



استفاده از تکنیک‌های هوش مصنوعی برای شناسایی میزان شدت بیماری کبد چرب غیر الکلی توسط شاخص‌های بالینی

مجتبی شهبابی^۱، حمید حسن‌پور^{۲*}

۱- دانشگاه صنعتی شاهرود- دانشکده کامپیوتر و فناوری اطلاعات- گروه هوش مصنوعی- دانشجوی ارشد.

۲- دانشگاه صنعتی شاهرود- دانشکده کامپیوتر و فناوری اطلاعات- گروه هوش مصنوعی- استاد.

تاریخ دریافت: ۱۳۹۴/۱۰/۳۰، تاریخ پذیرش: ۱۳۹۵/۳/۱۷

چکیده

مقدمه: بیماری کبد چرب غیر الکلی (NAFLD: Non-alcoholic fatty liver disease) یکی از شایع‌ترین بیماری‌های کبدی است که شدت آن دارای سطوح مختلفی می‌باشد. اخیراً، دستگاه فیبرواسکن به‌عنوان یک روش غیرتهاجمی برای اندازه‌گیری میزان ارتجاع‌پذیری کبد و در نتیجه چرب بودن آن مورد استفاده قرار می‌گیرد. هدف از این پژوهش، ارایه یک روش کم هزینه و ساده برای تشخیص این بیماری از طریق علائم بالینی می‌باشد. **مواد و روش‌ها:** در این تحقیق از یک مجموعه داده شامل ۷۲۶ بیمار استفاده شد که هر یک دارای عارضه کبد چرب با شدت مختلفی بودند. برای هر بیمار، شدت بیماری توسط دستگاه فیبرواسکن اندازه‌گیری و آزمایشات بالینی و سونوگرافی نیز انجام گرفت. سپس به‌منظور تعیین رابطه بین اطلاعات به‌دست آمده از بیماران و سطوح بیماری از شبکه‌های عصبی مصنوعی استفاده شده است. در نهایت، به کمک تکنیک‌های هوش مصنوعی روشی برای استخراج قانون از شبکه‌های عصبی مصنوعی برای نمایش ارتباط بین داده‌ها استفاده شده است.

نتایج: براساس نتایج به‌دست آمده از دستگاه فیبرواسکن، از بین ۷۲۶ بیمار موجود، ۵ مورد در کلاس F4، ۲۳ مورد در کلاس F3، ۱۳۲ مورد در کلاس F2، ۱۵۱ مورد در کلاس F1 و ۴۱۵ مورد در کلاس F0 (افراد سالم) قرار می‌گیرند. طبق روش پیشنهادی، دقت در شناسایی نمونه‌های هر کلاس به‌ترتیب: ۱۰۰٪ برای کلاس F4، ۹۹/۳۱٪ برای کلاس F3، ۹۳/۹۴٪ برای کلاس F2 و ۸۰/۵۸٪ برای کلاس F1 می‌باشد. بر این اساس این روش می‌تواند نمونه‌های دسته‌های F4 و F3 را به‌صورت ایده‌آل و نمونه‌های دسته‌های F2 و F1 را با دقت خوبی شناسایی کند. **نتیجه‌گیری:** نتایج حاصل از این تحقیق نشان می‌دهد که روش پیشنهادی قادر است علاوه بر صرف هزینه کمتر و قابلیت دسترسی آسان‌تر در شناسایی بیماری کبد چرب غیر الکلی (NAFLD)، نحوه تشخیص بیماری و شرایط هر سطح از بیماری را در قالب مجموعه قوانینی برای تشخیص بیماری (بدون ادامه نیاز به شبکه عصبی) برای پزشکان تعیین نماید.

واژه‌های کلیدی: تشخیص بیماری، بیماری کبد چرب غیر الکلی، پارامترهای بالینی، شبکه‌های عصبی مصنوعی، استخراج قانون.

*نویسنده مسئول: دانشگاه صنعتی شاهرود، دانشکده کامپیوتر و فناوری اطلاعات، تلفن: ۰۲۳-۳۲۳۰۰۲۵۱، نمابر: ۰۲۳-۳۲۳۰۰۵۲۱،

Email: h.hassanpour@shahroodut.ac.ir

ارجاع: شهبابی مجتبی، حسن‌پور حمید. استفاده از تکنیک‌های هوش مصنوعی برای شناسایی میزان شدت بیماری کبد چرب غیر الکلی توسط شاخص‌های بالینی. مجله دانش و تندرستی ۱۳۹۵؛ ۱۱(۳): ۶۹-۷۵.

مقدمه

در این پژوهش سعی شده است تا تنها با انجام آزمایشات ساده‌تری مانند آزمایش کامل خون (CBC: Complete blood count) و سونوگرافی بتوان سطوح بیماری NAFLD را همانند آزمایش فیبرواسکن تعیین نمود. همچنین در این پژوهش ابتدا از شبکه‌های عصبی مصنوعی به‌عنوان شناسایی‌کننده‌های سطوح بیماری (مطابق جدول ۱)، از طریق پارامترهای موجود استفاده شده است. سپس از روش FSRE جهت استخراج قوانین از این شبکه‌های عصبی برای تولید راه حل و روش شناسایی این بیماری استفاده گردیده است.

مواد و روش‌ها

برای انجام این پژوهش، از یک مجموعه داده شامل ۷۲۶ نمونه استفاده شده است. در این مجموعه از هر بیمار، ۱۷ ویژگی شامل اطلاعات شخصی (جنسیت (Sex)، سن (Age)، قد (Height) و وزن (Weight))، پارامترهای آزمایش خون (شامل کلسترول (T-CHOL)، کلسترول خوب (HDL)، کلسترول بد (LDL)، تری‌گلیسیرید (TG)، کراتینین (CRE)، تصفیه گلوبولی (GFR)، اوره (BU)، آنزیم کبدی (ALP)، گاما‌گلوتامیل ترانسفراز (GGT)، قند (FBS)، گلوبول سفید (WBC)، پلاکت (PLT) و نتیجه سونوگرافی (Ultra-Score)، جمع‌آوری گردیده است. همچنین نتایج حاصل از آزمایش فیبرواسکن برای هر بیمار به‌دست آمده و مطابق با گسسته‌سازی پیشنهاد شده در (۵)، به مقادیر گسسته تبدیل شدند. این مقادیر گسسته شده به‌عنوان خروجی مطلوب (Desire output) شناسایی‌کننده‌ها و معیار دسته‌بندی این بیماران مورد استفاده قرار گرفت. تمامی آزمایش‌ها در حضور پزشکان مربوطه انجام شده و مورد تأیید آنها قرار گرفته است. همچنین همه بیماران رضایت خود را در مورد استفاده از اطلاعات کلینیکی ثبت شده در پرونده پزشکی خود در امور پژوهشی و تحقیقاتی، به‌صورت شفاهی اعلام نمودند.

این مجموعه دارای داده‌های خام و ازدیاد ویژگی می‌باشد که نیاز به پیش‌پردازش و مدیریت دارد. بدین منظور ویژگی‌های موجود در این مجموعه پس از نرمال‌سازی (Normalization)، توسط روش بهره اطلاعاتی (Information gain) مورد بررسی قرار گرفتند. این روش، یک روش آماری جهت تعیین ارزش هر ویژگی در دسته‌بندی نمونه‌ها، بر اساس نتیجه مطلوب می‌باشد (۱۲). نتایج بهره اطلاعاتی نشان داد که این ویژگی‌ها به تنهایی قادر به شناسایی مناسب نمونه‌ها نبوده و اکثر این ویژگی‌ها نامرتبط می‌باشند. بنابراین ویژگی‌های BMI (Body mass index) و Forns_Score (۹) از ترکیب ویژگی‌های موجود و به‌منظور افزایش کارایی و دقت نهایی، طبق روابط (۱) و (۲) ایجاد گردیده و به مجموعه داده موردنظر اضافه شدند.

$$BMI = (\text{Weight_kg}) / (\text{Height_m}^2) \quad (1)$$

$$\text{Forns Score} = 7.811 - 3.131 \times \ln(\text{PLT}) + 0.781 \times \ln(\text{GGT}) + 3.467 \times \ln(\text{Age}) - 0.014 \times (\text{TCHOL}) \quad (2)$$

بیماری کبد چرب غیر الکلی یا NAFLD یکی از متداول‌ترین بیماری‌های کبدی است که باعث ایجاد سختی و عدم ارتجاع‌پذیری در سلول‌های کبدی می‌شود (۱). این بیماری خطرناک می‌تواند از یک "استاتوز ساده" (Simple steatosis) شروع شود و تا "سیروز" (Cirrhosis) کبدی (سرطان کبد) پیشروی کند (۲). میزان پیشروی این بیماری که از طریق شدت سختی به‌وجود آمده در سلول‌ها ایجاد می‌شود، سطوح مختلف این بیماری را تعیین می‌کند. این سطوح در جدول ۱ قابل مشاهده می‌باشد (۳-۵).

برای اولین بار، در سال ۱۸۸۳ از روش "نمونه‌برداری فیزیکی از بافت کبد (Liver biopsy)" به‌منظور شناسایی این بیماری استفاده شد (۴). اما تهاجمی بودن (Invasive) و خطرات موجود در زمان نمونه‌برداری موجب شد تا این روش جای خود را به روش‌های نوری و تجزیه و تحلیل تصویر بدهد (۴). امروزه فیبرواسکن به‌عنوان کم‌خطرترین روش غیرتهاجمی (Non-invasive) (۶) برای اندازه‌گیری میزان ارتجاع‌پذیری کبد مورد استفاده قرار می‌گیرد و می‌تواند سطوح این بیماری را نشان دهد. در کنار مزایایی همچون دقت بالا در شناسایی بیماری و کم‌خطر بودن این روش، عواملی مانند هزینه بالا و عدم دسترسی آسان به این دستگاه (در همه کارآمدتری باشند که دقت آنها همانند روش فیبرواسکن باشد (۴، ۷-۱۰)). محققینی همچون فورنس و همکاران (۹) در سال ۲۰۰۲، Lok و همکاران (۸) در سال ۲۰۰۵، انگولو و همکاران (۱۰) در سال ۲۰۰۷ و فوجیوارا و همکاران (۴) در سال ۲۰۰۷، هر کدام سعی نمودند تا با انجام آزمایشات بالینی مختلف بر روی افراد و مقایسه مقادیر به‌دست آمده از پارامترهای مختلف، فرمول و راه حلی را برای تشخیص این بیماری ارائه دهند. اما این روش‌ها تنها قادرند یک یا دو سطح خطرناک بیماری را از بین افراد سالم تشخیص دهند (عدم توانایی در تشخیص حالت ابتدایی و اولیه بیماری). همچنین برخی از روش‌ها نسبت به روش فیبرواسکن از دقت پایینی برخوردار هستند. امروزه از تکنیک‌های هوش مصنوعی مانند شبکه‌های عصبی مصنوعی، به‌دلیل نتایج مناسب و قابل قبول آن در زمینه‌های گوناگون پزشکی و تشخیص بیماری استفاده می‌گردد. دکتر پورنیک و همکاران (۱۱) در سال ۲۰۱۴ توانستند با به‌کارگیری یک شبکه عصبی مصنوعی، بیماری سیروز کبدی را در بین افراد تشخیص دهند. اما عملکرد این ابزار برای پزشکان به نوعی مبهم می‌باشد؛ زیرا تنها با دریافت مقادیر ورودی، خروجی موردنظر را تولید می‌نماید و راه حل و نحوه تشخیص بیماری را برای پزشکان مشخص نمی‌سازد. هدف از این پژوهش و نوآوری آن، استفاده از شبکه عصبی مصنوعی جهت تولید راه حل برای تشخیص بیماری کبد چرب غیر الکلی است که بتوان پس از استخراج راه حل (استخراج قانون)، بدون استفاده از شبکه عصبی مصنوعی این بیماری را تشخیص داد.

حذف نمونه‌های مشابه و تعیین قوانین عمومی برای هر دسته می‌باشد. در مرحله آخر نیز یکی از دسته‌های خروجی را به‌عنوان پاسخ پیش‌فرض مسئله در نظر می‌گیریم.

مراحل اول تا سوم این روش بر روی هر یک از چهار شبکه عصبی در نظر گرفته شده اعمال و نتیجه آن در قالب یک مجموعه قانون دسته‌بندی برای هر دسته از بیماری به‌دست آمد. همچنین دسته F0 به‌عنوان نتیجه پیش‌فرض این مسأله در نظر گرفته شد.

برای انجام عملیات آماری و آماده‌سازی داده‌ها از نرم‌افزار Weka و برای پیاده‌سازی شناسایی‌کننده‌ها و الگوریتم FSRE از نرم‌افزار Matlab استفاده شده است.

نتایج

از بین ۷۲۶ نمونه موجود در مجموعه داده استفاده شده، ۵ مورد در دسته F1، ۲۳ مورد در دسته F2، ۱۳۲ مورد در دسته F3، ۱۵۱ مورد در دسته F4 و ۴۱۵ مورد در دسته F0 (افراد سالم) می‌باشد. پس از آموزش هر شناسایی‌کننده توسط مجموعه داده موجود، نتیجه حاصل از آموزش هر کدام از آنها به‌دست آمد. این نتایج به‌ترتیب با مقادیر حساسیت (Sensitivity) (۳) و اختصاصی بودن (Specificity) (۴) مشخص می‌شود.

$$\text{Sensitivity} = TP / (TP + FN) \quad (3)$$

$$\text{Specificity} = TN / (TN + FP) \quad (4)$$

در این روابط منظور از TP مثبت صحیح (True positive)، FP مثبت کاذب (False positive)، TN منفی صحیح (True negative) و FN منفی کاذب (False negative) می‌باشد.

برای شناسایی‌کننده دسته F4 نتایج به‌دست آمده برابر است با ۱۰۰٪ و ۱۰۰٪، و برای شناسایی‌کننده دسته F3 برابر با ۱۰۰٪ و ۹۹/۲۸٪ است. این مقادیر نشان می‌دهد که شناسایی‌کننده‌های این دو دسته قادرند نمونه‌ها را به‌صورت ایده‌آل و کاملاً صحیح شناسایی کنند. همچنین نتایج به‌دست آمده برای شناسایی‌کننده دسته F2 برابر است با ۹۳/۶۰٪ و ۹۴/۹۶٪، و برای شناسایی‌کننده دسته F1 برابر ۸۰/۷۹٪ و ۸۲/۷۸٪ است. مشاهده می‌شود که شناسایی‌کننده‌های دسته‌های F1 و F2 نیز قادرند نمونه‌ها را به میزان مؤثر و قابل قبولی شناسایی کنند. دسته F0 به‌عنوان پاسخ پیش‌فرض این مسأله می‌باشد و به همین دلیل برای آن شناسایی‌کننده‌ای در نظر گرفته نشده است.

جدول ۱- سطوح مختلف بیماری NAFLD و معانی هر سطح

توضیحات	سطوح
Absence of fibrosis (No fibrosis)	F0
Perisinusoidal or portal involvement (Portal fibrosis without septa)	F1
Perisinusoidal and portal/peripoportal involvement (Portal fibrosis and few septa)	F2
Septal or bridging fibrosis (Numerous septa without cirrhosis)	F3
Cirrhosis	F4

همانطور که قبلاً نیز بیان شد، این مجموعه داده دارای تعداد زیادی ویژگی می‌باشد که استفاده از همه آنها باعث کاهش کارایی نتایج نهایی می‌گردد. پس به‌منظور اجتناب از ویژگی‌های اضافی در عملیات، هفت ویژگی برتر برای شناسایی هر دسته، بر اساس امتیاز بهره اطلاعاتی انتخاب شدند. ویژگی‌های انتخاب شده عبارتند از: سونوگرافی (Ultra_Score)، کلسترول خوب خون (HDL)، کلسترول بد خون (LDL)، تری‌گلیسرید (TG)، قند خون (FBS)، نسبت قد به وزن (BMI) و نتیجه روش فورنز (Forms_Score). ویژگی‌های انتخاب شده عموماً دارای مقادیر پیوسته هستند. مقادیر این ویژگی‌ها توسط روش ur-CAIM (۱۳) که یک روش هوشمند و بهینه در زمینه گسسته‌سازی است، به مقادیر گسسته تبدیل شد. سپس به‌منظور شناسایی نمونه‌های هر دسته، از ابزار شبکه عصبی مصنوعی استفاده گردید. این روش که اساس کار آن برگرفته از عملیات درون شبکه عصبی ذهن انسان است، به‌عنوان یکی از بهترین روش‌ها در زمینه دسته‌بندی شناخته می‌شود (۱۴ و ۱۵). در این روش ابتدا نیاز است تا شبکه عصبی ایجاد شده توسط یک مجموعه داده موجود و شناخته شده آموزش یافته تا رابطه بین نمونه‌های هر کلاس را به‌دست آورد. پس از مرحله آموزش می‌توان از این شبکه آموزش دیده برای دسته‌بندی نمونه‌های جدید (علاوه بر نمونه‌های مجموعه آموزش) استفاده نمود. نتایج تحقیقات انجام شده در سال‌های اخیر نشان می‌دهد که این ابزار می‌تواند حتی برای مسایل پیچیده مانند تشخیص انواع بیماری‌ها که رابطه خاصی برای آن وجود ندارد، نتایج بهینه‌ای به‌دست آورد (۱۶-۱۸). در این پژوهش، چهار شبکه عصبی مصنوعی چند لایه (Multi-layer perceptron (MLP)) به‌عنوان شناسایی‌کننده‌های دسته‌های بیماری (از F1 تا F4) ایجاد گردید. سپس این شناسایی‌کننده‌ها توسط زیرمجموعه‌ای از داده‌های موجود (۷۰٪) آموزش دیده و کارایی آن توسط تمام داده‌های موجود ارزیابی شد. به‌منظور ساده‌سازی ساختار این شبکه‌های آموزش دیده از عملیاتی مانند هرس کردن (Pruning) شبکه عصبی و تبدیل مقادیر پیوسته به گسسته در لایه میانی استفاده می‌شود. در نهایت از روش جدید FSRE جهت نمایش چگونگی انجام عملیات درون شبکه عصبی برای تولید نتایج خروجی استفاده شده است. روش FSRE یک روش جدید استخراج قانون مبتنی بر دیدگاه تجزیه‌ای (De-compositional approach) می‌باشد که سعی می‌کند با تجزیه شبکه عصبی به بخش‌های مختلف و تعیین رابطه جزیی بین این بخش‌ها، نهایتاً رابطه کلی برای تولید نتایج خروجی را به‌دست آورد. نحوه انجام این روش در ادامه آمده است:

در الگوریتم FSRE ابتدا برای هر نمونه، مقادیر گسسته در لایه مخفی (در مرحله ساده‌سازی شبکه عصبی) را به‌دست می‌آوریم؛ در مرحله دوم نمونه‌هایی را که دارای مقادیر گسسته مشابه هستند در یک دسته قرار می‌دهیم. مرحله سوم شامل ترکیب دسته‌های متعلق به یک کلاس خروجی،

جدول ۲- نتایج به دست آمده از مراحل مختلف عملیات

# Rules	Accuracy	قوانین استخراج شده		هرس سازی شبکه عصبی		آموزش شبکه عصبی		دسته
		Specificity*	Sensitivity**	Specificity*	Sensitivity**	Specificity*	Sensitivity**	
۵	%۸۰٫۵۸	%۷۵٫۶۵	%۹۹٫۳۴	%۸۱٫۲۱	%۸۴٫۷۶	%۸۲٫۷۸	%۸۰٫۷۹	F1
۳	%۹۳٫۹۴	%۹۳٫۷۷	%۹۳٫۹۳	%۹۳٫۷۷	%۹۳٫۹۳	%۹۳٫۶۰	%۹۴٫۹۶	F2
۳	%۹۹٫۳۱	%۹۹٫۲۸	%۱۰۰	%۹۹٫۲۸	%۱۰۰	%۹۹٫۲۸	%۱۰۰	F3
۱	%۱۰۰	%۱۰۰	%۱۰۰	%۱۰۰	%۱۰۰	%۱۰۰	%۱۰۰	F4

*Specificity = اختصاصی بودن

**Sensitivity = حساسیت

جدول ۳- مقایسه نتایج به دست آمده از روش پیشنهادی (FSRE) و چند روش دیگر

	J48	JRip	Forns_Score	FSRE	
دسته F1	%۶۷٫۰۷	%۷۶٫۷۲	%۷۰٫۱۱	%۸۰٫۵۸	F1
دسته F2	%۷۲٫۱۷	%۷۹٫۷۵	%۸۲٫۵۱	%۹۳٫۹۴	F2
دسته F3	%۹۳٫۵۲	%۹۴٫۷۶	%۹۱٫۳۲	%۹۹٫۳۱	F3
دسته F4	%۹۸٫۶۲	%۹۸٫۷۶	%۹۹٫۳۱	%۱۰۰	F4

پس از هرس سازی این شناسایی کننده‌ها، نتایج همانند مرحله قبل برای هر شناسایی کننده به دست آمد. این نتایج برای شناسایی کننده دسته F4 به ترتیب برابر با ۱۰۰٪ و ۱۰۰٪ و برای شناسایی کننده دسته F3 به ترتیب برابر با ۱۰۰٪ و ۹۹٫۲۸٪ بوده است. همچنین مقادیر ۹۳٫۹۳٪ و ۹۳٫۷۷٪ برای شناسایی کننده دسته F2 و مقادیر ۸۴٫۷۶٪ و ۸۱٫۲۱٪ برای شناسایی کننده دسته F1 به دست آمده است. مشاهده می شود که همانند مرحله قبل، شناسایی کننده‌های دسته‌های F3 و F4 قادرند با ساختار هرس شده همچنان نمونه‌ها را به صورت کاملاً صحیح شناسایی کنند. شناسایی کننده‌های دسته‌های F2 و F1 نیز قادرند تا در این حالت (با ساختار هرس شده) همچنان میزان مؤثر و قابل قبولی از نمونه‌ها را شناسایی نمایند.

در نهایت، با استفاده از الگوریتم FSRE قوانین و روابط مورد نظر از شناسایی کننده‌ها به دست آمد. سپس نتایج این مرحله نیز همانند مراحل قبل به ترتیب با مقادیر حساسیت و اختصاصی بودن محاسبه گردید. این نتایج برای دسته F4 با یک قانون برابر با ۱۰۰٪ و ۱۰۰٪ و برای دسته F3 با سه قانون برابر با ۱۰۰٪ و ۹۹٫۲۸٪ بود. همچنین مقادیر به دست آمده برای دسته F2 نیز توسط سه قانون به ترتیب ۹۳٫۹۳٪ و ۹۳٫۷۷٪ و برای دسته F1 با پنج قانون به ترتیب ۹۹٫۳۴٪ و ۷۵٫۶۵٪ می باشد. این نتایج مشخص می نماید که قوانین تولید شده برای دسته‌های F4 و F3 قادرند تا نمونه‌ها را همانند شناسایی کننده‌های پایه برای این دسته‌ها، به صورت ایده‌آل و کاملاً صحیح شناسایی نمایند. همچنین قوانین تولید شده برای دسته‌های F1 و F2 نیز قادرند تا نمونه‌ها را به میزان مؤثر و قابل قبولی شناسایی کنند.

(۵)

Accuracy = (Number of Correct Classified Instance) / (Number of All Instance)

در انتها میزان دقت این روش برای شناسایی نمونه‌های هر دسته از بیماری NAFLD طبق رابطه (۵) محاسبه گردید. این میزان برای شناسایی نمونه‌های دسته F4 برابر ۱۰۰٪ و برای شناسایی نمونه‌های دسته F3 برابر

۹۹٫۳۱٪ بوده است. همچنین برای دسته F2 میزان دقت ۹۳٫۹۴٪ و برای دسته F1 میزان دقت ۸۰٫۵۸٪ به دست آمد. تمامی نتایج ذکر شده در این قسمت به صورت مختصر در جدول ۲ آمده است. در این جدول ستون اول، نتایج به دست آمده از شبکه عصبی اولیه برای تشخیص بیماری NAFLD می باشد. ستون دوم نتایج به دست آمده از شبکه عصبی هرس شده و ستون سوم نتایج به دست آمده از قوانین تولید شده (روش تولید شده برای تشخیص این بیماری) توسط روش پیشنهادی است. با دقت در این نتایج مشاهده می شود که شبکه عصبی در حالت اولیه (اتصال کامل) نتایج بهتری را ارائه داده است؛ اما به دلیل نیاز به ساده سازی ساختار شبکه عصبی، دقت به دست آمده در حالت هرس شده کمی کاهش داشته است. همچنین در قسمت قوانین استخراج شده نیز به دلیل حذف برخی از قوانین کم ارزش، اندکی افت دقت (در سطوح اولیه F1 و F2 بیماری) نسبت به حالت هرس شده شبکه عصبی مشاهده می گردد. ستون Accuracy صحت نتایج به دست آمده از روش تولید شده برای شناسایی بیماری NAFLD نسبت به روش فیرواسکن نشان داده شده است. در واقع مقادیر این ستون نشان می دهند که روش جدید تولید شده تا چه میزان قادر است همانند روش فیرواسکن بیماران را در سطوح مختلف به درستی تشخیص دهد. همچنین در ستون آخر (#Rules) تعداد شرایط و قوانین موجود برای هر سطح از بیماری نشان داده شده است. مشاهده می شود که توسط روش تولید شده در تشخیص بیماری NAFLD، با تعداد اندک قوانین می توان درصد بالایی از بیماران در سطوح مختلف را شناسایی نمود.

نتایج نهایی تولید شده در قالب مجموعه قوانین دسته بندی برای شناسایی سطوح شدت بیماری NAFLD در ادامه آمده است:

• **Class F4:**

if (LDL < 216) and (FornsScore ≥ 115.01) then ClassF4

• **Class F3:**

if (UltraScore = [1 or 7]) and (LDL < 178.5) and (FronsScore ≥ 109.035) then ClassF3

در این پژوهش مشخص شد که روش پیشنهادی قادر است سطوح خطرناک‌تر بیماری (سطوح F3 و F4) را به‌صورت ایده‌آل شناسایی کند. همچنین در این روش نمونه‌های سطوح معمولی‌تر بیماری (سطوح F1 و F2) نیز به‌صورت قابل قبول شناسایی می‌شوند.

یکی از اصلی‌ترین مزیت‌های این روش نسبت به روش‌های موجود دیگر مانند (9) Forns، (8) Lok، (10) Angulo و (4) Fujiwara، شناسایی سطح ابتدایی بیماری (یعنی سطح F1) می‌باشد. این روش می‌تواند با دقت مناسبی وجود این سطح از بیماری را در بین بیماران شناسایی کند. این حالت منجر می‌شود که پزشکان بتوانند هرچه سریع‌تر عملیات مورد نظر جهت جلوگیری از پیشروی بیماری را انجام دهند. درحالی‌که، در روش‌های تشخیص بیماری پیشین (۴، ۷-۱۰)، تنها سطوح F2 تا F4 از بیماری تشخیص داده می‌شود و این زمانی است که بیماری به حالت خطرناک خود رسیده است.

مزیت دیگر این روش، تولید نتایج به‌صورت مجموعه‌ای از قوانین می‌باشد که در مقابل رابطه فرمولی (ریاضی) به‌دست آمده از روش‌های اخیر، از قابلیت درک بالاتری برخوردار می‌باشد. این قوانین تولید شده می‌تواند توسط پزشکان و مسئولین آزمایشگاه‌ها مورد بررسی و تأیید قرار گیرد.

در پایان نیز نتایج به‌دست آمده از روش پیشنهادی با برخی از دیگر روش‌های هم‌رده (تولید قانون و راه حل) برای تشخیص بیماری کبد چرب غیرالکلی مقایسه گردید. این نتایج که در جدول ۳ نشان داده شده است به‌ترتیب شامل نتایج روش پیشنهادی (FSRE)، روش Forns (۹)، روش JRip (۱۹) و روش J48 (۲۰) می‌باشد. این نتایج نشان می‌دهد که روش پیشنهادی با استفاده از شبکه عصبی توانسته است راه حل مناسب‌تری را برای تشخیص بیماری NAFLD (با استفاده از قوانین تولید شده) به‌دست آورد. همچنین به‌دلیل استفاده آسان‌تر و هزینه کمتر این روش نسبت به آزمایش فیبرواسکن، استفاده‌های بیشتری از این روش می‌توان نمود.

References

- Jamali RJA. Non-alcoholic fatty liver disease. Journal of Kashan University of Medical Sciences 2010;14:169-81. [Persian].
- Sug H, editors. Improving the prediction accuracy of liver disorder disease with oversampling. proceeding of the 6th wseas international conference on Computer Engineering and Application, and American conference on Applied Mathematics; 2012; Cambridge 2012.
- Cales P, Boursier J, Chaigneau J, Laine F, Sandrini J, Michalak S, et al. Diagnosis of different liver fibrosis characteristics by blood tests in non-alcoholic fatty liver disease. Liver International 2010;30:1346-54. doi: 10.1111/j.1478-3231.2010.02314.x
- Fujiwara Sh, Hongou Y, Miyaji K, Asai A, Tanabe T, Fukui H, et al. Relationship between liver fibrosis noninvasively measured by fibro scan and blood test. Bulletin of the Osaka Medical College 2007;53:93-105.
- Janda JM ASL, editors. The Entrobacteria. 2nd ed. Washington: ASM press;2006.
- Afdhal NH. Fibroscan (transient elastography) for the measurement of liver fibrosis. Gastroenterol Hepatol 2012;8:605-7.

if (UltraScore = [1 or 3 or 5 or 7]) and (LDL \geq 178.5) and (69.8 \leq FBS < 220.7) and (FornsScore \geq 109.035) then ClassF3

if (2 \leq UltraScore \leq 6) and (LDL < 178.5) and (69.8 \leq FBS < 220.7) and (FornsScore \geq 109.035) then ClassF3

• Class F2:

if (LDL \geq 71.5) and (FornsScore \geq 106.485) then ClassF2

if (3 \leq UltraScore \leq 5) and (LDL < 65.5) and (FornsScore \geq 106.485) then ClassF2

if (UltraScore = 1) and (65.5 \leq LDL < 71.5) and (FornsScore \geq 106.485) then ClassF2

• Class F1:

if (LDL \geq 105.5) and (TRIG < 275.5) and (BMI < 32.108) and (105.085 \leq FronsScore < 106.725) then ClassF1

if (LDL < 104.5) and (TRIG < 263) and (BMI < 32.108) and (105.085 \leq FronsScore < 106.725) then ClassF1

if (LDL < 104.5) and (TRIG \geq 275.5) and (BMI < 32.108) and (105.085 \leq FronsScore < 106.725) then ClassF1

if (LDL < 104.5) and (TRIG < 241.5) and (BMI \geq 32.667) and (105.085 \leq FronsScore < 106.725) then ClassF1

• Class F0:

Default is ClassF0

بحث

امروزه برای تشخیص بیماری‌ها (عمدتاً بیماری‌های جدید) استفاده از تکنیک‌های هوشمند کاربرد فراوانی دارند. اما اینگونه روش‌ها نحوه تشخیص بیماری و روش تشخیص آن را برای پزشکان مشخص نمی‌سازند. هدف از این پژوهش و نوآوری آن، استفاده از شبکه عصبی برای تولید راه حل برای شناسایی بیماری کبد چرب غیر الکلی (نه صرفاً شناسایی بیماری) بوده است. نتایج به‌دست آمده از این پژوهش نشان داد که با استفاده از قوانین تولید شده به تنهایی (بدون استفاده از شبکه عصبی) می‌توان سطوح مختلف این بیماری را با دقت مناسبی شناسایی نمود.

7. Wai Ch, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003;38:518-26. doi: 10.1053/jhep.2003.50346
8. Lok AS, Ghany MG, Goodman ZD, Wright EC, Everson GT, Sterling RK, et al. Predicting cirrhosis in patients with hepatitis c based on standard laboratory test: results of the halt-c cohort. *Hepatology* 2005;24:282-92. doi: 10.1002/hep.20772
9. Forns X, Ampurdanes S, Llovet JM, Aponte J, Quinto L, Martinez-Bauer E, et al. Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. *Hepatology* 2002;36:986-92. doi: 10.1053/jhep.2002.36128
10. Angulo P, Hui J, Marchesini G, Bugianesi E, George J, Farrel G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007;45:846-54. doi: 10.1002/hep.21496
11. Pournik O, Dorri S, Zabolinezhad H, Alavian SM, Eslami S. A diagnostic model for cirrhosis in patients with non-alcoholic fatty liver disease: an artificial neural network approach. *Med J Islam Repub Iran* 2014;28:116.
12. Roobeart D, Karakoulus G, editors. Information gain, correlation and support vector machine. In: Guyon I, Gunn S, editors. *Feature extraction*. Berlin: Springer-Verlag Berlin Heidelberg;2006.p.463-70. doi: 10.1007/978-3-540-35488-8
13. Cano A, Nguyen D, Ventura S, Cios K. ur-CAIM: improved CAIM discretization for unbalanced and balanced data. *Soft Computing* 2016;20:173-88. doi: 10.1007/s00500-014-1488-1
14. Antkowiak M. Artificial neural networks vs. Support vector machines for skin diseases recognition [dissertation]. Sweden: Umea Univ.;2006.
15. El Farissi I, Azizi M, Lanet J-L, Moussaoui M. Neural network vs. Bayesian network to detect Java card mutants. *AASRI Procedia* 2013;4:132-7. doi: 10.1016/j.aasri.2013.10.021
16. Ganesan N, Venkatesh K, Rama M, Palani AM. Application of neural networks in diagnosing cancer disease using demographic data. *International Journal of Computer Applications* 2010;1:76-85.
17. Jaleel JA, Salim S, Aswin R. Artificial neural network based skin cancer detection. *International Journal of Advanced Research in Electrical* 2012;1:200-5.
18. Khan IY, Zope P, Suralkar S. Importance of artificial neural network in medical diagnosis disease like acute nephritis disease and heart disease. *International Journal of Engineering Science and Innovative Technology (IJESIT)* 2013;2:210-7.
19. Cohen WW, editor. Fast effective rule induction. *Proceedings of the 12th international conference on machine learning;1995*; Tahoe City, California.
20. Quinlan JR. Learning with continuous classes. *5th Australian joint conference on artificial intelligence; 1992*: Singapore.



Using the Artificial Intelligence Techniques for Diagnosing of intensity of Non-Alcoholic Fatty Liver Disease by Clinical Parameters

Mojtaba Shahabi (M.Sc.)¹, Hamid Hassanpour (Ph.D.)^{1*}

1- Dept. of Artificial Intelligence, School of Computer and Information technology, Shahrood University of Technology, Shahroud, Iran.

Received: 20 January 2016, Accepted: 6 June 2016

Abstract:

Introduction: Non-Alcoholic Fatty Liver Disease (NAFLD) is one of the most prevalent liver diseases with several levels of severity. Recently, the FibroScan device has been used as a non-invasive method for measurement of elasticity of the liver and consequently its fattiness. The current study aimed at providing a low-cost and simple method for diagnosing of the disease through the clinical symptoms.

Methods: A collection of data obtained from 726 patients was used for conducting this study with each patient having a fatty liver disease with different levels of intensity. The severity of the disease for each patient was measured by FibroScan device along with medical tests and ultrasound monitoring. Then, the artificial neural networks were used for determination of the relationship between the data obtained from the patients and the intensity levels. Finally, by the aid of artificial intelligence techniques, a method is employed for extracting rules from artificial neural networks for representing the relationship between the data.

Results: According to the results obtained from FibroScan device, among the 726 patients, 5 were located in F4 class, 23 in F3 class, 132 in F2 class, 151 in F1, and 415 in F0 Class (Healthy people). According to the proposed method, the accuracy of diagnosis for various classes is as follows: 100% for F4 class, 99.31% for F3 class, 93.94% for F2 class, and 80.58% for F1 Class. Accordingly, this method can identify the samples in F4 and F3 classes with an ideal accuracy and the samples in F1 and F2 classes with a good accuracy.

Conclusion: Results in this research indicate that the proposed method can be used for diagnosing NAFLD and identifying its intensity levels with a lower costs and easier accessibility. It extracts the required rules for diagnosing the diseases that can be used by the physicians in their diagnosis.

Keywords: Disease diagnosis, Non-alcoholic fatty liver disease, Clinical parameters, Artificial neural network, Rule extraction.

Conflict of Interest: No

*Corresponding author: H. Hassanpour, Email: h.hassanpour@shahroodut.ac.ir

Citation: Shahabi M, Hassanpour H. Using the artificial intelligence techniques for diagnosing of intensity of non-alcoholic fatty liver disease by clinical parameters. Journal of Knowledge & Health 2016;11(3):69-75.