

تخمین قیمت نهایی در حراج های آنلاین مبتنی بر شبکه های عصبی

حمیدرضا یونس پور^{*}، مهدی براتی، مریم بیاتی

hr.yunespour@gmail.com

^{*}نویسنده مسئول، دریافت: ۱۴۰۱/۹/۵، بازنگری: ۱۴۰۱/۹/۲۴، پذیرش: ۱۴۰۱/۱۰/۵

کارشناسی ارشد مهندسی فناوری اطلاعات، گرایش تجارت الکترونیک، موسسه عالی علوم و فناوری سپاهان، اصفهان، ایران
استادیار، دانشکده مهندسی کامپیوتر، موسسه آموزش عالی راغب، اصفهان، ایران
استادیار، دانشکده فنی-مهندسی، موسسه آموزش عالی علوم و فناوری سپاهان، اصفهان، ایران

چکیده

در دهه گذشته اهمیت تجارت الکترونیک نسبت به قبل بدلیل تغییرات ایجاد شده در عادات های تجاری جامعه افزایش یافته است. رشد بی رویه اقتصاد و تکنولوژی، منجر به تولید داده های حجیم و در عین حال ارزان شده که به راحتی قابل دسترس هستند. در این بین موازی با پیشرفت های تجارت الکترونیک، اهمیت پیش بینی قیمت در حراج های آنلاین نیز افزایش یافته است. اگرچه پیش بینی قیمت نهایی در حراج های آنلاین مزایای بسیاری دارد، اما تاکنون روشی که بتواند در این خصوص عملکرد موفق نشان دهد ارائه نشده است. روش های ارائه شده عمدتاً مبتنی بر رگرسیون چندگانه هستند که نتایج امیدوار کننده ای ندارند. با توجه به اهمیت تخمین قیمت نهایی در حراج های آنلاین، در این تحقیق از شبکه های عصبی عمیق بازگشتی با واحد پنهان BiLSTM و شبکه پیشخور رگرسیون و درخت تصمیم و ترکیب روش های فوق استفاده شده است. در این روش برای کاهش جذر میانگین مربع خطا در پیش بینی قیمت نهایی از ترکیب هر یک از روش های فوق و تکنیک انتخاب ویژگی مبتنی بر الگوریتم فراابتکاری کرم شب تاب استفاده شده است. نتایج تجربی در دو پایگاه داده حراج آنلاین ترکیبی eBay نشان می دهد که استفاده از شبکه های عصبی و ترکیب روش های یادگیری ماشین سنتی و شبکه های عمیق در پیش بینی قیمت نهایی عملکرد موفق دارد.

۱- مقدمه

در دهه اخیر تجارت الکترونیک و خرید آنلاین به سرعت جایگزین تجارت معمولی شده است. عواملی همچون گسترش اینترنت، توسعه سیستم های پرداخت الکترونیک از اصلی ترین دلایل این پدیده محسوب می شوند [۱]. بازارهای آنلاین در چند سال گذشته به شدت مورد توجه و استقبال قرار گرفته اند. eBay Yahoo، بازار آمازون و سایر بازارهای آنلاین به نهادهای تجاری بسیار مهمی تبدیل شده اند و تخمین زده می شود که آن ها بالغ بر ۵۰ درصد از تجارت الکترونیک را تا سال ۲۰۲۲ کسب می کنند. حتی امروزه، eBay، یکی از بزرگترین بازارهای آنلاین است، که از ۱۵۰ میلیون عضو تشکیل شده که به طور معمول بیش از ۱۹ میلیون اقلام برای فروش در هر زمان عرضه می کند. در بین سال های ۲۰۰۰ تا ۲۰۰۵، ۲۴ میلیارد دلار کالا در eBay به فروش رسید [۲]. چنین بازارهای آنلاینی علاوه بر اینکه به افراد فرصت مناسبی برای خرید و فروش کالا می دهند، همچنین یک منبع جدید از داده ها را نیز تولید می کنند که می تواند برای تجزیه و تحلیل رفتار اقتصادی جمعی، بازاریابی محصول، تحقیقات بازار و ارائه خدمات به خریداران و فروشندگان در بازار آنلاین استفاده شود. عوامل بالا باعث شده سرمایه گذاری در شرکت های تجارت الکترونیک مطلوب باشد و ارزش بازار چنین شرکت هایی نیز افزایش یابد [۱]. تخمین زده می شود که ارزش چنین شرکت هایی تا سال ۲۰۲۲ به بیش از ۷۰۰ میلیارد دلار برسد و این بدان معناست که اهمیت تجارت

الکترونیک با شیب تندتری در حال افزایش است. افزایش اهمیت تجارت الکترونیک در سال های اخیر و همچنین افزایش اهمیت تخمین قیمت نهایی در حراج های آنلاین، انگیزه اصلی ارائه یک رویکرد تخمین و برآورد قیمت نهایی در حراج های آنلاین در این تحقیق است. در این تحقیق از تکنیک های هوش مصنوعی و همچنین روش های انتخاب ویژگی مبتنی بر الگوریتم های فرااکتشافی برای رسیدن به بهترین نتایج استفاده شده است.

۲- بیان دقیق مسئله تحقیق

در حراج های آنلاین، پیشنهاد دهندگان بارها پیشنهاد می دهند و قیمت پیشنهادی خود را افزایش می دهند. همه مناقصه گران می خواهند سود خود را در مزایده به حداکثر برسانند. یک مکانیسم تخمین و برآورد موثر قیمت نهایی می تواند رفاه اجتماعی را در حراج های آنلاین به حداکثر برساند [۳]. اخیراً محققان دریافته اند که استفاده از تکنیک های هوش مصنوعی و یادگیری ماشین در این زمینه می تواند عملکرد موفق داشته باشد. از جمله آخرین مطالعاتی که در خصوص پیش بینی قیمت نهایی در حراج آنلاین ارائه شده است، روش لی و همکاران^۱ در سال ۲۰۲۰ است [۴]. در این روش لی و همکاران همبستگی^۲ بین معیارهای فروش آنلاین و قیمت نهایی را محاسبه کرده اند و سپس با استفاده از مقادیر حاصل شده قیمت نهایی در حراج آنلاین در پایگاه داده eBay را با استفاده از رگرسیون چندگانه^۳ و فیلتر

^۳Multiple Regression

^۱Li et al.

^۲Correlation

- استفاده از تکنیک انتخاب ویژگی بسته‌بندی مبتنی بر خطای درخت تصمیم رگرسیون برای رسیدن به زیر مجموعه بهینه از ویژگی‌های موثر بر قیمت نهایی. در این بخش از الگوریتم فرااکتشافی کرم شب تاب برای انتخاب ویژگی‌ها استفاده شود.
- ترکیب نتایج مدل‌های فوق و بررسی نتایج تخمین قیمت نهایی در حالت گروهی. مطالعات نشان می‌دهد در تخمین قیمت نهایی در حراج‌های ترکیبی آنلاین تاکنون از روش‌های فوق استفاده نشده است.

۵-۱- اهداف تحقیق

اهداف تحقیق به شرح زیر است:

۱-۵-۱- هدف اصلی

- ارائه یک مدل بهینه تخمین قیمت نهایی در حراج‌های آنلاین با استفاده از شبکه‌های عصبی اعم از ساده و عمیق.
- ۱-۵-۲- اهداف فرعی
 - کاهش مقدار خطای جذر میانگین مربعات (RMSE) ۱۱ در پیش‌بینی قیمت نهایی حراج آنلاین در مقایسه با مدل پیشین. کاهش مقدار میانگین مطلق خطا (MAE) ۱۲ در پیش‌بینی قیمت نهایی حراج آنلاین در مقایسه با مدل پیشین. کاهش مقدار میانگین مطلق درصد خطا (MAPE) ۱۳ در پیش‌بینی قیمت نهایی حراج آنلاین در مقایسه با مدل پیشین.

۱-۶- سوالات تحقیق

سوال‌های تحقیق به شرح زیر است:

۱-۶-۱- سوال اصلی

- آیا استفاده از شبکه‌های عصبی و روش انتخاب ویژگی بسته‌بندی می‌تواند در تخمین قیمت نهایی در حراج آنلاین موفق عمل نماید؟
- ۱-۶-۲- سوالات فرعی
 - کدام یک از دو نوع معماری‌های پیشخور و بازگشتی در مقایسه با رگرسیون چندگانه می‌توانند موفق‌تر عمل نمایند؟
 - استفاده از مرحله انتخاب ویژگی مبتنی بر خطای تخمین، تا چه میزان می‌تواند باعث بهبود نتایج مدل شود؟
 - ۱-۶-۷- فرضیه‌های تحقیق

فرضیه‌های تحقیق به شرح زیر است:

 - استفاده از شبکه‌های عصبی با دو نوع معماری پیشخور و بازگشتی می‌تواند در مقایسه با روش‌های یادگیری ماشین بهترتر عمل نماید.
 - اگر از مرحله انتخاب ویژگی بسته‌بندی مبتنی بر خطای تخمین استفاده شود، انگاه خطای شبکه‌های عصبی کاهش بیشتری خواهد داشت.

۱-۸- متغیرهای تحقیق

متغیرهای وابسته و مستقل به شرح زیر است:

کالمن^۱ برآورد کرده‌اند. یکی از محدودیت‌های مدل ارائه شده توسط لی و همکارانش در عدم ارائه نتایج تقریباً مشابه در داده‌های مختلف است. به عبارت ساده در این مدل از پایگاهی به پایگاه دیگر پیش‌بینی قیمت نهایی با موفقیت همراه نیست. از طرف دیگر روش یادگیری ماشین^۲ رگرسیون اگرچه در داده‌هایی می‌تواند عملکرد قابل قبولی از خود نشان دهد اما نتایج آن در مقایسه با شبکه‌های عصبی^۳ و برخی دیگر از مدل‌های یادگیری ماشین مثل درخت تصمیم رگرسیون چندان امیدوارکننده نیست.

از اینرو با توجه به اهمیت تخمین قیمت نهایی در حراج‌های آنلاین در این تحقیق در جهت بهبود عملکرد روش لی و همکاران از چندین روش مختلف شامل شبکه‌های عصبی و شبکه‌های عمیق استفاده می‌شود. در این روش برای تخمین قیمت نهایی از شبکه‌های عصبی پیشخور رگرسیون و شبکه عصبی عمیق بازگشتی^۴ با واحد BiLSTM [۵] برای تخمین قیمت نهایی استفاده می‌گردد و به این صورت مقایسه‌ای نیز بین دو نوع شبکه فوق انجام می‌شود. هدف از انتخاب دو نوع شبکه در این تحقیق، یافتن بهترین معماری برای تخمین دقیق‌تر قیمت نهایی در حراج آنلاین است. در این روش برای بهبود روش و در راستای کاربردی شدن روش در داده‌های مختلف از تکنیک انتخاب ویژگی بسته‌بندی^۵ استفاده خواهد شد [۶]. این نوع از تکنیک‌های انتخاب ویژگی مبتنی بر خطای عملکرد الگوریتم هستند و می‌توانند زیر مجموعه بهینه از ویژگی‌ها را به خوبی ارائه دهند. استفاده از تکنیک انتخاب ویژگی مناسب علاوه بر اینکه پیچیدگی زمانی و محاسباتی مدل را کاهش می‌دهد منجر به بهبود عملکرد آن می‌شود. در این بخش از الگوریتم فرااکتشافی^۶ [۷] کرم شب تاب^۷ [۸] برای رسیدن به زیر مجموعه بهینه از ویژگی‌ها استفاده می‌شود [۹، ۱۰]. همچنین در تابع هزینه^۸ الگوریتم کرم شب‌تاب از خطای مبتنی بر الگوریتم درخت تصمیم رگرسیون^۹ بهره برده می‌شود تا به این صورت بهترین نتایج حاصل شود.

۱-۳- اهمیت و ضرورت تحقیق

ظهور حراج‌های آنلاین چالش‌های بسیاری با خود داشته است، اما در عین حال دانش حاصل از حراجی‌ها از بسیاری جهات، سودآور است؛ که باعث شده اخیراً فعالیت‌های زیادی جهت تجزیه و تحلیل داده‌های حاصل از حراجی‌ها طراحی و انجام شود. در بین فاکتورهای موثر در فروش، قیمت در مقایسه با دیگر فاکتورها از اهمیت بالایی برخوردار است. تخمین قیمت نهایی در حراج‌های آنلاین تأثیر بسزایی در موفقیت یک تجارت دارد. اکثر شرکت‌ها برای برآورد قیمت نهایی از روش‌های سنتی مثل آزمون و خطا استفاده می‌کنند. اما رشد روزانه حراج‌های آنلاین و داده‌های مربوط به آن‌ها عملاً استفاده از روش‌های سنتی را غیر ممکن کرده است [۱۱]. محققان اخیراً دریافته‌اند استفاده از روش‌های مبتنی بر تکنیک‌های هوش مصنوعی می‌تواند در این زمینه بسیار موفق عمل نماید. در این راستا با توجه به اهمیت برآورد قیمت نهایی در حراج‌های آنلاین و با توجه به تأثیر بالای این فاکتور در موفقیت یک تجارت، در این تحقیق یک روش پیش‌بینی قیمت نهایی آنلاین مبتنی بر هوش مصنوعی و شبکه‌های عصبی ارائه شده است.

۱-۴- جنبه نوآوری تحقیق

- در این تحقیق، یک روش تخمین قیمت نهایی در حراج آنلاین ارائه می‌شود که مبتنی بر شبکه‌های عصبی است. در این روش برخلاف اکثر روش‌های موجود از شبکه عصبی عمیق بازگشتی BiLSTM و شبکه‌های پیشخور^{۱۰} رگرسیون استفاده می‌شود. مطالعات مختلف در این زمینه نشان می‌دهد روش‌های مطرح غالباً مبتنی بر یکی از تکنیک‌های رگرسیون، رگرسیون فازی [۱۲]، رگرسیون چندگانه، رگرسیون نیمه خودکار [۱۳] و ترکیب آن با مدل‌های دیگری مثل کالمن فیلتر [۴] هستند و تاکنون روشی که از معماری‌های عمیق بازگشتی و شبکه‌های عصبی پیشخور برای پیش‌بینی قیمت نهایی حراج آنلاین استفاده کند ارائه نشده است. در نتیجه نوآوری این تحقیق در موارد زیر خلاصه می‌شود:
- استفاده از دو نوع شبکه عصبی بازگشتی و پیشخور و مقایسه عملکرد آن‌ها با هم.

^۲Firefly

^۳Cost Function

^۴Decision tree for regression

^۵Feedforward

^۶Root Mean Square Error

^۷Mean Absolute Error

^۸Mean Absolute Percentage Error

^۹Kalman Filter

^{۱۰}Machine Learning

^{۱۱}Neural Networks

^{۱۲}Recurrent Neural Networks

^{۱۳}W rapper

^{۱۴}Metaheuristic

الف) مستقل:

- ۳) پارامترهای شبکه‌های عصبی: در شبکه عصبی بازگشتی BiLSTM و شبکه پیشخور تعداد واحد پنهان و نوع تابع یادگیری، از جمله اصلی‌ترین ورودی‌ها هستند؛ که با آزمون مقدار دهی می‌شوند.

• خروجی‌ها:

خروجی‌های روش شامل، میانگین جذر مربعات خطا (RMSE)، میانگین مطلق درصد خطا (MAPE) و میانگین مطلق خطا (MAE) است. در فصل اول روابط این متغیرها بیان شد.

EbayID	Quantity	Price	PricePerC	StartingBi	SellerName	SellerClos	Category	PersonID	StartingBi	AvgPrice	EndDay
1.61E+11	1	27.25	1.4787	0.0537	petesandi	0.925926	73396	9174	0.99	18.4273	Tuesday
3.91E+11	1	43	1.4422	0.335	graphn4fu	0.990566	27278	11046	9.99	29.815	Sunday
2.01E+11	1	11	0.5969	0.5426	hondo194	0.422131	73396	9174	10	18.4273	Monday
2.71E+11	0	10	0.5426	0.5426	winemey	0.245098	73396	9174	10	18.4273	Monday
3.11E+11	0	9.99	1.9159	1.9159	chucksme	0.295359	73396	17956	9.99	5.2142	Monday
3.91E+11	1	33.01	0.8372	0.0251	firstclassa	0.664516	27276	22954	0.99	39.4247	Monday
2.61E+11	0	49.99	2.7128	2.7128	sure2mkst	0.044444	73396	9174	49.99	18.4273	Thursday
3.01E+11	0	1.99	0.8644	0.8644	wabtradir	0.267606	73406	30356	1.99	2.302	Tuesday
1.31E+11	0	9.99	1.998	1.998	ledgauges	0.038315	27278	31149	9.99	5	Tuesday
1.21E+11	1	180	3.483	3.483	emmyjoic	0.333333	27278	20913	180	51.6793	Thursday
2.21E+11	0	7.99	0.4846	0.4846	ambiethe	0.242105	27278	50319	7.99	16.485	Thursday
2.71E+11	1	63.99	1.2409	0.0191	cantonanc	0.905846	53	57704	0.99	51.5669	Friday
2.31E+11	0	49.99	2.7128	2.7128	obx1971	0.2	73396	9174	49.99	18.4273	Saturday
3.01E+11	0	17.99	0.5527	0.5527	espnnot0	0.157895	27276	68252	17.99	32.5466	Friday
3.81E+11	1	1.25	0.7589	0.601	the_sport	0.553381	73396	14046	0.99	1.647	Friday
2.61E+11	0	4.99	1.0588	1.0588	ethreedav	0.3	73396	8799	4.99	4.7126	Saturday
1.71E+11	1	38	0.4374	0.0113	itsmittytir	0.625	27275	249	0.99	86.8692	Saturday
1.81E+11	0	34.99	0.4027	0.4027	thingsofyf	0.041237	27275	249	34.99	86.8692	Sunday
3.21E+11	0	59	0.6791	0.6791	streuer42	0.5	27275	249	59	86.8692	Sunday

ب) وابسته

- خطای جذر میانگین مربعات^۲: این معیار از تفاوت بین مقدار قیمت تخمین زده شده توسط مدل و قیمت واقعی حاصل می‌شود [۱۷].
- میانگین مطلق درصد خطا^۴: این معیار، دقت تخمین قیمت نهایی توسط مدل را نشان می‌دهد [۱۷].
- میانگین مطلق خطا^۵: این معیار نیز از میانگین خطای مطلق بین جفت مشاهدات مربوط به یک رویداد یکسان، حاصل می‌شود. در جدول (۱-۱) متغیرها و نحوه محاسبه آن‌ها و نوع آن‌ها بیان شده است [۱۷].
- ۲-۳- ضرورت ارائه راه کار پیشنهادی
- در بین فاکتورهای موثر بر موفقیت حراج‌های آنلاین، قیمت نهایی اقلام از اهمیت بالاتری برخوردار است. تخمین قیمت نهایی اقلام در حراج‌ها دارای ابعاد بسیار متنوعی است. اکثر سازمان‌های مربوطه برای تخمین قیمت نهایی متکی به آزمون و خطا هستند. اما افزایش روزانه حجم داده‌های مربوط به حراج‌های آنلاین باعث شده تخمین قیمت نهایی با استفاده از روش‌های سنتی براحتی امکان‌پذیر نباشد. از اینرو محققان به روش‌های مبتنی یادگیری ماشین برای تخمین قیمت نهایی در عصر حاضر روی آوردند. با توجه به اهمیت تخمین قیمت نهایی در حراج‌های آنلاین و با توجه به اهمیت این فاکتور در موفقیت حراجی‌ها، در این تحقیق یک روش مبتنی بر شبکه‌های عصبی جهت تخمین قیمت نهایی ارائه شده است.

۳-۳- معیارهای اساسی و نیازمندی‌های راه‌کار پیشنهادی

ورودی و خروجی‌های مدل به شرح زیر است:

• ورودی‌ها:

- ۱) پایگاه داده حراج‌های آنلاین ترکیبی eBay: اولین ورودی مدل، دو پایگاه داده مربوط به اطلاعات حراج‌های آنلاین است. این پایگاه داده‌ها به صورت دو بخش آزمایش و آموزش قابل دریافت هستند. بخشی از داده‌های پایگاه داده اول در شکل‌های (۳-۱) و (۳-۲) ارائه شده است.
- ۲) ورودی‌های انتخاب ویژگی: در روش ارائه شده، یک گام انتخاب ویژگی مبتنی بر روش‌های فراکتشافی تعریف شده است، که در آن زیر مجموعه‌ای از ویژگی‌ها با استفاده از الگوریتم کرم شب تاب انتخاب می‌شود. در این گام تعداد ویژگی منتخب بین ۱ تا ۹ تعیین می‌شود و بهترین تعداد انتخاب می‌گردد. تعداد ویژگی، بیانگر ابعاد هر عضو از جمعیت اولیه کرم شب تاب است.

شکل (۳-۱) بخشی از پایگاه داده اول: داده‌های آموزش

EbayID	Quantity	Price	PricePerC	StartingBi	SellerName	SellerClos	Category	PersonID	StartingBi	AvgPrice	EndDay
4E+11	0	0.99	0.5892	0.5892	harryjeanl	0.128269	73409	8215	0.99	1.68	Sunday
3.51E+11	0	119	1.1427	1.1427	rrsponts23	0.116667	27260	34	119	104.133	Monday
3.81E+11	1	1.75	1.0416	0.4464	mojo640	0.531599	73409	8215	0.75	1.68	Thursday
3.01E+11	1	66	0.6338	0	mintsigna	0.971014	27260	34	0.01	104.133	Saturday
2.01E+11	0	9.99	0.5197	0.5197	realdealsi	0.187117	27285	28504	9.99	19.2203	Thursday
3.11E+11	0	29.99	0.6202	0.6202	bbcardsco	0.083521	27285	6904	29.99	48.3506	Friday
2.91E+11	1	14.99	1.7565	1.7565	a1a-collec	0.089041	27265	37724	14.99	8.534	Sunday
3.81E+11	0	4.95	1.1272	1.1272	rocky-moi	0.064801	73409	21663	4.95	4.3911	Monday
2.51E+11	0	25	0.6886	0.6886	njh23	0.085714	27260	4697	25	36.3052	Sunday
1.51E+11	0	0.99	0.6101	0.6101	qualitysig	0.029412	73396	16731	0.99	1.6225	Tuesday
1.21E+11	1	150	1.4404	1.4404	oldnewbc	0.232558	27260	34	150	104.133	Tuesday
1.61E+11	0	135	1.2964	1.2964	tony2136	0.032258	27260	34	135	104.133	Tuesday
1.81E+11	0	29.99	0.6202	0.6202	rtstew197	0.092647	27285	6904	29.99	48.3506	Tuesday
2.81E+11	1	97.89	0.6414	0.0064	livehorn	0.928962	27277	16917	0.99	152.6145	Thursday
2.81E+11	0	3.75	3.7878	3.7878	deadedes8	0.026059	73396	7989	3.75	0.99	Wednesday
2.71E+11	0	40	0.8272	0.8272	puppyseb	0.1	27285	6904	40	48.3506	Thursday
1.91E+11	0	2.45	1.8421	1.8421	bclaus	0.066667	73396	12652	2.45	1.33	Wednesday
1.91E+11	0	34.99	1.9577	1.9577	clemensh	0.108333	27295	30170	34.99	17.8725	Friday
2.51E+11	0	179	2.142	2.142	goalposts	0.038835	27277	7548	179	83.5666	Monday

شکل (۳-۲) بخشی از پایگاه داده اول: داده‌های آزمایش

^۲Root Mean Square Error (RMSE)

^۴Mean Absolute Percentage Error (MAPE)

^۵Mean Absolute Error (MAE)

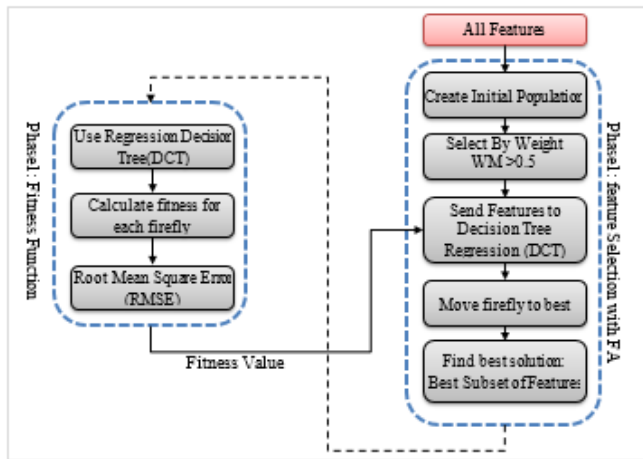
^۱Learning

Function

^۲Hidden Unit

۴-۳- چارچوب راهکار پیشنهادی

در این تحقیق یک مدل یادگیری ماشین و مبتنی بر شبکه های عصبی برای تخمین قیمت نهایی اقلام در حراج های آنلاین ترکیبی ارائه شده است. دیاگرام مدل پیشنهادی در شکل (۳-۳) ترسیم شده است. این مدل از سه گام اساسی انتخاب ویژگی بسته بندی مبتنی بر الگوریتم فراکتشافی کرم شب تاب و تابع برازش مبتنی بر خطای درخت تصمیم رگرسیون، مرحله تخمین قیمت نهایی توسط سه مدل انفرادی شبکه عصبی بازگشتی با واحد پنهان BiLSTM، شبکه عصبی پیشخور و درخت تصمیم و گام ترکیب نتایج با استفاده از تکنیک رای گیری اکثریت نرم تشکیل شده است. در ادامه این فصل سه گام فوق به صورت جزیی معرفی می شوند.

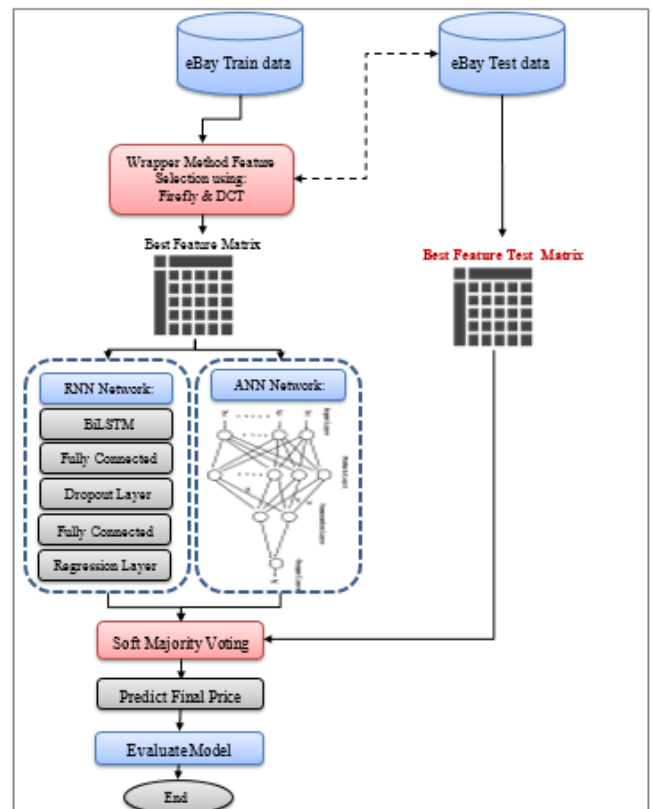


شکل (۳-۴) دیاگرام بخش انتخاب ویژگی

۴-۳-۵- انتخاب ویژگی بسته بندی مبتنی بر الگوریتم کرم شب تاب

اولین مرحله در روش پیشنهادی، مرحله انتخاب ویژگی است که در آن از الگوریتم کرم شب تاب برای تولید زیر مجموعه بهینه از ویژگی ها استفاده می شود. در الگوریتم های فراکتشافی ابتدا یک جمعیت اولیه تولید می شود که هر یک از اعضای این جمعیت اولیه یک راه حل برای مسئله هستند. با توجه به اینکه در این تحقیق از الگوریتم کرم شب تاب برای انتخاب ویژگی استفاده شده است در نتیجه هر عضو از جمعیت اولیه زیر مجموعه ای از کل ویژگی های پایگاه داده است. دیاگرام جزیی این بخش از مدل در شکل (۴-۳) ترسیم شده است.

است. پس از تولید هر عضو از جمعیت برای انتخاب ویژگی های برتر از حداقل آستانه بزرگتر از 0.5 استفاده می شود $MW > 0.5$ ، به عبارت ساده در بین ویژگی های ارائه شده در شکل زیر فقط ویژگی هایی انتخاب می شوند که مقداری بزرگتر از 0.5 دارند. به عنوان مثال ویژگی اول از بین 30 ویژگی اولین عضو از جمعیت اولیه در شکل زیر دارای مقدار 0.1847 است و از آنجایی که حداقل آستانه برابر با 0.5 تعریف شده است این ویژگی انتخاب می شود. حال برای هر یک از اعضای جمعیت اولیه فرآیند محاسبه برازش با استفاده از داده های **train** و **test** انجام می شود. ویژگی های منتخب از داده های ورودی انتخاب می شوند و به الگوریتم درخت تصمیم رگرسیون ارائه می شوند که مقدار تابع برازش برای آن ها محاسبه شود. در شکل (۳-۵) ستون دوم بیانگر مقدار تابع برازش برای هر عضو از جمعیت اولیه است. به عنوان مثال اولین عضو از جمعیت اولیه دارای مقدار برازش برابر با $1/0.108$ است. تابع برازش در مسئله این تحقیق RMSE یا جذر میانگین مربعات خطا است که از رابطه (۳-۱) محاسبه می شود.



شکل (۳-۳) دیاگرام مدل تخمین قیمت پیشنهادی

$$\varepsilon = \delta \times U \quad (-1, 1) \quad (4-3)$$

در رابطه (۳-۲) $x_{i,k}$ بیانگر بردار کرم شب تاب اول و $x_{j,k}$ بیانگر بردار کرم شب تاب دوم است. در واقع در فاصله اقلیدسی بین دو بردار، اختلاف بین مؤلفه‌های متناظر دو بردار محاسبه شده و جمع می‌شود. در رابطه (۳-۳)، β بیانگر مقدار جذابیت اولیه است که در مسئله این تحقیق برابر با ۲ در نظر گرفته شده است. e نماد نمایی یا Exponential است و گاما برابر با ۱ در نظر گرفته شده است. مقدار r نیز در رابطه (۳-۳) توسط رابطه (۳-۲) حاصل می‌شود. رابطه (۳-۴) مقدار اپسیلون را محاسبه می‌کند که یک بردار از اعدادی است که از حاصل ضرب توزیع یکنواخت پیوسته تصادفی بین $(-1, 1)$ و مقدار دلتا δ حاصل می‌شود. در مسئله این تحقیق مقدار دلتا برابر با ۰.۰۵ تعریف شده است. پس از محاسبه مقادیر فوق یک عضو جدید از جمعیت اولیه با استفاده از رابطه (۳-۵) تولید می‌شود. عضو جدید با ویرایش عضو جمعیت اولیه تولید می‌شود. در این رابطه x_j یک عضو از جمعیت اولیه با برازش بیشتر و x_i یک عضو از جمعیت اولیه با مقدار برازش کمتر است. که اختلافشان محاسبه شده و با مقدار β حاصل شده از رابطه (۳-۳) ضرب می‌شود. سپس مقدار فوق با حاصل ضرب آلفا و اپسیلون جمع می‌شود و مقدار نهایی با مقادیر عضو جمعیت اولیه با برازش کمتر x_i جمع می‌شود. در مسئله این تحقیق مقدار آلفا α برابر با ۰.۲ در نظر گرفته شده است. با این فرآیند کرم شب تاب با برازش کمتر به سمت کرم شب تاب با برازش بیشتر حرکت می‌کند.

$$newx_i = x_i + \beta(x_j - x_i) + \alpha \varepsilon \quad (5-3)$$

این بخش از روش با ذکر یک مثال بطور دقیق‌تر بررسی می‌شود. مطابق با جمعیت اولیه ارائه شده در شکل (۵-۳) دومین عضو از جمعیت دارای برازش ۲۱/۶۷ و اولین عضو از جمعیت دارای برازش ۱/۰۸ است در نتیجه طبق فرآیند فوق، دومین عضو از جمعیت اولیه باید به سمت اولین عضو از جمعیت اولیه مهاجرت کند. در شکل (۶-۳)، اولین و دوم عضو از جمعیت اولیه به ترتیب در سطر اول و دوم ارائه شده‌اند. در سطر چهارم اختلاف بین مؤلفه‌های متناظر دو عضو فوق محاسبه شده و در سطر پنجم نیز اختلاف بین مؤلفه‌ها به توان ۲ رسیده است. در سطر ششم مجموع مقادیر سطر پنجم حاصل شده و در سطر هفتم جذر مقدار سطر ششم حاصل شده است که بیانگر فاصله اقلیدسی بین این دو عضو است و برابر با ۲/۵۰ می‌باشد. این مقدار در واقع مقدار r_{ij} در رابطه (۲-۳) است. در سطر هشتم با استفاده از مقادیر حاصل شده از فاصله بین دو کرم شب تاب مقدار جذابیت محاسبه می‌شود. بردار اپسیلون در سطر یازدهم نشان داده شده است که از ضرب مقدار دلتا (برابر با ۰/۰۵) و بردار ۳۰ عضوی توزیع پیوسته یکنواخت بین $(-1, 1)$ که در سطر دهم نشان داده شده است، حاصل می‌شود. در سطر آخر جمعیت جدید نشان داده شده است. این جمعیت مطابق با فرمول (۵-۳) حاصل می‌شود. ابتدا مقدار بتا برابر با ۰/۰۳ در مقادیر سطر چهارم ضرب می‌شود. سپس آلفا برابر با ۰/۲ در مقادیر بردار اپسیلون در سطر یازدهم ضرب می‌شود. سپس مقدار این دو بردار حاصل شده با مقدار عضو دوم جمعیت ارائه شده در سطر دوم جمع می‌شود.

صعودی مرتب می‌شوند و اولین عضو از جمعیت که دارای کمترین مقدار خطا یا بیشترین مقدار برازش است به عنوان راه حل نهایی و بهترین زیر مجموعه از ویژگی‌ها انتخاب می‌شوند. در شکل (۳-۸) جمعیت تولید شده و مرتب شده براساس برازش در گردش اول از الگوریتم و در شکل (۳-۹) نمودار همگرایی الگوریتم کرم شب تاب در ۱۰۰ تکرار نشان داده شده است.

Position	Cost
1x30 double	1.0108
1x30 double	21.6752
1x30 double	21.7492
1x30 double	1.0713
1x30 double	25.4538

pop(1).Position										
pop(1).Position	1	2	3	4	5	6	7	8	9	10
	0.8147	0.9058	0.1270	0.9134	0.6324	0.0975	0.2785	0.5469	0.9575	0.9649
	11	12	13	14	15	16	17	18	19	20
	0.1576	0.9706	0.9572	0.4854	0.8003	0.1419	0.4218	0.9157	0.7922	0.9595
	21	22	23	24	25	26	27	28	29	30
	0.6557	0.0357	0.8491	0.9340	0.6787	0.7577	0.7431	0.3922	0.6555	0.1712

شکل (۵-۳) پنج عضو از جمعیت تصادفی اولیه و مقدار برازش هر یک از آن‌ها

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_{Pred}^i - y_{Truth}^i)^2} \quad (1-3)$$

در رابطه بالا y_{Pred}^i بیانگر پیش‌بینی‌های مدل و y_{Truth}^i بیانگر مقدار واقعی یا قیمت اصلی اقلام است. بطور دقیق در این گام برای محاسبه مقدار برازش برای هر یک از اعضای جمعیت اولیه، ابتدا با توجه به ویژگی‌های انتخاب شده (ویژگی‌هایی که دارای مقدار بیشتر از ۰/۵ هستند) ویژگی‌های منتخب از داده‌های آموزش و آزمایش جدا می‌شوند. سپس بخش انتخاب شده از داده‌های آموزش برای تشکیل یک مدل به درخت تصمیم رگرسیون داده می‌شود و پس از تکمیل آموزش آن توسط داده‌های آموزشی، داده‌های منتخب آزمایش به مدل داده می‌شود تا قیمت نهایی را برآورد کند. سپس با توجه به پیش‌بینی مدل و مقادیر حقیقی قیمت‌ها مقدار برازش برای هر عضو از جمعیت با توجه به رابطه (۱-۳) محاسبه می‌شود. فرآیند فوق برای کلیه اعضای جمعیت اولیه تکرار می‌شود. اگر فرض شود کاربر تعداد 5=Pop تعریف کرده باشد برای هر ۵ عضو جمعیت اولیه مطابق با مراحل فوق مقدار برازش محاسبه می‌گردد. به تعداد تکرار تعیین شده توسط کاربر برای تکرار الگوریتم کرم شب تاب، مهاجرت کرم‌های شب‌تاب به سمت بهینه سراسری انجام می‌شود. در الگوریتم کرم شب تاب کلیه کرم‌های شب‌تابی (راه‌حلی‌هایی) که دارای برازش کمتر هستند به سمت راه حل‌ها با برازش بیشتر حرکت می‌کنند. در هر گردش مراحل به شرح زیر انجام می‌شود. ابتدا با استفاده از فاصله اقلیدسی ارائه شده در رابطه (۳-۲)، فاصله بین کرم شب تاب با برازش کمتر و کرم شب تاب با برازش بیشتر محاسبه می‌شود. سپس با استفاده از رابطه (۳-۲) و (۴-۳) مقادیر β ، ε محاسبه می‌شود.

$$r_{ij} = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad (3-2)$$

$$\beta = \beta \cdot e^{-\gamma r} \quad (3-3)$$

Position	Cost
1x30 double	0.9831
1x30 double	1.0276
1x30 double	1.0289
1x30 double	1.1163
1x30 double	1.1225
1x30 double	1.1906
1x30 double	9.0654
1x30 double	22.0551
1x30 double	28.3378

شکل (۳-۸) جمعیت نهایی مرتب شده بر اساس براز

پس از انتخاب ویژگی‌های بهینه، داده‌های کاهش ابعاد داده شده به دو شبکه عصبی پیشخور و بازگشتی با واحد پنهان BiLSTM ارائه می‌شود که این دو شبکه آموزش داده شوند و بتوانند قیمت نهایی را در داده‌های آزمایش پیش‌بینی کنند. مطابق با دیاگرام مدل پیشنهادی در شکل (۳-۳) شبکه عصبی بازگشتی پیشنهاد شده در این تحقیق از ۶ لایه به شرح زیر تشکیل شده است:

لایه ورودی یا Sequence Input: اولین لایه در شبکه بازگشتی پیشنهادی وظیفه دریافت ورودی را بر عهده دارد. ورودی‌ها به صورت بردارهای عددی هستند.

لایه واحد پنهان: لایه دوم، نوع واحد پنهان تعریف شده در شبکه است و در شبکه پیشنهادی این لایه BiLSTM تعریف شده است. لایه‌ها پنهان در شبکه بازگشتی وابستگی بین داده‌های ورودی را فرا می‌گیرند. تعداد واحدهای پنهان در این لایه تعریف می‌شود. تابع فعال‌ساز این لایه در مدل پیشنهادی تانژانت هایپربولیک (Tanh) است. تابع فعال‌سازی دروازه‌ها سیگموئید (Sigmoid) است.

لایه تمام متصل یا Fully Connected: سومین لایه تمام متصل است. این لایه مشابه با شبکه‌های عصبی پیشخور است که خروجی لایه قبل را مناسب با خروجی شبکه کامپایل می‌کند. در این لایه تعداد کلاس داده‌ها تعیین می‌شود. در شبکه پیشنهادی تعداد کلاس داده برابر با ۵۰ در نظر گرفته شد. خروجی این لایه برابر با ۵۰ خواهد بود.

لایه حذف تصادفی Drop Out: لایه حذف تصادفی فرآیند منظم‌سازی شبکه را انجام می‌دهد. این لایه احتمال معینی را دریافت می‌کند و بر طبق آن، تعدادی از نرون‌ها و اتصالات آن‌ها را به صورت تصادفی غیر فعال یا خاموش می‌کند. وجود این لایه در شبکه به تعمیم‌پذیری بهتر شبکه کمک می‌کند و علاوه بر این بیش‌برازش^۱ را نیز کنترل می‌کند. با غیر فعال نمودن برخی از نرون‌ها و اتصالات آن‌ها، معماری شبکه به چند معماری کوچک‌تر تبدیل می‌شود شبکه‌ای با وزن کم به عنوان نمایند برتر انتخاب می‌شود [۱۸]. در شکل (۳-۱۰) مثالی از لایه حذف تصادفی با احتمال حذف ۵٪ یا ۵۰ درصد ارائه شده است. در مدل پیشنهادی احتمال حذف در لایه حذف تصادفی برابر ۲۰ درصد (۰/۲) تعریف گردید.

لایه رگرسیون: آخرین لایه خطای نصف میانگین مربع خطا^۲ را محاسبه می‌کند. در این لایه برآورد قیمت نهایی انجام می‌شود. این لایه دارای یک تابع خطا یا loss است که در مسئله این تحقیق جذر میانگین مربعات خطا (RMSE) است. در این لایه قیمت حقیقی داده‌ها با قیمت پیش‌بینی شده توسط شبکه مقایسه می‌شود، حاصل مقایسه فوق مقدار خطای شبکه

است که برای برورسانی وزن‌های لایه‌های شبکه استفاده می‌شود. فرآیند فوق تا رسیدن به حداکثر تعداد تکرار (epoch) مشخص شده برای شبکه ادامه می‌یابد. در شکل (۳-۱۱) نمودار مرحله آموزش شبکه عصبی بازگشتی پیشنهادی در پایگاه داده اول ارائه شده است.

پس از پایان آموزش شبکه عصبی بازگشتی، مرحله یادگیری شبکه عصبی پیشخور با داده‌های بخش آموزش انجام می‌شود. در شبکه پیشخور پیشنهادی تعداد نرون در لایه پنهان برابر با ۵۰ تعریف گردید. در شکل (۳-۱۲) مشخصات شبکه عصبی پیشخور تعریف شده در این تحقیق ارائه شده است. تابع فعال‌ساز در این شبکه تابع ReLU است و تابع بهینه‌ساز شبکه نیز تابع LBFGS^۳ است. BFGS با حافظه محدود (L-BFGS یا LM-BFGS) یک الگوریتم بهینه‌سازی در خانواده روش‌های شبه نیوتنی است که الگوریتم برودن-فلچر-گلدفارب-شانو (BFGS) را با استفاده از حافظه کامپیوتر محدود تقریب می‌زند. این الگوریتم یک روش بهینه‌سازی عددی است که به طور موثر برای تخمین پارامترها برای آموزش مدل‌های مختلف یادگیری ماشین استفاده می‌شود [۱۹]. پس از تشکیل معماری شبکه پیشخور رگرسیون با استفاده از داده‌های آموزش مرحله یادگیری شبکه کامل می‌شود. سپس از شبکه آموزش دیگه برای برآورد قیمت داده‌های بخش آزمایش استفاده می‌شود.

پس از تکمیل فرآیند یادگیری هر دو شبکه، از آن‌ها برای پیش‌بینی قیمت نهایی داده‌های بخش آزمایش استفاده می‌شود. خروجی این بخش از مدل دو بردار به طول داده‌های بخش آزمایش است که در واقع شامل مقادیر قیمت پیش‌بینی شده توسط دو شبکه هستند. مقادیر پیش‌بینی توسط دو شبکه با استفاده از تکنیک رای گیری اکثریت نرم با یکدیگر ترکیب می‌شوند و مقادیر قیمت نهایی برآورد شده توسط مدل را تشکیل می‌دهند.

پس از تشکیل بردارهای پیش‌بینی قیمت توسط دو شبکه عصبی برای رسیدن به قیمت نهایی مدل، خروجی دو شبکه با استفاده از تکنیک رای گیری اکثریت نرم ترکیب می‌شوند. تکنیک رای گیری اکثریت دارای دو حالت نرم و سخت است. در تکنیک رای گیری سخت که برای مسائل طبقه‌بندی استفاده می‌شود، کلاس نهایی داده برابر با کلاسی است که بیشترین رای را در بین طبقه‌بندی کننده‌ها کسب کرده باشد. تکنیک رای گیری اکثریت نرم در مسائل رگرسیون استفاده می‌شود. در این حالت مقدار نهایی داده‌ها از میانگین حسابی خروجی کلیه مدل‌ها بدست می‌آید. این بخش از مدل با استفاده از یک مثال شرح داده می‌شود. در شکل (۳-۱۳) یک نمونه از خروجی مدل نشان داده شده است. در این شکل برای وضوح بیشتر پیش‌بینی‌های انجام شده توسط مدل، مقدار قیمت پیش‌بینی شده فقط در ۱۲۰ نمونه داده از داده‌های بخش آزمایش ترسیم شده است. لازم به ذکر است مقادیر معیارهای ذکر شده در بالای شکل در کل داده‌های آزمایش محاسبه شده است. در این شکل به خوبی مشاهده می‌شود که پیش‌بینی مدل در حالت رای گیری اکثریت نرم بسیار دقیق است و خطای مدل در تخمین قیمت نهایی پایین است. در فصل آتی نتایج مدل بطور کامل گزارش می‌شود.

^۱ Limited-memory BFGS^۲ Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS)^۱ Over fitting^۲ Half-Mean-Squared-Error Loss

- جذر میانگین مربعات خطا^۱: این معیار با توجه به رابطه (۴-۱)، اختلاف بین قیمت پیش‌بینی شده و قیمت حقیقی را محاسبه می‌کند.

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_{Pred}^i - y_{Truth}^i)^2} \quad (۴-۱)$$

در رابطه فوق y_{Pred}^i نشان دهنده قیمت پیش‌بینی شده توسط مدل و y_{Truth}^i بیانگر قیمت حقیقی موجود در پایگاه است [۲۰].

- میانگین مطلق درصد خطا^۲: این معیار که از رابطه (۴-۲) محاسبه می‌شود، دقت پیش‌بینی قیمت را ارائه می‌دهد.

$$MAPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{y_{Truth}^i - y_{Pred}^i}{y_{Truth}^i} \right| \quad (۴-۲)$$

در رابطه بالا y_{Pred}^i نشان دهنده قیمت پیش‌بینی شده توسط مدل و y_{Truth}^i بیانگر قیمت حقیقی موجود در پایگاه است [۶۴].

- میانگین مطلق خطا^۳: معیاری با توجه به رابطه (۴-۳) میانگین خطای مطلق بین جفت مشاهدات را ارائه می‌دهد.

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_{Pred}^i - y_{Truth}^i| \quad (۴-۳)$$

در رابطه بالا y_{Pred}^i نشان دهنده قیمت پیش‌بینی شده توسط مدل و y_{Truth}^i بیانگر قیمت حقیقی موجود در پایگاه است [۶۴].

۴-۳- مشخصات پایگاه داده

در این فصل برای انجام آزمون‌ها، مشابه با مقاله پایه از دو مجموعه داده حراج آنلاین eBay استفاده شد. این دو مجموعه داده به صورت ۴ فایل با پسوند CSV. در بخش‌های آموزش و آزمایش مجزا بوده و از آدرس <https://cims.nyu.edu/~munoz/data> قابل دریافت هستند. در جدول (۴-۱) مشخصات ۴ فایل متعلق به دو مجموعه داده بیان شده است. در جدول (۴-۲) نیز تعدادی از ویژگی‌های مهم آن‌ها همراه با تعاریف هر یک بیان شده است [۲۱].

پیاده‌سازی مدل پیشنهاد شده و روش پایه لی و همکاران [۴] در محیط متلب نسخه ۲۰۲۰ انجام شده است. پارامترهای پیاده‌سازی شبکه عصبی عمیق بازگشتی و شبکه پیشخور و مرحله انتخاب ویژگی توسط الگوریتم کرم شب تاب در جدول (۴-۳) بیان شده است. سه

جدول (۳-۱) رای گیری اکثریت نرم بین شبکه‌ها

Case id	Forecast Price ANN	Forecast Price BiLSTM	Soft Majority Voting	Real Price	Error ^۲
۱	۲,۰	۲,۴۰	۲,۴۷۰	۲,۰	۰,۰۰۰۶
۲	۳,۹۸	۳,۹۱	۳,۹۴۰	۳,۹۰	۰,۰۰۰۰۲۵
۳	۵,۶۱	۵,۶۹	۵,۶۰	۵,۴۹	۰,۰۲۰۶
۴	۵,۰	۵,۱	۵,۳	۵,۰	۰,۰۴
۵	۹,۳۸	۹,۳۳	۹,۳۰۰	۹,۸۹	۰,۲۸۶
۶	۷,۳۹	۷,۴۳	۷,۴۱	۷,۴۱	۰
Sum	-	-	-	-	۰,۳۰۲۲۲
RMSE					۰,۰۹۳

در جدول بالا ۶ نمونه داد فرضی در نظر گرفته شده است که شبکه‌های عصبی پیشخور و بازگشتی، قیمت نهایی را برای ۶ داده مطابق با مقادیر ارائه شده در ستون‌های دوم و سوم ارائه داده‌اند. سپس در ستون چهارم از جدول فوق رای گیری اکثریت نرم بین مقادیر دو شبکه عصبی انجام شده است و میانگین مقادیر هر سطر حساب شده است. در ستون پنجم مقادیر فرضی حقیقی قیمت برای ۶ نمونه داده ذکر شده است و در نهایت در ستون آخر از جدول خطای بین پیش‌بینی نهایی مدل و قیمت حقیقی داده‌ها انجام شده و خطای مدل محاسبه شده است.

۳-۸- جمع بندی

در این فصل راهکار تخمین قیمت نهایی در حراج‌های ترکیبی آنلاین بطور کامل معرفی شد. روش ارائه شده دارای سه فاز اصلی شامل مرحله انتخاب ویژگی مبتنی بر یک روش فرااکتشافی، مرحله تخمین قیمت توسط شبکه‌های عصبی و مرحله رای گیری اکثریت نرم جهت ترکیب نتایج شبکه‌ها است. در فاز انتخاب ویژگی از الگوریتم فراابتکاری کرم شب‌تاب برای انتخاب زیر مجموعه بهینه از ویژگی‌ها استفاده می‌شود. این الگوریتم زیر مجموعه‌های تصادفی از ویژگی‌ها تولید می‌کند و زیر مجموعه‌ای که دارای خطای کمتری باشد را به عنوان بهترین ویژگی‌ها ارائه می‌دهد. در فاز تخمین قیمت توسط شبکه عصبی بازگشتی BiLSTM و شبکه رگرسیون پیشخور، داده‌های آموزش به ترتیب شبکه‌ها داده می‌شوند تا دو شبکه آموزش ببینند و سپس دو شبکه در داده‌های آزمایش پیش‌بینی قیمت را انجام می‌دهند. در مدل پیشنهادی یک شبکه بازگشتی ۶ لایه‌ای شامل لایه‌های واحد پنهان و حذف تصادفی و تمام متصل و رگرسیون ارائه شد. در نهایت پس از حصول پیش‌بینی‌های هر یک از شبکه‌ها، برای رسیدن به برآورد قیمت نهایی مدل با استفاده از تکنیک رای گیری اکثریت نرم و میانگین گیری از خروجی هر یک از شبکه‌ها قیمت نهایی را برآورد می‌کند. در پایان ارزیابی مدل با مقایسه قیمت‌های حقیقی اقلام و قیمت‌های پیش‌بینی شده توسط مدل انجام می‌شود. در فصل آتی نتایج این بخش گزارش می‌شود

معیارهای ارزیابی

برای بررسی عملکرد مدل ارائه شده در این تحقیق با توجه به مقاله پایه از سه معیار، جذر میانگین مربعات خطا (RMSE)، میانگین مطلق درصد خطا (MAPE) و میانگین مطلق خطا (MAE) استفاده شده است. تعاریف و روابط معیارهای بالا در زیر بیان شده است:

^۲Mean Absolute Error (MAE)

^۱Root Mean Square Error (RMSE)

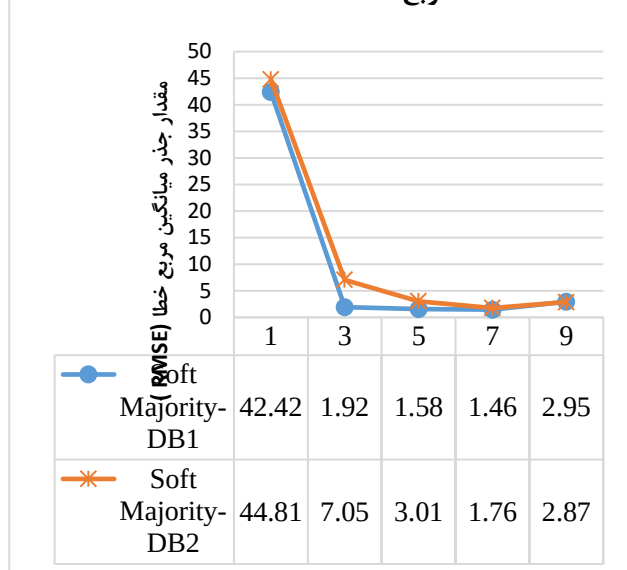
^۲Mean Absolute Percentage Error (MAPE)

متغیر مستقل تحقیق مقادیرشان بازه ای تعریف شده و در آزمون ها بهترین مقدار آن ها مشخص می شود.

جدول (۳-۴) پارامترهای پیاده سازی

نام پارامتر	مقدار تعیین شده
تعداد ویژگی	[۱-۹]
تعداد تکرار شبکه ها (Max Epoch)	۵۰
سایز MiniBatch	۱۰۲۴
تعداد واحد پنهان (Num. Hidden Unit)	۲۰۰
تعداد نرون در شبکه پیشخور	[۵۰-۲۰۰]
نرخ حذف تصادفی (Drop Out Rate)	۰/۲
تابع یادگیری سه شبکه (Optimizer Function)	[Adam, Rmsprop, SGDM]
نرخ یادگیری اولیه (initial Learning Rate)	۰/۰۱
Gradient Threshold	۱
تعداد تکرار الگوریتم کرم شب ناپ	۱۰۰
تعداد جمعیت اولیه	۲۵
مقدار γ ^۱	۱
مقدار β ^۲	۲
مقدار α ^۳	۰/۲

بررسی تاثیر تعداد ویژگی بر جذر میانگین مربع خطا



شکل (۱-۴) بررسی تعداد ویژگی منتخب در الگوریتم کرم شب تاب

نتایج نمودار بالا نشان می دهد که در تعداد ۱ ویژگی نتایج خطای مدل بسیار بالا است به این دلیل که تعداد یک ویژگی به هیچ صورتی نمی تواند دانش کافی برای پیش بینی قیمت فراهم کند. بهترین نتایج در تعداد ۷ ویژگی حاصل شده است. به این معنی که با تعداد ۷ ویژگی خطای مدل به حداقل ممکن رسیده است. با افزایش تعداد ویژگی به ۹ مجدداً نتایج خطای مدل روند افزایشی دارد به این معنی که در نظر گرفتن ویژگی های بیشتر نه تنها عملکرد مدل را در تخمین قیمت نهایی کاهش می دهد بلکه باعث افزایش پیچیدگی های محاسباتی مدل نیز می شود. در نتیجه تعداد نهایی ویژگی در هر دو پایگاه داده که باعث می شود مدل به بهترین نتایج دست یابد تعداد ۷ ویژگی است.

۴-۵- بررسی تاثیر تابع بهینه ساز در یادگیری شبکه بازگشتی

عملکرد شبکه عصبی بازگشتی استفاده شده در این مدل وابسته به تابع بهینه سازی شبکه است که در فرآیند یادگیری شبکه موثر است. در محیط متلب سه تابع بهینه ساز برای شبکه عصبی بازگشتی تعریف شده است که در شکل (۲-۴) نتایج مدل در دو پایگاه داده با سه تابع بررسی شده است تا بتوان بهترین تابع را برای مرحله یادگیری شبکه بازگشتی انتخاب نمود.

نتایج نشان می دهد در بین سه تابع، بهترین نتایج متعلق به تابع است که در اغلب مسائل که از شبکه های عمیق استفاده می شود عملکرد بهتری در مقایسه با دو بهینه ساز دیگر ارائه می دهد. تابع Adam در مقایسه با دو بهینه ساز دیگر دارای سرعت بالاتری است و از اینرو باعث می شود شبکه فرآیند یادگیری را با سرعت بیشتری انجام دهد. نتایج این آزمایش نشان می دهد در مدل ما در هر دو پایگاه داده حراج آنلاین، تابع Adam می تواند خطای شبکه بازگشتی BiLSTM را به حداقل ممکن برساند.

از بین سه متغیر مستقل این تحقیق در دو آزمایش فوق بهترین مقادیر برای دو پارامتر تعداد ویژگی و تابع بهینه ساز تعیین گردید. در آزمایش بعد سومین متغیر مستقل این تحقیق بررسی

^۳Mutation Coefficient^۱Light Absorption Coefficient^۲Attraction Coefficient Base Value

می‌شود و تاثیر آن بر روی اصلی‌ترین متغیر وابسته تحقیق یعنی جذر میانگین مربعات خطا گزارش می‌شود.

۴-۶- بررسی تاثیر تعداد نرون در لایه پنهان شبکه عصبی پیشخور

شبکه عصبی پیشخور از دیگر شبکه‌های مورد استفاده در مدل این تحقیق است. یکی از اصلی‌ترین پارامترهای موثر در عملکرد این شبکه تعداد نرون در لایه پنهان آن است. تعداد نرون در این شبکه باید بنحوی تعیین شود که پیچیدگی شبکه افزایش نداشته باشد و در عین حال بهترین خروجی از شبکه حاصل شود. برای بررسی این پارامتر آزمایشی طراحی شد که در آن تعداد نرون لایه پنهان شبکه عصبی پیشخور بین ۵۰ تا ۲۰۰ تعریف گردید و عملکرد شبکه بررسی شد. نتایج این آزمایش در شکل (۴-۳) بیان شده است.

شکل (۴-۳) بررسی تاثیر تعداد نرون پنهان بر عملکرد شبکه عصبی ANN

نتایج نمودار بالا نشان می‌دهد که حداقل خطا از تعداد ۵۰ نرون پنهان، حاصل شده است و با تعداد نرون‌های ۱۰۰ و ۲۰۰ نیز همین نتایج حاصل شده است؛ با این تفاوت که با افزایش تعداد نرون در لایه پنهان سرعت فرآیند آموزش شبکه به شدت کاهش می‌یابد به این دلیل که حجم محاسبات شبکه افزایش می‌یابد. از اینرو برای رسیدن به بهترین نتایج در شبکه عصبی پیشخور تعداد نرون ۵۰ در نظر گرفته شد.

در این مرحله با انجام سه گروه آزمون مختلف تاثیر متغیرهای مستقل تحقیق بر روی متغیرهای وابسته تحقیق بررسی گردید و در عین حال بهترین مقادیر برای پارامترهای مهم مدل نیز مشخص گردید. در آزمون‌های بعدی مقایسه بین مدل پیشنهاد شده و روش پایه انجام می‌شود. آزمون‌های مقایسه‌ای برای بررسی اهداف مطرح در این تحقیق، طراحی شده‌اند.

۴-۷- مقایسه نتایج در پایگاه داده اول

در آزمون‌های مقایسه‌ای حالت‌های مختلف مدل پیشنهادی بررسی شده است. در این آزمون‌ها نتایج هر یک از مدل‌ها شامل شبکه عصبی بازگشتی BiLSTM و شبکه عصبی پیشخور و درخت تصمیم (از آن در مرحله انتخاب ویژگی استفاده شده است) به صورت انفرادی گزارش شده است. سپس نتایج ترکیب‌های مختلف هر یک از این مدل‌ها نیز گزارش شده است. در نهایت نتایج کلیه روش‌های پیشنهادی فوق با روش پایه لی و همکاران [۴] مقایسه شده است. قبل از گزارش نتایج به عنوان نمونه، نمودار پیش بینی قیمت توسط هر یک از مدل‌ها و ترکیب آن‌ها در ۱۲۰ نمونه داده نشان داده می‌شود. لازم به ذکر است در کلیه این شکل‌ها مقادیر ذکر شده از متغیرهای ارزیابی بالای هر نمودار، در کل داده‌های آزمایشی محاسبه شده، اما برای وضوح شکل، نمودار فقط پیش‌بینی را در ۱۲۰ نمونه داده نشان می‌دهد. در شکل‌های (۴-۴) تا (۴-۱۱) نمونه‌های پیش‌بینی مدل‌ها نشان داده شده است. در شکل (۴-۱۲) مقایسه نتایج روش‌ها بیان شده و در جدول (۴-۴) مقادیر نتایج بیان شده است. در کلیه شکل‌ها FA مخفف الگوریتم انتخاب ویژگی کرم شب تاب و DT مخفف درخت تصمیم است.

نتایج گزارش شده نشان می‌دهد کلیه روش‌های پیشنهادی در مقایسه با روش لی و همکاران به نتایج بهتر و خطای کمتر در تخمین قیمت رسیده‌اند. در بین مدل‌های پیشنهاد شده بهترین عملکرد متعلق به درخت تصمیم است. پس از آن تکنیک رتبی گیری اکثریت نرم بین سه مدل شبکه عصبی پیشخور و بازگشتی BiLSTM و درخت تصمیم بهترین نتایج را دارد. در مقابل بدترین نتایج در بین مدل‌های پیشنهادی متعلق به شبکه عصبی پیشخور است در بین دو شبکه عصبی این تحقیق شبکه بازگشتی BiLSTM در مقایسه با شبکه پیشخور موفق‌تر بوده است و این به دلیل ساختار و معماری قویتر آن در مقایسه با شبکه پیشخور ساده است. اما بطور کلی هر ۷ روش پیشنهاد شده در این تحقیق در مقایسه با روش لی عملکرد بهتری داشته‌اند. این موفقیت به دلیل مرحله انتخاب ویژگی است. یکی از اصلی‌ترین مراحل در روش‌های یادگیری ماشین مرحله انتخاب ویژگی است که اگر به درستی انجام شود باعث می‌گردد نتایج مدل‌ها بنحو موثری بهبود داده شود. در مدل پیشنهاد شده نیز، استفاده

از الگوریتم کرم شب تاب و استفاده از یک روش انتخاب ویژگی بسته‌بندی که مبتنی بر خطای الگوریتم درخت تصمیم است باعث شده کلیه روش‌ها به نتایج مناسبی در تخمین قیمت دست یابند. از طرف دیگر تکنیک رای گیری اکثریت نرم در ۴ مدل پیشنهاد شده باعث موفقیت روش‌های گروهی در مقایسه با روش‌های انفرادی شده است. به عنوان نمونه ترکیب شبکه بازگشتی BiLSTM و شبکه پیشخور با تکنیک رای گیری اکثریت نرم باعث شده RMSE به ۲/۷۵ کاهش یابد در صورتی که این معیار در حالت انفرادی در شبکه ANN برابر با ۴/۱ و در شبکه BiLSTM برابر با ۳/۲۲ است. در نتیجه می‌توان گفت ترکیب به روش رای گیری اکثریت نرم در مدل این تحقیق، منجر به بهبود نتایج شده است. نتایج آزمون‌های فوق به صورت زیر خلاصه می‌شود:

- شبکه عصبی پیشخور در مقایسه با روش لی و همکاران به ترتیب در دو معیار MAE، RMSE توانسته است خطا را ۳/۳۹، ۸/۲۲۱ بهبود ببخشد اما در معیار MAPE موفق نبوده و ۰/۰۶ خطای بیشتری داشته است.
- شبکه بازگشتی BiLSTM در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۳/۶۷، ۹/۱۰ و ۱/۱۵ خطا را کاهش دهد.
- درخت تصمیم در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۴/۴۴، ۱۰/۷۷ و ۱/۱۸ خطا را کاهش دهد.
- ترکیب نرم شبکه بازگشتی و شبکه پیشخور در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۳/۶۷، ۹/۵۷ و ۰/۶۷ خطا را کاهش دهد.
- ترکیب نرم درخت تصمیم و شبکه بازگشتی در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۴/۲۶، ۱۰/۴۹ و ۱/۲۶ خطا را کاهش دهد.
- ترکیب نرم درخت تصمیم و شبکه پیشخور در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۳/۹۰، ۱۰/۰۹ و ۰/۷۸ خطا را کاهش دهد.
- ترکیب نرم درخت تصمیم، شبکه بازگشتی و شبکه پیشخور در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۴/۳۷، ۱۰/۸۶ و ۱/۰۵ خطا را کاهش دهد.

در پایگاه داده دوم نتایج مدل‌های ارائه شده در این تحقیق در مقایسه با روش لی و همکاران به مراتب بالاتر از پایگاه اول است. در این پایگاه مرحله انتخاب ویژگی بسته بندی مبتنی بر الگوریتم کرم شب تاب به طور چشمگیری باعث موفقیت مدل‌ها شده است. از طرف دیگر رای گیری اکثریت نیز در موفقیت مدل‌ها موثر بوده است. در این پایگاه ترکیب اکثریت نرم سه مدل بازگشتی، درخت تصمیم و شبکه پیشخور در مقایسه با کلیه مدل‌ها به بهترین نتایج و کمترین خطا رسیده است. در رتبه دوم ترکیب درخت تصمیم و شبکه پیشخور قرار می‌گیرد. در این پایگاه داده شبکه عصبی پیشخور در مقایسه با شبکه بازگشتی به نتایج بهتری رسیده است. در واقع می‌توان گفت عملکرد این دو شبکه به شدت وابسته به ماهیت داده‌ها است و هر دو شبکه در داده‌های مناسب خود می‌توانند بهترین نتایج را ارائه دهند. در این پایگاه بیشترین موفقیت در معیار RMSE رخ داده است. نتایج آزمون‌های فوق به صورت زیر خلاصه می‌شود:

- شبکه عصبی پیشخور در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است خطا را ۱۰/۳۶، ۱۹/۲۴ و ۱/۳۱ بهبود دهد.
- شبکه بازگشتی BiLSTM در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۹/۸۷، ۱۸/۸۵ و ۰/۹۷ خطا را کاهش دهد.
- درخت تصمیم در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۱۰/۴۵، ۱۸/۹۱ و ۱/۳۴ خطا را کاهش دهد.
- ترکیب نرم شبکه بازگشتی و شبکه پیشخور در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۱۰/۴۷، ۱۹/۹۱ و ۱/۳۵ خطا را کاهش دهد.

بر شبکه‌های عصبی و یادگیری ماشین سنتی برای پیش‌بینی قیمت نهایی در حراج‌های آنلاین ارائه شد. مدل پیشنهاد شده مبتنی بر سه الگوریتم شامل درخت تصمیم، شبکه عصبی عمیق بازگشتی BiLSTM و شبکه عصبی پیشخور است. در این روش برای بهبود عملکرد آن در تخمین قیمت نهایی و در راستای کاربردی شدن آن، از تکنیک انتخاب ویژگی بسته‌بندی مبتنی بر الگوریتم فراابتکاری کرم شب‌تاب استفاده گردید. نتایج ارزیابی مدل پیشنهاد شده بر روی دو پایگاه داده eBay نشان داد که کلیه روش‌های انفرادی و همچنین ترکیب نرم سه روش فوق دارای نتایج بهتری در مقایسه با روش پایه لی و همکاران است. در آزمون‌های مقایسه‌ای در هر دو پایگاه داده بیشترین موفقیت مدل‌ها در معیار RMSE بدست آمده است. در بین مدل‌های منفرد درخت تصمیم بهترین نتایج را دارد و پس از آن با توجه به ماهیت داده‌ها شبکه عصبی پیشخور و بازگشتی قرار می‌گیرند. در پایگاه داده اول مدل‌ها توانستند جذر مربعات خطا RMSE را بین ۸ تا ۱۰ و در پایگاه داده دوم بین ۱۹ تا ۲۰ کاهش دهند. معیار MAE را نیز در پایگاه داده اول مدل‌ها بین ۳ تا ۴ و در پایگاه داده دوم بین ۹ تا ۱۰ کاهش دادند. معیار MAPE را نیز مدل‌ها بین ۰/۵ تا ۰/۱ و در پایگاه اول و ۰/۹ تا ۱/۵ در پایگاه دوم در مقایسه با روش لی کاهش دادند. بطور کلی نتایج نشان می‌دهد مرحله انتخاب ویژگی به شدت منجر به کاهش خطای مدل‌ها در پیش‌بینی قیمت نهایی می‌شود. همچنین تکنیک رای گیری اکثریت نرم نیز به خوبی می‌تواند نتایج مدل‌های انفرادی را بهبود ببخشد.

۵-۲- کارهای آینده

با توجه به روش مطرح شده در این تحقیق و محدودیت‌های آن، کارهای آتی پیشنهادی به شرح زیر می‌باشند:

- استفاده از روش‌های یادگیری ماشین دیگر و شبکه‌های بازگشتی با واحدهای پنهان دیگر مثل LSTM و ترکیب آن‌ها با استفاده از تکنیک رای گیری اکثریت نرم می‌تواند باعث بهبود نتایج در مقایسه با مدل این تحقیق شود.
- استفاده از روش‌های فرااکتشافی جدیدتر مثل الگوریتم فرااکتشافی موش صحرایی در مقایسه با روش استفاده شده در این تحقیق می‌تواند منجر به انتخاب ویژگی‌های بهتر و در نتیجه به بود نتایج شود.
- بررسی مدل پیشنهاد شده در این تحقیق بر روی پایگاه داده‌های حراج‌های آنلاین دیگر از جمله کارهایی است که بررسی آن می‌تواند عملکرد مدل را به نحو بهتری نشان دهد.

فهرست منابع

- [1] E. Öz, "Empirical studies on price determinants of online auctions with machine learning applications," ۲۰۱۹.
- [2] R. Ghani and H. Simmons, "Predicting the end-price of online auctions," in *International workshop on data mining and adaptive modelling methods for economics and management*, 2004: Citeseer.
- [3] X. Li and H. Dong, "A Transaction Trade-Off Utility Function Approach for Predicting the End-Price of Online Auctions in IoT," *Wireless Communications and Mobile Computing*, vol. 2021, 2021.
- [4] X. Li, H. Dong, and S. Han, "Multiple linear regression with Kalman filter for predicting end prices of online auctions," in *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, 2020: IEEE, pp. ۱۸۲-۱۹۱.

- ترکیب نرم درخت تصمیم و شبکه بازگشتی در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۱۰/۴۶، ۱۹/۴۰ و ۱/۳۵ خطا را کاهش دهد.
 - ترکیب نرم درخت تصمیم و شبکه پیشخور در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۱۰/۶۸، ۲۰/۲۲ و ۱/۳۷ خطا را کاهش دهد.
 - ترکیب نرم درخت تصمیم، شبکه بازگشتی و شبکه پیشخور در مقایسه با روش لی و همکاران به ترتیب در سه معیار MAE، RMSE و MAPE توانسته است ۱۰/۶۹، ۲۰/۱۸ و ۱/۴۷ خطا را کاهش دهد.
- در جدول (۴-۶) خلاصه نتایج مقایسه مدل پیشنهادی و روش پایه بیان شده است. نمودارهای افقی در این جدول حاکی از آن است که اهداف این تحقیق حاصل شده‌اند. در این تحقیق هدف اصلی ارائه یک مدل تخمین قیمت نهایی در حراج‌های آنلاین با استفاده از شبکه‌های عصبی اعم از ساده و عمیق است که نتایج ارائه شده روی دو پایگاه داده حراج آنلاین نشان داد هدف اصلی این تحقیق حاصل شده است

۹- جمع بندی

در ابتدای این فصل مشخصات پایگاه داده‌ها و معیارهای ارزیابی بیان شد و در ادامه نتایج حاصل از آزمایش‌های مختلف گزارش شد. در این فصل ابتدا تاثیر متغیرهای مستقل تحقیق بر روی متغیرهای وابسته بررسی شد. در این راستا تاثیر مقادیر مختلف سه متغیر تعداد ویژگی، تعداد نرون پنهان در شبکه پیشخور و تابع بهینه‌ساز در شبکه عصبی بازگشتی بررسی گردید. نتایج این سه گروه از آزمایش‌ها نشان داد که در تعداد ۷ ویژگی بهترین دانش از پایگاه داده به شبکه‌های عصبی ارائه می‌شود. همچنین تعداد ۵۰ نرون مخفی برای شبکه عصبی پیشخور بدون افزایش پیچیدگی‌های محاسباتی می‌تواند منجر به حصول بهترین نتایج شود و در پایان بهترین تابع بهینه‌ساز برای یادگیری شبکه عصبی بازگشتی این تحقیق تابع Adam است که باعث می‌شود شبکه مرحله یادگیری را در کوتاهترین زمان و با حداقل خطا به پایان برساند. در دسته بعد از آزمایش‌ها در دو پایگاه داده حراج آنلاین نتایج مدل‌های این تحقیق و روش لی و همکاران [۴] مقایسه شد. نتایج این آزمایش‌ها نشان داد که در هر دو پایگاه داده مدل توانسته در مقایسه با روش لی به نتایج بهتری دست یابد و در پایگاه داده اول معیار RMSE را بین ۸ تا ۱۰ و در پایگاه داده دوم همین معیار را بین ۱۹ تا ۲۰ کاهش دهد. نتایج حاصل شده در این فصل نشان داد که کلیه اهداف مطرح شده در این تحقیق حاصل شده‌اند.

نتیجه‌گیری و کارهای آینده:

تجارت الکترونیک باعث شده، فرصت‌های جدیدی برای کسب و کارهای مختلف بوجود آید. حراج‌های آنلاین یکی از پر اهمیت‌ترین فعالیت در تجارت الکترونیک محسوب می‌شوند. قیمت نهایی از جمله فاکتورهای اساسی و موثر در حراج‌های آنلاین است که تخمین دقیق آن می‌تواند برای مصرف‌کننده در بازارهای تجارت آنلاین سود زیادی داشته باشد. رشد روزانه داده‌های مربوط به حراج‌های آنلاین و تشکیل داده‌های حجیم در این زمینه باعث شده استفاده از الگوریتم‌های یادگیری ماشین برای پیش‌بینی قیمت‌های نهایی مورد توجه قرار گیرد. تشکیل داده‌های حجیم و بزرگ از جمله عوامل اصلی برای استفاده از تکنیک‌های یادگیری ماشین به عنوان یک گزینه مناسب و هوشمند است. پیش‌بینی قیمت نهایی در حراج آنلاین بسیار سود آور است اما تاکنون مطالعات محدودی در این خصوص انجام شده است. در بین مطالعات محدود ارائه شده، تکنیک‌های یادگیری ماشین مثل رگرسیون یا k نزدیکترین همسایه و نسخه‌های بهبود یافته آن‌ها بسیار مورد استفاده قرار گرفته است. روش‌های یادگیری ماشین سنتی استفاده شده در روش‌های پیشین غالباً به صورت منفرد هستند و تاکنون مطالعاتی که مبتنی بر ترکیب روش‌های عمیق عصبی و یادگیری ماشین سنتی باشد در این زمینه مطرح نشده است. از اینرو با توجه به اهمیت تخمین قیمت در حراج‌های آنلاین در این تحقیق برای رفع محدودیت‌های مطرح در روش لی و همکاران [۴]، یک مدل ترکیبی مبتنی

- [5] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, pp. 1-41, 2018.
- [6] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70-79, 2018.
- [7] T. Dokeroglu, E. Sevinc, T. Kucukyilmaz, and A. Cosar, "A survey on new generation metaheuristic algorithms," *Computers & Industrial Engineering*, vol. 137, pp. 1-29, ۲۰۱۹.
- [8] X.-S. Yang and A. Slowik, "Firefly algorithm," in *Swarm Intelligence Algorithms*: CRC Press, 2020, pp. 163-174.
- [9] P. Agrawal, H. F. Abutarboush, T. Ganesh, and A. W. Mohamed, "Metaheuristic algorithms on feature selection: A survey of one decade of research (2009-۲۰۱۹)," *International Journal of Data Mining and Analytics*, vol. ۹, no. ۲, pp. ۲۶۷۶۶-۲۶۷۹۱, ۲۰۲۱.
- [10] B. H. Nguyen, B. Xue, and M. Zhang, "A survey on swarm intelligence approaches to feature selection in data mining," *Swarm and Evolutionary Computation*, vol. 54, pp. 1-27, 2020.
- [11] S. Bubeck, N. R. Devanur, Z. Huang, and R. Niazadeh, "Online auctions and multi-scale online learning," in *Proceedings of the 2017 ACM Conference on Economics and Computation*, 2017, pp. 497-514.
- [12] P. Kaur, M. Goyal, and J. Lu, "A price prediction model for online auctions using fuzzy reasoning techniques," in *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2014: IEEE, pp. 1311-1318.
- [13] N. H. Chan and W. W. Liu, "Modeling and forecasting online auction prices: a semiparametric regression analysis," *Journal of Forecasting*, vol. 36, no. 2, pp. 156-۱۶۴, ۲۰۱۷.
- [14] R. Aggarwal and R. Kumar, "Effect of training functions of artificial neural networks (ANN) on time series forecasting," *International Journal of Computer Applications*, vol. 109, no. 3, pp. 1-4, 2015.
- [15] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16-28, 2014.
- [16] J. Miao and L. Niu, "A survey on feature selection," *Procedia Computer Science*, vol. 91, pp. 919-926, 2016.
- [17] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, and V. A. e. Kamaev, "A survey of forecast error measures," *World applied sciences journal*, vol. 24, no. 24, pp. 171-176, 2013.
- [18] N. Johnson *et al.*, *Machine Learning for Materials Developments in Metals Additive Manufacturing*. 2020.
- [19] M. M. Najafabadi, T. M. Khoshgoftaar, F. Villanustre, and J. Holt, "Large-scale distributed l-bfgs," *Journal of Big Data*, vol. 4, no. 1, pp. 1-17, 2017.
- [20] L. Zhou, P. Zhao, D. Wu, C. Cheng, and H. Huang, "Time series model for forecasting the number of new admission inpatients," *BMC medical informatics and decision making*, vol. 18, no. 1, pp. 1-11, 2018.
- [21] J. Grossman, "Predicting-ebay-auction-sales-with-machine-learning," Working Paper, Published on Web. Research <http://jaygrossman.com/post/2013...>, ۲۰۱۳.