تخمین درصد چربی بدن افراد با استفاده از تکنیک های یاد گیری ماشین

رسمیه عسگری[۱]، محمد رضا ولی زاده[۲*]، کوروش جعفریان[۲]

۱) گروه کامپیوتر، دانشگاه آزاد اسلامی، واحد ایلام، ایلام، ایران

۲) گروه کامپیوتر، دانشکده فنی، دانشگاه ایلام، ایلام، ایران

۳) گروه تغذیه، دانشکده علوم تغذیه و رژیم شناسی، دانشگاه علوم پزشکی تهران ، تهران، ایران

تاریخ دریافت: ۹۵/۶/۲۰ تاریخ پذیرش:۹۵/۱۲/۱۵

چکیده

مقدمه: پزشکان با استفاده از دستگاه BIA (Bioelectrical Impedance Analysis) نسبت به محاسبه درصد چربی بدن افراد مبادرت می نمایند. مطالعه حاضر با هدف تخمین درصد چربی بدن افراد بدون استفاده از دستگاه صورت گرفته است. در این پژوهش سعی شده است که با استفاده از شبکه عصبی به پیش بینی میزان چربی بدن افراد پرداخته شود.

مواد و روش ها:داده مورد استفاده در این تحقیق اطلاعات بیماران مراجعه کننده به یکی از کلینیک های تغذیه درشهرتهران است. این مجموعه دارای ۴۰۰ رکورد است که از آنها برای آزمایش و ارزیابی شبکه عصبی پرسپترون چندلایه برای تخمین درصد چربی بدن افراد استفاده شده است. شبکه عصبی مورد استفاده دارای پنج نرون ورودی و ده نرون میانی است. هم چنین از روش ارزیابی متقاطع برای سنجش کارایی روش پیشنهادی این تحقیق استفاده شده است.

یافته های پژوهش: نتایج نشان دهنده ۲.۵ واحد خطا بر اساس روش ارزیابی متقاطع است که بیان کننده کارایی روش پیشنهادی است. نتایج آزمایش ها نشان می‌دهد که شبکه عصبی پیشنهادی برای تخمین درصد چربی بدن افراد دارای دقت میانگین ۹۳٪ می‌باشد. بنابراین روش پیشنهادی می تواند میزان دقیق درصد چربی بدن افراد را به خوبی تخمین بزند.

بحث و نتیجه گیری: نتایج حاصل از این تحقیق نشان می دهد که روش پیشنهاد شده به عنوان اولین روشی که از تکنیک های داده کاوی برای این منظور استفاده می کند، برای تخمین دقیق درصد چربی از دقت بالایی برخوردار بوده و می‌تواند به عنوان یک ابزار کارآمد مورد استفاده قرار گیرد. استفاده از راهکار پیشنهادی می تواند نیاز به دستگاه BIA یا روش های مشابه دیگر به منظور تخمین درصد چربی بدن را بر طرف نماید.

واژه های کلیدی: الگوریتم یاد گیری، درصد چربی بدن، داده‌کاوی، شبکه عصبی

* **نویسنده مسئول:** گروه کامپیوتر، دانشکده فنی، دانشگاه ایلام، ایلام، ایران

Email: mr.valizadeh@ilam.ac.ir

Copyright © 2017Journal of Ilam University of Medical Science. This is an open-access article distributed under the terms of the Creative Commons Attribution international 4.0 International License (https://creativecommons.org/licenses/by-nc/4.0/) which permits copy and redistribute the material, in any medium or format, provided the original work is properly cited.

[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]

مقدمه

چاقی در سراسر دنیا به طور هشدار دهنده ای در حال افزایش است و به یک همه گیری جهانی تبدیل شده است. چاقی در جهان در بین سال های ۱۹۸۰ تا ۲۰۱۳ بیش از دو برابر رشد داشته است (۱). علل مختلفی از جمله دسترسی آسان و فراوان به غذا های پر انرژی، کاهش سطح فعالیت بدنی و عوامل مشابه موجب افزایش شیوع چاقی شده اند. چاقی به طور مستقیم یا غیر مستقیم با بیماری های متعددی از قبیل: فشار خون بالا، بیماری های قلبی عروقی، دیابت، سکته مغزی، سرطان و مرگ زود رس مرتبط است. درصد چربی بدن یکی از معیار های مفید برای سنجش ریسک چاقی افراد و شناسایی چاقی نهفته به شمار می آید (۲). تخمین میزان درصد چربی بدن امروزه با استفاده از تجهیزات پزشکی مانند محاسبه مقاومت بیو الکتریک (Bioelectrical Impedance Analysis) صورت می گیرد. در صورتی که به نظر می رسد تخمین آن با استفاده از مشخصات افراد امکان پذیر باشد. به همین دلیل در این مقاله سعی شده تا با استفاده از تکنیک های یادگیری ماشین یک مدل کارآمد برای تخمین درصد چربی بدن ارائه شود. هدف ما از این تحقیق پیش بینی و تخمین درصد چربی بدن با استفاده از تکنیک های یاد گیری ماشین است؛ به طوری که بتوان با استفاده از راهکار های پیشنهادی در هزینه و زمان صرف شده برای تخمین درصد چربی بدن صرفه جویی کرد. داده کاوی فرآیندی است خودکار برای استخراج الگو هایی که دانش را بازنمایی می‌کنند، که این دانش به صورت ضمنی در پایگاه داده های عظیم، انباره داده و دیگر مخازن بزرگ اطلاعات ذخیره شده است. داده کاوی، به طور همزمان از چندین رشته علمی، نظیر: تکنو لوژی پایگاه داده، هوش مصنوعی، یادگیری ماشین، شبکه های عصبی، آمار، شناسایی الگو، سیستم های مبتنی بر دانش، بازیابی اطلاعات، محاسبات سرعت بالا و بازنمایی بصری داده بهره می برد (۳،۴) . شبکه عصبی، یکی از ابزارهای داده کاوی برای پیش بینی با طبقه بندی داده ها می - باشد و در این تحقیق به منظور پیش بینی میزان درصد چربی بدن افراد استفاده شده است. شبکه های عصبی مصنوعی یا به زبان ساده‌تر شبکه های عصبی، سیستم ها و روش های محاسباتی نوینی برای یاد گیری ماشینی، نمایش دانش و در انتها اعمال دانش به دست آمده در جهت پیش‌بینی پاسخ های خروجی از سامانه - های پیچیده هستند. ایده اصلی این گونه شبکه ها (تا حدودی) از شیوه کارکرد سیستم عصبی زیستی الهام گرفته شده اند و برای پردازش داده ها و اطلاعات به منظور یاد گیری و ایجاد دانش ارایه شده است. این سیستم از شمار زیادی عناصر پردازشی فوق العاده به هم پیوسته با نام نرون تشکیل شده که برای حل یک مسئله با هم هماهنگ عمل می‌کنند و توسط سیناپس ها (ارتباطات الکترو مغناطیسی) اطلاعات را منتقل می کنند. یکی از معروف ترین معماری های شبکه های عصبی پرسپترون است که از روش پیش خور استفاده می کند. پردازنده های شبکه (نرون ها) به چند لایه مختلف تقسیم می شوند؛ در این شبکه ها لایه اول، ورودی، لایه آخر، خروجی و لایه (های) میانی، لایه (های) پنهان نامیده می شوند. می توان این معماری را پر کاربرد ترین معماری شبکه های عصبی نامید. از شبکه عصبی در پزشکی به صورت گسترده ای استفاده شده است. جوتیسونی و همکارن درسال ۲۰۱۱ از شبکه عصبی برای پیش بینی بیماری قلبی استفاده کردند (۵). آن ها از مجموعه داده ای که راج کومار آماده کرده بود استفاده کردند (۶). برای ارزیابی سیستم پیشنهادی از روش ارزیابی متقاطع (Cross validation) استفاده شده که نتایج در مقایسه با سیستم های قبلی رضایت بخش بوده است. در (۷) از روش های مختلف از جمله شبکه عصبی برای پیش بینی سرطان سینه استفاده شده است. آن ها از روش ارزیابی متقاطع برای بررسی کارایی سیستم خود استفاده نمودند. صادق کارا و همکاران (۸) در تشخیص بیماری عصب بینایی برای تجزیه و تحلیل آزمایش ها از شبکه عصبی مصنوعی استفاده کردند. نتایج پایانی به عنوان سالم و بیمار طبقه بندی شدند. کومار و همکاران از جنگل های تصادفی وشبکه عصبی برای تشخیص بیماری سنگ کلیه استفاده کردند (۹). در این تحقیق با استفاده از شبکه عصبی چند لایه مدل چاقی افراد جامعه استخراج می شود. این کار با استفاده از داده

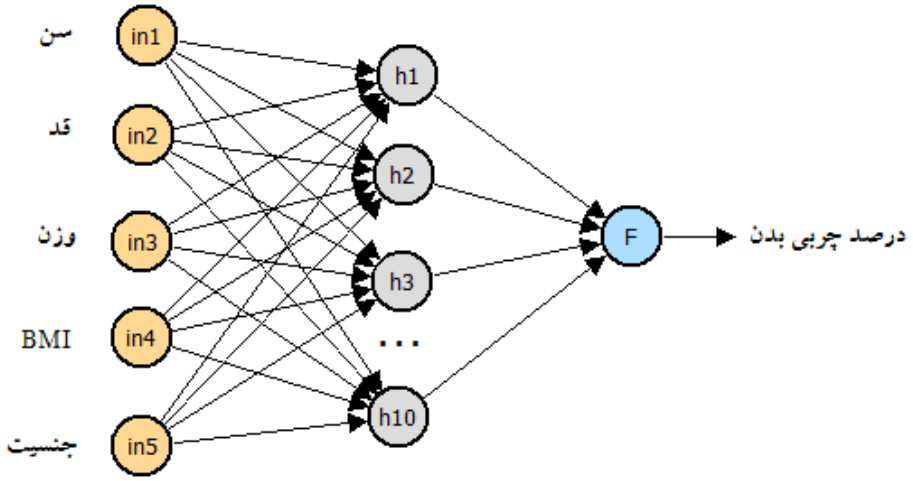
[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]

های جمع آوری شده از افراد مراجعه کننده به کلینیک تغذیه امکان پذیر شده است. نتایج نشان دهنده کارایی روش پیشنهادی است.

مواد و روش ها

داده های مورد استفاده در این تحقیق از بیماران مراجعه کننده به یکی از کلینیک ها تغذیه شهر تهران از تاریخ شهریور ۱۳۹۳ تا اسفند ۱۳۹۴ جمع آوری گردید. این پایگاه داده شامل ۴۰۰ رکورد داده مربوط به افراد مراجعه کننده بود. در پایگاه داده مورد استفاده، سعی شد که تمامی محدوده ها برای داده های موجود پوشش داده شود تا دقت مدل پیشنهادی به حداکثر برسد. به منظور تخمین میزان دقیق درصد چربی بدن، از یک شبکه عصبی پروسپترون با یک لایه‌ی مخفی استفاده شد. این شبکه عصبی در لایه ی مخفی خود دارای ۱۰ نرون بوده و تابع انتقال این لایه از نوع سیگموید است. تعداد نرون های لایه ورودی مطابق با تعداد خصوصیات هر نمونه و برابر با پنج تعیین گردید. هم چنین یک نرون در لایه خروجی قرار گرفت. ساختار این شبکه در شکل (۱) نمایش داده شده است.



شکل (۱)- ساختارشبکه عصبی پیشنهادی دراین تحقیق برای تخمین مقدار درصد چربی بدن

در مدل شبکه عصبی، پیش از آغاز عمل آموزش، داده های پایگاه داده نرمال سازی شدند. به منظور نرمال سازی داده ها، ابتدا صفت جنسیت که تنها صفت اسمی پایگاه داده بود با مقادیر عددی جایگزین گردید. مقدار «مذکر»، با عدد یک و «مؤنث» با عدد صفر جایگزین شدند. تمامی مقادیر پایگاه داده با استفاده از رابطه زیر نرمال سازی شدند.

رابطه (۱) : $Norm_i = \frac{X_i - \min_i}{\max_i - min_i}$

در رابطه (۱)، $\mathrm{Norm_i}$ مقدار نرمال شده صفت i ام، $\mathrm{X_i}$ مقادیر صفت i ام در پایگاه داده و $\mathrm{min_i}$ و $\mathrm{max_i}$ مقادیر ماکزیمم و مینیمم برای این صفت در پایگاه داده هستند. این رابطه تمامی مقادیر ورودی را به بازه‌ی [0,1] نگاشت می‌کند. پس از بهینه سازی داده ها، عمل آموزش شبکه عصبی انجام شد. الگوریتم آموزش شبکه عصبی با نزدیک کردن خطای خروجی شبکه به سمت صفر عمل یاد گیری شبکه را انجام داد. به منظور ارزیابی دقیق عملکرد مدل پیشنهادی از روش ارزیابی متقاطع (cross validation) که از معروف ترین روش های ارزیابی در حوزه داده کاوی است، استفاده شد. در این روش، داده ها به چندین بخش (به طور مثال n) شکسته شده و هر بار با n-1 بخش، آموزش داده شد و با یک بخش ارزیابی گردید. سپس میانگین نتایج مورد استناد قرار گرفت. در این تحقیق از روش ارزیابی متقاطع با n=10 استفاده شد. از پارامتر های مهمی که برای ارزیابی متقاطع محاسبه شد، میانگین مربعات خطا بود که با میانگین مجذور اختلاف بین مقدار درصد چربی واقعی بدن افراد و مقدار تخمین زده شده، برابر است که در مطالعه حاضر به وسیله رابطه (۲) محاسبه گردید:

رابطه (۲) : $MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$

[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]

در این رابطه، N برابر با تعداد نمونه های آزمایش شده (داده های تست در ده دور آزمایش در نظر گرفته شده است)، y_i برابر با درصد چربی تخمین زده شده برای نمونه i ام توسط روش پیشنهادی و $\hat{y}_i$ مقدار درصد چربی واقعی نمونه i ام می‌باشد. هم چنین معیار میانگین قدر مطلق خطا از طریق میانگین اختلاف بین مقدار درصد چربی واقعی بدن افراد و مقدار تخمین زده شده توسط روش پیشنهادی محاسبه شد و توسط رابطه (۳) تعیین گردید:

رابطه (۳): $MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$

هم چنین معیار مجموع مربعات خطا بیانگر مجذور خطای کلی تولید شده توسط سیستم است و با استفاده از رابطه (۴) محاسبه شد:

رابطه (۴): $SSE = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2$

قابل ذکر است که پیاده سازی مدل پیشنهادی و آزمایش آن در محیط نرم افزار MATLAB انجام شد.

یافته های پژوهش

پایگاه داده مورد استفاده برای ارزیابی عملکرد روش مطالعه پیشنهادی شامل ۴۰۰ رکورد داده در مورد اطلاعات جسمی افراد بود. که این اطلاعات در جدول (۱) نمایش داده شده است.

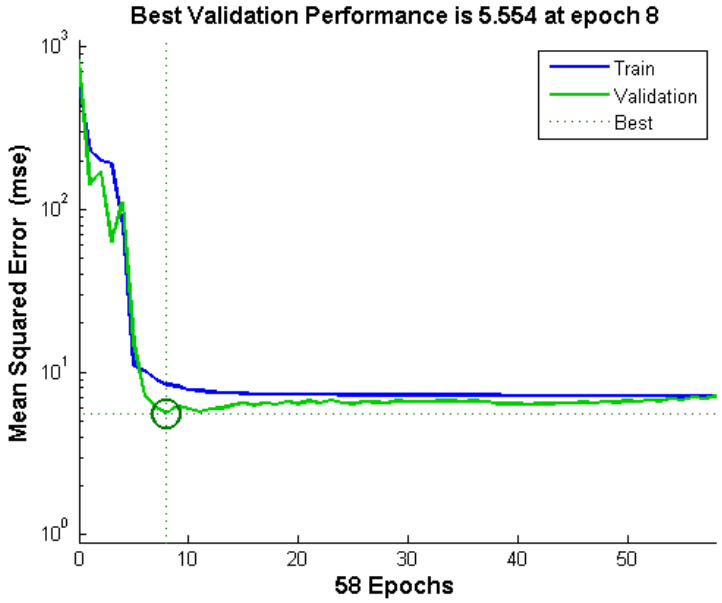
جدول (۱)- مشخصات پایگاه داده مورد استفاده برای آموزش وارزیابی شبکه عصبی

عنوان	بیشینه	کمینه	میانگین	انحراف معیار
سن (برحسب سال)	۷۷	۱۵	۳۵,۴	۱۰,۹
قد (سانتی متر)	۱۸۹	۱۴۳	۱۶۲,۴	۸,۱
وزن (کیلو گرم)	۱۶۳,۴	۳۹,۵	۸۰,۵	۱۷,۵
BMI (کیلو گرم/متر)	۵۰,۱	۱۴,۷	۳۰,۵	۶,۲
جنسیت	–	–	–	–
درصد چربی بدن	۵۲,۱	۲	۳۵,۲	۹,۷

در جدول شماره ۱ ، ستون عنوان مشخص کننده نام ویژگی های موجود در پایگاه داده و ستون بیشینه نشان دهنده بیش ترین مقدار موجود در پایگاه داده برای هر ویژگی است. هم چنین ستون کمینه کم ترین مقدار هر ویژگی در پایگاه داده را نشان می دهد و ستون میانگین مشخص کننده میانگین مقادیر ویژگی های نمایش داده شده است. سن فرد در زمان مراجعه به پزشک بر حسب سال شمسی بوده و مشخصات قد و وزن به ترتیب بر حسب سانتی متر و کیلو گرم می - باشند. BMI (Body Mass Index) نشان دهنده شاخص توده بدن است و با تقسیم وزن بدن بر مجذور قد هر فرد $\left(\frac{\text{وزن (کیلو گرم)}}{\text{قد}^2\text{ (متر)}}\right)$ محاسبه می شود. هم چنین پارامتر جنسیت یک متغیر اسمی است که افراد را در دو دسته مذکر و مؤنث تقسیم بندی می کند. در پایگاه داده مورد استفاده، جنسیت مؤنث به عدد ۰ و جنسیت مذکر با عدد ۱ مشخص شده است. هدف از انجام روش مطالعه پیشنهادی، رسیدن به کم ترین اختلاف بین میزان درصد چربی واقعی فرد و مقدار تخمین زده شده توسط مدل پیشنهادی می باشد. به همین منظور معیار های میانگین مربعات خطا و میانگین قدرمطلق خطا (Mean Absolute Error) برای ارزیابی کارایی مدل پیشنهادی، استفاده شد. میانگین مربعات خطا برای داده های آموزش و اعتبار سنجی در هر دوره تکرار آموزش شبکه عصبی پیشنهادی در نمودار (۱) نمایش داده شده است.

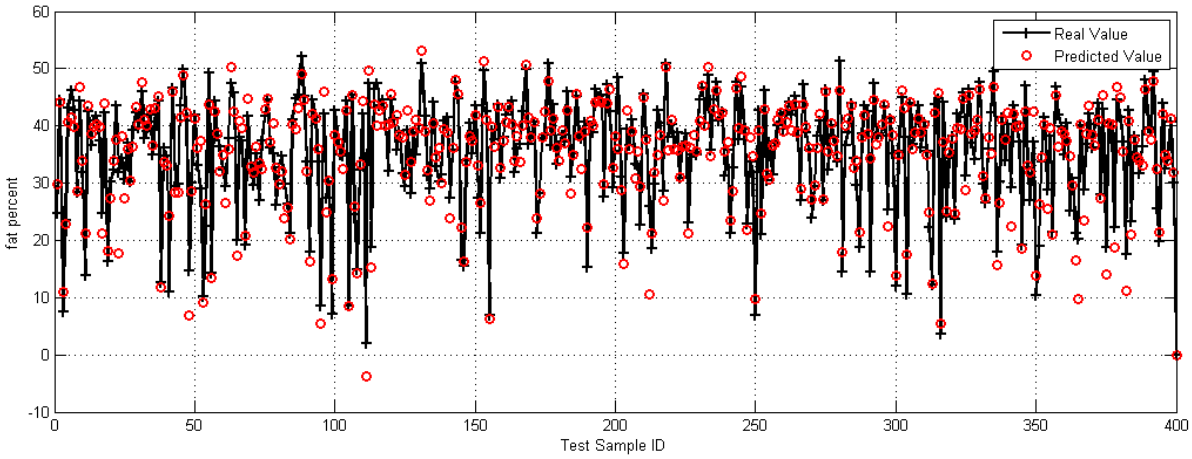
[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]



نمودار (۱)- نمودار تغییرات میانگین مربعات خطا در هر دوره تکرار آموزش شبکه عصبی

در این مطالعه، برای بهبود عملیات تخمین از داده های اعتبار سنجی نیز استفاده شده است. داده های اعتبار سنجی در حین عملیات آموزش و پس از انتهای هر دور آموزش، به منظور میزان پیشرفت فرآیند آموزش شبکه عصبی استفاده می شود. نمودار (۱) نشان می دهد که میانگین مربعات خطای شبکه عصبی پس از ۸ دور تکرار به مقدار کم ترین مقدار خود یعنی ۵،۵۵۴ رسیده است. پس از چرخه ۵۸ شبکه عصبی خطای بیش تری را برای داده های اعتبار سنجی تولید کرده و نتیجه آموزش خاتمه یافته است. در نمودار (۲) داده های مربوط به درصد چربی بدن افراد و مقدار تخمین زده شده توسط شبکه عصبی در تمام آزمایش ها نمایش داده شده است. این نتایج، حاصل ۱۰ بار تکرار عملیات آموزش و آزمون شبکه عصبی می باشد که در انتهای آزمون تمامی داده ها با هم ترکیب شده و نتایج تکرار های مختلف درکنار هم نمایش داده شده است.



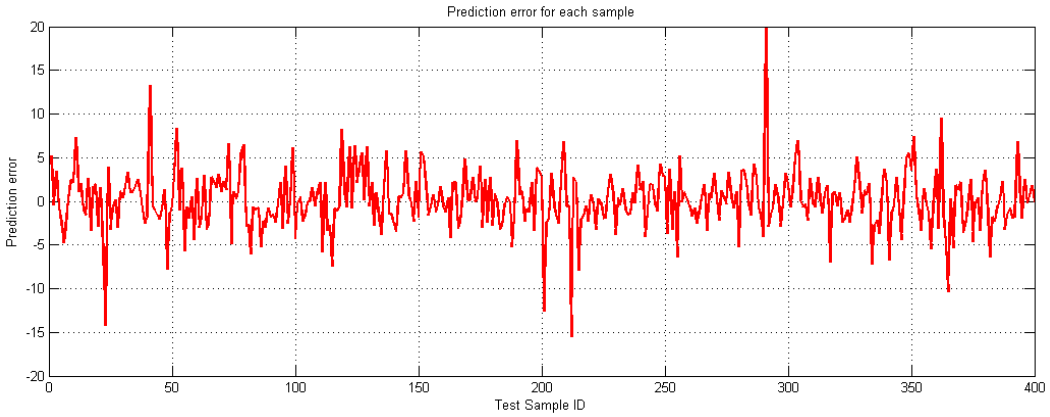
نمودار (۲)- داده های مربوط به درصد چربی بدن افراد و مقدار تخمین زده شده توسط شبکه عصبی

در نمودار (۲)، مقادیر واقعی به صورت نقاط مشکی و مقادیر خروجی شبکه عصبی به صورت دایره نمایش داده شده است. این نقاط بر هم منطبق بوده و میزان خطای شبکه عصبی در تخمین دقیق درصد چربی بدن نا چیز می باشد. برای محاسبه خطای تخمین، میزان واقعی درصد چربی نمونه ها را از مقدار تخمین زده

[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]

شده توسط شبکه عصبی کسر شده اند که مقادیر حاصل، خطای شبکه عصبی در تخمین میزان دقیق چربی بدن به ازای هر نمونه خواهد بود و در نمودار (۳) قابل مشاهده است.



نمودار (۳)- خطای تخمین درصد چربی بدن به ازای هر نمونه آزمون

نتایج حاصل از نمودار (۳) نشان می دهد که در بد ترین حالت شبکه عصبی مورد استفاده در مطالعه حاضر، درصد چربی بدن افراد را با ۱۹/۸ درصد اختلاف و به طور میانگین با اختلاف ۲/۵ واحد با دقت ۹۳٪ تخمین می زند. با نگاه دقیق تر به عملکرد روش پیشنهادی در تخمین درصد چربی بدن، از طریق محاسبه پارامتر های خطا امکان پذیر است که در جدول (۲) پارامتر های استخراج شده برای خطا در ۱۰ تکرار آزمایش انجام شدند.

جدول (۲)- خطای شبکه عصبی برای ۱۰ بار تکرار آزمایش

پارامتر	*مقدار*
میانگین خطا	۲,۵
میانگین مربعات خطا (MSE)	۱۱,۶۰
میانگین قدر مطلق خطا (MAE)	۲,۵
مجموع مربعات خطا (SSE)	۴۶۴۰

نتایج جدول (۲) نشان می‌دهد که میانگین خطای شبکه ۲,۵ واحد برای درصد چربی بدن است که با توجه به میانگین ۳۵,۲ واحدی درصد چربی بدن افراد، خطایی قابل قبول بوده و بنا بر این مدل پیشنهادی دارای دقت بسیار خوب (۹۳٪) می باشد.

بحث و نتیجه گیری

با آگاهی از این موضوع که وجود چربی بیش از حد استاندار در بدن انسان، عامل به وجود آورنده بسیاری از بیماری ها می باشد، لذا محاسبه ی درصد چربی بدن انسان بسیار با اهمیت خواهد بود که پزشکان با استفاده از تجهیزات خاص به محاسبه آن می پردازند. بنا بر این، ابزار پیش گویی کننده درصد میزان چربی، نقش مهمی در سلامت افراد دارند و در پیشگیری بسیاری از بیماری های مزمن موثر واقع می شوند. استفاده از هوش مصنوعی و ابزار یاد گیری ماشین برای پیش بینی درصد چربی بدن انسان، مطالعات اندکی در دنیا انجام شده است. در سال ۲۰۱۳ آقای کاپوسینک و همکاران از شبکه عصبی برای پیش بینی درصد چربی بدن انسان استفاده کردند که خصوصیت های ورودی شبکه آن ها سه مورد جنسیت، سن و BMI بود. درصد دقت کاراین مطالعه، میزان ۸۰,۴۳ ٪ تخمین داده شد (۱۰). در مطالعه حاضر ار پنج خصوصیت به عنوان ورودی شبکه استفاده و دقت کار به میزان ۹۳٪ به دست آمد. با وجود آن که تعداد نمونه های آقای کاپوسینک و همکاران خیلی بیش تر (در حدود ۲۷۰۰) از مطالعه اخیر است اما دارای دقت کم تری است که دلیل آن انتخاب تعداد خصوصیات ورودی شبکه است که در تحقیق حاضر، بیش تر است.

[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]

تحقیقات متعددی برای محاسبه درصد میزان پربی در دنیا انجام شده است اما در هیچ کدام، از روش های هوش مصنوعی استفاده نشده است. کوهن و همکاران از روش بایو امپدانس برای پیش بینی میزان چربی در دانش آموزان هندی استفاده کردند که این روش ار مقایسه با روش مطالعه حاضر، مستلزم ابزار و صرف هزینه است (۱۱). در مطالعه نورادیلاه و همکاران در کشور مالی محاسبه درصد میزان چربی ۱۶۰ دانش آموز بر اساس روش اندازه گیری ضخامت پوست و تحلیل امپدانس بیو الکتریک انجام شد که روشی بسیار وقت گیر و پر هزینه بود (۱۲). به منظور تخمین میزان دقیق درصد چربی افراد در روش مطالعه حاضر، از یک شبکه عصبی پروسپترون با یک لایه مخفی استفاده شد که توانست عمل تخمین درصد چربی بدن را با میانگین خطای ۲,۵ واحدی اندازه گیری کند. در این روش با اخذ داده های سن، قد، وزن، جنسیت و BMI از مریض، شبکه عصبی آموزش دیده درصد چربی را محاسبه کرد. نتایج این مطالعه نشان داد که مدل استفاده شده برای تخمین دقیق درصد چربی بدن انسان از دقت بسیار بالایی برخوردار است و به عنوان یک ابزار کار آمد اندازه گیری پیشنهاد می شود. با استفاده از این مدل، دیگر نیازی به استفاده از دستگاه BIA یا روش های مشابه دیگر، نیست. به منظوراندازه گیری میزان واقعی چربی بدن، ادامه تحقیقات مطالعه اخیر در آینده با استفاده از الگوریتم‌های خوشه بندی پیشنهاد می شود. هم چنین استفاده از سایر الگوریتم های یاد گیر مانند ماشین بردار پشتیبان یا شبکه های عصبی فازی (Fuzzy Neural Network) ، در زمینه سنجش درصد چربی بدن می توانند سودمند باشند. به طور کلی استفاده از تکنیک های یاد گیری در علم پزشکی بسیار مفید بوده و استفاده از آن ها تشخیص بیماری ها یا سیستم های توصیه گر پیشگیری از بیماری بر اساس مشخصات افراد، پیشنهاد می گردد.

References

1.Myint PK, Kwok CS, Luben RN, Wareham NJ, Khaw KT. Body fat percentage body mass index and waist-to-hip ratio as predictors of mortality and cardiovascular disease. Heart 2014 Oct;100:1613-9.

2.Mohammadi S, Shakerhosseini R, Rastmanesh R, Jafarian K, Amiri Z, Jahangir F. [Effects of melatonin supplementation on weight and body fat mass percentage in overweight or obese people]. JQUMS 2015; 19:24-31. (Persian)

3.Hand D, Mannila H, Smyth P. Principles of Data Mining.1th ed. MIT Publication.2001; P.128-15.

4.Han J, Kamber M, Pei J. Data Mining concepts and techniques. 3th ed. Elsevier Publication. 2012; P.368- 75.

5.Soni J, Ansari U, Sharma D, Soni S. Predictive data mining for medical diagnosis an overview of heart disease prediction. IJCSE 2011; 17:43-8.

6. Rajkumar A, Reena GS. Diagnosis of heart disease using data mining algorithm. GJCST 2010; 10:38-43.

7.Dursun D, Walker G, Amit, K. Predicting breast cancer survivability: a comparison of three data mining methods. Art Intell Med 2005; 34:113-127.

8.Kara S,Guvenb A, Oner AO. Utilization of artificial neural networks in the diagnosis of optic nerve diseases. Comp Bio Med 2006; 36:428-37.

9.Kumar M. Prediction of chronic kidney disease using random forest machine learning algorithm. Int J Comp Sci Mob Com 2016; 5:24-33.

10. Kupusinac A,Stokic E, Doroslovacki R, Predicting body fat percentage based on gender age and BMI by using artificial neural networks. Comput Meth Prog Biomed 2014; 113: 610-9.

11.Kehoe SH, Krishnaveni GV, Lubree HG, et al. Prediction of body fat percentage from skinfold and bio impedance measurements in Indian school children. European J Clin Nutrition 2011; 65:1263-70.

12. Noradilah MJ. Assessing body fat of children by skinfold thickness, bioelectrical impedance analysis, and dual energy x-ray absorptiometry a validation study among malay children aged 7 to 11 years. Asia Pac J Public Health 2016; 28: 74-84.

[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]

Stimated Body Fat Percentage using Mechine Learning Techniques

Asgari R[1], Valizadeh M[2], Jafarian K[3]*

(Received: September 10, 2016 Accepted: March 5, 2017)

Abstract

***Introduction*:** Doctors undertake calculation of body fat percentage by using BIA (Bioelectrical Impedance Analysis) equipment. In this study, we measured body fat percentage without using equipment. For this purpose, an artificial neural network has been used to estimate the exact amount of fat.

***Materials & methods*:** The sample was selected from patients admitted in a nutrition clinic in Tehran. 400 patients took part in this study. MLP neural network was used to estimate body fat percentage. The used neural network had five input neurons and ten neurons in the hidden layer. Also, cross validation method for evaluating the proposed method has been used.

***Findings*:** The proposed method is efficient because of the results that demonstrate 2.5 units error based on cross validation. The results of experiments show that the proposed neural network for estimating body fat percentages has an average accuracy of 93%. Therefore the proposed method can accurately estimate body fat percentage of people with very high accuracy.

***Discussion & conclusions*:** The results of this research show that the proposed method as the first method used in machine learning technique, can estimate fat percentage with high accuracy. This method can be used as a useful method without using BIA device.

***Keywords*:** Learning algorithm, Body fat percentage, Data mining, Neural network

1. Dept of Computer, Islamic Azad University, Ilam Branch, Ilam, Iran
2. Dept of Computer, Faculty of Engineering, Ilam University, Ilam, Iran
3. Dept of Nutrition, Faculty of Nutrition and Dietetics, Tehran University of Medical Sciences, Tehran, Iran

[Downloaded from sjimu.medilam.ac.ir on 2024-09-27]

[DOI: 10.29252/sjimu.25.4.171]