بسم الله الرحمن الرحيم

دانشگاه شهید باهنر کرمان

دانشکده ریاضی و کامپیوتر، بخش آمار

استفاده از روش های یادگیری ماشین
جهت کمک به کادر درمان در تشخیص و درمان بیماری COVID-19

مجری طرح:

دکتر ایوب شیخی

بهار 1400

فصل اول : کلیات پژوهش

1-1 مقدمه..... 1

1-2 عنوان پژوهش..... 2

1-3 بیان مسئله..... 2

1-4 اهمیت و ضرورت پژوهش..... 5

1-4-1 تحقیقات داخلی در زمینه روش های یادگیری ماشین برای تشخیص بیماری کرونا..... 8

1-5 اهداف پژوهش..... 11

1-5-1 هدف اصلی..... 11

1-5-2 اهداف فرعی..... 12

1-6 قلمرو پژوهش..... 12

1-6-1 قلمرو مکانی پژوهش..... 12

1-6-2 قلمرو زمانی پژوهش..... 12

1-6-3 قلمرو موضوعی پژوهش..... 12

1-7 تعریف واژه ها..... 13

1-7-1 داده..... 13

1-7-2 مجموعه داده..... 13

1-7-3 یادگیری ماشین..... 14

1-7-4 کرونا..... 14

1-8 خلاصه..... 16

فصل دوم : مروری بر ادبیات پژوهش و پیشینه تحقیق

2-1

مقدمه.....18

2-2 بخش اول : کووید-19.....19

2-2-1 تعریف کووید- 19 و مقایسه با سایر بیماری های تنفسی.....19

2-2-2 راه های انتقال.....20

2-2-3 نام شناسی..... 21

2-2-4 کشف.....21

2-2-5 تاریخچه کووید-19 در ایران.....22

2-3 بخش دوم : مجموعه داده.....23

2-3-1-مجموعه داده های کووید-19.....23

2-4 بخش سوم: یادگیری ماشین در کمک به تشخیص کووید-19.....25

2-4-1 انواع یادگیری ماشین.....25

2-4-1-1 یادگیری نظارت شده.....26

2-4-1-2 یادگیری نظارت نشده.....26

2-4-1-3 یادگیری عمیق و شبکه عصبی.....27

2-4-2 روش های تشخیص.....28

2-4-2-1 تصاویر قفسه سینه.....28

2-4-2-2 سرفه.....30

2-4-3 بات های هوشمند.....32

2-5 خلاصه.....37

فصل سوم : روش شناسی پژوهش

3-1 مقدمه.....39

3-2 بخش اول : روش کار.....39

3-3 بخش دوم : فرآیند پژوهش 53

3-4 بخش سوم :ابزار گردآوری.....53

3-5 خلاصه.....57

فصل چهارم : تجزیه و تحلیل داده ها

4-1

مقدمه.....59

4-2 بخش اول : روش پاکسازی داده.....60

4-3 بخش دوم : مدل و الگوریتم مورد استفاده در پژوهش 65

4-4 بخش سوم :تجزیه و تحلیل داده ها.....77

4-5 خلاصه.....99

فصل پنجم : نتیجه گیری

5-1 نتیجه گیری و آینده تحقیق 101

مراجع.....107

ضمیمه‌ها و پیوست‌ها

پرسشنامه.....115

واژگان.....117

فصل اول:
کلیات پژوهش

1-1 مقدمه

با شیوع ویروس کرونا در جهان و گسترش بی رویه آن، بیشتر کشورهای دنیا به دنبال راهی برای مهار این ویروس هستند. مکانیسم‌های زیادی از جمله آنالیز بالینی، تصاویر سی تی اسکن قفسه سینه، نتایج آزمایش خون و فن آوری های دیگر برای شناسایی این بیماری به کار گرفته شده است. در حالی که این راه حل‌ها وقت گیر و پر هزینه هستند و نیاز به دقت زیادی دارند.

فن‌آوری های مبتنی بر یادگیری ماشین و هوش مصنوعی در کنار استفاده از کلان داده ها و به همراه شبکه های نوین امروزی قابلیت های زیادی برای متخصصان بهداشت عمومی و سیاست گذاران ارائه می‌دهند تا بهداشت و سلامت مردم را از طریق مداخلات مبتنی بر زمینه درمان، گسترش دسترسی به اطلاعات بهداشتی و خدمات و افزایش مسئولیت‌پذیری فردی برای سلامت متحول کنند.

این روزها که جهان با بیماری کوید-۱۹ دست به گریبان است، از هر نوآوری و تکنولوژی برای مبارزه با این بیماری استفاده می‌شود. مانند بسیاری از عرصه‌های دیگر، عرصه خدمات بهداشت و درمان نیازمند پشتیبانی فناوری های جدیدی مانند هوش مصنوعی[۱](AI) ، اینترنت اشیا[۲](IoT) ، کلان داده‌ها[۳] و یادگیری ماشین[۴] است. به تازگی، تجزیه و تحلیل داده‌های مربوط به بیماری، آماده سازی داده‌ها، پیشگیری و مبارزه با بیماری‌های واگیرداری مانند کوید-۱۹، به یکی از اهداف هوش مصنوعی تبدیل شده است.

هوش مصنوعی و یادگیری ماشین نقش مهمی در درک بهتر، مقابله با بحران کوید-۱۹ و کشف واکسن کوید-۱۹ دارند. فناوری یادگیری ماشین به رایانه ها این امکان را می‌دهد تا دانشی که از مجموعه داده ها بدست می اورند و با تقلید از حجم زیادی از داده‌ها، الگو و سرعت انتشار بیماری ها را پیش‌بینی کنند. این فناوری نتیجه محور، برای

[1] Artificial Indigence
[2] Internet of Things
[3] Big Data
[4] Machine Learning

غربالگری مناسب، تجزیه و تحلیل، پیش بینی و ردیابی بیماران فعلی و بیماران احتمالی آینده مورد استفاده قرار می‌گیرد. هوش مصنوعی از اطلاعات حاصل از افراد مبتلا به کرونا، افراد بهبود یافته و فوتی به عنوان داده استفاده می‌کند.

1-2 عنوان پژوهش

استفاده از روش های یادگیری ماشین جهت کمک به کادر درمان در تشخیص و درمان بیماری COVID-19

1-3 بیان مسئله

شاید واقعاً هیچ اتفاقِ چندان جدید یا بحرانی در دنیا رخ نمی‌دهد؛ بشر پیش از این بارها شاهد جنگ‌ها، شیوع بیماری های واگیردار مختلف و ناتوانی‌اش در برابر طبیعت بوده است. در این میان، آنچه بحران‌ها را قابل تحمل‌تر می‌سازد راه‌ها و میان‌برهایی است که انسان‌ها با استفاده از امکانات موجودشان در مواجهه با مساله می‌یابند. بیماری‌های واگیردار در عصر انقلاب دیجیتالی همچنان همان بیماری‌های خطرناک‌اند، با این تفاوت که درمان، کنترل و پیشگیری از شیوع بیشتر به کمک فناوری بسیار سریع‌تر و موثرتر انجام می‌شود. بسیاری از تحلیلگران معتقدند به رغم اینکه در زمانه مبهمی به سر می‌بریم، می‌توان امید داشت این بحران منتهی به پیشرفت‌های بسیار شود. به گزارش گلوبال اینسایت(Global Insight)[5] ، ارزش بازار سلامت دیجیتالی جهان در سال ۲۰۱۹ حدود ۱۰۶ میلیارد دلار تخمین زده شده است، پیش‌بینی می‌شود با توجه به شیوع ویروس کرونا و افزایش سرعت استفاده

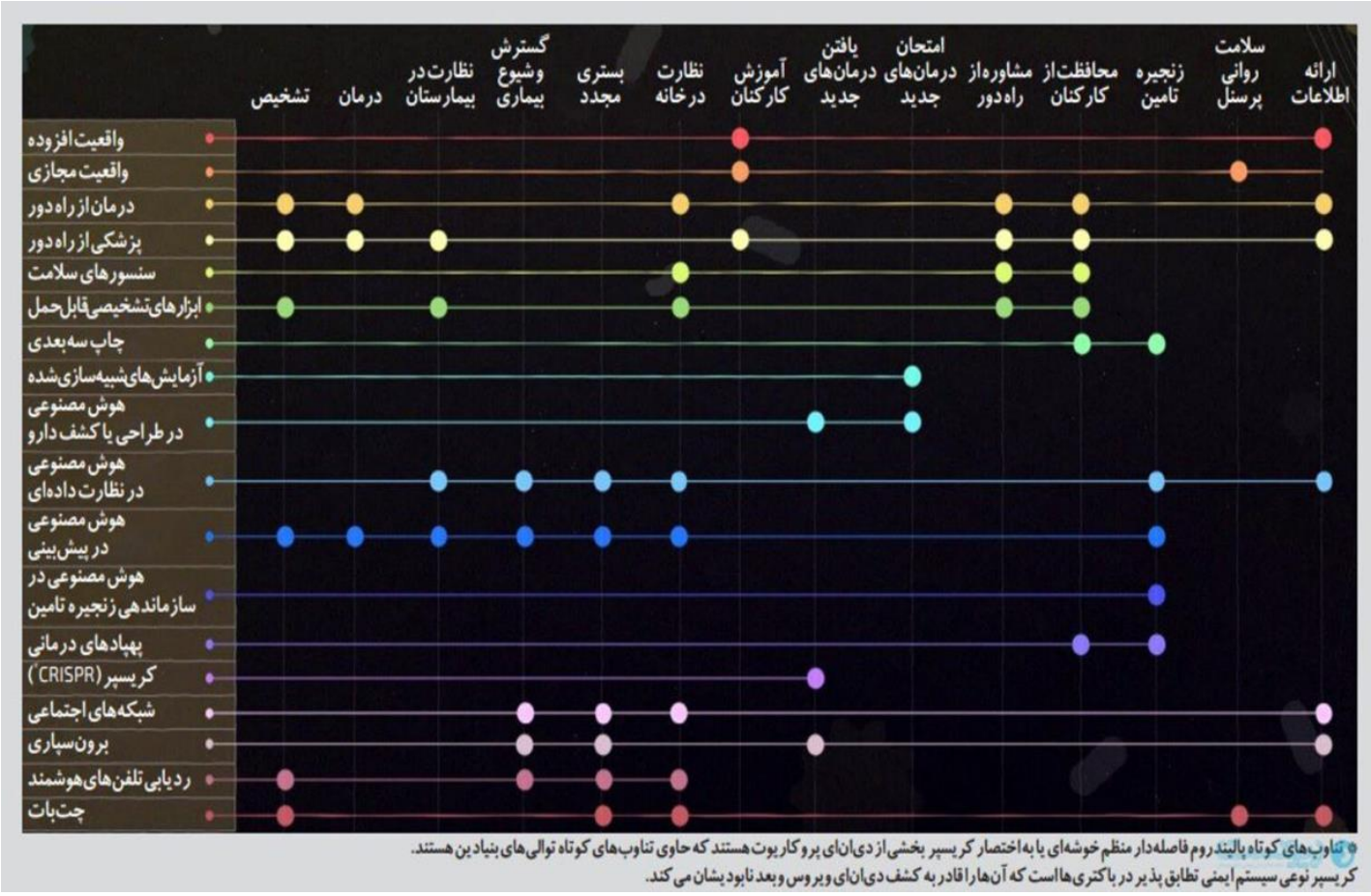
[5] https://ihsmarkit.com/research-analysis/nevermore-ignore-the-role-of-digital-health.html

از فناوری‌های دیجیتالی برای پیشگیری، درمان و مراقبت از بیماران، صنعت بهداشت و درمان بسیار زودتر از آنچه پیش‌بینی می‌شد تحول دیجیتالی را تجربه خواهد کرد.[1]

در عین حال که شیوع ویروس کرونا ادامه یافته و میلیون‌ها نفر در سراسر جهان به آن مبتلا شده‌اند و هر روز شاهد میزان مرگ و میر بالایی در سرتاسر جهان هستیم، کاربردهای فناوری در مبارزه با این بیماری نیز رو به افزایش است. فناوری‌های دیجیتالی به طور روزافزون در این زمینه به کار گرفته می‌شوند. با مروری بر نقش فناوری، سیاست‌گذاران و کادر درمان متوجه می‌شوند که چقدر ما در مبارزه با این ویروس جدید به فناوری وابسته‌ایم و کدام بخش‌ها و دوره‌های مراقبت و درمان را می‌توان با استفاده از فناوری‌های دیجیتالی ارتقا بخشید یا عملکردشان را تسهیل کرد. کادر درمان روزانه با تعداد بی‌سابقه‌ای از موارد ابتلای جدید روبه‌رو می‌شوند که این موضوع خود باعث افزایش فرسودگی و خستگی کادر درمان و مدافعان سلامت می شود. اقدامات برای حفظ فاصله اجتماعی با توجه به راحتی و شدت انتقال SARS-CoV-2 ضروری است. [2]

استفاده از فناوری های نوین بر پایه ی علم آمار و هوش مصنوعی راهکاری جدی برای کاهش خطر ابتلا و کمک به کادر درمان است؛ به طور مثال استفاده از اپلیکیشن درمان از راه دور آمول[6] از ماه ژانویه در ایالات متحده آمریکا ۱۵۸ درصد افزایش یافته و ملاقات‌های درمانی PlushCare نیز افزایش ۷۰ درصدی پیدا کرده است. قبل از پاندمی فقط یک نفر از هر ۱۰ نفر بیمار آمریکایی از خدمات درمان از راه دور استفاده می‌کرد. علاوه بر این در دوره پاندمی اپلیکیشن‌هایی مثل Doctor on Demand و HeyDoctor ارزیابی‌های رایگان کرونا را به خدمات از راه دور خود اضافه کرده‌اند. PlushCare و Everlywell هم چنین قصد دارند طرح کیت‌های آزمایش کرونا در خانه را راه‌اندازی کنند. پاندمی نیاز به درمان از راه دور را تقویت کرده است و احتمالاً حتی بعد از رفع این بحران، شاهد افزایش استفاده از این خدمات خواهیم بود.[3]

[6] Amwell



شکل 1-1 بررسی روش های تشخیص، درمان، نظارت و غیره در شناسایی و محدودسازی ویروس

منبع این تحقیق (شکل 1-1) سایت مدیکال فیوچریست[۷] بوده است از آنجا که ویروس جدید می‌تواند تعداد زیادی از افراد را آلوده کند، روش‌های تشخیص، درمان، نظارت، مشاوره و خیلی از چیز های دیگر که در جدول بالا آمده برای شناسایی ویروس و محدودسازی انتقالش کافی نیست. به همین دلیل دولت‌های سراسر جهان برای شناسایی و محدودسازی گسترش آن از فناوری استفاده می‌کنند و همان گونه که مشاهده می کنیم در جدول بالا هوش مصنوعی به کمک یادگیری ماشین جهت پیش بینی توانسته بهترین عملکرد را نسبت به سایرین داشته باشد .

[7] https://medicalfuturist.com/

هدف از این تحقیق جواب دادن به این سوال است که اگر احساس می‌کنید بیمار هستید یا تحت درمان قرار دارید، چه باید کرد؟ آیا باید به بیمارستان بروید و سلامت‌تان را بیشتر به خطر بیندازید؟ زیرا در حال حاضر بیمارستان‌ها پر از مبتلایان جدی است. آیا راهی وجود ندارد که بتوان مراجعات غیرضروری را کاهش داد؟

1-4 اهمیت و ضرورت پژوهش

در دوران بحران و با همه‌گیر شدن کوید-۱۹ یکی از نیازهای اصلی تحلیل و تفسیر داده است. دولت‌ها، نهادهای درمانی جهانی، مراکز تحقیقات علمی و صنعت، همگی در کنار هم به دنبال راه تازه‌ای برای جمع آوری، انباشته کردن و کار با اطلاعات هستند. ما به دیدن نتایج این کار در اخبار شبانه که آخرین آمار ابتلا و مرگ را به اطلاع می‌رساند، عادت کرده‌ایم. تکنولوژی‌های پیشرفته یکی ابزارهای مهم در مبارزه با این بحران محسوب می‌شوند و هوش مصنوعی از جمله تکنولوژی‌هایی است که در این مبارزه بسیار موثر خواهد بود. [4]

جمع‌آوری و تحلیل داده که در برخی صنایع کشور به کار گرفته می‌شود، می‌توانند از بحران کرونا جلوگیری، و آن را شناسایی و حل کند و ابزاری باشد برای پیش‌بینی موج‌های بعدی همه‌گیری ویروس، این در حالی است که استفاده از داده‌کاوی در مواجه با کرونا کمتر به کار گرفته شده است. هرچند اهمیت بخشی به نگه داری داده‌ها در ایران دیرتر در دیگر کشورها اتفاق افتاد (برای مثال در حالی اطلاعاتی دقیق از مسافران کشتی تایتانیک، گمشدگانی که بیش از پنجاه سال از نبودشان می‌گذرد در دیگر کشورها موجود است که از مهم‌ترین وقایع کشور ما همچون جنگ هشت ساله داده دقیقی موجود نیست) اما چند سالی است که تحلیل داده در صنعت، پزشکی، فرهنگ و دیگر حوزه‌های کشور ما نیز تا حدی جای خودش را پیدا کرده است. به گفته کارشناسان علم داده و هوش مصنوعی یکی از کاربردهای هوش مصنوعی و علم داده در حوزه تشخیص بیماری است. همچنین گفته می‌شود که یکی از چالش‌های این زمینه در ایران، نبود یا دسترسی نداشتن به داده است که (شاید) به دو دلیل عمده کیفیت پایین

سامانه های وزارت بهداشت و محرمانگی اطلاعات بیمار باشد. متخصصان علم داده داده می‌گویند که با توسعه فناوری اطلاعات در دانشگاه‌های علوم پزشکی و با به‌کارگیری متخصصان و فارغ‌التحصیلان این حوزه و بحث داده باز و محرمانه کردن داده می‌توان این دو مشکل عمده را حل کرد. و بیماری کرونا یک فرصت عالی برای بررسی دقیق‌تر استفاده از هوش مصنوعی به معنی واقعی در تشخیص بیماری است. کما اینکه تحلیل گران علم داده در جهان از پردازش تصاویر اشعه ایکس[8] برای تشخیص افراد مبتلا به ویروس کرونا استفاده کرده‌اند. در واقع ویروس کوید-۱۹ بر سیستم تنفسی افراد سالم تاثیر می‌گذارد و عکس شعه ایکس از قفسه سینه یکی از روش‌های مهم برای تشخیص ابتلای فرد به ویروس کرونا است. توسعه‌دهندگان با استفاده از دیتاست های موجود در kaggle.com می‌توانند مدل یادگیری ماشین مورد نظر جهت تشخیص فرد سالم از مبتلا به ویروس کرونا توسعه دهند. [5]

محققان دانشگاه فناوری و طراحی سنگاپور (SUTD) در تحقیقی در سال ۲۰۲۰ با استفاده از هوش مصنوعی اطلاعات مربوط به پاندمی کووید-۱۹ در سراسر جهان را تحلیل و زمان احتمالی پایان یافتن شیوع این بیماری در کشورهای مختلف جهان را پیش‌بینی کردند. (پیش‌بینی‌های محققان این دانشگاه درباره زمان پایان شیوع ویروس کووید-۱۹ در کشورهای مختلف جهان در وب سایت ddi.sutd.edu.sg قابل مشاهده است).[6]

اهمیت به داده‌ها در این دوران البته در کشور ما هم مورد توجه بوده است؛ برای مثال کمیته اپیدمیولوژی کووید- ۱۹ دانشگاه علوم پزشکی تهران، با همکاری تیمی از دانشجویان و با حمایت دانشکده بهداشت و موسسه ملی تحقیقات کشور به رصد همه‌گیری کووید-۱۹ و جمع‌آوری داده‌های مربوط و تهیه گزارش‌ها و راهنماهای علمی در این خصوص پرداخته است. با همه این ها البته در برخی بخش‌ها هنوز به نظر می‌رسد جمع‌آوری و تحلیل داده اهمیت خود را پیدا نکرده است. در بحث کرونا با استفاده از تحلیل داده اقدام‌های بسیاری می‌توان انجام داد، از لحاظ تعداد کسانی که مبتلا می‌شوند، از لحاظ تعداد افرادی که از میان خوب می‌شوند یا فوت می‌کنند و هم از لحاظ امکاناتی که هر روز در بیمارستان‌ها استفاده یا تخصیص داده می‌شود و... . همه این‌ها دارای اطلاعات،

[8] X-ray

شاخص‌ها و نرخ‌هایی است که در مدت زمان مشخصی جمع می‌شوند و قابل بررسی هستند. (این زمان لازم نیست بلند مدت یا کوتاه مدت باشد).

پیش از همه، این‌که زودتر از آن که اتفاقی به وقوع بپیوندد، ما می‌توانیم آن را تشخیص دهیم، سپس این‌که اگر اتفاقی افتاد، می‌توانیم آن را تشخیص دهیم که علتش چه بوده و چه چیز بیشترین تاثیر را بر آن داشته است، و دیگر این‌که متوجه می‌شویم از مجموعه‌ای از عناصر (مثلا افراد، تجهیزات و...) کدام یک خلاف حالت عادی رفتار می‌کنند و علت این تفاوت چیست تا اگر این تفاوت مشکل‌زاست، آن را برطرف کنیم؛ مثل پیدا کردن شهرهای قرمز، علت قرمز شدن آن‌ها و ارتباط شهرهای قرمز با یک‌دیگرچه بوده است. یا حتی از طریق آمارهای مربوط به مناطقی که در موج سوم و آن‌ها که در موج چهارم هستند، می‌توان جلوی موج چهارم را در جایی دیگر گرفت. بنابراین، تحلیل داده‌ها می‌تواند به میزان زیادی در جلوگیری از بحران، شناسایی بحران و حل آن به ما کمک کند.

اکثر الگوریتم‌های هوش مصنوعی در حال‌حاضر به سمت پیش‌بینی متمایل شده‌اند و سیستم ها با کمک هوش مصنوعی قرار است به دقت، زمان و مکان بحران‌های بعدی را پیش‌بینی کنند. تحقیقات در این زمینه مدت‌ها پیش آغاز شده است و در واقع هوش مصنوعی هشدارهایی درمورد بحران فعلی را در مراحل اولیه ارائه کرده بود. ابزاری به نام بلودات (Blue Dot) در تورنتو هنگام اسکن ۱۰۰هزار منابع اطلاعات دولتی و رسانه‌ای بود که در ۳۱ دسامبر سال ۲۰۱۹ احتمال شیوع بحرانی در ووهان چین را هشدار داد. [7][8][9]

هوش مصنوعی و یادگیری ماشینی علاوه بر حوزه درمان و شناسایی کرونا در زمینه‌های دیگری از جمله سرطان نیز در حال رشد هستند و امید می‌رود که سال آینده نه تنها برای کووید-۱۹ بلکه برای دیگر حوزه‌های درمان نیز سال بزرگی باشد. بنابراین هوش مصنوعی بدون شک جریان مهمی است که نحوه زندگی و دیگر حوزه های دانش را در آینده‌ای نزدیک دگرگون خواهد کرد.

1-4-1 تحقیقات داخلی در زمینه روش های یادگیری ماشین برای تشخیص بیماری کرونا

طرح تحقیقاتی مشترک بیمارستان سینا با دانشکده انفورماتیک دانشگاه علوم پزشکی شهید بهشتی در حال انجام است که این طرح منجر به طراحی یک نرم افزار می انجامد، با بررسی ۳ آیتم سی تی اسکن، صدای سرفه و علایم حیاتی بیمار، اطلاعات را پردازش و براساس آن احتمال ابتلا به بیماری کرونا را پیش بینی می کند. که این پژوهش هنوز جای کار دارد، به عنوان مثال می توان پارامتر هایی مانند نوار قلب و یا پارامترهای خونی را نیز برای پردازش و پیش بینی دقیق تر استفاده کرد.

دانشگاه های مطرح دنیا در حال پژوهش بر روی تشخیص کرونا از طریق سرفه به کمک روش های هوش مصنوعی هستند و که نتایج برخی از این پژوهش ها نیز به صورت مقاله منتشر گردیده است،از جمله دانشگاه هایی که بر روی این پروژه در حال کار هستند، دانشگاه MIT آمریکا، دانشگاه کمبریج انگلستان، دانشگاه EPFL سوئیس، دانشگاه کارنگی ملون آمریکا می باشند.

پژوهشکده فضای مجازی دانشگاه شهید بهشتی طرحی دیگر برای مقابله با ویروس کرونا را مطرح کرد. پیش بینی و تشخیص با تصاویر سی تا اسکن قفسه سینه از قابلیت پیشگویی و حساسیت بالایی در تشخیص کرونا برخوردارهستند.

همچنین برخی دیگر از بیماری‌های تنفسی علائمی مشابه با کرونا در این تصاویر دارند که باعث می‌شود عامل انسانی نتواند دقت زیادی در تفکیک بیماری کرونا از بیماری های مشابه داشته باشند. الگوریتم‌های هوش مصنوعی قادر هستند ضمن افزایش سرعت و یکنواختی تشخیص، دقت تفکیک کرونا از سایر بیماری‌های تنفسی را افزایش دهند.

طرحی که حاصل همکاری دانشگاه شهید بهشتی و بیمارستان سینا (دانشگاه علوم پزشکی تهران) است که برخی دیگر از بیماری‌های تنفسی علائمی مشابه با کرونا در این تصاویر دارند که باعث می‌شود عامل انسانی نتواند دقت

زیادی در تفکیک بیماری کرونا از بیماری‌های مذکور داشته باشد. الگوریتم‌های هوش مصنوعی قادرند ضمن افزایش سرعت و یکنواختی تشخیص، دقت تفکیک کرونا از سایر بیماری‌های تنفسی را افزایش دهند.

در حالت کلی در کاربردهای یادگیری ماشین و هوش مصنوعی در این زمینه با وجود کاربردهای متنوع هوش مصنوعی در مطالعات بالینی و خدمات بهداشتی، آنها به دو دسته عمده تقسیم می شوند: تجزیه و تحلیل داده های ساختاری، از جمله تصاویر، ژن ها و نشانگرهای زیستی، و تجزیه و تحلیل داده های غیر ساختاری، مانند یادداشت ها، مجلات پزشکی یا نظرسنجی بیماران برای تکمیل داده های ساختار یافته. رویکرد قبلی توسط الگوریتم های یادگیری ماشین و یادگیری عمیق تقویت می شود، در حالی که روش دوم به روش های تخصصی پردازش زبان طبیعی بستگی دارد.

به عنوان مثال :

هوش مصنوعی می تواند ویژگی ها را از حجم زیادی از داده های مراقبت های بهداشتی بیاموزد و سپس از تفسیر های به دست آمده برای کمک به تمرینات بالینی در طراحی درمان یا ارزیابی ریسک استفاده کند.

سیستم هوش مصنوعی می تواند اطلاعات مفیدی را از جمعیت زیادی از بیماران استخراج کند تا به استنباط زمان واقعی هشدار خطر سلامتی و پیش بینی نتیجه سلامتی کمک کند.

هوش مصنوعی می تواند کارهای تکراری، مانند تجزیه و تحلیل آزمایشات، اشعه ایکس، سی تی اسکن یا ورود اطلاعات را انجام دهد.

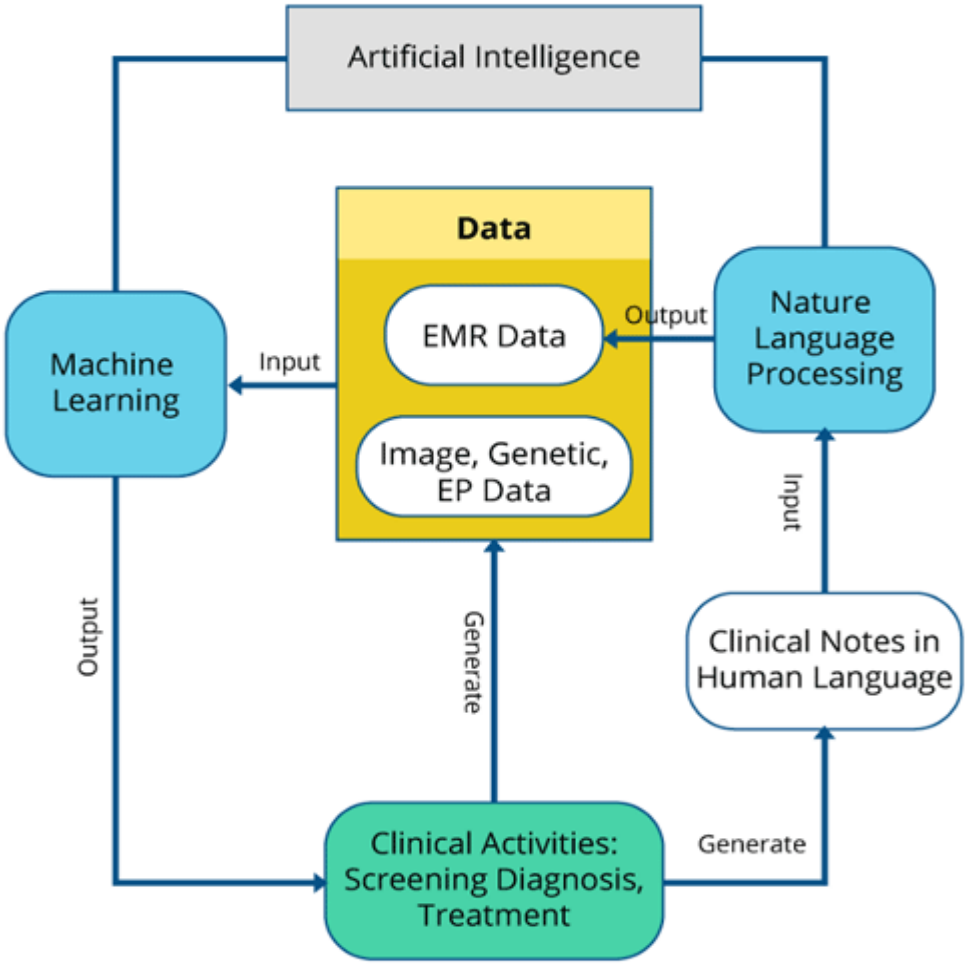
سیستم های هوش مصنوعی می توانند به کاهش خطاهای تشخیصی و درمانی کمک کنند که در عمل بالینی انسان اجتناب ناپذیر است.

هوش مصنوعی می تواند با ارائه اطلاعات پزشکی به روز از ژورنال ها، کتاب های درسی و روش های بالینی، به پزشکان کمک کند تا از مراقبت صحیح بیمار مطلع شوند.

هوش مصنوعی می تواند سوابق پزشکی را مدیریت کرده و عملکرد هر یک از موسسات و کل سیستم مراقبت های بهداشتی را تجزیه و تحلیل کند.

هوش مصنوعی می تواند به تولید داروی دقیق و داروهای جدید مبتنی بر پردازش سریع جهش ها و پیوند با بیماری کمک کند.

هوش مصنوعی می تواند مشاوره دیجیتال و خدمات نظارت بر سلامتی را در حد "پرستار دیجیتال" یا "ربات های بهداشتی" ارائه دهد.[18][19]



شکل1-2 عملکرد کلی هوش مصنوعی

1-5 اهداف پژوهش

امروزه تمامی کشورهای جهان می کوشند تا گسترش این ویروس را کنترل کنند. مکانیسم های زیادی از جمله آنالیز بالینی، پردازش تصاویر سی تی اسکن قفسه سینه، نتایج آزمایش خون، پردازش صوت بیماران کرونایی و فن آوری های هوش مصنوعی برای شناسایی این بیماری به کار گرفته شده است .در حالی که این راه حل ها وقت گیر و پر هزینه هستند و یا نیاز به دقت بیشتر دارند. فن آوری های مبتنی بر هوش مصنوعی در کنار استفاده از کلان داده ها و به همراه شبکه های نوین امروزی قابلیت های زیادی برای متخصصان بهداشت عمومی و سیاست گذاران ارائه می دهند تا بهداشت و سلامت مردم را از طریق مداخلات مبتنی بر زمینه درمان، گسترش دسترسی به اطلاعات بهداشتی و خدمات و افزایش مسئولیت پذیری فردی برای سلامت متحول نمایند. هدف اصلی این پژوهش نشان دادن اثربخشی هوش مصنوعی، کلان داده ها و شبکه های هوشمند برای مبارزه با بیماری همه گیرکووید-19 و بررسی راه حل های پیشرفته با استفاده از این فن آوری ها است. در این پژوهش ابتدا سعی شده است تا مروری بر فناوری های مبتنی بر هوش مصنوعی ، کلان داده و شبکه های نوین تاثیرگذار در تشخیص بیماری کرونا ارائه گردد و در ادامه تلاش دارد تا بتواند تاثیرگذاری استفاده مناسب از این فناوری ها را در تشخیص به موقع و سریع این بیماری همه گیر ارزیابی بکند و همچنین ارائه ی راهکارهای جدید مطابق بر قلمرو پژوهشی خود داشته باشد.

بنابراین اهداف پژوهش به شرح زیر است.

1-5-1 هدف اصلی

استفاده از روش های یادگیری ماشین جهت کمک به کادر درمان در تشخیص و درمان بیماری COVID-19 در استان کرمان

1-5-2 اهداف فرعی

- طراحی اپلیکیشن جهت شناسایی افراد مبتلا به ویروس کرونا در استان کرمان
- تشکیل مجموعه داده ای[9] از علائم بیماری کوید-19 (تب، لرز، سرفه، مشکلات گوارشی و...) بیماران کرونایی استان کرمان جهت تشخیص کرونا با استفاده از اپلیکیشن
- با استفاده از الگوریتم های یادگیری ماشین از قبیل پیش بینی، خوشه بندی و درخت تصمیم جهت تحلیل داده های بند 2
- محاسبه احتمال ابتلا افراد به بیماری کرونا با استفاده از داده های جمع آوری شده در بند 2

1-6 قلمرو پژوهش

1-6-1 قلمرو مکانی پژوهش

قلمرو مکانی پژوهش "داده و اطلاعات بیماران کرونایی استان کرمان" است.

1-6-2 قلمرو زمانی پژوهش

قلمرو زمانی پژوهش، از 16 فروردین تا 16 بهمن 1400 است.

1-6-3 قلمرو موضوعی پژوهش

قلمرو موضوعی پژوهش، شامل بررسی توانایی روشهای یادگیری ماشین کمک به کادر درمان استان کرمان

[9] Dataset

1-7 تعریف واژه‌ها

1-7-1 داده

داده در مفهوم، اطلاعات یا دانشی است که به صورت خام ارائه می شود و می توان آن را به شیوه ی مناسبی برای استفاده یا پردازش کدگذاری کرد. داده، به اطلاعات ذخیره شده گفته می شود که در آینده قصد استفاده از آن را داریم. معمولاً منظور از داده، اعداد هستند اما ممکن است به کلمات، صدا، عکس، ویدئو و هر حقیقت شناخته شده ی دیگری نیز داده گفته شود. به طور خلاصه، داده به اطلاعات و به ویژه حقایق یا اعدادی گفته می شود که به منظور بررسی یک یا چند ویژگی (متغیر) جمع آوری شده است و برای کمک به تصمیم گیری مورد استفاده قرار می گیرد.

1-7-2 مجموعه داده

به مجموعه ای مرتب از داده ها می گویند که با موضوعیت واحد جهت انجام کارها و پروژه های مربوط به علم داده استفاده می شود؛ مجموعه داده ها می توانند دسته ای از جداول، تصاویر، صدا و سایر اشیا باشند که ساده ترین مجموعه داده ای یک جدول است. داده ها در مجموعه داده اساساً بر اساس یک مدل خاص سازماندهی شده اند که به پردازش اطلاعات مورد نیاز کمک می کند.

یک کاربرد دیگر مجموعه داده ها نیز برای مقایسه بین روش های مختلف به این صورت است که به طور نمونه بر روی مجموعه داری A، دو روش یا الگوریتم مختلف را اجرا کرده و با توجه به نتایج می توان بر اساس معیارهای مختلف مانند سرعت دقت و پیچیدگی هر یک از روش ها را مقایسه کرد.

1-7-3 یادگیری ماشین

یادگیری ماشین شاخه‌ای از هوش مصنوعی است که به سیستم توانایی یادگیری خودکار و بهبود، بدون برنامه ریزی صریح را می‌دهد. یادگیری ماشین بر توسعه برنامه های رایانه ای متمرکز است. این برنامه ها می‌توانند به داده ها دسترسی داشته باشند و از آن‌ها برای اهداف مختلف استفاده کنند. فرایند یادگیری با مشاهدات و با استفاده از برخی داده‌ها مانند مثال‌ها، تجربه مستقیم یا دستورالعمل‌ها آغاز می‌شود. هدف از یادگیری ماشین این است که بر اساس داده‌هایی که اکنون به ماشین‌ها می‌دهیم، بتوانیم الگوهای موجود در آن‌ها را پیدا کنیم. در نهایت نیز به کمک این الگوها برای آینده تصمیمات بهتری بگیریم.

به عبارت دیگر هدف اصلی این است که به کامپیوترها اجازه دهیم بدون دخالت و کمک انسان به طور خودکار یاد بگیرند و اقدامات خود را بر اساس آن تنظیم کنند.

یادگیری ماشین تجزیه و تحلیل مقادیر عظیم داده را امکان پذیر می‌کند. اگرچه این علم به طور کلی نتایج سریعتر و دقیق‌تری را برای شناسایی فرصت‌های سودآور یا خطرات ارائه می‌دهد؛ اما ممکن است برای افزایش دقت و صحت آن به زمان و منابع بیش‌تری نیاز باشد. تلفیق یادگیری ماشین با هوش مصنوعی و فناوری‌های شناختی می‌تواند ماشین‌ها را برای پردازش حجم زیادی از اطلاعات آماده کند. در نظریه‌ی یادگیری ماشین علومی چون ریاضیات، آمار و علم رایانه به کار گرفته می شود.

1-7-4 کرونا

آنچه که به تازگی موجب ترس و بیماری بسیاری از هم وطن های ما شده است، ویروس جدیدی است که اولین بار در شهر ووهان چین شناسایی شد. این ویروس جدیدترین آن ها از خانواده ویروس های کرونا است. این ویروس

به نام COVID-19 نامگذاری شده است. نسخه های قدیمی تر این ویروس مانند سارس[10] هم در چین و کشورهای دیگر از جمله ایران تا حدودی همه گیر شدند. ویروس‌های کرونا گروه بزرگی از ویروس‌ها هستند که می‌توانند حیوانات و انسان‌ها را آلوده کنند و باعث بروز ناراحتی‌های تنفسی شوند؛ این ناراحتی‌ها ممکن است به اندازه سرما خوردگی خفیف یا به اندازه ذات الریه شدید باشند. همان گونه که در شکل 1-3 چرخه ی سرایت را مشاهده می کنید ویروس‌های کرونای حیوانی می توانند انسان‌ها را آلوده کنند و سپس بین آنها سرایت پیدا می‌کنند. شاید ویروس سارس (نشانگان تنفسی حاد) در سال 2002 الی 2003 را به خاطر بیاورید؛ این ویروس نمونه‌ای از ویروس کرونا بود که از حیوانات به انسان‌منتقل شد. یکی دیگر از موارد مهم و جدیدتر نژاد ویروس‌کرونا، مرس[11](نشانگان تنفسی خاورمیانه) نام دارد که در سال 2012 در خاورمیانه کشف شد و به گفته دانشمندان این ویروس ابتدا از شتر به انسان انتقال پیدا کرده است.

موارد زیر علائم ویروس کرونای جدید هستند:

تب یا لرز

سردرد

سرفه، گلودرد، انسداد بینی یا آبریزش بینی

از دست دادن ناگهانی حس چشایی یا بویایی

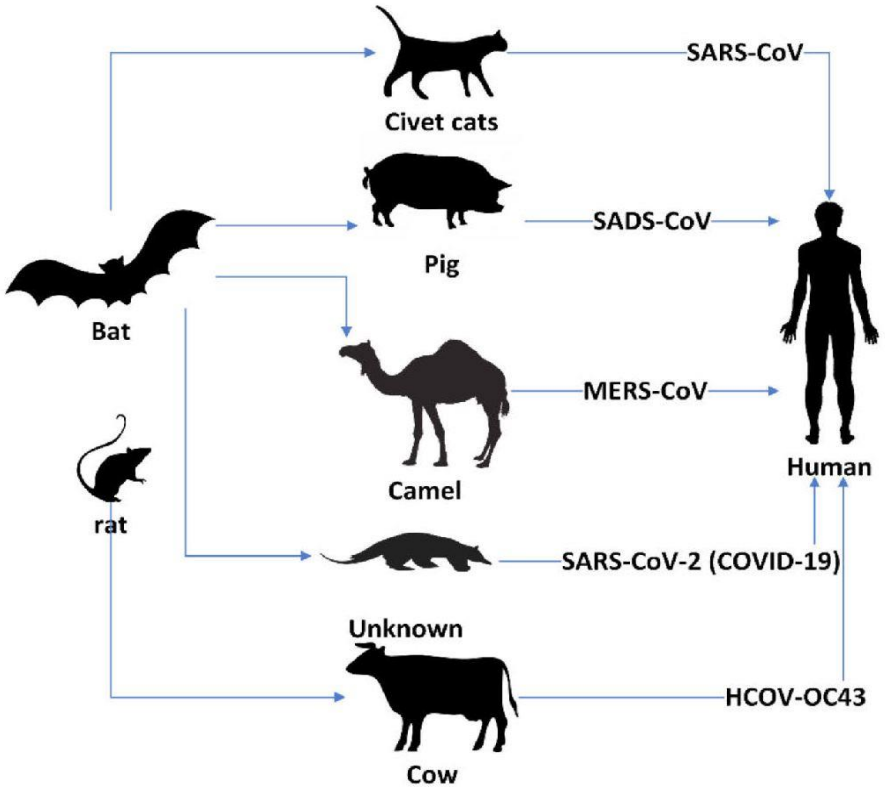
تنگی نفس یا مشکل در تنفس

حالت تهوع، استفراغ یا اسهال

سرگیجه، یا درد عضلات یا بدن

[10] SARS

[11] نوعی بیماری در ریه هاست که توسط کروناویروس ایجاد می شود. پزشکان اولین نمونه های مرس را در سال مشاهده کردند. این یک بیماری جدی است که می تواند مهلک باشد. بیشتر موارد ابتلا به مرس در کشورهای داخل یا اطراف شبه جزیره عربستان رخ داده اند. 2012 موارد گزارش شده در سایر مناطق همگی دارای سابقه در شبه جزیره عربستان بوده اند.



شکل 1-3 چرخه ی سرایت بیماری های کرونا از حیوانات به انسان ها

1-8 خلاصه

پژوهش حاضر با هدف استفاده از روش های یادگیری ماشین جهت کمک به کادر درمان در تشخیص و درمان بیماری کووید-19 در استان کرمان شکل گرفته است. هدف از تدوین این فصل که کلیات پژوهش نام دارد، ارائه‌ی یک تصویر کلی از این پژوهش است. بر این اساس در فصل اول این پژوهش، به کلیاتی در خصوص بیان مسئله، اهمیت و ضرورت پژوهش، اهداف پژوهش، قلمرو پژوهش و تعریف واژه‌ها پرداخته شده است که بدین ترتیب بستری را جهت اجرای سایر فصل‌های این پژوهش فراهم می‌آورد.

فصل دوم :

مروری بر ادبیات پژوهش و پیشینه تحقیق

2-1 مقدمه

کرونا ویروس ها خانواده بزرگی از ویروس ها هستند که از ویروس سرما خوردگی معمولی تا عامل بیماری های شدیدتری همچون سارس و کووید-19 را شامل می شوند. از ۴0 گونه مختلف خانواده کروناویروس، تاکنون هفت گونه منتقل شده به انسان کشف شده است [20] که موجب بروز بیماری هایی همچون خانواده سرما خوردگی در انسان میگردند. ویروس کرونا ویروس ۲۰۱۹ (کووید-19) از دسامبر ۲۰۱۹ در سراسر جهان رواج داشته است در ۱۱ مارس ۲۰۲۰، سازمان بهداشت جهانی این ارزیابی را انجام داده است که کووید-19را میتوان به عنوان یک بیماری همه گیرمعرفی کرد. [21]

مانند بسیاری از عرصه های دیگر، عرصه خدمات بهداشت و درمان نیز نیازمند پشتیبانی فناوری های جدیدی مانند هوش مصنوعی اینترنت اشیا، کلان داده ها و یادگیری ماشین است. هوش مصنوعی و یادگیری ماشین نقش مهمی در درک بهتر، مقابله با بحران کووید-19دارند. تکنیک های هوش مصنوعی می توانند یک ابزار غربالگری برای افراد ناقل در مقیاس بزرگ و در زمان واقعی در مهار گسترش کووید-19 تولید کنند. بیماری کووید-19 اولین بیماری اپیدمیک حاد تنفسی کرونا ویروسی نیست. در دو دهه گذشته، کرونا ویروس‌ها منجر به سه بیماری اپیدمیک حاد تنفسی به نام های کووید-19، سندرم حاد تنفسی یا بیماری سارس و سندرم تنفسی خاورمیانه یا مرس شده اند.

در این فصل از پژوهش، به ادبیات یادگیری ماشین در زمینه تشخیص بیماری کووید-19 پرداخته شده است. بر این اساس، ابتدا با استفاده از منابع موجود، تعریف، تاریخچه و زبانشناسی کووید-19، بیان شده است و سپس به تناسب عنوان بحث، پیشینه تحقیق محلی، داخلی و خارجی درباره این بیماری ارائه شده است.

2-2 بخش اول: کووید-19

2-2-1 تعریف کووید-19 و مقایسه آن با سایر بیماری های تنفسی

در پایان سال 2019، ویروس جدید کرونا کووید-19 به عنوان دلیل ایجاد یک گروه از مشکلات ذات الریه در ووهان[۱۲]، شهری در استان هوبئی[۱۳] چین شناخته شد؛ این بیماری به سرعت گسترش یافت و منجر به یک اپیدمی در سراسر چین و به دنبال آن در سایرکشورهای جهان شد.

طبق نتایج، ویروس جدید کرونا کووید-19 یا ویروس کرونای سندرم حاد تنفسی حاد2 [۱۴] از منشا خود در شهر ووهان چین به وسیله خفاش ها به بقیه مناطق جهان با سرعت گسترش یافته است[22] و براساس سن، سابقه بیماری و موارد دخیل دیگر باعث مرگ و میر بسیاری شده است و همچنین کمبود های پزشکی، هزینه های بالا، زمانبر بودن آزمایش ها، افراد ناقل از دلایل اصلی گسترش کرونا هستند. از طرفی مراجعه به پزشک و شلوغی بیمارستان ها خود دلیلی برای گسترش ویروس است و استفاده از روش های یادگیری ماشین می تواند راه حل مناسبی برای کمک به کادر درمان جهت حل مشکلات ذکرشده باشد.

در بسیاری از افراد در هنگام بروز تعدادی از علایمی مانند گلودرد، سرفه و تب بالا این سوال مطرح می شود که به کدام یک از بیماری های کووید-19، آنفولانزا و سرماخوردگی دچار شده شده اند. همه این بیماری ها توسط ویروس هایی ایجاد می شوند که دستگاه تنفسی را آلوده می کنند. همه مسری هستند و به راحتی از فردی به فرد دیگر سرایت می کنند. و برخی علائم مشابه ایجاد می کنند. بنابراین تشخیص آنها از یکدیگر دشوار است.

[12] Wuhan
[13] Hubei
[14] SARS-Cov2

علائم کووید-19 می تواند بسیار شبیه علائم سرماخوردگی یا آنفولانزا شامل شامل گلودرد، آبریزش بینی یا گرفتگی بینی و عطسه، تب، سردرد، بدن درد و مشکلات گوارشی باشد. اما یک علامت که بیشتر در کووید-19 اتفاق می افتد از دست دادن حس بویایی و چشایی است.

2-2-2 راه های انتقال

بر خلاف اغلب بیماری‌های عفونی که انتقال عوامل میکروبی آن هم زمان با بروز علائم بیماری اتفاق می‌افتد، این امر در مورد بیماری کووید-19 صدق نمی‌کند[23] زمان انتشار ویروس کووید-19 با عفونت های کرونا ویروسی انسانی سارس و مرس نیز متفاوت است و علی‌رغم آنکه انتشار ویروس های سارس و مرس همزمان با بروز علائم تنفسی (سرفه، عطسه، تب وغیره) آغاز می‌شود، زمان انتقال ویروس کووید-19، قبل از بروز علائم و در دوره کمون[15] بیماری نیز امکان پذیر است. مطالعات نشان داده اند که زمان پراکنش ویروس کرونا کووید-19به محیط از دستگاه تنفسی فرد ناقل، بطور متوسط از ۵/۲ روز قبل از بروز علائم شروع می شود و اوج آن نیم روز قبل از بروز علائم می‌باشد.[24]

محققان معتقدند در ٪۴۴ از موارد انتقال ویروس کووید-19 قبل از آنکه فرد از آلودگی خود مطلع شود، اتفاق می‌افتد. [26] از این رو، به نظر می‌رسد تحت شرایط کنونی که در محدودیت‌های روش‌های تشخیص بیماری کووید-19 مواجه به سر می بریم، انتشار ویروس پیش از بروز علائم اولیه آن اساسی ترین عامل گسترش جهانی بیماری کووید-19 می‌باشد.

امکان انتقال ویروس تا چند روز پس از بهبودی فرد، از بیماری کووید-19 نیز امکان پذیر است و عدم مشاهده یا کاهش علائم تنفسی و تب در بیماران نباید به منزله بهبودی کامل او و از بین رفتن خطر انتقال ویروس در نظر

[15] incubation period

گرفته شود. برای اطمینان از زمان غیر عفونی شدن افراد آلوده به ویروس کووید-19، انجام تست های تشخیصی مولکولی (RT-qPCR) بر روی نمونه های دستگاه تنفسی و سرولوژی خون ضروری است.[26] مرکز کنترل و پیشگیری بیماری ها بر اساس نتایج حاصل از تست های تشخیصی بیماری کووید-19، زمانی را که احتمال خطر انتقال ویروس کووید-19 از افراد آلوده به افراد سالم به حداقل ممکن می رسد، حداقل ۳ روز پس از بهبودی کامل علائم تنفسی همچون تب و تنگی نفس و حداقل ۱۰ روز پس از مشاهده این علائم، در نظر گرفته است.

2-2-3 نام شناسی

واژه «کروناویروس» از کلمه زبان لاتین" corōna" یا کلمه یونانی" κορώνη" به معنی تاج یا هاله گرفته شده است. این واژه به مشخصه ظاهری ویریون ها، شکل عفونی ویروس که در زیر میکروسکوپ الکترونی دیده میشود، اشاره دارد که حاشیه ای از سطح بزرگ و پیازدار داشته و یادآور تصویری از یک تاج سلطنتی، یا تاج خورشیدی است؛ از این رو کروناویروس را ویروس تاجدار نیز می نامند.

2-2-4 کشف

کرونا ویروس ها تا 8000 سال قبل از میلاد وجود داشته اند. نخستین کروناویروس های کشف شده مورد مربوط به ویروس برونشیت عفونی مرغی بوده است. سپس دونوع انسانی(۲۲۹ئی[16] و اوسی43[17]) که منجر به سرماخوردگی در انسان می شود نیز کشف شد.

[16] HCoV-229E
[17] HCoV-OC43

می‌توان این‌گونه نتیجه‌گیری کرد که تاکنون قابل‌پذیرش‌ترین فرضیه برای منشاء این ویروس، فراهم شدن امکان انتقال ویروس از حیوان‌های وحشی به انسان به‌عنوان میزبان، جهش آن در انسان و سرایت آن به جهان انسانی است. تا زمانی که فرضیه قابل‌اثبات دیگری نیامده، این فرضیه قانع‌کننده‌ترین فرضیه‌ای است که فراروی ماست.

محققان در آمریکا و بریتانیا صدها جهش ژنتیکی را در ویروسی که باعث بیماری کووید-۱۹ می شود شناسایی کرده اند، اما محققان هنوز نمی دانند این جهش‌ها چه نقشی در انتشار ویروس بازی می کنند و آیا جلوی موثر بودن یک واکسن بالقوه را خواهند گرفت یا خیر. جهش کردن، بخشی عادی از عملکرد یک ویروس است. سوال این است که کدام یک از این جهش ها بر شدت یا قابلیت سرایت ویروس تاثیر می گذارد.

2-2-5 تاریخچه کووید19 در ایران

در ایران، در تاریخ 29 بهمن 1398 ، به دنبال فوت 2 بیمار با عوارض تنفسی در بیمارستان کامکار قم، شایعاتی درباره مرگ این بیماران به علت کروناویروس شکل گرفت. اما دانشگاه علوم پزشکی قم با تکذیب شایعات مربوط به مبتلا بودن این دو بیمار فوت شده به کروناویروس، اعلام کرد که تاکنون هیچگونه شواهد تشخیصی مبنی بر ابتلا به بیماری کرونا دیده نشده است. در تاریخ 30 بهمن اولین موارد از کرونای جدید در شهر قم از طریق روابط عمومی وزارت بهداشت گزارش شد. رئیس مرکز اطلاع رسانی و روابط عمومی وزارت بهداشت اعلام کرد که نتایج آزمایش اولیه 2 مورد از موارد مشکوک را از نظر ابتلا به کروناویروس، مثبت گزارش شده است.[27] بیمارستان کامکار قم نیز بعنوان محل قرنطینه و مراجعه بیماران و موارد مشکوک به کرونا مشخص شد. در همین روز، رئیس دانشگاه علوم پزشکی قم اعلام کرد که در چهار روز گذشته شاهد گسترش بیماری های تنفسی در قم بودیم و طی این مدت 2 نفر در یکی از بیمارستانهای قم بر اثر بیماری تنفسی فوت کردند که امروز تست اولیه این دو فرد در مورد بیماری کرونا مثبت اعلام شد. در 2 اسفند، با اعلام رسمی فوت شدن چهار نفر، ایران پس از چین بیشترین تلفات ناشی از کروناویروس را داشته است .تا روز نهم اسفند، 16 کشور کانادا، لبنان، امارات، بحرین، کویت،

افغانستان، عراق، عمان، پاکستان، گرجستان، چین، استونی، نیوزلند، بلاروس، بریتانیا و آذربایجان اعلام کردند که افرادی مبتلا به کروناویروس را شناسایی کرده اند که از مبدأ ایران آمده اند . در تاریخ چهارم اسفند، وزیر بهداشت اعلام کرد که یکی از فوتی ها در قم بازرگانی بود که سفرهایی به چین داشته است. [26]

از دیروز تا امروز ۲۲ اردیبهشت ۱۴۰۰ و بر اساس معیارهای قطعی تشخیصی، ۱۶ هزار و ۴۰۹ بیمار جدید مبتلا به کووید-۱۹ در کشور ایران شناسایی شد که دو هزار و ۲۵۲ نفر از آنها بستری شدند که مجموع بیماران کووید-۱۹ در کشور ایران تا تاریخ 22 اردیبهشت به دو میلیون و ۷۰۷ هزار و ۷۶۱ نفر رسید.[27]

2-3 بخش دوم : مجموعه داده

2-3-1 مجموعه داده های کووید-19

از همان اوایل به وجود آمدن کووید-19 پژوهشگران در تلاش برای کمک به کادر درمان، شروع به تلاش و تحقیقات در این زمینه کردند؛ یکی از این تلاش ها در زمینه ایجاد بانک داده یا مجموعه داده مناسبی از بیماران شامل تصاویر قفسه سینه، تب، سن، جنسیت، فاکتورهای خونی، سرفه و سایر داده های مهم بود.

مجموعه داده های بزرگی مانند مجموعه داده تصاویر قفسه سینه، مجموعه داده سرفه افراد، مجموعه داده اطلاعات بالینی یا مجموعه داده ترکیبی ازاین اطلاعات ایجاد شده است اما به دلیل حفظ حریم خصوصی بیمار، در حال حاضر تعداد کمی از مجموعه تصاویر کووید-19 در دسترس است که در ادامه به بررسی برخی از آن ها پرداخته می شود.

رحیم زاده و همکاران در ایران مجموعه داده تصویر که شامل 48260 تصویر سی تی اسکن[18] از 282 فرد عادی و 15589 تصویر از 95 بیمار مبتلا به عفونت کووید-19 است، ایجاد کرده اند. [28]

تیمی از محققان دانشگاه قطر در دوحه و دانشگاه داکا در بنگلادش به همراه همکارانی از پاکستان و مالزی با همکاری تعدادی پزشک، یک مجموعه داده از تصاویر اشعه ایکس قفسه سینه برای موارد مثبت کووید-19 به همراه تصاویر ذات ویروسی و تصاویر طبیعی ایجاد کرده اند، این مجموعه داده دائما درحال توسعه و بروز رسانی است و جزو بهترین و پراستفاده ترین مجموعه داده ها است که در حال حاضر شامل تعداد 21 هزار و189 تصویر قفسه سینه است. [29]

مجموعه داده COVID-19-CT-CXR یکی دیگر از مجموعه داده های عمومی تصویر قفسه سینه شامل 1327 تصویر CT و 263 تصویر CXR (از 9 مه 2020) با متن مربوطه است. [30] مجموعه داده AIforCOVID بایگانی بزرگی از تصاویر پزشکی بیماران کووید-19 ایتالیایی است با مشرکت CDI (یک مرکز تشخیص بیماری‌ها در میلان، ایتالیا) ایجاد شده است که این مجموعه شامل شامل داده های جمع آوری شده از 820 بیمار توسط شش بیمارستان ایتالیا در بهار 2020 در اولین اورژانس کووید-19 است؛ این مجموعه داده شامل تصاویر CXR ، چندین ویژگی بالینی و نتایج بالینی است. [31]

مجموعه داده تصاویر قفسه سینه توسط کوهن[19] ، از دانشگاه مونترال گردآوری شده‌است؛ مجموعه‌ای شامل 5856 تصویر سی تی اسکن قفسه سینه که شامل داده های بیماری کووید-19 و داده های مربوط به بیماری های ریوی دیگر از جمله سارس، سندرم دیسترس حاد تنفسی[20] و غیره است این مجموعه داده علاوه بر تصاویر قفسه سینه بیماران اطلاعات بالینی را نیز شامل می شود . [32]

[18] CT Scan
[19] Joseph Paul Cohen
[20] ARDS

یک مجموعه داده دیگر مربوط به کووید-19، COVIDGR است که تحت همکاری نزدیک با بیمارستان Universitario Clínico San Cecilio، گرانادا در اسپانیا ایجاد شده است، این مجموعه داده همگن و متعادل که شامل سطوح شدت بیماری بر اساس RT-PCR و شامل 426 تصویر CXR مثبت و 426 تصویرCXR منفی است. [33]

مجموعه داده ChestX-ray8، مجموعه داده قابل دسترس دیگری است که شامل 108.948 تصویر اشعه ایکس با نمای جلویی 32717 بیمار منحصر به فرد است. [34]

در زمینه سرفه افراد مبتلا به کووید-19 مجموعه داده های کمتری نسبت به مجموعه داده تصاویر قفسه سینه و اطلاعات بالینی موجود است، NoCoCoDa یکی از مجموعه داده های دردسترس و نسبتا مناسب سرفه است که این مجموعه داده شامل 73 سرفه است که این سرفه ها حاصل از مصاحبه های رسانه های عمومی با بیماران کووید-19 است. [35] Coswara مجموعه داده دیگری در زمینه ویروس جدید کرونا است که بیش از 400 صدای سرفه مربوط به افراد مبتلا به کووید-19 دارد و میتواند به راحتی جهت پیش بینی و تشخیص کووید-19 استفاده گردد. [36]

2-4 بخش سوم: یادگیری ماشین در کمک به تشخیص کووید-19

2-4-2 انواع یادگیری ماشین

الگوریتم های یادگیری ماشین عمدتا در دو نوع نظارت شده و نظارت نشده دسته بندی می شوند که در ادامه به توضیح آن ها پرداخته می شود.

2-4-2-1 یادگیری با ناظر [21]

الگوریتم یادگیری ماشین نظارت شده می تواند از آن چه که در گذشته آموخته است و همچنین داده های جدید برچسب گذاری شده، برای پیش بینی آینده استفاده کند. این کار از آنالیز مجموعه داده های آموزشی شروع می شود، الگوریتم یادگیری یک عملکرد استنباطی تولید می کند تا پیش بینی های مربوط به مقادیر خروجی را انجام دهد. این نوع سیستم قادر است پس از آموزش های کافی برای هر داده جدیدی هدف مشخص کند. این الگوریتم یادگیری همچنین می تواند خروجی اش را با خروجی درست و از قبل تعیین شده مقایسه کند و خطاهای موجود را بیابد تا بر اساس آن مدل را اصلاح کند.

همانطور که از اسم این یادگیری مشخص هست، در یادگیری نظارت شده مدل به کمک یک ناظر آموزش می‌بینند. در واقع در این روش ورودی (x)و خروجی (y) مشخص است و هدف مدل پیدا کردن رابطه(.)f بین ورودی و خروجی می باشد(y=f(x)).

از جمله الگوریتم‌های متداول در یادگیری با نظارت می‌توان به رگرسیون لجستیک[22] ،بیز ساده[23]، ماشین‌های بردار پشتیبان[24] و غیره اشاره کرد.

2-4-2-2 یادگیری بدون ناظر[25]

زمانی از الگوریتم های یادگیری ماشین بدون نظارت استفاده می شود که اطلاعات مورد نیاز برای آموزش نه طبقه بندی شده باشد و نه برچسب زده باشد. یادگیری بدون نظارت چگونگی اینکه سیستم می تواند توصیف ساختار پنهان از داده های بدون برچسب استنباط کنند، مطالعه می کند. این نوع سیستم خروجی مناسب را

[21] supervised learning
[22] Logistic Regression
[23] Naïve Bayes
[24] Support Vector Machine
[25] unsupervised learning

مشخص نمی کند و تنها می تواند داده ها را کاوش کند و از داده های برچسب زده نشده ساختارهای پنهان را استنتاج کند.

در این دسته از یادگیری تنها ورودی (*) را داریم و خروجی مشخص نیست. در واقع اینجا ناظری وجود ندارد تا به الگوریتم در یادگیری کمک کند. هدف اصلی یادگیری بدون نظارت مدل کردن توزیع داده هست تا بتواند اطلاعات بیشتری در باره داده بدست آورد. برعکس یادگیری نظارت شده، هیچ ناظری وجود ندارد و مدل مجبور است خودش ساختار مخفی داده‌ی بدون برچسب را پیدا کند.

2-4-2-3 یادگیری عمیق و شبکه عصبی

شبکه عصبی[۲۶]، درونی‌ترین لایه علم شگفت انگیزِ هوش مصنوعی است. علمی که دنیای امروز ما را با چند دهه قبل بسیار متفاوت کرده و تکنولوژی بخش اعظمی از پیشرفت خود را مدیون آن است. شبکه های عصبی با ساده‌تر کردن زندگی انسان‌ها در زمینه های مختلفی مثل علم پزشکی، اقتصاد، مهندسی و غیره تفاوت های زیادی نسبت به شیوه زندگی در چند دهه پیش ایجاد کرده‌اند .شبکه عصبی، بنای علم یادگیری عمیق [۲۷]و یادگیری عمیق هم خود، پایه و اساس یادگیری ماشین است. این مفاهیم با هم، علمِ هوش مصنوعی را تشکیل می‌دهند .هدف کلی این است که یک سری اطلاعات را طوری به یک ماشین یا همان کامپیوتر بدهیم، که برای او قابل درک باشد و بتواند از آن در راستای اهداف بشر استفاده کند.

شبکه‌های عصبی مصنوعی با الهام گرفتن از عملکرد مغز انسان ایجاد می‌شوند و در موارد گوناگونی از پزشکی تا اقتصاد کاربرد دارند. مفهوم شبکه عصبی مثل این است که بخواهیم به یک کودک یاد بدهیم که چگونه از بین اشکال مختلف، شکل دایره را تشخیص دهد. به او چندین عکس از دایره ها در ابعاد و رنگ های مختلف نشان

[26] Neural Network
[27] Deep Learning

می‌دهیم. پس از مدتی یاد می‌گیرد که دایره چیست و می‌تواند از میان همه تصاویری که به او نشان داده می‌شود، دایره‌ها را تشخیص دهد. این دقیقا همان کاری‌ست که به کمک شبکه‌های عصبی برای آموزش به یک ماشین انجام می‌دهیم؛ آموزش دادن به ماشین نهایتا باعث ایجاد هوش مصنوعی در آن می‌شود.

2-4-3 روش های تشخیص

هوش مصنوعی به خصوص روش های یادگیری ماشین در تشخیص کووید-19 از روش های مختلفی از جمله پردازش تصاویر سی تی اسکن یا اشعه ایکس قفسه سینه، پردازش صدای سرفه افراد، سن، جنسیت، درجه تب و سایر فاکتورها مانند فاکتورهای خونی استفاده می کنند که در ادامه به برخی از کارهای انجام شده دراین زمینه اشاره می شود.

2-4-3-1 تصاویر قفسه سینه

رایج ترین نشانه‌های ابتلا به کووید-19 تب، خستگی و سرفه خشک است. ممکن است فرد مبتلا علائم دیگری از قبیل کوفتگی، آبریزش بینی، گلودرد یا اسهال نیز داشته باشد .اما یک پزشک چطور کرونا را تشخیص می‌دهد. مثلا بررسی دمای بدن ساده‌ترین راه تشخیص این بیماری است که اصلا هم دقیق نیست .یکی دیگر از راه‌های تشخیص بیماری، اسکن قفسه سینه است. یعنی از قفسه سینه عکسبرداری شده و سلامت ریه‌ها به دقت بررسی می‌شود.

از زمان پیدایش کووید-19 تا به کنون، پژوهشگران کارهای بسیار زیادی را بر روی تصاویر قفسه سینه با کمک روش های یادگیری ماشین انجام داده اند ؛ دانشجویان دانشکده حضرت فاطمه کرمان با استفاده از روش های

یادگیری عمیق (معماری VGG[28]) و براساس مجموعه داده 60 اسکن سه بعدی ریه CT (استفاده از مجموعه داده غیرایرانی) به دقت نسبت خوبی دست یافتند روش پیشنهادی میتواند علاوه براینکه قادر به تشخیص دقیق عفونت وتشخیص کووید-19 ازمواردغیر کووید-19 است، الزامات برچسب گذاری دستی تصاویر CT را به حداقل برساند. [37] قیومی زاده و همکاران در دانشگاه ولی عصر رفسنجان کرمان با پردازش 686 تصویرقفسه سینه شامل 331 تصویر سالم، 183تصویر کووید-19، 83 تصویر باکتریایی و79 تصویر ویروسی که از بیمارستان واسعی سبزوار جمع آوری شده است با استفاده از ترکیب معماری های مختلف (Vgg, AlexN ،CaffeNet) و انتخاب ویژگی بر اساس معیارهای مانند انتروپی و طبقه بندی بر اساس ماشین بردار پشتبان به دقت نزدیک به 99٪ رسیدند. [38]

عباسیان اردکانی و همکاران با استفاده از شبکه عصبی کانولوشنی[29] Resnet-101 با استفاده از مجموعه داده شامل 1020 تصویر سی تی اسکن به دقت 99.51٪ رسیدند. [39] همچنین تبریزچی از دانشگاه شهید باهنر کرمان به همراه همکاران یک ارزیابی کلی بر روی مدل های پرکاربرد در حوزه تشخیص ویروس کووید19 مانند شبکه عصبی چندلایه پرسپترون[30]، شبکه عصبی کانولوشنی، ماشین بردار پشتیبان و روش های بوستینگ[31]، با استفاده از 982 تصویر قفسه سینه انجام دادند که روش های شبکه عصبی کانولوشنی و ماشین بردار پشتیبان به تربیت با دقت های ٪0.97 و ٪0.99 دارای بالاتری دقت بودند.[40]

براساس یک بررسی کلی جهت تشخیص بیماری کووید-19 در دانشگاه شهید باهنر کرمان، شبکه های کانولوشنی عمیق تعداد بیشتری از مدل ها را نسبت به روش های غیر عصبی مبتنی بر شبکه اختصاص داده اند. علاوه بر این، شبکه های عمیق نتایج بهتری نسبت به سایر مدل های یادگیری ماشین به دست آوردند. با این حال، گسترش سریع کووید-19 و کمبود داده و اطلاعات بیماران ممکن است باعث خطا در این نتیجه گیری شده باشد. [41]

[28] Virtual Graphics Group
[29] Convolution Neural Network
[30] Multi-Layer Perceptron
[31] Boosting

معماری VGG در پردازش تصاویر یکی از معماری های با دقت بالا است، سیوی[۳۲] و آیدین[۳۳] در سال 2020 از کشور ترکیه با ارزیابی چهار معماری مختلف مربوط به پردازش تصویر (شبکه های CNN،VGG-19 VGG-16، inception V3) بر روی یک مجموعه داده شامل 657 تصویر قفسه سینه (219تصویر مربوط به افراد سالم، 219 تصویر افراد مبتلا به کووید19 و 219 تصویر مربوط به افراد مبتلا به ذات الریه) به دقت های[۳۴] مختلفی برای هر مدل دست یافتند که نتایج نشان دهنده برتری VGG-19 با دقت 95٪ بود.[42]

نارین[۳۵] و همکاران در سال 2020 برای افزایش دقت تشخیص علاوه بر استفاده از معماری های مختلف شبکه عصبی کانولوشنی (در این مقاله Resnet-50) از روش ماشین بردار پشتیبان استفاده کردند و به دقت نزدیک به ٪99 با استفاده از مجموعه داده شامل2905 تصویر قفسه سینه و خروجی 3 کلاسه نرمال، ذات الریه و کووید-19رسیدند.[43]

در کشور اندونزی نیز فیان یولیو سانتوسو36 و همکاران با تغییر در شبکه عصبی کانولوشنی به دقت 100٪ بر روی داده های آموزشی (مجموعه داده شامل 618 تصویر 256*256) و دقت 99.1٪ بر روی داده های تست دست یافتند. [44]

2-4-3-2 سرفه

سرفه یکی از دلایل اصلی در گسترش کووید-19 است و برخی افراد تنها سرفه می کنند و علائم دیگری ندارند و در واقع ناقل هستند و باعث گسترش خیلی سریع آن می شوند ؛ مطالعات نشان می دهد که در سرفه اطلاعات مهمی راجع به بیماری های مختلفی مانند بیماری سل [45] آسم، کرونا، سیاه سرفه، آلزایمر و غیره

[32] Mehmet Sevi
[33] İlhan AYDIN
[34] Accuracy
[35] Ali Narin
[36] Fian Yulio Santoso

نسبت به سایر صداهایی مانند نفس کشیدن یا شمارش، نهفته شده است [46] افراد بدون علامت ممکن است از نظر سرفه با افراد سالم متفاوت باشند. این تفاوت ها برای گوش انسان قابل رمزگشایی نیستند. اما با هوش مصنوعی می توان این اطلاعات را از طرق مختلفی مانند ضبط سرفه در افراد داوطلب (به علاوه دریافت اطلاعاتی شامل سن، جنسیت،کشور و غیره) از طریق سایت یا اپلیکیشن و به صورت آنلاین دریافت کند و با تبدیل فرکانسی MFCC و ورود فرکانس های صدا به یک معماری شبکه عصبی خاص، بیمار بودن افراد تشخیص داده شود ، شبکه عصبی ذکر شده از سه معماری یکسان به نام Resnet50 به صورت موازی استفاده می کند که هر یک از این معماری ها پردازش ویژگی های خاصی از صدای سرفه به عنوان مثال، پردازش احساس فردی که سرفه کرده است را انجام می دهند و در نهایت خروجی تجمیع شده فرد مبتلا به کووید-19 را با دقت ٪98.5 را تشخیص می دهد[47] علاوه براستفاده از سایت و دریافت اطلاعات اضافی ذکر شده، می توان از یک اپلیکیشن برای دریافت سرفه به مدت 3 ثانیه و تشخیص ویروس کرونا استفاده کرد، سرفه ضبط شده ابتدا توسط یک شبکه عصبی و براساس دیتاست های موجود(ESC- 50) باید تشخیص داده شود و اگر صدایی غیر از سرفه بود درخواست ضبط مجدد به کاربر ارسال شود و در صورت وجود سرفه، صدا وارد سه مدل موازی جهت تشخیص اینکه سرفه مربوط به ویروس کووید-19 یا سیاه سرفه یا برونشیت یا سرفه فرد عادی است، می شود و بر اساس نتایج هرسه مدل، خروجی به صورت : مبتلا، سالم و بدون نتیجه نمایش داده می شود.[48] علاوه بر مجموعه داده ESC- 50 می توان از مجموعه داده های دیگری مانند audioset1 ، ESC502 برای تشخیص سرفه خشک کرونا از سایر سرفه های مربوط به بیماری های مختلف استفاده کرد که با استفاده از یک شبکه عصبی کانولوشنی می توان با یک دقت منطقی ٪70.58 آن را تشخیص داد. [49]

آندره پرز[۳۷] و همکاران درسال 2021، با استفاده از مجموعه داده های موجود مانندCoswara یک وب سایت[۳۸] جهت تشخیص میزان عفونت کووید-19 را براساس سه شدت متفاوت و براساس ویژگی های صوتی و طبقه بندی

[37] Javier Andreu-Perez
[38] https://coughdetect.com

کننده شبکه عصبی مصنوعی عمیق با لایه های کانولوشن به نام DeepCough است ایجاد کرده اند؛ نرخ نتایج تشخیص کووید-19 به یک AUC امیدوار کننده 98.80 ٪ ، حساسیت[39] 96.43٪ و تشخیص پذیری[40] 96.20٪ رسیده است. [50]

2-4-4 بات های هوشمند

بات هوشمند (اپراتور هوشمند) یک نرم افزار هوش مصنوعی است که می‌تواند مکالمه یا چت یک کاربر به زبان طبیعی را از طریق پیام رسان، وب سایت و اپلیکیشن‌های تلفن همراه شبیه سازی کند. در واقع چت بات ها یک سرویس مبتنی بر چت هستند که به سوالات کاربران پاسخ می‌دهند و اطلاعات لازم را برای آن‌ها ارسال می‌کند. از این رو می‌توان گفت چت بات‌ها در حال حاضر یکی از پیشرفته‌ترین و امیدوار کننده‌ترین ابزار تعامل بین انسان و ماشین‌ها هستند. از دیدگاه تکنولوژیکی، یک چت بات تنها نشان دهنده تکامل سیستم پاسخگویی به سوال است که از پردازش زبان طبیعی بهره می‌برد. یعنی یک سیستم چت آنلاین با پشتیبانی ۲۴ ساعته که زبان انسان‌ها را درک می‌کند.

با افزایش محبوبیت پیام رسان‌ها، چت بات‌ها نیز نقش مهمتری به خود گرفتند و نحوه تعامل کسب و کارها با مشتریان را تغییر دادند. در واقع این چت بات‌ها به کسب و کارها این امکان را می‌دهد تا بدون صرف هزینه برای استخدام یک شخص، با مشتری ارتباط برقرار کنند. بسیاری از سوالات و موضوعاتی که برای مشتریان حائز اهمیت است، مشترک بوده و می‌توان به راحتی به آن‌ها پاسخ داد. در نتیجه این فناوری می‌تواند یک جایگزین برای نیروی انسانی باشد و به عنوان راهکاری برای صرفه جویی در هزینه و زمان استفاده شوند. چت بات فارسی نیز می‌تواند ابزاری توانمند برای تعامل با کاربران باشد. چت بات هوشمند برای تحقق یک هدف به یک فرآیند سه مرحله‌ای

[39] sensitivity
[40] specificity

پایبند است. این فرآیند سه مرحله‌ای حس، فکر و عمل است که می تواند هوشمندی یک اپراتور مجازی را تعریف کند.

مراجعه به پزشک و شلوغی بیمارستان ها خود دلیلی برای گسترش ویروس است و استفاده از سیستم های پشتیبانی از تصمیم[۴۱] برای تشخیص ویروس کووید-19 و ارائه پیشنهادات (مانند ماندن در خانه به مدت ۱۴ روز در صورت تشخیص سیستم به مبتلا بودن کاربر، ماسک زدن وغیره) مشابهه کاری که در کشور اندونزی انجام گرفته است می تواند راه حل مناسب تری قبل از مراجعه به پزشک باشد. [51]

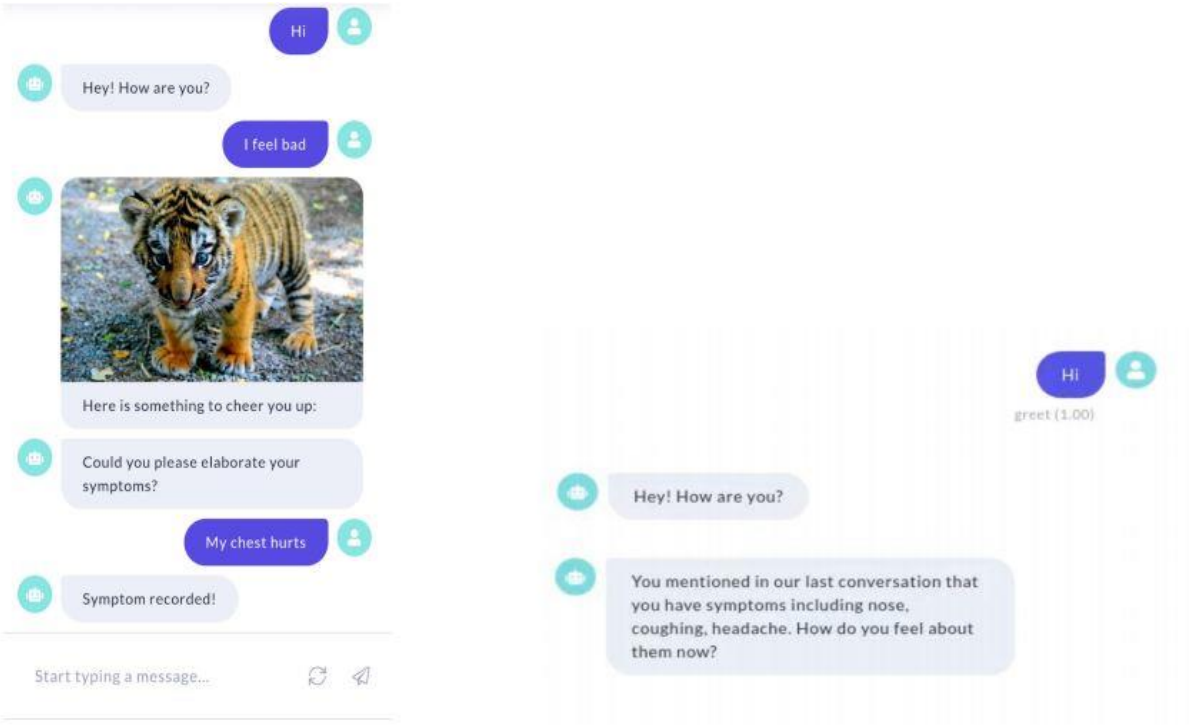
یک تیم تحقیقاتی در مراکش با استفاده از 300 تصویر رادیوگرافی قفسه سینه شامل 100 تصویر نرمال، 100 تصویر تایید شده کووید-19 و 100 تصویر ذات الریه جهت تشخیص زود هنگام مبتلا شدن افراد به ویروس، با استفاده از روش های یادگیری عمیق(معماری VGG) و پردازش تصویر ایجاد کرده اند که این مدل که دقت 92.5% را دارا است ، پس از استخراج ویژگی های تصویر قفسه سینه افراد مبتلا به کووید-19، پیشنهاداتی را برای افراد مبتلا ارائه می دهد. [52]

بسیاری از بیماران کووید-19پس از مبتلا شدن، علائم طولانی مدت از جمله خستگی، هذیان و سردرد را نشان دادند. تأثیر طولانی مدت این شرایط در سلامتی هنوز مشخص نیست. تهیه روشی برای پیگیری این بیماران برای نظارت بر وضعیت سلامت آنها برای حمایت از مداخله و درمان به موقع ضروری است و ایجاد یک چت بات می تواتد راه حل مناسبی جهت کمک به این نوع افراد باشد.

بات های هوشمند با استفاده از داده های متنی از ادبیات تحقیقات علمی موجود در مورد کووید-19 و بیماری های مربوطه آموزش داده می شوند تا نام علائم، داروها و اصطلاحات دیگر مرتبط با این بیماری را بیاموزند. با استفاده از این اصطلاحات مدل بر اساس روش های پردازش متن طبیعی[۴۲] مبتنی بر یادگیری ماشین آموزش می بیند تا

[41] system support decision
[42] Natural Language Processing

قادر به ایجاد یک محیط مکالمه برای تعامل با بیماران، پاسخگویی به سوالات بیماران و کاوش عمیق برای سنجش شرایط و ارائه راه حل مناسب به بیماران باشند. شکل زیر مکالمه آزمایشی بین بیمار و چت بات در محیط تلفن همراه و محیط رایانه شخصی را نشان می دهد. کادر آبی ورودی بیمار و کادر خاکستری پاسخ چت بات است.[53]



2-1 مکالمه آزمایشی بین چت بات و بیمار

بیماران مبتلا به خوردن داروهای کافی و مراقبت های بی موقع نیاز دارند. بنابراین برای غلبه بر مشکلات، ایجاد یک ربات به جای پرستار راه حل مناسبی است. ربات پرستار می تواند بیمار و مصرف دارو بسته به زمان را کنترل کند . ربات پرستار می تواند کمک زیادی به کاهش رفت و امد بیمار به مراکز بهداشتی و کاهش ارتباط با افراد و در نتیجه کاهش انتقال ویروس شود. [54]

داروهای بدون وقت برای بیماران می توانند مشکلات جدی بهداشتی برای بیماران در بیمارستان ها و خانه های سالمندان پیری ایجاد کنند. ایده پیشنهادی موروگان[۴۳] و همکاران، به نام AKM Med برای کمک به این بیماران

[43] Kalpana Murugan

در تأمین یا توصیه داروهای لازم به ترتیب صحیح در ساعات مشخصی از روز و بدون نیاز واقعی پرستار انسان است. این یک ربات قابل برنامه ریزی مبتنی بر میکروکنترلر نیمه مستقل است، که می تواند به بیماران یادآوری کند، قرص های خود را در فواصل زمانی مناسب مصرف کنند. پس از مصرف موفقیت آمیز قرص ها، وضعیت دارو از طریق یک پیامک برای نزدیکان بیمار ارسال می شود. [55]

احمد خان[۴۴] و همکاران در سال2021، از هوش مصنوعی و پردازش زبان طبیعی برای تهیه سیستمی استفاده کردند که با پیام رسان بسیار معروف واتساپ[45] ادغام شده است، این ربات برای ارائه مشاوره مناسب بر اساس علائم بیمار است. این ربات در ادامه به توصیه ها و اطلاعاتی در مورد نزدیکترین بیمارستان ها برای مراقبت های پزشکی از فرد بیمار می دهد، در واقع ورودی داده شده توسط کاربر علائمی است که توسط خود ربات درخواست می شود. بر اساس علائم منتقل شده به ربات، تصمیم گرفته می شود که موتور پردازش زبان طبیعی توصیه ای را برای کاربر مشخص کند یا نه. در صورت مطابقت علائم برای کاربر مثبت کووید-19، ربات نزدیکترین آدرس بیمارستان یا واحد بهداشتی را فراهم می کند. در صورت بروز شرایطی که ربات هیچ مشکلی با شرایط بهداشتی شناسایی نکند یعنی در واقع تشخیص حاد بودن وضعیت کاربر را ندهد برای اطمینان به کاربر توصیه می کند که خودش را در خانه قرنطینه کند و خود را برای مدت هفت روز یا بیشتر تحت نظر نگه دارد. [56]

Symptoma یک دستیار بهداشتی دیجیتال است که می تواند بیش از 20000 بیماری را با دقت بیش از 90٪ از هم متمایز کند. مارتین[۴۶] و همکاران در سال2020، با استفاده از مجموعه ای از موارد مختلف بالینی کووید-19، دقت Symptoma را برای شناسایی کووید-19 آزمایش کردند. نتایج نشان می دهد که Symptoma می تواند کووید-19 را با دقت 96.32٪ بر اساس فاکتور های بالینی تشخیص دهد. به طور خلاصه، Symptoma دقت بی نظیری را در شناسایی سیستماتیک موارد کووید-19 و در عین حال بررسی بیش از 20000 بیماری دیگر فراهم

[44] Nayyar Ahmed Khan
[45] WhatsApp
[46] Alistair Martin

می کند. علاوه بر این، Symptoma امکان وارد کردن متن رایگان، همراه با سوالات پیگیری بیماری خاص، به 36 زبان را فراهم می کند. در مجموع، این نتایج و قابلیت دسترسی به Symptoma این امکان را می دهد که ابزاری اساسی در مبارزه جهانی با کووید-19 باشد. [57]

بات هوشمند Medbot، بات هوشمند دیگری است که می تواند به عنوان یک پزشک مجازی شخصی عمل می کند و به طور مناسب برای تعامل با بیماران مانند انسان طراحی و آموزش دیده است. این بات مبتنی بر معماری بدون سرور است و خدمات پزشک را با ارائه اقدامات پیشگیرانه، درمان در خانه، جلسات مشاوره تعاملی، نکات مراقبت های بهداشتی و علائم شایع ترین بیماری ها در روستاهای هند جمع می کند. این برنامه مکالمه منجر به کاهش موانع دسترسی به مراکز بهداشتی درمانی و مشاوره های هوشمند از راه دور برای کمک به مراقبت به موقع و درمان با کیفیت، با کمک موثر به جامعه شده است. این بات دارای ویژگی های اضافی از جمله درمان های خانگی، توصیه های رژیم غذایی مبتنی بر مکان، سن و مشاوره، شماره تلفن های کمک اضطراری است و می تواند با یک برنامه پیام رسان در زمان واقعی مانندپیام رسان واتساپ مرتبط شود. [58]

بات های هوشمند علاوه بر رفتار کردن مانند یک پرستار یا پزشک مجازی، می توانند تجزیه و تحلیل احساسات عمومی را نسبت به کووید-19 را انجام دهند ؛ در یک روش پیشنهادی بات هوشمند به تجزیه و تحلیل احساسات عمومی افراد با هشتگ کووید-19 در توییتر می پردازد که با با استفاده از روش Naïve Bayes ، 91٪ دقت برای توییت های کوتاه و با استفاده از روش طبقه بندی رگرسیون لجستیک، 74٪ دقت برای توییت های کوتاه را حاصل می شود. [59]

2-5 خلاصه

اساس انجام هر پژوهشی، ادبیات و پیشینه ی آن پژوهش است که به عنوان یک چارچوب و پشتوانه ی علمی، پژوهش را در انجام آن هدایت و حمایت میکند. در این فصل که از آن به عنوان مروری بر ادبیات پژوهش و پیشینه تحقیق، نام برده شده است ابتدا بر اساس منابع موجود تعریف های لازم اورده شده است و سپس به صورت دسته بندی شده کارهای گذشته انجام شده در هر مرحله تشخیص ویروس کووید-19 آورده شده است.به طور کلی هدف از تدوین فصل حاضر، ارائه ادبیات پژوهش، مثال کاربردی، کارهای گذشته و همچنین جمع بندی کلی و بحث در خصوص آنها بوده است.

فصل سوم:
روش‌شناسی پژوهش

3-1 مقدمه

هدف تمام علوم، شناخت و درک دنیای پیرامون ما است. به منظور آگاهی از مسائل و مشکلات دنیای اجتماعی و تجاری، روش‌های علمی تغییرات قابل ملاحظه‌ای پیدا کرده‌اند. این روندها سبب شده است که برای بررسی مسائل مختلف، از روش‌های علمی (پژوهش) استفاده شود. پژوهش، عملی منظم، منطقی و دقیق، برای دست یافتن به حقایق، دانش نو و ارتباط آن با کل زندگی و همچنین درک روابط بین متغیرها است. پژوهش، بررسی کامل موضوع به گونه‌ای منظم و منسجم و بر اساس روش‌های عینی و غیر شهودی است که به منظور کسب اطلاعات یا کشف اصول وابسته به آن انجام می‌گیرد. پژوهش، فرایندی است که از طریق آن می‌توان درباره‌ی ناشناخته به جستجو پرداخت و نسبت به آن، شناخت لازم را کسب کرد. پژوهش یعنی کار منظم و پیگیر در کشف و فهم مسئله‌ای خاص، که به نظریه‌ای برای تعمیم منتهی شود. در واقع، پژوهش را می‌توان تلاش منظم و سازمان‌یافته برای طراحی و پیگیری راه‌حل، برای یک مسئله‌ی خاص توصیف کرد تا پاسخ‌هایی برای مسئله مورد علاقه در محیط کاری به دست آید.

3-2 روش کار

دستیابی به هدف‌های علم یا شناخت علمی میسر نخواهد بود، مگر زمانی که با روش‌شناسی[۴۷] درست صورت پذیرد. به عبارت دیگر، پژوهش، نه از حیث موضوع، بلکه از حیث روش است که اعتبار می‌یابد. روش پژوهش به معنای رویکرد هستی‌شناسی است. روش پژوهش به عنوان یک فرایند نظام‌مند، برای یافتن پاسخ یک پرسش یا راه‌حل یک مسئله است. انتخاب روش پژوهش بستگی به هدف، ماهیت و موضوع پژوهش و همچنین امکانات اجرایی آن دارد. بنابراین، هنگامی می‌توان در مورد بررسی و انجام یک پژوهش تصمیم گرفت که ماهیت موضوع پژوهش، هدف‌ها و نیز وسعت دامنه‌ی آن مشخص باشد. در این بخش به تشریح روش کار پرداخته می‌شود.

[47]Methodology

این پژوهش، از نوع پژوهش‌های توسعه‌ای-کاربردی[۴۸] است، زیرا به دنبال افزودن دانش در حوزه‌ی بیماری کرونا و افزایش اعتبار آن است. همچنین می‌توان گفت این پژوهش از حیث افق زمانی از نوع پژوهش‌های تک مقطعی است. زیرا به دنبال یک پدیده در طول زمان و یا مقاطع مختلف نبوده است. در صورتی که اطلاعات مورد نیاز پژوهش‌گر در منابعی ثبت و مدون شده باشد، برای آگاهی از آن اطلاعات، پژوهش‌گر به آن اطلاعات رجوع نموده و اطلاعات را جمع‌آوری می‌کند. از آنجا که غالباً چنین منابعی را در کتابخانه نگهداری می‌کنند به این روش جمع‌آوری اطلاعات، روش کتابخانه‌ای می‌گویند.

پژوهش توسعه‌ای، پژوهشی است که با هدف گسترش مرزهای دانش عمومی بشر صورت می‌گیرد. بنابراین، هر یک از موضوعاتی که در حوزه‌ی علوم انسانی و تجربی بتواند به گسترش این علوم کمک نماید و به شرح و تبیین آن بپردازد و وجوه ناشناخته‌ی آن را روشن کند، در حوزه پژوهش‌های توسعه‌ای جای می‌گیرد.

پژوهش کاربردی، پژوهش‌هایی با هدف برخورداری از نتایج یافته‌ها، برای حل مسائل موجود در سازمان هستند.

پژوهش تک مقطعی، نوعی بررسی است که در آن داده‌های اطلاعاتی، یک بار در طی یک دوره‌ی چند روزه یا چند هفته یا چند ماهه، به منظور پاسخ به پرسش یک پژوهش جمع‌آوری می‌شود.

در فصل چهارم پژوهش، در مورد گام های مدل و الگوریتم به کار رفته از نظر تئوری و کاربردی و همچنین به تجزیه و تحلیل داده ها به تفصیل بحث شده است. در این قسمت، به عنوان روش کار به روش های نرم افزاری و دیجیتالی که برای تشخیص و درمان بیماری کرونا اشاره می‌کنیم.

در این پژوهش تلاش شد انگیزه های هوش مصنوعی، کلان داده ها و شبکه ها ی نوین برای یافتن رویکرد های سریع و موثر که می تواند به طور موثر با بیماری کووید-۱۹ مبارزه کند بررسی گردد. این بررسی ها نشان میدهد تلاش های امیدوارکننده ای در حوزه پیشگیری و پیش بینی، تشخیص و درمان این بیماری با استفاده از این

[48]Development-Applied Research

تکنولوژی ها صورت پذیرفته است که همچنان در حال تکامل می باشد. امید است نتایج بررسی های انجام شده راه گشای کادر درمان در مبارزه با این بیماری باشد.

3-2-1 گام‌های کاربرد هوش مصنوعی و یادگیری ماشین

ظهور فن آوری های پیشرفته مدل سازی کامپیوتری و اتخاذ آن ها، در بسیاری از بخش های بین المللی منجر به بهبود و ارزیابی ریسک کیفیت سلامت منطقه ای و جهانی شده است. در میان آن ها، زمینه پزشکی به طور خاص تغییر کرده است، که در آن افزایش داده های سلامت و بهداشتی با توجه به در دسترس بودن فن آوری های مختلف مانند کلان داده، یادگیری ماشینی، اینترنت اشیاء، هوش مصنوعی و شبکه های هوشمند به طور قابل توجهی به ظرفیت های مدلسازی پیشگویانه کمک کرده است. در دسترس بودن ابزارهای محاسبه پیش گویانه و استفاده متعاقب آنها در مراقبت های بهداشتی در سراسر جهان تحولات قابل توجهی را در عملیات های پزشکی به ارمغان آورده است و انتظار میرود که چنین تغییراتی به بهبود دقت در این زمینه، به ویژه در صحت تشخیص، ادامه دهد.

هدف اصلی این پژوهش نشان دادن اثربخشی هوش مصنوعی، کلان داده ها و شبکه های هوشمند برای مبارزه با بیماری همه گیر کووید-19 و بررسی راه حل های پیشرفته در پیش بینی، تشخیص و درمان بیماری با استفاده از این فن آوری ها است.

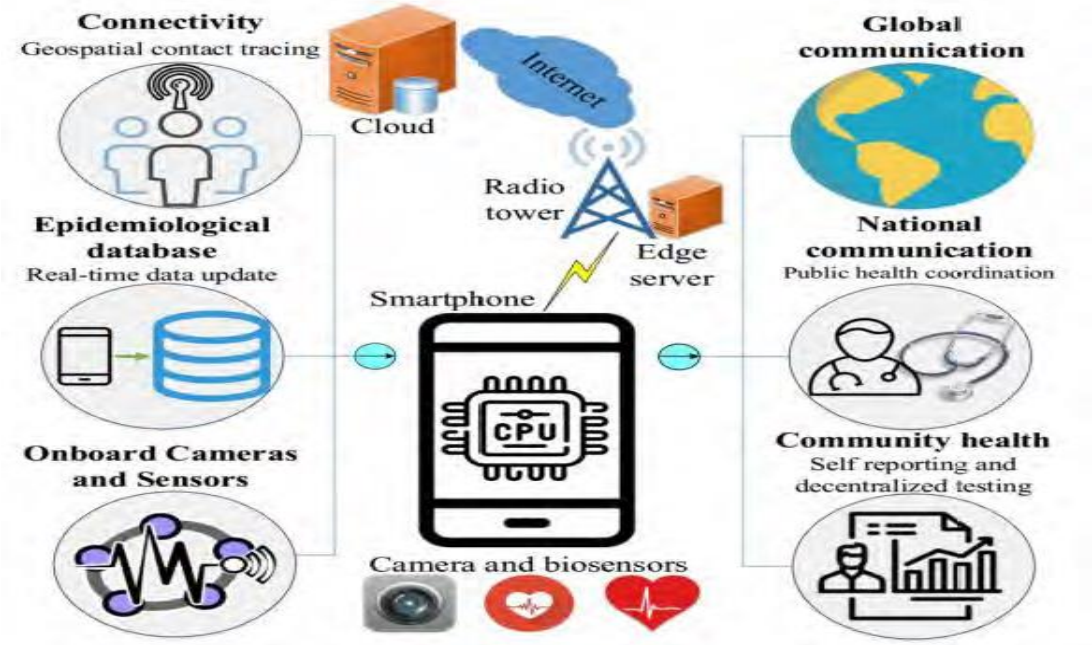
گام 1- هوش مصنوعی برای شناسایی و تشخیص ویروس کرونا

در بین بسیاری از شاخه های هوش مصنوعی، یادگیری ماشین و یادگیری عمیق دو رویکرد مهم است. به طور کلی، یادگیری ماشین به توانایی یادگیری و استخراج الگوهای معنی دار از داده ها اشاره می کند و عملکرد الگوریتم های مبتنی بر یادگیری ماشین و سیستم ها به شدت به ویژگی های نماینده وابسته هستند.

با توجه به پیشرفت های اخیر در زمینه تکنیک های مولکولی و محاسباتی و اطلاعات و فن آوری های ارتباطی (ICT)، هوش مصنوعی و کلان داده ها می توانند به کنترل حجم عظیم و بی سابقه داده های به دست آمده از نظارت سلامت عمومی، نظارت بر روند جاری، گزارش وضعیت عادی و به روز رسانی اطلاعات از موسسات و ادارات دولتی و اطلاعات استفاده از تاسیسات بهداشتی و درمانی کمک کنند. هدف این مقاله نشان دادن کاربردهای بالقوه هوش مصنوعی و کلان داده در آزمون سراسری جهانی برا ی مدیریت بیماری همه گیر است.

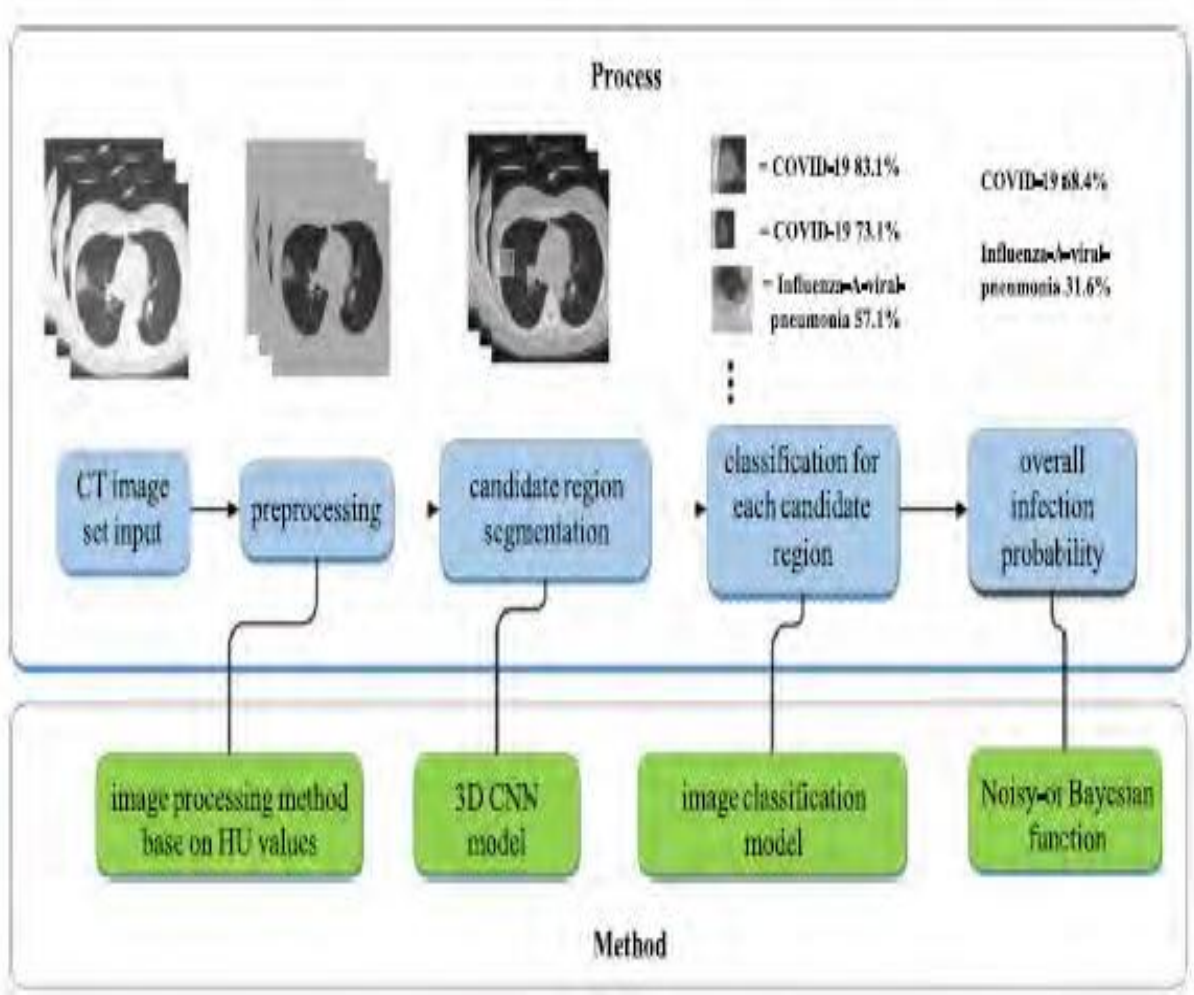
یک چارچوب مبتنی بر تلفن همراه برای تشخیص و نظارت کووید-۱۹ ، نشان داده شده است. اتصال بی سیم (wifi، 3G، 4G و5G) به طور گسترده تقریبا در تمام کشورها استقرار یافته اند. بنابراین هر فردی با یک تلفن همراه می تواند به اینترنت متصل شود و فعالیت های زیادی را به صورت آنلاین انجام دهد.

به عنوان مثال ، با استفاده از دوربین های تعبیه شده و حسگرهای زیستی ، تلفن همراه می تواند اطلاعات شخصی را جمع آوری کند ، مانند تصاویر اشعه ایکس و سی تی ، صدا ی سرفه و ضربان قلب ، که ممکن است قبل از ارسال به بانک داده برا ی آموزش رمزگذاری و فشرده شود.



شکل ۳-۱ هوش مصنوعی در تشخیص بیماری

هوش مصنوعی فن آوری پیشرفته در بسیاری از کاربردهای هوشمند در زمینه های مختلف است. برخی از نمونه های برجسته هوش مصنوعی، وسایل نقلیه مستقل، تشخیص پزشکی از راه دور در مراقبت های بهداشتی هستند. در مثالی دیگر در این راهکار از یک مدل ۳ بعدی شبکه عصبی کانولوشن برای بخش بندی کردن تیکه های تصویر چندگانه استفاده شده است. سپس از یک تصویر برای طبقه بندی تمام تیکه های تصویر به یکی از سه نوع: کووید-۱۹ ، آنفلوان ای آ- ویروسی و غیرمرتبط با عفونت استفاده شده است .در نهایت ، گزارش تحلیل کلی برای یک نمونه سی تی اسکن با استفاده از تابع Noisy یا بیز محاسبه شد.



شکل۳-۲ طبقه بندی و تشخیص بیماری براساس تصاویر

گام 2- کاربردهای اصلی هوش مصنوعی در همه گیر شدن ویروس کرونا

الف) تشخیص زود هنگام و تشخیص عفونت

هوش مصنوعی می تواند به سرعت علائم نامنظم را تجزیه و تحلیل کند و در نتیجه به بیماران و مسئولان مراقبت بهداشت هشدار دهد. این کار به فراهم کردن تصمیم گیری سریعتر کمک می کند که مقرون به صرفه است. هوش مصنوعی می تواند در تشخیص موارد آلوده به کمک فن آوری های تصویربرداری پزشکی مانند پرتو نگاری کامپیوتری (CT) ، تصویربرداری تشدید مغناطیسی (MRI) در بخش های بدن انسان مفید باشد.

ب) نظارت بر درمان

هوش مصنوعی می تواند یک پلتفرم هوشمند برای نظارت و پیش بینی اتوماتیک گسترش این ویروس ارائه دهد.

ج) کاهش حجم کار کارکنان بخش مراقبت های بهداشتی

با توجه به افزایش ناگهانی و گسترده تعداد بیماران در طی شیوع بیماری همه گیر کووید-۱۹ ، متخصصان مراقبت های بهداشتی حجم کاری بسیار بالایی دارند. در اینجا ، هوش مصنوعی برای کاهش حجم کار کارکنان بخش بهداشت و درمان مورد استفاده قرار میگیرد. این کار به تشخیص زودهنگام و ارائه درمان در مراحل اولیه با استفاده از رویکردهای دیجیتال و علوم تصمیمگیری کمک می کند و بهترین آموزش را برای دانشجویان و پزشکان در رابطه با این بیماری جدید ارائه می دهد. هوش مصنوعی می تواند بر مراقبت از بیماران آینده تاثیر بگذارد و به چالش های بالقوه بیشتری رسیدگی کند که فشارکاری پزشکان را کاهش می دهد.

د) پیش گیری از بیماری ها

هوش مصنوعی با کمک تحلیل داده های بلادرنگ می تواند اطلاعات را به روز رسانی کند که در پیشگیری از این بیماری مفید است. در شکل زیر روند عمومی شناسایی با استفاده از هوش مصنوعی را نشان می دهد که به پزشکان

عمومی کمک میکند تا علائم کووید-۱۹ را شناسایی کنند. دیاگرام زیر روند درمان بدون و با در نظر گرفتن هوش مصنوعی را با یک دیگر مقایسه می کند. این روند نحوه استفاده از هوش مصنوعی و یادگیری ماشین در گام های مهم درمان با دقت بالا را نشان می دهد و پیچیدگی و زمان صرف شده را کاهش می دهد. پزشک نه تنها بر درمان بیمار متمرکز می شود، بلکه بیماری را با استفاده از هوش مصنوعی کنترل می کند.

ت) استراتژی های درمانی

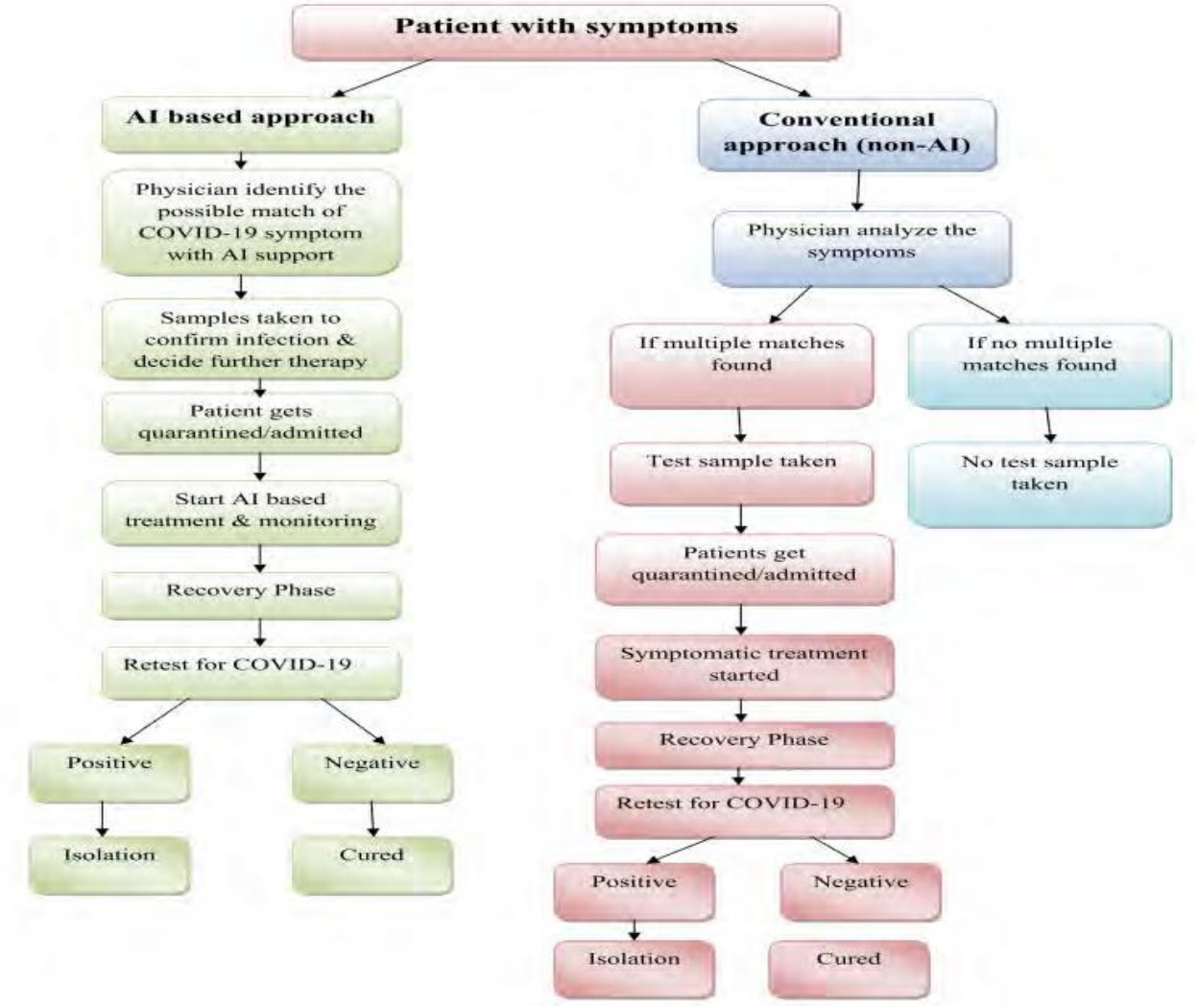
یک مدل شبکه عصبی مصنوعی کانولوشنی را برای تشخیص موارد کووید-۱۹ طراحی کردند. این مدل افراد را به سه دسته(نرمال، افراد آلوده شده، افراد آلوده نشده) طبقه بندی می کند ، کارکنان پزشکی با استفاده از این طبقه بندی کدام موارد باید با روش های استاندارد مورد آزمایش قرار گیرند(بین پرونده های نرمال و افراد آلوده شده) و کدام استراتژی های درمانی باید (بین پرونده های آلوده نشده) مورد آزمایش قرار گیرند.

ث) هزینه کم دقت بالا

یک مدل عمیق کانولوشنی به نام DeTraC (شامل لایه های تجزیه، انتقال و ترکیب) به منظور پردازش تصاویر عکس برداری شده ریه به جهت طبقه بندی بیماری کووید-۱۹ ارائه داده اند. هدف اصلی لایه تجزیه کاهش فضای ویژگی می باشد که منجر به بهبود کارایی اموزش میگردد. در حالی که لایه ترکیب وظیفه تعریف زیرکلاس های تولید شده در لایه های قبلی را به منظور تولید نتایج طبقه بندی نهایی بر عهده دارد. در کنار این دو لایه، لایه انتقال به منظور افزایش سرعت یادگیری، کاهش هزینه های محاسباتی و انطباق مدل یادگیری عمیق با مجموعه داده های کوچک پیشنهاد شده است.

پ) تشخیص ویروس با صدا

یک چارچوب مبتنی بر هوش مصنوعی به نام AI4COVID-19 پیشنهاد شده است به گونه ای از دانش متخصصان پزشکی بهره ببرند. داده های ورودی این سیستم سیگنال های صوتی و سرفه می باشد که توسط تلفن های هوشمند ضبط شده است. این سیستم نتایج بسیار امیدوارکننده ای بدست اورده است که بادقت ۹۷.۹۹ درصد بیماری را تشخیص می دهد.



شکل 3- 3 روند عمومی شناسایی بیماری با استفاده از هوش مصنوعی

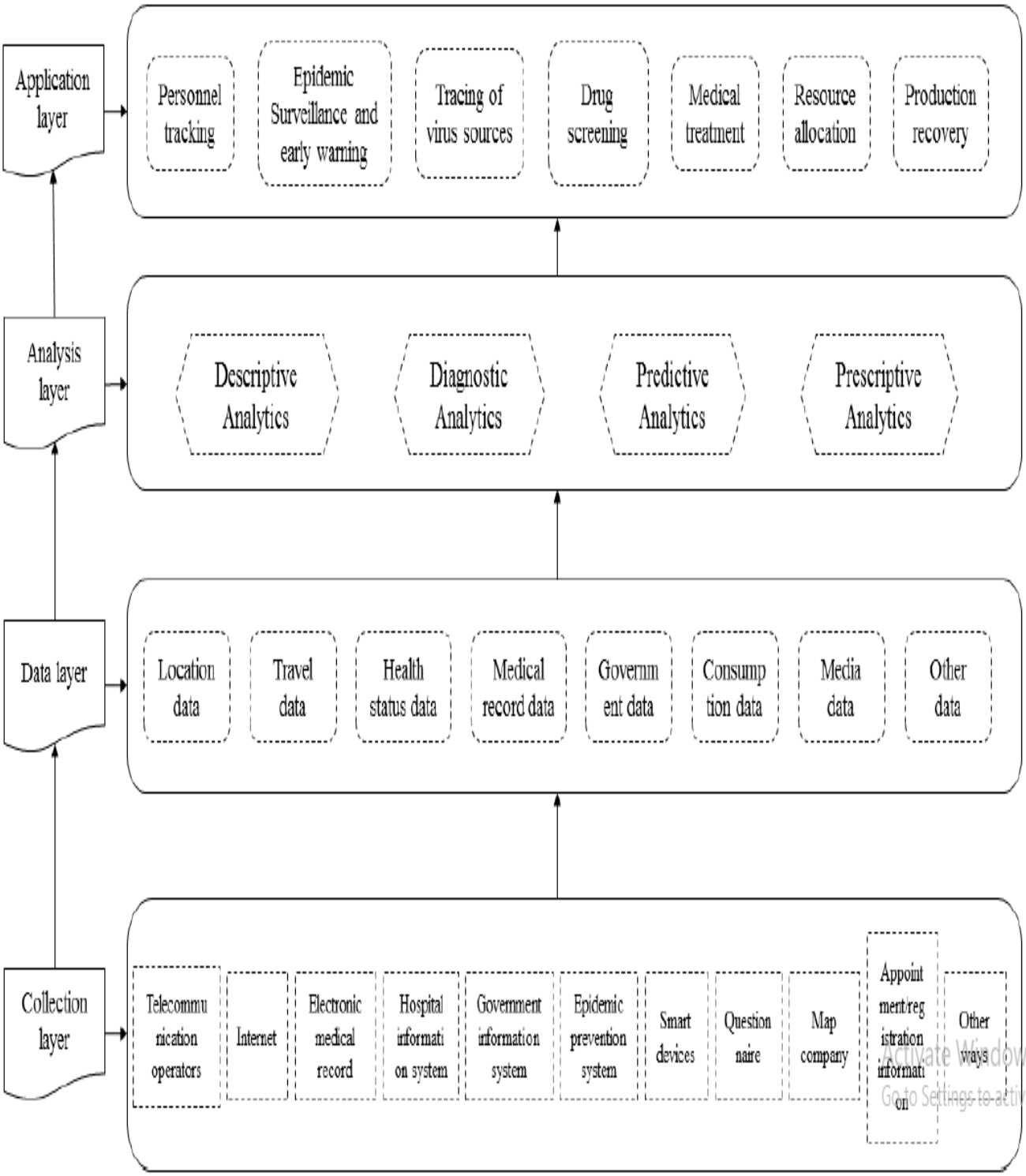
گام 3- کلان داده ها برای شناسایی و تشخیص ویروس کرونا

کلان داده ها به انواع داده ها با در نظر گرفتن حجم، تنوع، ارزش، صحت و پویایی آن ها اشاره دارد. این داده ها می تواند داده های ایجاد شده توسط ماشین و یا داده های شبکه های اجتماعی باشد.

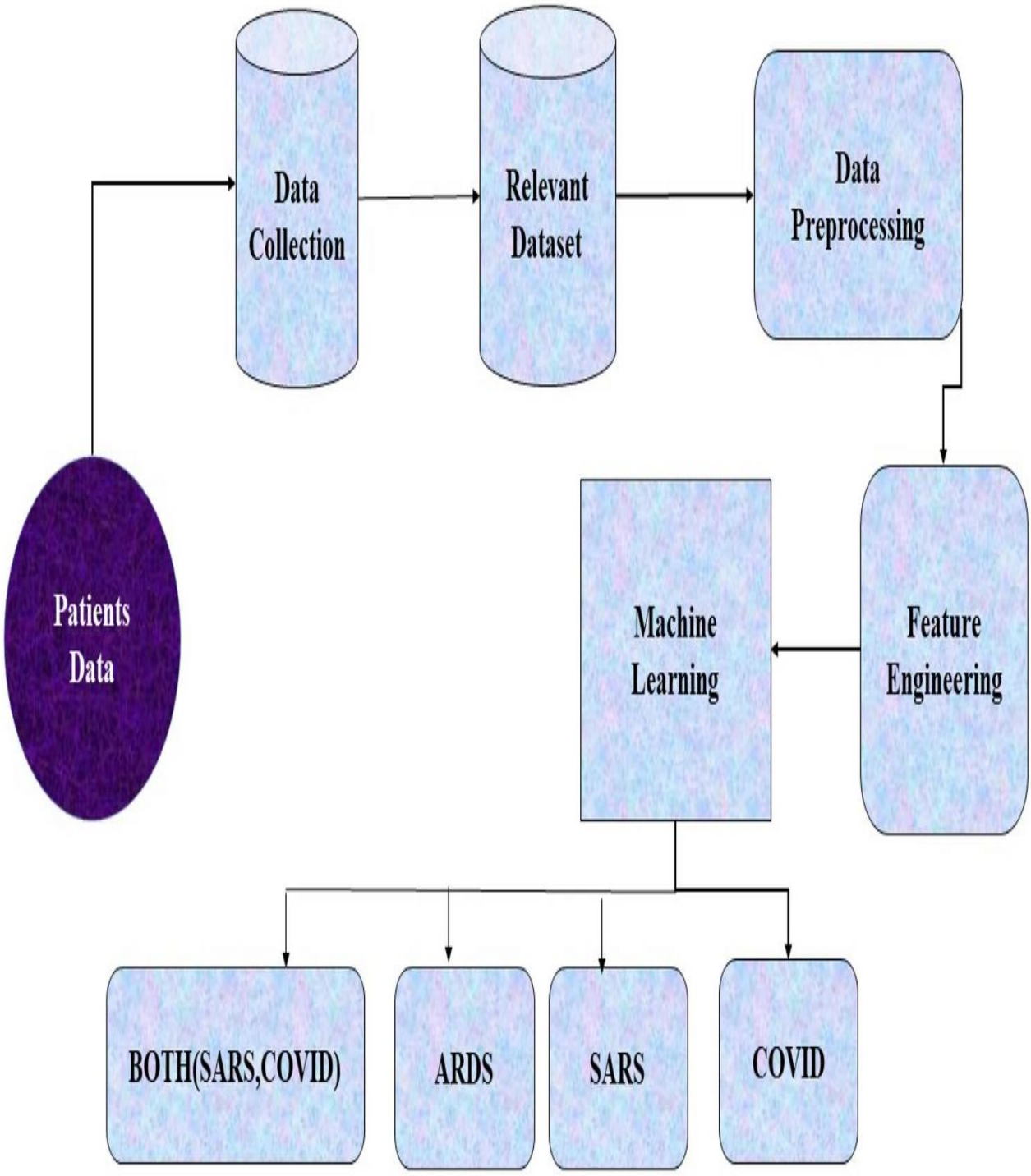
با در نظر گرفتن ذخیره سازی، تجزیه و تحلیل داده های بیماری کووید-۱۹ ، تحلیل کلان داده برای این بیماری عمدتا بوسیله تکنیک های کلیدی از علم کلان داده مشخص می گردد .این تکنیک ها شامل تحلیل مجموعه داده چند حوز ه ای(ترکیب داده های سری زمانی وتصاویر بیماری)، آنالیز عمیق(با استفاده از یادگیری ماشین و مدل مارکوف پنهان و شبکه عصبی عمیق) و محاسبات موازی انجام می شود.

یک مدل مفهومی مبتنی بر تکنولوژی کلان داده ها به منظور پیشگیری و کنترل بیماری کووید-۱۹ در چهار سطح جمع آوری ، داده ، تجزیه و تحلیل و کاربرد نشان میدهد.

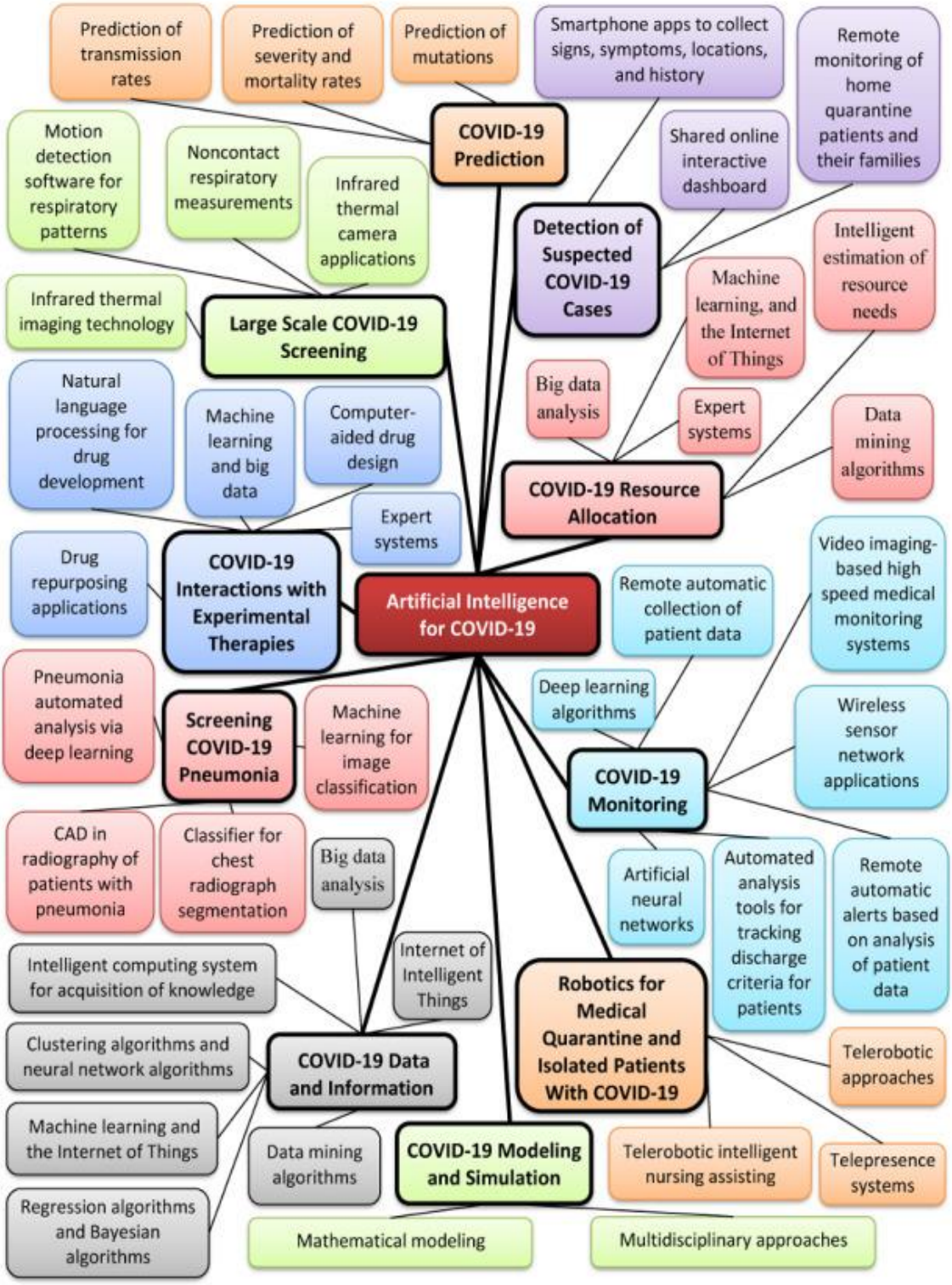
در پژوهشی از داده های مربوط به شیوع بیماری که از منابع موثق نظیرکمیسیون ملی بهداشت و شهرداری چین جمع اوری گردیده است استفاده شده است . این کلان داده می تواند به اجرای مدل سازی همه گیر برای تفسیر تجمعی کمک کند. تعداد افراد مبتلا به این بیماری در پنج منطقه مختلف از کشور چین مورد بررسی قرار گرفته است. این اطلاعات به محققان اجازه می دهد تا شبیه سازی ها را برای پیش بینی شیوع کووید-۱۹ انجام دهند. به عنوان مثال شناسایی مناطق با خطرات بالای همه گیر شدن و تشخیص جمعیت با افزایش موارد آلوده، که همه آن ها به موفقیت فعالیت های ضد همه گیر کمک می کنند.



شکل۳-۴ اجرای مدل سازی همه گیر برای تفسیر تجمعی



شکل ۳-۵ طبقه بندی بیماری های تنفسی



شکل 3- 6 یک نمای کلی از استفاده هوش مصنوعی در زمینه بیماری کووید-19

گام 4- برخی از روش های آماری مورد استفاده در این پژوهش

شبکه عصبی:

شبکه عصبی سیستم‌ها و روش‌های محاسباتی جدید هستند که برای یادگیری ماشینی، نمایش دانش و در انتها اعمال دانش جهت پیش‌بینی پاسخ‌های خروجی از سیستم پیچیده می‌باشد.

به طورکلی شبکه عصبی:

۱) روشی برای محاسبه است که بر پایه اتصال به هم پیوسته چندین واحد پردازشی ساخته می‌شود.

۲) شبکه از تعداد دلخواهی سلول یا گره یا واحد یا نرون تشکیل می‌شود که مجموعه ورودی را به خروجی ربط می‌دهند.

۳) شبکه عصبی مصنوعی روشی عملی برای یادگیری توابع گوناگون نظیر توابع با مقادیر حقیقی، توابع با مقادیر گسسته و توابع با مقادیر برداری می‌باشد.

۴) یادگیری شبکه عصبی در برابر خطاهای داده‌های آموزشی مصون بوده و اینگونه شبکه ها با موفقیت به مسائلی نظیر شناسایی گفتار، شناسایی و تعبیر تصاویر و یادگیری ربات اعمال شده است.

اجزای شبکه عصبی عبارتست از:

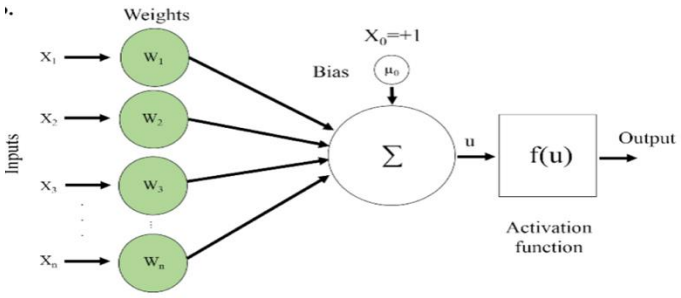
ورودی و خروجی ها: اعداد و ارقام در قالب یک یا چند متغیر، ورودی های یک شبکه عصبی را تشکیل می‌دهند. این ورودی ها پس از انجام تحلیل و پردازش‌های خاص به یک یا چند متغیر خروجی تبدیل می‌شوند. ورودی‌ها نقش متغیر مستقل و خروجی‌ها نقش متغیر وابسته را بر عهده دارند.

نرون ها: مهمترین جزء سیستم عصبی نرون‌ها هستند که به سه دسته نرون‌های ورودی، خروجی و پنهان تقسیم می‌شوند و در قالب لایه ورودی،لایه خروجی و لایه پنهان یا میانی قرار می‌گیرند. نرون‌ها یا واحدهای ورودی وظیفه دریافت داده‌های ورودی را بر عهده دارند. لایه‌های میانی و خروجی شامل واحدهای پردازش اطلاعات هستند که

در این واحدها عملیات جبری بر اطلاعات ورودی انجام و نتیجه آنها به صورت یک ورودی جدید به واحدهای دیگر در لایه های بعدی ارسال می‌شود.

وزن ها: متغیرهای مختلف ورودی شبکه ارزش‌های مختلف دارند که به کمک وزن‌ها به آنها اختصاص می‌یابد. این وزن‌ها که قبل از لایه خروجی و لایه پنهانی لحاظ می‌شوند با روش اعداد تصادفی تولید و در استفاده از شبکه تصحیح می‌شوند.

مفاهیم فوق در قالب شکل زیر که نمایی کلی از یک شبکه عصبی ساده است، نمایش داده شده است.



شکل 3-7 نمای کلی از یک شبکه عصبی

شبکه عصبی کانولوشن:

یک شبکه عصبی کانولوشنی(CNN) یک الگوریتم یادگیری عمیق بوده که قادر است یک ورودی تصویری را دریافت کند. به جنبه‌ها یا اشیای مختلف موجود در تصویر، اهمیت (وزن‌ها و بایاس‌های قابل یادگیری) اختصاص دهد و در نهایت بتواند یکی را از دیگری متمایز کند. پیش پردازش مورد نیاز در یک شبکه کانولوشنی، به مراتب کمتر از سایر الگوریتم‌های دسته‌بندی است .همچنین این شبکه‌ها، توانایی یادگیری فیلترها یا ویژگی‌ها را دارند، در حالی که در روش‌های اولیه، این فیلترها به صورت دستی طراحی و مهندسی می‌شدند.

معماری شبکه کانولوشنی، مشابه الگوی اتصال نورون‌ها در مغز انسان است و از قشر دیداری الهام گرفته شده است. نورون‌ها تنها در یک ناحیه محدود از میدان دید به محرک‌ها پاسخ می‌دهند که به آن، میدان پذیرش (Receptive Field) گفته می‌شود. مجموعه‌ای از این میدان‌ها، با هم همپوشانی دارند تا کل ناحیه‌ی بینایی را

پوشش دهند. یک شبکه کانولوشنی قادر است با استفاده از فیلترهای مربوطه، وابستگی‌های مکانی و زمانی یک تصویر را با موفقیت ضبط و ثبت کند. این معماری، به واسطه کاهش تعداد پارامترهای درگیر و امکان استفاده‌ی مجدد از وزن‌ها، سازگاری بهتری در مواجه با مجموعه داده های تصویری از خود نشان می‌دهد. به بیان دیگر، این شبکه می‌تواند به گونه‌ای آموزش ببیند که پیچیدگی تصاویر را بهتر درک کند.

در پژوهش انجام شده از شبکه عصبی کانولوشنی یک بعدی جهت طبقه بندی بر اساس داده های صوتی و سایر علائم کاربر استفاده شده است.

3-3 فرایند پژوهش

فرایند پژوهش، بیانگر مجموعه اقدامات متوالی، مرتبط و منظمی است که از ابتدا تا انتهای پژوهش برای دستیابی به اهداف مورد نظر انجام شده است. شکل 3-8 فرایند پژوهش را نشان می‌دهد.

3-4 ابزار گردآوری اطلاعات

گردآوری اطلاعات یکی از اصلی ترین بخش‌های هر کار پژوهشی را تشکیل می‌دهد .چنانچه این کار به شکل منظم وصحیح صورت پذیرد کار تجزیه و تحلیل و نتیجه گیری از داده ها با سرعت و دقت خوبی انجام خواهد شد. روش‌های گردآوری اطلاعات پژوهش به دو دسته کتابخانه‌ای و میدانی تقسیم می‌شود. در خصوص گردآوری اطلاعات مربوط به ادبیات موضوع و پیشینه پژوهش از روش‌های کتابخان‌ای و جهت جمع آوری اطلاعات برای تایید یا رد فرضیه‌های پژوهش از روش میدانی استفاده می‌شود. ابزارهای گردآوری اطلاعات به دو دسته کلی کتابخانه‌ای و میدانی قابل تقسیم است. روش‌های میدانی نیز شامل ابزارهایی مانند مشاهده، مصاحبه و پرسشنامه است.در چارچوب روش گردآوری اطلاعات، از ابزاری برای گردآوری اطلاعات مورد نیاز استفاده شده است که بررسی

اسناد و مدارک[۴۹] (فیش‌برداری) نامیده می‌شود. اسناد و مدارک مورد بررسی در این پژوهش، شامل آثار مکتوب و الکترونیکی از قبیل کتب، مجلات، مقالات و وبسایت‌ها است. این اقدام، مبنای طراحی الگوریتم های یادگیری ماشین بر اساس معیارهای پیشنهادی پژوهش‌گران بوده است.

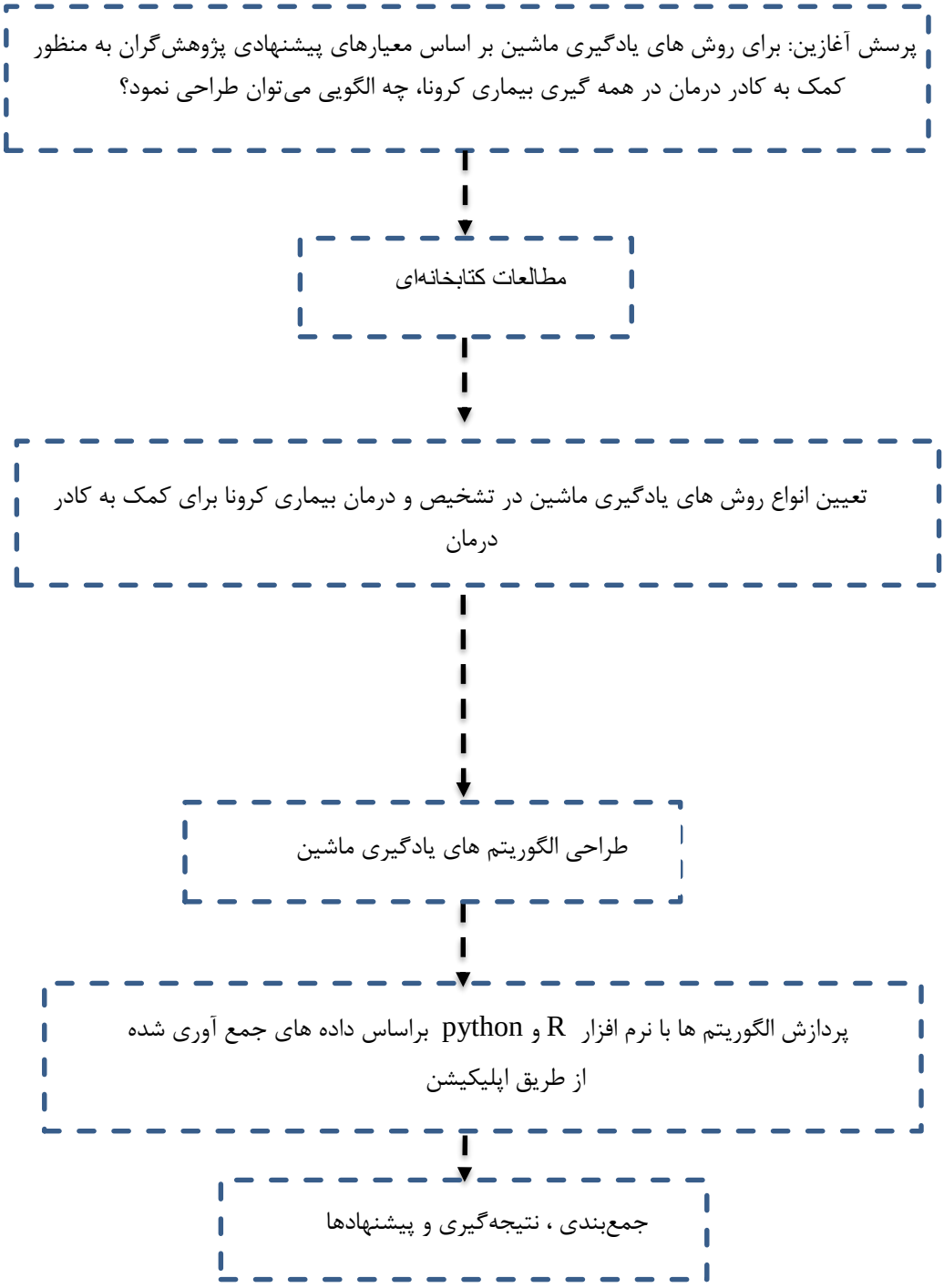
3-5 ابزار گردآوری داده

یکی از مهم ترین مراحل در روش شناسی هر پژوهش، جمع آوری داده ها یا Data Collecting است بطوری که در هر پژوهش باید برای رسیدن به اهداف و بررسی فرضیات و سوالات خود داده ها را جمع آوری کرده و سپس به آنالیز و تحلیل آن ها بپردازند. در این پژوهش برای جمع آوری داده ها از نرم افزار طراحی شده در این پروژه استفاده می شود برای توضیح مراحل جمع آوری داده ابتدا باید چند واژه را تعریف نمود.

پاکسازی داده

پاکسازی داده فرایندی است که طی آن، اجزای ناکامل، نادقیق، ناصحیح و ناسازگار داده شناسایی شده و اصلاح یا حذف می‌شوند. پاکسازی داده در آمار عبارت از فرایند حذف یا جانهی نقاط داده‌ای نامعتبر از یک مجموعه داده، بر اساس فرضیه‌هایی در مورد توزیع داده و با استفاده از تحلیل‌های آماری است.

[49]Studying of Documents

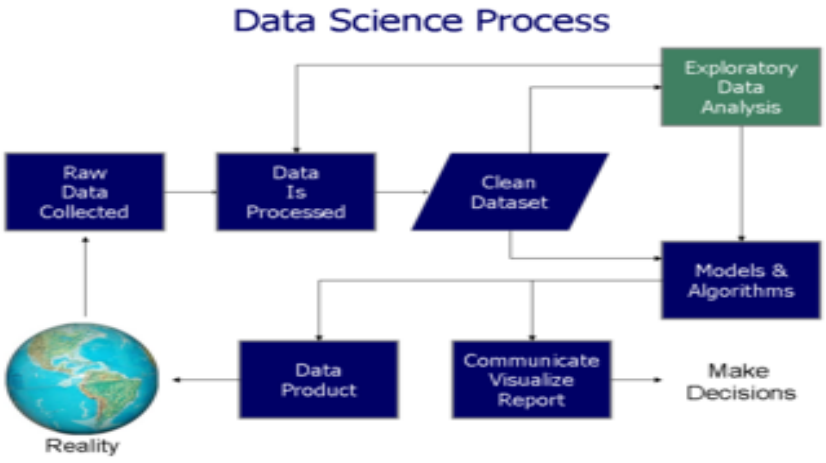


شکل ۳-۸ فرایند کلی پژوهش

کیفیت داده

کیفیت داده، عبارت است از تمام ویژگی‌ها و خصوصیاتی که لازم است داده داشته باشد تا هدف مورد نظر کاربر برآورده شود. داده‌ی باکیفیت، داده‌ای کامل، سازگار، بهنگام، دقیق، در دسترس، مبتنی بر استانداردها و دارای اعتبار است. تعاریف زیادی برای کیفیت داده وجود دارد، اما به طور کلی، داده‌ای دارای کیفیت بالا است که برای عملکرد مورد نظر خود مناسب باشد.

به طور کلی روش کار این اپلیکیشن به این صورت است؛ ابتدا کاربر با استفاده از اپلیکیشن طراحی شده اطلاعاتی را به سرور ارسال خواهد کرد(در همین زمان به مجموعه ای از داده های حیاتی پزشکی مربوط به کرونا و سایر پارامترها دست خواهیم یافت) بعد از دریافت اطلاعات کاربر از سرور ابتدا بر روی داده ها عملیات پیش پردازش داده از جمله روش های پاکسازی داده و تبدیل داده صورت خواهد گرفت. در مرحله ی بعد با استفاده از مدل طراحی شده مبتنی بر روش های یادگیری ماشین توسط پژوهشگران داده های پاکسازی شده با استفاده از آن مدل به تجزیه و تحلیل اطلاعات ارسالی کاربر می پردازد. و در نهایت الگوریتم طراحی شده از روی ۷۰ درصد داده های موجود و داده های ارسالی توسط کابر از طریق اپلیکیشن ، عملیات یادگیری را انجام می‌دهد و از روی ۳۰ درصد بقیه، خود را ارزیابی می‌کند و نتیجه‌ی تست را به کاربر ارسال خواهد کرد.



شکل 3-9 مراحل کلی کار علم داده

6-3 خلاصه

با توجه به کاربردهای فراوان هوش مصنوعی و داده کاوی در شیوع بیماری های واگیرداری مانند کرونا، ضروری است تا نیروهای متخصص مسلط به هوش مصنوعی و علوم پزشکی و استارت آپ های این حوزه برای کمک به بهبود شرایط کنونی وارد عمل شوند. همچنین داده های تولیدی در پزشکی و درمانی برای استفاده در فناوری های هوش مصنوعی باید تا حد امکان دیجیتالی شوند. در حال حاضر با توجه به شیوع ویروس کرونا باید از استارت آپ های پردازش متن در تحلیل داده های تولیدی در شبکه های اجتماعی استفاده کرد.

هوش مصنوعی و یادگیری ماشین ابزاری آینده نگر و مفید برای شناسایی عفونت‌های زودرسی مانند ویروس کرونا است و همچنین به نظارت بر وضعیت بیماران آلوده کمک می‌کند. با ایجاد الگوریتم های مفید می‌تواند به طور قابل توجهی تصمیم گیری را بهبود بخشد. این دانش نه تنها در درمان مبتلایان به کوید-۱۹ بلکه برای نظارت بر سلامتی آنها مفید است. کلان داده ها می‌تواند بحران کوید-۱۹ را در مقیاس‌های مختلف مانند عرصه‌های پزشکی، مولکولی و اپیدمیولوژیکی ردیابی کند. همچنین تسهیل تحقیقات در مورد این ویروس با استفاده از تجزیه و تحلیل داده‌های موجود مفید است. هوش مصنوعی و یادگیری ماشین می‌تواند در ایجاد رژیم‌های درمانی مناسب، راهکارهای پیشگیری، ساخت دارو و واکسن کمک کند.

فلذا همچنان در پیش بینی و درمان این ویروس از پتانسیل های بالقوه فناوری های ذکر شده استفاده های موثرتری می تواند صورت پذیرد و استفاده از ظرفیت کامل این فناوری ها مستلزم اقدامات تکمیلی در این حوزه می باشد. امید است نتایج بررسی های انجام شده راه گشای کادر درمان در مبارزه با این بیماری باشد.

فصل چهارم:
تجزیه و تحلیل داده ها

4-1 مقدمه

در حال حاضر جهان با یک بحران همه گیر در خصوص کرونا ویروس ۲۰۱۹ مواجه است سرعت انتقال ومقیاس جهانی این ویروس به قدری وسیع است که نیاز است از فناوری های هوشمند انقلاب چهارم یا هوش مصنوعی استفاده شود. در واقع با استفاده از هوش مصنوعی می توانیم سرعت تشخیص را به شدت افزایش دهیم؛ در حال حاضر هوش مصنوعی و یادگیری ماشین به دلیل قابلیت بالای استخراج موفقیت های بزرگی را در زمینه ی تصویر برداری پزشکی و همچنین افزایش دقت تشخیص ، توسعه ی رویکردهای جدید درمانی و همچنین درک بهتر الگوی گسترش ، کسب کرده است. در حال حاضر همه دولت ها نیاز به یک بانک داده بزرگ و تحقیقاتی روزآمد دارند که به طور مداوم مقالات علمی مرتبط را مستند کند تا پروژه های تحقیقاتی جدید را که به داده های زمان واقعی نیاز دارند سریع تر کنند . بیماران می توانند با روش های یادگیری ماشین تجزیه و تحلیل شوند تا نه تنها ویژگی های قابل اعتماد را شناسایی کنند بلکه انجام یک پیش بینی و تشخیص خطر آماده سازی در برابر بیماری های همه گیر آینده را فراهم کنند و بتواند روند درمان آن ها را بهبود بخشد و هزینه های درمان را برای تشخیص بیماری کاهش دهد. در واقع هوش مصنوعی از طریق شبیه سازی محیط و متغیرها ، روند تست را سرعت می بخشد.

در فصل های گذشته در مورد مفهوم یادگیری ماشین و هوش مصنوعی و کاربرد های آن در کرونا، بحث شد. اقدامات انجام گرفته در این فصل، به سه بخش تقسیم شده است. ابتدا به صورت مختصر روش‌های پاکسازی داده جهت پردازش داده ها در این پژوهش مورد استفاده قرار گرفت بحث شده است. سپس مدل و الگوریتم مورد استفاده در این پژوهش برای بحث پردازش صوت مورد مطالعه قرار گرفته است. در نهایت، در بخش سوم به تجزیه و تحلیل داده ها که توسط پژوهشگران از نرم افزار طراحی شده استخراج شده، بحث گردیده است.

4-2 بخش اول: روش‌های پاکسازی داده

4-2-1 انواع روش‌های پاکسازی داده‌ی آلوده

در این قسمت روش‌های پاکسازی شامل پاکسازی داده‌ی پرت[۵۰]، داده‌ی گمشده[۵۱]، داده‌ی نامربوط، داده‌ی تکراری[۵۲]، استانداردسازی[۵۳] داده و مقیاس‌سازی[۵۴] داده بیان شده است. پاکسازی داده با توجه به نوع مسئله و نوع داده، شامل تکنیک‌های متفاوتی است. روش‌های متفاوت مزایا و نقاط ضعف مربوط به خود را دارند. در پاکسازی داده، داده‌ی نادرست ممکن است حذف، تصحیح و یا جانهی شود.

4-2-2-1 پاکسازی داده‌های پرت

مقادیر پرت، مقادیری هستند که به طور معناداری از سایر مشاهدات فاصله دارند. قانونی برای تعریف داده‌ی پرت وجود ندارد. اگر متغیری دارای توزیع نرمال باشد، هر مقداری که بیشتر از سه برابر انحراف استاندارد از میانگین فاصله داشته باشد به عنوان داده‌ی پرت شناخته می‌شود. داده‌ی پرت ممکن است خطا یا یک داده‌ی پرت واقعی باشد، یا امکان دارد به علت طبقه‌بندی اشتباه به وجود آمده باشد. همچنین، داده‌ی پرت می‌تواند به دلیل خطای دستورالعمل رخ دهد. به عنوان مثال، اگر به جای 100.00، مقدار 10000 ثبت شود، داده‌ی ثبت شده با مقدار اصلی تفاوت فاحشی دارد. قبل از اجرای هر تحلیل آماری، بسیار مهم است که وجود داده‌ی پرت در مجموعه داده بررسی شود. زیرا وجود آن منجر به تحریف تحلیل و در نتیجه نتایج غیرقابل اعتماد و تصمیم‌گیری نادرست می‌شود. داده‌ی پرت را نباید حذف کرد مگر آن که دلیل خوبی برای این کار وجود داشته باشد.

---

[50] Outliers
[51] Missing values
[52] Duplicate data
[53] Standardize
[54] Scaling

معمولاً برای تشخیص تک متغیره‌ی داده‌ی پرت از نمودار جعبه‌ای استفاده می‌شود. داده‌هایی که از چارک اول یا سوم نمودار جعبه‌ای، بیش از 1/5 برابر فاصله‌ی میان چارکی فاصله داشته باشند به عنوان داده‌ی پرت تلقی می‌شوند. یک روش دیگر برای شناسایی داده‌ی پرت در مجموعه داده، استفاده از رگرسیون است. به طور کلی، در تحلیل رگرسیونی سه نوع داده‌ی پرت وجود دارد. نوع اول آن، داده‌ی پرتی است که مؤلفه‌ی x آن به طور معناداری از سایر xهای مجموعه داده فاصله دارد. نوع دوم آن، داده‌ی پرتی است که مؤلفه‌ی y آن به طور معناداری از سایر yهای مجموعه داده فاصله دارد. نوع سوم داده‌های پرتی هستند که مؤلفه‌های x و y هر دو به طور معناداری از سایر مؤلفه‌های مجموعه داده فاصله دارند.

در ادامه، انواع روش‌های برخورد با داده‌ی پرت معرفی شده است:

1. داده پرت کنار گذاشته شود: معمولاً داده‌ی پرت توزیع داده را چوله می‌کند. بنابراین منطقی به نظر می‌رسد که هنگام تحلیل داده‌ها آن را نادیده بگیریم.
2. داده پرت با مقادیر ثابتی مانند چندک اول یا سوم جانهی شود.
3. داده‌ی پرت با مقدار NA جانهی شود و متعاقباً طی عملیات جانهی داده‌ی گمشده، مقدار مناسبی به آن تخصیص داده شود.

برای بررسی داده‌های پرت با نرم‌افزار R ابتدا برای بررسی تک متغیره نمودار جعبه‌ای هر یک از متغیرها را رسم می‌کنیم و برای بررسی دو متغیره‌ی این مجموعه داده، نمودار پراکندگی را به همراه خط برازش رسم می‌کنیم. همچنین برای شناسایی داده‌های پرت در رگرسیون می‌توانیم نمودار فواصل کوک[55] را رسم کنیم. در فواصل کوک، اثر تمام مقادیر برازش شده به طور همزمان اندازه‌گیری می‌شود. با استفاده از تابع scores() از بسته نرم‌افزار outliers می‌توان برای داده‌های یک متغیر، امتیاز z، t، chisq، iqr و mad را محاسبه نمود و مقادیر فراتر از

[55]Cook’s distances

یک آستانه‌ی تعیین شده را به عنوان داده‌های پرت شناسایی کرد. بسته نرم‌افزارهای مفید دیگر برای تعیین مقادیر پرت extremevalues، SeleMix و tsoutliers هستند.

4-2-2-2 پاکسازی داده‌های گمشده

از آنجا که جمع‌آوری داده‌های کامل در پژوهش‌های عملی چندان مقدور نیست، مقادیر گمشده اغلب در تمام پژوهش‌ها وجود دارند و مشکل اساسی در تجزیه و تحلیل این داده‌ها به شمار می‌روند. مقادیر گمشده حجم اطلاعات را کاهش می‌دهد و موجب عدم تطابق نمونه و جامعه می‌شود. میزان این گمشدگی می‌تواند در نتایج تحلیل به صورت متفاوت اثرگذار بوده و منجر به نتیجه‌گیری اشتباه شود. هر چه درصد مقادیر گمشده افزایش یابد، به تبع آن میزان اریبی برآوردها نیز افزایش می‌یابد. از طرف دیگر، استفاده از روش‌های برآورد پارامترها نیازمند داده‌های کامل است. لذا در هنگام برخورد با داده‌های ناکامل، استفاده از روش‌های مناسب جایگذاری ضروری می‌شود. در این خصوص موضوع اساسی برای یک تحلیل‌گر، تشخیص پارامترهای مناسب مدل جانهی است. زیرا که نامشخص بودن این پارامترها می‌تواند دلیلی بالقوه برای اریبی باشد.

پیشنهاد می‌شود قبل از استانداردسازی و نرمال‌سازی داده، برخورد مناسب با داده‌ی گمشده صورت گیرد. هنگام مواجه شدن با داده‌ی گمشده ابتدا بایستی مکانیسم آن را به درستی تشخیص داد. قبل از آن که در مورد نوع برخورد با داده‌ی گمشده تصمیم‌گیری شود، لازم است گام‌های زیر طی شود.[39]

1) بررسی دقیق بودن کدگذاری داده
2) شناسایی متغیرهای شامل مقادیر گمشده
3) شناسایی الگوی گم‌شدگی در مجموعه داده
4) بررسی همبستگی میان داده‌های گمشده و مشاهده شده
5) برخورد مناسب با داده‌های گمشده

برای بررسی قالب‌بندی صحیح متغیرها دستورات متفاوتی مانند class() و str() در نرم‌افزار R وجود دارد. با استفاده از دستور vis_dat از بسته نرم‌افزار visdat می‌توان نوع متغیرها و پراکندگی مقادیر گمشده در داده را بررسی کرد. همچنین برای بررسی تعداد فاکتورها و تشخیص صحیح بودن سطوح آن‌ها می‌توان از دستوراتی مانند str() و تابع ff_glimpse() از بسته نرم‌افزار finalfit استفاده کرد. و برای رسم این نمودار از دستور vis_miss در بسته نرم‌افزار visdat استفاده شده است.

برای شناسایی توزیع گم‌شدگی در مجموعه داده می‌توان نمودار missing_plot() (از بسته نرم‌افزار finalfit) را رسم کرد و آن را مورد بررسی قرار داد. در این نمودار، محور x نشان دهنده‌ی تعداد مشاهدات است و محور y اسامی متغیرها را نشان می‌دهد. با استفاده از این نمودار می‌توان وابستگی بین گم‌شدگی و مشاهدات و همچنین رابطه‌ی گم‌شدگی در متغیرهای مختلف را نیز بررسی کرد.

برای شناسایی الگوی گم‌شدگی داده، بسته نرم‌افزارهای متعددی وجود دارد. با استفاده از تابع aggr در بسته نرم‌افزار VIM و تابع marginplot در بسته نرم‌افزار VIM می‌توان الگوی گم‌شدگی در داده را مورد بررسی قرار داد.

علاوه بر آن، با استفاده از تابع gg_miss_upset از بسته نرم‌افزار naniar می‌توان نمودار دیگری برای تشخیص الگوی گمشدگی در داده رسم کرد.

4-2-2-3 پاکسازی داده‌های نامربوط

با استفاده از دانش زمینه‌ای و اطلاعات گذشته باید داده‌های نامربوط را شناسایی نموده و از مجموعه داده حذف کرد. داده‌ی نامربوط[۵۶] به داده‌ای گفته می‌شود که در حقیقت به آن نیاز نیست و در مسئله‌ای که می‌خواهیم آن را حل کنیم به کار نمی‌رود. به عنوان مثال، فرض کنید یک مجموعه داده مربوط به سلامت عمومی را تحلیل می‌کنیم. متغیر شماره‌ی تماس، یک متغیر نامربوط است. به طور مشابه، اگر در تحلیل داده‌ی مذکور، فقط شهر مخصوصی مد نظر ما باشد سایر شهرها موارد نامربوط به حساب می‌آیند. همچنین اگر بخواهیم فقط بیمارانی را مورد مطالعه قرار دهیم که تحت عمل جراحی قرار گرفته‌اند، سایر بیماران داده‌ی نامربوط محسوب می‌شوند.

4-2-2-4 پاکسازی داده‌های تکراری

بعضی از انواع آلودگی، هزینه‌ی زیادی را به سازمان‌ها تحمیل می‌کنند. داده‌های تکراری یکی از این موارد و از رایج‌ترین خطاها در مجموعه داده می‌باشد. برای پاکسازی داده‌ی تکراری ابتدا باید آن‌ها را شناسایی نمود و پس از اطمینان از زائد بودن بایستی آن‌ها را از مجموعه داده حذف کرد. در حذف موارد تکراری باید بسیار احتیاط کرد تا اطلاعات مهم از دست نروند. با استفاده از تابع duplicated از بسته نرم‌افزار data.table به سادگی می‌توان سطرهای تکراری در مجموعه داده را شناسایی کرد که البته مجموعه داده مورد استفاده ما شامل این مورد یعنی داده های تکراری نمی باشد.

4-2-2-5 پاکسازی داده‌های نامعتبر

پاکسازی داده‌های نامعتبر شامل پاکسازی نوع متغیر، طول مقادیر متغیر و حدود متغیر می‌باشد. در ادامه روش‌های پاکسازی این خطاها بررسی شده است.

[56] Irrelevant data

1) بررسی نوع متغیر

برای تعیین ساختار و نوع متغیرها توابع متعددی مانند تابع ()str، ()glimpse (از بسته نرم‌افزار tibble) و تابع ()ff_glimpse (از بسته نرم‌افزار finalfit) در نرم‌افزار R وجود دارد. تابع ()ff_glimpse پس از تفکیک متغیرهای پیوسته و رده‌ای، علاوه بر نمایش نوع متغیرها آماره‌های مهم را نیز برای هر کدام از این دو گروه متغیر خلاصه می‌کند.[39] پس از شناسایی نوع متغیرها، اگر متغیری با نوع مورد نظر مطابقت نداشته باشد بایستی نوع آن را به نوع مورد نظر تبدیل کرد. در هنگام تبدیل نوع متغیر ممکن است با خطا یا هشدار مواجه شویم. مطالعه‌ی کامل هشدارها و خطاها شدیداً توصیه می‌شود.

2) بررسی حدود متغیر

برای بررسی حدود متغیر تابع ()range در نرم‌افزار R موجود است. علاوه بر آن می‌توان توابع ()min، ()max و ()summary را نیز به کار برد.

3) خطای طول مقادیر موجود در متغیر

برای بررسی طول مقادیر یک متغیر می‌توان از بسته نرم‌افزار stringr استفاده کرد. تابع ()str_length در این بسته نرم‌افزار تعداد کاراکترهای رشته را برمی‌گرداند.

3-4 بخش دوم: مدل و الگوریتم مورد استفاده در این پژوهش

دستاورد جدید در تشخیص غیر حضوری بیماری کرونا، ارائه اپلیکیشن "تست کووید 19" با کارایی بهتر برای کاربران عادی است. این نرم افزار ازهوش مصنوعی قدرت گرفته و همان کاری که پزشکان در ابتدا ی مراجعه به آنها به منظورحدس ابتلا یا عدم ابتلا به کووید 19 انجام می دهند را تکرار میکند؛ برنامه تست کووید 19 به صدای سرفه شما گوش میکند.

در واقع نرم افزار "تست کووید 19" یک اپلیکیشن بوده که روی هرگوشی هوشمند با میکروفون سالم، قابل استفاده است. ایده پشت توسعه آن بسیار ساده است. بعد از ورود به اپلیکشن داخل میکروفون گوشی سرفه می کنید. شبیه به یک پزشک حرفه ای، که قادر است از روی نحوه سرفه کرونایی بودن یا نبودن آن را تشخیص بدهد، اپلیکیشن مذکور نیز بعد از ارزیابی نوع سرفه، نتیجه را اعلام خواهد کرد. این برنامه در واقع بر اساس صدای سرفه،احتمال ابتلا به ویروس کرونا را اعلام میکند.

بر اساس اعلام سازمان بهداشت جهانی، سرفه خشک یکی از متداول ترین نشانه های بیماری کووید 19است، به طوری که67,7 درصد مبتلایان به آن دچار خواهند شد. همچنین نباید از بیماران فاقد علائم بیماری اما ناقل کروناویروس چشم پوشی کرد. با این حال در اختیار داشتن یک ابزار اعلام شانس ابتلا به کرونا تنها از روی سرفه، در بحران کنونی، کمک زیادی خواهد کرد. افراد بعد از پی بردن به شانس بالای ابتلا، خود را قرنطینه خانگی کرده و از انتشار ویروس جلوگیری می کنند. هدف اصلی از توسعه چنین اپلیکیشنی، کاهش حجم مراجعه حضوری به پزشکان است. به این ترتیب از حضور افراد زیادی در مراکز آلوده پزشکی جلوگیری شده و به رعایت طرح فاصله گذاری اجتماعی کمک خواهد شد. در نظر داشته باشید هوش مصنوعی، در هر سیستمی، به مرور زمان خبره تر میشود. به عبارت دیگر این اپلیکیشن به مرور زمان که داده های وسیعی جمع آوری خواهد کرد به مراتب دقت بالاتری پیدا خواهد کرد و با عملکرد بالاتری قادر به تشخیص بیماری صرفا از روی صدای سرفه فرد خواهد شد. همچنین با توجه به علائم سرفه ی خشک طبق اعلام سازمان بهداشت جهانی در مورد بیماری covid19 می توان این نرم افزار را حتی برای تجزیه و تحلیل صداهای سرفه در تشخیص شرایط تنفسی مانند سیاه سرفه، آسم و ذات الریه به شرط جمع آوری داده دقیق تر و بیشتر نیز موفقیت آمیز دانست .

یکی دیگر از برنامه های تشخیص کرونا از سرفه ی بیماران برنامه COVID-19 Sounds می باشد که توسط محققان دانشگاه کمبریج توسعه داده شده است. این برنامه به عنوان وبسایت Sounds COVID-19 در دسترس قرار دارد. در این راستا همه می توانند با مراجعه به این سایت صدای دم و بازدم و سرفه خود را ضبط کنند. این اطلاعات صوتی که در سرورهای دانشگاه کمبریج ذخیره می شود، می تواند به توسعه الگوریتم های یادگیری

ماشین کمک کند که برای شناسایی خودکار بیماری"کووید 19" مورد استفاده قرار می گیرند. محققان پروژه دقت این الگوریتم را 85 درصد اعلام کردند و امیدوارند تا حد امکان صدای سرفه و تنفس افراد زیادی را جمع آوری کنند تا بتوانند به دقت قابل توجه ای برسند. اما اپلکیشن ارائه شده توسط پژوهش ما اولین اپلیکیشن داخل ایران می باشد که جهت تشخیص کووید-19 طراحی شده است.

4-3-1 مراحل انجام مدل به کار رفته:

فلوچارت زیر مراحل کامل کار اپلیکیشن پیشنهادی که آن را اپلیکیشن کووید ایران(ICA[57]) نامگذاری کرده ایم را مشخص می کند. برنامه گوشی هوشمند(اپلیکیشن) با فشار دادن دکمه از شما درخواست می کند تا صدا سرفه را ضبط نمائید. با فشار دادن دکمه تشخیص، صداهای ضبط شده بررسی می شوند و در صورت سکوت بودن و مشکل داشتن ، برنامه به کاربر اعلام می نماید که سرفه خود را مجددا ضبط نماید. در صورتی که مشکلی در ضبط صوت سرفه وجود نداشت، کاربر باید سایر علائم خود مانند تب ، سرفه، بدن درد و.... را در صورت وجود مشخص نماید و از آنجا که اپلیکیشن همزمان داده های مربوط به بیماری کووید-19 را جهت ایجاد یک مجموعه داده جمع آوری می کند کاربر باید اعلام نماید که قبلا به پزشک مراجعه کرده یا تست انجام داده است یا خیر؛ علاوه بر این گزینه های نامبرده در زیر شامل علائم بیماری کووید-19 است پس از ارسال اطلاعات، کاربر می تواند جهت کمک به جمع آوری مجموعه داده و کمک به تیم تحقیقاتی سایر سوالاتی که بیشتر مربوط به ارتباطات کاربر و استفاده از وسایل نقلیه عمومی است را نیز پاسخ دهد.

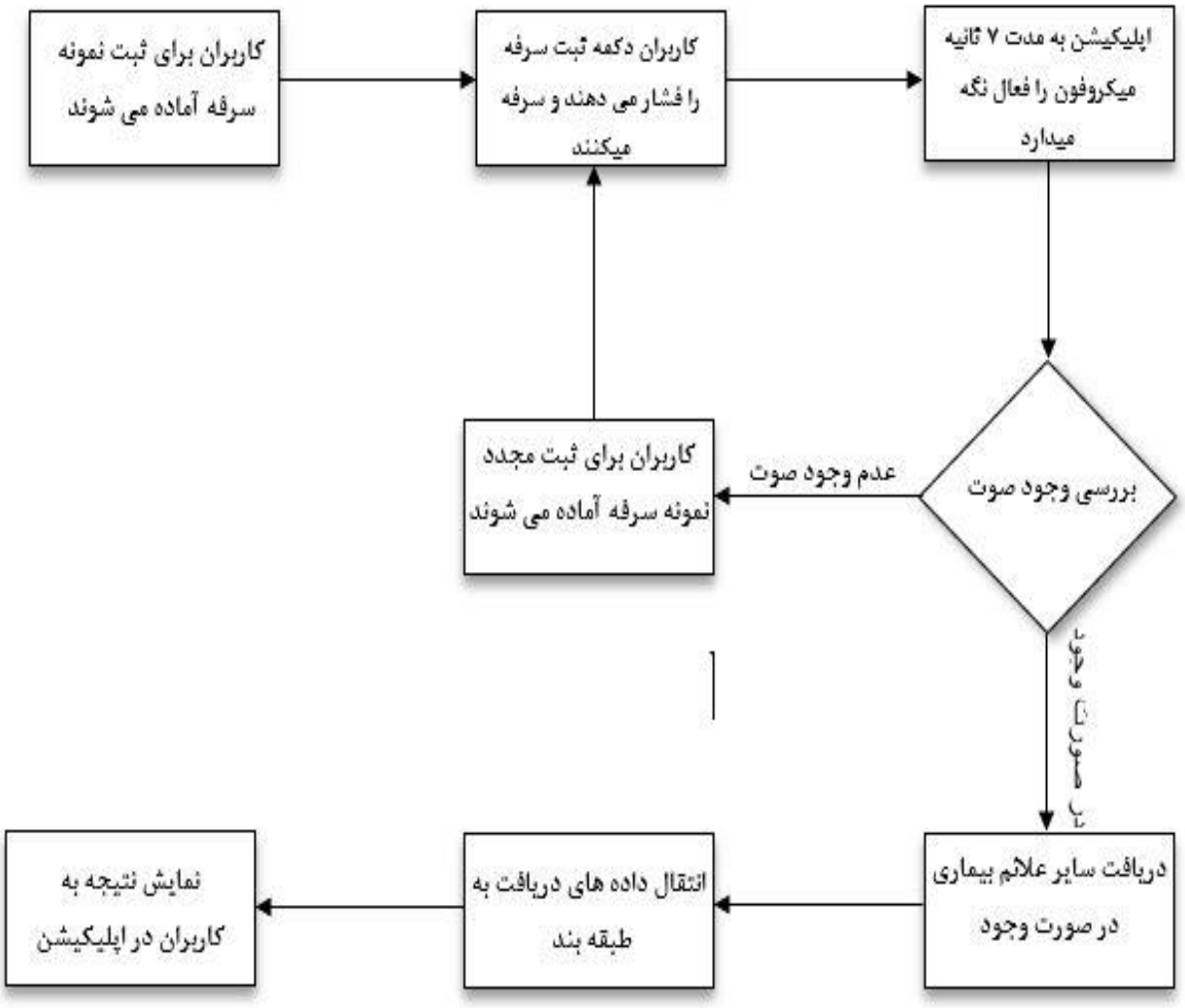
علائمی که ممکن است در یک فرد مبتلا به کووید-19 وجود داشته باشد و در اپلیکیشن آورده شده است به صورت زیر می باشد:

[57] Iran Covid App

1. تب
2. لرز
3. خستگی مفرط
4. درد مفاصل
5. سردرد
6. تنگی نفس
7. درد قفسه سینه
8. گلودرد
9. بی اشتهایی
10. سرفه خلط دار
11. سرفه خشک
12. آبریزش بینی
13. خارش گلو
14. ناراحتی معده و اسهال
15. از بین رفتن حس بویایی و چشایی

پس از دریافت هر داده مربوط به کاربر، داده وارد یک مدل تشخیص و پیش بینی کووید-19 مبتنی بر یادگیری عمیق می شود و در صورت تشخیص کووید-19 براساس داده های ورودی کاربر، پیغام مبنی بر اینکه کاربر با احتمال مشخص شده مبتلا به کووید-19 است به او نمایش داده می شود؛ در صورتی که مدل ارائه شده کاربر را سالم تشخیص داد به دلیل اهمیت تشخیص زودهنگام و جلوگیری از انتقال بیماری از فرد به افراد دیگر و افزایش قابلیت اطمینان اپلیکیشن گزینه ی "آیا قبلا تست کرونا داده اید یا به پزشک مراجعه نموده اید" بررسی

می شود و اگر کاربر پاسخ مثبت داده بود به اون پیغامی مبنی بر احتیاط داشتن و مدتی در خانه بمانید نمایش داده می شود و اما اگر مدل فرد را سالم تشخیص داد و همچنین کاربر پاسخ منفی یا خالی به گزینه گفته شده داده بود فرد سالم تشخیص داده می شود و پیغامی مبنی بر سالم بود به آن نمایش داده می شود.



شکل 1-4 فلوچارت مراحل کامل کار اپلیکیشن

4-3-2 بحث نظری و پیش پردازش مدل به کار رفته در پژوهش :

1- حذف سکوت :

قبل از شروع تشخیص، قسمت بی صدا از صدای ضبط شده باید حذف شود تا تداخل ایجاد شده در مراحل بعدی کاهش یابد و سرعت محاسبات بهبود یابد. برای حذف قسمت بی صدا، با تعیین یک مقدار آستانه که با آزمون و خطا به دست آمد، قسمت های بی صدا از قسمت های صدا حذف می شود.

2- استخراج ویژگی های صوتی :

داده های دریافت شده به صورت سیگنال های صوتی و داده های صفر و یک یا بولی[58] هستند و به دلیل تفاوت نوع داده ها باید از داده های صوتی ویژگی های مناسب و کارآمد در تشخیص بیماری استخراج شود و به صورت همزمان با داده های بولی وارد مدل ارائه شده بشوند.

داده هایی که از سیگنال های صوتی استخراج شده است به صورت زیر می باشند :

- نرخ عبور صفر (ZCR[59]): تعداد دفعاتی که علامت سیگنال صوتی از مثبت به منفی و بالعکس تغییر می کند. ZCR می تواند مولفه فرکانس غالب سیگنال را تخمین بزند و به صورت زیر محاسبه می شود:

$$Z\,(t)\;=\;\frac{1}{N-1}\sum_{t\,=1}^{N-1}\pi[y_i(t)y_i(t-1)]$$

به صورتی که تابع $\Pi[A]$ به صورت زیر است :

$$\pi[A]=\begin{cases}1, & A<0 \text{ اگر}\\ 0, & \text{در غیر اینصورت}\end{cases}$$

[58] Boolean
[59] Zero-Crossing Rate

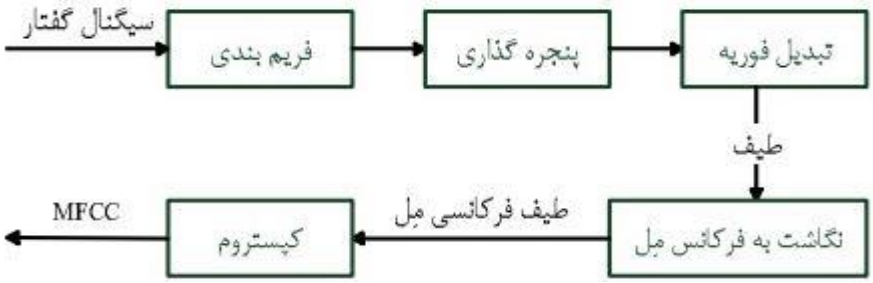
- مرکز طیفی (SC[60]): میانگین انرژی طیفی که تغییرات فرکانس سیگنال و محتوای فاز سیگنال را در طول زمان نشان می دهد. همچنین، SC به ما امکان می دهد که انرژی یک طیف در کدام فرکانس متمرکز است .این مانند یک میانگین وزنی است:

$$f_c = \frac{\sum_k S(k) f(k)}{\sum_k S(k)}$$

که در آن S(k) قدر طیفی در فرکانس k است، f(k)فرکانس در k است.

- ضرایب فرکانسی مل کپستروم MFCC[61] ایده اصلی در استخراج ضرایب MFCC خاصیت گوش انسان در دریافت و فهم گفتار است و همین مسئله این ضرایب را بـه ابـزاری قدرتمنـد در تمامی حوزه هایی پردازش و بازشناخت گفتار تبدیل کرده است. تعـداد ضرایب مورد استفاده در بازشناسی گفتار معمولاً بین 9تـا 13ضـریب تغییر میکند. ضریب صفرم نشاندهنده انـرژی اسـت که از آن بعنوان ویژگی شیمر یاد میشود. برای استخراج این ضرایب در مرحلـه اول، تبـدیل فوریه بر سیگنال اعمال میشود. سپس توان طیف بـه دسـت آمـده بـه مقیاس مل نگاشته میشود و از توان در هر فرکانس مل، لگاریتم گرفته میشود. در مرحله آخر طیف لگاریتمیِ مل به حـوزه زمـان برگردانـده میشود. نتیجه این تبدیلات نمایش کپسترال طیف سیگنال صوتی است که ویژگیهای طیفی یک فریم از سـیگنال صوتی را بـه خـوبی نشـان میدهد.

[60] Spectral Centroid
[61] mel frequency cepstral coefficients



شکل 2-4 استخراج ویژگی صوتی MFCC

- ویژگی SR[62] طیفی فرکانسی است که در زیر آن درصد مشخصی از کل انرژی طیفی، مشخص می شود. اساساً این ویژگی برای تمایز صداهای صوتی از صداهای بدون صدا استفاده می شود.

- spectral bandwidth پهنای باند طیفی مرتبه p را محاسبه می کند:

$$\left(\sum_k s(k)(f(k) - f_c)^{\,p}\right)^{\frac{1}{p}}$$

که در آن S(k) قدر طیفی k است، f(k) فرکانس در k، و fc مرکز طیفی است. وقتی p=2، این مانند یک انحراف معیار وزنی است.

- chroma_stft یک کروماگرام را از یک طیف شکل موج یا توان محاسبه کنید.
- انرژی جذر متوسط مربع[۶۳] rms ؛ انرژی یک سیگنال با بزرگی کل سیگنال مطابقت دارد. برای سیگنال های صوتی، این تقریباً با میزان بلندی سیگنال مطابقت دارد. انرژی در یک سیگنال به صورت زیر تعریف می شود:

[62] Spectral roll off
[63] root mean square

$$\sum_{n}[x(n)]^2$$

که انرژی جذر متوسط مربع نیز به صورت زیر محاسبه می شود که x نشان دهنده سیگنال صوتی است:

$$\sqrt{\frac{1}{N}\sum_{n}[x(n)]^2}$$

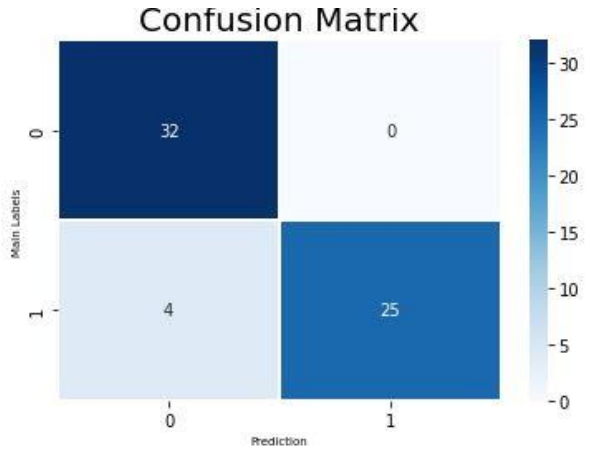
پس از استخراج ویژگی های مناسب از سیگنال صوتی از تمام ویژگی های میانگین گرفته می شود.

3- متعادل سازی[۶۴] مجموعه داده :

یکی از مسائل مهمی که کاربران مبتدی هنگام برخورد با مجموعه داده های نامتعادل با آن مواجه می شوند، به معیارهای مورد استفاده برای ارزیابی مدل آنها مربوط می شود. استفاده از معیارهای ساده‌تر مانند accuracy می‌تواند گمراه‌کننده باشد. در یک مجموعه داده با کلاس‌های بسیار نامتعادل، اگر طبقه‌بندی‌کننده همیشه رایج‌ترین کلاس را بدون انجام هیچ‌گونه تحلیلی از ویژگی‌ها «پیش‌بینی» کند، باز هم نرخ دقت بالایی خواهد داشت که بدیهی است اشتباه است. یک راه مناسب برای ارزیابی نتایج با استفاده از یک ماتریس درهم ریختگی[۶۵] است که پیش‌بینی‌های صحیح و نادرست را برای هر کلاس نشان می‌دهد. در ردیف اول، ستون اول نشان می دهد که چند کلاس 0 به درستی پیش بینی شده اند و ستون دوم، چند کلاس 0 به عنوان 1 پیش بینی شده است. در ردیف دوم، توجه می کنیم که تمام ورودی های کلاس 1 به اشتباه به عنوان کلاس 0 پیش بینی شده اند.بنابراین،

[64] Balancing
[65] Confusion matrix

هر چه مقادیر قطری ماتریس سردرگمی بالاتر باشد، بهتر است، که نشان‌دهنده بسیاری از پیش‌بینی‌های صحیح است.

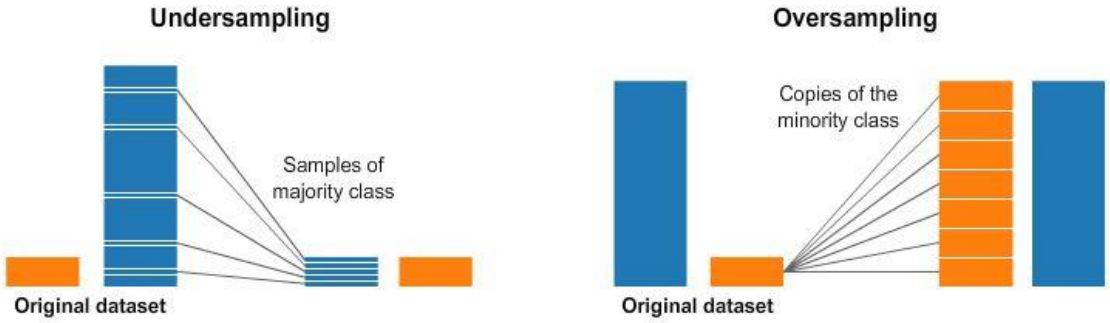


شکل 3-4 ماتریس درهم ریختگی مدل پیشنهادی

تعداد داده های جمع آوری شده مربوط به افراد مبتلا به کووید-19 نسبت به داده های افراد سالم کمتر است و قبل از ورود به مدل ارائه شده جهت تشخیص بیماری باید متعادل سازی گردد.

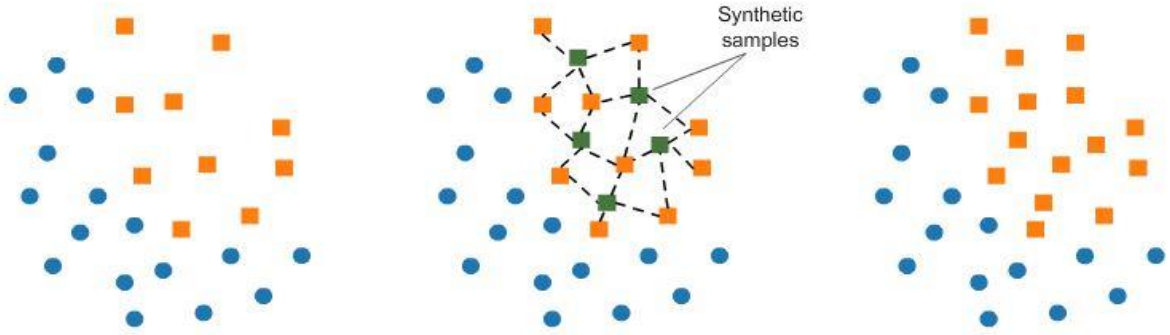
روشی که به طور گسترده برای برخورد با مجموعه داده های بسیار نامتعادل پذیرفته شده است، نمونه گیری مجدد نامیده می شود. این شامل حذف نمونه‌ها از کلاس اکثریت (نمونه‌گیری کمتر[۶۶]) و/یا افزودن نمونه‌های بیشتر از کلاس اقلیت (نمونه‌گیری بیش از حد[۶۷]) است.

[66] Under sampling
[67] Over sampling



شکل 4-4 روش های متوازن سازی مجموعه داده

با توجه به اینکه یکی از دلایل افزایش عملکرد مدل های یادگیری عمیق افزایش داده است بنابراین باید از روش های نمونه گیری بیش از حد استفاده می گردد؛ روش های مختلفی در این زمینه موجود است که پس از انجام آزمایشات مختلف به صورت آزمون و خطا، روش SMOTE انتخاب گردید؛ SMOTE (تکنیک نمونه برداری بیش از حد) از ترکیب عناصر برای کلاس اقلیت، بر اساس عناصر موجود تشکیل شده است. این به طور تصادفی کار می کند و یک نقطه از کلاس اقلیت را انتخاب می کند و k نزدیک ترین همسایه ها را برای این نقطه محاسبه می کند. نقاط مصنوعی بین نقطه انتخاب شده و همسایگان آن اضافه می شوند.



شکل 4-5 روش SMOTE جهت متوازن سازی مجموعه داده

پس از پیش پردازش داده ها ، وارد یک مدل یادگیری عمیق می شوند و مدل پس از آموزش و ارزیابی عملکرد نتایج به صورت یک پیغام سالم ، مبتلا به کووید-19 و نیاز به احتیاط به کاربر نمایش داده می شود.

4-3-3پکیج های مورد استفاده :

:Pandas -1

یک کتابخانه «متن‌باز» است که کارایی بالا، ساختاری با قابلیت استفاده آسان و ابزارهای تحلیل داده برای زبان برنامه‌نویسی پایتون را فراهم می‌کند. در واقع، می‌توان گفت پانداس یک کتابخانه قدرتمند برای تحلیل، پیش‌پردازش و بصری‌سازی داده‌ها است. همچنین برای کارهای مربوط به «علم داده» در پایتون استفاده می شود.

:Numpy -2

یک کتابخانه برای زبان برنامه نویسی پایتون است. با استفاده از این کتابخانه امکان استفاده از آرایه‌ها و ماتریس‌های بزرگ چند بعدی فراهم می‌شود. هم‌چنین می‌توان از تابع‌های ریاضیاتی سطح بالا بر روی این آرایه‌ها استفاده کرد.

:Librosa -3

لیبروسا یک بسته پایتون برای پردازش موسیقی و فایل‌های صوتی است که به کاربر این امکان را می‌دهد تا فایل‌های صوتی را در نوت بوک به عنوان آرایه NumPy برای تحلیل و دستکاری وارد کنند. برای بخش عمده پیش‌پردازش‌ها، امکان استفاده از تابع Librosa’s load() وجود دارد که به طور پیش‌فرض، نرخ

نمونه‌برداری را به 22.05 KHz تغییر می‌دهد، داده‌ها را نرمال می‌کند، بنابراین طیف مقادیر عمیق بیت بین ۱- و ۱ و کانال‌های صوتی را به صورت مونو (Mono) مسطح می‌کند.

:Keras -4

یک کتابخانه رایگان منبع باز قدرتمند و با کاربرد آسان برای توسعه و ارزیابی مدل های یادگیری عمیق است کراس دو کتابخانه یادگیری ماشین عددی Theano و TensorFlow را پوشش می دهد این امکان را می دهد فقط در چند خط کد، مدل های شبکه عصبی را تعریف و آموزش داد.

:Scikit Learn -5

از کتابخانه‌های متن‌باز، مفید، پرکاربرد و قدرتمند در زبان برنامه‌نویسی پایتون است که برای اهداف یادگیری ماشین به کار می‌رود. این کتابخانه ابزارهای کاربردی زیادی به منظور یادگیری ماشین و مدل‌سازی آماری داده‌ها همچون طبقه‌بندی(classification) ، رگرسیون، خوشه‌بندی و کاهش ابعاد فراهم می‌کند. این کتابخانه که به طور عمده توسط زبان پایتون ارائه شده، بر پایه‌ی کتابخانه‌های Numpy ، Scipyو Matplotlib طراحی شده است.

4-4 بخش سوم: تجزیه و تحلیل داده ها:

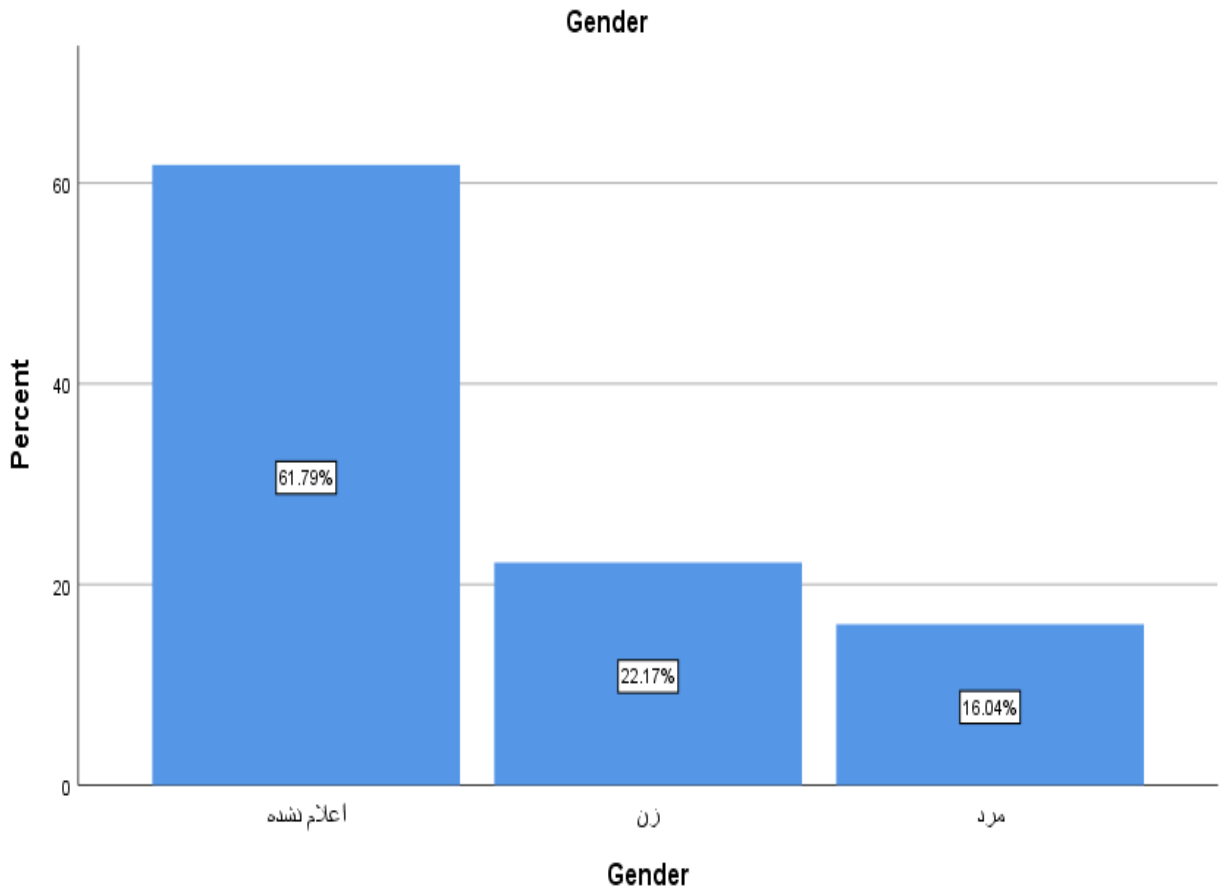
در این بخش سعی شده است تا با استفاده از داده های بدست آمده از اپلیکیشن طراحی شده ، ویژگی های جمعیت شناختی و پارامترهای دیگر تحقیق مورد بررسی، در قالب جداول و نمودار های آماری توصیف گردد.

- **توزیع پاسخ دهندگان بر حسب جنسیت:**

در جدول زیر توزیع پاسخ دهندگان بر حسب جنسیت نشان داده شده است . ملاحظه می گردد که 16 درصد پاسخ دهنگان را مردان و 22 درصد را زنان تشکیل می دهند.

جنسیت

		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیرها	اعلام نشده	131	61.8	61.8	61.8
	زن	47	22.2	22.2	84.0
	مرد	34	16.0	16.0	100.0
	کل	212	100.0	100.0	

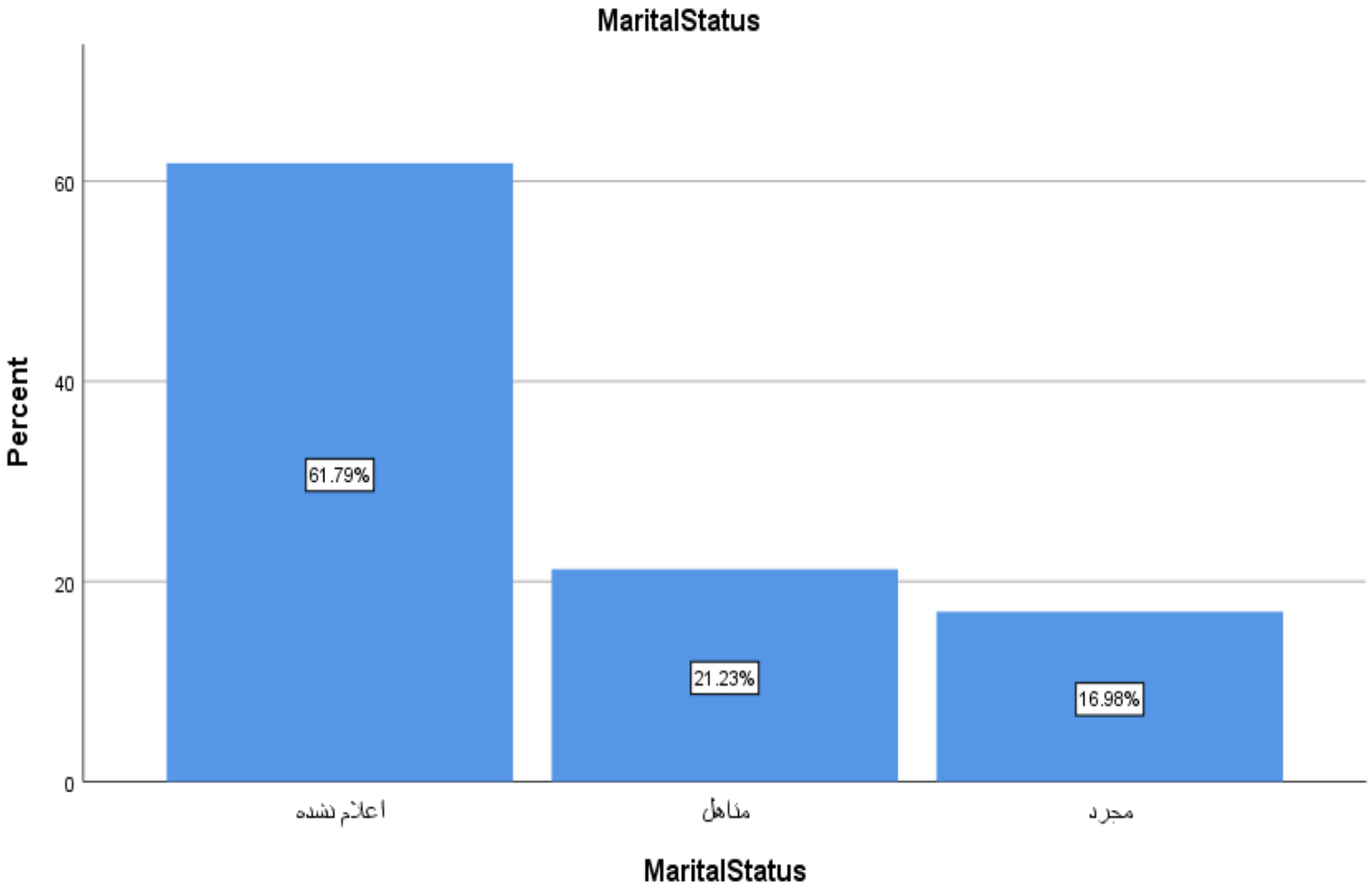


شکل6-4 توزیع پاسخ دهندگان بر حسب وضعیت تاهل

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت تاهل نشان داده شده است . ملاحظه می گردد که 21.2 درصد پاسخ دهنگان را متاهل و 17 درصد را مجردان تشکیل می دهند.

وضعیت تاهل

		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیرها	اعلام نشده	131	61.8	61.8	61.8
	متاهل	45	21.2	21.2	83.0
	مجرد	36	17.0	17.0	100.0
	کل	212	100.0	100.0	

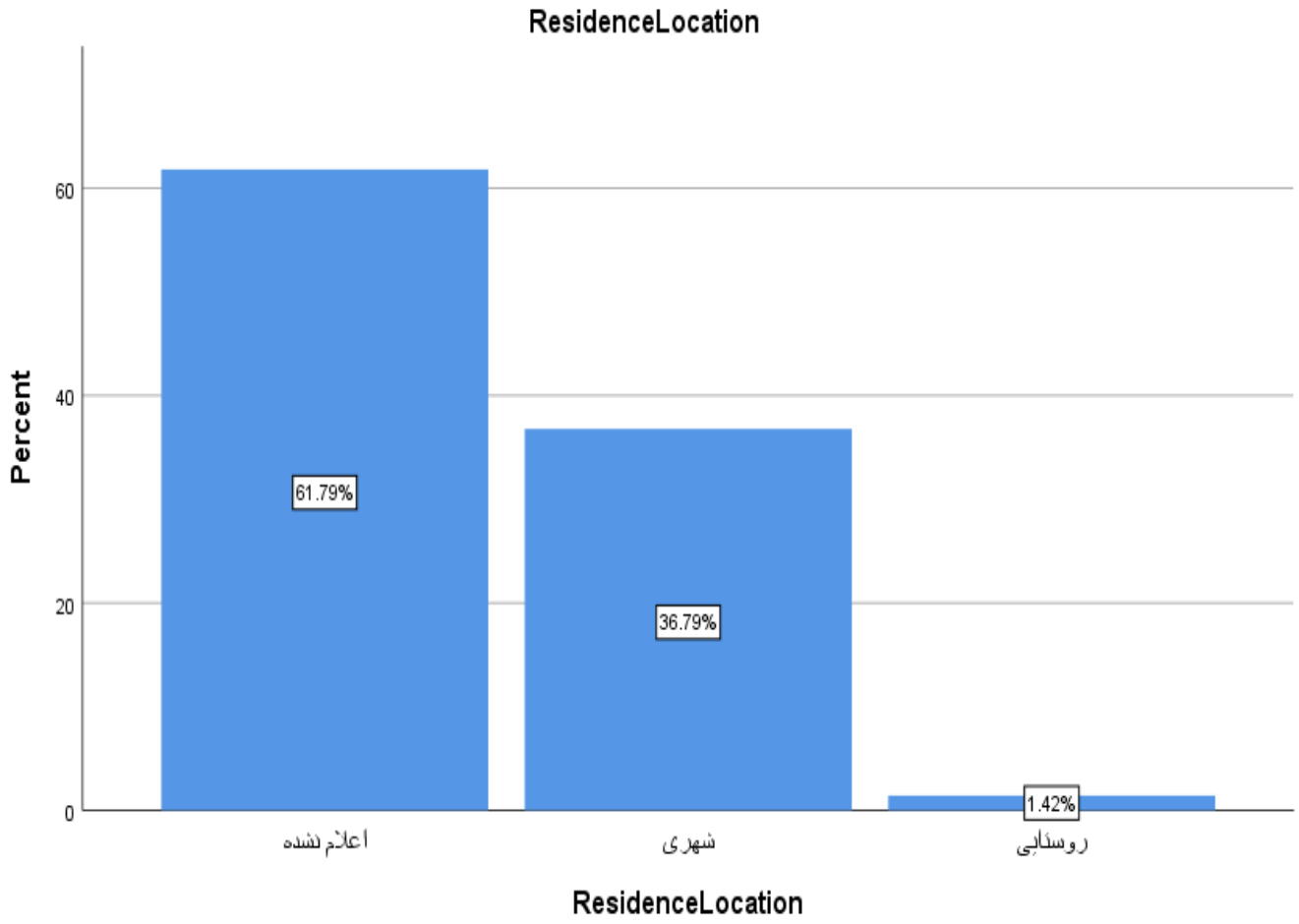


- **توزیع پاسخ دهندگان بر حسب وضعیت محل سکونت:**

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت محل سکونت نشان داده شده است . ملاحظه می گردد که 36.8 درصد پاسخ دهنگان را شهری و 1.4 درصد را روستائیان تشکیل می دهند.

محل سکونت

		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیرها	اعلام نشده	131	61.8	61.8	61.8
	شهری	78	36.8	36.8	98.6
	روستایی	3	1.4	1.4	100.0
	کل	212	100.0	100.0	

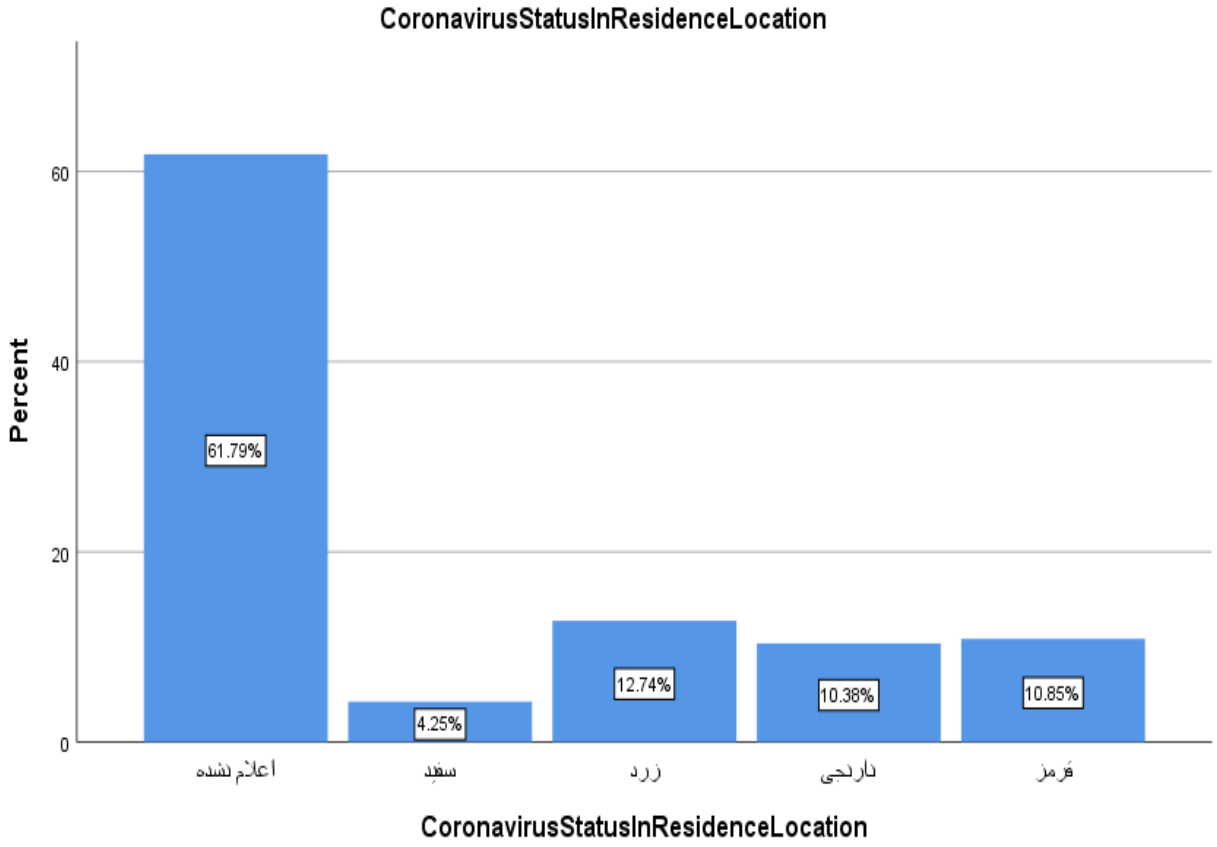


- **توزیع پاسخ دهندگان بر حسب وضعیت رنگ بندی کرونایی در محل سکونت:**

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت رنگ بندی کرونایی در محل سکونت نشان داده شده است . ملاحظه می گردد که 4.2 درصد پاسخ دهنگان رنگ بندی کرونایی شهر خود را سفید و 12.7 درصد رنگ زرد و 10.4 درصد رنگ نارنجی و در اخر 10.8 درصد رنگ قرمز را اعلام کرده اند.

رنگ بندی کرونایی محل سکونت

		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیرها	اعلام نشده	131	61.8	61.8	61.8
	سفید	9	4.2	4.2	66.0
	زرد	27	12.7	12.7	78.8
	نارنجی	22	10.4	10.4	89.2
	قرمز	23	10.8	10.8	100.0
	کل	212	100.0	100.0	

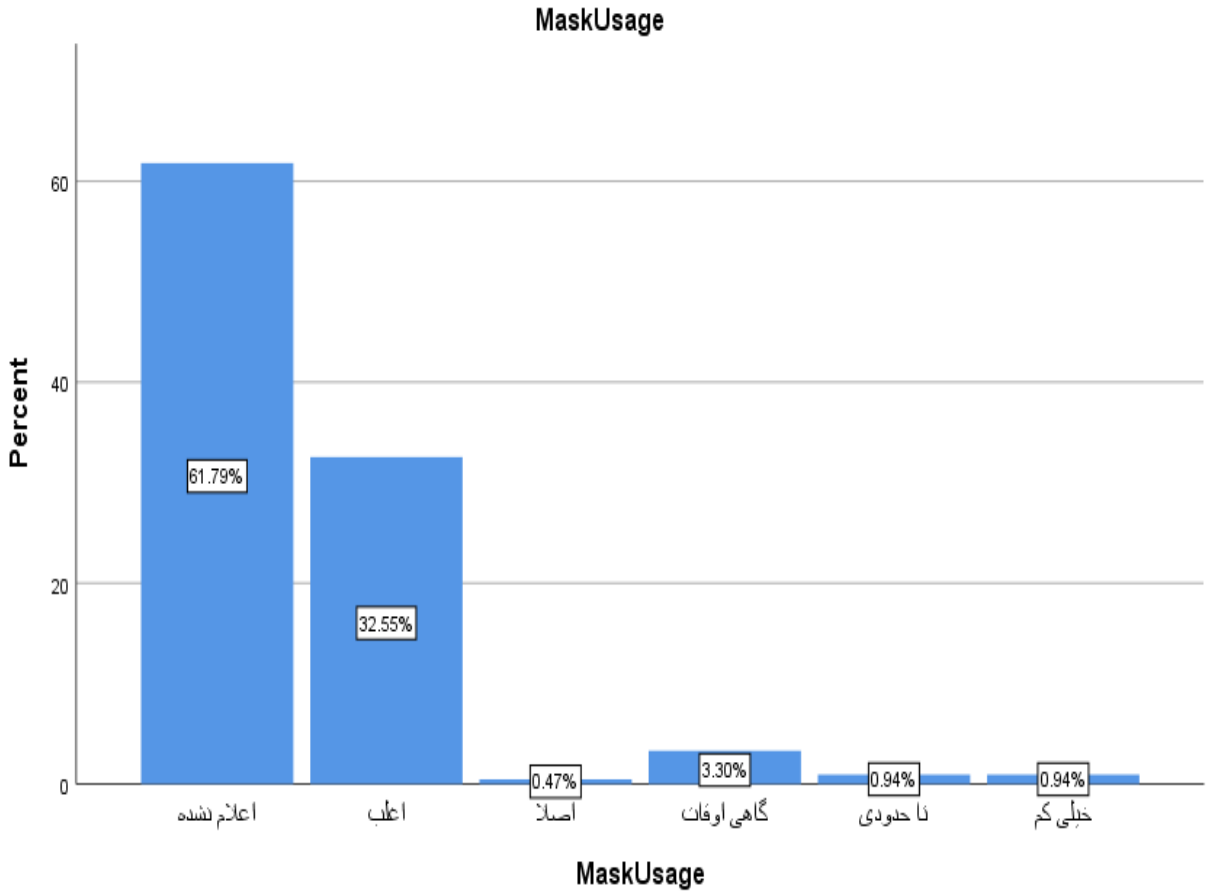


- **توزیع پاسخ دهندگان بر حسب وضعیت استفاده از ماسک:**

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت استفاده از ماسک نشان داده شده است . ملاحظه می گردد که 32.5 درصد پاسخ دهنگان اغلب از ماسک استفاده کرده و 0.5 درصد از پاسخ دهندگان اصلا از ماسک استفاده نمی کنند.

استفاده از ماسک

		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیرها	اعلام نشده	131	61.8	61.8	61.8
	اغلب	69	32.5	32.5	94.3
	اصلا	1	.5	.5	94.8
	گاهی اوقات	7	3.3	3.3	98.1
	تا حدودی	2	.9	.9	99.1
	خیلی کم	2	.9	.9	100.0
	کل	212	100.0	100.0	

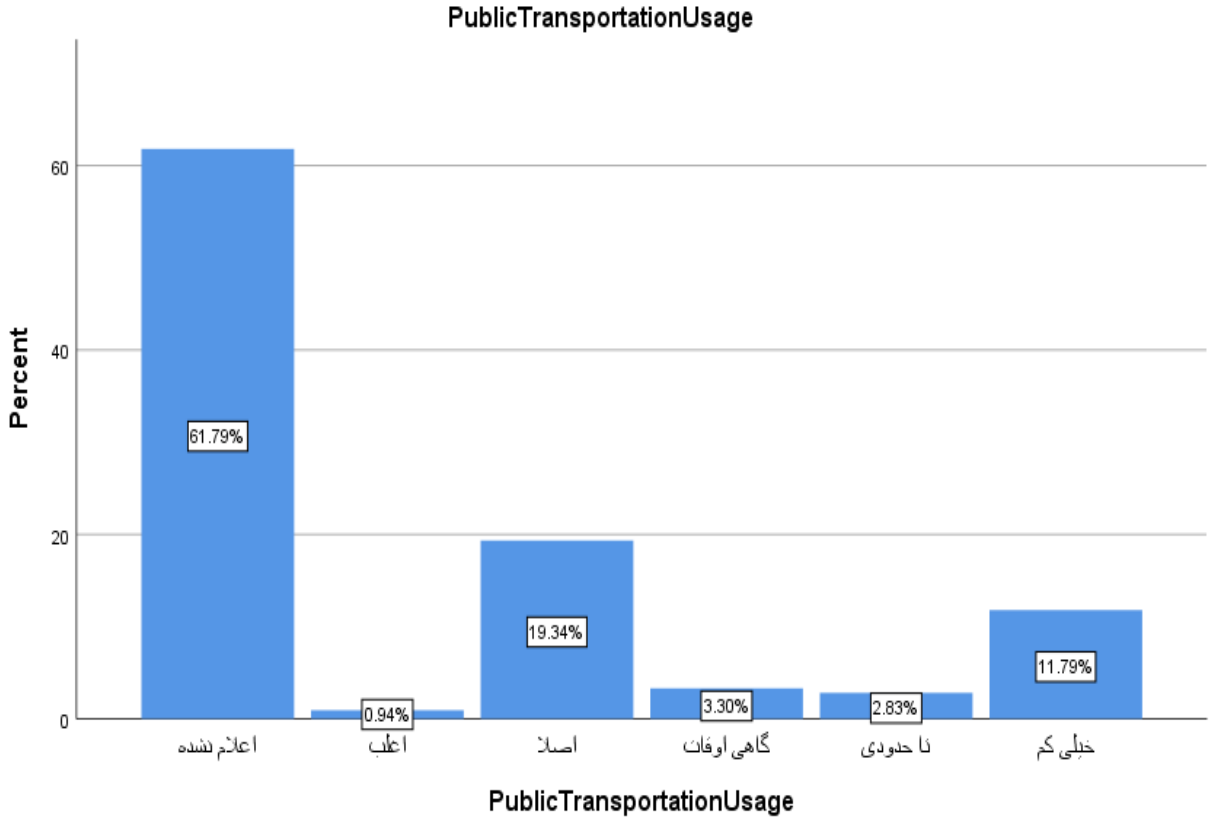


- **توزیع پاسخ دهندگان بر حسب استفاده از حمل و نقل عمومی:**

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت استفاده از حمل و نقل عمومی نشان داده شده است . ملاحظه می گردد که 0.9 درصد پاسخ دهنگان اغلب از حمل و نقل عمومی استفاده کرده و 19.3 درصد از پاسخ دهندگان اصلا از حمل و نقل عمومی استفاده نمی کنند.

استفاده از حمل و نقل عمومی

		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیرها	اعلام نشده	131	61.8	61.8	61.8
	اغلب	2	.9	.9	62.7
	اصلا	41	19.3	19.3	82.1
	گاهی اوقات	7	3.3	3.3	85.4
	تا حدودی	6	2.8	2.8	88.2
	خیلی کم	25	11.8	11.8	100.0
	کل	212	100.0	100.0	

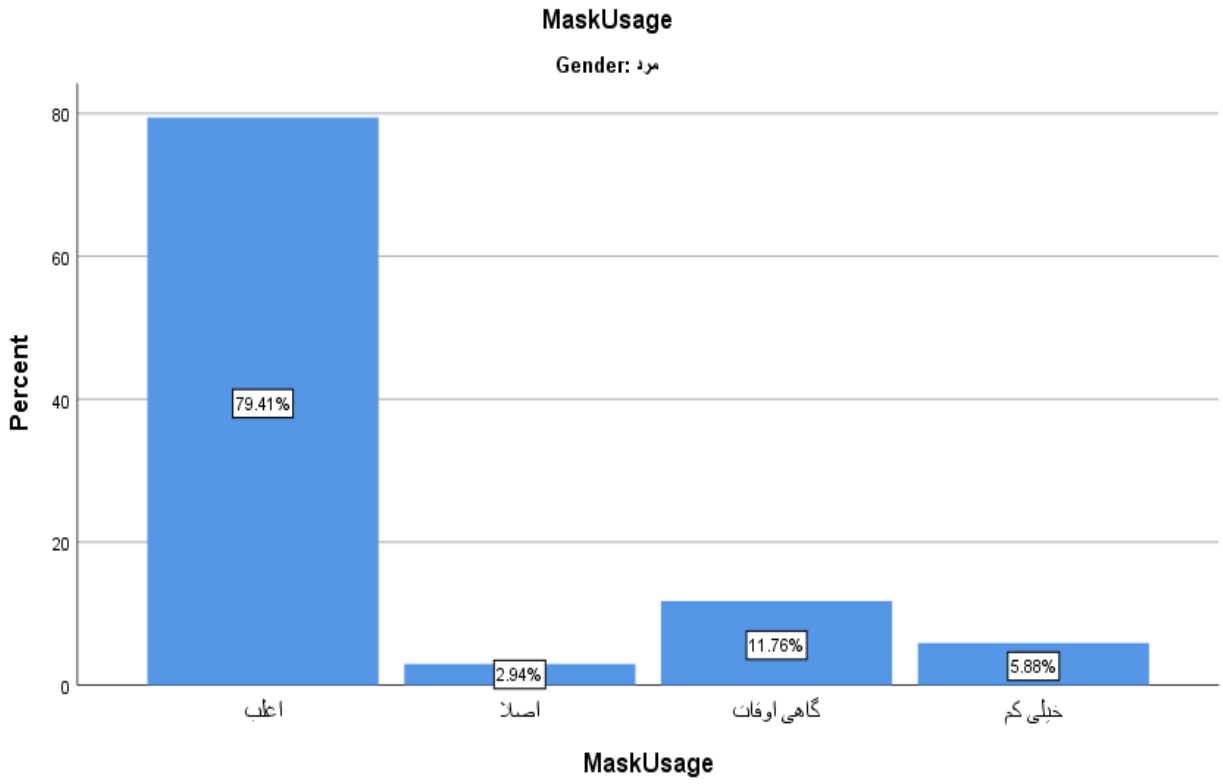


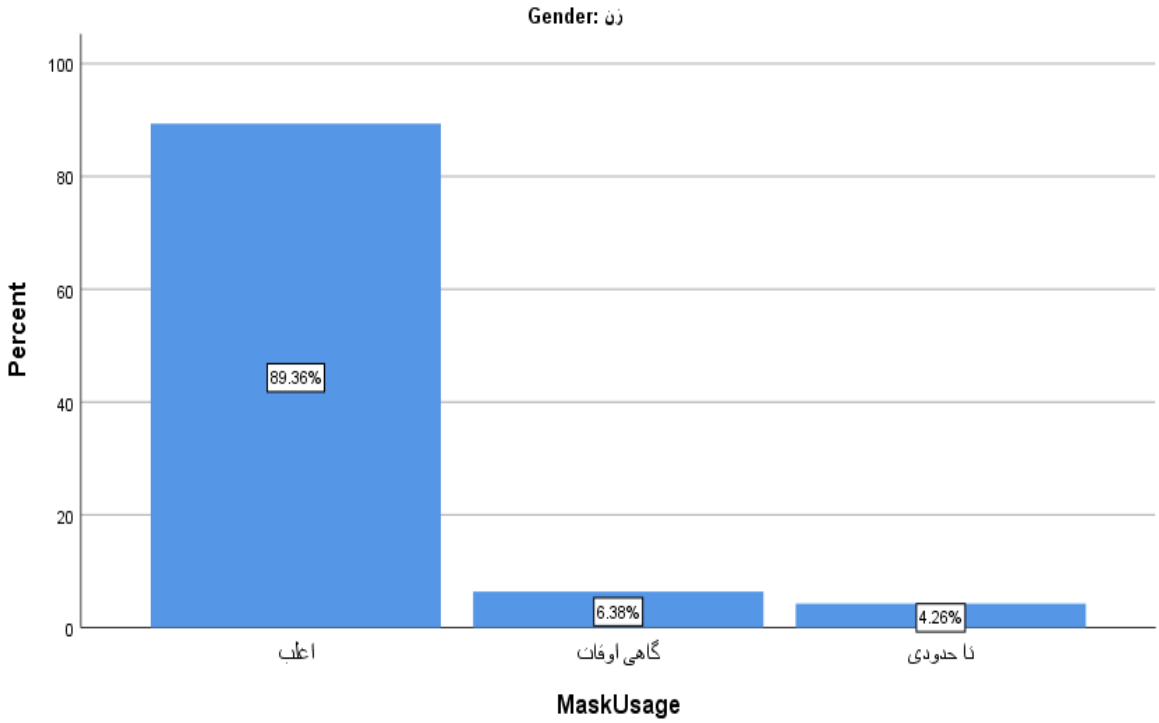
- **توزیع پاسخ دهندگان بر حسب وضعیت استفاده از ماسک به تفکیک جنسیت:**

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت استفاده از ماسک به تفکیک جنسیت نشان داده شده است . ملاحظه می گردد که در بین افرادی که از ماسک استفاده می شود 89.4 درصد از زنان اغلب از ماسک استفاده کرده و 79.4 درصد از مردان اغلب از ماسک استفاده کرده و 2.9 درصد از مردان اصلا از ماسک استفاده نکرده که این متغیر در بین زنان وجود ندارد و این را بازگو می کند که زنان بیشتر التزام به استفاده از ماسک را دارند.

استفاده از ماسک

جنسیت			فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
اعلام نشده	متغیر	اعلام نشده	131	100.0	100.0	100.0
زن	متغیرها	اغلب	42	89.4	89.4	89.4
		گاهی اوقات	3	6.4	6.4	95.7
		تا حدودی	2	4.3	4.3	100.0
		کل	47	100.0	100.0	
مرد	متغیرها	اغلب	27	79.4	79.4	79.4
		اصلا	1	2.9	2.9	82.4
		گاهی اوقات	4	11.8	11.8	94.1
		خیلی کم	2	5.9	5.9	100.0
		کل	34	100.0	100.0	


MaskUsage
Gender: مرد
Percent
80
60
40
20
0
79.41%
2.94%
11.76%
5.88%
اغلب
اصلا
گاهی اوقات
خیلی کم
MaskUsage

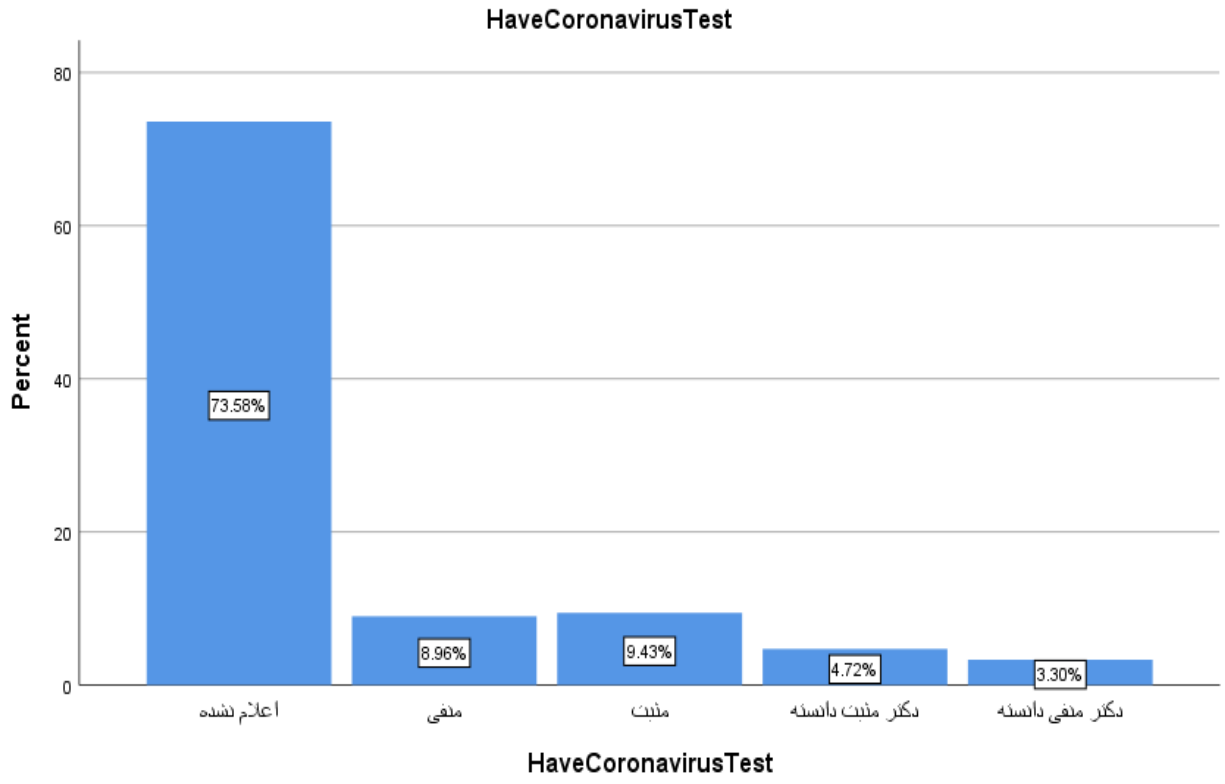

Gender: زن
Percent
100
80
60
40
20
0
89.36%
6.38%
4.26%
اغلب
گاهی اوقات
نا حدودی
MaskUsage

- **توزیع پاسخ دهندگان بر حسب وضعیت کرونایی افراد:**

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت کرونایی افراد و آیا تست انجام گرفته یا خیر و ایا از نظر پزشک بدون انجام تست مبتلا یا سالم باشد نشان داده شده است . ملاحظه می گردد که در بین افرادی که از آن ها تست گرفته شده 9 درصد منفی و 9.4 درصد هم مثبت بوده اند. همچنین در بین این افراد به صورت تجربی از سوی پزشک 4.7 درصد از افراد مثبت بوده اند و 3.3 درصد نیز منفی اعلام شده است.

وضعیت کرونایی افراد از روی تست PCR

		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیر ها	اعلام نشده	156	73.6	73.6	73.6
	منفی	19	9.0	9.0	82.5
	مثبت	20	9.4	9.4	92.0
	دکتر مثبت دانسته	10	4.7	4.7	96.7
	دکتر منفی دانسته	7	3.3	3.3	100.0
	کل	212	100.0	100.0	

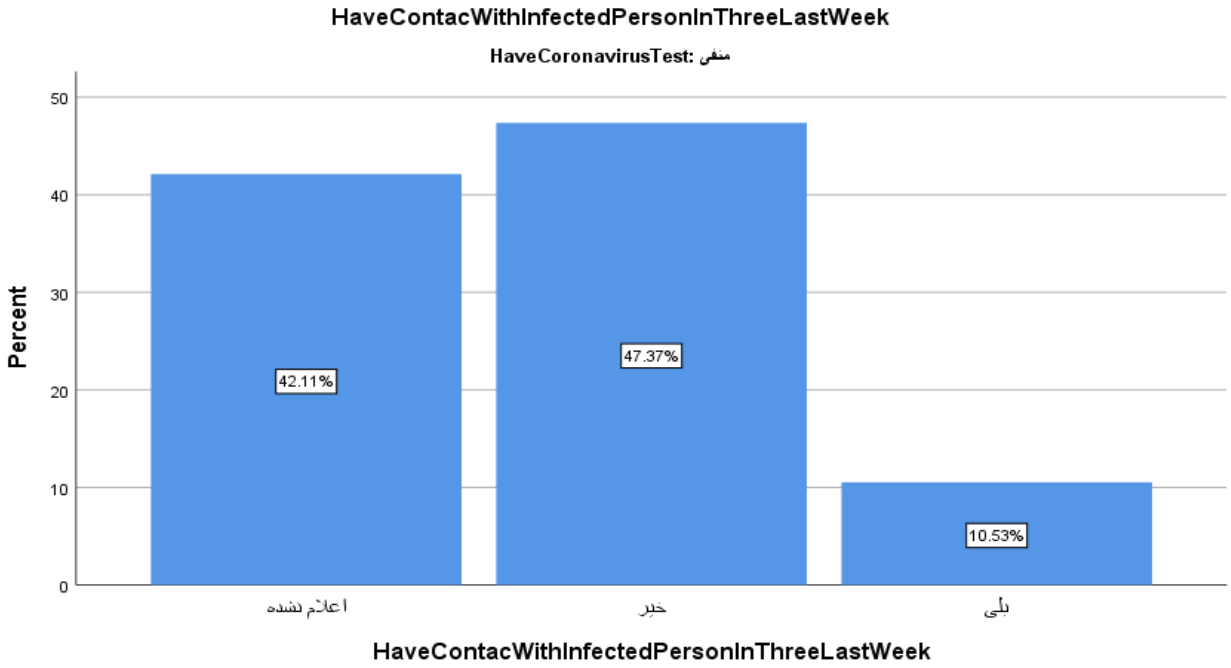


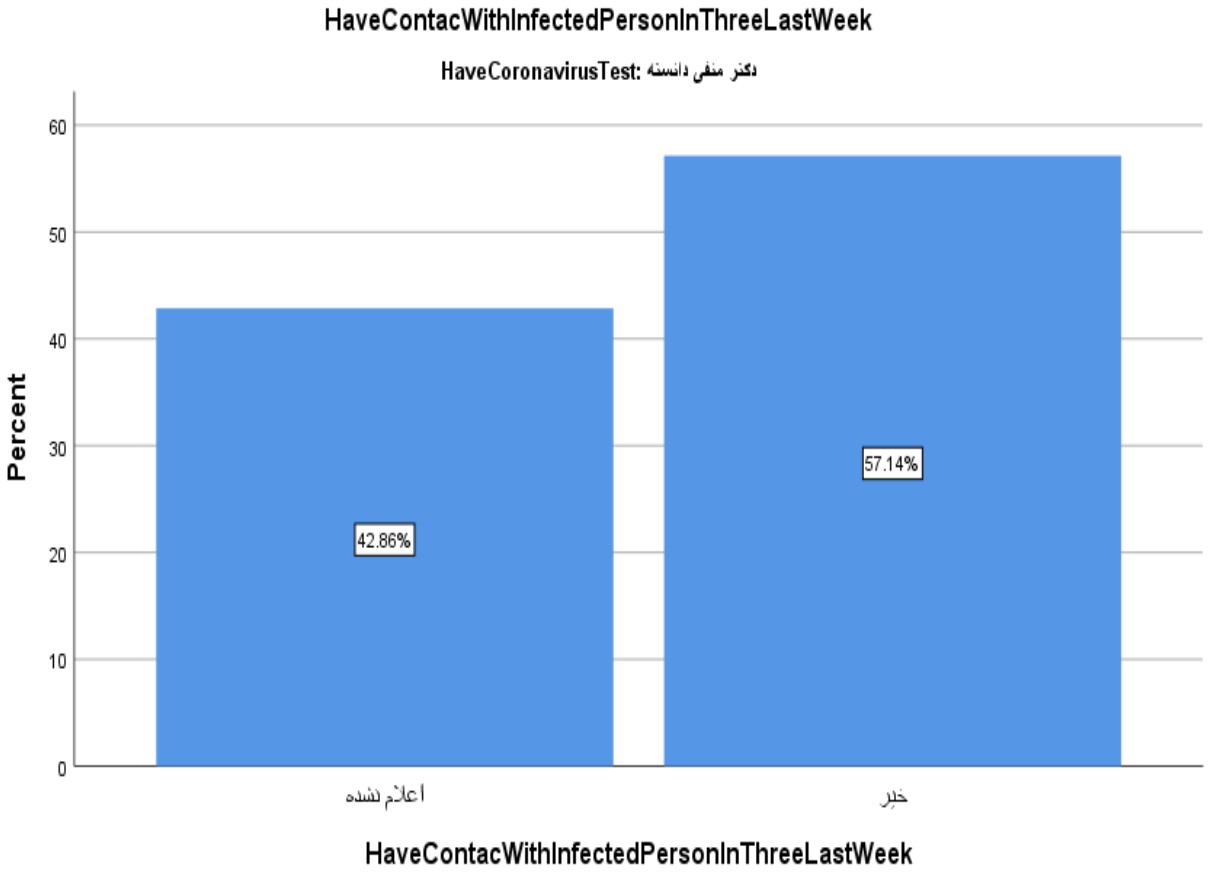
● **توزیع پاسخ دهندگان بر حسب وضعیت کرونایی افراد و آیا با فرد کرونایی در ارتباط بوده یا خیر:**

در جدول زیر توزیع پاسخ دهندگان بر حسب وضعیت کرونایی افراد و آیا با فرد مبتلا به کرونا در یک هفته ی گذشته در تماس بوده اند نشان داده شده است . ملاحظه می گردد که در بین افرادی که تست آنها منفی بوده است 47.4 درصد با فرد کرونایی در ارتباط نبوده اند و 10.5 درصد با فرد کرونایی ارتباط داشته اند. همچنین در بین افرادی که تست آنها مثبت بوده است 30 درصد با فرد کرونایی ارتباط نداشته و 5 درصد از افراد اعلام کرده اند که ارتباط داشته اند با فرد کرونایی، همچنین در بین افرادی که دکتر اعلام کرده که به کرونا مبتلا نیستند 57.1 درصد آنها اعلام کرده بودند که با فرد کرونایی ارتباط نداشته اند و هیچ کسی اعلام نداشته که در ارتباط بوده است.

آیا با فرد مبتلا به کرونا در ارتباط بوده است

وضعیت کرونایی فرد			فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
منفی	متغیرها	اعلام نشده	8	42.1	42.1	42.1
		خیر	9	47.4	47.4	89.5
		بلی	2	10.5	10.5	100.0
		Total	19	100.0	100.0	
مثبت	متغیرها	اعلام نشده	13	65.0	65.0	65.0
		خیر	6	30.0	30.0	95.0
		بلی	1	5.0	5.0	100.0
		Total	20	100.0	100.0	
دکتر مثبت دانسته	متغیرها	اعلام نشده	3	30.0	30.0	30.0
		خیر	3	30.0	30.0	60.0
		بلی	4	40.0	40.0	100.0
		Total	10	100.0	100.0	
دکتر منفی دانسته	متغیرها	اعلام نشده	3	42.9	42.9	42.9
		خیر	4	57.1	57.1	100.0
		Total	7	100.0	100.0	


HaveContacWithInfectedPersonInThreeLastWeek
HaveCoronavirusTest: منفی
Percent
50
40
30
20
10
0
42.11%
47.37%
10.53%
اعلام نشده
خیر
بلی
HaveContacWithInfectedPersonInThreeLastWeek

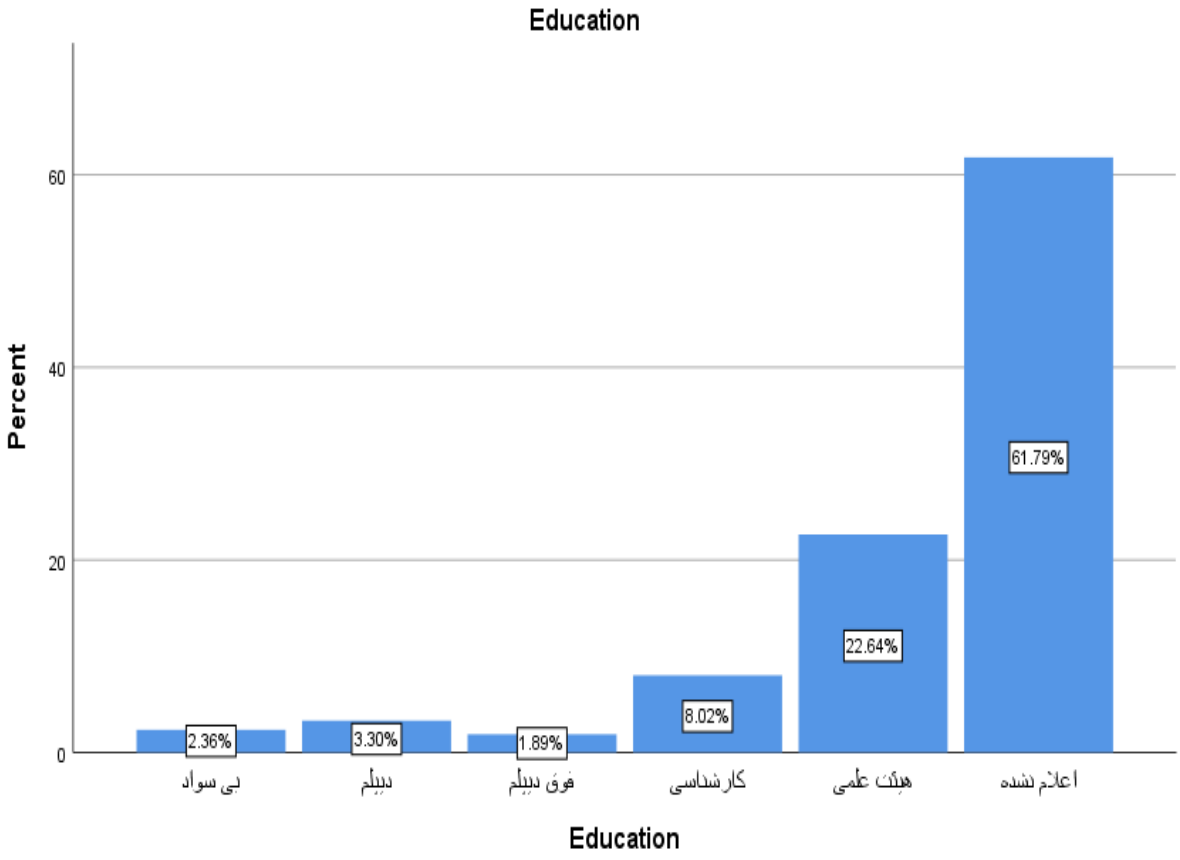

HaveContacWithInfectedPersonInThreeLastWeek
HaveCoronavirusTest: دکتر منفی دانسته
Percent
60
50
40
30
20
10
0
42.86%
57.14%
اعلام نشده
خیر
HaveContacWithInfectedPersonInThreeLastWeek

- **توزیع پاسخ دهندگان بر حسب میزان تحصیلات:**

در جدول زیر توزیع پاسخ دهندگان بر حسب میزان تحصیلات نشان داده شده است . ملاحظه می گردد که 22 درصد از پاسخ دهندگان دکترا و عضو هیئت علمی بوده اند و 2.4 درصد از آن ها بی سواد ثبت کرده اند.

تحصیلات

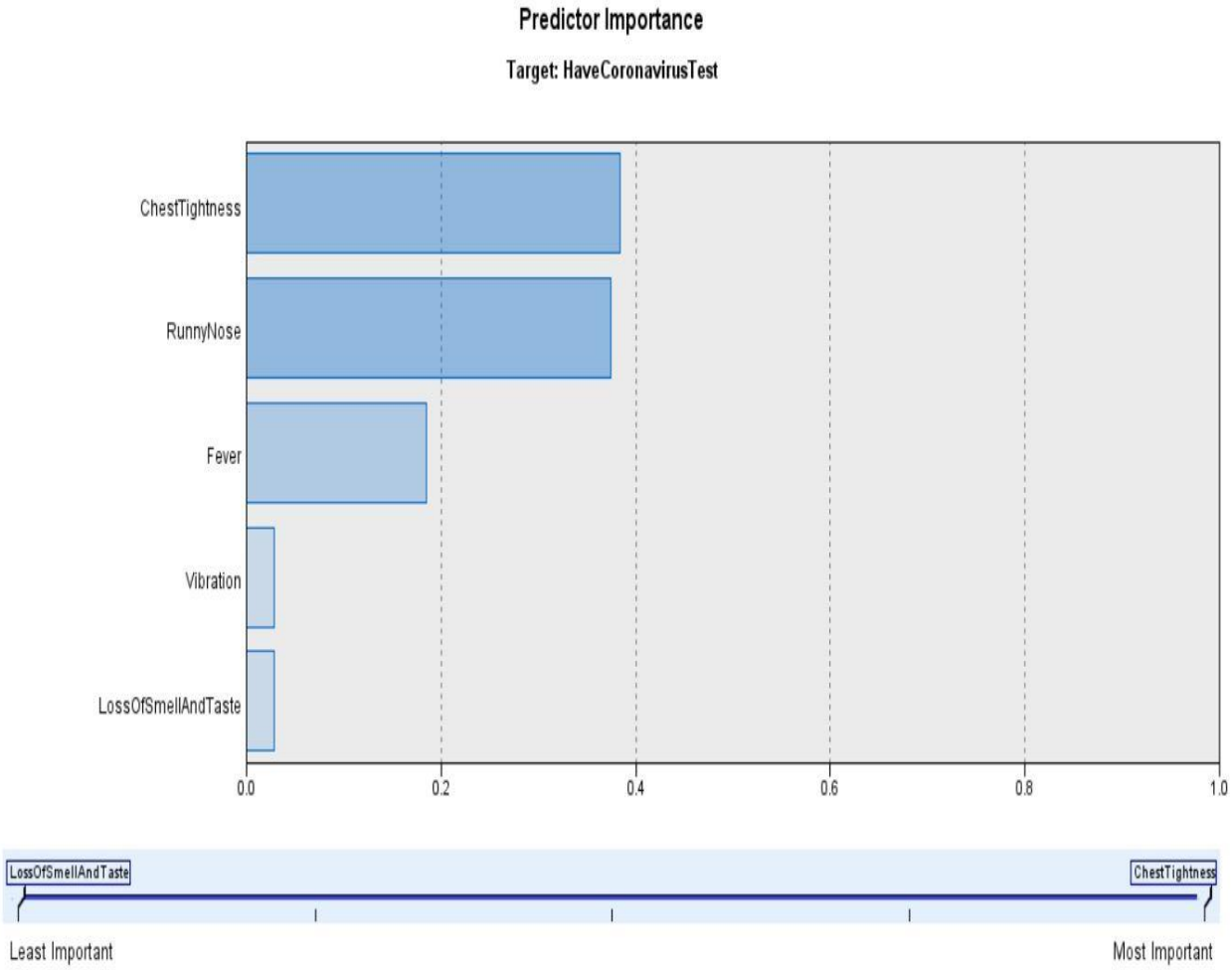
		فراوانی	درصد فراوانی	درصد فراوانی معتبر	درصد فراوانی تجمعی
متغیرها	بی سواد	5	2.4	2.4	2.4
	دیپلم	7	3.3	3.3	5.7
	فوق دیپلم	4	1.9	1.9	7.5
	کارشناسی	17	8.0	8.0	15.6
	هیئت علمی	48	22.6	22.6	38.2
	اعلام نشده	131	61.8	61.8	100.0
	کل	212	100.0	100.0	



- **رایج ترین علائم در بین افراد کرونایی که اعلام کرده اند تست کرونا انجام دادند:**

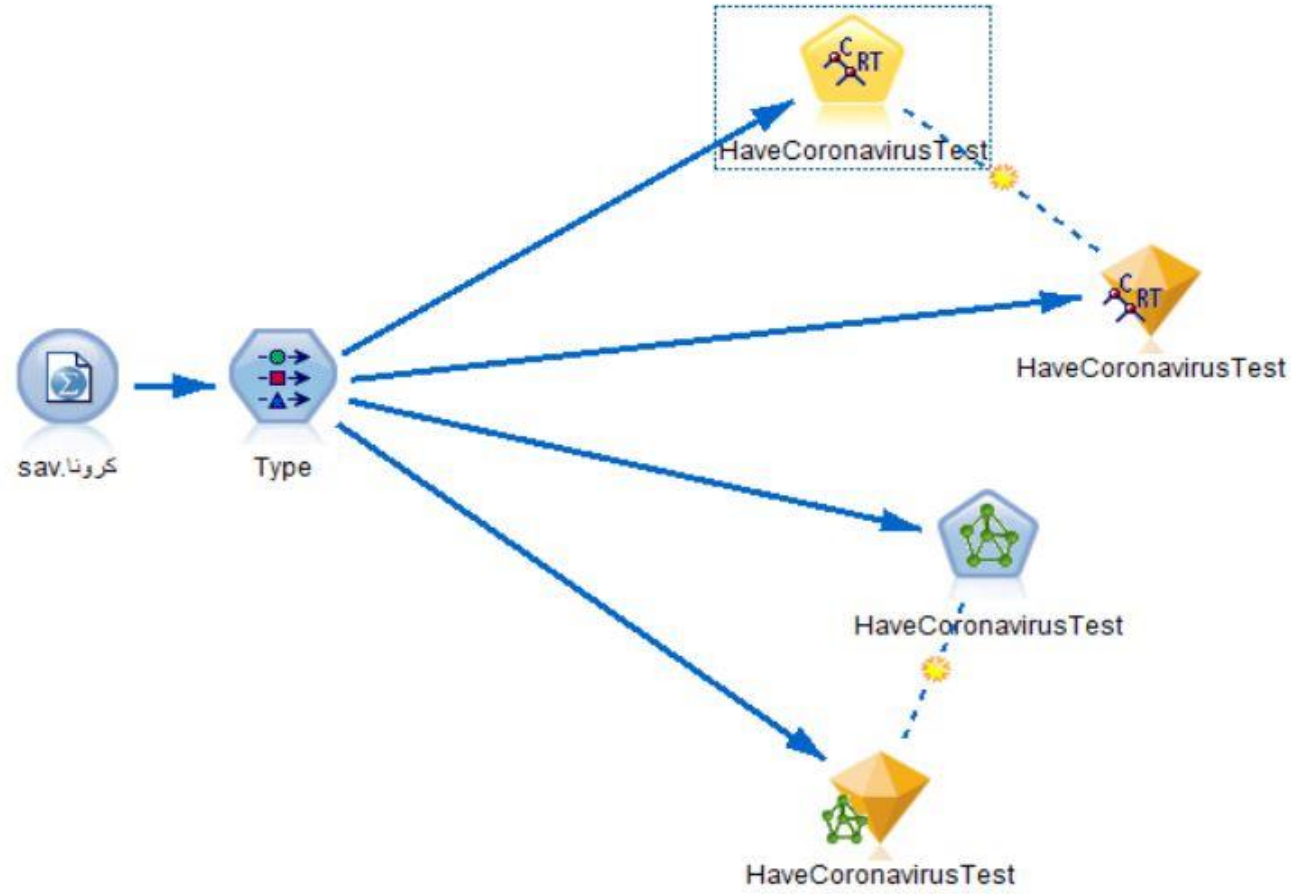
با استفاده از روش های آماری، مشخص گردید که در بین افرادی که تست کرونا انجام داده اند و یا به کرونا مبتلا شده و به پزشک مراجعه داشته اند به ترتیب بیشترین و مهم ترین علائم به صورت زیر بیان شده است:

1- تنگی نفس
2- آبریزش بینی
3- تب
4- لرز
5- از دست دادن حس بویایی و چشایی



Nodes	Importance

Taste And Smell Of Loss	0.0287
Vibration	0.0287
Fever	0.1843
Nose Runny	0.3748
Tightness Chest	0.3835

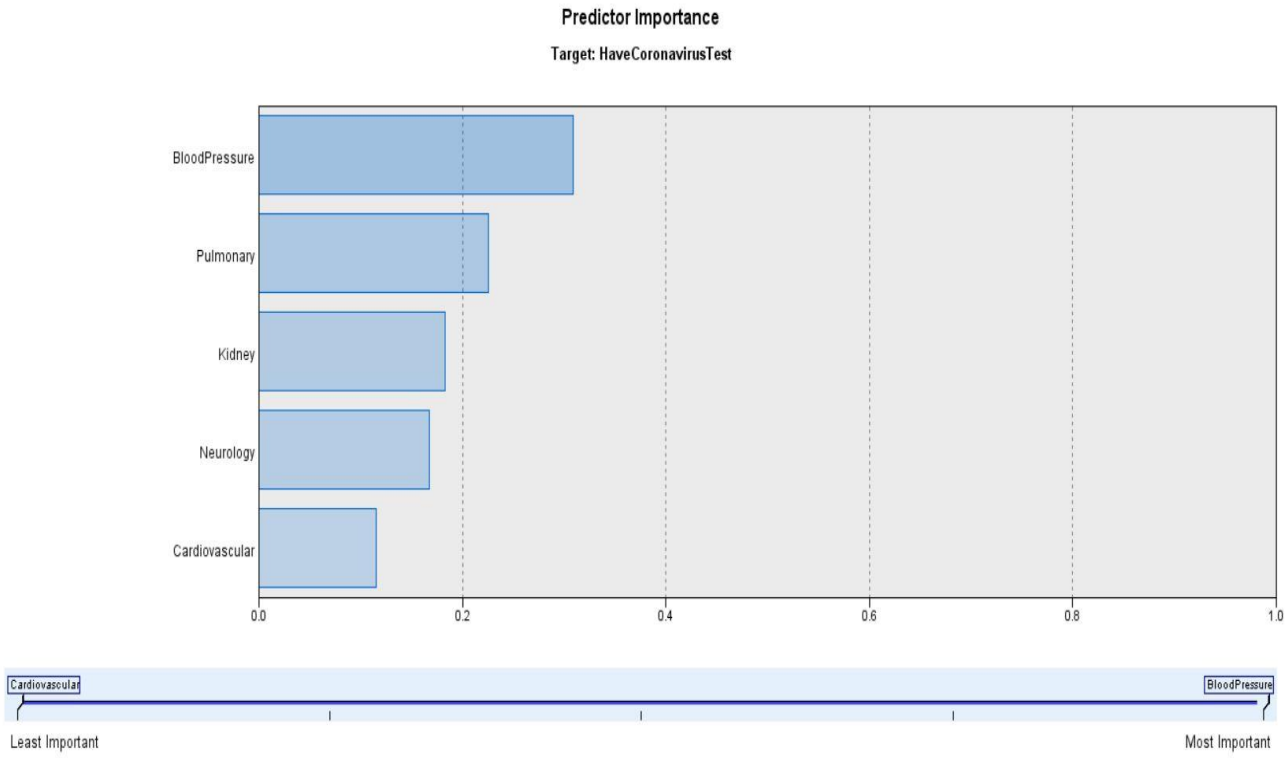


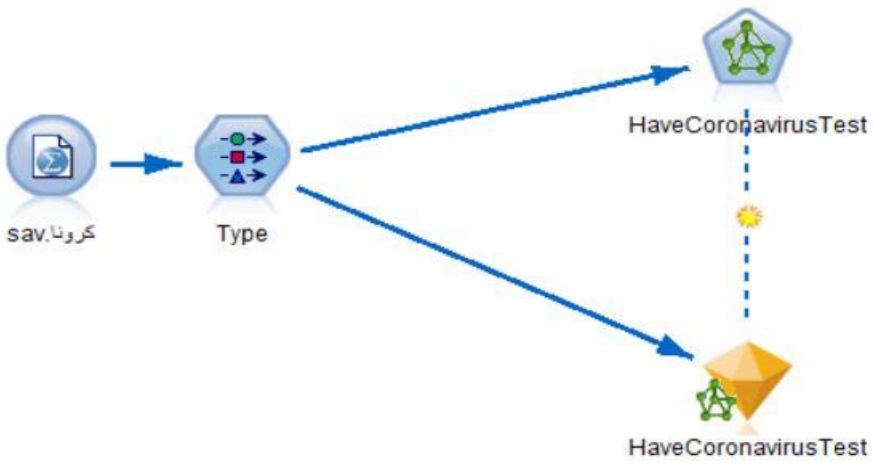
همچنین در جدول فوق مشاهده می شود که تنگی نفس با درصد اهمیت 38.3 بیشترین درصد در بین علائم را دارد و از دست دادن حس بویایی و چشایی با 2.8 درصد کمترین مقدار را در بین علائم به ثبت رسانده است.

- **بیشترین بیماری زمینه ای در بین پاسخ دهندگان:**

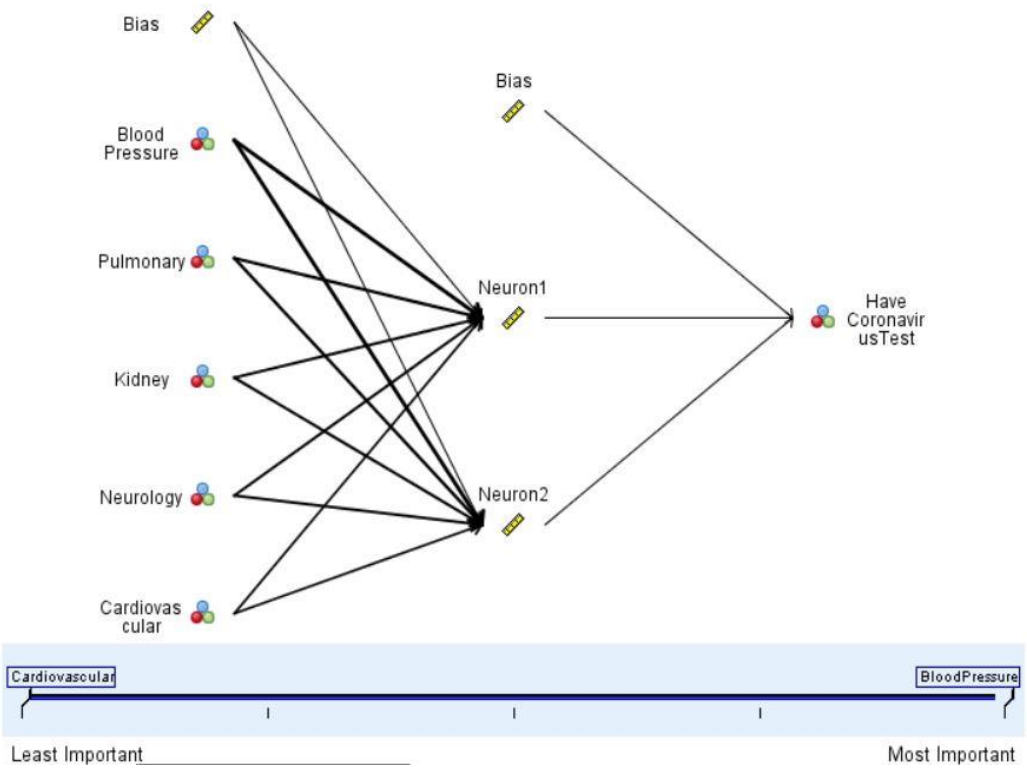
با استفاده از روش شبکه ی عصبی بیشترین بیماری زمینه ای در بین پاسخ دهندگان و همچنین بیشترین بیماری زمینه ای در بین پاسخ دهندگانی که اعلام داشته اند تست کرونا داده و یا به تشخیص پزشک کرونایی سنجیده شده به شرح زیر می باشد:

1- فشار خون
2- ریوی
3- کلیوی
4- اعصاب
5- قلبی-عروقی




HaveCoronavirusTest
sav.كرونا
Type
HaveCoronavirusTest

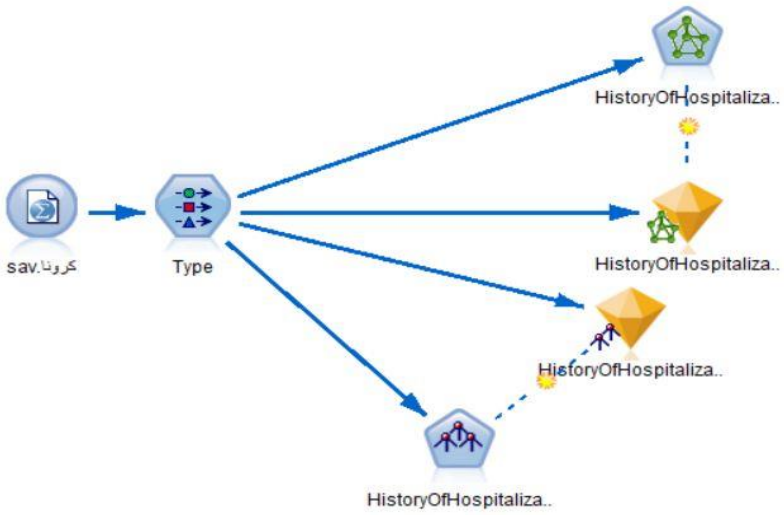
Network

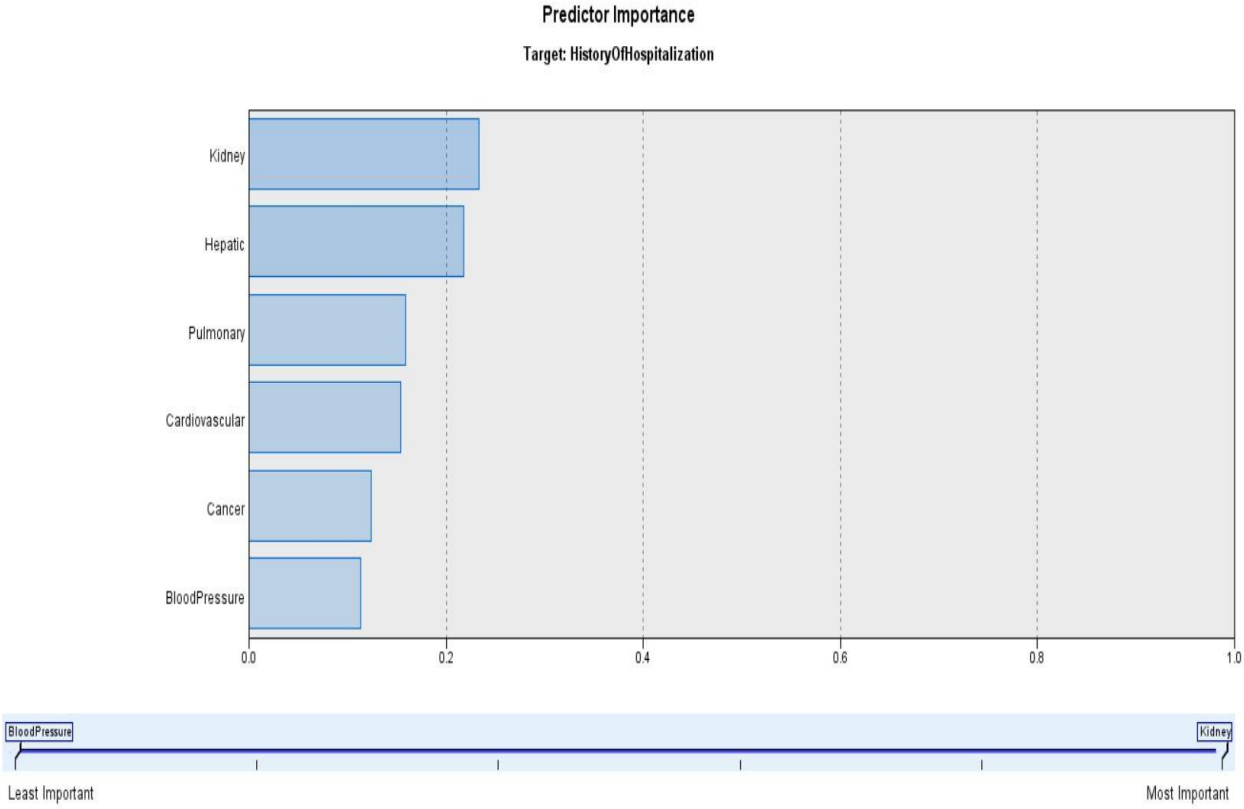

Bias
Blood Pressure
Pulmonary
Kidney
Neurology
Cardiovascular
Bias
Neuron1
Neuron2
Have Coronavirus Test
Cardiovascular
BloodPressure
Least Important
Most Important

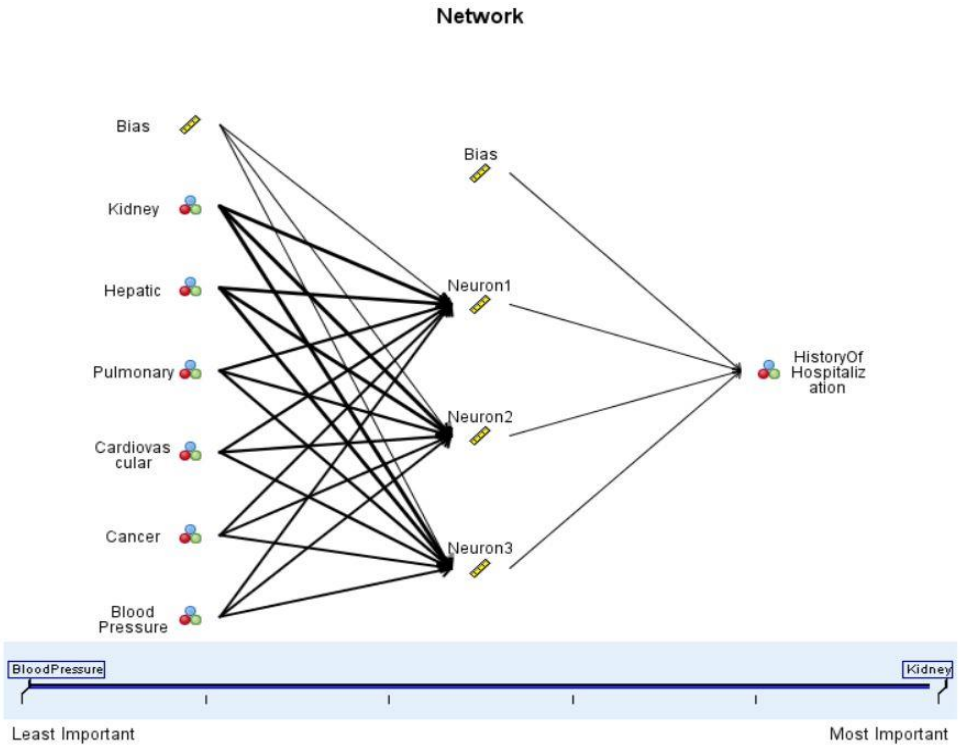
- **بیشترین بیماری زمینه ای که منجر به بستری در بیمارستان بوده است:**

با استفاده از روش های آماری مشخص گردید که کسانی که سابقه ی بستری در بیمارستان به دلیل کرونا را داشته اند بیشتر به بیماری زمینه ای زیر دچار بوده اند:

1-کلیوی
2- کبدی
3- ریوی
4- قلبی-عروقی
5- سرطان
6- فشار خون







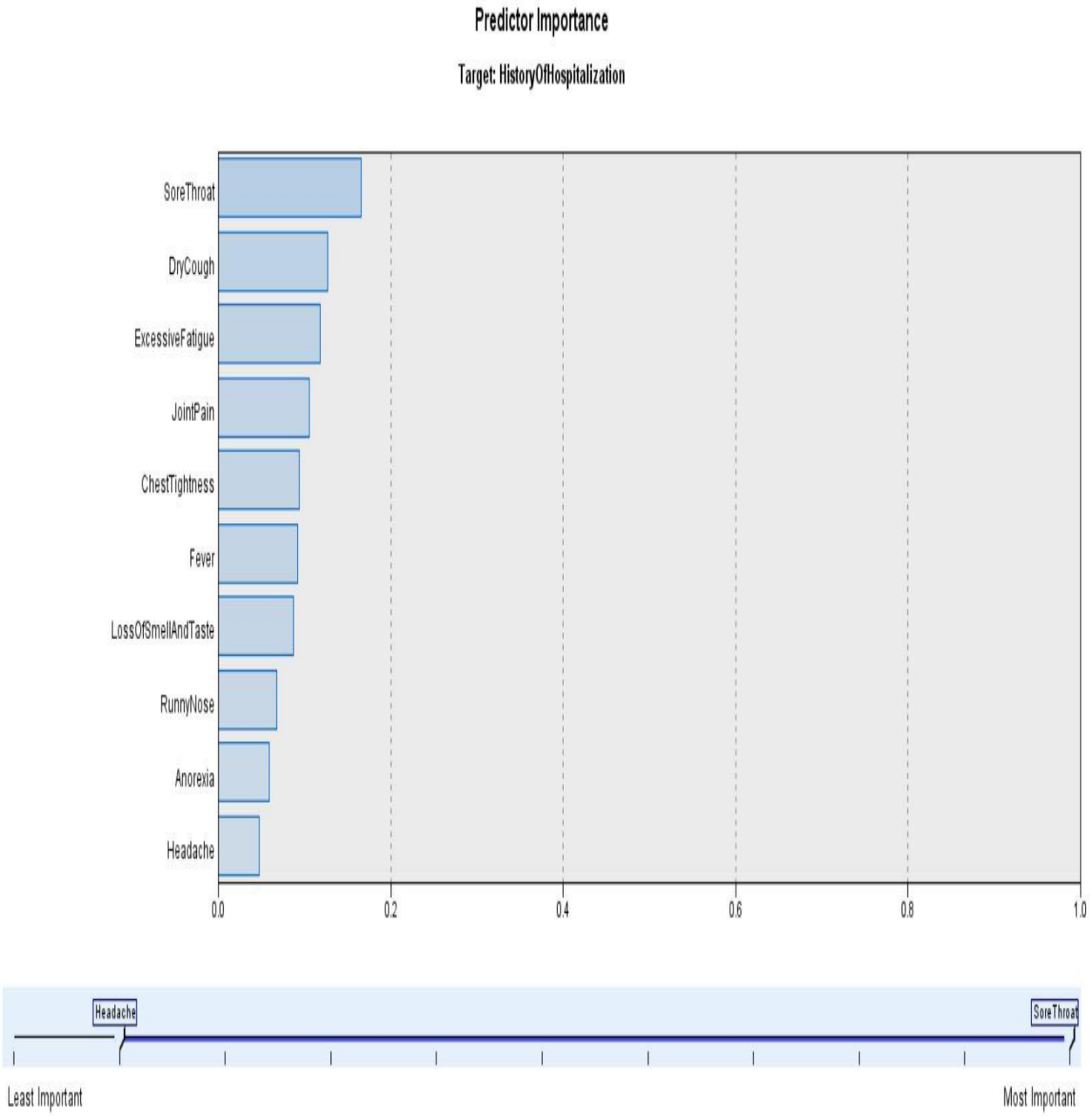
Nodes	Importance
Blood Pressure	0.115
Cancer	0.122
Cardiovascular	0.158
Pulmonary	0.165
Hepatic	0.224
Kidney	0.231

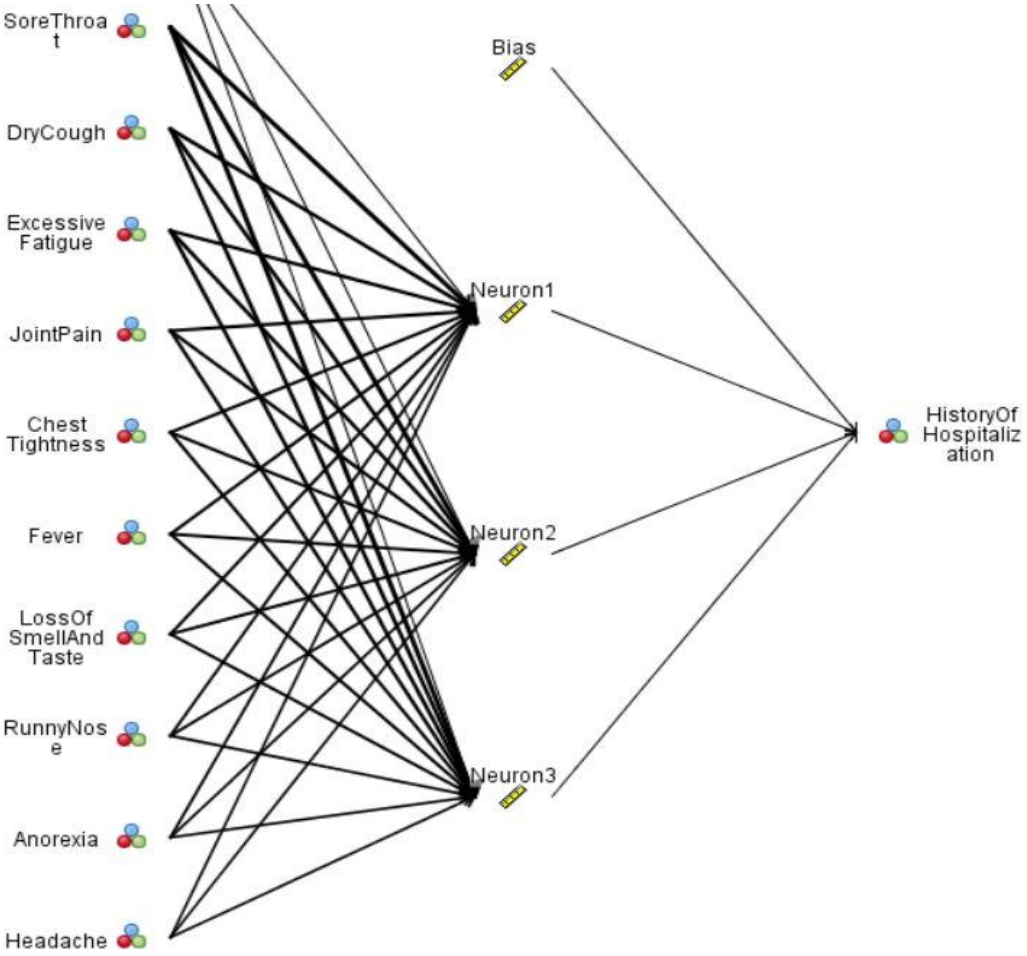
همچنین در جدول فوق مشاهده می شود که بیماری زمینه ای کلیوی با درصد اهمیت 23 بیشترین درصد در بین کسانی که سابقه بستری به علت کرونا را داشته اند دارد و از بیماری زمینه ای فشار خون با 11 درصد کمترین مقدار را در این بین به ثبت رسانده است.

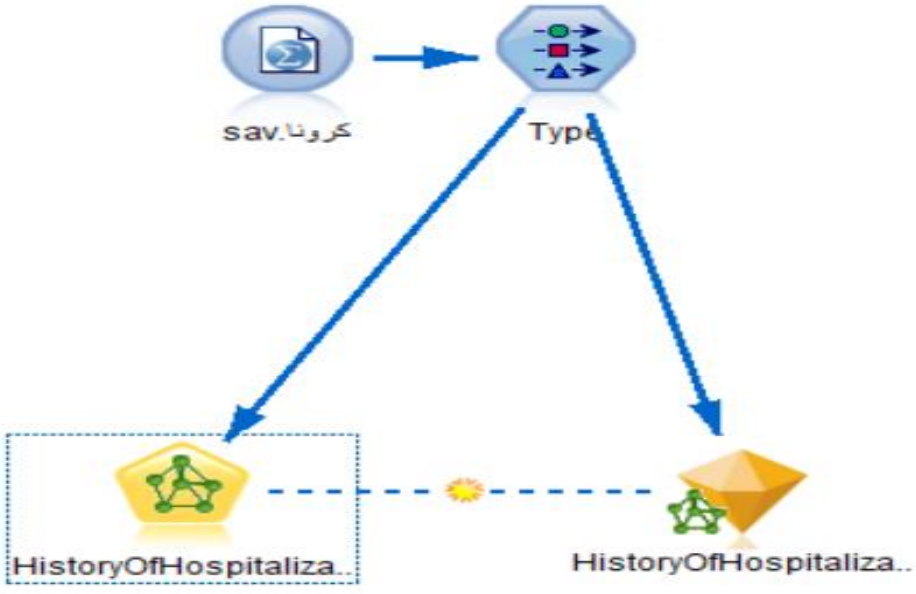
- **بیشترین علائمی که منجر به بستری در بیمارستان بوده است:**

با استفاده از روش شبکه عصبی مشخص گردید که کسانی که سابقه ی بستری در بیمارستان به دلیل کرونا را داشته اند بیشتر به علائم زیر به ترتیب اهمیت دچار بوده اند:

1- گلو درد
2- سرفه خشک
3- خستگی مفرط
4- بدن درد
5- تنگی نفس
6- تب
7- از دست دادن حس بویایی و چشایی
8- آبریزش بینی
9- بی اشتهایی
10- سردرد


Predictor Importance
Target: HistoryOfHospitalization
SoreThroat
DryCough
ExcessiveFatigue
JointPain
ChestTightness
Fever
LossOfSmellAndTaste
RunnyNose
Anorexia
Headache
0.0
0.2
0.4
0.6
0.8
1.0
Headache
SoreThroat
Least Important
Most Important


SoreThroat
DryCough
Excessive Fatigue
JointPain
Chest Tightness
Fever
LossOf SmellAnd Taste
RunnyNose
Anorexia
Headache
Bias
Neuron1
Neuron2
Neuron3
HistoryOf Hospitalization


sav.كرونا
Type
HistoryOfHospitaliza..
HistoryOfHospitaliza..

Nodes	Importance
Headache	0.051
Anorexia	0.060
Nose Runny	0.070
Taste And Smell Of Loss	0.091
Fever	0.092
Tightness Chest	0.095
Pain Joint	0.110
Fatigue Excessive	0.120
Cough Dry	0.130
Throat Sore	0.170

همان گونه که در جدول فوق مشاهده می شود علامت گلو درد با درصد اهمیت 17 بیشترین درصد در بین کسانی که سابقه بستری به علت کرونا را داشته اند را دارد و علامت سردرد با 5 درصد کمترین مقدار را در این بین به ثبت رسانده است. بنابراین حائز اهمیت است که علائمی که باید بیشتر مورد توجه قرار گیرد گلودرد، سرفه خشک،تنگی نفس، تب و خستگی و بدن درد است.

5-4 خلاصه فصل:

در این فصل از پژوهش ابتدا به معرفی کار پژوهش و در ادامه به مدل و الگوریتم مورد استفاده در این فصل پرداخته شد. در ادامه ی این فصل به بحث نظری و روند فعالیت نرم افزار صحبت مفصلی شدو همچنین از نرم افزار های مورد استفاده و معرف یبسته های آن پرداخته شد. و در اخر به تجزیه و تحلیل داده های بدست آمده از نرم افزار پرداختیم ایم.

فصل پنجم :

نتیجه گیری و آینده تحقیق

5-1 مقدمه

در این فصل پس از مرور مختصری بر موضوع تحقیق، ابتدا به معرفی نتایج حاصل از انجام تحقیق پرداخته و سپس پیشنهاداتی را برای تحقیقات آتی ارائه خواهیم کرد .

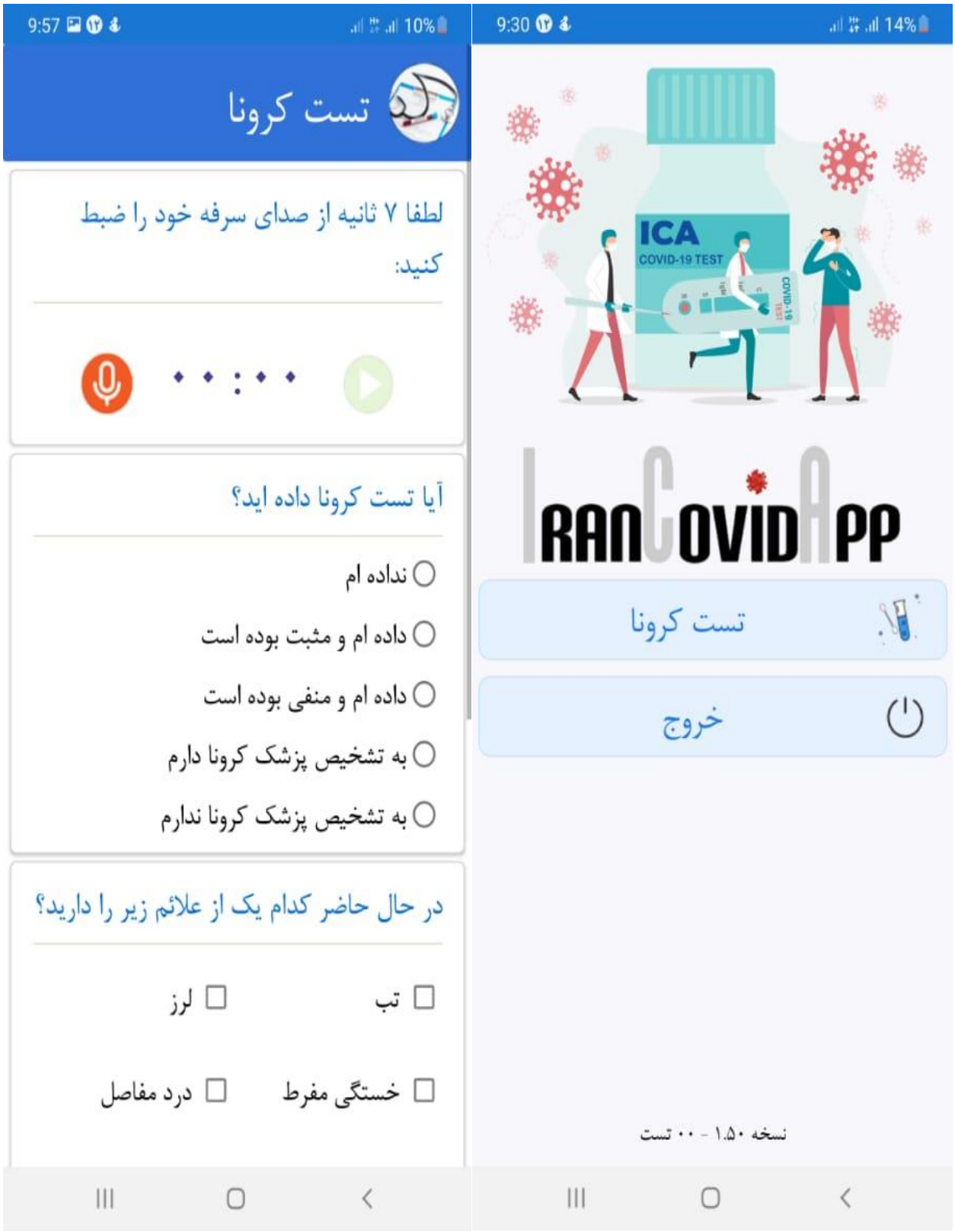
5-2 نتیجه گیری

در اواخر 2019، یک بیماری ویروسی جدید، با نام کووید-19، ناشی از کروناویروس SARS-CoV-2، از چین آغاز و خیلی زود به یک پاندمی تبدیل شد. کووید-19 عامل تهدیدکننده سلامت انسان‌ها و مسبب خسارات جانی و مالی جوامع بشری است .لذا محققین در تلاش برای شناسایی بهتر و کنترل سریعتر آن هستند.

گرچه، گسترش تکنولوژی های پزشکی، موجب کاهش میزان مرگ ومیر ناشی از این بیماری شده است و به دلیل تشخیص زودهنگام و درمان بهینه، میزان بهبودی افزایش یافته است. اما میزان بهبودی بیماران در این نوع بیماری ها بستگی به نوع بیماری و وسعت آسیب وارده به دستگاه تنفسی دارد. تشخیص زودهنگام نیز به یک روش تشخیص صحیح و قابل اعتماد نیاز دارد که به پزشکان اجازه میدهد که عفونت های مختلف را از هم شناسایی کنند. بنابراین یافتن یک روش تشخیص صحیح و موثر در بروز این بیماری، بسیار با اهمیت است.

روش های یادگیری ماشین علاوه بر اینکه میتوانند در تشخیص سریع این بیماری ها موثر باشند، میتوانند از طریق شناسایی عوامل موثر در بروز انواع بیماری های تنفسی، به پیشگیری از این بیماری ها کمک کنند و به این ترتیب باعث ارتقای سلامت جامعه شده و از تحمیل هزینه های سنگین ناشی از تشخیص نادرست، بر بیماران جلوگیری کنند. این در حالی است که امروزه استفاده از سیستم های طبقه بندی، میتوانند به کاهش خطایی که ممکن است توسط کارشناسان کم تجربه اتفاق بیفتد، کمک کنند و همچنین امکان بررسی داده های پزشکی را در زمان کوتاه تر و با جزئیات بیشتر فراهم می کنند.

روش های یادگیری ماشین میتوانند دانش نهفته در مجموعه داده را استخراج و در پیشگیری، تشخیص و معالجه ی این بیماری به پزشک و بیمار کمک کنند .مهمترین اصلی که در ارتباط با یادگیری ماشین وجود دارد، این است که هیچ روشی برای تمام داده ها مناسب نخواهد بود. به عبارت دیگر هیچگاه نمیتوان ادعا کرد که یک الگوریتم یا روش برای تمام داده ها مناسب بوده و نتیجه ی مطلوبی ارائه میدهد. بلکه نتایج حاصل از روش های یادگیری ماشین، وابستگی بسیاری به مجموعه داده ی مورد استفاده دارند اما تلاش برآن است که مدلی استفاده شود که با اکثریت سازگار باشد و عملکرد خوبی ارائه دهد . در این تحقیق ابتدا در صدد برآمدیم تا یک مدل یادگیری ماشین را روی مجموعه داده ی جمع آوری شده پیاده کنیم. براساس نتایج به دست آمده، در ابتدا روش ارائه شده از مشکل دقت پائین رنج میبرد که دلیل اصلی این امر نیز نامتوازن بودن داده های مورد استفاده ی جمع آوری شده در این تحقیق بود. به همین جهت، به بررسی روش های ارائه شده برای مسائل نامتوازن پرداخته شد. از میان روش های مختلف برای یادگیری در داده های نامتوازن، در این تحقیق از روش نمونه برداری استفاده شد و مشکل نامتوازن بودن داده ها به خوبی حل شد. در این پژوهش روش با دقت قابل توجه ای از یادگیری ماشین برای تحلیل داده های جمع آوری شده ی سرفه ی بیماران مبتلا به کووید 19 ارائه گردید همچنین یک بانک داده ای از علائم پزشکی جمع آوری شده است و در انتها در این پژوهش یک اپلیکیشن شناسایی بیماری کووید 19 با استفاده از سرفه ی افراد نیز طراحی گردید که میتواند به جامعه پزشکی و همچنین مردم استان در تشخیص به مرور زمان کمک شایانی کند. قابل ذکر است این اپلیکیشن اولین نرم افزار تحت اندروید است که در ایران طراحی گردیده و سومین اپلیکیشن طراحی شده بعد از دانشگاه ام آی تی و کمبریج است که با درصد قابل توجه ای کرونا را از صدای سرفه ی افراد تشخیص می دهد.


9:30
14%
ICA
COVID-19 TEST
IRANCOVIDAPP
تست کرونا
خروج
نسخه ۱.۵۰ - ۰۰ تست
9:57
10%
تست کرونا
لطفا ۷ ثانیه از صدای سرفه خود را ضبط کنید:
آیا تست کرونا داده اید؟
نداده ام
داده ام و مثبت بوده است
داده ام و منفی بوده است
به تشخیص پزشک کرونا دارم
به تشخیص پزشک کرونا ندارم
در حال حاضر کدام یک از علائم زیر را دارید؟
تب
لرز
خستگی مفرط
درد مفاصل

شکل5-1 اپلیکیشن طراحی شده پیشنهادی

5-3 پیشنهادات

در این بخش، پیشنهادها در قالب چندین جنبه مختلف ارائه شده است. جنبه اول در ارتباط با مجموعه داده ی مورد استفاده در تحقیق می باشد و جنبه دوم، در مورد عملیات داده کاوی و استفاده از روش های یادگیری ماشین و یادگیری عمیق مختلف که میتوان روی مجموعه ی داده انجام داد و همچنین پیش بینی تعداد مبتلایان به صورت روزانه و در آخر ارائه نسخه های مختلف اپلیکیشن است.

5-3-1 مجموعه داده

پیشنهاد می شود در تحقیقات بعدی از مجموعه داده ای استفاده شود که در آن، مقادیر ویژگی های داده، یکنواختی کمتری داشته و دارای تنوع بیشتری باشند همچنین در ارتباط با توزیع داده ها، چنانچه توزیع داده ها متوازن باشد، نتایج بسیار بهتری حاصل خواهد شد. چنانچه تعداد بیماری کووید-19 در مجموعه ی داده بیشتر باشد، قطعا نتایج بهتری در تشخیص و طبقه بندی به دست خواهد آمد.

5-3-2 یادگیری ماشین

به منظور درک بهتر ویژگی های ارائه شده روی مجموعه داده ی بیماری و ارائه مدلی با عملکرد بهتر پیشنهاد می شود که علاوه بر مدل ارائه شده، از مدل های موجود به صورت ترکیب با استفاده از ویژگی های مختلف قابل استخراج برای برگزیدن روشی با سرعت و دقت بالاتر استفاده شود.

5-3-3 پردازش تصاویر قفسه سینه

در افراد گاهی ممکن است ناهنجاری‌هایی در مراحل ابتدایی مبتلا شدن به ویروس کووید-19 وجود داشته باشد که با مشاهده پزشک قابل دیدن نباشند اما این ناهنجاری ها با استفاده از مدل های یادگیری عمیق به خوبی قابل تشخیص است و علاوه بر تشخیص زودهنگام باعث می شود در برخی هزینه ها و زمان صرفه جویی شود و فرد مبتلا باعث انتقال ویروس به سایر افراد به صورت ناخواسته نشود.

مدل ارائه شده فعلی براساس پردازش صدای سرفه و سایر علائم این بیماری مانند تب، لرز، بدن درد و غیره کار می کند، انتظار می رود که استفاده از تصاویر قفسه سینه افراد باعث می شود یک سیستم پیش بینی زودهنگام کووید-19 و با قابلیت اطمینان بالاتر ایجاد شود که علاوه بر تشخیص می توان اولین مجموعه داده ی کامل کووید-19 ایران را ایجاد کرد تا دیگر پژوهشگران نیز بتوانند از آن به عنوان یک مجموعه داده داخلی و با کیفیت در کارهای تحقیقاتی خود استفاده کنند. البته قابل ذکر است که اعمال این روش پیشنهادی نیازمند همکاری افراد در استفاده از اپلیکیشن و همچنین همکاری علوم پزشکی در صورت امکان، برای تهیه داده ها است.

5-3-4 پیش بینی تعداد مبتلایان روزانه

همانگونه که از آمار چند روز اخیر نیز پیداست، تعداد مبتلایان کرونا یا اُمیکرون روزانه به صورت درصدی در حال افزایش یا کاهش است و گاهی این ارقام و درصدها برای افراد جامعه استرس آور یا برعکس خوشایند و درنتیجه منجر به افکار اشتباه در مورد پایان کامل کووید-19 و رعایت نکردن نکات ایمنی می شود؛ بنابراین به نظر می رسد ارائه سیستم یا مدلی که منجر به پیش بینی تعداد مبتلایان به ویروس به صورت روزانه شود می تواند

باعث شود که افراد اگاهی بیشتری نسبت به شرایط منطقه و کشور داشته باشند و باعث کاهش برخی مشکلات شود ، وجود چنین سیستمی نیازمند تجزیه و تحلیل داده های ماه های اخیر می باشد که باید در درسترس قرار گیرد تا بتوان خروجی مناسبی وصحیح داشته باشد.

5-3-5 اپلیکیشن سازگار با iOS

به دلیل اینکه امروزه افراد زیادی علاوه بر سیستم عامل اندروید از iOS هم استفاده می کنند و این اپلیکیشن در حال حاضر با اندروید سازگاری دارد؛ تلاش بر این است که در آینده نسخه مختص iOS نیز ارائه شود.

5-3-6 مراقبت از کادر درمان در برابر عفونت های بیمارستانی و کرونا

می توان با استفاده از روش های یادگیری ماشین و هوش مصنوعی به کمک کادر درمان آمده و از مبتلا شدن به کرونا و عفونت های بیمارستانی کادر جلوگیری کرد. به عنوان مثال می توان با استفاده از روش های یادگیری ماشین به پیش بینی عفونت های دیگر بیمار در بدو ورود به بیمارستان رسید و یا اینکه در روز ها و هفته های بعدی با توجه به علائم بیماران احتمال بدحالی و انتقال به بخش مراقبت های ویژه را پیش بینی کرد.

مراجع

[1] https://ihsmarkit.com/research-analysis/nevermore-ignore-the-role-of-digital-health.html

[2] https://militarymedj.ir/article-1-2650-fa.pdf

[3] https://time.com/5805953/home-covid-19-test-everlywell/
https://insights.digitalmediasolutions.com/articles/rise-at-home-medical-testing

[4] https://medicalfuturist.com/

[5] https://www.kaggle.com/c/csc532

[6] https://ddi.sutd.edu.sg/

[7] 5. Alvaro V. A Sales Forecasting System Based on Fuzzy Neural Model for Time Series Forecasting, IEEE Transaction on Neural Net Work 2000; 11: 1402-1412.

[8] https://www.forbes.com/sites/bernardmarr/2020/09/21/the-4-top-artificial-intelligence-trends-for-2021/?sh=76fca90e1c2a

[9] https://bluedot.global/

[10] Armstrong,js Collopy f. Error Measures for Generalizing about Forecasting Methods: Empirical Comparation, International Journal of Forecasting 2007; 8: 69-84.

[11] Pham Q, Nguyen DC, Huynh-The T, Hwang W, Pathirana PN. Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts. IEEE Access. 2020;

8:130820–39.

[12] Allam Z, Dey G, Jones DS. Artificial intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future Urban health policy internationally. AI. 2020;1(2):156–65.

[13] Kavadi DP, Patan R, Ramachandran M, Gandomi AH. Partial derivative Nonlinear Global Pandemic Machine Learning prediction of COVID 19. Chaos, Solitons & Fractals [Internet]. 2020; 139:110056. Available from: http://www.sciencedirect.com/science/article/pii/S0960077920304537.

[14] Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, et al. Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19. IEEE Rev Biomed Eng. 2020;1.

[15] Shaw R, Kim Y, Hua J. Governance, technology and citizen behavior in pandemic: Lessons from COVID-19 in East Asia. Prog disaster Sci. 2020;100090.

[16] Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos, Solitons & Fractals [Internet]. 2020; 139:110059. Available from: http://www.sciencedirect.com/science/article/pii/S0960077920304562.

[17] Allam Z, Jones DS. On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management. In: Healthcare. Multidisciplinary Digital Publishing Institute; 2020. p. 46.

[18] https://medium.com/sciforce/top-ai-algorithms-for-healthcare-aa5007ffa330

[19] https://www.sbu.ac.ir/web/news/-/%D8%AA%D8%B4%D8%AE%DB%8C%D8%B5-%DA%A9%D8%B1%D9%88%D9%86%D8%A7-%D8%A7%D8%B2-%D8%B7%D8%B1%DB%8C%D9%82-%D8%B3%D8%B1%D9%81%D9%87-%D8%A8%D9%87-%DA%A9%D9%85%DA%A9-%D8%B1%D9%88%D8%B4-%D9%87%D8%A7%DB%8C-%D9%87%D9%88%D8%B4-%D9%85%D8%B5%D9%86%D9%88%D8%B9%DB%8C

[20] "Coronavirus History: Origin and Evolution." WebMD, https://www.webmd.com/lung/coronavirus-history. Accessed 24 May 2021.

Ohannessian, Robin, et al. “Global Telemedicine Implementation and Integration Within [21] Health Systems to Fight the COVID-19 Pandemic: A Call to Action.” JMIR Public Health and Surveillance, vol. 6, no. 2, Apr. 2020, p. e18810. DOI.org (Crossref), doi:10.2196/18810.

[22] Buoro, Sabrina, et al. “Papa Giovanni XXIII Bergamo Hospital at the Time of the COVID-19 Outbreak: Letter from the Warfront….” International Journal of Laboratory Hematology, vol. 42 Suppl 1, June 2020, pp. 8–10. PubMed, doi:10.1111/ijlh.13207.

[23] Morens, David M., and Anthony S. Fauci. “Emerging Pandemic Diseases: How We Got to COVID-19.” Cell, vol. 182, no. 5, Sept. 2020, pp. 1077–92. DOI.org (Crossref), doi: 10.1016/j.cell.2020.08.021.

[24] He, X., Lau, E.H.Y., Wu, P. et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 26, 672–675 (2020). https://doi.org/10.1038/s41591-020-0869-5

[25] https://www.kmtmed.com/Fa/blog/page/8270/تست-های-تشخیصی-ویروس-کرونا

[26] رجمندنیا، روح اله و عاطفی راد، مازیار،1399،کاربرد هوش مصنوعی در تشخیص ویروس کرونا COVID19،دومین کنفرانس علمی پژوهشی مکانیک، برق، کامپیوتر و علوم مهندسی موناکو،، https://civilica.com/doc/1133783

[27] https://www.mehrnews.com/news/5210678/شناسایی-۱۶۴۰۹-بیمار-جدید-کرونایی-۳۶۶-نفر-دیگر-فوت-شدند

[28] Rahimzadeh, Mohammad, et al. “A Fully Automated Deep Learning-Based Network for Detecting COVID-19 from a New and Large Lung CT Scan Dataset.” Biomedical Signal Processing and Control, vol. 68, July 2021, p. 102588. DOI.org (Crossref), doi: 10.1016/j.bspc.2021.102588.

[29] Chowdhury, Muhammad E. H., et al. “Can AI Help in Screening Viral and COVID-19 Pneumonia?” IEEE Access, vol. 8, 2020, pp. 132665–76. DOI.org (Crossref), doi:10.1109/ACCESS.2020.3010287.

[30] Peng, Yifan, et al. “COVID-19-CT-CXR: A Freely Accessible and Weakly Labeled Chest X-Ray and CT Image Collection on COVID-19 From Biomedical Literature.” IEEE Transactions on BigData, vol.7, no.1, Mar.2021, pp.3–12. DOI.org (Crossref), doi:10.1109/TBDATA.2020.3035935.

[31] Soda, Paolo, et al. “AIforCOVID: Predicting the Clinical Outcomes in Patients with COVID-19 Applying AI to Chest-X-Rays. An Italian Multicentre Study.” ArXiv:2012.06531 [Cs, Eess], Dec. 2020. arXiv.org, http://arxiv.org/abs/2012.06531.

[32] Cohen, Joseph Paul, et al. “COVID-19 Image Data Collection: Prospective Predictions Are the Future.” ArXiv:2006.11988[Cs, Eess, q-Bio], Dec. 2020. arXiv.org, http://arxiv.org/abs/2006.11988.

[33] Tabik, S., et al. “COVIDGR Dataset and COVID-SDNet Methodology for Predicting COVID-19 Based on Chest X-Ray Images.” IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 12, Dec. 2020, pp. 3595–605. DOI.org (Crossref), doi:10.1109/JBHI.2020.3037127.

[34] Wang, Xiaosong, et al. “ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases.” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, pp. 3462–71. DOI.org (Crossref), doi:10.1109/CVPR.2017.369.

[35] Cohen-McFarlane, Madison, et al. “Novel Coronavirus Cough Database: NoCoCoDa.” IEEE Access, vol. 8, 2020, pp. 154087–94. IEEE Xplore, doi:10.1109/ACCESS.2020.3018028.

[36] Sharma, Neeraj, et al. “Coswara — A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis.” Interspeech 2020, ISCA, 2020, pp. 4811–15. DOI.org (Crossref), doi:10.21437/Interspeech.2020-2768.

[37] معصومی، زهره و محمدی بازرگان، معظمه و ابراهیمی، زینب و رام، الهام،1399،یادگیری عمیق نظارت ضعیف برای تشخیص و طبقه بندی عفونت COVID19 و طبقه بندی تصاویرCT ،پنجمین کنفرانس ملی محاسبات نرم در مهندسی برق و کامپیوتر،اردبیل،، https://civilica.com/doc/1143212

[38] Rezaee, Khosro, et al. "A Hybrid Deep Transfer Learning Based Approach for COVID-19 Classification in Chest X-Ray Images." 2020 27th National and 5th International Iranian Conference on Biomedical Engineering (ICBME), IEEE, 2020, pp. 234–41. DOI.org (Crossref), doi:10.1109/ICBME51989.2020.9319426.

[39] Ardakani, Ali Abbasian, et al. "Application of Deep Learning Technique to Manage COVID-19 in Routine Clinical Practice Using CT Images: Results of 10 Convolutional Neural Networks." Computers in Biology and Medicine, vol. 121, June 2020, p. 103795. DOI.org (Crossref), doi: 10.1016/j.compbiomed.2020.103795.

[40] Tabrizchi, Hamed, et al. "Rapid COVID-19 Diagnosis Using Deep Learning of the Computerized Tomography Scans." 2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE), IEEE, 2020, pp. 000173–78. DOI.org (Crossref), doi:10.1109/CANDO-EPE51100.2020.9337794.

[41] Montazeri, Mahdieh, et al. "Machine Learning Models for Image-Based Diagnosis and Prognosis of COVID-19: Systematic Review." JMIR Medical Informatics, vol. 9, no. 4, Apr. 2021, p. e25181. DOI.org (Crossref), doi:10.2196/25181.

[42] Sevi, Mehmet, and Ilhan Aydin. "COVID-19 Detection Using Deep Learning Methods." 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), IEEE,2020, pp.1–6. DOI.org (Crossref), doi:10.1109/ICDABI51230.2020.9325626.

[43] Narin, Ali. "Detection of Covid-19 Patients with Convolutional Neural Network Based Features on Multi-Class X-Ray Chest Images." 2020 Medical Technologies Congress (TIPTEKNO), IEEE, 2020, pp. 1–4. DOI.org (Crossref), doi:10.1109/TIPTEKNO50054.2020.9299289.

[44] Santoso, Fian Yulio, and Hindriyanto Dwi Purnomo. “A Modified Deep Convolutional Network for COVID-19 Detection Based on Chest X-Ray Images.” 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE, 2020, pp. 700–04. DOI.org (Crossref), doi:10.1109/ISRITI51436.2020.9315479.

[45] Miranda, Igor D. S., et al. “A Comparative Study of Features for Acoustic Cough Detection Using Deep Architectures .” 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 2601–05. DOI.org (Crossref), doi:10.1109/EMBC.2019.8856412.

[46] Alsabek, Mohammed Bader, et al. “Studying the Similarity of COVID-19 Sounds Based on Correlation Analysis of MFCC.” 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), IEEE, 2020, pp. 1–5. DOI.org (Crossref), doi:10.1109/CCCI49893.2020.9256700.

[47] J. Laguarta, F. Hueto and B. Subirana, "COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings," in IEEE Open Journal of Engineering in Medicine and Biology, vol. 1, pp. 275-281, 2020, doi: 10.1109/OJEMB.2020.3026928.

[48] Imran, Ali, et al. “AI4COVID-19: AI Enabled Preliminary Diagnosis for COVID-19 from Cough Samples via an App.” Informatics in Medicine Unlocked, vol. 20, 2020, p. 100378. DOI.org (Crossref), doi: 10.1016/j.imu.2020.100378.

[49] Bansal, Vipin, et al. “Cough Classification for COVID-19 Based on Audio Mfcc Features Using Convolutional Neural Networks.” 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), IEEE, 2020, pp. 604–08. DOI.org (Crossref), doi:10.1109/GUCON48875.2020.9231094.

[50] Andreu-Perez, Javier, et al. “A Generic Deep Learning Based Cough Analysis System from Clinically Validated Samples for Point-of-Need Covid-19 Test and Severity Levels.” IEEE Transactions on Services Computing, 2021, pp. 1–1. DOI.org (Crossref), doi:10.1109/TSC.2021.3061402.

[51] Henderi, et al. “Model Decision Support System for Diagnosis COVID-19 Using Forward Chaining: A Case in Indonesia.” 2020 8th International Conference on Cyber and IT Service Management (CITSM), IEEE, 2020, pp. 1–4. DOI.org (Crossref), doi:10.1109/CITSM50537.2020.9268853.

[52] Qjidaa, M., et al. “Development of a Clinical Decision Support System for the Early Detection of COVID-19 Using Deep Learning Based on Chest Radiographic Images.” 2020 International Conference on Intelligent Systems and Computer Vision (ISCV), IEEE, 2020, pp. 1–6. DOI.org (Crossref), doi:10.1109/ISCV49265.2020.9204282.

[53] Lei, Hannah, et al. “COVID-19 Smart Chatbot Prototype for Patient Monitoring.” ArXiv:2103.06816 [Cs], Mar. 2021. arXiv.org, http://arxiv.org/abs/2103.06816.

[54] Manikandan, P., et al. “Smart Nursing Robot for COVID-19 Patients.” 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), IEEE, 2021, pp. 839–42. DOI.org (Crossref), doi:10.1109/ICACITE51222.2021.9404698.

[55] Murugan, Kalpana, et al. “Medicine Distribution Robot and Human Less Intervention for Covid-19 Affected People (AKM MED ASSISTIVE BOT).” IOP Conference Series: Materials Science and Engineering, vol. 1049, no. 1, Jan. 2021, p. 012013. DOI.org (Crossref), doi:10.1088/1757-899X/1049/1/012013.

[56] Khan, Nayyar Ahmed, and Jawad Albatein. “COVIBOT- An Intelligent WhatsApp Based Advising Bot for Covid-19.” 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), IEEE, 2021, pp. 418–22. DOI.org (Crossref), doi:10.1109/ICCIKE51210.2021.9410801.

[57] Martin, A., Nateqi, J., Gruarin, S. et al. An artificial intelligence-based first-line defence against COVID-19: digitally screening citizens for risks via a chatbot. Sci Rep 10, 19012 (2020). https://doi.org/10.1038/s41598-020-75912-x.

[58] Bharti, Urmil, et al. “Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19.” 2020 5th International Conference on Communication and Electronics Systems (ICCES), IEEE, 2020, pp. 870–75. DOI.org (Crossref), doi:10.1109/ICCES48766.2020.9137944.

[59] Ramya, B.N., Shetty, S.M., Amaresh, A.M. et al. Smart Simon Bot with Public Sentiment Analysis for Novel Covid-19 Tweets Stratification. SN COMPUT. SCI. 2, 227 (2021). https://doi.org/10.1007/s42979-021-00625-5

پیوست ها

پرسشنامه پژوهش

الف- مشخصات فردی

جنسیت: مرد ○ زن ○
سن:
وضعیت تاهل: مجرد ○ متاهل ○
محل سکونت: شهر ○ روستا ○ عشایر ○
وزن (کیلوگرم):
قد (سانتی متر):
شغل:
وضعیت فعلی کرونا در محل سکونت شما: زرد ○ نارنجی ○ قرمز ○

ب- وضعیت ابتلا به کرونا

آیا شغل شما مربوط به بهداشت و درمان است؟ بله ● خیر ●
آیا در سه هفته اخیر سابقه سفر داشته اید؟ بله ● خیر ●
آیا در اماکن تجمعی در سه هفته اخیر مراجعه نموده اید؟ بله ● خیر ●
آیا در سه هفته اخیر با بیمار قطعی کرونا تماس داشته اید؟ بله ● خیر ●
آیا در سه هفته اخیر با بیمار با علائم مشکوک به کرونا تماس داشته اید؟ بله ● خیر ●
میزان استفاده از وسیله حمل و نقل عمومی به چه میزان است؟ همیشه ● اغلب ● گهگاهی ● تا حدی ● خیلی کم ● اصلا ●
میزان استفاده شما از ماسک چقدر است؟ همیشه ● اغلب ● گهگاهی ● تا حدی ● خیلی کم ● اصلا ●
سابقه ابتلا به کووید 19 داشته اید؟ بله ● خیر ●
سابقه بستری به دلیل کووید 19 داشته اید؟ بله ● خیر ●
بیماری زمینه ای: دیابت ● سرطان ● ریوی ● قلبی-عروقی ● کبدی ● کلیوی ● بیماری خونی ● فشار خون ● بیماری مغز و اعصاب ● هیچکدام ●
شما دارای کدام یک از علائم زیر می باشید؟ تب ● لرز ● سرفه خشک ● درد مفاصل ● سرفه خلط دار ● خستگی ● ابریزش بینی ● سردرد ● گلودرد ● درد قفسه سینه ● از بین رفتن بویایی و چشایی ● بی اشتهایی ● اسهال ● حالت تهوع یا استفراغ ● تنگی نفس ● هیچکدام ●

واژگان

فارسی	انگلیسی
سرفه	Cough
کرونا	Coronavirus
تب	Fever
درد مفاصل	Joint's Pain
سرفه خلط دار	Sputum Cough
لرز	Vibration
سردرد	Headache
تنگی نفس	Shortness Of Breath
خستگی مفرط	Excessive Fatigue
تنگی و درد فقسه سینه	Chest Tightness and Pain
آبریزش بینی	Runny Nose
هیچکدام	None
گلو درد	Sore Throat
سرفه خشک	Dry Cough
از بین رفتن بویایی و چشایی	Loss Of Smell and Taste
درد قفسه سینه	Chest Pain
بی اشتهایی	Anorexia

ناراحتی معده و اسهال	Stomach Upset And Diarrhea
خارش گلو	Itchy Throat
مستندات	Documentation
پزشکی	Medical
پزشک	Doctor
عکس	Picture
سی تی	Computed Tomography
نوار	Electrocardiogram (EGG)
سایر	Other
نتیجه	Result
تحصیلات	Education
بیسواد	Illiterate
ابتدایی	Primary School
راهنمایی	Junior School
دبیرستان	High School
دانشگاهی	University
جنسیت	Gender
سن	Age
وضعیت تاهل	Marital Status
مجرد	Single

متاهل	Married
محل سکونت	Residence Location
شهر	City
روستا	Village
عشایر	Nomads
وزن	Weight
قد	Height
شغل	Job
وضعیت	Status
بهداشت و درمان	Healthcare
مربوط	Related
سفر	Trip
حمل و نقل عمومی	Public Transportation
استفاده	Usage
ماسک	Mask

History Of Infection	سابقه ابتلا
Underlying Disease	بيماري زمينه اي
Hospitalization	بستري
Crowded Places	اماكن تجمعی(شلوغ)
Infected Person	فرد آلوده

سامانه کوید

Suspicious Person	فرد مشكوك
Diabetes	ديابت
Cancer	سرطان
Pulmonary	ريوي
Cardiovascular	قلبی – عروقی
Hepatic	كبدي
Kidney	كليوي
Neurology	مغز و اعصاب
None	هيچكدام
Blood Pressure	فشار خون
Hematologic Disease	بيماري خونی
Mostly	اغلب
Sometimes	گهگاهی

تا حدي	ToSomeExtent
خیلی کم	VeryLittle
اصلا	NotAtAll