tự axit amin ứng cử viên của vi sinh vật (peptit) và dữ liệu nhãn hoạt động từ dữ liệu biểu mô té bào B được cung cấp trong IEDB. Các protein kháng thể được trình bày bị hạn chế đối với IgG tạo thành loại được ghi nhận nhiều nhất trong IEDB. Để thuận tiện, tổ chức y tế đã loại trừ các bản ghi đại diện cho các phép đo định lượng khác nhau về hoạt tính của kháng thể đối với cùng một peptit khỏi các thí nghiệm. Dữ liệu epitope thu được từ IEDB tương ứng với năm loại hoạt động: "Tích cực-Cao", "Tích cực-Trung bình", "Tích cực-Thấp", "Tích cực" và "Tiêu cực".

Để đối phó với đại dịch COVID-19, Nhà Trắng và liên minh các nhóm nghiên cứu hàng đầu đã chuẩn bị Bộ dữ liệu nghiên cứu mở COVID-19 (CORD-19). CORD-19 là một nguồn tài nguyên của hơn 200.000 bài báo học thuật, trong đó có hơn 100.000 bài báo có nội dung đầy đủ, về COVID-19, SARS-CoV-2 và các coronavirus liên quan. Bộ dữ liệu có sẵn này được cung cấp cho cộng đồng nghiên cứu toàn cầu để áp dụng những tiến bộ gần đây trong xử lý ngôn ngữ tự nhiên và các kỹ thuật AI khác để tạo ra những hiểu biết mới nhằm hỗ trợ cuộc chiến đang diễn ra chống lại căn bệnh truyền nhiễm này. Ngày càng có nhiều cấp bách đối với các phương pháp tiếp cận này do các tài liệu về coronavirus mới tăng tốc nhanh chóng, khiến cộng đồng nghiên cứu y khoa khó theo kịp.

CHƯƠNG 2

THIẾT KẾ VÀ CÀI ĐẶT

Đề tài niêm luận gồm năm nhiệm vụ được chia làm hai giai đoạn chính:

Giai đoạn 1: Dự đoán biểu hiện té bào B-cell/SARS

- Nhiệm vụ 1: Dự đoán SARS với dữ liệu B-cell
- Nhiệm vụ 2: Dự đoán Covid-19 với dữ liệu B-cell và SARS

Giai đoạn 2: Phân tích đại dịch Covid-19

- Nhiệm vụ 3: Những gì đã biết về sự lây truyền, ủ bệnh và ổn định của Covid-19 đối với môi trường
- Nhiệm vụ 4: Số trường hợp nhiễm bệnh, tổng số người chết và đại dịch phát triển ở các quốc gia trên thế giới
- Nhiệm vụ 5: Tạo bảng nghiên cứu liên quan đến Covid-19 và thống kê câu hỏi liên quan đến bệnh nhân

GIAI ĐOẠN 1: DỰ ĐOÁN BIỂU HIỆN TÉ BÀO B-CELL/SARS

A. NHIỆM VỤ 1: DỰ ĐOÁN SARS VỚI DỮ LIỆU B-CELL

I. Yêu cầu của nhiệm vụ:

- Ước tính vùng biểu mô của SARS-CoV bằng cách chỉ phân tích tập dữ liệu té bào B.
- Vấn đề cần giải quyết: phân loại peptit thành hai loại là kháng thể có đặc tính cảm ứng (dương tính) và kháng thể không có đặc tính cảm ứng (âm tính).
- Sử dụng dữ liệu được cung cấp, dự đoán hóa trị kháng thể của dữ liệu SARS-CoV. Nó có giá trị mục tiêu, cần thử nghiệm và đào tạo bình thường để phát triển mô hình.

II. Ý tưởng giải quyết vấn đề:

Thực hiện huấn luyện và tìm ra giá trị đích của kháng thể trên tập B_cell và so sánh lại với kết quả ở tập covid.

III. Tập dữ liệu:

1. Sơ lược về tập dữ liệu:

Tập dữ liệu gồm các thuộc tính: parent_protein_id, protein_seq, start_position, end_position, peptide_seq, chou_fasman, emini, kolaskar_tongaonkar, parker, isoelectric_point, aromacity, hydrophobicity, stability, target.

Trong đó target là cột nhãn.

parent_protein_id	protein_seq	start_position	end_position	peptide_seq	chou_fasman	emini	kolaskar_tongaonkar	parker
0	A2T3T0	MDVLYSLSKTLKARDKIVEGTLYSNVSDLIQQFNQMIITMNGNEF...	161	165	SASFT	1.016	0.703	1.018 2.22
1	F0V2I4	MTIHKAINGFGRIGRLLFRNLLSSQGVQVVNDVVDIKVLTHLL...	251	255	LCLKI	0.770	0.179	1.199 -3.86
2	O75508	MVATCLQVVGFTVFGWIGIVTTSTNDWWTCGYTIPTCRKLDE...	145	149	AHRET	0.852	3.427	0.960 4.28
3	O84462	MTNSISGYQPTVTTSTSSTTSASGAGSGSLGASSVTTANATVTQTA...	152	156	SNYDD	1.410	2.548	0.936 6.32
4	P00918	MSHHWGYGKHNGPEHWKDFPIAKGERQSPVIDHTAKYDPSLKP...	85	89	DGTYR	1.214	1.908	0.937 4.64

Hình 19. Hình ảnh tập dữ liệu B-cell

2. Tiền xử lý tập dữ liệu:

Các cột parent_protein_id, peptide_seq và protein_seq không phải giá trị liên lục và không có ý nghĩa nên sẽ xóa đi 3 cột đó.

Sau khi xóa ta được tập dữ liệu như sau:

start_position	end_position	chou_fasman	emini	kolaskar_tongaonkar	parker	isoelectric_point	aromaticity	hydrophobicity	stability	target
0	161	1.016	0.703	1.018	2.22	5.810364	0.103275	-0.143829	40.273300	1
1	251	0.770	0.179	1.199	-3.86	6.210876	0.065476	-0.036905	24.998512	1
2	145	0.852	3.427	0.960	4.28	8.223938	0.091787	0.879227	27.863333	1
3	152	1.410	2.548	0.936	6.32	4.237976	0.044776	-0.521393	30.765373	1
4	85	1.214	1.908	0.937	4.64	6.867493	0.103846	-0.578846	21.684615	1

Hình 20. Tập dữ liệu B-cell sau khi tiền xử lý

IV. Nội dung:

1. Cài đặt thư viện:

Các thư viện cần thiết cho hai nhiệm vụ:

Numpy

Pandas

Seaborn

Matplotlib

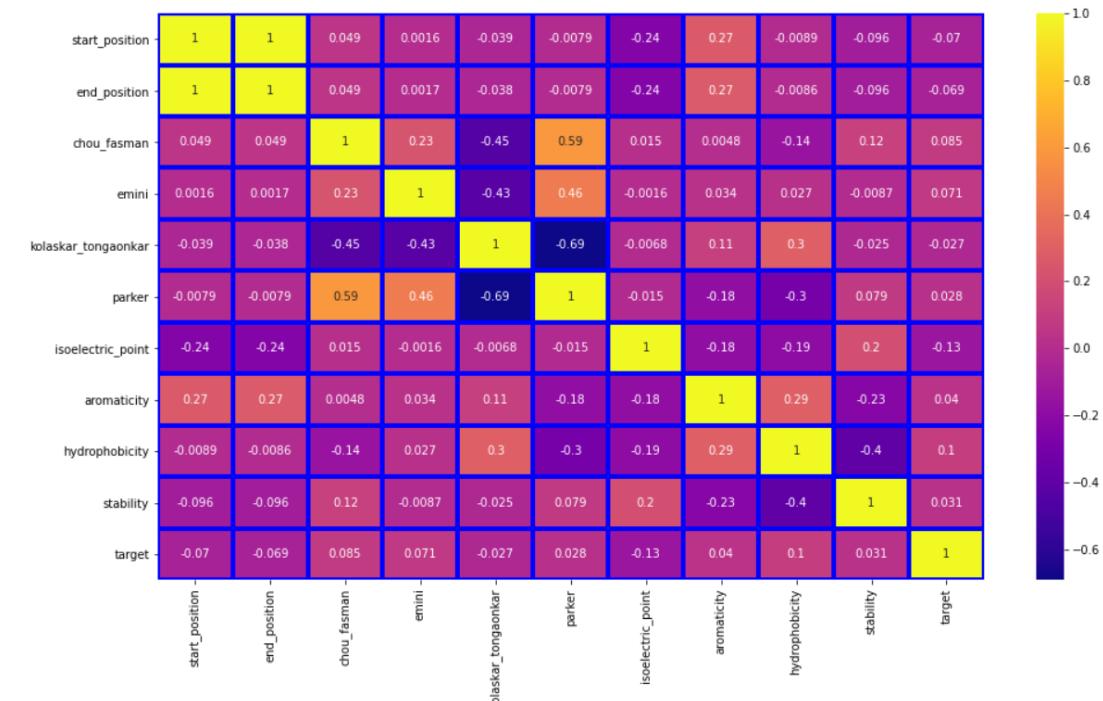
Plotly

Sklearn

2. Phân tích dữ liệu gen:

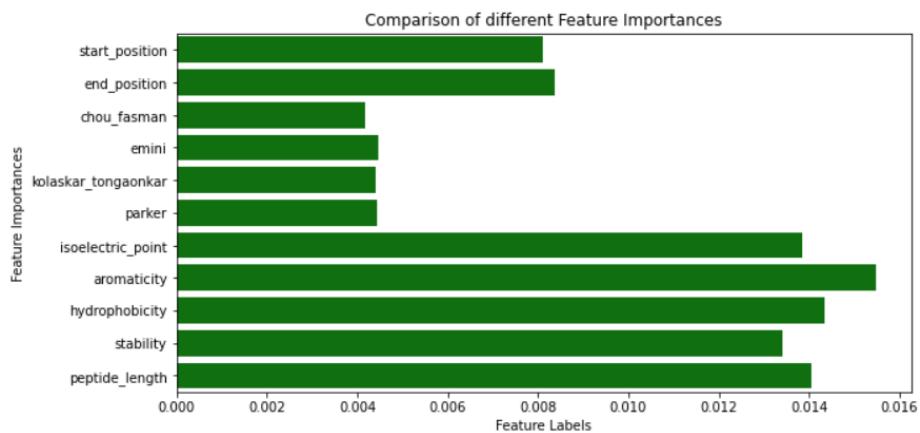
Hiển thị dữ liệu thành bảng đồ để quan sát và xem thuộc tính nào quan trọng nhất.

Xem xét dựa vào chỉ số 1 của đường chéo, xem xét điểm giao nhau giữa hai cột, giá trị nào gần với 1 nhất thể hiện độ quan trọng cao và ngược lại. Như đã phân tích ở trên, ta thấy các giá trị của các thuộc tính khi giao nhau ở gần 1 đa phần đều nằm trong ngưỡng cho phép để thuận tiện cho việc tạo kháng thể chống lại virus corona. Ta sẽ dựa vào các giá trị đó để các nhà khoa học điều chỉnh lượng phù hợp và thông số thích hợp để tạo ra kháng thể tốt nhất.

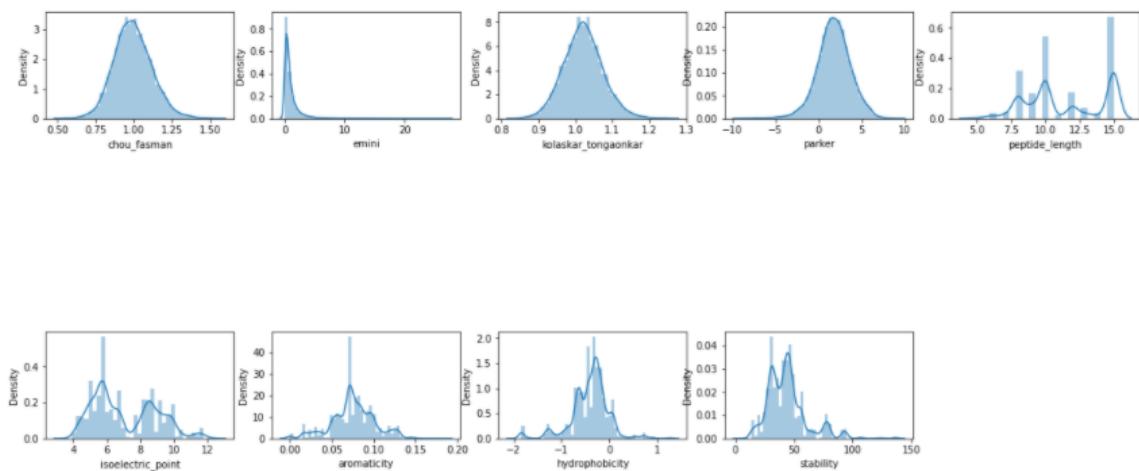


Hình 21. Biểu đồ y sinh thuộc tính quan trọng

Xem xét lại qua biểu đồ tầm quan trọng của thuộc tính. Ta thấy rõ hơn các thuộc tính **'stability'**, **'hydrophobicity'**, **'aromaticity'** và **'isoelectric_point'** có sự liên kết tốt hơn hết trong việc truyền thông tin cho thấy tầm quan trọng của dữ liệu.



Hình 22. Biểu đồ các thuộc tính quan trọng

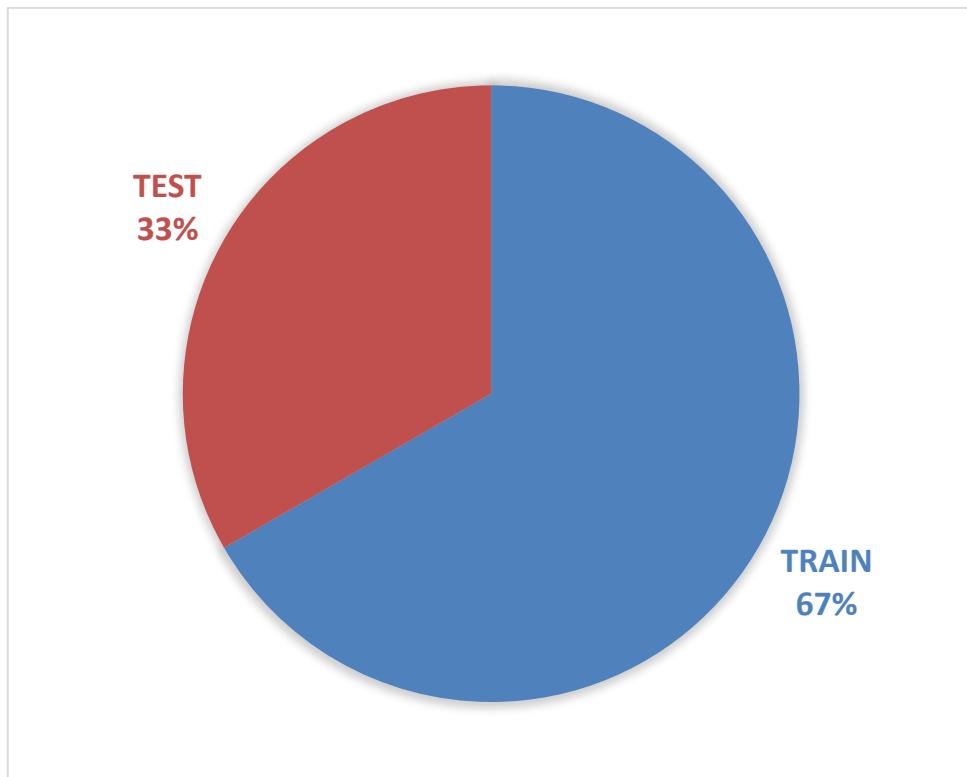


Hình 23. Biểu đồ dữ liệu gen

Theo như các mô hình tương quan của các thuộc tính cho thấy các thuộc tính đều nằm trong ngưỡng cho phép và phần lớn giá trị thuộc tính đều có thể chồng lại được sự tác động và liên kết của virus. Đặc biệt là ở 'stability', 'hydorophobicity', 'aromaticity' và 'isoelectric_point' cho ta thấy rõ nhất về điều đó khi giá trị luôn cao ở các khoảng gần ngưỡng cho phép.

3. Cài đặt giải thuật:

Tiến hành chia tập dữ liệu thành 3 phần test và 7 phần train. Thực hiện trên 3 giải thuật và so sánh độ chính xác để tìm ra giải thuật tốt nhất để huấn luyện.



Hình 24. Biểu đồ phân chia tập dữ liệu train-test

Đề tài dùng 3 giải thuật rừng ngẫu nhiên, cây quyết định, hồi quy Logistic để huấn luyện và tìm ra độ chính xác do giá trị của các biến thuộc giá trị liên tục và phụ thuộc sự kéo theo để đưa ra được nhãn 0 hay 1.

Độ chính xác thử nghiệm lần đầu:

	Hồi Quy Logistic Logistic Regression	Cây Quyết Định Decision Tree	Rừng ngẫu nhiên Random Forest
Độ chính xác Accuracy	73.05%	99.98%	99.63%

Bảng 1. Độ chính xác của các giải thuật

Ở đây ta thấy cây quyết định và rừng ngẫu nhiên có độ chính xác hầu như tương đương nhau, do đó ta cần xác thực lại lần nữa xem giải thuật nào tốt nhất trong 2 giải thuật này

Độ chính xác- Accuracy	
Cây Quyết Định Decision Tree	Rừng ngẫu nhiên Random Forest
83.17%	83.27%
84.91%	86.05%
87.01%	88.93%
88.46%	90.48%
90.4%	91.8%
91.32%	92.44%
92.17%	93.32%
92.73%	94.03%
93.07%	94.42%
93.20%	94.58%

Bảng 2. Độ chính xác của hai giải thuật sau các lần lặp

Sau khi duyệt lại qua 10 lần lặp ta thấy Rừng ngẫu nhiên sẽ tốt hơn cho giải thuật này. [117] Rừng ngẫu nhiên tốt hơn vì nó là thuật toán bao đóng giúp Tăng cường làm giảm phương sai, và cũng làm giảm sự thiên vị. Nó làm giảm phương sai vì sử dụng nhiều mô hình (đóng bao). Nó làm giảm sự thiên vị bằng cách đào tạo mô hình tiếp theo bằng cách cho biết những lỗi mà các mô hình trước đó đã gây ra (phản tăng cường). Nó có thể được áp dụng hiệu quả cho các bộ dữ liệu quy mô lớn cho cả phân loại cũng như phân tích hồi quy làm cho nó có thể mở rộng dựa trên việc tăng kích thước dữ liệu. Do đó, đây là một thuật toán hoàn hảo để thực hiện phân tích dự đoán trên Dữ liệu lớn. [116]

Sau khi huấn luyện các giá trị đã tương đối ổn nhưng để chính xác hơn sẽ kiểm tra và so sánh lại khi thực hiện test ở tập dữ liệu test đã cho.

❖ **Covid.csv:**

Ở tập covid này ta sẽ xử lý các bước xử lý dữ liệu tương tự như tập dữ liệu bcell do các thuộc tính đều giống nhau

start_position	end_position	chou_fasman	emini	kolaskar_tongaonkar	parker	isoelectric_point	aromaticity	hydrophobicity	stability	
0	1	5	0.948	0.280	1.033	-2.720	6.03595	0.10929	-0.138642	31.377603
1	2	6	1.114	0.379	1.070	-0.580	6.03595	0.10929	-0.138642	31.377603
2	3	7	1.106	0.592	1.108	-1.300	6.03595	0.10929	-0.138642	31.377603
3	4	8	1.324	0.836	1.053	1.440	6.03595	0.10929	-0.138642	31.377603
4	5	9	1.326	1.004	0.968	2.440	6.03595	0.10929	-0.138642	31.377603
...	
20307	1258	1277	0.947	0.406	1.063	0.155	6.03595	0.10929	-0.138642	31.377603
20308	1259	1278	0.937	0.353	1.061	0.355	6.03595	0.10929	-0.138642	31.377603
20309	1260	1279	0.960	0.647	1.047	0.645	6.03595	0.10929	-0.138642	31.377603
20310	1261	1280	0.960	0.449	1.058	0.540	6.03595	0.10929	-0.138642	31.377603
20311	1262	1281	0.957	0.306	1.067	0.360	6.03595	0.10929	-0.138642	31.377603

Hình 25. Tập dữ liệu Covid.csv

4. So sánh và kết luận:

	Train	Test
0	23.51%	85.98%
1	9.8%	14.02%

Bảng 3. So sánh và kết luận hóa trị kháng sinh

Các dự đoán cho thấy rằng hóa trị kháng thể âm tính khoảng 23.5% và dương tính 0nghĩa là phần lớn các kháng thể sẽ chống lại sự gắn kết của virus như SARS-CoV, điều này sẽ làm giảm số trường hợp mắc bệnh.

B. NHIỆM VỤ 2: DỰ ĐOÁN COVID-19 VỚI DỮ LIỆU B-CELL VÀ SARS

I. Yêu cầu của nhiệm vụ:

- Uớc tính vùng biểu mô của protein đột biến SARS-CoV2 (nguyên nhân gây ra bệnh coronavirus 2019 (COVID-19)) bằng cách sử dụng bộ dữ liệu té bào B và SARS-CoV. Có thể sử dụng thông tin về sự giống nhau giữa chuỗi SARS-CoV và SARS-CoV2.
- Vấn đề đặt ra: phân loại peptide thành hai loại trong nhiệm vụ này: kháng thể có đặc tính cảm ứng (dương tính) và kháng thể không có đặc tính cảm ứng (âm tính).
- Sử dụng dữ liệu được cung cấp, dự đoán hóa trị kháng thể của dữ liệu SARS-CoV2. Nó không có giá trị mục tiêu, vì vậy đây là điểm thách thức của nhiệm vụ này.

II. Ý tưởng giải quyết vấn đề:

Kết hợp 2 tập dữ liệu sars và B_cell, sau đó tiến hành tập huấn và tìm ra giá trị đích của kháng thể trên tập B_cell và so sánh lại với kết quả ở tập covid.

III. Sơ lược về các tập dữ liệu:

Cách xử lý tập dữ liệu tương tự như các bước đã làm ở nhiệm vụ 1.

IV. Nội dung:

1. Cài đặt thư viện, phân tích dữ liệu gen, cài đặt giải thuật:

Ở nhiệm vụ này, tập dữ liệu có sự kết nối giữa hai tập dữ liệu riêng biệt có cùng các thuộc tính. Do đó các thư viện cài đặt, phân tích dữ liệu gen và tiến hành lại các giải thuật ở nhiệm vụ 1 để xem xét sự thay đổi.

2. So sánh và kết luận:

Các dự đoán cho thấy rằng hóa trị kháng thể âm tính khoảng 22.8% và 53.23%, có nghĩa là phần lớn các kháng thể ở nhiệm vụ này cũng sẽ chống lại sự gắn kết của virus như SARS-CoV, điều này sẽ làm giảm số trường hợp mắc bệnh, giúp tạo kháng thể thuốc nhanh chóng hơn.

	Train	Test
0	22.83%	53.23%
1	7.18%	46.76%

Bảng 4. Kết luận về hóa trị kháng sinh

Ở cả hai nhiệm vụ cho thấy giá trị âm tính đều lớn hơn giá trị dương tính, điều này cho thấy lợi thế của việc tạo ra kháng thể chống lại sự phát triển của covid corona.

GIAI ĐOẠN 2: PHÂN TÍCH ĐẠI DỊCH COVID-19

Các thư viện được sử dụng:

- Numpy
- Pandas
- Sklearn
- Resgetion
- Pylot
- Seaborn
- Matplotlib
- BeautifulSoup
- Requests
- Time
- Json

C. NHIỆM VỤ 3: NHỮNG GÌ ĐÃ BIẾT VỀ SỰ LÂY TRUYỀN, Ủ BỆNH VÀ ÔN ĐỊNH CỦA COVID-19 ĐỐI VỚI MÔI TRƯỜNG

I. Yêu cầu của nhiệm vụ:

- Tính thời gian của sự lây truyền, tốc độ truyền, tốc độ phục hồi, tỷ lệ tử vong trên toàn thế giới việc kiểm soát dịch bệnh
- Lịch sử tự nhiên của vi rút và sự lây lan của vi rút từ người bị nhiễm, sự toàn cầu của SARS Cov 2_Bản đồ toàn cầu COVID-19
- Sự toàn cầu của SARS Cov 2_Bản đồ toàn cầu COVID-19
- Độ tuổi dễ lây bệnh và các yếu tố liên quan
- Mối liên quan giữa tốc độ tăng trưởng, R0 (t), lưới ngăn chặn và giảm thiểu.
- Các biện pháp ngăn chặn và giảm thiểu

II. Ý tưởng giải quyết vấn đề:

Để trả lời cho các câu hỏi ở nhiệm vụ này, đầu tiên tiến hành tính toán tốc độ truyền bệnh của virus corona theo thời gian ở mỗi châu lục và một số quốc gia khác nhau.

Công việc sẽ được tiến hành như sau:

- Gọi tốc độ truyền SARS Cov 2 (β), tỷ lệ hồi phục (γ) và tỷ lệ tử vong (δ)
- Đánh giá ngắn về tỷ lệ tiếp xúc và tỷ lệ tiếp xúc thích hợp
- Mô hình SIRF với tỷ lệ chuẩn được điều chỉnh
- Thực tế: SIRF gần đúng
- Ước lượng β (t), γ (t), δ (t)
- Xem toàn cầu SARS Cov 2
- Xác định β (t), γ (t), δ (t) ở mỗi lục địa
- Lập biểu đồ tốc độ truyền, tỷ lệ phục hồi và tỷ lệ tử vong cho từng châu lục

- Diễn giải dịch bệnh kiểm soát ở lục địa Châu Đại Dương
- Tìm R₀ theo thời gian
- Nghiên cứu điển hình: Trung Quốc

1. Tỷ lệ tiếp xúc và tỷ lệ tiếp xúc thích hợp

Tỷ lệ tiếp xúc U (N) là số lượng cá thể bị lây nhiễm tiếp xúc trên một đơn vị thời gian. Giả sử rằng xác suất lây nhiễm của mỗi lần tiếp xúc là β_0 , thì tốc độ tiếp xúc thích hợp là $\beta_0U (N)$.

Tỷ lệ tiếp xúc thích hợp trung bình của một cá thể bị nhiễm bệnh với một người nhạy cảm là $\beta_0U (N) SN$. Tỷ lệ này được gọi là tỷ lệ nhiễm trùng. Khi đó, tổng số trẻ mới mắc bệnh được nhiễm bởi tất cả các cá thể trong nhóm bị nhiễm trên một đơn vị thời gian, tại thời điểm t là $(\beta_0U (N) SN) I$, được gọi là tỷ lệ mắc bệnh.

Nếu $U (N) = kN$, tức là tốc độ tiếp xúc tỷ lệ với tổng quy mô quần thể, tỷ lệ mắc bệnh là $\beta (t) S (t) I (t)$, trong đó $\beta = \beta_0k$ được gọi là hệ số truyền (tốc độ truyền). Loại tỷ lệ này được gọi là tỷ lệ song tuyế

Nếu $U (N) = k'$, tức là tốc độ tiếp xúc là một hằng số trong trường hợp này, thì tỷ lệ mắc trở thành βISN , trong đó $\beta = \beta_0k'$, và nó được gọi là tỷ lệ mắc chuẩn.^[115]

2. Mô hình SIRF với tỷ lệ chuẩn được điều chỉnh

Biết rằng covid19 có nhiều biến số quan trọng, có bốn biến số: Cơ sở xác nhận (TotalpositiveCases), Cơ sở xác nhận hiện tại (Cơ sở hiện tại), Đã khôi phục và Cơ sở tử vong. Làm thế nào có thể có được phương trình hệ thống động lực học cho các biến này? Để trả lời câu hỏi này, đề tài sẽ sử dụng Mô hình SIRF với tỷ lệ mắc chuẩn:

Mô hình SIRF với tỷ lệ mắc bệnh chuẩn là một mô hình kinh điển trong dịch tễ học, nó bao gồm 04 quần thể con, những người nhạy cảm S, những người mắc bệnh I và những người đã hồi phục R, những trường hợp tử vong F:

- Chất nhạy cảm
- Nhiễm trùng
- Phục hồi
- Tử vong

Nhiễm trùng có thể trở thành nhiễm trùng, và nhiễm trùng có thể được phục hồi hoặc Tử vong, nhưng không có sự chuyển đổi nào khác được xem xét. Dân số $N = S + I + R + F$ không đổi. Mô hình mô tả chuyển động giữa các lớp bằng hệ phương trình vi phân.

$$\frac{dS}{dt} = -\beta I \frac{S}{N}, \quad \frac{dI}{dt} = \beta I \frac{S}{N} - (\gamma + \delta)I, \quad \frac{dR}{dt} = \gamma I \quad \frac{dF}{dt} = \delta I.$$

Trong đó β là tốc độ truyền, γ là tốc độ phục hồi, δ là tỷ lệ tử vong và $R_0 = \beta / (\gamma + \delta)$.

3. Thực tế: SIRF gần đúng

Trong bối cảnh của sars cov 2 trên thế giới, cần điều chỉnh mô hình SIRF cho phù hợp với dữ liệu của mình để có thể đưa ra một số ước lượng gần đúng về hành vi của bệnh và xác định tốc độ lây truyền và các yếu tố khác. Nếu coi (N) là số lượng dân cư trên một bề mặt cố định nào đó (Km^2) tại thời điểm t . Biết rằng tồn tại một số trường hợp được xác nhận là dân số và số trường hợp không được xác nhận.

quy mô dân số = tổng số trường hợp tích cực + tổng số trường hợp tiêu cực

➔ tổng số trường hợp tích cực = hiện tại + (hồi phục + tử vong)

Do đó, có thể nghĩ quy mô dân số như sau:

quy mô dân số = tổng số lượng phủ định + số lượng hiện tại + số người hồi phục + số người chết (1)

Từ (1) có thể thực hiện một số nhận dạng:

- Quy mô dân số có thể là tổng Dân số (N).
- Tổng số trường hợp tiêu cực có thể là một Cảm ứng (S)
- Số trường hợp tích cực hiện tại có thể là một Infective (I)
- Phục hồi + chết có thể là một Cá nhân được hồi phục (R) + Tử vong (F)

Có thể hiểu như sau:

$$S = N - S_c \rightarrow \frac{S}{N} = 1 - \frac{S_c}{N} \text{ if } \frac{S_c}{N} << 1$$

Ta có $S \approx N$ và Mô hình SIRF với tỷ lệ mắc chuẩn trở thành:

$$\frac{dI}{dt} = (\beta - \gamma - \delta)I, \quad \frac{dR}{dt} = \gamma I \quad \frac{dF}{dt} = \delta I$$

Công thức tính ước lượng $\beta(t)$, $\gamma(t)$, $\delta(t)$ [9]

$$\beta(t) = \frac{\text{số lượng bệnh nhân hiện tại hàng ngày}}{\text{số lượng bệnh nhân covid được xác nhận tích lũy 19 tại thời điểm t}}$$

$$\gamma(t) = \frac{\text{số lượng bệnh nhân covid phục hồi hàng ngày 19 tại thời điểm t}}{\text{số lượng tích lũy được 19 bệnh nhân covid xác nhận tại thời điểm t}}$$

$$\delta(t) = \frac{\text{số người chết hàng ngày là 19 bệnh nhân tại thời điểm t}}{\text{số lượng tích lũy được 19 bệnh nhân covid xác nhận tại thời điểm t}}$$

III. Sơ lược về các tập dữ liệu:

1. Sơ lược về tập dữ liệu:

Nhiệm vụ được tiến hành phân tích trên các tập **covid_19_data.csv**, **COVID 19 Containment measures data.csv**, **PatientInfo.csv**, **Case.csv** và **time_series_covid_19_confirmed.csv**

2. Tiền xử lý tập dữ liệu:

Các tập dữ liệu được dùng nguyên bản không qua bước tiền xử lý. Tuy nhiên, tập dữ liệu **time_series_covid_19_confirmed.csv** phải qua bước tiền xử lý để có output tốt hơn.

Các tính năng có giá trị tìm kiếm là

- Tên quốc gia.
- Số người bị nhiễm
- Ngày

Tại đó họ đã bị ảnh hưởng trong nước. Ở đây đề tài không tập trung vào việc dự đoán cho từng quốc gia vì nó có thể thiên vị về một số quốc gia, kết quả là dự đoán có thể rất lớn hoặc quá nhỏ do có tác dụng ngoại lệ.

Xóa các cột Tỉnh / Bang, Vĩ độ và Kinh độ vì chúng làm cho dữ liệu bị thu hẹp và dữ liệu cho các cột đó có thể bị thiếu đối với một số quốc gia.

Country/Region	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	1/28/20	1/29/20	1/30/20	1/31/20	2/1/20	2/2/20	2/3/20	2/4/20	2/5/20	2/6/20	2/7/20	2/8/20	2/9/20
0	Afghanistan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Albania	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Algeria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Andorra	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Angola	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	Antigua and Barbuda	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	Argentina	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	Armenia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	Australia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	Australia	0	0	0	0	3	4	4	4	4	4	4	4	4	4	4	4	4	4
10	Australia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Australia	0	0	0	0	0	0	1	3	2	3	2	2	3	3	4	5	5	5
12	Australia	0	0	0	0	0	0	0	0	0	1	2	2	2	2	2	2	2	2
13	Australia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	Australia	0	0	0	0	1	1	1	1	2	3	4	4	4	4	4	4	4	4
15	Australia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Hình 26. Tập dữ liệu time_series_covid_19_confirmed.csv sau tiền xử lý

Tiếp theo định hình lại dữ liệu theo yêu cầu, tức là tạo chỉ mục dưới dạng ngày tháng và tên cột là tên quốc gia

Để có được tổng số người bị nhiễm mỗi ngày trên thế giới, tổng hợp số người bị nhiễm từ tất cả các quốc gia và phân nhóm họ theo ngày

Trong cột quốc gia / khu vực, có dữ liệu trùng lặp, do đó, trước tiên cần tính tổng tất cả quốc gia và sau đó lấy số lượng tất cả những người bị nhiễm

11/11/20	91752
11/12/20	91783
11/13/20	91807
11/14/20	91828
11/15/20	91850

Hình 27. Gom dữ liệu theo ngày

Sau quá trình làm sạch và biến đổi dữ liệu, tập dữ liệu sau cùng sẽ có dạng như sau:

	Afghanistan	Albania	Algeria	Andorra	Angola	Antigua and Barbuda	Argentina	Armenia	Australia	Austria	Azerbaijan	Bahamas	Bahrain	Bangladesh	Barbados	Belarus	Belgium	Belize
11/11/20	42609	25801	64257	5567	12953	131	1273356	110548	27675	172380	68594	7060	84042	425353	249	109357	515391	4520
11/12/20	42795	26211	65108	5616	13053	131	1284519	112680	27676	181642	70216	7124	84192	427198	249	110455	520393	4596
11/13/20	42969	26701	65975	5725	13228	133	1296378	114383	27682	191228	71580	7163	84349	428965	249	111622	525012	4715
11/14/20	43035	27233	66819	5725	13374	134	1304846	115855	27711	198291	73429	7163	84523	430496	249	112870	531280	4783
11/15/20	43240	27830	67679	5872	13451	134	1310491	117337	27749	203956	75688	7186	84703	432333	250	114185	535939	4861

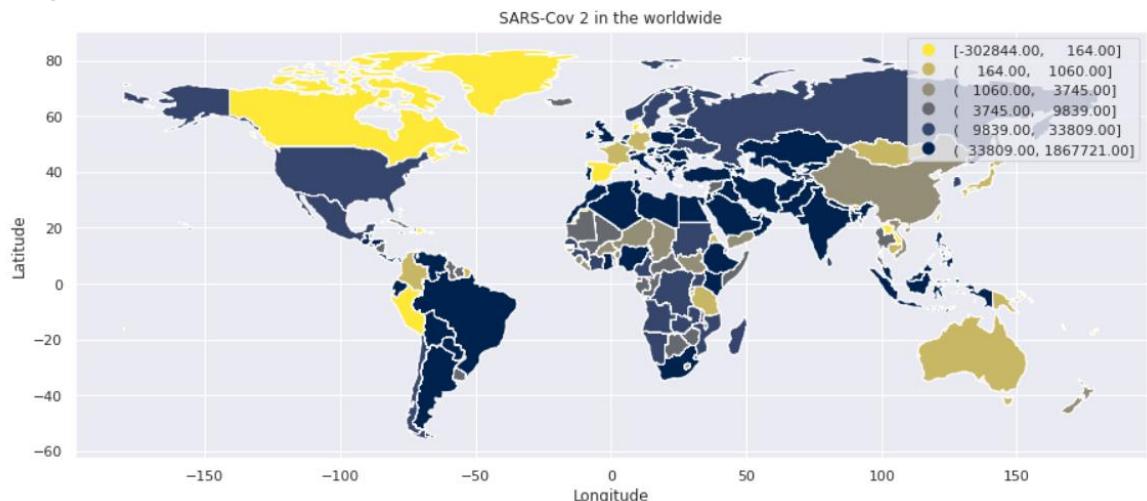
Hình 28. Tập dữ liệu time_series_covid_19_confirmed.csv sau khi tiền xử lý gom nhóm

IV. Nội dung:

1. Sự toàn cầu của SARS Cov 2

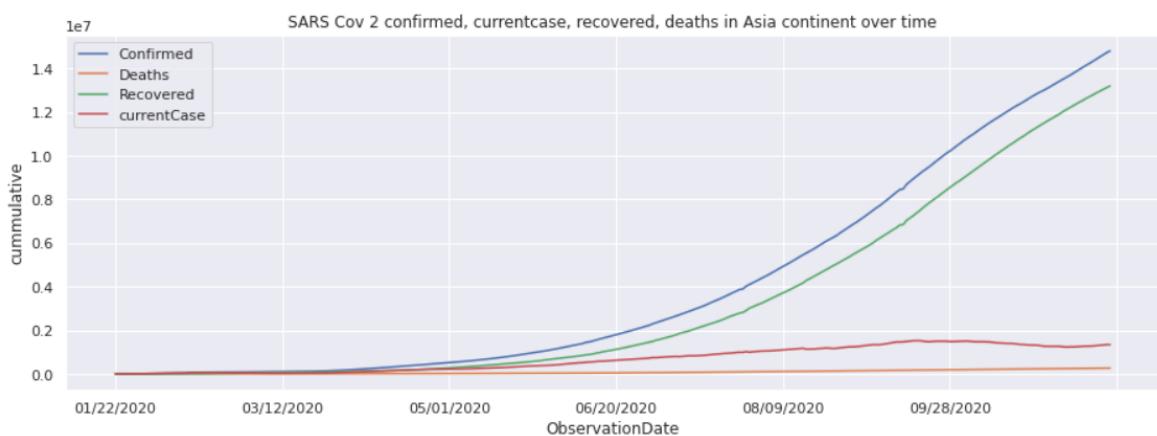
1.1. Bản đồ toàn cầu hóa COVID-19

Hiển thị không gian địa lý cho thấy sự lan rộng của Covid-19 trên toàn thế giới.



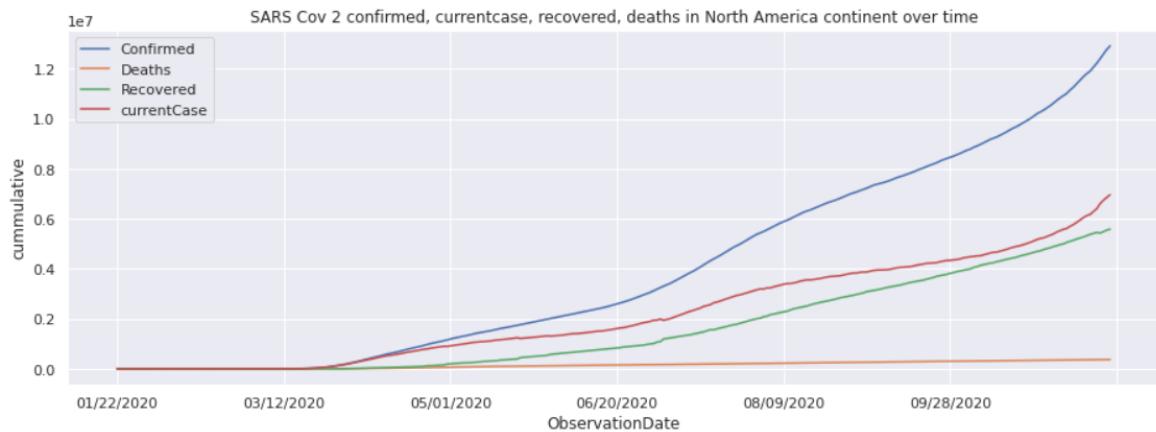
Hình 29. Sự lan rộng của covid 19 trên toàn thế giới

Theo đó, có thể thấy rằng phần lớn covid lan rộng và bắt đầu mạnh dần tại Trung Quốc, Tây Âu và Hoa Kỳ, điều này đã dẫn đến toàn bộ châu lục và nhiều quốc gia mới gặp nhiều khó khăn trong việc tiếp cận và nghiên cứu do tốc độ lâu lan của covid. Theo bản đồ cho thấy rằng covid đang hoành hành nặng nề ở các nước Tây Âu, Hoa Kỳ và Trung Quốc và có xu hướng ngày càng tăng. Điều này làm hạn chế qua trình nghiên cứu của các nhà khoa học muốn tiếp cận và tìm cách ngăn chặn chúng.



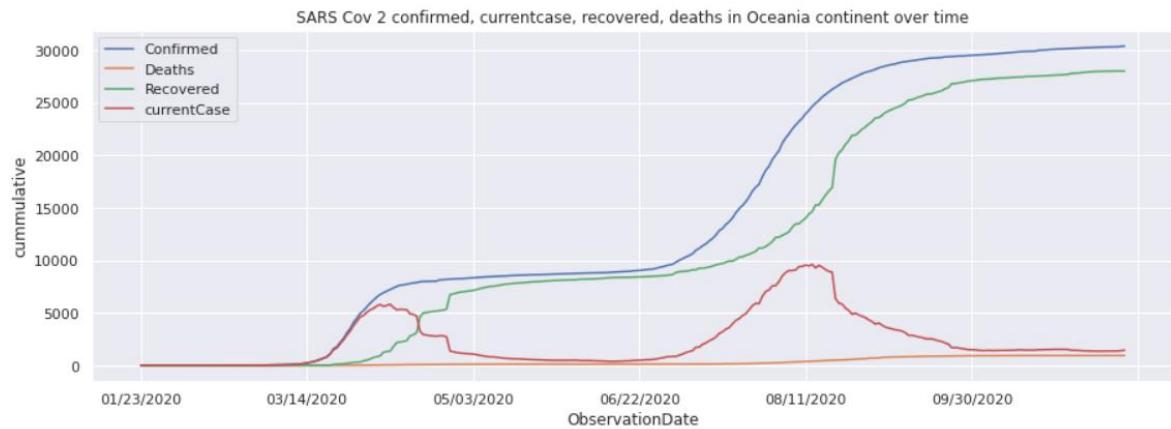
Hình 30. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước Asian

Ở các nước Asian bắt đầu phát tán covid nặng lên từ gần cuối tháng 05/1/2020 và số trường hợp xác nhận tăng lên nhanh chóng theo thời gian và luôn cao mặc dù tỉ lệ phục hồi vẫn tương đối cao.



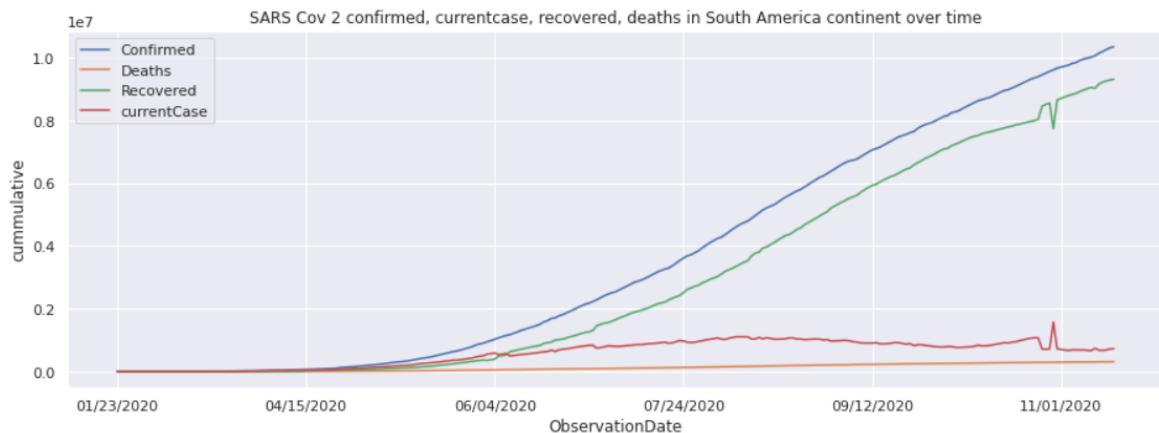
Hình 31. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước phía Bắc America

Ở các nước phía Bắc America có độ phát hiện bệnh với sự xác nhận bệnh nhanh hơn khoảng giữa đầu tháng 3 là đã phát hiện và cho thấy ở các nước này, tỉ lệ tử vong cao hơn so với tỉ lệ phục hồi và tỉ lệ phát hiện bệnh luôn cao hơn rất nhiều so với các tỉ lệ còn lại.



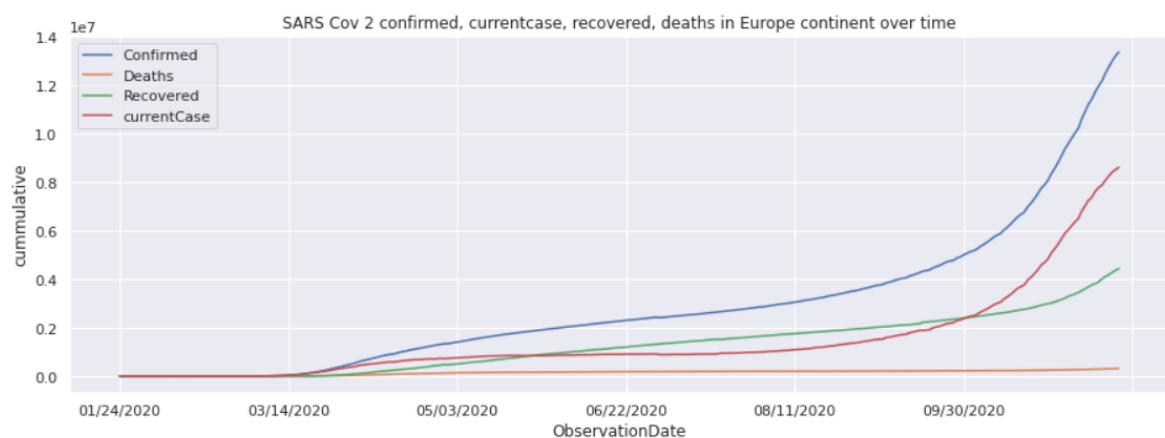
Hình 32. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước khu vực Oceania

Ở các nước khu vực Oceania có sự biến đổi bất thường ở tỷ lệ bệnh và phục hồi. Giữa ba tháng 3 và tháng 5 tỉ lệ bệnh hiện tại phát hiện nhanh trong khi tỉ lệ phục hồi lại rất thấp không đáng kể, nhưng từ tháng 5 trở lên ta thấy tỉ lệ phục hồi đã tăng cao cho thấy các nước ở khu vực này đã có sự phát triển trong việc y tế phòng chống. Tuy nhiên, việc đó đã gặp bất lợi khi vào khoảng tháng 8 tỉ lệ bệnh lại tăng cao thất thường và giảm xuống nhanh chóng



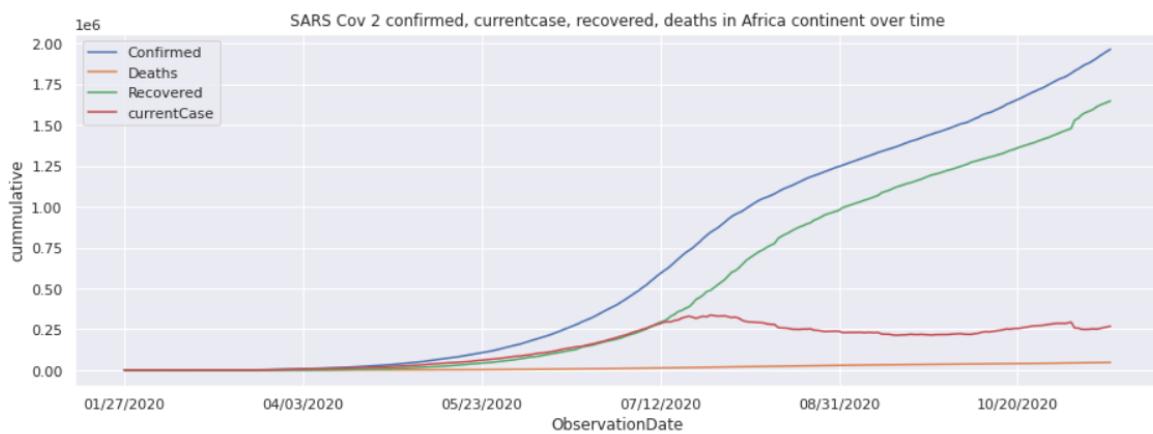
Hình 33. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước phía nam America

Nếu ở các nước trên có sự biến đổi đột ngột hay tỉ lệ bệnh và bình phục không có sự chênh lệch nhiều thì ở khu vực phía nam hầu như có sự chuẩn bị và chống bệnh tốt hơn khi tỉ lệ bình phục rất cao trong khi tỉ lệ bệnh rất thấp không đáng kể.



Hình 34. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở các nước Europe

Các nước Europe không có sự phát triển nhiều khi tỉ lệ bệnh và phục hồi đang xen nhau và bệnh luôn chiếm ưu thế hơn về tổng thể.



Hình 35. Biểu đồ thể hiện tỷ lệ tử vong, xác nhận và phục hồi ở khu vực Africa
Ở khu vực Africa có sự tiến bộ ở đầu tháng 7 về việc phòng chống và chữa bệnh khi phát hiện bệnh càng xuống thấp theo thời gian và phục hồi tăng nhanh.

1.2. Xác định $\beta(t)$, $\gamma(t)$, $\delta(t)$

Trước khi tìm $\beta(t)$, $\gamma(t)$, $\delta(t)$, cần xem tỷ lệ giữa số bệnh nhân được xác nhận ở một châu lục và quy mô dân số ở châu lục đó tại thời điểm t.

Số trường hợp dương tính trên mỗi quy mô dân số ở mỗi châu lục vào thời điểm hiện tại: 15/11/2020

Châu lục	Số trường hợp dương tính
Africa	0.001616
Asia	0.000149
Europe	0.000696
North America	0.000547
Oceania	0.000153
South America	0.001225

Bảng 5. Bảng số trường hợp dương tính

	ObservationDate	name	continent	Confirmed	Deaths	Recovered	currentCase
0	01/22/2020	China	Asia	547.0	17.0	28.0	502.0
1	01/22/2020	Japan	Asia	2.0	0.0	0.0	2.0
2	01/22/2020	South Korea	Asia	1.0	0.0	0.0	1.0
3	01/22/2020	Taiwan	Asia	1.0	0.0	0.0	1.0
4	01/22/2020	Thailand	Asia	2.0	0.0	0.0	2.0

Hình 36. Tốc độ tử vong, xác nhận và bình phục của các nước

1.3. Tốc độ truyền, tốc độ phục hồi, tỷ lệ tử vong trên toàn thế giới

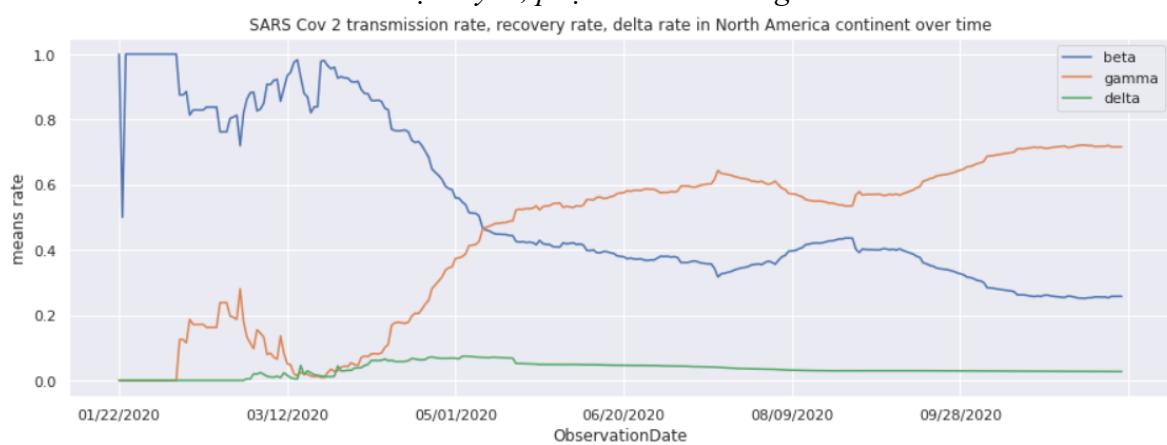
	ObservationDate	continent	beta	gamma	delta
0	01/22/2020	Asia	0.983547	0.010238	0.006216
1	01/22/2020	North America	1.000000	0.000000	0.000000
2	01/23/2020	Asia	0.740610	0.005869	0.003521
3	01/23/2020	North America	0.500000	0.000000	0.000000
4	01/23/2020	Oceania	0.000000	0.000000	0.000000

Hình 37. Beta, gamma, delta ở các nước

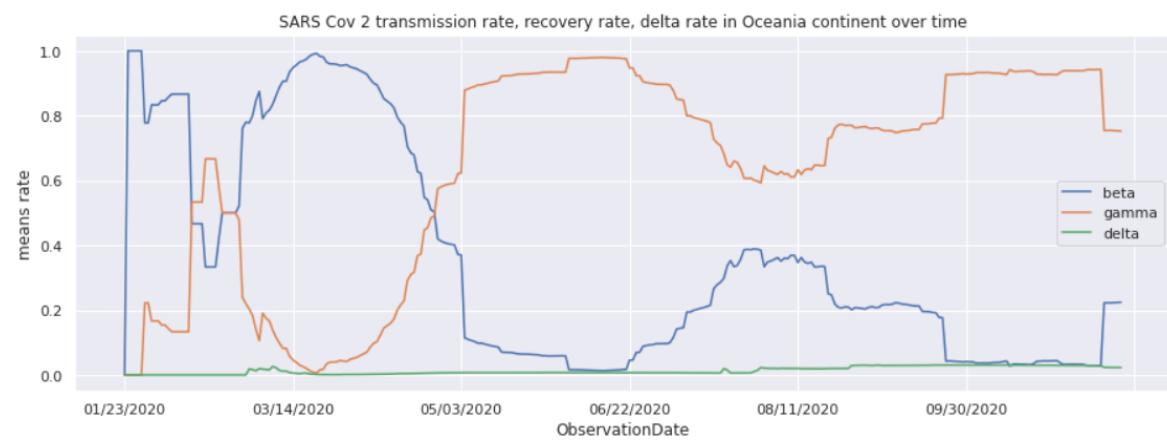
Xem các biểu đồ cho thấy tỉ lệ tử vong (delta) tương đối thấp và không có sự biến động nhiều ở hầu hết các nước và khu vực. Tuy nhiên tỉ lệ truyền bệnh (beta) và tỉ lệ phục hồi (gamma) có sự biến đổi theo thời gian nhưng đều có điểm chung: ở giai đoạn đầu tốc độ truyền bệnh tương đối nhanh và cao nhưng càng về sau thì tỉ lệ phục hồi bệnh lại cao và chiếm ưu thế hơn. Từ đó thấy rằng ở hầu như các nước không ít thì nhiều đã có sự đột phát trong lĩnh vực y tế phòng chống covid.



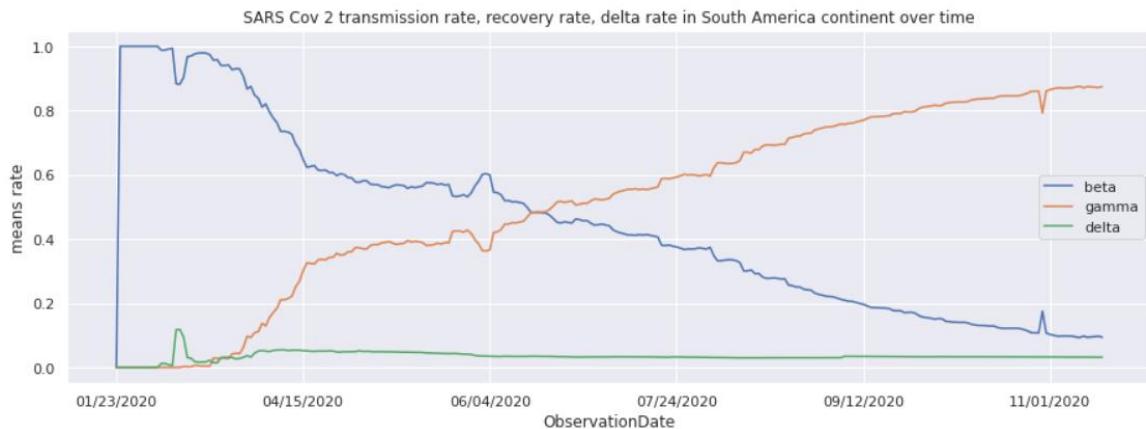
Hình 38. Tốc độ truyền, phục hồi và tử vong ở các nước Asian



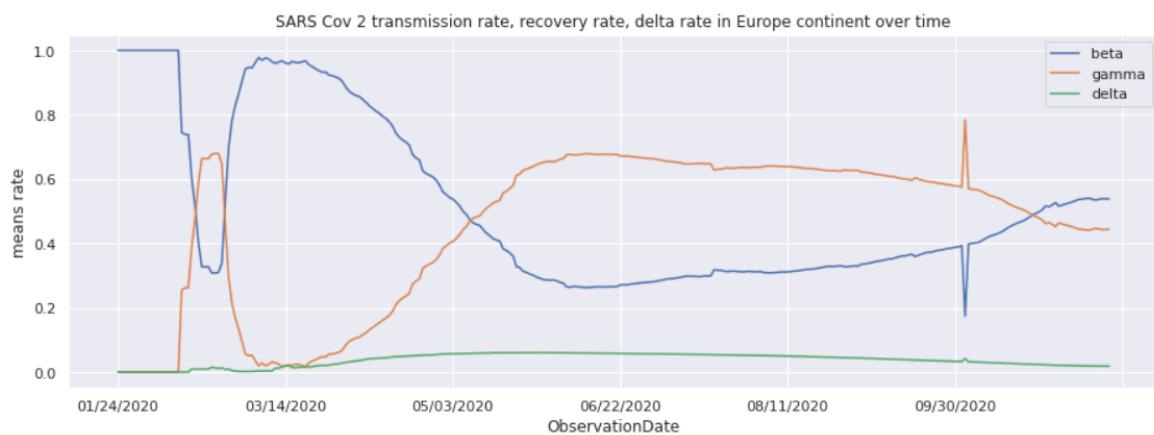
Hình 39. Tốc độ truyền, phục hồi và tử vong ở các nước phía bắc America



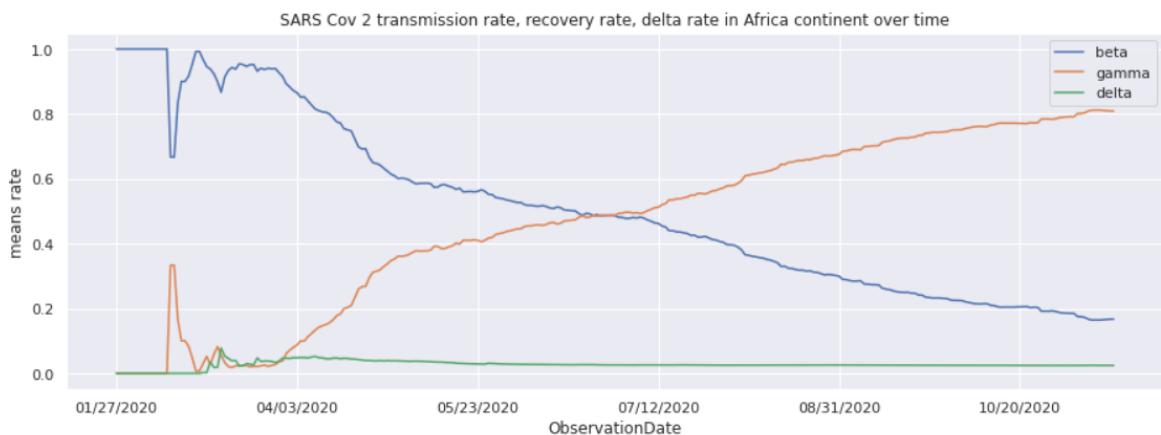
Hình 40. Tốc độ truyền, phục hồi và tử vong ở các nước Oceania



Hình 41. Tốc độ truyền, phục hồi và tử vong ở các nước phía nam America

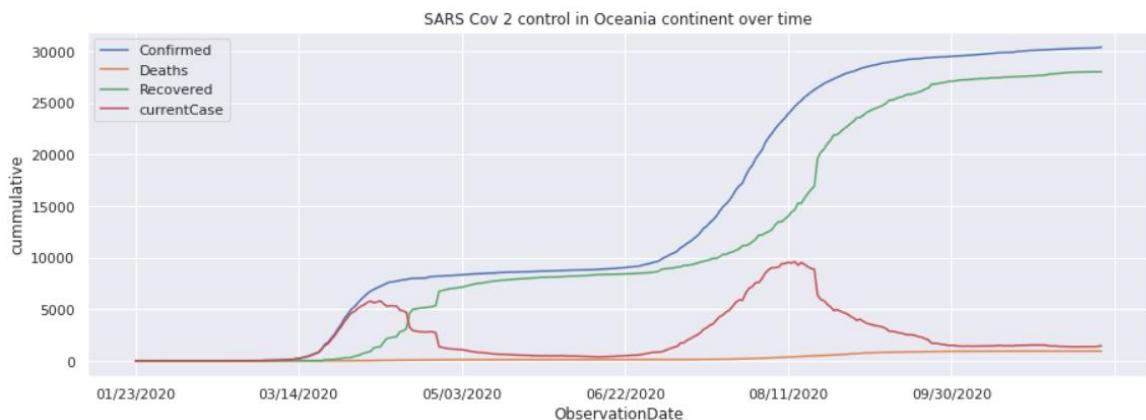


Hình 42. Tốc độ truyền, phục hồi và tử vong ở các nước Europe



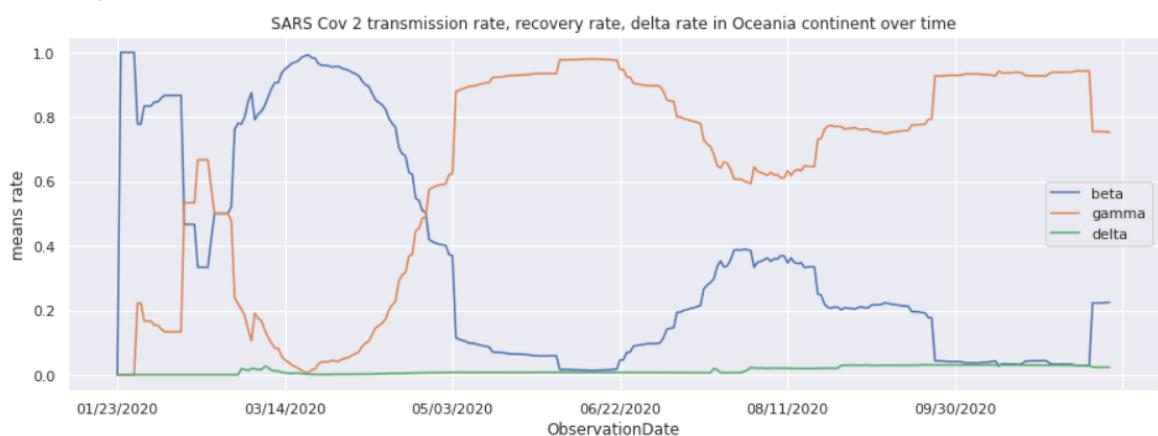
Hình 43. Tốc độ truyền, phục hồi và tử vong ở các nước Africa

1.4. Diễn giải việc kiểm soát dịch bệnh ở lục địa Châu Đại Dương



Hình 44. Diễn giải kiểm soát dịch bệnh ở lục địa Oceania

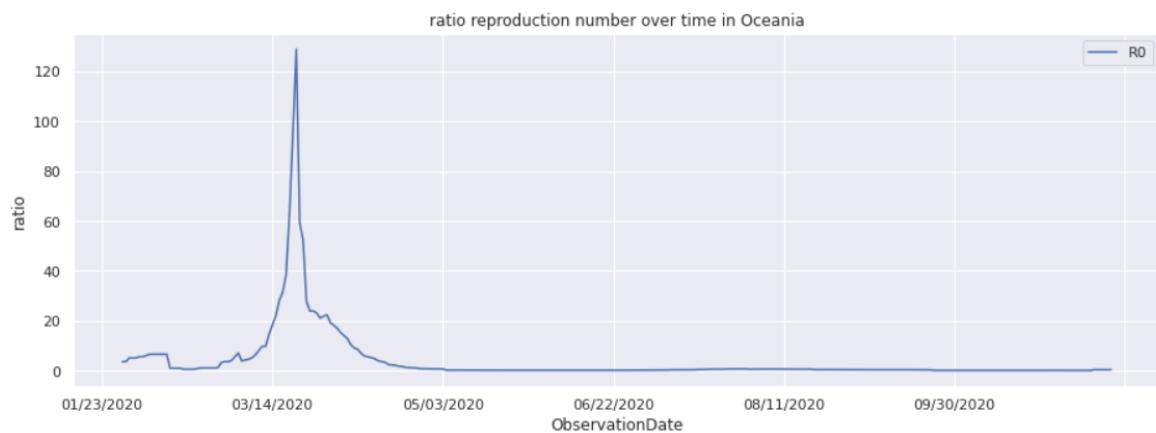
Hiện tại, Châu Đại Dương đang kiểm soát tốt dịch bệnh. Điều đó có nghĩa là tốc độ truyền giảm và tốc độ hồi phục cũng tăng tỷ lệ tử vong theo thời gian.



Hình 45. Tốc độ lan truyền, tử vong và phục hồi ở lục địa Oceania

Có thể thấy rằng nếu $\beta(t) - (\gamma(t) + \delta(t)) < 0$ thì dịch bệnh được kiểm soát nếu không sẽ bùng phát trở lại. Vì vậy, tính R_0 theo công thức $R_0(t) = \beta(t) / (\gamma(t) + \delta(t))$ để xem bệnh kiểm soát ở lục địa Châu Đại Dương.

1.5. Tìm R_0 theo thời gian

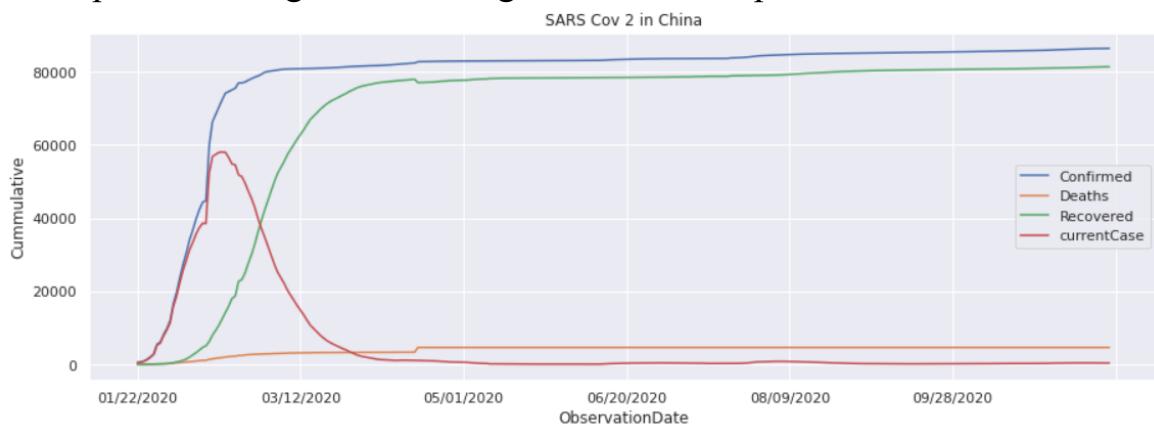


Hình 46. Chuyển động của R_0 theo thời gian ở Oceania

R_0 ở châu Đại Dương đã dần ổn định theo thời gian, điều này đã chứng minh được rằng châu Đại Dương kiểm soát dịch tốt như đã nói ở trên.

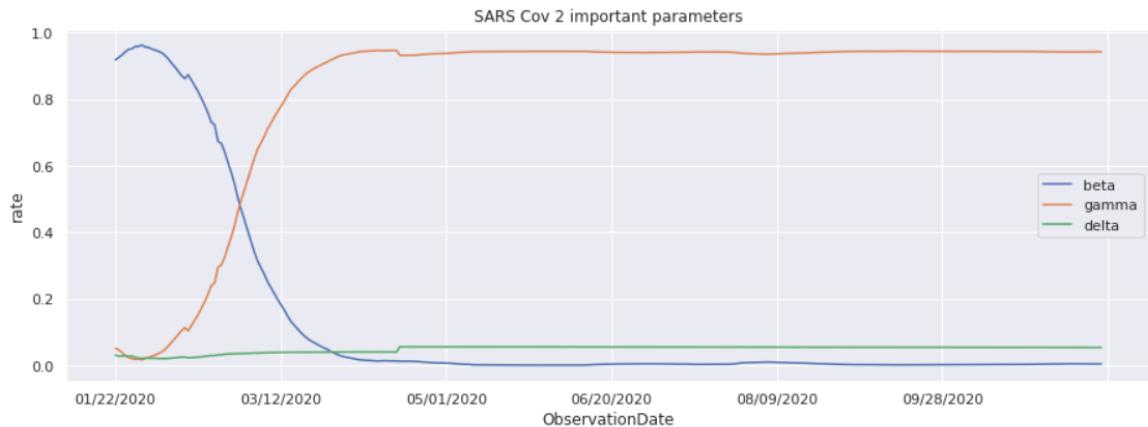
1.6. Nghiên cứu điển hình: Trung Quốc

Tiến hành xem xét ở quốc gia Trung Quốc về tốc độ truyền bệnh, tỉ lệ phục hồi cũng như khả năng nhiễm bệnh và phục hồi, ...

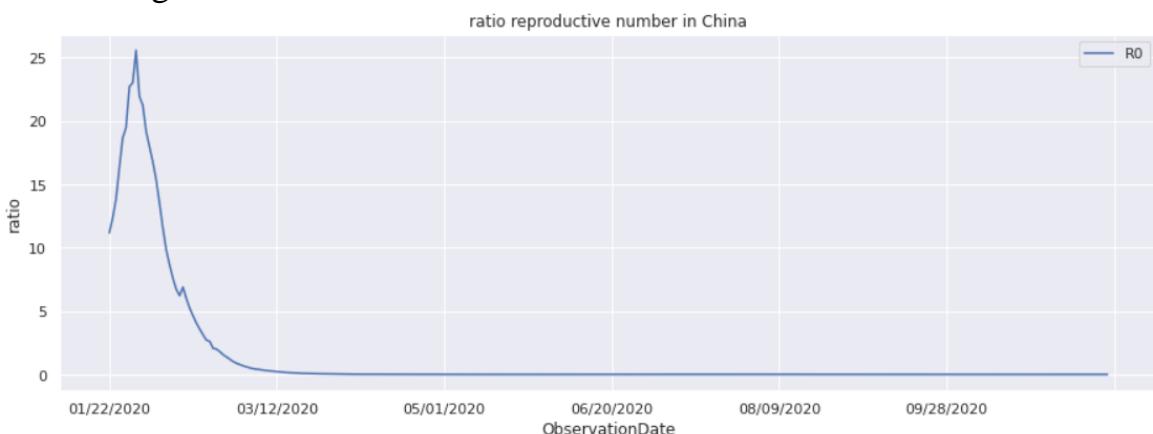


Hình 47. Diễn giải kiểm soát dịch bệnh ở Trung Quốc

Tỉ lệ tử vong và mắc bệnh hiện tại trong giai đoạn này rất thấp, chỉ cao ở giai đoạn tháng 1 đến tháng 3. Xem chung quy tỉ lệ này vẫn thấp hơn nhiều so với tỉ lệ phục hồi bệnh.



Hình 48. Tốc độ lan truyền, tử vong và phục hồi ở Trung Quốc
Ở biểu đồ tốc độ truyền bệnh cũng vậy, tỉ lệ phục hồi rất cao so với tốc độ truyền và tử vong, điều này là một bước tiến lớn trong công cuộc đẩy lùi covid 19 ở quốc gia này mặc dù vẫn còn gặp nhiều khó khăn trong công tác.



Hình 49. Chuyển động của R0 theo thời gian ở Trung Quốc
Như đã nói ở trên, tỉ lệ R0 cũng đã nói lên được Trung Quốc đã kiểm soát tốt được bệnh theo thời gian thực.

2. Sự ảnh hưởng của COVID-19:

Tiếp theo đề tài xem xét về độ tuổi dễ lây bệnh để hiểu rõ hơn về sự nhiễm bệnh và lây lan của covid 19.

Để phân tích được việc đó đề tài tiến hành trên tập dữ liệu PatientInfo.csv và Case.csv

2.1. Các trường hợp nhiễm bệnh nhiều nhất ở Hàn Quốc:

Tỉnh Daegu có nhiều trường hợp được xác nhận nhất. Rất dễ hiểu tại sao Daegu là tỉnh có nhiều trường hợp được xác nhận khi nó được mệnh danh là một trong hai ổ dịch lớn nhất của Hàn Quốc về covid 19. Số ca nhiễm COVID-19 đã tăng hơn gấp đôi chỉ trong vài ngày vừa qua, hầu hết các trường hợp dương tính được xác định đều có liên quan tới nhà thờ ở Daegu - thành phố lớn thứ tư của Hàn Quốc và tại Bệnh viện Daenam ở quận Cheongdo.

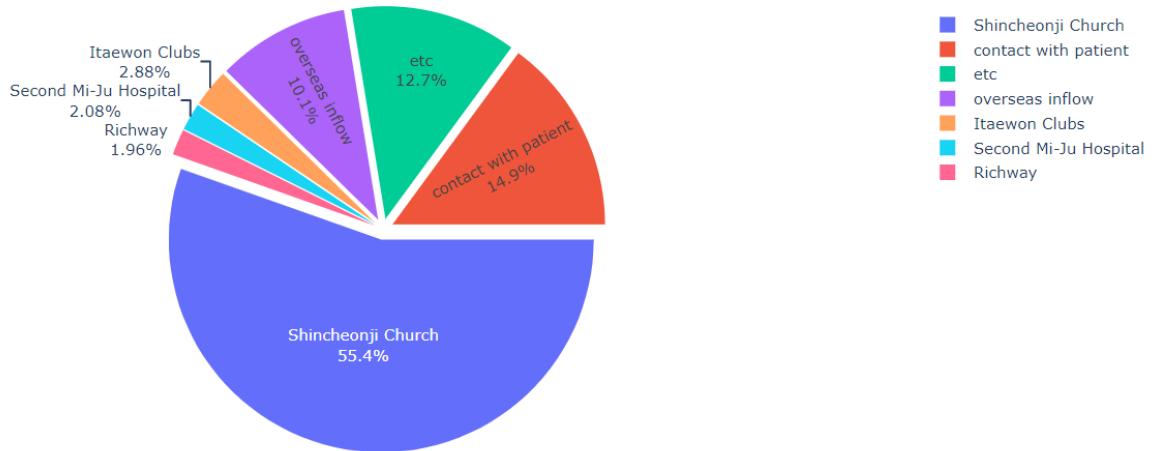


Hình 50. Các trường hợp xác nhận nhiều nhất ở Hàn Quốc
Tỉnh Jeju-do có ít trường hợp được xác nhận nhất.



Hình 51. Các trường hợp xác nhận ít nhất ở Hàn Quốc

Trường hợp nhiễm trùng Shincheonji Church (Tân Thiên Địa) có số trường hợp được xác nhận nhiều nhất với 5213 trường hợp được phát hiện.

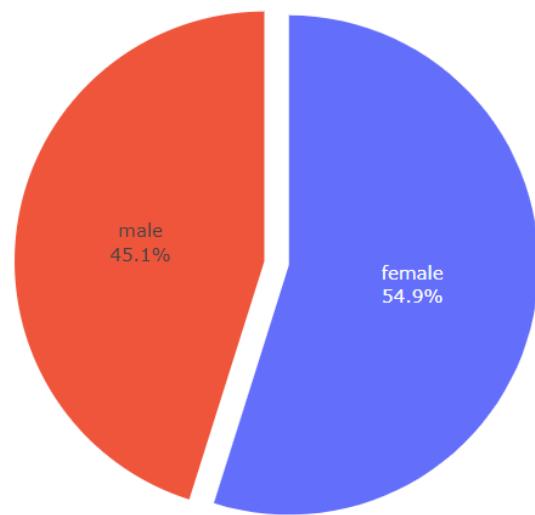


Hình 52. Các trường hợp nhiễm trùng nhiều nhất ở Hàn Quốc

2.2. Sự ảnh hưởng của COVID-19 đối với giới tính:

Nam giới ít bị nhiễm vi-rút Covid-19 hơn nữ giới với 1825 trên tổng số 2218 trường hợp. Trên thực tế, sự khác biệt về tỷ lệ tử vong do COVID-19 giữa nam và nữ là không đáng kể, tuy nhiên có thể thấy rằng ở nam giới khi nhiễm phải bệnh tỉ lệ phục hồi sẽ tùy thuộc vào đội tuổi do tốc độ đào thải của kháng thể không ổn định ở mọi lứa tuổi. Trong khi đó nữ giới lại duy trì được khả năng này ở hầu hết các lứa tuổi và khả năng nhiễm covid “nặng” rất ít so với nam giới.

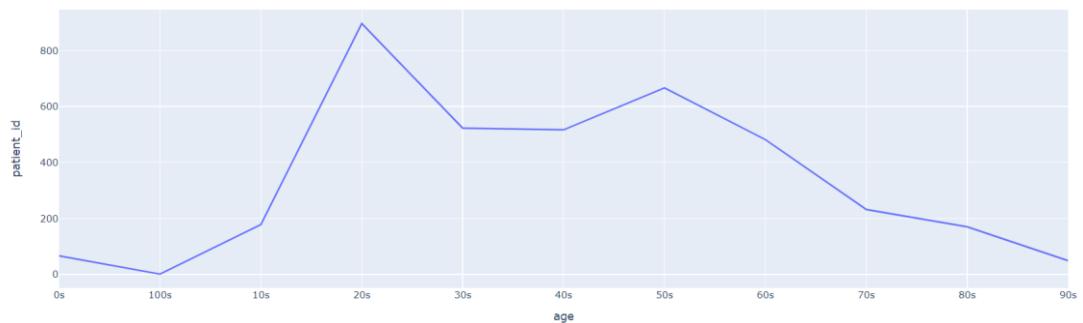
Chú ý rằng ở thời điểm này số bệnh nhân nam giới và nữ giới không đồng đều nhau, điều này cũng cho thấy được nữ giới có khả năng chống bệnh tốt hơn nam giới như đã nói ở trên.



Hình 53. Ảnh hưởng của covid đối với giới tính

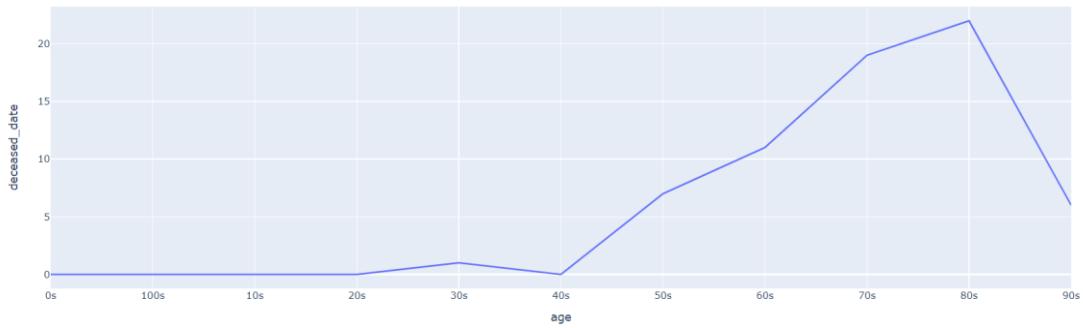
2.3. Sự ảnh hưởng của COVID-19 đối với các nhóm tuổi:

- Theo dữ liệu ta thấy nhóm người trong độ tuổi khoảng 20 tuổi đến 50 tuổi dễ bị nhiễm vi-rút Covid-19 nhất, đặc biệt là ở độ tuổi 50 này họ thường đã có bệnh tim ẩn trong người nên càng dễ mắc phải bệnh hơn.



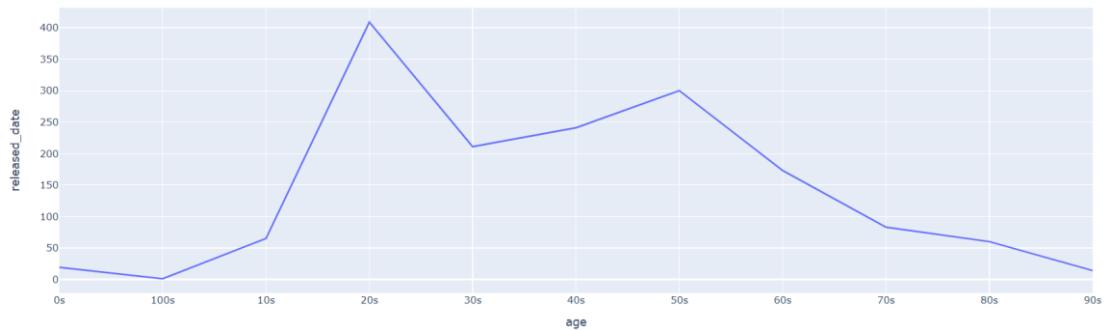
Hình 54. Nhóm tuổi dễ nhiễm bệnh

- Dễ dàng thấy được nhóm người cao tuổi (70 tuổi – 90 tuổi) có tỷ lệ tử vong cao nhất. Trường hợp này dễ dàng hiểu được khi trên thực tế, những người cao tuổi một phần do sức khỏe không tốt hơn lớp trẻ và họ cũng đã mang một số bệnh và bệnh tiềm ẩn trong người, vì thế không thể nào chống lại được covid.



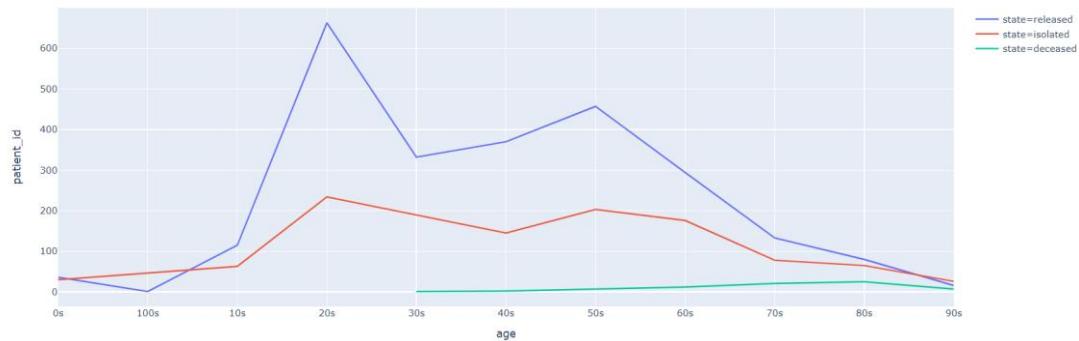
Hình 55. Nhóm tuổi dễ tử vong

- Nếu nói nhóm tuổi từ 20 – 50 tuổi dễ mắc bệnh nhất thì đây cũng là độ tuổi có khả năng phục hồi nhanh nhất cũng với lứa tuổi từ 20 trở xuống. Dễ dàng thấy được khi trên thực tế số người chết do covid rất ít thấy ở độ tuổi này và trẻ em đa phần là có biểu hiện nhẹ và được chữa trị nhanh chóng. Nghĩ ở phương diện y học, nhóm tuổi này là nhóm tuổi có sức đề kháng nguồn cao, đặc biệt là trẻ em, vì thế họ có thể chống lại covid 19 rất hiệu quả.



Hình 56. Nhóm tuổi có khả năng dễ phục hồi

- Cuối cùng, mô hình về tình trạng bệnh và độ tuổi của họ cho thấy rằng số người được phát hiện (released) và cách ly (isolated) ở hầu hết mọi lứa tuổi và tập trung cao ở độ tuổi 20 – 50 tuổi. Tuy nhiên, số người tử vong do covid 19 chỉ nằm ở khoảng 30 tuổi trở lên và hầu như tập trung ở 80 – 90 tuổi.



Hình 57. Tương quan giữa tình trạng bệnh và nhóm tuổi

3. Mối liên quan giữa tốc độ tăng trưởng, R0 (t), lưới ngăn chặn và giảm thiểu.

Mối tương quan giữa lưới giảm thiểu, ngăn chặn và tốc độ tăng trưởng, R0 (t). Để làm được điều này, đề tài tiến hành quan sát chiến lược của từng quốc gia (Trung Quốc, Mỹ, Đức, Hàn Quốc, Ý, Pháp ...) để chống lại covid19 và tác động của chiến lược này đến tốc độ tăng trưởng và R0 (t). Cân ba tốc độ tăng trưởng R0 (t).

- Tỷ lệ tăng trưởng trường hợp tích cực
- Tốc độ tăng trưởng phục hồi
- Tốc độ tăng trưởng tử vong

Quan hệ giữa tỷ lệ tăng trưởng trường hợp dương hiện tại với R0 (t)

Tỷ lệ tiếp xúc và tỷ lệ mắc bệnh thích hợp, Mô hình SIRF với tỷ lệ mắc bệnh chuẩn được điều chỉnh.

3.1. Các biện pháp ngăn chặn và giảm thiểu của Trung Quốc

Trước:

Country	Date Start	Description of measure implemented
150 China	2019-12-18 00:00:00	Seven cases of COVID-19 documented in Wuhan
151 China	2019-12-25 00:00:00	Medical staff in two Wuhan hospitals were suspected of contracting the virus; first reports of isolation being used
152 China	2019-12-30 00:00:00	Wuhan CDC confirmed 27 cases of viral pneumonia of unknown cause; all hospitals required to report all cases of unknown origin and related to Huanan Seafood Market and give patients proper treatment;
153 China	2020-01-01 00:00:00	Huanan Seafood Market closed by Jianghan District's Health Agency and Administration for Market Regulation for 'environment improvement'; workers in hazmat suits inspected market and collected samples
157 China	2020-01-03 00:00:00	Research on virus commenced; 44 cases reported, 120 close contacts started being monitored
166 China	2020-01-08 00:00:00	Scientists announce new coronavirus
205 China	2020-01-10 00:00:00	First two patients outside Wuhan admitted to hospital in Shenzhen
206 China	2020-01-12 00:00:00	700 close contacts of initial 41 infections closely monitored with no new infections; WHO published initial guidance on travel, lab testing and medical investigation
154 China	2020-01-15 00:00:00	Wuhan Municipal Health Commission stated on its website that there was no clear indication of human to human transmission but did not rule out the possibility, although maintained that the risk of human to human transmission was low; total case number 41
156 China	2020-01-20 00:00:00	Case numbers rose to 136; Beijing and Shenzhen reported first cases
155 China	2020-01-20 00:00:00	city formed specialized command for epidemic control (CEC) to upgrade measures to cope with the epidemic including enhanced protection over the medical workers;

Hình 58. Các biện pháp ngăn chặn và giảm thiểu của Trung Quốc trước dịch bệnh

Trước khi các biện pháp được ban hành, thấy rằng Trung Quốc đã có các trường hợp và liên tục tăng lên:

- Bảy trường hợp COVID-19 được ghi nhận ở Vũ Hán
- Nhân viên y tế ở hai bệnh viện ở Vũ Hán bị nghi nhiễm virus; báo cáo đầu tiên về sự cô lập được sử dụng
- Vũ Hán CDC xác nhận 27 trường hợp viêm phổi do virus không rõ nguyên nhân; tất cả các bệnh viện được yêu cầu báo cáo tất cả các trường hợp không rõ nguồn gốc và liên quan đến Chợ Hải sản Huanan và điều trị thích hợp cho bệnh nhân
- Nghiên cứu về vi rút bắt đầu; 44 trường hợp được báo cáo, 120 liên hệ thân thiết bắt đầu được theo dõi
- Các nhà khoa học công bố coronavirus mới
-

Sau khi các trường hợp lần lượt xuất hiện và ngày càng tăng lên, Trung Quốc đã lần lượt đưa ra các lệnh khẩn để phòng chống và ngăn chặn covid lây lan đối với trong nước và khách du lịch:

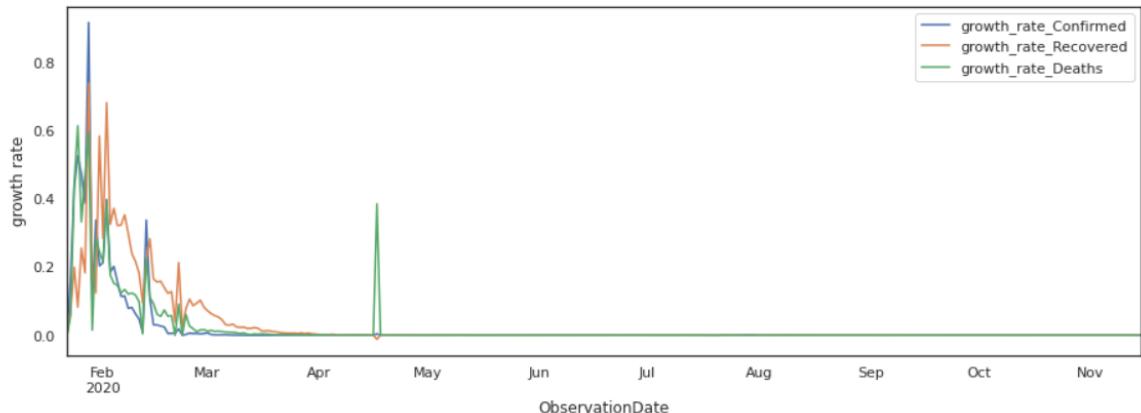
- Vũ Hán yêu cầu cách ly 14 ngày đối với những bệnh nhân xuất viện từ COVID-19.
- Ước tính này là một ước tính gần đúng và được đề cập trong trang 9 của báo cáo. Không có ngày chính xác cho thời gian thử nghiệm được cung cấp.
- Để đối phó với tình hình dịch bệnh ở Hàn Quốc và nhu cầu mua vé máy bay từ Hàn Quốc đến các thành phố Sơn Đông gia tăng đáng kể, Qingdao hiện sẽ cách ly tất cả những người có tiền sử đi lại các khu vực có dịch, Uy Hải hiện sẽ cách ly tất cả những người nhập cảnh từ Nhật Bản và Hàn Quốc trong 14 ngày, Thẩm Dương hiện sẽ kiểm tra nhiệt độ của tất cả hành khách đi vào từ thành phố và Đại Liên hiện sẽ đưa tất cả người nước ngoài vào cơ chế kiểm soát dịch bệnh của nó.
- Bắc Kinh yêu cầu du khách nước ngoài đến từ các khu vực bị ảnh hưởng bởi COVID-19 phải được cách ly trong 14 ngày.
- Giờ đây, Quảng Châu sẽ kiểm tra SARS-CoV-19 của tất cả các bệnh nhân nội trú.

- Yên Đài, Sơn Đông hiện sẽ kiểm tra miễn phí tất cả những người nhập cảnh vào Trung Quốc từ thành phố này để tìm SARS-CoV-2.

Country	Date Start	Description of measure implemented
258 China	2020-02-22 00:00:00	Wuhan requires a 14-day quarantine for patients discharged from COVID-19.
1189 China	2020-02-24 00:00:00	This estimate is an approximation and is mentioned in page 9 of the report. No exact date for the testing period is provided.
259 China	2020-02-25 00:00:00	In response the epidemic situation in South Korea and an apparent significant surge of demand for flight tickets from South Korea to Shandong cities, Qingdao will now quarantine all persons entering with a travel history of areas with the epidemic.[692] Weihai will now quarantine all persons entering from Japan and South Korea for 14 days.[693] Shenyang will now test the temperature of all passengers entering from the city.[694] and Dalian will now include all foreigners into its epidemic control mechanism.
263 China	2020-02-25 00:00:00	Beijing required foreign travellers arriving from areas affected by COVID-19 to be quarantined for 14 days.
261 China	2020-02-25 00:00:00	Guangzhou will now test all inpatients for SARS-CoV-19.
264 China	2020-02-26 00:00:00	Yantai, Shandong will now test all persons entering China from the city for SARS-CoV-2 for free.

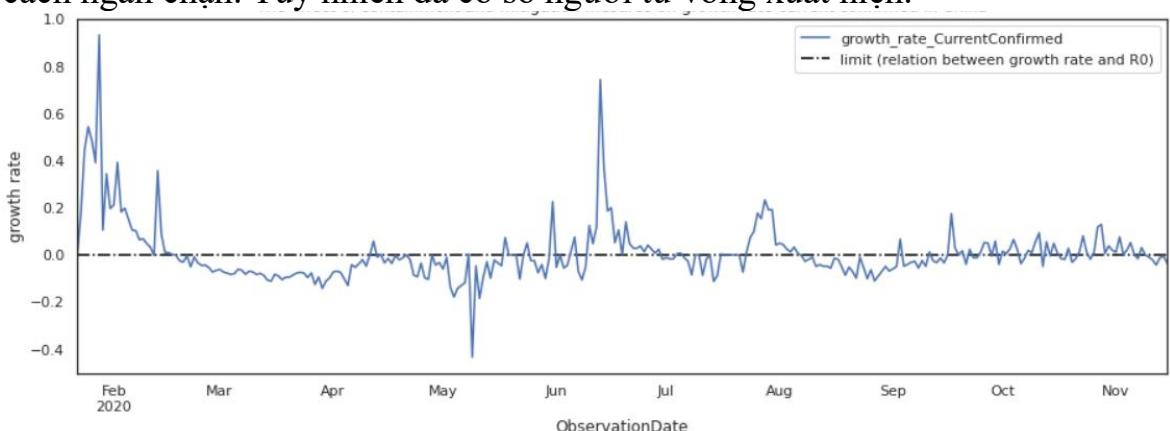
Hình 59. Các biện pháp ngăn chặn và giảm thiểu của Trung Quốc sau dịch bệnh

3.2. Tỉ lệ tăng trưởng



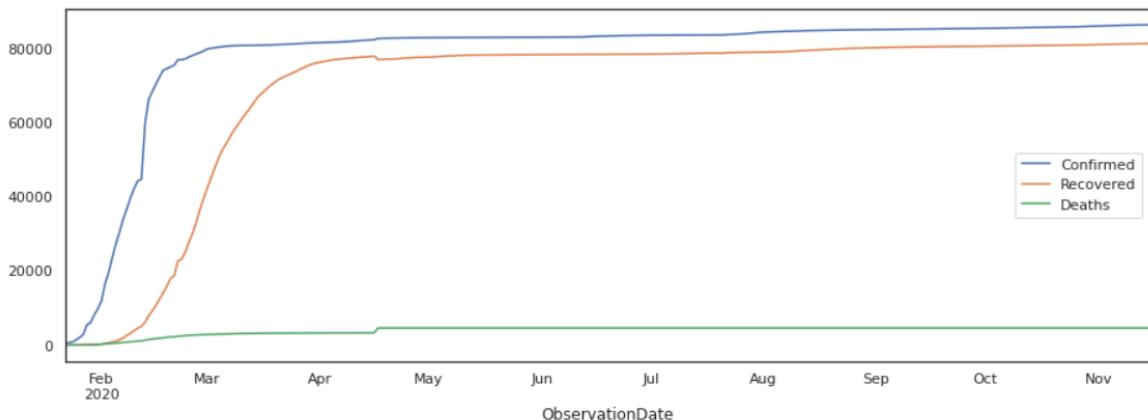
Hình 60. Tình hình dịch bệnh ở Trung Quốc giai đoạn đầu

Tình hình dịch bệnh ở Trung Quốc trong giai đoạn này thực sự rất bất ổn, tình hình phát hiện bệnh, phục hồi luôn biến động không rõ ràng, có thể thấy ở giai đoạn đầu này mọi thứ về covid vẫn còn mơ hồ nên chưa có sự đột phá trong cách ngăn chặn. Tuy nhiên đã có số người tử vong xuất hiện.



Hình 61. Hiệu quả của việc phòng chống đối với dịch bệnh ở Trung Quốc

Các biện pháp tuy không có đột phá nhiều nhưng cũng có tác động nhất định trong việc phòng chống và tạo lối ngăn chặn covid 19 qua biểu đồ tác động của các biện pháp ngăn chặn và giảm thiểu đối với tốc độ tăng trưởng hiện nay đã được khẳng định ở Trung Quốc khi số trường hợp xác nhận covid hiện tại đã giảm xuống, không cao như lúc đầu.



Hình 62. Tỷ lệ tử vong, phục hồi và xác nhận ở Trung Quốc sau khi áp dụng lối ngăn chặn

Cuối cùng, qua sự phân tích và những biện pháp phòng chống mà Trung Quốc đưa ra đã cho thấy có sự hy vọng rõ rệt cho công tác chống covid sau này:

- Tình trạng phục hồi đã cao rất nhiều gần như có thể vượt qua tình hình bệnh phát hiện đang nghỉ ngơi

Tình trạng người tử vong giảm xuống rất nhiều và có sự bình ổn, không còn bất ổn nữa.

D. NHIỆM VỤ 4: SỐ TRƯỜNG HỢP NHIỄM BỆNH, TỔNG SỐ NGƯỜI CHẾT VÀ ĐẠI DỊCH PHÁT TRIỂN Ở CÁC QUỐC GIA TRÊN THẾ GIỚI

I. Yêu cầu của nhiệm vụ:

- Top 10 quốc gia có số ca tử vong được xác nhận, hồi phục và tử vong trung bình
- Phân phối các trường hợp đã xác nhận
- Các triệu chứng
- Tốc độ đại dịch covid 19 phát triển trên thế giới
- Tỷ lệ hồi phục và tử vong trên toàn cầu và ở Trung Quốc
- Sự tương quan giữa thời gian kể từ trường hợp đầu tiên và sự phục hồi
- Tuổi tác
- Dự đoán số người chết theo thời gian trên thế giới

II. Sơ lược về các tập dữ liệu:

1. Sơ lược các tập dữ liệu:

Nhiệm vụ được tiến hành trên các tập dữ liệu *covid_19_data.csv*, *COVID19_open_line_list.csv*, *COVID19_line_list_data.csv* và *time_series_covid_19_confirmed.csv*.

2. Tiền xử lý tập dữ liệu:

Các tập dữ liệu được sử dụng không thông qua bước tiền xử lý dữ liệu

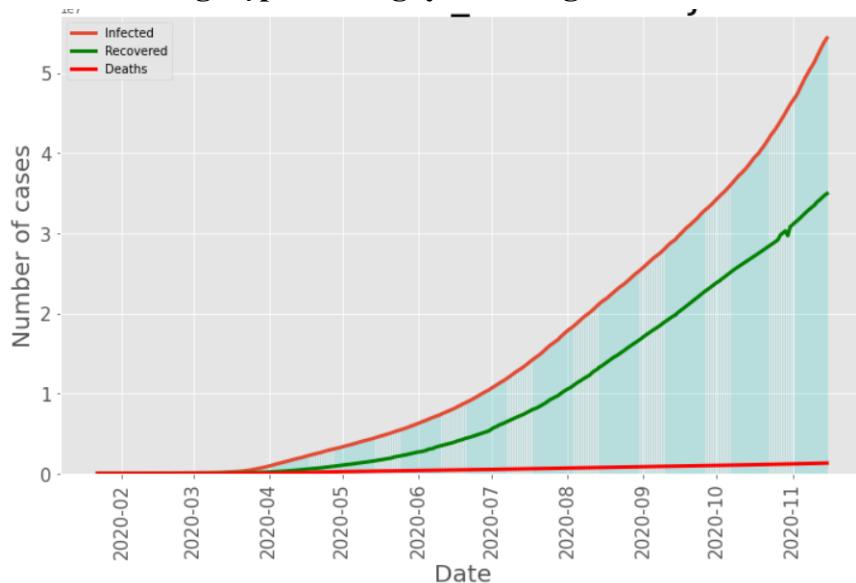
III. Nội dung:

1. Số trường hợp nhiễm bệnh:

Bệnh Coronavirus 2019 (COVID-19) là một bệnh truyền nhiễm do chủng vi rút "coronavirus 2 gây hội chứng hô hấp cấp tính nghiêm trọng" (SARS-CoV-2). Căn bệnh này được xác định lần đầu tiên vào năm 2019 tại Vũ Hán, Trung Quốc và kể từ đó đã lây lan toàn cầu, dẫn đến đại dịch coronavirus 2019–2020. Bệnh thường lây từ người này sang người khác qua các giọt đường hô hấp tạo ra khi ho và hắt hơi. Thời gian từ khi tiếp xúc đến khi bắt đầu các triệu chứng thường từ 2 đến 14 ngày, trung bình là 5 ngày.

Tổ chức Y tế Thế giới (WHO) tuyên bố virus coronavirus 2019–20 bùng phát là một đại dịch và là Tình trạng khẩn cấp về sức khỏe cộng đồng cần quan tâm của quốc tế (PHEIC).

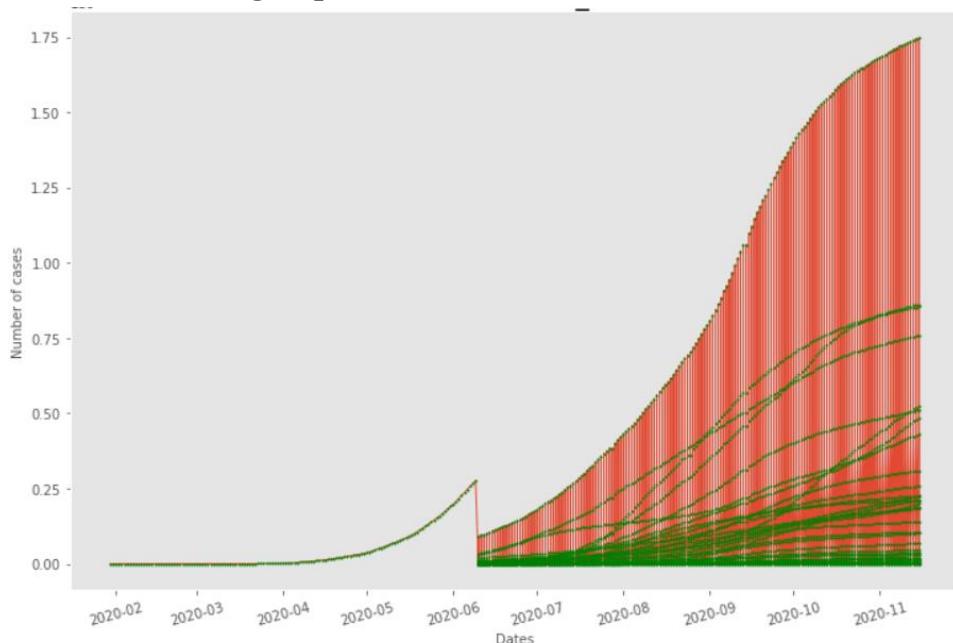
1.1. Các trường hợp kể từ ngày 22 tháng 1 năm 2020



Hình 63. Các trường hợp nhiễm bệnh từ ngày 22/1/2020

Kể từ ngày 22 tháng 1 năm 2020 sau khi đại dịch bùng phát cho thấy tốc độ lây nhiễm của virus không có dấu hiệu giảm, luôn tăng và tăng một cách đột biến ngày càng cáo vượt xa tốc độ phục hồi theo thời gian mặc dù lúc đầu (tháng 4 năm 2020) tốc độ lây nhiễm và bình phục hầu như không thay đổi nhiều – không xa nhau. Tuy nhiên, từ tháng 7 tốc độ lây nhiễm lan truyền rất nhanh, vượt xa tốc độ phục hồi. Tỉ lệ người chết trong thời gian đầu này không nhiều, không đáng kể so với tỷ lệ phục hồi.

1.2. Các trường hợp ở An Độ

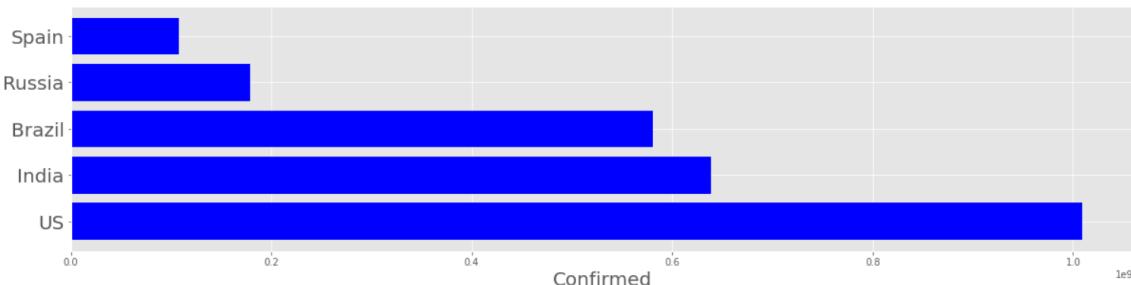


Hình 64. Các trường hợp nhiễm bệnh ở Ấn Độ

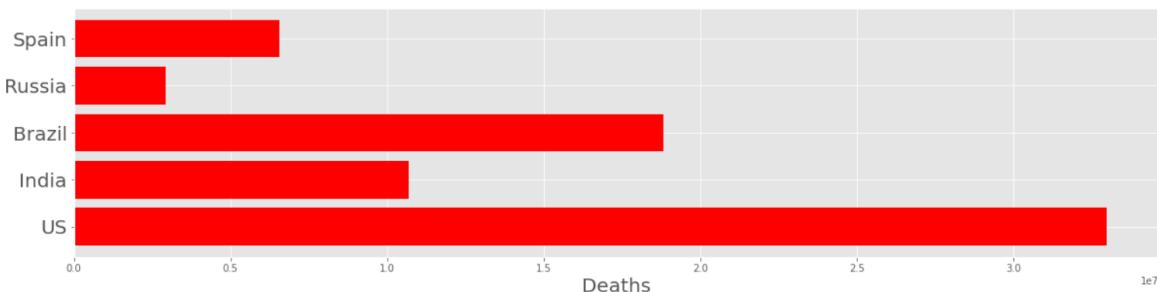
Trường hợp đầu tiên của đại dịch coronavirus ở Ấn Độ được báo cáo vào ngày 30 tháng 1 năm 2020, bắt nguồn từ Trung Quốc.

Sau ngày 3 tháng 3 năm 2020, các trường hợp nhiễm bệnh tăng dần ở Ấn Độ. Tính đến ngày 20 tháng 3 năm 2020, Bộ Y tế và Phúc lợi Gia đình đã xác nhận tổng số 223 trường hợp mắc và 4 trường hợp tử vong trong cả nước.

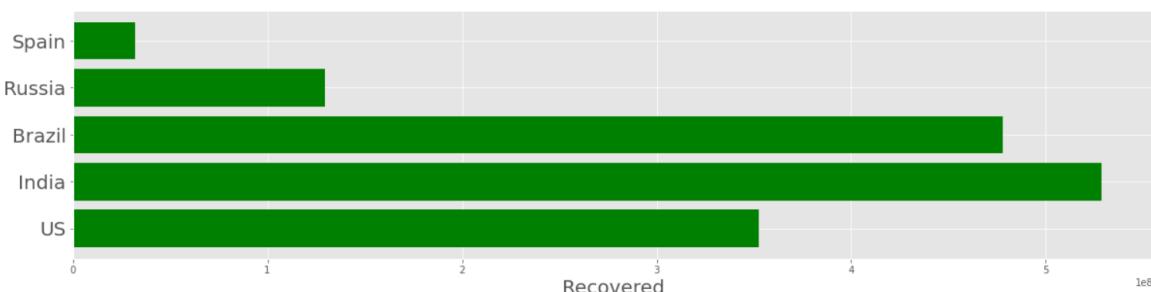
1.3. Top 5 quốc gia có số ca được xác nhận, phục hồi và tử vong cao nhất:



Hình 65. Top 5 quốc gia có số ca đã xác nhận cao nhất



Hình 66. Top 5 quốc gia có số ca tử vong cao nhất



Hình 67. Top 5 quốc gia có số ca phục hồi cao nhất

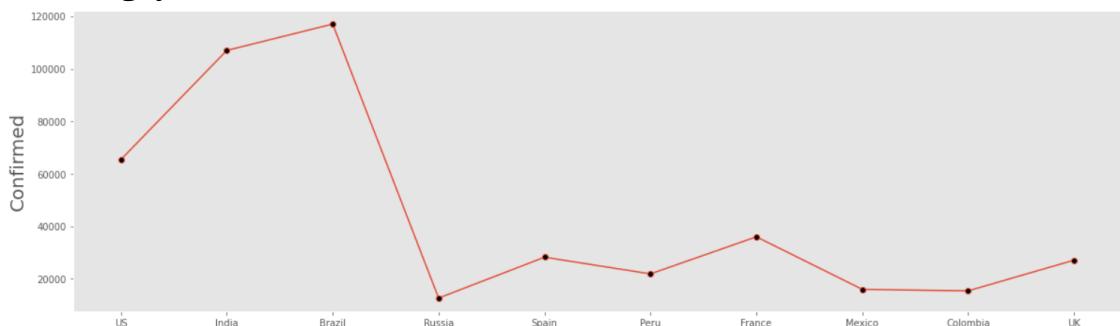
Trong thời điểm hiện tại, Hoa Kỳ là nước đứng đầu về quốc gia có số ca tử vong được xác nhận và tử vong cao nhất trong top 5 quốc gia hàng đầu về đại dịch covid. Điều này dễ nhận biết được khi Hoa Kỳ đã được xác nhận ngày 20 tháng 1 năm 2020, ở một người đàn ông 35 tuổi đã trở về từ Vũ Hán, Trung Quốc năm ngày trước đó [11], và gần đây nhất Hoa Kỳ đã lập kỷ lục khi số nạn nhân thiệt mạng vì dịch Covid-19 trong 24 tiếng đồng hồ đã lại vượt

kỷ lục vào hôm qua 16/12/2020 với trên 3.700 trường hợp, trong lúc số người bị nhiễm mới cũng vượt quá 250 000 ca theo số liệu của Đại Học Johns Hopkins. Đây là lần thứ ba trong một tuần lễ mà số ca tử vong vượt quá mức 3000 người chết [12].

Ấn Độ là nước có tỉ lệ phục hồi cao nhất trong 5 quốc gia đứng hàng đầu về covid khi và tỉ lệ người chết cũng không đứng trong top 2 thế giới và có số người chết ít nhất là nước Nga.

1.4. Top 10 quốc gia có số ca được xác nhận, hồi phục và tử vong trung bình

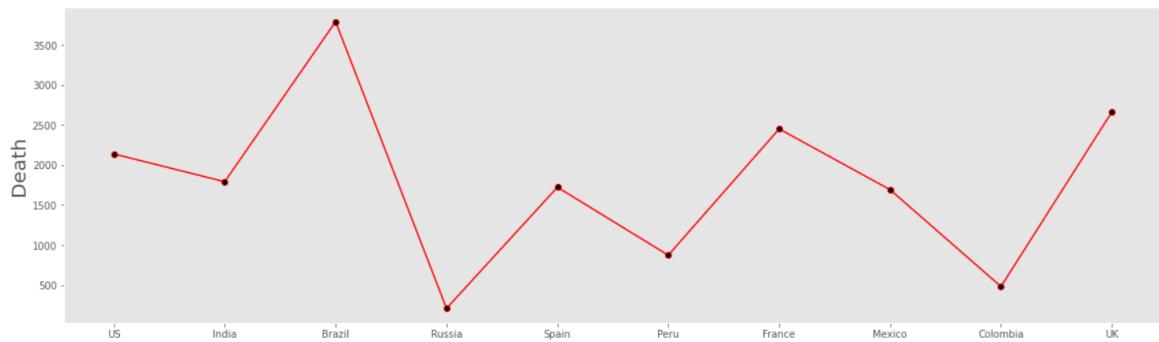
Brazil là nước đứng đầu về số ca tử vong, phục hồi và được xác nhận trong 10 quốc gia ở khoảng trung bình. Khá dễ nhận định được điều này khi WHO đã nhận định rằng: “**Các ca mắc Covid-19 tại Brazil không còn ở cấp số nhân nhưng quốc gia Nam Mỹ này đang ở “tâm của cuộc chiến” khi số ca mắc mới và số người thiệt mạng vẫn tăng cấp độ hàng nghìn mỗi ngày.**”



Hình 68. Top 10 quốc gia có số ca được xác nhận trung bình



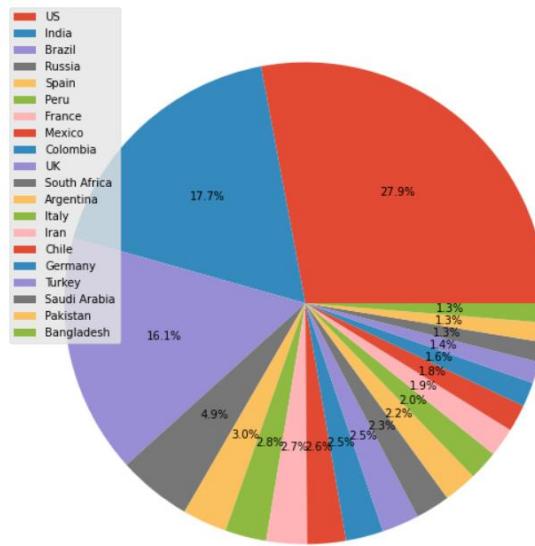
Hình 69. Top 10 quốc gia có số ca phục hồi trung bình



Hình 70. Top 10 quốc gia có số ca tử vong trung bình

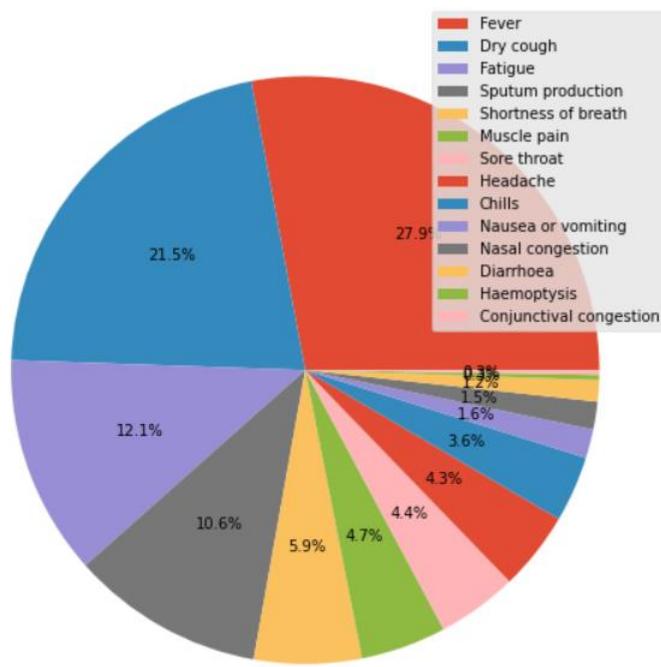
1.5. Phân phối các trường hợp đã xác nhận

Khi xem xét chung về các yếu tố ảnh hưởng của covid 19, ba quốc gia chịu ảnh hưởng nhiều nhất về mặt số người xác nhận nhiễm covid 19 là Hoa Kỳ, Ấn Độ và Brazil khi số người xác nhận ngày một tăng và đón khi tăng lên theo cấp số nhân.



Hình 71. Phân phối các trường hợp đã xác nhận

❖ Các triệu chứng



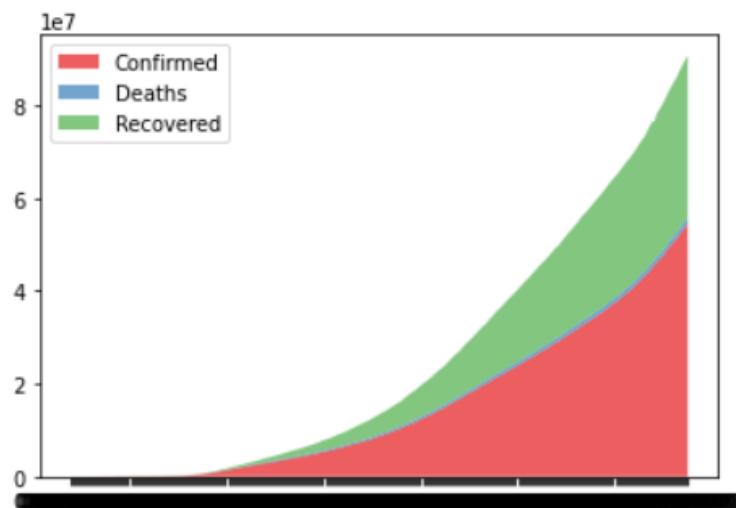
Hình 72. Phân phối các triệu chứng của covid 19

Mặc dù những người bị nhiễm vi-rút có thể không có triệu chứng, nhiều người phát triển các triệu chứng giống như cúm bao gồm sốt, ho và khó thở. Ngoài 3 triệu chứng chính trên đôi khi các bệnh nhân cũng có các triệu chứng phụ khác kèm theo như xuất đờm, mệt mỏi, đau đầu, ớn lạnh, ...

2. Đại dịch phát triển trên thế giới:

2.1. Tốc độ đại dịch covid 19 phát triển trên thế giới

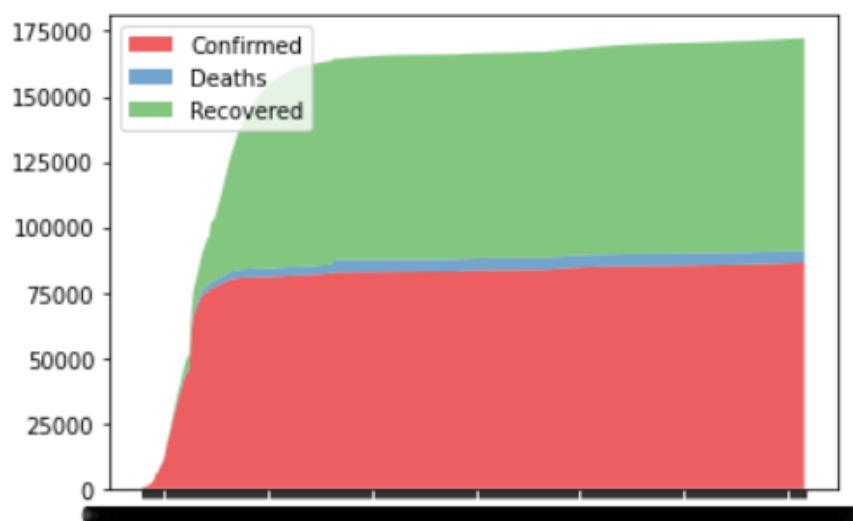
Các trường hợp được xác nhận vẫn đang tăng lên. Cũng có thể thấy sự gia tăng các trường hợp được phục hồi khi những người ăn xin đang chữa bệnh và sự gia tăng số người chết khi có nhiều người bị nhiễm bệnh.



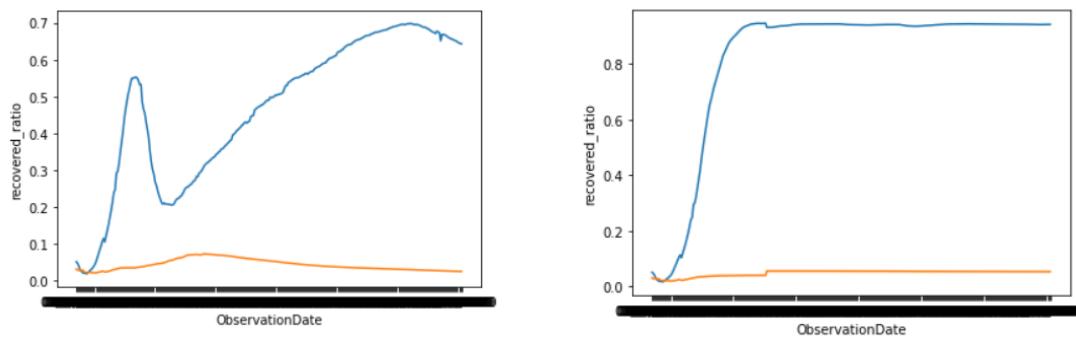
Hình 73. Tốc độ đại dịch covid phát triển trên thế giới
Trọng điểm Trung Quốc:

2.2. Tỷ lệ hồi phục và tử vong trên toàn cầu và ở Trung Quốc

Trung Quốc hầu như không có ca bệnh mới và những ca bệnh đã được thu hồi hầu như giống như đã xác nhận.



Hình 74. Tỷ lệ phục hồi và tử vong của Trung Quốc

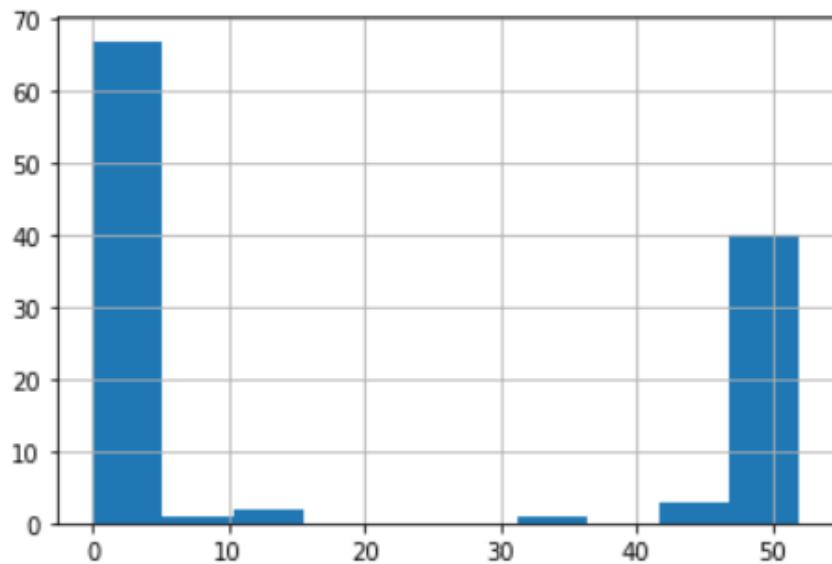


Hình 75. So sánh với Trung Quốc

Trung Quốc có tỷ lệ phục hồi là 80% và đang tăng lên, có nghĩa là họ đã kiểm soát được mức độ lây lan. Tỷ lệ phục hồi trên toàn cầu đang tăng lên nhưng các quốc gia mới bị nhiễm bệnh có thể là lý do đằng sau sự suy giảm khả năng phục hồi mới này. Tỷ lệ tử vong ổn định khoảng 3%.

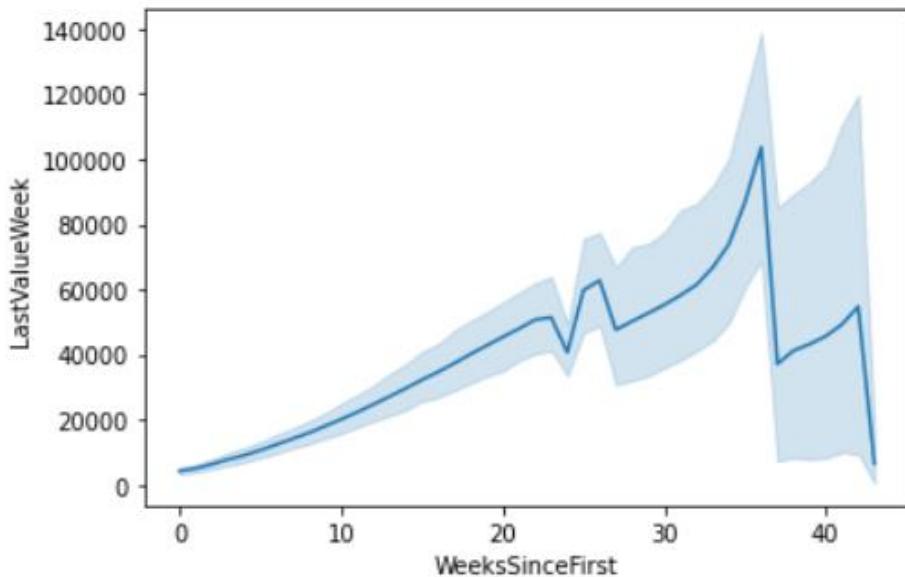
2.3. Sự tương quan giữa thời gian kể từ trường hợp đầu tiên và sự phục hồi

	Confirmed	TimeSinceFirst	RecoveredRatio	DeathsRatio
Confirmed	1.000000	0.187524	-0.030503	0.010619
TimeSinceFirst	0.187524	1.000000	0.062056	-0.132754
RecoveredRatio	-0.030503	0.062056	1.000000	-0.033199
DeathsRatio	0.010619	-0.132754	-0.033199	1.000000



Hình 76. Sự tương quan giữa thời gian kể từ trường hợp đầu tiên và sự phục hồi

Thời gian kể từ trường hợp đầu tiên và sự phục hồi có mối tương quan và nhiều khu vực cách trường hợp đầu tiên ít hơn một tuần. Vì vậy thời gian lan truyền là một yếu tố quan trọng.



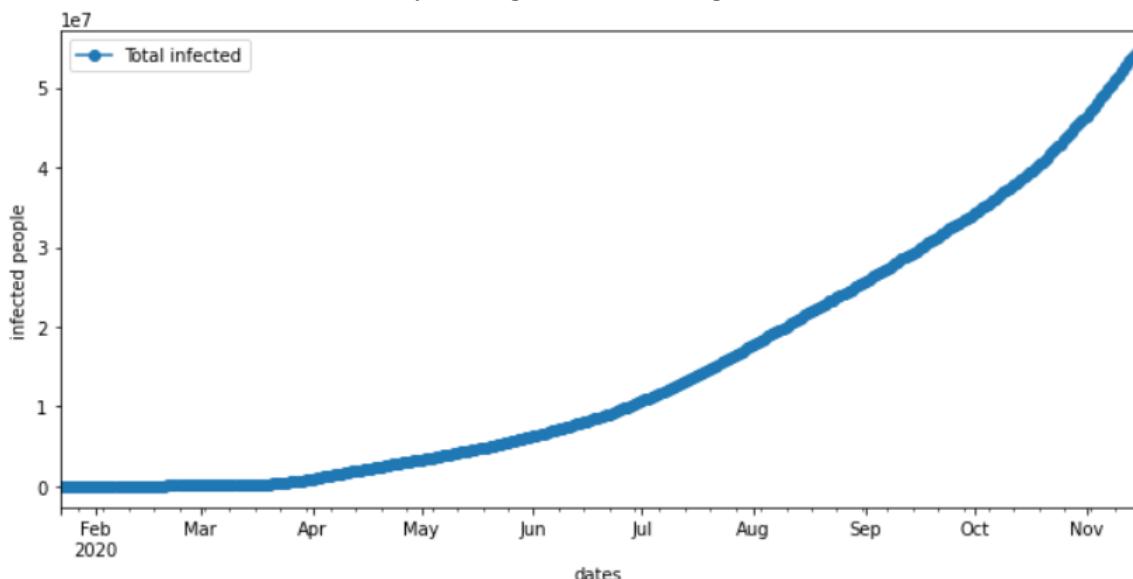
Hình 77. Trường hợp không đồng đều

Các trường hợp tiến hóa không đồng đều ở những nơi khác nhau, chủ yếu là do quy mô dân số, số đo lưu vực, v.v. mà không thể đo lường bằng dữ liệu.

3. Dự đoán số người chết theo thời gian trên thế giới:

3.1. Số người nhiễm bệnh theo thời gian:

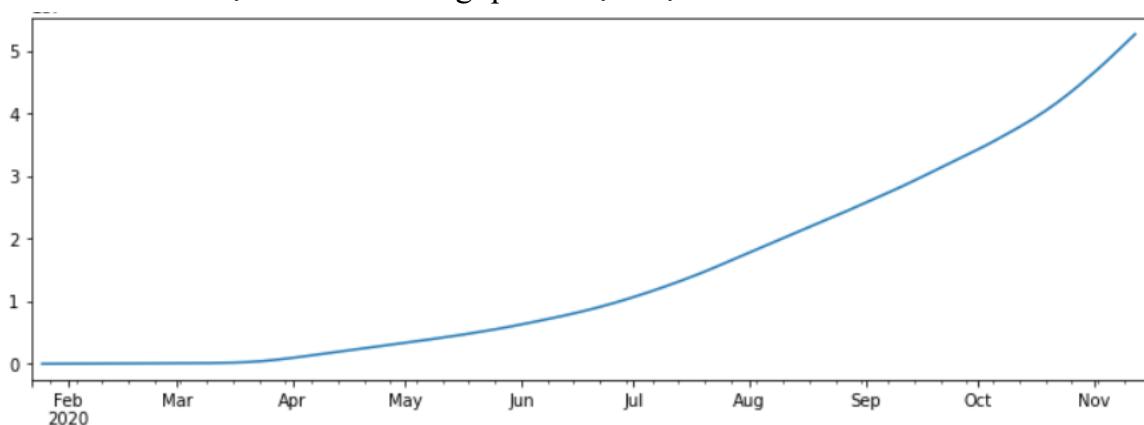
Số người nhiễm bệnh luôn tăng trong thời điểm hiện tại và chưa có dấu hiệu giảm do trong giai đoạn này là giai đoạn đầu của đại dịch covid 19 nên các biện pháp phòng chống chưa được thực thi hiệu quả cũng như chưa tìm ra được thuốc điều trị hay kháng sinh để chống lại virus.



Hình 78. Số trường hợp nhiễm bệnh theo thời gian

3.2. Xu hướng nhiễm bệnh theo thời gian:

Hiển nhiên, số người nhiễm bệnh không giảm thì xu hướng nhiễm bệnh cũng sẽ luôn tăng. Qua đó có thể nhận định số người nhiễm bệnh và xu hướng nhiễm bệnh có mối tương quan tỉ lệ thuận với nhau.



Hình 79. Xu hướng nhiễm bệnh theo thời gian

4. Cài đặt giải thuật:

Giải thuật được sử dụng ở nhiệm vụ này là hồi quy tuyến tính. Sử dụng giải thuật này vì để dự đoán giá trị của biến mục tiêu liên tục (số người trên thế giới ở các hạng mục) [118]

Độ chính xác của giải thuật đối với trường hợp dự đoán Covid-19 này là 69.98%

```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=0)
model=LinearRegression()
model.fit(x_train,y_train)
y_pred=model.predict(x_test)
from sklearn.metrics import r2_score

print(' R2 Score : ',r2_score(y_test, y_pred))
```

E. NHIỆM VỤ 5: TẠO BẢNG NGHIÊN CỨU LIÊN QUAN ĐẾN COVID-19 VÀ THỐNG KÊ CÂU HỎI LIÊN QUAN ĐẾN BỆNH NHÂN

I. Yêu cầu của nhiệm vụ:

Tạo bảng nghiên cứu liên quan đến covid 19

Cụ thể, tìm hiểu về:

1/Chúng ta biết gì về chẩn đoán và coronavirus?

2/Những tiến bộ mới trong chẩn đoán SARS-COV-2

Tạo bảng thống kê câu hỏi liên quan đến bệnh nhân:

Ở phần này sẽ khai thác các câu trả lời từ tóm tắt văn học về:

1/ Thời gian vi rút phát tán sau khi bệnh khởi phát

2/ Thời gian ủ bệnh ở các nhóm tuổi khác nhau

3/ Thời kỳ ủ bệnh của vi rút là gì?

4/ Tỷ lệ bệnh nhân không có triệu chứng

5/ Bệnh nhi không có triệu chứng

6/ Lây truyền không có triệu chứng trong thời gian ủ bệnh

7/ Lịch sử tự nhiên của vi rút từ người bị nhiễm

8/ Thời gian rụng của virus trung bình là gì?

9/ Khoảng thời gian dài nhất của virus là bao lâu?

10/Biểu hiện của COVID-19 bao gồm nhưng không giới hạn ở bệnh cơ tim có thể xảy ra và ngừng tim

11/Tải lượng vi rút liên quan như thế nào đến biểu hiện bệnh, bao gồm khả năng xét nghiệm chẩn đoán dương tính?

II. Sơ lược về các tập dữ liệu:

1. Sơ lược về tập dữ liệu:

Nhiệm vụ được thực hiện trên tập dữ liệu metadata.csv cùng hai tệp giải nén đính kèm để định dạng ASTRACT của tập dữ liệu

2. Tiền xử lý tập dữ liệu:

Quá trình làm sạch giữ liệu:

- Xóa đi các cột dữ liệu thừa và giữ lại các cột sau ['title','publish_time','journal','url','abstract','doi','cord_uid']
- Kiểm tra rỗng
- Tách từng từ trong cột ABSTRACT
- Tách các từ trong cột trừu tượng và tạo một cột mới

Sau khi qua quá trình làm sạch dữ liệu, có được bảng dữ liệu sau và bắt đầu tiến hành nhiệm vụ ở phần này:

Unnamed: #		title	publish_time	journal	url	abstract	doi	cord_uid	words
0	0	Clinical features of culture-proven Mycoplasma...	2001-07-04	BMC Infect Dis	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	OBJECTIVE: This retrospective chart review des...	10.1186/1471-2334-1-6	ug7v899j	['OBJECTIVE', 'This', 'retrospective', 'chart...', '...
1	1	Nitric oxide: a pro-inflammatory mediator in l...	2000-08-15	Respir Res	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5...	Inflammatory diseases of the respiratory tract...	10.1186/rr14	02tnwd4m	['Inflammatory', 'diseases', 'of', 'the', 'res...', '...
2	2	Surfactant protein-D and pulmonary host defense	2000-08-25	Respir Res	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5...	Surfactant protein-D (SP-D) participates in th...	10.1186/rr19	ejv2xln0	['Surfactant', 'protein', 'D', 'SP', 'D', 'par...', '...
3	3	Role of endothelin-1 in lung disease	2001-02-22	Respir Res	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5...	Endothelin-1 (ET-1) is a 21 amino acid peptide...	10.1186/rr44	2b73a28n	['Endothelin', '1', 'ET', '1', 'is', 'a', '21...', '...
4	4	Gene expression in epithelial cells in respons...	2001-05-11	Respir Res	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5...	Respiratory syncytial virus (RSV) and pneumoni...	10.1186/rr61	9785vg6d	['Respiratory', 'syncytial', 'virus', 'RSV', '...', '...

Hình 80. Tập dữ liệu metadata.csv sau khi tiền xử lý dữ liệu

III. Nội dung:

Tạo bảng thống kê nghiên cứu liên quan đến covid 19:

Sơ lược về các khái niệm cơ bản về chẩn đoán liên quan đến COVID19:

'Các xét nghiệm vi rút cho biết liệu bạn hiện có bị nhiễm SARS-CoV-2, loại vi rút gây ra COVID-19 hay không. Kết quả xét nghiệm dương tính có nghĩa là bạn bị nhiễm trùng (CDC). ' Xét nghiệm RT-PCR được sử dụng rộng rãi và có thể cho thấy tỷ lệ mắc bệnh trong dân số (Trung tâm An ninh Y tế Johns Hopkins).

'Xét nghiệm máu kháng thể kiểm tra máu của bạn bằng cách tìm kiếm các kháng thể, cho biết liệu bạn có bị nhiễm vi rút (CDC) trước đó hay không.' Các xét nghiệm huyết thanh học giúp định lượng tốt hơn số trường hợp mắc COVID-19, bao gồm cả các trường hợp không có triệu chứng và được tiên lượng sai (Trung tâm An ninh Y tế Johns Hopkins).

Các loại xét nghiệm huyết thanh học bao gồm: Xét nghiệm chẩn đoán nhanh (RDT), Xét nghiệm hấp thụ miễn dịch liên kết với enzym (ELISA), Xét nghiệm trung hòa và xét nghiệm miễn dịch phát quang (Trung tâm An ninh Y tế Johns Hopkins).

Nguồn:

Tiếp theo, tìm hiểu từng thuật ngữ và tạo bảng nghiên cứu về covid 19

Các thuật ngữ sau liên quan đến covid 19: chẩn đoán corona (diagnostics), corona cho công cụ chẩn đoán (diagnostic tolos), kiểm tra (test), COVID, FDA, kiểm tra virus (viral test). Đây là các thuật ngữ thường được tìm thấy nhiều nhất khi liên quan đến vấn đề nghiên cứu liên quan đến covid 19.

1. Các chẩn đoán về virus corona:

Khi tìm kiếm với thuật ngữ chẩn đoán về corona, có được 126 bản tóm tắt liên quan đến chẩn đoán. Trong đó gồm 126 bản tóm tắt là các bài báo cáo và nghiên cứu về virus corona như *Hội nghị chuyên đề quốc tế lần thứ 36 về chăm sóc chuyên sâu, Phổ phát hiện mầm bệnh đường hô hấp, Hội chứng Hô hấp Trung Đông Corona, Bảo tồn bộ gen RNA và cấu trúc thứ cấp, Phương pháp tiếp cận qRT-PCR được tối ưu hóa, ...*^{[25][26]} Các bài báo trình bày về Xét nghiệm Xpert Xpress Cúm / RSV được FDA công bố để phát hiện nhanh bệnh cúm A (FluA), cúm B (FluB) và vi rút hợp bào hô hấp (RSV). Nghiên cứu này nhằm đánh giá

hiệu quả lâm sàng của nó so với xét nghiệm RT-PCR đã được FDA Trung Quốc (CFDA-PCR) phê duyệt. Các bệnh phẩm mũi họng được thu thập từ bệnh nhân và được xét nghiệm song song với nhau. Các kết quả trái ngược đã được kiểm tra bằng PCR thời gian thực do phòng thí nghiệm phát triển để phân giải. Tải lượng virus trong mẫu được định lượng bằng PCR kỹ thuật số dạng giọt.

	title	publish_time	journal	url	abstract	doi	cord_uid	words
2770	36th International Symposium on Intensive Care...	2016-04-20	Crit Care	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5...	P001 - Sepsis impairs the capillary response w...	10.1186/s13054-016-1208-6	1virmqhdp	[P001, Sepsis, impairs, the, capillary, respon...
14569	452. Spectrum of Respiratory Pathogens Detected...	2018-11-26	Open Forum Infect Dis	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6...	BACKGROUND: Respiratory tract infections (RTI)...	10.1093/ofid/ofy210.461	ijjpxao	[BACKGROUND, Respiratory, tract, infections, R...
14572	739. Middle East Respiratory Syndrome Coronavi...	2018-11-26	Open Forum Infect Dis	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6...	BACKGROUND: A deadly zoonotic Middle East resp...	10.1093/ofid/ofy210.746	2g0zq9dh	[BACKGROUND, A, deadly, zoonotic, Middle, East...
15383	HRCT imaging features in representative Import...	2020-02-11	Precis Clin Med	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	With the spread of novel coronavirus (2019-nCo...	10.1093/pcmed/pbaa004	95l3zka6	[With, the, spread, of, novel, coronavirus, 20...
16315	The Domestic Cat, <i>Felis catus</i> , as a Model of H...	2008	Sourcebook of Models for Biomedical Research	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	The domestic cat, currently the most frequent ...	10.1007/978-1-59745-285-4_25	80l5muaz2	[The, domestic, cat, currently, the, most, fre...
...
359345	RNA genome conservation and secondary structur...	2020-04-06	bioRxiv	https://www.ncbi.nlm.nih.gov/pubmed/32511306/...	As the COVID-19 outbreak spreads, there is a g...	10.1101/2020.03.27.012906	c8lf7a9bq	[As, the, COVID, 19, outbreak, spreads, there,...
359574	Optimized qRT-PCR Approach for the Detection o...	2020-06-20	Int J Mol Sci	https://doi.org/10.3390/ijms21124396; https://...	The novel coronavirus SARS-CoV-2 is the causat...	10.3390/ijms21124396	d65njv7b	[The, novel, coronavirus, SARS, CoV, 2, is, th...

Hình 81. Các chuẩn đoán về virus corona

2. Tìm kiếm thông qua tóm tắt với từ CORONA cho CÔNG CỤ (diagnostic tolos):

Chuẩn đoán sớm cho thấy rằng nhiễm trùng SARS-CoV-2 có liên quan chặt chẽ với bệnh tim mạch. Hơn nữa, nó cũng thúc đẩy sự phát triển của các rối loạn tim mạch như suy tim, hội chứng mạch vành cấp tính, chấn thương cơ tim và huyết khối tĩnh mạch.

Do tính chất đại dịch của đợt bùng phát COVID-19 và sự gia tăng tỷ lệ mắc bệnh trên toàn cầu, các nỗ lực nhất quán đang diễn ra để kiềm chế tình hình và quản lý bệnh nhân bị nhiễm bệnh. Mặc dù đã có sự gia tăng theo cấp số nhân trong các nghiên cứu, nhưng vẫn chưa có phương pháp điều trị hoặc vắc xin hiệu quả nào cho COVID-19. Tuy nhiên, một số thử nghiệm lâm sàng đã được đề xuất hoặc đang tiếp tục. [27][28]

Các chuyên gia nghiên cứu lâm sàng nhận thấy rằng các điều chỉnh thử nghiệm lâm sàng liên quan đến COVID-19 tác động tích cực đến nhiều khía cạnh của việc tiến hành nghiên cứu. Những người có kinh nghiệm tốt nhất-cụ thể đối với các thay đổi liên quan đến COVID-19 và nói chung có nhiều khả năng khuyến nghị những điều chỉnh này tiếp tục trong tương lai.

204201	CRISPR-cas systems based molecular diagnostic...	2020	J Drug Target	NaN	Emerging infectious diseases, the persistent p...	NaN	5uv9kaxf	[Emerging, infectious, diseases, the, persiste...
210504	Automated detection of COVID-19 cases using de...	2020	Comput Biol Med	NaN	The novel coronavirus 2019 (COVID-2019), which...	NaN	6aj8zsto	[The, novel, coronavirus, 2019, COVID, 2019, w...
220791	Nano-enabled biosensing systems for intelligent...	2020	Mater. Today Chem.	NaN	Biosensors are emerging as efficient (sensitiv...	NaN	opjfs7w	[Biosensors, are, emerging, as, efficient, sen...
260876	A Review of Salivary Diagnostics and Its Poten...	2020-04-17	Cureus	https://doi.org/10.7759/cureus.7708; https://w...	Saliva is an exocrine secretion produced from ...	10.7759/cureus.7708	awy3aans	[Saliva, is, an, exocrine, secretion, produced...
267271	Pathogenicity and pathogenesis of a United Sta...	2015-03-25	Virology	https://www.ncbi.nlm.nih.gov/pubmed/25817405/...	Porcine deltacoronavirus (PDCoV) was first ide...	10.1016/j.virol.2015.03.024	p5kjnfyf	[Porcine, deltacoronavirus, PDCoV, was, first,...
277065	Aspects of nano-enabling biosensing systems fo...	2020-06-05	Mater Today Chem	https://www.sciencedirect.com/science/article/...	Biosensors are emerging as efficient (sensitiv...	10.1016/j.mtchem.2020.100306	4vjf6sgs	[Biosensors, are, emerging, as, efficient, sen...
304092	CRISPR-cas systems based molecular diagnostic...	2020-05-26	J Drug Target	https://doi.org/10.1080/1061186x.2020.1769637;...	Emerging infectious diseases, the persistent p...	10.1080/1061186x.2020.1769637	q52mia0e	[Emerging, infectious, diseases, the, persiste...
312191	Automated detection of COVID-19 cases using de...	2020-04-28	Comput Biol Med	https://doi.org/10.1016/j.combiomed.2020.1037...	The novel coronavirus 2019 (COVID-2019), which...	10.1016/j.combiomed.2020.103792	yolo8bb3	[The, novel, coronavirus, 2019, COVID, 2019, w...

Hình 82. Tìm kiếm thông qua tóm tắt với từ CORONA cho CÔNG CỤ (diagnostic tolols)

3. Tìm kiếm thông qua tóm tắt với từ kiểm tra (test):

Virus đại diện cho một trong những tác nhân môi trường chính gây ra bệnh tật và bệnh tật cho con người. Tuy nhiên, khả năng chẩn đoán nhiễm virus bị giới hạn bởi khả năng và phạm vi phát hiện. Ở đây mô tả một số công nghệ mới nỗi cung cấp thông tin chẩn đoán virus nhanh chóng và / hoặc chất lượng cao. Hai công nghệ, chẩn đoán dựa trên CRISPR mới và một công cụ giải trình tự DNA di động, phù hợp duy nhất để tăng số lượng tác nhân vi rút được phân tích, ngay cả trong các cơ sở chăm sóc. [119]. Cách cảm biến sinh học không có nhãn nanophotonic có thể góp phần chẩn đoán nhanh chóng và đại trà các bệnh nhiễm vi rút đường hô hấp [120]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
290	Emerging Technologies for the Detection of Rab...	2009-09-29	PLoS Negl Trop Dis	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2...	The diagnosis of rabies is routinely based on ...	10.1371/journal.pntd.0000530	d0eur1hq	[The, diagnosis, of, rabies, is, routinely, ba...
554	A Simple, Inexpensive Device for Nucleic Acid ...	2011-05-09	PLoS One	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	BACKGROUND: Molecular assays targeted to nucle...	10.1371/journal.pone.0019738	qol5poan	[BACKGROUND, Molecular, assays, targeted, to, ...
611	Identification of a Highly Conserved H1 Subtyp...	2011-08-19	PLoS One	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	Subtype specificity of influenza A virus (IAV)...	10.1371/journal.pone.0023374	ff2zzadol	[Subtype, specificity, of, influenza, A, virus...]
763	Isothermal Amplification Using a Chemical Heat...	2012-02-23	PLoS One	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	BACKGROUND: To date, the use of traditional nu...	10.1371/journal.pone.0031432	hviale7ay	[BACKGROUND, To, date, the, use, of, tradition...
779	Use and Evaluation of Molecular Diagnostics fo...	2012-03-08	Clinical Infectious Diseases	https://academic.oup.com/cid/article-pdf/54/su...	Comprehensive microbiological testing will be ...	10.1093/cid/cir1060	z5p3a64x	[Comprehensive, microbiological, testing, will...
...
360785	Molecular Diagnosis of Diarrhea: Current Status...	2011-11-26	Curr Infect Dis Rep	https://doi.org/10.1007/s11908-011-0223-7; ht...	Determining the microbiologic etiology of ente...	10.1007/s11908-011-0223-7	tlpncodox	[Determining, the, microbiologic, etiology, of...]
361549	How Nanophotonic Label-Free Biosensors Can Con...	2020-08-07	ACS Sens	https://www.ncbi.nlm.nih.gov/pubmed/32786383/...	[Image: see text] The global sanitary crisis c...	10.1021/acssensors.0c01180	p2cm0gnd	[, Image, see, text, The, global, sanitary, cr...
361570	10 Wochen alter Säugling mit Sepsis Verdacht un...	2020-06-03	Monatsschr Kinderheilkd	https://doi.org/10.1007/s00112-020-00942-8; ht...	A young male infant was referred to the emerge...	10.1007/s00112-020-00942-8	m8fl96zb	[A, young, male, infant, was, referred, to, th...

360785	Molecular Diagnosis of Diarrhea: Current Status...	2011-11-26	Curr Infect Dis Rep	https://doi.org/10.1007/s11908-011-0223-7; ht...	Determining the microbiologic etiology of ente...	10.1007/s11908-011-0223-7	tlpnodox	[Determining, the, microbiologic, etiology, of...]
361549	How Nanophotonic Label-Free Biosensors Can Con...	2020-08-07	ACS Sens	https://www.ncbi.nlm.nih.gov/pubmed/32786383; ht...	[Image: see text] The global sanitary crisis c...	10.1021/acssensors.0c01180	p2cm0gnd	[, Image, see, text, The, global, sanitary, cr...
361570	10 Wochen alter Säugling mit Sepsisverdacht un...	2020-06-03	Monatsschr Kinderheilkd	https://doi.org/10.1007/s00112-020-00942-8; ht...	A young male infant was referred to the emerge...	10.1007/s00112-020-00942-8	m8ff96zb	[A, young, male, infant, was, referred, to, th...
361821	Improving Evidence Based Bronchiolitis Care	2018-02-06	Clin Pediatr Emerg Med	https://doi.org/10.1016/j.cpem.2018.02.003; ht...	Bronchiolitis is the number one cause of hospit...	10.1016/j.cpem.2018.02.003	59wx6ws	[Bronchiolitis, is, the, number, one, cause, o...
361915	Point-of-care CRISPR/Cas nucleic acid detectio...	2020-07-26	Biosens Bioelectron	https://www.sciencedirect.com/science/article/...	With the trend of moving genomic test from (me...	10.1016/j.bios.2020.112445	muy75w79	[With, the, trend, of, moving, genomic, test , ...]

Hình 83. Tìm kiếm thông tin qua từ kiểm tra (test)

4. Tìm kiếm thông qua tóm tắt với từ COVID:

Các biện pháp cấm cửa, hạn chế đi lại, cách ly, cách xa xã hội và các biện pháp y tế công cộng nghiêm ngặt khác đang góp phần vào việc trì hoãn sự lây lan của bệnh nhiễm trùng, nhưng một loại vắc xin an toàn và hiệu quả, phương pháp điều trị hiệu quả, chẩn đoán tại điểm chăm sóc và các sản phẩm y tế khác đang rất tuyệt vời cần thiết bởi vì trên thực tế, các chính phủ có thể không thể gia hạn các biện pháp này trong một khoảng thời gian không xác định. Để ngăn chặn sự lây lan theo cấp số nhân của SARS-CoV-2, các biện pháp bảo vệ khác nhau là cần thiết để ngăn ngừa nhiễm trùng giọt, nhưng cũng ngăn ngừa lây truyền qua các bề mặt bị ô nhiễm. Các biện pháp bảo vệ để ngăn chặn đại dịch bao gồm các quy định về khoảng cách (1,5 m), hạn chế tiếp xúc và đeo khẩu trang [344]. Những biện pháp này làm giảm đáng kể thực hành lâm sàng hàng ngày trong phẫu thuật chỉnh hình và chấn thương [288][14][15][16][17][18][19]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
24726	SARS-CoV-2-Pandemie und ihre Auswirkungen auf ...	2020-05-14	Knie J.	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	With the outbreak of the COVID-19 pandemic, in...	10.1007/s43205-020-00062-z	q4gozwtv	[With, the, outbreak, of, the, COVID , 19, pand...
25447	Multifunctional nano-magnetic particles assist...	2020-05-18	NaN	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	The enduring outbreak of the corona virus dise...	10.1080/14328917.2020.1769350	2v1pqnjt	[The, enduring, outbreak, of, the, corona, vir...
27886	Olfactory Device for Large Scale Pre-screening...	2020-06-26	Trans Indian Natl	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	Recent pandemic caused by COVID-19 has motivat...	10.1007/s41403-020-00126-6	f0jv0fyw	[Recent, pandemic, caused, by, COVID , 19, has...
28495	COVSACK: an innovative portable isolated and s...	2020-07-05	Trans Indian Natl	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	The Corona Virus Disease 2019 (COVID-19) is on...	10.1007/s41403-020-00139-1	68jypaaaw	[The, Corona, Virus, Disease, 2019, COVID , 19,...
31641	Nanoparticle-Based Strategies to Combat COVID -19	2020-08-25	ACS Appl Nano Mater	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	[Image: see text] Coronavirus disease 2019 (CO...	10.1021/acsanm.0c01978	rrsmc250	[, Image, see, text, Coronavirus, disease, 201...

Hình 84. Tìm kiếm thông tin qua từ COVID

5. Tìm kiếm thông tin tóm tắt với từ FDA:

Công dụng của FDA là giám sát và đánh giá chất lượng thực phẩm, dược phẩm có phù hợp với tiêu chí nhập khẩu không. Nó được quy định về chất lượng có thể kể đến gồm:

- Thực phẩm
- Thuốc lá
- Thực phẩm chức năng
- Sản phẩm bổ sung chế độ ăn uống
- Dược phẩm theo toa hoặc không theo toa
- Vắc-xin
- Truyền máu
- Thiết bị y tế
- Thiết bị phát bức xạ điện tử
- Các sản phẩm liên quan đến Thú y

Các bài báo ở mục này nhằm thể hiện quan điểm của FDA về lâm sàng thiết bị chẩn đoán, đánh giá lâm sàng và xác nhận nhanh chóng, giấy phép sử dụng khẩn cấp của FDA Hoa Kỳ, chẩn đoán đồng hành: các chiến lược mới nổi hay thử nghiệm dựa trên bảng điều khiển Syndromic trong lâm sàng [20][21][22][23][24]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
37562	Companion diagnostics: emerging strategies and...	2012	Expert review of molecular diagnostics	https://doi.org/10.1586 erm.12.49; https://www...	This article summarizes the broad messages fro...	10.1586 erm.12.49	mwzhy7lb	[This, article, summarizes, the, broad, messag...
53802	United States FDA's emergency use authorization...	2015	Expert review of molecular diagnostics	https://doi.org/10.1586/14737159.2015.1077117...	The Ebola outbreak that took hold in West Afr...	10.1586/14737159.2015.1077117	qb4zizor	[The, Ebola, outbreak, that, took, hold, in, W...
61124	Syndromic Panel-Based Testing in Clinical Micr...	2018	Clinical microbiology reviews	https://doi.org/10.1128/cmr.00024-17; https://...	The recent development of commercial panel-bas...	10.1128/cmr.00024-17	gsprsbud	[The, recent, development, of, commercial, pan...
71305	Diagnosis of SARS-CoV-2: A Review on the Curre...	2020-09-28	Acta virologica	https://doi.org/10.4149/av_2020_402; https://w...	The novel severe acute respiratory syndrome co...	10.4149/av_2020_402	07lu55ah	[The, novel, severe, acute, respiratory, syndr...
262445	FDA Perspectives on Diagnostic Device Clinical...	2011-05-01	Clin Infect Dis	https://doi.org/10.1093/cid/cir056; https://ww...	Two pathways are described for submission to F...	10.1093/cid/cir056	uijwq8gx	[Two, pathways, are, described, for, submissio...
281748	Validation of a Single-step, Single-tube Rever...	2020-04-30	bioRxiv	https://doi.org/10.1101/2020.04.28.067363	2.IntroductionThe SARS-CoV-2 pandemic of 2020 ...	10.1101/2020.04.28.067363	wiyvrplj	[2, IntroductionThe, SARS, CoV, 2, pandemic, o...
288894	Clinical assessment and validation of a rapid ...	2020-05-18	NaN	https://doi.org/10.1101/2020.05.12.20095638; h...	Amid the enduring COVID-19 pandemic, there is ...	10.1101/2020.05.12.20095638	m7gzgtl	[Amid, the, enduring, COVID, 19, pandemic, the...
295945	Molecular detection and point-of-care testing ...	2015-09-14	Expert Rev Mol Diagn	https://www.ncbi.nlm.nih.gov/pubmed/26367243/	Ultrahigh sensitivity and specificity assays t...	10.1586/14737159.2015.1079776	3wsihzhk	[Ultrahigh, sensitivity, and, specificity, ass...
296280	The sensitivity of respiratory tract specimens...	2020-07-04	medRxiv	https://doi.org/10.1101/2020.07.02.20144543; h...	BACKGROUND: Highly sensitive, non-invasive, an...	10.1101/2020.07.02.20144543	9afukke4	[BACKGROUND, Highly, sensitive, non, invasive,...
296901	Comparison of Four Molecular In Vitro Diagnost...	2020-04-22	NaN	https://doi.org/10.1101/2020.04.17.20069864; h...	The novel human coronavirus SARS-CoV-2 was fir...	10.1101/2020.04.17.20069864	2bf3xznl	[The, novel, human, coronavirus, SARS, CoV, 2,...
299197	Comparison of Four Molecular In Vitro Diagnost...	2020-07-23	J Clin Microbiol	https://doi.org/10.1128/jcm.00743-20; https://...	Severe acute respiratory syndrome coronaviru ...	10.1128/jcm.00743-20	b0qtzzac	[Severe, acute, respiratory, syndrome, coronav...

Hình 85. Tìm kiếm thông tin qua từ FDA

6. Tìm kiếm thông qua tóm tắt với từ Viral Test:

Đại dịch COVID-19, do virus SARS-CoV-2 gây ra, đặt ra những mối đe dọa nghiêm trọng đối với cả nền kinh tế và sức khỏe toàn cầu. Màn hình chẩn đoán chủ yếu được sử dụng để phát hiện SARS-CoV-2 là các kỹ thuật phân tử, chẳng hạn như các xét nghiệm khuếch đại axit nucleic. So sánh các phương pháp chẩn đoán bằng nhiệt hiện tại và mới nổi cho COVID-19 và phác thảo các kỹ thuật phân tử và huyết thanh học hiện đang được sử dụng để phát hiện nhiễm SARS-CoV-2, trong quá khứ hoặc hiện tại, ở bệnh nhân. Thảo luận về nghiên cứu đang diễn ra về kỹ thuật bằng nhiệt, xét nghiệm phát hiện qua trung gian CRISPR và chẩn đoán điểm chăm sóc có tiềm năng sử dụng trong phát hiện SARS-CoV-2. Thủ nghiệm vi rút quy mô lớn trong một đại dịch toàn cầu đặt ra những thách thức độc đáo, chủ yếu là nhu cầu đồng thời về nguồn cung cấp thử nghiệm, thiết bị bền và nhân sự ở nhiều khu vực trên toàn thế giới, với mỗi khu vực có nhu cầu kiểm tra khác nhau khi đại dịch tiến triển. Các công nghệ bằng nhiệt chi phí thấp được mô tả trong bài đánh giá này cung cấp một phương tiện đầy hứa hẹn để giải quyết những nhu cầu này và đáp ứng nhu cầu toàn cầu về việc kiểm tra những người có triệu chứng, cũng như cung cấp một phương tiện khả thi để kiểm tra định kỳ những người không có triệu chứng, cung cấp một phương tiện tiềm năng để cho phép mở cửa trở lại một cách an toàn và cung cấp giám sát sớm các đợt bùng phát. [29][30][31]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
138164	Clinical features, diagnostics, and outcomes o...	2020	EClinicalMedicine	NaN	Background: Most data on the clinical presenta...	NaN	I4mg5vg5	[Background, Most, data, on, the, clinical, pr...
212704	Isothermal SARS-CoV-2 Diagnostics: Tools for E...	2020	ACS synth. biol	NaN	The COVID-19 pandemic, caused by the SARS-CoV...	NaN	ez66h8mq	[The, COVID, 19, pandemic, caused, by, the, SA...
258386	Isothermal SARS-CoV-2 Diagnostics: Tools for E...	2020-09-23	ACS Synth Biol	https://doi.org/10.1021/acssynbio.0c00359; ht...	[Image: see text] The COVID-19 pandemic, cause...	10.1021/acssynbio.0c00359	a6p6ma8r	[, Image, see, text, The, COVID, 19, pandemic, ...
261284	More targeted use of oseltamivir and in-hospital...	2019-04-07	J Clin Virol	https://www.sciencedirect.com/science/article/...	BACKGROUND: Immunocompromised adults are more ...	10.1016/j.jcv.2019.04.003	3cwh5yh7	[BACKGROUND, Immunocompromised, adults, are, m...
300743	Clinical features, diagnostics, and outcomes o...	2020-05-06	medRxiv	https://www.ncbi.nlm.nih.gov/pubmed/32511488; ...	BACKGROUND: Emerging data on the clinical pres...	10.1101/2020.05.02.20082461	jgwqqr5	[BACKGROUND, Emerging, data, on, the, clinical...
313095	Clinical features, diagnostics, and outcomes o...	2020-08-26	EClinicalMedicine	https://doi.org/10.1016/j.eclinm.2020.100518; ...	BACKGROUND: Most data on the clinical presenta...	10.1016/j.eclinm.2020.100518	76w0p6pk	[BACKGROUND, Most, data, on, the, clinical, pr...
361821	Improving Evidence Based Bronchiolitis Care	2018-02-06	Clin Pediatr Emerg Med	https://doi.org/10.1016/j.cpem.2018.02.003; ht...	Bronchiolitis is the number one cause of hospit...	10.1016/j.cpem.2018.02.003	59lwx6ws	[Bronchiolitis, is, the, number, one, cause, o...

Hình 86. Tìm kiếm thông tin qua từ Viral Test

Tạo bảng câu hỏi liên quan đến bệnh nhân:

Trước tiên, để trả lời được các câu hỏi trên, xem qua sơ lượt các thuật ngữ về chúng trong 111 bài viết **ABSTRACTS** với kích thước mẫu của 51012 mục nhập.

	title	publish_time	journal	url	abstract	doi	cord_uid	words
50	Markers of exacerbation severity in chronic ob...	2006-05-10	Respir Res	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1...	BACKGROUND: Patients with chronic obstructive ...	10.1186/1465-9921-7-74	utgik4af	[BACKGROUND, Patients, with, chronic, obstruct...
717	True versus False Parasite Interactions: A Rob...	2012-01-03	PLoS One	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	BACKGROUND: Multiple infections are common in ...	10.1371/journal.pone.0029618	bbjmod05	[BACKGROUND, Multiple, infections, are, common...
781	Common Variants in CDKN2B-AS1 Associated with...	2012-03-12	PLoS One	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	BACKGROUND: To date, only a small portion of t...	10.1371/journal.pone.0033389	zmhobi0r	[BACKGROUND, To, date, only, a, small, portion...
923	The VNTR Polymorphism of the DC-SIGNR Gene and...	2012-09-05	PLoS One	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	BACKGROUND: Dendritic cell-specific intercellu...	10.1371/journal.pone.0042972	nebf6u0z	[BACKGROUND, Dendritic, cell, specific, inter...
1142	Carriage of Mycoplasma pneumoniae in the Upper...	2013-05-14	PLoS Med	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	Mycoplasma pneumoniae is thought t...	10.1371/journal.pmed.1001444	d6awwygy	[BACKGROUND, Mycoplasma, pneumoniae, is, though...
...
361602	How Should Clinicians Interpret Imprecise Tria...	2020-06-05	NaN	http://medrxiv.org/cgi/content/short/2020.05.2...	As the COVID-19 pandemic progresses, research...	10.1101/2020.05.26.20114009	jyktuvyy	[As, the, COVID, 19, pandemic, progresses, res...
361804	The use of intravenous immunoglobulin gamma fo...	2020-10-21	BMC Infect Dis	https://doi.org/10.1186/s12879-020-05507-4; ht...	BACKGROUND: Coronavirus disease 2019 (COVID-19...	10.1186/s12879-020-05507-4	6gssaxcl	[BACKGROUND, Coronavirus, disease, 2019, COVID...
361993	High and increasing prevalence of SARS-CoV-2 s...	2020-10-14	NaN	https://doi.org/10.1101/2020.10.12.20211227; h...	Background REACT-1 is quantifying prevalence o...	10.1101/2020.10.12.20211227	x6ykak5k	[Background, REACT, 1, is, quantifying...

Hình 87. Các thuật ngữ trong 111 bài viết ABSTRACTS

1. Thời gian vi rút phát tán sau khi bệnh khởi phát:

Các bài báo nói về việc phát hiện virus ở người và virus corona ở những người nhận ghép tế bào gốc tạo máu di gen, việc thực hiện giám sát trong 1 năm và mô tả tiền sử tự nhiên của các bệnh nhiễm trùng này. Cuối cùng đưa lên thời gian tiêu diệt virus (3 tuần) và các số liệu để chứng minh cho việc đó. [40][41][42][43][44][45][46][47][48]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
141758	Case Report: Viral Shedding for 60 Days in a W...	2020	Am J Trop Med Hyg	NaN	Novel coronavirus disease (COVID-19) caused by...	NaN	1vba9i42	[Novel, coronavirus, disease, COVID, 19, caused...
156014	SARS-CoV-2 Shedding from Asymptomatic Patients...	2020	Am J Trop Med Hyg	NaN	The ongoing pandemic COVID-19, caused by SARS-...	NaN	r5qyroo3	[The, ongoing, pandemic, COVID, 19, caused, by...
226801	Kinetics of Viral Clearance and Antibody Produ...	2020	J. pediatr. (Rio J.)	NaN	OBJECTIVES: To improve understanding of transi...	NaN	q2jj7ohe	[OBJECTIVES, To improve, understanding, of, t...
226802	Kinetics of Viral Clearance and Antibody Produ...	2020	J. pediatr. (Rio J.)	NaN	OBJECTIVES: To improve understanding of transi...	NaN	q2jj7ohe	[OBJECTIVES, To, improve, understanding, of, t...
267480	SARS-CoV-2 Shedding from Asymptomatic Patients...	2020-05-13	Am J Trop Med Hyg	https://www.ncbi.nlm.nih.gov/pubmed/32406369/; ht...	The ongoing pandemic COVID-19, caused by SARS-...	10.4269/ajtmh.20-0279	g948h572	[The, ongoing, pandemic, COVID, 19, caused, by...
270131	Genetic and antigenic heterogeneity of GI-1/Ma...	2020-09-01	Poult Sci	https://doi.org/10.1016/j.psj.2020.08.037; ht...	Four GI-1/Massachusetts-type (GI-1/Mass-type) ...	10.1016/j.psj.2020.08.037	li1vr0b	[Four, GI, 1, Massachusetts, type, GI, 1, Mass...
272361	Kinetics of viral clearance and antibody produ...	2020-08-07	NaN	http://medrxiv.org/cgi/content/short/2020.08.0...	Objectives: To improve understanding of transi...	10.1101/2020.08.06.20162446	6sa6oc6u	[Objectives, To, improve, understanding, of, t...

Hình 88. Thời gian virus phát tán sau khi khởi phát

2. Thời gian ủ bệnh ở các nhóm tuổi khác nhau

Để làm rõ về thời gian ủ bệnh ở các nhóm tuổi, vấn đề chia làm hai thành phần nhỏ về tuổi gồm tuổi khác nhau và nhóm tuổi:

2.1. Tuổi khác nhau:

Khi tìm hiểu theo độ tuổi khác nhau, đặc điểm dịch tễ học của virus corona mới về thời gian ủ bệnh sẽ lâu hơn khi nó ở người cao tuổi cũng như tình trạng nhiễm bệnh sẽ nhanh hơn. Sự tái phát bệnh sau khi khỏi bệnh cũng chiếm nguy cơ cao hơn các độ tuổi còn lại. [49][50][51][51]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
226457	Epidemiological features of new coronavirus in...	2020	Problemy Osobopospasnykh Infektsii	NaN	The review analyses the information on the epi...	NaN	sc1ic51i	[The, review, analyses, the, information, on, ...]
226747	Probable longer incubation period for elderly ...	2020	Risk Manage. Healthc. Policy	NaN	Background: Factors associated with the incubation period for elderly ...	NaN	u1fde7ok	[Background, Factors, associated, with, the, i...]
268572	Estimate the incubation period of coronavirus ...	2020-02-29	NaN	https://doi.org/10.1101/2020.02.24.20027474	Motivation: Wuhan pneumonia is an acute infection...	10.1101/2020.02.24.20027474	vkgnwxzc	[Motivation, Wuhan, pneumonia, is, an, acute, ...]
294601	Probable Longer Incubation Period for Elderly ...	2020-08-11	Risk Manag Healthc Policy	https://doi.org/10.2147/rmhps257907; https://...	BACKGROUND: Factors associated with the incubation period for elderly ...	10.2147/rmhps257907	9djaqk36	[BACKGROUND, Factors, associated, with, the, ...]
316492	Effects of latency and age structure on the dy...	2020-04-29	NaN	https://doi.org/10.1101/2020.04.25.20079848; h...	In this paper we develop an SEIR-type model of the effects of latency and age structure on the dynamics of the disease.	10.1101/2020.04.25.20079848	rv06wb3x	[In, this, paper, we, develop, an, SEIR, type,...]
351134	Relapsing and remitting viral diseases of the ...	1984-08-31	Trends in Neurosciences	https://www.sciencedirect.com/science/article/...	Abstract Virus infections can recur intermittently.	10.1016/s0166-2236(84)80191-5	m7mhyrp8	[Abstract, Virus, infections, can, recur, inter...]

Hình 89. Thời gian ủ bệnh ở độ tuổi khác nhau

2.2. Nhóm tuổi:

Khi tìm hiểu theo nhóm tuổi đã thấy được sự ảnh hưởng của nhóm tuổi đối với đại dịch lần này. Ngoài ra các bài báo cũng cho thấy một phần sự lan truyền của đại dịch song song việc nhận thức của việc cung cấp y tế, các chi tiết về tác động của đại dịch đối với các vấn đề bệnh tật [52][53][54][55][56][57][58][59][60]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
7259	The Influence of Age on Steroidogenic Enzyme A...	1987-06-01	Endocrinology	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	The ability of isolated adrenocortical cells to...	10.1210/endo-120-6-2521	4niom5mw	[The, ability, of, isolated, adrenocortical, c...
24086	Clinical Characteristics and Differential Clin...	2020-04-30	Coronavirus Disease 2019 (COVID-19)	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	Novel Coronavirus Disease (COVID-19) has become...	10.1007/978-981-15-4814-7_6	nzjz0q8b	[Novel, Coronavirus, Disease, COVID, 19, has, ...]
34167	Demographic factors, attitude and knowledge of...	2020-09-30	NaN	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	This study investigated demographic factors, ...	10.35241/emeraldopenres.13800.2	nk7u1o9a	[This, study, investigated, demographic, facto...]
35025	Insights of COVID-19 pandemic impact on anesthe...	2020-11-11	Ain-Shams J Anesthesiol	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	ABSTRACT: New corona virus disease COVID-19 is...	10.1186/s42077-020-00110-w	uyyis9e3	[ABSTRACT, New, corona, virus, disease, COVID,...]
39036	Household secondary attack rate of COVID-19 an...	2020-06-17	The Lancet. Infectious diseases	https://doi.org/10.1016/s1473-3099(20)30471-0; ...	BACKGROUND As of June 8, 2020, the global repo...	10.1016/s1473-3099(20)30471-0	gt5pq29h	[BACKGROUND, As, of, June, 8, 2020, the, globa...]
50620	[Study on transmission dynamic of 15 clusters ...	2020-05-13	Zhonghua liu xing bing xue za zhi = Zhonghua I...	https://doi.org/10.3760/cma.j.cn112338-2020033...	Objective: To describe the basic characteristi...	10.3760/cma.j.cn112338-20200330-00466	5zzqt1vp	[Objective, To, describe, the, basic, characte...]
52108	Epidemiological parameters of COVID-19 and its...	2020-10-01	Euro surveillance : bulletin Europeen sur les ...	https://doi.org/10.2807/1560-7917.es.2020.25.4...	BackgroundThe natural history of disease in pa...	10.2807/1560-7917.es.2020.25.40.2000250	as7x9up6	[BackgroundThe, natural, history, of, disease,...]
60323	Data sharing during COVID-19 pandemic:	2020-08-26	Expert review of gastroenterology &	https://doi.org/10.1080/17474124.2020.1815533; ...	INTRODUCTION In light of the viral	10.1080/17474124.2020.1815533	hmxih0i0	[INTRODUCTION, In, light, of, the, ...]

Hình 90. Thời gian ủ bệnh ở nhóm tuổi khác nhau

3. Thời kỳ ủ bệnh của vi rút là gì?

Thời gian ủ bệnh là khoảng thời gian từ khi bị nhiễm cho tới khi khởi phát các triệu chứng bệnh. Thời gian ủ bệnh ước tính hiện nay trong khoảng 2-12 ngày, và thời gian này sẽ được cập nhật chính xác hơn khi có thêm dữ liệu. Dựa trên thông tin về các bệnh do virus corona khác như MERS và SARS, thời gian ủ bệnh của 2019-nCoV có thể dưới 14 ngày. Để tính toán thời gian ủ bệnh, các nhà khoa học phân tích dữ liệu từ hơn 180 báo cáo các ca Covid-19 tại những khu vực không có tình trạng truyền nhiễm trong cộng đồng, bệnh nhân lây nhiễm sau khi ra nước ngoài. [61][62][63][64][65][66]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
24086	Clinical Characteristics and Differential Clin...	2020-04-30	Coronavirus Disease 2019 (COVID-19)	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	Novel Coronavirus Disease (COVID-19) has become...	10.1007/978-981-15-4814-7_6	nzjz0q8b	[Novel, Coronavirus, Disease, COVID, 19, has, ...]
24559	Recommendations for control and prevention of ...	2020-04-23	NaN	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	The outbreak of Coronavirus Disease 2019 (COVI...	10.1136/wjps-2020-000124	yump76qu	[The, outbreak, of, Coronavirus, Disease, 2019...]
24663	Role of Nasopharyngeal lactate dehydrogenase a...	2020-04-30	Indian J Med Sci	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	COVID-19 pandemic is affecting almost every co...	10.25259/ijms_25_2020	xc6sz085	[COVID, 19, pandemic, is, affecting, almost, e...
25447	Multifunctional nano-magnetic particles assist...	2020-05-18	NaN	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	The enduring outbreak of the corona virus dise...	10.1080/14328917.2020.1769350	2v1pqnjt	[The, enduring, outbreak, of, the, corona, vir...
31914	Estimating the Impact of Daily Weather on the ...	2020-09-17	Earth Syst Environ	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	The COVID-19 pandemic has spread obstreperous...	10.1007/s41748-020-00179-1	kcs6oukj	[The, COVID, 19, pandemic, has, spread, obstre...
...
360549	Update on cerebrovascular manifestations of CO...	2020-10-20	Neurol Sci	https://doi.org/10.1007/s10072-020-04837-0; ht...	The novel member of coronaviruses family, seve...	10.1007/s10072-020-04837-0	nr8fu0ze	[The, novel, member, of, coronaviruses, family...]
360697	Bacteremia and Blood Culture Utilization Durin...	2020-05-06	NaN	http://medrxiv.org/cgi/content/short/2020.05.0...	A surge of patients with coronavirus disease 2...	10.1101/2020.05.05.2008044	b5kw494g	[A, surge, of, patients, with, coronavirus, di...
361210	Estimating the generation interval for COVID-1...	2020-03-08	NaN	https://doi.org/10.1101/2020.03.05.20031815	Background: Estimating key infectious disease ...	10.1101/2020.03.05.20031815	cq6mvr9	[Background, Estimating, key, infectious, dise...

Hình 91. Hình E/IV/3: Thời kỳ ủ bệnh của virus

4. Tỷ lệ bệnh nhân không có triệu chứng

Thời gian từ khi một người phơi nhiễm với vi-rút cho đến khi họ có các triệu chứng (thời gian ủ bệnh). Nay giờ một người có thể bị nhiễm vi-rút trong 2-14 ngày trước khi họ cảm thấy mắc bệnh và một số người không bao giờ cảm thấy mắc bệnh.

Tỷ lệ bệnh nhân không có triệu chứng chiếm khoảng 78% trên tổng số. Điều này được dựa trên mô hình nhân khẩu học và mô hình cấu trúc theo độ tuổi cho covid 19. [67][68][69][70][71][72][73]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
36747	Coronavirus disease 2019-Historical context, v...	2020-08-05	Human vaccines & immunotherapeutics	https://doi.org/10.1080/21645515.2020.1787068;...	The current Coronavirus Disease 2019 (COVID-19) ...	10.1080/21645515.2020.1787068	lwdms33s	[The, current, Coronavirus, Disease, 2019, COV...]
44774	COVID-19 Infection Pandemic: From the Frontlin...	2020-10-16	Journal of the American College of Nutrition	https://doi.org/10.1080/07315724.2020.1779147;...	The infection caused by COVID-19 (i.e. corona ...	10.1080/07315724.2020.1779147	lsl0j6cp	[The, infection, caused, by, COVID, 19, i, e, ...]
58273	Dynamical evolution of COVID-19 in Italy with ...	2020-07-14	IEEE journal of biomedical and health informatics	https://doi.org/10.1109/jbhi.2020.3009038; ht...	The present work deals with an Ordinary Differ...	10.1109/jbhi.2020.3009038	npigomoe	[The, present, work, deals, with, an, Ordinary...]
65671	Addressing COVID-19 Among People Experiencing ...	2020-06-09	Public health reports	https://doi.org/10.1177/0033354920936227; http...	People experiencing homelessness are at high r...	10.1177/0033354920936227	cuemif94l	[People, experiencing, homelessness, are, at, ...]
73097	Why the immune system fails to mount an adapti...	2020-04-01	Transplant International : official journal of...	https://doi.org/10.1111/tri.13611; https://www...	The Covid-19 pandemic has already affected man...	10.1111/tri.13611	z2n51cet	[The, Covid, 19, pandemic, has, already, affec...]
...
355476	Clinical Performance Evaluation of a SARS-CoV...	2020-11-20	Int J Infect Dis	https://api.elsevier.com/content/article/pi/S...	OBJECTIVES: The true prevalence and seropositi...	10.1016/j.ijid.2020.11.164	i7kvv2xe	[OBJECTIVES, The, true, prevalence, and, serop...]

Hình 92. Tỷ lệ bệnh nhân không có triệu chứng

5. Bệnh nhi không có triệu chứng

[74][75][76][77][78] Các nghiên cứu đã báo cáo virus bocavirus ở người (HBoV) ở trẻ em bị nhiễm trùng đường hô hấp, nhưng chỉ thỉnh thoảng ở người lớn. Khi tìm kiếm HBoV DNA trong dịch hút mũi họng (NPA) từ người lớn bị đợt cấp của viêm phế quản mãn tính hoặc viêm phổi, từ trẻ em nhập viện vì từ trẻ em không có triệu chứng trong mùa đông 2002-2003 ở Canada. HBoV được phát hiện trong NPA của 1 (0,8%) của 126 người lớn có triệu chứng, 31 (13,8%) trong số 225 trẻ em có triệu chứng, và 43 (43%) trong số 100 trẻ em không có triệu chứng được phẫu thuật chọn lọc. Một loại virus khác được phát hiện ở 22 (71%) trong số 31 NPA dương tính với HBoV từ những trẻ có triệu chứng. Hai nhóm HBoV đã được xác định.

	title	publish_time	journal	url	abstract	doi	cord_uid	words
205	Human Bocavirus Infections in Hospitalized Chi...	2008-02-23	Emerg Infect Dis	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2...	Studies have reported human bocavirus (HBoV) ...	10.3201/eid1402.070851	uj7iyawj	[Studies, have, reported, human, bocavirus, HBo...
1142	Carriage of Mycoplasma pneumoniae in the Upper...	2013-05-14	PLoS Med	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	BACKGROUND: Mycoplasma pneumoniae is thought ...	10.1371/journal.pmed.1001444	d6awwygy	[BACKGROUND, Mycoplasma, pneumoniae, is, thoug...
10221	Impact of intestinal parasites on microbiota a...	2020-04-19	Parasit Vectors	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	BACKGROUND: Approximately 30% of children wor...	10.1186/s13071-020-04073-7	zfln4bk0	[BACKGROUND, Approximately, 30, of, children, ...]
31520	Variety of radiological findings in a family w...	2020-09-09	Egypt J Radiol Nucl Med	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	INTRODUCTION: In December 2019, a case series ...	10.1186/s43055-020-00293-0	k16gsa7e	[INTRODUCTION, In, December, 2019, a, case, se...
37098	Clinical utility of PCR for common viruses in ...	2014	Pediatrics	https://doi.org/10.1542/peds.2013-3042; https:...	BACKGROUND Acute respiratory illness (ARI) acc...	10.1542/peds.2013-3042	6ig9wmxm	[BACKGROUND, Acute, respiratory, illness, ARI,...]
...
352056	Outcome of children with rhinovirus detection ...	2018-10-19	Pediatr Transplant	https://doi.org/10.1111/petr.13301; https://ww...	Rhinoviruses are commonly detected in symptom... ...	10.1111/petr.13301	6yr14yk3	[Rhinoviruses, are, commonly, detected, in, sy...
352098	Respiratory viral infection in lower airways o...	2009-12-09	Acta Paediatr	https://www.ncbi.nlm.nih.gov/pubmed/20003105/...	Aim: The aim of this study was to determine if...	10.1111/j.1651-2227.2009.01627.x	p762ixkh	[Aim, The, aim, of, this, study, was, to, dete...
355553	SARS-CoV-2 Infections Among Children in the Bi...	2020-11-03	Clin Infect Dis	https://doi.org/10.1093/cid/ciaa1693; https://...	BACKGROUND: Children with SARS-CoV-2 infection...	10.1093/cid/ciaa1693	lkbv06dd	[BACKGROUND, Children, with, SARS, Cov, 2, inf...

Hình 93. Bệnh nhi không có triệu chứng 1

Một nghiên cứu dài hạn trong 15 tháng về bệnh tiêu chảy và nhiễm virus rota (RV) đã được tiến hành đồng thời ở trẻ sơ sinh và trẻ mới biết đi tại các trung tâm chăm sóc ban ngày (DCC) và tại một phòng khám nhi khoa lớn ở Houston. Ở những trẻ nhiễm RV có triệu chứng trong các bệnh DCC, 42% có RV được xác định trong mẫu phân trong vòng 2 ngày trước khi bị tiêu chảy. Ba mươi tám trẻ em DCC có nhiều hơn một đợt nhiễm RV, nhưng chỉ có năm em bị hai đợt nhiễm RV có triệu chứng. Tiêu chảy do RV thường gặp ở trẻ em trong các DCC, thường xảy ra thành dịch do cùng một chủng, và mắc bệnh song song trong cộng đồng. [79][80][81][82][83]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
142	La Crosse virus infectivity, pathogenesis, and...	2008-02-11	Virol J	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2...	BACKGROUND: La Crosse virus (LACV), family Bun...	10.1186/1743-422x-5-25	efxldnr	[BACKGROUND, La, Crosse, virus, LACV, family, ...]
1170	Correlation between Dengue-Specific Neutralizi...	2013-06-13	PLoS Negl Trop Dis	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3...	Although heterotypic secondary infection with ...	10.1371/journal.pntd.0002274	asghqx6q	[Although, heterotypic, secondary, infection, ...]
3976	Introducing a New Algorithm for Classification...	2019-04-26	JMIR Res Protoc	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6...	BACKGROUND: There is a need to better distingui...	10.2196/12705	d4fymkg9	[BACKGROUND, There, is, a, need, to, better, d...
4030	Clinical correlation of influenza and respirat...	2019-09-03	PLoS One	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6...	Acute respiratory tract infections are a major...	10.1371/journal.pone.0220908	1sl4e5v	[Acute, respiratory, tract, infections, are, a...
8562	Rotavirus in infant-toddler day care centers: ...	2006-03-29	J Pediatr	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	A 15-month prospective longitudinal study of d...	10.1016/s0023-3476(06)80624-3	kysrdfj5	[A, 15, month, prospective, longitudinal, stud...
...
357854	Heterogeneity in transmissibility and shedding...	2020-10-15	Nan	https://doi.org/10.1101/2020.10.13.20212233 ; h...	A growing number of studies provide insight in...	10.1101/2020.10.13.20212233	2ez9vnrt	[A, growing, number, of, studies, provide, ins...

Hình 94. Bệnh nhi không có triệu chứng 2

6. Lây truyền không có triệu chứng trong thời gian ủ bệnh

Khi virus phát triển đủ lớn cả về số lượng và cấu trúc, chúng sẽ phá vỡ tế bào để tràn ra ngoài và mỗi con virus sẽ tìm cách chui vào một tế bào mới. Cứ như thế, chu trình phát triển được lặp lại nhiều lần và ngày càng đi sâu hơn vào cơ thể. Đây gọi là giai đoạn ủ bệnh.

Đến một lúc nào đó, tuỳ thuộc vào sức khỏe mỗi người, số lượng virus đủ lớn sẽ phá vỡ cơ chế tự bảo vệ của con người, các triệu chứng sẽ phát ra, Covid-19 bắt đầu gây ra những ảnh hưởng xấu đến các bộ phận trong cơ thể. Đây là giai đoạn phát bệnh. Như vậy, trong giai đoạn ủ bệnh, virus đã có trong cơ thể người bệnh nhưng lại không có biểu hiện triệu chứng nên người khác không thể biết.

Giai đoạn ủ bệnh này nếu có xét nghiệm mẫu bệnh phẩm PCR thì kết quả có thể âm tính hoặc dương tính tuỳ theo lượng virus trong cơ thể. Có nhiều trường hợp xét nghiệm có kết quả âm tính nhưng vẫn có thể lây nhiễm cho những người tiếp xúc gần, vì vậy, nếu là một trong hai đối tượng: nghi ngờ tiếp xúc gần với người có nguy cơ mắc bệnh hoặc người đi về từ vùng dịch đều cần khai báo y tế để được cách ly theo dõi sức khỏe, phát hiện bệnh kịp thời để được điều trị và đặc biệt, không lây nhiễm ra cộng đồng. [84][85][86][87][88][89][90][91][92]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
34258	Cross-Sectional Seroepidemiologic Study of Cor...	2020-10-02	Int J Environ Res Public Health	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	(1) Background: Along with an increasing risk ...	10.3390/ijerph17197223	5br4faov	[, 1, Background, Along, with, an, Increasing...
39036	Household secondary attack rate of COVID-19 an...	2020-06-17	The Lancet. Infectious diseases	https://doi.org/10.1016/s1473-3099(20)30471-0;...	BACKGROUND As of June 8, 2020, the global rep...	10.1016/s1473-3099(20)30471-0	gt5pq29h	[BACKGROUND, As, of, June, 8, 2020, the, globa...
57582	Emergency management for preventing and contro...	2020-03-05	The British journal of dermatology	https://doi.org/10.1111/bjd.19011; https://www...	As of Feb 15, 2019, the novel coronavirus (201...	10.1111/bjd.19011	z74suepj	[As, of, Feb, 15, 2019, the, novel, coronaviru...
124249	Household secondary attack rate of COVID-19 an...	2020	Lancet Infect Dis		BACKGROUND: As of June 8, 2020, the global rep...		NaN	[BACKGROUND, As, of, June, 8, 2020, the, globa...
158680	Transmission of COVID-19 in the terminal stage...	2020	Int J Infect Dis		We report a familial cluster of 2019 novel cor...		p116erdu	[We, report, a, familial, cluster, of, 2019, n...
165333	Viral dynamics in asymptomatic patients with C...	2020	Int J Infect Dis		Data are limited on the viral load, viral shed...		thftzde3	[Data, are, limited, on, the, viral, load, vir...
165334	Viral dynamics in asymptomatic patients with C...	2020	Int J Infect Dis		Data are limited on the viral load, viral shed...		thftzde3	[Data, are, limited, on, the, viral, load, vir...
231210	Follow up investigation of asymptomatic COVID-...	2020	Epidemiol Health		OBJECTIVES: The objective of the study was to ...		op7ohlv3	[OBJECTIVES, The, objective, of, the, study, w...

Hình 95. Lây truyền không có triệu chứng trong thời gian ủ bệnh

7. Lịch sử tự nhiên của virus từ người bị nhiễm

Các ca nhiễm đầu tiên được phát hiện ở một chợ bán động vật và cá ở thành phố Vũ Hán, tỉnh Hồ Bắc của Trung Quốc từ cuối tháng 12 năm 2019. Trong khu chợ này có khoảng 1000 quầy bán gà, mèo, gà lôi, dơi, con duí (marmota), rắn độc, hươu đốm, các bộ phận của thỏ và nhiều loài động vật hoang dã khác, do đó dẫn đến sự nghi ngờ rằng nguyên nhân gây bệnh có thể là một loại coronavirus mới được bắt nguồn từ động vật. Việc dịch bệnh ở Vũ Hán có nguồn gốc liên quan đến một thị trường lớn chuyên bán hải sản và động vật để tiêu thụ đã dẫn đến khả năng bệnh bắt nguồn từ động vật. Điều này dẫn đến nỗi lo ngại rằng dịch bệnh sẽ tương tự như sự bùng phát SARS trước đó, thậm chí một mối lo ngại trầm trọng hơn trước bởi một lượng lớn người dân dự kiến sẽ đi du lịch vào dịp Tết Nguyên Đán này khiến cho dịch bệnh dễ dàng lây lan nhanh chóng ở nhiều nơi trên thế giới, bắt đầu từ ngày 25 tháng 1 năm 2020.

Chủng virus mới gây ra các trường hợp đã được xác nhận là SARS-CoV-2 (tên gọi tạm thời là coronavirus mới (2019-nCoV)), là chủng thứ bảy trong họ coronavirus được biết đến có khả năng lây bệnh cho người, với trình tự bộ gen được báo cáo là giống đến 79,5% so với trình tự bộ gen của SARS-CoV và có sự tương đồng đến 96% với các chủng coronavirus lây nhiễm ở loài dơi móng ngựa. [93][94][95][96][97][98][99]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
6751	Characterization of viral genomic mutations in...	2019-07-13	Virus Genes	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	Since February 2013, human infections with the...	10.1007/s11262-019-01678-8	9gjifk2w	[Since, February, 2013, human, infections, wit...
71931	Critically Ill Patients With the Middle East R...	2017	Critical care medicine	https://doi.org/10.1097/ccm.0000000000002621; ...	OBJECTIVES To describe patient characteristics...	10.1097/ccm.0000000000002621	xt2jgb8y	[OBJECTIVES, To, describe, patient, characteri...
90483	Critically Ill Patients With the Middle East R...	2017	Critical care medicine	https://doi.org/10.1097/ccm.0000000000002621; ...	OBJECTIVES To describe patient characteristics...	10.1097/ccm.0000000000002621	xt2jgb8y	[OBJECTIVES, To, describe, patient, characteri...
129599	Factors associated with prolonged viral shedi...	2020	Eur. respir. j.	NaN	BACKGROUND: The duration of viral shedding is ...	NaN	akipevw2	[BACKGROUND, The, duration, of, viral, sheddin...
137602	Fecal viral shedding in COVID-19 patients: Cil...	2020	Virus Res	NaN	BACKGROUND: To investigate the clinical signif...	NaN	n2hbmh7g	[BACKGROUND, To, investigate, the, clinical, s...
156422	Liver Injury in Critically Ill and Non-critica...	2020	Front. Med.	NaN	Background: Liver injury commonly occurs in pa...	NaN	owynryqc	[Background, Liver, injury, commonly, occurs, ...
157068	Analysis of factors affecting the prognosis of...	2020	Epidemiol Infect	NaN	The clinical characteristics of patients with ...	NaN	qpyki7y7	[The, clinical, characteristics, of, patients,...

Hình 96. Lịch sử tự nhiên của virus từ người nhiễm bệnh

8. Thời gian rụng trung bình của virus là gì?

[100][101][102][103][104][105][106][107][108][109] Các bài báo nêu ra đặc điểm bộ gen, nguồn nhiễm và cách thức lây nhiễm của virus trong cộng đồng:

Đặc điểm bộ gen: vi rút có vỏ bao, hạt vi rút hình tròn hoặc bầu dục, thường là đa diện với đường kính 60-140nm. Đặc điểm di truyền của vi rút SARS-CoV-2 khác với SARS và MER-CoV, chỉ 85% trình tự gen của vi rút giống với chủng gây SARS.

Nguồn truyền nhiễm cộng đồng: Người mắc bệnh có triệu chứng từ mức độ nhẹ đến nặng là nguồn truyền nhiễm chính lây lan dịch bệnh trong cộng đồng. Người mang vi rút hoàn toàn không có triệu chứng đường như có khả năng lây truyền thấp. Việc đánh giá vai trò lây truyền của người mang vi rút không triệu chứng vẫn đang được tiếp tục nghiên cứu, đánh giá thêm.

Cách thức lây truyền: Bệnh có thể lây trực tiếp từ người sang người qua hô hấp hoặc hít phải những giọt bắn từ mũi hoặc miệng của người bị COVID-19 phát tán khi ho, hắt hơi hoặc thở ra. Nếu hít hoặc nuốt phải những giọt bắn này từ người bị nhiễm SARS-CoV-2 thì sẽ có nguy cơ bị nhiễm bệnh. Đây là lý do tại sao phải cách xa người bệnh hơn 2 mét. Đến thời điểm này, hình thức này được coi là đường lây lan chính của bệnh. Bệnh có thể lây do người lành tiếp xúc với các bề mặt có SARS-CoV-2 trên. Những giọt bắn văng xa tới 2 mét do người bệnh phát tán khi ho, hắt hơi, thở ra, rơi xuống các vật thể và các bề mặt xung quanh người. Nếu hít phải những giọt bắn này từ người bị nhiễm SARS-CoV-2 thì sẽ có nguy cơ bị nhiễm bệnh. Đây là lý do tại sao phải cách xa người bệnh hơn 2 mét hoặc phải đeo khẩu trang để hạn chế giọt bắn văng ra xa.

	title	publish_time	journal	url	abstract	doi	cord_uid	words
6751	Characterization of viral genomic mutations in...	2019-07-13	Virus Genes	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7...	Since February 2013, human infections with the...	10.1007/s11262-019-01678-8	9gjffk2w	[Since, February, 2013, human, infections, wit...]
71931	Critically Ill Patients With the Middle East R...	2017	Critical care medicine	https://doi.org/10.1097/ccm.0000000000002621; ...	OBJECTIVES To describe patient characteristics...	10.1097/ccm.0000000000002621	xt2gb8y	[OBJECTIVES, To, describe, patient, character...]
90483	Critically Ill Patients With the Middle East R...	2017	Critical care medicine	https://doi.org/10.1097/ccm.0000000000002621; ...	OBJECTIVES To describe patient characteristics...	10.1097/ccm.0000000000002621	xt2gb8y	[OBJECTIVES, To, describe, patient, character...]
129599	Factors associated with prolonged viral shedi...	2020	Eur. respir. j	NaN	BACKGROUND: The duration of viral shedding is ...	NaN	akipeew2	[BACKGROUND, The, duration, of, viral, sheddin...]
137602	Fecal viral shedding in COVID-19 patients: Cll...	2020	Virus Res	NaN	BACKGROUND: To investigate the clinical signif...	NaN	n2hbmh7g	[BACKGROUND, To, investigate, the, clinical, s...]
156422	Liver Injury in Critically Ill and Non-critica...	2020	Front. Med.	NaN	Background: Liver injury commonly occurs in pa...	NaN	owynryqc	[Background, Liver, injury, commonly, occurs, ...]
157068	Analysis of factors affecting the prognosis of...	2020	Epidemiol Infect	NaN	The clinical characteristics of patients with ...	NaN	qpyki7y7	[The, clinical, characteristics, of, patients,...]
160953	Clinical course and treatment efficacy of COVI...	2020	Transbound. emerg. dis. (Internet)	NaN	Currently, COVID-19 has been reported in nearl...	NaN	c0k3jk9w	[Currently, COVID, 19, has, been, reported, in

Hình 97. Thời gian rụng trung bình của virus

9. Khoảng thời gian dài nhất của virus là bao lâu?

Trên lý thuyết, Covid-19 có khả năng sống sót trong môi trường bên ngoài cơ thể tương đối bền vững, nhất là ở điều kiện nhiệt độ thấp, thời tiết mát và lạnh, ẩm – thích hợp nhất là dưới 25°C – giống đặc tính chung cho cả họ virus corona.

Dù khả năng sinh tồn của virus này trong điều kiện nhiệt độ cao bị hạn chế, tuy nhiên, vẫn chưa có nghiên cứu đầy đủ, chính thức nào chứng minh về khả năng tồn tại của Covid-19 ở môi trường ngoài cơ thể với nhiệt độ cao. Với khí hậu ở Việt Nam, cụ thể như ở miền Nam (TP. HCM) thời gian gần đây, nhiệt độ buổi trưa khoảng 34-35°C khiến cho virus bị hạn chế sinh trưởng và lây lan, đồng nghĩa với việc giảm nguy cơ dịch bùng phát mạnh hơn. Nhưng ngược lại, ở miền Bắc nước ta với đặc điểm thời tiết lạnh và mưa nhiều, độ ẩm cao trong không khí là điều kiện thuận lợi cho Covid-19 bùng phát mạnh.

Hơn nữa, Covid-19 có thể tồn tại đến 3 – 4 ngày trên bề mặt các vật liệu kim loại, gỗ, vải, giấy, da tay, 4 – 5 ngày trong nước bọt và dịch tiết hô hấp của bệnh nhân. Chúng có thể sống và tồn tại với hoạt lực cao ở nhiệt độ thấp 4 – 20°C trong 5 ngày, chỉ mất khả năng lây nhiễm ở môi trường ngoài sau 30 phút (ở 56°C). Tia cực tím (tia UV) và các hóa chất y tế có tác dụng khử trùng ở nồng độ thường thì có thể tiêu diệt được virus trong vòng 60 phút. [110][111][112][113][114]

	title	publish_time	journal	url	abstract	doi	cord_uid	words
107078	Clinical characteristics associated with long-term...	2020	American Journal of translational research	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7319408/	BACKGROUND To delineate the clinical character...	NaN	mzoq2z78	[BACKGROUND, To, delineate, the, clinical, cha...
135730	Clinical characteristics associated with long-term...	2020	American Journal of Translational Research		Background: To delineate the clinical character...	NaN	w37f2ze0	[Background, To, delineate, the, clinical, cha...
141758	Case Report: Viral Shedding for 60 Days in a W...	2020	Am J Trop Med Hyg		NaN Novel coronavirus disease (COVID-19) caused by...	NaN	1vba9i42	[Novel, coronavirus, disease, COVID-19, cause...
187000	Repeated COVID-19 relapse during post-discharge...	2020	J. microbiol. immunol. infect		NaN A case who revealed the longest duration of vi...	NaN	n8h9gl7c	[A, case, who, revealed, the, longest, duratio...
232032	Clinical course and risk factors for mortality...	2020	Lancet		NaN BACKGROUND: Since December, 2019, Wuhan, China...	NaN	9819n3zf	[BACKGROUND, Since, December, 2019, Wuhan, Chi...
264924	Clinical course and risk factors for mortality...	2020-03-11	Lancet	https://www.sciencedirect.com/science/article/pii/S0140673620305663	BACKGROUND: Since December, 2019, Wuhan, China...	10.1016/j.lane.2020.07.017	k36rymkv	[BACKGROUND, Since, December, 2019, Wuhan, Chi...
295516	Repeated COVID-19 relapse during post-discharge...	2020-08-06	J Microbiol Immunol Infect	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3280768/	Abstract A case who revealed the longest durat...	10.1016/j.jmii.2020.07.017	6hg2e230	[Abstract, A, case, who, revealed, the, longes...

Hình 98. Khoảng thời gian dài nhất của virus

PHẦN KẾT LUẬN

1. Kết quả đạt được:

- ❖ Dự đoán được các giá trị kháng thể té bào B-cell covid 19 với độ chính xác tương đối cao góp phần giúp tạo ra kháng thể chống lại virus một cách nhanh chóng
- ❖ Phân tích đại dịch covid-19 ở một số mặt cơ bản:
 - Sự lây truyền, ủ bệnh và ổn định đối với môi trường _ phát họa cơ bản lên được bản đồ thế giới
 - Tốc độ lây truyền, phục hồi và tử vong trên thế giới và diễn hình ở một số nước nhất định
 - Số trường hợp nhiễm bệnh và số người chết trên thế giới
 - Sự nhiễm bệnh và lây truyền ở giới tính và các nhóm tuổi khác nhau
 - Tìm được sự tương quan giữa thời gian kể từ trường hợp đầu tiên và sự phục hồi
 - Thống kê được các trường hợp nhiễm bệnh ở một số quốc gia, lý do nhiễm bệnh và cách chống lại covid-19 ở các quốc gia đó
 - Lập bảng thống kê trả lời các câu hỏi liên quan đến bệnh nhân và các nghiên cứu liên quan đến covid

2. Những hạn chế:

- ❖ Mặc dù cơ bản đã hoàn thành được bài phân luận nhưng vẫn còn những hạn chế.
- ❖ Chưa trả lời được các câu hỏi về covid-19 như:
 - Những tiến bộ mới trong chẩn đoán SARS-COV-2 do tập dữ liệu vẫn khá nhỏ và những hiểu biết của bản thân về covid vẫn còn hạn chế
 - Dự đoán số người chết theo thời gian trên thế giới (chưa đưa ra được con số cụ thể chính xác)
 - Biểu hiện của COVID-19 ở bệnh cơ tim có thể xảy ra và ngừng tim
 - Tải lượng vi rút liên quan như thế nào đến biểu hiện bệnh, bao gồm khả năng xét nghiệm chẩn đoán dương tính?

3. Hướng phát triển:

- ❖ Nâng cao kiến thức về các mảng y tế, thuật ngữ chuyên của đại dịch Covid 19 để có thể hoàn thành các câu hỏi còn gian giỡ.
- ❖ Thu thập thêm nhiều số liệu và dữ liệu để làm rõ hơn các vấn đề liên quan đến đại dịch covid 19

TÀI LIỆU THAM KHẢO

- [1] [ĐẠI CƯƠNG VỀ CORONAVIRUS](#)
- [2] [Cấu trúc bộ gen của vi rút SAR CoV-2](#)
- [3] [Vắc xin peptide dựa trên Epitope được dự đoán chống lại bệnh coronavirus mới do SARS-CoV-2 gây ra - ScienceDirect](#)
- [4] [Một nghiên cứu về tin học miễn dịch để dự đoán các biểu thể trong màng bọc protein của SARS-CoV-2 | bioRxiv](#)
- [5] [Điểm đẳng điện - Wikipedia, \[https://en.wikipedia.org/wiki/Isoelectric_point\]\(https://en.wikipedia.org/wiki/Isoelectric_point\)](#)
- [6] [Đặc điểm của Coronavirus và Sars-Cov-2](#)
- [7] [Đánh giá ảnh hưởng của tính ky nước / tính ưa nước bề mặt đến khả năng rửa trôi chất ô nhiễm của than sinh học trong xử lý nước - ScienceDirect](#)
- [8] [Tính ổn định của SARS-CoV-2 trong các điều kiện môi trường khác nhau - Vi khuẩn Lancet, Những hiểu biết mới về tính ổn định bề mặt SARS-CoV-2 và tính nhạy cảm với nhiệt độ](#)
- [9] [Nguồn: Zhien Ma, Jia Li - Mô hình động lực học và phân tích dịch tỉ-Công ty xuất bản khoa học thế giới \(2009\)](#)
- [10] [CDC. \[https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/testing-in-us.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Ftesting-in-us.html\]\(https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/testing-in-us.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Ftesting-in-us.html\)](#)
- [11] [^ a ^ Holshue ML, DeBolt C, Lindquist S, Lofy KH và đồng nghiệp \(tháng 3 năm 2020\). “First Case of 2019 Novel Coronavirus in the United States”. *N. Engl. J. Med.* **382** \(10\): 929–936. PMID 32004427. doi:10.1056/NEJMoa2001191](#)
- [12] <https://www.rfi.fr/vi/qu%E1%BB%91c-t%E1%BA%BF/20201217-covid-19-hoa-k%E1%BB%B3-l%E1%BA%A1i-b%E1%BB%8B-k%E1%BB%B7-l%E1%BB%A5c-h%C6%A1n-3-700-ca-t%E1%BB%AD-vong-trong-v%C3%B2ng-24-gi%E1%BB%9D>
- [13] [Johns Hopkins Center for Health Security. <https://www.centerforhealthsecurity.org/resources/COVID-19/serology/Serology-based-tests-for-COVID-19.html>](https://www.centerforhealthsecurity.org/resources/COVID-19/serology/Serology-based-tests-for-COVID-19.html)
- [15] [SARS-CoV-2-Pandemie und ihre Auswirkungen auf Orthopädie und Unfallchirurgie: „Booster“ für die Telemedizin](#)
- [16] [Olfactory Device for Large Scale Pre-screening for COVID-19](#)
- [17] [Multifunctional nano-magnetic particles assisted viral RNA-extraction protocol for potential detection of COVID-19](#)

- [18] COVSACK: an innovative portable isolated and safe COVID-19 sample collection kiosk with automatic disinfection
- [19] Treatment of the novel COVID-19: Why Costa Rica's proposal for the creation of a global pooling mechanism deserves serious consideration?
- [20] Companion diagnostics: emerging strategies and issues in pharmaceutical development
- [21] United States FDA's emergency use authorization of Ebola virus diagnostics: current impact and lessons for the future
- [22] Syndromic Panel-Based Testing in Clinical Microbiology
- [23] Laboratory diagnosis of SARS-CoV-2 - A review of current methods
- [24] Experience, Perceptions, and Recommendations Concerning COVID-19-Related Clinical Research Adjustments
- [25] Designing and Interpreting COVID-19 Diagnostics: Mathematics, Visual Logistics, and Low Prevalence
- [26] Comparison of the Xpert Flu/RSV XC and Xpress Flu/RSV Assays
- [27] A Web-Based Platform on Coronavirus Disease-19 to Maintain Predicted Diagnostic, Drug, and Vaccine Candidates
- [28] Experience, Perceptions, and Recommendations Concerning COVID-19-Related Clinical Research Adjustments
- [29] Clinical features, diagnostics, and outcomes of patients presenting with acute respiratory illness: A retrospective cohort study of patients with and without COVID-19
- [30] Isothermal SARS-CoV-2 Diagnostics: Tools for Enabling Distributed Pandemic Testing as a Means of Supporting Safe Reopenings
- [31] More targeted use of oseltamivir and in-hospital isolation facilities after implementation of a multifaceted strategy including a rapid molecular diagnostic panel for respiratory viruses in immunocompromised adult patients
- [40] SARS-CoV-2 Shedding from Asymptomatic Patients: Contribution of Potential Extrapulmonary Tissue Reservoirs
- [41] Kinetics of Viral Clearance and Antibody Production Across Age Groups in Children with Severe Acute Respiratory Syndrome Coronavirus 2 Infection
- [42] Clinical sequelae of COVID-19 survivors in Wuhan, China: a single-centre longitudinal study
- [43] Clinical sequelae of COVID-19 survivors in Wuhan, China: a single-centre longitudinal study
- [44] Characterization, pathogenicity and protective efficacy of a cell culture-derived porcine deltacoronavirus

- [45] Case Report: Viral Shedding for 60 Days in a Woman with COVID-19
- [46] Human rhinovirus and coronavirus detection among allogeneic hematopoietic stem cell transplantation recipients
- [47] Mechanisms of rhinovirus-induced asthma
- [48] Human rhinovirus and coronavirus detection among allogeneic hematopoietic stem cell transplantation recipients
- [49] Epidemiological Characteristics of New Coronavirus Diseases (COVID-19): Features of Risk Factors and Clinical Features of the Child Population
- [50] Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020
- [51] Relapsing and remitting viral diseases of the nervous system
- [51] Probable Longer Incubation Period for Elderly COVID-19 Cases: Analysis of 180 Contact Tracing Data in Hubei Province, China
- [52] The Influence of Age on Steroidogenic Enzyme Activities of the Rat Adrenal Gland: Enhanced Expression of Cholesterol Side-Chain Cleavage Activity*
- [53] Demographic factors, attitude and knowledge of persons with special needs towards COVID-19 in Nigeria: Implications for counselling and social policy [version 2; peer review: 2 approved]
- [54] Household Secondary Attack Rate of COVID-19 and Associated Determinants
- [55] Epidemiological parameters of COVID-19 and its implication for infectivity among patients in China, 1 January to 11 February 2020
- [56] Data sharing during COVID-19 pandemic: what to take away
- [57] Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study
- [58] Public and healthcare providers awareness of Coronavirus (COVID-19) in Qassim Region, Saudi Arabia
- [59] Modeling the dynamics of the COVID-19 population in Australia: A probabilistic analysis
- [60] Significantly longer Covid-19 incubation times for the elderly, from a case study of 136 patients throughout China
- [61] Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease. Centers for Disease Control and Prevention
- [62] Bacteremia and Blood Culture Utilization during COVID-19 Surge in New York City
- [63] Importance of clinical epidemiology research in studies on respiratory diseases

- [64] Delivering urgent urological surgery during the COVID-19 pandemic in the UK: outcomes from our initial 52 patients
 - [65] Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020
 - [66] Epidemiological and clinical based study on four passages of COVID-19 patients: intervention at asymptomatic period contributes to early recovery
 - [67] Explaining the Bomb-Like Dynamics of COVID-19 with Modeling and the Implications for Policy
 - [68] Coronavirus disease 2019—Historical context, virology, pathogenesis, immunotherapy, and vaccine development
 - [69] The Coronavirus Disease 2019 (COVID-19) Pandemic
 - [70] Dynamical evolution of COVID-19 in Italy with an evaluation of the size of the asymptomatic infective population
 - [71] Addressing COVID-19 Among People Experiencing Homelessness: Description, Adaptation, and Early Findings of a Multiagency Response in Boston
 - [72] Why the immune system fails to mount an adaptive immune response to a COVID-19 infection
 - [73] Clinical performance evaluation of a SARS-CoV-2 Rapid Antibody Test for determining past exposure to SARS-CoV-2
 - [74] Human Bocavirus Infections in Hospitalized Children and Adults
 - [75] Carriage of Mycoplasma pneumoniae in the upper respiratory tract of symptomatic and asymptomatic children: an observational study
 - [76] Impact of intestinal parasites on microbiota and cobalamin gene sequences: a pilot study
 - [77] Variety of radiological findings in a family with COVID-19: a case report
 - [78] Clinical utility of PCR for common viruses in acute respiratory illness
 - [79] La Crosse virus infectivity, pathogenesis, and immunogenicity in mice and monkeys
 - [80] Correlation between Dengue-Specific Neutralizing Antibodies and Serum Avidity in Primary and Secondary Dengue Virus 3 Natural Infections in Humans
 - [81] Introducing a New Algorithm for Classification of Etiology in Studies on Pediatric Pneumonia: Protocol for the Trial of Respiratory Infections in Children for Enhanced Diagnostics Study
 - [82] Clinical correlation of influenza and respiratory syncytial virus load measured by digital PCR
 - [83] Rotavirus in infant-toddler day care centers: epidemiology relevant to disease control strategies
-

- [84] Cross-Sectional Seroepidemiologic Study of Coronavirus Disease 2019 (COVID-19) among Close Contacts, Children, and Migrant Workers in Shanghai
- [85] Household Secondary Attack Rate of COVID-19 and Associated Determinants
- [86] Emergency management for preventing and controlling nosocomial infection of the 2019 novel coronavirus: implications for the dermatology department
- [87] Transmission of COVID-19 in the terminal stages of the incubation period: A familial cluster
- [88] Viral dynamics in asymptomatic patients with COVID-19
- [89] The Impact of Model Building on the Transmission Dynamics under Vaccination: Observable (Symptom-Based) versus Unobservable (Contagiousness-Dependent) Approaches
- [90] What We Learned about COVID-19 So Far? Notes from Underground
- [91] Consideration for the asymptomatic transmission of COVID-19: Systematic Review and Meta-Analysis
- [92] Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study
- [93] Ultrafast Sample Placement on Existing Trees (UShER) Empowers Real-Time Phylogenetics for the SARS-CoV-2 Pandemic
- [94] Toroviruses of Animals And Humans: A Review
- [95] Urological implications of SARS CoV-19
- [96] Analytical insights of COVID-19 pandemic
- [97] COVID-19: Review and hematologic impact
- [98] Genomic characterization and phylogenetic analysis of SARS-CoV-2 in Italy
- [99] Biological and social aspects of Coronavirus Disease 2019 (COVID-19) related to oral health
- [100] Characterization of viral genomic mutations in novel influenza A (H7N9)-infected patients: the association between oseltamivir-resistant variants and viral shedding duration
- [101] Critically Ill Patients With the Middle East Respiratory Syndrome: A Multicenter Retrospective Cohort Study
- [102] Factors Associated With Prolonged Viral RNA Shedding in Patients with Coronavirus Disease 2019 (COVID-19)
- [103] Fecal viral shedding in COVID-19 patients: Clinical significance, viral load dynamics and survival analysis
- [104] Liver Injury in Critically Ill and Non-critically Ill COVID-19 Patients: A Multicenter, Retrospective, Observational Study

- [105] Analysis of factors affecting the prognosis of COVID-19 patients and viral shedding duration
- [106] Clinical course and treatment efficacy of COVID-19 near Hubei Province, China: A multicentre, retrospective study
- [107] Viral Transmission and Clinical Features in Asymptomatic Carriers of SARS-CoV-2 in Wuhan, China
- [108] Identification of RT-PCR-Negative Asymptomatic COVID-19 Patients via Serological Testing
- [109] Positive SARS-CoV-2 RNA recurs repeatedly in a case recovered from COVID-19: dynamic results from 108 days of follow-up
- [110] Clinical characteristics associated with illness perception in psoriasis
- [111] Case Report: Viral Shedding for 60 Days in a Woman with COVID-19
- [112] Repeated COVID-19 relapse during post-discharge surveillance with viral shedding lasting for 67 days in a recovered patient infected with SARS-CoV-2
- [113] Virologic and clinical characteristics for prognosis of severe COVID-19: a retrospective observational study in Wuhan, China
- [114] A special case of COVID-19 with long duration of viral shedding for 49 days
- [115] Trích từ: Zhiien Ma, Jia Li - Dynamical Modeling and Anaylsis of Epidemics-World Scientific Publishing Company (2009)
- [116] Rùng ngẫu nhiên như một phân loại: Một giải pháp dựa trên Spark
- [117] <https://qastack.vn/datascience/23789/why-do-we-need-xgboost-and-random-forest>
- [118]https://vi.wikipedia.org/wiki/H%e1%bb%93i_quy_tuy%e1%ba%bfn_t%c3%adnh
- [119] Emerging technologies for the detection of viral infections
- [120] How Nanophotonic Label-Free Biosensors Can Contribute to Rapid and Massive Diagnostics of Respiratory Virus Infections: COVID-19 Case
- [288] Petersen W, Bierke S, Karpinski K, Häner M. Đại dịch coronavirus và ảnh hưởng của nó đối với phẫu thuật chỉnh hình và chấn thương: phẫu thuật, rủi ro và cách phòng ngừa? Knie J. 2020 doi: 10.1007 / s43205-020-00052-1.
- [344] Viện Robert Koch (2020) Hồ sơ SARS-CoV-2 về Bệnh do Coronavirus-2019 (COVID-19).
https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html#doc13776792bodyText2. Truy cập ngày 5 tháng 4 năm 2020