

پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور دو چرخ محرک با استفاده از شبکه عصبی مصنوعی و مقایسه آن با رگرسیون

عباس روحانی^۱، ایرج رنجبر^۲، یحیی عجب‌شیرچی^۲، *محمدحسین عباسپورفرد^۳ و مصطفی‌ولیزاده^۴

^۱دانشجوی دوره دکتری گروه ماشین‌های کشاورزی، دانشگاه تبریز، دانشیار گروه ماشین‌های کشاورزی، دانشگاه تبریز،

^۲استادیار گروه ماشین‌های کشاورزی، دانشگاه فردوسی مشهد، ^۳استاد گروه زراعت، دانشگاه تبریز

تاریخ دریافت: ۸۷/۸/۱۸؛ تاریخ پذیرش: ۸۸/۴/۸

چکیده

هدف هر شرکت زراعی، رسیدن به بیشترین سود می‌باشد. یکی از عوامل مؤثر در دستیابی به این هدف، مدیریت جایگزینی ماشین‌های مختلف به‌ویژه تراکتور است. بنابراین باید بتوان هزینه‌های تعمیر و نگهداری تراکتور را با دقت بسیار بالای پیش‌بینی کرد. هدف از این تحقیق پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور با استفاده از شبکه عصبی مصنوعی و روش رگرسیون و مقایسه عملکرد این دو تکنیک می‌باشد. در این مطالعه از داده‌های واقعی ۶۰ تراکتور دو چرخ محرک موجود در کشت و صنعت آستان قدس رضوی در این مطالعه استفاده شده است. تحلیل رگرسیونی مجموعه داده‌ها در فاز آموزش انجام، و مدل‌های پلی‌نوم میان درجه دوم، درجه سوم و مدل‌های توان ارزیابی شدند، که پلی‌نوم میان درجه سوم براساس معنی‌داری ضرایب رگرسیون و ضریب تبیین به‌عنوان شکل تابعی انتخاب گردید. همچنین شبکه عصبی توسط الگوریتم یادگیری پس از انتشار با ضریب ممتوم و تعداد گره‌ها در لایه مخفی متأثر از دقت پیش‌بینی هزینه است. در این بررسی مدل بهینه شبکه عصبی با ده نرون در لایه مخفی پیدا شد. به‌ترتیب مقادیر بهینه برای آهنگ یادگیری و ضریب ۰/۷ و ۰/۸ می‌باشند. عملکرد مدل شبکه عصبی و مدل رگرسیونی توسط مجموعه داده‌ها در فاز تست ارزیابی شدند. مدل بهینه شبکه عصبی قادر به پیش‌بینی مقادیر هزینه‌های تراکتور در فاز تست به‌ترتیب با متوسط قدرمطلق درصد خطا و ریشه متوسط مربعات خطای کمتر از ۲/۸۲ درصد و ۰/۵۲ در مقابل ۹/۵ درصد و ۱/۳۹ برای مدل رگرسیونی بود.

واژه‌های کلیدی: شبکه عصبی، رگرسیون، هزینه تعمیر و نگهداری

مقدمه

تصمیمات جایگزینی صحیح شوند زیرا بیشتر مدل‌های موجود برای پیش‌بینی هزینه‌های تعمیر و نگهداری دقت به نسبت پایینی هستند (میشل، ۱۹۹۸). از آنجا که تصمیم‌گیری جایگزینی به‌عنوان مهم‌ترین عامل مدیریتی ماشین‌های مزرعه به اطلاع از هزینه‌های آینده نیازمند است بنابراین داشتن ابزارهای پیش‌بینی‌کننده دقیق

تخمین دقیق هزینه‌های تعمیر و نگهداری برای تصمیم‌گیری‌های اقتصادی مانند جایگزینی یکی از دغدغه‌های همیشگی مدیر مزرعه می‌باشد. به‌ندرت مدل‌های اقتصادی یافت می‌شوند که بتوانند منجر به

* - مسئول مکاتبه: hossein_abbaspour@yahoo.co.uk

پس خوردی جریان دارند. گره‌های ورودی متغیرهای مستقل را دریافت، و گره‌های خروجی متغیرهای وابسته را پیش‌بینی می‌کنند. گره‌های مخفی همراه با تابع تبدیل غیرخطی اطلاعات دریافتی از گره‌های ورودی را پردازش می‌کنند. یادگیری شبکه از طریق تنظیم اتصالات سیناپسی بین نرون‌های لایه مخفی انجام می‌گیرد. الگوریتم پس‌انتشار رایج‌ترین روش یادگیری شبکه می‌باشد. در این روش وزن‌ها براساس کمینه‌سازی مجموع کل مربعات اختلاف بین خروجی مدل و خروجی واقعی انجام می‌گیرد. بعد از آن که یادگیری شبکه با موفقیت صورت گرفت، شبکه عصبی آموزش دیده را می‌توان در پیش‌بینی داده‌های جدید استفاده کرد و قابلیت تعمیم‌پذیری آن را ارزیابی نمود. شبکه عصبی چند لایه پرسپترون قادر است به خوبی روابط غیرخطی بین متغیرها را پیدا کند حتی اگر طبیعت دقیق روابط آنها نامشخص باشد (هیکن، ۱۹۹۴). با توجه به دلایل ذکر شده در بالا و نیز مشکلات موجود در به کارگیری مدل‌های رگرسیونی موجود برای پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور، از جمله پیشنهاد دو مدل برای تخمین این هزینه‌ها سبب شده که بیشتر کاربران از نتایج و مدل‌های پیشنهاد شده استقبال خوبی نداشته باشند و با بررسی مزایای مدل‌های عصبی از جمله مهم‌ترین آنها گنجاندن نتایج نهایی در قالب یک نرم‌افزار برای استفاده ساده‌تر از آن و نیز به روز کردن مدل با ایجاد بانک داده‌های مربوط به وضعیت تعمیر و نگهداری ماشین‌ها و قابلیت کاربردی آن در سطح کلان و نیز سطوح مدیریتی پایین می‌تواند جایگزین بسیار مناسبی برای مدل‌های رگرسیونی باشد. اهداف این مقاله شامل چگونگی به کارگیری شبکه عصبی در پیش‌بینی هزینه تعمیر و نگهداری تراکتور، پیدا کردن بهترین مقادیر پارامترهای شبکه و مدل‌های رگرسیونی و نیز ارزیابی قابلیت شبکه عصبی در مقابل تکنیک رگرسیون برای این مورد خاص است.

ضروری به نظر می‌رسند (تلسانگ، ۲۰۰۵). تعیین ظرفیت و تعداد بهینه ماشین‌های مورد نیاز و برنامه‌ریزی کاری آنها و همچنین فراهم کردن قطعات یدکی و تعیین زمان بهینه جایگزینی به منظور حفظ قابلیت دسترسی ماشین از عوامل ضرورت نیاز به پیش‌بینی می‌باشد (کیم، ۱۹۸۹). تکنیک‌های رگرسیونی به دلیل داشتن پشتوانه ریاضی بسیار خوب در پیش‌بینی هزینه‌های تعمیر و نگهداری ماشین‌های کشاورزی از سال ۱۹۷۰ استفاده شدند (باورز و هانت، ۱۹۷۰). تحلیل رگرسیونی تکنیکی است که رابطه بین یک متغیر وابسته نسبت به یک یا چند متغیر مستقل را پیدا می‌کند (لیمان و لانگنکر، ۲۰۰۱). یکی از فرضیه‌های اولیه در این تحلیل مشخص کردن رابطه تابع نسبت به متغیرها در ابتدا می‌باشد. همچنین در بیشتر موارد از روش خطی‌سازی در تحلیل توزیع داده‌های غیرخطی به کمک مدل رگرسیون خطی استفاده می‌شود و اگر عمل خطی‌سازی به درستی انجام نشود، اعتبار مدل به دست آمده خدشه‌دار خواهد شد (کومار، ۲۰۰۵). بیشتر این موارد به‌عنوان عیب‌های این تکنیک برشمرده می‌شود. شبکه عصبی مصنوعی به دلیل داشتن ویژگی‌های خاص مانند یادگیری مستقیم از روی مثال‌های واقعی، بی‌نیازی به برآوردهای ویژگی‌های آماری داده‌های اولیه و نیز از همه مهم‌تر برای این زمینه خاص تعیین روابط غیرخطی با مرتبه بالا بین متغیرهای موجود (وکیل - باغمیشه، ۲۰۰۲؛ گوپتا و همکاران، ۲۰۰۳)، می‌تواند به‌عنوان یک جایگزین برای روش رگرسیونی در پیش‌بینی هزینه‌ها باشد. شبکه عصبی حاصل شبیه‌سازی سیستم عصبی مغز انسان بوده و قادر به یادگیری هر مسأله براساس تجربه‌های به دست آمده از آن می‌باشد. بنابراین شبکه عصبی بدون وابستگی به فرضیاتی مانند شکل اولیه تابع و توزیع‌های احتمالی و تعداد متغیرها، مدل نهایی را با یادگیری داده‌های خام اولیه به دست می‌آورد (ولنتروف، ۱۹۹۵). در این مطالعه تنها از شبکه عصبی چند لایه پرسپترون، استفاده شده است. در این شبکه داده‌ها همواره از لایه ورودی به سمت لایه خروجی بدون هیچ

مواد و روش‌ها

داده‌های هزینه‌های تعمیر و نگهداری مزرعه‌ای: امکان دستیابی به داده‌های آزمایشگاهی که برای این مطالعه مناسب باشد وجود ندارد ولی داده‌های مزرعه‌ای تصویر واقعی از چگونگی تغییرات هزینه‌های تعمیر و نگهداری در طول زمان را نشان می‌دهند. برای انجام این تحقیق از داده‌های هزینه‌های تعمیر و نگهداری ماهانه در طی ۱۸ سال مربوط به ۶۰ تراکتور دو چرخ محرک فعال در مزرعه نمونه آستان قدس رضوی در استان خراسان رضوی در چهار نوع تراکتور جان‌دیر ۳۱۴۰، جان‌دیر ۴۴۵۰، فرگوسن ۲۸۵ و فیات ۴۴۵ به ترتیب به تعداد ۲۸، ۵، ۱۷ و ۱۰ استفاده گردید. این داده‌ها شامل هزینه تعمیرات (هزینه قطعات یدکی و دستمزد تعمیرات)، هزینه روغن (هزینه روان‌سازها و فیلترهای روغن)، هزینه سوخت (هزینه گازوئیل و فیلترهای سوخت)، هزینه تعمیر و نگهداری (مجموع هزینه تعمیرات، روغن و سوخت) و سال خرید و ساخت هر تراکتور می‌باشند.

استانداردسازی داده‌ها: در ابتدا و قبل از محاسبه هزینه جمعی باید اثر تورم بر هزینه‌ها را تعدیل کرد (اسکونزاد، ۲۰۰۳). به دلیل متفاوت بودن هزینه‌های تعمیر و نگهداری و نیز قیمت خرید اولیه انواع تراکتورهای مورد مطالعه و نیز متفاوت بودن طبیعت داده‌های مربوط به هر نوع تراکتور به علت به کارگیری آنها در عملیات زراعی مختلف، تمام این هزینه‌ها توسط شاخص زیر استانداردسازی شد (میشل، ۱۹۹۸).

$$CCI_t = \frac{\sum C_t}{PP_t} \times 100 \quad (1)$$

در اینجا CCI_t شاخص هزینه جمعی در زمان t ، C_t هزینه تعمیر و نگهداری در زمان t و PP_t قیمت خرید اولیه تراکتور است، که این شاخص در طول عمر تقویمی تراکتور همواره روند افزایشی یا ثابت دارد. بعد از استاندارد کردن هزینه‌ها، شاخص‌های جمعی به صورت شاخص هزینه تعمیراتی جمعی (CCI_{repair})، شاخص

هزینه روغن جمعی (CCI_{oil})، شاخص هزینه سوخت جمعی (CCI_{fuel}) محاسبه گردید. این شاخص‌ها به عنوان خروجی شبکه عصبی و نیز متغیرهای وابسته در روش رگرسیونی استفاده شد. اغلب در تصمیم‌گیری اقتصادی مدیریتی ماشین به منظور تعیین عمر مفید از شاخص هزینه تعمیر و نگهداری استفاده می‌شود. شاخص هزینه تعمیر و نگهداری جمعی برابر است با مجموع سه شاخص هزینه تعمیرات، روغن و سوخت.

طول عمر: به طور عمده متغیر مستقل در مدل رگرسیونی هزینه جمعی و ورودی شبکه عصبی، عمر تراکتور است. برای طول عمر سه تعریف عمر تقویمی، عمر بر حسب واحدهای تولید محصول و ساعات کارکرد جمعی وجود دارد. از بین آنها ساعات کارکرد جمعی^۱ مناسب‌ترین تعریفی است که می‌توان از عمر تراکتور داشت (میشل، ۱۹۹۸). ساعات کارکرد جمعی تعداد ساعاتی که تراکتور به طور فیزیکی کار کرده را نشان می‌دهد و نیز تغییرات بسیار زیاد هزینه‌های تعمیر و نگهداری را به خوبی در طول زمان تعدیل می‌کند. به دلیل سالم نبودن ساعت شمار تراکتورها، تعداد ساعات کارکرد هر تراکتور براساس تعداد تعویض روغن موتور محاسبه گردید. برای دستیابی به عملکرد بهتر مدل‌بندی این هزینه‌ها توسط شبکه عصبی و نیز تکنیک رگرسیونی، ساعات کارکرد جمعی بر حسب ۱۰۰ ساعت محاسبه شد.

فرضیات مدل‌سازی: در ابتدای تحلیل هزینه‌های تعمیر و نگهداری باید فرضیه‌های زیر را در نظر گرفت: (۱) اطمینان به داده‌های موجود، از آنجا که امکان برگشت به گذشته برای تصدیق مقادیر هزینه وجود ندارد باید به مقادیر هزینه‌های ثبت شده از طرف شرکت اطمینان کرد. (۲) هزینه تعمیر و نگهداری تراکتور در ابتدای عمر آن صفر است. این فرضیه کاملاً قابل قبول و ضروری است زیرا هزینه تعمیرات احتمالی قبل از به کارگیری تراکتور توسط شرکت سازنده پرداخت می‌شود. براساس این

1- Cumulative Hours of Usage (CCU)

در اینجا x داده خام اولیه، x_n داده نرمالیزه شده، x_{\min} و x_{\max} به ترتیب مقادیر بیشینه و کمینه داده‌های اولیه و r_{\min} و r_{\max} به ترتیب حد بالایی و پایینی دامنه تغییرات داده‌های تبدیل شده است.

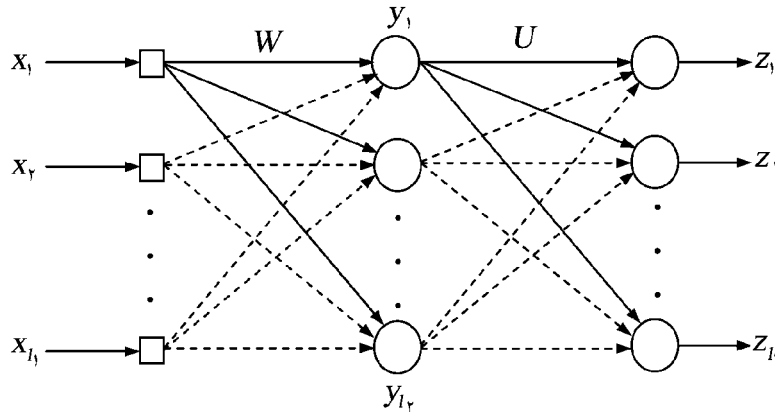
شبکه عصبی پرسپترون چندلایه: در این مطالعه از شبکه چند لایه پرسپترون استفاده شد. اگر تعداد نرون‌های لایه مخفی و نیز داده‌های آموزش به مقدار کافی باشد آن‌گاه شبکه چند لایه پرسپترون به‌طور واقعی می‌تواند هر تابع را با هر دقت دلخواه تقریب بزند (کوال و همکاران، ۲۰۰۵). ساختار شبکه چند لایه پرسپترون با یک لایه مخفی در شکل ۱ نشان داده شده است. شبکه چند لایه پرسپترون با دریافت بردار ورودی X^q ، بردار خروجی Z^q برای هر $q (q=1, \dots, Q)$ تولید خواهد کرد. هدف اتخاذ پارامترهای صحیح شبکه به منظور دستیابی به خروجی واقعی Z^q که تا جای ممکن نزدیک به خروجی مطلوب d^q متناظر با خود باشد. برای آموزش شبکه از الگوریتم پس‌انتشار با نرخ یادگیری کاهشی^۱ استفاده شد. برتری الگوریتم در مقابل با الگوریتم پس‌انتشار اصلی^۲ به‌عنوان یک الگوریتم رایج در آموزش شبکه چند لایه پرسپترون دارای برتری‌هایی از قبیل سهولت در تنظیم پارامترهای یادگیری شبکه به سبب پایین آوردن حساسیت شبکه به مقادیر آهنگ یادگیری و عامل ممتوم، کاهش زمان یادگیری از طریق پایین آوردن تعداد تکرارهای مورد نیاز برای یادگیری و بهبود رفتار شبکه در طول آموزش می‌باشد. کد کامپیوتری این الگوریتم در محیط برنامه‌نویسی مطلب^۳ تهیه شد.

فرضیه ضریب عرض از مبدأ مدل‌های رگرسیونی (β) صفر خواهد شد و همچنین نباید تخمین هزینه‌ها توسط شبکه عصبی کمتر از صفر باشد. (۳) ساعات کارکرد جمعی تنها متغیر مستقل برای مدل‌های رگرسیونی و ورودی شبکه عصبی است. اگرچه ممکن است متغیرهای زیادی در تخمین هزینه‌های تعمیر و نگهداری تراکتور مؤثر باشند ولی به دلیل وجود شرایط یکسان به کارگیری آنها از جمله سطح مدیریتی، شرایط آب و هوایی و نیز تا حدودی یکسان بودن سطح مهارت کاربران می‌توان سایر متغیرها را ثابت در نظر گرفت.

آماده‌سازی داده‌ها برای شبکه عصبی: در ابتدا به‌طور تصادفی میانگین وزنی داده‌های هزینه‌های تعمیر و نگهداری ماهانه ۶۰ تراکتور همراه با ساعات کارکرد جمعی آنها طی ۱۸ سال به دو دسته مجموعه آموزش با ۱۴۴ عضو ($\frac{2}{3}$ کل داده‌ها) و مجموعه تست با ۷۲ عضو ($\frac{1}{3}$ کل داده‌ها) تقسیم‌بندی شد. چنانچه این تقسیم‌بندی منجر به نتایج مطلوب نشود، می‌توان این مرحله را دوباره تکرار کرد (زانگ و فو، ۱۹۹۸). مرحله بعد از آماده‌سازی داده‌ها، نرمالیزه کردن آنها می‌باشد تا مجموعه داده‌های آموزش و تست دارای توزیع آماری تقریباً یکنواختی شوند. همچنین باید این تبدیل در دامنه تغییرات مناسبی انجام، تا شبکه به نقطه بهینه همگرا شود و از آنجا که تابع فعالیت انتخاب شده برای نرون‌های لایه مخفی شبکه سیگموئید است بنابراین بهترین دامنه تبدیل داده‌ها $[0/9]$ می‌باشد (وکیل - باغمیشه، ۲۰۰۲). در این بررسی از نرمالیزاسیون خطی برای تبدیل داده‌ها استفاده شد:

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \times (r_{\max} - r_{\min}) + r_{\min} \quad (2)$$

- 1- Backpropagation With Declining Learning-Rate Factor
- 2- Basic Backpropagation
- 3- Matlab



شکل ۱- ساختار شبکه چند لایه پرسپترون با یک لایه مخفی (وکیل- باغمیشه، ۲۰۰۲).

الگوریتم پس‌انتشار با نرخ یادگیری کاهشی: این الگوریتم نسخه اصلاح‌شده الگوریتم پس‌انتشار اصلی می‌باشد (وکیل- باغمیشه و پاوسیک، ۲۰۰۱). این الگوریتم آموزش با آهنگ یادگیری (η) و عامل ممتوم (α) ثابت و به نسبت بزرگ شروع می‌کند و قبل از آنکه شبکه ناپایدار و یا همگرایی آن کند شود، هر T تکرار $(3 \leq T \leq 5)$ آهنگ یادگیری و ضریب ممتوم را از طریق تصاعد حسابی به صورت یکنواخت کاهش داده می‌شود تا زمانی که این پارامترها به X درصد (۵ درصد) مقادیر اولیه خودشان برسند. η و به طور مشابه α را با استفاده از رابطه زیر کاهش داده شد:

$$u_{jk}(n+1) = u_{jk}(n) - \eta \times \frac{\partial E}{\partial u_{jk}} + \alpha(u_{jk}(n) - u_{jk}(n-1))$$

$$k = 1, \dots, l_p, j = 1, \dots, l_p, i = 1, \dots, l_p$$

$$w_{ij}(n+1) = w_{ij}(n) - \eta \times \frac{\partial E}{\partial w_{ij}} + \alpha(w_{ij}(n) - w_{ij}(n-1))$$

در اینجا w_{ij} اتصال وزنی بین گره‌های i و j ، u_{jk} اتصال وزنی بین گره‌های j و k ، مقادیر اولیه این وزن‌ها به‌طور تصادفی از دامنه مقادیر $[-0.25, 0.25]$ انتخاب می‌شود. l_p و l_p به ترتیب تعداد نرون‌ها در لایه مخفی و لایه خروجی می‌باشند. η و α به ترتیب آهنگ یادگیری و فاکتور ممتوم هستند و مقادیر آنها در فاصله $[0, 1]$ قرار دارند و n شماره تکرار الگوریتم $(n = 1, \dots, N)$. زمانی الگوریتم متوقف می‌شود که مجموع کل مربعات خطا کوچک‌تر از مقدار آستانه (0.0001) مقدار آستانه در نظر گرفته شده برای مطالعه حاضر است) باشد. جزئیات بیشتر این الگوریتم را می‌توان در وکیل- باغمیشه و پاوسیک (۲۰۰۳) مشاهده کرد.

در اینجا n_1 ، η_n و η به ترتیب نقطه شروع الگوریتم پس‌انتشار با نرخ یادگیری کاهشی، آهنگ یادگیری در n امین جمله از تصاعد حسابی و مقدار اولیه آهنگ یادگیری می‌باشند.

تابع هزینه به‌کار رفته در این الگوریتم مجموع کل مربعات خطا^۱ می‌باشد و به صورت زیر محاسبه می‌شود:

$$\eta_n = \eta + (x-1) \frac{n\eta T}{Q - n_1} \quad (3)$$

$$TSSE = \sum_q \sum_k (d_k^q - z_k^q), \quad q = 1, \dots, Q \quad (4)$$

1- Total Sum-Squared Error (TSSE)

نرم‌افزاری SPSS انجام گرفت. این دو روش براساس معیارهای مربوط به معادله‌های (۷) و (۸) و R^2 با هم مقایسه شدند.

نتایج و بحث

تنظیم پارامترهای شبکه: براساس قضیه تقریب عمومی، شبکه عصبی با یک لایه مخفی و با تعداد کافی نرون در آن لایه قادر به تقریب هر تابع پیوسته دلخواه است (هیکین، ۱۹۹۴). یک لایه مخفی برای شبکه عصبی انتخاب شد. لایه ورودی و لایه خروجی به ترتیب شامل ۲ و ۳ نرون می‌باشند زیرا ورودی‌های شبکه شامل عامل بایاس و ساعات کارکرد تجمعی و خروجی‌های شبکه شامل شاخص‌های هزینه تعمیر، روغن و سوخت هستند. برای تعیین مناسب‌ترین توپولوژی شبکه (تعداد نرون‌ها در لایه مخفی) از روش آزمون و خطا استفاده گردید. در طول آزمون آهنگ یادگیری، عامل ممتوم و تعداد تکرارها ($\alpha=0.001$, $\eta=0.5$, $n_1=9500$) ثابت نگه داشته شد. این فرآیند برای تعداد نرون‌های لایه مخفی ۵ تا ۱۵ انجام پذیرفت. در نهایت بهترین نتایج براساس معیارهای عملکردی در ۱۰ نرون در لایه مخفی شبکه به‌دست آمد. مقادیر بهینه پارامترهای η و α نیز از طریق آزمون و خطا انتخاب گردیدند. عملکرد شبکه برای تمام ترکیبات مقادیر آهنگ یادگیری و عامل ممتوم در بازه تغییراتی [۰/۵ ۰/۹۹] برای η و [۰/۵ ۰/۹۹] برای α با هم مقایسه شدند. نتایج نشان می‌دهند که بهترین عملکرد در سطح $\alpha=0.8$ و $\eta=0.7$ به‌دست می‌آید. تعداد بهینه تکرارهای مورد نیاز شبکه در ۲۵۰۰۰ و همراه با نقطه شروع الگوریتم در ۲۳۰۰۰ به‌دست آمد و همچنین در انتها مقادیر η و α به ترتیب به ۰/۳۵ و ۰/۰۴ رسیدند. ساختار شبکه چند لایه پرسپترون استفاده شده در این مطالعه در شکل ۲ نشان داده شده است.

معیارهای ارزیابی عملکرد شبکه عصبی: برای ارزیابی قابلیت شبکه عصبی چند لایه پرسپترون و تکنیک رگرسیونی در پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور از معیارهای متوسط قدرمطلق درصد خطا^۱، ریشه متوسط مربعات خطا^۲ و ضریب تبیین معادله خطی رگرسیونی بین مقادیر پیش‌بینی شده توسط شبکه عصبی و مقادیر واقعی آنها استفاده شدند.

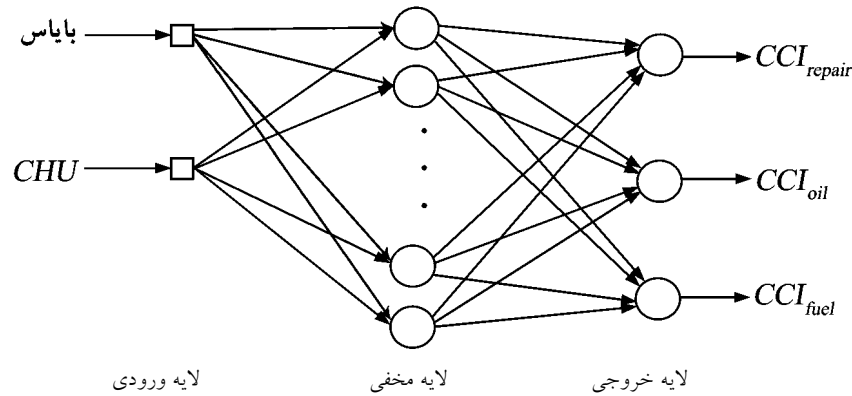
$$RMSE = \sqrt{\frac{\sum_{j=1}^n \sum_{i=1}^m (d_{ji} - p_{ji})^2}{nm}} \quad (۷)$$

$$MAPE = \frac{1}{nm} \sum_{j=1}^n \sum_{i=1}^m \left| \frac{d_{ji} - p_{ji}}{d_{ji}} \right| \times 100 \quad (۸)$$

در اینجا d_{ji} مؤلفه i ام از خروجی واقعی مربوط به پترن j ام، p_{ji} مؤلفه i ام از خروجی پیش‌بینی شده توسط شبکه مربوط به پترن j ام، \bar{d} و \bar{p} به ترتیب متوسط خروجی‌های واقعی و پیش‌بینی شده توسط شبکه و n و m به ترتیب تعداد پترن‌ها و تعداد متغیرهای خروجی هستند.

تحلیل رگرسیونی: برای آنکه بتوان نتایج به‌دست آمده از روش شبکه عصبی را با روش‌های رگرسیونی مقایسه کرد. از همان مجموعه داده‌های آموزش شبکه برای انتخاب بهترین مدل رگرسیونی از بین مدل‌های توصیه شده خطی، درجه دوم، درجه سوم، نمایی و توانی (موریس، ۱۹۸۸؛ روتز، ۱۹۸۷؛ فولس، ۱۹۹۹) و همچنین برای ارزیابی اعتبار مدل‌های انتخاب شده برای هر شاخص هزینه تعمیر و نگهداری از همان مجموعه تست شبکه عصبی استفاده شد (کوال و همکاران، ۲۰۰۵). ورودی شبکه به‌عنوان متغیر مستقل و خروجی‌های شبکه به‌عنوان متغیرهای مدل رگرسیونی در نظر گرفته شدند. انتخاب مدل‌های رگرسیونی براساس معنی‌داری ضرایب رگرسیونی و پارامتر F و بزرگی ضریب تبیین در محیط

1- Mean Absolute Percentage Error
2- Root Mean Squared Error



شکل ۲- ساختار شبکه عصبی چند لایه پرسپترون استفاده شده برای پیش‌بینی هزینه‌های تعمیر و نگهداری.

مقادیر واقعی شاخص هزینه روغن و سوخت و مقادیر پیش‌بینی شده آنها براساس مدل رگرسیونی در مقایسه با شبکه عصبی چند لایه پرسپترون بسیار چشم‌گیر می‌باشند. بنابراین شبکه عصبی در مقایسه با تکنیک رگرسیونی برای پیش‌بینی هزینه‌های تعمیر و نگهداری از شایستگی بیشتری برخوردار است.

برای بررسی بیشتر قابلیت هر دو روش رگرسیونی و شبکه عصبی، میانگین، واریانس و توزیع آماری مقادیر واقعی و مقادیر پیش‌بینی شده آنها از مجموعه داده‌های تست از نظر آماری مقایسه گردید. در اینجا فرضیه صفر بر تساوی میانگین، واریانس و توزیع آماری هر دو سری داده دلالت دارد. هر فرضیه در سطح احتمال ۹۵ درصد به کمک پارامتر p تست گردید. بنابراین اگر p محاسبه شده برای هر مقایسه بیشتر از $0/05$ باشد، فرضیه صفر را نمی‌توان رد کرد. به ترتیب برای مقایسه میانگین، واریانس و توزیع آماری از آزمون F ، t و کولموگروف-اسمیرنو استفاده شد. مقادیر p محاسبه شده برای هر سه شاخص هزینه در جدول ۳ نشان داده شده است. این نتایج نشان می‌دهند که میانگین، واریانس و توزیع آماری مقادیر واقعی و پیش‌بینی شده شاخص هزینه تعمیر توسط مدل رگرسیونی و مدل عصبی اختلاف معنی‌داری با هم ندارند و هر دو روش دارای قابلیت بالایی برای پیش‌بینی شاخص هزینه تعمیر هستند. میانگین و واریانس شاخص هزینه سوخت در سطح احتمال ۹۵ درصد در روش

تحلیل آماری: در اینجا تنها به آوردن نتایج به دست آمده از ارزیابی مدل‌های رگرسیونی انتخاب شده و نیز شبکه عصبی آموزش دیده اکتفا می‌کنیم زیرا معتبر بودن این نتایج دلیل بر موفقیت‌آمیز بودن فاز آموزش برای هر دو روش مورد بررسی می‌باشد. نتایج تحلیل رگرسیونی در فاز آموزش نشان می‌دهد که ضرایب رگرسیونی و نیز پارامتر F در تمام مدل‌ها برای سه شاخص هزینه در سطح احتمال ۱ درصد معنی‌دار شدند. اما ضرایب تبیین متفاوتی داشتند که براساس این معیار، مدل درجه سوم نسبت به سایر مدل‌های رگرسیونی تغییرات داده‌های آموزش سه شاخص هزینه تعمیر، روغن و سوخت را به طور مطلوبی در طول عمر تراکتور توجیه می‌کرد (جدول ۱). همچنین وزن‌های شبکه عصبی به منظور پیش‌بینی هم‌زمان هر سه شاخص هزینه طی فاز آموزش محاسبه می‌شوند. هدف از این مرحله ارزیابی ویژگی تعمیم‌پذیری و شایستگی مدل‌های رگرسیونی و مدل عصبی انتخاب شده است. بنابراین مدل‌ها، با استفاده از مجموعه داده‌هایی غیر از مجموعه داده‌های آموزش (مجموعه داده‌های تست) ارزیابی گردیدند. برخی ویژگی‌های آماری داده‌های استفاده شده در فاز تست همراه با مقادیر پیش‌بینی شده آنها توسط مدل عصبی و رگرسیونی در جدول ۲ نشان داده شده است. تفاوت زیادی بین مقادیر پیش‌بینی شده شاخص هزینه تعمیر و مقادیر واقعی آنها برای دو مدل رگرسیونی و عصبی ملاحظه نمی‌شود. ولی تفاوت بین

مطالعه می‌باشد. بنابراین شبکه عصبی توانایی تخمین دقیق هر سه شاخص هزینه تعمیر و نگهداری را به طور هم‌زمان دارد. اگرچه برای هر شاخص هزینه تعمیر و نگهداری یک مدل رگرسیونی پیدا شد ولی مدل رگرسیونی تنها قادر به تخمین دقیق شاخص هزینه تعمیر می‌باشد.

رگرسیونی اختلاف معنی‌داری با هم دارند، این نشان از عدم توانایی مدل رگرسیونی در پیش‌بینی شاخص هزینه سوخت دارد و نیز مقادیر p مدل رگرسیونی شاخص هزینه روغن و سوخت بسیار کمتر از مدل عصبی آنها می‌باشند. این دلیلی دیگر بر بالا بودن توانایی شبکه عصبی در مقایسه با تکنیک رگرسیونی برای موضوع مورد

جدول ۱- نتایج تحلیل رگرسیونی داده‌های مجموعه آموزش.

R^2	F	ضرایب رگرسیونی			نوع مدل	شاخص هزینه
		β_2	β_1	β_0		
۰/۹۹۴	۱۲۰۳۵**	-	-۰/۰۰۷**	۲/۵۱**	درجه دوم	
۰/۹۹۹	۲۸۷۱۴**	-۱/۱×۱۰ ^{-۵} **	-۰/۰۳۱**	۳/۵۷**	درجه سوم	CCI _{repair}
۰/۹۹۶	۳۲۳۳۲**	-	۹/۲۴۳**	۰/۶۴۶**	توانی	
۰/۹۹۸	۷۵۴۵۹**	-	-۰/۲۲۹**	۱۵/۴۵**	درجه دوم	
۰/۹۹۹	۲۰۳۵۳۹**	۰/۰۱**	-۰/۵۹۵**	۱۷/۸۴**	درجه سوم	CCI _{oil}
۰/۹۹۷	۴۶۳۹۱**	-	۱۹/۴**	۰/۸۲۱**	توانی	
۰/۹۹۸	۴۰۸۷۱**	-	-۰/۳۹۹**	۱۹/۰۷*	درجه دوم	
۰/۹۹۹	۲۵۶۹۷**	۰/۰۱۸**	-۰/۹۲۸**	۲۲/۴**	درجه سوم	CCI _{fuel}
۰/۹۸۳	۸۳۸۱**	-	۲۱/۱۶**	۰/۸۳۲**	توانی	

توضیحات: مدل درجه دوم: $y = \beta_0 x + \beta_1 x^2$ ؛ مدل درجه سوم: $y = \beta_0 x + \beta_1 x^2 + \beta_2 x^3$ ؛ مدل توانی: $y = \beta_0 x^\beta$ ؛ x ساعات کارکرد جمعی بر حسب ۱۰۰ ساعت، y هزینه جمعی بر حسب درصدی از قیمت اولیه. ** معنی‌دار در سطح احتمال یک درصد، R^2 : ضریب تبیین.

جدول ۲- ویژگی‌های آماری مقادیر واقعی و پیش‌بینی شده شاخص‌های هزینه تعمیر و نگهداری در فاز تست.

ویژگی‌های آماری								نوع روش	شاخص هزینه
sum	ske	kur	max	min	std	var	av		
۴۴۸۶/۲	۰/۱۷۳	۱/۶	۱۴۳/۱۳	۰/۱۳۲	۴۷/۹	۲۲۹۴/۷	۶۲/۳۱	Reg و MLP	dv
۴۴۸۷/۹	۰/۱۷۷	۱/۵۹	۱۴۲/۵۳	۰/۱۳۸	۴۷/۸	۲۲۸۹/۵	۶۲/۳۳	MLP	CCI _{repair}
۴۴۹۰/۹	۰/۱۷۵	۱/۵۹	۱۴۲/۳۶	۰/۰۶۷	۴۷/۹	۲۲۹۶/۴	۶۲/۳۷	Reg	pv
۷۲۵/۱	۰/۲۴	۱/۷۴	۲۳/۲۲	۰/۰۶۱	۷/۲۴	۵۲/۳۷	۱۰/۱	Reg و MLP	dv
۷۲۵/۳	۰/۲۴۲	۱/۷۴	۲۳/۱۷	۰/۰۸۶	۷/۲۳	۵۲/۲۱	۱۰/۱	MLP	CCI _{oil}
۶۹۶	۰/۲۰۱	۱/۷۳	۲۱/۷۸	۰/۰۶	۶/۸۱	۴۶/۳۸	۹/۷	Reg	pv
۶۵۳/۱	۰/۴۸۹	۱/۹۶	۲۳/۰۴	۰/۰۷۲	۷/۰۵	۴۹/۷۹	۹/۰۷	Reg و MLP	dv
۶۵۳/۹	۰/۴۹۰	۱/۹۶	۲۳/۰۷	۰/۱۲۸	۷/۰۵	۴۹/۷۸	۹/۰۸	MLP	CCI _{fuel}
۵۸۴/۹	۰/۴۰	۱/۹۳	۲۰/۱۵	۰/۰۵۳	۶/۱	۳۶/۹۵	۸/۱	Reg	pv

توضیحات: av: میانگین، var: واریانس، std: انحراف معیار، min: کمینه، max: بیشینه، kur: کشیدگی، ske: چولگی، sum: مجموع، dv: مقادیر واقعی داده‌ها، pv: مقادیر پیش‌بینی شده توسط مدل، Reg: مدل رگرسیونی، MLP: شبکه عصبی.

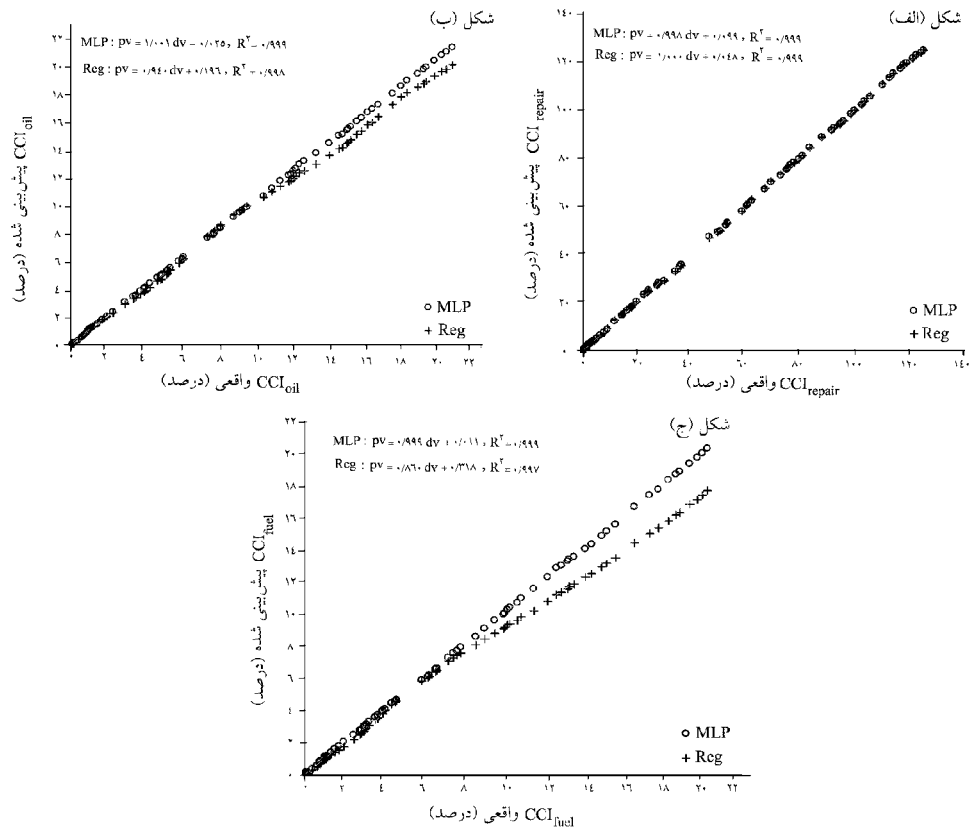
جدول ۳- مقایسه‌های آماری مقادیر واقعی و پیش‌بینی شده شاخص‌های هزینه تعمیر و نگهداری در فاز تست.

مقایسه توزیع	نوع تحلیل آماری		نوع روش	شاخص هزینه
	مقایسه واریانس	مقایسه میانگین		
۱/۰۰	۰/۹۸۵	۰/۹۹۹	MLP	CCI_{repair}
۱/۰۰	۰/۹۹۸	۰/۹۹۴	Reg	
۱/۰۰	۰/۹۹۰	۰/۹۹۸	MLP	CCI_{oil}
۰/۹۹۳	۰/۶۱۰	۰/۷۳۱	Reg	
۱/۰۰	۰/۹۹۹	۰/۹۹۳	MLP	CCI_{fuel}
۰/۷۴۱	۰/۲۱۱	۰/۳۸۹	Reg	

هم دارای شیب نزدیک‌تر به ۱ و عرض از مبدأ کوچک‌تری می‌باشد.

مقایسه براساس معیارهای عملکردی: در اینجا عملکرد دو روش مورد بحث از نظر معیارهای متوسط قدرمطلق درصد خطا و نیز ریشه متوسط مربعات خطا در دو فاز آموزش و تست با هم مقایسه شود (جدول ۴). با توجه به نتایج درج شده در این جدول، متوسط قدرمطلق درصد خطا و ریشه متوسط مربعات خطا در هر دو روش برای فاز تست بیشتر از فاز آموزش است زیرا داده‌های به‌کار گرفته شده در فاز تست کاملاً برای شبکه عصبی و مدل رگرسیونی تازگی دارند. مقایسه این دو روش در هر دو فاز آموزش و تست براساس این دو معیار نشان از برتری مطلق شبکه عصبی چند لایه پرسپترون در مقابل مدل رگرسیونی دارد زیرا متوسط قدرمطلق درصد خطا و ریشه متوسط مربعات خطای شبکه عصبی برای شاخص هزینه تعمیر، روغن و سوخت به ترتیب ۱۷/۶۰ و ۷۲ درصد کمتر از مدل رگرسیونی هستند. برتری دیگر شبکه عصبی پیش‌بینی هم‌زمان هر سه شاخص هزینه می‌باشد ولی در روش رگرسیونی باید برای هر شاخص یک مدل برازش شود زیرا در تکنیک رگرسیون تنها یک متغیر خروجی یا وابسته همراه با یک یا چند متغیر ورودی یا مستقل در نظر گرفت.

شکل ۳، همبستگی و رابطه رگرسیونی بین شاخص‌های هزینه واقعی و پیش‌بینی شده توسط مدل رگرسیونی و شبکه عصبی نشان می‌دهند. بهترین نتایج براساس این شکل‌ها وقتی به دست می‌آید که معادله خطی مابین شاخص هزینه واقعی و شاخص هزینه پیش‌بینی شده توسط مدل‌های برآورد شده علاوه بر داشتن ضرایب تبیین بالا دارای کمترین عرض از مبدأ (نزدیک به صفر) و شیب نزدیک به ۱ باشد ($p_v = 1/0.0 dv + 0/0.0$). کاملاً واضح است که ضرایب تبیین بین داده‌های واقعی و پیش‌بینی شده توسط شبکه عصبی چند لایه پرسپترون بسیار بالا می‌باشند ($R^2 = 0/999$). ولی ضریب تبیین برای مدل‌های رگرسیونی تعمیرات، روغن و سوخت به ترتیب برابر با ۰/۹۹۹، ۰/۹۹۸ و ۰/۹۹۷ می‌باشند. مقایسه رابطه خطی رگرسیونی بین شاخص هزینه واقعی تعمیر و نگهداری و پیش‌بینی شده توسط شبکه عصبی چند لایه پرسپترون و مدل رگرسیونی نشان از برتری مدل رگرسیونی دارد زیرا مدل رگرسیونی علاوه بر داشتن شیب ۱، عرض از مبدأ کوچک‌تری نسبت به مدل شبکه عصبی دارد. انجام این مقایسه برای شاخص هزینه روغن و سوخت دلالت بر توانایی شبکه عصبی نسبت به مدل رگرسیونی دارد زیرا شبکه عصبی برای این دو شاخص



شکل ۳- نمودار پراکندگی مقادیر واقعی و پیش‌بینی شده CCI_{repair} (شکل الف)، CCI_{oil} (شکل ب) و CCI_{fuel} (شکل ج) توسط شبکه عصبی و مدل رگرسیونی از مجموعه داده تست.

جدول ۴- مقایسه عملکرد شبکه عصبی و روش رگرسیونی در پیش‌بینی هزینه تعمیر و نگهداری تراکتور.

معیار عملکردی شبکه				نوع روش	شاخص هزینه
RMSE		MAPE			
فاز تست	فاز آموزش	فاز تست	فاز آموزش		
۰/۵۱۶	۰/۴۷۷	۲/۸۲	۲/۳۵	MLP	CCI_{repair}
۰/۶۲۱	۰/۵۷۷	۳/۷۲	۲/۹۷	Reg	
۰/۰۵۷	۰/۰۵۱	۲/۲	۱/۲۴	MLP	CCI_{oil}
۰/۶۳۸	۰/۵۹۲	۵/۲۱	۴/۱۲	Reg	
۰/۰۵	۰/۰۵۲	۲/۶۲	۱/۳۲	MLP	CCI_{fuel}
۱/۳۹	۱/۳۴	۹/۵۱	۹/۰۶	Reg	

منابع

1. Bowers, W., and Hunt, D.R. 1970. Application of mathematical formula to repair cost data. Transactions of the ASAE, 13: 806-809.
2. Fuls, J. 1999. The Correlation of repair and maintenance costs of agricultural machinery with operating hours management policy and operator skills for South Africa, available at <http://www.arc.agric.za>.
3. Gupta, M.M., Jin, J., and Homma, N. 2003. Static and Dynamic Neural Networks: From Fundamentals to Advanced Theory. John Wiley & Sons, Inc., Hoboken, New Jersey, 751p.
4. Haykin, S. 1994. Neural networks: A comprehensive foundation. McMillan College Publishing Company, New York, 837p.
5. Kaul, M., Hill, R.L., and Walthall, C. 2005. Artificial neural networks for corn and soybean yield prediction. Agriculture System, 85: 1-18
6. Kim, Y.H. 1989. A forecasting methodology for maintenance cost of long-life equipment. Doctoral thesis. University of Alabama.
7. Kumar, U.A. 2005. Comparison of neural networks and regression analysis: A new insight. Expert systems with applications, 29: 424-430.
8. Lyman, O., and Longnecker, M. 2001. An introduction to statistical methods and data analysis. R R. Donnelley & Sons, Inc./Willard, United States of America, Pp: 532-825.
9. Mitchell, Z.W. 1998. A Statistical Analysis of Construction Equipment Repair Costs Using Field Data & the Cumulative Cost Model. Ph.D. Thesis, Faculty of the Virginia Polytechnic Institute and State University.
10. Morris, J. 1988. Estimation of Tractor Repair and Maintenance Costs. Journal of Agricultural Engineering Research, 41: 191-200.
11. Oskounjad, M.M. 2003. Engineering economic evolution of industrial project. Amirkabir industrial Univ. Press, 417p.
12. Rotz, C.A. 1987. A Standard Model for Repair Costs of Agricultural Machinery. Applied Engineering in Agriculture, 3: 1. 3-9.
13. Telsang, M. 2005. Production management. S.chand & Company LTD, India, 476p.
14. Vakil-Baghmisheh, M.T., and Pavešic, N. 2001. Back-propagation with declining learning rate. P 297-300. Proceeding of the 10th Electrotechnical and Computer Science Conference, Portorož, Slovenia.
15. Vakil-Baghmisheh, M.T. 2002. Farsi Character Recognition Using Artificial Neural Networks. PhD Thesis, Faculty of Electrical Engineering, University of Ljubljana.
16. Vakil-Baghmisheh, M.T., and Pavešic, N. 2003. A Fast simplified fuzzy ARTMAP network. Neural Processing Letters, 17: 273-301.
17. Veelenturf, L.P.J. 1995. Analysis applications of artificial neural networks. Simon & Schuster international group, United States of America, 121p.
18. Zhang, Y.F., and Fuh, J.Y.H. 1998. A neural network approach for early cost estimation of packaging products. Comput Ind Eng., 34: 433-50.

Prediction of two-wheel drive tractor repair and maintenance costs using artificial neural network in comparing with regression

A. Rohani¹, I. Ranjbar², Y. Ajabshir², *M.H. Abbaspour-fard³ and M. Valizadeh⁴

¹Ph.D. Student, Dept. of Farm Machinery Engineering, University of Tabriz, Iran, ²Associate Prof., Dept. of Farm Machinery Engineering, University of Tabriz, Iran, ³Assistant Prof., Dept of Farm Machinery Engineering, Ferdowsi University of Mashhad, Iran, ⁴Professor, Dept. of Agronomy, University of Tabriz, Iran

Abstract

The highest profit is the ultimate goal of any farming firm. The management of machine replacement (specifically tractor) is one of the most critical factors for achieving this goal. Therefore, accurate prediction of tractor repair and maintenance costs is a must. The objective of this research was to predict tractor repair and maintenance costs using neural network and regression methods and comparing their performance for these predictions. The study was conducted using empirical data on 60 two-wheel drive tractors from Astan Ghodse Razavi agro-industry. The results indicated that both neural network and regression models have high rather accurate prediction of repair cost but the performance of regression model in comparison with neural network model for prediction of oil and fuel costs was very low with much less accuracy.

Keywords: Neural network; Regression; Repair and maintenance cost

*- Corresponding Author; Email: hossein_abbaspour@yahoo.co.uk