

مرور و مقایسه الگوریتم‌های شبکه عصبی بازگشتی عمیق LSTM و GRU در مدل‌سازی داده‌های سری زمانی نرخ ارز

مرضیه یراقی

دانشجوی کارشناسی ارشد، گروه کامپیوتر - دانشگاه آزاد اسلامی واحد دولت‌آباد - اصفهان - ایران
پست الکترونیکی: bahare.yaraghi@gmail.com

اعظم ربیعی*

استادیار، گروه کامپیوتر - دانشگاه آزاد اسلامی واحد دولت‌آباد - اصفهان - ایران
پست الکترونیکی: azrabiee@gmail.com

چکیده

مربعات خطا برای پیش‌بینی تغییرات نرخ ارز ۲۰ درصد کاهش یافت.

واژه‌های کلیدی: شبکه عصبی بازگشتی، پیش‌بینی، داده‌های سری زمانی، نرخ ارز، یادگیری عمیق

این مقاله به مرور و مقایسه دو شبکه عصبی بازگشتی LSTM و GRU در مدل‌سازی داده‌های سری زمانی نرخ ارز می‌پردازد. این دو الگوریتم، ساختار خاصی از شبکه‌های عصبی بازگشتی عمیق هستند که به علت داشتن سلول حافظه و دروازه‌های کنترلی، توانایی کنترل جریان اطلاعات و تعیین زمان بهینه برای به خاطر سپردن و فراموش کردن دارند. در این گزارش پیش‌بینی نرخ دلار آمریکا به ریال ایران با استفاده از شبکه عصبی بازگشتی ارزیابی شد. دادگان مورد آزمایش، نرخ برابری دلار آمریکا به ریال با ۸۶۸۷ رکورد از ابتدای سال ۱۳۷۱ تا آخر سال ۱۳۹۶ است. داده‌های خام نرخ ارز، بین صفر و یک نرمال‌سازی می‌شوند و با معیارهای اندازه‌گیری دقت پیش‌بینی نتایج شبکه عصبی بازگشتی LSTM با شبکه عصبی بازگشتی GRU در کنار یکدیگر مقایسه شدند. هدف اصلی مقایسه عملکرد دو شبکه عصبی بازگشتی LSTM و GRU در پیش‌بینی نرخ ارز می‌باشد. در آزمایش‌های مربوط به شبکه GRU، در مقایسه با LSTM، جذر میانگین

۱- مقدمه

پیش‌بینی سری‌های زمانی یک مسئله مهم کاربردی است. منظور از یک سری زمانی مجموعه‌ای از داده‌های آماری است که در فواصل زمانی مساوی و منظمی جمع‌آوری شده باشند. هدف اصلی در برپا کردن یک سری زمانی، معمولاً پیش‌بینی مقادیر آینده با بهره‌گیری از داده‌های گذشته می‌باشد [۱]. دنباله نرخ ارز در روزهای متوالی، یک نمونه سری زمانی است زیرا در فواصل زمانی مساوی و متوالی جمع‌آوری شده‌اند [۲]. در واقع، نسبت واحد پول ملی بر یک واحد پول خارجی را نرخ ارز می‌نامند. این نرخ یکی از متغیرهای مهم برای سنجش وضعیت اقتصادی هر کشوری حساب می‌شود [۱]. تا امروز سری‌های زمانی موضوع تحقیقات گسترده‌ای

بوده و با روش‌های مختلفی از جمله روش‌های آماری و روش‌های مبتنی بر یادگیری ماشین مدل‌شده‌اند [۳]. ماهیت روش‌های آماری در پیش‌بینی این داده‌های سری زمانی، به گونه‌ای است که در محیط‌هایی با تغییرات کم به خوبی قادر به پیش‌بینی هستند؛ اما در محیط‌های دارای اغتشاش نمی‌توانند تقریب خوبی از تغییرات محیطی را تخمین بزنند [۴]. از این رو، به روش‌های هوشمندانه‌تری نیاز است که قابلیت استخراج الگوهای مخفی در داده‌های سری زمانی، حتی در فواصل بلندمدت را داشته باشد.

شبکه‌های عصبی بازگشتی توانایی بالقوه خوبی برای پیش‌بینی سری‌های زمانی از خود نشان داده‌اند [۵]. شبکه عصبی بازگشتی به دلیل مزایای آن در مقایسه با روش‌های اقتصادسنجی در موضوع‌های متنوعی همچون مدل‌سازی، شناخت الگو و پیش‌بینی سری‌های زمانی استفاده شده و نتایج مفیدی داشته است [۳]. در واقع، شبکه عصبی بازگشتی دارای قابلیت است که آن‌ها را برای مدل‌سازی داده‌های سری زمانی قدرتمند می‌سازد [۴]. اصلی‌ترین ویژگی شبکه عصبی بازگشتی حالت مخفی آن است که اطلاعات یک توالی را ذخیره می‌کند. همچنین حتماً نیاز نیست در هر گام زمانی، یک خروجی و یا حتماً یک ورودی داشته باشد. [۵].

از ضعف‌های این شبکه این است که توانایی مدیریت و یادگیری وابستگی‌های بلندمدت را ندارند. به منظور حل مسئله ناپدید شدن گرادیان و مشکل فراموشی حافظه، ساختار خاصی از شبکه‌های عصبی بازگشتی مانند GRU [۶] و LSTM [۷] با واحدهای فراموشی، پیشنهاد شده است که به علت داشتن سلول حافظه، توانایی کنترل اطلاعات خاص را دارند و بدین ترتیب با در اختیار داشتن دروازه‌های کنترلی، زمان بهینه را برای فراموش کردن اطلاعات تعیین می‌کنند. در این مقاله از شبکه‌های عصبی بازگشتی LSTM و GRU برای استخراج الگوی مناسب برای مدل‌سازی پیش‌بینی نرخ ارز استفاده می‌شود. استخراج الگو از داده‌های سری‌زمانی مربوط به نرخ

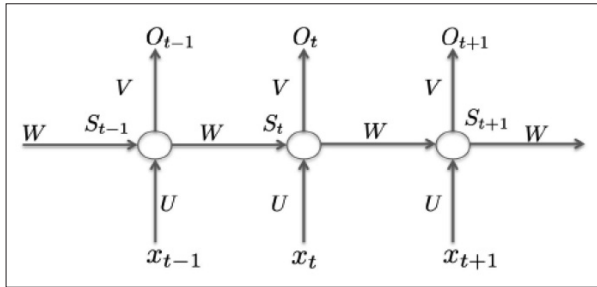
ارز با کمک شبکه‌های عصبی بازگشتی و مقایسه شبکه‌های بازگشتی LSTM و GRU در مدل‌سازی داده‌های سری زمانی از اهداف اصلی این تحقیق می‌باشد. این دو شبکه در پیش‌بینی نرخ ارز با دو معیار جذر میانگین مربعات خطا و میانگین قدر مطلق خطا در این مقاله مقایسه شده‌اند. در ادامه، بخش ۱ به روش تحقیق می‌پردازد. در این بخش، دو الگوریتم مد نظر این تحقیق، با ذکر جزئیات، معرفی شده‌اند. بخش ۲، مراحل پیاده‌سازی شامل مدل طراحی شده، معماری شبکه‌ها برای مقایسه، نوع و ساختار داده‌ها و معیارهای ارزیابی را شرح می‌دهد. نتایج آزمایش‌ها در بخش ۴ و نهایتاً جمع‌بندی و نتیجه‌گیری در بخش ۵ ذکر شده‌اند.

۲- روش تحقیق

مراحل انجام تحقیق در شکل (۱) نشان داده شده است. در این مطالعه با استفاده از داده‌های مربوط به روزهای گذشته، هفته یا ماه، نوسان قیمت در روزهای آتی پیش‌بینی می‌گردد. بدین منظور نرخ ارز دلار آمریکا مربوط به گذشته از پایگاه اطلاع‌رسانی بانک مرکزی جمع‌آوری و تحلیل و بررسی روی آن‌ها انجام شد. جزئیات بیشتر از پایگاه داده در بخش ۲-۳ شرح داده شده است.

داده‌های مدنظر این تحقیق، پیش‌پردازش و نرمال‌سازی شده‌اند؛ سپس با استخراج الگوی مناسب مبتنی بر شبکه‌های عصبی بازگشتی سعی بر پیش‌بینی نرخ ارز صورت گرفت. طبق شکل (۱)، پس از جمع‌آوری و نرمال‌سازی داده‌ها، جهت ارزیابی نهایی روش‌های به‌کار گرفته شده در این تحقیق، داده‌ها به دو دسته آموزش و آزمایش تقسیم شدند. آموزش شبکه‌های عصبی با دسته اول انجام شد. داده‌های آزمایش، نمونه‌های جدید و دیده‌نشده برای ارزیابی روش‌ها هستند.

به‌طور کلی، شبکه‌های عصبی بازگشتی در بخش ۲-۱ معرفی شده‌اند. سپس بخش‌های ۲-۲ و ۲-۳ به معرفی دو شبکه عصبی بازگشتی مد نظر این تحقیق، یعنی LSTM و



شکل ۲: شبکه عصبی بازگشتی مانند دنباله‌ای از شبکه‌های عصبی پیشخور است که در آن خروجی هر مرحله زمانی به عنوان ورودی به مرحله بعدی وارد می‌شود [۶].

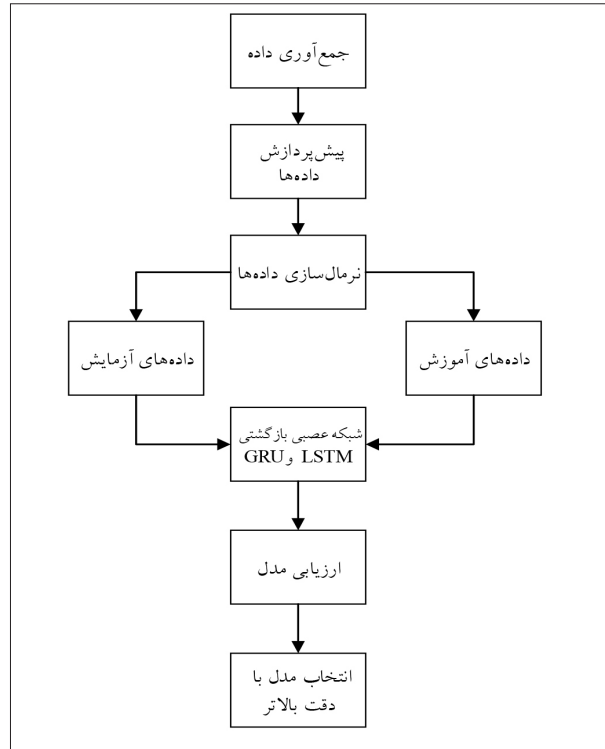
لایه‌های ماقبل آن وابسته است. در واقع، این شبکه‌ها دارای حافظه هستند که اطلاعات مربوط به داده‌های دیده شده را ذخیره می‌کند. این شبکه‌ها را می‌توان به صورت چندین نسخه یکسان از یک شبکه عصبی در نظر گرفت که هر کدام اطلاعاتش را به شبکه بعدی منتقل می‌کند. شکل (۲) وضعیت شبکه عصبی بازگشتی در صورت باز کردن حلقه‌ی بازگشتی را نشان می‌دهد.

۲-۲ شبکه عصبی بازگشتی LSTM

شبکه عصبی بازگشتی LSTM برای اولین بار توسط هوچریترو و اسمیدوبر سال ۱۹۹۷ برای توسعه شبکه‌های عصبی بازگشتی معرفی شدند [۷]. البته تعداد زیادی از محققان در بهبود این شبکه‌ها نقش داشتند. شبکه‌های عصبی بازگشتی ساده ابزاری بسیار قوی محسوب می‌شوند اما همان‌طور که اشاره شد از برخی مشکلات اساسی مانند محو گرادیان^۱ و عدم یادگیری توالی‌های بلند رنج می‌برند.

شبکه بازگشتی ساده در به خاطر سپردن توالی‌های بلند با مشکل مواجه است و قادر به نگهداری اطلاعات به مدت طولانی نیست [۹]. این مشکلات سبب شده است تا در اکثر پژوهش‌های اخیر از نسخه‌های توسعه‌یافته این شبکه استفاده شود. بنابراین شبکه‌های LSTM، نوع خاصی از شبکه‌های عصبی بازگشتی هستند که توانایی یادگیری وابستگی‌های بلندمدت را دارند [۱۰].

شکل (۳) یک نمونه از بلوک LSTM با یک سلول را



GRU می‌پردازند. نهایتاً مقایسه نظری این دو مدل در بخش ۲-۴ شرح داده شده است.

۲-۱ معرفی شبکه عصبی بازگشتی

شبکه عصبی بازگشتی معروف‌ترین مدل برای آموزش داده‌های متوالی است. شبکه عصبی بازگشتی گسترش یافته شبکه عصبی پیشخور می‌باشد. برخلاف شبکه عصبی پیشخور، شبکه عصبی بازگشتی دارای اتصالات چرخه‌ای است که آن‌ها را برای مدل‌سازی متوالی قدرتمند می‌سازد [۸]. این شبکه‌ها شامل یه حلقه بازگشتی هستند که منجر می‌شود اطلاعاتی که از لحظات قبلی به دست آمده، از بین نروند و در شبکه باقی بمانند [۶].

در تئوری این‌طور به نظر می‌رسد که این شبکه‌ها می‌توانند اطلاعات موجود در یک دنباله طولانی را ضبط و از آن‌ها استفاده کنند اما در عمل این‌طور نیست و بسیار محدود هستند، به این صورت که فقط اطلاعات چند گام قبل را ضبط می‌کنند.

به عبارت دیگر، شبکه‌های عصبی بازگشتی به این علت بازگشتی نامیده می‌شوند که خروجی هر لایه به محاسبات

1- Vanishing Gradient

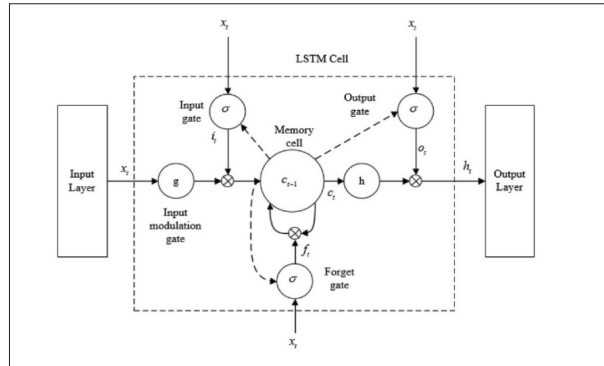
هنگام محاسبه حالت سلولی، این نسبت بر معادله (۴) تاثیر می‌گذارد. دروازه فراموشی حافظه قبل h_{t-1} را منتقل می‌کند. نسبت فراموشی در معادله (۲) محاسبه می‌شود که برای معادله (۴) استفاده می‌شود. دروازه خروجی تعیین می‌کند که آیا خروجی از سلول حافظه منتقل شود یا خیر. معادله‌های (۳) و (۵) این روند را نشان می‌دهند [۱۱]. با استفاده از شبکه LSTM، می‌توان با توجه به سه دروازه مشکلات محو گرادیان و گرادیان انفجاری ۲ حل کرد.

۲-۳ شبکه عصبی بازگشتی GRU

شبکه عصبی بازگشتی GRU توسط چو و همکاران در سال ۲۰۱۴ پیشنهاد شده است [۱۲]. این شبکه، نوع خاصی از شبکه‌های عصبی بازگشتی هستند که توانایی یادگیری وابستگی‌های بلندمدت را دارند. شبکه عصبی بازگشتی GRU با تنظیم انواع مختلف واحدهای چرخه، مشکل وابستگی به انواع مقیاس‌های زمانی را حل کرده است. شبکه عصبی بازگشتی GRU مانند سازوکار LSTM عمل می‌کند و فرق آن این است که به جای سه دروازه از دو دروازه استفاده می‌کند (شکل ۴) که در نتیجه از نظر سرعت بهبود نسبتاً خوبی خواهد داشت [۱۳].

همان‌طور که در شکل (۴) مشاهده می‌شود شبکه عصبی GRU دارای دو دروازه به‌روزرسانی و بازنشانی است. دروازه به‌روزرسانی تصمیم می‌گیرد تا چه مقداری واحد عصبی به روزرسانی شود و دروازه بازنشانی تصمیم می‌گیرد که چه مقدار از واحد عصبی قبلی را فراموش کند h_{t-1} آخرین حالت و ورودی است، W و U ماتریس‌های وزن برای اتصالات دروازه هستند. b بردار بایاس است که در حین آموزش مقداردهی می‌شوند [۱۴].

ورودی مدل به صورت دنباله $x=(x_1, x_2, \dots, x_t)$ بیان می‌شود بنابراین x_t بردار ورودی و h_t بردار خروجی مربوط به زمان t است، که طبق رابطه (۶) محاسبه می‌شود. کل فرایندی که در یک مرحله زمانی در یک سلول انجام می‌شود، محاسبه جمع کردن حالت موجود



شکل ۳: بلوک LSTM با یک سلول [۱۱]

نشان می‌دهد یک شبکه LSTM مشابه یک شبکه عصبی بازگشتی استاندارد است، با این تفاوت که واحدهای جمع‌کننده نرون‌ها در لایه مخفی توسط بلوک‌های حافظه جایگزین می‌شوند.

در شکل (۳) دروازه‌های ضربی به سلول حافظه LSTM امکان نگهداری و دسترسی به اطلاعات در دوره‌های زمانی بلند مدت را می‌دهد. بنابراین مشکل محو شدن گرادیان کاهش می‌یابد. به عنوان مثال تا زمانی که دروازه ورودی بسته باقی بماند، تابع فعالیت سلول توسط ورودی‌های جدید که به شبکه می‌رسند باز نوشته نخواهد شد؛ در نتیجه با باز کردن دروازه خروجی، در لحظات طولانی مدت آینده دنباله، برای شبکه قابل دسترس است.

در شکل (۳)، نمادهای c, o, f, i و h به ترتیب معرف ورودی، فراموشی، خروجی، حافظه و حالت سلول هستند. روابط زیر در هر سلول برقرار است:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (۱)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (۲)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (۳)$$

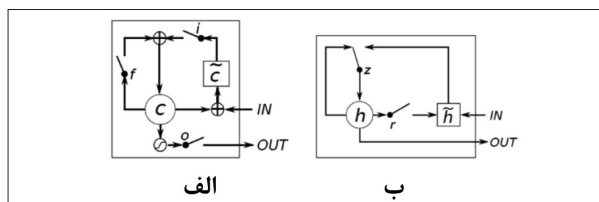
$$c_t = f_t c_{t-1} + i_t \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (۴)$$

$$h_t = o_t \tanh(c_t) \quad (۵)$$

در روابط فوق، σ معرف تابع سیگموند است. x_t و بردار ورودی و W و U ماتریس‌های وزن برای اتصالات دروازه مشخص شده‌اند. همان‌طور که قبلاً اشاره شد هر سلول LSTM، سه دروازه جریان اطلاعات را کنترل می‌کنند. دروازه ورودی، نسبت ورودی را تعیین می‌کند.

جدول ۱: مقایسه دو شبکه عصبی بازگشتی LSTM و GRU

GRU	LSTM	معیار ارزیابی
کل حافظه در شبکه عرضه می‌گردد.	مقدار حافظه عرضه شده توسط واحدهای دیگر شبکه بوسیله دروازه خروجی کنترل می‌شود.	عرضه حافظه کنترل شده
جریان اطلاعات را از فعال سازی قبلی کنترل می‌کند.	هیچ کنترل جداگانه‌ای برای میزان جریان اطلاعات از مرحله زمان قبلی وجود ندارد.	محاسبه حافظه جدید
پارامترهای کمتری دارد و بنابراین سریعتر آموزش داده می‌شود.	با یک دروازه اضافی پیچیدگی بالاتری دارد	پیچیدگی نسبت به عملکرد
دارای دو دروازه به‌روزرسانی و بازنشانی است.	دارای سه دروازه ورودی، خروجی و فراموشی است.	تعداد دروازه



شکل ۵: شکل الف نشان‌دهنده شبکه عصبی بازگشتی LSTM و شکل ب نشان‌دهنده شبکه عصبی بازگشتی GRU [۶]

کند، بلکه به طور خودکار تعادل زمان مطلوب برای پیش‌بینی را نیز تعیین می‌کند. در سال‌های اخیر، GRU با موفقیت در شناسایی دست خط، شناسایی حرکت انسان و کنترل ربات و غیره مورد استفاده قرار گرفته است [۱۷].

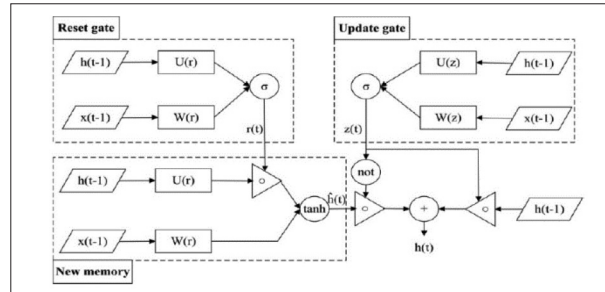
۲-۴- مقایسه شبکه‌های عصبی بازگشتی LSTM و GRU

تفاوت‌های میان این دو شبکه به طور خلاصه در جدول ۱ ارائه شده است [۳].

برای شفافیت بیشتر و مقایسه بهتر بین شبکه‌ها در شکل (۵) هر دو شبکه کنار یکدیگر رسم شده‌اند.

در شکل (۵) الف، نمادهای f ، i و o به ترتیب دروازه‌های ورودی، فراموشی و خروجی هستند و C و \tilde{C} به ترتیب محتوای سلول حافظه و سلول حافظه جدید را مشخص می‌کنند. در شکل (۵) ب، r و z به ترتیب دروازه‌های بازنشانی و به‌روزشده هستند و h و \tilde{h} به ترتیب تابع فعال‌سازی و تابع فعال‌سازی منتخب هستند.

اگرچه باهدانو و همکاران در سال ۲۰۱۴ نشان دادند



شکل ۴: منطق محاسباتی Gated Recurrent Unit [۴]

و همچنین محاسبه حالت به‌روزرسانی است. مدل GRU در هر محاسبه یک بار تمام حالت‌ها را به روز می‌کند زیرا نمی‌تواند طیف وسیعی از به‌روزرسانی‌ها را در هر حالت کنترل کند [۱۵].

$$h_t = (1 - z_t)h_{t-1} + z_t\tilde{h}_t \quad (6)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h(r_t h_{t-1}) + b_h) \quad (7)$$

در همان زمان، بردار دروازه به‌روزرسانی z_t تصمیم می‌گیرد تا چه میزان از محتوای واحد فعال را به‌روزرسانی کند. بردار z_t به صورت زیر محاسبه می‌شود:

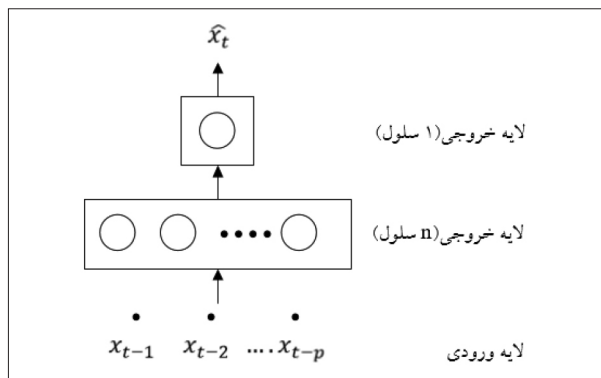
$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (8)$$

بردار r_t ، دروازه بازنشانی است هنگامی که دروازه بسته است، اجازه می‌دهد که سلول، گذشته را فراموش کند. پس در نتیجه هرکدام از بردارهای z و r یک بردار نرمال شده هستند که مقادیر آن‌ها بین صفر و یک است. مقدار صفر یعنی هیچ مقداری عبور نمی‌دهد و مقدار یک یعنی کل مقدار را عبور می‌دهد. بردار طبق رابطه (۹) زیر محاسبه می‌شود:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (9)$$

دروازه به‌روزرسان، کنترل می‌کند که چه میزان از حالت گذشته، مهم هستند. واحدهایی با وابستگی‌های کوتاه مدت با دروازه بازنشانی فعال خواهند شد و واحدهایی با وابستگی‌های بلند مدت با دروازه به‌روزرسان فعال خواهند شد [۱۶].

شبکه عصبی بازگشتی GRU می‌تواند از واحد حافظه در یک شبکه برای مقابله با هر توالی داده ورودی استفاده کند؛ بنابراین، توانایی یادگیری سری‌های زمانی را دارد. GRU نه تنها می‌تواند سری‌های زمانی طولانی را مرور



شکل ۶: معماری شبکه عصبی بازگشتی

مقدار آینده می‌باشد. x_2, x_1, \dots, x_T کل مشاهدات از نرخ ارز در مجموعه آموزشی است. هر نمونه ورودی، از یک پنجره متحرک با طول ثابت p تشکیل شده است. در این تحقیق از یک شبکه با یک لایه ورودی، یک لایه مخفی و یک لایه خروجی استفاده شد (شکل ۶). اولین نمونه آموزشی از x_2, x_1, \dots, x_p به عنوان ورودی و x_{p+1} به عنوان خروجی تشکیل می‌شود و در نمونه آموزش دوم x_3, x_2, \dots, x_{p+1} به عنوان ورودی و x_{p+2} به عنوان هدف خروجی در نظر گرفته می‌شود؛ بنابراین $T-p$ نمونه آموزشی در کل فرایند آموزش وجود دارد و هدف آموزش شبکه عصبی این است که مناسب‌ترین وزن‌ها (پارامترهای اتصالات شبکه) را به گونه‌ای پیدا کند که معیارهای ارزیابی بهینه شود.

شکل ۶ نمایی از ساختار شبکه عصبی بازگشتی این تحقیق می‌باشد. این مدل یک شبکه عصبی عمیق است. ساختار این مدل دارای سه لایه است که شامل یک لایه ورودی، یک لایه مخفی و یک لایه خروجی می‌باشد. لایه ورودی لایه‌ای است که ورودی را دریافت می‌کند و لایه اول شبکه است. لایه خروجی لایه‌ای است که خروجی را تولید می‌کند، یا همان لایه پایانی شبکه است. لایه‌هایی که کار پردازش را انجام می‌دهند، لایه‌های مخفی در شبکه هستند. این لایه‌ها عملیات خاصی را بر روی داده‌هایی که به آن‌ها ارسال شده انجام می‌دهند و خروجی تولید شده خود را به لایه‌های بعدی ارسال می‌کنند. فقط لایه‌های ورودی و خروجی قابل دیدن هستند و لایه‌های میانی مخفی هستند. هر لایه شامل چندین سلول است و هر سلول در هر لایه،

که این دو نوع شبکه، با توجه به آزمایش‌های اولیه خود در ترجمه ماشین مشابه یکدیگر رفتار می‌کنند، ولی معلوم نیست که آیا این امر به کاربردهایی غیر از ترجمه ماشین نیز قابل تعمیم است یا خیر. این مسئله باعث شد که در این تحقیق، مقایسه‌ی جامع‌تری بین واحد LSTM و GRU در مورد پیش‌بینی سری زمانی نرخ ارز انجام شود.

۳- روش پیشنهادی

در این تحقیق، با در اختیار داشتن نرخ واقعی ارز گذشته $(x_1, x_2, \dots, x_{t-1})$ و با طراحی یک شبکه عصبی بازگشتی به پیش‌بینی نرخ ارز آینده (\hat{x}_t) پرداخته می‌شود. از آنجا که مقادیر واقعی نرخ ارز در دسترس هستند، یادگیری به صورت با سرپرست انجام می‌شود. بازه‌های زمانی می‌توانند به صورت روزانه، هفتگی و ماهانه در نظر گرفته شوند و \hat{x}_t از طریق فرمول زیر محاسبه می‌شود که در آن $1 \leq t \leq T$ است:

$$\hat{x}_t = RNN([x_{t-p}, x_{t-p+1}, \dots, x_{t-1}]) \quad (10)$$

در مرحله اول جمع‌آوری داده‌های نرخ ارز به صورت خام انجام گرفت؛ سپس داده‌ها پاک‌سازی شدند و بعد از انتخاب داده‌های مناسب به عنوان متغیر ورودی، داده‌ها با نسبت ۸۰ به ۲۰ به دو بخش داده‌های آموزش و داده‌های آزمون تقسیم شد. نهایتاً شبکه‌های عصبی بازگشتی GRU و LSTM با کمک زبان برنامه‌نویسی پایتون پیاده‌سازی شد. یادگیری این شبکه‌ها روی داده‌های آموزش، انجام شد و نتایج پیش‌بینی روی داده‌های آزمون در هر دو شبکه با یکدیگر مقایسه شدند. همچنین معیارهای اندازه‌گیری جذر میانگین مربعات خطا و میانگین قدر مطلق خطا به منظور بررسی اثر بخشی شبکه عصبی بازگشتی پیشنهادی در پیش‌بینی نرخ ارز ارزیابی شدند.

۳-۱ ساختار شبکه عصبی طراحی شده

پیش‌بینی نرخ ارز یک پیش‌بینی سری زمانی است. داده‌های ورودی این شبکه، نرخ واقعی ارز در گذشته به صورت روزانه، هفتگی و ماهانه هستند و خروجی آن

جدول ۲: مقدار نرخ ارز دلار آمریکا به ریال در تاریخ ماهانه

تاریخ	مقدار نرخ ارز دلار آمریکا به ریال
۱۳۷۱/۱/۵	۱۴۳۹
۱۳۷۱/۲/۵	۱۴۵۲
۱۳۷۱/۳/۵	۱۴۳۹
۱۳۷۱/۴/۵	۱۴۳۹

LSTM و GRU پیاده‌سازی و نتایج هر دو شبکه با یکدیگر مقایسه و ارزیابی می‌شوند.

در بخش بعدی، از نظر دقت پیش‌بینی مقایسه شدند. هدف اصلی مقایسه عملکرد دو شبکه عصبی بازگشتی LSTM و GRU با اندازه‌گیری معیار جذر میانگین مربعات خطاها در پیش‌بینی داده‌های سری زمانی نرخ ارز می‌باشد.

۳-۳ معیار ارزیابی

رایج‌ترین راه برای اندازه‌گیری دقت مدل‌های پیش‌بینی، اندازه‌گیری معیار جذر میانگین مربعات خطاهای پیش‌بینی است. حال هر مدلی که جذر میانگین مربعات خطای کمتری را به خود اختصاص دهد به‌عنوان مدل برتر پیش‌بینی انتخاب می‌گردد. این معیار به‌صورت زیر محاسبه می‌شود:

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (x_t - \hat{x}_t)^2}{T}} \quad (11)$$

که در آن، x_t و \hat{x}_t به ترتیب مقادیر واقعی و پیش‌بینی شده نرخ ارز در زمان t هستند. علاوه بر معیار فوق، معیار میانگین قدر مطلق خطا نیز در این تحقیق محاسبه شد. مزیت استفاده از این شاخص این است که وابسته به مقیاس نیست و امکان مقایسه قدرت پیش‌بینی را برای سری‌هایی که مقیاس متفاوت دارند نیز فراهم می‌کند. این شاخص به‌صورت زیر تعریف می‌شود:

$$MAE = \frac{\sum_{t=1}^T \left| \frac{x_t - \hat{x}_t}{x_t} \right|}{T} \quad (12)$$

که در این رابطه، T تعداد داده‌های ورودی برای انتخاب فرایند پیش‌بینی هستند. برای هر الگوریتم پیشنهادی معیارهای اندازه‌گیری جذر میانگین مربعات خطا و میانگین قدر مطلق خطا به‌دست آورده و در یک جدول ثبت می‌شوند.

۴- نتایج آزمایش‌ها

همان‌طور که در بخش ۳-۲ ذکر شد، پایگاه داده مورد نظر نرخ ارز دلار آمریکا به ریال ایران از سال ۱۳۷۱ تا آخر سال ۱۳۹۶ به‌صورت روزانه می‌باشد. در شکل ۷ تغییرات

به تمام سلول لایه بعد متصل است. در این تحقیق لایه ورودی دارای یک متغیر ورودی شامل نرخ واقعی ارز در گذشته $(x_{t-1}, x_{t-2}, \dots, x_{t-p})$ می‌باشد. و لایه خروجی متغیر مشخصی برای مقادیر واقعی نرخ ارز حاضر (\hat{x}_t) است. لایه مخفی شامل n سلول است، که با توجه به نوع داده‌های ورودی تعداد سلول متفاوت می‌باشد. معماری داخلی لایه مخفی با توجه به نوع شبکه انتخابی، متفاوت هستند، لایه مخفی به ترتیب شامل چهار لایه در LSTM و سه لایه در GRU هستند و از آنجا که شبکه با بیش از دو لایه عمیق خوانده می‌شود بنابراین دو شبکه عصبی بازگشتی LSTM و GRU استفاده شده در این تحقیق عمیق به حساب می‌آیند [۱۸].

۳-۲ مجموعه داده‌ها

در این مطالعه تمام داده‌های نرخ ارز از پایگاه داده "Pacific ExchangeRateService" و پایگاه اطلاع رسانی بانک مرکزی بارگیری شده‌اند. وبگاه <http://fx.sauder.ubc.ca/data.html> راه دستیابی به پایگاه داده مذکور است. نرخ ارز دلار آمریکا به ریال با ۸۶۸۷ رکورد از سال ۱۳۷۱ تا پایان سال ۱۳۹۶ جمع‌آوری شد. داده‌های جمع‌آوری شده در سه بخش روزانه، هفتگی و ماهانه تقسیم‌بندی شدند. جدول ۲ با دو ستون تاریخ و مقدار نرخ ارز، چند رکورد از این داده را به‌عنوان نمونه نشان می‌دهد. داده‌های خام نرخ ارز با استفاده از کتابخانه MinMaxScaler نرمال‌سازی شدند تا بین ۰ و ۱ قرار بگیرند.

در ادامه، شبکه عصبی بازگشتی منتخب با استفاده از پایتون پیاده‌سازی شد و با افزودن کتابخانه‌های دیگری از جمله، mean_squared_error، mean_absolute_error،

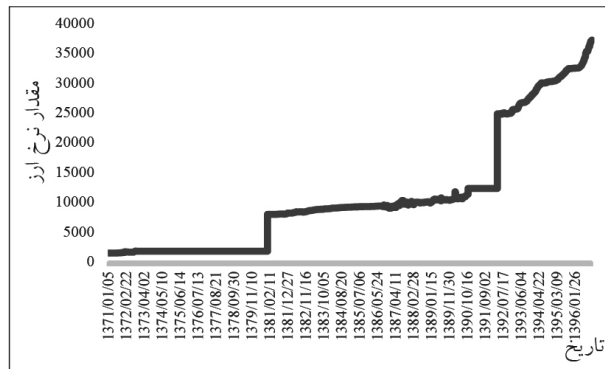
جدول ۳: ارزیابی نتایج طول پنجره برای نرخ ارز دلار آمریکا به ریال ایران (میانگین با ۹۵ درصد بازه اطمینان)

طول پنجره (p)	روزانه	هفتگی	ماهانه
۳	۲±۳۸	۱۰۰±۵	۱۷۰±۲
۵	۳۰±۱	۸۱±۲	۲۲۰±۳
۱۰	۱۸۶±۳	۲۱۴±۲	۲۳۰
۲۰	۲۰۴±۲	۱۶۸±۳	۵۹۸±۵

نشود. نتایج حاصل از این آزمایش، در جدول ۳ ثبت شده است. رکوردی که کمترین خطا را داشته باشد (مقادیر پررنگ در جدول)، به عنوان طول بهینه انتخاب می‌شود. لازم به ذکر است برای دقت بیشتر فرایند محاسبات طول پنجره در سطح اطمینان ۹۵ درصد با ده بار اجرا صورت گرفته است. نتایج نشان می‌دهد که برای این پایگاه داده، در پیش‌بینی روزانه و هفتگی، طول پنجره ۵ بهینه است و برای نرخ ماهانه عدد ۳ بهتر عمل کرده است.

۴-۲ ارزیابی نتایج LSTM و GRU

هر دو الگوریتم GRU و LSTM با تعداد سلول ۴، ۸، ۱۶ اجرا شدند و معیارهای اندازه‌گیری خطا برای هر دو شبکه محاسبه و برای مقایسه این دو شبکه نمودارهای میله‌ای برای هر دو خطای جذر میانگین مربعات خطا و میانگین قدر مطلق خطا رسم شد. درصد داده‌های آموزش به آزمایش در همه حالت‌ها ۸۰ درصد گرفته شد. شکل ۸ و ۹ خروجی آزمایش‌های مختلف با تعداد سلول متفاوت را به ترتیب برای جذر میانگین مربعات خطا و میانگین قدر مطلق خطا نشان می‌دهد. این نمودارها دارای سه بخش روزانه، هفتگی و ماهانه است. در هر شکل محور X تعداد سلول آزمایش را در نرخ‌های روزانه، هفتگی و ماهانه نشان می‌دهد و محور Y نشان‌دهنده مقدار خطاها می‌باشد. همان‌طور که از نمودار مشخص است در نرخ روزانه مقدار خطاها برای هر دو شبکه با افزایش تعداد سلول کاهش یافته است. این افزایش سلول تا جایی که خطا به طور ناگهانی افزایش پیدا کرد، ادامه داده شد. در نهایت در ۶۴ سلول به علت افزایش خطا، آزمایش‌ها متوقف شد و



شکل ۷: منحنی تغییرات روزانه نرخ ارز دلار آمریکا به ریال ایران از سال ۱۳۷۱ تا سال ۱۳۹۶

این نرخ نشان داده شده است. جهش‌های نمودار بیانگر این است که در سال‌های ۱۳۸۰ و ۱۳۹۲ نرخ ارز دچار افزایش ناگهانی قیمت بوده و در مابقی موارد با سیر صعودی یکنواختی همراه بوده است که البته این شیب صعودی در آخر نمودار بیشتر شده است.

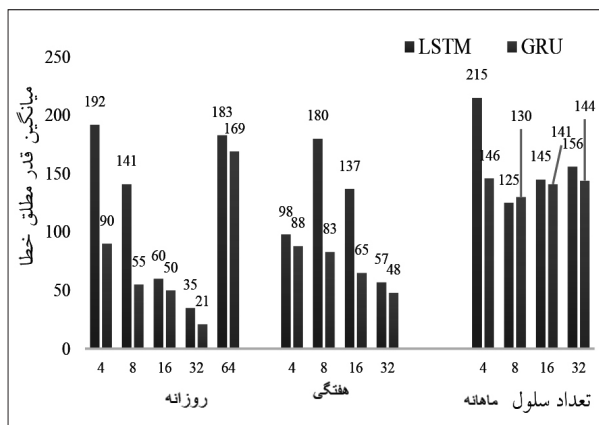
برای بررسی دقیق‌تر داده‌ها به سه دسته نرخ روزانه، هفتگی و ماهانه دسته‌بندی شدند و روی هر دسته جداگانه اجراها آزمایش و نتایج خطاها بررسی شد. در تمام آزمایش‌ها این تحقیق، از اعتبارسنجی متقاطع^۳ جهت پیش‌گیری از بیش‌برازش^۴ استفاده شده است. در ادامه، تعیین طول پنجره بهینه با آزمون و خطا برای هر سه دسته نرخ روزانه، هفتگی و ماهانه در بخش ۴-۱ شرح داده شده است. نهایتاً نتایج آزمایش‌های تجربی به‌طور مفصل در بخش ۲-۳ ذکر شده و دو الگوریتم با یکدیگر مقایسه شده‌اند.

۴-۱ محاسبه طول پنجره بهینه

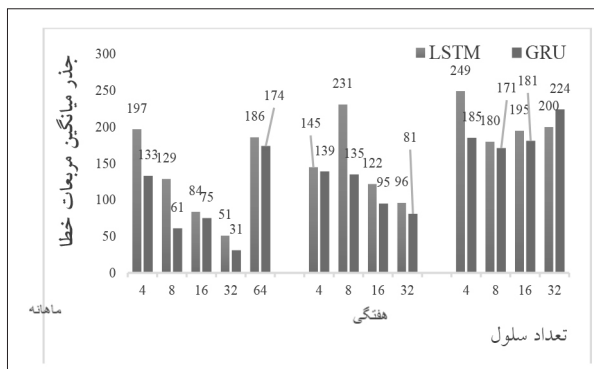
مرحله اول به دست آوردن طول پنجره بهینه (p) است. به علت نابرابری تعداد رکوردها در نرخ‌های روزانه، هفتگی و ماهانه، طول پنجره نیز متغیر است. برای به دست آوردن طول پنجره بهینه جذر میانگین مربعات خطای الگوریتم GRU روی داده‌های روزانه، هفتگی و ماهانه با طول‌های ۳، ۵، ۱۰ و ۲۰ محاسبه شدند. به علت افزایش ناگهانی خطا، افزایش طول پنجره بیشتر از ۲۰ ادامه داده

3- cross-validation

4- overfitting



شکل ۹: نمودار میله‌ای میانگین قدر مطلق خطا بر تعداد سلول برای نرخ‌های ارز روزانه، هفتگی و ماهانه دلار آمریکا در برابر ریال ایران



شکل ۸: نمودار میله‌ای جذر میانگین مربعات خطا بر تعداد سلول برای نرخ‌های ارز روزانه، هفتگی و ماهانه دلار آمریکا در برابر ریال ایران

برای بقیه داده‌های هفتگی و ماهانه، آزمایش تا ۳۲ سلول انجام شد.

نتایج داده‌های هفتگی برای تعداد ۴، ۸، ۱۶ و ۳۲ سلول مخفی آزمایش و در ستون دوم رسم شد. به همین ترتیب، در ستون سوم نتایج خطاهای داده‌های ماهانه مشاهده می‌شود. مقدار خطاهای مشاهده شده در نمودارها نشان می‌دهد نرخ روزانه و هفتگی با افزایش تعداد سلول تا ۳۲ در هر دو شبکه با مقدار خطای کمتری مواجه شده‌اند و بهبود یافته‌اند.

در نرخ ماهانه نیز برای تعداد ۴، ۸، ۱۶ و ۳۲ سلول الگوریتم‌ها آزمایش شدند. از نتایج می‌توان استنباط کرد که در داده‌های ماهانه به علت کم شدن تعداد رکوردها، مقدار خطاها با تعداد سلول کمتر کاهش یافته است، این امر به علت کاهش تعداد رکورد داده‌ها رخ داده است که منجر می‌شود آزمایش با افزایش سلول با سیر صعودی خطا مواجه شود. در همه مراحل برتری نسبی شبکه بازگشتی GRU مشخص است.

با توجه به نمودارها میزان خطا در نرخ ارز روزانه کمتر از دو نرخ هفتگی و ماهانه است که این نتیجه با توجه به نوع داده‌های موجود که سری زمانی هستند، صحیح به نظر می‌رسد. بنابراین ساختار این شبکه‌ها به صورتی است که اطلاعات خیلی دور را به خوبی یاد می‌گیرند.

از شکل‌های (۸) و (۹) استنباط شد که در نرخ روزانه در ۳۲ سلول با کمترین خطا تغییرات نرخ ارز پیش‌بینی

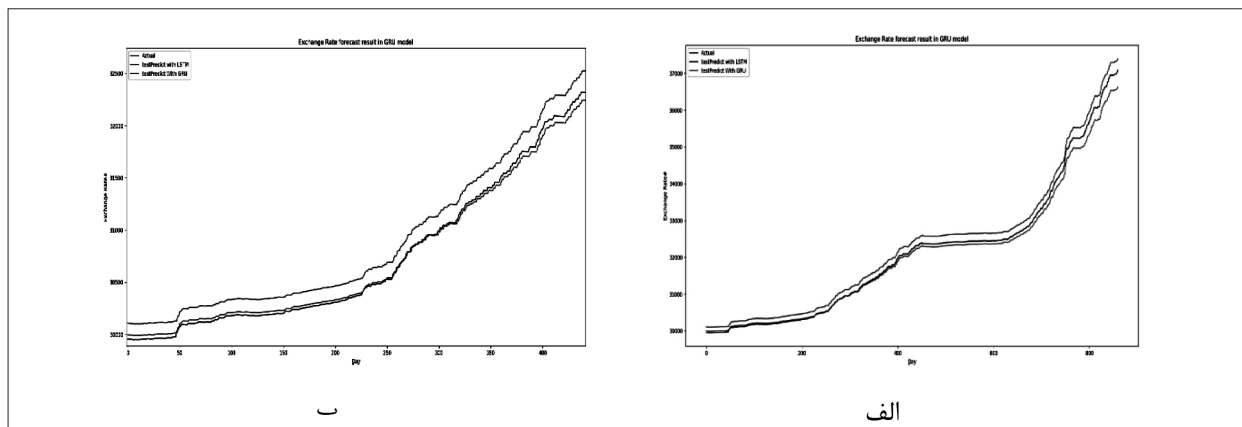
جدول ۴: ارزیابی نتایج خطاها در سطح اطمینان ۹۵ درصد

معیار مقایسه	جذر میانگین مربعات خطا	میانگین قدر مطلق خطا
تعداد سلول	۳۲	۳۲
شبکه عصبی بازگشتی LSTM	$51/28 \pm 1/6$	$34/55 \pm 0/6$
شبکه عصبی بازگشتی GRU	$33/48 \pm 1/2$	$22/24 \pm 0/57$

شدند. بنابراین جدول (۳) نتایج جذر میانگین مربعات خطا و میانگین قدر مطلق خطا را برای ۳۲ سلول در سطح اطمینان ۹۵ درصد برای اجراهای شبکه عصبی بازگشتی LSTM و GRU نشان می‌دهد. طبق مقادیر عددی جدول (۴)، جذر میانگین مربعات خطا در شبکه GRU ۲۱ درصد و میانگین قدر مطلق خطا ۲۰ درصد کاهش یافته و این برتری شبکه GRU در این آزمایش‌ها را نشان می‌دهد.

برای مقایسه بهتر پیش‌بینی‌های انجام شده روی داده‌های آزمایشی نتایج به صورت نمودار نیز در شکل (۱۰) رسم شده و با مقدار نرخ ارز واقعی مقایسه شده‌اند. محور X شاخص روز و محور Y مقدار نرخ را نشان می‌دهد. در نمودار رسم شده شکل (۱۰)، منحنی نرخ واقعی با نمودار آبی رنگ مشخص شده است. رنگ قرمز پیش‌بینی‌های مجموعه داده‌های آزمایش با شبکه عصبی بازگشتی LSTM و رنگ سبز با شبکه عصبی بازگشتی GRU را نشان می‌دهد.

شکل (۱۰) ب، برتری نسبی شبکه عصبی بازگشتی GRU را در یک نمای بزرگ‌نمایی شده نشان می‌دهد.



شکل ۱۰: الف) منحنی پیش‌بینی تغییرات نرخ ارز روزانه دلار آمریکا به ریال ایران با دو شبکه عصبی بازگشتی LSTM و GRU که روی داده‌های آزمایشی انجام شده است (ب) نمای بزرگ‌نمایی شده

دست یافت که هر دو شبکه LSTM و GRU تغییرات نرخ روزانه را با کمترین مقدار خطا پیش‌بینی کردند و ثابت کرد این شبکه‌ها توانایی آموزش و مدیریت وابستگی‌های بلند مدت سری زمانی را دارند.

علاوه بر این، عملکرد و کارایی محاسباتی GRU با LSTM مقایسه شد و نتایج نشان داد شبکه عصبی بازگشتی GRU دقیق‌تر عمل می‌کند؛ زیرا این شبکه در مقایسه با LSTM بهتر آموزش دیده است و عملکرد بهتری در کاهش مقدار خطاها داشت. در واقع، با استفاده از این شبکه در مقایسه با LSTM جذر میانگین مربعات خطا و میانگین قدر مطلق خطا برای پیش‌بینی تغییرات نرخ ارز به مقدار قابل ملاحظه‌ای کاهش یافت.

این تحقیق فقط از رفتار گذشته نرخ ارز برای پیش‌بینی روند آتی آن استفاده کرد. در حالی که پارامترهای متعددی در تعیین نرخ ارز در هر کشوری تاثیرگذار هستند که نقش و میزان اهمیت هر یک در اقتصاد متفاوت هستند. پیشنهاد می‌شود در آینده از تاثیر این پارامترهای موثر مانند تاریخ، صادرات، و واردات، قیمت و میزان عرضه نفت، قیمت طلا و نقش مصرفی یا سرمایه آن، عوامل سیاسی، سیاست‌های اقتصادی و بحران اقتصادی و تحریم‌ها بر قیمت نرخ ارز نیز بررسی شود.

همچنین، در تحقیق بعدی، سعی بر این است که در یک ساختار شبکه عصبی متمرکز، با بهره‌گیری از الگوهای

همان‌طور که مشخص است منحنی سبز رنگ با منحنی آبی رنگ که مقادیر حقیقی را نشان می‌دهد، نزدیکی بیشتری دارد که این نشان‌دهنده برتری پیش‌بینی شبکه GRU است. در شکل (۱۰) الف، اوایل آزمایش منحنی سبز رنگ با دقت بالایی پیش‌رفته است و هر چه به آخر نمودار نزدیک می‌شود کمی از دقت پیش‌بینی کاسته شده است و فاصله منحنی سبز رنگ از منحنی آبی رنگ بیشتر شده و این نشان‌دهنده این است که در توالی‌های بلندمدت ممکن است شبکه GRU نسبت به شبکه LSTM توانایی کمتری در یادگیری وابستگی‌های بلندمدت داشته باشد، که نیاز است در مراحل آتی به آن پرداخته شود.

۵- نتیجه‌گیری

این تحقیق به دنباله مقایسه شبکه‌های عصبی بازگشتی عمیق LSTM و GRU در پیش‌بینی سری زمانی نرخ ارز بود. آزمایش‌های این تحقیق، با به دست آوردن طول پنجره بهینه و تعداد مرحله مناسب، توانست مقدار خطا را به حداقل برساند و نشان دهد این شبکه‌ها پتانسیل زیادی برای مدل‌سازی و پیش‌بینی داده‌های سری زمانی از جمله نرخ ارز دارند.

همچنین در این پژوهش با مقایسه مقدار جذر میانگین مربعات خطا و میانگین قدر مطلق خطا برای پیش‌بینی تغییرات نرخ‌های روزانه، هفتگی و ماهانه به این نتیجه

- [9]Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," IEEE transactions on neural networks, vol. 5, no. 2, pp. 157-166, 1994.
- [10]F. Liu, L. Zhang, and Z. Jin, "Modeling programs hierarchically with stack-augmented LSTM," Journal of Systems and Software, vol. 164, p. 110547, 2020.
- [11]J. Kim, J. Kim, H. L. T. Thu, and H. Kim, "Long short term memory recurrent neural network classifier for intrusion detection," in Platform Technology and Service (PlatCon), 2016 International Conference on, 2016, pp. 1-5: IEEE.
- [12]K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," arXiv preprint arXiv:1409.1259, 2014.
- [13]J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv preprint arXiv:1412.3555, 2014.
- [14]Z. Zhang, X. Pan, T. Jiang, B. Sui, C. Liu, and W. Sun, "Monthly and Quarterly Sea Surface Temperature Prediction Based on Gated Recurrent Unit Neural Network," Journal of Marine Science and Engineering, vol. 8, no. 4, p. 249, 2020.
- [15]B. Liu, C. Fu, A. Bielefeld, and Y. Q. Liu, "Forecasting of Chinese Primary Energy Consumption in 2021 with GRU Artificial Neural Network," Energies, vol. 10, no. 10, p. 1453, 2017.
- [16]Z. Wu and S. King, "Investigating gated recurrent neural networks for speech synthesis," arXiv preprint arXiv:1601.02539, 2016.
- [17]X. Zhang, F. Shen, J. Zhao, and G. Yang, "Time Series Forecasting Using GRU Neural Network with Multi-lag After Decomposition," in International Conference on Neural Information Processing, 2017, pp. 523-532: Springer.
- [18]H. Salehinejad, J. Baarbe, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent Advances in Recurrent Neural Networks," arXiv preprint arXiv:1801.01078, 2017.

روزانه، هفتگی و ماهانه، به صورت مجتمع، پیش‌بینی را بهبود بخشید.

مراجع

- [1]A. Shakeel, T. Tanaka, and K. Kitajo, "Time-Series Prediction of the Oscillatory Phase of EEG Signals Using the Least Mean Square Algorithm-Based AR Model," Applied Sciences, vol. 10, no. 10, p. 3616, 2020.
- [2]H. Chacón, E. Kesici, and P. Najafirad, "Improving Financial Time Series Prediction Accuracy using Ensemble Empirical Mode Decomposition and Recurrent Neural Networks," IEEE Access, 2020.
- [3]D. Shah, H. Isah, and F. Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," International Journal of Financial Studies, vol. 7, no. 2, p. 26, 2019.
- [4]R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," in Chinese Association of Automation (YAC), Youth Academic Annual Conference of, 2016, pp. 324-328: IEEE.
- [5]R. J. Frank, N. Davey, and S. P. Hunt, "Time series prediction and neural networks," Journal of Intelligent & Robotic Systems, vol. 31, no. 1, pp. 91-103, 2001.
- [6]R. Rana, "Gated recurrent unit (gru) for emotion classification from noisy speech," arXiv preprint arXiv:1612.07778, 2016.
- [7]S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.
- [8]S. Haidong, C. Junsheng, J. Hongkai, Y. Yu, and W. Zhan-tao, "Enhanced deep gated recurrent unit and complex wavelet packet energy moment entropy for early fault prognosis of bearing," Knowledge-Based Systems, vol. 188, p. 105022, 2020.

جدیدترین کتاب

از انتشارات انجمن انفورماتیک ایران

منتشر شد!

مهارت‌های نرم

برای تهیه کتاب با دفتر انجمن انفورماتیک ایران

تماس بگیرید ۶۶۴۱۲۸۶۱

چاپ اول



مهارت‌های نرم

راهنمای زندگی تولیدکنندگان نرم‌افزار

نوشته جان سانمز
ترجمه ابراهیم نقیب‌زاده مشایخ

انجمن انفورماتیک ایران

علوم رایانشی / زمستان ۱۳۹۹