

## تشخیص ناباروری مردان از روی عوامل محیطی و سبک زندگی با استفاده از الگوریتم‌های هوش مصنوعی

اعظم اصیلیان بیدگلی، حسین ابراهیم‌پور کومله، سیدجلال‌الدین موسوی‌راد\*

گروه مهندسی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه کاشان

\* نویسنده مسؤول: دانشگاه کاشان، دانشکده مهندسی برق و کامپیوتر، آزمایشگاه تحقیقاتی مهندسی کامپیوتر، بلوار قطب راوندی، کاشان، ایران، تلفن: ۰۹۳۵۴۶۶۳۳۳۴  
پست الکترونیک: jalalmoosavirad@gmail.com

دریافت: ۹۴/۱۱/۳ پذیرش: ۹۵/۲/۱۴

### چکیده

**مقدمه:** مشکل ناباروری به خصوص در میان مردان یکی از مسائلی است که در دهه‌های اخیر به آن توجه خاصی شده است. ناباروری در مردان می‌تواند از عوامل مختلفی ناشی شود. تحقیقات گسترده‌ای در خصوص تأثیر عوامل محیطی و سبک زندگی افراد بر روی کیفیت اسپرم مردان و ناباروری آن‌ها انجام شده است. در این بین روش‌های هوش مصنوعی به عنوان یک سیستم تصمیم‌یار هوشمند می‌توانند کمک شایانی در تشخیص ناباروری مردان نمایند.

**روش کار:** در این پژوهش با استفاده از الگوریتم‌های داده‌کاوی در هوش مصنوعی سعی شده است تا عوامل تأثیرگذار بر روی کیفیت اسپرم مردان، و در نتیجه تشخیص ناباروری آن‌ها بررسی و بهترین روش به این منظور معرفی گردد. داده‌های مورد استفاده مربوط به ۱۰۰ داوطلب از دانشجویان دانشگاه آلیکانته بوده است. بعد از متوازن کردن داده‌ها، روش‌های مختلف داده‌کاوی مثل شبکه‌ی عصبی با هدف یافتن بهترین طبقه‌بند مورد بررسی قرار گرفته است.

**یافته‌ها:** روش پیشنهادی با الگوریتم‌های مختلف داده‌کاوی مورد بررسی قرار گرفت تا بتوان تحلیل بهتری از کارایی سایر الگوریتم‌ها به دست آورد. در بین طبقه‌بندهای مورد بررسی، ماشین بردار پشتیبان با استفاده از اعتبارسنجی منقطع ۱۰ لایه، توانست بهترین کارایی را با دقت ۹۵/۱۵٪ کسب کند. روش پیشنهادی، دارای کارایی رقابتی نسبت به سایر روش‌ها می‌باشد. نتیجه‌گیری: نتایج نشان داد که روش پیشنهادی می‌تواند با استفاده از عوامل محیطی و سبک زندگی ناباروری مردان را پیش‌بینی نماید.

**کل واژگان:** ناباروری مردان، هوش مصنوعی، الگوریتم‌های داده‌کاوی، ماشین بردار پشتیبان، سیستم تصمیم‌یار

### مقدمه

هر یک از شاخص‌ها اثر داشته باشند. از جمله این عوامل می‌توان به عادات و شیوه‌ی زندگی، و همچنین وضعیت سلامتی جسمی مردان اشاره کرد (۲ و ۳). نمونه‌های مختلفی در مورد هر یک از این عوامل می‌توان نام برد؛ به عنوان مثال صدمات و آسیب‌های جدی در تصادفات یا حوادث مختلف، بیماری‌های دوران کودکی مانند سرخک، آبله‌مرغان، فلج اطفال؛ و عاداتی مانند سیگار کشیدن یا مصرف الکل از جمله این موارد هستند. بنابراین بررسی تأثیر این عوامل بر کیفیت اسپرم، و نهایتاً میزان اثرگذاری آن بر ناباروری مردان می‌تواند یکی از موضوعات مهم

یکی از نیازهای انسان، تولید مثل و داشتن فرزند می‌باشد؛ و وجود مشکلات و موانع برای رفع این نیاز همیشه یکی از دغدغه‌های پژوهشگران بوده است. مشکل ناباروری در مردان یکی از مسائلی مهم در این زمینه طی دو دهه‌ی اخیر می‌باشد (۱). کیفیت اسپرم در مردان با شاخص‌های مختلفی شناخته می‌شود که عبارتند از: غلظت، میزان حرکت و شکل اسپرم. کم شدن هر یک از شاخص‌های اول و دوم، و همچنین شکل نامتناسب اسپرم‌ها در کیفیت آن تأثیرگذار بوده و باعث کم شدن احتمال باروری مردان می‌گردد. عوامل مختلفی می‌توانند بر روی

مجله تحقیقات نظام سلامت حکیم

شده است، اما در مطالعات مذکور همچنان مشکل عدم توازن وجود دارد. پرسپترون چند لایه برای بررسی عوامل مذکور بر شاخص‌های اثرگذار در کیفیت اسپرم شامل حرکت، غلظت و شکل اسپرم به طور جداگانه استفاده شده است (۱۶). ونگ<sup>۱</sup> و همکارانش (۱۷) یک روش براساس درختان تصمیم ارایه داده‌اند که مشکل عدم توازن داده‌ها را مورد توجه قرار می‌دهد. این مقاله به بررسی طبقه‌بندهای مختلف از جمله ماشین بردار پشتیبان<sup>۲</sup>، درخت تصمیم<sup>۳</sup>، شبکه عصبی مصنوعی<sup>۴</sup>، نزدیک‌ترین همسایگی<sup>۵</sup> و روش‌های ترکیبی از جمله آدابوست<sup>۶</sup> و تجمیع خودراه‌انداز<sup>۷</sup> (۱۸) پرداخته و سعی می‌کند با به کار گرفتن روش نمونه‌گیری خودراه‌انداز، مشکل عدم توازن داده‌ها را حل کند. بنابراین مطالعه‌ی حاضر با ارایه روش‌هایی نوین در آزمایشات، نسبت به پژوهش‌های قبلی دارای کارایی بهتری در تشخیص ناباروری مردان خواهد بود.

### روش کار

داده‌های مورد استفاده در این پژوهش از مرکز سیستم‌های هوشمند و یادگیری ماشین دانشگاه کالیفرنیا به دست آمده است (۱۹). داده‌ها مربوط به ۱۰۰ داوطلب از دانشجویان دانشگاه آلیکانت<sup>۸</sup> در سن بین ۱۸ تا ۳۶ سال در سال ۲۰۱۲ می‌باشد که عادات زندگی و وضعیت جسمانی آن‌ها بررسی شده، و بردار ویژگی مربوط به هر یک را تشکیل داده است. تعداد این ویژگی‌ها ۹ عدد می‌باشد. هر کدام از این افراد در یکی از طبقات نرمال (منفی - بدون احتمال ناباروری) و یا غیرنرمال (مثبت - با احتمال ناباروری) طبقه‌بندی شدند. ۸۸ نفر از آن‌ها در طبقه نرمال و ۱۲ نفر دیگر در طبقه غیرنرمال می‌باشند. توصیف ویژگی‌هایی که تأثیر آن‌ها بر روی کیفیت اسپرم بررسی شده است در جدول ۱ نشان داده شده است.

**متوازن کردن داده‌ها:** در این پژوهش برای حل مشکل عدم توازن داده‌ها از روشی به نام خودراه‌انداز<sup>۹</sup> برای نمونه‌گیری و ایجاد توزیع یکسان بین هر دو طبقه داده‌ایی استفاده شده است. به عبارت دیگر روش خودراه‌انداز، یک روش آماری برای

در علم پزشکی باشد. در سال‌های اخیر علوم داده‌کاوی و یادگیری ماشین، به عنوان ابزاری قدرتمند در سیستم‌های تشخیص در کاربردهای مختلف پزشکی، مورد توجه قرار گرفته‌اند (۴-۶). در زمینه‌ی بررسی اثرگذاری عوامل مختلف بر کیفیت اسپرم مردان، داده‌کاوی وظیفه‌ی دسته‌بندی داده‌ها را انجام می‌دهد. به این منظور مجموعه داده‌هایی از افراد مختلف جمع‌آوری، و در مورد هر فرد، عادات زندگی و وضعیت سلامتی او نیز مشخص می‌گردد. طبقه‌بند ساخته شده با استفاده از ابزار داده‌کاوی سعی می‌کند افراد را به دو مجموعه "نرمال" و "غیرنرمال" تفکیک کند. طبقه‌بندهای مختلفی برای این منظور به کار گرفته می‌شوند. ساختار کلی طبقه‌بندها از دو مرحله‌ی آموزش و آزمایش تشکیل یافته است. نمونه‌هایی که طبقه‌بند با آن کار می‌کند به عنوان بردار ویژگی یاد می‌شود، و شامل برداری از عوامل مورد بررسی در پژوهش می‌باشد. داده‌های مربوط به عادات زندگی و وضعیت سلامتی مجموعه‌ای از داوطلبان جمع‌آوری می‌شود. در ابتدا طبقه‌بند با استفاده از بخشی از این داده‌ها آموزش داده می‌شود، سپس طبقه‌بند آموزش دیده باید قادر به دسته‌بندی داده‌های شناخته نشده باشد. کارایی یک طبقه‌بند با میزان دقت آن در تشخیص درست داده‌ها ارزیابی می‌شود.

یکی از مشکلات عمده در طبقه‌بندی داده‌های ناباروری، وجود عدم توازن می‌باشد. مشکل هنگامی پیش می‌آید که تعداد نمونه‌های موجود در یک طبقه بسیار بیشتر از طبقه دیگر باشد. از آن جایی که تعداد افراد مبتلا به یک بیماری و یا به اصطلاح طبقه "مثبت" داده‌ها در پایگاه داده، بسیار کمتر از دسته نرمال (طبقه منفی) می‌باشد؛ این مسأله به طور کلی در اکثر داده‌های پزشکی به چشم می‌خورد. البته این مسأله باعث بهتر یادگرفتن داده‌های منفی توسط طبقه‌بند شده، و نهایتاً منجر به تشخیص غلط یک داده در کلاس مثبت به عنوان داده منفی می‌شود. روش‌های متفاوتی برای حل مسأله عدم توازن داده‌ها وجود دارد (۷-۹). نمونه‌گیری مجدد یکی از معروف‌ترین راه‌حل‌ها برای این مشکل می‌باشد. با استفاده از این روش، نمونه‌های کمتر در مجموعه‌ی داده‌ها تکرار شده، و سعی می‌شود توزیع یکسانی از هر دو گروه داده‌ایی به وجود آید.

در طبقه‌بندی داده‌های ناباروری و بررسی کیفیت اسپرم با استفاده از ویژگی‌های ذکر شده، کارهای مختلفی صورت گرفته است. داده‌های ناباروری در (۱۰) سه طبقه‌بند مختلف به نام‌های درختان تصمیم (۱۱)، پرسپترون چندلایه (۱۲ و ۱۳)، و ماشین بردار پشتیبان (۱۴ و ۱۵) برای طبقه‌بندی داده‌ها به کار گرفته

<sup>1</sup> Wang H.

<sup>2</sup> Support Vector Machine

<sup>3</sup> Decision Tree

<sup>4</sup> Artificial Neural Network

<sup>5</sup> K-nearest Neighbor

<sup>6</sup> Adaboost

<sup>7</sup> Bootstrap Aggregation

<sup>8</sup> University of Alicante

<sup>9</sup> Bootstrapping

نمونه‌گیری درون یک نمونه است. به این منظور داده‌های آموزشی با تعداد کمتر (در این جا طبقه‌ی غیر نرمال) با جایگزینی به صورت تصادفی نمونه‌برداری می‌شوند. منظور از جایگزینی این است که وقتی یک نمونه از مجموعه داده‌ها انتخاب شد، ممکن است دوباره نیز این نمونه انتخاب گردد، و در حقیقت این نمونه به مجموعه‌ی داده‌ها بازگردانده می‌شود. بنابراین ممکن است یک نمونه‌ی داده‌ای چند بار در بین

مجموعه‌ی داده‌ها تکرار شود. این عمل در به دست آمدن توزیع مناسبی از طبقه‌ی کوچک‌تر نسبت به داده‌های طبقه‌ی بزرگ‌تر کمک می‌کند. بنابراین طبقه‌بند در تشخیص داده‌های با تعداد کمتر، دچار مشکل کمتری خواهد شد. بعد از نمونه‌گیری و به دست آمدن دو طبقه داده‌ای متوازن، طبقه‌بند با استفاده از داده‌های آموزشی به گونه‌ای آموزش داده می‌شود که حداقل میزان خطا در تشخیص داده‌های آزمایشی حاصل شود.

جدول ۱- توصیف ویژگی‌های به کار رفته برای تشخیص ناباروری مردان

ویژگی	مقدار اصلی	مقدار نرمال شده
فصل انجام آزمایش	(۱ زمستان ۲ بهار ۳ تابستان ۴ پاییز)	(۰/۳۳، ۰/۳۳، ۰/۳۳، ۰/۳۳)
سن انجام پرسش	۱۸-۳۶	(۰/۰)
بیماری‌های دوران کودکی	(۱) بله (۲) خیر	(۰/۰)
تصادف یا حادثه	(۱) بله (۲) خیر	(۰/۰)
عمل جراحی	(۱) بله (۲) خیر	(۰/۰)
تب با درجه بالا در یک سال اخیر	(۱) بله- کمتر از سه ماه پیش (۲) بله- بیش از سه ماه پیش (۳) خیر	(-۰/۰، ۰/۰)
میزان مصرف الکل	(۱) چند بار در روز (۲) روزی یک بار (۳) چند بار در هفته (۴) هفته‌ای یک بار (۵) به ندرت یا اصلاً	(۰/۰)
میزان سیگار کشیدن	(۱) هرگز (۲) بعضی مواقع (۳) روزانه	(-۰/۰، ۰/۰)
میزان نشستن در طول روز	(۱) ۱-۱۶ ساعت	(۰/۰)

است. در ادامه، انواع طبقه‌بندی‌هایی که برای تشخیص داده‌های ناباروری از آن‌ها استفاده شده است، مورد بررسی قرار گرفته است. شبکه‌ی عصبی مصنوعی: شبکه‌ی عصبی مصنوعی<sup>۱۰</sup> یکی از ابزارهای قدرتمند در شناسایی الگو و داده‌کاوی می‌باشد که عملکرد آن از مغز انسان الهام گرفته شده است. شبکه‌های عصبی از تعدادی لایه با اجزایی به نام نرون تشکیل شده است. بین نرون‌ها ارتباطات وزن‌داری وجود دارد. متداول‌ترین ساختار برای یک شبکه‌ی عصبی مصنوعی شبکه‌ای با سه لایه می‌باشد. تعداد نرون‌های لایه‌ی اول برابر با تعداد ویژگی‌ها، یا ابعاد بردار ویژگی می‌باشد. لایه‌های میانی (مخفی) متناسب با مسأله دارای نرون هستند.

لایه‌ی آخر نیز به تناسب تعداد طبقه‌های داده، دارای نرون می‌باشد. در مورد داده‌های ناباروری با توجه به وجود دو طبقه، یک نرون برای طبقه‌بندی آن‌ها در لایه‌ی آخر کافی می‌باشد. به منظور یادگیری شبکه، نمونه‌های آموزشی از لایه‌ی اول به شبکه داده می‌شود و وزن‌های شبکه بر اساس این داده‌ها در هر دور تکرار فرایند یادگیری، به گونه‌ای تغییر می‌کنند که کمترین خطای کارایی در طبقه‌بندی داده‌ها رخ دهد.

**ماشین بردار پشتیبان:** ماشین بردار پشتیبان<sup>۱۱</sup> یکی از معروف‌ترین و قدرتمندترین طبقه‌بندی‌هاست که در تحقیقات مختلفی از آن استفاده می‌شود. این طبقه‌بند سعی دارد خط یا صفحه‌ی جداکننده‌ی بهینه بین داده‌ها با بیشترین حاشیه

**طبقه‌بندی:** در روش‌های مختلف داده‌کاوی، طبقه‌بندی‌ها به عنوان ابزاری برای تعیین طبقه داده‌های مختلف به کار می‌روند. در تشخیص داده‌های پزشکی، مجموعه داده‌ها به دو طبقه نرمال (منفی) و غیرنرمال (مثبت) تقسیم می‌شوند. وظیفه‌ی طبقه‌بند تشخیص طبقه‌ای است که یک بردار ویژگی به آن تعلق دارد. به این منظور در مرحله‌ی آموزش یک طبقه‌بند، ابتدا بخشی از داده‌ها به عنوان داده‌های آموزشی به طبقه‌بند داده شده تا بتواند آن‌ها را یاد بگیرد، و پس از آن قادر باشد تا داده‌هایی را که در آینده می‌بیند تشخیص دهد. بعد از مرحله‌ی آموزش، مرحله‌ی آزمایش طبقه‌بند آغاز می‌شود. در این مرحله با بخش دیگری از داده‌ها طبقه‌بند آموزش دیده، مورد آزمایش قرار می‌گیرد تا بتوان کارایی آن را ارزیابی نمود. داده‌های مورد استفاده در طبقه‌بند به صورت یک بردار از ویژگی‌ها نمایش داده می‌شود. به عنوان مثال در مورد داده‌های ناباروری، برداری شامل عوامل موثر بر کیفیت اسپرم (عوامل مورد بررسی) برای هر فرد تشکیل می‌شود.

مطابق با اصل "نهار مجانی وجود ندارد"<sup>(۲۰)</sup> هیچ الگوریتم طبقه‌بندی بهترین نیست؛ به این معنا که ممکن است برای یک مجموعه داده‌ها یک الگوریتم طبقه‌بند بهترین باشد، و برای مجموعه داده‌های دیگر الگوریتم دیگر، و یافتن بهترین الگوریتم طبقه‌بند معمولاً با سعی و خطا به دست می‌آید. در این مطالعه الگوریتم‌های طبقه‌بند مختلف مورد بررسی قرار گرفته‌اند. سپس کاراترین طبقه‌بندی‌ها به عنوان طبقه‌بند نهایی انتخاب شده

<sup>10</sup> Artificial Neural network

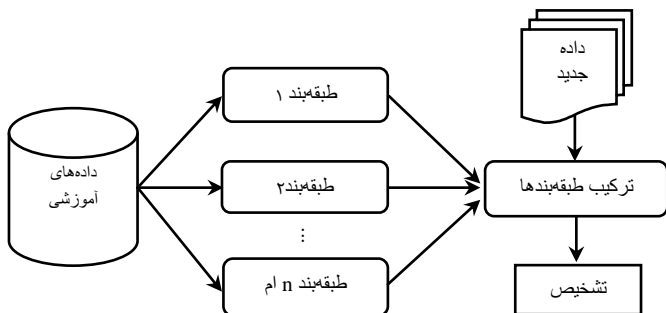
<sup>11</sup> Support Vector machine

در این شاخص منظور از عملگر  $|D|$  تعداد اعضای  $D$  می‌باشد. ویژگی با کمترین مقدار شاخص Gini می‌تواند بهترین ویژگی برای نود تصمیم گیرنده باشد.

**نزدیک‌ترین همسایگی:** طبقه‌بند  $K$  نزدیک‌ترین همسایگی<sup>۱۳</sup> بر اساس نزدیک‌ترین همسایگان یک الگو در فضای ویژگی تصمیم‌گیری می‌کند، و بعد از آرایه‌ی یک الگوی ناشناخته به صورت زیر تصمیم‌گیری می‌کند:  $k$  نزدیک‌ترین همسایگان الگوی داده شده را در فضا بدون توجه به طبقه آن‌ها مشخص می‌کند. طبقه‌ی هر یک از این الگوهای درون همسایگی تعیین می‌شوند، و الگوی داده شده به طبقه‌ای تعلق خواهد داشت که بیشترین تعداد از الگوهای همسایه متعلق به آن باشد.

**ترکیب طبقه‌بندها:** یک طبقه‌بند ممکن است روی مجموعه‌ای از داده‌ها به خوبی کارایی داشته باشد، اما روی داده‌های دیگر دچار شکست شود. بنابراین با ترکیب چندین طبقه‌بند می‌توان خطر انتخاب طبقه‌بند نامناسب را کم کرد. دو دسته‌ی کلی طبقه‌بندهای ترکیبی وجود دارد: دسته‌ی اول طبقه‌بندهای ترکیبی وزن‌دار هستند. در این روش‌ها در هر بار تکرار الگوریتم، یک طبقه‌بند پایه روی داده‌ها عمل کرده و بر اساس عملکرد آن، وزنی به داده‌ها اختصاص می‌یابد تا در تکرار بعدی، داده‌های با خطای بیشتر، وزن بیشتری برای یادگیری داشته باشند. در آخر نیز برای تصمیم‌گیری در مورد طبقه‌ی یک نمونه، رأی‌گیری وزن‌دار بین کلیه طبقه‌بندها انجام می‌گیرد. نمونه‌ای از این نوع طبقه‌بند که در آزمایشات این پژوهش مورد استفاده قرار گرفت، آدابوست نام دارد.

دسته‌ی دیگر طبقه‌بندهای ترکیبی، بدون وزن هستند. در این گروه تمامی طبقه‌بندها برای تعیین طبقه یک داده، به صورت یکسان در رأی‌گیری مشارکت می‌کنند. نمونه‌ای از این طبقه‌بندها، تجمع خودراه‌انداز می‌باشد که در آزمایشات به کار گرفته شده است. در شکل ۱ چگونگی ترکیب طبقه‌بندها نمایش داده شده است.



شکل ۱- چگونگی ترکیب طبقه‌بندها

(فاصله) در طبقه‌های مختلف را بیابد. ماشین بردار پشتیبان با نگاشت داده‌ها به فضایی بالاتر باعث می‌شود تا داده‌ها به صورت خطی از یکدیگر جدا شوند. بنابراین برای کاربردهایی که طبقه داده‌های آن به صورت خطی از یکدیگر جدایی‌پذیر نیستند، می‌تواند مناسب باشد. به این منظور از تابعی به نام هسته استفاده می‌شود تا نمونه‌ها به یک فضا با ابعاد بیشتر نگاشت شوند. انتخاب تابع هسته در کارایی ماشین بردار پشتیبان نقش به‌سزایی دارد. تابع پایه شعاعی یکی از توابع هسته است که کارایی خوبی در ماشین بردار پشتیبان دارد. این تابع به صورت زیر تعریف می‌شود:

$$K(x, y) = \exp\left(\frac{-\|x - y\|^2}{2\delta^2}\right)$$

در این تابع  $x$  و  $y$  دو بردار ویژگی در بین داده‌ها هستند، و  $\delta$  پارامتری آزاد می‌باشد. همچنین منظور از  $\| \cdot \|$  فاصله‌ی اقلیدسی بین دو نقطه  $x$  و  $y$  است.

**درخت تصمیم:** درخت تصمیم<sup>۱۴</sup> طبقه‌بندی با ساختار درخت می‌باشد. هر نود برگ در یک درخت تصمیم نشان دهنده‌ی یک طبقه است. نودهای میانی به عنوان نودهای تصمیم‌گیرنده عمل می‌کنند. در هر نود، سوآلی در خصوص یکی از ویژگی‌های بردار ویژگی می‌شود و با توجه به جواب داده شده یکی از فرزندان نود مورد نظر پیمایش خواهند شد. این روند ادامه پیدا می‌کند تا در نهایت به یکی از برگ‌های درخت رسیده که نشان‌دهنده‌ی طبقه داده مورد نظر می‌باشد. درختان تصمیم به دلیل سادگی در پیاده‌سازی در مسایل داده‌کاوی بسیار مورد توجه می‌باشند. مواردی همچون انتخاب ویژگی‌های مورد پرسش در هر نود، شرط خاتمه، و معیار شکست یک نود؛ شاخص‌هایی هستند که در ساختار یک درخت تصمیم مهم می‌باشند. یکی از معیارهای معروف برای انتخاب ویژگی مورد پرسش شاخص Gini می‌باشد که به صورت زیر تعریف می‌شود:

$$Gini(D) = 1 - \sum_{i=1}^m p_i^2$$

که در آن  $p_i$  احتمال تعلق یک نمونه به طبقه  $i$  ام و  $m$  تعداد طبقه‌ها می‌باشند. هنگامی که کلیه داده‌ها ( $D$ ) توسط ویژگی  $A$  به دو گروه  $D_1$  و  $D_2$  تقسیم می‌شوند. شاخص Gini برای ویژگی  $A$  به صورت زیر تعریف می‌شود:

$$Gini_A(D) = \frac{|D_1|}{|D|} \times Gini(D_1) + \frac{|D_2|}{|D|} \times Gini(D_2)$$

<sup>13</sup> K-nearest neighbor

<sup>14</sup> Decision tree

## روش اعتبارسنجی منقطع K لایه

در پژوهش حاضر، برای ارزیابی الگوریتم‌های طبقه‌بندی از روش اعتبارسنجی منقطع K لایه استفاده شده است (۶). در این روش کل مجموعه داده‌ها به k قسمت مساوی تقسیم می‌شود. در هر بار انجام آموزش و آزمایش طبقه‌بند، یکی از بخش‌ها به عنوان داده آزمایشی و k-1 بخش دیگر به عنوان داده آموزشی به کار گرفته می‌شوند. در نهایت با استفاده از میانگین‌گیری بین طبقه‌بندهای به دست آمده، طبقه‌بند نهایی به دست می‌آید.

**معیارهای ارزیابی:** به منظور ارزیابی کارایی طبقه‌بندها، معیارهای متفاوتی وجود دارد. این معیارها شامل دقت، AUC<sup>۱۴</sup> و ماتریس اغتشاش می‌باشد که تعریف هر یک در ادامه آورده شده است. دقت یک طبقه‌بند به معنای نرخ تشخیص آن می‌باشد که به صورت زیر تعریف می‌شود:

$$Accuracy = \frac{TP + TN}{P + N}$$

TP<sup>۱۵</sup>: تعداد نمونه‌های مثبت که صحیح تشخیص داده شده‌اند.

TN<sup>۱۶</sup>: تعداد نمونه‌های منفی که صحیح تشخیص داده شده‌اند.

FP<sup>۱۷</sup>: تعداد نمونه‌های مثبت که به اشتباه منفی تشخیص داده شده‌اند.

FN<sup>۱۸</sup>: تعداد نمونه‌های منفی که به اشتباه مثبت تشخیص داده شده‌اند.

معیار دقت در حالتی که طبقه‌ها در یک مجموعه داده متوازن باشند معیار مناسبی می‌باشد، اما در مواقعی که توزیع داده‌ها یکسان نباشد معیارهای بهتری برای ارزیابی طبقه‌بندها وجود دارد. در داده‌های پزشکی این مشکل معمولاً وجود دارد. در این پژوهش نیز تعداد داده‌های طبقه مثبت بسیار کمتر می‌باشد، در این صورت حتی اگر طبقه‌بندی کلیه داده‌ها را منفی تشخیص دهد، باز هم مقدار دقت آن عدد نسبتاً خوبی خواهد بود؛ پس این معیار نمی‌تواند به تنهایی یک طبقه‌بند را ارزیابی کند. معیار دیگری که برای ارزیابی طبقه‌بندها از آن استفاده می‌شود AUC می‌باشد. AUC به مساحت ناحیه‌ی زیر منحنی ROC<sup>۱۹</sup> گفته می‌شود. ROC، نموداری است که میزان معیار صحت (نرخ TP) یک طبقه‌بند را در مقابل فراخوانی (نرخ FP) آن نمایش می‌دهد. حداکثر مقدار معیار AUC عدد یک می‌باشد. ماتریس اغتشاش نیز یکی دیگر از معیارهایی است که اطلاعات خوبی از کارایی

یک طبقه‌بند نمایش می‌دهد. معیار مذکور یک ماتریس ۲ در ۲ می‌باشد که هر یک از عناصر آن شامل یکی از شاخص‌های گفته شده در بخش قبلی (TP, TN, FP, FN) خواهد بود.

## نتایج

برای ارزیابی کارایی روش پیشنهادی، آزمایشاتی بر روی مجموعه داده‌های توضیح داده شده در قسمت قبل (از مرکز سیستم‌های هوشمند و یادگیری ماشین دانشگاه کالیفرنیا) انجام شده و نتایج آن در این قسمت ارائه شده است. این آزمایشات در نرم‌افزار ریپد ماینر<sup>۲۰</sup> و متلب<sup>۲۱</sup> پیاده‌سازی شد. معمولاً برای مقایسه الگوریتم‌های طبقه‌بندی از روش اعتبارسنجی منقطع k لایه استفاده می‌کنند (۶).

در این روش کل مجموعه داده‌ها به k قسمت مساوی تقسیم می‌شود. در هر بار انجام آموزش و آزمایش طبقه‌بند، یکی از بخش‌ها به عنوان داده ی آزمایشی و k-1 بخش دیگر به عنوان داده‌ی آموزشی به کار گرفته می‌شوند. در نهایت با استفاده از میانگین‌گیری بین طبقه‌بندهای به دست آمده، طبقه‌بند نهایی به دست می‌آید.

**مطالعه مقایسه‌ای:** در آزمایشات انجام شده نمونه‌گیری به اندازه‌ای تکرار شد که تعداد داده‌های طبقه مثبت به نصف داده‌های طبقه منفی افزایش پیدا کند. در این حالت توزیع نسبتاً معقولی برای هر دو طبقه وجود خواهد داشت. تعداد تکرار بیشتر باعث یادگیری زیاد داده‌های آموزشی خواهد شد، که این امر منجر به عدم تشخیص داده‌های ناشناخته توسط طبقه‌بند خواهد گردید. شاخص دیگری که در آزمایشات باید تعیین شود مقدار k در روش اعتبارسنجی منقطع است. در این پژوهش این شاخص ۱۰ در نظر گرفته شد. واضح است که با بالا رفتن مقدار k تعداد نمونه‌هایی که برای مرحله‌ی آموزش استفاده می‌شود بیشتر خواهد بود. بنابراین طبقه‌بند کارایی بهتری را از خود نشان می‌دهد.

جدول ۲ ماتریس اغتشاش را برای هر یک از طبقه‌بندها نشان می‌دهد. همان طور که در جدول دیده می‌شود ماشین بردار پشتیبان تنها یک داده‌ی منفی و سه داده‌ی مثبت را به اشتباه تشخیص داده است، و بقیه‌ی داده‌ها با استفاده از این طبقه‌بند به درستی تشخیص داده شده‌اند.

<sup>14</sup> Area Under the Curve

<sup>15</sup> True Positive

<sup>16</sup> True Negative

<sup>17</sup> False Positive

<sup>18</sup> False Negative

<sup>19</sup> Receiver Operating Characteristic

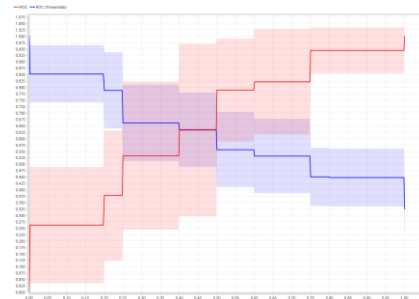
<sup>20</sup> Rapidminer

<sup>21</sup> Matlab

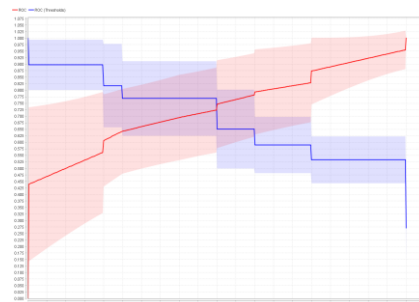
جدول ۲- ماتریس اغتشاش طبقه‌بندی‌های مختلف

طبقه‌بند	واقعی مثبت	واقعی منفی
ماشین بردار پشتیبان	۴۱	۲
پیش‌بینی مثبت	۳	۸۶
پیش‌بینی منفی	۱۹	۳
درخت تصمیم	۲۵	۸۵
پیش‌بینی مثبت	۴۱	۷
پیش‌بینی منفی	۳	۸۱
شبکه عصبی	۲۹	۹
پیش‌بینی مثبت	۱۵	۷۹
پیش‌بینی منفی	۳۷	۲
تجمیع خودراه‌انداز	۷	۸۶
پیش‌بینی مثبت	۳۸	۸
پیش‌بینی منفی	۶	۸۰

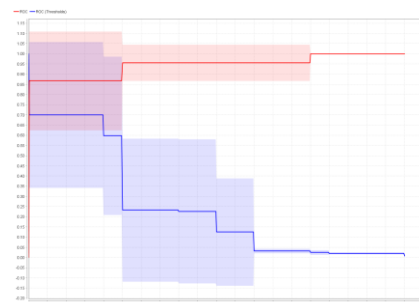
نمودار ROC مربوط به طبقه‌بندی‌های مختلف در شکل ۲ نشان داده شده است. همان طور که گفته شد سطح زیر این منحنی نشان دهنده AUC می‌باشد که هر چقدر به ۱ نزدیک‌تر باشد، نشانگر کارایی بالاتر طبقه‌بند می‌باشد.



شبکه عصبی



آدا بوست



تجمیع خودراه‌انداز

شکل ۲- نمودار ROC مربوط به طبقه‌بندی‌های مختلف (محور افقی نرخ FP و محور عمودی نرخ TP را نشان می‌دهد)

جدول ۳ مقدار دقت و AUC را برای تمام طبقه‌بندها نشان می‌دهد. طبقه‌بند ماشین بردار پشتیبان بالاترین مقدار دقت را نسبت به دیگر طبقه‌بندها دارد. اما به طور کلی طبقه‌بندی‌های درخت تصمیم و شبکه‌ی عصبی نیز توانسته‌اند نتایج قابل قبولی را ارائه کنند. نکته‌ی قابل توجه در مورد طبقه‌بندی‌های ترکیبی آن است که هر چند این طبقه‌بندها سرعت اجرای کمتری نسبت به طبقه‌بندی‌های غیرترکیبی دارند، اما در مقایسه با طبقه‌بند پایه‌ی خود از دقت تقریباً یکسانی برخوردار هستند. به همین دلیل استفاده از آن‌ها در این کاربرد مقرون به صرفه نمی‌باشد.

مانند معیار دقت، ماشین بردار پشتیبان از مقدار AUC بالاتری برابر با ۰/۹۸ نیز برخوردار است. همان طور که گفته شد در خصوص داده‌هایی مانند ناباروری که دارای مشکل عدم توازن داده‌ای هستند، معیار AUC نسبت به دقت ملاک معتبرتری می‌باشد. همچنین در ارزیابی طبقه‌بندها با استفاده از این معیار، اختلاف AUC برای ماشین بردار پشتیبان نسبت به طبقه‌بندی‌های دیگر قابل توجه می‌باشد.

جدول ۳- مقادیر دقت و AUC برای طبقه‌بندی‌های مختلف

طبقه‌بند	مقدار دقت طبقه‌بند (درصد)	مقدار AUC طبقه‌بند
ماشین بردار پشتیبان	۹۵/۱۵	۰/۹۸
درخت تصمیم	۸۳/۴۱	۰/۷۵
نزدیک‌ترین همسایگی	۹۲/۳۱	۰/۵۰
شبکه عصبی	۸۱/۷۰	۰/۶۶
تجمیع خودراه‌انداز	۹۳/۰۸	۰/۹۳
آدا بوست	۸۹/۳۶	۰/۸۹

برای ارزیابی بهتر، جدول ۴ روش پیشنهادی را با سایر روش‌ها در ادبیات موضوع مقایسه کرده است. همان طور که مشاهده می‌شود در تمام موارد روش پیشنهادی توانسته است مقدار دقت و AUC بهتری را ارائه نماید که نشانگر کارایی بهتری الگوریتم پیشنهادی نسبت به سایر روش‌ها است.

جدول ۴- مقایسه روش پیشنهادی با سایر روش‌های ارائه شده

مرجع	مقدار دقت طبقه‌بند	مقدار AUC طبقه‌بند
(۱۰)	٪۸۶	-
(۱۶)	٪۸۲	-
(۱۷)	-	۰/۸۰۵۱
روش پیشنهادی	٪۹۵	۰/۹۸

## بحث

سبک زندگی افراد تأثیر زیادی بر روی نحوه کارکرد بدن انسان دارد. عواملی مثل سیگار کشیدن، مصرف الکل، و یا حتی نحوه نشستن می‌تواند باعث ایجاد مشکلاتی در افراد شود. هدف از این مطالعه بررسی این عوامل جهت تشخیص ناباروری مردان بوده است. نتایج این مطالعه نشان داد که عوامل محیطی و سبک زندگی تأثیر مستقیمی بر روی ناباروری مردان دارد. این

تابستان ۹۵، دوره نوزدهم، شماره دوم، پیاپی ۷۳

همچنین حل مشکل عدم توازن داده‌های پزشکی به دلیل وجود تعداد بسیار بیشتر داده‌های مثبت نسبت به داده‌های منفی با استفاده از روش نمونه‌گیری خودرانداز مورد توجه قرار گرفت. آزمایشات انجام شده نشان داد که طبقه‌بند ماشین بردار پشتیبان از کارایی خوبی برای تشخیص ناباروری مردان از روی عوامل محیطی برخوردار است.

## کاربرد در تصمیم‌های مرتبط با سیاست‌گذاری در نظام سلامت

افزایش ناباروری یکی از موضوعاتی است که میزان آن در ایران در حال افزایش است. تجربیات و شواهد نشان می‌دهد که روش‌های هوش مصنوعی در کشورهای مختلف دنیا روز به روز جهت کنترل و پیشگیری بیماری‌ها گسترش می‌یابند. همچنین به دلیل گسترش دانش و پیچیده‌تر شدن تصمیم‌گیری، سیستم‌های هوش مصنوعی برای تصمیم‌گیران نظام سلامت بسیار مفید و با اهمیت ارزیابی می‌شوند. مطالعه‌ی حاضر به بررسی عوامل محیطی و سبک زندگی در تشخیص ناباروری پرداخت. نشان داد که با استفاده از سبکی که یک فرد زندگی می‌کند می‌توان ناباروری را تشخیص داد. به عبارت دیگر با تغییر سبک زندگی می‌توان به افزایش کنترل ناباروری کمک شایانی نمود. مخاطب اصلی این پژوهش شورای معاونین وزارت بهداشت، درمان و آموزش پزشکی؛ اعضای کمیسیون بهداشت و درمان مجلس شورای اسلامی؛ و مراکز تحقیقاتی ناباروری هستند تا با تلاش در تغییر سبک زندگی به کنترل این امر کمک نمایند.

## تشکر و قدردانی

این مطالعه با حمایت مالی دانشگاه کاشان به انجام رسیده است، بنابراین نویسندگان بر خود لازم می‌دانند تا از مدیریت محترم دانشگاه جهت حمایت‌های مالی تقدیر و تشکر نمایند.

## References

- 1- Splingart C, Frapsauce C, Veau S, Barthelemy C, Royère D, Guérif F. Semen Variation in a Population of Fertile Donors: Evaluation in a French Center Over a 34 Year Period. *International Journal of Andrology* 2012; 35(3):467-74.
- 2- Giwercman A, Giwercman YL. Environmental Factors and Testicular Function. *Best Practice & Research Clinical Endocrinology & Metabolism* 2011; 25(2):391-402.
- 3- Martini AC, Molina RI, Estofán D, Senestrari D, De Cuneo MF, Ruiz RD. Effects of Alcohol and Cigarette Consumption on Human Seminal Quality. *Fertility and Sterility* 2004; 82(2):374-7.
- 4- Biglarian A, Hajizadeh E, Kazemnejad A. Comparison of Artificial Neural Network and Cox Regression Models in Survival Prediction of Gastric Cancer Patients. *koomesh* 2010;11(3):215-20.
- 5- Montazeri M, Montazeri M. Machine Learning Models for Predicting the Diagnosis of Liver Disease. *koomesh* 2014; 16(1):53-9.

نتایج قبلاً در بعضی مطالعات به اثبات رسیده است ولی استفاده از روش‌های موجود خطی برای پیش‌بینی کار مشکلی است و بنابراین نیاز به استفاده از مدل‌هایی است که بتواند الگو و رابطه بین عوامل محیطی و ناباروری را به دست آورد. به دلیل این که لزوماً مجموعه مشخصه‌های مربوط به عوامل محیطی خطی نمی‌باشند بنابراین مدل‌های غیرخطی به خصوص روش‌های هوش مصنوعی مثل شبکه‌های عصبی و ماشین بردار پشتیبان در این زمینه بسیار قوی‌تر از مدل‌های خطی عمل می‌کنند و از این رو قادر هستند تا دقت تشخیص را افزایش داده و امکان برنامه‌ریزی‌های لازم را فراهم نمایند.

همچنین این مطالعه نشان داد که طبقه‌بندی ماشین بردار پشتیبان AUC بهتری نسبت به سایر روش‌ها دارد هر چند که طبقه‌بند‌های درخت تصمیم و شبکه عصبی نیز AUC قابل قبولی از خود ارایه دادند. همچنین از دیدگاه متریک دقت نیز ماشین بردار پشتیبان بالاترین دقت را ارایه نمود. نکته دیگر این است که هر چند طبقه‌بند‌های ترکیبی نتایجی مشابه با ماشین بردار پشتیبان ارایه نمودند اما به دلیل ترکیبی بودن، در حالت کلی سرعت اجرای کمتری نسبت به طبقه‌بندی‌های پایه دارند بنابراین در اینجا ماشین بردار پشتیبان ترجیح داده شده است. با توجه به مطالب مذکور، مشخص است که مطالعه حاضر در گام‌های اولیه در زمینه تشخیص ناباروری مردان با استفاده از عوامل محیطی و سبک زندگی می‌باشد و لازم است مطالعات آینده تکمیل کننده آن باشد تا بتوان به نتایج قطعی‌تر دست یافت.

## نتیجه‌گیری

در این پژوهش به بررسی استفاده از روش‌های داده‌کاوی در تشخیص ناباروری مردان از روی عوامل محیطی و سبک زندگی افراد پرداخته شد. عاداتی همچون سیگار کشیدن، نشستن زیاد در طول روز و صدمات جسمی در اثر حوادث مختلف می‌توانند بر کیفیت اسپرم اثرگذار باشند. هدف این پژوهش تشخیص این تأثیر با استفاده از طبقه‌بند‌های مختلف و مقایسه آن‌ها بود.

- 6- Mousavirad SJ, Ebrahimpour Komleh H. A New Intelligent Hepatitis Diagnosis Using Principal Component Analysis and Classifiers Fusion. *koomesh* 2015;16(2):149-58.
- 7- Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. A Review on Ensembles for the Class Imbalance Problem: Bagging, Boosting, and Hybrid-based Approaches. *IEEE Transactions on Systems, Man, and Cybernetics* 2012; 42(4):463-84.
- 8- Ganganwar V. An Overview of Classification Algorithms for Imbalanced Datasets. *International Journal of Emerging Technology and Advanced Engineering* 2012; 2(4):42-7.
- 9- He H, Garcia E. Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering* 2009; 21(9):1263-84.
- 10- Gil D, Girela JL, De Juan J, Gomez-Torres MJ, Johnsson M. Predicting Seminal Quality with Artificial Intelligence Methods. *Expert Systems with Applications* 2012; 39(16):12564-73.
- 11- Polat K, Güneş S. A Novel Hybrid Intelligent Method Based on C4. 5 Decision Tree Classifier and One-against-all Approach for Multi-class Classification Problems. *Expert Systems with Applications* 2009; 36(2):1587-92.
- 12- Mehrotra K, Mohan CK, Ranka S. *Elements of Artificial Neural Networks*. USA: MIT press; 1997:457.
- 13- Ripley BD. *Pattern Recognition and Neural Networks*. England: Cambridge University Press; 1996:368.
- 14- Conforti D, Guido R. Kernel Based Support Vector Machine via Semi Definite Programming: Application to Medical Diagnosis. *Computers & Operations Research* 2010; 37(8):1389-94.
- 15- Vapnik VN, Kotz S. *Estimation of Dependences Based on Empirical Data*. USA: Springer-verlag; 1982:785.
- 16- Girela JL, Gil D, Johnsson M, Gomez-Torres MJ, De Juan J. Semen Parameters Can be Predicted from Environmental Factors and Lifestyle Using Artificial Intelligence Methods. *Biology of Reproduction* 2013; 88(4):99.
- 17- Wang H, Xu Q, Zhou L. Seminal Quality Prediction Using Clustering-Based Decision Forests. *Algorithms* 2014; 7(3):405-17.
- 18- Bishop CM. *Pattern Recognition and Machine Learning*. USA: Springer; 2006:349.
- 19- Blake C, Merz C. UCI Repository of Machine Learning Databases, [Cited 2016 June 20], Available from: <https://archive.ics.uci.edu/ml/datasets/Fertility>
- 20- Wolpert DH, Macready WG. No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation* 1997; 1(1):67-82.



## Male Infertility Prediction from Environmental Factors and Lifestyle Using Artificial Intelligence Algorithms

Asilian Bidgoli A (PhD), Ebrahimpour-Komleh H (PhD), Mousavirad SJ\* (PhD)

*Department of Computer Engineering, University of Kashan, Kashan, Iran*

**Original Article**

**Received: 23 Jan 2016, Accepted: 3 May 2016**

### Abstract

**Introduction:** Infertility is one of the most important problems, especially among males which has received special attention recently. Male infertility can be affected by different factors. There is now a large body of evidence to support the effects of life styles and environmental factors on semen quality. Data mining methods in artificial intelligence, as a decision support system, could be helpful in medical diagnosis of male infertility.

**Methods:** The influencing factors of seminal quality, and as a result ability to detect the infertility in males was assessed in this study using data mining algorithms in artificial intelligence. The dataset of 100 volunteers among students of University of Alicante were used. After data balancing different classifiers such as neural network were used to find the best classifier to predict the male infertility.

**Results:** The proposed algorithm is evaluated with different data mining algorithms to find a better analytical application of the algorithm. Support vector machine had the best accuracy (%95.15) to predict proposed infertility in male compared to the other classifiers. The proposed algorithm has a competitive accuracy compared to other algorithms.

**Conclusion:** The proposed algorithm is able to predict male infertility from lifestyle and environmental factors using data mining algorithms.

**Key words:** Male Infertility, Artificial intelligence, Data mining algorithms, Support vector machine, Support decision system

---

### Please cite this article as follows:

Asilian Bidgoli A, Ebrahimpour-Komleh H, Mousavirad SJ. Male Infertility Prediction from Environmental Factors and Lifestyle Using Artificial Intelligence Algorithms. *Hakim Health Sys Res* 2016; 19(2): 72- 80.

---

---

\*Corresponding Author: Department of Computer Engineering, University of Kashan, Kashan, Iran.  
Tel: +98- 935- 4626334, E-mail: [jalalmoosavirad@gmail.com](mailto:jalalmoosavirad@gmail.com)