مجله دانشگاه علوم پزشکی تربت حیدریه
دوره ۱۱. شماره ۲. تابستان ۱۴۰۲

مقاله پژوهشی

تشخیص اختلال طیف اوتیسم با بهره‌گیری از تحلیل مولفه های اصلی جهت استخراج بهترین ویژگی‌ها با استفاده از شبکه‌های عصبی مصنوعی

حبیبه هوشمند[۱]، نجمه خردنیا[۲]، عباس قدرتی[۳]، آتنا عبیدی[۴*]

۱. دانش آموخته کارشناسی، دانشکده فنی مهندسی، دانشگاه تربت حیدریه
۲. دانش آموخته کارشناسی، دانشکده فنی مهندسی، دانشگاه تربت حیدریه
۳. گروه پرستاری، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی تربت حیدریه،ایران
۴. گروه کامپیوتر واحد بوشهر، دانشگاه آزاد اسلامی، بوشهر، ایران

چکیده

زمینه و هدف: طیف اوتیسم یکی از اختلالات روان‌شناختی کودکان محسوب می‌شود. تشخیص به‌موقع و با دقت این اختلال، اهمیت فراوانی در تأمین مراقبت و درمان مناسب کودکان دارد. هدف اصلی این تحقیق، تأکید بر اهمیت ویژگی‌های مرتبط با بیماری اوتیسم و تشخیص آن با استفاده از یک مدل هوشمند است، چراکه برخی از این ویژگی‌ها از درجه اولویت بالاتری برخوردارند.

روش‌ها: بدین منظور، از روش تحلیل مولفه اصلی(PCA)[۱] برای اولویت‌بندی ویژگی‌ها استفاده شد و پس از استخراج ویژگی های بهینه، با استفاده از شبکه های عصبی مصنوعی، به تشخیص خودکار بیماری پرداخته شده است.

نتایج: داده‌های مورد استفاده در این مطالعه از مجموعه‌داده Kaggle جمع‌آوری شده‌اندکه شامل ۱۰۵۴ فرد بوده، که از این تعداد ۷۲۸ نفر مبتلا به اوتیسم و ۳۲۶ نفر سالم بوده‌اند. بررسی‌های این مطالعه نشان می دهد که حذف تدریجی ویژگی‌ها و تقلیل از ۱۸ به ۱۲ ویژگی، می‌تواند به حصول همان دقت در تشخیص طیف اوتسیم با استفاده از شبکه‌های عصبی مصنوعی، منجر شود.

نتیجه‌گیری: کاهش تعداد ویژگی‌ها در مدل‌های هوش مصنوعی برای تشخیص اوتیسم، ضمن کمک به بهبود و بهینه‌سازی فرآیند تشخیص بیماری، می تواند منجر به کاهش استرس والدین و حفظ حریم خصوصی آنها بدلیل تعداد کمتر سوالات شده و در نهایت منجر به تولید مدل‌هایی با عملکرد بهتر و تفسیرپذیرتر شود.

کلید واژه‌ها:
اختلال طیف اوتیسم، شبکه عصبی مصنوعی، تحلیل مولفه اصلی، استخراج ویژگی

تمامی حقوق نشر برای دانشگاه علوم پزشکی تربت حیدریه محفوظ است.

1.Principal Component Analysis

*آدرس نویسنده مسئول: بوشهر، دانشگاه آزاد اسلامی، گروه کامپیوتر واحد بوشهر
آدرس پست الکترونیک: atena.abidi@gmail.com

مقدمه

اختلال طیف اوتیسم (ASD)[2] یک اختلال توسعه‌ای عصبی است که شیوه ارتباط و تعامل افراد با دیگران، همچنین رفتار و یادگیری آنها را تحت تأثیر قرار می‌دهد(۱). علائم و نشانه‌ها در زمان کودکی ظاهر می‌شوند. این شرایط برای تمام زندگی ادامه دارند و به طور کامل قابل درمان نیستند(۲). شکاف بین مهارت‌های عملکردی ذهنی و انطباقی این افراد اغلب زیاد است به طوری که بسیاری از افراد مبتلا به ASD دچار اختلال ذهنی و یا اختلال زبانی هستند. بنابراین تشخیص اوتیسم در سنین پایین بسیار حائز اهمیت است(۳). با توجه به اهمیت رفتار انطباقی برای کارکرد روزمره در خانه و مدرسه و نیز هزینه‌های اجتماعی و مراقبت در طول عمر این افراد، اهمیت تشخیص زودهنگام و مداخله در اوتیسم آشکار می‌شود و بدین وسیله بهبود رفتار شناختی، تطبیقی و کاهش شدت اوتیسم صورت می‌پذیرد(٤). درسال‌های اخیر روش شبکه‌های عصبی مصنوعی(Artifical Neural Network ANNs) در زمینه‌های مختلف علوم پزشکی و روانشناسی از جمله پیش بینی، تشخیص و طبقه‌بندی بیماری‌ها مورد استفاده قرار گرفته است(٥). در حال حاضر سیستم‌های هوشمند در برخی از حوزه‌های پزشکی استفاده می‌شوند و به عنوان ابزار کمکی برای افراد خبره در تشخیص، تعیین نیازمندی‌ها و برطرف نمودن مشکلات با حداقل میزان خطا بکار گرفته می‌شود به طوری که شاهد بهبود تصمیم‌گیری افراد متخصص در زمینه‌های گوناگون بوده‌ایم(٦).برخی از محققان از ANNsبرای تشخیص بیماری سرطان پستان با استفاده از معماری های مختلف شبکه عصبی، اقدام کردند(۷). Bennett و همکاران در پژوهش خود ثابت کردند که استفاده از شبیه سازی هوش مصنوعی می‌تواند تصمیمات بهینه‌ای را حتی در محیط‌های پیچیده و غیر اطمینان اتخاذ کند(۸). Alvarez و همکاران در پژوهشی با عنوان مفید بودن ANNs در تشخیص و درمان سندروم آپنه-هیپوپنه (Apnea-Hypopnea) از روش شبکه‌های عصبی چند لایه (perceptron Mmultilayer) استفاده کردند(۹). بعدها Thabtah و همکاران یک برنامه کاربردی تلفن همراه به نام ASDTests را براساس ابزار Q-CHAT و AQ-10 توسعه دادند که در تشخیص زود هنگام ASD به کار می‌رود. آن‌ها همچنین داده‌های ASD را با استفاده از این برنامه جمع آوری کردند و این داده‌ها را به عنوان مجموع داده منبع باز در Kaggle و مخزن یادگیری ماشین دانشگاه کالیفرنیا (UCI)ذخیره کردند(۱۰).مطالعات متعددی انجام شده است که سعی در تشخیص و پیش بینی اختلال ASD در کودکان با استفاده از روش‌های یادگیری ماشین دارند. Thabtah و Peebles یک الگوریتم بر مبنای قوانین ML(RML) برای ارزیابی صفات ASD پیشنهاد داده اند و مشاهده شده است که این الگوریتم به طبقه بندی‌ها کمک می‌کند که عملکرد خود را بهبود بخشد(۱۱).همچنین، مطالعه دیگری توسط Thabtah و همکاران نیز انجام شده است که در آن از یک روش هوش محاسباتی (CI) به نام تجزیه و تحلیل متغیر (VA) استفاده شد و همبستگی بین مولفه های مختلف هوش محاسباتی نسبت به یکدیگر را نشان دادندو در آن از ماشین بردار پشتیبان (SVM)، درخت تصمیم (DT) و رگرسیون لجستیک (LR) جهت پیش بینی و تشخیص اختلال طیف اوتیسم استفاده گردید(۱۲،۱۳). Duda و همکاران داده‌های ASD را با طبقه بندی‌های مختلف تجزیه و تحلیل کردند و دریافتند که ٥ مورد از ٦٥ ویژگی کافی بودند تا ASD را از اختلال کمبود توجه و (ADHD) تمایز دهد(١٤).

در این پژوهش تلاش شده است تا با استخراج ویژگی‌های مهم اوتیسم به کمک الگوریتم تحلیل مولفه اصلی، تشخیص اختلال طیف اوتیسم به وسیله شبکه عصبی MLP انجام شود. از آنجایی که برخی از این ویژگی‌ها توسط آزمون و برخی دیگری توسط اندازه‌گیری ثبت می‌شود لذا می‌توان با حجم کمتری از سوالات و ویژگی‌ها به نتایجی با دقت مطلوب رسید. با استفاده از روش‌های استخراج و انتخاب ویژگی‌ها، ویژگی‌های مناسب برای تشخیص

1. Autism Spectrum Disorder[2]

اوتیسم مشخص شده و توسط مدل شبکه عصبی MLP استفاده شده‌اند.

روش‌ها

۱) داده ها

در حال حاضر، یادگیری ماشین به منظور تشخیص بیماری‌های مختلف از جمله افسردگی و اختلال طیف اوتیسم (ASD) استفاده شده است. هدف اصلی استفاده از تکنیک‌های یادگیری ماشین، بهبود دقت تشخیص و کاهش زمان تشخیص یک بیماری است تا دسترسی سریعتر به خدمات بهداشتی فراهم شود. با توجه به اینکه فرآیند تشخیص یک بیمار شامل پیدا کردن کلاس مناسب ASD و No-ASD بر اساس ویژگی‌های ورودی مورد استفاده قرار گرفته، این فرآیند می‌تواند به عنوان فرم طبقه‌بندی در یادگیری ماشین مورد استفاده قرار گیرد. می‌توانیم تحقیقات شناسایی ASD را در دو پایگاه داده "تحقیقات بر مبنای کلیپ ویدیویی" و "تحقیقات بر مبنای پرسشنامه AQ " به شرح ذیل دسته‌بندی کنیم (۱۵).

۲) تحقیقات بر اساس کلیپ ویدیویی

محققان Tariq و همکارانشان فرضیه را بر این اساس بیان کرده‌اند که استفاده از تحلیل یادگیری ماشین بر روی ویدیوهای خانگی می‌تواند زمان تشخیص را افزایش دهد. این پژوهشگران با تحلیل سطح آیتم از دو ابزار تشخیص استاندارد، طبقه‌بندی‌های یادگیری ما شین را برای پراکندگی، قابلیت تفسیر و دقت بهینه شده، ارائه داده‌اند. آن‌ها هشت مدل یادگیری ماشین را برای ا ستفاده در ۱۶۲ فیلم خانگی دو دقیقه‌ای از کودکان با و بدون ASD مورد بررسی قرار دادند. علاوه بر این، یک پورتال وب موبایل برای ارزیابی ۳۰ ویژگی رفتاری (مانند تماس چشم، لبخند اجتماعی و غیره) که توسط هشت مدل یادگیری ماشین مختلف برای شناسایی ASD استفاده می‌شود، ایجاد شده است. نتایج نشان داد که با استفاده از آزمون اعتبارسنجی متقابل و اعتبار سنجی مستقل پیشین، دقت ۹۴٪ برای هر مورد به د ست آمده است. با این حال، این روش نیز زمان بر است زیرا ویدیو برای امتیاز گذاری بر اساس ۳۰ سوال باید ضبط و ارزیابی شود (۱۶).

۳) تحقیقات بر اساس پرسشنامه AQ

" شاخص طیف اوتیسم (AQ) "یک روش انتخاب شده توسط Baron-Cohen و همکارانش برای غربالگریASD است (۱۷).بعدها، Allison و همکاران با استفاده از رویکرد شاخص تمایز (DI) تعداد موارد ابزارهای Q-CHAT و AQ را به ۱۰ در مقابل ۵۰ کاهش دادند. آنها را به پنج حوزه شامل توجه به جزئیات، توجه به تغییرات، ارتباط، تخیل و مهارت های اجتماعی تقسیم کردند(۱۸).Thabtah و همکاران به منظور بهبود رفتار شناختی، تطبیقی و کاهش شدت اوتیسم برنامه ASDTests را توسعه دادند که از ابزارهایی برای غربالگری و شناسایی عوامل خطر اختلال اسپکتروم اوتیسم استفاده می‌شود. این برنامه یک امتیاز محاسبه می‌کند که در بازه ۰ تا ۱۰ قرار دارد و امتیاز فرد اگر بیشتر از ٦ باشد پیش بینی اختلال اسپکتروم اوتیسم مثبت را نشان می‌دهد. دیتا ست مورد استفاده در این پژوهش شامل ۱۰۵٤ رکورد اطلاعات مندرج در Kaggle و سایت UCI ML می‌باشد. این دیتا شامل اطلاعاتی هستند که با استفاده از ابزار ASDTests جمع‌آوری شده‌اند (۱۹،۲۰). از این تعداد ۷۲۸ نفر مبتلا به اوتیسم و ۳۲٦ نفر سالم بوده‌اند. جدول ۱ و ۲ توضیح کوتاهی از ویژگی‌های مختلف مجموعه داده‌های استفاده شده در این مطالعه را نشان می‌دهند.

ویژگی‌ها:

برای تشخیص اوتیسم از مجموعه داده‌های Q-Chat-10 استفاده شد. این مجموعه داده شامل ده سوال است که در آنها به سوالات احتمالی پاسخ داده می شود. پاسخ‌های ممکن برای هر سوال ("همیشه"، "معمولاً"، "گاهی اوقات"، "به ندرت" و "هرگز") به مقادیر "۱" و "۰" در نظر گرفته می‌شوند. برای سوالات A1 تا A9 ، اگر پاسخ "گاهی اوقات"، "به ندرت" یا "هرگز" باشد، مقدار "۱" به آن سوال اختصاص داده می شود. اما برای سوالA10 ، اگر پاسخ "همیشه"، "معمولاً" یا "گاهی اوقات" باشد، مقدار "۱" به آن سوال اختصاص داده می شود.

در نهایت، برای هر فرد، امتیاز کلی با جمع امتیازهای ده سؤال محاسبه می‌شود. علاوه بر این، ویژگی‌های دیگر در مجموعه داده‌ها از صفحه "ارسال" در برنامه غربالگریASDTests جمع‌آوری می شوند. متغیر کلاس بر ا ساس امتیازی که فرد در طی فرآیند غربالگری با ا ستفاده از اپلیکیشن ASDTests کسب می‌کند، به صورت خودکار به هر فرد اختصاص داده می‌شود.

پردازش و تجزیه و تحلیل داده ها

همانطور که جدول۱ نشان می‌دهد دیتاست‌های ASD مورد استفاده به طور عمده شامل ۱۸ ویژگی هستند(۲۱). تمام دیتاست‌ها دارای ده ویژگی باینری هستند که نشان‌دهنده پاسخ به سوالات (A1 تا A10) و متغیرهای دسته‌ای مانند جنسیت، نژاد، زردی، ASD مثبت در خانواده، سن، کلاس ASD، امتیاز ASD و تکمیل کننده آزمون هستند. مقدار کلاس در فرآیند جمع‌آوری داده‌ها بر اساس پاسخ‌های سوالات A1 تا A10 -AQ) (10 اختصاص داده شده است. مقدار کلاس"No" نمره نهایی روش AQ-10 کمتر یا مساوی ۷ باشد. در غیر این صورت، اختصاص داده می شود "Yes" که نشان می‌دهد که فرد دارای اختلال اسپکترم اوتیسم است. برای پردازش اطلاعات به روش PCA همه داده ها به عدد تبدیل شد تا بتوان پردازش درستی روی داده‌ها انجام داد.

مرحله ۱: استخراج و انتخاب ویژگی‌ها

در این مرحله، ما با هدف تشخیص اوتیسم، ۱۸ ویژگی مختلف را بررسی کرده‌ایم. این ویژگی‌ها شامل پاسخ به ۱۰ سوال و پارامترهای وضعیت فرد نظیر سن، جنسیت، نژاد، سابقه خانوادگی و امتیاز ۱۰ سوال متناسب با اوتیسم هستند. همچنین، وجود یرقان هنگام تولد و اطلاعات آزمون دهنده نیز مورد بررسی قرار گرفته است. در مرحله بعد، ۱۸ ویژگی را در ۲ دیتاست ایجاد کردیم. یک دیتاست شامل ۱۷ ویژگی مختلف است و دیتاست دیگر وضعیت اوتیسم فرد می‌باشد. با ترکیب داده‌های این دو دیتاست، ما قادر بودیم تا عملکرد سیستم طراحی شده را در تشخیص اوتیسم ارزیابی نماییم.

مرحله ۲: پیش پردازش داده ها

در این مرحله ، پیش پردازش دقیق روی داده های مورد استفاده انجام شده و سپس به تحلیل و طبقه بندی دقیق می‌پردازیم. دیتاست‌های ASD شامل چندین رکورد با مقادیر نامعلوم هستند. هم چنین برخی از ویژگی ها نشان دهنده اطلاعات اضافه است (کدام فرد آزمون را تکمیل کرده است)و به ASD مربوط نیستند. بنابراین، قبل از اعمال طبقه‌بندی، دیتاست‌ها باید پیش پردازش شوند و داده ها را می‌توان با کدهای خاصی تنظیم کرد تا بتوان نتیجه مطلوبی به دست آورد.

مرحله ۳: اعمال روش PCA

در این پژوهش روش PCA برای تجزیه و تحلیل کاهش ابعاد به کار برده شده است، تا داده‌ها به صورتی بررسی و تحلیل شود که الگوها و ارتباط‌های مهم در آنها ظاهر شوند. در این مرحله، ابتدا داده‌ها جمع‌آوری شد و سپس با استفاده از روش PCA، ابعاد داده‌ها را کاهش داده و بر ا ساس مجموعه‌ای از مولفه‌های اصلی تشکیل دهنده آنها، تحلیل صورت گرفته است. این روش، که در اصل به جمع‌آوری و بررسی اطلاعات اصلی داده‌ها تمرکز دارد، به ما این اجازه را می‌دهد تا با یک نگاه سریع، جزئیات مهم و ارتباطات بین داده‌ها را شناسایی کنیم. اطلاعاتی که از این تحلیل نتیجه گرفته شد، تعبیری قابل توجه و مفید از الگوها و روابط در داده‌ها ارائه می‌دهد. بر اساس این اطلاعات، ما قادر به شناسایی داده‌هایی هستیم که در ارتباط با یکدیگر بیشترین تراکم و ارتباط را دارند. این به آن معنا است که این داده‌ها می‌توانند نقش مهمی در تشخیص و درک اوتیسم ایفا کنند.

مرحله ۴: تجزیه و تحلیل عملکرد

با استفاده از نرم افزار SPSS و روش PCA بر روی داده هایی که شامل ۱۰ سوال مرتبط با الگوهای اوتیسم و امتیازهایی که به‌دست می‌آید، همراه با مشخصات فردی مانند سن، جنسیت، نژاد و سابقه خانوادگی، وجود یرقان هنگام تولد، تکمیل کننده آزمون تحلیل انجام شده است. از روش تحلیل ماتریس همبستگی (correlation matrix)، که یک ابزار آماری برای نشان دادن اینکه دو یا چند متغیر چقدر قوی و در چه جهتی با هم مرتبط

هستند،استفاده شد و نتایج نشان داد که پارامتر تکمیل کننده آزمون، یرقان، سابقه خانوادگی، نژاد و سن با بیش از ٤ عدد منفی در سطر و ستون، رابطه منفی قوی‌ای با هم دارند.به عبارت دیگر تغییر در این متغیرها ممکن است به تغییراتی منفی در نتایج آزمون و امتیازها منجر شود. با توجه به نتایج، تنها متغیری که بدون رابطه منفی است، امتیاز کسب شده است. این نتیجه مشخص می‌کند که امتیاز کسب شده در آزمون بیشترین تأثیر را بر دیگر متغیرها دارد و بر مبنای آن می‌توان پیش‌بینی کرد. جدول Communities نشان می‌دهد که امتیاز به دست آمده از پاسخ به ۱۰ سوال تنها پارامتر نزدیک به یک است. در تحلیل PCA معیار Total Variance Explained معمولا به صورت درصدی است که نشان می‌دهد که تاثیرات روش PCA چه میزان از تغییرات وارد شده در داده ها را تفسیر می‌کند. در این معیار نمونه‌های مورد بررسی ٥ عامل مقدار بیشتر از یک دارند و این ٥ عامل ٥٥ درصد واریانس را به خودشان اختصاص دادند. در تحلیل ها این نتایج حاصل می‌شود که در روش PCA مقدار واریانس هر متغیر از واریانس کل به چه صورت بوده و امتیاز به خوبی توسط عوامل اصلی تفسیر می‌شود. جدول۳ وضعیت ماتریس همبستگی را نشان میدهد.

جدول٤ ضرایب فاکتورهای اصلی را نمایش می‌دهد و هر فاکتور میزان تاثیر متغیرهای مختلف را نشان می‌دهد. هر سطر نشان دهنده یک متغیر ورودی است و هر ستون نشان دهنده جزئیات PCA است مقادیر در هر سلول جدول نشان می‌دهد که هر متغیر چه میزان به هر عنصر وابسته است. این مقادیر به عنوان بارهای فاکتور شناخته می شوند. مقادیر بالا و پایین یک نشان دهنده تاثیر بالا و پایین یک متغیر توسط آن فاکتور است. در قسمت قبل ٥ فاکتور انتخاب شد و در جدول٤ همانطور که مشاهده می‌شود برای این ٥ فاکتور امتیاز کسب شده به عدد ۱ نزدیک تر بوده و این نشان می‌دهد که تاثیر زیادی در یک متغیر توسط آن فاکتور وجود دارد.

مرحله ٥: انتخاب بهترین ویژگی

کاهش تعداد ویژگی‌ها در داده‌ها، دارای مزایای مهمی می باشد از قبیل: ۱-کاهش هزینه و زمان۲-جلوگیری از بیش برازش۳-افزایش تفسیرپذیری٤- انطباق با داده‌های واقعی٥-حفظ حریم خصوصی٦-کاهش استرس و نگرانی. در این پژوهش نیز با توجه به اهمیت تشخیص سریع و دقیق بیماری اوتیسم، ما به دنبال بهینه سازی تعداد ویژگی‌ها به منظور افزایش دقت مدل ما شینی هستیم. ا ستفاده از تمام ویژگی‌ها ممکن ا ست منجر به بیش برازش شود. با استفاده از الگوریتم‌ها و روش‌های مناسب، سعی می‌کنیم بیشترین اطلاعات را از تعداد کمی ویژگی استخراج کنیم و بر اساس آنها، مدل‌های ماشینی که قادر به تشخیص اوتیسم با دقت و قابل اعتماد هستند، طراحی و ارزیابی کنیم. در این قسمت، تمام داده‌ها به صورت عددی تبدیل شدند تا از آن‌ها در روش PCA بهره‌برداری شود. استفاده از این روش به منظور کاهش ابعاد داده ها و معرفی ویژگی های با اهمیت بالاتر برای اعمال به شبکه عصبی صورت گرفت. جدول ٤ ترتیب اولویت هر ویژگی در بالاترین سطر را نشان می‌دهد .به عبارت دقیق‌تر داده‌هایی که بیشترین ارتباط با یکدیگر را داشتند، در بالاترین سطر قرار گرفتند. امتیاز به ۱۰ سوال با ارزش ۰/۹۹٤در ماتریس جزئیات(جدول٤) نشان می‌دهد که این سوالات دارای وزن و اهمیت فوق العاده بالایی هستند و می‌توانند به تنهایی نقش مهمی در تشخیص یا ارزیابی الگوهای اوتیسم ایفا کنند. در واقع، این امتیاز نسبتاً نزدیک به ۱ نشان می‌دهد که این سوالات بسیار قدرتمند بوده و با توجه به پاسخ‌ها به طور قابل توجهی می‌توانند به تشخیص درست الگوهای اوتیسم کمک کنند. جدول ٤، با ارزش‌های ۰/٦٩٣، ۰/٦٧٣، ۰/٦٦٨و۰/٦٦۰ برای سوال‌های ۹، ٦، ٥ و ٤ به ترتیب، نشان می‌دهد که این سوالات نقش مهمی در تشخیص و ارزیابی الگوهای اوتیسم دارند. این ارزش‌ها نشان دهنده اهمیت قابل توجهی است که این سوالات در تمیز دادن الگوهای اوتیسم دارند. وجود امتیاز بالا نسبت به یک بیانگر این است که آن سوال قدرت تمیز بالایی دارد و ارائه اطلاعات مهمی در تشخیص الگوهای اوتیسم دارد. اطلاعات به دست آمده از جدول٤ با بررسی ارزش‌های سوالات، قادر است

ما را در فهم بهتری از اهمیت و نقش هر سوال در تشخیص الگوهای اوتیسم یاری دهد. نمودار شکل ۲ توزیع داده‌ها را در فضای چرخیده نمایش می‌دهد و به کمک آن می‌توان الگوهای مهمی را شناسایی کرد. با بررسی وزن‌ها در هر محور می‌توان عواملی که بیشترین تاثیر را در توزیع داده‌ها دارند، شناسایی کرد. این نمودار می‌تواند مشخصات و شاخص‌های مختلف را در بعضی از گروه‌ها تمایز دهد. این نمودار بر روی محورهای جزئیات PCA قرار دارد که از چرخش عناصر در روش PCA نتیجه می‌شود. با توجه به اینکه این نمودار در بازه ۱- و ۱ قرار دارد نشان می‌دهد که متغیرهای آنالیز شده در محدوده مشخصی از مقادیر خود قرار دارند و نتایج PCA محدود به این بازه می‌شود. همانطور که در این نمودار نشان می‌دهد مجموع امتیاز با ۱۰ سوال مرتبط بوده و ممکن است یک الگو یا ویژگی مشترک را نشان دهد. از طرفی ویژگی‌های سابقه خانوادگی و سن و جنسیت و نژاد و تکمیل کننده آزمون تفاوت‌های قابل توجهی را نسبت به باقی ویژگی‌ها نشان می‌دهند. همانطور که در نمودار مشخص است مولفه‌هایی که در بالاترین سطح قرار دارند دارای مقادیر واریانس بالاتری هستند و ویژگی‌هایی هستند که بیشترین وزن را در تشکیل مولفه اصلی دارند و نقش مهم تری در تحلیل داده‌ها ایفا می‌کنند.

مرحله ۶: آموزش و آزمون شبکه عصبی

همان‌طور که بیان شد ۱۸ ویژگی داشتیم که از این مجموعه ۱۷ ویژگی به عنوان ورودی و ویژگی باقی مانده که همان کلاس ASD است به عنوان هدف به شبکه عصبی اعمال می شود. در این مرحله، به بررسی و بهبود قابلیت تشخیص اختلال اوتیسم (ASD) با استفاده از یک شبکه عصبی با دو لایه مخفی و اعمال روش PCA میپردازیم.

تحلیل مؤلفه‌های اصلی(Principal Component Analysis) تبدیلی در فضای برداری است که آنالیز مجموعه داده های بزرگ با تعداد زیادی از ابعاد یا ویژگی ها را جهت افزایش تفسیرپذیری داده‌ها با حفظ حداکثر مقدار اطلاعات و تجسم داده های چند بعدی فراهم می‌کند. به عبارتی، تحلیل مؤلفه‌های اصلی(PCA) یک تکنیک آماری برای کاهش ابعاد یک مجموعه داده میباشد. طبقه‌بندی دقیق‌تر و دقت بالاتر این تشخیص می‌تواند به شناسایی زودهنگام و درمان بهتر این اختلال کمک کند. در این قسمت، مراحل ارزیابی دقیق شبکه و کاهش بعد ویژگی‌ها با استفاده از PCA توضیح داده می شود. این روش به ما اجازه می‌دهد تا تعداد ویژگی‌ها را با استفاده از تحلیل اجزاء اصلی تعیین کنیم و سپس ویژگی‌ها را به طوری کاهش دهیم که در صد خطای تشخیص به حداقل برسد. به این ترتیب الگوهای مهم و قابل استفاده در داده‌ها شناسایی می‌شود و اطمینان حاصل می‌شود که ویژگی‌های مهمی در تشخیص اوتیسم وجود دارند. معماری شبکه عصبی مورد استفاده در این شبکه در شکل۳ نشان داده شده است. به منظور اطمینان از صحت عملکرد مدل، از روش اعتبارسنجی متقابل (k_fold) استفاده شده است. به این ترتیب که داده‌ها به ۱۰ بخش تقسیم می‌شوند تا در هر مرحله آموزش و آزمون به طور مناسب انجام شود. سپس، عملکرد شبکه در قبال آموزش با ۹ بخش در هر بخش آزمون محاسبه می‌شود و نتایج در آرایه‌ای ذخیره می‌شود. همچنین، در هر آزمون، می‌توان نمودارهایی از برچسب‌ها و پیش‌بینی‌ها رسم کرد تا بتوانیم نتایج را بهبود داده و به صورت بصری تحلیل کنیم. در نهایت، با تحلیل مقدار کمینه، میانه عملکرد شبکه در تمام بخش‌های آزمون، دقت نهایی و کارایی تشخیص و پیش‌بینی اوتیسم مورد بررسی قرار می‌گیرد. با توجه به نتایج حاصل از روش کاهش بعد PCA، ترتیب اهمیت و تأثیر ویژگی‌ها در تشخیص اختلال اوتیسم در جدول ٤ قابل مشاهده است. بر اساس این نتایج، از آنجایی که همه ویژگی‌ها در همان ابتدا به شبکه عصبی اعمال و عملکردی با خطای صفر و دقت صددرصد حاصل شد، لذا برای کاهش ویژگی‌ها اقدام به حذف یک به یک ویژگی‌ها بر اساس درجه کمترین تاثیرش شده است. نتایج حاصل از آزمون در جدول ٥ نشان داده شده است. جدول زیر نشان می‌دهد که با حذف ٦ ویژگی از۱۷ ویژگی می‌توان به همان دقت کل رسید. اما حذف تعداد ویژگی‌های بیشتر، منجر به افزایش مقدار خطا خواهد شد. از این تحلیل

می‌توان نتیجه گرفت که برای دســـتیابی به عملکرد بدون خطا، حداقل ۱۲ ویژگی (که در جدول ٤ ذکر شده است) لازم است و لذا بکارگیری ترتیب ویژگی های برتر در تشـــخیص اختلال اوتیســـم حائز اهمیت اســـت و حذف برخی از ویژگی‌ها باعث کاهش دقت شبکه عصبی‌می‌شود.

مرحله ۷: نمایش اهمیت ویژه مجموعه داده ها

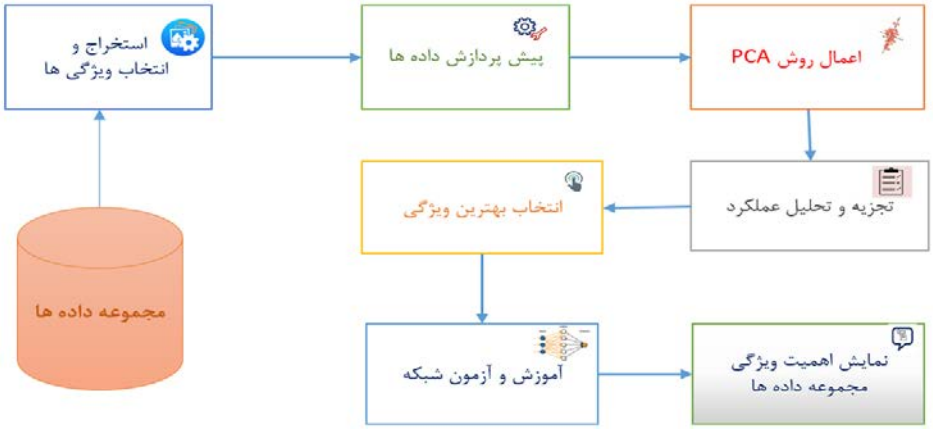
در این مقاله، ابتدا مجموعه داده‌های مورد بررســـی وارد SPSS شد و سپس به تجزیه و تحلیل با استفاده از روش PCA پرداخته شـد. ابتدا، ماتریس کواریانس داده‌ها محاسـبه شـد تا ارتباطات بین متغیرها را بررســـی کند. ســـپس، مقادیر ویژه و بردارهای ویژه ماتریس کواریانس محاســـبه شـــدند. مقادیر ویژه نشـــان می‌دهند که هر مؤلفه چقدر از تفاوت در داده ها را توضـــیح می‌دهد. به طور کلی، مقادیر ویژه بزرگتر به معنای اهمیت بیشتر مؤلفه اســـت. هر مؤلفه می‌تواند به عنوان یک الگوی معرفی‌کننده برای ســـاختار داده‌ها در نظر گرفته شـــود. مؤلفه‌های با مقادیر ویژه بزرگتر می‌توانند برای توصـــیف تفاوت‌های مهم در داده‌ها مورد اســـتفاده قرار گیرند. با بررســـی ضـــرایب بارگذاری، می‌توانید متغیرهایی که بیشترین وابستگی به هر مؤلفه را دارند را شناسایی کنید و اهمیت آن‌ها را در توصیف و تفسیر الگوهای اوتیسم مورد سنجش قرار دهد. استفاده از روش PCA در این مقاله به ما کمک می‌کند تا ســـاختار داده ها را درک کنیم و ارتباطات بین متغیرها را بررســـی کنیم. این روش به ما امکان می‌دهد تا متغیرها را بر اساس اهمیت آن‌ها در توضیح و تفسیر الگوهای اوتیسم مرتب کنیم، که برای تحلیل‌های آینده و بررسی عوامل کلیدی در رفتار و شـناخت افراد مبتلا به اوتیسم بسـیار مفید است.

جدول ۱. شرح ویژگی‌های مندرج در ابزار ASDTests

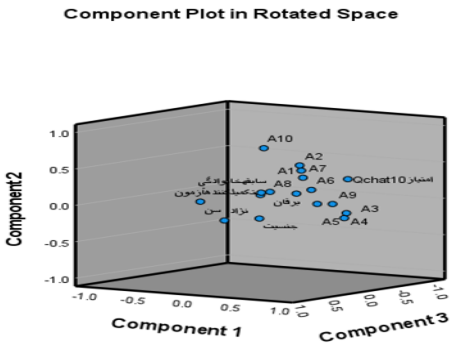
توضیحات	نوع	ویژگی
کودک نوپا (ماه)، کودکان، نوجوانان و بزرگسالان (سال)	عدد	سن
مرد یا زن	رشته	جنسیت
فهرست قومیت های رایج در متن	رشته	نژاد
آیا مورد با زردی متولد شده است	بله یا خیر	متولد شده با زردی
آیا یکی از اعضای نزدیک خانواده دارای PDD است	بله یا خیر	عضو خانواده با PDD
والدین، خود، مراقب، کادر پزشکی، پزشک و غیره.	رشته	چه کسی در حال تکمیل آزمون است
برای جزئیات Q1 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A1:Q1
برای جزئیات Q2 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A2:Q2
برای جزئیات Q3 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A3:Q3
برای جزئیات Q4 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A4:Q4
برای جزئیات Q5 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A5:Q5
برای جزئیات Q6 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A6:Q6
برای جزئیات Q7 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A7:Q7
برای جزئیات Q8 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A8:Q8
برای جزئیات Q9 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A9:Q9
برای جزئیات Q10 جدول ۲ را ببینید	باینری(۰،۱)	پاسخ A10:Q10
برای جزئیات به جدول ۲ مراجعه کنید	عدد	نتایج امتیازات
کودکان نوپا، کودکان، نوجوانان یا بزرگسالان مبتلا به اوتیسم	بله یا خیر	ASD

جدول ۲. جزئیات نگاشت متغیرها

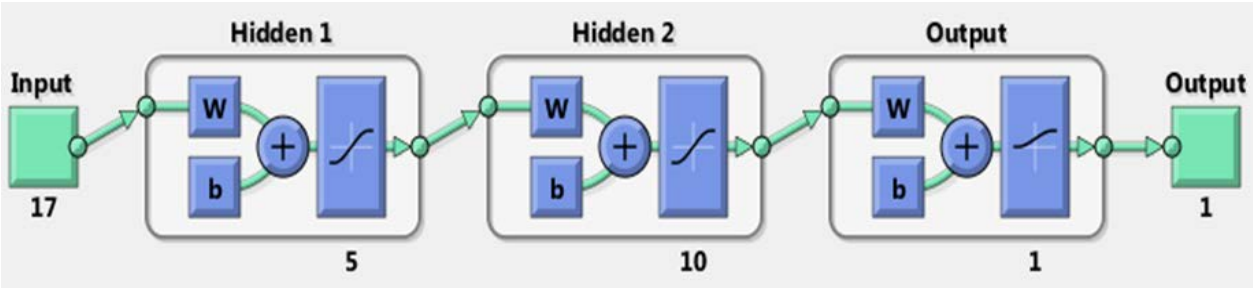
متغیر ها	نوپا ۱۸ تا ۳۶ ماهه ، ویژگی AQ -10	کودک ۴ تا ۱۱سال ، ویژگی AQ -10	نوجوان ۱۲ تا ۱۶سال ، ویژگی AQ -10	بزرگسال ۱۶ سال به بالا ، ویژگی AQ -10
Q1	آیا فرزندتان وقتی نام او را صدا می کنید به شما نگاه می کند؟	او اغلب به صداهای کوچکی توجه می‌کند که دیگران نمی‌توانند آن‌ها را	او بسیار الگوها را در اشیا و همچنین مسایل مختلف مشاهده می‌کند	او اغلب صداهای کوچکی را متوجه می‌شود که دیگران نمی‌توانند آن را بشنوند
Q2	ارتباط چشمی با فرزندتان چقدر برای شما آسان است؟	او معمولاً بیشتر روی تصویر کلی تمرکز می‌کند و به جزئیات کمتر توجه	او معمولاً به تصاویر کلی بیشتر توجه می‌کند و به جزئیات توجه کمتری دارد	او به طور معمول بیشتر بر کلیات تمرکز می‌کند تا جزئیات کوچک
Q3	آ یا فرزند شما با اشاره نشان می دهد که چیزی می خواهد؟	در گروه‌های اجتماعی، او به راحتی می‌تواند پیگیری صحبت‌های چندین نفر	در یک گروه اجتماعی، او به راحتی می‌تواند گفتگوهای مختلف افراد را پیگیری کند	او انجام چند کار به طور همزمان را آسان می‌یابد
Q4	آیا فرزندتان علاقه اش را با شما به اشتراک می گذارد؟	او به راحتی می‌تواند بین فعالیت‌های مختلف جابجا شود	اگر وقفه‌ای رخ دهد، او به سرعت می‌تواند به کاری که قبلاً در حال انجام آن بود،	اگر وقفه‌ای رخ دهد، او می‌تواند به سرعت به کاری که در حال انجام آن بود بازگردد
Q5	آیا فرزند شما تخیلی بازی می‌کند؟	او نمی‌داند چگونه با هم‌سن و سالان خود صحبت را ادامه دهد	او اغلب متوجه می‌شود که نمی‌داند چگونه بحث را به پیش ببرد	او به راحتی می‌توانم در حین صحبت کردن با شخص دیگر، پشت پرده های بیان شده را بفهمد
Q6	آیا فرزند شما هنگامی که شما به جایی نگاه می‌کنید، نگاه شما را	او در گفتگوهای اجتماعی ماهر است	او در چت و گفتگوهای کوتاه اجتماعی ماهر است	او می تواند بفهمد که شخصی که به او گوش می‌دهد علاقه مند نیست و خسته می‌شود
Q7	اگر شما یا عضوی ازخانواده ناراحت باشد، آیا فرزند شما تمایل	وقتی داستانی می خواند ، نمی‌تواند باورپذیری و احساسات شخصیت‌های	وقتی او کوچک‌تر بود، از بازی کردن با دیگر کودکان در بازی‌های تخیلی لذت می	وقتی یک داستان را می‌خواند، پیدا کردن نیت شخصیت‌ها برایش سخت است
Q8	آیا می‌توانید کلمات اولیه فرزند خود را شرح دهید؟	او قبلاً از بازی کردن در بازی‌های خیالی با بچه‌های دیگر لذت می‌برد	برای او بسیار سخت است تصور کند که چگونه به جای شخصی دیگری باشد و در	او از جمع‌آوری اطلاعات درباره دسته‌های مختلف چیزها لذت می‌برد
Q9	آیا فرزند شما از حرکات ساده استفاده می‌کند؟	او به راحتی می‌تواند با نگاه به صورت فرد، درک کند که آن شخص در حال فکر	او در موقعیت اجتماعی راحت عمل می‌کند	او به راحتی می‌تواند با نگاه کردن به چهره شخص، حس و احساسات او را درک کند
Q10	آیا فرزند شما بدون هدف و بدون دلیل خاصی به سمت چیزی	او به سختی دوستان جدید پیدا می کند	او به سختی دوستان جدید می‌یابد	برای او کشف و درک نیت و قصد افراد مشکل است



شکل ۱. روند تشخیص ASD در مراحل اولیه



شکل۲. نمودار جزئیات در فضای چرخیده.



شکل ۳. معماری شبکه عصبی چند لایه مورد استفاده در این پژوهش

جدول ۳. ماتریس همبستگی

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	سن	Qchat-10-امتیاز	جنسیت	نژاد	یرقان	سابقه خانوادگی	تکمیل کننده آزمون
A1	1	0.463	0.244	0.248	0.285	0.373	0.333	0.208	0.322	0.131	0.041	0.614	0.112	0.008	-0.006	0.053	0.038
A2	0.463	1	0.207	0.266	0.259	0.309	0.294	0.252	0.269	0.157	0.038	0.592	0.067	-0.015	0.046	0.035	0.047
A3	0.244	0.207	1	0.547	0.349	0.298	0.211	0.244	0.379	0.004	-0.11	0.593	0.063	0.04	-0.007	0.041	0.019
A4	0.248	0.266	0.547	1	0.348	0.348	0.319	0.252	0.428	0.017	-0.088	0.643	0.105	-0.008	0.027	-0.037	-0.004
A5	0.285	0.259	0.349	0.348	1	0.415	0.349	0.317	0.442	0.045	0.018	0.65	0.13	0.093	0.025	-0.022	0.021
A6	0.373	0.309	0.298	0.348	0.415	1	0.362	0.281	0.416	0.072	0.066	0.659	0.025	0.067	-0.005	0.047	0.018
A7	0.333	0.294	0.211	0.319	0.349	0.362	1	0.257	0.361	0.175	0.021	0.619	0.045	0.067	0.048	-0.024	0.001
A8	0.208	0.252	0.244	0.252	0.317	0.281	0.257	1	0.293	0.082	0.081	0.543	0.077	0.041	0.003	-0.037	0.043
A9	0.322	0.269	0.379	0.428	0.442	0.416	0.361	0.293	1	0.029	-0.049	0.672	0.087	0.027	0.013	-0.011	-0.026
A10	0.131	0.157	0.004	0.017	0.045	0.072	0.175	0.082	0.029	1	-0.013	0.29	-0.05	-0.047	0.027	0.007	0.038
سن	0.041	0.038	-0.11	-0.088	0.018	0.066	0.021	0.081	-0.049	-0.013	1	0.001	0.129	0.2	-0.042	0.033	0.013
Qchat-10-امتیاز	0.614	0.592	0.593	0.643	0.65	0.659	0.619	0.543	0.672	0.29	0.001	1	0.113	0.046	0.029	0.009	0.033
جنسیت	0.112	0.067	0.063	0.105	0.13	0.025	0.045	0.077	0.087	-0.05	0.129	0.113	1	0.035	0.038	0.059	0.023
نژاد	0.008	-0.015	0.04	-0.008	0.093	0.067	0.067	0.041	0.027	-0.047	0.2	0.046	0.035	1	-0.06	-0.041	0.032
یرقان	-0.006	0.046	-0.007	0.027	0.025	-0.005	0.048	0.003	0.013	0.027	-0.042	0.029	0.038	-0.06	1	0.078	-0.046
سابقه خانوادگی	0.053	0.035	0.041	-0.037	-0.022	0.047	-0.024	-0.037	-0.011	0.007	0.033	0.009	0.059	-0.041	0.078	1	-0.014
تکمیل کننده آزمون	0.038	0.047	0.019	-0.004	0.021	0.018	0.001	0.043	-0.026	0.038	0.013	0.033	0.023	0.032	-0.046	-0.014	1

جدول ۴. ماتریس جزئیات

	Component				
	۱	۲	۳	۴	۵
Qchat-10-امتیاز	۰/۹۹٤				
A9	۰/٦۹۳				
A6	۰/٦۷۳				
A5	۰/٦٦۸				
A4	۰/٦٦۰				
A7	۰/٦۱۹				
A1	۰/٦۱۳		-۳۰٦		
A3	۰/٦۰٥		-۰/۳٤۸		
A2	۰/٥۸۱		۰/۳۸٤		
A8	۰/٥۳۱				
سن		-۷۷٥			
نژاد		-٥۷۳	-۰/۳۳۸		
A10			۰/٦٦۱		
سابقه خانوادگی				۰/٦۳۷	
جنسیت		۰/۳۱٦		۰/٥۲٤	
یرقان				۰/٥۱٦	
تکمیل کننده آزمون					۰/۸۰۸

جدول۵. درصد صحت و خطای داده ها با درنظر گرفتن ویژگی‌های متفاوت

تعداد ویژگی‌ها	۱۷	۱۶	۱۵	۱۴	۱۳	۱۲	۱۱	۱۰	۹
درصد صحت	۱۰۰ درصد	۱۰۰ درصد	۱۰۰ درصد	۱۰۰ درصد	۱۰۰ درصد	۱۰۰ درصد	۹۹/٦ درصد	۹۹/۷ درصد	۹٥/٥ درصد
درصد خطا	۰ درصد	۰ درصد	۰ درصد	۰ درصد	۰ درصد	۰ درصد	۰/٤ درصد	۰/۳ درصد	٤/٥ درصد
Performance	$۷/۳ \times ۱۰^{-۷}$	$۳/۷ \times ۱۰^{-۷}$	$٤/٤ \times ۱۰^{-۷}$	$۳/۱ \times ۱۰^{-۷}$	$۷/٦ \times ۱۰^{-۷}$	$٤/۲ \times ۱۰^{-۷}$	۰/۱۷	۰/۰۱۷	۰/۱٥
میانگین خطا	$۳/٤ \times ۱۰^{-۷}$	$۳/٥ \times ۱۰^{-۷}$	$٤/٦ \times ۱۰^{-۷}$	۳×۱۰^{-۷}	$۳/٥ \times ۱۰^{-۷}$	۰/۰۲	۰/۰٤	۰/۰۳	$۲/۹ \times ۱۰^{-۷}$
تعداد تکرار	۳۷	۳٤	٤۱	۳۱	۳۷	٦۱	۱۷	۳٥	۱۳
Train Data	۱۷×۹٤۹	۱٦×۹٤۹	۱٥×۹٤۹	۱٤×۹٤۹	۱۳×۹٤۹	۱۲×۹٤۹	۱۱×۹٤۹	۱۰×۹٤۹	۹×۹٤۹

نتایج

مقاله حاضر به طور جامع به بررسی یک روش تشخیص اوتیسم با استفاده از یک مجموعه ویژگی می‌پردازد. در ابتدا، با استفاده از تحلیل مؤلفه های اصلی ، ویژگی‌ها در مقیاس تنظیم شده، ارائه شده اند. سپس با استفاده از یک شبکه عصبی MLP، تشخیص اوتیسم انجام می‌شود. آموزش شبکه عصبی با استفاده از داده‌های آموزشی صورت گرفته است تا بتواند نمونه جدیدی را تشخیص دهد که آیا مبتلا به اوتیسم است یا خیر. نتایج این مطالعه طیفی از یافته‌های مهم و قابل توجه را برای تشخیص اوتیسم ارائه می‌دهد. نشان می‌دهد که استفاده از شبکه‌های عصبی مصنوعی و تحلیل مولفه اصلی در تشخیص اختلال طیف اوتیسم دارای دقت و کارایی بسیار بالاست. یکی از نتایج مهم این مطالعه، توانایی بالای روش پیشنهادی در تمایز اشخاص مبتلا به اوتیسم از افراد سالم است. این نتایج نشان می‌دهد که روش پیشنهادی می‌تواند به عنوان یک ابزار قدرتمند و قابل اعتماد در تشخیص اوتیسم استفاده شود و می‌تواند به مؤسسات بهداشتی و متخصصان روانشناسی در تشخیص اولیه و درمان این اختلال کمک کند. در حین این مطالعه، با استفاده از تحلیل مولفه اصلی، ویژگی‌های مهم در تشخیص اوتیسم شناسایی شدند. این ویژگی‌ها می‌توانند به عنوان نشانگرهای مؤثر و حیاتی در تشخیص این اختلال مورد استفاده قرار گیرند. این نتایج نشان می‌دهند که استفاده از تحلیل مولفه اصلی، به میزان قابل توجهی در بهبود دقت و عملکرد روش‌های دیگر تشخیصی در تشخیص اوتیسم موثر است. دیگر یافته‌های مهم این مطالعه مربوط به عوامل مؤثر در تشخیص اوتیسم است. با تحلیل داده‌ها و ویژگی‌های استخراج شده، مشخص شد که برخی از عوامل مؤثر مانند نحوه پاسخگویی به سوالات مرتبط با اوتیسم، سطح تعامل اجتماعی و وجود الگوهای خاص بر روی داده‌ها، تاثیر قابل توجهی در تشخیص این اختلال دارند. این نتایج می‌توانند به متخصصان در ارزیابی و بررسی علائم و نشانه‌های اوتیسم کمک کنند و برای تشخیص دقیق‌تر و موثرتر این اختلال مورد استفاده قرار بگیرند. این یافته‌ها نشان می‌دهند که رویکرد پیشنهادی قابلیت کاربرد گسترده‌تری دارد و می‌تواند در تحقیقات بیشتر و درمان افراد مبتلا به اوتیسم مورد استفاده قرار گیرد.

نتیجه‌گیری

در این مطالعه، تلاش‌هایی برای تشخیص اختلال طیف اوتیسم با استفاده از شبکه‌های عصبی مصنوعی و تحلیل مولفه اصلی انجام شد. با به کارگیری یک مجموعه داده کامل که شامل ۱۰ سوال مرتبط با تشخیص اوتیسم و همراه با اطلاعات شخصی بود، میزان دقت در تشخیص اوتیسم به بالاترین سطح رسید. با استفاده از این روش ، ویژگی‌ها بر اساس اهمیت و نقش آن‌ها در تشخیص اوتیسم، مرتب شدند. سپس با استفاده از روش k-fold cross-validation، داده‌ها به شکلی بهینه تقسیم و شبکه‌های عصبی با استفاده از ویژگی‌های کاهش یافته، آموزش داده شدند. نتایج به دست آمده نشان داد که با استفاده از ۱۲ ویژگی منتخب، می‌توان به دقت ۱۰۰ درصد در تشخیص اوتیسم دست یافت.

کاهش تعداد ویژگی ها در داده ها، به طور کلی دارای چندین مزیت مهم است که می‌توان به آنها اشاره کرد:

۱. کاهش هزینه و زمان: با کاهش تعداد ویژگی‌ها، محاسبات مورد نیاز برای آموزش و استفاده از مدل‌های هوش مصنوعی به شدت کاهش می‌یابد. این به معنای کاهش هزینه و زمان مورد نیاز برای اجرای مدل و تشخیص بیماری است.

۲. جلوگیری از بیش برازش: در صورتی که تعداد زیادی ویژگی در مدل استفاده شود، ممکن است مدل به داده‌ها بیش از حد بخواند و برای داده‌های جدید عملکرد ضعیفی داشته باشد. با کاهش تعداد ویژگی‌ها، خطر بیش برازش، کاهش می‌یابد و عملکرد مدل بهبود می‌یابد.

۳. افزایش تفسیرپذیری: تفسیرپذیری مدل‌های هوش مصنوعی، به معنای قابلیت تفسیر و توضیح عملکرد آنها است. با کاهش تعداد ویژگی‌ها، ممکن است بتوانیم واضح تر توضیح دهیم که این ویژگی‌ها چگونه تاثیرگذار هستند و چطور مدل تصمیم‌گیری می‌کند.

٤. انطباق با داده‌های واقعی: معمولاً در برخی موارد، داده‌های واقعی به صورت کامل و شامل تمام ویژگی‌ها در دسترس نیستند. با کاهش تعداد ویژگی‌ها، می‌توانیم مدل‌های خود را بهتر با داده‌های واقعی هماهنگ کنیم و برای ورودی‌های ناقص‌تر نیز عملکرد مناسبی داشته باشیم.

٥. حفظ حریم خصوصی:کاهش تعداد ویژگی‌ها می‌تواند به والدین در مسائل حساسی که ممکن است با تشخیص اوتیسم برای فرزندانشان مطرح شود، کمک کند و منجر به کاهش اطلاعات مرتبط با وضعیت فرزندان شود. این امر می‌تواند نگرانی‌های والدین را در خصوص افشای اطلاعات حساس و شخصی فرزندان کاهش دهد.

٦. کاهش استرس و نگرانی: اگر برای تشخیص اوتیسم از تعداد سوالات زیادی استفاده شود، می‌تواند نگرانی‌ها و استرس والدین را بیشتر کند. با کاهش ویژگی‌ها، ممکن است اطمینان بیشتری حاصل شود و والدین با خیال راحت‌تری مسائل مرتبط با تشخیص اوتیسم را مدیریت کنند.

تشکر و قدردانی

بدینوسیله نویسندگان از کلیه افرادی که در مراحل نگارش این مقاله همکاری کردند، کمال تشکر و قدردانی را دارند.

تضاد منافع

در این پژوهش هیچ گونه تعارض منافعی توسط نویسندگان گزارش نشده است.

مشارکت نویسندگان:

(۱) مفهوم پردازی و طراحی مطالعه، یا جمع آوری داده ها، یا تجزیه و تحلیل و تفسیر داده ها : همه نویسندگان

(۲) تهیه پیش نویس مقاله یا بازبینی آن جهت تدوین محتوای اندیشمندانه: همه نویسندگان

(۳) تایید نهایی دستنوشته پیش از ارسال به مجله: همه نویسندگان

References

1. Thabtah F, Abdelhamid N, Peebles D: A machine learning autism classi_cation based on logistic regression analysis. Health Information Science and Systems 7(1), 1(11 (2019

2. Wiggins L.D, Reynolds A, Rice C.E, Moody E.J, Bernal P, Blaskey L, Rosenberg S.A, Lee L.C, Levy S.E: Using Standardized Diagnostic Instruments to Clas- sify Children with Autism in the Study to Explore Early Development. Journal of Autism and Developmental Dis- orders 45(5), 1271(1280 (2015))

3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5rd ed. Washington, DC: American Psychiatric Association; 2013.

4. Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics. 2010; 125(1):e17-e23.

5. Abedian Iman, Ali Ayoobi, hamid reza ghaffari and iman zabbah. "Diagnosis of diabetes by using a data mining method based on native data." (2019): 1-14.

6. Zabbah I, Maroosi A, Ebrahimpour R. Intelligent diagnosing COVID-19 Disease Using a Combination of Deep Features and Analyzing Original Component. Advanced Signal Processing. 2022 Apr 22.

7. Maroosi A, Zabbah I, Mogharebi M, Yasrebi SE. Improving Diagnosis of Breast Cancer Disease Using Adaptive Neuro-Fuzzy Inference System. Karafan Quarterly Scientific Journal. 2022 Nov 22;19(3):377-91.

8. Bennett CC, Hauser K. Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach. Artificial Intelligence in Medicine. 2013; 57(1):9-19.

9. Álvarez D, Cerezo-Hernández A, López-Muñiz G, Álvaro-De Castro T, Ruiz-Albi T, Hornero R, et al. Usefulness of artificial neural networks in the diagnosis and treatment of sleep apneahypopnea

syndrome. In: SZ Assefa, M Diaz-Abad, SM Scharf. Sleep Apnea-Recent Updates, London: InTech; 2017.

10. F Thabtah, F Kamalov, and K Rajab, ``A new computational intelligence approach to detect autistic features for autism screening," Int. J. Med. Inform., vol. 117, pp. 112_124, Sep. 2018.

11.F Thabtah and D Peebles, ``A new machine learning model based on induction of rules for autism detection," Health Inform. J., 2019, Art.

12.F. Thabtah, ``Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," Informat. Health Social Care, vol. 44, no. 3, pp. 278_297, 2018.

13.M. A. Hossain, S. M. S. Islam, J. M. Quinn, F. Huq, and M. A. Moni, ``Machine learning and bioinformatics models to identify gene expres- sion patterns of ovarian cancer associated with disease progression and mortality," J. Biomed. Inform., vol. 100, Oct. 2019, Art.

14.M. Duda, R. Ma, N. Haber, and D. P. Wall, ``Use of machine learning for behavioral distinction of autism and ADHD," Transl. Psychiatry, vol. 6, no. 2, p. e732, 2016.

15. Islam, M.R., Kabir, M.A., Ahmed, A., Kamal, A.R.M., Wang, H., Ulhaq, A.: Depression detection from social network data using machine learning techniques. Health information science and systems 6(1), 8 (2018).

16. Parikh, M.N., Li, H., He, L.: Enhancing diagnosis of autism with optimized machine learning models and per- sonal characteristic data. Frontiers in Computational Neuroscience 13 (2019)

17. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E.: The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathemati- cians. Journal of Autism and Developmental Disorders31(1), 5(17 (2001).)

18.C. Allison, B. Auyeung, and S. Baron-Cohen, ``Toward brief `red _ags' for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls," J. Amer. Acad. Child Adolescent Psychiatry, vol. 51, no. 2, pp. 202_212, 2012.

19.Autism Screening Data for Toddlers. Accessed: Sep. 10, 2018. [Online]. Available: https://www.kaggle.com/fabdelja/autism-screening-for- Toddlers

20.UCI Machine Learning Repository: Autism Screening Adult Data Set. Accessed: Sep. 10, 2018. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Autism+Screening+Adult.

21. Thabtah, F.: ASD Dataset (accessed January 13, 2020).

Journal of Torbat Heydariyeh University of Medical Sciences
2023, Volume 11, No. 2: 30-45.

Original article

Diagnosis of Autism Spectrum Disorder Using Principal Component Analysis for Feature Extraction and Artificial Neural Networks

Habibe Hooshmand[1], Najmeh Kheradnia[2], Abbas GHodrati[3], Atena Abidi [4]*

1. BSc Graduates, Faculty of Engineering, Torbat Heydarieh University, Iran

2. BSc Graduates, Faculty of Engineering, Torbat Heydarieh University, Iran

3. Assistant Professor, Department of Nursing, School of Nursing and Midwifery, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran

4. Department of Computer Engineering, Bushehr Branch, Islamic Azad University, Bushehr, Iran

Corresponding author: Department of Computer Engineering, Bushehr Branch, Islamic Azad University, Bushehr, Iran

Keywords:

Autism spectrum disorder, Artificial neural networks, Principal component analysis, Feature extraction

Abstract

Background & Aim: Autism spectrum disorder is one of the psychological disorders in children. Timely and accurate diagnosis of this disorder is of great importance in providing appropriate care and treatment for children. The main objective of this research is to emphasize the significance of features related to autism and their diagnosis using an intelligent model, as some of these features have higher priority.

***Methods*:** For this purpose, the principal component analysis (PCA) method was employed to prioritize the features, and after extracting the optimal features, automatic disease diagnosis was performed using artificial neural networks.

***Results*:** The data used in this study were collected from the Kaggle dataset, including 1054 individuals, out of which 728 were diagnosed with autism and 326 were healthy. The results of this study indicate that the gradual elimination of features and reduction from 18 to 12 features can lead to achieving the same accuracy in diagnosing the autism spectrum using artificial neural networks.

***Conclusion*:** Reducing the number of features in artificial intelligence models for autism diagnosis not only improves and optimizes the diagnostic process but also helps in reducing parental stress and preserving their privacy due to the reduced number of questions. Ultimately, this leads to the generation of models with better performance and interpretability.

How to Cite this Article: Hooshmand H, Kheradnia N, GHodrati A, Abidi A . Diagnosis of Autism Spectrum Disorder Using Principal Component Analysis for Feature Extraction and Artificial Neural Networks. Journal of Torbat Heydariyeh University of Medical Sciences. 2023;11(2):30-45.

Copyright © 2020 Journal of Torbat Heydariyeh University of Medical Sciences. This is an open-access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/) which permits copy and redistribute the material just in noncommercial usages, provided the original work is properly cite.