

## تشخیص شدت بیماری کبد چرب غیرالکلی بر مبنای هوش مصنوعی تفسیرپذیر

## چکیده

دریافت: ۱۴۰۳/۰۷/۰۵ ویرایش: ۱۴۰۳/۰۷/۱۰ پذیرش: ۱۴۰۳/۰۹/۲۳ آنلاین: ۱۴۰۳/۱۰/۰۱

حامد زمانیان، احمد شالباف\*

گروه فیزیک و مهندسی پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

**زمینه و هدف:** ابتلا به بیماری کبد چرب غیرالکلی، به دلایلی چون چاقی مفرط، دیابت نوع دوم و ناهمگونی‌های هورمونی به‌طور قابل‌توجهی شیوع پیدا کرده است. هدف از این مقاله، ارائه یک الگوریتم کاربردی براساس الگوریتم‌های تفسیرپذیر یادگیری ماشین برای تشخیص افراد مبتلا به چربی کبد غیرالکلی و تعیین شدت و گستردگی بیماری به کمک اطلاعات بالینی می‌باشد.

**روش بررسی:** در این راستا، مجموع ۱۸۱ بیمار مبتلا به درجه‌های مختلف بیماری کبد چرب غیرالکلی، مورد بررسی قرار گرفتند. این مجموعه داده در بازه زمانی بهمن ۱۳۸۹ تا دی سال ۱۳۹۸ در بیمارستان دانشگاه ایپایم با تشخیص ابتلا یا عدم ابتلا به کبد چرب براساس نتیجه ارزیابی پاتولوژیکی نمونه کبدی جمع‌آوری شده است. برای این دادگان، از روش الگوریتم‌های انتخاب ویژگی برای تعیین ویژگی‌های غالب و روش‌های طبقه‌بندی مختلف برای تعیین سطح شدت شاخص‌های شدت کبد چرب استفاده شد. همچنین برای تفسیر بهتر کارکرد روش‌های مورد استفاده، این اطلاعات به کمک الگوریتم‌های هوش مصنوعی تفسیرپذیر، تحلیل و بررسی شد.

**یافته‌ها:** استفاده از روش کلاس‌بندی XGBoost به‌همراه روش انتخاب ویژگی رو به‌جلو، منجر به دقت  $0.69/23 \pm 0.05$  برای شاخص شدت کبد چرب گردید. بررسی این دادگان توسط الگوریتم‌های تفسیرپذیری هوش مصنوعی نشان داد که متغیرهای سن، BMI، HDL، LDL، HbA1c، و GPT به‌ازاء سه کلاس تعیین شده و شاخص‌های GPT، سن، BMI، HDL، HbA1c، LDL، تری‌گلیسرید، و کلسترول در تعیین کارایی مدل کلی، جزو اثرگذارترین متغیرهای بالینی هستند که در طبقه‌بندی شدت بیماری کبد نقش مهمی دارند.

**نتیجه‌گیری:** بررسی‌های این مطالعه به‌خوبی نشان داد که می‌توان تخمین قابل‌قبولی از شدت ابتلا به بیماری کبد چرب به‌دست آورد که همین موضوع می‌تواند در تعیین درجه ابتلا برای اقدامات درمانی بعدی بیماران کمک‌کننده باشد و از تحمیل هزینه‌های درمانی به سیستم درمانی و بیمار جلوگیری نماید.

**کلمات کلیدی:** هوش مصنوعی تفسیرپذیر، دادگان کلینیکی، یادگیری ماشین، کبد چرب غیرالکلی.

\* نویسنده مسئول: تهران، ولنجک، بلوار دانشجو، کوچه کودکیار، دانشگاه علوم پزشکی شهید بهشتی، دانشکده پزشکی، گروه فیزیک و مهندسی پزشکی.  
تلفن: ۰۲۱-۲۲۴۳۹۹۴۱  
E-mail: shalbf@sbumu.ac.ir

## مقدمه

همراه با آسیب سلولی هپاتیسیت، التهاب قابل توجه کبدی همراه است. این بیماری، زیرمجموعه‌ای از کبد چرب غیرالکلی به‌شمار می‌رود که از منظر پاتولوژیکی، با تغییرات التهابی هپاتیتی و بالونی شدن سلول‌ها به همراه و یا بدون فیبروز، روی می‌دهد.<sup>۱</sup> ابتلای به

عارضه کبد چرب استئاتوسیس-هپاتیتی یک فرمت پاتولوژیکی کبد چرب غیرالکلی است که با شواهدی از سطوح بالای استئاتوسیس

ورودی براساس یک آموزش اولیه است. این مدل سعی می‌کند تا کورولیشن اطلاعات بالینی با برچسب‌های تعریف شده را به دست آورده و ریسک پیشرفت بیماری‌های مربوطه استخراج گردد.<sup>۹</sup> گروه Han Ma از پارامترهای بالینی و خونی برای تعیین ابتلا استفاده نمود که در نهایت الگوریتم رگرسیون لجستیک بالاترین دقت را در کلاس‌بندی مورد نظر نشان داد.<sup>۱۰</sup> گروه تحقیقاتی Wu در همین راستا با بهره‌گیری از الگوریتم‌های یادگیری هوشمند، براساس داده‌های کلینیکی به تخمین کبد چرب پرداخت.<sup>۱۱</sup> همچنین در مطالعه دیگری، برای بررسی و تمیز دادن بیماران مبتلا به چربی ساده و پیشرفته، از ۱۱ داده کلینیکی استفاده شده و یک شبکه عصبی برای تعیین کلاس‌بندی بین سطوح بیماری مورد استفاده قرار گرفت.<sup>۱۲</sup> در یک مطالعه دیگر، گروه کانبی از ویژگی‌های بالینی یک گروه تحقیقی شامل ۱۶۴ بیمار مختلف با چاقی مفرط برای تعیین شدت ابتلا به بیماری استفاده نمود.<sup>۱۳</sup> شاخص‌های مهمی مثل سن، اندیس حجمی بدن، گلوتامیل ترانسفراز گاما، هموگلوبین A1c، آلانین و غیره در تشخیص بیماری نقش بسزایی را نشان دادند. تیم دوچرتی از اطلاعات بالینی حدود ۷۰۴ بیمار داوطلب برای تخمین بهترین مدل ممکن استفاده نمود.<sup>۱۴</sup> در این کار، ۱۴ ویژگی برتر معرفی شده به کمک الگوریتم انتخاب ویژگی از روش حذف بازگشتی، در کنار الگوریتم‌های کلاس‌بندی مختلف مورد سنجش قرار گرفت. اما اکثر این مطالعات، به صورت دوکلاسه و بدون برآورد درستی از منطق بکارگیری هریک از شاخص‌ها صورت گرفته است و محدودیت‌ها در بکارگیری چند متغیر رایج، فاحش است. براین اساس، در این مطالعه، پس از معرفی بانک داده مورد استفاده و الگوریتم‌های یادگیری ماشین و هوش مصنوعی تفسیرپذیر، نتایج بکارگیری این دو زمینه ارزیابی را ارائه و در ادامه، به شرح و توضیح کاملتر نتایج می‌پردازیم و در پایان، جمع‌بندی کلی مطالعه ارائه می‌گردد.

## روش بررسی

بانک داده کلینیکی مورد مطالعه این کار، از مجموع داده‌های ارزیابی شده در انتخاب گردید.<sup>۱۵</sup> این مجموعه داده از یک تعداد بیمار تحت مطالعه ۱۸۱ نفره (شامل ۸۱ مرد و ۱۰۰ زن) در رنج سنی ۲۰ تا ۷۹ سال انتخاب گردید که براساس بررسی‌های پاتولوژیکی صورت

کبد چرب شدید به‌طور جدی‌تر، به دلیل گسترش ریسک سیروز کبدی و حتی بیماری‌های قلبی-عروقی، در موارد درمانی کلینیکی مدنظر قرار می‌گیرد.<sup>۲</sup> فاکتورهای متعددی موجب فرم‌های شدیدتر بیماری می‌گردند که از آن جمله می‌توان به سن، چاقی مفرط، عدم تعادل گلوکز یا دیابت نوع دوم، و تعداد موارد سندرم متابولیک اشاره نمود.<sup>۳-۵</sup> تشخیص دقیق بیماری نیازمند یک پاتولوژیست باتجربه کبد است.<sup>۶</sup> با در نظر گرفتن ماهیت تهاجمی تحلیل پاتولوژی بیماری، در حال حاضر، مارکرهای خونی متعددی نیز مثل آمینوترانسمینانس و گلوتامیل ترانسفراز گاما برای این منظور کاربرد دارند.<sup>۷،۸</sup> با وجود این، این پارامترها به تنهایی نمی‌توانند معیار مناسبی برای تعیین سطح استئاتوسیس باشد. در حالت عملی، درجه‌بندی بیماری‌های کبدی مربوط به ویژگی‌های پاتولوژیکی است که در فرآیند درمانی بیشترین کاربرد را دارند و در پیشرفت بیماری حائز اهمیت هستند.

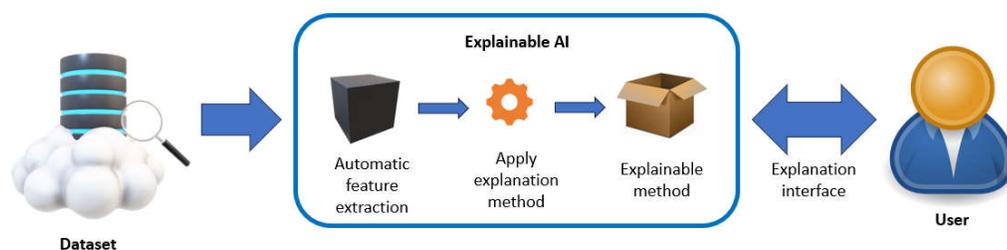
در سال‌های اخیر، هوش مصنوعی تفسیرپذیر توجهات زیادی را به خود جلب نمود و تلاش کرد تا به کاربران امکان تفسیر زنجیره منطقی را برای تصمیم‌گیری نهایی می‌دهد.<sup>۹</sup> برای افزایش توصیف‌پذیری مدل چندین روش ارائه شده است که در دو دسته توصیفات محلی و عمومی می‌توان دسته‌بندی نمود، توصیفات عمومی یک فهم کلی از نحوه عملکرد مدل مورد بحث در ازاء تمام فرایندهای تخمینی ارائه می‌کند. روش‌هایی مثل تفسیر شکلی کلی از مکانیزم‌های مبتنی بر تئوری بازی برای مقادردهی میزان مشارکت هر ویژگی در تخمین نهایی مدل استفاده می‌کنند. دسته بعدی، توصیفات محلی هستند که بر روی خود تصمیم‌گیری و دلیل گرفته شدن چنین تصمیمی تمرکز می‌کنند. تفسیر محلی مدل روش رایج برای تخمین محلی مدل با فرایندهای ساده تفسیر تحلیلی است که به ارائه دید کلی از تاثیر ویژگی‌های به‌خصوص در یک تخمین داده شده می‌پردازد. تفسیر محلی در توصیف هر کلاس پیش‌بینی یک سری مشخصه‌ها را گزارش می‌کند. در مقابل، تفسیر شکل کلی اثر هر ویژگی را بر روی تخمین‌های مدل یادگیری ماشین بر روی بانک داده نشان می‌دهد. این رویکرد تفسیری، اهمیت ویژگی‌ها و تاثیرگذاری آنها را بر روی خروجی کلی با هم ترکیب می‌کند. در نمودار جمع‌بندی این روش تفسیری، ویژگی‌ها براساس درجه اهمیت خود مرتب می‌شوند. هدف اصلی بکارگیری شبکه‌های هوش مصنوعی، دستیابی به یک مکانیزم خودکار برای تشخیص برچسب داده‌های

اطلاعات متقابل<sup>۱۸</sup> و<sup>۱۹</sup> روش انتخاب ویژگی رو به جلو بر مبنای میزان کارایی متغیرهای انتخاب شده برای کلاس‌بندی توسط الگوریتم مدنظر تصمیم می‌گیرد. روش اطلاعات متقابل در واقع اندازه میزان اطلاعاتی است که یک متغیر در مورد متغیر دیگری بدست می‌دهد. این ویژگی می‌تواند برای میزان رابطه یک ویژگی با برجسب خروجی در فرآیند انتخاب ویژگی مورد استفاده قرار گیرد.

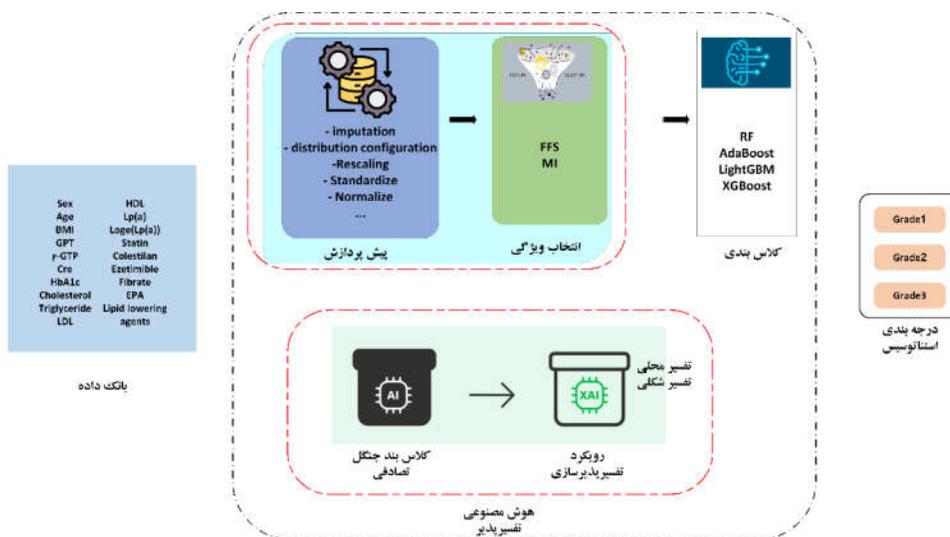
الگوریتم‌های یادگیری ماشین: الگوریتم‌های کلاس‌بندی به دو دسته کلی روش‌های یادگیری ماشین و رویکردهای یادگیری عمیق دسته‌بندی می‌شوند. در این مطالعه، برای دادگان کلینیکی، مجموعه‌ای از الگوریتم‌های با پایه ریاضیاتی مختلف برای فرآیند کلاس‌بندی مورد استفاده قرار گرفت که از مهمترین آنها می‌توان به جنگل

گرفته، مبتلا به بیماری تشخیص داده شده بودند. این بیماران در بیمارستان دانشگاه ایپایم مورد مطالعه قرار گرفتند. تمام این بیماران انتخاب شده، براساس یک فرآیند لاپاروسکوپی، تحت بیوپسی کبدی قرار گرفتند و نمونه برداشته شده پس از پیش پردازش‌های لازم، توسط دو پاتولوژیست مجرب، بدون اطلاع از سایر اطلاعات جمع‌آوری شده از بیماران، مورد بررسی قرار گرفت.<sup>۱۷</sup> نتیجه این بررسی، منجر به ارائه یک برجسب استاندارد برای همه نمونه‌های تحت مطالعه در ارزیابی کارایی روند کلاس‌بندی شد.

روش‌های انتخاب ویژگی: در این کار، الگوریتم‌های مختلف انتخاب ویژگی مورد استفاده و پیاده‌سازی قرار گرفت. این الگوریتم‌ها عبارتند از: انتخاب ویژگی رو به جلو و انتخاب ویژگی بر پایه



شکل ۱: شماتیک کلی ساختار رویکردهای هوش مصنوعی تفسیرپذیر



شکل ۲: شمای کلی فرایند دنبال شده در مطالعه پیش‌رو، برای کلاس‌بندی و مطالعه تفسیری

پردازنده CPU i7 5900 HS و کارت گرافیک Nvidia GeForce 3060 با ۱۶ گیگابایت حافظه RAM بکار گرفته شد.

## یافته‌ها

برای بانک داده، نتایج در قالب جدول ۱ ارائه شد. مطابق با این جدول، بهترین نتیجه به ازای بکارگیری روش کلاس‌بندی جنگل تصادفی به‌دست آمد. مطابق این آرایش، بالاترین دقت مدل جنگل تصادفی به ازای تعداد ۸۰ درخت تصمیم‌گیری به‌دست آمده است. نتیجه ارزیابی توسط الگوریتم‌های هوش مصنوعی تفسیرپذیر، برای عملکرد هریک از ویژگی‌ها در کلاس‌های مختلف در برجسب استاتوستیس با استفاده از روش توصیفی محلی و کلی در شکل ۳ آورده شده است.

## بحث

در این کار، برای دادگان کلینیکی بانک داده، الگوریتم پیشنهادی به‌صورت اتوماتیک می‌تواند به تخمین درجه تجمع چربی استاتوستیس با دقت قابل قبولی پرداخته و یک گزارش اولیه از وضعیت بیمار پیش از استفاده از روش‌های تخمینی دیگر تهاجمی در اختیار پزشک متخصص قرار دهد. بر این اساس، روش تقویت‌کننده افراطی توانسته است تا با دقت ۶۹/۲۳٪ به تعیین سطح ابتلا پردازد. الگوریتم پیشنهادی در این کار مزایای متعددی نسبت به سایر مطالعات ارائه می‌نماید. این پاسخ در شرایطی به‌دست آمده است که امکان دسترسی به سایر اطلاعات تکمیلی بیماران وجود نداشته است. از طرفی دیگر، روش تخمین تقویت‌کننده افراطی به‌عنوان یک الگوریتم مقاوم، با اتکا به جمع‌بندی نهایی از نظر درخت‌های طراحی شده، می‌تواند گزینه مناسبی برای دستیابی به یک نتیجه قابل قبول به‌شمار آید. این مجموعه داده، مطابق با، با هدف بررسی رابطه بین ریسک حمله قلبی و کبد چرب غیرالکلی جمع‌آوری گردید و تنها چند شاخص اولیه این فرآیند برای بررسی در نظر گرفته شد.<sup>۱۶</sup> مطالعه نتایج الگوریتم‌های تفسیری، نشان می‌دهد که مطابق با شکل ۳-ب، ویژگی‌های مورد استفاده به‌ترتیب اولویت، گلوتامیت پایروویت ترانس‌مینانس، تری‌گلیسرید، اندیس BMI، هموگلوبین

تصادفی، کلاس‌بند تقویت‌کننده وفقی، کلاس‌بند تقویت‌کننده افراطی و الگوریتم تقویت‌کننده گرادیان اشاره نمود.<sup>۲۰-۲۳</sup>

روش‌های هوش مصنوعی تفسیرپذیر: یکی دیگر از ابزارهای مورد استفاده در فرایند تحلیل دادگان، رویکردهای هوش مصنوعی تفسیرپذیر است. رویکردهای هوش مصنوعی تفسیرپذیر با بررسی روش پاسخگویی هوش مصنوعی به این نیازها پاسخ می‌دهد. پیچیدگی مدل‌های هوش مصنوعی، اغلب آنها را در قالب جعبه سیاه دسته‌بندی می‌کند و درک نحوه عملکرد داخلی آنها را برای کاربر مشکل می‌سازد. لذا یک بده‌بستان بین کارایی مدل و تفسیرپذیری آن برقرار است.<sup>۲۴</sup> یک‌شماره کلی از این ساختارها را ارائه می‌کند. در بررسی دادگان کلینیکی، ابتدا مقادیر گم‌شده به کمک روش خوشه‌بندی k میانگین بازایی و رنج متغیرهای مورد استفاده تکمیل گردید. پس از آن، به ازای دادگان با مقادیر نرمال شده، یک فرایند انتخاب ویژگی بر روی داده‌ها پیاده‌سازی شد. به‌دلیل عدم تعادل بین کلاس‌های مختلف، از روش داده‌افزایی سنتزی بر روی بانک داده‌های مورد مطالعه استفاده شد تا حداکثر بهبود در عملکرد معیارهای کارایی در اختیار قرار داده شود. از طرفی، به کمک یک مکانیزم نمونه‌برداری، تعداد اعضای کلاس‌ها مجدداً مورد ارزیابی قرار گرفت و فرایند وزن‌دهی به کلاس‌های مختلف، متناسب با تعداد اعضای آنها در نظر گرفته شد. تمام در این فرایند، با مدنظر قراردادن الگوریتم با بالاترین کارایی به‌دست آمده، وابستگی خروجی ارزیابی شده و متغیرهای ورودی مورد سنجش قرار گرفتند. همچنین مطابق با رویکرد پیشنهادی، برای جبران محدودیت بانک داده، یک تقسیم‌بندی داده با در نظر گرفتن استراتژی اعتبارسنجی k خوشه‌ای برای آموزش و اعتبارسنجی مدل در نظر گرفته شد تا همه مدل‌های مختلف انتخاب داده آموزش و تست در آرایش‌های مختلف انتخاب گردد. برجسب‌های استاندارد برای کلاس‌های مختلف، به کمک نتیجه ارزیابی پاتولوژیکی بافت کبدی نمونه‌ها صورت گرفت. در این بررسی، به کمک الگوریتم تفسیر محلی مدل، نسبت متغیرهای ورودی با هریک از کلاس‌ها، و رویکرد تفسیر شکلی کلی مورد سنجش قرار گرفت. شکل ۲ شمایی کلی از فرایند پیش‌رو را نشان می‌دهد. پیاده‌سازی و اجرای این پروژه در ازای قالب کاری Tensorflow انجام پذیرفت. توسعه همه مدل‌ها، در محیط برنامه‌نویسی PyCharm Community Edition 2023.2.7 و در بستر

جدول ۱: مقایسه نتیجه عملکرد الگوریتم‌های مختلف کلاس‌بندی درجه استناتوسیس به ازای روش‌های مختلف انتخاب ویژگی در بانک داده

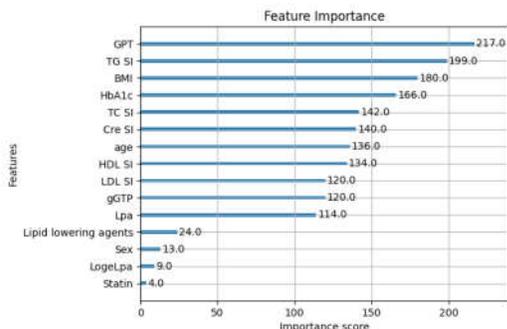
مدل مورد استفاده	روش انتخاب ویژگی	دقت	صحت	بازایی	F1-score
RF	تمام ویژگی‌ها	۴۹/۷۳±۹/۴۴	۴۸/۵۴±۷/۶۳	۴۹/۷۳±۹/۴۴	۴۹/۱۳±۷/۷۲
	FFS	۶۱/۷۵±۵/۹۷	۶۶/۸۱±۶/۲۵	۶۵/۷۵±۵/۹۷	۶۲/۳۷±۵/۱۰
	MI	۵۱/۶۹±۱/۵۶	۵۰/۴۱±۲/۶۸	۵۱/۶۹±۱/۵۶	۵۰/۰۳±۱/۷۴
AdaBoost	تمام ویژگی‌ها	۵۰/۹۳±۷/۸۹	۵۱/۶۴±۶/۰۳	۵۵/۰۱±۲/۴۱	۵۳/۲۷±۴/۳۲
	FFS	۵۱/۴۲±۶/۴۰	۵۲/۱۱±۵/۳۰	۵۱/۴۲±۶/۴۰	۵۰/۷۲±۵/۳۹
	MI	۵۳/۱۸ ± ۵/۲۹	۵۴/۹۹±۶/۳۶	۵۳/۱۸±۵/۲۹	۵۳/۱۳±۵/۱۸
LightGBM	تمام ویژگی‌ها	۵۰/۳۱±۶/۲۲	۴۴/۱۶±۱۲/۸۳	۵۲/۰۶±۶/۰۵	۴۷/۷۸±۶/۶۸
	FFS	۵۲/۰۱±۵/۶۵	۴۶/۱۴±۱۱/۶۴	۵۲/۰۱±۵/۶۵	۴۸/۲۹±۸/۷۵
	MI	۴۷/۹۴±۶/۸۶	۴۸/۰۲±۶/۹۶	۴۷/۹۴±۶/۸۵	۴۶/۹۰±۵/۹۳
XGBoost	تمام ویژگی‌ها	۶۷/۱۱±۴/۲۷	۶۴/۰۰±۴/۵۰	۶۷/۲۳±۴/۳۰	۶۵/۴۱±۴/۳۷
	FFS	۶۹/۲۳±۵/۵۰	۶۷/۳۱±۵/۶۰	۶۹/۴۱±۵/۶۰	۶۷/۳۲±۵/۶۰
	MI	۶۴/۰۶±۶/۴۲	۶۰/۹۱±۶/۲۱	۶۴/۱۸±۶/۱۳	۶۱/۴۲±۶/۱۸

مطابق با نمودار شکل ۳-ج، شاخص‌های مختلف با اثرگذاری‌های سازنده و یا مخرب، با ضریب مقدار متفاوتی در نتیجه نهایی تخمین مشارکت داشته‌اند. مقدار هریک از این شاخص‌ها، در شکل ۳-د ارائه شده است. مقادیری که با رنگ نارنجی در این جدول نشان داده شده‌اند، در حداکثر مقدار نهایی قرار داشتند که بنا به میزان اهمیت پارامتر، در تخمین نهایی و قرارگیری در کلاس به‌دست آمده، اثرگذار بودند. با وجود مزایای متعدد، بکارگیری هوش مصنوعی تفسیرپذیر با چالش‌هایی نیز همراه است که بایست مورد توجه قرار گیرند. از مهمترین دلایل بهبود نتیجه کارایی روش XGBoost نسبت به سایر روش‌های درختی مثل روش جنگل تصادفی و تقویت‌کننده تطبیقی، می‌توان به استفاده این الگوریتم از پارامترهای تنظیم‌گر Regularization برای جلوگیری از Overfitting و همچنین محاسبه چندگانه موازی برای تسریع در عملیات آموزش اشاره نمود. این الگوریتم به‌طور خودکار، فرایند تنظیم درخت‌های طراحی شده را

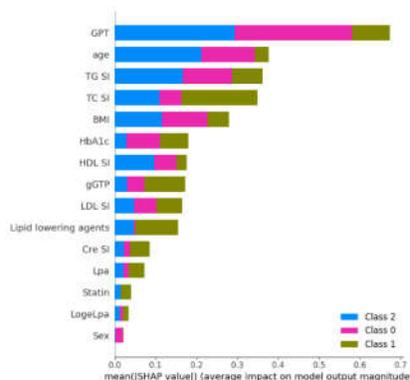
A1c، کلسترول، کراتین، سن، لیپوپروتئین دانسیته بالا و لیپوپروتئین دانسیته پایین هستند. ضریب اهمیت هریک از این ویژگی‌ها، در میزان اثرگذاری آنها در کلاس‌های مختلف نیز به‌صورت محلی در شکل ۳-الف دیده می‌شود. مطابق با شکل ۳-الف، میزان تاثیر هریک از این مشخصه‌های اندازه‌گیری شده در کلاس‌های مختلف متفاوت است، به‌طوری‌که شاخص گلوتامیت پایروویت ترانسمینانس، سن و تری‌گلیسرید در دسته‌بندی در کلاس صفر و دو اثر قابل‌توجهی ایفا می‌کند، ولی برای کلاس یک، کلسترول، گلوتامیت پایروویت ترانسمینانس، هموگلوبین A1c، عامل کاهنده چربی، و gGTP اثربخشی بالاتری خواهد داشت. مطابق با ۰-ج و د، مدل بهینه به‌دست آمده را برای نمونه داده اول مورد آزمایش قرار دادیم تا اثر شاخص‌های مختلف را در تخمین نهایی آن بسنجیم. این داده براساس تست پاتولوژی در کلاس‌بندی ۱ تقسیم شده بود که تخمین صورت گرفته نیز مطابق با شکل ۳-ج، همین نتیجه را به‌دست داد.

الگوریتم‌هایی مثل SHAP را بیشتر نمایان می‌سازد. ولی به هرحال، به‌دلیل همین توانایی‌ها، به‌طور رایج‌تری برای کلاس‌بندی‌های چندکلاس‌ه مورد استفاده قرار می‌گیرد. جدول ۲، مقایسه‌ای نتایج به‌دست آمده از مطالعات مختلف را به ازای دادگان کلینیکی نشان

برعهده می‌گیرد که به میزان زیادی، نتیجه بهینه‌تری نسبت به بکارگیری الگوریتم‌های پیش پردازش در استفاده برای سایر رویکردهای محاسباتی است. اما این الگوریتم از درخت‌های با میزان عمق بالاتری استفاده می‌کند که نیاز برای تفسیر به کمک



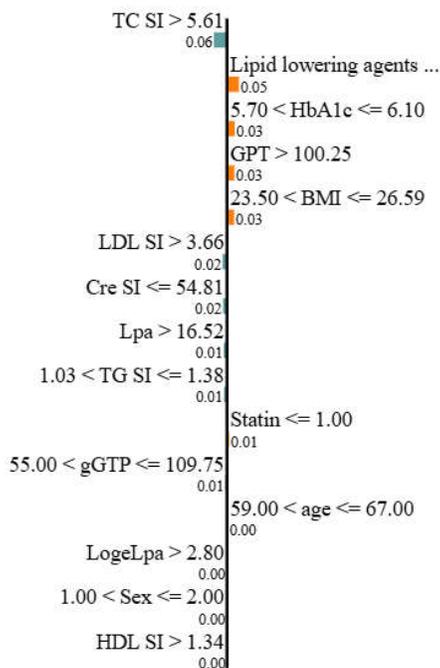
(ب)



(الف)

مقدار	ویژگی
۵/۸۷	TC SI
۱/۰۰	Lipid lowering agents
۶/۱۰	HbA1c
۱۱۸/۰۰	GPT
۲۳/۶۶	BMI
۳/۹۶	LDL SI
۴۵/۹۷	Cre SI
۱۷/۱۰	Lpa
۱/۱۲	TG SI
۱/۰۰	Statin
۵۹/۰۰	gGTP
۶۱/۰۰	Age
۲/۸۴	LogeLpa
۲/۰۰	Sex
۱/۵۰	HDL SI

(د)



(ج)

شکل ۳: مقایسه عملکرد ویژگی‌های مختلف و نوع اثرگذاری آنها در کلاس‌های مختلف بانک داده به ازاء تفسیر کلی (الف و ب) و تفسیر به ازاء یک نمونه انتخاب شده (ج و د).

جدول ۲: مقایسه مطالعات مختلف در کلاس بندی سطح شاخص پاتولوژیکی براساس دادگان کلینیکی

مطالعه	تعداد نمونه	الگوریتم های مورد استفاده	زمینه کلاس بندی	یافته ها
Razmpour F. و همکاران <sup>۳۶</sup>	۲۴۰ مرد	KNN, SVM with RBF, Gaussian process, RF, NN, AdaBoost, naïve Bayes	Fatty liver, steatosis stages, fibrosis stages	Acc (fatty liver) 0.82. Acc (steatosis) 0.52. Acc (fibrosis) 0.57
Ji W. و همکاران <sup>۳۷</sup>	۳۰۴۱۴۵ مراجعه کننده	XGBoost	NAFLD, no NAFLD	Acc 0.88 Prec 0.80 Sen 0.89 F1-score 0.88
Birjandi M. و همکاران <sup>۳۸</sup>	۱۶۰۰ کاندید	Classification tree	NAFLD, no NAFLD	AUC 0.95 Acc 0.75 Sen 0.73 Spec 0.77
Sowa J. و همکاران <sup>۳۹</sup>	۱۲۶ بیمار	RF	Stage of fibrosis	AUC 0.75 Acc 0.79 Sen 0.60 Spec 0.77
مطالعه ما	۱۸۱ نمونه (۱۰۰ مرد، ۸۱ زن)	RF, AdaBoost, LightGBM, XGBoost	Steatosis stages	Acc 0.69 Pre 0.67 Recall 0.69 F1-score 0.67

مصنوعی تفسیرپذیر و الگوریتم های مختلف آن، به خوبی مشاهده می شود. در کنار آن، طراحی یک چارچوب جدید و کاملاً خودکار بر اساس الگوریتم های پیشرفته برای تشخیص فرایند از روی داده های موجود، می تواند در روش های بالینی بسیار مفید و کمک کننده باشد. مطابق با ارزیابی صورت گرفته، متغیرهای GPT و سن و کلسترول و تری گلیسرید و اندیس BMI اثر غالب را در عملکرد مدل کلاس بندی ایفا می کنند که با نتایج تجربی سازگاری دارد.

کد اخلاق: فرایند جمع آوری داده در این مطالعه، براساس کد اخلاق با شماره شناسایی ۱۰۱۲۰۰۴، ۱۷۰۹۰۰۸ در کمیته اخلاق بیمارستان دانشگاه ایهامیم (طبق راهنماهای کد اخلاق هلسینکی اصلاح شده در سال ۱۹۸۳) مورد تایید قرار گرفته است.

سپاسگزاری: این مقاله منتج از پایان نامه تحت عنوان "تشخیص شدت بیماری کبد چرب غیرالکلی مبتنی بر روش های یادگیری عمیق براساس تصاویر سونوگرافی و اطلاعات بالینی" مقطع دکترای تخصصی رشته مهندسی پزشکی در سال ۱۴۰۳ با کد ۲۷۳۲۸ می باشد که در دانشگاه علوم پزشکی شهید بهشتی به انجام رسیده است.

## References

- Chitturi, S., et al., NASH and insulin resistance: insulin hypersecretion and specific association with the insulin resistance

می دهد. یکی از مهمترین نقاط چالش، ایجاد تعادل بین کارایی و توصیف پذیری مدل است. به علاوه، توصیفات حاصل از روش های هوش مصنوعی تفسیرپذیر گاهی اوقات می توانند منجر به گمراهی و یا ساده انگاری بیش از حد گردند. متناسب با الگوریتم های کلاس بندی مورد استفاده، میزان توان تفسیرپذیری نتایج بهبود و یا تخریب می گردد. در نتیجه، بهره گیری از مدل های مقاوم برای دستیابی به توصیفات دقیق و با مفهوم از اهمیت بالایی برخوردار است.

جمع بندی: مطالعه حال حاضر نشان داد که نتایج به دست آمده از رویکرد یادگیری ماشین XGBoost این توانایی را دارد تا با دقت بالاتری، در هر بانک داده با هر تعداد نمونه تحت مطالعه، به خصوص داده های بزرگ، نتیجه بهتری نسبت سایر روش های مشابه و یا با ساختار محاسباتی متفاوت ارائه نماید. این کارایی، در چهار شاخصه دقت، صحت، بازیابی و F1-score مورد مقایسه قرار گرفت. نتیجه ارزیابی نشان داد که براساس اولویت بندی شاخصه ها در کلاس های مختلف، اثر تجربی و تحلیل های آزمایشگاهی سازگاری در کلاس های مختلف را نشان می دهند. این ویژگی، از نتایج بکارگیری هوش

- syndrome. *Hepatology*, 2002. 35(2): p. 373-379. <https://doi.org/10.1053/jhep.2002.30692>

2. Ekstedt, M., et al., Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology*, 2006. 44(4): p. 865-873. <https://doi.org/10.1002/hep.21327>
3. Larter, C.Z., et al., A fresh look at NASH pathogenesis. Part 1: the metabolic movers. *Journal of gastroenterology and hepatology*, 2010. 25(4): p. 672-690. <https://doi.org/10.1111/j.1440-1746.2010.06253.x>
4. Ayonrinde, O.T., et al., Gender-specific differences in adipose distribution and adipocytokines influence adolescent nonalcoholic fatty liver disease. *Hepatology*, 2011. 53(3): p. 800-809. <https://doi.org/10.1002/hep.24097>
5. Gan, L., S. Chitturi, and G.C. Farrell, Mechanisms and implications of age-related changes in the liver: nonalcoholic fatty liver disease in the elderly. *Current gerontology and geriatrics research*, 2011. 2011(1): p. 831536. <https://doi.org/10.1155/2011/831536>
6. Brunt, E.M., et al., Portal chronic inflammation in nonalcoholic fatty liver disease (NAFLD): A histologic marker of advanced NAFLD—Clinicopathologic correlations from the nonalcoholic steatohepatitis clinical research network†. *Hepatology*, 2009. 49(3): p. 809-820. <https://doi.org/10.1002/hep.22724>
7. Nanji, A.A., S.W. French, and J.B. Freeman, Serum alanine aminotransferase to aspartate aminotransferase ratio and degree of fatty liver in morbidly obese patients. *Enzyme*, 1986. 36(4): p. 266-269. <https://doi.org/10.1159/000469304>
8. Giboney, P.T., Mildly elevated liver transaminase levels in the asymptomatic patient. *American family physician*, 2005. 71(6): p. 1105-1110. PMID: 15791889
9. Garcia-Carretero, R., R. Holgado-Cuadrado, and Ó. Barquero-Pérez, Assessment of classification models and relevant features on nonalcoholic steatohepatitis using random forest. *Entropy*, 2021. 23(6): p. 763. <https://doi.org/10.3390/e23060763>
10. Ma, H., et al., Application of machine learning techniques for clinical predictive modeling: a cross-sectional study on nonalcoholic fatty liver disease in China. *BioMed research international*, 2018. 2018(1): p. 4304376. <https://doi.org/10.1155/2018/4304376>
11. Wu, C.-C., et al., Prediction of fatty liver disease using machine learning algorithms. *Computer methods and programs in biomedicine*, 2019. 170: p. 23-29. <https://doi.org/10.1016/j.cmpb.2018.12.032>
12. Newsome, P.N., et al., FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. *The Lancet Gastroenterology & Hepatology*, 2020. 5(4): p. 362-373. [https://doi.org/10.1016/s2468-1253\(19\)30383-8](https://doi.org/10.1016/s2468-1253(19)30383-8)
13. Okanou, T., et al., Artificial intelligence/neural network system for the screening of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Hepatology Research*, 2021. 51(5): p. 554-569. <https://doi.org/10.1111/hepr.13628>
14. Canbay, A., et al., Non-invasive assessment of NAFLD as systemic disease- A machine learning perspective. *PLOS ONE*, 2019. 14(3): p. e0214436. <https://doi.org/10.1371/journal.pone.0214436>
15. Docherty, M., et al., Development of a novel machine learning model to predict presence of nonalcoholic steatohepatitis. *Journal of the American Medical Informatics Association*, 2021. 28(6): p. 1235-1241. <https://doi.org/10.1093/jamia/ocab003>
16. Konishi, K., et al., Advanced fibrosis of non-alcoholic steatohepatitis affects the significance of lipoprotein(a) as a cardiovascular risk factor. *Atherosclerosis*, 2020. 299: p. 32-37. <https://doi.org/10.1016/j.atherosclerosis.2020.02.026>
17. Kleiner, D.E., et al., Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*, 2005. 41(6): p. 1313-21. <https://doi.org/10.1002/hep.20701>
18. Marcano-Cedeño, A., et al. Feature selection using Sequential Forward Selection and classification applying Artificial Metaplasticity Neural Network. in *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*. 2010. <http://dx.doi.org/10.1109/IECON.2010.5675075>
19. Vergara, J.R. and P.A. Estévez, A review of feature selection methods based on mutual information. *Neural Computing and Applications*, 2014. 24(1): p. 175-186. <https://doi.org/10.1007/s00521-013-1368-0>
20. Shaik, A.B. and S. Srinivasan. A Brief Survey on Random Forest Ensembles in Classification Model. in *International Conference on Innovative Computing and Communications*. 2019. Singapore: Springer Singapore. [http://dx.doi.org/10.1007/978-981-13-2354-6\\_27](http://dx.doi.org/10.1007/978-981-13-2354-6_27)
21. Cao, Y., et al., Advance and Prospects of AdaBoost Algorithm. *Acta Automatica Sinica*, 2013. 39(6): p. 745-758. [https://doi.org/10.1016/S1874-1029\(13\)60052-X](https://doi.org/10.1016/S1874-1029(13)60052-X)
22. Chen, T. and C. Guestrin. Xgboost: A scalable tree boosting system. in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016. <http://dx.doi.org/10.1145/2939672.2939785>
23. Ke, G., et al., Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 2017. 30.
24. Deeks, A., The judicial demand for explainable artificial intelligence. *Columbia Law Review*, 2019. 119(7): p. 1829-1850. <https://www.jstor.org/stable/26810851>
25. Escalante, H.J., et al., Explainable and interpretable models in computer vision and machine learning. 2018: Springer.
26. Razmpour, F., et al., Application of machine learning in predicting non-alcoholic fatty liver disease using anthropometric and body composition indices. *Scientific reports*, 2023. 13(1): p. 4942. <https://doi.org/10.1038/s41598-023-32129-y>
27. Ji, W., et al., A machine learning based framework to identify and classify non-alcoholic fatty liver disease in a large-scale population. *Frontiers in Public Health*, 2022. 10: p. 846118. <https://doi.org/10.3389/fpubh.2022.846118>
28. Birjandi, M., et al., Prediction and diagnosis of non-alcoholic fatty liver disease (NAFLD) and identification of its associated factors using the classification tree method. *Iranian Red Crescent Medical Journal*, 2016. 18.(11) <https://doi.org/10.5812/ircmj.32858>
29. Sowa, J.-P., et al., Novel algorithm for non-invasive assessment of fibrosis in NAFLD. *PloS one*, 2013. 8(4): p. e62439. PMID: 23638085.

## Diagnosing the severity of non-alcoholic fatty liver disease based on explainable artificial intelligence

Hamed Zamanian Ph.D.  
Ahmad Shalbaf Ph.D.\*

Department of Physics and  
Biomedical Engineering, School of  
Medicine, Shahid Beheshti  
University of Medical Sciences,  
Tehran, Iran.

\*Corresponding author: Department of  
Physics and Biomedical Engineering,  
School of Medicine, Shahid Beheshti  
University of Medical Sciences, Koodak  
Yar St., Daneshjou Blvd., Velenjak,  
Tehran, Iran.  
Tel: +98-21-22439941  
E-mail: Shalbaf@sbmu.ac.ir

### Abstract

Received: 26 Sep. 2024 Revised: 01 Oct. 2024 Accepted: 13 Dec. 2024 Available online: 21 Dec. 2024

**Background:** Nonalcoholic fatty liver disease (NAFLD) represents a growing global health burden, strongly associated with rising rates of obesity, diabetes, and metabolic syndrome. This study introduces a machine learning framework to precisely diagnose NAFLD, classify disease severity, and stratify risk using routine clinical data. Our model improves early detection and risk prediction, supporting evidence-based clinical decisions. Leveraging predictive analytics, this scalable approach identifies high-risk patients and enables personalized interventions. The data-driven strategy optimizes NAFLD management by extracting maximal value from standard healthcare records, delivering both clinical and operational advantages.

**Methods:** This study examined 181 NAFLD patients across disease stages. The dataset was compiled from February 2010 to January 2019 at Eheim University Hospital, comprising general volunteers who were diagnosed with or without fatty liver based on histopathological evaluation of liver biopsy samples. Forward selection and mutual information identified predictive features, applied in classification models (e.g., random forest) to assess steatosis severity. Explainable AI (XAI) improved model interpretability. Combining robust feature selection, machine learning, and XAI ensured accurate, clinically actionable NAFLD severity evaluation.

**Results:** The XGBoost classifier with forward feature selection attained a classification accuracy of 69.23%±5.5% for steatosis severity. Interpretability analysis highlighted age, Body Mass Index (BMI), High-Density Lipoprotein (HDL), Low-Density Lipoprotein (LDL), A1c Hemoglobin (HbA1c), and glutamate pyruvate transaminase (GPT) as the most impactful variables across three severity classes. Furthermore, GPT, age, BMI, HDL, HbA1c, LDL, triglycerides, and cholesterol were critical to model performance, emphasizing their diagnostic significance in NAFLD progression. These findings suggest their utility in clinical assessments and risk stratification.

**Conclusion:** This study developed a machine learning model for accurate NAFLD diagnosis and severity stratification using routine clinical data. Accessible biomarkers reliably predicted disease progression, enabling gastroenterologists to facilitate early intervention. This cost-effective approach reduces healthcare costs while improving outcomes through precision medicine. Implementing such predictive tools in clinical practice could optimize resource allocation and enhance long-term NAFLD management. The framework supports timely diagnostics and targeted therapies, advancing patient-centered care.

**Keywords:** artificial intelligence, data collection, machine learning, non-alcoholic fatty liver.

Copyright © 2025 Zamanian et al. Published by Tehran University of Medical Sciences.



This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International license (<https://creativecommons.org/licenses/by-nc/4.0/>). Non-commercial uses of the work are permitted, provided the original work is properly cited.