

Hybrid Artificial Neural Network with Meta-heuristic Algorithms for Predicting Sugarcane Yield

NEGAR HAFEZI^{1*}, HOUSHANG BAHRAMI¹, MOHAMMAD JAVAD SHEIKH DAVOODI¹, SEYED ENAYATALLAH ALAVI²

1. Biosystems Engineering Department, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran

2. Computer Engineering Department, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran

(Received: Oct. 22, 2019- Revised: May. 10, 2020- Accepted: June. 13, 2020)

ABSTRACT

In this study, to predict sugarcane yield, an Artificial Neural Network based on meta-heuristic algorithms is used as an efficient method to estimate crop yield based on actual data. In order to predict sugarcane yield, effective parameters such as plant characteristics, electrical conductivity of soil and water, maximum temperature and average hours of sunshine during the growing season and on a time scale of seven years were used. Accordingly, four hybrid models were used to build neural networks which including of artificial neural network based on Back Propagation (BP) algorithm, combining neural network with Genetic Algorithm (GA), combining neural network with Particle Swarm Optimization (PSO) algorithm and finally a new approach as combining neural network with GA-PSO. The results show that the neural network performance can be improved using meta-heuristic algorithms and can be increased significantly the prediction power. The Mean Square Error (MSE) and correlation coefficient (R) in the hybrid method of neural network with GA-PSO were obtained 0.00057 and 0.91457 respectively on the test data, which show the superiority of this method to other patterns. In addition, cross validation test of the proposed model by K-Nearest Neighbors showed the accuracy of training and test data to predict sugarcane yield has been equal to 98.5% and 95.5%, respectively.

Keywords: Sugarcane, Yield, Artificial Neural Network, Meta-heuristic Algorithms, Cross Validation

هیبرید شبکه عصبی مصنوعی با الگوریتم‌های فرااکتشافی در پیش‌بینی عملکرد نیشکر

نگار حافظی^{۱*}، هوشنگ بهرامی^۱، محمد جواد شیخ داودی^۱، سید عنایت‌اله علوی^۲

۱. گروه مهندسی بیوسیستم، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران

۲. گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه شهید چمران اهواز، اهواز، ایران

(تاریخ دریافت: ۱۳۹۸/۷/۳۰ - تاریخ بازنگری: ۱۳۹۹/۲/۲۱ - تاریخ تصویب: ۱۳۹۹/۳/۲۴)

چکیده

در این پژوهش برای پیش‌بینی عملکرد نیشکر از شبکه عصبی مصنوعی مبتنی بر الگوریتم‌های فرااکتشافی به عنوان یک روش کارآمد که تخمین عملکرد محصول را بر اساس داده‌های واقعی به دست می‌آورد، استفاده شده است. به منظور پیش‌بینی عملکرد نیشکر از پارامترهای مؤثری نظیر مشخصات گیاه، هدایت الکتریکی آب و خاک، حداکثر درجه حرارت و میانگین ساعات آفتابی در طول دوره رشد محصول و در مقیاس زمانی هفت ساله استفاده گردید. بر همین اساس از چهار الگوی آموزشی در ساخت شبکه‌های عصبی استفاده شد که به ترتیب شامل شبکه عصبی مصنوعی مبتنی بر الگوریتم پس انتشار خطا، ترکیب شبکه عصبی با الگوریتم ژنتیک، ترکیب شبکه عصبی با الگوریتم ازدحام ذرات و در آخر رویکرد نوین ترکیب شبکه عصبی با الگوریتم مرکب ژنتیک- ازدحام ذرات بوده‌اند. نتایج بدست آمده از پژوهش، حاکی از آن است که استفاده از الگوریتم‌های فرااکتشافی سبب بهبود عملکرد شبکه عصبی شده و قدرت پیش‌بینی را به طور محسوسی افزایش داده است. مقادیر میانگین مربعات خطا و ضریب همبستگی برای روش ترکیبی شبکه عصبی، الگوریتم ژنتیک و الگوریتم ازدحام ذرات به ترتیب $0/00057$ و $0/91457$ برای داده‌های آزمایش بدست آمد که برتری این روش را نسبت به سایر الگوها نشان می‌دهد. علاوه بر این، اعتبارسنجی متقاطع مدل پیشنهادی به روش k نزدیک‌ترین همسایه نشان داد که صحت داده‌های آموزشی و آزمایشی به منظور پیش‌بینی عملکرد نیشکر به ترتیب برابر با $98/5$ و $95/5$ درصد بوده است.

واژه‌های کلیدی: نیشکر، عملکرد، شبکه عصبی مصنوعی، الگوریتم‌های فرااکتشافی، اعتبارسنجی متقاطع

مقدمه

نیشکر سهم عظیمی از اراضی کشاورزی استان خوزستان را به خود اختصاص داده است. عملکرد نیشکر تحت تأثیر عوامل مختلفی قرار دارد اما از آنجایی که امکان بررسی هم‌زمان اثرات مدیریتی، آب، خاک و اقلیم روی گیاه نیشکر در روش میدانی وجود ندارد؛ بنابراین معرفی مدلی موفق و قابل اعتماد کمک شایانی برای این مشکل است (Golabi et al., 2013).

با پیشرفت علوم مختلف استفاده از روش‌های فرااکتشافی مانند روش‌های مبتنی بر هوش مصنوعی برای حل مسائل پیچیده، پژوهشگران را برای استفاده از این روش‌ها در مدل‌سازی فرآیند تصمیم‌گیری ترغیب کرده است. یکی از این پیشرفت‌ها در زمینه هوش مصنوعی، شبکه‌های عصبی مصنوعی^۱ (ANN) هستند که به عنوان ابزار قدرتمندی در پردازش اطلاعات غیرخطی، قادر به انجام موفقیت‌آمیز اعمالی مانند تقریب توابع غیرخطی، طبقه‌بندی الگوها، تشخیص الگوها و پیش‌بینی هستند

(AmirEntezari, 2008). یکی از مهم‌ترین کاربردهای شبکه‌های عصبی مصنوعی، طراحی مدلی است که بتوان براساس آن مقدار یک یا چند متغیر وابسته را به کمک متغیرهای مستقل پیش‌بینی کرد.

اولین گام در آموزش یک شبکه عصبی، ارائه الگوهای است که شبکه بتواند با استفاده از آن‌ها آموزش داده شود. الگوریتم‌های بسیاری به منظور آموزش وزن‌های شبکه عصبی وجود دارد. از جمله این الگوریتم‌ها می‌توان به الگوریتم آموزش پس انتشار^۲ اشاره کرد. این الگوریتم الگوی آموزشی ورودی را جلو برده سپس خطای مربوطه را محاسبه کرده و به عقب انتشار داده و مجدداً وزن‌ها را تنظیم می‌کند. در استفاده از این روش احتمال به دام افتادن شبکه در نقاط مینیمم محلی بسیار بالاست. برای حل این مشکل، استفاده از الگوریتم‌های فرااکتشافی به منظور بهبود دقت آموزش شبکه‌های عصبی به شدت مورد توجه قرار گرفته است (Kia, 2011). الگوریتم‌های فرااکتشافی، روش‌های هوشمندی هستند که بر پایه‌ی فرآیند جست‌وجوی تصادفی و تئوری

* نویسنده مسئول: n-hafezi@stu.scu.ac.ir

نشان داد که الگوریتم ترکیبی با دقت بالایی توانسته است به پیش‌بینی عملکرد زردآلو بپردازد. به طوری که مقدار ریشه میانگین مربعات خطا در دو مرحله آموزش و آزمایش داده‌ها به ترتیب برابر با ۱/۷۳۷ و ۲/۳۲۹ بدست آمد. در بخش تحلیل حساسیت، نتایج حاکی از آن بود که از بین متغیرهای مورد بررسی، شش پارامتر تعداد دفعات آبیاری در دوره‌ی رشد (۲۹٪)، فاصله‌ی درختان در یک ردیف (۱۳٪/۴)، فاصله ردیف‌ها (۱۳٪)، میزان منیزیم برگ (۵/۵٪)، مقدار نیتروژن برگ (۵٪) و نوع کاشت (۴٪) توانسته‌اند تقریباً ۷۰٪ تغییرات عملکرد زردآلو را بیان نمایند. در پژوهشی (Bagheri and Sohrabi, 2018) اثر پارامترهای اقلیمی را بر روی عملکرد محصول جو دیم و آبی مورد مطالعه قرار دادند. نتایج نشان داد که حساسیت شبکه عصبی ساخته شده به متغیرهای ورودی مدل (عوامل اقلیمی مورد نظر) برای جو دیم نسبت به جو آبی بالاتر بوده است. در نتیجه این شبکه برای پیش‌بینی عملکرد جو دیم قابل اطمینان‌تر بود. (Hosseini et al., 2007) به برآورد عملکرد گندم دیم با استفاده از شبکه عصبی مصنوعی و رگرسیون چند متغیره پرداختند. در این پژوهش از داده‌های هواشناسی ۱۰ ساله به منظور تخمین عملکرد گندم دیم استفاده شد. نتایج نشان داد که مدل شبکه عصبی مصنوعی، مقدار عملکرد گندم دیم را قبل از برداشت محصول با قابلیت برآورد بسیار خوب پیش‌بینی نمود. نتایج روابط رگرسیونی نیز نشان داد که دو عامل بارندگی سالیانه و حداقل رطوبت نسبی هوا (برحسب درصد) در ماه خرداد ۹۴/۳٪ تغییرات عملکرد گندم دیم را در منطقه توضیح می‌دهند. (Hosseini et al., 2016) در پژوهش خود به این نتیجه رسیدند که استفاده از الگوریتم‌های فرااکتشافی ازدحام ذرات و ژنتیک نسبت به مدل رگرسیون چند متغیره نتایج مطلوب‌تری در پیش‌بینی مقاومت مکانیکی خاک داشته است.

شوری آب و خاک از عمده‌ترین مشکلات کشاورزی در نواحی خشک و نیمه خشک دنیاست. در این نواحی کافی نبودن آب با کیفیت، وجود گرما و اقلیم بسیار خشک غالباً علت اصلی شوری می‌باشند که تولید گیاهان به خصوص نیشکر را در این نواحی محدود می‌کند. میزان تولید محصولات کشاورزی صرف نظر از وابستگی به خصوصیات ژنتیکی رقم، شرایط خاک، آب و فعالیت‌های مدیریتی، به شدت تحت تأثیر عوامل اقلیمی در طول دوره رشد گیاه قرار دارد. به عبارت دیگر تأثیر عناصر و عوامل اقلیمی بر روی جنبه‌های مختلف فعالیت بشری به ویژه بخش کشاورزی امری بدیهی است؛ از میان عناصر اقلیمی پارامترهای دمایی و تابش خورشید جزء مهم‌ترین عواملی هستند که نقش تعیین‌کننده بر تولید محصول نیشکر دارند. بدین منظور هدف از

احتمالات و با الهام از فرآیندهای تکاملی محیط پیرامون در زمینه‌های مختلف شکل گرفته‌اند. به بیان دیگر، روش‌های فرااکتشافی فضای کاوش را به صورت نقطه به نقطه جست‌وجو می‌کنند و در سیر فرآیند بهینه‌یابی با پیروی از منطق الگوریتم به سوی نقطه بهینه حرکت می‌کنند (Yang, 2008). تعدادی از معروف‌ترین الگوریتم‌های فرااکتشافی عبارتند از: الگوریتم ژنتیک، ازدحام ذرات، رقابت استعماری، الگوریتم کلونی مورچگان و در حال حاضر ترکیب اجزای مختلف الگوریتم‌ها به یکی از گرایش‌های موفق در بهینه‌سازی تبدیل شده است. تکنیک ترکیب الگوریتم‌ها یک استراتژی کارآمد و انعطاف‌پذیر برای حل مسائل دنیای واقعی با اندازه بزرگ را فراهم می‌کند که در آن نقاط قوت هر دو طرف ترکیب، با یکدیگر ادغام شده و ساختار مؤثری را ایجاد می‌کند (Jeong et al., 2009). هر دو الگوریتم ژنتیک و ازدحام ذرات از جمله الگوریتم‌های جمعیت محور هستند که طرفداران و منتقدان خاص خود را دارند. الگوریتم ژنتیک توانمندی بالایی در کاوش همزمان همه فضای جست‌وجو را دارد اما قدرت جست‌وجوی محلی نسبتاً ضعیفی دارد. در آن طرف، الگوریتم ازدحام ذرات قدرت جست‌وجوی محلی بالایی دارد. این الگوریتم توانمندی بالایی در حفظ اطلاعات با ارزش تکرارهای قبل خود را دارد. ولی در عوض در الگوریتم ژنتیک، اطلاعات یک کروموزوم در صورتی که به نسل بعد خود منتقل نشود، از بین می‌رود (Settles and Soule, 2005). بنابراین در این پژوهش سعی شده است که با ترکیب دو الگوریتم از قدرت جست‌وجوی سراسری (کلی) و محلی قوی همزمان استفاده کرد.

طبق تحقیقات انجام شده استفاده از شبکه عصبی برای تخمین عملکرد محصولات زراعی نتایج مطلوبی را به همراه داشته و در مقایسه با برخی روش‌های دیگر از دقت بالاتری برخوردار بوده است. (Pourmohammadali et al., 2019) در پژوهشی اثرات خصوصیات خاک، کیفیت آب، گیاه و رویکردهای مدیریتی را بر عملکرد پسته‌ی رفسنجان با استفاده از شبکه عصبی مصنوعی بهینه شده با الگوریتم ژنتیک و مدل رگرسیون خطی چندگانه مورد بررسی قرار دادند. نتایج نشان داد که مدل رگرسیونی تنها به میزان ۲۸٪ توانسته است تغییرات عملکرد محصول را پیش‌بینی نماید؛ در حالی که دقت مدل ترکیبی شبکه عصبی و الگوریتم تکاملی ژنتیک برای پیش‌بینی عملکرد پسته، براساس متغیرهای ورودی، به میزان ۹۰٪ بوده است. (Esfandiarpour- Boroujeni et al., 2019) در مطالعه‌ای به تخمین عملکرد زردآلو با استفاده از الگوریتم ترکیبی ازدحام ذرات، رقابت استعماری^۱ و رگرسیون برداری پشتیبان^۲ پرداختند. نتایج اعتبارسنجی مدل

مواد و روش‌ها

به منظور پیش‌بینی عملکرد نیشکر از یکسری اطلاعات زراعی، خصوصیات خاک، آب و هواشناسی مربوط به هفت سال برداشت متوالی از آبان ماه سال ۱۳۹۰ تا اردیبهشت ماه سال ۱۳۹۷ استفاده شد. داده‌های مورد نظر از واحد کشت و صنعت دعبل خزاعی واقع در ۳۰ کیلومتری جنوب شهرستان اهواز و شرق رودخانه کارون تهیه شد. برای ساخت شبکه‌های عصبی مصنوعی در این پژوهش به تعداد ۲۰۱۰ سری داده مورد استفاده قرار گرفت که شامل مجموع تعداد مزارع منتخب برحسب کلاس محصول نیشکر (تازه کشت، راتون سال اول تا چهارم) در دوره‌ی زمانی مورد مطالعه بود. در جدول (۱) خلاصه توصیف آماری عوامل مؤثر در تخمین عملکرد نیشکر ارائه شده است.

مطالعه حاضر، پیش‌بینی عملکرد نیشکر براساس پارامترهای مؤثری نظیر کلاس محصول، رقم نیشکر، هدایت الکتریکی خاک، هدایت الکتریکی آب آبیاری، حداکثر درجه حرارت و متوسط ساعات آفتابی در طی دوره رشد است (Heydarnejadi, 2016; Ahmadvand, 2009) که با توجه به حجم بالای داده‌ها و پیروی آن‌ها از روند غیرخطی، تکنیک ترکیبی الگوریتم‌های فرااکتشافی و شبکه عصبی پرسپترون چند لایه برای تخمین بهتر عملکرد نیشکر ارائه شده است. در این پژوهش تمرکز بر روی بهینه ساختن وزن‌های شبکه عصبی پرسپترون چند لایه با الگوریتم‌های ژنتیک، ازدحام ذرات و ترکیب این دو الگوریتم هوشمند است. از آنجایی که تعداد وزن یال‌ها و بایاس‌ها در شبکه عصبی مصنوعی بالاست این الگوریتم‌ها در یافتن بهترین مقادیر برای وزن یال‌ها و بایاس‌ها می‌توانند مورد استفاده قرار گیرند.

جدول ۱- توصیف آماری پارامترهای منتخب برای مدل شبکه عصبی

نام متغیر	دامنه تغییرات	میانگین	انحراف معیار
کلاس محصول	[۱-۵]	۳	۱/۳۱
رقم	[۱-۳]	۲	۰/۷۴
هدایت الکتریکی خاک (ds/m)	[۳/۴-۱۲/۴]	۳۰/۹۸	۱۵/۶۹
هدایت الکتریکی آب آبیاری (μs/cm)	[۰-۳/۸]	۲/۴۵	۰/۴۳
حداکثر درجه حرارت (°C)	[۷/۷۷-۵۳/۲۸]	۳۴/۸۹	۱۰/۳۴
میانگین ساعات آفتابی (h)	[۰-۱۳/۳]	۸/۴۲	۳/۴۲
عملکرد نیشکر (t/ha)	[۱۵/۹۰-۱۵۸/۴۸]	۷۱/۹۲	۲۲/۶۰

در یک محدوده عددی خاصی قرار گیرد. هر شبکه عصبی از تعدادی نورون که به اصطلاح گره نامیده می‌شود، تشکیل شده است. این گره‌ها بسته به نوع شبکه، به روش‌های مختلف توسط اتصالات وزن دار (یال) با یکدیگر در ارتباط هستند. هر شبکه داده‌های ورودی را گرفته و خروجی متناظر با آن را تولید می‌نماید؛ اما برای این که خروجی تولید شده خروجی صحیحی باشد، باید شبکه از قبل آموزش دیده باشد.

آموزش شبکه به معنی تعیین وزن‌ها و پارامترهای شبکه است، به طوری که خروجی شبکه به خروجی واقعی بسیار نزدیک باشد. بدین منظور باید پارامترها و وزن‌های بهینه را بدست آورد. طبق پژوهش‌های انجام شده یک شبکه پرسپترون با یک لایه میانی (پنهان) قادر به تقریب زدن هر تابع غیرخطی می‌باشد (Menhaj, 2008).

شکل (۱) نورونی با n ورودی را نشان می‌دهد. بردار ورودی با X نمایش داده شده است. اسکالرهایی $(j=1,2,3,\dots,n)$ عناصر x_j بردار X هستند. مجموعه‌ی سیناپس‌های w عناصر ماتریس W را تشکیل می‌دهند.

برای دو متغیر توصیفی کلاس و رقم محصول مقادیر کمی تعریف شده است؛ به عنوان مثال در مورد کلاس محصول مزارعی که محصول آن‌ها تازه کشت (پلنت) بوده و نیز مزارعی که رقم نیشکر کشت شده در آن‌ها از نوع زودرس بوده مقدار یک فرض شده است. کمی‌سازی برای سایر کلاس‌های محصول (راتون سال‌های اول تا چهارم) و ارقام نیشکر (میان‌رس و دیررس) نیز به همین ترتیب انجام شد.

شبکه عصبی مصنوعی

شبکه‌های عصبی زیستی، مجموعه‌ای بسیار عظیم از پردازشگرهای کوچکی به نام نورون هستند که به صورت موازی و هماهنگ برای حل مسئله عمل می‌کنند. هر نورون از دو بخش تشکیل شده است. در بخش نخست، مجموع وزنی مقادیر ورودی محاسبه می‌شود. در بخش دوم، خروجی بخش نخست در یک تابع ریاضی قرار گرفته و خروجی نورون محاسبه می‌شود. این تابع ریاضی، اصطلاحاً تابع فعالیت یا آستانه نام دارد که عملکرد آن شبیه یک فیلتر غیرخطی است و موجب می‌شود تا خروجی نورون

آموزش، ۱۰٪ به اعتبارسنجی و ۲۰٪ باقی به آزمایش شبکه اختصاص داده شد. به کمک کدنویسی در نرم‌افزار MATLAB نسخه‌ی ۲۰۱۷، شبکه‌های عصبی با شش نورون در لایه ورودی تشکیل شدند. لایه خروجی نیز شامل یک نورون که همان پیش‌بینی میزان عملکرد نیشکر بود، در نظر گرفته شد. پس از ساخت شبکه و انتخاب بهترین شبکه براساس آزمایش و خطا، تعداد نورون‌های لایه پنهان نیز مشخص گردید. سپس تنظیم وزن‌ها و اعمال مقادیر آستانه‌ی شبکه بر روی تمام ورودی‌های شبکه انجام گرفت. برای یافتن بهترین شبکه‌ی توصیف‌کننده عملکرد نیشکر، از ترکیب شبکه عصبی چند لایه پرسپترون با قواعد یادگیری پس انتشار خطا (BP)، الگوریتم ژنتیک (GA)، الگوریتم بهینه‌سازی ازدحام ذرات (PSO) و الگوریتم ترکیبی ژنتیک- ازدحام ذرات (GA-PSO) استفاده گردید. در نهایت شبکه‌ای که کمترین خطای مدل‌سازی را به همراه داشت، انتخاب شد.

آموزش شبکه عصبی مصنوعی

آموزش شبکه عصبی، شامل الگوریتم‌هایی برای بهینه‌سازی وزن یال‌ها و بایاس‌های شبکه است. هدف از آموزش شبکه عصبی حداقل نمودن تابع هزینه است. در شبکه‌های عصبی اغلب از مقدار میانگین مربعات خطا (MSE) به عنوان معیار سنجش خطا استفاده می‌شود که در رابطه (۴) نشان داده شده است.

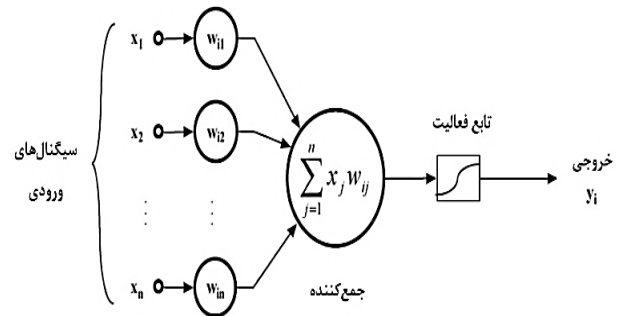
$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (\text{رابطه ۴})$$

که در آن، n تعداد داده‌ها در مرحله آموزش شبکه، y_i خروجی واقعی مربوط به نمونه‌ی i ام و \hat{y}_i خروجی مطلوب شبکه از ورودی i ام است. خطای MSE بعد از هر بار تکرار محاسبه می‌گردد. در واقع این خطا، تابع هدفی است که باید بهینه شود. روند آموزش زمانی متوقف می‌شود که MSE به اندازه‌ی کافی کوچک شده باشد و یا یکی از شرایط توقف رخ داده باشد. هر چه معیار خطای پیش‌بینی کمتر باشد نمایانگر پیش‌بینی دقیق‌تر است.

معیار دیگر برای تعیین عملکرد شبکه آموزش دیده، ضریب همبستگی می‌باشد که نشانگر میزان بازده شبکه است. این ضریب با R نمایش داده می‌شود و با استفاده از رابطه (۵) محاسبه می‌گردد.

$$R = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (\text{رابطه ۵})$$

که در آن \bar{y} میانگین مقادیر خروجی واقعی می‌باشد. بهترین جواب برای مدل هنگامی ایجاد خواهد شد که R و MSE به ترتیب



شکل ۱- ساختار یک نورون مصنوعی با چند ورودی

در این حالت، w یک بردار سطری با عناصر w_{ij} که $i = \{1, 2, \dots, k\}$ و k نورون در لایه خروجی است. هر عنصر از بردار ورودی x در عنصر متناظر از w ضرب می‌شود. این ورودی‌ها می‌توانند ورودی‌های شبکه عصبی و یا سیگنال خروجی از سایر نورون‌های موجود در شبکه باشند هر یک از ورودی‌ها در وزن متناظر خود ضرب شده و سپس با یکدیگر جمع می‌گردند تا سیگنال net_j^h تولید شود که h معرف شماره لایه می‌باشد همچنین نورون می‌تواند یک جمله بایاس b (ب) نیز داشته باشد. مقدار این سیگنال از طریق رابطه (۱) محاسبه می‌شود:

$$net_j^h = \sum_{i=1}^n x_i w_{ij} + b_i = w \underline{x} + b \quad (\text{رابطه ۱})$$

سرانجام خروجی نورون به شکل زیر خواهد بود:

$$y_i = f(w \underline{x} + b) \quad (\text{رابطه ۲})$$

که در آن f تابع فعالیت می‌باشد.

پیش پردازش

قبل از آموزش شبکه عصبی، داده‌های ورودی به آن نرمال می‌شوند. هدف از نرمال‌سازی آن است که داده‌ها به اعدادی بین صفر و یک تبدیل شوند. برای ورودی‌های نزدیک به صفر و یک، تغییرات وزن نورون‌ها حداقل خواهد بود؛ زیرا در این اعداد عناصر پردازش‌گر (نورون‌ها) به دلیل شکل توابع آستانه نظیر تابع سیگموئیدی (تانژانت/ لگاریتم سیگموئیدی)، کند عمل می‌کنند ولی برای مقادیر ورودی‌های نزدیک به نیم (۰/۵) پاسخ نورون‌ها به سیگنال ورودی سریع‌تر خواهد بود (Kumar et al., 2002). با در نظر گرفتن این واقعیت، نرمال‌سازی داده‌ها به نحوی صورت گرفت که میانگین سری داده‌ها برابر ۰/۵ گردد. بدین منظور از رابطه (۳) برای نرمال‌سازی استفاده شد.

$$x_{norm} = 0.5 \left(\frac{x_0 - \bar{x}}{x_{max} - x_{min}} \right) + 0.5 \quad (\text{رابطه ۳})$$

که در آن x_{norm} مقدار نرمال شده ورودی x_0 ، x_{max} میانگین داده‌ها، x_{min} به ترتیب حداکثر و حداقل داده‌ها می‌باشد. پس از نرمال‌سازی، داده‌ها به سه دسته آموزش، اعتبارسنجی و آزمایش تقسیم‌بندی شدند. ۷۰٪ داده‌ها به مرحله

به یک و صفر میل کنند.

یادگیری فرااکتشافی

روش یادگیری فرااکتشافی از این جهت که توانایی دستیابی به بهینه‌ی سراسری در مسائلی با فضای وسیع، پیچیده، چندحالتی و غیر قابل تشخیص را دارد، مورد توجه است. در این روش نیاز به اطلاعاتی در رابطه با تابع خطا نبوده، و بنابراین در مواردی که این اطلاعات موجود نباشد و یا دستیابی به آن‌ها بسیار پرهزینه باشد، استفاده از این روش می‌تواند مفید باشد. یادگیری فرااکتشافی قابلیت کاربرد در هر دو فضای نظارت شده و غیر نظارت شده را دارد. دو نمونه از پرکاربردترین الگوریتم‌های فرااکتشافی الگوریتم ژنتیک و ازدحام ذرات است.

الگوریتم ژنتیک

الگوریتم‌های ژنتیک از مجموعه‌ای از جواب‌های اولیه با برازندگی^۱ معین شروع می‌کنند و مجموعه‌ای از جواب‌ها را تولید می‌کنند. مجموعه جواب‌ها در هر مرحله جمعیت (نسل) نامیده می‌شوند. از این رو، در یک فرآیند تکاملی با شروع از یک جمعیت اولیه، در هر تکرار جمعیت بهتری به دست می‌آید. پس از ایجاد جمعیت اولیه، حال با استفاده از عملگر انتخاب (چرخ رولت^۲، ترتیبی^۳ و ...) تعدادی از کروموزوم‌ها برای زاد و ولد و ایجاد نسل بعدی انتخاب می‌شوند. هر کروموزوم نشان‌دهنده یک نقطه در فضای جستجو و یک راه‌حل ممکن برای مسأله موردنظر است. خود کروموزوم‌ها (راه‌حل‌ها) از تعداد ثابتی ژن (متغیر) تشکیل شده‌اند؛ در شبکه عصبی به ازای هر یک از یال‌ها، بایاس‌ها و وزن متناظر با آن، یک ژن اختصاص داده می‌شود که مقادیر تصادفی در بازه (۱-۱) دارند. در گام بعدی با عملگر برش یا ادغام^۴، فرزندان^۵ ساخته می‌شوند. این عملگر روی همه والدین انجام نمی‌شود بلکه با احتمال برش (Pc) صورت می‌گیرد. به‌منظور ایجاد نسل جدید از رابطه (۶) استفاده می‌شود.

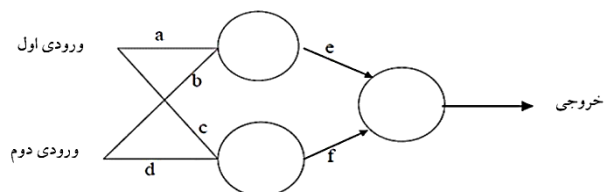
$$S_{new} = \alpha \times S_{in} + (1 - \alpha) \times S_{jn} \quad (\text{رابطه } ۶)$$

که در آن، α یک عدد تصادفی در فاصله‌ی [۰،۱] می‌باشد. S_{in} و S_{jn} به ترتیب متغیر nام در کروموزوم مادر و پدر هستند. عملگر جهش^۶ به طور معمول بعد از عملگر برش اعمال می‌شود. این عملگر برای تنوع بخشیدن به جمعیت و ایجاد نقاط جستجوی جدید اعمال می‌شود که از رابطه (۷) می‌توان آن را محاسبه کرد (Rezaei and Ranjbaran, 2009). پس از آن فرآیند تولید تا زمانی که جواب مورد نظر حاصل شود، ادامه می‌یابد.

$$M_{new} = M_i + \sigma N(0,1) \quad (\text{رابطه } ۷)$$

در رابطه فوق، M_i نشان‌دهنده‌ی متغیر انتخاب شده تصادفی از یک کروموزوم و σ معرف انحراف استاندارد در توزیع نرمال $N(0,1)$ است.

در پژوهش حاضر، براساس متغیرهای ورودی، پس از تعیین بهترین ساختار شبکه از لحاظ تعداد لایه پنهان، تعداد نورون در لایه‌های پنهان و خروجی و نیز توابع فعالیت بکار رفته در این دو لایه، حال برای آموزش شبکه عصبی با الگوریتم ژنتیک، وزن یال‌ها و بایاس‌های شبکه به عنوان جمعیت اولیه (مجموعه راه‌حل‌ها) وارد الگوریتم ژنتیک می‌شوند. در واقع در این روش بهینه‌سازی، مجهولات مسأله همان وزن و بایاس‌های نورون‌های موجود در لایه‌های مختلف شبکه می‌باشند (شکل ۲). بعد از آن برازندگی کروموزوم‌های جمعیت اولیه که بر مبنای کمینه کردن خطا (اختلاف بین داده‌های واقعی و داده‌های بدست آمده توسط شبکه عصبی) است، بررسی می‌شود. در ادامه شرط پایان الگوریتم که تعداد جمعیت می‌باشد، بررسی می‌شود. اگر این شرط برقرار نشده باشد عملگرهای الگوریتم ژنتیک مانند انتخاب که در این پژوهش از نوع چرخ رولت بکار رفته است. دیگر عملگرهای برش و جهش اعمال می‌شوند و در نهایت جمعیت جدیدی به وجود می‌آید و وارد چرخه تکرار شده و تا زمانی که شرط پایان برقرار شود، ادامه پیدا می‌کند. سپس الگوریتم ژنتیک متوقف می‌شود و پارامترهای بهینه شده به شبکه اعمال می‌شوند و شبکه آموزش می‌بیند.

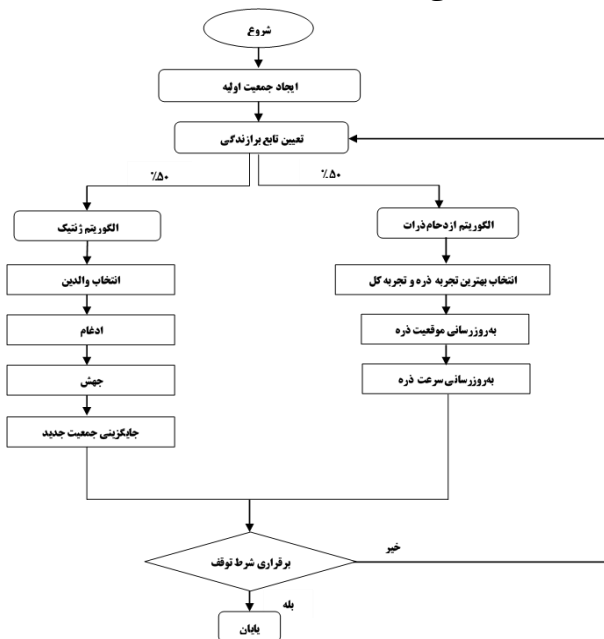


شکل ۲- وزن و بایاس‌های شبکه به عنوان ژن در کروموزوم

الگوریتم ازدحام ذرات

روش ازدحام ذرات یک روش سراسری بهینه‌سازی است که با استفاده از آن می‌توان با مسائلی که جواب آن‌ها یک نقطه یا سطح در فضای n بعدی می‌باشد، برخورد نمود. در چنین فضایی، فرضیاتی مطرح و یک سرعت ابتدایی به ذرات اختصاص داده می‌شود، همچنین کانال‌های ارتباطی بین ذرات در نظر گرفته می‌شود. سپس این ذرات در فضای پاسخ حرکت می‌کنند، و نتایج بدست آمده بر مبنای یک ملاک برازندگی پس از هر بازه‌ی زمانی

جمعیت اولیه (وزن یال‌ها و بایاس‌های شبکه) با k کروموزوم را به صورت تصادفی در فضای جستجو ایجاد می‌کند. شایستگی هر ذره توسط تابع برازندگی محاسبه می‌شود و ذرات بر اساس شایستگی مرتب می‌شوند. نیمی از افراد برتر جمعیت (بر اساس شایستگی)، به عنوان والدین انتخاب می‌شود؛ عملگرهای ژنتیکی برش (ادغام) و جهش بر روی این والدین اعمال می‌شوند؛ سپس از بین جمعیت ایجاد شده (والدین و فرزندان) در الگوریتم ژنتیک، به تعداد جمعیتی که در ابتدا به GA ارسال شده بود، افراد شایسته‌تر انتخاب شده و جایگزین همان نیمی شایسته انتخاب شده از جمعیت می‌شوند.



شکل ۳- فلوچارت ترکیب الگوریتم ژنتیک و ازدحام ذرات

برای هر کروموزوم باید مقدار تابع برازندگی با خطای کوچک در مجموعه داده‌های ورودی، مینیمم شود که این مقدار مینیمم از طریق رابطه (۱۱) قابل محاسبه است. از این رو، عملگرهای الگوریتم ازدحام ذرات برای به‌روزرسانی سرعت و موقعیت نیمی از جمعیت با شایستگی کمتر اعمال می‌شوند. در نهایت، پس از چندین تکرار، کروموزومی که کمترین برازندگی را داشته باشد به عنوان وزن بهینه انتخاب می‌شود. در غیر این صورت این فرآیند بار دیگر تکرار می‌شود.

$$\hat{y}_i = \text{sign}(\sum_{j=1}^S X_{i,j} W_{k,j}) \quad (\text{رابطه } 10)$$

$$Fitness_k = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|^2}{N} \quad (\text{رابطه } 11)$$

در این رابطه N و S به ترتیب تعداد نمونه‌های ورودی فاز آموزش و تعداد جواب‌های مساله می‌باشند. y_i مقدار هدف مربوط به نمونه‌ی i ام و \hat{y}_i مقدار خروجی مطلوب شبکه از ورودی i ام، sign تابع علامت که در بازه $[-1, 1]$ تغییر می‌کند، $X_{i,j}$ مقدار ورودی i ام در ژن k ام و $W_{k,j}$ وزن ژن k ام از کروموزوم k ام است.

محاسبه می‌شود. با گذشت زمان، ذرات به سمت ذراتی که دارای ارزش برازندگی بالاتری هستند و در گروه ارتباطی یکسانی قرار دارند، شتاب می‌گیرند. جهت حرکت هر ذره متأثر از دو مقدار بهترین مکان ملاقات شده توسط خود ذره (P_{best}) و بهترین مکان ملاقات شده توسط همه‌ی ذرات (G_{best}) تاکنون است. موقعیت (X) و سرعت (V) هر ذره در هر تکرار توسط رابطه (۸) و (۹) بدست می‌آید (Eberhart and Kennedy, 1995).

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (\text{رابطه } 8)$$

$$v_{ij}(t+1) = w \times v_{ij}(t) +$$

$$c_1 r_{1j}(t) \times (p_{ij,best}(t) - x_{ij}(t)) + c_2 r_{2j}(t) \times (p_{j,gbest}(t) - x_{ij}(t))$$

معادله فوق حالت زبُعدی را نشان می‌دهد. هر بُعد نشان‌دهنده یک خروجی بهینه مورد انتظار از الگوریتم می‌باشد که در آن، $x_i(t)$: موقعیت ذره i ام در زمان t ام، w : وزن اینرسی، c_1 و c_2 به ترتیب بهترین شاخص یادگیری ذره و بهترین شاخص یادگیری کل یا گروه می‌باشد که به صورت پیش فرض برای این شاخص‌ها مقدار $[2-1]$ اختیار می‌شود. r_1 و r_2 ضرایبی هستند که به طور تصادفی بین صفر و یک در نظر گرفته می‌شوند.

در پژوهش حاضر به منظور ترکیب شبکه عصبی با الگوریتم ازدحام ذرات، ابتدا ساختار شبکه تعیین می‌شود سپس وزن یال‌ها و بایاس‌های شبکه به عنوان ذرات اولیه وارد الگوریتم ازدحام ذرات شده و برازندگی آن‌ها که همان کمینه خطای مدل‌سازی است (اختلاف بین داده‌های واقعی و داده‌های شبکه) بررسی می‌شود. در ادامه اگر شرط توقف که تعداد ذرات است برقرار نباشد موقعیت و سرعت ذرات به ترتیب توسط روابط (۸) و (۹) به‌روز می‌شوند. اگر موقعیت جدید ذرات بهتر از تجربه شخصی آن‌ها باشد جایگزین می‌شوند. همین روال برای تجربه گروه ذرات نیز صادق است. سپس برازندگی آن‌ها بررسی شده و جمعیت جدید وارد چرخه تکرار الگوریتم شده تا شرط توقف برقرار گردد. سرانجام الگوریتم ازدحام ذرات متوقف شده و پارامترهای بهینه شده به شبکه اعمال می‌شود و شبکه آموزش می‌بیند.

رویکرد پیشنهادی

در این تحقیق علاوه بر روش‌هایی که پیش از این گفته شد، از یک رویکرد نوین به منظور بهینه‌سازی فرآیند آموزش شبکه عصبی مصنوعی استفاده شده است. این تکنیک به گونه‌ای است که در آن کلیه‌ی عملگرهای هر دو الگوریتم ژنتیک و ازدحام ذرات، هم‌زمان بر روی جمعیت اولیه اعمال می‌شوند. در شکل (۳) فلوچارت اجرای الگوریتم ترکیبی GA-PSO ارائه شده در این تحقیق نمایش داده شده است.

همانطور که در شکل (۳) مشاهده می‌شود این الگوریتم،

اعتبارسنجی مدل

پس از مقایسه و ارزیابی مدل‌ها با استفاده از آماره‌های میانگین مربعات خطا و ضریب همبستگی، در انتهای پژوهش برای بررسی صحت نتایج پیش‌بینی مدل پیشنهادی از روش اعتبارسنجی متقاطع k لایه (K-fold Cross Validation) استفاده می‌شود. در این روش داده‌ها به k زیر مجموعه مجزا تقسیم شده و k بار آموزش و ارزیابی بر روی آنها انجام می‌شود. به این صورت که هر بار یکی از زیر مجموعه‌ها برای ارزیابی مدل نگه داشته شده و $k-1$ زیر مجموعه دیگر، برای آموزش مدل استفاده می‌شود. این فرآیند k بار تکرار شده به طوری که هر یک از زیر مجموعه‌ها دقیقاً یک بار برای آموزش و یک بار برای ارزیابی به کار می‌روند. در نهایت، نتیجه این k تکرار برای دستیابی به یک برآورد نهایی میانگین‌گیری می‌شود (Rodriguez et al., 2010). به طور معمول روش اعتبارسنجی پنج لایه یا ده لایه برای مسائل مدل‌سازی و پیش‌بینی پیشنهاد می‌شود. در پژوهش حاضر از روش اعتبارسنجی متقاطع k نزدیک‌ترین همسایه‌ها (K-Nearest Neighbors) با تعداد ۱۰ لایه (تعداد همسایه‌ها) برای برآورد سه شاخص صحت، حساسیت و ویژگی مدل پیشنهادی استفاده شد. در روش ارزیابی نزدیک‌ترین همسایه حداقل فاصله اقلیدسی نمونه مورد بررسی تا نمونه‌های موجود دیگر برای تعیین k نزدیک‌ترین همسایه‌ها مورد استفاده قرار می‌گیرد و نمونه را متعلق به دسته‌ای می‌داند که بیشترین آرا را در بین k نزدیک‌ترین همسایه داشته باشد (Thanuja et al., 2011).

نتایج و بحث

یکی از ویژگی‌های شبکه‌های عصبی عدم وجود قانون تأیید شده برای معماری شبکه‌ای آن‌هاست؛ معیار مناسبی برای تعیین تعداد لایه و تعداد نورون در لایه میانی، نوع تابع فعالیت برای لایه میانی وجود نداشته و تنها راه‌حل استفاده از روش آزمون و خطا است.

به طور کلی، در طراحی شبکه‌های عصبی، دو پارامتر بحرانی وجود دارد. این دو پارامتر به ترتیب اهمیت، تعداد گره‌های ورودی و پس از آن تعداد گره‌ها یا نورون‌های لایه پنهان معرفی شده‌اند. نورون‌های پنهان به شبکه عصبی اجازه می‌دهند الگوهای غیرخطی درون داده‌ها را کشف نماید. به لحاظ تئوری محدودیتی در تعداد نورون‌های درون لایه پنهان وجود ندارد؛ اما به ندرت دیده می‌شود که این تعداد از دو برابر تعداد گره‌های ورودی فراتر رود. برخی از محققان برای تعداد نورون‌های پنهان، از رابطه $2n+1$ ، برخی از $2n$ ، برخی از n و برخی دیگر از $n/2$ استفاده نموده‌اند که n تعداد گره‌های ورودی است. در پژوهش حاضر برای تعیین حداقل و حداکثر تعداد بهینه نورون در لایه پنهان (x) از رابطه $n/2+1 < x < 2n+1$ استفاده شده است. در این پژوهش با توجه به ۶ پارامتر ورودی، حداقل و حداکثر تعداد نورون در لایه پنهان به ترتیب ۴ و ۱۳ نورون مورد استفاده قرار گرفته است. همچنین برای لایه خروجی نیز یک نورون وابسته به طبیعت مسأله مورد نظر که پیش‌بینی عملکرد نیشکر می‌باشد، لحاظ شده است. پس از آزمون پارامترهای مختلف شبکه عصبی، حال بهترین نوع شبکه برای پیش‌بینی عملکرد نیشکر که کمترین میزان خطا و بیشترین ضریب همبستگی را داشته است، شبکه عصبی پرسپترون دو لایه با توابع فعالیت بکار رفته از نوع تانژانت سیگموئید در لایه‌های پنهان و خروجی انتخاب شد. مطلوب‌ترین چیدمان شبکه به صورت ۱-۹-۶ (شش نورون به ازای متغیرهای ورودی، نه نورون در لایه پنهان و یک نورون در لایه خروجی) بدست آمد. برای آموزش شبکه‌ها در این الگو از الگوریتم کلاسیک لوبنبرگ مارکوآرت (Levenberg-Marquardt) به دلیل سرعت بالا در مقایسه با سایر الگوریتم‌های آموزش پس انتشار خطا (BP) استفاده گردید. نتایج حاصل از سعی و خطای شبکه عصبی مصنوعی با ساختارهای مختلف در جدول (۲) ارائه شده است.

جدول ۲- پارامترهای اصلی و مقادیر شاخص‌های ارزیابی شبکه عصبی مصنوعی مبتنی بر الگوریتم پس انتشار

نوع الگو	تابع فعالیت لایه پنهان	تابع فعالیت لایه خروجی	حداقل تعداد بهینه نورون در لایه پنهان	حداکثر تعداد بهینه نورون در لایه پنهان	MSE		R	
					اعتبار	آموزش	اعتبار	آموزش
LM-ANN	Logsig	Purelin	۴	۱۳	۰/۰۱۰۴۵	۰/۰۱۰۱۴	۰/۰۱۱۲۷	۰/۶۴۳۳۸
	Tansig	Tansig	۴	۱۳	۰/۰۰۴۲۹	۰/۰۰۴۷۶	۰/۰۰۴۳۹	۰/۶۵۵۹۴
	Tansig	Logsig	۴	۱۳	۰/۰۱۰۲۵	۰/۰۱۰۰۴	۰/۰۱۰۹۰	۰/۴۱۵۵۴
	Tansig	Purelin	۴	۱۳	۰/۰۲۷۱۵	۰/۰۲۵۶۹	۰/۰۲۵۶۲	۰/۶۱۸۷۶
	Logsig	Logsig	۴	۱۳	۰/۰۱۴۰۰	۰/۰۱۳۵۵	۰/۰۱۴۲۹	۰/۳۳۴۱۱
	Logsig	Tansig	۴	۱۳	۰/۰۱۲۳۶	۰/۰۱۳۰۹	۰/۰۱۲۹۸	۰/۶۲۶۳۴

روش‌های بهینه‌سازی که دستیابی به وزن یال‌ها و بایاس‌های مطلوب شبکه عصبی را میسر می‌سازند، استفاده شده است.

در اجرای الگوهای شبکه عصبی بهینه از الگوریتم‌های فرااکتشافی ژنتیک، ازدحام ذرات و ترکیبی از این دو به عنوان

ترکیبی شبکه عصبی، الگوریتم ژنتیک و ازدحام ذرات (GA- PSO-ANN) و سپس به ترتیب الگوی ترکیب شبکه عصبی با الگوریتم ازدحام ذرات (GA-ANN) و الگوی ترکیب شبکه عصبی با الگوریتم ژنتیک (GA-ANN) در پیش‌بینی عملکرد نیشکر موفق‌تر عمل کرده است. ضریب همبستگی الگوی ترکیبی شبکه عصبی با الگوریتم GA-PSO در مرحله آموزش شبکه برابر با ۰/۹۲۵۳۲ بدست آمد. نتایج جدول (۵) نشان می‌دهد که اثربخشی استفاده از الگوریتم‌های فرااکتشافی در یادگیری و تصحیح خطاهای تخمین عملکرد محصول نیشکر نسبت به شبکه عصبی مبتنی بر الگوریتم پس انتشار خطا بالاتر بوده است. همچنین از نتایج جدول (۵) می‌توان دریافت که بین دو الگوریتم فرااکتشافی مورد مطالعه، عملکرد الگوریتم ازدحام ذرات به منظور کاهش خطای شبکه در قیاس با الگوریتم ژنتیک بهتر بوده است. صحت این نتیجه با نتایج پژوهشی (Keynia and Heydari (2014 نیز هم‌خوانی دارد؛ آنها به منظور پیش‌بینی توان بادی از ترکیب شبکه عصبی المان و دو الگوریتم فرااکتشافی ازدحام ذرات و ژنتیک استفاده کردند. نتایج نشان داد که استفاده از این الگوریتم‌ها به ویژه الگوریتم ازدحام ذرات در افزایش عملکرد مدل مؤثر بوده و به قدری کامل و کارآمد می‌باشد که می‌توان از آن به عنوان یک روش هوشمند در بسیاری از مسائل استفاده کرد.

بهینه‌سازی شبکه عصبی با بهترین توپولوژی ۱-۹-۶ توسط این الگوریتم‌ها مستلزم تعیین پارامترهای اصلی الگوریتم ژنتیک و الگوریتم ازدحام ذرات است (جدول ۳ و ۴).

جدول ۳- پارامترهای اصلی الگوریتم ژنتیک

جمعیت اولیه	احتمال تقاطع	احتمال جهش	تعداد نسل
۵۰	۰/۷	۰/۲	۱۰۰۰

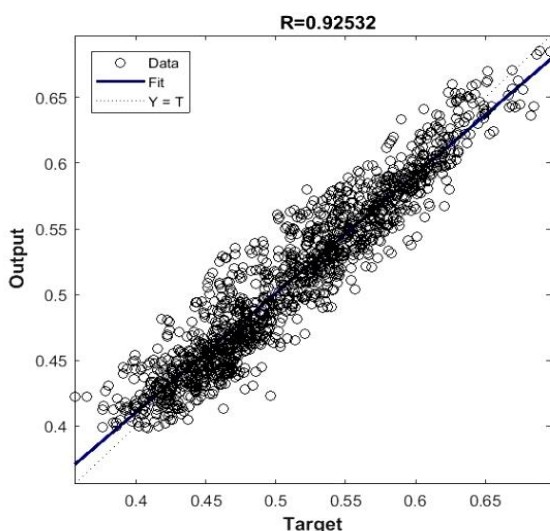
جدول ۴- پارامترهای اصلی الگوریتم ازدحام ذرات

وزن اینرسی	تعداد ذرات	شاخص یادگیری ذره	شاخص یادگیری کل	تعداد تکرار
۰/۷۳	۵۰	۲/۰۵	۲/۰۵	۱۰۰۰

پس از تنظیم پارامترهای الگوریتم فرااکتشافی، وزن یال‌ها و بایاس‌های شبکه عصبی محاسبه می‌گردد و بدین طریق شبکه عصبی برای پیش‌بینی آماده می‌شود. در نهایت، پس از سه مرتبه اجرای الگوریتم‌ها، مقادیر محاسبه شده‌ی معیار میانگین مربعات خطا (MSE) و ضریب همبستگی (R) برای سه الگوی آموزشی شبکه عصبی بهینه شده با الگوریتم‌های فرااکتشافی (GA, PSO, GA-PSO) در جدول (۵) ارائه شده است. همانطور که از نتایج جدول (۵) مشخص است، ابتدا الگوی

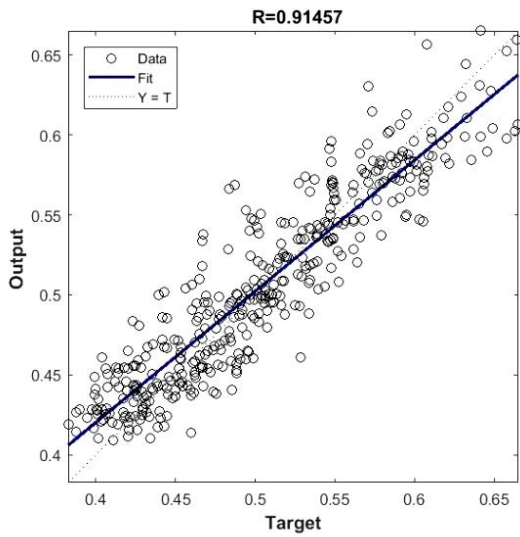
جدول ۵- مقادیر شاخص‌های ارزیابی شبکه عصبی بهینه‌سازی شده با الگوریتم‌های فرااکتشافی

نوع الگو	MSE		R	
	آموزش	اعتبار	آموزش	اعتبار
GA-ANN	۰/۰۰۲۰۸	۰/۰۰۵۳۴	۰/۰۰۵۷۲	۰/۶۵۴۷۱
PSO-ANN	۰/۰۰۱۵۱	۰/۰۰۴۰۱	۰/۰۰۵۰۶	۰/۷۸۰۴۲
GA-PSO-ANN	۰/۰۰۰۴۵	۰/۰۰۰۶۷	۰/۹۲۵۳۲	۰/۸۹۸۶۷

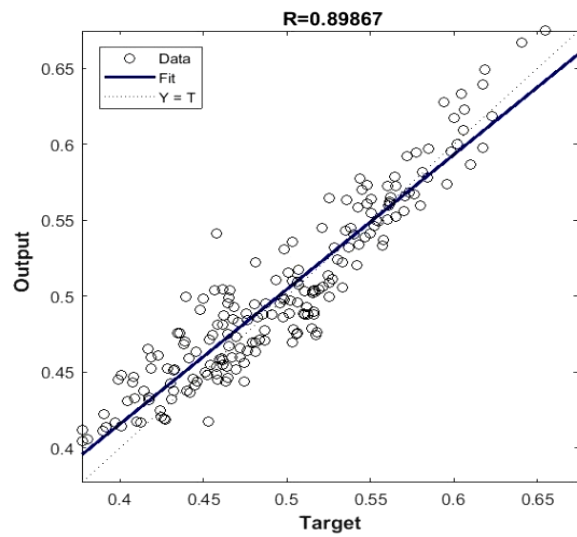


شکل ۴- ضریب همبستگی مرحله آموزش شبکه GA-PSO-ANN

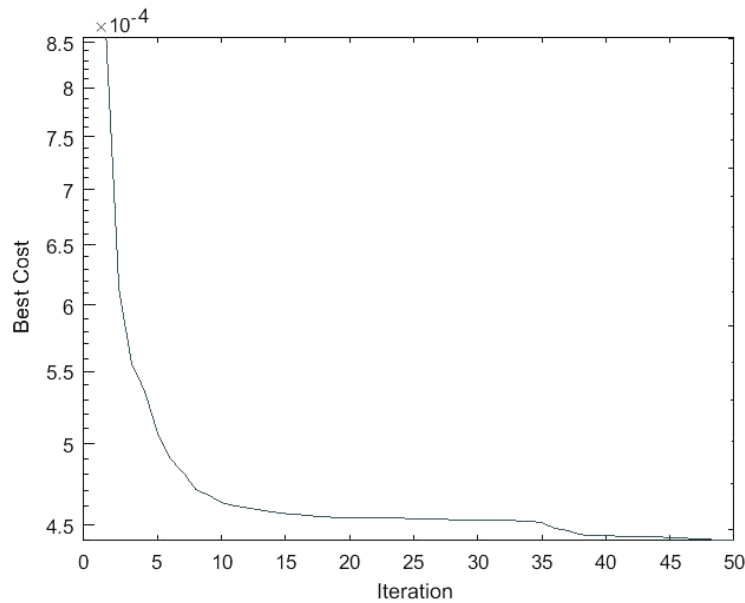
در شکل‌های (۴-۶) ضریب همبستگی به ترتیب برای مراحل آموزش، اعتبارسنجی و آزمایش شبکه عصبی بهینه شده با الگوریتم ترکیبی GA-PSO ارائه شده است. هر چقدر خط ضریب همبستگی بر خط نقطه‌چین (Y=T) منطبق‌تر باشد (نشانهگر برابری مقدار پیش‌بینی شده توسط شبکه عصبی با مقدار تابع هدف) عملکرد شبکه عصبی بهتر خواهد بود. در شکل (۷) نمودار مربوط به همگرایی تابع هزینه در مقابل ماکزیمم تعداد تکرار الگوریتم نمایش داده شده است. این نمودار در ۵۰ تکرار از الگوریتم GA-PSO برای آموزش شبکه عصبی پرسپترون با توپولوژی منتخب، به خوبی سیر نزولی کاهش خطای شبکه را تا رسیدن به میزان بهینه نشان می‌دهد. همانطور که از شکل پیداست، میزان تابع هزینه ۰/۰۰۰۴۵ بدست آمد و از تکرار ۴۵ به بعد تغییری در مقدار هزینه حاصل نشده است.



شکل ۶- ضریب همبستگی مرحله آزمایش شبکه GA-PSO-ANN



شکل ۵- ضریب همبستگی مرحله تعیین اعتبار شبکه GA-PSO-ANN



شکل ۷- همگرایی تابع هزینه براساس تعداد تکرار در مرحله آموزش شبکه GA-PSO-ANN

برای اثبات قابلیت مدل پیشنهادی در پیش‌بینی عملکرد محصول نیشکر در مقایسه با سایر مدل‌های گفته شده، فرآیند اعتبارسنجی متقاطع با ۱۰ لایه بر روی مجموعه داده‌های آموزشی، آزمایشی و اعتبارسنجی مدل شبکه GA-PSO اعمال شد؛ به طوری که هر قسمت دارای ترکیبی از تمام متغیرهای ورودی مؤثر در پیش‌بینی عملکرد نیشکر می‌باشد. نتایج روش k نزدیک‌ترین همسایه برای ارزیابی شاخص‌های صحت، حساسیت و ویژگی مدل پیشنهادی در جدول (۷) ارائه شده است. شاخص صحت بیانگر این نکته است که مدل به درستی مقدار پیش‌بینی را هم ارز مقدار واقعی آن تخمین زده است؛ یا به عبارتی دیگر، مقدار اندازه‌گیری شده چقدر به مقدار واقعی نزدیک است. شاخص

با توجه به چیدمان مطلوب شبکه عصبی انتخاب شده برای پیش‌بینی عملکرد نیشکر که به صورت ۱-۹-۶ می‌باشد؛ در جدول (۶) می‌توان مقادیر ضرایب وزنی یال‌ها و بایاس‌های مدل پیشنهادی (GA-PSO-ANN) را مشاهده نمود. به این صورت که ماتریس وزن یال‌های شبکه برای بردار ورودی به لایه پنهان یک ماتریس ۹×۶ هسین (اتصال ۶ نورون ورودی به ۹ نورون لایه پنهان) و برای لایه پنهان به لایه خروجی به صورت یک ماتریس ۱×۹ هسین (اتصال ۹ نورون لایه پنهان به ۱ نورون لایه خروجی) بوده است. همچنین ماتریس بایاس شبکه برای بردار ورودی به لایه پنهان به شکل یک ماتریس ۹×۱ و برای لایه پنهان به لایه خروجی در قالب یک ماتریس ۱×۱ می‌باشد.

ثبات رفتار متغیرهای توصیف‌کننده مدل در طول زمان است. نتایج جدول (۷) نشان می‌دهد که مدل شبکه عصبی بهبود یافته با الگوریتم ترکیبی GA-PSO در تخمین عملکرد نیشکر دقیق‌تر و حساس‌تر بوده است.

حساسیت بیانگر ارزیابی خوش‌بینانه از یک مدل پیش‌بینی است که تخمین بالاتری از مقدار واقعی ارائه می‌کند؛ اگر دو مدل میزان دقتشان یکسان باشد مدلی انتخاب می‌گردد که شاخص حساسیت آن بیشتر باشد. اما، شاخص ویژگی ارزیابی واقع‌بینانه از انجام یک پیش‌بینی قابل اتکا را فراهم ساخته و نشان‌دهنده

جدول ۶- ضرایب وزنی یال‌ها و بایاس‌های مدل GA-PSO-ANN به ازای ۹ نورون موجود در لایه پنهان شبکه

N9	N8	N7	N6	N5	N4	N3	N2	N1	ضرایب وزنی یال‌ها
-۱/۰۴۶	۱/۸۳۴	۰/۱۵۱	-۰/۰۱۴	-۱/۷۵۰	-۰/۴۵۵	۰/۵۳۳	-۰/۰۱۴	۰/۰۶۴	وزن ورودی‌ها به لایه پنهان
-۰/۵۳۶	-۰/۹۷۶	۰/۸۱۱	-۰/۰۴۳	۱/۶۳۶	۲/۸۰۹	۲/۳۸۸	-۰/۰۶۶	۰/۱۱۱	
-۲/۰۱۰	-۲/۶۸۳	۴/۲۰۴	-۰/۱۶۰	-۱/۳۹۴	-۰/۰۵۹	۱/۱۱۷	۰/۱۰۵	۰/۵۲۹	
-۴/۴۹۴	-۲/۲۳۶	-۴/۱۱۰	۰/۳۵۳	۱/۱۹۰	۰/۷۵۳	-۱/۰۶۲	-۰/۷۹۲	-۰/۳۳۰	
-۱/۴۱۱	۴/۸۹۷	-۰/۳۲۶	۰/۰۹۵	-۲/۹۳۶	۰/۴۹۹	۲/۳۳۲	-۰/۳۸۹	-۰/۰۸۹	
۲/۷۱۴	-۰/۶۰۹	۱/۶۵۷	۰/۰۴۱	-۲/۰۴۳	۰/۷۵۴	-۰/۱۱۵	-۰/۱۱۴	۰/۰۶۱	
۰/۱۱۵	۰/۰۹۱	-۰/۳۶۱	-۳/۸۸۹	۰/۰۸۸	-۰/۱۴۱	-۰/۰۸۷	-۰/۹۳۷	۱/۵۴۶	وزن لایه پنهان به لایه خروجی
N9	N8	N7	N6	N5	N4	N3	N2	N1	مقادیر بایاس شبکه
-۰/۱۳۲	۰/۷۷۴	-۰/۶۴۳	-۰/۱۷۹	۰/۸۴۹	-۰/۸۴۳	۴/۶۵۸	-۱/۰۲۵	-۲/۷۳۴	بایاس ورودی به لایه پنهان
-۲/۳۹۸									بایاس لایه پنهان به لایه خروجی

تعیین اعتبار و امتحان شبکه ساخته شده، استفاده گردید. دو معیار میانگین مربعات خطا و ضریب همبستگی به عنوان توابع عملکرد برای هر چهار مدل انتخاب شدند. عملکرد الگوهای مدل‌سازی بر روی داده‌های آموزش نشان داد که کمترین میزان خطا به الگوی ترکیبی شبکه عصبی، الگوریتم ژنتیک و ازدحام ذرات اختصاص داشته است. این الگو قادر است ۸۵/۶۲ درصد از تغییرات پارامترهای مؤثر در مدل‌سازی را توسط متغیر عملکرد نیشکر توضیح دهد.

سپاسگزاری

نویسندگان این مقاله از واحد کشت و صنعت نیشکر دعبل خزاعی که اطلاعات موردنیاز را در اختیار قرار داده‌اند کمال تشکر و امتنان را دارند.

هیچگونه تعارض منافع بین نویسندگان وجود ندارد.

REFERENCES

- Ahmadvand, M. (2009). Modeling the effect of water table fluctuations on sugarcane yield using artificial neural network and fuzzy logic (Case Study: Mirza Koochak-Khan Agro-Industry Co.). Master of Science Thesis. Shahid Chamran University of Ahvaz. (In Farsi).
- AmirEntezari, K. (2008). Review of new methods for training neural networks using artificial

جدول ۷- نتایج اعتبارسنجی متقاطع به روش KNN=10 بر روی مجموعه داده‌های آموزشی، آزمایشی و اعتبارسنجی مدل GA-PSO-ANN

ویژگی	حساسیت	صحت	داده‌های آموزشی
۰/۹۲۳	۰/۹۹۶	۰/۹۸۵	داده‌های آموزشی
۰/۷۵۶	۱/۰۰۰	۰/۹۵۵	داده‌های آزمایشی
۰/۹۰۹	۰/۹۸۷	۰/۹۶۴	داده‌های اعتبارسنجی

نتیجه‌گیری

در این مقاله پیش‌بینی عملکرد محصول نیشکر با استفاده از داده‌های جمع‌آوری شده از کشت و صنعت دعبل خزاعی مورد توجه قرار گرفته است. مدل تخمین‌گر برای عملکرد نیشکر با استفاده از شبکه عصبی مصنوعی و الگوریتم‌های فرااکتشافی نظیر ژنتیک و ازدحام ذرات تعیین شد. داده‌های آموزش برای ساخت مدل هوش مصنوعی و تنظیم وزن یال‌ها و بایاس‌های شبکه عصبی استفاده شدند و داده‌های اعتبارسنجی و آزمایش نیز برای

intelligence algorithms. Master of Science Thesis. University of Tabriz. (In Farsi).

- Bagheri, A. and Sohrabi, N. (2018). Predicting yield of rainfed and irrigated barley (*Hordeum vulgare* L.) in Kermanshah by Artificial Neural Network approach (Case study Kermanshah, Iran). *Journal of Agroecology*. 10(2): 516-528. (In Farsi).
- Eberhart, R. and Kennedy, J. (1995). A new optimizer

- using particle swarm theory. Proceedings of the Sixth International Symposium on Micro Machine and Human Science. Nagoya, Japan, pp: 39-43.
- Esfandiarpour-Boroujeni, I., Karimi, E., Shirani, H., Esmaeilzadeh, M. and Mosleh, Z. (2019). Yield prediction of apricot using a hybrid particle swarm optimization- imperialist competitive algorithm-support vector regression (PSO-ICA-SVR) method. *Scientia Horticulturae*, 257, 1-12.
- Golabi, M., Karami, B. and Albaji, M. (2013). Sensitivity analysis of sugarcane yield using artificial neural networks. 4th National Conference on Irrigation and Drainage Network Management. Faculty of Water Sciences Engineering. Shahid Chamran University of Ahvaz, pp: 1917-1924. (In Farsi).
- Heydarnejadi, S.Z. (2016). Effects of climate change on net irrigation requirement and yield of sugarcane in south of Ahvaz (Case study: Amir Kabir Agro-Industry Co.). Master of Science Thesis. Shahid Chamran University of Ahvaz. (In Farsi).
- Hosseini, M.T., Siosemarde, A., Fathi, P. and Siosemarde, M. (2007). Application of artificial neural network and multiple regression for estimating dry farming wheat yield in Ghorveh region, Kurdistan province. *Agricultural research: water, soil and plant in agriculture*, 7(1), 41-54. (In Farsi).
- Hosseini, M., Movahedi-Naeini, S.A., Dehghani, A.A. and Khaledian, Y. (2016). Estimation of soil mechanical resistance parameter by using particle swarm optimization, genetic algorithm and multiple regression methods. *Soil and Tillage Research*, 157, 32-42.
- Jeong, S., Hasegawa, S., Shimoyama, K. and Obayashi, S. (2009). Development and investigation of efficient GA/PSO hybrid algorithm applicable to real-world design optimization, *IEEE Computational Intelligence Magazine*, pp. 36-44.
- Kia, S.M. (2011). *Neural Networks in MATLAB*. Kian Publication. Tehran, 232 pages. (In Farsi).
- Keynia, F. and Heydari, A. (2014). The combination of Particle Swarm Optimization algorithm and artificial neural network to forecast wind power. 4th Annual Clean Energy Conference (ACEC2014). Kerman. (in Farsi).
- Kumar, M., Raghuwanshi, N.S., Wallender, W.W. and Pruitt, W.O. (2002). Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128(4), 224-233.
- Menhaj, M.B. (2008). *Fundamentals of neural networks (computational intelligence)*. Volume 1, 8th Edition. Publishing Center, Amirkabir University of Technology (Tehran Polytechnic). 715 pages. (in Farsi).
- Pourmohammadali, B., Hosseini, S.J., Salehi, M.H., Shirani, H. and Esfandiarpour Boroujeni, I. (2019). Effects of soil properties, water quality and management practices on pistachio yield in Rafsanjan region, southeast of Iran. *Agricultural Water Management*, 213, 894-902.
- Rezaei, A.R. and Ranjbaran, S. (2009). *Functional training of genetic algorithm in MATLAB software* (2nd ed.), Farhang-e-Matin Publication. Tehran. 144 pages. (In Farsi)
- Rodriguez, J. D., A. Perez, and J. A. Lozano. (2010). A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceeding of the international joint conference on artificial intelligence, 32: 569-575.
- Settles, M. and Soule, T. (2005). Breeding swarms: A GA/PSO hybrid. the Genetic and Evolutionary Computation Conference (GECCO-2005): 161-168.
- Thanuja, V., Venkateswarlu, B. and Anjaneyulu, G. S. G. N. (2011). Applications of Data Mining in Customer Relationship Management. *Journal of Computer and Mathematical Sciences*, 2(3): 423-433.
- Yang, X.S. (2008). *Nature Inspired Metaheuristic Algorithm*. 2nd Edition, Luniver Press, UK. 128 pages.