

## Evaluation and comparison of machine learning models in predicting spirometric indices based on individual characteristics in Fasa county

Sareh Rafatmagham<sup>1</sup>

Neda Jafari<sup>\*2</sup>

Azizollah Dehghan<sup>3</sup>

Mohammad Roshanzamir<sup>4</sup>

1- Department of Internal Medicine, School of Medicine, Fasa University of Medical Sciences, Fasa, Iran

2- Department of Internal Medicine, School of Medicine, Fasa University of Medical Sciences, Fasa, Iran ,

3- School of Health, Fasa University of Medical Sciences, Fasa, Iran

4- Department of Computer Engineering, School of Engineering, Fasa University, Fasa, Iran

\* Corresponding Author:  
nedajafari.med@gmail.com

### Abstract

**Introduction:** Chronic respiratory diseases impose a substantial burden on public health. Accurate prediction of spirometric indices based on individual characteristics can enhance the diagnosis and management of these conditions. This study aimed to evaluate and compare the performance of machine learning models in predicting spirometric indices using individual features in the population of Fasa County.

**Methods:** This cross-sectional observational study was conducted on 5,450 individuals who attended pulmonary clinics between spring 2024 and autumn 2025. After excluding incomplete records, 4,615 eligible participants were included in the final analysis. Individual characteristics included age, sex, height, weight, and smoking status. Spirometric indices comprised forced vital capacity (FVC), forced expiratory volume in one second (FEV1), FEV1/FVC ratio, forced expiratory flow at 25–75% (FEF25–75), and peak expiratory flow (PEF). Five machine learning algorithms—Linear Regression, K-Nearest Neighbors, Support Vector Machine, Random Forest, and Gradient Boosting—were evaluated using mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination ( $R^2$ ). The relative importance of individual features was assessed using the feature importance output of the Gradient Boosting model. All ethical principles were observed, and the study was approved under ethics code IR.FUMS.REC.1404.174.

**Results:** The Gradient Boosting model demonstrated the lowest error rates and the highest  $R^2$  values across all spirometric indices, indicating superior performance compared to the other models. Age, height, and weight were the most influential predictors, whereas sex and smoking status had relatively lower contributions to model performance.

**Conclusion:** The Gradient Boosting model effectively captured nonlinear relationships and complex interactions between individual characteristics and spirometric indices. These findings suggest that machine learning approaches can support the development of localized predictive models and improve the interpretation of spirometry results in native populations.

**Keywords:** Spirometry, Respiratory Function Tests, Machine Learning, Pulmonary Disease, Chronic, Cross-Sectional Studies

**How to cite this article:** Rafatmagham S, Jafari N, Dehghan A, Roshanzamir M. Evaluation and comparison of machine learning models in predicting spirometric indices based on individual characteristics in Fasa county. Alborz University Medical Journal 2026; 15 (1) : 1-16

## ارزیابی و مقایسه عملکرد مدل‌های یادگیری ماشین در پیش‌بینی شاخص‌های اسپرومتری مبتنی بر ویژگی‌های فردی در شهرستان فسا

### چکیده

**مقدمه:** بیماری‌های تنفسی مزمن بار بالایی بر سلامت جامعه دارد و پیش‌بینی دقیق شاخص‌های اسپرومتری با استفاده از ویژگی‌های فردی می‌تواند در تشخیص و مدیریت این بیماری‌ها مؤثر باشد. هدف این مطالعه، ارزیابی و مقایسه عملکرد مدل‌های یادگیری ماشین در پیش‌بینی شاخص‌های اسپرومتری بر اساس ویژگی‌های فردی در جمعیت شهرستان فسا است.

**روش کار:** این مطالعه مشاهده‌ای مقطعی از ۵۴۵۰ مراجعه‌کننده به مطب ریه در بازه بهار ۱۴۰۳ تا پاییز ۱۴۰۴، پس از حذف داده‌های ناقص، ۴۶۱۵ فرد واجد شرایط وارد تحلیل نهایی شد. ویژگی‌های فردی شامل سن، جنسیت، قد، وزن و وضعیت سیگاری بودن بود و شاخص‌های اسپرومتری نیز شامل FEV1، FEV1/FVC، FEF25-75، FVC، PEF و PEF بود. همچنین، پنج الگوریتم یادگیری ماشین شامل رگرسیون خطی، الگوریتم کی-نزدیک‌ترین همسایه، ماشین بردار پشتیبان، جنگل تصادفی و تقویت گرادیان بر اساس معیارهای RMSE، MSE، و R<sup>2</sup> ارزیابی شدند. اهمیت نسبی ویژگی‌ها نیز با خروجی feature\_importances مدل تقویت گرادیان تحلیل شد.

**یافته‌ها:** مدل تقویت گرادیان در تمامی شاخص‌های اسپرومتری کمترین خطا و بیشترین ضریب تعیین را داشت و عملکرد برتری نسبت به سایر مدل‌ها نشان داد. سن، قد و وزن بیشترین تأثیر را بر پیش‌بینی شاخص‌ها داشت، در حالی که جنسیت و وضعیت سیگاری بودن تأثیر کمتری داشت.

**نتیجه‌گیری:** مدل تقویت گرادیان قادر است روابط غیرخطی و تعاملات پیچیده بین ویژگی‌های فردی و شاخص‌های اسپرومتری را شناسایی کند. این مطالعه نشان می‌دهد که استفاده از یادگیری ماشین می‌تواند در ایجاد مدل‌های پیش‌بینی محلی و بهبود تفسیر نتایج اسپرومتری در جمعیت‌های بومی مفید باشد.

**واژه‌های کلیدی:** یادگیری ماشین، شاخص‌های اسپرومتری، ویژگی‌های فردی، شهرستان فسا.

ساره رفعت مقام<sup>۱</sup>

ندا جعفری<sup>۲\*</sup>

عزیزالله دهقان<sup>۳</sup>

محمد روشن ضمیر<sup>۴</sup>

۱. گروه داخلی، دانشکده پزشکی، دانشگاه علوم

پزشکی فسا، فسا، ایران

۲. گروه داخلی، دانشکده پزشکی، دانشگاه علوم

پزشکی فسا، فسا، ایران

۳. دانشکده بهداشت، دانشگاه علوم پزشکی

فسا، فسا، ایران

۴. گروه مهندسی کامپیوتر، دانشکده مهندسی،

دانشگاه فسا، فسا، ایران

\* نویسنده مسئول:

nedajafari.med@gmail.com

## مقدمه

بیماری های تنفسی مزمن از جمله آسم، «بیماری مزمن انسدادی ریه ۱» و فیبروز ریوی، بار قابل توجهی بر سیستم های بهداشتی و کیفیت زندگی افراد در سراسر جهان تحمیل می کند (۱). پایش، ارزیابی و تحلیل دقیق عملکرد ریه با استفاده از اسپرومتری به عنوان یک تست استاندارد و غیرتهاجمی، نقش حیاتی در مدیریت این بیماری ها، پیشگیری از پیشرفت آن ها و بهبود پیامدهای بالینی دارد (۲). در این راستا اسپرومتری ابزاری کلیدی در ارزیابی حجم ها و جریان های هوایی قابل اندازه گیری در طی مانورهای دم و بازدم اجباری به شمار می رود و شاخص های حاصل از آن اطلاعات ارزشمندی در مورد وضعیت عملکرد تنفسی فرد ارائه می دهد و مبنای مهمی برای پیش بینی و برآورد وضعیت عملکرد ریوی به شمار می آید (۳). اگرچه این روش قادر به اندازه گیری حجم های غیر قابل تخلیه مانند «حجم باقی مانده ۲» و «ظرفیت کل ریه ۲» نیست.

شناخت دقیق عوامل فردی مؤثر بر شاخص های اسپرومتری مانند سن، جنسیت، قد، وزن و وضعیت استعمال دخانیات، می تواند دقت تفسیر یافته های این آزمون را افزایش دهد و به طراحی مداخلات پیشگیرانه و درمانی مؤثرتر کمک کند. این موضوع در جمعیت هایی که مقادیر مرجع بومی برای «معادلات پیش بینی ۴» در دسترس نیست، اهمیت بیشتری دارد، زیرا استفاده از معادلات و استانداردهای غیربومی می تواند منجر به «کلاسه بندی نادرست ۵» یا برآورد شدت بیماری شود.

مطالعه های متعدد نشان داده است که با افزایش سن، به دلیل تغییرات ساختاری در ریه ها و کاهش الاستیسیته بافت ریه و قفسه سینه،  $FVC^6$  و  $FEV_1^7$  به صورت تدریجی کاهش می یابد (۴ و ۵). این کاهش در عملکرد ریه یک فرایند فیزیولوژیک مرتبط با پیری است. همچنین به طور کلی، مردان به دلیل داشتن جثه بزرگتر و حجم ریوی بیشتر، مقادیر مطلق بالاتری از  $FVC$  و  $FEV_1$  را در مقایسه با زنان نشان می دهند (۶). قد نیز یکی از قوی ترین عوامل تعیین کننده حجم ریه ها است. افراد بلندقدتر به

طور معمول ریه های بزرگتری دارند و در نتیجه، مقادیر بالاتری از  $FVC$  و  $FEV_1$  را نشان می دهند (۷). این موضوع به خوبی در معادلات پیش بینی نرمال برای شاخص های اسپرومتری لحاظ شده است (۸). همچنین، چاقی و افزایش BMI می تواند با محدود کردن حرکت دیافراگم و قفسه سینه، منجر به کاهش حجم های ریوی مانند  $FVC$  و  $FEV_1$  شود (۹). افزون بر آن، التهاب مزمن ناشی از چاقی ممکن است بر عملکرد راه های هوایی کوچک تأثیر گذاشته و منجر به کاهش  $FEF_{25-75}$  شود (۱۰). از سوی دیگر، استعمال دخانیات یک عامل مهم شناخته شده برای آسیب به ساختار و عملکرد ریه ها است. مواد شیمیایی موجود در دخانیات می تواند منجر به التهاب مزمن، تخریب بافت ریه (آمفیزم)، افزایش تولید مخاط و تنگی راه های هوایی شود که همگی به کاهش شاخص های اسپرومتری منجر می شود (۱۱ و ۱۲). تأثیر سیگار بر  $FEF_{25-75}$ ، به دلیل حساسیت این شاخص به تغییر در راه های هوایی کوچک، قابل توجه است (۱۳). مطالعه های دیگری نیز نشان داده است که ویژگی های فردی مانند سن، جنسیت، قد و وزن به عنوان قوی ترین عوامل دموگرافیک، به گونه ای طبیعی بر مقادیر نرمال شاخص های اسپرومتری تأثیر می گذارد (۱۴). همچنین، عوامل محیطی و رفتاری مانند استعمال دخانیات به گونه مستقیم با کاهش عملکرد ریه و افزایش خطر ابتلا به بیماری های تنفسی مرتبط است (۱۵). درک دقیق چگونگی تعامل این عوامل فردی با شاخص های اسپرومتری در یک جمعیت خاص می تواند به ایجاد مقادیر مرجع منطقه ای، شناسایی گروه های در معرض خطر و بهبود تفسیر نتایج اسپرومتری کمک کند. اگرچه اذعان می شود که عوامل بالینی و محیطی متعددی می تواند بر شاخص های اسپرومتری تأثیر گذار باشد، اما مطالعه های متعددی در سطح جهانی نیز از ویژگی های دموگرافیک به عنوان پیش بینی کننده های اصلی در مدل های اولیه استفاده کرده اند. هدف این پژوهش، خلق یک ابزار تشخیص کامل نیست، بلکه کشف قابلیت و کارایی یادگیری ماشین در تحلیل الگوهای موجود در داده های پایه ای و فراگیر است. با وجود این، پژوهش جامعی که به پیش بینی با طیف وسیعی از ویژگی های فردی و شاخص های اصلی اسپرومتری با استفاده از روش های پیشرفته یادگیری ماشین در جمعیت شهرستان فسا پرداخته باشد، مشاهده نشد. شهرستان فسا با ویژگی های نیمه صنعتی خود، ترکیبی از مواجهه با آلاینده های کشاورزی و صنعتی سبک دارد که می تواند بر سلامت ریه تأثیر گذار باشد. از طرفی،

8 Forced Expiratory Flow between 25% and 75% of FVC

- 1 COPD
- 2 RV
- 3 TLC
- 4 Reference Equations
- 5 Misclassification
- 6 Forced Vital Capacity
- 7 Forced Expiratory Volume in 1 Second

تشخیص و پایش این بیماری‌ها، تأثیر عوامل فردی بر شاخص‌های اسپیرومتری و محدودیت‌های روش‌های سنتی در تحلیل این روابط، انجام پژوهشی که با بهره‌گیری از روش‌های یادگیری ماشین این ارتباط را در جمعیت شهرستان فسا بررسی کند، امری ضروری و ارزشمند است.

این پژوهش عملکرد مدل‌های مختلف را مقایسه و نقاط قوت و ضعف هر روش را ارزیابی می‌کند. هدف اصلی این پژوهش، تعیین و مدل‌سازی با مجموعه‌ای از ویژگی‌های فردی شامل سن، جنسیت، قد، وزن و وضعیت استعمال دخانیات و شاخص‌های کلیدی اسپیرومتری شامل  $PEF^1$  و  $FEF_{25-75}$ ،  $FEV_1/FVC$ ،  $FVC$ ،  $FEV_1$  در جمعیت مراجعه‌کننده به مطب تخصصی ریه در شهرستان فسا و تعیین بهینه‌ترین مدل از نظر دقت پیش‌بینی در بین الگوریتم‌های مورد مقایسه است. هدف فرعی این پژوهش نیز ارزیابی قدرت پیش‌بینی مدل‌های یادگیری ماشین در تخمین شاخص‌های اسپیرومتری بر اساس ویژگی‌های فردی و شناسایی قوی‌ترین ارزیاب در بین روش‌های یادگیری ماشین است و همچنین، بررسی اهمیت نسبی ویژگی‌ها در بهترین مدل انتخاب شده برای درک بهتر سازوکارهای فیزیولوژیک در این جمعیت است. لازم به تأکید است که پژوهش حاضر با هدف تشخیص بیماری انجام نشده است، بلکه تمرکز آن صرفاً بر پیش‌بینی شاخص‌های عملکرد ریوی و تحلیل تأثیر ویژگی‌های فردی بر آن‌ها با بهره‌گیری از الگوریتم‌های یادگیری ماشین است.

### روش بررسی

این پژوهش یک مطالعه مشاهده‌ای از نوع مقطعی است که به صورت «پس‌نگر» و بر پایه تحلیل ثانویه داده‌های سرشماری از پرونده‌های بیماران انجام شد. معیار ورود شامل بیماران مراجعه‌کننده به مطب ریه در شهرستان فسا در بازه زمانی بهار ۱۴۰۳ تا پاییز ۱۴۰۴ است که نتایج تست اسپیرومتری و اطلاعات جمعیت‌شناختی آن‌ها در پرونده موجود است. داده‌ها به صورت ثانویه از پرونده‌های بایگانی شده استخراج شد. معیار خروج نیز شامل کامل نبودن اطلاعات پرونده بیمار یا عدم دسترسی به اطلاعات آنها است. همچنین، از کلیه تست‌های گرفته شده از بیماران، «بهترین تست‌ها»<sup>۲</sup> به تشخیص نرم‌افزار دستگاه اسپیرومتری Winspiro Spirolab برند MIR Italy انتخاب شد. انتخاب بهترین تست نیز بر

الگوی مصرف دخانیات در فسا، شامل مصرف سیگار و قلیان و مواجهه با دود غیرمستقیم، ممکن است منجر به تغییرات متفاوت در عملکرد ریه نسبت به سایر مناطق شود.

روش‌های آماری سنتی مانند رگرسیون خطی، باوجود سادگی، در شناسایی تعاملات پیچیده و اثرات غیرخطی محدودیت دارد. در نتیجه، روابط بین ویژگی‌های فردی و شاخص‌های اسپیرومتری می‌تواند ماهیتی غیرخطی و چندبعدی داشته باشد. در مقابل، الگوریتم‌های یادگیری ماشین قادر است بدون اتکا به این فرضیات، الگوهای پنهان و تعامل‌های چندمتغیره را شناسایی کند و پیش‌بینی‌هایی با دقت بالاتر ارائه دهد؛ بنابراین، پژوهش حاضر به دنبال ارزیابی دقیق قابلیت یادگیری ماشین در مدل‌سازی این روابط پیچیده است.

مطالعه حاضر با ارائه مدلی مبتنی بر یادگیری ماشین که قادر به پیش‌بینی شاخص‌های اسپیرومتری بر اساس ویژگی‌های فردی است، به دانش موجود در زمینه فیزیولوژی تنفس و عوامل مؤثر بر عملکرد ریه می‌افزاید. همچنین، بررسی اهمیت نسبی هر یک از ویژگی‌های فردی در پیش‌بینی شاخص‌های اسپیرومتری می‌تواند درک عمیق‌تری از سازوکارهای دخیل در عملکرد ریه فراهم کند. افزون بر آن، این پژوهش می‌تواند کاربردهای عملی متعددی داشته باشد، زیرا ایجاد مدل‌های پیش‌بینی محلی برای شاخص‌های اسپیرومتری می‌تواند به پزشکان در تفسیر دقیق‌تر نتایج این بررسی در جمعیت شهرستان فسا کمک کند و از تفسیرهای نادرست ناشی از استفاده از مقادیر مرجع عمومی جلوگیری نماید. شناسایی عوامل فردی با بیشترین تأثیر بر عملکرد ریه نیز می‌تواند به طراحی برنامه‌های پیشگیرانه و مداخلات بهداشتی هدفمند برای گروه‌های در معرض خطر در این منطقه منجر شود. به‌عنوان نمونه، اگر مشخص شود که استعمال دخانیات در این جمعیت تأثیر بسیار قوی بر کاهش  $FEF_{25-75}$  دارد، برنامه‌های ترک دخانیات می‌تواند اولویت بالاتری پیدا کند. افزون بر آن، مدل‌های یادگیری ماشین می‌تواند الگوهای پیچیده‌ای را در داده‌ها شناسایی کند که ممکن است با روش‌های آماری سنتی قابل تشخیص نباشد. این امر می‌تواند به شناسایی افراد مستعد کاهش عملکرد ریوی و ارائه مداخلات به‌موقع کمک کند. در نتیجه با درک بهتر عوامل مؤثر بر عملکرد ریه در سطح جامعه، می‌توان برنامه‌های ارتقای سلامت و پیشگیری از بیماری‌های تنفسی را به طور مؤثرتری طراحی و اجرا کرد.

با توجه به بار بالای بیماری‌های تنفسی، نقش حیاتی اسپیرومتری در

- 1 Peak Expiratory Flow
- 2 Retrospective
- 3 Best

برای بررسی موضوع، پژوهشگر پس از کسب اجازه های لازم و کد تأیید اخلاق (IR.FUMS.REC.1404.174) و هماهنگی با پزشک مربوطه، اطلاعات لازم را جمع آوری کرد. افزون بر آن، اجبار برای استفاده از خدمات اسپرومتری به هیچ عنوان وجود نداشته و مراجعه کنندگان با رضایت آگاهانه به این تست پرداخته اند. همچنین، تمامی اطلاعات شرکت کنندگان به صورت محرمانه نگهداری شده و فقط برای اهداف پژوهش استفاده شد. برای حفظ حریم خصوصی، اطلاعات با استفاده از کد جایگزین شد و فقط پژوهشگران به کلید کد دسترسی داشتند. شرکت در این پژوهش هیچ گونه خطر یا آسیب جسمی یا روانی برای شرکت کنندگان ندارد، زیرا از داده های موجود استفاده می شود و مداخله ای روی شرکت کنندگان انجام نمی شود. با رعایت این ملاحظات اخلاقی، حقوق و سلامت شرکت کنندگان در این پژوهش تضمین خواهد شد. سنجش عملیاتی متغیرها نیز به شیوه زیر انجام شد.

**متغیرهای مستقل:** ویژگی های فردی مراجعه کنندگان در پژوهش حاضر شامل سن، جنسیت، قد، وزن و سیگاری بودن یا نبودن است که سن به سال بیان می شود. جنسیت نیز به صورت صفر و یک بیان میشود که صفر برای زنان و یک برای مردان. همچنین، قد به سانتی متر ارائه میشود. وزن نیز به کیلوگرم بیان میشود. افزون بر آن، وضعیت استعمال دخانیات به صورت صفر و یک بیان می شود که صفر برای افرادی که استعمال دخانیات ندارند و یک برای افرادی که به استعمال دخانیات می پردازند. از آنجاکه اطلاعات مربوط به مدت زمان مصرف دخانیات، تعداد نخ سیگار در روز و شاخص بسته - سال در پرونده ها ثبت نشده بود، وضعیت استعمال دخانیات تنها به صورت یک متغیر دودویی (بلی/خیر) در مدل ها وارد شد که این موضوع می تواند به عنوان یکی از محدودیت های پژوهش حاضر در نظر گرفته شود.

افزون بر مقیاس بندی متغیرهای پیوسته، متغیرهای طبقه ای دوتایی (جنسیت و وضعیت استعمال دخانیات) که به صورت صفر و یک کدگذاری شده است، برای استفاده در مدل های رگرسیون خطی و ماشین بردار پشتیبان به عنوان «متغیرهای مجازی ۱» در نظر گرفته شد تا از تحمیل فرض ترتیبی یا پیوسته بودن جلوگیری شود.

**متغیرهای وابسته:** عامل های مورد بررسی اسپرومتری نیز شامل موارد زیر است که جدول ۱ آن را نشان می دهد.

اساس رعایت معیارهای پذیرش و تکرارپذیری طبق دستورالعمل های استاندارد ATS/ERS صورت گرفت.

دلیل حدود ۱۸ ماهه بودن دوره بررسی شامل محدودیت های زمانی و منابع، ثبات نسبی ویژگی ها و پروتکل ها، کاهش اثرات عوامل مخدوش کننده فصلی یا دوره ای و دسترسی به داده ها است. در راستای اطمینان بخشی به نمونه گیری و کاهش خطای نمونه گیری و دقت نظر بیشتر، کل مراجعه کنندگانی که اطلاعات آنها برای انجام مطالعه موجود باشد (اطلاعات کامل مورد نیاز برای پژوهش شامل اطلاعات جمعیت شناختی و نتایج تست اسپرومتری) به عنوان نمونه در نظر گرفته شد. به این معنا که هیچ گونه نمونه گیری احتمالی به عمل نیامد و به روش شمارش کامل از کل داده های واجد شرایط و در دسترس استفاده و جامعه و نمونه پژوهش یکسان در نظر گرفته شد. با این حال، به منظور آموزش، اعتبارسنجی و ارزیابی عملکرد مدل های یادگیری ماشین، داده ها در ادامه به مجموعه های آموزش و آزمون تقسیم شدند که این امر صرفاً یک مرحله فنی در مدل سازی بوده و به معنای نمونه گیری آماری جدید نیست. باتوجه به اینکه حجم جامعه پژوهش بیش از ۵۰۰۰ تست است، تمامی موارد واجد شرایط وارد مطالعه می شود و بدین ترتیب حجم نمونه تضمین خواهد شد. همچنین، با عنایت به در دسترس بودن داده های کافی و استفاده از کل جامعه واجد شرایط، نیازی به برآورد حجم نمونه جداگانه نیست.

برای کاهش «تورش ۱» انتخاب و اطمینان از همگنی داده ها در راستای هدف اصلی پژوهش، یعنی مقایسه متدولوژیک مدل های پیش بینی، استفاده از داده های حاصل از چند مطب توصیه نمی شود، زیرا هر مرکز ممکن است از دستگاه های اسپرومتری با برند، مدل، کالیبراسیون و الگوریتم های اندازه گیری متفاوتی استفاده کند که این ناهمگنی می تواند منجر به واریانس ناخواسته در مقادیر ثبت شده شاخص های اسپرومتری شود و دقت نتایج را تحت تأثیر قرار دهد. افزون بر این، تفاوت در شرایط محیطی، نحوه نگهداری و فرکانس کالیبراسیون دستگاه ها در مراکز مختلف نیز می تواند منبع دیگری از خطا و نوسان داده ها باشد. بر این اساس، تمرکز بر یک مطب فوق تخصصی که از دستگاه اسپرومتری ثابت، کالیبره شده و دارای پروتکل های اجرایی یکنواخت استفاده می کند، امکان کنترل بهتر متغیرهای مخدوش کننده و افزایش صحت و قابلیت اطمینان داده ها را فراهم می سازد. این رویکرد برای دستیابی به اهداف فنی و مقایسه ای پژوهش حاضر از اهمیت ویژه ای برخوردار است.

## 2 Dummy Variables

## 1 Bias

جدول ۱- نماد، عنوان و تعریف عملیاتی شاخص‌های اسپرومتری

نماد	عنوان	تعریف عملیاتی
FVC	ظرفیت حیاتی اجباری	حجم کل هوایی که فرد پس از یک دم عمیق، با حداکثر تلاش و با سرعت تا آخر بازدم می‌تواند از ریه‌های خود خارج کند.
FEV <sub>1</sub>	حجم بازدم اجباری در یک ثانیه	حجم هوایی که فرد در اولین ثانیه از بازدم اجباری (همان تستی که برای FVC انجام می‌شود) می‌تواند از ریه‌های خود خارج کند.
FEV <sub>1</sub> /FVC	نسبت حجم بازدمی اجباری در یک ثانیه به ظرفیت حیاتی اجباری	تقسیم نسبت‌های ذکر شده که به صورت درصد بیان می‌شود.
FEF25-75	جریان بازدمی اجباری بین ۲۵ تا ۷۵ درصد ظرفیت حیاتی اجباری	میانگین سرعت جریان هوا در میانه بازدم اجباری، یعنی زمانی که بین ۲۵٪ تا ۷۵٪ از کل حجم بازدمی (FVC) خارج شده است.
PEF	حداکثر جریان بازدمی	حداکثر سرعتی که فرد می‌تواند هوا را در ابتدای بازدم اجباری از ریه‌های خود خارج کند.

ماشین را فراهم می‌کند. پایتون همچنین قابلیت کار با داده‌های پیچیده، انجام پیش‌پردازش، نرمال‌سازی، تقسیم داده به مجموعه‌های آموزش و آزمون و محاسبه شاخص‌های ارزیابی مدل مانند RMSE و R<sup>2</sup> را در یک محیط یکپارچه دارد. افزون بر این، بسیاری از مطالعات مشابه در زمینه تحلیل داده‌های اسپرومتری و مدل‌سازی پیش‌بینی‌کننده از پایتون به‌عنوان بستر اصلی استفاده کرده‌اند، بنابراین، انتخاب پایتون در این پژوهش بر اساس توان محاسباتی بالا، پشتیبانی گسترده از الگوریتم‌های یادگیری ماشین، سهولت پیاده‌سازی و استاندارد بودن آن در پژوهش‌های داده‌محور پزشکی انجام شده است.

در این پژوهش از ۵ الگوریتم مرتبط با یادگیری ماشین استفاده شد که در ادامه توضیح داده می‌شود.

#### ۱- «ماشین بردار پشتیبان»

یکی از روش‌های پیشرفته و قوی آماری در زمینه تشخیص و طبقه‌بندی صحیح پیامدها بر اساس عوامل خطر مرتبط که اخیراً کاربرد فراوانی در علوم پزشکی پیدا کرده است، روش ماشین بردار پشتیبان است (۲۳). فرایند ساخت مدل شامل دو مرحله آموزش و آزمون است که در مرحله آموزش، مدل الگو و روابط موجود بین پیامد و پیشین‌ها را یاد می‌گیرد و در انتهای مرحله آموزش، قابلیت تعمیم‌دهی مدل برازش داده شده توسط داده‌های آزمون مورد ارزیابی قرار می‌گیرد. ماشین بردار پشتیبان پیش‌بینی‌های خود را با استفاده از ترکیبات خطی و غیرخطی بر روی مجموعه‌ای از داده‌های آموزش به نام بردارهای پشتیبان انجام می‌دهد. باتوجه به مزایایی

#### 2 Support Vector Machines - SVMs

بر اساس پژوهش‌های پیشین، رابطه بین شاخص‌های اسپرومتری در تست‌های اسپرومتری مانند FEV<sub>1</sub> و FVC با ویژگی‌های فردی مانند سن و قد به‌عنوان متغیر مستقل، غیرخطی است که استفاده از شیوه‌های آماری خاص خود را طلب می‌کند (۱۶). باین‌حال، اکثر مطالعه‌های قبلی در ایران از رگرسیون خطی چندگانه برای ارائه معادلات مرجع اسپرومتری استفاده کرده‌اند (۱۹، ۱۸، ۱۷ و ۲۰). روش‌های یادگیری ماشین در تجزیه و تحلیل داده‌های ریوی برای پیش‌بینی و برآورد وضعیت عملکرد ریوی و طبقه‌بندی بهتر بیماری دارای قدرت توضیح‌دهندگی بهتری از روش‌های سنتی است (۲۱). در این راستا، پیشرفت‌های اخیر در حوزه هوش مصنوعی مسیرهای نوینی را در مراقبت‌های بهداشتی گشوده است و امکان استخراج بینش‌های عمیق‌تری از داده‌های موجود را فراهم می‌کند؛ از جمله شناسایی الگوهایی که به‌سادگی از طریق تفسیر انسانی یا مدل‌سازی آماری سنتی قابل تشخیص نیست. در این میان، روش‌های یادگیری ماشین به‌عنوان یکی از کارآمدترین رویکردها در این زمینه شناخته می‌شود و در مطالعه‌های مختلف مرتبط با ارزیابی عملکرد ریوی نیز مورد استفاده قرار گرفته است (۲۱، ۱۶ و ۲۲).

تحلیل داده‌ها در این پژوهش با استفاده از نرم‌افزار «پایتون» انجام شد، زیرا این زبان یکی از قدرتمندترین و پرکاربردترین ابزارها در حوزه یادگیری ماشین، تحلیل آماری و علم داده است. کتابخانه‌های متعددی مانند TensorFlow و Scikit-learn, Pandas, NumPy, Matplotlib در پایتون، امکان پیاده‌سازی دقیق، مقایسه و ارزیابی مدل‌های یادگیری

#### 1 Python

با الگوریتم های سنتی، جنگل تصادفی دارای ویژگی های برجسته ای است که باعث گسترده شدن دامنه کاربرد آن شده است. از جمله این ویژگی ها می توان به دقت بالا در پیش بینی، کاهش احتمال «بیش برآزش»<sup>۵</sup>، توانایی مدیریت داده های بزرگ و پر بعد، قابلیت ارزیابی اهمیت ویژگی ها، انعطاف پذیری در کاربردهای مختلف و مقاومت نسبت به داده های آلوده به خطا اشاره کرد (۲۷).

### پیش پردازش داده ها و آموزش مدل های یادگیری ماشین

«تنظیمات اولیه»<sup>۱</sup> مدل ها بر اساس مقادیر پیش فرض کتابخانه Scikit-learn و با انجام آزمون مقادیر متداول صورت گرفت. برای هر مدل یادگیری ماشین، هایپر پارامترهای اصلی و مقادیر بررسی شده به شرح زیر بودند:

**رگرسیون خطی:** از مقادیر پیش فرض استفاده شد و تنظیم خاصی نیاز نبود.

**الگوریتم کی - نزدیک ترین همسایه:** تعداد همسایگان K در بازه ۳ تا ۱۵ مورد آزمون قرار گرفت و بهترین مقدار با استفاده از کمترین خطای اعتبارسنجی متقابل انتخاب شد.

**ماشین بردار پشتیبان:** از «کرنل شعاعی»<sup>۲</sup> استفاده شد و پارامترهای C و  $\gamma$  با روش جستجوی شبکه ای در مقادیر متداول [۱,۰,۱,۰,۱] و [۱,۰,۱,۰,۱] بهینه شدند.

**جنگل تصادفی:** تعداد درخت ها بین ۱۰۰ تا ۳۰۰ و بیشینه عمق درخت بین ۵ تا ۳۰ بررسی شد و بهترین ترکیب بر اساس معیار RMSE روی مجموعه اعتبارسنجی انتخاب شد.

**تقویت گرادیان:** نرخ یادگیری در بازه ۰/۰۱ تا ۰/۲ و تعداد تخمین گر ها بین ۱۰۰ تا ۳۰۰ مورد آزمون قرار گرفت. انتخاب بهترین مقدار بر اساس کمترین خطای RMSE روی مجموعه اعتبارسنجی انجام شد.

برای شناسایی نقاط پرت، ابتدا با استفاده از روش «محدوده بین چارکی»<sup>۸</sup> مقادیر خارج از محدوده  $Q1-1.5 \times IQR, Q3+1.5 \times IQR$  به عنوان نقاط پرت شناسایی شدند. بسته به شدت و نوع نقاط پرت، تصمیم به حذف آنها یا اعمال تبدیل داده گرفته شد؛ به عنوان مثال، در متغیرهایی با

که این روش نسبت به روش های کلاسیک آماری دارد و نیاز به پیش فرض خاصی در مورد داده ها نیست، شناسایی الگوی ارتباط از روی داده های گروه آموزش صورت می گیرد و از روی الگوی تشخیص داده شده توسط مدل، داده های جدید ارزیابی می شود. عدم اهمیت به ساختار داده ها و رده ها، قابلیت استفاده برای داده ها با ابعاد بالاتر و استفاده از اصول ریسک ساختاری، به منظور کاهش خطای مدل از مزایای دیگر این روش است (۲۴).

### ۲- «تقویت گرادیان»<sup>۱</sup>

تقویت گرادیان یک الگوریتم یادگیری ماشین قدرتمند و کارآمد است که در هر دو حوزه رگرسیون و طبقه بندی کاربرد دارد. این روش به ویژه برای مجموعه داده های بزرگ و پیچیده مناسب است، زیرا با ترکیب تدریجی مدل های ضعیف تر، کارایی مدل نهایی را بهبود بخشیده و درعین حال مصرف حافظه را بهینه می کند (۲۵).

### ۳- «الگوریتم کی - نزدیک ترین همسایه»<sup>۲</sup>

نزدیک ترین همسایه یکی از روش های ساده و درعین حال مؤثر در حوزه یادگیری ماشین است که بر پایه شباهت میان نمونه ها عمل می کند. در این روش، طبقه بندی یا پیش بینی بر اساس نزدیک ترین نمونه ها در فضای ویژگی انجام می شود؛ به طوری که موارد مشابه در مجاورت یکدیگر قرار گرفته و موارد غیرمشابه در فواصل دورتری از هم واقع می شود (۲۶).

### ۴- «رگرسیون خطی»<sup>۳</sup>

رگرسیون خطی یک روش پایه و پرکاربرد در یادگیری ماشین و آمار است که برای مدل سازی رابطه خطی بین یک متغیر وابسته یا پاسخ و یک یا چند متغیر مستقل یا پیش بین استفاده می شود. هدف از رگرسیون خطی، یافتن بهترین خط (در حالت یک متغیر مستقل) یا بهترین ابر صفحه (در حالت چند متغیر مستقل) است که بتواند رابطه بین این متغیرها را توصیف کند و مقادیر متغیر وابسته را بر اساس مقادیر متغیرهای مستقل پیش بینی کند.

### ۵- «جنگل تصادفی»<sup>۴</sup>

جنگل تصادفی یک الگوریتم یادگیری ماشین و یک روش ترکیبی بر پایه مجموعه ای از درختان تصمیم است. این الگوریتم به صورت گسترده در مسائل طبقه بندی، پیش بینی و حتی رگرسیون استفاده می شود. در مقایسه

5 Overfitting

6 Hyperparameters

7 RBF

8 IQR

1 Gradient Boosting

2 K-nearest Neighbors Algorithm

3 Linear Regression

4 Random Forest

گزارش شد.

در ادامه، برای استانداردسازی فرایند ارزیابی و مقایسه، یک تابع ارزیابی تعریف شد تا محاسبه تمامی معیارهای عملکرد مدل در یک روال واحد انجام گیرد. سپس، پنج مدل یادگیری ماشین شامل رگرسیون خطی، الگوریتم کی - نزدیکترین همسایه، تقویت گرادیان، جنگل تصادفی و ماشین بردار پشتیبان، با استفاده از داده‌های آموزش برازش داده شده و عملکرد آن‌ها بر روی مجموعه تست با سه معیار اصلی رگرسیون مورد ارزیابی قرار گرفت. این سه معیار عبارت است از «میانگین مربعات خطا»، «ریشه میانگین مربعات خطا» و «ضریب تعیین<sup>۵</sup>». در معیارهای میانگین مربعات خطا و ریشه میانگین مربعات خطا، مقادیر کمتر (نزدیک به صفر) نشان‌دهنده خطای کمتر و عملکرد بهتر مدل است، درحالی‌که در معیار ضریب تعیین، مقادیر بالاتر (نزدیک به یک) نشان‌دهنده قدرت توضیحی بهتر مدل است.

#### یافته‌ها

از ۵۴۵۰ جمعیت مورد مطالعه، ۲۴۹۸ نفر مرد و ۲۹۵۲ نفر زن بوده‌اند. همچنین، ۲۲۶ نفر استعمال کننده دخانیات و ۵۲۲۴ نفر بدون استعمال دخانیات بوده‌اند. افزون بر آن، میانگین سن  $۲۱/۱۸ \pm ۴۷/۶۳$ ، میانگین قد  $۱۳/۷۶ \pm ۱۶۲/۱۸$  و میانگین وزن  $۱۸/۹۲ \pm ۶۷/۰۱$  بوده است. در ادامه تابع `df.corr` در کتابخانه `Pandas` برای محاسبه ضریب همبستگی پیرسون بین ستون‌های عددی استفاده شد. ضریب پیرسون عددی بین -۱ تا ۱ است که شکل ۱ میزان و جهت رابطه خطی بین متغیرها را نشان می‌دهد.

توزیع نرمال نقاط پرت شدید حذف شدند و در متغیرهای دارای توزیع نامتقارن از تبدیل لگاریتمی یا ریشه دوم برای کاهش اثر آن‌ها استفاده شد. همچنین، لازم است داده‌های عددی مقیاس‌بندی شود تا تأثیر متغیرهایی با دامنه‌های بزرگ‌تر بر نتایج مدل کاهش یابد. برای این منظور، از روش استانداردسازی استفاده شد که هر مقدار  $x$  از هر متغیر عددی به صورت زیر تبدیل شد:

$$\frac{x - \mu}{\sigma} = std^x$$

که در آن:

$\mu$  میانگین و  $\sigma$  انحراف معیار همان متغیر در مجموعه آموزش است. با این تبدیل، تمامی متغیرها دارای میانگین صفر و انحراف معیار یک شدند و آموزش مدل‌ها از تأثیر دامنه‌های مختلف متغیرها مستقل شد.

پس از آماده‌سازی داده‌ها، مجموعه داده به دو بخش اصلی تقسیم شد: مجموعه آموزش که ۸۰ درصد داده‌ها برای آموزش مدل‌های یادگیری ماشین اختصاص می‌یابد. مدل‌ها با استفاده از این مجموعه داده، الگوها و روابط بین ویژگی‌های فردی و شاخص‌های اسپیرومتري را یاد می‌گیرند. مجموعه تست که ۲۰ درصد باقیمانده داده‌ها برای ارزیابی عملکرد نهایی مدل آموزش دیده استفاده می‌شود. این مجموعه داده هرگز در مرحله آموزش استفاده نمی‌شود و برای ارزیابی توانایی مدل در تعمیم به داده‌های جدید و دیده نشده به کار می‌رود.

برای اطمینان از نماینده بودن هر دو مجموعه (آموزش و تست)، از روش تقسیم تصادفی استفاده شد، تا اطمینان حاصل شود که مجموعه آموزش و تست هر دو نماینده داده‌های کل جمعیت باشند، به خصوص اگر توزیع متغیر وابسته نامتقارن باشد. پس از توسعه مدل‌ها، می‌توان روش‌های انتخاب ویژگی را برای شناسایی مهم‌ترین ویژگی‌های فردی که بیشترین تأثیر را بر شاخص‌های اسپیرومتري دارند، به کار برد. این امر موجب بهبود دقت مدل می‌شود. روش‌هایی مانند `RFE` از طریق مدل‌هایی مانند جنگل تصادفی یا تقویت گرادیان انجام می‌شود.

برای جلوگیری از وابستگی عملکرد مدل‌ها به یک تقسیم‌بندی خاص از داده‌ها و افزایش پایداری ارزیابی، از روش «اعتبارسنجی متقابل ۵-لایه‌ای<sup>۶</sup>» استفاده شد. در این روش کل داده‌ها به ۵ بخش تقسیم شد و مدل پنج بار آموزش دید؛ هر بار چهار بخش برای آموزش و یک بخش برای اعتبارسنجی استفاده شد. میانگین امتیازها نیز به عنوان عملکرد نهایی مدل‌ها

1 Recursive Feature Elimination

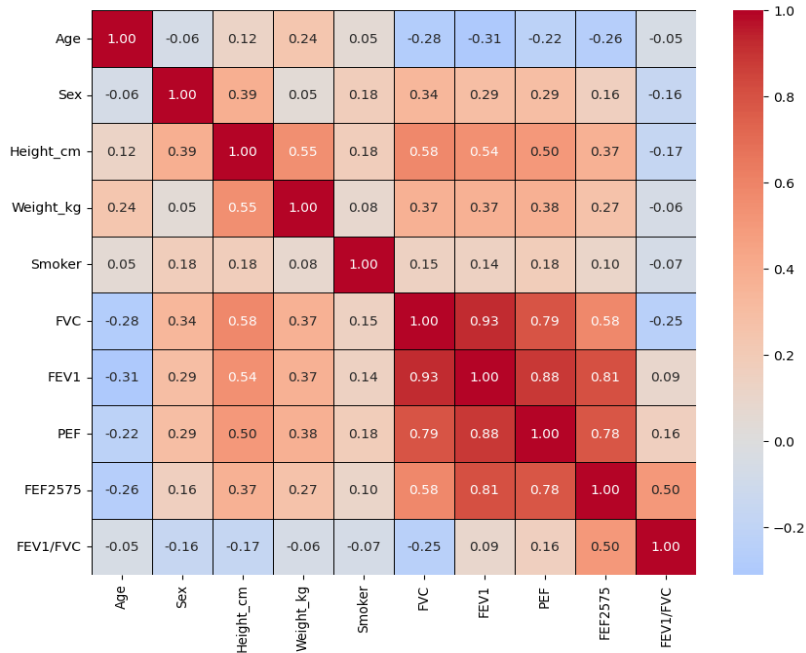
2 5-Fold Cross-Validation

3 MSE

4 RMSE

5 R2-Score

## Correlation Matrix



شکل ۱- ماتریس همبستگی پیرسون بین ویژگی های فردی و شاخص های اسپرومتری

روابط فیزیولوژیک پایه را مخدوش می سازند. با حذف ۸۳۵ رکورد که در شاخص های اسپرومتری یا ویژگی های فردی به عنوان پرت شناسایی شد، حجم نهایی نمونه مورد مطالعه به ۶۶۱۵ رکورد کاهش یافت. اگرچه وجود نقاط پرت ممکن است نمایانگر موارد بالینی خاص باشد، اما هدف این مطالعه مدل سازی روابط فیزیولوژیک در جمعیت عمومی است. داده های بسیار دور از الگوی فیزیولوژیک می تواند الگوریتم ها را دچار بیش برآزش کند و به همین دلیل حذف آن ها برای افزایش کیفیت مدل، منطقی است. مقایسه اولیه نتایج مدل ها قبل و بعد از حذف نقاط پرت نشان داد که انجام این مرحله موجب کاهش محسوس میزان خطای پیش بینی و افزایش ثبات عملکرد مدل ها شده است؛ بنابراین حذف داده های پرت نه تنها باعث کاهش حجم نمونه مؤثر نشد، بلکه کیفیت، قابلیت تعمیم و اعتبار نتایج نهایی را به طور معنی داری افزایش داد. نسبت نمونه حذف شده نیز کمتر از ۲۰ درصد بوده و همچنان حجم نمونه نهایی برای تحلیل های یادگیری ماشین کاملاً مناسب و قابل اتکا باقی مانده است.

همچنین، برای آماده سازی متغیرهای طبقه ای (جنسیت و وضعیت استعمال دخانیات) که به صورت باینری (۰ و ۱) کدگذاری شده اند، اطمینان حاصل شد که این متغیرها برای استفاده در مدل های رگرسیون

تحلیل ماتریس همبستگی پیرسون نشان می دهد سن با  $FEV_1$  و  $FVC$  همبستگی منفی متوسط دارد. این یافته با تغییرات فیزیولوژیک ناشی از پیری سازگار است؛ به طوری که با افزایش سن، الاستیسیته بافت ریه کاهش یافته و حجم های ریوی افت می کند. همبستگی مثبت قوی میان قد و  $FVC/FEV_1$  نیز نشان می دهد که حجم ریوی رابطه ساختاری با ابعاد قفسه سینه دارد. در مقابل، وزن و وضعیت سیگاری بودن همبستگی خطی ضعیف نشان دادند؛ این امر بیانگر آن است که اثر این دو متغیر ممکن است غیرخطی یا وابسته به تعامل با سایر متغیرها باشد؛ مسئله ای که مدل های درختی توانایی بیشتری در کشف آن دارند. این یافته، توجیه کننده استفاده از مدل های غیرخطی در این پژوهش است.

در مرحله پیش پردازش، با هدف تضمین کیفیت داده ها و کاهش تأثیر نویز ناشی از مقادیر غیرعادی، از روش محدوده بین چارکی برای شناسایی و حذف نقاط پرت استفاده شد. این اقدام بر اساس ضرورت تمرکز مدل ها بر الگوهای فیزیولوژیک مرتبط با ویژگی های فردی در جمعیت عمومی و حذف موارد شدید بالینی که خارج از هدف پیش بینی ویژگی های فردی است صورت گرفت، زیرا برون یابها می توانند عمدتاً ناشی از خطای اندازه گیری، دمش غیراستاندارد یا بیماری های حاد باشند که مدل سازی

متمركزسازی فرایند، یک «تابع پوششی<sup>۲</sup>» به نام evaluate تعریف شد تا وظیفه تولید پیش‌بینی‌ها و محاسبه کلیه معیارهای ارزیابی عملکرد مدل را برای هر الگوریتم به‌صورت خودکار اجرا کند. سپس ۵ مدل ذکر شده با ۳ روش مورد اشاره ارزیابی می‌شود. جدول ۲ مقایسه عملکرد مدل‌های یادگیری ماشین بر اساس معیار میانگین مربعات خطا را نشان می‌دهد که هر چه کمتر باشد بهتر است.

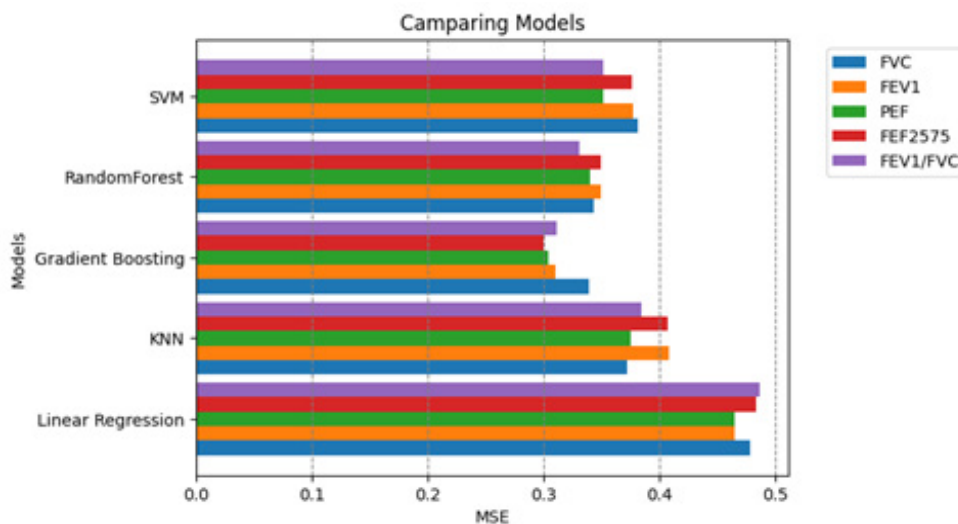
خطی و ماشین بردار پشتیبان به‌درستی به‌عنوان متغیرهای مجازی تفسیر شود، تا از تحمیل یک رابطه خطی غیرمنطقی جلوگیری شود. سپس مجموعه‌داده نهایی به نسبت ۸۰ درصد (۳۶۹۲ رکورد) به داده آموزش و ۲۰ درصد (۹۲۳ رکورد) به داده تست اختصاص داده شد. این تقسیم‌بندی نیز با استفاده از روش تقسیم تصادفی انجام شد. به‌منظور حاکمیت نظم متدولوژیک، جلوگیری از «تکرار کد» و

جدول ۲- مقایسه عملکرد مدل‌های یادگیری ماشین بر اساس معیار میانگین مربعات خطا

FEV1/FVC	FEF25-75	PEF	FEV <sub>1</sub>	FVC	عوامل
۰/۴۸۷۱	۰/۴۸۳۶	۰/۴۶۴۹	۰/۴۶۴۷	۰/۴۷۸۳	رگرسیون خطی
۰/۳۸۴۱	۰/۴۰۷۴	۰/۳۷۵۷	۰/۴۰۷۸	۰/۳۷۲۱	الگوریتم کی - نزدیک‌ترین همسایه
۰/۳۱۱۹	۰/۳۰۰۱	۰/۳۰۴۱	۰/۳۱۰۳	۰/۳۳۹۱	تقویت گرادیان
۰/۳۳۱۱	۰/۳۴۹۶	۰/۳۳۹۸	۰/۳۴۹۶	۰/۳۴۳۱	جنگل تصادفی
۰/۳۵۱۵	۰/۳۷۶۲	۰/۳۵۱۶	۰/۳۷۶۹	۰/۳۸۱۲	ماشین بردار پشتیبان

تحلیل نمودار مقایسه مدل‌ها (شکل ۲) حاکی از آن است که مدل تقویت گرادیان در معیار میانگین مربعات خطا، کمترین میزان خطا را داشته و رگرسیون خطی، بیشترین میزان خطا را نشان داده است.

شکل ۲- نمایش نموداری مقایسه عملکرد مدل‌ها بر اساس معیار میانگین مربعات خطا



## 2. Wrapper Function

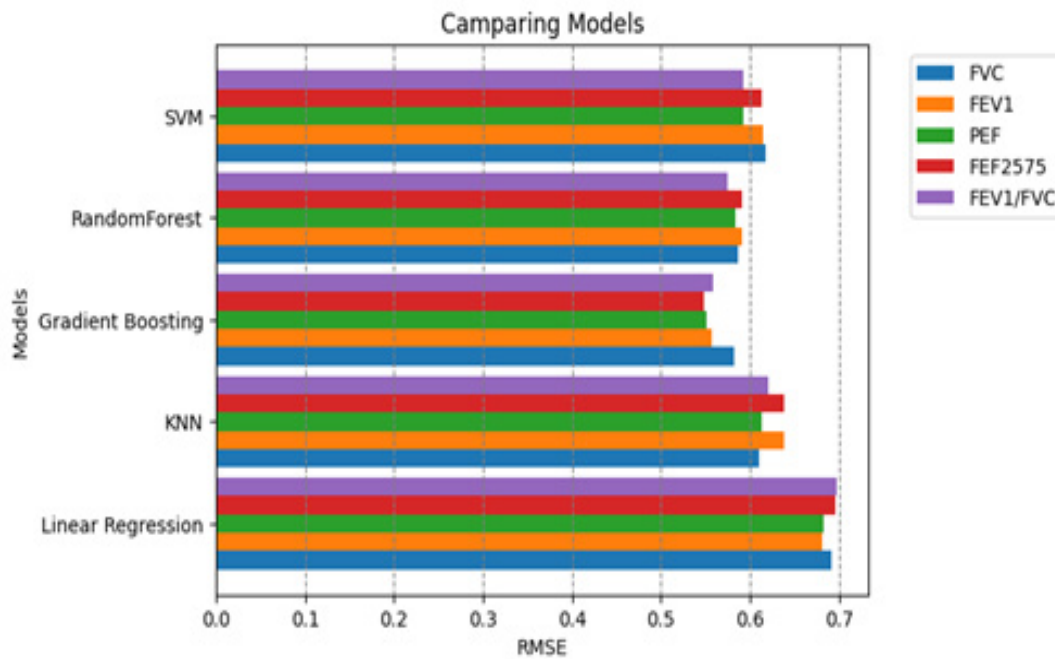
## 1. Code Redundancy

جدول ۳ مقایسه عملکرد مدل های یادگیری ماشین بر اساس معیار ریشه میانگین مربعات خطا را نشان می دهد.

جدول ۳- مقایسه عملکرد مدل های یادگیری ماشین بر اساس معیار ریشه میانگین مربعات خطا

عوامل	FVC	FEV1	PEF	FEF25-75	FEV1/FVC
رگرسیون خطی	۰/۶۹۱۶	۰/۶۸۱۷	۰/۶۸۱۹	۰/۶۹۵۴	۰/۶۹۷۹
الگوریتم کی - نزدیک ترین همسایه	۶۱۰/۱۰	۰/۶۳۸۶	۰/۶۱۳۱	۰/۶۳۸۲	۰/۶۱۹۷
تقویت گرادیان	۵۸۲۳/۰	۰/۵۵۷۱	۰/۵۵۱۴	۰/۵۴۷۸	۰/۵۵۸۴
جنگل تصادفی	۰/۵۸۵۸	۰/۵۹۱۳	۰/۵۸۲۹	۰/۵۹۱۲	۰/۵۷۵۴
ماشین بردار پشتیبان	۰/۶۱۷۴	۰/۶۱۳۹	۰/۵۹۳۰	۰/۶۱۳۳	۰/۵۹۲۹

شکل ۳ نمایش نموداری مقایسه عملکرد مدل ها بر اساس معیار ریشه میانگین مربعات خطا را نشان می دهد.



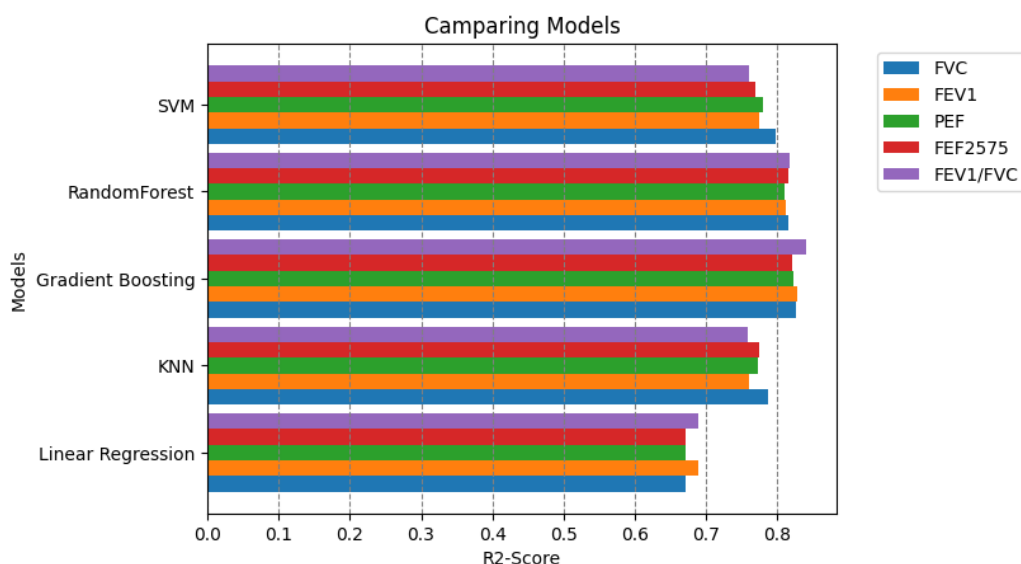
شکل ۳- نمایش نموداری مقایسه عملکرد مدل ها بر اساس معیار ریشه میانگین مربعات خطا

جدول ۴، مقایسه عملکرد مدل‌های یادگیری ماشین بر اساس معیار ضریب تعیین را نشان می‌دهد و هر چه قدر به عدد ۱ نزدیک تر باشد بهتر است.

جدول ۴- مقایسه عملکرد مدل‌های یادگیری ماشین بر اساس معیار ضریب تعیین

عوامل	FVC	FEV <sub>1</sub>	PEF	FEF <sub>25-75</sub>	FEV <sub>1</sub> /FVC
رگرسیون خطی	۰/۶۷۱۶	۰/۶۸۸۳	۰/۶۷۱۳	۰/۶۷۲۱	۰/۷۹۸۶
الگوریتم کی - نزدیک‌ترین همسایه	۰/۷۸۷۵	۰/۷۵۹۶	۰/۷۷۲۵	۰/۷۷۴۸	۰/۲۹۵۷
تقویت گرادیان	۰/۸۲۶۹	۰/۸۲۸۱	۰/۸۲۳۱	۰/۸۲۰۶	۰/۱۱۴۸
جنگل تصادفی	۰/۸۱۴۷	۰/۸۱۲۵	۰/۸۰۹۷	۰/۸۱۴۷	۰/۸۱۸۱
ماشین بردار پشتیبان	۰/۷۹۷۲	۰/۷۷۳۸	۰/۷۸۰۱	۰/۷۶۸۴	۰/۷۶۰۴

شکل ۴، نمایش نموداری مقایسه عملکرد مدل‌ها بر اساس معیار ضریب تعیین را نشان می‌دهد.



شکل ۴- نمایش نموداری مقایسه عملکرد مدل‌ها بر اساس معیار ضریب تعیین

گرادیان توانایی بالایی در یادگیری این الگوهای پیچیده دارد. مدل جنگل تصادفی و ماشین بردار پشتیبان نیز عملکرد قابل قبولی ارائه کرد، ولی همچنان نسبت به تقویت گرادیان دقت پایین تری داشت. به طور میانگین، ریشه میانگین مربعات خطا مدل تقویت گرادیان نسبت به رگرسیون خطی بین ۱۸ تا ۲۵ درصد کمتر بود که نشان‌دهنده دقت به مراتب بالاتر آن در برآزش روابط میان متغیرها است. این اختلاف عملکرد بیانگر آن است که

نتایج ارزیابی مقایسه‌ای پنج مدل یادگیری ماشین بر اساس سه معیار ذکر شده، برتری قاطع مدل تقویت گرادیان را در پیش‌بینی تمامی شاخص‌های اسپیرومتری تأیید کرد. این مدل در تمام شاخص‌ها کمترین مقدار میانگین مربعات خطا و ریشه میانگین مربعات خطا و بالاترین مقدار ضریب تعیین را به دست آورد. مدل‌های خطی به دلیل فرض ساده خطی قادر به استخراج این الگوها نیست، درحالی‌که مدل‌های درختی و به‌ویژه تقویت

این عوامل نقش کلیدی در حجم و عملکرد ریوی دارد. در مقابل، جنسیت و وضعیت سیگاری بودن تأثیر کمتری نشان می دهد که ممکن است به دلیل مقیاس بندی باینری یا اثرات غیرخطی و تعامل با سایر ویژگی ها باشد.

### بحث

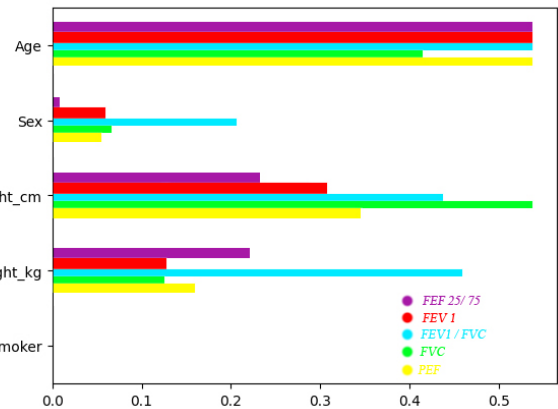
هدف این پژوهش، ارزیابی و مقایسه عملکرد مدل های یادگیری ماشین در پیش بینی شاخص های اسپیرومتری بر اساس ویژگی های فردی در جمعیت شهرستان فسا بود، امری که می تواند دانش موجود در زمینه «معادلات مرجعی محلی» و «پیش بینی عملکرد ریوی بدون تست مستقیم» را گسترش دهد. یافته های اصلی نشان داد که مدل تقویت گرادیان در مقایسه با سایر الگوریتم ها دقت پیش بینی بالاتری دارد. این مدل توانست روابط پیچیده، احتمالاً غیرخطی و تعاملی بین ویژگی های فردی و شاخص های اسپیرومتری را کشف کند. خروجی `feature_importances` نیز نشان داد که سن، قد و وزن سه متغیر با تأثیر قابل توجه و عمده بر پیش بینی شاخص ها است، درحالی که جنسیت و وضعیت سیگاری بودن تأثیر به نسبت کمتری داشت.

افزایش سن با کاهش قدرت الاستیک ریه، کاهش حجم ریوی و تغییرات ساختاری قفسه سینه همسو است؛ موضوعی که با فیزیولوژی ریه سازگار است. قد به عنوان یک شاخص ساختاری بدن بر حجم ریه تأثیر دارد: افراد بلندتر معمولاً ریه بزرگتری دارند و بنابراین FVC و FEV<sub>1</sub> بالاتری دارند. وزن ممکن است با محدودیت مکانیکی قفسه سینه و دیافراگم، به ویژه در افراد با چاقی یا اضافه وزن، بر عملکرد ریوی تأثیر بگذارد؛ یا اینکه وزن نماینده ترکیبی از ترکیب بدنی، چربی و توده عضلانی باشد که بر ریتم تنفس و ظرفیت ریوی اثرگذار است. اثر کم تر «جنسیت» و «وضعیت سیگاری بودن» در مدل می تواند به مقیاس بندی باینری، تعداد کم افراد سیگاری، یا تعامل با سایر متغیرها بستگی داشته باشد؛ یعنی شاید در این جمعیت سیگاری ها کمتر یا با ویژگی متفاوتی باشند؛ یا تأثیر سیگار در داده ما با متغیرهای دیگر خنثی شده باشد.

نتایج ما در اهمیت «سن» و «قد» با بسیاری از مطالعات اسپیرومتری همخوانی دارد. به عنوان نمونه، در مطالعه ای با جمعیت ایرانی، سن و قد به عنوان اصلی ترین پیش بین های FEV<sub>1</sub> و FVC گزارش شده است (۲۸). همچنین در گروه های جوان تر و حتی کودکان در پژوهشی از پاکستان، وزن نیز رابطه معناداری با شاخص های ریوی مانند FVC و FEV<sub>1</sub> داشته است (۲۹)، اما در برخی مطالعات، وزن یا BMI تأثیر منفی یا ناچیز بر عملکرد

روابط میان ویژگی های فردی و شاخص های اسپیرومتری ماهیتی غیرخطی داشته و شامل تعاملات چندمتغیره مانند اثر هم زمان سن، قد و جنس است؛ روابطی که مدل های خطی قادر به استخراج آن ها نیست. در مقابل، مدل های درختی مانند جنگل تصادفی و به ویژه تقویت گرادیان، با ساختار مرحله ای خود توانایی بیشتری در شناسایی این الگوهای پیچیده دارند. مدل ماشین بردار پشتیبان نیز در رتبه بعدی عملکرد قرار گرفت و به دلیل استفاده از «کرنل ۱» قادر به مدل سازی الگوهای غیرخطی است.

خروجی `feature_importances` مدل تقویت گرادیان که اهمیت نسبی هر یک از ویژگی های فردی را در پیش بینی شاخص های اسپیرومتری نشان می دهد، در شکل ۵ ارائه شده است.



شکل ۵- خروجی `feature_importances` مدل تقویت گرادیان برای

هر یک از شاخص های اسپیرومتری

این یافته ها اهمیت نسبی هر یک از ویژگی های فردی سن، جنسیت، قد، وزن و وضعیت سیگاری بودن را در پیش بینی شاخص های اسپیرومتری با استفاده از مدل تقویت گرادیان نشان می دهد. اعداد بالاتر نشان دهنده تأثیر بیشتر آن ویژگی بر پیش بینی شاخص مورد نظر است. همانگونه که مشاهده می شود، ویژگی های سن، قد و وزن در اکثر شاخص ها بیشترین اهمیت را دارد، درحالی که جنسیت و وضعیت سیگاری بودن تأثیر کمتری دارند. این یافته ها نشان می دهد که مدل تقویت گرادیان توانسته است روابط غیرخطی و پیچیده بین ویژگی های فردی و شاخص های اسپیرومتری را شناسایی کند. اهمیت بالای سن، قد و وزن با دانش فیزیولوژی ریه همخوانی دارد، زیرا

### 1 Kernel

کرنل در ماشین بردار پشتیبان تابعی است که داده ها را به فضای بالاتر می برد تا الگوهای غیرخطی قابل تشخیص شوند، بدون آن که این تبدیل به صورت واقعی محاسبه شود.

مرکزی، عوامل محیطی از جمله میزان آلودگی هوا و شرایط اقلیمی را نیز در مدل‌های خود وارد کنند. همچنین ترکیب داده‌های جمعیت‌شناختی با داده‌های بالینی، تصویربرداری و بیومارکرها می‌تواند به ساخت مدل‌هایی دقیق‌تر و کاربردی‌تر منجر شود.

### نتیجه‌گیری

پژوهش حاضر نشان داد که مدل تقویت‌گرایان به‌عنوان یک الگوریتم یادگیری ماشین مؤثر، می‌تواند با دقت مناسب شاخص‌های کلیدی اسپیرومتری را بر اساس ویژگی‌های فردی در جمعیت محلی پیش‌بینی کند. اهمیت سن، قد و وزن در مدل، تأکیدی بر نقش ساختاری و فیزیولوژیک این عوامل در عملکرد ریوی است، امری که با یافته‌های مطالعات مرجع همسو است. در عین حال، نتایج این مطالعه چالش‌ها و تفاوت‌هایی را نسبت به بعضی مطالعه‌های پیشین نشان می‌دهد که می‌تواند به ویژگی‌های جمعیت، روش ثبت داده، مقیاس‌بندی متغیرها و نیز استفاده از مدل‌های پیشرفته نسبت داده شود.

### اعلان‌ها

**تشکر و قدردانی:** از همه افرادی که در تهیه این پژوهش مشارکت داشته‌اند، تشکر می‌شود.

**حمایت مالی:** این پژوهش بدون هرگونه حمایت مالی انجام شده است.  
**تعارض منافع:** نویسندگان گواهی می‌دهند که در این مطالعه تعارض منافی نداشته‌اند.

**ملاحظات اخلاقی:** در این پژوهش کلیه اصول اخلاقی رعایت شد و کد اخلاق IR.FUMS.REC.1404.174 نیز اخذ شده است.

**استفاده از هوش مصنوعی:** نویسندگان اعلام می‌دارند که در فرایند نگارش این مقاله از هیچ‌گونه فناوریهای هوش مصنوعی برای تولید محتوا، ایده، تجزیه و تحلیل و سایر موارد استفاده نشده است.

**مشارکت نویسندگان:** ساره رفعت مقام (تهیه طرح اولیه پژوهش، تدوین مبانی نظری، گردآوری داده‌های ثانویه و نگارش پیش‌نویس مقاله)، ندا جعفری (بازنگری طرح پژوهش، گردآوری داده‌های اولیه و نظارت علمی بر اجرای پژوهش)، عزیزالله دهقان (همکاری در اصلاح و تحلیل بخش آماری)، محمد روشن ضمیر (نظارت تخصصی بر بخش یادگیری ماشین). کلیه نویسندگان نسخه نهایی مقاله را مطالعه کرده و در تکمیل و تأیید آن مشارکت داشته‌اند.

ریوی با برخی شاخص‌ها داشت. به‌عنوان نمونه، مطالعه روی افراد مبتلا به آسم نشان داد چاقی با کاهش معنی‌دار  $FEV_1$  و  $FVC$  همراه است (۳۰). این تناقض‌ها با نتایج پژوهش ما که وزن اهمیت نسبی بالایی دارد، ممکن است به ویژگی‌های جمعیت مطالعه غیرآسمی. عمومی، تفاوت سبک زندگی، چاقی یا توزیع BMI و نیز تفاوت در مدل آماری و الگوریتم مورد استفاده بازگردد؛ بنابراین، یافته‌های ما مطابق با معادلات مرجع سنتی و استاندارد، نه تنها تأکید بر نقش مهم سن و قد دارد، بلکه وزن را نیز به‌عنوان یک متغیر با تأثیر واقعی در پیش‌بینی شاخص‌های ریوی در جمعیت فسا مطرح می‌کند. این موضوع شاخص است برای اینکه معادلات مرجع باید برای جمعیت محلی با ویژگی‌های جمعیتی، اقلیمی و رفتاری خاص طراحی شود. از جمله محدودیت‌های این پژوهش می‌توان به تک‌مرکزی بودن مطالعه، استفاده از داده‌های ثانویه، ثبت ساده متغیر وضعیت سیگار به‌صورت باینری و تفاوت‌های احتمالی در کالیبراسیون و دقت دستگاه‌های اسپیرومتری اشاره کرد. افزون بر این، محدودیت زمانی پژوهشگر و دسترسی محدود به منابع داده‌ای و مالی، امکان گسترش حجم نمونه و افزودن متغیرهای بیشتر را کاهش داد.

با وجود این محدودیت‌ها، پژوهش حاضر گامی مهم در جهت ادغام روش‌های نوین یادگیری ماشین با سلامت تنفسی در ایران محسوب می‌شود و می‌تواند مبنایی برای پژوهش‌های گسترده‌تر، کاربردی‌تر و سیاست‌گذارانه در آینده باشد.

بر اساس نتایج به‌دست‌آمده، استفاده از مدل‌های یادگیری ماشین، به‌ویژه الگوریتم تقویت‌گرایان، می‌تواند به عنوان ابزاری کارآمد در طراحی معادلات مرجع محلی و حتی به عنوان یک ابزار غربالگری اولیه برای شناسایی افراد در معرض خطر اختلالات ریوی مورد استفاده قرار گیرد. این ابزار می‌تواند در محیط‌های کم‌برخوردار که دسترسی به تجهیزات پیشرفته محدود است، به پزشکان و مراکز بهداشتی کمک کند تا بر اساس ویژگی‌های فردی، تصویر اولیه‌ای از وضعیت عملکرد ریوی افراد به دست آورند. همچنین توصیه می‌شود سیاست‌گذاران سلامت از به‌کارگیری مقادیر مرجع غیربومی پرهیز کرده و به تدوین استانداردهای محلی مبتنی بر داده‌های بومی توجه ویژه‌ای داشته باشند.

به پژوهشگران آینده نیز پیشنهاد می‌شود مطالعات چندمرکزی در مناطق مختلف ایران انجام دهند و متغیرهای بیشتری مانند شاخص دقیق مصرف دخانیات (بسته - سال)، نوع شغل، سطح فعالیت بدنی، شاخص چاقی

## References

- World Health Organization. (2024). Chronic obstructive pulmonary disease (COPD). [https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-\(copd\)](https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd))
- Pellegrino, R., Viegi, G., Brusasco, V., Crapo, R. O., Burgos, F., Casaburi, R., ... Wanger, J. (2005). Interpretative strategies for lung function tests. *European Respiratory Journal*, 26(5), 948–968.
- Miller, M. R., Hankinson, J., Brusasco, V., Burgos, F., Casaburi, R., Coates, A., ... Wanger, J. (2005). Standardisation of spirometry. *European Respiratory Journal*, 26(2), 319–338.
- Knudson, R. J., Slatin, R. C., Lebowitz, M. D., & Burrows, B. (1976). The maximal expiratory flow-volume curve: Normal standards, variability, and effects of age, smoking, and body size. *American Review of Respiratory Disease*, 113(5), 587–600.
- Sharma, G., & Goodwin, J. (2016). Effect of aging physiology on respiratory function and performance. *Comprehensive Physiology*, 6(4), 1603–1631.
- Crapo, R. O., Morris, A. H., & Gardner, R. M. (1981). Reference spirometric values using techniques and equipment that meet ATS recommendations. *American Review of Respiratory Disease*, 123(6), 659–664.
- Zapletal, A., Samanek, M., & Paul, T. (1987). Lung function in children and adolescents: Methods and reference values. *Progress in Respiration Research*, 22, 83–112.
- Quanjer, P. H., Stanojevic, S., Cole, T. J., Baur, X., Hall, G. L., Culver, B. H., ... Stocks, J. (2012). Multi-ethnic reference values for spirometry for the 3–95-year age range: The global lung function 2012 equations. *European Respiratory Journal*, 40(6), 1324–1343.
- Jones, R. L., & Nzekwu, M. M. (2006). The effects of body mass index on lung volumes. *Chest*, 130(3), 827–833.
- McClean, K. M., Kee, F., Young, I. S., & Elborn, J. S. (2008). Obesity and the lung: *Epidemiology. Thorax*, 63(7), 649–654.
- Janoff, A. (1985). Elastase in tissue injury. *Annual Review of Physiology*, 47(1), 545–558.
- Berg, K., & Wright, J. L. (2016). The pathology of chronic obstructive pulmonary disease: Progress in the 20th and 21st centuries. *Archives of Pathology & Laboratory Medicine*, 140(12), 1423–1428.
- Martin, R. R., Lindsay, D., Despas, P., Bruce, D., Leroux, M., Anthonisen, N. R., & Macklem, P. T. (1975). The early detection of airway obstruction. *American Review of Respiratory Disease*, 111(2), 119–125.
- Quanjer, P. H., Tammeling, G. J., Cotes, J. E., Pedersen, O. F., Peslin, R., & Yernault, J. C. (1993). Lung volumes and forced ventilatory flows. *European Respiratory Journal*, 6(Suppl. 16), 5–40.
- Doll, R., & Peto, R. (1976). Mortality in relation to smoking: 20 years' observations on male British doctors. *British Medical Journal*, 2(6012), 1525–1536.
- Loeloe, M. S., Sefidkar, R., Tabatabaei, S. M., Mehrparvar, A. H., & Jambarsang, S. (2025). Machine learning-based spirometry reference values for the Iranian population: A cross-sectional study from the Shahedieh PERSIAN cohort. *Frontiers in Medicine*, 12, 1480931.
- Golshan, M., Nematbakhsh, M., Amra, B., & Crapo, R. O. (2003). Spirometric reference values in a large Middle Eastern population. *European Respiratory Journal*, 22(3), 529–534.
- Razi, E., Mousavi, S. G. A., & Akbari, H. (2005). Spirometric standards for healthy Iranians dwelling in the centre of Iran. *Tanaffos*, 4(4), 19–26.
- Etemadinezhad, S., & Alizadeh, A. (2011). Spirometric reference values for healthy adults in Mazandaran Province, Iran. *Jornal Brasileiro de Pneumologia*, 37(5), 615–620.
- Aloosh, O., Torkashvand, M., Torkashvand, A., &

- Mohammadi, N. (2022). Evaluation of spirometric values in a healthy population referred to spirometry centers in Hamedan City, Iran. *Journal of the Iranian Medical Council*, 5(4), 661–667.
21. Taloba, A. I., & Matoog, R. T. (2025). Detecting respiratory diseases using machine learning-based pattern recognition on spirometry data. *Alexandria Engineering Journal*, 113, 44–59.
22. Helgeson, S. A., Quicksall, Z. S., Johnson, P. W., Lim, K. G., Carter, R. E., & Lee, A. S. (2025). [Article in JMIR AI].
23. Fatahi, M. (2015). Support vector machines: A survey (Master's thesis/Doctoral dissertation). Razi University.
24. Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics & Proteomics*, 15(1), 41–45.
25. Chen, H., Li, X., Feng, Z., Wang, L., Qin, Y., Skibniewski, M. J., et al. (2023). Shield attitude prediction based on Bayesian-LGBM machine learning. *Information Sciences*, 632, 105–129.
26. Shahrabi, J., & Zare, A. (2013). *Data mining with Clementine*. Tehran: Academic Jihad Press.
27. Liu, Y., Wang, Y., & Zhang, J. (2012). New machine learning algorithm: Random forest. In *Proceedings of ICICA 2012* (pp. 246–252).
28. Sahebi, L., Rahimi, B., Shariat, M., et al. (2022). Normal spirometry prediction equations for the Iranian population. *BMC Pulmonary Medicine*, 22, 472. <https://doi.org/10.1186/s12890-022-02273-8>
29. Sadiq, S., Rizvi, N. A., Soleja, F. K., & Abbasi, M. (2019). Factors affecting spirometry reference range in growing children. *Pakistan Journal of Medical Sciences*, 35(6), 1587–1591.
30. Alqarni, A. A., Aldhahir, A. M., Siraj, R. A., Alqahtani, J. S., Alshehri, H. H., Alshamrani, A. M., et al. (2023). Prevalence of overweight and obesity and their impact on spirometry parameters in patients with asthma: A multicentre retrospective study. *Journal of Clinical Medicine*, 12(5), 1843.