



پردیس فنی و مهندسی  
دانشکده مهندسی کامپیوتر

پایان نامه برای دریافت درجه کارشناسی ارشد  
در مهندسی کامپیوتر - نرم افزار

بهبود عملکرد پیش بینی سپسیس با استفاده از الگوریتم شبکه عصبی بازگشتی

استاد راهنما:

دکتر محمدرضا پژوهان

استاد مشاور:

دکتر مهدی رضاییان

نگارش:

پریسا اعتمادی نژاد

دی ماه ۱۴۰۰

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

## تقدیم

به پاس گرمای امیدبخش وجودشان که در این روزگاران، بهترین پشتیبان است، به پاس قلب‌های  
بزرگشان که فریادرس است و به پاس محبت‌های بی‌دریغشان، این پایان‌نامه را به پدر، مادر و همسر  
عزیزم تقدیم می‌کنم.

## سپاسگذاری

با سپاس از استاد راهنمای محترم و ارجمندم آقای دکتر محمدرضا پژوهان و همچنین استاد مشاور محترم آقای دکتر مهدی رضاییان، که با صبر و حوصله فراوان در این تحقیق مرا همراهی کردند و در هر شرایطی پشتیبان من بودند.

از سرکار خانم سمیه اعتمادی‌نژاد، پرستار با سابقه بخش مراقبت‌های ویژه، که همواره دقت نظر و مشاوره ایشان در روال تحقیق راهگشا بوده‌است بسیار سپاسگزارم.

## چکیده

یکی از عوامل های مرگ و میر در بخش مراقبت های ویژه سسپیس یا همان عفونت خون است. لذا تشخیص زودهنگام آن و انجام اقدامات لازم برای بیمار در بخش مراقبت های ویژه، از اهمیت بالایی برخوردار است. هدف از این پایان نامه، طراحی یک سیستم تصمیم یار، برای پیش بینی شروع سسپیس در بیماران بستری شده در بخش مراقبت های ویژه است؛ برای این منظور از الگوریتم شبکه عصبی بازگشتی برای پیش بینی شروع سسپیس استفاده می شود. با استفاده از این الگوریتم و مقدار ویژگی های در دسترس از بیماران، پزشک می تواند پیش بینی کند که آیا در این بیمار سسپیس شروع می شود یا خیر.

در این پایان نامه، تشخیص شروع سسپیس در پایگاه داده MIMIC-III با تمرکز بر دو موضوع متفاوت مورد بررسی قرار گرفته است. برای بررسی این دو موضوع، ابتدا مقادیر ویژگی های داده های آموزش، بر اساس تعریف سسپیس به برجسب، ابتدا به سسپیس/عدم ابتلا به سسپیس، نگاشت شده و تشخیص شروع سسپیس به یک مسئله دسته بندی تبدیل می شود. سپس ۴ الگوریتم دسته بندی شناخته شده؛ با تمرکز بر الگوریتم های شبکه عصبی بازگشتی LSTM، برای مدلسازی مورد استفاده قرار گرفت؛ که نتایج، نمایانگر قابلیت این روش ها در دسته بندی بیماران است. بعد از مدلسازی با الگوریتم های شبکه عصبی بازگشتی LSTM، در موضوع اول ابتدا مدلسازی در بازه های زمانی مختلف، برای دو بازه زمانی گذشته نگر و پیش بینی کننده در نظر گرفته شد. بعد از تعیین بهترین بازه زمانی در موضوع دوم، از بین ویژگی هایی که برای مدلسازی در نظر گرفته شده است؛ با استفاده از روش های انتخاب ویژگی، بهترین ترکیب ویژگی انتخاب شده است با ویژگی های انتخاب شده، مجدداً مدلسازی انجام شده و در نهایت بهترین ترکیب ویژگی، برای بهترین بازه های زمانی، ارائه شده است.

نتایج این پژوهش نشان می دهد که برای پیش بینی زودهنگام شروع سسپیس، بهترین بازه زمانی نگاه به گذشته ۱۲ ساعت و بهترین بازه زمانی پیش بینی ۶ ساعت قبل از شروع سسپیس است. در این پژوهش الگوریتم شبکه عصبی بازگشتی LSTM با مقدار AUROC، ۰/۹۲۵۲، بهترین الگوریتم برای پیش بینی سسپیس در این بازه های زمانی می باشد. همچنین نتایج این پژوهش نشان می دهد که ویژگی های جنسیت و میزان تنفس، تاثیر کمی در پیش بینی سسپیس دارند و بهترین مقدار برای پر کردن داده های از دست رفته در هر سری زمانی مقدار میانگین می باشد.

**واژه های کلیدی:** پیش بینی بیماری، داده سری زمانی چند متغیره، سسپیس، شبکه عصبی بازگشتی،



## فهرست مطالب

عنوان	صفحه
فصل ۱- مقدمه.....	۱
۱-۱- استفاده از تعاریف سپسیس.....	۳
۲-۱- استفاده از سیستمهای تصمیمیار پزشکی.....	۶
۱-۲-۱- علل توسعه سیستمهای تصمیم یار پزشکی.....	۶
۲-۲-۱- انواع سیستمهای تصمیمیار پزشکی.....	۷
۱-۲-۲-۱- سیستمهای تصمیمیار پزشکی از جنبه ی نوع کاربری.....	۷
۲-۲-۲-۱- سیستمهای تصمیم یار پزشکی از جنبهی نوع مداخله.....	۸
۳-۲-۲-۱- سیستمهای تصمیمیار پزشکی از جنبهی استفاده از دانش.....	۹
۳-۱- تعریف مسئله.....	۹
فصل ۲- مروری بر تحقیقات گذشته.....	۱۲
فصل ۳- آمادهدسازی دادهها.....	۲۷
۱-۳- معرفی پایگاهداده MIMIC-III.....	۲۸
۲-۳- مشکلات پایگاهداده MIMIC-III.....	۳۸
۳-۳- مراحل انتخاب مجموعه داده.....	۳۹
فصل ۴- روشهای پیشنهادی.....	۴۵
۱-۴- روشهای ارزیابی بازههای زمانی پیشبینی و نگاه به گذشته.....	۴۷
۱-۱-۴- روش پیشنهادی.....	۴۷
۱-۱-۱-۴- شبکههای عصبی بازگشتی و LSTM.....	۵۳
۲-۱-۱-۴- الگوریتم درخت تصمیم.....	۵۹
۳-۱-۱-۴- الگوریتم جنگل تصادفی.....	۶۰
۴-۱-۱-۴- شبکه عصبی پرسپترون چند لایه.....	۶۰
۲-۴- روشهای انتخاب مجموعه ویژگی.....	۶۰
۱-۲-۴- روشهای انتخاب ویژگی بسته‌بند.....	۶۱
۱-۱-۲-۴- انتخاب ویژگی مستقیم.....	۶۲
۲-۱-۲-۴- انتخاب ویژگی معکوس.....	۶۲
۳-۴- روشهای پرکردن دادههای از دست رفته.....	۶۳

فصل ۵- ارزیابی روشهای پیشنهادی.....	۶۴
۵-۱- ارزیابی بازههای زمانی پیشبینی و نگاه به گذشته.....	۶۵
۵-۲- ارزیابی روشهای انتخاب ویژگی.....	۷۲
۵-۳- ارزیابی روشهای پرکردن دادههای از دست رفته.....	۷۴
فصل ۶- نتیجه گیری و پیشنهادات.....	۷۶
۶-۱- پیشنهادات ادامه کار.....	۷۷
ضمائم.....	۷۹
فهرست مراجع.....	۸۱



## فهرست جداول

صفحه	عنوان
۵.....	جدول ۱-۱: خلاصه تغییر تعریف سپسیس.....
۱۵.....	جدول ۱-۲: عبارت‌های استفاده شده در مدل کینگلی و همکاران.....
۱۶.....	جدول ۲-۲: مقایسه نتایج مدل‌های مختلف با مدل کینگلی و همکاران در مجموعه دادگان MIMIC-III.....
۱۶.....	جدول ۳-۲: مقایسه نتایج مدل‌های مختلف با مدل کینگلی و همکاران در مجموعه دادگان eIR.....
۱۸.....	جدول ۴-۲: مقایسه الگوریتم اینسایت با الگوریتم شبکه عصبی بازگشتی.....
۲۱.....	جدول ۵-۲: مقایسه الگوریتم‌های مختلف در پژوهش ژآوو و همکاران.....
۲۲.....	جدول ۶-۲: مقایسه الگوریتم پیشنهادی بارتون و همکاران با سایر الگوریتم‌های یادگیری ماشین.....
۲۳.....	جدول ۷-۲: مقایسه الگوریتم پژوهش سباستین و همکاران با الگوریتم کالورت.....
۲۴.....	جدول ۸-۲: مقایسه الگوریتم پیشنهادی فاگرستروم و همکاران با سایر الگوریتم‌ها.....
۲۶.....	جدول ۹-۲: خلاصه سایر پیشینه‌های تحقیق.....
۳۲.....	جدول ۱-۳: انواع مختلف داده در پایگاه‌داده MIMIC-III.....
۴۲.....	جدول ۲-۳: ویژگی‌های استفاده شده بعنوان ورودی مدل و کدهای شناساگر آن‌ها از جدول d_items از پایگاه‌داده.....
۴۴.....	جدول ۳-۳: لیست متغیرهای استفاده شده در اندازه‌گیری درجه SOFA و جدول منبع در پایگاه‌داده.....
۶۱.....	جدول ۱-۴: مقایسه متغیرهای مختلف برای الگوریتم‌های منتخب یادگیری ماشین و انتخاب بهترین آن‌ها.....
۶۹.....	جدول ۱-۵: مقایسه نتایج پیاده‌سازی حالت‌های مختلف شبکه LSTM.....
۷۴.....	جدول ۲-۵: مقایسه الگوریتم‌های یادگیری ماشین با شبکه LSTM.....

عنوان	صفحه
شکل ۱-۲: معماری مدل کینگلی و همکاران	۱۴
شکل ۲-۲: گراف اطلاعات پزشکی استفاده شده بعنوان یک منبع مورد اعتماد	۱۵
شکل ۳-۲: مقایسه الگوریتم اینسایت با شبکه عصبی بازگشتی در بازه‌های زمانی مختلف	۱۸
شکل ۴-۲: طرح کلی پژوهش ژاوو و همکاران	۲۰
شکل ۱-۳: نمودار زمانی یک بیمار در پایگاه داده MIMIC-III و کدهای شناساگر مربوطه	۳۱
شکل ۲-۳: روند استخراج داده از پایگاه داده MIMIC-III	۳۲
شکل ۳-۳: شمای جدول chartevents در پایگاه داده MIMIC-III	۳۴
شکل ۴-۳: شمای جدول inpuvents_cv در پایگاه داده MIMIC-III	۳۶
شکل ۵-۳: شمای جدول inpuvents_mv در پایگاه داده MIMIC-III	۳۷
شکل ۶-۳: ارتباط جدول labevents و d_labitems در پایگاه داده MIMIC-III	۳۸
شکل ۷-۳: شمای جدول outpuvents در پایگاه داده MIMIC-III	۳۸
شکل ۸-۳: ارتباط جدول diagnoses_icd و d_icd_diagnoses در پایگاه داده MIMIC-III	۳۹
شکل ۹-۳: نحوه محاسبه مقدار درجه SOFA	۴۵
شکل ۱-۴: مراحل روش پیشنهادی	۴۹
شکل ۲-۴: روند استخراج مجموعه دادگان برای هر زوج مرتب با بازه زمانی پیش‌بینی n ساعت	۵۱
شکل ۳-۴: نمونه داده استفاده شده برای استخراج داده مورد نیاز	۵۲
شکل ۴-۴: مفهوم بازه زمانی نگاه به گذشته و پیش‌بینی	۵۳
شکل ۵-۴: زوج مرتب‌های بررسی شده در پژوهش	۵۴
شکل ۶-۴: طول بازه‌های بستری در بیمارستان برای دو دسته شروع سپسیس و عدم شروع سپسیس	۵۴
شکل ۷-۴: الف- شبکه بازگشتی المان ب- شبکه بازگشتی جردن	۵۶

- شکل ۴-۸: مشکل اختفای گرادیان در شبکه‌های بازگشتی ..... ۵۷
- شکل ۴-۹: بلوک LSTM ..... ۵۸
- شکل ۴-۱۰: مدل پیشنهادی ..... ۶۰
- شکل ۴-۱۱: روال کلی انتخاب ویژگی بسته‌بند ..... ۶۴
- شکل ۵-۱: نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار دقت ..... ۷۱
- شکل ۵-۲: مقایسه نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار AUC-ROC ..... ۷۲
- شکل ۵-۳: مقایسه نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار صحت ..... ۷۲
- شکل ۵-۴: مقایسه نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار CWA ..... ۷۳
- شکل ۵-۵: نمودار مقایسه تاثیر ویژگی‌های مختلف در پیاده‌سازی شبکه LSTM به روش forward stepwise ..... ۷۵
- شکل ۵-۶: نمودار مقایسه تاثیر ویژگی‌های مختلف در پیاده‌سازی شبکه LSTM به روش RFE ..... ۷۶
- شکل ۵-۷: مقایسه روش‌های پرکردن داده‌های از دست رفته ..... ۷۷



## فصل ۱- مقدمه

سپسیس<sup>۱</sup> یک نوع بیماری است؛ که وقتی بدن نمی‌تواند در مقابل عفونت دوام بیاورد اتفاق می‌افتد. این عفونت می‌تواند باعث تغییراتی در بدن شود و به عضوهای مختلف بدن آسیب برساند و در صورتی که این بیماری شدید شود، منجر به مرگ می‌شود. در مرحله اولیه‌ی بیماری، درمان سپسیس با استفاده از مصرف زیاد آنتی‌بیوتیک‌ها، نسبتاً آسان است. با این حال، تشخیص بیماری در این مرحله نیز یک چالش است. در مراحل بعدی، تشخیص بیماری بسیار آسان‌تر است، اما درمان آن بسیار سخت است [۱]. این بیماری با خطر بالای عوارض، افزایش طول زمان بستری در بیمارستان، هزینه بالای پزشکی و مرگ همراه است [۲].

امروزه سپسیس به یکی از دلایل عمده مرگ در بیمارستان برای بیماران بخش مراقبت‌های ویژه تبدیل شده است بطور متوسط، سپسیس هر ساله ۱۱ میلیون نفر را می‌کشد؛ بیشتر این مرگ‌ها، در کشورهای کم درآمد یا با درآمد متوسط<sup>۲</sup> اتفاق می‌افتد. در این کشورها مراقبت از بیماران درگیر عفونت، با چالش‌هایی نظیر کمبود کارکنان درمانی آموزش دیده، عدم کفایت خدمات تشخیص عفونت، زیرساخت‌های قدیمی و حتی دسترسی محدود به آب و برق در برخی مکان‌ها روبرو است. رعایت نکات بهداشت عمومی، بسیاری از بیماری‌های عفونی را کاهش می‌دهد، اما رعایت این نکات در برابر سپسیس کارایی کمی دارد [۲]. در کشورهای پر درآمد، از میزان شدت بیماری<sup>۳</sup> معمولاً برای شناسایی بیمارانی که بیشتر در معرض خطر هستند استفاده می‌شود. با این حال همه این‌ها، در کشورهای توسعه یافته و پردرآمد تایید شده‌اند و بسیاری از آن‌ها به پشتیبانی آزمایشگاه‌های پیچیده نیاز دارند؛ که ممکن است در کشورهای کم‌درآمد یا با درآمد متوسط موجود نباشد. علاوه بر این بیماری‌های همراه با سپسیس، در کشورهایی با درآمد کم یا درآمد متوسط، متفاوت است [۲].

با توجه به مرگ‌ومیر بالای بیماران سپسیس در ICU<sup>۴</sup>، هر چه زودتر این عفونت در بیماران تشخیص داده‌شود بهتر است. شناسایی سریع‌تر و دقیق بیماران سپسیس با خطر بالای مرگ، می‌تواند به پزشکان در ICU کمک کند تا تصمیمات بالینی بهتری بگیرند. که به نوبه خود می‌تواند؛ نتایج بالینی آن‌ها را بهبود بخشد [۳]. امروزه در زمینه بهداشت و درمان، حجم زیادی از دادگان تولید می‌شود؛ که اغلب در سیستم‌های جداگانه پراکنده هستند. دادگان علائم حیاتی، نتایج تست‌های آزمایشگاهی، یادداشت‌های پیشرفت بیماران و داروهای مصرفی در پرونده الکترونیکی بیمار ذخیره می‌شود. این یک چالش برای پزشکان است؛ زیرا آن‌ها غرق در اطلاعات زیادی هستند؛ که برای تصمیم‌گیری ابتدا باید داده‌ها را قبل از استفاده، جمع‌آوری و درک کنند.

<sup>1</sup> Sepsis

<sup>2</sup> Low-income and Middle-Income Countries (LMIC)

<sup>3</sup> severity scores

<sup>4</sup> Intensive care units

با استفاده از ابزارهای تشخیصی و پیش‌بینی، برای پزشکان دشوار است که بیماران مبتلا به سپسیس را زود تشخیص دهند و بهترین روش درمان را برای هر بیمار پیش‌بینی کنند. یکی از دلایلی که برای این مسئله می‌توان بیان کرد این است که سپسیس یک سندرم ناهمگن است؛ به این معنی که بیماران با علائم مختلف ممکن است به سپسیس مبتلا باشند. تقریباً از هر پنج بیمار که با شک به سپسیس به بخش اورژانس مراجعه می‌کنند؛ هیچ اختلالی در عملکرد اعضای آن‌ها وجود ندارد؛ در حالی که این بیماران ظرف ۴۸ ساعت پس از پذیرش، در عملکرد اعضای آن‌ها اختلال ایجاد می‌شود [۱].

## ۱-۱- استفاده از تعاریف سپسیس

تعریف اولیه سپسیس در سال ۱۹۹۱ ارائه شد، که به آن سپسیس یک گفته می‌شد. واکنش التهابی عمومی بدن<sup>۱</sup>؛ یعنی هرگاه به هر علتی بدن در مقابل عوامل تهدیدکننده سیستم<sup>۲</sup> ایمنی، واکنش نشان دهد؛ به گونه‌ای که هنگام معاینه بیمار یا آزمایشات بالینی اولیه، این معیارها وجود داشته باشد، SIRS گفته می‌شود؛ این معیارها عبارتند از: (۱) افزایش درجه حرارت بدن به بیش از ۳۸ درجه سانتیگراد یا کاهش دمای بدن به زیر ۳۶ درجه. (۲) افزایش تعداد ضربان قلب به بیش از ۹۰ ضربان در دقیقه. (۳) افزایش تعداد تنفس بیمار به بیش از ۲۰ تنفس در دقیقه و یا اینکه طی اندازه‌گیری فشار، گاز دی‌اکسید کربن خون<sup>۳</sup>، کمتر از ۳۲ باشد. (۴) افزایش تعداد گلبول‌های سفید خون به بیش از ۱۲۰۰۰ سلول در میلی لیتر مکعب خون و یا کاهش آن به کمتر از ۴۰۰۰ در هر میلی لیتر مکعب خون یا اینکه تعداد گلبول‌های سفید نارس<sup>۴</sup> خون، بیش از ۱۰٪ کل گلبول‌های سفید باشد. گلبول‌های سفید نارس زمانی وارد گردش خون می‌شوند، که نیاز بدن به گلبول سفید، بیش از میزان تولید آن در مغز استخوان باشد. در این حالت مغز استخوان که کار ساخت و پرورش گلبول‌های سفید را به عهده دارد؛ مجبور می‌شود؛ این گلبول‌ها را که هنوز باید در مغز استخوان بمانند تا پرورش کامل یابند را روانه گردش خون کند. این حالت که باندمی نیز خوانده می‌شود؛ نشان دهنده این است که بدن با یک عفونت جدی روبرو است [۴].

تعریف سپسیس یک در سال ۲۰۰۱ به تعریف سپسیس دو بروزرسانی شد [۵]. در این تعریف، شرح معیارهای تعریف سپسیس یک گسترش یافتند، ولی موضوع اصلی این معیارها، تغییری نکردند. در حقیقت تعریف سپسیس بیش از دو دهه است، که بدون تغییر باقی مانده است. در سال ۲۰۱۶ تعریف جدید از سپسیس ارائه شد [۶]. در سپسیس سه به جای بررسی معیارهای SIRS از درجه SOFA برای

<sup>1</sup> Systemic Inflammatory Response Syndrome (SIRS)

<sup>2</sup> System

<sup>3</sup> Pressure Carbon Dioxide (PCO2)

<sup>4</sup> Band Cell

تشخیص بیماران سپسیس، استفاده شده‌است در این تعریف بیماران عفونی که درجه SOFA بیشتر از ۲ دارند، بعنوان بیمار مبتلا به سپسیس شناسایی می‌شوند [۷].

در جدول ۱-۱ بصورت خلاصه معیارهای تعریف‌های سپسیس نشان داده شده‌است.

جدول ۱-۱: خلاصه تغییر تعریف سپسیس [۸]

Older definitions		Newer definition: Sepsis 3	
Sepsis 1	Sepsis 2	Definition	Clinical Criteria
<p>Systemic inflammatory response syndrome (SIRS) = the systemic inflammatory response to a variety of severe clinical insults.</p> <ol style="list-style-type: none"> <li>1. Temperature <math>&gt;38^{\circ}\text{C}</math> or <math>&lt;36^{\circ}\text{C}</math>;</li> <li>2. Heart rate <math>&gt;90</math> beats per minute;</li> <li>3. Respiratory rate <math>&gt;20</math> breaths per minute or <math>\text{PaCO}_2 &lt; 32</math> mmHg; and</li> <li>4. White blood cell count <math>&gt;12,000/\text{cu mm}</math>, <math>&lt;4,000/\text{cu mm}</math>, or <math>&gt;10\%</math> immature (band) forms.</li> </ol> <p>Sepsis = the systemic response to infection, manifested by two or more of the SIRS criteria as a result of infection Severe sepsis = sepsis associated</p>	<p><b>Diagnostic criteria for sepsis Infection</b> Documented or suspected and some of the following: <b>General parameters</b> Fever (core temperature <math>&gt;38.3^{\circ}\text{C}</math>) Hypothermia (core temperature <math>&lt;36^{\circ}\text{C}</math>) Heart rate <math>&gt;90</math> bpm or <math>&gt;2</math> SD above the normal value for age Tachypnea: <math>&gt;30</math> bpm Altered mental status Significant edema or positive fluid balance (<math>&gt;20</math> mL kg<sup>-1</sup> over 24 h) Hyperglycemia (plasma glucose <math>&gt;110</math> mg dL<sup>-1</sup> or 7.7 mM L<sup>-1</sup>) in the absence of diabetes <b>Inflammatory parameters</b> Leukocytosis (white blood cell count <math>&gt;12,000/\mu\text{L}</math>) Leukopenia (white blood cell count <math>&lt;4,000/\mu\text{L}</math>) Normal white blood cell count with <math>&gt;1\%</math> immature forms Plasma C reactive protein <math>&gt;2</math> SD above the normal value Plasma procalcitonin <math>&gt;2</math> SD above the normal value</p>	<p><b>Screening for Sepsis</b> qSOFA (quick Sequential Organ Failure Assessment) scoring system Accordingly, an increase of 2 or more in the qSOFA score should create a suspicion of sepsis and organ dysfunction.  Sepsis is life-threatening organ dysfunction caused by a dysregulated host response to infection</p>	<p><b>qSOFA</b> 1. Altered mental status (GCS score <math>&lt;15</math>) 2. Systolic blood pressure <math>&lt;100</math> mmHg 3. Respiratory rate <math>&gt;22/\text{min}</math>  If 2..3 of these 3 criteria are positive, the QSFA would be positive!  Sepsis and vasopressor therapy needed to elevate <math>\text{MAP} \geq 65</math> mm Hg and lactate <math>&gt;2</math> mmol L<sup>-1</sup> (18 mg dL<sup>-1</sup>) despite adequate fluid</p>



جدول ۱-۱: خلاصه تغییر تعریف سپسیس (ادامه) [۸]

with organ dysfunction, hypoperfusion, or hypotension. Hypoperfusion and perfusion abnormalities may include, but are not limited to lactic acidosis, oliguria, or an acute alteration in mental status. Septic shock = sepsis-induced with hypotension despite adequate fluid resuscitation along with the presence of perfusion abnormalities that may include, but are not limited to, lactic acidosis, oliguria, or an acute alteration in mental status. Patients who are receiving inotropic or vasopressor agents may not be hypotensive at the time that perfusion abnormalities are measured.	<p><b>Hemodynamic parameters</b> Arterial hypotension (systolic blood pressure &lt;90 mmHg, mean arterial pressure &lt;70, or a systolic blood pressure decrease &gt;40 mmHg in adults or &lt;2 SD below normal for age) Mixed venous oxygen saturation &gt;70% Cardiac index &gt;305 l min<sup>-1</sup>m<sup>-2</sup></p> <p><b>Organ dysfunction parameters</b> Arterial hypoxemia (PaO<sub>2</sub>/FiO<sub>2</sub> &lt;300) Acute oliguria (urine output &lt;0.5 ml kg<sup>-1</sup> h<sup>-1</sup> or 45 mM L<sup>-1</sup> for at least 2 h) Creatinine increase ≥0.5 mg dL<sup>-1</sup> Coagulation abnormalities (international normalized ratio &gt;1.5 or activated partial thromboplastin time &gt;60 s) Ileus (absent bowel sounds) Thrombocytopenia (platelet count &lt;100000/μL) Hyperbilirubinemia (plasma total bilirubin &gt;4 mg dL<sup>-1</sup> or 70 mmol L<sup>-1</sup>)</p> <p><b>Tissue perfusion parameters</b> Hyperlactatemia (&gt;3 mmol L<sup>-1</sup>) Decreased capillary refill or mottling</p>	Septic shock is a subset of sepsis in which underlying circulatory and cellular/metabolic abnormalities are profound enough to substantially increase mortality	resuscitation Sepsis and vasopressor therapy needed to elevate MAP ≥6.5 mm Hg and lactate >2 mmol L <sup>-1</sup> (18 mg dL <sup>-1</sup> ) despite adequate fluid resuscitation
--	--	---	---

برخی از ابزارهای امتیازدهی برای ارزیابی شدت سپسیس ایجاد شده است. در میان آن‌ها ابزار فیزیولوژیک حاد ساده<sup>۱</sup>، ارزیابی فیزیولوژی حاد و سلامت مزمن<sup>۲</sup> و درجه SOFA اغلب در ICU بعنوان ابزارهای

<sup>1</sup> Simplified Acute Physiology Score 2 (SAPS2)

<sup>2</sup> Acute Physiology and Chronic Health Evaluation (APACHE)

ارزیابی شدت استفاده می‌شوند، همچنین علائم حیاتی بیمار، نتایج آزمایشگاهی و گزارشات دموگرافیک فاکتورهایی هستند، که در سیستم‌های تعیین شدت میزان بیماری سپسیس استفاده می‌شوند، با این حال عملکرد این معیارها به شدت کاهش یافته‌است؛ زیرا این معیارها با گذشت زمان تغییر کرده‌است.

این‌ها ابزاری برای تشخیص زودهنگام سپسیس نیست و با استفاده از آن‌ها، فقط می‌توان مشخص کرد، که آیا در همان لحظه بیمار به سپسیس مبتلا است یا خیر. به منظور تشخیص سپسیس، پزشکان به تغییرات فیزیولوژیکی بیماران توجه می‌کنند، با این حال وابستگی‌های زمانی و تغییرات فیزیولوژیکی پنهان، ممکن است نشانه‌ای از سپسیس قبل از اختلال عملکرد ارگان‌های تهدید کننده حیات باشد، که برای برآورده کردن معیارهای تشخیص فعلی سپسیس، مورد نیاز است. از آنجا که در نظر گرفتن همه ی این آثار و یافته‌ها برای هر بیمار از نظر انسانی امکان پذیر نیست، باید ابزار تجزیه و تحلیل خودکار، ایجاد شود [۸] و سیستم‌هایی برای پیش‌بینی مرگ و میر بیماران مبتلا به سپسیس در بیمارستان در بیمار ICU طراحی شود [۳].

## ۱-۲- استفاده از سیستم‌های تصمیم‌یار پزشکی

یکی از روش‌هایی که برای حل چالش‌های این حوزه مورد استفاده قرار می‌گیرد، سیستم تصمیم‌یار پزشکی می‌باشد. سیستم تصمیم‌یار پزشکی با تحلیل داده‌های به دست آمده از بیماران، به پزشک در روند تشخیص و درمان، یاری می‌رساند. با استفاده از این سیستم‌های تصمیم‌یار برای هر بیمار، بطور خاص پیشنهادی به پزشک ارائه می‌شود.

### ۱-۲-۱- علل توسعه سیستم‌های تصمیم‌یار پزشکی

بطور کلی، انگیزه‌ی توسعه‌ی سیستم‌های تصمیم‌یار پزشکی را می‌توان بصورت موارد زیر بیان کرد [۹]:

- حجم اطلاعات: کم بودن اطلاعات مناسب ممکن است، اثرات منفی در درمان بیمار داشته‌باشد، اما داشتن داده‌ای زیاد که گیج‌کننده می‌شود نیز، می‌تواند همان اثر منفی را در درمان داشته‌باشد. با استفاده از سیستم تصمیم‌یار می‌توان، اطلاعات مفیدی از حجم زیاد داده‌ای جمع‌آوری شده به دست آورد.
- سوالات بی‌پاسخ: طبق بررسی‌های انجام شده، به ازای هر دو بیمار، یک سوال برای پزشک ایجاد می‌شود و پزشکان، تنها به دنبال پاسخ ۳۳٪ از سوال‌هایشان می‌روند. دو فاکتور مهم در ترغیب پزشک در یافتن پاسخ، باور پزشک به این است که پاسخ مشخصی برای سوال وجود دارد و

دیگری حیاتی بودن مشکل بیمار است [۱۰]. این پرسش‌های بی پاسخ می‌توانند، انگیزه‌های قوی برای توسعه‌ی سیستم‌های تصمیم‌یار پزشکی باشند.

- کمبود زمان: به علت کمبود زمان، تمایل پزشکان به پذیرش اولین تشخیصی که تمام مشاهدات را تأیید می‌کند، بدون در نظر گرفتن سایر احتمالات، زیاد می‌باشد. این مسأله می‌تواند، عامل مهمی برای توسعه‌ی سیستمی باشد، که در سریع‌ترین زمان به پزشک در روند تشخیص یاری برساند.

- خطاهای پزشکی و اتفاقات ناشی از آن‌ها: کوهن<sup>۱</sup> و همکاران در [۱۱] اشاره کرده اند که خطاهای پزشکی بیش از تصادفات وسایل نقلیه و سرطان سینه، تلفات جانی به همراه داشته‌است طبق تحقیقات، بسیاری از اتفاقات ناشی از خطاهای پزشکی، قابل اجتناب است، در حالی که سالانه میلیون‌ها رخداد از این دست گزارش می‌شود. در پاسخ به این مسأله، سیستم‌های تصمیم‌یار می‌توانند، نقش مهمی را در کاهش این اتفاقات داشته باشند.

با توجه به انگیزه‌های ذکر شده برای توسعه‌ی این سیستم‌ها، در ادامه به بررسی انواع سیستم‌های تصمیم‌یار پزشکی می‌پردازیم.

## ۱-۲-۲- انواع سیستم‌های تصمیم‌یار پزشکی

سیستم‌های تصمیم‌یار پزشکی را می‌توان از جنبه‌های مختلفی دسته‌بندی کرد.

### ۱-۲-۲-۱- سیستم‌های تصمیم‌یار پزشکی از جنبه‌ی نوع کاربری

در این حالت سیستم‌های تصمیم‌یار پزشکی به دو دسته‌ی سیستم‌های هشدار و سیستم‌های پیشبینی وضعیت بیمار<sup>۲</sup> تقسیم می‌شوند [۹]. از جمله سیستم‌های هشدار می‌توان، به این موارد اشاره کرد: سیستم‌های یادآور<sup>۳</sup> مناسب برای ماموگرافی، سیستم‌های اخطاردهنده<sup>۴</sup>، سیستم‌هایی برای تعیین تداخلات دارویی، میزان مجاز دارو و آلرژی‌های احتمالی و سیستم‌هایی برای تشخیص تصاویر پزشکی و تفسیر آن‌ها.

دسته‌ی دیگر، سیستم‌های پیشبینی وضعیت بیمار می‌باشد. از جمله مواردی که می‌توان در این دسته ذکر کرد، سیستم تصمیم‌یار برای محاسبه‌ی احتمال حمله‌ی قلبی با استفاده از پارامترهایی همچون وزن، میزان تحرک و برخی مقادیر آزمایش‌ها می‌باشد.

---

<sup>۱</sup> Kohn

<sup>۲</sup> Prognosis Systems

<sup>۳</sup> Reminder Systems

<sup>۴</sup> Warning Systems

از دیگر نمونه‌های این نوع سیستم‌ها، سیستم APACHE<sup>1</sup> می‌باشد، که برای دسته‌بندی شدت بیماری به کار می‌رود. این سیستم براساس یک سری متغیر، امتیازی به هر بیمار می‌دهد. امتیاز بالاتر، نشانه‌ی شدید بودن بیماری است.

برخی دیگر از سیستم‌ها که در این دسته قرار می‌گیرند، سیستم‌های تصمیم‌یار تشخیصی<sup>2</sup> هستند، که در انتخاب تشخیص مناسب از بین گزینه‌های موجود به پزشک یاری می‌رسانند. در واقع براساس پارامترهای مشخصی، تشخیص‌های احتمالی به کاربر پیشنهاد داده می‌شود و کاربر با تعامل با سیستم، نتایج به دست آمده را بهبود می‌بخشد.

#### ۱-۲-۲- سیستم‌های تصمیم‌یار پزشکی از جنبه‌ی نوع مداخله

یک نوع دیگر دسته‌بندی سیستم‌های تصمیم‌یار پزشکی، براساس نوع مداخله‌ای است که این سیستم‌ها در پزشکی دارند. در این حالت سیستم‌های تصمیم‌یار پزشکی به دو دسته‌ی منفعل<sup>3</sup> و فعال<sup>4</sup> تقسیم می‌شوند [۱۲، ۱۳].

**سیستم‌های تصمیم‌یار پزشکی منفعل:** در این سیستم‌ها، پزشک درخواست خود را در سیستم ایجاد کرده و منتظر می‌ماند تا سیستم پاسخ بدهد. این سیستم‌ها شامل سیستم‌های مشاوره‌ای و سیستم‌های انتقادی هستند.

در سیستم مشاوره‌ای، کاربر اطلاعات مربوط به بیمار را در سیستم وارد می‌کند و سیستم پیشنهاد درمانی و یا تشخیصی را به پزشک می‌دهد. در سیستم انتقادی، اطلاعاتی در ارتباط با وضعیت بیمار و درمان پیشنهادی پزشک به سیستم، داده می‌شود و سیستم در پاسخ، درمان پیشنهادی پزشک را نقد می‌کند.

**سیستم‌های تصمیم‌یار پزشکی فعال:** دسته‌ی دیگر، سیستم‌های فعال هستند که به طور خودکار، شروع به کار کرده و بدون ایجاد درخواست از طرف کاربر، با پزشک تعامل دارند.

این سیستم‌ها، شامل سیستم‌های هشدار و سیستم‌های یادآور خودکار می‌باشند. در سیستم‌های هشدار، براساس وضعیت بیمار هشدارهایی داده می‌شود. این هشدارها، گاهی برای اعلام مقادیر غیرطبیعی و گاهی برای نمایش تغییرات غیرطبیعی در پارامترهای بیمار می‌باشد. سیستم‌های یادآور خودکار نیز، برای جلوگیری از معاینات اضافی و تسهیل روند درمان توسط پزشک به کار می‌روند.

<sup>1</sup> Acute Physiologic and Chronic Health Evaluation

<sup>2</sup> Diagnostic Decision Support System

<sup>3</sup> Passive

<sup>4</sup> Active

### ۱-۲-۳- سیستم‌های تصمیم‌یار پزشکی از جنبه‌ی استفاده از دانش

دسته‌بندی دیگری که برای سیستم‌های تصمیم‌یار پزشکی وجود دارد این است که آیا سیستم مبتنی بر دانش بوده یا نه، که از رویکردهای شناسایی الگو و یادگیری ماشین استفاده می‌کند [۱۲].

**سیستم‌های مبتنی بر دانش:** بسیاری از سیستم‌های تصمیم‌یار پزشکی مبتنی بر دانش، از حوزه‌ی تحقیقاتی سیستم‌های خبره بوجود آمده‌اند؛ که در این حوزه، هدف، ساختن برنامه کامپیوتری است که تفکر انسان را شبیه‌سازی کند. در این نوع سیستم‌های تصمیم‌یار، هدف، شبیه‌سازی روال تصمیم‌گیری خبره نمی‌باشد، بلکه هدف، یاری پزشک در تصمیم‌گیری است. در این سیستم‌ها به جای یک پاسخ، اطلاعاتی در اختیار کاربر قرار داده می‌شود و کاربر از بین آن‌ها اطلاعات غلط و بی‌فایده را کنار می‌گذارد. در واقع کاربر با سیستم تعامل دارد و نقش فعالی در سیستم ایفا کرده‌است [۱۲].

سیستم تصمیم‌یار پزشکی مبتنی بر دانش، شامل قواعدی به شکل جملات اگر-آنگاه<sup>۱</sup> می‌باشند و داده‌ها با این قواعد در ارتباط هستند [۱۲]. برای مثال، اگر بیماری حداقل دو مورد از معیارهای SIRS را داشته باشد، آنگاه ممکن است به سپسیس مبتلا باشد. [۱۴]

سیستم‌های مبتنی بر دانش، مشتمل بر سه بخش عمده هستند:

- ۱- پایگاه دانش، که شامل قواعد سیستم است.
- ۲- موتور استنتاج، که وظیفه ترکیب قواعد را با داده‌های مربوط به بیمار را بر عهده دارد.
- ۳- مکانیسم ارتباطی، برای نشان‌دادن قواعد به کاربران.

**سیستم‌های غیرمبتنی بر دانش<sup>۲</sup>:** برخلاف سیستم‌های مبتنی بر دانش، این سیستم‌ها از تکنیک‌های هوش مصنوعی برای یادگیری و کشف الگوها در داده‌های پزشکی استفاده می‌کنند. از انواع سیستم‌های غیر مبتنی بر دانش می‌توان، به سیستم‌هایی که با شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک ساخته می‌شوند، اشاره کرد [۱۲].

### ۱-۳- تعریف مسئله

یکی از حوزه‌های فعال تحقیق در زمینه‌ی انفورماتیک پزشکی و سیستم‌های تصمیم‌یار پزشکی، پزشکی اضطراری<sup>۳</sup> است. برای بیماران بستری در ICU، روند تشخیص و درمان، باید به موقع و در زمان معقولی انجام شود [۱۳].

با توجه به توسعه سیستم‌های تصمیم‌یار پزشکی، در این مطالعه، هدف، ارائه یک سیستم تصمیم‌یار برای پیش‌بینی سپسیس و یاری رساندن به پزشک، به منظور تشخیص زودهنگام سپسیس است. برای بهبود

<sup>۱</sup> If-Then Rules

<sup>۲</sup> Non-knowledge-based systems

<sup>۳</sup> Emergency Medicine

نتایج تشخیص سپسیس ضروری است، که زمان تشخیص و دقت پیش‌بینی سپسیس را بهبود بخشیم. در بررسی مسائل مربوط به بهبود پیش‌بینی سپسیس، یک سری چالش‌ها وجود دارد، که می‌توان گفت تقریباً تمامی پژوهش‌های این حوزه، در محور این چالش‌ها حرکت می‌کنند. بطور کلی، چالش‌هایی که در پیش‌بینی سپسیس بررسی می‌شود، شامل:

۱- کاهش زمان نگاه به گذشته و کاهش زمان پیش‌بینی شروع سپسیس

۲- استفاده از حداقل تعداد پارامتر برای پیش‌بینی

۳- چالش پرکردن داده‌های از دست رفته

هر کدام از این چالش‌ها و چالش‌های دیگری که در حوصله این پژوهش نمی‌گنجد، یک مسئله قابل بررسی در بهبود پیش‌بینی سپسیس می‌باشد. که در ادامه، توضیح مختصری درمورد هر کدام، داده خواهد شد.

با توجه به اینکه همواره، از بیماران طی بازه‌های زمانی مختلف مثلاً، هر یک ساعت یک بار مقادیر مختلفی ذخیره می‌شود، لذا، اینکه بتوان با پردازش کمترین اطلاعات، بیشترین دقت را در پیش‌بینی سپسیس بدست آورد، حائز اهمیت می‌باشد. در این پژوهش بازه‌های زمانی مختلف برای نگاه به گذشته بیمار و بازه‌های زمانی مختلف قبل از شروع سپسیس<sup>۱</sup> بررسی خواهد شد. با مشخص شدن بازه‌زمانی مناسب، اینکه بتوان از کمترین اطلاعات بیمار، بهترین نتیجه را برای پیش‌بینی سپسیس بدست آورد مهم است. در این پژوهش بعد از مشخص کردن بهترین بازه‌زمانی مشاهده و بهترین بازه‌زمانی پیش‌بینی، با توجه به معیارهای ارزیابی سعی می‌کنیم مهم‌ترین ویژگی‌های موجود را استخراج کنیم. بعد از استخراج ویژگی‌های اساسی بایستی مقادیر از دست رفته برای بیماران مدیریت شود، که با روش‌های مختلف این پیش پردازش انجام خواهد شد. بطور کلی، نمی‌توان بهترین حالت را بطور دقیق مشخص کرد، لذا با توجه به میزان اهمیت هر موضوع، می‌توان بهترین انتخاب را در هر چالش داشت. برای ارزیابی حالت‌های مختلف هر کدام از این چالش‌ها، از مدل‌های پیش‌بینی هوش مصنوعی که در ادامه بررسی خواهند شد استفاده می‌شود. مدل‌های پیش‌بینی هوش مصنوعی که نشان داده شده اند، برای تشخیص و پیش‌بینی بیماری در زمینه‌های مختلف پزشکی، مفید است که برای بیماران مبتلا به سپسیس نیز سودمند است [۱]. داده‌های مورد استفاده از نسخه ۱/۴ پایگاه داده MIMIC-III<sup>۲</sup> انتخاب شده است.

<sup>۱</sup> Sepsis Onset

<sup>۲</sup> Medical Information Mart for Intensive Care-III

## ساختار کلی پایان نامه

این پایان نامه در ۶ فصل تنظیم شده است در فصل دوم، تحقیقاتی که با هدف پیش بینی سپسیس انجام شده است، بررسی می شوند. در فصل سوم پایگاه داده MIMIC-III بعنوان یک مرجع بسیار مناسب معرفی شده و چالش های کار با این پایگاه داده معرفی شده است. در این فصل همچنین، نحوه استخراج مجموعه داده های مورد نظر، قدم به قدم توضیح داده می شود. در فصل چهارم از الگوریتم های داده کاوی، برای پیش بینی سپسیس استفاده می شود و بصورت ویژه به بررسی شبکه عصبی بازگشتی، برای پیش بینی سپسیس در بیماران استفاده می شود. در فصل پنجم، ارزیابی و تحلیل نتایج انجام می شود. در نهایت در فصل ششم، نتیجه گیری و پیشنهاداتی برای ادامه تحقیق ارائه خواهد شد.

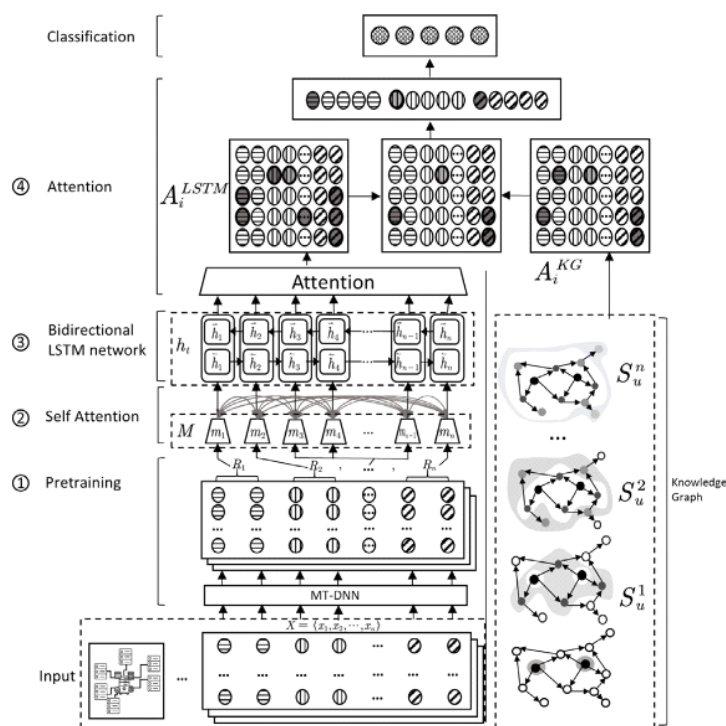
## فصل ۲- مروری بر تحقیقات گذشته



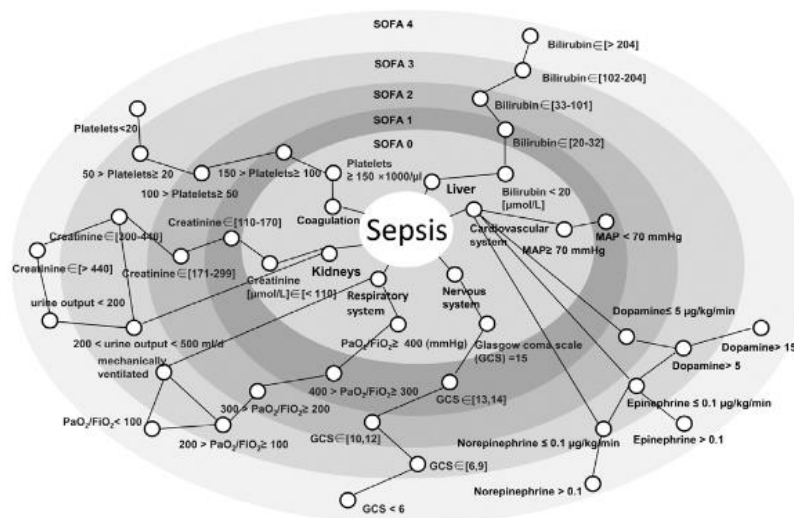
با توجه به اهمیت موضوع پیش‌بینی سپسیس در ICU، و با توجه به حساسیت بخش ICU، تحقیقات مختلفی با هدف بهبود پیش‌بینی سپسیس انجام شده‌است

در تحقیق کینگ لی و همکاران، یک مدل یادگیری عمیق مطابق با شکل ۱-۲ برای پیش‌بینی شدت سپسیس بیماران، طراحی شده‌است در این تحقیق از یک گراف اطلاعات پزشکی مطابق با شکل ۲-۲، بعنوان یک منبع قابل اعتماد و قوی از اطلاعات جانبی استفاده شده‌است این پژوهش از الگوریتم شبکه عصبی و تجزیه و تحلیل همزمان اندام‌های مختلف بدن و بصری سازی آن در در مدل، استفاده کرده‌است، که با استفاده از این الگوریتم، توانسته وضعیت بیمار را به موقع منعکس کند.

در این پژوهش از داده بزرگسالان مجموعه دادگان واقعی MIMIC-III و eIR استفاده شده‌است بعد از پاکسازی دادگان، ۱۷۰۸۳ پذیرش از دادگان MIMIC-III و ۷۹۰۷۳ پذیرش از دادگان eIR و ۴۸ ویژگی برتر استفاده شده‌است ابتدا ۸۰ درصد از پذیرش‌ها، از دادگان MIMIC-III در مرحله آموزش مدل و ۲۰ درصد در مرحله آزمایش مدل استفاده شده است، سپس این مدل را در دادگان MIMIC-III آموزش داده شده‌است و در دادگان eIR تست کرده‌است.



شکل ۱-۲) معماری مدل کینگ لی و همکاران [۱۵]



شکل ۲-۲) گراف اطلاعات پزشکی استفاده شده بعنوان یک منبع مورد اعتماد [۱۵]

جدول ۲-۱) عبارتهای استفاده شده در مدل کینگ لی و همکاران

Notation	Definition
$x_i$	A quarter hour of illness severity score in ICU stay record
$R_i$	Each organ system feature
$a_{t_i,d}$	Record of the $t_i$ -th time ICU stay in the d-th organ system feature
$h_t$	The hidden state of Bi-LSTM
$\vec{h}_t$	The hidden state of forward LSTM network
$\overleftarrow{h}_t$	The hidden state of backward LSTM network
$\mathcal{G}$	The medical knowledge graph
$(b, r, e)$	Head entity – Relation – Tail entity
$A_i^{KG}$	The knowledge graph entity attention matrix
$A_i^{LSTM}$	The LSTM entity attention matrix

در جدول ۲-۱ عبارتهای استفاده شده در مدل این پژوهش و مفهوم آنها نشان داده شده است در این مدل، یک مکانیسم پیش آموزش با استفاده از شبکه LSTM<sup>۱</sup> دو طرفه، استفاده شده است این مکانیسم از گراف دانش پزشکی استفاده کرده است مدل این پژوهش از ۵ جزء اصلی تشکیل شده است:

<sup>۱</sup> Long Short-Term Memory

- پیش آموزش
- self-attention
- شبکه دو طرفه LSTM
- گراف دانش پزشکی
- Attention

جدول ۲-۲) مقایسه نتایج مدل‌های مختلف با مدل کینگ لی و همکاران در مجموعه دادگان MIMIC-III [۱۵]

	Models	F1
Traditional model	Support vector machine	68.93
	Random forest	71.53
	Decision tree	72.30
	Linear discriminant analysis	71.22
	XGboost	63.34
Neural model	GRU	83.05
	RNN	86.90
	PRNN	80.41
	RNN+ATT	89.90
<b>Our model</b>		<b>93.15</b>
- pretaining		91.66
- pretaining - Self-Attention		90.79
- pretaining - Self-Attention-Attention		89.17

جدول ۲-۳) مقایسه نتایج مدل‌های مختلف با مدل کینگ لی و همکاران در مجموعه دادگان eIR [۱۵]

Models	F1
GLM	82.02
XGboost	83.19
RNN	85.41
<b>Our model</b>	<b>88.32</b>
- pretaining	86.47
- pretaining - Self-Attention	84.95
- pretaining - Self-Attention-Attention	82.07

این مدل نه تنها، ویژگی‌های منحصر بفرد سیستم‌های اندام انسانی را با استفاده از دادگان سری زمانی EHR<sup>۱</sup> یاد می‌گیرد، بلکه می‌تواند از روابط زمانی بین سیستم‌های اندام انسانی استفاده کند. مهمترین

<sup>۱</sup> Electronic Health Record

ویژگی این مدل این است که می‌تواند، شرایط بیماران ICU را با استفاده از درجه SOFA، پیش‌بینی کند. همانطور که در تصاویر بالا مشخص است، این مدل در این پژوهش به  $F1\text{-score}$  ۱۵/۹۳٪ در مجموعه دادگان MIMIC-III دست یافته است، که نسبت به سایر مدل‌ها آورده شده در جدول ۲-۲ بالاتر است. ولی میزان این معیار با توجه به جدول ۳-۲ در دادگان eIR برای این مدل کاهش یافته است [۱۵].

در تحقیق اسچرپ و همکاران از روش شبکه عصبی بازگشتی<sup>۱</sup> برای پیش‌بینی شروع سپسیس استفاده شده است و عملکرد این روش با الگوریتم اینسایت<sup>۲</sup> [۱۶]. مقایسه شده است در این پژوهش از مجموعه دادگان MIMIC-III استفاده شده است مجموعه دادگان استفاده شده در این پژوهش، مربوط به بزرگسالان در بخش مراقبت‌های ویژه، که در زمان پذیرش هیچ علائمی نداشته‌اند است. در این مقاله از معیار سنجش AUROC<sup>۳</sup> استفاده شده است نتایج این پژوهش نشان می‌دهد که در بازه زمانی ۳ ساعت قبل از شروع سپسیس، AUROC شبکه به ۸۱٪، حساسیت ۹۰٪ و تشخیص پذیری ۴۷٪ است، در حالی که در الگوریتم insight این مقادیر به ترتیب ۷۲٪، ۳۱/۱٪ بوده است علاوه بر بازه زمانی ۳ ساعت برای نگاه به گذشته بیمار، بازه‌های زمانی ۶ و ۱۲ ساعت نیز بررسی شد. تحقیقات این پژوهش نشان می‌دهد که طول بازه ی زمانی نگاه به گذشته بیمار، در عملکرد کلاس بندی پیش‌بینی شروع سپسیس، تاثیر زیادی دارد. با توجه به نتایج این پژوهش، یک شبکه عصبی بازگشتی می‌تواند شروع سپسیس را با اطمینان پیش‌بینی کند. در جدول ۴-۲ خلاصه‌ای از مقایسه نتایج این پژوهش و الگوریتم اینسایت نشان داده شده است [۱۴].

---

<sup>1</sup> Recurrent Neural Network

<sup>2</sup> Insight

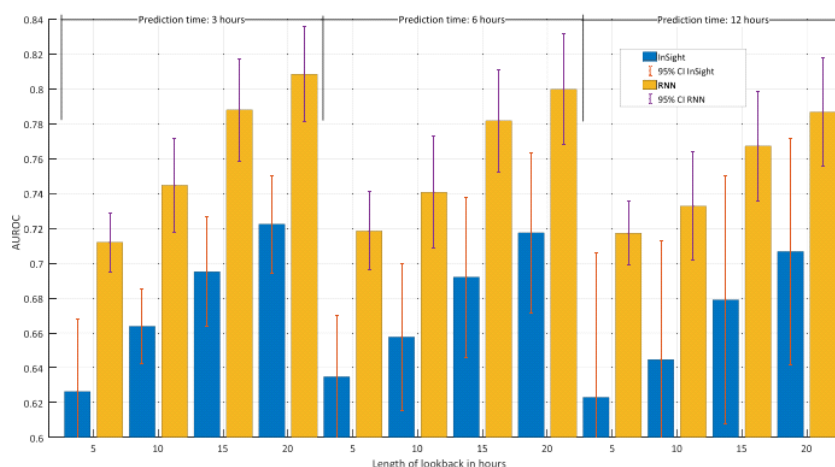
<sup>3</sup> Area Under the Receiver Operating Characteristic

جدول ۲-۴) مقایسه الگوریتم اینسایت با الگوریتم شبکه عصبی بازگشتی [۱۴]

Computed AUROCs		
PT	RNN (CI)	InSight (CI)
0 accepted interpolations		
3	0.81 (0.79, 0.83)	0.72 (0.71, 0.74)
6	0.79 (0.77, 0.82)	0.71 (0.70, 0.71)
12	0.76 (0.73, 0.79)	0.68 (0.66, 0.69)
5 accepted interpolations		
3	0.81 (0.78, 0.84)	0.72 (0.69, 0.75)
6	0.80 (0.79, 0.83)	0.72 (0.67, 0.76)
12	0.79 (0.76, 0.82)	0.71 (0.64, 0.77)

Computed Specificities in %		
PT	RNN (CI)	InSight (CI)
0 accepted interpolations		
3	46.9 (39.9, 53.9)	31.4 (30.4, 32.4)
6	45.3 (38.1, 52.6)	29.1 (26.6, 31.5)
12	38.8 (31.7, 45.8)	23.9 (21.6, 26.2)
5 accepted interpolations		
3	47.0 (43.1, 50.8)	31.1 (24.8, 37.5)
6	44.9 (35.3, 54.6)	32.5 (24.4, 40.6)
12	46.3 (40.5, 52.1)	34.1 (27.4, 40.7)



شکل ۲-۳) مقایسه الگوریتم اینسایت با شبکه عصبی بازگشتی در بازه‌های زمانی مختلف [۱۴]

با توجه به شکل ۲-۳، شبکه عصبی بازگشتی عملکرد بالایی نسبت به الگوریتم اینسایت دارد؛ چرا که شبکه عصبی بازگشتی از مزیت نگاه به گذشته سود می‌برد، اما با افزایش بازه‌زمانی پیش‌بینی، عملکرد هر دو مدل کاهش می‌یابد. با توجه به نتایج ۲۰ ساعت نگاه به گذشته این شکل می‌توان نتیجه گرفت که RNN<sup>۱</sup> از نظر معیار AUROC از الگوریتم اینسایت بهتر است [۱۴].

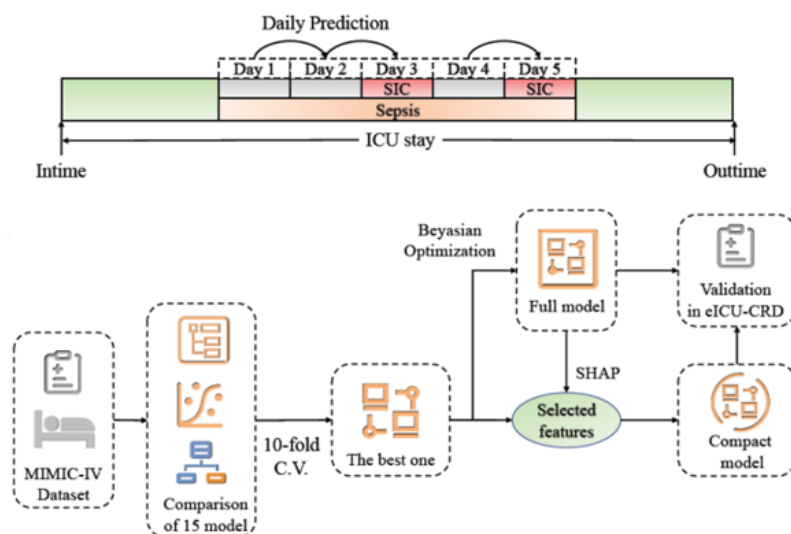
نعمتی و همکاران، یک الگوریتم هوش مصنوعی خبره سپسیس برای پیش‌بینی شروع سپسیس را مورد ارزیابی قرار دادند. آن‌ها از اطلاعات بیش از ۳۱۰۰۰ پذیرش در ICU از دو بیمارستان در مرحله آموزشی و از اطلاعات ۵۲۰۰۰ بیمار در ICU از دادگان MIMIC-III در مرحله آزمایشی استفاده کردند. آن‌ها بیمارانی که طبق تعریف سپسیس ۳ به سپسیس مبتلا بودند و یا کمتر از ۴ ساعت بعد از بستری شدن، علائم نشان دادند را حذف کردند؛ که در نتیجه ۲۷۰۰۰ بیمار را در مرحله آزمایشی و ۴۲۰۰۰ بیمار را در مرحله آموزشی مورد بررسی قرار دادند.

داده‌های سری زمانی مربوط به علائم حیاتی بیماران و داده‌های پرونده پزشکی الکترونیکی استخراج شد. مجموعه‌ای از ۶۵ متغیر، هر یک ساعت یک بار برای بیماران محاسبه شد و برای پیش‌بینی شروع سپسیس به الگوریتم داده شد، تا شروع سپسیس را در T ساعت بعد محاسبه کند؛ زمانی که T مقادیر ۴ و ۶ و ۸ و ۱۲ می‌گیرد. معیار AUROC در این ۴ بازه‌زمانی برای این الگوریتم، بین ۷۹٪ و ۸۸٪ است. این پژوهش به این نتیجه رسید که با افزایش طول بازه زمانی گذشته نگر، میزان AUROC و دقت و تشخیص پذیری کاهش می‌یابد [۱۷].

انعقاد خون ناشی از سپسیس<sup>۲</sup>، نشان دهنده افزایش مرگ و میر و سپسیس شدید است. در این پژوهش یک مدل یادگیری ماشین طراحی شده است آن‌ها از دادگان MIMIC IV، مربوط به بیماران مبتلا به سپسیس که سن آن‌ها بالای ۱۸ سال بوده است و بیشتر از ۲۴ ساعت در ICU بستری بوده اند، استفاده کرده‌اند.

<sup>۱</sup> Recurrent Neural Network (RNN)

<sup>۲</sup> Sepsis-Induced Coagulopathy (SIC)



شکل ۲-۴) طرح کلی پژوهش ژاوو و همکاران [۱۸]

شکل ۲-۴، طرح کلی این پژوهش را نشان داده‌است ابتدا از بیماران بصورت روزانه، ارزیابی لازم انجام شد. اگر فرد طبق اطلاعات بدست آمده در آن روز، به SIC مبتلا نباشد، خطر ابتلا به SIC در روز بعد، توسط مدل ما مشخص می‌شود و پیش‌بینی متوقف می‌شود. به محضی که SIC تشخیص داده شود و بیمار کاملاً درمان شود، مجدداً این روند شروع می‌شود.

با استفاده از روش k-fold cross validation عملکرد این ۱۵ مدل مقایسه شده‌است یکی از بهترین مدل‌ها که دقت و AUROC بیشتری داشت انتخاب شده‌است، سپس با استفاده از روش بهینه سازی بیزی<sup>۱</sup>، پارامترهای اصلی تشخیص داده‌شد. ۱۵ ویژگی بر اساس مقادیر SHAP<sup>۲</sup> و دسترس‌پذیری به داده‌های پزشکی، یک مدل مناسب بر اساس ۱۵ ویژگی انتخاب شده و پیاده‌سازی شده است و با استفاده از دیتا eICU-CRD<sup>۳</sup> ارزیابی می‌شود [۱۸].

از بین ویژگی‌های در دسترس، هشتاد و هشت پیش‌بینی کننده بالقوه، استخراج شد و ۱۵ مدل یادگیری ماشین مختلف، برای تعیین ریسک روزانه SIC بررسی شد و قوی ترین مدل بر اساس دو معیار دقت و AUROC انتخاب شد. بعد از انجام پیش پردازش‌های لازم، یک مدل مناسب بر اساس ۱۵ ویژگی انتخاب شد، که با توجه به اهمیت و دسترس پذیری این ویژگی‌ها توسعه داده شد. از ۱۱۳۶۲ نمونه در MIMIC-IV، ۵۹/۰ آنها یعنی ۶۷۴۴ بیمار در طول سپسیس، SIC داشتند. اطلاعات این ۶۷۴۴ بیمار، بصورت

<sup>۱</sup> Bayesian Optimization

<sup>۲</sup> Shapley Additive Explanations

<sup>۳</sup> Electronical ICU Collaborative Research Database

روزانه استخراج شد و در نهایت ۱۶۱۸۳ نمونه برای پردازش تعیین شد. مدلی به نام بوستینگ طبقه‌بندی شده<sup>۱</sup> دارای بزرگترین AUROC، یعنی ۰/۸۶ بود [۱۸]. در جدول ۲-۵ نتایج حاصل از پیاده‌سازی الگوریتم‌های مختلف نشان داده شده‌است.

جدول ۲-۵) مقایسه الگوریتم‌های مختلف در پژوهش ژآوو و همکاران [۱۸]

	Model	Accuracy	AUC
1	CatBoost Classifier	0.913 ( $\pm$ 0.004)	0.841 ( $\pm$ 0.025)
2	Light Gradient Boosting	0.912 ( $\pm$ 0.005)	0.835 ( $\pm$ 0.024)
3	Extreme Gradient Boosting	0.912 ( $\pm$ 0.004)	0.837 ( $\pm$ 0.025)
4	Gradient Boosting Classifier	0.911 ( $\pm$ 0.005)	0.832 ( $\pm$ 0.023)
5	Extra Trees Classifier	0.911 ( $\pm$ 0.002)	0.819 ( $\pm$ 0.032)
6	Random Forest Classifier	0.909 ( $\pm$ 0.002)	0.760 ( $\pm$ 0.022)
7	Ridge Classifier	0.908 ( $\pm$ 0.003)	0.753 ( $\pm$ 0.031)
8	Logistic Regression	0.908 ( $\pm$ 0.002)	0.746 ( $\pm$ 0.030)
9	K Neighbors Classifier	0.904 ( $\pm$ 0.001)	0.611 ( $\pm$ 0.040)
10	Ada Boost Classifier	0.902 ( $\pm$ 0.003)	0.804 ( $\pm$ 0.029)
11	Linear Discriminant Analysis	0.902 ( $\pm$ 0.003)	0.796 ( $\pm$ 0.027)
12	Multi-Level Perceptron	0.883 ( $\pm$ 0.004)	0.754 ( $\pm$ 0.022)
13	Decision Tree Classifier	0.861 ( $\pm$ 0.003)	0.593 ( $\pm$ 0.019)
14	SVM – RBF Kernel	0.859 ( $\pm$ 0.004)	0.777 ( $\pm$ 0.015)
15	Naive Bayes	0.805 ( $\pm$ 0.005)	0.756 ( $\pm$ 0.031)

بارتون و همکاران در یک پژوهش، یک ابزار یادگیری ماشین را برای تشخیص و پیش‌بینی سپسیس، ارزیابی کردند. آن‌ها عملکرد آن را با روش‌های موجود مقایسه کردند. در این پژوهش از داده‌های BIDMC<sup>۲</sup> و UCSF<sup>۳</sup> استفاده شده‌است آن‌ها ابتدا مدل را بر روی داده‌های UCSF آموزش داده‌اند و بر روی داده‌های BIDMC آزمایش کرده‌اند. برای ارزیابی مدل، از معیار AUROC، که یک معیار اساسی از دقت است، استفاده شده‌است این پژوهش، عملکرد الگوریتم‌های یادگیری ماشین را در زمان شروع سپسیس، ۲۴ و ۴۸ ساعت قبل از شروع سپسیس بررسی کرده‌اند. در جدول ۲-۶ نتایج الگوریتم‌های یادگیری ماشین پیشنهادی این پژوهش را با سایر سیستم‌های موجود مقایسه شده‌است با توجه به جدول ۲-۶، الگوریتم‌های یادگیری ماشین پیشنهادی این پژوهش، دقت بسیار بالایی نسبت به سایر روش‌های پیش‌بینی سپسیس دارد [۱۹].

<sup>1</sup> Categorical Boosting

<sup>2</sup> Beth Israel Deaconess Medical Center

<sup>3</sup> University Of California, San Francisco



جدول ۲-۶) مقایسه الگوریتم پیشنهادی بارتون و همکاران با سایر الگوریتم‌های یادگیری ماشین [۱۹]

	MLA (t = ۰)	MLA (t = -۱)	MLA (t = -۴۸)	SIRS (t = ۰)	MEWS (t = ۰)	SOFA (t = ۰)	qSOFA (t = ۰)
AUROC (SD)	۸۸/۰ (۰۰۸/۰)	۸۴/۰ (۰۴/۰)	۸۳/۰ (۰۴/۰)	۶۶/۰	۶۱/۰	۷۲/۰	۶۰/۰
Sensitivity	۸۰/۰	۸۰/۰	۸۴/۰	۷۰/۰	۵۲/۰	۷۸/۰	۳۷/۰
Specificity	۷۸/۰	۷۲/۰	۶۶/۰	۵۱/۰	۷۲/۰	۵۹/۰	۸۱/۰
DOR	۷۹/۱۴	۶۹/۱۳	۱۵/۱۳	۴۴/۲	۸۱/۲	۲۰/۵	۶۰/۲
LR +	۷۶/۳	۵۷/۲	۸۶/۲	۴۳/۱	۸۶/۱	۹۲/۱	۰۰/۲
LR -	۲۶/۰	۲۸/۰	۲۱/۰	۵۶/۰	۶۶/۰	۳۷/۰	۷۷/۰

DOR, diagnostic odds ratio; LR+, positive likelihood ratio; LR-, negative likelihood ratio; MEWS, modified early warning score

در پژوهش سباستین و همکاران یک مدل طبقه بندی خطی را ارزیابی شده است و آن را با استفاده از یک شبکه عصبی عمیق بهبود بخشیده است. این پژوهش بر روی پایگاه داده MIMIC-III انجام شده است و ویژگی‌ها بعنوان تاثیر گذارترین عامل در پیش‌بینی سپسیس بیان شده است برای تشخیص اینکه کدام ویژگی‌ها برای پیش‌بینی دقیق سپسیس، مهم هستند، در سه بازه زمانی ۳ و ۶ و ۱۲ ساعت، یک ویژگی را بعنوان ورودی به مدل داده است و از معیار گرادیان احتمال کلاس مثبت شبکه، استفاده کرده است گرادیان بیشتر نشان دهنده حساسیت بیشتر است، که بعنوان یک واسط برای پیش‌بینی سپسیس استفاده می‌شود. با توجه به نتایج این تحقیق، پارامترهایی که بیشترین تاثیر را در پیش‌بینی سپسیس دارند، شامل دما، WBC<sup>۱</sup>، ضربان قلب و تعداد است.

نتایج این پژوهش با پژوهش نعمتی و همکاران مقایسه شد است و با اینکه به نتایج یکسانی رسیده است، ولی از تعداد ویژگی‌های کمتری استفاده کرده است در پژوهش نعمتی و همکاران از ۶۵ ویژگی برای پیش‌بینی استفاده شده است؛ در حالی که در این پژوهش از ۸ ویژگی استفاده شده است در شکل زیر،

<sup>۱</sup> White Blood Cell

نتایج این دو پژوهش با هم مقایسه شده است [۲۰]. شکل ۲-۷ مقایسه نتایج حاصل از این پژوهش را با نتایج حاصل از الگوریتم کالورت و همکاران نشان داده است

جدول ۲-۷) مقایسه الگوریتم پژوهش سباستین و همکاران با الگوریتم کالورت [۱۶]

Sensitivity, specificity, AUPRC and AUROC of the method using a linear classification model compared to the deep neural network. Calculated area under curves show 95% lower-upper confidence intervals.

ahead					
<b>Linear (Calvert et al.)</b>	3	0.74	0.75	0.45 [0.451- 0.457]	0.82 [0.817- 0.820]
	6	0.73	0.72	0.43 [0.429- 0.436]	0.80 [0.796- 0.799]
	12	0.72	0.70	0.36 [0.360- 0.366]	0.77 [0.770- 0.773]
<b>Deep neural network</b>	3	0.80	0.76	0.49 [0.483- 0.490]	0.86 [0.855- 0.857]
	6	0.78	0.74	0.46 [0.454- 0.460]	0.83 [0.828- 0.831]
	12	0.74	0.72	0.38 [0.376- 0.382]	0.80 [0.797- 0.800]

کوک و همکاران به منظور رفع محدودیت‌هایی که برای تشخیص سپسیس در بیمارستان‌ها وجود دارد، یک ابزار تشخیص خودکار سپسیس پیشنهاد داده اند. آن‌ها یک شبکه عصبی عمیق DTCN<sup>۱</sup> برای پیش‌بینی سپسیس پیشنهاد داده اند. در این پژوهش از داده‌های دو بیمارستان استفاده شده است تعداد رکوردهای بیمارستان اول و دوم به ترتیب، ۱۷۹۰ و ۱۱۴۲ می‌باشد. در این پژوهش برای پرکردن داده‌های از دست رفته، از روش رگرسیون گوسین<sup>۲</sup> و برای استخراج ویژگی از تعریف سپسیس<sup>۳</sup> استفاده شده است نتایج این پژوهش نشان می‌دهد، می‌توان در هر بازه زمانی قبل از شروع سپسیس، با دقت AUROC به ترتیب، ۹۸ و ۰/۹۸ سپسیس را پیش‌بینی کرد [۲۱].

فاگروستروم و همکاران برای تشخیص سریع شوک سپتیک، از یک نوع شبکه عصبی LSTM بنام LiSep LSTM استفاده کرده‌اند. این نوع شبکه عصبی زمانی کاربردی و مناسب است، که ویژگی‌های در دسترس بصورت سری زمانی باشند. آن‌ها از اطلاعات پزشکی ۵۹۰۰۰ بیمار در پایگاه داده MIMIC-II استفاده

<sup>۱</sup> Deep Temporal Convolution Network

<sup>۲</sup> Gaussian Process Regression

کرده‌اند. نتایج نشان می‌دهد، که این نوع شبکه عصبی می‌تواند با مقدار AUROC، ۰/۸۳ تا ۴۰ ساعت قبل از شروع سپسیس را پیش‌بینی کند. نتایج این مدل شبکه عصبی با پنج الگوریتم پرکاربرد شبکه عصبی، مقایسه شده‌است.

جدول ۸-۲) مقایسه الگوریتم پیشنهادی فاگستروم و همکاران با سایر الگوریتم‌ها [۲۲]

Model Number	AUROC (95% CI)	HBO Median (IQR)
LiSep LSTM	0.83 (0.82, 0.84)	48 (20.0, 135.0)
TREWScore	0.83 (0.81, 0.85)	28.2 (10.6, 94.2)
InSight	0.83 (0.80, 0.86)	<3* (N/A)
Multitask LSTM	0.85 (N/A)	N/A (N/A)
SepLSTM	0.93 (N/A)	<3* (N/A)
Liu <i>et al.</i> pre-shock RNN	0.93 (N/A)	7.0 (N/A)

جدول ۸-۲، نتایج این الگوریتم را با سایر الگوریتم‌های یادگیری ماشین مقایسه می‌کند. در این پژوهش نتایج الگوریتم Sep LSTM و نتایج پژوهش ردیف آخر، یکسان می‌باشد؛ اما مسئله‌ای که باعث برتری LSTM Sep می‌باشد مقدار HBO<sup>۱</sup> مربوط به هر کدام از مدل‌ها است. همچنین نتایج مدل TREWScore و LiSep LSTM تقریباً یکسان هستند؛ اما مدل ما از ویژگی‌های کمتری استفاده می‌کند و سریعتر می‌تواند سپسیس را پیش‌بینی کند [۲۲].

یکی از عملکردهای مراقبتی رایج در ICU، آزمایش کشت خون می‌باشد. این آزمایش به منظور تشخیص زودهنگام عفونت خون انجام می‌شود. چون عفونت خون موجب سپسیس و سپسیس موجب نارسایی اندام و حتی مرگ می‌شود، بنابراین تشخیص زودهنگام آن یک مسئله چالش برانگیز است. از آنجایی که دریافت نتیجه آزمایش کشت خون به زمان نیاز دارد، پیش‌بینی نتایج آزمایش کشت خون از اهمیت بالایی برخوردار است. در تحقیق استینکیستی و همکاران، یک روش جدید برای پیش‌بینی نتیجه آزمایش کشت خون بررسی شده‌است این روش مبتنی بر یک مدل محاسباتی است؛ که از ۹ پارامتر بالینی اندازه‌گیری شده در طول زمان، استفاده می‌کند.

<sup>۱</sup> Hour Before Onset

در این پژوهش از یک شبکه عصبی LSTM دو طرفه<sup>۱</sup> استفاده شده است آن‌ها از یک پایگاه داده با ۲۱۷۷ پذیرش در ICU بیمارستانی در بلژیک استفاده کردند. نتایج روش آن‌ها به مقدار AUROC، ۰/۸۲ دست یافته است. این پژوهش نشان داد که با کاهش کمی از AUROC می‌توان، زمان پیش‌بینی را افزایش داد. همچنین این مدل بطور دقیق، نتیجه آزمایش‌های کشت خون را پیش‌بینی می‌کند؛ لذا می‌توان از آن در ICU به عنوان یک سیستم هشدار اولیه برای تشخیص بیماران در معرض خطر عفونت خون استفاده کرد [۲۳].

لوریتسن و همکاران، سه مدل مختلف را برای تشخیص زودهنگام سپسیس مقایسه کرده‌اند. مدل اول مورد بررسی در این پژوهش، یک مدل GB-vital<sup>۲</sup> است. در این مدل از ویژگی‌های علائم حیاتی برای پیش‌بینی استفاده شده است مدل دوم یک مدل MLP<sup>۳</sup> غیرمتوالی، با تعداد بسیار زیاد ویژگی است. مدل سوم یک CNN-LSTM است. ویژگی‌های استفاده شده در مدل اول و سوم، بطور یکسان انتخاب شده‌اند. نتایج این پژوهش نشان می‌دهد، که مدل اول سه ساعت قبل از شروع سپسیس با AUROC، ۰/۷۸۹ می‌تواند سپسیس را پیش‌بینی کند. در همین زمان برای مدل دوم و سوم به ترتیب با AUROC، ۰/۸۲۵ و ۰/۸۵۵ می‌توان سپسیس را پیش‌بینی کرد. با توجه به نتایج بدست آمده، نمی‌توان دقیقاً مشخص کرد، که در این مسئله چه روشی بهتر است؛ چرا که با کاهش مقدار کمی از معیارهای ارزیابی می‌توان، زمان پیش‌بینی سپسیس را افزایش داد و اینکه کدام مدل بعنوان مدل بهتر انتخاب شود بستگی به موضوع‌های مورد توجه در این مسئله دارد. در این پژوهش مدل سوم، عملکرد بهتری داشته است [۲۴].

همانطور که گفته شد، پژوهش‌های انجام شده در این زمینه بسیار زیاد است؛ لذا توضیح کامل همه ی آن‌ها در این پژوهش نمی‌گنجد. در جدول ۲-۹ بعضی از پیشنهادها تحقیق مربوط به این پژوهش، بصورت مختصر آورده شده است

---

<sup>۱</sup> Bidirectional Long Short-Term Memory Neural Network

<sup>۲</sup> Gradient Boosting

<sup>۳</sup> Multi Layer Perceptron

جدول ۲-۹) خلاصه سایر پیشینه‌های تحقیق

Authors	Title	Dataset	Methodology	Predictors	Outcome	Sepsis definition
Calvert et al. [۱۶]	A computational approach to early sepsis detection	MIMIC II	InSight	Systemic Inflammatory Response Syndrome (SIRS)	prediction of sepsis up to three hours prior to a patient's first five-hour Systemic Inflammatory Response Syndrome (SIRS) episode	I/II
Mao et al. [۲۵]	Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department	Dataset from University of San Francisco (UCSF)	machine-learning algorithm with gradient tree boosting	six vital sign measurements	detection of sepsis and severe sepsis	I/II
Kam et al. [۲۶]	Learning representations for the early detection of sepsis with deep neural networks	MIMIC II	InSight	Systemic Inflammatory Response Syndrome (SIRS)	early stage of sepsis	I/II
Desautels et al. [۲۷]	Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach	MIMIC-III	InSight	vitals, peripheral capillary oxygen saturation, Glasgow Coma Score, and age	validate a sepsis prediction method	III
Shashikumar et al. [۲۸]	Early sepsis detection in critical care patients using multiscale blood pressure and heart rate dynamics	Emory affiliated hospital	multivariate modeling approach	high-resolution blood pressure (BP) and heart rate (HR) times series dynamics	early prediction of sepsis	III
Masson, et al. [۲۹]	Presepsin (soluble CD۱۴ subtype) and procalcitonin levels for mortality prediction in sepsis: data from the Albumin Italian Outcome Sepsis trial	A multicentre, randomised Albumin Italian Outcome Sepsis trial, ۱۰۰ patients	Cox regression model	Presepsin level, procalcitonin level and some covariates	۲۸-day/ICU/۹۰-day mortality	II
Adrie et al. [۳۰]	Model for predicting short-term mortality of severe sepsis	A multicentre database including data from ۱۲ ICUs, ۲۲۶۸ patients	Generalised linear model	SAPS II and LOD scores at ICU admission, septic shock, multiple organ failure, comorbidities, procedures, agents, bacteraemia and sources of infection	۱۴-day mortality within ICU stay	II

جدول ۲-۹) خلاصه سایر پیشینه‌های تحقیق (ادامه)

Ripoll, V.J.R. et al. [۳۱]	Sepsis mortality prediction with the Quotient Basis Kernel	MIMIC II	Support vector machines (SVMs), LR, SAPS	SOFA and SAPS scores at ICU admission	ICU mortality	II
Fang W-F et al. [۳۲]	Development and validation of immune dysfunction score to predict ۲۸-day mortality of sepsis patients	Sepsis patients admitted to ICU at a hospital in Taiwan, ۱۵۱ patients	LR	Monocyte HLA-DR* expression, plasma G-CSF* level, plasma IL*-۱۰ level, and serum SeMo* ratio	۲۸-day mortality	III
Xie, Y. et al. [۳۳]	Using clinical features and biomarkers to predict ۶۰-day mortality of sepsis patients	Protocol-based care in early septic shock trial, around ۵۳۰ patients	LR	Clinical features and biomarkers obtained during the first ۲۴ h of hospital admission	۶۰-day mortality	Not mentioned
Poucke, S.V. et al. [۳۴]	Scalable predictive analysis in critically ill patients using a visual open data analysis platform	MIMIC II	Naïve Bayes, LR, RF, AdaBoost, Bagging, Stacking, SVM	Demographics, comorbidities, types of care unit, platelet count	ICU mortality	NA
Zhang, Z. & Hong, Y [۳۵]	Development of a novel score for the prediction of hospital mortality in patients with severe sepsis: the use of electronic healthcare records with LASSO regression	MIMIC-III	LASSO, LR	Demographics, clinical and laboratory variables recorded during the first ۲۴ h in ICU	Hospital mortality	II
aylor, R.A. et al. [۳۶]	Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach	Adult ED* visits over ۱۲ months, ۴۶۷۶ patients	RF, CART, LR	Demographics, previous health status, ED health status, ED services rendered and operational details	Hospital mortality	II
Pregernig, A. et al. [۳۷]	Prediction of mortality in adult patients with sepsis using six biomarkers: a systematic review and meta-analysis	۴۴ articles in English	Qualitative analysis, meta-analysis	Angiopietin ۱ and ۲, HMGB۱, sRAGE, sTREM-۱, suPAR	۲۸-day/۳۰-day/ICU/hospital/۹۰-day mortality	

## فصل ۳- آماده‌سازی داده‌ها

در این تحقیق از نسخه ۱/۴ پایگاه داده MIMIC-III استفاده شده است. داده‌های MIMIC-III شامل اطلاعات پزشکی بیش از ۴۰ هزار بیمار در بخش مراقبت‌های ویژه می‌باشد. مجموعه دادگان MIMIC-III در یک دوره ۱۱ ساله، از سال ۲۰۰۱ و از مرکز پزشکی<sup>۱</sup> BIDMC جمع‌آوری شده است.

### ۳-۱- معرفی پایگاه داده MIMIC-III

MIMIC-III یک پایگاه داده بزرگ با دسترسی آزاد مربوط به بیماران بستری شده در بخش مراقبت‌های ویژه است. این پایگاه داده شامل اطلاعاتی مانند ویژگی‌های دموگرافیکی، علائم حیاتی، نتایج آزمایشگاهی، مشاهدات و یادداشت‌های ارائه شده توسط مسئولین مراقبتی، مراقبت‌های انجام شده، گزارش‌های تصویربرداری، مرگ و میر، زمان ترخیص از بیمارستان و طول مدت اقامت بیمار است. دسترسی آزاد این پایگاه داده باعث شده است، که این پایگاه داده بسیار بهبود پیدا کند. با این حال برای دسترسی به این پایگاه داده، باید آزمونی به منظور تایید هویت محقق و ارتباط صحیح موضوع او داده شود. داده‌های این پایگاه داده در طول مراقبت‌های معمول بیمارستانی بدست آمده است؛ بنابراین هیچ تداخلی با گردش کار مراقبین سلامت نداشته است.

مجموعه دادگان MIMIC-III از بخش‌های مختلف، شامل اطلاعات مربوط به سیستم‌های مراقبت‌های ویژه، اطلاعات پرونده‌های الکترونیکی سلامت<sup>۲</sup> و اطلاعات محرمانه مرگ و میر سازمان تامین اجتماعی جمع‌آوری شده است

برای جمع‌آوری اطلاعاتی که از سیستم‌های موجود در بخش مراقبت‌های ویژه بدست می‌آید، از دو سیستم مجزا استفاده می‌شود؛ یکی سیستم اطلاعات بالینی Philips CareVue و دیگری iMDsoft MetaVision ICU می‌باشد. برای تمایز داده‌های ذخیره شده توسط هر کدام از این سیستم‌ها، به ترتیب از پسوند‌های CV و MV استفاده شده است.

این سیستم منابعی برای داده‌های بالینی، مانند اندازه‌گیری‌های فیزیولوژیکی تایید شده توسط پرستار (بعنوان مثال مستندات ساعتی ضربان قلب فشار خون شریانی یا تعداد تنفس، گزارش‌های روند درمانی بیمار ارائه شده توسط مراقبان سلامت و داروهای مصرفی بیماران در طول بستری در بیمارستان) است. سایر اطلاعات بیمار، از جمله ویژگی‌های دموگرافیکی و مرگ و میر داخل بیمارستانی، نتایج آزمایشگاه‌ها،

<sup>1</sup> Beth Israel Deaconess Medical Center

<sup>2</sup> EMR (Electronic Medical Record)



خلاصه تریخیص و اطلاعات تصویر برداری بیمار و اطلاعات بین المللی، مانند کدهای ICD-9<sup>۱</sup>، DRG<sup>۲</sup> و CPT<sup>۳</sup> از سیستم پرونده سلامت بیمار و آزمایشگاه جمع آوری شده است.

محتوای پایگاه داده MIMIC-III از داده های واقعی بدست آمده است؛ بنابراین این داده ها حاوی اطلاعات شخصی افراد نیز می باشد؛ بنابراین قبل از اینکه داده ها در پایگاه داده MIMIC-III گنجانده شوند یک سری تغییرات روی آن ها اعمال شده است این تغییرات به منظور غیرقابل شناسایی کردن<sup>۴</sup> داده ها انجام شده است این تغییرات طبق استاندارد HIPAA<sup>۵</sup> اعمال شده است تغییرات اجرا شده روی این پایگاه داده، به شرح زیر است:

۱. اطلاعات شخصی بیمار مانند نام و نام خانوادگی، شماره تلفن و آدرس حذف شده است.
  ۲. تمام تاریخ ها به تاریخ هایی در آینده تبدیل شده اند. مثلا "۰۵:۰۰:۰۰-۲۱۱۵" تاریخ و زمان بستری یک بیمار در بیمارستان است. ماه و روز به همان صورت اصلی حفظ شده و سال بستری تبدیل به سالی در آینده بین سال ۲۱۰۰ و ۲۲۰۰ شده است.
  ۳. برای هر بیمار، تمام تاریخ های ICU بصورت یکسانی تبدیل شده اند؛ تا فواصل زمانی حفظ شوند.
  ۴. روزهای هفته و فصل های سال به همان صورت باقی مانده اند.
  ۵. بیمارانی که در طول بستری ۹۰ ساله شدند، از پایگاه داده حذف شدند.
  ۶. بیمارانی که بیش از ۸۹ سال سن دارند، در اولین بستری، سنشان با مقدار بیشتر از ۳۰۰ سال ذخیره شده است یکی از علت های چنین تبدیلی ممکن است این باشد، که برای بیماران در این بازه سنی، با کنار هم قرار دادن اطلاعاتی که برای آن ها ذخیره شده است، احتمال بیشتری دارد که شناسایی شوند. در واقع، این تبدیل احتمالا برای محفوظ ماندن هویت این افراد انجام شده است.
  ۷. از آنجایی که تبدیل تاریخ ها بصورت تصادفی انجام شده است، تحقیقاتی که در آن سال بستری مورد نیاز است، با این داده ها امکان پذیر نیست [۳۸].
- یک بیمار ممکن است، چندین بار در تاریخ های مختلف بستری شود. از آنجایی که اطلاعات بیمار در پایگاه داده، باید با کد<sup>۶</sup> شناساگر به هم مرتبط شوند، سه کد شناساگر برای هر بیمار برای داده های بالینی تعریف می شوند.

<sup>1</sup> Current Procedural Terminology (CPT) codes, 9th Edition

<sup>2</sup> Diagnosis-Related Group (DRG) codes,

<sup>3</sup> Current Procedural Terminology (CPT) codes

<sup>4</sup> Deidentification

<sup>5</sup> Health Insurance Portability and Accountability Act

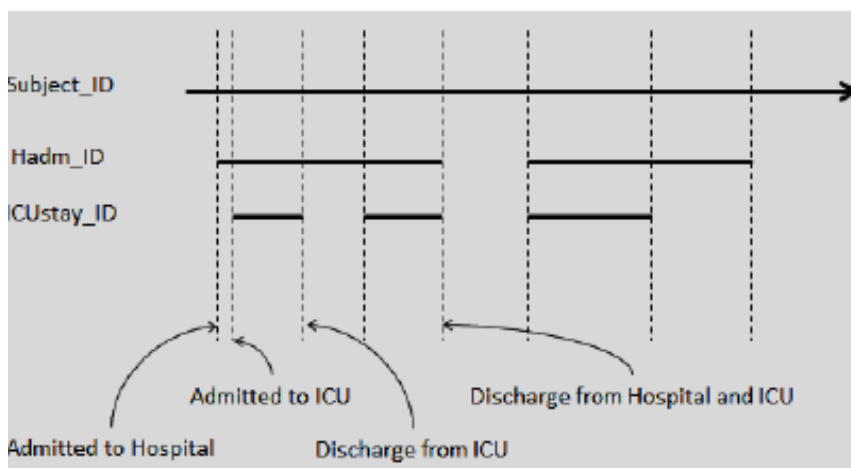
<sup>6</sup> Code

۱- کد شناساگر subject\_id : این کد برای شناسایی یک بیمار است و در واقع همان شماره رکورد پزشکی است.

۲- کد شناساگر hadm\_id : این کد برای مشخص کردن یک بستری بیمار در بیمارستان است. هر بیمار ممکن است چندین hadm\_id داشته باشد، که مرتبط با subject\_id آن بیمار است.

۳- کد شناساگر icustay\_id : به ازای هر دوره ی بستری در ICU، که بیمار تحت مراقبت‌های ویژه قرار می‌گیرد، یک کد شناساگر برای آن دوره ی بستری بیمار تعریف می‌شود. هر بیمار ممکن است چندین icustay\_id داشته باشد.

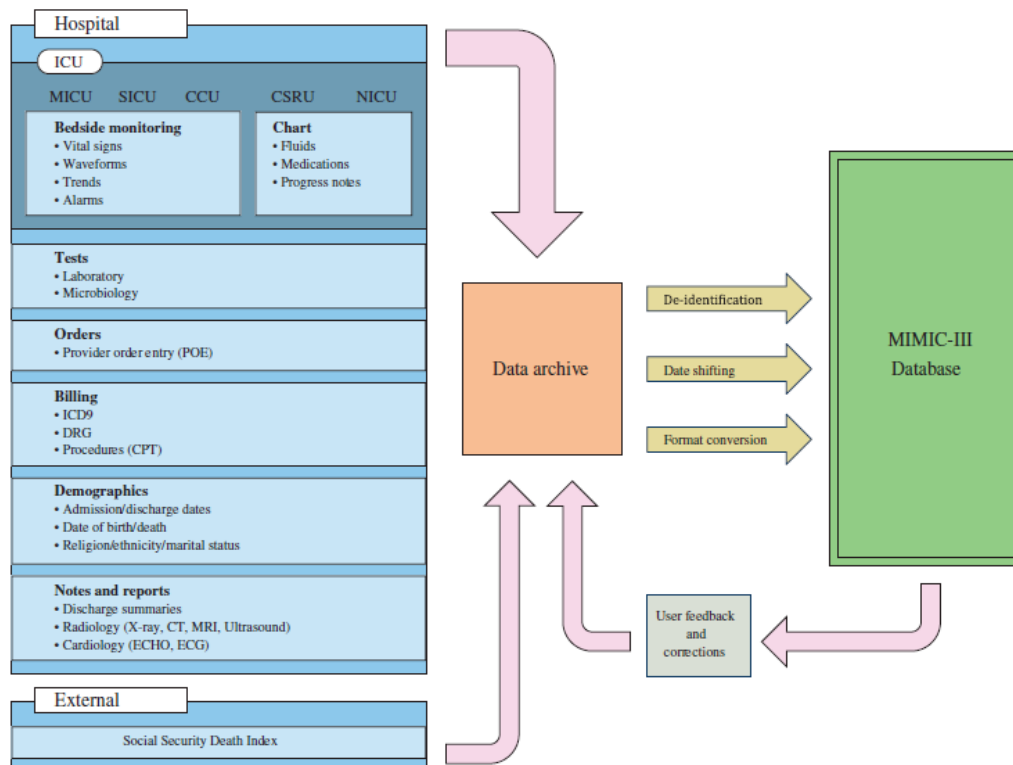
برای بهتر مشخص شدن این سه کد، نمودار زمانی یک بیمار فرضی در شکل ۱-۳ نشان داده شده است. در این شکل همانطور که مشخص است، یک بیمار با یک subject\_id ممکن است چندین hadm\_id داشته باشد. به ازای هر کدام از hadm\_id ها، یک یا چند icustay\_id وجود دارد، که بازه زمانی آن‌ها از بازه بستری در بیمارستان خارج نمی‌شود. داده‌ای که با مجوز دسترسی، برای استفاده در اختیار ما قرار گرفت، شامل اطلاعات ۳۸۵۹۷ بیمار و ۴۹۷۸۵ پذیرش است [۳۸].



شکل ۱-۳) نمودار زمانی یک بیمار در پایگاه داده MIMIC-III و کدهای شناساگر مربوطه [۳۸]

جدول ۳-۱) انواع مختلف داده در پایگاه داده MIMIC-III [۳۸]

Design Type(s)	Data integration objective
Measurement Type(s)	Demographics, clinical measurement, intervention, Billing, Medical History Dictionary, Pharmacotherapy, clinical laboratory test, medical data
Technology Type(s)	Electronic Medical record, Medical Record, Electronic Billing System, medical Coding Process Document, Free Text Format
Factor Type(s)	
Sample Characteristic(s)	Homo sapiens



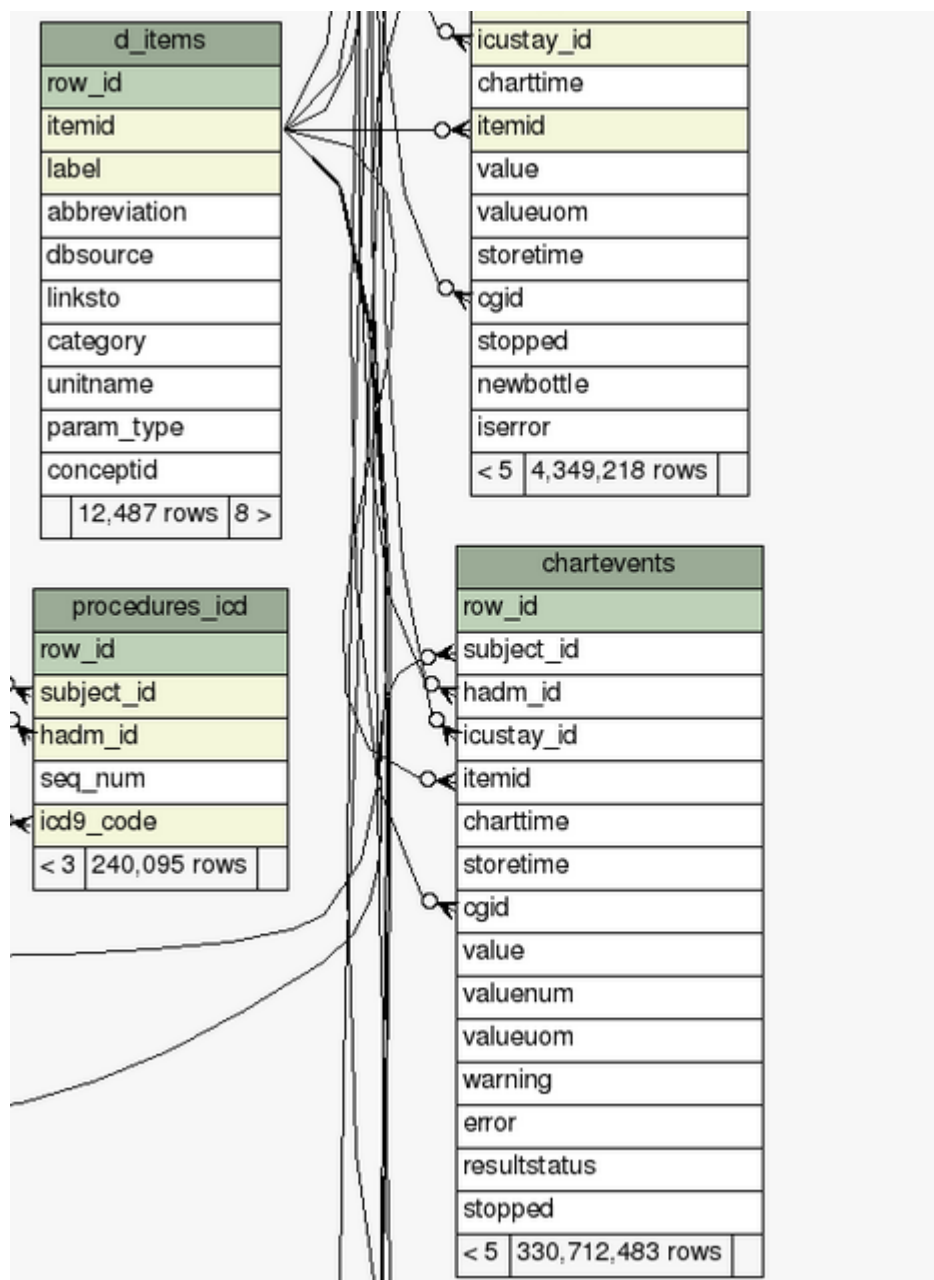
شکل ۳-۲) روند استخراج داده از پایگاه داده MIMIC-III [۳۸]

پایگاه داده MIMIC-III، یک پایگاه داده رابطه‌ای با ۲۶ جدول است. در جدول‌هایی که برای استفاده از پایگاه داده MIMIC-III، برای دانلود در دسترس قرار می‌گیرد، اطلاعات بیمار در جداول مختلف ثبت شده است. جدول ۱-۳ اطلاعات موجود در این پایگاه داده را به تفکیک نوع آن‌ها نشان می‌دهد. این اطلاعات از طریق کدهای شناساگر شکل ۱-۳، به هم مرتبط می‌شوند. برای مثال، اطلاعات مربوط به پذیرش بیماران در جدول ADMITTION و اطلاعات مربوط به ترخیص بیمار در جدول CALLOUT آمده است. با استفاده از جداولی که از این پایگاه داده در دسترس قرار گرفته است، می‌توان پایگاه داده رابطه‌ای مربوط به بخش داده‌های بالینی MIMIC-III را در سیستم مدیریت پایگاه داده‌ی PostgreSQL ایجاد کرد.

نمودار ارتباط موجودیت‌های این پایگاه داده و توضیح مختصری در مورد جداول این پایگاه داده، در ضمیمه آورده شده است. در اینجا، تنها قسمت‌هایی از این پایگاه داده که در این تحقیق مورد استفاده قرار گرفته اند توضیح داده می‌شود.

در پایگاه داده‌ی رابطه‌ای MIMIC-III برای هر کدام از موارد پزشکی مختلف (مثل آزمایش‌ها و داروها) دو نوع جدول وجود دارد. یک جدول که شامل آیتم‌ها و کدشناساگر مربوط به آن‌ها و احیاناً توضیحی در مورد آیتم‌ها است. جدول دیگر شامل مقادیر اندازه‌گیری شده برای این آیتم‌ها، برای بیماران مختلف در زمانهای مختلف است.

در شکل ۳-۳، جدول chartevents نشان داده شده است. در این جدول اطلاعات مراقبی بیماران که در ICU ثبت می‌شود. در جدول d\_items به ازای پارامترهایی که در ICU اندازه‌گیری می‌شوند، کدهایی تخصیص داده شده است. جدول chartevents با جدول d\_items به وسیله‌ی این کدها مرتبط می‌شود. در جدول chartevents یک مقدار بنام chartitem وجود دارد، که با استفاده از این مقدار، می‌توان مقدار هر کدام از ویژگی‌های علائم حیاتی بیمار را در ساعت مشخص بدست آورد، در واقع این ویژگی، زمان اندازه‌گیری یک ویژگی یا زمان رخداد یک موضوع را بیان می‌کند. ما از این جدول برای بدست آوردن سری زمانی از مقادیر ویژگی‌های علائم حیاتی بیمار و مقادیر لازم برای اندازه‌گیری SOFA، در بازه‌های زمانی یک ساعته استفاده شده است.



شکل ۳-۳ شمای جدول chartevents در پایگاه داده MIMIC-III

در این پژوهش برای بدست آوردن مقدار کلاس یک سری زمانی، از علائم حیاتی از معیار نمره دهی SOFA استفاده شده است برای بدست آوردن مقدار SOFA، به اندازه گیری بعضی از ویژگی های علائم حیاتی و مقادیر آزمایشگاهی نیاز داریم. نحوه بدست آوردن این کلاس، بطور کامل، در بخش بعدی این فصل توضیح داده شده است برای بدست آوردن مقادیر مربوط به

مقیاس کمای گلاسکو<sup>۱</sup>، فشار متوسط شریانی<sup>۲</sup>، میزان اکسیژن<sup>۳</sup> داده شده به بیمار توسط ونتیلاتور<sup>۴</sup>، وضعیت استفاده از ونتیلاتور<sup>۵</sup> از جدول chartevents، برای مقادیر میزان کراتینین<sup>۶</sup>، فشار نسبی اکسیژن در خون شریانی<sup>۷</sup>، میزان بیلی روبین<sup>۸</sup>، میزان پلاکت خون<sup>۹</sup> از جدول labevents، میزان دوپامین<sup>۱۰</sup>، میزان دوبوتامین<sup>۱۱</sup>، میزان اپی نفرین<sup>۱۲</sup>، میزان نور اپی نفرین<sup>۱۳</sup> از جدول های inputevents\_mv و inputevents\_cv و برای بدست آوردن میزان خروجی ادرار<sup>۱۴</sup> از جدول outputevents استفاده می شود. در شکل های ۳-۴، ۳-۵، ۳-۶ و ۳-۷ جداول مورد استفاده در این اندازه گیری همراه با فیلدهای آن ها نشان داده شده است.

---

<sup>1</sup> GCS (Glasgow Coma Scale)

<sup>2</sup> MAP (Mean Arterial Pressures)

<sup>3</sup> FIO2( Fractional Inspired Oxygen)

<sup>4</sup> Ventilator

<sup>5</sup> Ventilation status

<sup>6</sup> Creatinine

<sup>7</sup> PAO2( Partial Pressure of Oxygen in Arterial Blood)

<sup>8</sup> Bilirubin

<sup>9</sup> Platelets

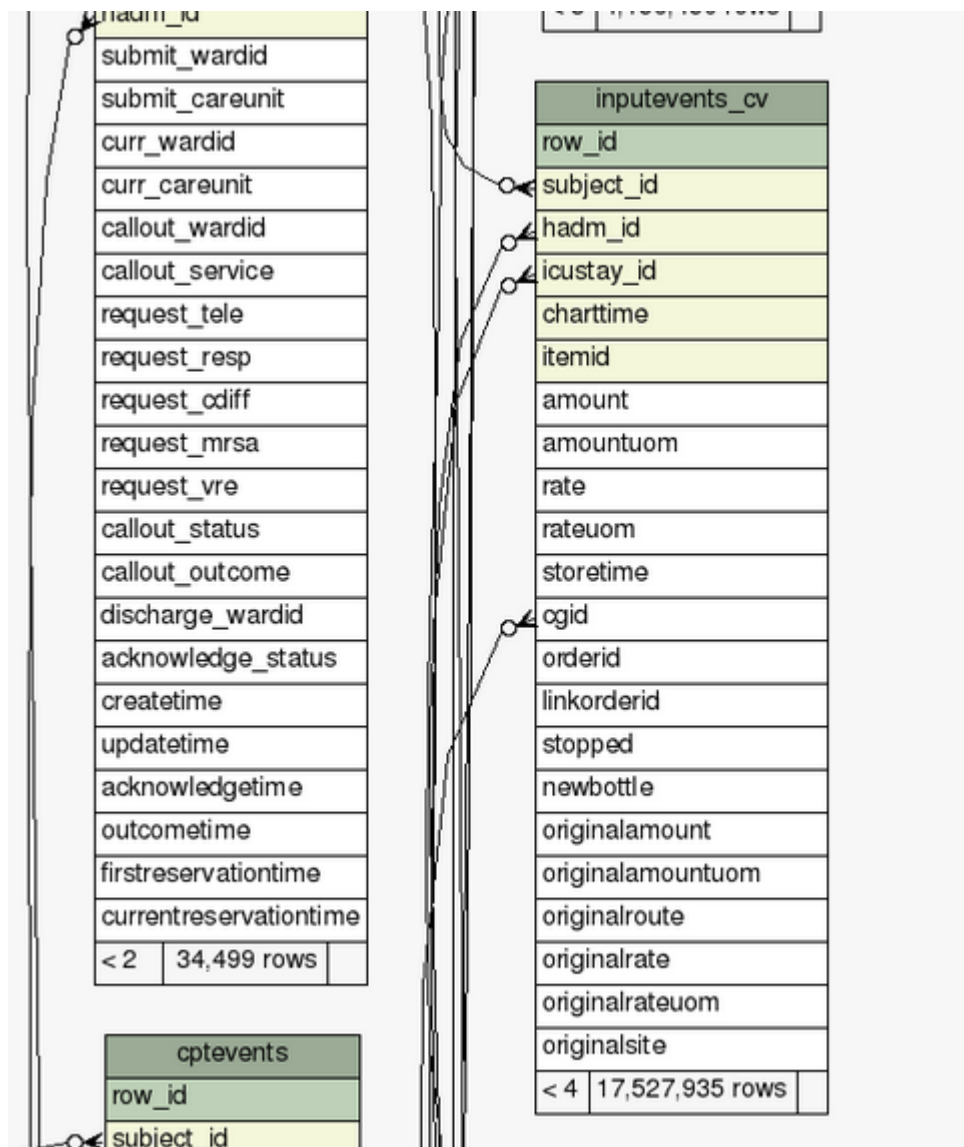
<sup>10</sup> Dopamine

<sup>11</sup> Dobutamine

<sup>12</sup> Epinephrine

<sup>13</sup> Norepinephrine

<sup>14</sup> Urine output

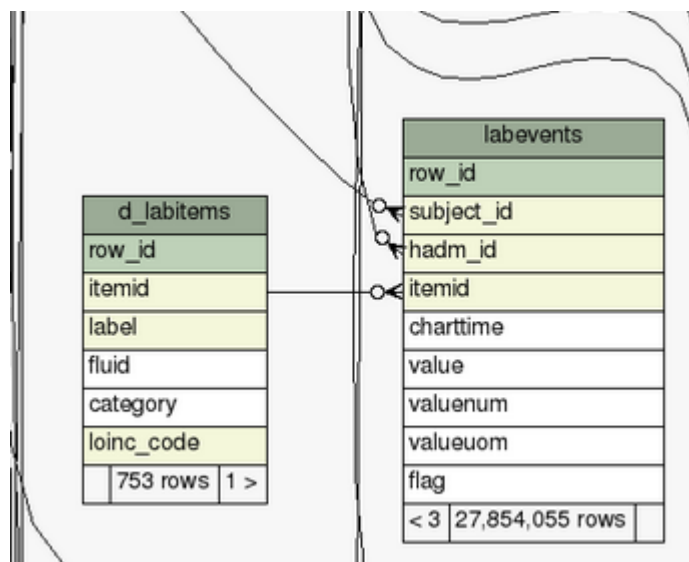


شکل ۳-۴) شمای جدول inputevents\_cv در پایگاه داده MIMIC-III

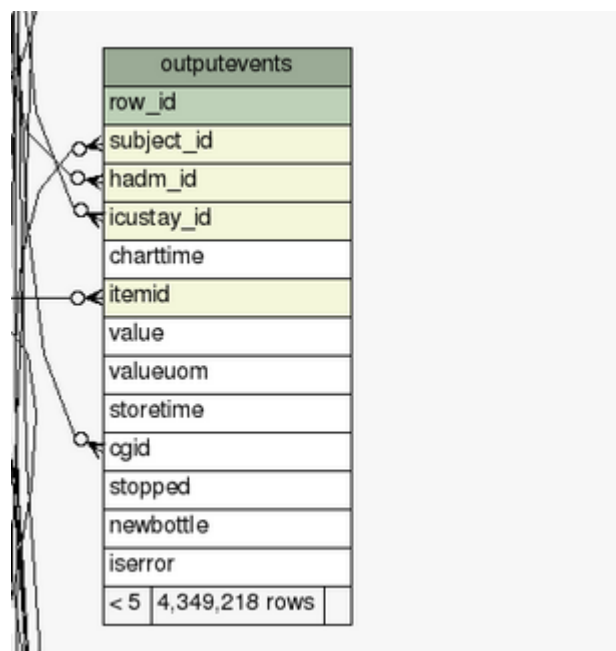
inputevents_mv		
row_id		
subject_id		
hadm_id		
icustay_id		
starttime		
endtime		
itemid		
amount		
amountuom		
rate		
rateuom		
storetime		
cgid		
orderid		
linkorderid		
ordercategoryname		
secondaryordercategoryname		
ordercomponenttypedescription		
ordercategorydescription		
patientweight		
totalamount		
totalamountuom		
isopenbag		
continueinnextdept		
cancelreason		
statusdescription		
comments_editedby		
comments_canceledby		
comments_date		
originalamount		
originalrate		
< 5	3,618,991 rows	

شکل ۳-۵) شمای جدول inputevents\_mv در پایگاه داده MIMIC-III





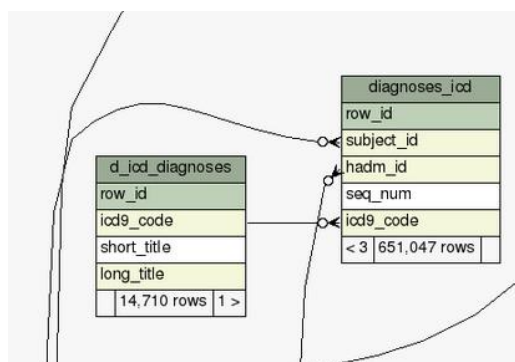
شکل ۳-۶) ارتباط جدول labevents و d\_labitems در پایگاه داده MIMIC-III



شکل ۳-۷) شمای جدول outputevents در پایگاه داده MIMIC-III

برای بدست آوردن کلاس آن دسته از سری‌های زمانی که تشخیص کلاس آن‌ها از طریق مقدار SOFA امکان پذیر نبود، از کد ICD-9 استفاده شده‌است در جدول d\_icd\_diagnoses برای هر تشخیص، یک کد در نظر گرفته شده‌است. از این مقادیر در جدول diagnoses\_icd استفاده شده‌است؛ بطوری که برای

هر بستری بیمار در بیمارستان، یک یا چند تشخیص ثبت می‌شود. نحوه استخراج مقادیر کلاس سری زمانی‌های خواسته شده با استفاده از این جدول، در بخش بعدی همین فصل توضیح داده خواهد شد. در شکل ۳-۸ نحوه ارتباط این دو جدول نشان داده شده است.



شکل ۳-۸) ارتباط جدول diagnoses\_icd و d\_icd\_diagnoses در پایگاه داده MIMIC-III

### ۳-۲- مشکلات پایگاه داده MIMIC-III

با وجود پیچیدگی و انبوه بودن داده‌های بخش ICU، در این پایگاه داده سعی شده است، که تا حد امکان خوانایی وجود داشته باشد و بتوان با پرس و جوهای مناسب، داده‌های مورد نظر را به دست آورد. اما با این وجود، با توجه به این نکته که MIMIC-III از داده‌های واقعی جمع‌آوری شده است و محیطی که، داده‌ها جمع‌آوری شده یک محیط تحت کنترل برای پژوهش نبوده است؛ بنابراین در داده‌ها مشکلاتی وجود دارد [۳]: پنج مسأله مهم در این داده‌ها به شرح زیر است:

۱. مقادیر نامعلوم<sup>۱</sup> در داده‌ها وجود دارد. بین هر تاریخ بستری و ترخیص این احتمال وجود دارد، که بعضی از داده‌ها از دست بروند. دلایل ایجاد مقادیر نامعلوم میتواند این باشد، که بیمار برای گرفتن اسکن یا جابجایی الکترودها از دستگاه جدا شود یا سیستمی که اطلاعات را جمع‌آوری می‌کند، دچار مشکل شود و یا اینکه داده‌های جمع‌آوری شده خراب شوند [۳۹].
۲. در داده‌ها ممکن است، تناقض وجود داشته باشد. علت آن وجود اختلال در داده‌ها، اشتباهات تنظیم زمان و بازنویسی مقادیر توسط افراد با مقادیری که به نظر آن‌ها صحیح است، می‌باشد. به همین جهت، همیشه نمی‌توان به این داده‌ها اطمینان کرد. بنابراین برای استفاده از این داده‌ها در طراحی سیستم‌های تصمیم‌یار، باید هوشیار بود و با احتیاط بیشتری داده‌ها را استخراج کرد [۳۸].

<sup>۱</sup> Miss Values

۳. یک سیگنال یا پارامتر ممکن است، با اسامی مختلفی ذخیره شده باشد. در بعضی از آیتم‌ها تفاوتی وجود ندارد؛ ولی برای آن‌ها دو کد شناساگر تعریف شده است همچنین گاهی اشکالات تایپی در این آیتم‌ها وجود دارد. همچنین مقادیر هر پارامتر ممکن است، در بیشتر از یک جدول در پایگاه داده ذخیره شود. مثلا، آزمایشگاه بیمارستان نتایج آزمایش را در جدول labevents ذخیره می‌کند و پرستاران بعضی از این مقادیر را برای راحتی کارشان، در جدول chartevents کپی می‌کنند. آزمایش‌هایی که در ICU انجام می‌شود، نیز در جدول chartevents ذخیره می‌شوند. بنابراین، مقادیری از آزمایش‌ها در جدول labevents وجود دارند، که در جدول chartevents درج نشده اند و برعکس. برای استفاده از مقادیر آزمایش‌ها در طراحی سیستم تصمیم‌یار، باید در هر دو جدول به دنبال مقادیر مورد نظر بود [۱].
۴. ممکن است، در نام بیماران اشتباهاتی رخ دهد؛ مثلا، نام و کد شناساگر نادرست در دستگاه کنار بیمار وارد شود. حتی، در مواردی ممکن است، پارامترهای بیمار جدید برای بیمار قبلی که به آن دستگاه متصل بوده است، ثبت شود [۳۸].
۵. ممکن است، در تبدیل واحدها اشتباهی رخ داده باشد؛ مثلا، درجه حرارت بدن به واحد فارنهایت ذخیره شده است، در صورتی که در واقع واحد آن سانتی گراد باشد [۳۸].

با توجه به مشکلاتی که برای پایگاه داده MIMIC-III وجود دارد، جهت استفاده از این پایگاه داده برای طراحی سیستم‌های تصمیم‌یار، باید با دقت بیشتری، پرس و جوهای مناسب را نوشت. در ادامه، مراحل انتخاب مجموعه دادگان، بصورت گام به گام، توضیح داده خواهد شد.

### ۳-۳- مراحل انتخاب مجموعه داده

در این تحقیق راه‌های مختلف بهبود پیش‌بینی سپسیس، با استفاده از شبکه عصبی بازگشتی مورد بررسی قرار گرفت. با توجه به اینکه در پیش‌بینی سپسیس، با استفاده از الگوریتم‌های یادگیری ماشین، عوامل مختلفی از جمله مدت زمان نگاه به گذشته<sup>۱</sup>، مدت زمان پیش‌بینی<sup>۲</sup> و ویژگی‌های<sup>۳</sup> مورد استفاده در الگوریتم تاثیر دارند، می‌خواهیم تاثیر هر یک از این عوامل را بررسی کنیم.

در ابتدا، تاثیر تغییر بازه‌ی زمانی را، برای دو زمان نگاه به گذشته و زمان پیش‌بینی بررسی می‌کنیم. برای این مسئله بایستی ویژگی‌های مورد نظر را، برای پیش‌بینی در بازه‌های زمانی خواسته شده، بصورت

<sup>۱</sup> Observation Window

<sup>۲</sup> Prediction Window

<sup>۳</sup> Predictors

سری زمانی از پایگاه داده استخراج کرد. در این مرحله، نمونه‌هایی که سری زمانی کوتاه‌تر از مجموع بازه زمانی نگاه به گذشته و بازه زمانی پیش‌بینی داشته باشند، از دیتابیس حذف می‌شوند. در بررسی مسئله دوم، که نحوه انتخاب ویژگی و پرکردن داده‌های از دست رفته است، با توجه به مسئله سر ریز کردن در مسائل یادگیری ماشین، بعد از یافتن بهترین نتیجه در راه حل مسئله اول زمانی بعنوان بهترین زمانی برای پیش‌بینی سسپیس در نظر گرفته می‌شود، که با توجه به آن زمان، به بررسی سایر عوامل تاثیر گذار در پیش‌بینی سسپیس پرداخته می‌شود.

مراحلی که برای انتخاب داده‌های مورد نظر از پایگاه داده MIMIC-III باید انجام شود، بصورت زیر است: ابتدا این پایگاه داده را به محیط BigQuery اضافه می‌کنیم. BigQuery محیطی برای اجرای کوئری‌های پیچیده در پایگاه داده‌های بزرگ، مانند، MIMIC-III است. سپس، برای انتخاب داده‌گان مراحل زیر انجام می‌شود:

مرحله اول) در این مرحله، ابتدا باید یک سری زمانی از علائم حیاتی بیمار استخراج شود. هر یک از علائم حیاتی بیمار در پایگاه داده، یک کد مخصوص به خود دارد. لیست این کدها در جدول d\_items ذخیره شده است با ارتباط این جدول با جدول chartevents می‌توان، در یک زمان مشخص، مقدار یک ویژگی را برای یک بیمار بدست آورد. در جدول ۲-۳ می‌توان، نام متغیرهای مورد استفاده در این پژوهش و مقدار کد شناساگر آن‌ها را مشاهده کرد. در این تحقیق ما به سری زمانی از مقادیر ضربان قلب<sup>۱</sup>، سرعت تنفس<sup>۲</sup>، دمای بدن<sup>۳</sup>، حداقل فشار خون<sup>۴</sup>، حداکثر فشار خون<sup>۵</sup>، میانگین فشار خون<sup>۶</sup>، فشار اشباع اکسیژن<sup>۷</sup> و قند خون<sup>۸</sup> نیاز داریم. همچنین، در کنار این مقادیر دو مقدار ثابت سن<sup>۹</sup> و جنسیت<sup>۱۰</sup> نیز ثبت می‌شود.

---

<sup>1</sup> Heart Rate

<sup>2</sup> Respiratory Rate

<sup>3</sup> Temperature

<sup>4</sup> Systolic Blood Pressure

<sup>5</sup> Diastolic Blood Pressure

<sup>6</sup> Mean Blood Pressure

<sup>7</sup> Spo2

<sup>8</sup> Glucose

<sup>9</sup> Age

<sup>10</sup> Gender

جدول ۳-۲) ویژگی‌های استفاده شده بعنوان ورودی مدل و کدهای شناساگر آن‌ها از جدول d\_items از پایگاه داده

متغیرهای ورودی ( نام فارسی )	متغیرهای ورودی ( نام انگلیسی )	کدهای شناساگر از جدول d_items
ضربان قلب	Heart Rate	۲۲۰۰۴۵, ۲۱۱
نرخ تنفس ( تعداد تنفس در دقیقه )	Respiratory Rate	۲۲۴۶۹۰, ۲۲۰۲۱۰, ۶۱۵, ۶۱۸
دمای بدن	Temperature	۲۲۳۷۶۲, ۶۷۶
حداقل فشار خون	Systolic blood pressure	۲۲۰۰۵۰, ۲۲۰۱۷۹, ۶۷۰۱, ۴۵۵, ۴۴۲, ۵۱
حداکثر فشار خون	diastolic blood pressure	۲۲۰۰, ۲۲۰۱۸۰, ۸۵۵۵, ۸۴۴۱, ۸۴۴۰, ۸۳۶۸۸۳۶۸ ۵۱
میانگین فشار خون	mean blood pressure	۲۲۵۳۱۲, ۲۲۰۱۸۱, ۲۲۰۰۵۲, ۴۴۳, ۶۷۰۲, ۵۲, ۴۵۶
فشار اشباع اکسیژن	Spo2	۲۲۰۲۷۷, ۶۴۶
قندخون	glucose	۲۲۰۶۲۱, ۲۲۵۶۶۴, ۳۷۴۴, ۳۷۴۵, ۱۵۲۹, ۸۱۱, ۸۰۷ ۲۲۶۵۳۷,
سن	age	-
جنسیت	gender	-

برای بدست آوردن سری زمانی، از مقادیر ویژگی‌های بالا و مقادیر ثابت سن و جنسیت، یک پرس و جو در محیط BigQuery اجرا شده است. این پرس و جو روی جدول chartevents اعمال می‌شود و مقادیر این ویژگی‌ها همراه با زمان ثبت آن‌ها استخراج شده است. در استخراج مقادیر این ویژگی‌ها، باید از کدهای شناساگر آن‌ها استفاده کرد. زمان ثبت هر کدام از ویژگی‌ها به عنوان زمان مرجع برای استخراج آن ویژگی قرار می‌گیرد. در زمان استخراج یک ویژگی ممکن است، هیچ مقداری در آن زمان مرجع، برای آن ویژگی وجود نداشته باشد. در این شرایط نزدیکترین مقدار از نظر زمانی به زمان مرجع، به عنوان مقدار مورد نظر برای آن ویژگی انتخاب می‌شود.

مرحله دوم) برای بدست آوردن اینکه یک بیمار در یک زمان خاص به یک بیماری مشخص مبتلا شده است یا نه، رایج ترین راه استفاده از کدهای ICD-9 می‌باشد، اما در این پژوهش به طور مطلق از این کد استفاده نشده است همانطور که در فصل اول در رابطه با روش‌های تشخیص سپسیس صحبت شد، یکی از این روش‌ها استفاده از درجه SOFA است. برای بدست آوردن درجه SOFA برای یک بیمار بستری شده در ICU، به یک مقدار اکتفا نمی‌کنیم و یک سری زمانی از مقادیر SOFA را استخراج می‌کنیم.

برای استخراج سری زمانی از مقادیر SOFA برای همه ی بیمارانی که از مرحله اول بدست آمد، نیاز به یک سری از مقادیر علائم حیاتی و آزمایشگاهی داریم. متغیرهایی که در این اندازه گیری استفاده می شود، شامل مقیاس کمای گلاسکو، فشار متوسط شریانی، میزان اکسیژن داده شده به بیمار توسط ونتیلاتور، وضعیت استفاده از ونتیلاتور، میزان کراتینین، فشار نسبی اکسیژن در خون شریانی، میزان بیلی روبین، میزان پلاکت خون، میزان دوپامین، میزان دوبوتامین، میزان اپی نفرین، میزان نور اپی نفرین و میزان خروجی ادرار است. میزان هر کدام از این ویژگی ها در یک جدول از پایگاه داده MIMIC-III قرار دارد. در جدول ۳-۳ منبعی که هر کدام از این ویژگی ها از آن استخراج شده را می توان دید. با استفاده از مقادیری که برای اندازه گیری SOFA استخراج شده، طبق شکل ۳-۹ مقدار آن را برای هر بیمار برای یک روز کامل هر یک ساعت یک بار بدست می آوریم. با انجام این کار در واقع، یک سری زمانی از مقادیر استخراج SOFA برای بیمار استخراج می شود.

اگر در هنگام محاسبه مقدار SOFA هر کدام از متغیرهای مورد نیاز برای این اندازه گیری بعنوان داده از دست رفته باشد، دو راهکار در نظر گرفته شده است یک سری زمانی از مقادیر آن ویژگی در زمان های قبل و بعد از آن ساعت استخراج شد و با توجه به نظر خبره، یک مقدار برای آن بطوری که به مقدار آن ویژگی در ساعت قبل و بعد نزدیک باشد، در نظر گرفته شد. اگر مقادیر در ساعت های قبلی و بعدی نامعلوم باشد، مقدار صفر برای آن در نظر گرفته شده است.

جدول ۳-۳ لیست متغیرهای استفاده شده در اندازه گیری درجه SOFA و جدول منبع در پایگاه داده

متغیر اندازه گیری SOFA	جدول منبع در پایگاه داده
مقیاس کمای گلاسکو فشار متوسط شریانی میزان اکسیژن داده شده به بیمار توسط ونتیلاتور وضعیت استفاده از ونتیلاتور	CHARTEVENTS
میزان کراتینین فشار نسبی اکسیژن در خون شریانی میزان بیلی روبین میزان پلاکت خون	LABEVENTS
میزان دوپامین میزان دوبوتامین میزان اپی نفرین میزان نور اپی نفرین	INPUTEVENTS_MV and INPUTEVENTS_CV
میزان خروجی ادرار	OUTPUTEVENTS

System or organ and measure	SOFA score				
	0	1	2	3	4
Respiratory:					
P <sub>a</sub> O <sub>2</sub> /FiO <sub>2</sub> , mmHg	≥400	300-399	200-299	100-199 with respiratory support	<100 with respiratory support
Coagulation:					
Platelets, × 10 <sup>3</sup> /μL	≥150	100-149	50-99	20-49	<20
Liver:					
Bilirubin, μmol/L (mg/dL)	<20 (1.2)	20-32 (1.2-1.9)	33-101 (2.0-5.9)	102-204 (6.0-11.9)	>204 (12.0)
Circulatory:					
Mean arterial pressure, mm Hg	≥70	<70	Low dose dopamine or any dose dobutamine	Low-medium dose noradrenalin or adrenalin; medium dose dopamine	High dose noradrenalin, adrenalin, or dopamine
Central nervous system:					
Glasgow Coma Scale score	15	13-14	10-12	6-9	<6
Renal:					
Creatinine, μmol/L (mg/dL)	<110 (1.2)	110-170 (1.2-1.9)	171-299 (2.0-3.4)	300-440 (3.5-4.9)	>440 (5.0)
Urine output, mL/day	–	–	–	<500	<200

\*Our recommendation applies to patients with an infection and a SOFA score of ≥2.  
P<sub>a</sub>O<sub>2</sub> = partial pressure of oxygen (arterial). FiO<sub>2</sub> = fraction of inspired oxygen.

### شکل ۳-۹) نحوه محاسبه مقدار درجه SOFA

مرحله سوم) در این مرحله برای هر بستری بیمار دو سری زمانی در دسترس است؛ یکی سری زمانی مربوط به درجه SOFA و دیگری سری زمانی مربوط به علائم حیاتی بیمار. ابتدا مقادیر سری زمانی درجه SOFA بیمار مورد بررسی قرار گرفته است و در زمانی که مقدار SOFA بیشتر از مقدار دو شده باشد، بعنوان زمان شروع سپسیس برای آن بستری بیمار ثبت می شود. حال زمان دو روز قبل و یک روز بعد از این زمان استخراج شده است. اگر زمان دو روز قبل از شروع سپسیس کمتر از زمان بستری بیمار باشد، ما زمان بستری بیمار را برای در نظر گرفتن شروع سری زمانی علائم حیاتی انتخاب می کنیم. به همین صورت اگر زمان یک روز بعد از شروع سپسیس بعد از زمان ترخیص باشد، ما زمان ترخیص را بعنوان زمان پایان سری زمانی علائم حیاتی انتخاب می کنیم. حال باید با استفاده از مقادیر زمان شروع و پایان، ثبت سری زمانی علائم حیاتی بازه زمانی نگاه به گذشته<sup>۱</sup> و بازه زمانی پیش بینی<sup>۲</sup> را استخراج کنیم.

<sup>1</sup> Observation Window

<sup>2</sup> Prediction Window

در این پژوهش برای دو مقدار بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی مقدارهای مختلف تست شده است که در فصل پنجم توضیح داده خواهد شد.

برای مثال اگر بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی به ترتیب ۶ و ۱۲ ساعت در نظر گرفته شود، باید بیمارانی را انتخاب کنیم که در بازه‌زمانی که سری زمانی علائم حیاتی از آن‌ها ثبت شده است، ۱۸ یا بیشتر از ۱۸ باشد، در غیراینصورت آن بستری بیمار برای این مرحله تست مدل، حذف می‌شود.

مرحله چهارم) با توجه به مقدار SOFA در زمان شروع سپسیس می‌توان، کلاس آن بستری بیمار را مشخص کرد و با توجه بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی و دادگان علائم حیاتی که در مرحله اول بدست آمد، یک کلاس برای آن بستری بیمار و آن علائم حیاتی مشخص کرد. زمانی که مقدار SOFA روند صعودی از مقدار دو پیدا کند، آن مقدار بعنوان زمان شروع سپسیس ثبت می‌شود و در واقع به آن بیمار در آن بستری کلاس مثبت داده می‌شود و اگر در هیچ زمانی مقدار این متغیر از دو بیشتر نشود، به آن بیمار در آن بستری مقدار کلاس منفی داده می‌شود. در حال حاضر ما یک سری زمانی از علائم حیاتی همراه با کلاس تشخیص آن‌ها داریم.

مرحله پنجم) آن‌هایی که درجه SOFA برای آن‌ها ثبت نشده بود، با استفاده از کدهای ICD-9 درمورد ابتلا به سپسیس بیمار در آن بستری تصمیم‌گیری شد. کدهای ICD\_9 که برای سپسیس در این پایگاه داده وجود دارد، دو مقدار ۵۲.۷۸۵ و ۹۲.۹۹۵ است. با توجه به این دو مقدار در جدول diagnoses ICD می‌توان، برای هر بستری بیمار وضعیت ابتلا یا عدم ابتلا به سپسیس را مشخص کرد.

مرحله ششم) با توجه به نظر خبره، کودکان از این دیتاست حذف شدند؛ زیرا تعداد این بیماران باعث می‌شد، که تعادل بین کلاس منفی و کلاس مثبت از بین برود. در واقع بیمارانی که زمان بستری شدن آن‌ها بیشتر از یک سال بعد از تولدشان بود را، بعنوان بیماران بزرگسال استخراج می‌کنیم. در نهایت بعد از انجام تمام مراحل بالا به ۴۶۴۷ بستری با کلاس مثبت و ۴۶۸۵۰ بستری با کلاس منفی بدون اعمال محدودیت طول سری زمانی دست می‌یابیم.



## فصل ۴- روش‌های پیشنهادی

در این تحقیق، هدف، تشخیص شروع سپسیس بیماران بستری شده در ICU می باشد. در تحقیق اسپرپ و همکاران [۱۴]، که بر روی بیماران مبتلا به سپسیس انجام شده است؛ در ابتدا ویژگی های پیش بینی کننده را برای پیش بینی سپسیس انتخاب کرده اند و با استفاده از کدهای ICD-9 به بیماران با علائم حیاتی مشخص برچسب شروع شدن سپسیس و عدم شروع سپسیس زده اند [۱۴]. در این حالت، با نداشتن مقادیر علائم حیاتی به برچسب شروع سپسیس/عدم شروع سپسیس، تشخیص سپسیس تبدیل به یک مسئله دسته بندی خواهد شد.

با توجه به تحقیقات گذشته ای که در زمینه سپسیس انجام شده است، می توان گفت که در پژوهش های این موضوع، دو چالش اصلی وجود دارد. چالش اول مربوط به زمان پیش بینی سپسیس می باشد، که در این مورد دو بازه زمانی در نظر گرفته می شود. اولین بازه زمانی مربوط به بازه زمانی نگاه به گذشته و دومین بازه زمانی مربوط به بازه زمانی پیش بینی است. در حل این چالش در واقع با نگاه کردن به اطلاعات بیمار در بازه زمانی نگاه به گذشته شروع سپسیس در یک بیمار حداکثر قبل از بازه ی زمانی پیش بینی، مشخص می شود. در این چالش هدف کوتاه کردن بازه زمانی نگاه به گذشته و افزایش بازه زمانی پیش بینی قبل از شروع سپسیس است. چالش دوم مربوط به انتخاب ویژگی های کارآمد در این پیش بینی است. عوامل مختلفی مانند دسترسی پذیری به اطلاعات پزشکی، استفاده از حداقل تعداد ویژگی در پیش بینی، در انتخاب این ویژگی ها وجود دارد. در این چالش هدف استفاده از کمترین و تاثیرگذارترین تعداد ویژگی است. در حل چالش دوم موضوع روش پرکردن داده های از دست رفته نیز اهمیت دارد.

در چالش اول عملکرد ۴ الگوریتم بصورت کلی بررسی شد اما با توجه به اینکه هدف اصلی این پژوهش بررسی الگوریتم های شبکه عصبی بازگشتی است، یک نوع شبکه عصبی بازگشتی بصورت جزئی بررسی شده است چهار الگوریتم اول عبارتند از دسته بند درخت تصمیم، دسته بند جنگل تصادفی و دسته بند پرسپترون چند لایه است. شبکه عصبی استفاده شده در این پژوهش الگوریتم LSTM است.

این نکته باید ذکر شود که در طراحی سیستم های تصمیم یار پزشکی یکی از نکات مهم این است که خروجی سیستم باید برای پزشکان ارزش بالینی داشته باشد، کیفیت مراقبت های پزشکی را بالا ببرد و عوارض ناشی از مراقبت های پزشکی را کاهش دهد. سیستم تصمیم یار باید با روند کاری پزشک همخوانی داشته باشد و اطلاعاتی که در اختیار می گذارد برای پزشک مفید باشد. همچنین هر پزشک ممکن است عادات کاری متفاوتی داشته باشد [۳۹].

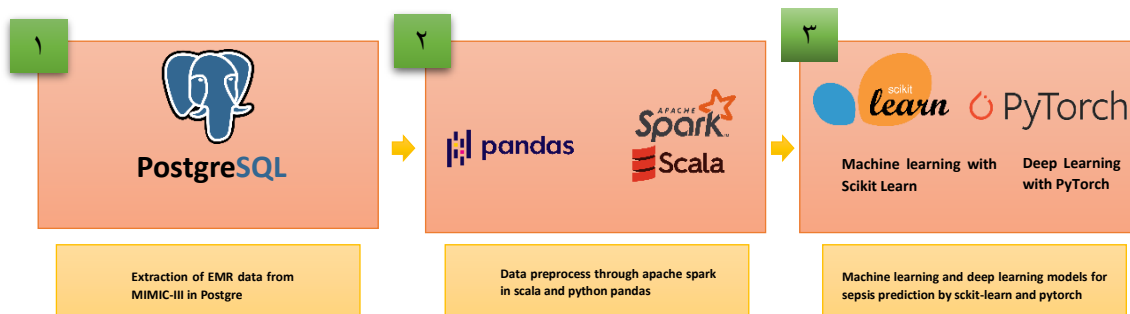
در پژوهش های قبلی همگی با استفاده از کدهای ICD-9 مقدار کلاس را برای پیش بینی استفاده کرده بودند و از تعداد ویژگی های بالایی استفاده کرده اند، حتی در بعضی از پژوهش ها اگر از معیار SOFA برای تعیین کلاس استفاده کرده باشند آخرین مقدار SOFA را برای تعیین کلاس بیمار در نظر گرفته اند. اما در این پژوهش برای تعیین کلاس نه تنها از مقدار SOFA استفاده شده؛ بلکه به مقدار آخر درجه

SOFA اکتفا نشده و یک سری زمانی از مقادیر SOFA را در نظر گرفتیم. با توجه به نکات بالا معمولاً پزشکان و مسئولان سلامت به دلایل مختلف بعضی از ویژگی‌های بیماران را ذخیره نمی‌کنند. این اتفاق ممکن است دلایلی از جمله خراب شدن دستگاه ثبت اطلاعات، عدم وجود زمان کافی و ... داشته باشد. لذا در چالش دوم بعد از انتخاب بهترین زمان، از ویژگی‌های استفاده شده که تعداد کمتری داده‌های از دست رفته را داشته باشد. در ادامه روش پیشنهادی برای پیاده‌سازی الگوریتم‌ها گفته شده است.

#### ۱-۴- روش‌های ارزیابی بازه‌های زمانی پیش‌بینی و نگاه به گذشته

##### ۱-۱-۴- روش پیشنهادی

همانطور که گفته شد هدف پیش‌بینی شروع سپسیس در بیماران بخش مراقبت‌های ویژه است تا پزشک بتواند بر اساس آن راحت‌تر تصمیم بگیرد. روش پیشنهادی برای رسیدن به این هدف شامل سه مرحله اصلی است. این مراحل در شکل ۱-۴ نشان داده شده است:



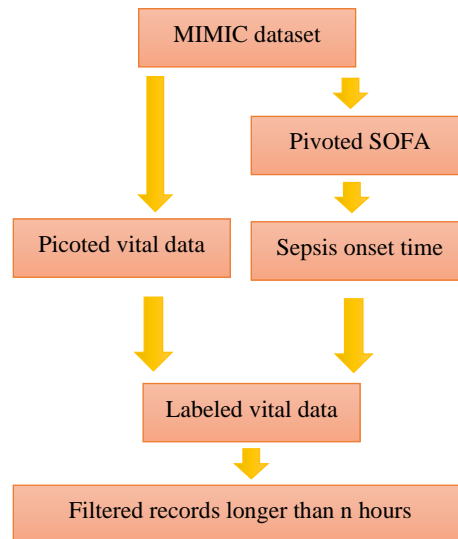
شکل ۱-۴) روند اجرای روش پیشنهادی

#### گام ۱- استخراج دادگان

در تحقیق اسچرپ و همکاران برای دسته‌بندی بیماران به دو کلاس مثبت و منفی یا همان شروع سپسیس/عدم شروع سپسیس از کدهای ICD-9 استفاده شده است. همانطور که در فصل سوم گفته شد، کدهای ICD-9 منحصربفردی هستند که زمانی که بیمار مرخص می‌شود، طبق تشخیص پزشک ثبت می‌شود. در این تحقیق بر خلاف روش‌های پیشین، برای دسته‌بندی بیماران از معیار SOFA استفاده شده است.

وقتی با استفاده از کدهای ICD-9 برای تعیین کلاس در مسائل دسته‌بندی استفاده می‌شود در واقع فقط بر اساس تشخیص پزشک مقدار کلاس مشخص شده‌است؛ ولی وقتی از معیارهای SOFA استفاده می‌شود تعیین کلاس شروع سپسیس/ عدم شروع سپسیس فقط براساس تشخیص پزشک نیست. همانطور که در فصل قبل گفته شد برای تعیین مقدار درجه SOFA از ویژگی‌های علائم حیاتی، مقادیر آزمایشگاهی و ... استفاده می‌شود. استفاده از این معیار به ما کمک می‌کند که نه تنها کلاس مقادیر ورودی را تعیین کنیم بلکه می‌توان زمان شروع سپسیس را تشخیص داد و بر اساس آن مقادیر بازه‌زمانی پیش‌بینی و مقادیر بازه‌زمانی گذشته نگر را بدست آورد. در این پژوهش زمان شروع و پایان این دو بازه‌زمانی برای برای همه ی بیماران یکسان نیست.

با استفاده از این معیار از تاثیر پارامترهای بیشتری برای پیش‌بینی استفاده می‌شود اما این در حالی است که برای پیش‌بینی اینکه یک بیمار سپسیس در او شروع می‌شود یا نه نیاز به همه ی این متغیرها نیست ؛ عبارت دیگر برای ساخت مدل باید ابتدا دسته بیماران مشخص شود. برای تعیین دسته بیماران از معیار درجه SOFA استفاده شده‌است برای بدست آوردن مقدار درجه SOFA نیاز به یک سری علائم-حیاتی و مقادیر آزمایشگاهی داریم. لیست کامل این مقادیر آزمایشگاهی در جدول ۳-۳ آمده‌است حال وقتی دسته بیمار مشخص شد و با استفاده از آن مدل را ساختیم، از همان ویژگی‌ها بعنوان پیش‌بینی کننده استفاده نمی‌شود بلکه از ده ویژگی دیگر برای پیش‌بینی استفاده می‌شود. با این روش با استفاده از مقادیر تعداد کمتری ویژگی از تاثیر سایر ویژگی‌ها نیز بهره برده‌ایم. در شکل ۴-۲ مراحل برچسب زدن به مجموعه دادگان نشان داده شده‌است. در مرحله آخر در این شکل بسته به اینکه چه بازه‌زمانی را می‌خواهیم بررسی کنیم دادگانی که سری زمانی کوتاهتری دارند حذف شده‌است.



شکل ۴-۲) روند استخراج مجموعه دادگان برای هر زوج مرتب با بازه زمانی پیش‌بینی  $n$  ساعته

## گام ۲- پیش پردازش دادگان

بعد از دسته‌بندی بیماران به دو دسته شروع سپسیس یا عدم شروع سپسیس باید مقادیر ویژگی‌های پیش‌بینی کننده را برای آن بستری بیمار استخراج کنیم. بعد از استخراج ای مقادیر و تعیین درجه SOFA می‌توان به این مقادیر کلاس صفر و یک اختصاص داد. در حال حاضر تعداد ۵۱۴۹۷ رکورد از بستری بیماران در ساعت‌های مختلف ( نحوه انتخاب زمان شروع و پایان این ساعت در فصل قبل توضیح داده شد) ثبت شده‌است برای نمونه برای بستری با شماره بستری ۲۰۰۲۷، در ساعت ۱۵ روز فعلی تا ساعت ۹ روز بعد هر یک ساعت یک بار ثبت شده است پس برای این بستری ۱۸ رکورد در مجموعه دادگان موجود است که هر رکورد کلاس مخصوص به خود را دارد. شکل ۴-۳ نمونه از این دیتاست را نشان می‌دهد. نحوه استخراج مقادیر ویژگی‌های پیش‌بینی کننده و ساخت جدول ۴-۳ در فصل سه گفته شده‌است.

icustay_id	record_time	heartrate	sysbp	diasbp	meanbp	resprate	tempc	spo2	glucose	age	gender	sepsis_label
200277	1/9/2143 15:00	85.5	114	59	71	14	36.72222	100	73	22.9875	1	0
200277	1/9/2143 16:00	96	110	66	77	20	36.72222	99	73	22.9875	1	0
200277	1/9/2143 17:00	96	105	58	69	20	36.72222	100	73	22.9875	1	0
200277	1/9/2143 18:00	89	105	60	70	17	36.72222	99	73	22.9875	1	0
200277	1/9/2143 20:00	100	105	60	70	14	37.22222	99	73	22.9875	1	0
200277	1/9/2143 21:00	94	109	58	70	17	37.33333	100	73	22.9875	1	0
200277	1/9/2143 22:00	89	105	65	74	17	37.38889	100	73	22.9875	1	0
200277	1/9/2143 23:00	100	118	72	82	24	37.33333	100	73	22.9875	1	0
200277	1/10/2143 0:00	85	102	54	65	12	37.33333	99	73	22.9875	1	0
200277	1/10/2143 1:00	81	101	61	69	21	37.33333	100	73	22.9875	1	0
200277	1/10/2143 2:00	71	105	66	75	16	37.33333	100	73	22.9875	1	0
200277	1/10/2143 3:00	86	103	56	66	18	37.33333	100	73	22.9875	1	0
200277	1/10/2143 4:00	80.5	109	58	68	18	37	100	73	22.9875	1	0
200277	1/10/2143 5:00	83	106	64	74	19	37	99	73	22.9875	1	0
200277	1/10/2143 6:00	92	113	67	77	17	37	100	73	22.9875	1	0
200277	1/10/2143 7:00	100	107	69	77	20	37	100	73	22.9875	1	0
200277	1/10/2143 8:00	92	107	67	74	18	37.27778	98	73	22.9875	1	0
200277	1/10/2143 9:00	91	100	54	65	18	37.27778	97	73	22.9875	1	0
200324	11/13/2161 23:00	85	102.3333	66	87.33333	13.5	36	100	142.5	65.4279	0	0
200324	11/14/2161 0:00	85	102.3333	66	87.33333	13.5	36	100	76	65.4279	0	0
200324	11/14/2161 1:00	89.25	105.25	70.25	93	14.75	36.22222	100	76	65.4279	0	0
200324	11/14/2161 2:00	89.5	105.75	65.5	107.5	15.75	35.51111	100	82	65.4279	0	0

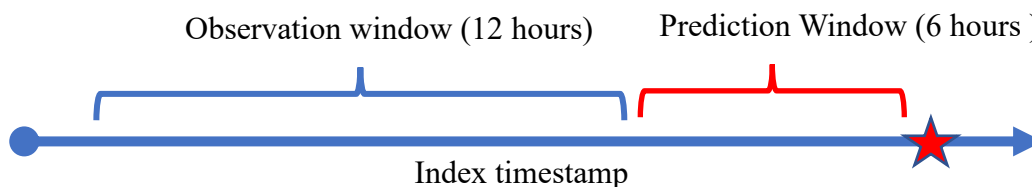
شکل ۴-۳) نمونه داده استفاده شده برای استخراج داده مورد نیاز

حال باید دادگان را برای استفاده در مدل، آماده‌سازی کنیم. برای آماده‌سازی دادگان ابتدا کل دادگان بصورت تصادفی به داده آموزشی، آزمایشی و ارزیابی تقسیم شده‌است در این تقسیم بندی به ترتیب به دسته آموزشی، آزمایشی و ارزیابی ، ۸۰ ، ۲۰ و ۱۰ درصد دادگان را اختصاص می‌دهیم.

در این پژوهش یک از چالش‌های مورد بررسی تغییر در بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی است. با توجه به اینکه می‌خواهیم مقادیر مختلفی را برای این دو بازه‌زمانی بررسی کنیم، نمی‌توان از یک مجموعه داده ثابت برای دادن به مدل و ارزیابی آن استفاده کرد پس با توجه به مقادیری که این دو بازه‌زمانی می‌گیرند، مجموعه دادگان مختلف به روش زیر ساخته می‌شود.

برای ساخت مجموعه دادگان برای هر زوج‌مرتب بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی باید ابتدا مقادیر این دو بازه را تعیین کرد. بعنوان مثال ما مقادیر ۱۲ و ۶ را به ترتیب برای بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی در نظر می‌گیریم. حال برای هر بستری بیمار ابتدا باید به کلاس‌های آن‌ها نگاه کنیم، اگر کلاس‌های این بستری‌ها همواره صفر باشد لذا باید آخرین ساعت ثبت اطلاعات را بعنوان مبدا زمانی در نظر گرفت و به داده ۱۲ ساعت قبل آن نگاه کرد و داده قبل از ۱۲ ساعت را کنار می‌گذاریم. در این انتخاب الزاما ۱۲ رکورد انتخاب نمی‌شود. اگر در یک بستری بیمار مقدار کلاس از ساعتی به بعد ۱ شود ۶ ساعت قبل از ساعتی که مقدار کلاس ۱ شده بعنوان مبدا زمانی در نظر گرفته می‌شود و داده ۱۲ ساعت قبل این مبدا زمانی استخراج می‌شود و سایر دادگان در این مرحله کنار گذاشته می‌شود. حال هر رکورد را دو بخش می‌کنیم یکی مقادیر ویژگی‌ها و دیگری مقدار کلاس‌ها و برای هر شماره

بستری یک سری زمانی از کلاس‌ها و مقادیر ویژگی‌ها تشکیل می‌دهیم. شکل ۴-۴ نحوه ارتباط این دو بازه‌زمانی با شروع سپسیس را نشان می‌دهد.

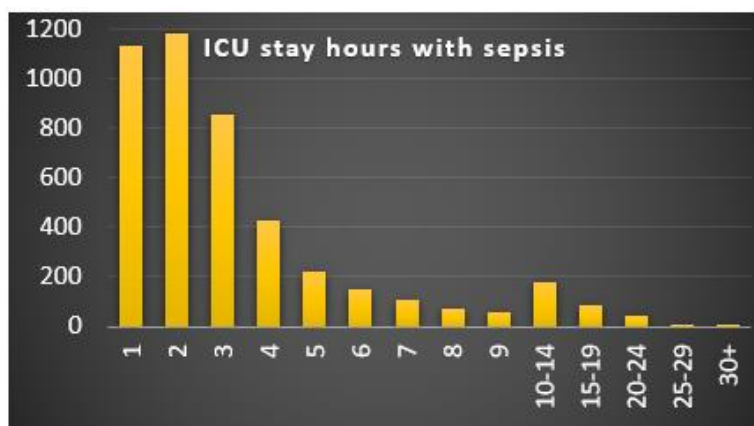


شکل ۴-۴) مفهوم بازه‌زمانی نگاه به گذشته و پیش‌بینی

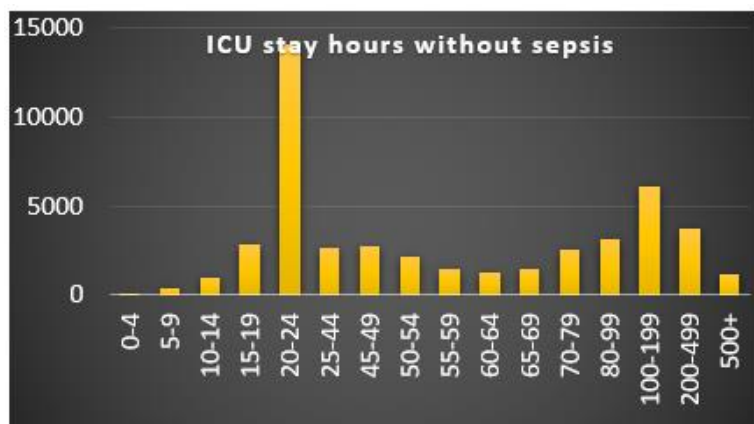
این روند را برای تمامی زوج‌مرتب‌ها، برای سه دسته آموزشی و آزمایشی و ارزیابی اجرا می‌کنیم و مجموعه داده هر بخش را بصورت جداگانه ذخیره می‌کنیم. در شکل ۴-۶ تعداد ساعات بستری بیماران به تفکیک برچسب شروع سپسیس و عدم شروع سپسیس نشان داده شده‌است همانطور که مشخص است اکثر بستری‌های بیماران کمتر از ۲۴ ساعت به طول انجامیده‌است؛ لذا با توجه به مسئله توضیح داده شده، برای انتخاب سری‌زمانی‌هایی که پیش از این گفته شد نمی‌توان بازه‌زمانی را برای هر دو بازه-زمانی بیشتر از ۲۴ ساعت در نظر گرفت. در این پژوهش در فواصل زمانی ۳ ساعته از ۳ ساعت تا ۲۴ ساعت برای هر دو بازه‌زمانی بررسی شده‌است بعبارتی دیگر برای دو بازه‌زمانی گذشته نگر و پیش‌بینی به ترتیب زوج‌مرتب‌های مشخص شده در شکل ۴-۵ در نظر گرفته شده‌است. وقتی مجموعه دادگان هر زوج‌مرتب مشخص شد، باید در این مرحله دو مرحله نرمال‌سازی بر روی دادگان انجام دهیم. ابتدا با استفاده از روش حداقل- حداکثر نرمال‌سازی تمامی دادگان را در بازه صفر و یک نرمال می‌کنیم. سپس مقادیر از دست‌رفته را با مقدار میانگین برای هر شماره بستری پر می‌کنیم. در ادامه روش‌های مختلف پرکردن داده‌های از دست‌رفته را بررسی می‌کنیم و نشان خواهیم‌داد که بهترین روش برای پرکردن داده‌های از دست‌رفته میانگین برای هر شماره بستری است.

	۳	۶	۹	۱۲	۱۵	۱۸	۲۱	۲۴
۳	*	*	*	*	*	*	*	*
۶	*	*	*	*	*	*	*	*
۹	*	*	*	*	*	*	*	*
۱۲	*	*	*	*	*	*	*	*
۱۵	*	*	*	*	*	*	*	*
۱۸	*	*	*	*	*	*	*	*
۲۱	*	*	*	*	*	*	*	*
۲۴	*	*	*	*	*	*	*	*

شکل ۴-۵) زوج‌مرتب‌های بررسی شده در پژوهش



الف) طول بازه‌های بستری در بیمارستان برای بیماران با شروع سپسیس



ب) طول بازه‌های بستری در بیمارستان برای بیماران با عدم شروع سپسیس

شکل ۴-۶) الف) طول بازه‌های بستری در بیمارستان برای بیماران با شروع سپسیس ب) طول بازه‌های بستری در بیمارستان برای بیماران با عدم شروع سپسیس



### گام ۳- مدلسازی

بعد از آماده‌سازی دادگان باید مجموعه دادگان هر زوج‌مرتب را به مدل مورد بررسی بدهیم. با توجه به اینکه هدف این پژوهش بهبود پیش‌بینی سپسیس با استفاده از الگوریتم‌های شبکه عصبی بازگشتی است؛ ابتدا ما عملکرد یکی از معروف‌ترین الگوریتم‌های شبکه عصبی بازگشتی را برای تمام زوج‌مرتب‌های گفته‌شده در بخش قبل بررسی می‌کنیم. الگوریتم شبکه عصبی استفاده شده در این پژوهش الگوریتم شبکه عصبی بازگشتی LSTM است. سپس بهترین زوج‌مرتب را بر اساس تمامی معیارهای ارزیابی استخراج می‌کنیم. داده‌های زوج‌مرتب برتر را به سایر الگوریتم‌های دسته‌بند می‌دهیم و عملکرد آن‌ها را با شبکه عصبی بازگشتی LSTM مقایسه می‌کنیم. در ادامه به معرفی شبکه‌های عصبی بازگشتی و LSTM می‌پردازیم.

#### ۴-۱-۱-۱- شبکه‌های عصبی بازگشتی و LSTM

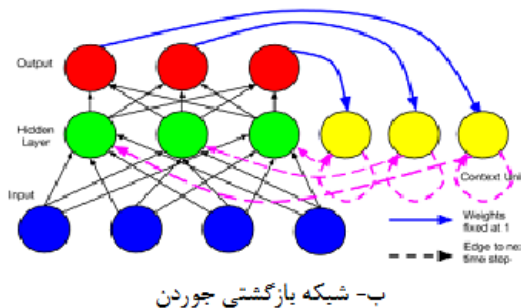
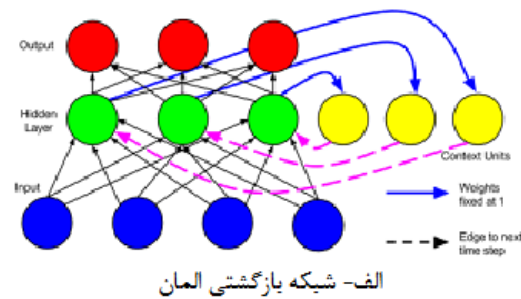
شبکه‌های عصبی بازگشتی در واقع سامان‌های پویایی هستند که وضعیت<sup>۱</sup> درونی خود را در خلال گام‌های زمانی یک فرایند حفظ می‌کنند. این قابلیت به دلیل ارتباطات حلقوی است که در میان نورون‌های سطوح بالا با نورون‌های لایه‌های پایین‌تر وجود دارد. در برخی از مدل‌ها، این ارتباطات در میان نورون‌های هم‌لایه یا حتی در ارتباط یک نورون با خودش نیز دیده می‌شود. این ارتباطات بازگشتی به شبکه اجازه می‌دهد که داده‌های مربوط به گام‌های گذشته را تا مراحل بعدی در خود حفظ کند. به این ترتیب شبکه‌های بازگشتی درواقع از نوعی حافظه برخوردار هستند و در نتیجه مدل محاسباتی آن‌ها از شبکه‌های عصبی پیشخور<sup>۲</sup> قدرتمندتر می‌باشد.

شبکه‌های بازگشتی از انواع کمی متصل تا کاملاً متصل گسترده شده‌اند. دو نوع از مشهورترین شبکه‌های بازگشتی سنتی، به وسیله المان [۴۱] و جوردن [۴۲] معرفی شده‌است شبکه المان شبیه یک شبکه عصبی معمولی است که از سه لایه تشکیل شده و علاوه بر اتصالات رایج، خروجی لایه پنهان آن در سلول‌های خاصی که سلول‌های زمینه نام دارند ذخیره می‌شود. در هر گام از فرایند آموزش، خروجی‌های سلول‌های زمینه<sup>۳</sup> که از گام قبل باقی مانده‌اند به همراه ورودی‌های گام جاری مجدداً به لایه پنهان تزریق می‌شوند. شبکه‌های جوردن، ساختاری شبیه شبکه‌های المان دارند و تنها تفاوت آن‌ها در این نکته است که محتویات سلول‌های زمینه به جای لایه پنهان، از طریق لایه خروجی تأمین می‌شود. شکل ۴-۷ نمونه‌هایی از این دو شبکه را نشان می‌دهد.

<sup>1</sup> State

<sup>2</sup> Feed Forward Network

<sup>3</sup> Context Cell



شکل ۴-۷) الف- شبکه بازگشتی المان ب- شبکه بازگشتی جوردن [۴۳]

برای آموزش یک شبکه بازگشتی می‌توان از الگوریتم معمول انتشار پسگرد<sup>۱</sup> استفاده کرد. به این شکل که خروجی سلول‌های زمینه به مجموعه ورودی‌های لایه پنهان اضافه بشوند تا ساختار شبکه در بستر زمان گسترش یافته و از حالت بازگشتی خارج شود.

### انفجار و اختفای گرادیان

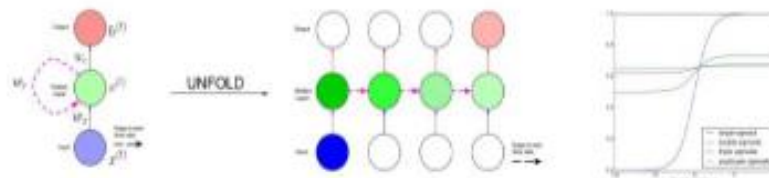
شبکه‌های بازگشتی سنتی از دو مشکل اختفا<sup>۲</sup> و انفجار<sup>۳</sup> گرادیان دارند. می‌دانید که در هر مرحله از آموزش یک شبکه عصبی بازگشتی تعداد زیادی ضرب انجام می‌پذیرد. وجود مجموعه بزرگی از وزن‌های بزرگتر از یک، و انجام ضرب‌های یاد شده، موجب می‌شود شبکه به سمت اوزان بسیار بزرگ یا همان انفجار گرادیان هدایت شود. در نتیجه این پدیده، حرکت کل شبکه به سمت نقاط مطلوب از کنترل خارج می‌شود و شبکه با گام‌های بسیار بزرگ در اطراف نقاط جواب نوسان می‌کند. البته مشکل انفجار گرادیان معمولاً از طریق محدود کردن وزن‌های شبکه قابل حل است؛ اما مشکل اختفای گرادیان که مثل انفجار از فراوانی تعداد ضرب‌ها در فرایند یادگیری یک شبکه عصبی بزرگ ناشی می‌شود، مشکل بدخیمی است که در نتیجه وجود مجموعه بزرگی از وزن‌های کوچک‌تر از یک ایجاد می‌گردد. در این حالت، گرادیان خطا به تدریج کوچک و کوچک‌تر می‌شود، در نتیجه انتشار پسگرد خطا و اصلاح وزن‌های

<sup>۱</sup> Backpropagation

<sup>۲</sup> Vanishing Gradient

<sup>۳</sup> Exploding Gradient

شبکه در اندازه‌های بسیار کوچک صورت می‌پذیرد، تا جایی که نهایتاً انتشار خطا متوقف شده و شبکه عملاً حساسیت خود را نسبت به وقوع خطا از دست می‌دهد. از این نقطه به بعد، بدون اینکه شبکه به جواب یا حتی اطراف آن همگرا شده باشد، وزن‌ها دیگر هیچ تغییری نمی‌کنند. این مشکل به صورت تصویری در شکل ۴-۸ نشان داده شده است.



شکل ۴-۸) مشکل اختفای گرادیان در شبکه‌های بازگشتی [۴۳]

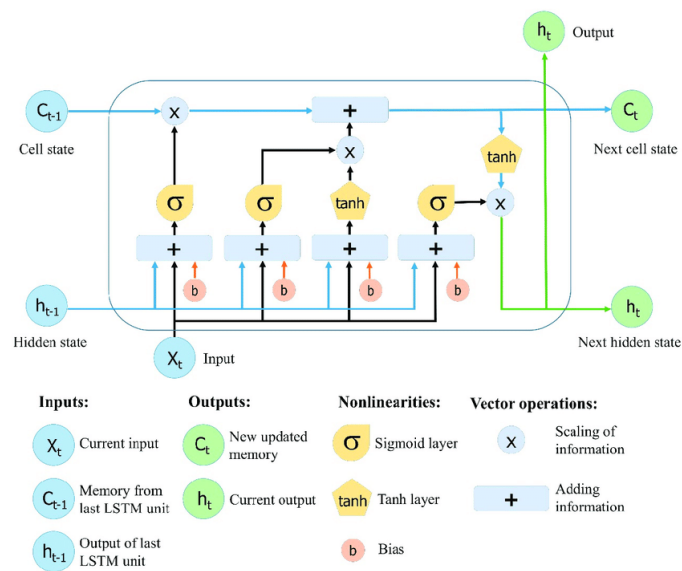
مجموعه این مشکلات باعث می‌شد تا شبکه‌های بازگشتی قدیمی امکان پردازش و پویا وابستگی‌های دور و دراز را نداشته باشند و مثل مدل‌های مارکوف به یکی دو سه گام قبل کفایت کنند. در سال ۱۹۹۷ سپ هواختر و هوشیدر [۴۴] سعی کردند با ابداع مفهوم درگاه‌های ورودی و خروجی مشکل اختفای گرادیان در شبکه‌های بازگشتی را حل کنند. تلاش ایشان به ارائه نوعی از شبکه‌های بازگشتی به نام حافظه طولانی منجر گردید. این ساختار که بعدها با LSTM کوتاه مدت با اضافه شدن درگاه فراموشی<sup>۱</sup> کاملتر شد [۴۵] در حل بسیاری از مسائل پیچیده، از جمله پیش‌بینی با استفاده از داده‌های سری زمانی مورد استفاده قرار گرفته است.

## ساختار شبکه‌های LSTM

شبکه‌های LSTM در واقع به جای نورون از واحدهای بزرگتر و پیچیده‌تری به نام بلوک حافظه تشکیل شده است. اصلی‌ترین مفهوم در مورد هر بلوک حافظه، وضعیت بلوک است که برآیند پیچیده و غیرخطی از ورودی بلوک، خروجی بلوک‌های مجاور، وضعیت بلوک در گام‌های پیشین، و وضعیت بلوک قبلی می‌باشد. هر بلوک حافظه از یک سلول حافظه<sup>۲</sup> و تعدادی درگاه تشکیل شده است. سلول حافظه وضعیت جاری بلوک را نگهداری می‌کند و درگاه‌ها شارش اطلاعات به سمت داخل و خارج مدیریت می‌کنند. شکل ۴-۹ ساختار یک بلوک حافظه LSTM را نشان می‌دهد.

<sup>۱</sup> Forget Gate

<sup>۲</sup> Memory Cell



شکل ۹-۴) بلوک LSTM [۴۴]

چنانکه از شکل ۹-۴ برمی آید، هر درگاه درواقع یک لایه شبکه عصبی با تابع فعال سازی سیگمید است که خروجی آن در انتها به یک عملگر ضرب عنصری<sup>۱</sup> ختم می شود. در معماری هوشیدر [۴۴] درگاهها، ورودی جاری بلوک ( $x_t$ ) و خروجی بلوک قبلی ( $h_{t-1}$ ) را بعنوان ورودی دریافت می نمایند و با عبور آنها از پالایه یک لایه سیگمیدی، عددی را در مقام خروجی تولید می کنند که در فاصله ۰ تا ۱ قرار دارد و مقدار آن معرف میزان تأثیرگذاری درگاه است. درگاه ورودی تعیین می کند که کدام داده ها باید در سلول حافظه بلوک ذخیره شوند، درگاه فراموشی تصمیم می گیرد که این داده ها تا چه زمانی باید در بلوک باقی بمانند و درگاه خروجی مشخص مینماید که کدام داده ها باید از بلوک خارج شوند. لازم به ذکر است که نویسندگان مرجع [۴۵] با ملاحظه کاستی های موجود در بلوک های LSTM زمان خود، وضعیت جاری بلوک را نیز به مجموعه ورودی های هر درگاه اضافه کردند و با نامگذاری اتصالات مذکور به اتصالات آفتابگردان<sup>۲</sup> راجح ترین نوع بلوک های LSTM یعنی LSTM آفتابگردانی<sup>۳</sup> را پدید آوردند.

<sup>۱</sup> Pointwise

<sup>۲</sup> Peephole

<sup>۳</sup> Peephole LSTM

## آموزش بلوک‌های LSTM

یک بلوک LSTM را می‌توان بعنوان یک تابع F در نظر گرفت که با دریافت ورودی جاری ( $x_t$ ) وضعیت موجود ( $c_{t-1}$ ) و خروجی بلوک قبلی ( $h_{t-1}$ )، خروجی بلوک جاری ( $h_t$ ) را تولید می‌کند.

$$h_t = F(x_t, h_{t-1}, c_{t-1}) \quad (۱-۴)$$

تابع F از زمان دریافت ورودی تا لحظه تولید خروجی مراحل زیر را طی می‌کند.

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} * c_{t-1} + b_i) \quad (۲-۴)$$

$$\hat{c} = \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \quad (۳-۴)$$

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} * c_{t-1} + b_f) \quad (۴-۴)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t \quad (۵-۴)$$

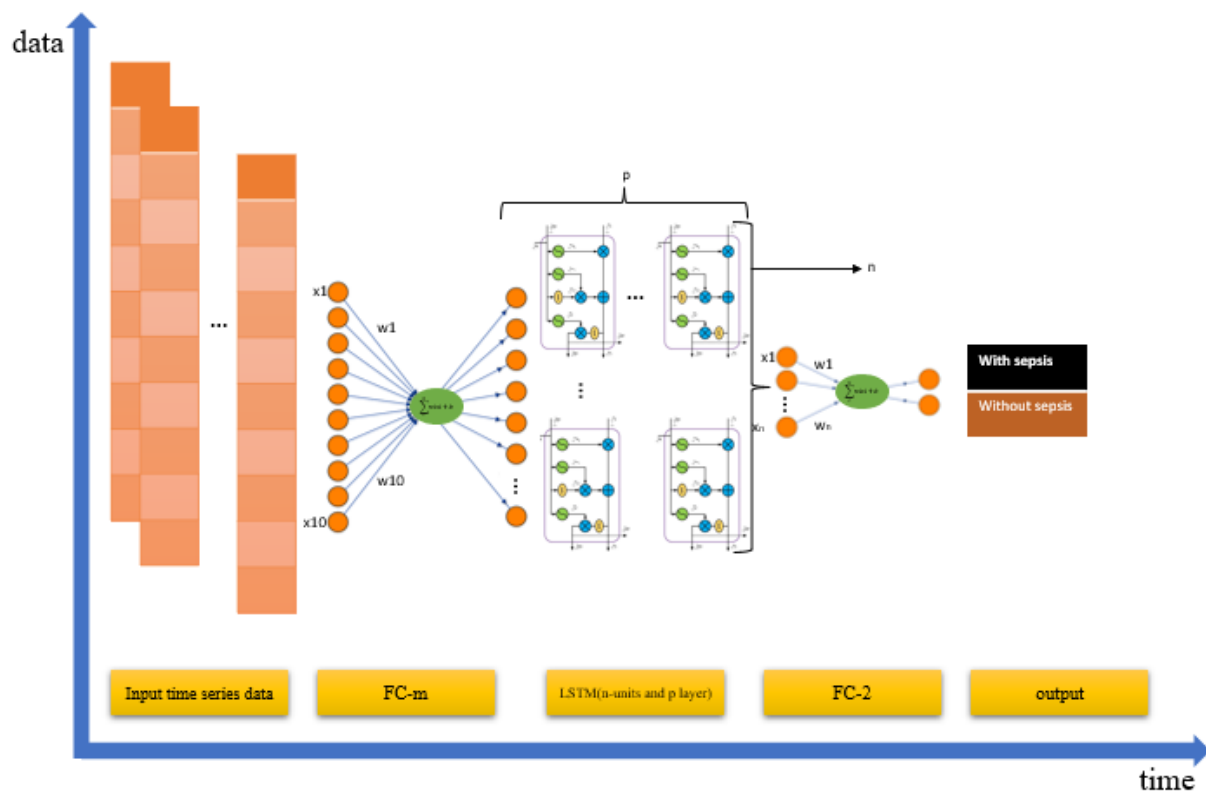
$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} * c_{t-1} + b_o) \quad (۶-۴)$$

$$h_t = o_t \odot \tanh(c_t) \quad (۷-۴)$$

در عبارات فوق،  $i_t$ ،  $f_t$ ،  $o_t$  به ترتیب معرف درگاه‌های ورودی، فراموشی و خروجی هستند. W و b پارامترهای مدل یعنی وزن‌های اتصالات و بایاس‌ها می‌باشند. Tanh،  $\sigma$  و  $\odot$  نیز توابع تانژانت هائپربولیک و سیگمید و عملگر ضرب عنصری را نشان می‌دهند.

با توجه به توضیحات مربوط به شبکه LSTM که در بخش قبلی گفته شد در این بخش به تشریح شبکه استفاده شده در این پژوهش می‌پردازیم. در شکل ۴-۱۰ مدل استفاده شده برای مدلسازی مجموعه دادگان این پژوهش نشان داده شده‌است. این مدل در دو بعد زمان و داده قرار دارد، به این معنا که برای تعیین کلاس هر نمونه در مجموعه دادگان ما یک سری مراحل روی آن‌ها انجام گیرد. در قسمت سمت چپ این مدل، مجموعه دادگان سری‌زمانی را داریم. که بسته به اینکه برای هر شماره بستری چند ساعت اطلاعات ثبت شده باشد، باید به همان تعداد ستون ۱۰ بخشی بعنوان ورودی سری‌زمانی به مدل داده شود. بعنوان مثال اگر برای یک بستری اطلاعات سه ساعت بستری او با توجه به محدودیت‌های گفته شده برای ما قابل قبول باشد، سه تا از این ستون‌ها را بعنوان ورودی به مدل داده می‌شود. در مرحله بعد برای کاهش تعداد ویژگی‌ها از یک شبکه عصبی پرسپترون چند لایه استفاده شده است که با اعمال یک تابع فعالساز روی مقادیر ورودی تعداد ویژگی‌ها را به m ویژگی کاهش می‌دهد. مقدار m

ویژگی می‌تواند متغیر باشد در واقع مقدار  $m$  در این مدل به عنوان یک پارامتر مورد ارزیابی در نظر گرفته می‌شود. اصلی‌ترین بخش این مدل شبکه LSTM استفاده شده در این مدل است. با توجه به معماری شبکه‌های LSTM دو متغیر  $n$  و  $p$  در این بخش از مدل بعنوان متغیر مورد ارزیابی مدل در نظر گرفته می‌شود. متغیر  $p$  تعداد لایه‌های شبکه‌ی LSTM و متغیر  $n$  تعداد بلوک‌های شبکه‌ی LSTM است. در نهایت مقادیر بدست آمده از شبکه LSTM به یک شبکه عصبی پرسپترون چند لایه با دو گره خروجی برای تعیین کلاس شروع سپسیس/عدم شروع سپسیس داده می‌شوند و کلاس نمونه شماره بستری مورد نظر مشخص می‌شود.



شکل ۴-۱۰) مدل پیشنهادی

بعد از انتخاب بهترین زوج مرتب سری زمانی باید دادگان این زوج مرتب را به سایر الگوریتم‌های یادگیری ماشین منتخب داد و نتایج آن را بررسی کرد. در پیاده‌سازی سه الگوریتم درخت تصمیم، جنگل تصادفی و شبکه عصبی چند لایه پرسپترون برخی از متغیرهای این الگوریتم‌ها مورد بررسی قرار داده شده‌است و

بهترین مقدار برای این متغیرها مشخص شده است. با استفاده از بهترین مقدار برای این متغیرها، الگوریتم‌ها برای زوج مرتب منتخب پیاده‌سازی شده و نتایج آن‌ها با هم و با الگوریتم اصلی این پژوهش مقایسه شده است. جدول ۴-۱ نتایج پیاده‌سازی این الگوریتم‌ها در مقادیر مختلف برای متغیرها و بهترین مقدار آن‌ها را نشان می‌دهد.

جدول ۴-۱) مقایسه متغیرهای مختلف برای الگوریتم‌های منتخب یادگیری ماشین و انتخاب بهترین آن‌ها

	<i>Parameters_space</i>	<i>Best_params</i>	<i>Best ROC-AUC</i>
<i>Decision Tree</i>	Max_depth=[۱,۱۰]	Max_depth=۵	۰/۶۵۳
<i>Random Forest</i>	N_estimators=[۱۰,۲۰۰] Max_depth=[۱,۱۰]	N_estimators=۵۰ Max_depth=۹	۰/۷۳۴
<i>MLP</i>	Hidden_layes_sizes= [(۱۰), (۵۰), (۱۰۰), (۱۰,۱۰), (۱۰,۱۰,۱۰)]	Hidden_layer_sizes=۱۰	۰/۸۶۰

#### ۴-۱-۱-۲- الگوریتم درخت تصمیم

روش‌های اصلی داده‌کاوی دو دسته هستند: توصیفی و پیش‌بینانه. وظایف توصیفی خواص عمومی داده‌ها را مشخص می‌کند و هدف آن پیدا کردن الگوهای قابل تفسیر توسط انسان برای داده‌ها است. وظایف پیش‌بینانه، پیش‌بینی رفتار آینده آن‌ها است و منظور از آن بکارگیری چند متغیر در پایگاه داده برای پیش‌بینی مقادیر آینده یا ناشناخته دیگر متغیرها می‌باشد. یکی از عملکردهای پیش‌بینی، دسته‌بندی است. دسته‌بندی فرایند یافتن مدلی است که با تشخیص دسته‌ها یا مفاهیم داده می‌تواند دسته ناشناخته اشیا دیگر را پیش‌بینی کند. در واقع دسته‌بندی یک تابع یادگیری است که یک قلم داده را به یکی از دسته‌های از قبل تعریف شده نگاشت می‌کند. یکی از روش‌های متداول دسته‌بندی درخت تصمیم می‌باشد. این درخت ساختاری شبیه به فلوچارت دارد که بالاترین گره، ریشه درخت است. هر شاخه بیانگر خروجی‌های آزمایش و گره‌های برگ دسته یا توزیع دسته‌ها را نمایش می‌دهد [۴۶]. قوانین ایجادشده توسط درخت تصمیم بصورت اگر و آنگاه بیان می‌شوند.

#### ۴-۱-۱-۳- الگوریتم جنگل تصادفی

جنگل تصادفی یا جنگل‌های تصمیم تصادفی یک روش یادگیری ترکیبی برای دسته‌بندی، رگرسیون می‌باشد، که بر اساس ساختاری متشکل از شمار بسیاری درخت تصمیم، بر روی زمان آموزش و خروجی کلاس‌ها (کلاس‌بندی) یا برای پیش‌بینی‌های هر درخت به شکل مجزا، کار می‌کنند. جنگل‌های تصادفی برای درختان تصمیم که در مجموعه‌ی آموزشی دچار بیش‌برازش می‌شوند، مناسب هستند. عملکرد جنگل تصادفی معمولاً بهتر از درخت تصمیم است، اما این بهبود عملکرد تا حدی به نوع داده هم بستگی دارد [۴۷].

#### ۴-۱-۱-۴- شبکه عصبی پرسپترون چند لایه

پرسپترون چند لایه<sup>۱</sup> دسته‌ای از شبکه‌های عصبی مصنوعی پیشخور است. یک پرسپترون چندلایه شامل حداقل سه لایه گره است: یک لایه ورودی، یک لایه پنهان و یک لایه خروجی. به جز گره‌های ورودی، هر گره یک نورون است که از یک تابع فعال‌سازی غیر خطی استفاده می‌کند پرسپترون چندلایه از تکنیک یادگیری نظارت شده به نام بازپرداخت برای آموزش استفاده می‌کند. لایه‌های متعدد آن و فعال‌سازی غیر خطی آن پرسپترون چندلایه را از یک پرسپترون خطی متمایز می‌کند. در واقع می‌تواند داده‌هایی را متمایز کند که به صورت خطی قابل تفکیک نیستند.

اگر یک پرسپترون چند لایه، تابع فعال‌سازی خطی در تمام نورون‌ها داشته باشد، در واقع با این تابع خطی ورودی‌های وزن‌دار هر نورون را ترسیم می‌کند. سپس با استفاده از جبر خطی نشان می‌دهد که هر عددی مربوط به لایه‌ها را می‌توان به یک مدل ورودی-خروجی دو لایه کاهش داد. در پرسپترون چندلایه، برخی از نورون‌ها از یک تابع فعال غیرخطی استفاده می‌کنند که برای مدل‌سازی فرکانس پتانسیل‌های عمل یا شلیک نورون‌های بیولوژیکی توسعه داده شده است [۴۸].

#### ۴-۲- روش‌های انتخاب مجموعه ویژگی

یکی از چالش‌های قابل بررسی در مسائل پیش‌بینی بیماری، مجموعه ویژگی‌های مورد استفاده در مسئله دسته‌بندی است. با توجه به اینکه ذخیره سازی مقادیر تعداد زیادی از ویژگی‌های بیماران برای مسئولان درمان کار راحتی نیست و همواره داده‌های کمی از بیماران در دسترس می‌باشد؛ لذا استفاده از حداقل

---

<sup>۱</sup> Multilayer Perceptron



تعداد ویژگی با بیشترین دقت در پیش‌بینی از اهمیت بالایی برخوردار است. در این بخش می‌خواهیم اهمیت ویژگی‌هایی که تا این بخش از پژوهش برای پیش‌بینی استفاده کرده‌ایم را بررسی کنیم. برای بررسی اهمیت هر کدام از ویژگی‌ها از دو روش Stepwise و روش حذف بازگشتی ویژگی<sup>۱</sup> استفاده شده‌است. در تعیین ویژگی‌های پر اهمیت در پیش‌بینی، ملاک مدلسازی، بهترین مدل از بین ۴ مدل دسته‌بند می‌باشد. در ادامه به شرح استفاده از این دو روش می‌پردازیم.

از دیدگاه طبقه‌بندی، روش‌های انتخاب ویژگی را می‌توان در چهار دسته زیر طبقه‌بندی کرد:

- روش‌های فیلتر<sup>۲</sup>
- روش‌های بسته‌بند<sup>۳</sup>
- روش‌های تعبیه‌شده<sup>۴</sup>
- روش‌های ترکیبی<sup>۵</sup>

#### ۴-۲-۱ روش‌های انتخاب ویژگی بسته‌بند

همانطور که در شکل ۴-۱۱ نشان داده شده است، این دسته از روش‌های انتخاب ویژگی، در هر مرحله، زیرمجموعه‌ای از ویژگی‌ها در فضای ویژگی را انتخاب می‌کند و عملکرد الگوریتم یادگیری ماشین روی این زیر مجموعه را می‌سنجند. از نتیجه عملکرد الگوریتم یادگیری ماشین، برای ارزیابی زیرمجموعه انتخاب شده از فضای ویژگی استفاده می‌شود. به عنوان نمونه، مشخص می‌شود که کدام یک از زیر مجموعه‌های انتخاب شده، بیشترین تاثیر را در افزایش دقت یک مسأله دسته‌بندی خواهند داشت. از این روش مدل‌سازی و جستجوی ویژگی برای انتخاب ویژگی جهت آموزش انواع روش‌های یادگیری ماشین استفاده می‌شود. روش‌های انتخاب ویژگی بسته‌بند بطور کلی به دو دسته گام‌به‌گام<sup>۶</sup> و روش حذف بازگشتی ویژگی تقسیم می‌شوند. روش گام‌به‌گام دو حالت انتخاب ویژگی مستقیم<sup>۷</sup> و انتخاب ویژگی معکوس<sup>۸</sup> دارد که در ادامه این سه روش توضیح داده خواهد شد.

---

<sup>۱</sup> Recursive Feature Elimination

<sup>۲</sup> Filter

<sup>۳</sup> Wrapper

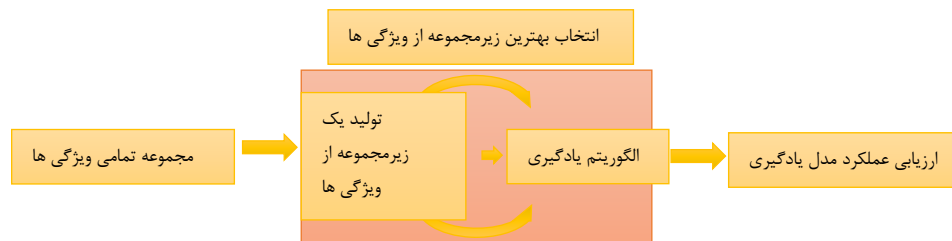
<sup>۴</sup> Embedded

<sup>۵</sup> Hybrid

<sup>۶</sup> Stepwise

<sup>۷</sup> Forward Feature Selection

<sup>۸</sup> Backward Feature Selection



شکل ۴-۱۱) روال کلی انتخاب ویژگی بسته‌بند

#### ۴-۲-۱-۱- انتخاب ویژگی مستقیم

در روش‌های انتخاب ویژگی مستقیم، ابتدا یک زیر مجموعه تهی از ویژگی‌ها ساخته می‌شود. سپس در هر مرحله، ویژگی‌هایی که بهترین عملکرد را برای مدل یادگیری به ارمغان می‌آورند، به این زیر مجموعه اضافه می‌شوند.

#### ۴-۲-۱-۲- انتخاب ویژگی معکوس

در روش‌های انتخاب ویژگی معکوس، ابتدا تمامی ویژگی‌ها در زیر مجموعه حضور دارند. سپس در هر مرحله، بدترین ویژگی‌ها از زیر مجموعه حذف می‌شوند ویژگی‌هایی که حذف آن‌ها، باعث ایجاد کمترین کاهش در عملکرد، دقت روش یادگیری ماشین می‌شوند.

#### روش حذف بازگشتی ویژگی

روش حذف بازگشتی ویژگی، یک روش حریصانه<sup>۱</sup> برای انتخاب ویژگی است. در این روش، ویژگی‌ها به طور بازگشتی و با در نظر گرفتن مجموعه‌های کوچک و کوچک‌تر از ویژگی‌ها انتخاب می‌شوند. در این روش، ویژگی‌ها بر اساس مرتبه حذف شدن آن‌ها از فضای ویژگی رتبه‌بندی می‌شوند [۴۹].

<sup>1</sup> Greedy

#### ۴-۳- روش‌های پرکردن داده‌های از دست رفته

همواره در داده‌های پزشکی موضوع پرکردن داده‌های از دست‌رفته یک موضوع قابل بحث می‌باشد چرا که به علت‌های مختلف که در ابتدای همین فصل گفته شد بخش زیادی از داده‌های پزشکی ذخیره نمی‌شوند. برای پرکردن داده‌های از دست‌رفته روش‌های مختلفی وجود دارد. در این پژوهش ۴ روش معروف برای پرکردن داده‌های از دست‌رفته استفاده شده‌است. مقادیر از دست‌رفته مجموعه دادگان زوج‌مرتب منتخب را با مقدار میانه، حداکثر مقدار، حداقل مقدار، مقدار صفر و میانگین پر کرده و مجموعه دادگان حاصل به شبکه عصبی منتخب داده شده‌است. و نتایج آن با هم مقایسه شده‌است.

## فصل ۵- ارزیابی روش‌های پیشنهادی

در این فصل به ارزیابی نتایج اجرای روش‌های مطرح شده در فصل قبل می‌پردازیم. داده‌های مورد بررسی در فصل سوم توضیح داده شده‌اند. در این تحقیق به پیش‌بینی شروع سپسیس در بیماران بخش مراقبت‌های ویژه با تمرکز بر دو موضوع مورد بررسی قرار گرفت که عبارتند از تعیین بهترین بازه‌زمانی برای دو بازه‌زمانی نگاه به گذشته و بازه‌ی زمانی پیش‌بینی و تعیین بهترین مجموعه ویژگی پیش‌بینی کننده.

## ۵-۱- ارزیابی بازه‌های زمانی پیش‌بینی و نگاه به گذشته

همانطور که در فصل قبل گفته شد، در این ارزیابی ۴ روش دسته‌بندی، مورد ارزیابی قرار گرفته‌است. برای پیاده‌سازی دسته‌بندها از زبان برنامه نویسی پایتون<sup>۱</sup> نسخه ۹.۳ استفاده شده‌است.

برای این کار مقادیر ورودی را به مدل داده شده‌است و برچسب صفر و یک را برای مقادیر ویژگی‌های ورودی تعیین شده‌است. برای دسته‌بند جنگل تصادفی از تابع `randomforestclassifier` استفاده شده- است برای دسته‌بند درخت تصمیم از تابع `decisiontreeclassifier` استفاده شده‌است برای `MLP` از تابع `MLPclassifier` با ۳ عصب در لایه پنهان و توابع انتقال لگاریتمی و تانژانت استفاده شده‌است. در این الگوریتم‌ها که همه از پکیج `skit-learn` استفاده شده است و با دادن مقادیر ورودی، کلاس صفر و یک برای نمونه‌ها مشخص می‌شود. برای پیاده‌سازی الگوریتم شبکه عصبی بازگشتی `LSTM` از متد `LSTM` از کتابخانه پایترچ<sup>۲</sup> استفاده شده‌است. معیارهایی که در این تحقیق برای ارزیابی دسته‌ها مورد استفاده قرار گرفته‌است عبارتند از دقت<sup>۳</sup>، صحت<sup>۴</sup>، `CWA` و `AUC-ROC`.

فرمول معیار حساسیت بصورت

$$Sensitivity = \frac{TP}{TP + FN} \quad (۵-۱)$$

تعریف می‌شود. در رابطه‌ی بالا `TP` تعداد مواردی است که دسته‌بند به درستی شروع سپسیس را مثبت تشخیص داده‌است و `FN` تعداد مواردی است که دسته‌بند یک مورد شروع سپسیس را به اشتباه عدم شروع سپسیس تشخیص داده‌است. معیار حساسیت در واقع قابلیت دسته‌بند در تشخیص درست شروع سپسیس را نشان می‌دهد.

فرمول معیار تشخیص بصورت

---

<sup>۱</sup> Python

<sup>۲</sup> Pytorch

<sup>۳</sup> Accuracy

<sup>۴</sup> Precision

$$Specificity = \frac{TN}{TN + FP} \quad (2-5)$$

در رابطه ی بالا TN تعداد مواردی است که دسته‌بند به درستی پیش‌بینی شروع سپسیس را عدم شروع سپسیس تشخیص داده‌است. معیار تشخیص قابلیت دسته‌بند در پیش‌بینی درست عدم شروع سپسیس را نشان می‌دهد.

در سیستم‌های پزشکی، اما، ترکیب وزن دار معیار حساسیت و تشخیص اهمیت دارد [۴۹]. به عبارت دیگر با وجود اینکه هدف پیش‌بینی زودهنگام شروع سپسیس است و مقدار معیار تشخیص باید بالا باشد، اما پیش‌بینی زودهنگام شروع سپسیس باید بگونه ای باشد که FN به حداقل برسد زیرا در این مسئله برچسب اشتباه شروع سپسیس به عدم شروع سپسیس بسیار خطرناک‌تر از زدن برچسب شروع سپسیس به عدم شروع سپسیس است. بنابراین معیار اصلی مورد استفاده در این تحقیق برای ارزیابی دسته‌بندها از CWA استفاده شده است این معیار ترکیبی وزن دار از حساسیت و تشخیص است که در (۴۹) ارائه شده‌است برای دسته‌بندی دودویی معیار CWA به فرم

$$CWA = w * sensitivity + (1 - w) * specificity \quad (3-5)$$

تعریف می‌شود. در رابطه ی بالا W وزنی است که برای کلاس مثبت یا همان برچسب شروع سپسیس در نظر گرفته می‌شود. در این تحقیق برای W مقدار ۰/۹ در نظر گرفته شده‌است.

فرمول معیار دقت بصورت

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4-5)$$

تعریف می‌شود.

فرمول معیار صحت بصورت

$$Precision = \frac{TP}{TP + FP} \quad (5-5)$$

تعریف می‌شود.

منحنی AUC-ROC<sup>1</sup> به معنی سطح زیر منحنی مشخصه‌ی عملکرد سیستم است. این منحنی یکی از مهم‌ترین معیارهای ارزیابی برای بررسی کردن عملکرد هر مدل طبقه‌بندی به حساب می‌آید. این معیار

<sup>1</sup> Area Under the Curve - Receiver Operating Characteristics

همچنین به شکل  $^1\text{AUROC}$  نیز نوشته می‌شود به معنی ناحیه ی زیر منحنی مشخصه عملکرد سیستم است. منحنی  $\text{AUC-ROC}$  معیار سنجش عملکرد برای مسائل طبقه بندی است.  $\text{ROC}$  یک منحنی احتمالاتی هست و  $\text{AUC}$  درجه یا معیار را بازنمایی می‌کند. این معیار میگوید چه مقدار مدل قادر است بین کلاس‌ها تمایز ایجاد کند.  $\text{AUC}$  بالاتر مدل در پیش‌بینی کردن  $\cdot$  به عنوان  $\cdot$  و  $\cdot$  به عنوان  $\cdot$ ، بهتر است. به تناسب، با  $\text{AUC}$  بالاتر، مدل در تمایز ایجاد کردن بین بیماران مبتلا به بیماری و بیماران رو به بهبود (بدون بیماری) بهتر عمل می‌کند. منحنی  $\text{ROC}$  به کمک دو محور  $\text{TPR}^2$  و  $\text{FPR}^3$  کشیده می‌شود. به طوریکه محور  $x$  برای  $\text{FPR}$  و محور  $y$  برای  $\text{TPR}$  در نظر گرفته شده‌است [۵۰].

### پیاده‌سازی مدل LSTM پروژه

همانطور که در فصل قبل گفته شد، برای پیاده‌سازی مدل مربوط به این پژوهش که هسته آن شبکه عصبی بازگشتی  $\text{LSTM}$  است، ۳ متغیر وجود دارد. متغیر اول ( $m$ ) مربوط به تعداد نودهای خروجی شبکه عصبی چند لایه پرسپترون اول، متغیر دوم ( $n$ ) مربوط به تعداد بلوک‌های استفاده شده در شبکه عصبی  $\text{LSTM}$ ، متغیر سوم ( $p$ ) مربوط به تعداد لایه‌های این شبکه می‌باشد. با بررسی حالت‌های مختلف پیاده‌سازی این مدل برای بازه‌های زمانی مختلف نتایج آن بصورت جدول ۵-۱ بدست آمد.

همانطور که در جدول ۵-۱ آورده شده‌است، بهترین ترکیب برای  $m$ ،  $n$  و  $p$  به ترتیب ۸، ۴ و ۲ است. بنظر می‌آید که استفاده از ۳ لایه در این مدل، آن را بسیار قدرتمند می‌کند تا حدی که سریعاً به بیش‌برازش منجر می‌شود. از طرفی ترکیب ۶، ۶ و ۲ از قدرت کافی برای بازنمایی اطلاعات برخوردار نیست. همچنین یک ترکیب دیگر از پارامترها یعنی ۸، ۴ و ۱ مقدار بالایی برای شاخص  $\text{CWA}$  بدست می‌دهد، اما در این پژوهش ترکیب ۸، ۴ و ۲ ترجیح داده شده‌است. در این تصمیم تنها معیار انتخاب عملکرد مدل روی داده‌های آزمون نبوده است، چرا که دقت مدل روی داده‌های آموزشی این امکان را فراهم می‌آورد که ترکیب‌های مشکوک به بیش‌برازش، شناسایی شوند.

<sup>1</sup> Area Under the Receiver Operating Characteristics

<sup>2</sup> True Positive Rate

<sup>3</sup> False Positive Rate

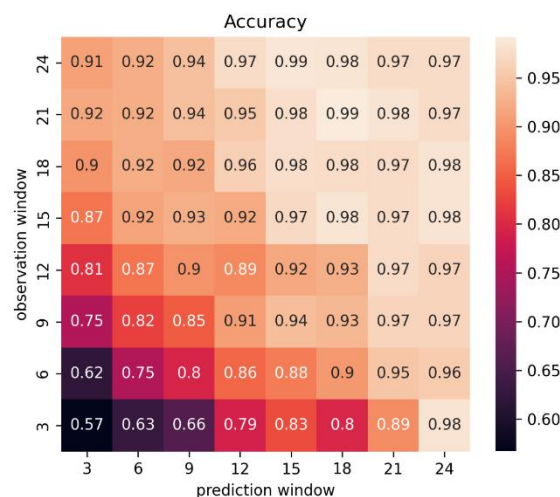
جدول ۵-۱) مقایسه نتایج پیاده‌سازی حالت‌های مختلف شبکه LSTM

	<i>Accuracy Train data</i>	<i>Accuracy test data</i>	<i>Precision Train data</i>	<i>Precision test data</i>	<i>CWA train data</i>	<i>CWA test data</i>
<i>LSTM-۶۶۲</i>	۰/۸۰۵۶	۰/۸۰۳۶	۰/۸۷۵۲	۰/۸۵۱۱	۰/۷۸۹۳	۰/۷۷۳۰
<i>LSTM-۸۴۱</i>	۰/۸۲۸۸	۰/۸۰۳۶	۰/۹۰۶۲	۰/۹۲۳۱	۰/۸۴۷۱	۰/۸۳۸۷
<i>LSTM-۸۴۲</i>	۰/۸۳۵۶	۰/۸۲۱۴	۰/۹۱۴۵	۰/۹۰۷۰	۰/۸۴۲۲	۰/۸۴۰۳
<i>LSTM-۸۴۳</i>	۰/۹۳۵۸	۰/۸۱۲۵	۰/۹۷۲۵	۰/۹۰۴۸	۰/۷۰۲۵	۰/۷۹۳۴
<i>LSTM-۹۵۲</i>	۰/۹۲۷۸	۰/۸۱۲۵	۰/۹۵۳۱	۰/۹۷۲۲	۰/۸۱۲۲	۰/۸۰۱۸

بعد از مشخص شدن مدل و درک آن باید تمام داده‌های هر زوج مرتب، به منظور مدل‌سازی به این مدل داده شود و بهترین زوج مرتب بازه‌زمانی مشخص شود. در این میان باید به نکته‌ای توجه شود که به شمول مدل مربوط است. هر اندازه جمع طول بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیشینی طولانی‌تر باشد داده‌های کمتری امکان دارند که به مدل داده شوند. چرا که باید داده‌هایی که طول کوتاه‌تری از مجموع این پنجره‌ها دارند را کنار بگذاریم. در نتیجه نوعی دادوستد میان کیفیت مدل و شمول آن وجود دارد. علاوه بر این باید به عدم قطعیت نتایج دقت شود، چراکه هرچقدر شمول مدل کمتر باشد، مدل روی نمونه‌های کمتری ارزیابی می‌شود و در نتیجه تخمین‌های ما از نتایج، واریانس بالاتری خواهد داشت و عدم قطعیت در مورد آن بالاتر خواهد بود.

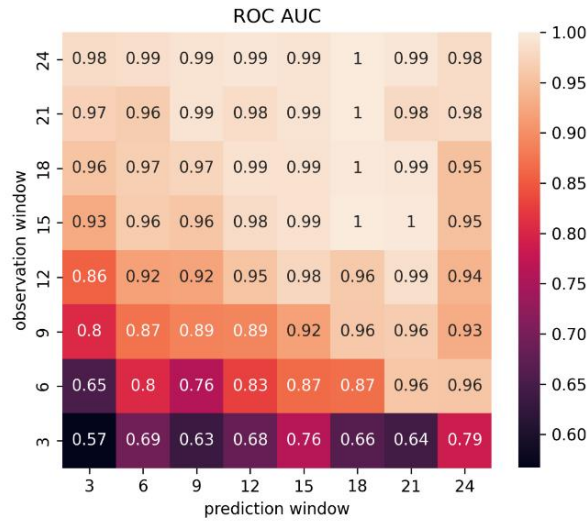
در ادامه، ماتریس‌هایی از شاخص‌های ارزیابی مدل بر روی داده‌های تست برای انتخاب بهترین طول بازه‌زمانی نگاه به گذشته و طول بازه‌زمانی پیشینی ارائه می‌دهیم. این شاخص‌ها علاوه بر داده‌های آزمون برای داده‌های آموزشی هم محاسبه شده‌اند اما تصاویر مرتبط با این شاخص‌ها در این بخش از پایان نامه آورده نشده‌است در ادامه زوج مرتب (۶ و ۱۲) بیانگر عدد اول یعنی ۶ بعنوان بازه‌زمانی پیش-بینی و عدد دوم یعنی ۱۲ بعنوان بازه‌زمانی مشاهده تعریف می‌شود.





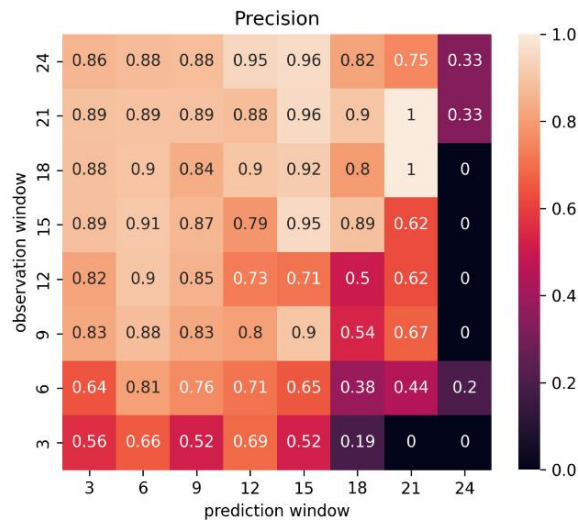
شکل ۵-۱) نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار دقت

همانطور که در شکل ۵-۱ مشاهده می‌کنید، دقت مدل برای ترکیب‌های مختلف از طول بازه‌های زمانی نگاه به گذشته و پیش‌بینی اندازه‌گیری شده‌است. پایین‌ترین دقت مربوط به مدل با زوج مرتب (۳ و ۳) است. درحالی که این مدل بالاترین شمول را دارد؛ چرا که نمونه‌ای بخاطر کوتاه بودن سری زمانی از مجموعه داده‌ها حذف نشده‌است. در هر خانه از شکل ۵-۱ دقت دو تصمیم وجود دارد: اول اینکه طول بازه‌زمانی نگاه به گذشته را افزایش داده و دوم اینکه طول بازه زمانی پیش‌بینی افزایش داده شود. اگر بر این اساس تصمیم گرفته شود، می‌توان دید که بهترین گام برای افزایش دقت مدل افزایش طول بازه‌زمانی پیش‌بینی است. پس از آن بهترین گزینه تا جایی که طول بازه‌زمانی نگاه به گذشته برابر ۱۲ شود، افزایش طول بازه‌زمانی پیش‌بینی است. بر اساس این معیار، در این پژوهش تصمیم گرفته‌شد که طول بازه‌زمانی پیش‌بینی را برابر ۶ و بازه‌زمانی نگاه به گذشته را برابر ۱۲ قرار داده‌شود.



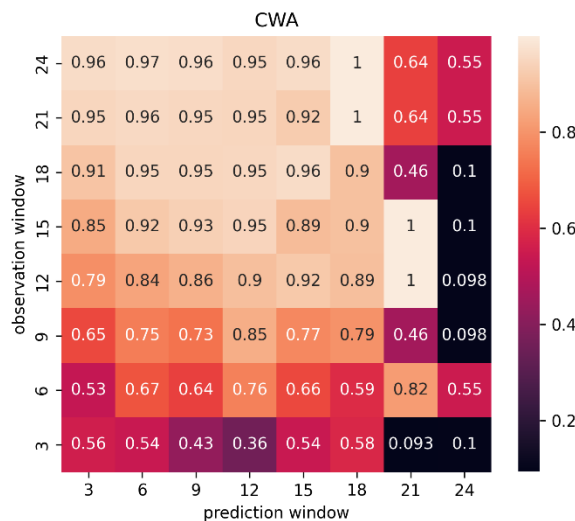
شکل ۵-۲) مقایسه نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار AUC-ROC

با توجه به اینکه در خانه‌های قطری شکل ۵-۲، مقدار شمول ثابت می‌ماند، شیوه مناسب برای انتخاب بهترین طول بازه‌های زمانی، مقایسه مقادیر معیار ارزیابی بصورت قطری است. با مقدار بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی مقدار شمول تقریباً برابر ۸۷٪ است. در قطر بعدی، شمول به صورت چشمگیری به مقدار ۵۳٪ کاهش می‌یابد.



شکل ۵-۳) مقایسه نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار صحت

دادوستد میان شاخص‌های عملکرد و شمول در بسیاری از پروژه‌های یادگیری ماشین و یادگیری عمیق مشاهده می‌شود و تنها مختص به این پیاده‌سازی از روش پیشنهادی ما نمی‌شود. با توجه به شکل ۳-۵ و ۴-۵ یکی از گزینه‌هایی که می‌توان به جای انتخاب زوج مرتب (۶ و ۱۲) داشت، زوج مرتب (۳ و ۱۲) است. اما با توجه به اینکه هزینه تشخیص اشتباه در این حوزه بسیار زیاد است، داشتن Precision بالاتر مطلوب تر است و باعث می‌شود گزینه زوج مرتب (۶ و ۱۲) بهتر از (۳ و ۱۲) باشد.



شکل ۴-۵) مقایسه نتایج پیاده‌سازی شبکه LSTM در بازه‌های زمانی مختلف بر اساس معیار CWA

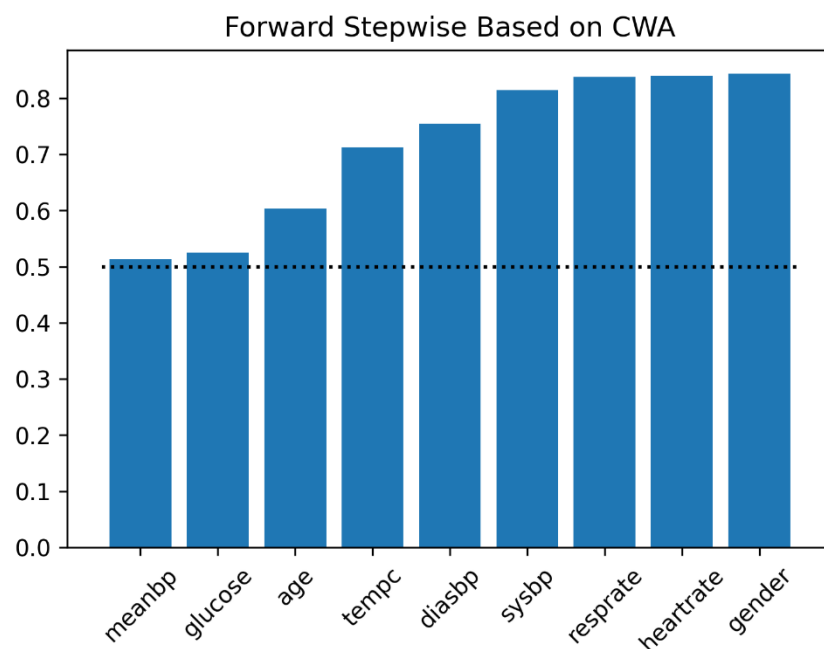
بعد از انتخاب بهترین زوج مرتب، مجموعه دادگان این زوج مرتب را به سایر الگوریتم‌های یادگیری ماشین مورد بررسی در این پژوهش داده و نتایج آن‌ها با هم و با سایر الگوریتم‌ها مقایسه شده‌است. نتایج حاصل از اجرای این ۴ الگوریتم دسته‌بند در جدول ۵-۲ آمده‌است. با توجه به این نتایج، می‌توان مشاهده کرد که استفاده از شبکه عصبی بازگشتی LSTM بهترین شیوه برای مدل‌سازی این داده‌ها است. این نتایج با دانش قبلی ما مبنی بر اینکه داده خاصیت سری زمانی دارد منطبق است. چراکه شبکه عصبی چندلایه علی‌رغم قدرت بالایی که در بازنمایی اطلاعات دارد، نسبت به شاخص‌های بدست آمده برای LSTM خیلی ضعیف‌تر است. علاوه بر این، شاخص‌های بدست آمده برای Decision Tree دقت پایینی را ثبت کرده‌اند. علاوه بر این، جدول ۵-۲ نشان می‌دهد که قدرت بازنمایی این مدل حتی با استفاده از Random Forest افزایش چشمگیری نداشته است و بسرعت موجب بیش‌برازش شده‌است.

جدول (۵-۲) مقایسه الگوریتم‌های یادگیری ماشین با شبکه LSTM

	<i>Accuracy train data</i>	<i>Accuracy test data</i>	<i>Precision train data</i>	<i>Precision test data</i>	<i>AUC- ROC train data</i>	<i>AUC- ROC test data</i>	<i>CWA train data</i>	<i>CWA test data</i>
<i>Decision Tree</i>	۰/۶۳۲۴	۰/۶۶۵۲	۰/۷۱۲۲	۰/۷۰۱۱	۰/۶۸۲۲	۰/۶۶۳۲	۰/۶۱۵۴	۰/۶۳۰۱
<i>Random Forest</i>	۰/۶۸۵۰	۰/۶۷۸۶	۰/۶۹۸۷	۰/۷۰۶۵	۰/۶۴۵۸	۰/۶۷۷۰	۰/۶۷۵۲	۰/۶۵۵۳
<i>MLP</i>	۰/۷۷۷۲۱	۰/۷۰۹۸	۰/۷۳۶۸	۰/۷۱۸۴	۰/۷۶۵۰	۰/۰۷۰۹۲	۰/۶۸۹۸	۰/۶۹۸۲
<i>LSTM</i>	۰/۸۹۲۳	۰/۸۷۰۵	۰/۸۸۸۹	۰/۹۰۱۰	۰/۹۴۲۷	۰/۹۲۵۲	۰/۸۳۳۷	۰/۸۴۰۳

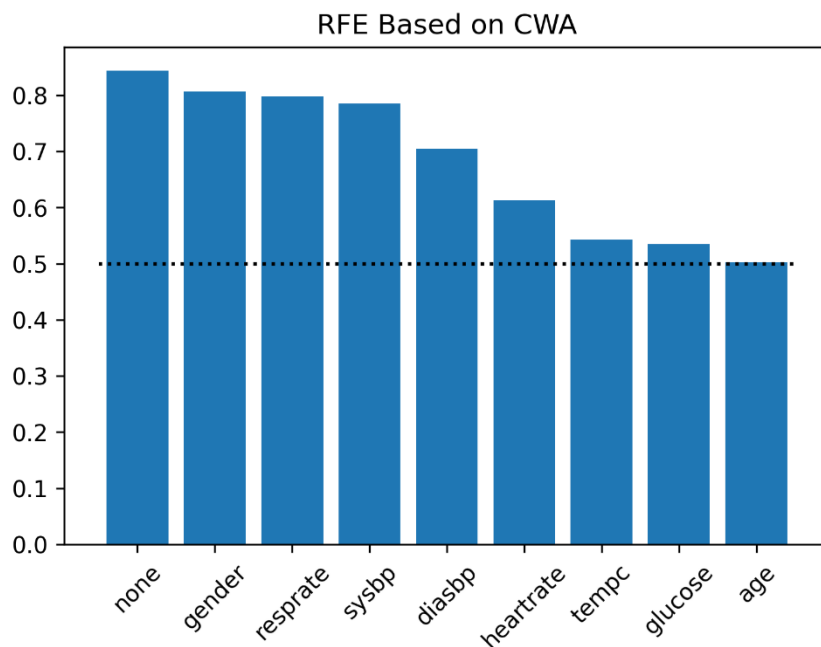
## ۵-۲- ارزیابی روش‌های انتخاب ویژگی

برای انتخاب بهترین مجموعه ویژگی از دو روش انتخاب ویژگی REF و Stepwise استفاده شده‌است در پیاده‌سازی این دو روش ملاک تصمیم‌گیری، الگوریتم دسته‌بند منتخب می‌باشد. نتایج حاصل از پیاده‌سازی این دو الگوریتم در نمودارهای ۵-۵ و ۵-۶ نشان داده شده‌است.



شکل (۵-۵) نمودار مقایسه تاثیر ویژگی‌های مختلف در پیاده‌سازی شبکه LSTM به روش forward stepwise

خط نقطه چین در شکل‌های ۵-۵ و ۵-۶ نشان‌دهنده‌ی مدل تصادفی هستند. همانطور که مشخص است، در روش Stepwise مهمترین ویژگی‌ها از چپ به راست لیست شده‌اند. اما دقت مدل تا قبل از افزودن ویژگی Tempc افزایش چشمگیری نداشته‌است بنظر می‌آید که عدم افزایش دقت مدل مربوط به ناکافی بودن اطلاعات برای بازنمایی مناسب در فضای پسین باشد. همچنین نتایج نشان می‌دهد که استفاده از تنها ۸ ستون برای بدست آوردن مقدار خوبی از شاخص cwa کافی باشد.

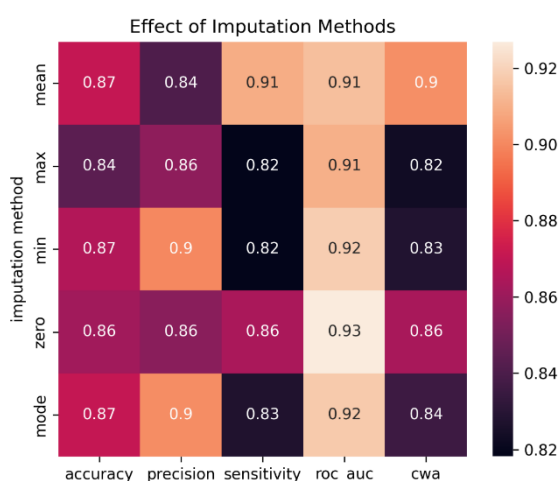


شکل ۵-۶) نمودار مقایسه تاثیر ویژگی‌های مختلف در پیاده‌سازی شبکه LSTM به روش RFE

در شکل ۵-۶ عبارت none به معنی شاخص cwa در حالی است که هیچ یک از ویژگی‌ها از ورودی حذف نشده‌است. این نمودار نشان می‌دهد که کم اهمیت‌ترین ویژگی‌ها gender و resprate هستند. این نتایج در نمودار ۵-۶ نیز تایید می‌شود (به ترتیب آخرین و یکی پیش از آخرین ویژگی). با توجه به اینکه از روش‌های مختلف برای اندازه‌گیری اهمیت ویژگی‌ها استفاده شده است، طبیعی است که نتایج کاملاً یکسان نباشد. اما نتایج بدست آمده در مورد اهمیت ویژگی‌ها از این دو نمودار تا حد خوبی توافق دارند.

### ۵-۳- ارزیابی روش‌های پرکردن داده‌های از دست رفته

بخشی از مجموعه دادگان زوج‌مرتب منتخب بعنوان داده از دست‌رفته ثبت شده‌اند. همانطور که فصل قبل گفته‌شد، در این پژوهش از روش‌های مختلف برای پرکردن داده‌های از دست رفته استفاده شده‌است. با توجه به شکل ۵-۷، بهترین روش برای پرکردن داده‌های از دست‌رفته، روش پرکردن با میانگین اطلاعات فرد است. نکته جالب توجه در این نمودار شباهت بسیار زیاد روش min و max است، که بدترین روش‌های پرکردن داده‌های از دست‌رفته هستند؛ چرا که ممکن است در اطلاعات فرد min و max داده‌های پرت محسوب شوند. یک آماره دیگر که کمی بهتر از min و max عملکرد است، mode بوده است که به داده‌های پرت حساسیت ندارد. آخرین روش مورد بررسی پرکردن با مقدار صفر بوده است که در نمودار با عبارت zero نشان‌داده شده‌است این روش نسبت به روش میانگین، از مساحت زیر نمودار عملکرد شاخص یا ROC - AUC بالاتری برخوردار است اما این افزایش دو درصدی را در ازای کاهش چشمگیر شاخص‌های دیگر بدست آمده‌است.



شکل ۵-۷) مقایسه روش‌های پرکردن داده‌های از دست رفته

داده‌های مورد بررسی برای دسته‌بندی نسبت به روش‌های پیشین متفاوت است. اما با این حال در این تحقیق، مقادیر بدست آمده برای معیار حساسیت نسبت به روش‌های پیشین میزان ۱۰ درصد بهتر است. بطور کلی با بررسی روش‌های دسته‌بندی شناخته‌شده مشخص شد که این روش‌ها قابلیت پیش‌بینی

شروع سپسیس را دارند اما ذکر این نکته اهمیت دارد که در مقالات گذشته و همچنین در این تحقیق نرخ FN همچنان به میزان بالایی است و باید تاکید کرد که این خطای سیستم بسیار هزینه گزافی خواهدداشت . چرا که عدم تشخیص شروع سپسیس خطر جانی برای بیماران بستری در ICU دارد.

## فصل ۶- نتیجه گیری و پیشنهادات



با توجه به اینکه در ICU میزان مرگ و میر بیماران در اثر ابتلا به سپسیس زیاد است و اینکه عفونت خون یا همان سپسیس در بیماران بصورت ناگهانی شروع نمی‌شود؛ بلکه با بررسی تغییرات برخی مقادیر علائم حیاتی و مقادیر آزمایشگاهی در دسترس، می‌توان زمان شروع آن را پیش‌بینی کرد. با پیش‌بینی به‌هنگام سپسیس می‌توان از عوارض ناشی از آن مانند از کارافتادن اندام‌ها و در نهایت مرگ و میر جلوگیری کرد. بنابراین طراحی یک سیستم تصمیم‌یار برای تشخیص زودهنگام شروع سپسیس می‌تواند بسیار مفید باشد.

در این تحقیق پیرو پژوهش‌های پیشین، ابتدا داده‌های مورد بررسی بر اساس مراحل که در بخش مراحل انتخاب مجموعه داده فصل سوم گفته شده، به برچسب ابتلا به سپسیس و عدم ابتلا به سپسیس نگاشت می‌شوند. با این کار مسأله پیش‌بینی شروع سپسیس به یک مسأله دسته‌بندی تبدیل شده است؛ لذا می‌توان از روش‌های دسته‌بندی برای مدلسازی استفاده کرد، در این تحقیق ۴ روش شناخته شده دسته‌بندی مورد ارزیابی قرار گرفت. نتایج ارزیابی نشان داد که این روش‌ها قابلیت پیش‌بینی برچسب‌ها را دارند. در پژوهش‌های پیشین مربوط به پیش‌بینی سپسیس میزان دقت این پیش‌بینی به حد مطلوب نرسیده است. با بررسی و تعمق بیشتر در مسأله، دو نکته به ذهن رسید.

نکته اول این است که انتخاب ویژگی‌های پیش‌بینی کننده، احتیاج به دقت بیشتری دارد. بنابراین با توجه به تحقیقات پیشین ده ویژگی انتخاب می‌شود و با استفاده از روش‌های انتخاب ویژگی که در فصل چهارم گفته شد، بهترین مجموعه ویژگی‌ها انتخاب شدند و سپس مدلسازی‌های لازم انجام شد.

نکته دوم این است که با توجه به تحقیقات پیشین بازه‌های زمانی محدودی برای بازه‌زمانی نگاه به گذشته و بازه‌زمانی پیش‌بینی استفاده شده است در این پژوهش ما بازه‌های زمانی مختلف را برای دو بازه‌زمانی گفته شده بررسی می‌کنیم و با توجه به نکات گفته شده در فصل ۴ یک تعادل زمانی بین این دو بازه بدست می‌آوریم.

در بررسی این دو نکته از ۴ الگوریتم دسته‌بندی استفاده شده است. بصورت ویژه عملکرد یکی از الگوریتم‌های شبکه عصبی بازگشتی بنام LSTM مورد بررسی قرار گرفته است. نتایج نشان می‌دهد که می‌توان با دقت مطلوبی شروع سپسیس را در بیماران بخش ICU پیش‌بینی کرد.

## ۶-۱- پیشنهادات ادامه کار

برای بهبود مدلسازی در آینده و همچنین بالا بردن احتمال پذیرش این سیستم تصمیم‌یار توسط متخصص، موارد زیر پیشنهاد می‌شود:

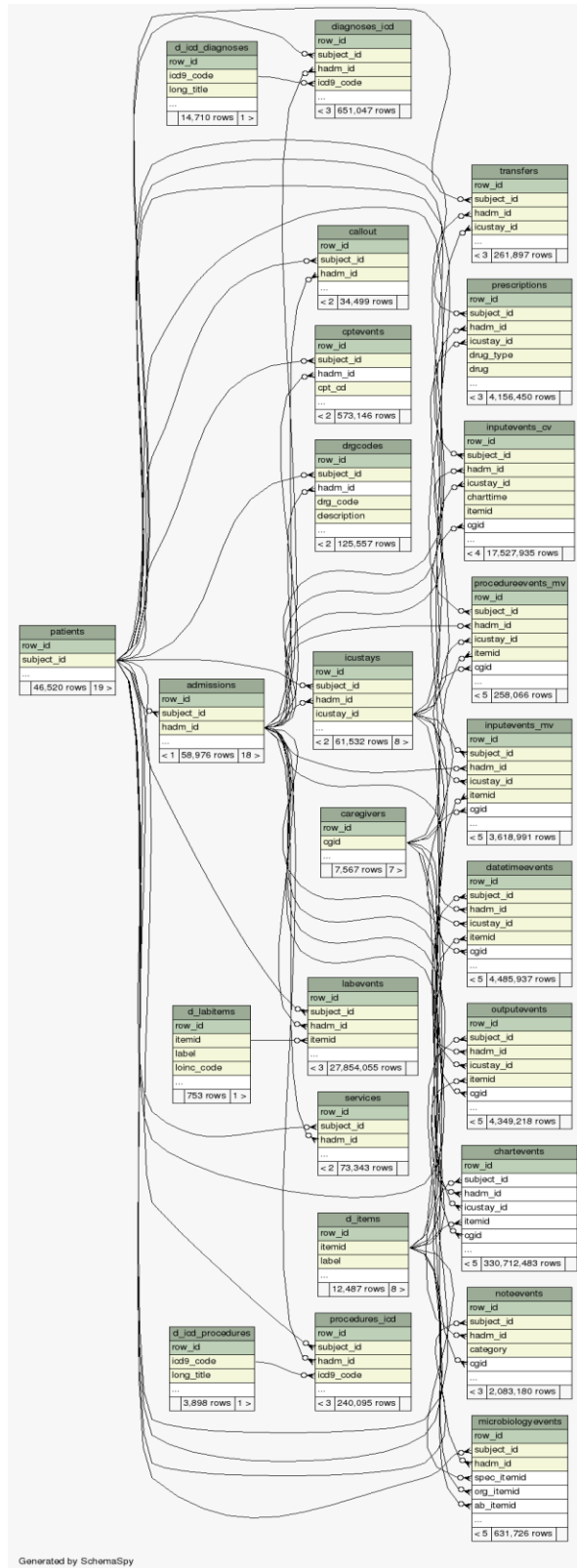
- ۱- پیش از انجام پرس وجوها و استخراج داده‌ها، در پایگاه داده MIMIC-III یا هر مجموعه داده واقعی دیگر کاوش بیشتری برای فهم مشکلات و اختلالات احتمالی انجام شود.
- ۲- با استفاده از کدهای ICD-9 ثبت شده برای بیماران، می‌توان از بیماری‌های همراه بیمار در بازه‌های زمانی بستری در بیمارستان بعنوان ویژگی‌های پیش‌بینی کننده استفاده کرد.
- ۳- برای اینکه احتمال اختلال در داده‌های محلی کم شود، داده‌ها در یک محیط کنترل شده برای پژوهش و در دوره‌های مختلف جمع‌آوری شوند.
- ۴- می‌توان از داروهای مصرفی بیماران در بخش مراقبت‌های ویژه بعنوان ویژگی‌های پیش‌بینی کننده استفاده کرد.
- ۵- در تعیین کلاس‌های بیماران بستری شده در بخش مراقبت‌های ویژه در این پژوهش از SOFA استفاده شده است که می‌توان از سایر تعریف‌ها و معیارهای تشخیص سپسیس استفاده کرد.
- ۶- در مسائل مربوط به پیش‌بینی بیماری انتخاب مجموعه ویژگی‌های پیش‌بینی کننده همیشه بعنوان یک مسئله باز است. لذا استفاده از روش‌های مختلف انتخاب ویژگی و انتخاب مجموعه‌های مختلف در پیش‌بینی سپسیس می‌تواند باعث پیشرفت این مسئله شود.
- ۷- در این پژوهش بازه‌های زمانی مختلف برای دو بازه زمانی پیش‌بینی کننده و گذشته نگر در نظر گرفته شده است ولی با این وجود تغییر این بازه‌های زمانی می‌تواند باعث پیشرفت در پیش‌بینی سپسیس شود.
- ۸- در مجموعه دادگان MIMIC-III مجموعه داده‌های از دست‌رفته همواره قابل توجه است، لذا استفاده از روش‌های پرکردن داده‌های از دست رفته می‌تواند باعث تغییر در دقت پیش‌بینی سپسیس و حتی سایر بیماری‌ها شود.
- ۹- برای اینکه سیستم در عمل قابل استفاده باشد و بعد از مدتی کنار گذاشته نشود، یکی از مواردی که در طراحی سیستم تصمیم‌یار اهمیت زیادی دارد، نحوه جمع‌آوری و دریافت ورودی‌هاست. به این خاطر، نیاز است که تا حد امکان سیستم به صورت خودکار مقادیر ورودی را از سیستم‌های موجود در ICU دریافت کند و فقط پیش از انجام عملیات پیش‌بینی، لیست مقادیر ورودی را به متخصص نمایش دهد تا در صورت نیاز و در صورت وجود تناقض در متغیرها اصلاحات لازم انجام شود.
- ۱۰- مسأله‌ی دیگر بروز شدن سیستم تصمیم‌یار بعد از جمع‌آوری داده‌های جدید در بازه‌های زمانی معین می‌باشد. این مسأله نیز می‌تواند نقش مهمی در پذیرش و استفاده کاربردی سیستم ایفا کند.

## ضمائم

### لیست جدول‌های موجود در پایگاه داده MIMIC-III:

Table name	Description
ADMISSIONS	Every unique hospitalization for each patient in the database (defines HADM_ID).
CALLOUT	Information regarding when a patient was cleared for ICU discharge and when the patient was actually discharged.
CAREGIVERS	Every caregiver who has recorded data in the database (defines CGID).
CHARTEVENTS	All charted observations for patients.
CPTEVENTS	Procedures recorded as Current Procedural Terminology (CPT) codes.
D_CPT	High level dictionary of Current Procedural Terminology (CPT) codes.
D_ICD_DIAGNOSES	Dictionary of International Statistical Classification of Diseases and Related Health Problems (ICD-9) codes relating to diagnoses.
D_ICD_PROCEDURES	Dictionary of International Statistical Classification of Diseases and Related Health Problems (ICD-9) codes relating to procedures.
D_ITEMS	Dictionary of local codes (ITEMIDs) appearing in the MIMIC database, except those that relate to laboratory tests.
D_LABITEMS	Dictionary of local codes (ITEMIDs) appearing in the MIMIC database that relate to laboratory tests.
DATETIMEEVENTS	All recorded observations which are dates, for example time of dialysis or insertion of lines.
DIAGNOSES_ICD	Hospital assigned diagnoses, coded using the International Statistical Classification of Diseases and Related Health Problems (ICD) system.
DRGCODES	Diagnosis Related Groups (DRG), which are used by the hospital for billing purposes.
ICUSTAYS	Every unique ICU stay in the database (defines ICUSTAY_ID).
INPUTEVENTS_CV	Intake for patients monitored using the Philips CareVue system while in the ICU, e.g., intravenous medications, enteral feeding, etc.
INPUTEVENTS_MV	Intake for patients monitored using the iMDSoft MetaVision system while in the ICU, e.g., intravenous medications, enteral feeding, etc.
OUTPUTEVENTS	Output information for patients while in the ICU.
LABEVENTS	Laboratory measurements for patients both within the hospital and in outpatient clinics.
MICROBIOLOGYEVENTS	Microbiology culture results and antibiotic sensitivities from the hospital database.
NOTEEVENTS	Deidentified notes, including nursing and physician notes, ECG reports, radiology reports, and discharge summaries.
PATIENTS	Every unique patient in the database (defines SUBJECT_ID).
PRESCRIPTIONS	Medications ordered for a given patient.
PROCEDUREEVENTS_MV	Patient procedures for the subset of patients who were monitored in the ICU using the iMDSoft MetaVision system.
PROCEDURES_ICD	Patient procedures, coded using the International Statistical Classification of Diseases and Related Health Problems (ICD) system.
SERVICES	The clinical service under which a patient is registered.
TRANSFERS	Patient movement from bed to bed within the hospital, including ICU admission and discharge.

## نمودار ارتباطات موجودیتهای پایگاه داده MIMIC-III :



## فهرست مراجع

- [1] M. Schinkel, K. Paranjape, R. S. Nannan Panday, N. Skyttberg, and P. W. B. Nanayakkara, "Clinical applications of artificial intelligence in sepsis: A narrative review," *Comput. Biol. Med.*, vol. 115, no. 103488, p. 103488, 2019.
- [2] M. M. Minn, "The comparative ability of commonly used disease severity scores to predict death or a requirement for ICU care in patients hospitalised with possible sepsis in Yangon, Myanmar," *International Journal of Infectious Diseases*, vol. 104, pp. 543–550, 2021.
- [3] G. Kong, K. Lin, and Y. Hu, "Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU," *BMC Med. Inform. Decis. Mak.*, vol. 20, no. 1, p. 251, 2020.
- [4] R. C. Bone, "Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis," *The ACCP/SCCM Consensus Conference Committee*, vol. 101, pp. 1644–1655, 1992.
- [5] M. M. Levy, M. P. Fink, J. C. Marshall, E. Abraham, D. Angus, D. Cook, J. Cohen, S. M. Opal, J. L. Vincent, and G. Ramsay, "2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference," *Intensive Care Med.*, vol. 29, no. 4, pp. 530–538, 2003.
- [6] M. Singer, "The third international consensus definitions for sepsis and septic shock (sepsis-3)," *JAMA*, vol. 315, no. 8, pp. 801–810, 2016.
- [7] J. L. Vincent, R. Moreno, J. Takala, S. Willatts, A. De Mendonça, H. Bruining, C. K. Reinhart, P. Suter, and L. G. Thijs, "The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine," *Intensive Care Med.*, vol. 22, no. 7, pp. 707–710, 1996.
- [8] F. Gül, M. K. Arslantaş, İ. Cinel, and A. Kumar, "Changing definitions of sepsis," *Turk. J. Anaesthesiol. Reanim.*, vol. 45, no. 3, pp. 129–138, 2017.
- [9] M. Moore and K. A. Loper, "An introduction to clinical decision support systems," *J. electron. resour. med. libr.*, vol. 8, no. 4, pp. 348–366, 2011.
- [10] A. Wulff, S. Montag, M. Marschollek, and T. Jack, "Clinical decision-support systems for detection of systemic inflammatory response syndrome, sepsis, and septic shock in critically ill patients: A systematic review," *Methods Inf. Med.*, vol. 58, no. S 02, pp. e43–e57, 2019.
- [11] L. T. Kohn, J. M. Corrigan, and M. S. Donaldson, *To Err Is Human: Building a Safer Health System*, vol. 627. National Academies Press, 2000.
- [12] E. S. Berner and T. J. La Lande, "Overview of clinical decision support systems," in *Health Informatics*, Cham: Springer International Publishing, 2016, pp. 1–17.
- [13] S. Hajioff, "Computerized decision support systems: an overview," *Health Informatics J.*, vol. 4, no. 1, pp. 23–28, 1998.
- [14] M. Scherpf, F. Gräßer, H. Malberg, and S. Zaunseder, "Predicting sepsis with a recurrent neural network using the MIMIC-III database," *Comput. Biol. Med.*, vol. 113, no. 103395, p. 103395, 2019.

- [15] Q. Li, L. Li, J. Zhong, and L. Frank Huang, "Real-time sepsis severity prediction on knowledge graph deep learning networks for the intensive care unit," *J. Vis. Commun. Image Represent.*, no. 102901, p. 102901, 2020.
- [16] J. S. Calvert, D. A. Price, U. K. Chettipally, C. W. Barton, M. D. Feldman, J. L. Hoffman, M. Jay, and R. Das, "A computational approach to early sepsis detection," *Comput. Biol. Med.*, vol. 74, pp. 69–73, 2016.
- [17] S. Nemati, A. Holder, F. Razmi, M. D. Stanley, G. D. Clifford, and T. G. Buchman, "An interpretable machine learning model for accurate prediction of sepsis in the ICU," *Crit. Care Med.*, vol. 46, no. 4, pp. 547–553, 2018.
- [18] Q. Y. Zhao, L. P. Liu, J. C. Luo, Y. W. Luo, H. Wang, Y. J. Zhang, R. Gui, J. C. Tu, G. W. and Z. Luo, "A machine-learning approach for dynamic prediction of sepsis-induced coagulopathy in critically ill patients with sepsis: an integrated analysis of the MIMIC-IV and eICU-CRD databases," *Research Square*, 2020.
- [19] C. Barton, U. Chettipally, Y. Zhou, Z. Jiang, A. Lynn-Palevsky, S. Le, J. Calvert, and R. Das, "Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs," *Comput. Biol. Med.*, vol. 109, pp. 79–84, 2019.
- [20] S. P. Oei, R. J. G. van Sloun, M. van der Ven, H. H. M. Korsten, and M. Mischi, "Towards early sepsis detection from measurements at the general ward through deep learning," *Intelligence-Based Medicine*, vol. 5, no. 100042, p. 100042, 2021.
- [21] C. Kok, V. Jahmunah, S. L. Oh, X. Zhou, R. Gururajan, X. Tao, K. H. Cheong, R. Gururajan, F. Molinari and U. R. Acharya, "Automated prediction of sepsis using temporal convolutional network," *Comput. Biol. Med.*, vol. 127, no. 103957, p. 103957, 2020.
- [22] J. Fagerström, M. Bång, D. Wilhelms, and M. S. Chew, "LiSep LSTM: A machine learning algorithm for early detection of septic shock," *Sci. Rep.*, vol. 9, no. 1, p. 15132, 2019.
- [23] T. Van Steenkiste, J. Ruyssinck, L. De Baets, J. Decruyenaere, F. De Turck, F. Ongenaë and T. Dhaene, "Accurate prediction of blood culture outcome in the intensive care unit using long short-term memory neural networks," *Artif. Intell. Med.*, vol. 97, pp. 38–43, 2019.
- [24] S. Meyer Lauritsen, M. Ellersgaard Kalør, E. L. Kongsgaard, K. Meyer Lauritsen, M. Johansson Jørgensen, J. Lange and B. Thiesson, "Early detection of sepsis utilizing deep learning on electronic health record event sequences," *Artif. Intell. Med.*, vol. 104, no. 101820, p. 101820, 2020.
- [25] Q. Mao, M. Jay, J. L. Hoffman, J. Calvert, C. Barton, D. Shimabukuro, L. Shieh, U. Chettipally, G. Fletcher, Y. Kerem and Y. Zhou, "Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU," *BMJ Open*, vol. 8, no. 1, p. e017833, 2018.
- [26] H. J. Kam and H. Y. Kim, "Learning representations for the early detection of sepsis with deep neural networks," *Comput. Biol. Med.*, vol. 89, pp. 248–255, 2017.
- [27] T. Desautels, J. Calvert, J. Hoffman, M. Jay, Y. Kerem, L. Shieh, D. Shimabukuro, U. Chettipally, M. D. Feldman, C. Barton and D. J. Wales, "Prediction of sepsis in the Intensive Care unit with minimal electronic health record data: A machine learning approach," *JMIR Med. Inform.*, vol. 4, no. 3, p. e28, 2016.

- [28] S. P. Shashikumar, M. D. Stanley, I. Sadiq, Q. Li, A. Holder, G. D. Clifford and S. Nemati, "Early sepsis detection in critical care patients using multiscale blood pressure and heart rate dynamics," *J. Electrocardiol.*, vol. 50, no. 6, pp. 739–743, 2017.
- [29] S. P. Shashikumar, M. D. Stanley, I. Sadiq, Q. Li, A. Holder, G. D. Clifford and S. Nemati, "Presepsin (soluble CD14 subtype) and procalcitonin levels for mortality prediction in sepsis: data from the Albumin Italian Outcome Sepsis trial," *Crit. Care*, vol. 18, no. 1, p. R6, 2014.
- [30] C. Adrie, A. Francais, A. Alvarez-Gonzalez, R. Mounier, E. Azoulay, J. R. Zahar, C. Clec'h, D. Goldgran-Toledano, L. Hammer, A. Descorps-Declere and S. Jamali, "Model for predicting short-term mortality of severe sepsis," *Crit. Care*, vol. 13, no. 3, p. R72, 2009.
- [31] V. J. R. Ripoll, A. Vellido, E. Romero, and J. C. Ruiz-Rodríguez, "Sepsis mortality prediction with the quotient basis kernel," *Artificial intelligence in medicine*, vol. 61, no. 1, pp. 45–52, 2014.
- [32] W. F. Fang, I. S. Douglas, Y. M. Chen, C. Y. Lin, H. C. Kao, Y. T. Fang, C. H. Huang, Y. T. Chang, K. T. Huang, Y. H. Wang and C. C. Wang, "Development and validation of immune dysfunction score to predict 28-day mortality of sepsis patients," *PLoS One*, vol. 12, no. 10, p. e0187088, 2017.
- [33] Y. Xie, J. Yabes, R. Parker, and G. Clermont, "1485: Using clinical features and biomarkers to predict 60-day mortality of sepsis patients: Using clinical features and biomarkers to predict 60-day mortality of sepsis patients," *Crit. Care Med.*, vol. 46, no. 1, pp. 726–726, 2018.
- [34] S. V. Poucke, Z. Zhang, M. Schmitz, M. Vukicevic, M. V. Laenen, L. A. Celi and C. D. Deyne, "Scalable predictive analysis in critically ill patients using a visual open data analysis platform," *PloS one*, vol. 11, no. 1, 2016.
- [35] Z. Zhang and Y. Hong, "Development of a novel score for the prediction of hospital mortality in patients with severe sepsis: the use of electronic healthcare records with LASSO regression," *Oncotarget*, vol. 8, no. 30, pp. 49637–49645, 2017.
- [36] R. A. Taylor, J. R. Pare, A. K. Venkatesh, H. Mowafi, E. R. Melnick, W. Fleischman and M. K. Hall, "Prediction of in-hospital mortality in emergency department patients with sepsis: A local big data-driven, machine learning approach," *Acad. Emerg. Med.*, vol. 23, no. 3, pp. 269–278, 2016.
- [37] A. Pregernig, M. Müller, U. Held, and B. Beck-Schimmer, "Prediction of mortality in adult patients with sepsis using six biomarkers: a systematic review and meta-analysis," *Ann. Intensive Care*, vol. 9, no. 1, p. 125, 2019.
- [38] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-Wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi and R. G. Mark, "MIMIC-III, a freely accessible critical care database," *Sci. Data*, vol. 3, no. 1, p. 160035, 2016.
- [39] O. T. Abdala and M. Saeed, "Estimation of missing values in clinical laboratory measurements of ICU patients using a weighted K-nearest neighbors' algorithm," in *Computers in Cardiology*, 2004, 2005.
- [40] D. E. Ong, M. Frize, J. Gilchrist, E. Bariciak, and C. M. Ennett, "Usefulness analysis of a Clinical Data Repository design," in *2013 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2013.

- [41] J. Elman, "Finding structure in time\* 1," *Cognitive Science*, vol. 14, no. 1, pp. 179–211, 1990.
- [42] M. I. Jordan, "Serial order: A parallel distributed processing approach," in *Neural-Network Models of Cognition - Biobehavioral Foundations*, Elsevier, 1997, pp. 471–495.
- [43] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv [cs. LG]*, 2015.
- [44] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [45] F. A. Gers and J. Schmidhuber, "Recurrent nets that time and count," in *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, 2000.
- [46] J. Han and M. Kamber, chapter 1: introduction: *Data Mining: Concepts and Techniques*. Morgan Kaufman Publisher, 2006.
- [47] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 2020.
- [48] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer, 2009.
- [49] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*, vol. 112. New York: springer, 2013.
- [50] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, 1988.





## **Abstract**

Sepsis is one of the leading causes of death in the intensive care unit. Therefore, early diagnosis and taking the necessary measures for the patient in the intensive care unit is of great importance. The aim of this dissertation is to design a decision-making system to predict the onset of sepsis in patients admitted to the intensive care unit; For this purpose, the recursive neural network algorithm is used to predict the onset of sepsis. Using this algorithm and the number of features available to patients, the physician can predict whether sepsis will begin in this patient.

In this paper, the detection of the onset of sepsis in the MIMIC-III database is examined with a focus on two different topics. To address these two issues, the values of the characteristics of the training data are first mapped based on the definition of sepsis labeled, sepsis / non-sepsis, and the onset of sepsis. Then 4 known classification algorithms; Focusing on LSTM recursive neural network algorithms, it was used for modeling; The results show the capability of these methods in classifying patients. After modeling with LSTM recursive neural network algorithms, in the first issue, modeling at different time intervals was first considered for both retrospective and predictive time periods. After determining the best time interval in the second topic, among the features considered for modeling; Using the feature selection methods, the best feature combination is selected. With the selected features, re-modeling is performed and finally the best feature combination is presented for the best time intervals.

The results of this study show that the best time to look forward to the onset of sepsis is 12 hours and the best time to predict sepsis is 6 hours before the onset of sepsis. In this study, LSTM recursive neural network algorithm with AUROC value of 0.9252 is the best algorithm for predicting sepsis in these time periods. The results of this study also show that gender characteristics and respiration rate have little effect on predicting sepsis and the best value for filling in the missing data in each time series is the mean value.

**Keywords:** Predictive disease, Multivariate time series, Sepsis, Recurrent neural network, MIMIC-III.



**Yazd University**  
**Campus of Engineering**

**Thesis Submitted in Partial Fulfillment of the Requirement  
for the Master Degree in Computer Engineering**

**Title**

**Improving the performance of predicting sepsis using the  
recurrent neural network algorithm**

**Supervisor**

**Dr. M. R. Pajoohan**

**Advisor**

**Dr. M. Rezaeian**

**By**

**Parisa Etemadi-Nejad**

**January 2022**