Journal of Health and Biomedical Informatics 2024; 11(2): 115-130
doi:10.34172/jhbmi.2024.18
Medical Informatics Research Center, Kerman University of Medical Sciences

Original article

Intelligent Diagnosis of Larynx Cancer Using Machine Learning Methods

Fereshteh Arad[1,2], Seyyed Mohammad Mousavi [3,4], Soodeh Hosseini [1,5]*, Maryam Amizade[1,6], Ayyub Sheikhi[1,7]

1. Mahani Math Center, Afzalipour Research Institute, Shahid Bahonar University of Kerman, Kerman, Iran
2. Ph.D. Student of Statistics, Department of Statistics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran
3. Medical Informatics Research Center, Institute for Future Studies in Health, Kerman University of Medical Sciences, Kerman, Iran
4. Ph.D. Student of Medical Informatics, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran
5. Ph.D of Computer Engineering, Associate Professor, Department of Computer Science, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran
6. Ph.D of Otolaryngologists, Associate Professor, Clinical Research Development Unit, Shafa Hospital, Kerman University of Medical Sciences, Kerman, Iran
7. Ph.D of Statistics, Associate Professor, Department of Statistics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran

ARTICLE INFO:

Article History:

Received: 12 Mar 2024

Accepted: 3 Aug 2024

Published: 21 Sep 2024

Corresponding Author:

Soodeh Hosseini

Email:
so_hosseini@uk.ac.ir

Citation: Arad F, Mousavi S M, Hosseini S, Amizade M, Sheikhi A. Intelligent Diagnosis of Larynx Cancer Using Machine Learning Methods. Journal of Health and Biomedical Informatics 2024; 11(2): 115-30. [In Persian]

Abstract

Introduction: Larynx cancer can be benign or malignant based on various factors. This research aimed to provide a machine learning-based model to improve the diagnosis of individuals with larynx cancer.

Method: In the first step, the voices of the people who visited the medical centers (including the sounds (A), (E), and (O)) were recorded and considered as a data set. In the second step, the data were classified into three classes (benign cancer, malignant cancer, and healthy) by a specialist. In the third step, the data cleaning was done. In the fourth step, the features related to sound were extracted from the data. In the fifth step, five machine learning models including SVM, Decision Tree, Naïve Bayes, MLP, and Random Forest were implemented on the data set. Finally, the performance of the models was evaluated using evaluation criteria such as accuracy, F-score, and other evaluation criteria.

Results: The results of the implementation showed that the SVM model had a higher accuracy than other models for the sound (A) and sound (O) with an accuracy of 0.818, and the sound (E) with an accuracy of 0.818 in the model MLP had the highest accuracy.

Conclusion: The present study evaluated machine learning models for the diagnosis of laryngeal cancer based on audio data. The results showed that the use of the SVM model for the diagnosis of laryngeal cancer can help diagnose this disease more accurately and provide reliable results.

Keywords: Laryngeal cancer, Machine learning, Deep learning, Artificial intelligence

© 2024 The Author(s); Published by Kerman University of Medical Sciences. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cite

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

مجله انفورماتیک سلامت و زیست پزشکی دوره یازدهم، شماره ۲، سال ۱۴۰۳، ص ۱۳۰-۱۱۵
doi: 10.34172/jhbmi.2024.18
مرکز تحقیقات انفورماتیک پزشکی، دانشگاه علوم پزشکی کرمان

مقاله پژوهشی

تشخیص هوشمند سرطان حنجره با استفاده از روش‌های یادگیری ماشین

فرشته آراد [۱،۲]، سید محمد موسوی [۳،۴]، سوده حسینی [۱،۵]*، مریم عامی‌زاده [۱،۶]، ایوب شیخی [۱،۷]

۱. پژوهشکده ریاضی ماهانی، پژوهشگاه افضلی پور، دانشگاه شهید باهنر کرمان، کرمان، ایران
۲. دانشجو دکتری آمار، بخش آمار، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان، کرمان، ایران
۳. مرکز تحقیقات انفورماتیک پزشکی، پژوهشکده آینده پژوهی، دانشگاه علوم پزشکی کرمان، کرمان، ایران
۴. دانشجوی دکتری انفورماتیک پزشکی، دانشکده مدیریت و اطلاع رسانی پزشکی، دانشگاه علوم پزشکی کرمان، کرمان، ایران
۵. دکتری مهندسی کامپیوتر، دانشیار، بخش علوم کامپیوتر، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان، کرمان، ایران
۶. دکتری تخصصی گوش و حلق و بینی، دانشیار، واحد توسعه تحقیقات بالینی، بیمارستان شفا، دانشگاه علوم پزشکی کرمان، کرمان، ایران
۷. دکتری آمار، دانشیار، بخش آمار، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان، کرمان، ایران

اطلاعات مقاله

سابقه مقاله

دریافت: ۱۴۰۲/۱۲/۲۲
پذیرش: ۱۴۰۳/۵/۱۳
انتشار برخط: ۱۴۰۳/۶/۳۱

نویسنده مسئول:
سوده حسینی
ایمیل:
so_hosseini@uk.ac.ir

ارجاع: آراد فرشته، موسوی سید محمد، حسینی سوده، عامی‌زاده مریم، شیخی ایوب. تشخیص هوشمند سرطان حنجره با استفاده از روش‌های یادگیری ماشین. مجله انفورماتیک سلامت و زیست پزشکی ۱۴۰۳؛ ۱۱(۲): ۳۰-۱۱۵.

مقدمه: سرطان حنجره بر اساس عوامل مختلف می‌توانند خوش‌خیم یا بدخیم باشند. هدف این پژوهش بهبود مدلی مبتنی بر یادگیری ماشین جهت ارتقاء تشخیص افراد درگیر با سرطان حنجره می‌باشد.

روش کار: گام نخست، صداهای افرادی که به مراکز درمانی مراجعه کرده‌اند (شامل آواهای (آ)، (ای)، (او)) ضبط شده و به عنوان مجموعه داده در نظر گرفته شده‌اند؛ در گام دوم داده‌ها توسط پزشک متخصص به سه کلاس سرطان خوش‌خیم، سرطان بدخیم و سالم تقسیم شده‌اند؛ در گام سوم مرحله پاکسازی داده‌ها انجام شده است؛ در گام چهارم ویژگی‌های مرتبط با صدا از داده‌ها استخراج و در گام پنجم، ۵ مدل یادگیری ماشین SVM، Decision Tree، Naïve Bayes، MLP و Random Forest بر روی مجموعه داده‌ها پیاده‌سازی شده است و در آخر با استفاده از معیارهای ارزیابی مانند دقت، F-score و دیگر معیارهای ارزیابی، عملکرد مدل‌ها ارزیابی شده است.

یافته‌ها: نتایج پیاده‌سازی نشان داد که مدل SVM برای آوای (آ) و آوای (او) با دقت ۰/۸۱۸، دقت بالاتری نسبت به سایر مدل‌ها از خود نشان داده است، آوای (ای) با دقت ۰/۸۱۸ در مدل MLP بالاترین دقت را دارا می‌باشد.

نتیجه‌گیری: پژوهش حاضر، به ارزیابی مدل‌های یادگیری ماشین برای تشخیص سرطان حنجره بر اساس داده‌های صوتی پرداخته است، نتایج نشان داد که استفاده از مدل SVM برای تشخیص سرطان حنجره می‌تواند با دقت بیشتری در تشخیص این بیماری کمک کند و ارائه نتایج قابل اعتمادی را فراهم آورد.

کلیدواژه‌ها: سرطان حنجره، یادگیری ماشین، یادگیری عمیق، هوش مصنوعی

[Downloaded from jhbmi.ir on 2026-09-07]

[DOI: 10.34172/jhbmi.2024.18]

مقدمه

سرطان حنجره یکی از انواع سرطان‌های حلق و حنجره است که ممکن است در بخش‌های مختلف حنجره ایجاد شود. این نوع سرطان معمولاً به تدریج و با آغاز از سلول‌های غددی یا پوششی حنجره شکل می‌گیرد. عواملی مانند مصرف تنباکو و الکل، عفونت با ویروس‌های هموفیلوس آنفلوآنزا و پاپیلوما ویروس انسانی (ایدز)، تغذیه نامناسب و تماس با مواد شیمیایی می‌توانند عوامل خطرساز برای ابتلاء به سرطان حنجره باشند. به طورکلی سرطان حنجره به دو صورت خوش‌خیم یا بدخیم ظاهر می‌شود و نیازمند بررسی و درمان توسط پزشک متخصص حنجره می‌باشد. شناسایی زودرس و درمان مناسب سرطان حنجره از اهمیت بالایی برخوردار است تا از پیشرفت بیشتر بیماری جلوگیری شود و بهبود پیش آگهی برای بیمار فراهم شود. امروزه با استفاده از تکنولوژی‌های پردازش سیگنال صوتی و هوش مصنوعی، می‌توان با تجزیه و تحلیل الگوهای صوتی افراد، بیماری‌های مختلف را تشخیص داد. این روش‌های پیشرفته برای تشخیص بیماری‌هایی استفاده می‌شوند که ممکن است به پزشکان در تشخیص اولیه بیماری‌ها، ارائه نتایج دقیق و سریع و کاهش پیشرفت بیماری کمک کنند. یکی از مهم‌ترین مباحث حوزه کامپیوتر و برنامه‌نویسی هوش مصنوعی (Artificial Intelligent) است. هوش مصنوعی در اصل یک ماشین یا سیستم است که عملکردی نزدیک به انسان را دارد و می‌تواند برنامه‌ریزی کند، یاد بگیرد، استدلال کند، قدرت حل مسئله داشته باشد و در نهایت با توجه به شرایط بهترین عملکرد را داشته باشد [۱].

یادگیری ماشین (Machine Learning) شاخه‌ای از هوش مصنوعی است که به سیستم‌ها توان یادگیری خودکار و پیشرفت می‌دهد. در واقع یادگیری ماشین به زبان ساده جزء مهم حوزه رو به رشد علم داده است. از طریق استفاده از روش‌های آماری، الگوریتم‌ها برای طبقه‌بندی یا پیش‌بینی آموزش داده می‌شوند و بینش‌های کلیدی را در پروژه‌های داده‌کاوی آشکار می‌کنند. این بینش‌ها متعاقباً تصمیم‌گیری را در برنامه‌ها و کسب وکارها هدایت می‌کنند و به‌طور ایده آل بر معیارهای رشد کلیدی تأثیر می‌گذارند. در یادگیری ماشین، پروسه یادگیری از طریق داده‌ها انجام می‌شود. پس از فرآیند اولیه یادگیری، ماشین قادر است به‌تنهایی و بدون دخالت انسان الگوهایی در میان داده‌های جدید کشف کند. هدف اولیه از به کارگیری یادگیری ماشین این است که کامپیوترها توانایی یادگیری خودکار بدون دخالت یا کمک انسان را پیدا کنند و بر اساس آموخته‌های خود اقداماتی انجام دهند [۲]. بر اساس رویکردهای یادگیری، الگوریتم‌های یادگیری ماشین را می‌توان به چهار نوع تقسیم کرد [۳].

- الگوریتم‌های یادگیری ماشین با رویکرد «یادگیری نظارت شده (Supervised Learning)»
- الگوریتم‌های یادگیری ماشین با رویکرد «یادگیری نظارت نشده (Unsupervised Learning)»
- الگوریتم‌های یادگیری ماشین با رویکرد «یادگیری نیمه نظارت شده (Semi-Supervised Learning)»
- الگوریتم‌های یادگیری ماشین با رویکرد «یادگیری تقویتی (Reinforcement Learning)»

همچنین از الگوریتم‌های یادگیری ماشین می‌توان به الگوریتم درخت تصمیم [۴]، الگوریتم جنگل تصادفی [۵]، الگوریتم ماشین بردار پشتیبان [۶]، الگوریتم بیز ساده [۷]، و.... اشاره نمود. یکی دیگر از زیر شاخه‌های حوزه یادگیری ماشین، یادگیری عمیق تلقی می‌شود. هدف یادگیری عمیق (Deep Learning) طراحی سیستم‌های کامپیوتری هوشمند است که بتوانند مشابه انسان درباره موضوعی خاص، راه حل ارائه کنند و مفاهیم جدیدی را یاد بگیرند. این حوزه از فناوری، شاخه‌ای مهم در علم داده است، زیرا اصلی‌ترین مباحث این شاخه آمار و مدل‌سازی برای پیش‌بینی مسائل مختلف را شامل می‌شود مهندسان علم داده با استفاده از روش‌های یادگیری عمیق می‌توانند جمع‌آوری، تجزیه و تحلیل و تفسیر حجم عظیمی از داده‌ها را سریع‌تر و آسان‌تر انجام دهند. مدل‌های یادگیری عمیق از ساختارهای لایه به لایه‌ای تشکیل شده‌اند که به آن‌ها شبکه عصبی (Neural Network) گفته می‌شود. از انواع شبکه عصبی می‌توان به شبکه‌های عصبی پرسپترون چند لایه (Perceptron Multi-Layer)، شبکه‌های عصبی شعاعی (Radial Basis Functions)، شبکه‌های عصبی کوهونن (Self-Organizing Map) و شبکه عصبی هاپفیلد (Hopfield) و..... اشاره کرد [۸،۹]. Al-Dhief و همکاران [۱۰] ادغام فناوری ابری با اینترنت اشیاء (Internet of Things) را پیشنهاد کردند، به منظور دستیابی به عملکرد بهتر برای یک چهارچوب یکپارچه، پیوسته و فراگیر، IoT کاربردهای زیادی در بخش مراقبت‌های بهداشتی دارد. اکثر محققان بر روی آسیب‌شناسی صدا تمرکز

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

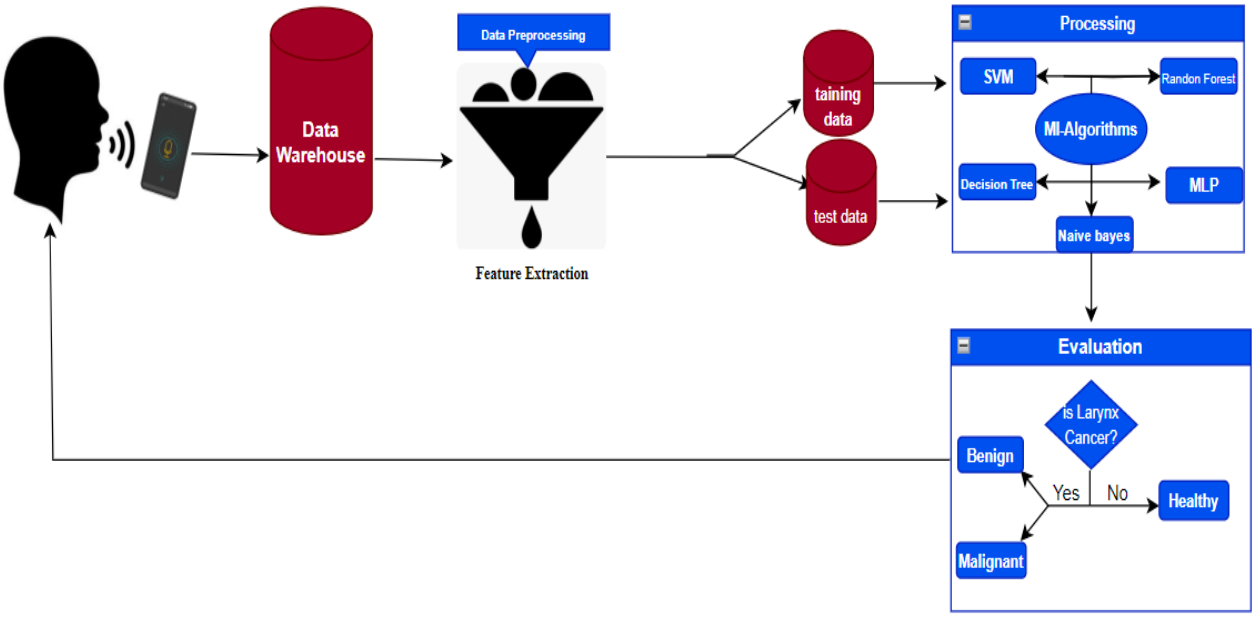
کرده‌اند و یافته‌های آن‌ها فقط تشخیص صدای طبیعی (سالم) یا آسیب‌شناسی است که در آن یافته‌ها مطالعاتی برای تشخیص بیماری خاصی مانند سرطان حنجره وجود ندارد. همچنین یک بررسی گسترده از تکنیک‌ها و مطالعات پیشرفته در چهارچوب IoT و الگوریتم‌های یادگیری ماشین مورد استفاده در مراقبت‌های بهداشتی به طور کلی و در سیستم‌های نظارت آسیب شناسی صوتی به طور خاص ارائه می‌دهند[۱۰]. صیادی و همکاران [۱۱] اشاره به این موضوع دارند که نیاز به تشخیص اختلالات گفتاری وجود دارد. یادگیری ماشین به عنوان حوزه‌ای از هوش مصنوعی، محققان، پزشکان و بیماران را قادر می‌سازد تا این مسائل را حل کنند. هدف از این مطالعه طبقه‌بندی و مقایسه روش‌های یادگیری ماشین در تشخیص بیماری‌های گفتار محور بود. در این بررسی سیستماتیک، یک جستجوی جامع برای نشریات در پایگاه‌های اطلاعاتی اسکوپوس، IEEE، Web of Science، PubMed و Cochrane از سال ۲۰۰۲ تا ۲۰۲۲ انجام شد. از ۵۳۳ نتیجه جستجو، ۴۸ مقاله براساس معیارهای شایستگی انتخاب شدند. یافته‌های آن‌ها نشان می‌دهد که تشخیص بیماری‌های گفتار محور با استفاده از سیگنال‌های گفتاری به فرهنگ، زبان و محتوای گفتار، جنسیت، سن، لهجه و عوامل دیگر بستگی دارد. استفاده از مدل‌های یادگیری ماشینی بر روی صداهای گفتار یک مسیر امیدبخش برای بهبود تشخیص و درمان بیماری‌های گفتاری در راستای داروهای پیشگیرانه و شخصی است. Mesallam و همکاران [۱۲] به این موضوع اشاره دارند که پایگاه داده اختلالات صوتی یک عنصر ضروری در انجام تحقیقات در زمینه تشخیص و طبقه‌بندی اختلالات صوتی خودکار است. قومیت بر ویژگی صوتی افراد تأثیر می‌گذارد، بنابراین لازم است با جمع‌آوری نمونه‌های صدای گروه قومیتی، بانک اطلاعاتی ایجاد شود. این امر شانس دستیابی به یک راه حل جهانی برای تشخیص دقیق و قابل اعتماد اختلالات صوتی را با درک ویژگی‌های یک گروه محلی افزایش می‌دهد. با انگیزه چنین ایده‌ای، یک پایگاه داده آسیب‌شناسی صدای عربی (AVPD) در این مطالعه با ضبط سه واکه، گفتار در حال اجرا و کلمات مجزا طراحی و توسعه یافته است. برای هر نمونه ثبت شده، شدت ادراکی نیز ارائه شده است که جنبه منحصر به فرد AVPD است. Sharma و Gupta [۱۳] در پژوهش خود به بررسی کلیه تکنیک‌های پردازش ویژگی‌های صدا و طبقه‌بندی آن‌ها در گروه‌های خاص می‌پردازند به طوری که بیان می‌کنند ویژگی‌های صدا با استفاده از برخی از تکنیک‌های پردازش سیگنال دیجیتال استخراج می‌شوند. این ویژگی‌ها شامل اطلاعاتی در مورد سلامت مجرای صدا و اعضایی که در تولید گفتار همکاری می‌کنند، هستند. این ویژگی‌ها نمایانگر صدای خاصی هستند و می‌توانند برای تمیز دادن صدای افراد سالم و ناسالم استفاده شوند. تجزیه و تحلیل حوزه زمان، طیف، سپستر و موج شکافی گلاوتال برای استخراج ویژگی‌های صدا استفاده می‌شود. این ویژگی‌ها سپس با استفاده از تکنیک‌های طبقه‌بندی مختلف مانند کوانتیزاسیون بردار، پیچیدگی زمان پویا، ماشین بردار پشتیبان، مدل مخلوط گوسی و شبکه عصبی مصنوعی به گروه‌ها دسته‌بندی می‌شوند. Kim و همکاران [۱۴] به بررسی این که آیا تجزیه و تحلیل خودکار سیگنال صدا می‌تواند برای تمایز بین بیماران مبتلا به سرطان حنجره از افراد سالم استفاده شود. ویژگی‌ها را با استفاده از بسته نرم‌افزاری برای تجزیه و تحلیل گفتار در زبان‌شناسی (PRAAT) استخراج کرده و ضرایب کپسترال فرکانس مل (Mel-frequency cepstral coefficients) را از نمونه‌های صدای یک صدای واول /a:/ محاسبه کرده‌اند و روش پیشنهادی خود را با شش الگوریتم آزمون کردند: ماشین بردار پشتیبان، افزایش گرادیانی بیشینه، ماشین گرادیانی تقویت شده نور، شبکه عصبی مصنوعی، شبکه عصبی کانولوشن یک بعدی و شبکه عصبی کانولوشن دو بعدی و در ادامه عملکردهای آن‌ها به ازای دقت، حساسیت و اختصاصیت ارزیابی و نتیجه با عملکرد انسان مقایسه شده است. مجموعاً چهار داوطلب، دو نفر از آن‌ها لرنگولوژیست آموزش دیده، به همان فایل‌ها امتیاز دادند. شبکه عصبی کانولوشن یک بعدی دقت بالای ۸۵٪ و سطوح حساسیت و خصوصیت ۷۸٪ و ۹۳٪ را نشان داد دو لرنگولوژیست دقت ۶۹/۹٪ و سطوح حساسیت ۴۴٪ را داشتند. تجزیه و تحلیل خودکار سیگنال‌های صدا می‌توانست افراد مبتلا به سرطان حنجره را از افراد سالم با ویژگی‌های تشخیصی بالاتر از آن‌هایی که توسط چهار داوطلب تشخیص داده شده‌اند، تمایز دهد. هدف این مطالعه بررسی امکان استفاده از تحلیل خودکار سیگنال‌های صوتی برای تشخیص سرطان حنجره در بیماران است. به طور خاص، محققان این تحقیق به بررسی اینکه آیا می‌توان با استفاده از ویژگی‌های استخراج‌شده از صدا و الگوریتم‌های یادگیری ماشین، بیماران مبتلا به سرطان حنجره را از افراد سالم تمایز داد، پرداخته‌اند.

روش کار

در ابتدا مجموعه داده‌ای از صوت افراد جمع‌آوری شد، بدین صورت که صدای فرد مراجعه کننده به مرکز درمانی (آواهای (آ)، (ای)، (او)).

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

ضبط شد، سپس داده‌ها توسط پزشک متخصص در سه کلاس ۱- سرطان خوش‌خیم ۲- سرطان بدخیم ۳- سالم برچسب گذاری شده‌اند در ادامه در مرحله پاکسازی داده‌ها (اصواتی که دارای کیفیت پایین ضبط صدا یا مدت زمان کوتاه بودند از مجموعه داده حذف شدند) استخراج ویژگی انجام گرفت و مدل‌های یادگیری ماشین بر اساس داده‌های جمع‌آوری مورد آموزش و آزمون قرار گرفتند و در نهایت دقت مدل‌ها ارزیابی شد. با توجه به شکل ۱ ابتدا داده جدید که صدای فرد ضبط شده (آواهای مدنظر) به مخزن داده‌های صوتی اضافه می‌شود، با طی کردن مراحل پیش‌پردازش و پردازش، طبقه‌بندی داده صوتی انجام و فرد در یکی از سه گروه قرار می‌گیرد، بنابراین می‌تواند درمان خود را با توجه به گروهی که در آن قرار می‌گیرد، شروع کند.



شکل ۱: روند کلی روش کار

مجموعه داده صوتی جمع‌آوری شده

برای ارزیابی روش مدنظر این پژوهش، ابتدا روی بانک داده جمع‌آوری شده که شامل اصوات ضبط شده (آواهای (آ)، (ای) و (او)) ۱۰۹ نفر از افراد مراجعه کننده به مراکز درمانی شهر کرمان که دارای سرطان حنجره (خوش‌خیم و بدخیم) و یا سالم می‌باشد، مرحله پاکسازی داده انجام گرفت، به طوری که جامعه آماری این تحقیق بعد از پاکسازی داده‌ها شامل اصوات ضبط شده (آواهای (آ)، (ای) و (او)) ۱۰۵ نفر از افراد مراجعه کننده به مراکز درمانی بود. برای آزمودن مدل‌های خود دو حالت در نظر گرفته شد، حالت اول مبنای آزمودن تشخیص در نظر گرفته شد در این صورت ۸۰ درصد داده‌ها که شامل ۲۵۲ داده است در گروه آموزش قرار داده شد و ۲۰ درصد که ۶۳ داده را شامل می‌شود در گروه آزمون بود، حالت دوم مبنای آزمودن صوت ضبط شده (آواهای (آ)، (ای) و (او)) قرار داده شد در این صورت ۹۰ درصد داده‌ها که شامل ۹۴ داده است گروه آموزش در نظر گرفته شد و ۱۰ درصد که شامل ۱۱ داده است در گروه آزمون قرار گرفت.

طبقه‌بندی صوت افراد

ساختار این پژوهش بدین صورت است که ابتدا آواهای ((آ)، (ای) و (او)) فردی که به مراکز درمانی مراجعه می‌کند به مجموعه داده ارسال شد. سپس آواهای ارسال شده در مجموعه داده مورد پیش‌پردازش قرار گرفتند و سپس در مرحله پردازش توسط مدل‌های یادگیری ماشین مورد آزمایش قرار گرفتند. در مرحله پیش پردازش استخراج ویژگی انجام شد، هر سیگنال صوتی ویژگی‌های بسیاری دارد، اما برای حل مسئله‌ای خاص باید ویژگی‌های مرتبط با آن مسئله استخراج شوند. برای اطمینان از یکنواختی داده‌ها، موج صوتی با نرخ نمونه‌برداری ۲۲

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

مجله انفورماتیک سلامت و زیست پزشکی دوره یازدهم، شماره ۲، سال ۱۴۰۳، ص ۱۳۰-۱۱۵
doi: 10.34172/jhbmi.2024.18
مرکز تحقیقات انفورماتیک پزشکی، دانشگاه علوم پزشکی کرمان

کیلوهرتز در فرآیند استخراج ویژگی برای کاربردهای صوتی استفاده شد. هفت ویژگی طیفی Chromagram، Root Mean Squared، Spectral Centroid، Spectral bandwidth، Spectral Rolloff، Zero Crossing Rate، Mel-Frequency Cepstral Coefficients، از صدای نمونه‌برداری شده با استفاده از کتابخانه librosa [۱۶] پایتون استخراج شده‌اند. در ادامه مدل با هایپر پارامترهای در نظر گرفته شده در این مرحله پردازش بیان شده‌اند.

مدل Support Vector Machine با C=۳۰ و kernel=rbf، پارامتر C در ماشین بردار پشتیبان به کنترل تعادل بین خطای آموزش و حاشیه کمک می‌کند، زیرا می‌تواند جریمه اشتباهات داده‌های طی فرآیند آموزش را تعیین کند. به طور خاص، مقدار کوچک‌تری از C اجازه حاشیه بزرگ‌تری را می‌دهد که ممکن است منجر به بیشتر شدن اشتباهات در داده‌های آموزش شود. از سوی دیگر، مقدار بزرگ‌تری از C بیشتر بر کاهش خطای آموزش تأکید می‌کند که ممکن است منجر به حاشیه باریک‌تری شود.

مدل Decision Tree با Criterion = entropy, max_depth=۹ که Criterion تابع اندازه‌گیری کیفیت تقسیم درخت می‌باشد. معیارهای قابل پشتیبانی در این بخش عبارت خواهند بود از "gini" برای انتخاب تصادفی در مجموعه دیتا و "entropy" برای نرخ متوسط اطلاعات تولید شده منابع اطلاعاتی دیتا.

مدل Random Forest با ۵۰، ۱۰۰، ۲۰۰، ۵۰۰ = Estimators.

مدل Multi-Layer Perceptron با Hidden_layer_sizes= ۳۰۰ و نرخ یادگیری جهت آموزش شبکه برابر با ۰/۳، تعداد تکرارهای ۱۰۰۰ و learning-rate برابر با adaptive.

یافته‌ها

در این مطالعه، نتایج مربوط به ارزیابی عملکرد و دقت ۵ مدل مختلف در دو حالت مختلف از مجموعه داده‌ها مورد بررسی قرار گرفت. یکی از مراحل حیاتی پس از ایجاد و پیاده‌سازی یک مدل یا الگوریتم، ارزیابی دقت، صحت و کارآیی آن می‌باشد. در این پژوهش، معیارهای ارزیابی مدل‌ها تشریح شده و نتایج ارزیابی‌ها جهت مقایسه و تحلیل ارائه شد. این اطلاعات می‌تواند به تحلیل عمیق‌تر عملکرد مدل‌ها و انتخاب بهترین مدل برای مسئله مورد بررسی کمک کند.

معیارهای ارزیابی روی مدل‌های یادگیری ماشین

معیارهای متنوعی برای ارزیابی کارایی الگوریتم‌ها وجود دارد. در انتخاب معیارهای ارزیابی کارایی می‌بایست دلایل قانع کننده وجود داشته باشد چرا که چگونگی اندازه‌گیری و مقایسه کارایی الگوریتم‌ها، کاملاً به معیارهایی که انتخاب می‌کنید وابسته است. همچنین چگونگی وزن دادن به اهمیت ویژگی‌های مختلف در نتایج، کاملاً تحت تأثیر معیارهایی است که انتخاب می‌کنید.

دقت(Accuracy)

این متداول‌ترین معیار کارایی برای الگوریتم‌ها است. می‌توان آن را در قالب تعداد پیش‌بینی‌های صحیح صورت گرفته به نسبت همه پیش‌بینی‌های صورت گرفته، تعریف شود. با کمک رابطه زیر و با استفاده از ماتریس درهم ریختگی (Confusion Matrix) می‌توان آن را محاسبه کرد.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (۱-۱)$$

صحت(Precision)

صحت استفاده شده در بازیابی اسناد می‌تواند به صورت تعداد اسناد صحیح بازگردانده شده توسط مدل یادگیری ماشین، تعریف شود. با کمک رابطه زیر و با استفاده از ماتریس درهم ریختگی می‌توان آن را به سادگی محاسبه کرد.

$$Precision = \frac{TP}{TP+FP} \quad (۲-۱)$$

یادآوری یا حساسیت (Sensitivity)

یادآوری را می‌توان در قالب تعداد مثبت‌های بازگردانده شده توسط یادگیری ماشین تعریف کرد. با کمک رابطه زیر و با استفاده از ماتریس درهم ریختگی می‌توان آن را به سادگی محاسبه کرد.

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

$$Recall = \frac{TP}{TP+FN} \qquad (۳-۱)$$

اختصاصی بودن (Specificity)

در مقابل یادآوری، اختصاصی بودن را می‌توان در قالب تعداد منفی‌های بازگردانده شده توسط الگوریتم یادگیری ماشین تعریف کرد. با کمک رابطه زیر و با استفاده از ماتریس درهم ریختگی می‌توان آن را به سادگی محاسبه کرد.

$$Specificity = \frac{TN}{TN+FP} \qquad (۴-۱)$$

امتیاز F1(F-score)

این امتیاز، میانگین هارمونیک از دقت و یادآوری را به ما می‌دهد. به بیان ریاضی، امتیاز F1 میانگین وزن دار از دقت و یادآوری است. بهترین مقدار F1، یک و بدترین مقدار صفر است. با کمک رابطه زیر می‌توان امتیاز F1 را محاسبه کرد.

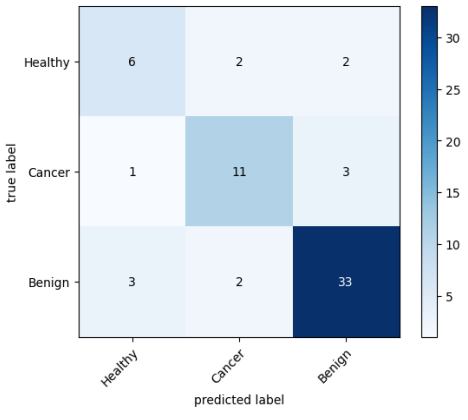
$$F1 = \frac{2*Recall*Precision}{Precision+Recall} \qquad (۵-۱)$$

ماتریس درهم ریختگی

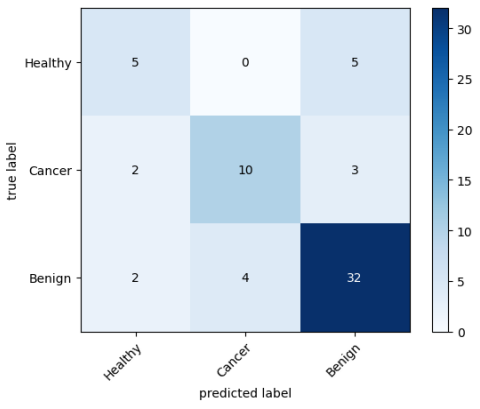
زمانی که خروجی شامل دو نوع کلاس و یا بیشتر باشد، این ساده ترین راه برای اندازه گیری کارایی یک مسئله طبقه‌بندی است. ماتریس درهم ریختگی چیزی جزء یک جدول با دو بعد نیست. مقدار واقعی (Actual Value) و پیش بینی شده، هر دو بعد دارای مثبت–صحیح (True Positive)، منفی–صحیح (True Negative)، مثبت–غلط (False Positive) و منفی–غلط (False Negative) است.

نتایج ارزیابی روش پیشنهادی روی مجموعه داده جمع‌آوری شده

با پیاده‌سازی پنج مدل یادگیری ماشین SVM، MLP، Naive Bayes، Decision Tree و Random Forest روی مجموعه داده بیان شده نتایج زیر حاصل شد.



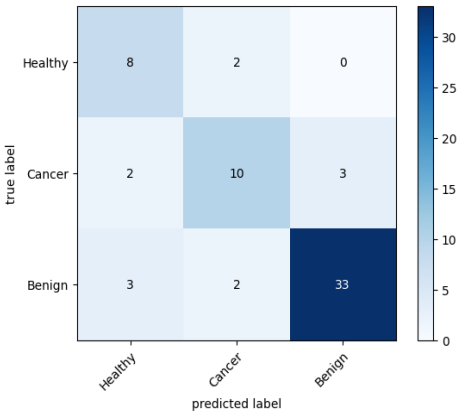
الف) نتایج ماتریس درهم ریختگی MLP



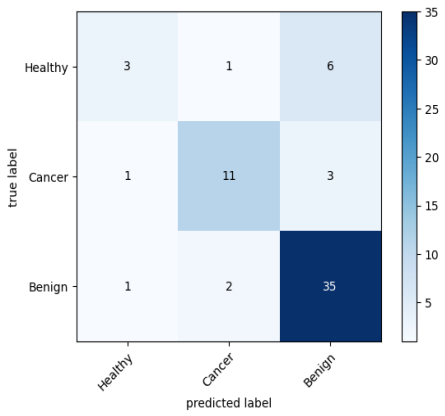
ب) نتایج ماتریس درهم ریختگی مدل Decision Tree

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

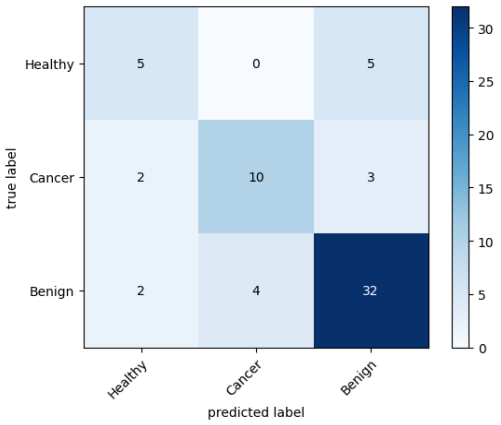
مجله انفورماتیک سلامت و زیست پزشکی دوره یازدهم، شماره ۲، سال ۱۴۰۳، ص ۱۳۰-۱۱۵
doi: 10.34172/jhbmi.2024.18
مرکز تحقیقات انفورماتیک پزشکی، دانشگاه علوم پزشکی کرمان



ج) نتایج ماتریس درهم ریختگی مدل SVM



د) نتایج ماتریس درهم ریختگی مدل Random Forest



ذ) نتایج ماتریس درهم ریختگی مدل Naïve bayes

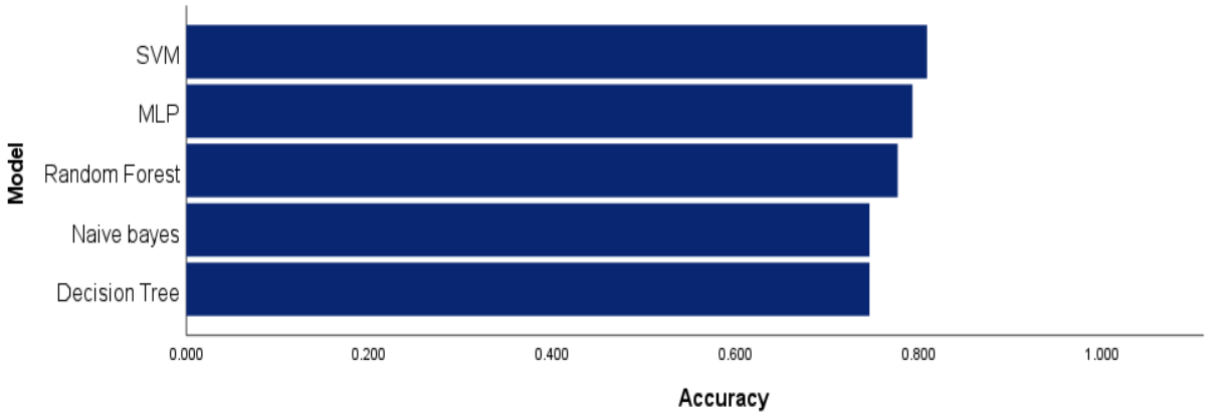
شکل ۲: نتایج ماتریس درهم ریختگی (حالت اول)

همان‌طور که مشاهده می‌کنید با توجه به داده‌های حالت اول در نظر گرفته شده از بانک داده جمع آوری شده [۱۷]، در جدول ۲ خروجی مقادیر Precision، Recall، F-score و Accuracy به ازای هر کدام از پنج مدل یادگیری ماشین SVM، Decision Tree، Naïve bayes، MLP، RandomForest آورده شده است.

جدول ۱: خروجی Precision، Recall، F-score و Accuracy به ازای هر پنج مدل یادگیری ماشین (حالت اول)

رویکرد	Precision	Recall	F-score	Accuracy
SVM	۰/۸۲	۰/۸۱	۰/۸۱	۰/۸۰۹
Decision Tree	۰/۷۴	۰/۷۵	۰/۷۴	۰/۷۴۶
Naïve bayes	۰/۷۴	۰/۷۵	۰/۷۴	۰/۷۴۶
MLP	۰/۷۹	۰/۷۹	۰/۷۹	۰/۷۹۳
Random Forest	۰/۷۶	۰/۷۸	۰/۷۶	۰/۷۷۷

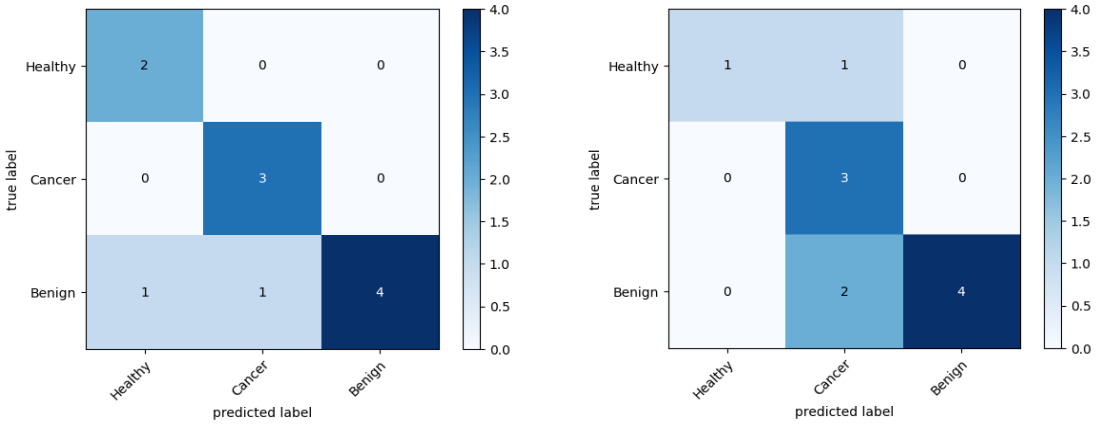
[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]



شکل ۳: معیار ارزیابی دقت به ازای هر پنج مدل یادگیری ماشین (حالت اول)

همان‌طور که در جدول ۱ و شکل ۳ مشاهده می‌کنید در حالت اول در نظر گرفته شده با توجه به معیار دقت، مدل SVM بهتر از سایر مدل‌ها عمل می‌نماید. در ادامه ماتریس درهم ریختگی و نتایج خروجی Precision، Recall، F-score و Accuracy برای هر پنج مدل یادگیری ماشین در حالت دوم در نظر گرفته شده از مجموعه داده بیان شده‌اند.

٤-١: نتایج ماتریس درهم ریختگی MLP

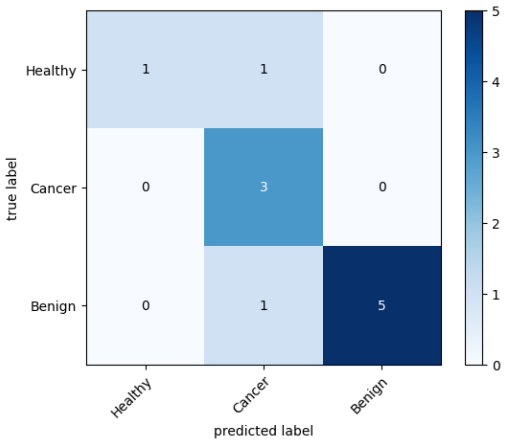


ب) آوای (ای)　　الف) آوای (آ)

[Downloaded from jhbmi.ir on 2026-09-07]

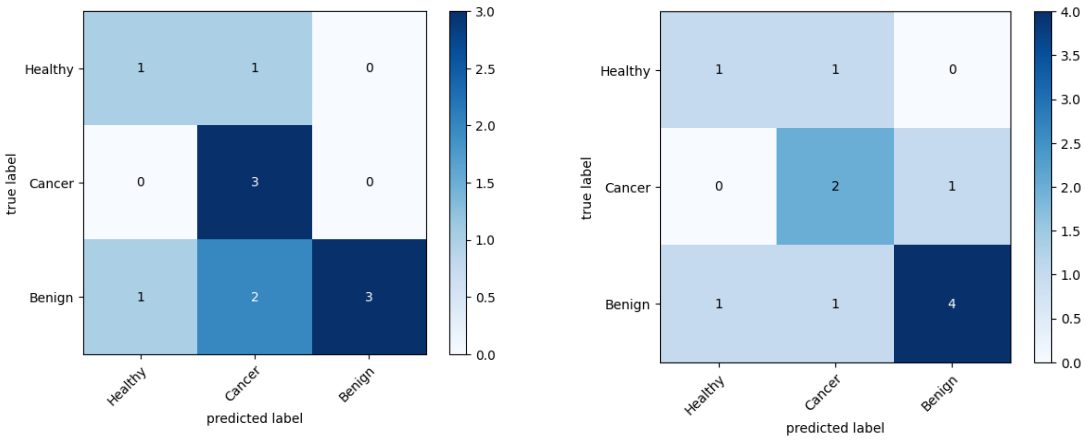
[DOI: 10.34172/jhbmi.2024.18]

مجله انفورماتیک سلامت و زیست پزشکی دوره یازدهم، شماره ۲، سال ۱۴۰۳، ص ۱۳۰-۱۱۵
doi: 10.34172/jhbmi.2024.18
مرکز تحقیقات انفورماتیک پزشکی، دانشگاه علوم پزشکی کرمان

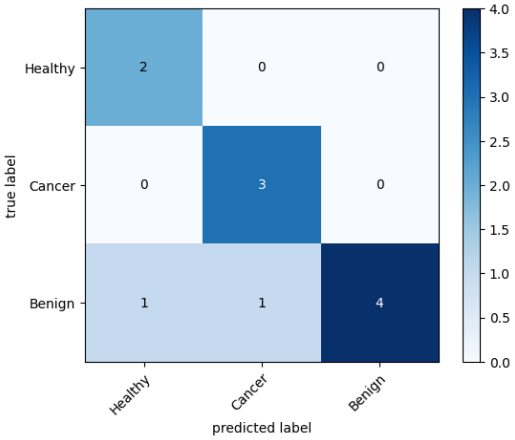


ج) آوای (او)

٤-٢: نتایج ماتریس درهم ریختگی مدل Decision Tree



ب) آوای (ای)

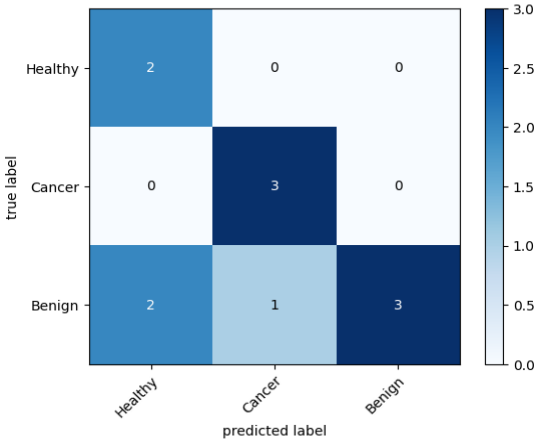


ج) آوای (او)

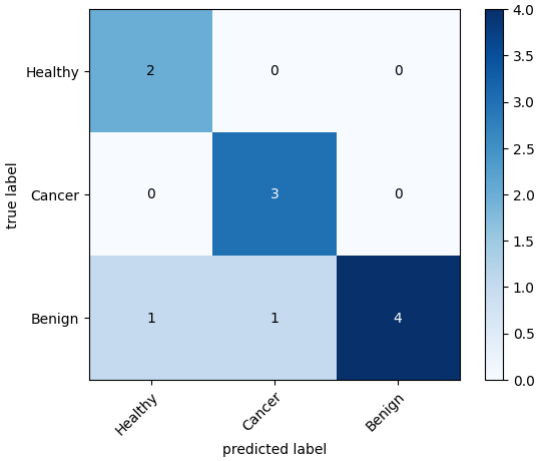
٤-٣: نتایج ماتریس درهم ریختگی مدل SVM

[Downloaded from jhbmi.ir on 2026-09-07]

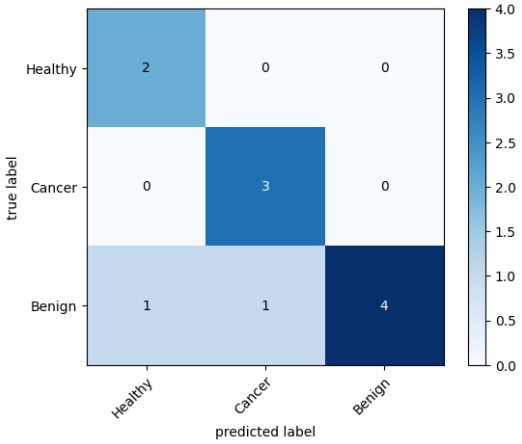
[DOI: 10.34172/jhbmi.2024.18]



ب) آوای (ای)



الف) آوای (آ)



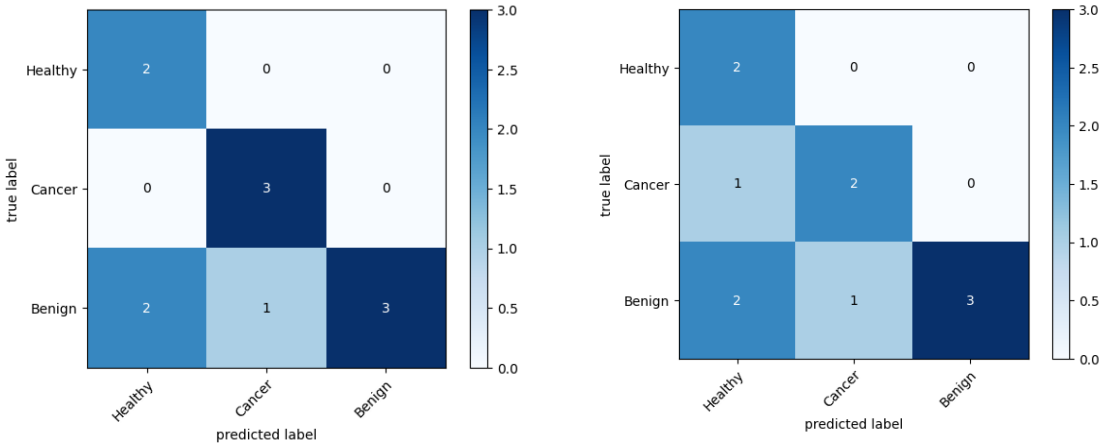
ج) آوای (او)

٤-٤: نتایج ماتریس درهم ریختگی مدل Naïve bayes

[Downloaded from jhbmi.ir on 2026-09-07]

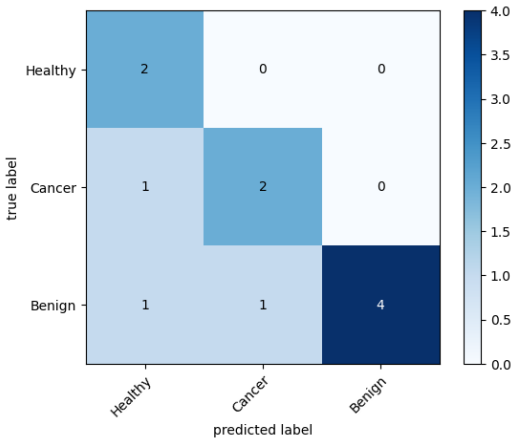
[DOI: 10.34172/jhbmi.2024.18]

مجله انفورماتیک سلامت و زیست پزشکی دوره یازدهم، شماره ۲، سال ۱۴۰۳، ص ۱۳۰-۱۱۵
doi: 10.34172/jhbmi.2024.18
مرکز تحقیقات انفورماتیک پزشکی، دانشگاه علوم پزشکی کرمان



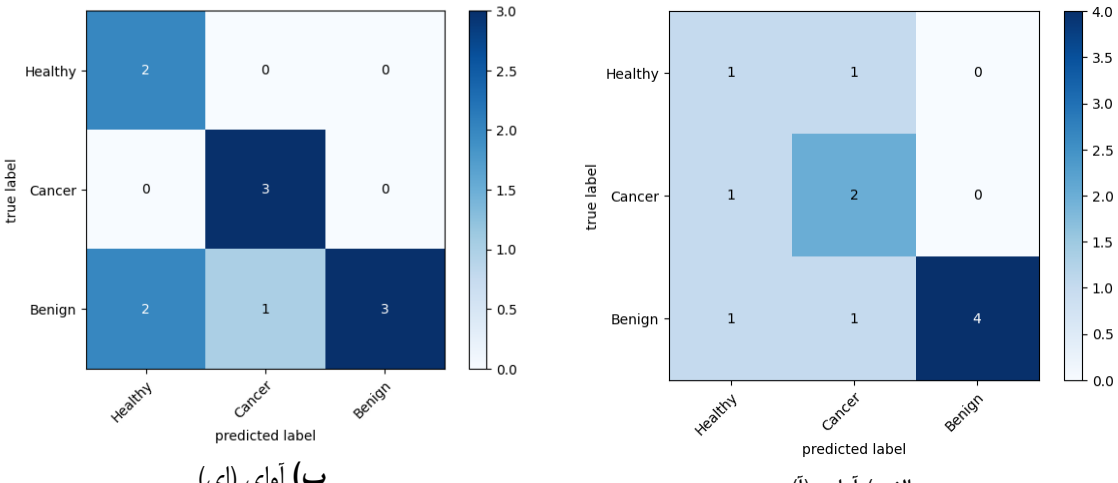
ب) آوای (ای) الف) آوای (آ)

شکل ٤: نتایج ماتریس درهم ریختگی (حالت دوم)



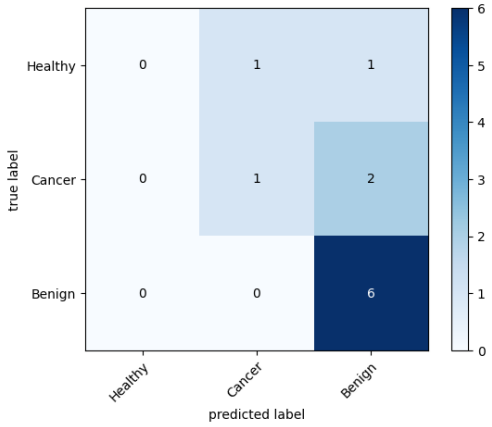
ج) آوای (او)

٤-٥: نتایج ماتریس درهم ریختگی مدل Random Forest



ب) آوای (ای) الف) آوای (آ)

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]



ج) آوای (او)

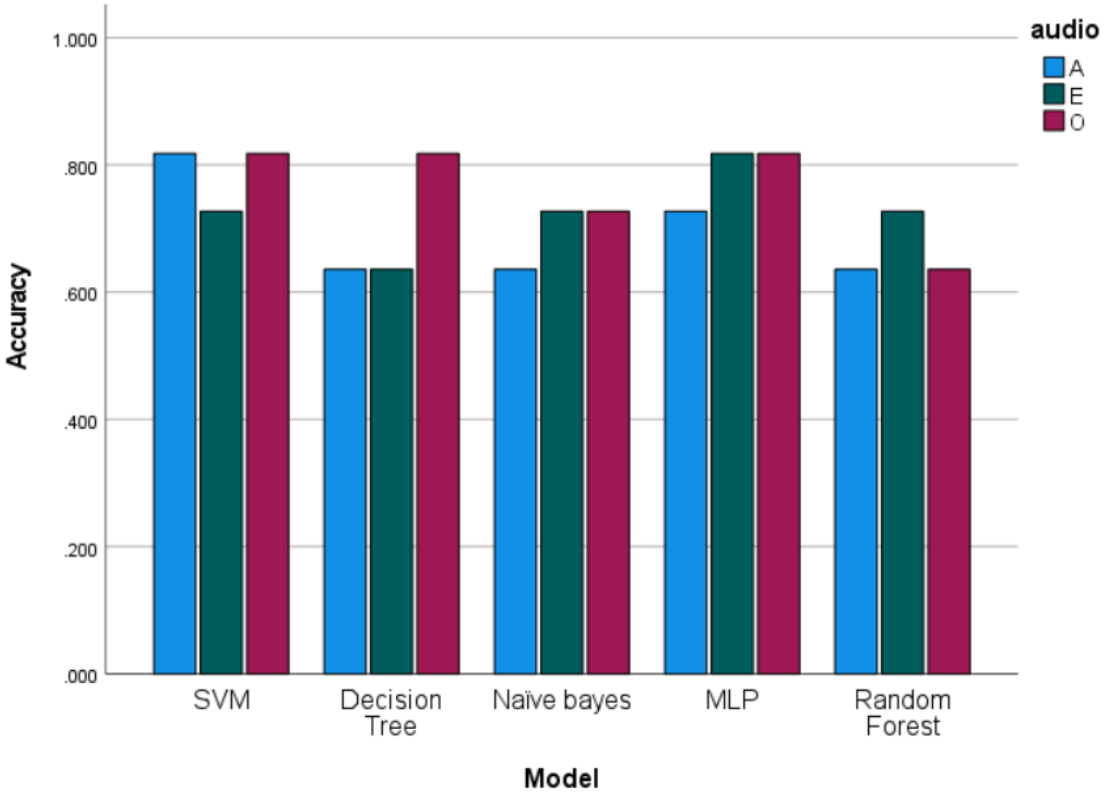
[Downloaded from jhbmi.ir on 2026-09-07]

[DOI: 10.34172/jhbmi.2024.18]

مجله انفورماتیک سلامت و زیست پزشکی دوره یازدهم، شماره ۲، سال ۱۴۰۳، ص ۱۳۰-۱۱۵
doi: 10.34172/jhbmi.2024.18
مرکز تحقیقات انفورماتیک پزشکی، دانشگاه علوم پزشکی کرمان

جدول ۲: خروجی Precision، Recall، F-score و Accuracy به ازای هر پنج مدل یادگیری ماشین (حالت دوم)

رویکرد	Sound	Under sampling	Precision	Recall	F-score	Accuracy
SVM	A	Yes	۰/۸۷	۰/۸۲	۰/۸۲	۰/۸۱۸
	E	Yes	۰/۸۴	۰/۷۳	۰/۷۲	۰/۷۲۷
	O	Yes	۰/۸۷	۰/۸۲	۰/۸۲	۰/۸۱۸
Decision Tree	A	No	۰/۶۶	۰/۶۴	۰/۶۴	۰/۶۳۶
	E	No	۰/۷۷	۰/۶۴	۰/۶۴	۰/۶۳۶
	O	Yes	۰/۸۷	۰/۸۲	۰/۸۲	۰/۸۱۸
Naïve bayes	A	Yes	۰/۸۰	۰/۶۴	۰/۶۵	۰/۶۳۶
	E	Yes	۰/۸۴	۰/۷۳	۰/۷۲	۰/۷۲۷
	O	Yes	۰/۸۲	۰/۷۳	۰/۷۴	۰/۷۲۷
MLP	A	No	۰/۸۶	۰/۷۳	۰/۷۴	۰/۷۲۷
	E	Yes	۰/۸۷	۰/۸۲	۰/۸۲	۰/۸۱۸
	O	Yes	۰/۸۹	۰/۸۲	۰/۸۲	۰/۸۱۸
RandomForest	A	Yes	۰/۷۴	۰/۶۴	۰/۶۶	۰/۶۳۶
	E	Yes	۰/۸۴	۰/۷۳	۰/۷۲	۰/۷۲۷
	O	No	۰/۵۰	۰/۶۴	۰/۵۵	۰/۶۳۶



شکل ۵: میزان دقت به ازای هر پنج مدل یادگیری ماشین و هر آوا (حالت دوم)

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

همان‌طور در جدول ۲ و شکل ۵ نمایان است و با توجه به معیار دقت، آوای A در مدل SVM بیشترین دقت را دارا می‌باشد آوای E در مدل MLP دارای بیشترین دقت است و آوای O در مدل‌های SVM، Decision Tree و MLP بالاترین دقت را دارا است.

بحث و نتیجه‌گیری

استفاده از روش‌های پیشرفته تشخیصی، مانند تجزیه و تحلیل خودکار سیگنال‌های صدا با استفاده از الگوریتم‌های یادگیری ماشین، می‌تواند به زودتر و آسان‌تر شدن تشخیص سرطان حنجره کمک کند. این روش‌ها امکان ارائه نتایج دقیق و سریع را فراهم می‌کنند که به پزشکان کمک می‌کند تا بیماری را در مراحل اولیه تشخیص دهند و درمان مناسب را به موقع شروع کنند. این نوع تشخیص زودهنگام و دقیق می‌تواند به کاهش مرگ و میر ناشی از سرطان حنجره کمک کند و به بیماران امکان درمان مؤثرتر و موفق‌تر را فراهم آورد. در این پژوهش مدلی مبتنی بر یادگیری ماشین جهت تشخیص سرطان حنجره ارائه شده است. تحقیق حاضر به دلیل آن که بر اساس یک مجموعه داده منحصر به فرد جمع‌آوری شده است و تاکنون هیچ تحقیق مشابهی با استفاده از این مجموعه داده انجام نشده است و نشان از ارزش و اهمیت بیشتر تحقیق در ارتقای دانش در این زمینه است. همان‌طور که بیان شد مجموعه داده در نظر گرفته شده بعد از پاکسازی داده‌ها شامل ۱۰۵ نفر از افراد مراجعه کننده به مراکز درمانی شهر کرمان بوده است نتایج پیاده‌سازی ۵ مدل (SVM، Decision Tree، Naïve bayes، MLP، RandomForest) در دو حالت در نظر گرفته نشان داد که در حالتی که تشخیص مد نظر است با توجه به خروجی جدول ۱، مدل SVM با دقت ۰/۸۰۹ از سایر مدل‌ها دقت بالاتری دارد به عبارتی SVM بهتر از سایر مدل‌ها می‌تواند به تشخیص سرطان حنجره با توجه به صوت افراد کمک کند و در حالتی که آواهای (آ)، (ای) و (او)) مد نظر باشد، با توجه به جدول ۲، آوای A در مدل SVM با دقت ۰/۸۱۸ بیشترین دقت را دارا می‌باشد، آوای E با دقت ۰/۸۱۸ در مدل MLP دارای بیشترین دقت است و آوای O با دقت ۰/۸۱۸ در مدل‌های SVM، Decision Tree و MLP بالاترین دقت را دارا است. استفاده از یادگیری ماشین و الگوریتم‌های هوش مصنوعی، می‌تواند ویژگی‌های صوتی افراد را تحلیل کرده و الگوهای خاصی را شناسایی کند. این الگوها می‌توانند به عنوان نشانگرهای زودهنگام بیماری‌ها، به ویژه سرطان، استفاده شوند. با استفاده از این روش، می‌توان بهبود قابل توجهی در تشخیص زودرس بیماری‌ها و افزایش نرخ بازماندگی بیماران داشت. از این روش می‌توان در حوزه‌های پزشکی و تشخیصی بهره برد و بهبود معیارهای بهداشتی و درمانی را به ارمغان آورد و در آخر به دلیل این که تحقیق حاضر با توجه به مجموعه داده جمع آوری شده مشابه سایر تحقیقات نبوده است امکان مقایسه با سایر روش‌های دیگر وجود نداشته است.

پیشنهادات

هدف مطالعات آینده را می‌توان استفاده از دیگر ابزارهای تشخیص سرطان حنجره مانند، استروبوسکپی، نیزوآندوسکپی، آندوسکپی معده در کنار صوت گرفته شده از افراد قرار داد به طوری که دقت مدل افزایش یابد و همچنین پیشنهاد می‌شود یک مجموعه داده بزرگ‌تر جمع‌آوری گردد تا دقت مدل‌های پیشنهادی در این تحقیق را مورد بررسی قرار دهد.

تعارض منافع

نویسندگان اظهار داشتند که هیچ‌گونه تضاد منافعی وجود ندارد.

کد اخلاق

کد اخلاق این مطالعه از دانشگاه علوم پزشکی کرمان با شماره IR.KMU.REC.1403.063 گرفته شده است.

سهم مشارکت نویسندگان

کلیه نویسندگان سهم یکسانی در انجام مقاله داشته‌اند.

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]

References

[1]. Sarmah SS. Concept of artificial intelligence, its impact and emerging trends. International Research Journal of Engineering and Technology 2019;6(11):2164-8.
[2]. Zhang Y. New Advances in Machine Learning. Intechopen; 2010.
[3]. Alatrany AS, Hussain AJ, Mustafina J, Al-Jumeily D. Machine learning approaches and applications in genome wide association study for Alzheimer's disease: A systematic review. IEEE Access 2022;10:62831-47. doi**:** 10.1109/ACCESS.2022.3182543
[4]. Maimon O, Rokach L, editors. Data Mining and Knowledge Discovery Handbook. New York: Springer; 2005.
[5]. Zhou Y, Qiu G. Random forest for label ranking. Expert systems with Applications 2018;112:99-109.
[6]. Rani A, Kumar N, Kumar J, Sinha NK. Machine learning for soil moisture assessment. In Deep learning for sustainable agriculture. Academic Press; 2022. p. 143-68.
[7]. Chaudhuri K.D. Building Naive Bayes Classifier from Scratch to Perform Sentiment Analysis. Analytics Vidya.com. 2022. [2023 Aug 16]. Available from: https://www.analyticsvidhya.com/blog/2022/03/building-naive-bayes-classifier-from-scratch-to-perform-sentiment-analysis.
[8]. Nazzal JM, El-Emary IM, Najim SA, Ahliyya A. Multilayer perceptron neural network (MLPs) for analyzing the properties of Jordan oil shale. World Applied Sciences Journal. 2008;5(5):546-52.
[9]. Hamedi S, Kordrostami Z, Yadollahi A. Artificial neural network approaches for modeling absorption spectrum of nanowire solar cells. Neural Computing and Applications 2019;31(12):8985-95. doi:10.1007/s00521-019-04406-3
[10]. Al-Dhief FT, Latiff NM, Malik NN, Salim NS, Baki MM, Albadr MA, Mohammed MA. A survey of voice pathology surveillance systems based on internet of things and machine learning algorithms. IEEE Access 2020;8:64514-33. doi: 10.1109/ACCESS.2020.2984925
[11]. Sayadi M, Varadarajan V, Langarizadeh M, Bayazian G, Torabinezhad F. A systematic review on machine learning techniques for early detection of mental, neurological and laryngeal disorders using patient's speech. Electronics 2022;11(24):4235. https://doi.org/10.3390/electronics11244235
[12]. Mesallam TA, Farahat M, Malki KH, Alsulaiman M, Ali Z, Al-Nasheri A, et al. Development of the arabic voice pathology database and its evaluation by using speech features and machine learning algorithms. Journal of healthcare engineering. 2017;2017(1):8783751. https://doi.org/10.1155/2017/8783751
[13]. Saloni, Sharma RK, Gupta AK. Disease detection using voice analysis: A review. International Journal of Medical Engineering and Informatics 2014;6(3):189-209. https://doi.org/10.1504/IJMEI.2014.063173
[14]. Kim H, Jeon J, Han YJ, Joo Y, Lee J, Lee S, et al. Convolutional neural network classifies pathological voice change in laryngeal cancer with high accuracy. J. Clin Med 2020, 9(11): 3415. https://doi.org/10.3390/jcm9113415.
[15]. Gour GB, Udayashankara V, Badakh DK, Kulkarni YA. Voice-Disorder Identification of Laryngeal Cancer Patients. International Journal of Advanced Computer Science and Applications. 2020;11(11). doi: 10.14569/IJACSA.2020.0111145
[16]. McFee B, Raffel C, Liang D, Ellis DP, McVicar M, Battenberg E, Nieto O: librosa: Audio and music signal analysis in python. In: SciPy: 2015; 2015: 18-24.
[17]. Larynx_Cancer code [cited 2024 May 14]. Available from: https://github.com/muhammadmo/Larynx_cancer.

[Downloaded from jhbmi.ir on 2026-09-07]
[DOI: 10.34172/jhbmi.2024.18]