

تشخیص بیماری بروسولوزیس در شهرستان رفسنجان با استفاده از شبکه‌های عصبی مصنوعی آتوانکو در عمیق

حسین قیومی‌زاده^۱، علی فیاضی^{۲*}، مصطفی دانائیان^۳، آلاء سعیدی^۴

• پذیرش مقاله: ۱۳۹۸/۳/۱۱

• دریافت مقاله: ۱۳۹۷/۹/۲۵

مقدمه: بیماری تب مالت (بروسولوزیس) یکی از مهم‌ترین بیماری‌های عفونی مشترک بین انسان و دام محسوب می‌شود. با توجه به اندمیک بودن تب مالت و وجود گزارش‌های متعددی از موارد انسانی و حیوانی آن، میزان شیوع تب مالت انسانی در شهرستان رفسنجان در طول ۳ سال (۱۳۹۷-۱۳۹۵) تعیین گردید. هدف از این مطالعه یافتن یک تطابق و همسان‌سازی روشی خودکار و هوشمندانه با حساسیت کم بر مبنای شبکه عصبی است، که قادر به تشخیص دقیق تب مالت باشد.

روش: در این مطالعه توصیفی-تحلیلی موارد بروسولوز انسانی بر اساس شاخص‌های جنس، سن، بارداری، سابقه تماس با دام و استفاده از مواد غیرپاستوریزه لبنی، پارامترهای آزمایشگاهی Wright و 2ME در طول ۳ سال از شهرستان رفسنجان واقع در جنوب ایران تحلیل شد. داده‌ها به دو زیر مجموعه آموزش (۸۰٪) و آزمون (۲۰٪) تقسیم شدند. از روش شبکه عصبی مصنوعی آتوانکو در عمیق برای آموزش هر زیر مجموعه استفاده شد.

نتایج: روش آتوانکو در عمیق به صحت ۹۰/۸۴٪، حساسیت ۹۴/۶۱٪ و ویژگی ۵۰٪ در پیش‌بینی موارد بیماری بروسولوزیس در مجموعه داده‌های آزمون دست یافت، نتایج تجربی، عملکرد دقیق شبکه عصبی را در تشخیص بیماری بروسولوز نشان داد.

نتیجه‌گیری: مدل شبکه عصبی مصنوعی عمیق می‌تواند به عنوان یک روش کارآمد و هوشمند برای تشخیص موارد بروسولوز انسانی به کار گرفته شود. با این حال مطالعه و پژوهش‌های بیشتری برای طراحی مدل‌های دیگر از شبکه‌های عصبی مصنوعی بر اساس یادگیری عمیق جهت تشخیص سایر بیمارهای عفونی مورد نیاز است.

کلیدواژه‌ها: بروسولوز انسانی، رفسنجان، شبکه عصبی مصنوعی، روش خودرمنزگار عمیق، تشخیص

ارجاع: قیومی‌زاده حسین، فیاضی علی، دانائیان مصطفی، سعیدی آلاء. تشخیص بیماری بروسولوزیس در شهرستان رفسنجان با استفاده از شبکه‌های عصبی مصنوعی آتوانکو در عمیق. مجله انفورماتیک سلامت و زیست پزشکی ۱۳۹۸؛ ۶(۴): ۲۹۸-۳۰۸.

۱. دکتری مهندسی پزشکی، استادیار، گروه مهندسی برق، دانشکده فنی و مهندسی، دانشگاه ولی عصر (عج) رفسنجان، رفسنجان، ایران
۲. دکتری مهندسی کنترل، استادیار، گروه مهندسی برق، دانشکده فنی و مهندسی، دانشگاه ولی عصر (عج) رفسنجان، رفسنجان، ایران
۳. دکتری مهندسی مخابرات، استادیار، گروه مهندسی برق، دانشکده فنی و مهندسی، دانشگاه ولی عصر (عج) رفسنجان، رفسنجان، ایران
۴. کارشناس ارشد حشره‌شناسی پزشکی مبارزه با ناقلین، گروه مبارزه با بیماری‌ها، معاونت بهداشتی، دانشگاه علوم پزشکی رفسنجان، رفسنجان، ایران

* **نویسنده مسئول:** رفسنجان، میدان امام خمینی، دانشگاه ولی عصر (عج)، دانشکده فنی و مهندسی، گروه مهندسی برق

• Email: a.fayazi@vru.ac.ir

• شماره تماس: ۰۳۴۳۱۳۱۲۳۹۸

مقدمه

بروسلوزیس یک بیماری با درگیری چند ارگان (multi-systemic) است که در موارد شدید و حاد خود را به صورت یک بیماری تبار (ناگهانی و یا تدریجی) نشان می‌دهد؛ ولی در صورت عدم تشخیص و درمان به موقع و مناسب، بیماری ادامه یافته و می‌تواند به سمت یک بیماری مزمن و ناتوان کننده با عوارض شدید و گاهی مرگبار پیشرفت نماید [۱]. معمولاً شدت بیماری در کسانی که به دفعات با دام‌ها در تماس هستند و به طور دائمی با باکتری بروسلا تماس دارند خفیف و ملایم است. این بیماری به علت ایجاد سقط جنین در دام، کاهش تولید شیر، عقیمی و نازایی دام‌های مبتلا و همچنین به علت ابتلای انسان به بیماری تب مالت، همواره از دو بعد اقتصادی و بهداشتی مورد توجه قرار می‌گیرد [۲،۳]. تب مالت بیماری است که در جوانان بیشتر دیده می‌شود و در افراد با سن بالا، فراوانی کمتری دارد [۴].

سودمندترین آزمون بررسی پاسخ درمانی در بروسلوز، اندازه‌گیری عیار آزمون 2ME است؛ لذا بدین منظور نیز به کار می‌رود. در ایران حداقل تیترا برای Wright ۱/۸۰ و برای 2ME نیز ۱/۴۰ در نظر گرفته شده است. هر چند تعداد کمی از بیماران ممکن است به طور خودبه‌خودی خوب شوند؛ اما تشخیص و درمان زود هنگام و طولانی‌مدت، در بیماری تب مالت یک اصل است، چراکه هر چه درمان دیرتر شروع شود احتمال بروز عوارض و عود (حتی با درمان مناسب) بیشتر می‌شود. بر اساس مطالعات تأخیر بیش از ۳۰ روز در شروع درمان، با احتمال بیشتری با عوارض بیماری همراه خواهد بود. تقریباً هر عضوی از بدن ممکن است به این باکتری آلوده شود که بسته به نوع عضو مبتلا، علائم بسیار متنوعی ایجاد می‌شود. علائم آن به سایر بیماری‌های عفونی شبیه است و گاهی پزشک را در تشخیص دقیق بیماری به اشتباه می‌اندازد. از این رو است که به بیماری تب مالت لقب بیماری هزارچهره را داده‌اند [۱].

در مطالعه نبوی و همکاران تحت عنوان «ویژگی‌های اپیدمیولوژی، بالینی و پاراکلینیکی کودکان و نوجوانان مبتلا به تب مالت در استان کهگیلویه و بویراحمد واقع در جنوب غربی ایران طی سال‌های ۱۳۹۴-۱۳۸۸» مشخص گردید از ۸۱ نفر فرد مبتلا به تب مالت تعداد ۴۰ نفر (۴۹/۴ درصد) مرد و بقیه (۵۰/۶ درصد) زن بودند. ۶۷ نفر (۸۲/۷ درصد) از بیماران ساکن روستا و عشایر و ۱۴ نفر دیگر (۱۷/۳ درصد) ساکن شهر بودند. میانگین سن مبتلایان در زمان تشخیص $11/30 \pm 4/73$ سال و

شایع‌ترین گروه سنی ۱۸-۱۱ سال با ۶۱/۷ درصد بوده است. شایع‌ترین راه انتقال بیماری مصرف محصولات لبنی مشکوک و تماس هم زمان با دام به ترتیب با (۸۵/۲ درصد) و (۸۱/۵ درصد) بوده است. حدود ۸۲/۸ درصد از کل بیماران زمان تشخیص تا درمان آن‌ها بین صفر تا سه ماه (بروسلوز حاد) و حدود ۱۷/۲ درصد بین سه ماه تا یک سال (بروسلوز تحت حاد) بوده است [۵].

در مطالعه محمدیان و همکاران در استان اصفهان در سال ۱۳۹۱، از تعداد ۵۷۵ مورد بیمار مبتلا به تب مالت شناسایی شده ۶۹/۹ درصد از بیماران مرد و ۳۰/۱ درصد زنان تشکیل می‌دادند. همچنین زنان خانه‌دار دامدار بیشترین و کودکان کمترین نسبت بیماران بودند [۶]. در مطالعه پاکزاد و همکاران با عنوان «بررسی اپیدمیولوژی تب مالت در شمال و شمال غرب ایران» مشخص گردید بیشتر افراد مبتلا به بیماری روستایی بودند و مردها بیشتر از زن‌ها دچار بیماری شده بودند، بیشترین گروه سنی درگیر، گروه سنی ۴۴-۲۵ سال بود [۷].

در مطالعه‌ای که با عنوان «خصوصیات اپیدمیولوژیک و روند بروز بیماری تب مالت در استان خراسان رضوی در سال‌های ۱۳۸۸-۱۳۹۲» انجام شد ۸۵٪ بیماران ساکن روستا و ۱۵٪ ساکن شهر بودند و مردان با ۵۶/۹ درصد بیشترین موارد را به خود اختصاص دادند. شغل‌های خانه‌دار (۳۳/۹٪) و کشاورز-دامدار (۲۷/۸٪) بیشترین موارد بیماری را داشتند [۸]. در مطالعه هاجری و همکاران در شهرستان تیران و کرون طی سال‌های ۱۳۸۳-۱۳۹۲ نشان داد که ۶۵/۸ درصد مبتلایان تب مالت را مردان و ۳۴/۲ درصد را زنان تشکیل می‌دادند. همچنین ۸۶/۶ درصد در مناطق روستایی و ۱۳/۴ درصد مناطق شهری زندگی می‌کردند. سابقه تماس با دام در ۸۲/۵ درصد موارد وجود داشت [۹].

شناسایی الگو به دلیل نیاز مبرم به یادگیری ماشینی و هوش مصنوعی در مسائل عملی به یک مسئله بسیار مهم تبدیل شده است. یادگیری عمیق، یک شبکه ساختار سلسله مراتبی است که ساختار مغز انسان را شبیه‌سازی می‌کند تا ویژگی‌های داده‌های ورودی داخلی و خارجی را استخراج کند. یادگیری عمیق مبتنی بر الگوریتم‌هایی با استفاده از شبکه عصبی چندلایه از قبیل شبکه‌های عصبی عمیق، عصبی کانولوشن، باور عمیق، عصبی بازگشتی و آتوانکودر انباشت شده می‌باشند. این الگوریتم‌ها به کامپیوترها و ماشین‌ها اجازه می‌دهند تا دنیای ما را به خوبی مدل کنند تا هوش را به نمایش بگذارند. در آتوانکودر، لایه پنهان، نمایش بهتری نسبت به ورودی خام

آزمایش 2ME، اطلاعات مربوط به تماس با دام همچنین اطلاعات مربوط به سابقه استفاده از لبنیات غیرپاستوریزه بود. قابل ذکر است که آزمایش Wright برای تأیید حضور آنتی‌بادی علیه بروسلا است که در این روش آزمایشگاه در سرم بیمار ابتدا آزمایش سریعی را روی نمونه انجام داد که در این کار از ماده‌ای به نام رزبنگال (که حاوی آنتی ژن بروسلا است) استفاده شد. پس از مجاور کردن سرم بیمار با رزبنگال در اثر واکنش بین آنتی‌ژن با آنتی‌بادی اختصاصی آن، توده‌ای حاصل از این واکنش ایجاد شد که با چشم قابل مشاهده است. به این نوع واکنش آنتی‌ژن و آنتی‌بادی، به طور کلی آگلوتیناسیون گویند. اگر این تست مثبت شود برای تعیین میزان تیتراژ آنتی‌بادی باید از آزمایش تیتراسیون که به رایت لوله‌ای مشهور است، استفاده کرد. همچنین آزمایش 2ME پس از مثبت شدن رایت صورت می‌گیرد تا کلاس آنتی‌بادی را تشخیص دهد. در این تست ابتدا به وسیله ماده‌ای به نام ۲- مرکاپتواتانول، پیوندهای دی‌سولفیدی موجود در ساختمان Igm و در نتیجه کل Igm‌های احتمالی موجود در سرم بیمار را تخریب می‌کنند. سپس با این سرم بدون Igm، آزمایش را به همان طریقی که در رایت گفته شد، ادامه داده و در انتها آنتی‌ژن مخصوص 2ME را به همه لوله‌ها می‌افزایند. پس از ۲۴ ساعت انکوباسیون در ۳۷ درجه، آخرین لوله‌ای که در آن آگلوتیناسیون مشاهده شد به عنوان تیتراژ 2ME گزارش شد. مهم‌ترین کاربرد این آزمایش، تشخیص افتراقی بین بروسلاز فعال از غیرفعال در فردی است که تظاهرات بالینی بیماری را دارد؛ ولی کشت او منفی و تیتراژ وی نیز پایین است.

در این بخش طراحی SAE را برای طبقه‌بندی رقم معرفی می‌شود. خودرمن‌نگار یک شبکه عصبی مصنوعی است که برای یادگیری کدگذاری کارآمد مورد استفاده قرار می‌گیرد. هدف یک خودرمن‌نگار، یادگیری بازنمایش فشرده‌شده، برای یک مجموعه داده است. خودرمن‌نگارها از سه لایه یا بیشتر، استفاده می‌کنند. که شامل یک لایه ورودی و تعدادی لایه مخفی خیلی کوچک‌تر که رمزنگاری را تشکیل خواهند داد و در نهایت یک لایه خروجی که در آن هر نورون همان معنای مشابه لایه ورودی را دارد. خودرمن‌نگار تابعی است که عمدتاً از دو بخش تشکیل می‌گردد، یک بخش انکدر هست که یک تابع استخراج ویژگی بوده و بردار ویژگی را از روی ورودی‌ها محاسبه می‌نماید. بخش دیگر، یک دیکدر است که در واقع از یک تابع احتمال مشخص، مدل‌های احتمالاتی تعریف شده و برای حداکثر نمودن تشابه داده‌ها، آموزش داده می‌شوند. معمولاً

اصلی ارائه می‌دهد و لایه پنهان همیشه تراکم داده ورودی است که ویژگی‌های مهم ورودی هستند.

هدف این پژوهش استفاده از یک شبکه عمیق خودرمن‌نگار برای ایجاد یک سیستم تشخیص یادگیری عمیق برای کلاستر بندی کردن افراد دارای تب مالت است. از آنجایی که مطالعات اپیدمیولوژیک با بررسی عوامل مهم و خطرناک، به منظور کنترل بیماری تب مالت و اقدامات پیشگیرانه مؤثر هستند؛ لذا مطالعه حاضر جهت نیل به هدف کنترل بیماری در گروه‌های در معرض خطر انجام شد. در بیشتر تحقیقات قبلی از روش نیمه نظارتی یا با ناظر استفاده شده است که نقش کاربر انسانی در آن‌ها بسیار حائز اهمیت است.

روش

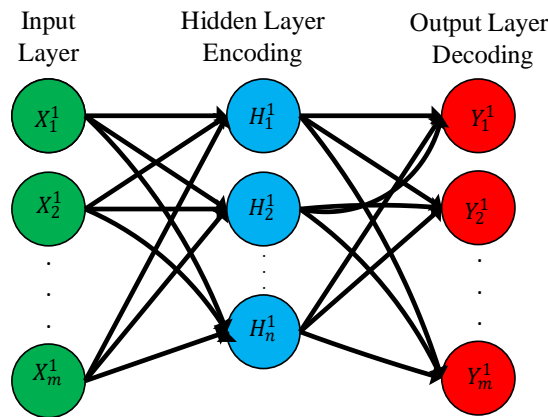
این مطالعه از نوع توصیفی-تحلیلی بود. تعداد کل افراد شرکت‌کننده در این مطالعه ۱۷۶ نفر بود. در این مطالعه به دلیل نامعلوم بودن حجم جامعه، حجم نمونه بر اساس رابطه زیر به دست آمد:

$$n = \frac{p(1-p)Z^2}{d^2} \quad (1)$$

که در آن p مشخص کننده نسبت مبتلایان به بیماری بروسلاز، Z معرف سطح اطمینان، d حاشیه خطا (بازه اطمینان) است. با توجه به مطالعات انجام شده توسط واحد بیمارهای معاونت بهداشتی دانشگاه علوم پزشکی رفسنجان نسبت مبتلایان به بیماری بروسلاز در شهرستان رفسنجان از ۱۳/۲۵ درصد تجاوز نمی‌کند؛ بنابراین در این مطالعه نسبت مبتلایان به بیماری بروسلاز، سطح اطمینان و حاشیه خطا به ترتیب ۰/۱۳۲۵، ۱/۹۶ (سطح اطمینان ۹۵٪) و ۰/۰۵ (حاشیه خطای ۵٪) انتخاب شد. با قرار دادن مقادیر بالا در رابطه ۱ حجم نمونه برای این مطالعه نزدیک ۱۶۱ نفر به دست آمد. مشخصات دموگرافیک و اپیدمیولوژی بیماران مراجعه کننده برحسب جنسیت در طول ۳ سال در جدول ۱ نمایش داده شده است. در بین بیماران بروسلازی شرکت‌کننده، ۵۳ نفر زن و ۱۰۸ نفر مرد بودند که دامنه سنی زنان ۲ تا ۸۳ سال و مردان ۲ تا ۸۱ سال بود. مشخصات دموگرافیک بیماران در جدول ۱ نمایش داده شد. اطلاعاتی بیماران شامل سن، جنس، بارداری، اطلاعات مربوط به آزمایش Wright، اطلاعات مربوط به

یک شبکه عصبی اتوانکودر را مشاهده نمود.

خودرمزنگارها برای کاهش بعد و یا استخراج ویژگی مورد استفاده قرار می‌گیرند. در شکل ۱ می‌توان ساختار مربوط به



شکل ۱: ساختار sparse auto-encoder

بوده و m تعداد نورون‌های مخفی و l تعداد sparse auto-encoder در واقع لایه خروجی، بردار ورودی I_m^l را به لایه پنهان H_m^l به کمک تابع غیرخطی S نگاشت می‌کند [۱۰]:

الگوریتم‌های sparse auto-encoder پیشنهادی با استفاده از تابع موجود در رابطه ۲ آموزش داده می‌شود که در آن ورودی با x_n^l ، لایه پنهان با H_m^l و لایه خروجی با Y_n^l مشخص می‌باشد همچنین n تعداد ورودی یا نورون خروجی

$$H_m^l = s \left(\sum_{i=0}^n (w_i * X_i^l) + b_m \right) \quad (2)$$

تابع سیگموئید به صورت رابطه ۳ تعریف می‌شود:

قابل ذکر است که در این رابطه w_i پارامترهای (یا وزن‌های) مرتبط با اتصال بین واحد ورودی و واحد پنهان است. b_m بایاس در لایه پنهان هستند. $S(v)$ معرف تابع سیگموئید است.

$$s(v) = \frac{1}{1 + e^{-v}} \quad (3)$$

جدول ۱: مشخصات دموگرافیک و اپیدمیولوژی بیماران مراجعه کننده برحسب جنسیت در طول ۳ سال (۹۷-۱۳۹۵)

متغیر	مرد (درصد) تعداد	زن (درصد) تعداد	تعداد کل (درصد) تعداد
سن (سال)	≤۲۴	۸ (۲۴/۴)	۳۳ (۱۰۰)
	۲۵-۴۴	۱۵ (۲۴/۲)	۶۲ (۱۰۰)
	۴۵-۶۴	۲۳ (۵۴/۸)	۴۲ (۱۰۰)
	≥۶۵	۱۱ (۴۵/۸)	۲۴ (۱۰۰)
محل سکونت	شهرستان رفسنجان	۳۶ (۲۸/۳)	۱۲۷ (۱۰۰)
	شهرستان انار	۲۱ (۶۱/۸)	۳۴ (۱۰۰)
وضعیت اشتغال	کشاورز و دامدار	۱۸ (۲۲/۵)	۸۰ (۱۰۰)
	خانه‌دار	۳۱ (۱۰۰)	۳۱ (۱۰۰)
	دانش آموز و دانشجو	۵ (۳۵/۷)	۱۴ (۱۰۰)
	سایر مشاغل	۴ (۱۱/۲)	۳۶ (۱۰۰)
سابقه تماس با دام	داشته	۲۸ (۲۴/۱)	۱۱۶ (۱۰۰)
	نداشته	۲۵ (۵۵/۶)	۴۵ (۱۰۰)
سابقه استفاده از فرآورده‌های لبنی غیرپاستوریزه	داشته	۳۸ (۳۳/۳)	۱۱۴ (۱۰۰)
	نداشته	۱۵ (۳۱/۹)	۴۷ (۱۰۰)

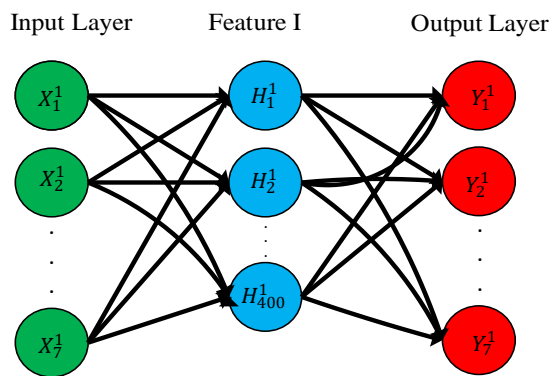
می‌توان از تابع فعالیت‌ساز tanh نیز استفاده نمود که نسبت به سیگموئید خصوصیت zero-center خروجی را داشت؛ اما هر دو تابع مشکلات تقریباً یکسانی دارند اعم از مشکل اشباع، مشکل آپدیت وزن‌ها در ناحیه اشباع (گرادیان صفر)، نیاز به هزینه زیاد initialization در جهت آپدیت وزن‌ها و در نهایت هزینه بر بودن محاسبه برای (Graphics

می‌توان از تابع فعالیت‌ساز tanh نیز استفاده نمود که نسبت به سیگموئید خصوصیت zero-center خروجی را داشت؛ اما هر دو تابع مشکلات تقریباً یکسانی دارند اعم از مشکل اشباع، مشکل آپدیت وزن‌ها در ناحیه اشباع (گرادیان صفر)، نیاز به هزینه زیاد initialization در جهت آپدیت وزن‌ها و در نهایت هزینه بر بودن محاسبه برای (Graphics

$$y_n^l = s \left(\sum_{j=0}^m (\hat{w}_j * H_j^l) + b_n \right) \quad (4)$$

مالت معرفی شد. اولین sparse autoencoder حاوی لایه ورودی است تا ویژگی‌های اولیه H_m^l ورودی خام که در شکل ۲ نشان داده شده است را یاد بگیرد [۱۰].

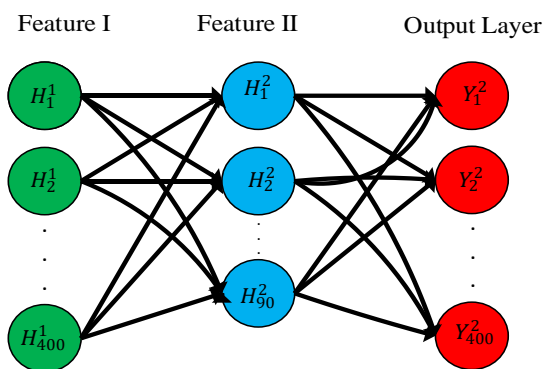
که در آن \hat{w}_j پارامترهای (وزن‌های) مرتبط با اتصال بین واحد پنهان و واحد خروجی هستند. b_n بایاس در لایه خروجی هستند. S تابع سیگموئید در معادله است. در این پژوهش طراحی Stacked Auto-encoders برای دسته‌بندی تب



شکل ۲: اولین sparse auto-encoder

ویژگی‌های اولیه که به‌عنوان ورودی خام به نسل بعدی auto-encoder برای استفاده از ویژگی‌های ثانویه مورد استفاده قرار می‌گیرد، نشان داده شد.

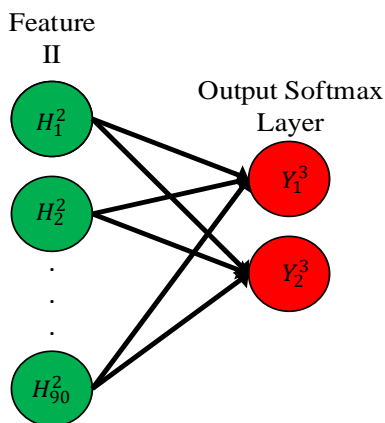
اولین sparse auto-encoder ویژگی اصلی (ویژگی I) را تولید می‌کند. ویژگی اصلی H_m^I لایه ورودی را به دومین sparse auto-encoder آموزش داده شده، می‌دهد که ویژگی‌های ثانویه (II) را تولید می‌کنند. در شکل ۳،



شکل ۳: دومین sparse auto-encoder

برچسب‌های رقمی که در شکل ۴ نشان داده شد، رفتار می‌کند.

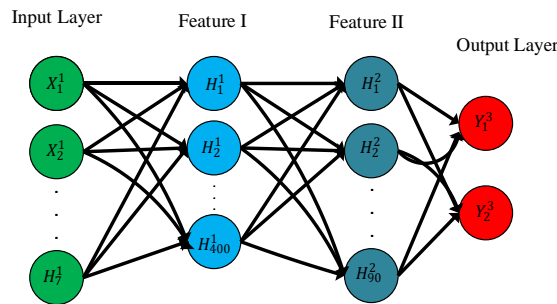
سپس، ویژگی ثانویه به‌عنوان لایه ورودی برای یک طبقه‌بندی کننده softmax برای نگاشت ویژگی‌های ثانویه به



شکل ۴: طبقه‌بندی کننده softmax

autocoder ترکیب شدن که در شکل ۵ نشان داده شد.

در نهایت، sparse auto-encoder اول و دوم با طبقه بندی کننده softmax برای تولید سه لایه از



شکل ۵: طرح پیشنهادی معماری autoencoder

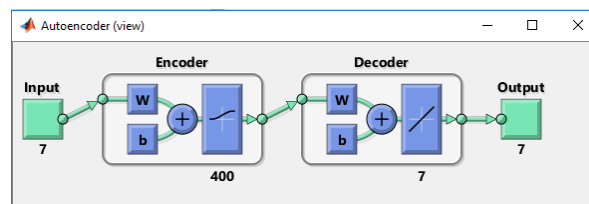
با دام (۱: تماس داشته، ۰: تماس نداشته)، اطلاعات مربوط به سابقه استفاده از لبنیات غیرپاستوریزه (۱: سابقه دارد، ۰: سابقه ندارد) بود.

آتوانکودر کشف کرد که چگونه ویژگی‌های از هر ورودی از دیتاهای بیمار را یاد بگیرد. آزمایش‌ها در محیط برنامه‌نویسی Matlab نسخه ۲۰۱۸ انجام شد. آتوانکودر با استفاده از جعبه ابزار یادگیری عمیق توسط Matlab انجام شد. اولین sparse auto encoder با لایه ورودی ۷، یک لایه پنهان از اندازه ۴۰۰ و لایه خروجی اندازه ۷ در شکل ۶ نشان داده شد. اولین sparse auto encoder برای تولید ۴۰۰ ویژگی اولیه آموزش داده شد.

Stacked Auto-encoders دارای دو لایه پنهان (ویژگی‌های اولیه و ثانویه) و لایه خروجی (طبقه بندی کننده softmax) هستند که قادر به طبقه بندی ارقام تب مالت می‌باشند.

نتایج

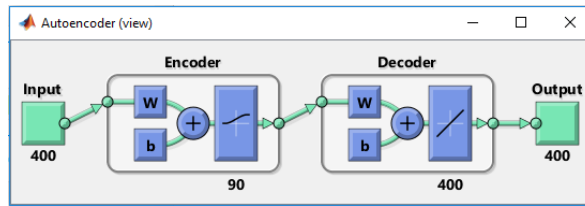
Stacked Auto-encoders شامل یک رمزگذار با لایه‌های ورودی ۷ بود. ورودی‌هایی که در ایجاد شبکه استفاده شده است شامل سن، جنس (۱: مرد، ۲: زن)، بارداری (۱: باردار بوده، ۰: باردار نبوده)، اطلاعات مربوط به آزمایش Wright، اطلاعات مربوط به آزمایش 2ME، اطلاعات مربوط به تماس



شکل ۶: اولین sparse auto encoder پیشنهاد شده

autoencoder برای تولید ۹۰ ویژگی ثانویه آموزش دیده است.

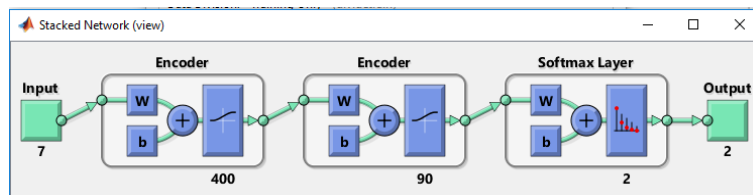
لایه دوم sparse autoencoder با لایه ورودی ۴۰۰، یک لایه پنهان با اندازه ۹۰ و لایه خروجی به اندازه ۴۰۰ که در شکل ۷ نشان داده شد، طراحی شد. دومین sparse



شکل ۷: دومین sparse autoencoder پیشنهادی

پیشنهادی توسط یک لایه ورودی از اندازه ۷، لایه‌های پنهان با ۴۰۰ ویژگی اولیه، ۹۰ ویژگی ثانویه و لایه خروجی با ۲ برجسب را نشان داد. Stack auto encode پیشنهادی نشان داده شده در شکل ۸، ۱۴۲ داده‌های آزمایش از تصویر را آموزش می‌دهند.

Sparse auto encoder اول و دوم از L2 regularization برای یادگیری نمایش اسپارس استفاده نمودند. Regularization تأثیر یک L2 regularizer را برای وزن شبکه (و نه بایاس‌ها) کنترل نمود. در نهایت، این ۹۰ ویژگی، لایه softmax نشان داده شده در شکل ۴ را تغذیه کردند. لایه softmax برای تولید دو کلاس خروجی آموزش



شکل ۸: مدل شماییک stack auto encode پیشنهادی

نتایج مربوط به حساسیت و ویژگی به دست آمده در داده‌های آموزش و آزمون در جدول ۲ نشان داده شد. متغیرهای SE و ACC به ترتیب حساسیت و صحت تشخیصی مدل بودند و متغیرهای True Positive (TP)، True Negative (TN)، False Positive (FP) و False Negative (FN) به ترتیب تعداد مثبت‌های درست، منفی‌های درست، مثبت‌های نادرست و منفی‌های نادرست هستند.

برای ایجاد نتایج بهتر، پس از اتمام این فاز آموزش، تنظیم خوب با استفاده از پس انتشار می‌تواند برای بهبود نتایج با تنظیم پارامترهای تمام لایه‌ها در یک‌زمان انجام گردد. autoencoder را با ۱۴۲ داده آموزش تغذیه شد. نگاهت آموخته‌شده توسط قسمت رمزگذار یک auto encoder می‌تواند برای استخراج ویژگی‌های داده مفید باشد. پس از به دست آمدن پارامترهای وزنی در شبکه آموزش دیده، ۳۷ داده را به عنوان آزمون به شبکه اعمال شد.

جدول ۲: نتایج مربوط به حساسیت و ویژگی داده‌ها

نوع داده	SE	ACC	SPE	FN	TN	TP	FP	تعداد
آموزش	۹۴٪/۱۶۱	۹۰٪/۱۸۴	۵۰٪	۷	۶	۱۲۳	۶	۱۴۲
آزمون	۶۱٪	۸۱٪/۰۸	۲۰٪	۳	۱	۲۹	۴	۳۷

بحث و نتیجه گیری

عملکرد مدل پیشنهادی در سیستم‌های تشخیص ناهنجاری تب مالت توسط سه طبقه‌بند کننده شامل ماشین بردار پشتیبان،

تجزیه و تحلیل متقارن درجه دوم و الگوریتم K-نزدیک‌ترین همسایگی با استفاده از cross validation پنج‌گانه مورد ارزیابی قرار گرفت (جدول ۳).

جدول ۳: ارزیابی مدل پیشنهادی با سایر طبقه‌کننده‌ها

مدل	Accuracy
شبکه عصبی اتوانکودر	۹۰/۸۴٪
ماشین بردار پشتیبان	۸۸٪
تجزیه و تحلیل متقارن درجه دوم	۹۰٪
الگوریتم K-نزدیک‌ترین همسایگی	۸۸/۵٪

بزرگ‌تر شدن بانک اطلاعاتی این نقص کم کم برطرف شد. از دیگر معایب این روش می‌توان به وقت‌گیر بودن آن اشاره نمود؛ بنابراین کارهای بیشتری برای بهبود استحکام مدل، حد تشخیص و کارایی مرتب‌سازی برای تشخیص سرعت بالا مورد نیاز است. برای بهبود بیشتر دقت، قابلیت اطمینان و سرعت تشخیص این رویکردها، به‌ویژه برای ارزیابی آنلاین، توسعه الگوریتم‌های جدید و مؤثر metrics chemo ضروری است. علاوه بر این، ترکیبات فناوری‌های مختلف می‌توانند به خوبی بر معایب یک فن منفرد فائق آیند که می‌تواند راه دیگری نیز باشد.

تشکر و قدردانی

نویسندگان مقاله مراتب تقدیر و تشکر خود را از مسئولین محترم دانشگاه علوم پزشکی رفسنجان و همکاران شاغل در آن شبکه به دلیل مساعدت‌های بی‌دریغ‌شان طی اجرای مطالعه، اعلام می‌نمایند.

تعارض منافع

بدین وسیله نویسندگان تصریح می‌نمایند که هیچ‌گونه تضاد منافی در خصوص پژوهش حاضر وجود ندارد.

همان‌گونه که در جدول ۳ مشاهده شد، تمامی الگوریتم‌های به کار گرفته در سیستم تشخیص ناهنجاری تب مالت، دارای دقت بسیار بالایی بودند. ولیکن، دقت مدل پیشنهادی با ۹۰/۸۴ درصد، نسبت به سایر مدل‌ها اندکی بیشتر است. البته با توجه به مورد مطالعه، هر یک از این مدل‌ها می‌تواند به عنوان طبقه‌کننده با دقت مناسب مورد استفاده قرار گیرد.

بیماری‌های مشترک بین انسان و حیوان یکی از مهم‌ترین مشکلات بهداشتی و درمانی به ویژه در مناطق روستایی است، یکی از شایع‌ترین این بیماری‌ها تب مالت است. در این پژوهش، اثربخشی یادگیری عمیق stacked برای تشخیص تب مالت ارائه شد.

در کاربرد timereal- بودن، به‌عنوان هدف نهایی برای تشخیص تب مالت، به کمک شبکه‌های عصبی عمیق هنوز در مراحل اولیه تحقیقات قرار دارد، علت آن هم دسترسی به داده‌های کم جهت آموزش شبکه‌های عصبی بود. یکی از معایب این مدل می‌تواند پایین بودن نسبی ویژگی تشخیصی باشد که به دلیل کم بودن نمونه‌های منفی نسبت به نمونه‌های مثبت بود که با افزایش تعداد نمونه‌های منفی قابل حل خواهد بود. یکی دیگر از معایب مهم این روش و سایر روش‌های عمیق وابستگی آن‌ها به بانک اطلاعاتی مورد استفاده بود که با

References

1. Shirzadi M, Guideline for the Diagnosis and Treatment of Brucellosis. Tehran: Andishmand; 2013. [In Persian]
2. Abdussalam M, Fein D. Brucellosis as a world problem. Developments in Biological Standardization 1976; 31: 9-23.
3. Taheri Soodejani M, Lotfi M, Ghaderi A, Reisi A, Mohammadzadeh M. Epidemiology of brucellosis in

- Shahr-e- Kord during the years 2010 to 2014. Pars of Jahrom University of Medical Sciences 2016;14(1):1-7. [In Persian] doi: 10.29252/jmj.14.1.1
4. Heidari M, Nazer M, Kheirollahi A, Birjandi M, Zareie H. Frequency of epididymo-orchitis in hospitalized patients with acute scrotum at Shohadaye Ashayer Hospital, Khorramabad, Iran. J Pak Med Assoc 2012;62(1):44-6.

5. Nabavi M, Hatami H, Jamali Arand H. Epidemiological, Clinical and Paraclinical Surveys of Children and Adolescents with Brucellosis in Kohgiluyeh and Boyer Ahmad Province in Southwest of Iran during 2009-2015. *Armaghane Danesh* 2018; 23(2):236-52. [In Persian]
6. Mohammadian M, Salehiniya H, Kazaei S, Ramazanpour J, Mohammadian-Hafshejani A. Epidemiological characteristics and incidence rate of brucellosis in Isfahan province. *Journal of Isfahan Medical School* 2015; 33(355): 1784-95. [In Persian]
7. Pakzad R, Barati M, Moludi J, Barati H, Pakzad I. Epidemiology of Brucellosis in the North and North-West Iran. *Paramedical Sciences and Military Health* 2016; 11(1):17-23. [In Persian]
8. Hashtarkhani S, Akbari M, Jarahi L, Etmnani K. Epidemiological characteristics and trend of incidence of human brucellosis in Razavi Khorasan province. *Medical Journal of Mashhad University of Medical Sciences* 2015; 58(9): 531-8. [In Persian]
9. Hajari A, Shams M, Afrooghi S, Fadaei Nobari R, Abaspoor Najafabadi R. Brucellosis in Tiran & Karvan: Study of the Disease in the Recent 10-Year Period (2004-2013). *Military Caring Sciences* 2016; 3(2):100-6. [In Persian] doi: 10.18869/acadpub.mcs.3.2.100
10. Loey M, El-Sawy A, EL-Bakry H. Deep Learning Autoencoder Approach for Handwritten Arabic Digits Recognition. arXiv 2017.

The Diagnosis of Brucellosis in Rafsanjan City Using Deep Auto-Encoder Neural Networks

Ghayoumi Zadeh Hossein¹, Fayazi Ali^{2*}, Danaeian Mostafa³, Saeidi Alae⁴

• Received: 25 Dec, 2018

• Accepted: 1 Jun, 2019

Introduction: Brucellosis is considered as one of the most important common infectious diseases between humans and animals. Considering the endemic nature of brucellosis and the existence of numerous reports of human and animal cases of brucellosis in Iran, the incidence of human brucellosis in Rafsanjan city was determined in the last 3 years (2016–2018). The main objective of this study was to find an automated consistent and intelligent method with low sensitivity based on a neural network which is capable of accurate detection of brucellosis disease.

Methods: In this descriptive analytic study, cases of human brucellosis in Rafsanjan, south of Iran, were analyzed based on sex, age, pregnancy, history of contact with livestock and the use of non-pasteurized dairy products, Right Laboratory parameters and 2ME during 3 years (2016–2018). Data were split into two subsets of train (80%) and test (20%). The artificial neural network approach of the deep auto-encoder was used to train each subset.

Results: The deep auto-encoder method achieves 94.61% sensitivity, 90.84% accuracy and 50% specificity in the diagnosis of brucellosis over the experimental data sets. The experimental results also showed the excellent performance of the proposed artificial neural network.

Conclusion: The deep artificial neural network model can be used as an efficient and intelligent method to detect the human cases of brucellosis. However, further studies are required to design other models of artificial neural networks based on deep learning to detect other infectious diseases.

Keywords: Human Brucellosis, Rafsanjan, Artificial neural network, Deep auto-encoder method

• **Citation:** Ghayoumi Zadeh H, Fayazi A, Danaeian M, Saeidi A. The Diagnosis of Brucellosis in Rafsanjan City Using Deep Auto-Encoder Neural Networks. Journal of Health and Biomedical Informatics 2020; 6(4): 298-308. [In Persian]

1. Ph.D. in Biomedical Engineering, Assistant Professor, Electrical Engineering Dept., Faculty of Engineering, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran
2. Ph.D. in Control Engineering, Assistant Professor, Electrical Engineering Dept., Faculty of Engineering, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran
3. Ph.D. in Communication Engineering, Assistant Professor, Electrical Engineering Dept., Faculty of Engineering, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran
4. M.Sc. of Medical Entomology, Fighting Diseases Dept., Deputy of Health, Rafsanjan University of Medical Sciences, Rafsanjan, Iran

***Correspondence:** Electrical Engineering Dept., Faculty of Engineering, Vali-e-Asr University of Rafsanjan, Imam Khomeini Square, Rafsanjan, Iran

• **Tel:** 03431312398

• **Email:** a.fayazi@vru.ac.ir