مقایسه مدل شبکه عصبی مصنوعی و رگرسیون کاکس در پیش بینی بقای بیماران مبتلا به سرطان پستان

مسعود صالحی*[1]، محمود رضا گوهری[1]، نسیم وهابی[1]، فرید زایری[2]، سیدحسین یحیی زاده[3]، محمدرضا کفاشیان[4]

1) گروه آمار و ریاضی، مرکز تحقیقات علوم مدیریت اقتصاد دانشکده مدیریت و اطلاع رسانی پزشکی، دانشگاه علوم پزشکی ایران
2) گروه آمار زیستی، مرکز تحقیقات پروتئومیکس، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی تهران
3) گروه پرتودرمانی - آنکولوژی، مرکز تحقیقات بالینی، بیمارستان عمومی میلاد(سازمان تامین اجتماعی) تهران
4) گروه فیزیولوژی، دانشکده پزشکی، دانشگاه علوم پزشکی ایلام

تاریخ دریافت: 91//1/18 تاریخ پذیرش: 92/2/16

چکیده

مقدمه: امروزه انواع سرطان یکی از مهم ترین عوامل مرگ و میر در دنیا و سرطان پستان از شایع ترین آن ها در زنان میان سال می باشد. میزان بقای پس از تشخیص و درمان در این بیماران یکی از شاخص های مهم در کنترل بیماری است. در این مطالعه دو مدل رگرسیون کاکس و شبکه عصبی مصنوعی در پیش بینی بقای بیماران سرطان پستان با یکدیگر مقایسه شده اند.

مواد و روش ها: داده های این پژوهش که از نوع مطالعات بقا است، از پرونده تمامی 344 زن مبتلا به سرطان پستان که در سال های 1384 تا 1391 برای درمان به بیمارستان شهید فیاض بخش تهران مراجعه کرده بودند استخراج شده است. وضعیت بقا به عنوان متغیر وابسته در نظر گرفته شد. مقایسه پیش بینی های دو مدل با استفاده از سطح زیر منحنی مشخصه عملکرد و صحت پیش بینی کل و به کمک نرم افزارهای R و Minitab انجام شد.

یافته های پژوهش: میانگین ± انحراف معیار سن زمان تشخیص بیماران 10/93±49/9 سال و میانه بقا برابر با 44/6 ماه بود و تا پایان مطالعه واقعه مرگ برای تعداد 45 نفر(13/1 درصد) از آن ها رخ داد. سطح زیر منحنی مشخصه عملکرد مدل کاکس و شبکه عصبی به ترتیب برابر 87/6 درصد و 75/4 درصد و هم چنین صحت پیش بینی کل برای مدل کاکس و شبکه عصبی به ترتیب برابر 89/42 و 77/68 محاسبه گردید.

بحث و نتیجه گیری: صحت پیش بینی شبکه در تشخیص وضعیت بقای زنان مبتلا به سرطان پستان بالاتر بود و بنا بر این این مدل برای پیش بینی بقای بیماران سرطان پستان پیشنهاد می شود.

واژه های کلیدی: سرطان پستان، بقا، رگرسیون کاکس، شبکه عصبی مصنوعی

***نویسنده مسئول:** گروه آمار و ریاضی، دانشکده مدیریت و اطلاع رسانی پزشکی، دانشگاه علوم پزشکی تهران

Email: salehi74@yahoo.com

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

مقدمه

انواع سرطان ها را می توان به عنوان مهم ترین بیماری های مزمن گریبان گیر بشر امروزه دانست،که عوامل فراوانی در بروز آن نقش داشته و در دهه های اخیر در بسیاری از نقاط دنیا میزان مرگ و میر بالایی را نشان می دهند،(۲،۱). بار جهانی سرطان ها در سالیان آتی روندی رو به رشد داشته و انتظار می رود تعداد موارد جدید آن از ده میلیون نفر در سال 2000 میلادی به 15 میلیون نفر در سال 2020 افزایش یابد که در حدود 60 درصد از این موارد در کشورهای در حال توسعه رخ خواهند داد،(۳،۴). گزارش ها و مطالعات مختلف نشان می دهند که سرطان پستان شایع ترین سرطان در میان زنان ایرانی است،(8-4). سرطان پستان یکی از شایع ترین عوامل مرگ و میر زنان میان سال(نه جوان تر از 25 سال) هم در کشورهای توسعه یافته و هم در کشورهای در حال توسعه است، از این رو نرخ شیوع آن در 50 سالگی و بالاتر به سرعت افزایش می یابد. در سالیان اخیر نرخ مرگ و میر سرطان پستان در تمامی کشورهای جهان به سرعت افزایش یافته است،(۱۰،۹). سرطان پستان، پنجمین سرطان کشنده پس از سرطان معده، لوسمی، ریه و برونش، کبد و مجاری صفراوی در زنان ایرانی به شمار می رود که باعث مرگ 1200 زن در هر سال می گردد و میزان آن 3/4 در هـــر صـــد هـــزار زن ایرانی است.(11،12)

میزان بقای پس از تشخیص و درمان در بیماران سرطانی یکی از شاخص های مهم در کنترل بیماری و ارزیابی روش های درمانی مورد استفاده است. روش رگرسیون کاکـــس یکی از پرکاربردترین روش های آماری تحلیل بقا بوده که دارای پیش فرض هایی برای مدل سازی است و می توان از مهم ترین آن ها به پیش فرض متناسب بودن مـــخاطـــرات و هم چنین استقلال زمان هـــای رخداد پیـــشامدها اشاره نمود. برقراری این پیـــش فرض هـــا در غالـــب داده های گردآوری شده محدودیت هایی را برای برازش این مدل و مدل های مشابه ایجاد می کند،(15-13). از این رو مدل هایی که به این پیش فرض ها بستگی نداشته باشند در دهه های اخیر توسعه یافته اند که یکی از مهم ترین آن ها را مدل شبکه عصبی مصنوعی دانست،(16). امروزه شبکـــه های عصبی مصنوعی در تحلیل داده های پزشکی مربوط به تشخیص بالینی، تجزیه و تحلیل تصاویر پزشکی، پیش بینی بقا در انواع سرطان ها و جراحی ها و ... به کار گرفته می شوند.(22-17)

روزنـــبلات در سال 1958 مـــیلادی شـــبکه پرسپترون را معرفی کرد که تشکیل شده از لایه هایی است که اجزای ساده پردازش مرتبط با یکدیگر به نام نرون را شـــامل می شـــود. به عبارت دیگر نرون کوچک ترین بخش یک شبکه عصبی مصنوعی است که به عنوان یک واحد پردازش اطلاع اساس شبکه عصبی را تشکیل می دهد. هم چنین یک لایه در شبکه عصبی مصنوعی شامل نرون های است که در یک سطح قرار دارند. هر شبکه عصبی حداقل دارای سه لایه ورودی، میانی و خروجی است و هر لایه وزنی دارد که نشان دهنده میزان تاثیر دو نرون بر یکدیگر است. هر نرون دارای یک تابع فعالیت است که در آموزش شبکه عصبی به کار گرفته می شود. در شبکه پرسپترون با استفاده از الگوریتم یادگیری پس انتشار خطا، میانگیـــن مربعـــات خطای خروجی به کمک روش های عددی حداقل می شود.(۱۶،۲۳،۲۴)

با توجه به اهمیت و شیوع بالای سرطان پستان در زنان و هم چنین اهمیت زیاد ارزیابی میزان بقا در این بیماران و تعیین عوامل موثر بر بقای بیماران سرطان پستان، در این مـــقاله به مـــقایسه دو مدل رگرسیون کاکس و شبکه عصبی مصنوعی در پیش بینی میزان بقای این بیماران می پردازیم.

مواد و روش ها

داده های این پژوهـــش که مطالعه ای از نوع مطالعات بقا است، از پرونده تمامی 344 زن مبتلا به سرطان پستان که در سال های 1384 تا 1391 برای درمان به بیمارستان شهید فیاض بخش تهران مراجعه کرده بودند استخراج شده و به همراه آخرین وضعیت آن ها تا پایان مطالعه در فرم ثبت اطلاع ذخیره گردید. متغیرهای ثبت شده برای برازش مـــدل بقا در این بیماران پس از عمل جراحی شامل مدت زمان بقا پس از عمل جراحی، سن زمان تشخیص بیمار، اندازه تومور،

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

درجه بدخیمی تومور، تعداد گره های درگیر، گیرنده پروژسترون(PR)، گیرنده استروژن(ER) و گیرنده فاکتور رشد اپیدرمی انسانی(HER2) بود که پس از استخراج در یک صفحه گسترده Excel وارد شدند. شرط ورود بیماران به مطالعه، نداشتن متاستاز به هنگام پذیرش در اولین مراجعه و قرار گرفتن تحت عمل جراحی ماستکتومی رادیکال یا جراحی نگهدارنده پستان و پس از آن دریافت درمان های کمکی بود. تجزیه و تحلیل داده ها نیز با استفاده از مدل رگرسیون کاکس و مدل شبکه عصبی مصنوعی به کمک نرم افزارهای R و Minitab انجام شد.

برای برازش شبکه عصبی مصنوعی، داده ها به دو بخش آموزش و آزمایش شبکه تقسیم شدند که داده های آموزش شبکه در حدود 70 درصد(240 مشاهده) از داده های مطالعه را تشکیل دادند. همگنی منحنی های بقای داده های بخش آموزش و آزمایش به کمک آزمون کاپلان-مایر ارزیابی شده و در ادامه یک شبکه عصبی مصنوعی پرسپترون سه لایه آموزش داده شد. معماری شبکه آموزش دیده متشکل از سه لایه ورودی، میانی و خروجی بوده که لایه ورودی متشکل از متغیرهای مستقل مورد بررسی، لایه دوم تحت عنوان لایه میانی یا مخفی و لایه خروجی شامل رخداد پیشامد مرگ و یا عدم رخداد آن با دو نرون است. برای انتخاب بهترین ساختار شبکه عصبی از ملاک اطلاع شبکه (Network Information Criterion: NIC) و سطح زیر منحنی مشخصه عملکرد (Receiver Operating Characteristic Curve: ROC Curve) استفاده شد. در نهایت پس از انتخاب بهترین ساختار، شبکه مذکور با داده‌های مرحله دوم که در آموزش شبکه نقشی نداشته اند، مورد ارزیابی و اعتبار سنجی واقع شد. پیش بینی های حاصل از این مرحله با پیش بینی‌های حاصل از مدل رگرسیون کاکس و به کمک منحنی مشخصه عملکرد مورد مقایسه قرار گرفت. شبکه عصبی پرسپترون چند لایه (Multi -Layer Perceptron: MLP) به طور کلی دارای تعدادی نود با تابع فعالیت مشخص در لایه های مختلف است. هر نود با وزن مشخص خود، خروجی تمامی نودهای لایه های قبلی را دریافت کرده و با توجه به تابع فعالیت در لایه مذکور، خروجی ها را به لایه بعدی ارسال می نماید. خروجی i ام در یک شبکه عصبی MLP با یک لایه ورودی، میانی و خروجی و تعداد گره های متفاوت در هر لایه، به صورت زیر به دست می آید.

$$\hat{y} = f_2\left[b_0 + \sum_{j=1}^{M} w_j f_1\left(b_{j0} + \sum_{s=1}^{p} x_{is} w_{js}\right)\right] \qquad i = 1,2,\ldots,n$$

که در آن f_1 و f_2 به ترتیب توابع فعالیت لایه میانی و خروجی، bj_o و b_o به ترتیب میزان اریبی نودهای لایه میانی و خروجی، W_j وزن متناسب با گره j ام لایه میانی، W_{js} وزن متناسب با ورودی X_{is} در گره j ام، n تعداد مشاهدات، p تعداد نودهای لایه ورودی و M تعداد نودهای لایه میانی است.

یکی از اهداف مهم در شبکه عصبی یافتن وزن متناسب با لایه های مختلف و در واقع برآورد پارامترهای شبکه عصبی است. الگوریتم پس انتشار خطا روشی مرسوم برای محاسبه وزن های شبکه است که از دو مسیر رفت و برگشت قابل محاسبه است و این رفت و برگشت تا دستیابی به بهترین برآورد از پارامترهای شبکه تکرار می گردد و به عنوان فرایند آموزش از آن یاد می شود.

یافته های پژوهش

در این مطالعه همگروهی تعداد 344 زن مبتلا به سرطان پستان مورد بررسی قرار گرفتند که میانگین±انحراف معیار سن زمان تشخیص این بیماران 10/93±49/9 سال، میانگین±انحراف معیار وزن 15/60±73/1 کیلوگرم، میانگین±انحراف معیار قد 6/49±156/1 سانتی متر به دست آمد. هم چنین، میانه بقا در این بیماران برابر با 44/6 ماه با حداقل دو ماه و حداکثر 78 ماه محاسبه شد و تا پایان مطالعه واقعه مرگ برای تعداد 45 نفر(13/1 درصد) از آن ها رخ داد. توزیع فراوانی سایر متغیرهای مستقل مورد بررسی در این مطالعه در جداول شماره 1 و 2 ارائه شده است.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

جدول شماره 1. توزیع فراوانی متغیرهای وضعیت بیماری در زنان مبتلا به سرطان پستان مراجعه کننده به بیمارستان شهید فیاض بخش از سال 1384 تا 1391

متغیر	سطح	تعداد	درصد
درجه بدخیمی تومور	I	43	12/5
	II	177	51/5
	III	124	36/0
اندازه تومور	کمتر از دو سانتیمتر	62	18/0
	بین دو تا پنج سانتیمتر	169	49/1
	بیش از پنج سانتیمتر	87	25/3
	تهاجم به پوست و قفسه سینه	26	7/6
تعداد گره های درگیر	صفر	108	31/4
	یک	92	26/7
	دو	72	20/9
	سه و بیشتر	31	14/4
	نامشخص	41	11/9
متاستاز	دارد	14	4/1
	ندارد	276	80/2
	نامشخص	54	15/7

جدول شماره 2. توزیع فراوانی وضعیت گیرنده های پروژسترون، استروژن و فاکتور رشد اپیدرمی انسانی در زنان مبتلا به سرطان پستان مراجعه کننده به بیمارستان شهید فیاض بخش از سال 1384 تا 1391

متغیر	سطح	تعداد	درصد
گیرنده استروژن (ER)	مثبت	213	61/9
	منفی	110	32/0
	نامشخص	21	6/1
گیرنده پروژسترون(PR)	مثبت	207	60/2
	منفی	116	33/7
	نامشخص	21	6/1
گیرنده فاکتور رشد اپیدرمی انسانی(HER2)	مثبت	103	29/9
	منفی	212	61/6
	نامشخص	29	8/4

برای مدل سازی شبکه عصبی مصنوعی پرسپترون سه لایه، با توجه به مجموعه داده های گروه آموزشی و متغیرهای مستقل ورودی تعداد 387 ساختار مبنی بر 6 تا 15 نود در لایه میانی، با اندازه حرکت 0/80 تا 0/95، نرخ یادگیری 0/01 تا 0/40 با الگوریتم آموزش شیب توام مقیاس شده (Scaled Conjugate Gradiant) و تابع پیوند لوجستیک مورد ارزیابی قرار گرفت. پس از بررسی تمامی مدل های ممکن برای ساختار شبکه عصبی سه لایه در نهایت شبکه با 24 نود ورودی، 9 نود میانی، 2 نود خروجی، نرخ یادگیری 0/05 و الگوریتم پس انتشار خطا به عنوان بهترین شبکه عصبی انتخاب شد. ملاک اطلاع شبکه 795/3 درصد و میزان پیش بینی های صحیح مدل 89/42 به دست آمد. هم چنین، مدل رگرسیونی کاکس نیز به کمک روش پسرو به داده ها برازش داده شد و معیار اطلاع آکائیکه در مدل نهایی برابر 1063/9 به دست آمد. در جدول شماره 3، نتایج حاصل از اجرای دو مدل رگرسیونی کاکس و شبکه عصبی مصنوعی برای تعیین میزان اهمیت متغیرهای مستقل مورد بررسی بر اساس ضریب تاثیر آن ها (اهمیت استاندارد شده در شبکه عصبی و مقدار احتمال در رگرسیون کاکس) ارائه شده است.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

جدول شماره 3. نتایج رگرسیون کاکس و شبکه عصبی مصنوعی برای بررسی عوامل پیش آگهی دهنده بقای زنان مبتلا به سرطان پستان مراجعه کننده به بیمارستان شهید فیاض بخش از سال 1384 تا 1391

مدل شبکه عصبی مصنوعی		مدل کاکس	
متغیر مرتب شده	اهمیت استاندارد شده	متغیر مرتب شده	مقدار احتمال
متاستاز	100/0	تعداد گره های درگیر	0/005
درجه بد خیمی تومور	88/8	درجه بدخیمی تومور	0/028
تعداد گره های درگیر	78/6	گیرنده فاکتور رشد اپیدرمی انسانی(HER2)	0/039
گیرنده فاکتور رشد اپیدرمی انسانی(HER2)	63/1	متاستاز	0/044
اندازه تومور	55/6	سن زمان تشخیص	0/118
گیرنده استروژن (ER)	19/2	قد	0/387
گیرنده پروژسترون(PR)	16/9	گیرنده استروژن (ER)	0/447
سن زمان تشخیص	13/6	وزن	0/746
قد	4/2	گیرنده پروژسترون(PR)	0/818
وزن	3/4	اندازه تومور	0/855

در جدول شماره 3، با توجه به اهمیت استاندارد شده بالای 20 درصد در شبکه عصبی مصنوعی سه لایه به کار گــرفته شده، متغیرهای متاستاز، درجه بدخیمی تومور، تعداد گره های درگیر، گیرنده فاکتور رشد اپیدرمی انسانی و اندازه تومور، بیشترین اهمیت را نشان دادند در حالی که در مدل رگرسیونی کاکس متغیرهای تعداد گره های درگیر، درجه بدخیمی تومور، گیرنده فاکتور رشد اپیدرمی انسانی و متاستاز با توجه به مقدار احتمال کمتر از 0/05 بیشترین اهمیت را دارا بودند.

برای بررسی صحت پیش بینی های به دست آمده در دو مدل، معیارهای سطح زیر منحنی مشخصه عملکرد و صحت کلاس بندی به کار گرفته می شوند. مقدار صفر تا 0/5 سطح زیر منحنی مشخصه عملکرد نشان دهنده کلاس بندی تصادفی و مقادیر بیش از 0/5 تا یک، بیانگر توان مدل در تشخیص سطوح مختلف است. سطح زیر منحنی مشخصه عملکرد برای مدل شبکه عصبی مصنوعی و رگرسیون کاکس به ترتیب برابر 87/6 درصد و 75/4 درصد به دست آمد. صحت کلاس بندی نیز بیان کننده نسبتی از مواردی است که در هر گروه به درستی دسته بندی شده اند و طبیعتا مقادیر بزرگ تر این معیار نشان دهنده بالاتر بودن قدرت تشخیصی مدل است. صحت کلاس بندی دو مدل در جدول شماره 4 ارائه شده است.

جدول شماره 4. صحت طبقه بندی مدل های رگرسیون لجستیک و شبکه عصبی مصنوعی با استفاده از مجموعه داده آزمایشی زنان مبتلا به سرطان پستان مراجعه کننده به بیمارستان شهید فیاض بخش از سال 1384 تا 1391

وضعیت	مشاهده شده	پیش بینی شده رگرسیون کاکس		پیش بینی شده شبکه عصبی مصنوعی	
		تعداد	درصد	تعداد	درصد
مرگ	88	71	81/63	84	95/15
بقا	16	11	67/31	13	81/60
کل	104	82	77/68	97	89/42

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

با توجه به جدول فوق، درصد پیش بینی کل برای مدل کاکس و شبکه عصبی مصنوعی به ترتیب برابر 89/42 و 77/68 است. بنا بر این میزان پیش بینی های صحیح صورت گرفته توسط شبکه عصبی مصنوعی بالاتر از مدل رگرسیون کاکس بوده و این نشان دهنده آن است که شبکه عصبی دسته بندی صحیح تری نسبت به رگرسیون کاکس انجام می دهد. هم چنین با توجه به مقادیر کلی گزارش شده در جدول شماره 4، میزان پیش بینی های نادرست (Incorrect Prediction) برای شبکه عصبی و مدل رگرسیون کاکس به ترتیب برابر با 10/85 و 22/32 بوده و نشان دهنده بالاتر بودن خطای مدل رگرسیون کاکس نسبت به شبکه عصبی مصنوعی می باشد.

بحث و نتیجه گیری

امروزه بیماری های سرطانی از جمله مهم ترین عوامل مرگ و میر و مشکلات بهداشتی در جهان به شمار می روند،(۱،۲). در ایران سرطان پستان جزو شایع ترین سرطان ها در میان زنان و اولین علت مرگ و میر ناشی از سرطان در آن ها است که تاکنون مطالعات فراوانی در این زمینه انجام شده است،(8-4). در مطالعه حاضر از دو مدل رگرسیون کاکس و شبکه عصبی مصنوعی سه لایه پرسپترون برای بررسی عوامل موثر بر بیماران سرطان پستان استفاده شد. نتایج هر دو مدل نشان داد که متغیرهای متاستاز، تعداد گره های درگیر، درجه بدخیمی تومور و گیرنده فاکتور رشد اپیدرمی انسانی اهمیت بالایی در پیش آگهی مرگ این بیماران برخوردار بوده اند. خدابخشی و همکاران نیز در مطالعه که به منظور تعیین عوامل موثر بر بقای بدون بیماری سرطان پستان انجام دادند، نشان دادند که متغیرهای تعداد گره های درگیر، درجه بدخیمی تومور و گیرنده فاکتور رشد اپیدرمی انسانی اثر آماری معنا داری بر بقای این بیماران داشته اند،(26). هم چنین یغمایی و همکاران نیز در مطالعه مشابهی در شهر سمنان نشان دادند که متغیرهای درجه بدخیمی تومور و سن تاثیر معناداری بر بقای بیماران سرطان پستان دارند،(27). بختیاری و حاجیانی نیز در مطالعه دیگری در شهر بابلسر بقای پنج ساله بیماران سرطان پستان را مورد بررسی قرار داده و نتایج ارائه شده نشان دهنده معنادار بودن اثر متغیرهای سن، درجه بدخیمی تومور و تعداد گره های درگیر بر بقای بیماران بود. گرچه برخلاف نتایج به دست آمده در مطالعه حاضر، متغیر اندازه تومور اثر معناداری را بر بقای این بیماران نشان نداد،(28). مطالعات مختلف دیگری نیز میزان اهمیت و معناداری اثر دو متغیر درجه بدخیمی تومور و گیرنده فاکتور رشد اپیدرمی انسانی را نشان داده اند.(۲۹،۳۰)

به علاوه در این مطالعه نتیجه حاصل از مقایسه مدل رگرسیونی کاکس و شبکه عصبی مصنوعی سه لایه پرسپترون نشان دهنده برتری نسبی قدرت تشخیص شبکه عصبی مصنوعی در پیش بینی بقای بیماران سرطان پستان بود. مطالعات متعددی در زمینه مقایسه شبکه عصبی مصنوعی با سایر مدل های کلاسیک آماری انجام شده است. از این بین می توان به مطالعه بیگلریان و همکاران در رابطه با مقایسه شبکه عصبی مصنوعی و رگرسیون کاکس برای پیش بینی بقای بیماران سرطان معده اشاره نمود که نتایج آن نشان دهنده بالاتر بودن قدرت تشخیصی شبکه عصبی مصنوعی بود،(31). گوهری و همکاران نیز برای بررسی عوامل پیش آگهی دهنده در بقای بیماران سرطان کولورکتال، دو مدل رگرسیون کاکس و شبکه عصبی مصنوعی را مورد مقایسه قرار دادند که نتایج این مطالعه نیز حاکی از برتری شبکه عصبی مصنوعی نسبت به مدل رگرسیون کاکس بود،(32). چین و همکاران نیز مدل شبکه عصبی مصنوعی، رگرسیون لوجستیک و درخت تصمیم را برای پیش بینی عوارض پس از جراحی در بیماران مبتلا به سرطان معده مورد استفاده قرار داده و نشان دادند که شبکه عصبی مصنوعی روش بهتری برای پیش بینی این عوارض در مقایسه با دو روش دیگر است،(20). گرچه نتایج مطالعه حاضر و سایر موارد عنوان شده همگی نشان دهنده برتری مدل شبکه عصبی مصنوعی نسبت به رگرسیون کاکس و سایر روش های آماری بودند، برخی از محققین نیز خلاف این موضوع را در مطالعات خود نشان داده و عنوان کردند که بهتر است شبکه عصبی

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

مصنوعی جایگزین مدل های کلاسیک آماری نشده و هر دو رویکـرد در کـنار هـم مورد استفاده قرار بگیرند.(33)

سپاسگزاری

این مقاله حاصل بخشی از طرح تحقیقاتی تحت عنوان«کارگیری شبکه عصبی مصنوعی بازخور برای پیش بینی بقـای بیـماران سرطان پستان» مصوب دانشـگاه عـلوم پزشـکی و خـدمات بهـداشتی درمـانـی تـهـران در سـال 1391 بـا کد 90-04-136-15979 می باشد که با حمایت دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران اجرا شده است.

References

1-Yavari P, Abadi A, Mehrabi Y. [Mortality and changing epidemiological trends in Iran during 1979-2001]. Hakim Res J 2003;6 7-14.(Persian)

2-Hosseini SM, Hassannejad R, Khademolghorani S, Tabatabaeian M, Mokarian F. Identification of patterns of breast cancer metastasis among women referred to Isfahan Seyedoshohada Center, Iran, between 1999 and 2009 by association rules and ordinal logistic regression. J Health Sys Res 2011;7:746-62.

3-National cancer registry report. Tehran cancer administration. [Non-communicable diseases sector]. Iranian Center for Diseases Control and Prevention; 2005.(Persian)

4-Saki A, Hajizadeh E, Tehranian N. [Evaluating the risk factors of breast cancer using the analysis of tree models]. J Gonabad Uni Med Sci 2011;17:60-9.(Persian)

5-Jami M, Tavassoli M, Hemmati S. [Association of the length of CA dinucleotide repeat in the epidermal growth factor receptor with risk and age of breast cancer onset]. J Isfahan 2008;88:22-30.(Persian)

6-Babazadeh SH, Andalib A, Amouheidari A, Tabatabaeian M, Emami H, Adibi A, et al. A study of the long-term trend of epidemiological and clinical indices of breast cancer in Isfahan. J Isfahan Med School 2012;29:1831-43.

7-Deputy Cancer Office. Iranian annual of national cancer registration report (2005-2006) Tehran, Iran: Islamic Republic of Iran. Ministry of Health and Medical Education, Health Deputy; 2006.

8-Parkin DM, Fernandez LM. Use of statistics to assess the global burden of breast cancer. Breast J 2006;12:70-80.

9-Brunicardi FC, Schwartz SI. Schwartz's principles of surgery. New York: McGraw Hill; 2010.

10-Fuladi N, Amani F, Harghi AS, Nayebyazdi N. Five year survival of women with breast cancer in Ardabil, north-west of Iran. Asian Pac J Cancer Prev 2011;12:1799-801.

11-Akbari ME, Khayamzadeh M, Khoshnevis SJ, Nafisi N, Akbari A. Five & ten years survival in breast cancer patients mastectomies vs. breast conserving surgeries, personal experience. Iran J Cancer Prevent 2008;1:53-7.

12-Maleki D. [The demographic charachterestics and clinical presentation of patients with breast cancer in Urmia, which screening program may be appropriate?] Urmia M J 2010;21:273-277.(Persian)

13-Kutner MH, Nachtsheim CJ. Neter J. Applied linear regression models. 4th ed. New York: McGraw Hill; 2004.

14-Lee ET, Wang JW. Statistical methods for survival data analysis. 3rd ed. New Jersey: John Wiley & Sons; 2003.

15-Jobson DJ. Applied multivariate data analysis: categorical and multivariate methods. New York: Springer; 1992.

16-Warner B, Manavendra M. Understanding neural networks as statistical tools. Am Stat 1996;50:284-93.

17-Baxt WG, Skora J. Prospective validation of artificial neural networks trained to identify acute myocardial infarction. Lancet 1996;347:5-12.

18-Mobley BA, Schecheer E, Moore WE, McKee PA, Eichner JE. Prediction of coronary artery enosis by artificial networks. Artific Intel Med. 2000;18:187-203.

19-Lai KC, Chiang HC, Chen WC, Tsai FJ, Jeng LB. Artificial neural network-based study can predict gastric cancer staging. Hepatogastroenterol 2008;55:1859-63.

20-Chien CW, Lee YC, Ma T, Lee TS, Lin YC, Wang W, et al. The application of artificial neural networks and decision tree

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

model in predicting post-operative complication for gastric cancer patients. Hepatogastroenterol 2008;55:1140-5.
21-Gohari MR, Biglarian A, Bakhshi E. Artificial neural network to determine the prognostic factors in colorectal cancer patients. Asian Pac J Cancer Prev 2011;12: 1469-72.
22-Biglarian A, Bakhshi1 E, Gohari MR, Khodabakhshi R. Artificial neural network for prediction of distant metastasis in colorectal cancer. Asian Pac J Cancer Prev 2012;13:927-30.
23-Kay JW, Titterington DM. Statistics and neural networks: Advanced at the interface. Oxford: Oxford University Press; 1999.
24-Ripely BD, Ripely RM. Neural networks as statistical methods in survival analysis. In: Artificial neural networks: prospects for medicine. Dybowsky R, Gant V, editors. Landes Biosciences Publishers; 1998.
25-Anderson A. An introduction to neural network. Cambridge: MIT press; 1995.
26-Khodabakhshi R, Gohari MR, Moghadamifard Z, Foadzi h, Vahabi N. Disease-free survival of breast cancer patients and identification of related factors. Razi J Med Sci 2011;18:27-33.
27-Yaghmaei S, Bani Hashemi G, Ghorbani R. [Survival rate following treatment of primary breast cancer in Semnan, Iran (1991-2002).] J Semnan Uni Med Sci 2008; 9:111-17.
28-Bakhtiari A, Hajian K. Five year disease-free survival rate in breast cancer patients in Rajaii Hospital, Babolsar. J School Public Health 2007;5:53-60.
29-Rondeau V, Mathoulin-Pélissier S, Tanneau L, Sasco AJ, MacGrogan G, Debled M. Separate and combined analysis of successsive dependent outcomes after breast-conservation surgery: recurrence, metastases, second cancer and death. BMC cancer 2010;10:697-709.
30- Gohari MR, Mahmoudi M, Mohammed K, Pasha E, Khodabakhshi R. Recurrence in breast cancer. Saudi Med J 2006;27:1187-93.
31-Biglarian A, Hajizadeh E, Kazemnejad A. Comparison of artificial neural network and Cox regression models in survival prediction of gastric cancer patients. J Semnan Uni Med Sci 2010;11:205-11.
32-Gohari MR, Biglarian A, Bakhshi E, Pourhoseingholi MA. Use of an artificial neural network to determine prognostic factors in colorectal cancer patients. Asian Pac J Cancer Prev 2011;12:1469-72.
33-Sargent DJ. Comparison of artificial neural networks with other statistical approaches: results from medical data sets. Cancer 2001;15;91:1636-42.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

Comparison of Artificial Neural Network and Cox Regression Models in Survival Prediction of Breast Cancer Patients

Salehi M[1], Gohari M.R[1], Vahabi N[1], Zayeri F[2], Yahyazadeh S.H[3], Kafashian M.R[4]*

(Received: 6 Feb. 2013 Accepted: 6 May. 2013)

Abstract

Introduction: Nowadays, cancer diseases are the most important causes of death worldwide and the breast cancer is the most important of them within women. Assessing the survival of the patients is one of the most important indices of controlling the cancer. This study aimed to make a comparison between the prediction of artificial neural networks (ANN) and Cox regression models for the breast cancer survival.

Materials & Methods: The data of the survival study gathered from 344 breast cancer patients between 2005 and 2012 that registered at the Fayyazbakhsh hospital, Tehran, Iran. The status of survival was considered as a dependent variable. Area under receiver operative characteristic curve (AUROC) and classification accuracy were used for the comparison of artificial neural networks and Cox regression models. Data analysis was performed by R and Minitab software.

Findings: The age of participants expressed as Mean ± SD, was 49.9±10.93 years and the median of survival was 44.6 months. Up to the end study, 45 (13.1) were died. Results showed that AUROC for ANN and Cox regression were 87.6% and 75.4%, respectively. In addition the classification accuracy of ANN and Cox regression were calculated as 89.42 and 77.68, respectively.

Discussion & Conclusion: According to the results, the total classification accuracy of the ANN was better than those of the Cox regression; therefore, the ANN model is suggested to predict the survival status of breast cancer disease and also is suggested for diagnostic goals.

Keywords: breast cancer, survival, cox regression, artificial neural networks

1. Dept of Statistics & Mathematics, Management & Economics Research Center, School of Health Management & Information Sciences, Tehran University of Medical Sciences, Tehran, Iran
2. Dept of Biostatistics, Proteomics Research Center, School of Paramedical Sciences, Shahid Beheshti, University of Medical Sciences, Tehran, Iran
3. Dept of Radiotherapy and Oncology, Clinical Research Center, Milad General Hospital, Tehran, Iran
4. Dept of Physiology, Faculty of Medicine, Ilam University of Medical Sciences, Ilam, Iran
** (corresponding author)*

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]