



موضوع تحقیق : پیش بینی سری های زمانی در بازار مالی با استفاده از شبکه
عصبی حافظه کوتاه مدت بلند مدت

محقق: اشکان بویری کناری

فهرست مطالب:

۱- بیان مسئله	۲
۲- تعریف داده ها	۶
۳- معماری شبکه های عصبی	۱۲
۴- نتیجه گیری	۳۲
۵- منابع	۳۶

۱. بیان مسئله

امروزه دانشمندان داده و مهندس ها با استفاده از ابزار یادگیری ماشین و هوش مصنوعی می توانند برای پیش بینی آینده از آن ها استفاده کنند. در میان ابزارهای یادگیری ماشین، ابزارهایی هستند مناسب برای داده های زمانی هستند. داده های زمانی؛ داده هایی هستند که تغییرات را در واحد زمان ثبت و نمایش می دهند. داده های زمانی هستند که به صورت طبیعی با آن ها در زندگی روزمره آشنا هستیم مانند فیلم و موزیک. در این مثال ما

از یک الگوریتم خاص از خانواده ی شبکه های عصبی بازگشتی^۱ استفاده می کنیم با نام ال اس تی ام^۲ است. عدم قطعیت در بازار سرمایه به معنای تفاوت مقادیر مورد انتظار و مقادیری است که در واقعیت اتفاق می افتند. طراحی روش های تحلیل و پیش بینی مختلف در بازار سرمایه نیز به دلیل بالا بودن این مقدار و نیاز به دانستن قیمت ها در آینده با قطعیت بیشتر یا عدم قطعیت کمتر بوده است. برای کسب سود در بازار سرمایه، سرمایه گذاران همواره به دنبال پیدا کردن سهم مناسب جهت سرمایه گذاری و قیمت مناسب برای خرید و فروش بوده اند و لذا تمام مدل های پیش بینی مطرح شده همواره به دنبال پاسخ دادن به سه سوال اساسی بوده اند؛ چه سهمی، در چه محدوده زمانی و در چه قیمتی خریداری شود و یا به فروش برسد. قبل از بررسی پاسخ های داده شده به این سوالات، باید به سوال جدی تری پاسخ داد. از جمله اینکه آیا پیش بینی بازارهای مالی ممکن است؟! همچنین، در ادامه باید به این موضوع پرداخته شود که در صورت پیش بینی پذیر بودن بازار سرمایه، باید به بررسی ابعاد مختلف بازار سرمایه و متدهایی که در هر زمینه برای پیش بینی ارائه شده است، پرداخت. در ادامه باید بررسی کرد که چه متدهایی کارایی لازم برای این پیش بینی را دارند و آیا ترکیب این متدها به صورت کلی ممکن است یا خیر. در ادامه خواهیم دید که می توان ابزار به کار گرفته شده در پیش بینی تمام ابعاد بازار سرمایه را در سه دسته کلی متدهای تکنیکال، متدهای بنیادین و متدهای ریاضی، شامل متدهای کلاسیک سری زمانی و رگرسیون و متدهای هوش مصنوعی قرار داد. در این پژوهش، با بررسی تمام موارد بالا و امکان سنجی تلفیق متدهای به کارگرفته شده جهت پیش بینی قیمت، به سوالات مطرح شده پاسخ داده خواهد شد و برای اولین بار، به پیش بینی دو قیمت برای دوره های جلوتر پرداخته می شود؛ قیمت بالا و قیمت پایین سهام. به این وسیله، سفته بازان می توانند با به کارگیری این متد، با دقت قابل قبولی به پیش بینی قیمت پرداخته و از طریق نوسان گیری، کسب سود کنند.

بیت کوین مستقیماً بین طرفین معامله انتقال می یابد و تراکنش های آن نیازمند گذر از یک موسسه مالی (نهاد مرکزی واسط) نیست. بیت کوین معروف ترین رمز ارز است که از فناوری زنجیره بلوکی استفاده می کند. این واحد پولی نوپا در سال ۲۰۰۹ توسط ساتوشی ناکاموتو ابداع شد که به نظر می رسد نامی مستعار باشد. البته این امکان هم وجود دارد که ایشان عضو گروهی بزرگ تر از مبتکران ایده بیت کوین باشند. در حال حاضر بیت کوین با مبادلات دنیای واقعی تطابق و سازگاری بیشتری پیدا کرده است و بعضاً معاملات مهمی مثل خرید خانه یا ماشین نیز با بیت کوین انجام می شود. حجم معاملات با بیت کوین در آغاز سال ۲۰۱۹ بسیار بالا بوده و شروعی رویایی داشته

^۱Recurrent Neural Network

^۲LSTM

است. این مسئله باعث تحیر افراد فعال در بازار سرمایه شده است. به همین دلیل تاجران و سرمایه‌گذاران تلاش می‌کنند تا بیشترین سود را از این وضعیت ببرند. هر چه حجم مبادلات بیت‌کوین افزایش پیدا کند، بالطبع قیمت آن نیز افزایش پیدا خواهد کرد. علاوه بر این، تعداد بیت‌کوین بسیار محدود بوده و تنها ۲۱ میلیون بیت‌کوین را می‌توان استخراج کرد. این در حالی است که طبق آمار، تا تاریخ ۱۰ می ۲۰۲۰ تعداد ۱۸,۳۷۲,۱۲۵ بیت‌کوین استخراج شده و تنها ۲,۶۲۷,۸۷۵ بیت‌کوین دیگر را می‌توان استخراج کرد و این یعنی احتمالاً بعد از استخراج این مقدار، ارزش بیت‌کوین افزایش پیدا خواهد کرد. بیت‌کوین به عنوان یک رمز ارز، فاقد شکل فیزیکی است به طوری که انجام تحلیل آن به نحوی غیرممکن شده است. در نتیجه، بسیاری از سرمایه‌گذاران، پیوسته با استفاده از شاخص‌های تجاری تحلیل فنی (الگوهای هندسی ایجاد شده از قیمت‌های پیشین و حجم معاملات) برای درک و پیش‌بینی روند آتی قیمت بیت‌کوین تلاش می‌کنند. هدف از این تحقیق تعیین چگونگی پیش‌بینی صحت و دقت قیمت بیت‌کوین با استفاده از الگوریتم‌های یادگیری ماشین است. علاوه بر این، بیت‌کوین به طرز شگفت‌آوری غیرقابل پیش‌بینی‌تر از استانداردهای مختلف پولی مانند دلار آمریکا است. هم‌چنین، بیت‌کوین به عنوان یک رمز ارز اصلی، پیش‌رو بوده که انتظار می‌رود که رشد بیشتری داشته باشد و این امر الهام بخش تحقیقات در این زمینه است. پیش‌بینی قیمت بیت‌کوین، یک فرصت را برای کسب سود از طریق خرید و فروش رمز ارز ایجاد می‌کند. قیمت بیت‌کوین بر خلاف بازار سهام، به وقایع تجاری یا مداخله دولت بستگی ندارد. پیش‌بینی ارزش بیت‌کوین از جنس مسائل چند متغیره است. روش‌های قدیمی پیش‌بینی رمزارزها، مانند مدل‌های هموارسازی نمایی هولت وینترس، به فرضیات خطی متکی هستند و برای موثر بودن، نیازمند داده‌هایی هستند که بتوان آنها را به روند، فصلی و نویز تقسیم کرد. این روش برای کارهایی مانند پیش‌بینی فروش محصولات، مناسب است که اثرات فصلی در آن وجود دارد. به دلیل فقدان چنین ویژگی در بازار بیت‌کوین (یا حداقل ندرت آن) و نوسانات زیاد قیمت بیت‌کوین، این روش‌ها برای این کار چندان مؤثر نیستند. با توجه به پیچیدگی کار، داده کاوی، راه‌حلی جالب و فناورانه را ایجاد می‌کند.

تکنیک‌های مختلفی برای پیش‌بینی روندهای بازار سهام به کار گرفته شده است. با این حال، به دلیل پیچیدگی بازار سهام، نتایج کاملاً رضایت‌بخش نیستند. بسیاری از رویکردها یا فاقد تعریف واضح و معقول از روند هستند یا از منحصر به فرد بودن ویژگی زمان در داده‌های سهام غفلت می‌کنند و با آن‌ها مانند سایر ویژگی‌ها رفتار می‌کنند. از آنجا که بورس یک سیستم پویا، غیرخطی و با ابعاد بالا است که عوامل مختلفی مانند اقتصاد، سیاست‌ها، روان‌شناسی و غیره را درگیر می‌کند که پیش‌بینی آن را دشوار می‌کند. برخی از تئوری‌ها، به عنوان مثال، نظریه بازار کارآمد بیان می‌کند که دارایی‌ها، در یک بازار با جریان آزاد اطلاعات، همیشه به ارزش منصفانه خود معامله

می‌شوند و همه اطلاعات موجود را منعکس می‌کنند، به این معنی که قیمت سهام فقط تحت تأثیر اطلاعات جدید است. بنابراین، عملکرد بهتر از بازار با انتخاب سهام یا زمان بندی بازار غیرممکن به نظر می‌رسد. با این حال محققان دریافته‌اند که اگرچه پیش‌بینی حرکت سهام در کوتاه‌مدت دشوار است، اما در بلندمدت هنوز قابلیت پیش‌بینی وجود دارد. هوش مصنوعی در آینده نزدیک تمام جنبه های زندگی بشر را متحول می‌سازد، از آنجا که حوزه معاملات اقتصادی و بازار بورس یکی از حوزه های پیچیده می‌باشد که نیازمند تحلیل های دقیق، هوشمندانه و منطقی است به همین سبب استفاده از هوش مصنوعی میتواند درهای جدیدی به روی دنیای اقتصاد باز کند. با وجود چنین سیستم های ریسک معاملات کاهش پیدا می‌کند، در نتیجه معامله گر با اطمینان بیشتری معاملات خود را انجام دهد. هوش مصنوعی در زمینه معاملات اقتصادی و بازار بورس میتواند این کاربرد ها را داشته باشد: کمک به داشتن فرآیندی اتوماتیک و دقیق تر برای معاملات، پیش بینی روند پیش رو به کمک تجزیه و تحلیل داده های گذشته، با نظارت بر بازار و تجربه و تحلیل آن این سیستم های میتوانند فرصت های معاملاتی برای داشتن خرید و فروش های کارآمد برای معامله گران ایجاد کنند.

استفاده از تکنولوژی و فناوری های به روز همیشه می‌تواند ما را یک قدم جلوتر از دیگران قرار دهد، به ویژه وقتی صحبت از زمینه هایی مثل بورس و سرمایه گذاری باشد. دو نمونه از بهترین تکنولوژی های امروز در بازار بورس هوش مصنوعی و یادگیری ماشین نام دارند که به تازگی پا به این عرصه گذاشته‌اند. با استفاده از این دو عامل و پلتفرم های وابسته به آن می‌توان بازدهی معاملات را افزایش داد، در هزینه و وقت صرفه جویی کرد و نهایتاً درصد ریسک را تا حد قابل توجهی کاهش داد. در سال های اخیر استفاده از هوش مصنوعی و یادگیری ماشین یکی از امیدوار کننده ترین ابزار ها برای تجزیه و تحلیل و پیش بینی داده های سری زمانی که شامل داده های بازار های مالی می‌شود بوده است، پیش بینی بازار سهام عملیاتی است که تلاش می‌کند ارزش آینده سهام یک شرکت یا سایر نماد های مالی معامله شده در بازار های مالی را تعیین کند (پیش بینی کند). پیش بینی موفقیت آمیز قیمت آینده سهام، سود سرمایه گذار را به حداکثر می‌رساند. در عصری زندگی می‌کنیم که تکنولوژی تا ریزترین بخش های زندگی فردی و اجتماعی انسان وارد شده است و اجتناب از آن امکان پذیر نیست. از جمله بازارهایی که چند سالی می‌شود به اجتناب ناپذیر بودن این حقیقت پی برده‌اند بازارهای مالی هستند. ورود بازار سرمایه به عصر تکنولوژی با معاملات الگوریتمی اتفاق افتاد، همه افراد از متضرر شدن در بازار بیزارند. آگاهی از بستر ایمن و مطمئن سرمایه گذاری برای افراد ضروری است. الگوریتم های یادگیری ماشین با استفاده از داده های بسیار زیادی که متخصصین در اختیارش قرار میدهند آموزش می‌بیند. سپس با توجه به یک چهارچوب منطقی به شما می‌گوید که به عنوان مثال این شرکت سابقه ی خوبی دارد یا نه. این مورد حتی در جهت شناسایی کلاهبرداران نیز مورد

استفاده قرار می‌گیرد که شما را از معرض مورد تقلب قرار گرفتن ایمن می‌سازد. یکی از مهم‌ترین ویژگی‌های هوش مصنوعی و یادگیری ماشین این است که نیازی به مداخله انسانی ندارد. این موضوع سرعت ما را در زندگی روزمره به شدت افزایش می‌دهد، مدیریت زمان ما را معنادارتر و دیدگاه ما را بزرگ‌تر می‌کند. به بیان بهتر به کارگیری هوش مصنوعی و یادگیری ماشین مانند استخدام کارمندی است که کاملاً دقیق بوده، به صورت مستمر در حال یادگیری است، خسته نمی‌شود، قدرت پردازش خوبی دارد و می‌تواند سفارشی سازی شود. آیا واقعاً انسانی به این شکل وجود دارد؟ مسلماً خیر. سیستم‌های مبتنی بر هوش مصنوعی و یادگیری ماشین کاملاً بر اساس منطق تصمیم‌گیری می‌کند و فاقد هرگونه احساسی می‌باشند از این رو بدون در نظر گرفتن خرافات و احساسات به ارائه‌ی نتایج از داده‌ها زمان دار که به دقت آن می‌افزاید، اقدام می‌کند، استفاده از هوش مصنوعی سبب کاهش ریسک و صرفه جویی در زمان می‌شود.

استفاده از روش‌های غیر کلاسیک در شناسایی مدل و پیش‌بینی رفتار سیستم‌های پیچیده، مدتهاست در محافل علمی و حتی حرفه‌ای متداول و معمول شده است. در بسیاری از سیستم‌های پیچیده و خصوصاً غیر خطی که مدل سازی و به دنبال آن پیش‌بینی و کنترل آنها از طریق روش‌های کلاسیک و تحلیلی امری بسیار دشوار و حتی بعضاً غیر ممکن می‌نماید، از روش‌های غیر کلاسیک که از ویژگی‌هایی همچون هوشمندی، مبتنی بر معرفت و خبرگی برخوردار هستند، استفاده می‌شود. شبکه‌های عصبی، یکی از این روش‌های بدیع و در حال تحول است که در موضوعات متنوعی از قبیل الگوسازی، شناخت الگو، خوشه بندی و پیش‌بینی به کار رفته و نتایج مفیدی داشته است. در این مقاله، از شبکه‌های عصبی در پیش‌بینی سری‌های زمانی داده‌های اقتصادی استفاده کرده ایم.

مسئله‌ای که می‌خواهیم به حل آن بپردازیم، مسئله‌ی پیش‌بینی قیمت بیت کوین بر پایه‌ی اطلاعات ۲۴ ساعت گذشته از قیمت آن است.

۲- تعریف داده‌ها

در این پژوهش ما با کمک کتابخانه یاهو فایننس^۳ که یک رابط بین پایتون و سایت یاهو فایننس می‌باشد داده‌های قیمت بیت کوین از تاریخ ۲۰۱۴/۰۹/۱۷ تا تاریخ ۲۰۲۰/۰۷/۰۸ جمع‌آوری کرده ایم باتوجه به آنکه هر زمان که را اجرا کنیم قیمت بیت کوین را تا آن روز برای ما جمع‌آوری می‌کند. یاهو فایننس بخشی از شرکت یاهو می‌باشد.

^۳Yahoo Finance

باشد که اخبار مالی، داده ها و تفسیر از جمله قیمت سهام، بیانیه های مطبوعاتی، گزارش های مالی و ... را ارائه می دهد. هم چنین برخی از ابزارهای آنلاین را برای مدیریت مالی شخصی ارائه می دهد. ما در این پژوهش سعی کرده ایم شبکه های را آموزش دهیم که قیمت بسته شدن سهام در هر روز را پیش بینی کنند داده ها به صورت داده های فایل که به صورت فایل متنی هستند که داده ها با کاما از هم جدا شده اند.

ما برای خواندن داده ها از ورودی از کتابخانه ی پانداز^۴پایتون استفاده می کنیم.

برای شرح داده های تحقیق و کسب اطلاعات از دیتاست از تابع `head()`, `info()` و توابع همبستگی داده استفاده کرده ایم. نمودار های برای توصیف ستون های داده ها با استفاده از کتابخانه های مت پلات لیب^۵ و سیبورن^۶ و پلاتی^۷ رسم کرده ایم در ادامه هر کدام از این نمودار ها آورده شده اند و شرح مختصری از آنچه که به نمایش گذاشته اند بیان شده است.

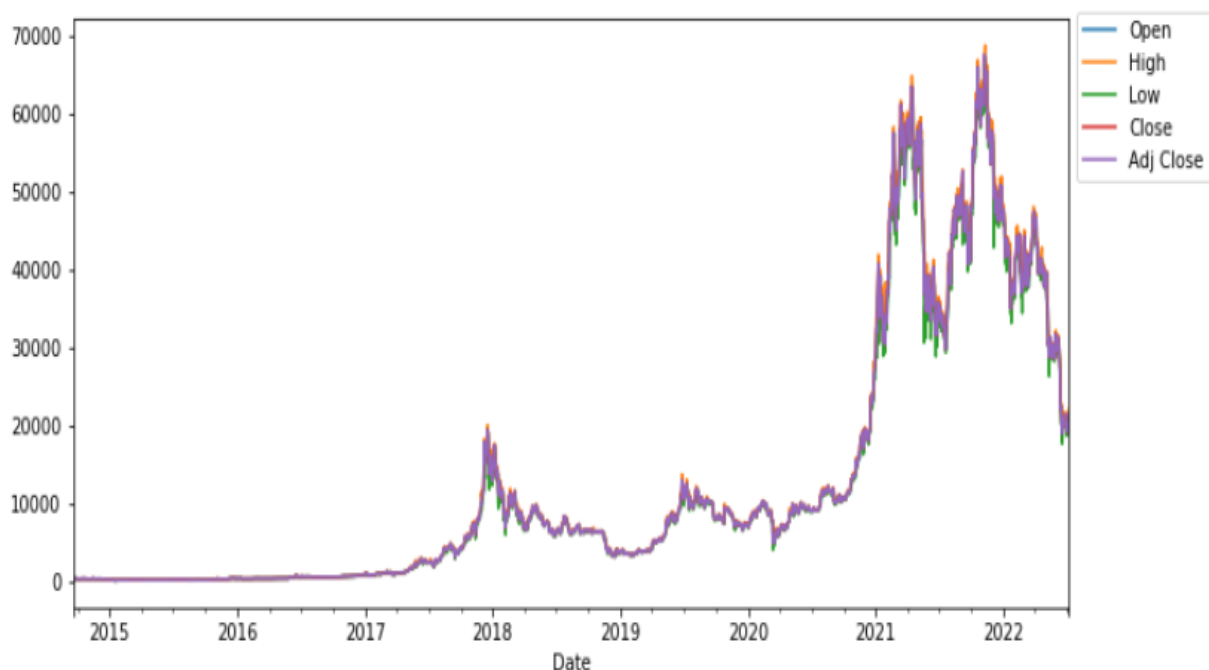
شرح ستون های مختلف داده ها:

نام ستون	توضیحات
Index	تاریخ روزی که اطلاعات اوراق بهادر در آن ثبت شده اند
Open	قیمت باز شدن سهام در آن روز
Low	کمترین مقداری که سهام در آن روز باز شده است
High	بیشترین مقداری که سهام در آن روز باز شده است
Close	قیمت باز شدن اوراق در آن روز
Adj close	قیمت بسته شدن تعدیل شده در آن روز

^۴Pandas
^۵Matplotlib
^۶Seaborn
^۷plotly

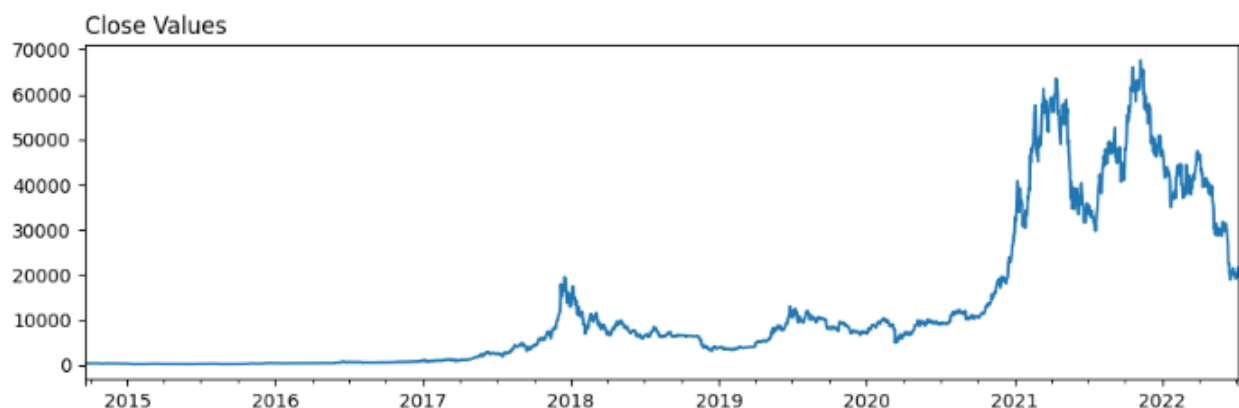
میزان حجم معاملات در آن روز	Vloun
-----------------------------	-------

برای به خط کشیدن روند داه های تحقیق از نمودار زیر استفاده شده است. که مقادیر ویژگی های مهم بیت کوین از انتهای سال ۲۰۱۴ تا اواسط سال ۲۰۲۰ را نشان می دهد. که روند تغییرات تقریبا مشابه هم بوده است.

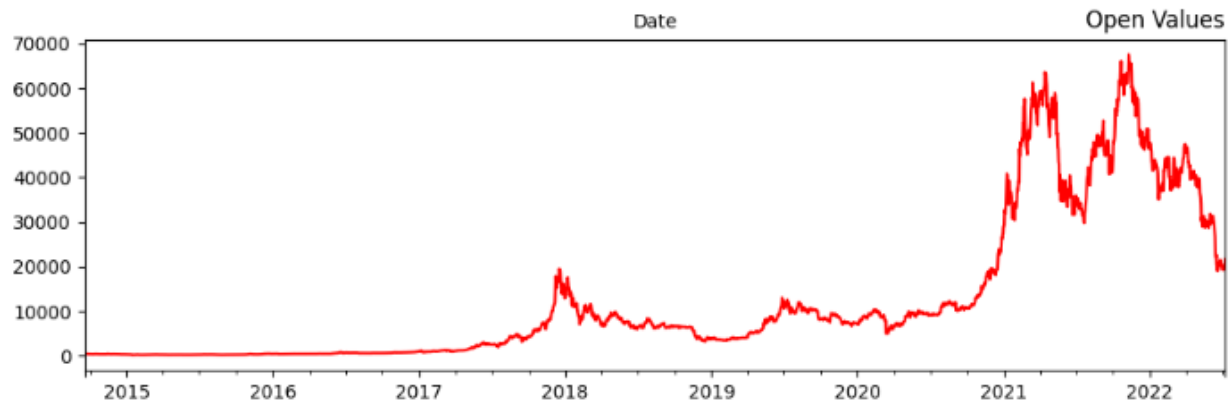


شکل ۲-۱ روند تغییر مقادیر ستون های (open,close,high,adj close,vloun)

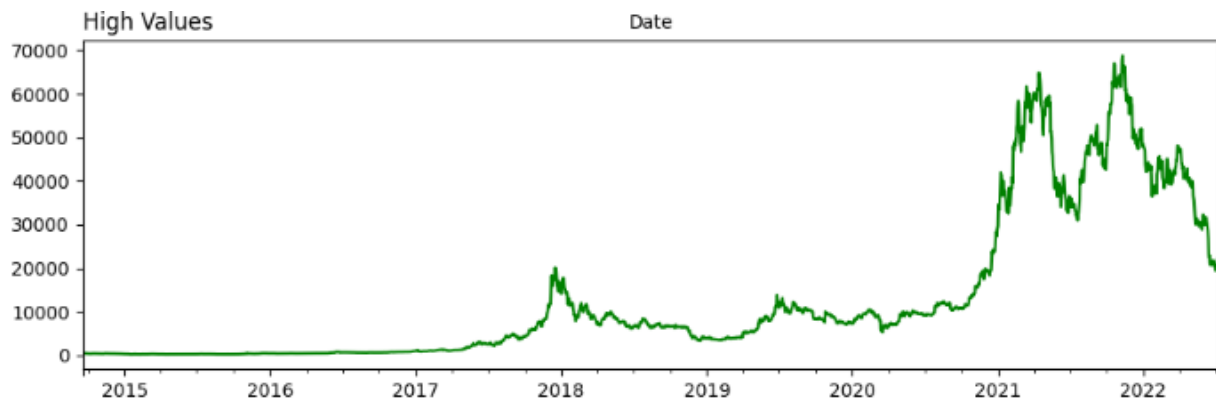
در نمودار های زیر روند تغییر مقادیر ستون های مختلف دیتاست در طی زمان را مشاهده می کنیم:



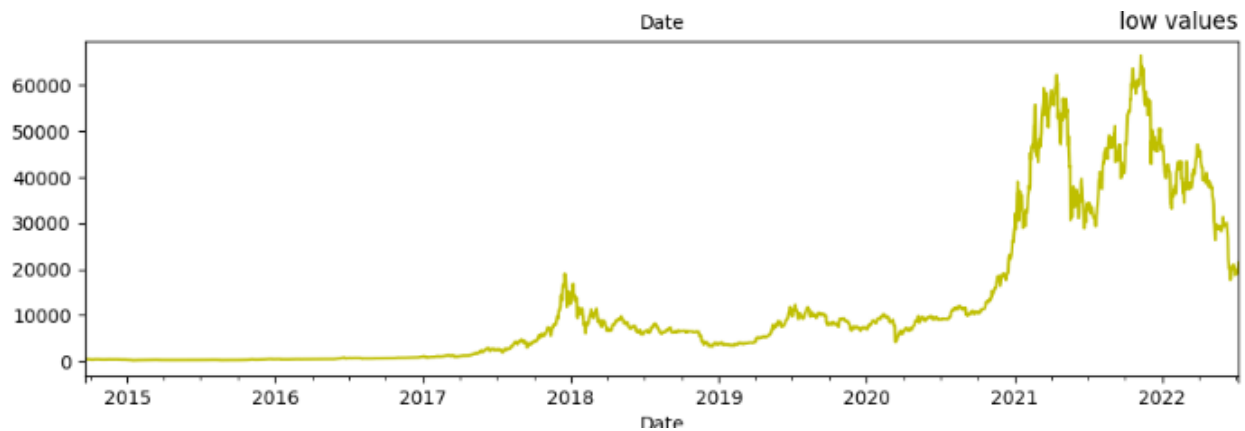
شکل ۲-۲: نمودار روند تغییرات مقدار بسته شده بیت کوین از سال ۲۰۱۴ تا سال ۲۰۲۰



شکل ۲-۳: نمودار روند تغییرات مقدار باز شده بیت کوین از سال ۲۰۱۴ تا سال ۲۰۲۰

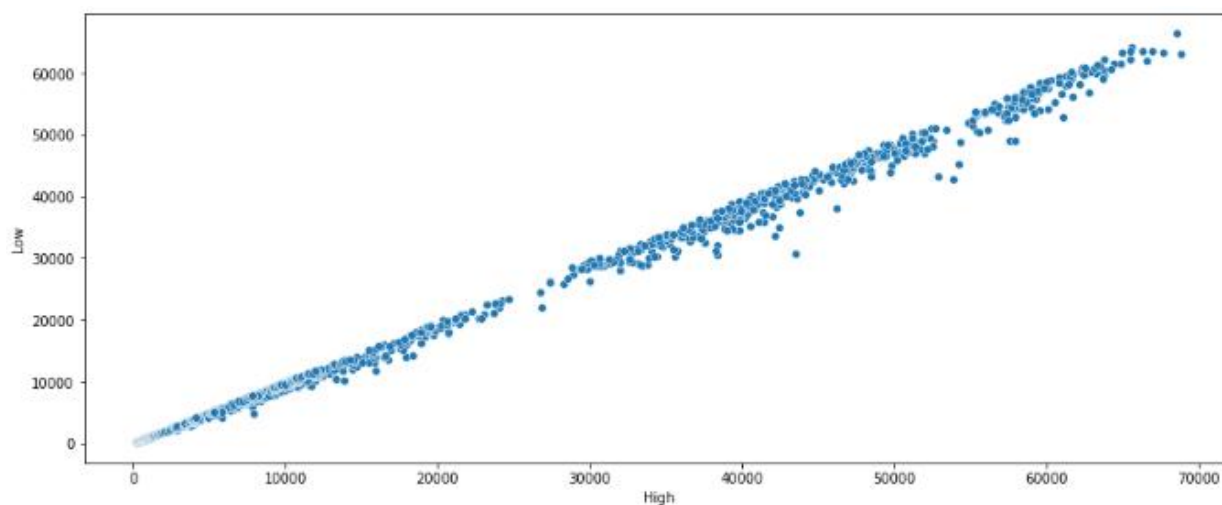


شکل ۲-۴: نمودار روند تغییرات با بیشترین قیت بیت کوین از سال ۲۰۱۴ تا سال ۲۰۲۰

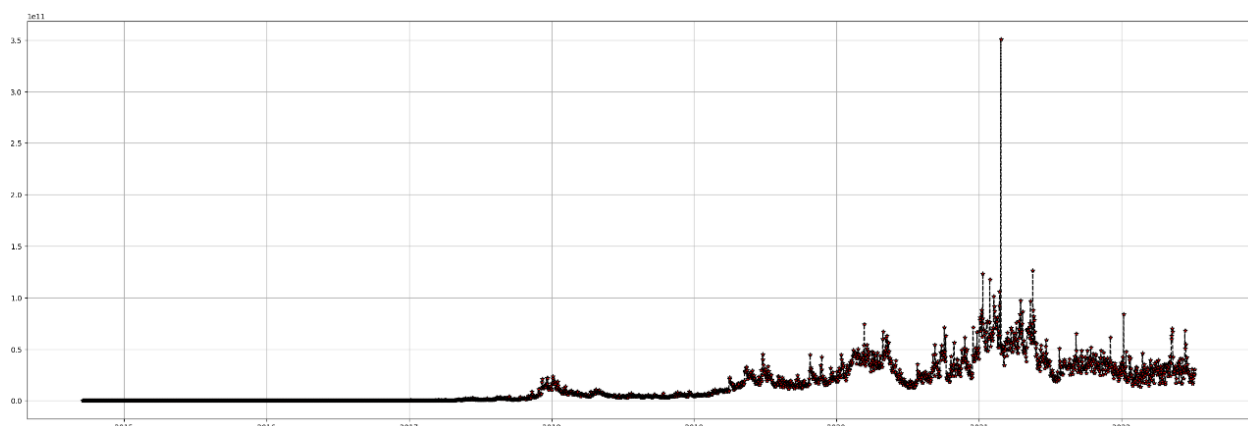


شکل ۲-۵: نمودار روند تغییرات با کم ترین قیت بیت کوین از سال ۲۰۱۴ تا سال ۲۰۲۰

توصیف وابستگی بین کم ترین مقدار قیمت در حجم معاملات و بیشترین مقدار قیمت در حجم معاملات بیت کوین را نشان می دهد.

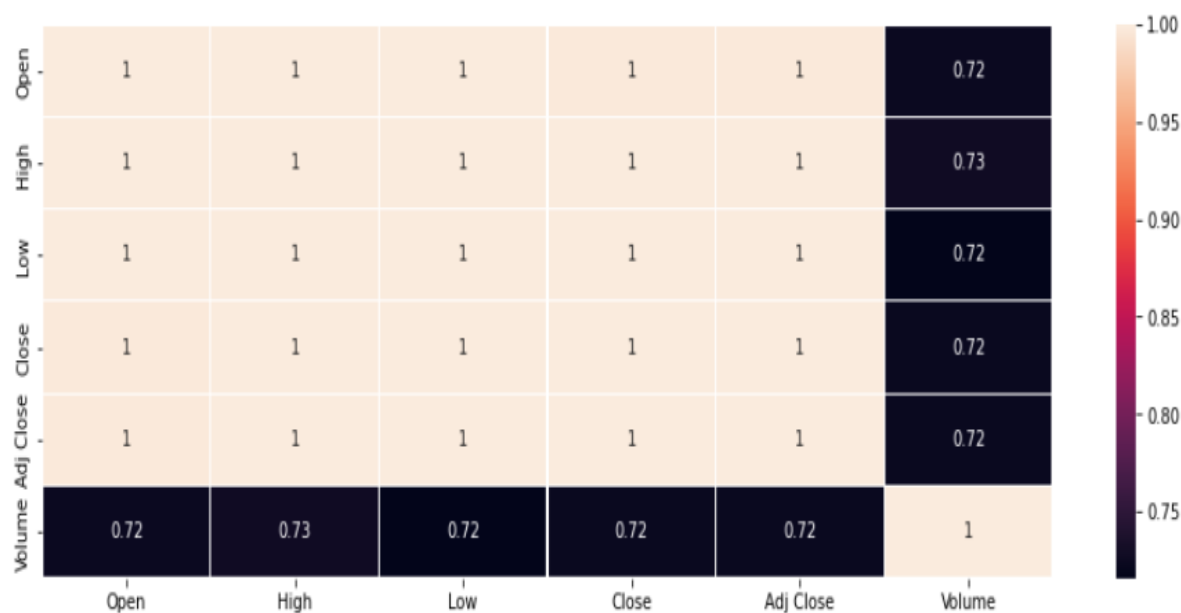


شکل ۲-۵: نمودار اسکتر پلات بین دو ویژگی کم ترین حجم معاملات و بیشترین مقدار حجم معامله



شکل ۲-۶: نمودار گرید توصیف حجم معاملات بیت کوین در سری زمانی سال ۲۰۱۴ تا سال ۲۰۲۰

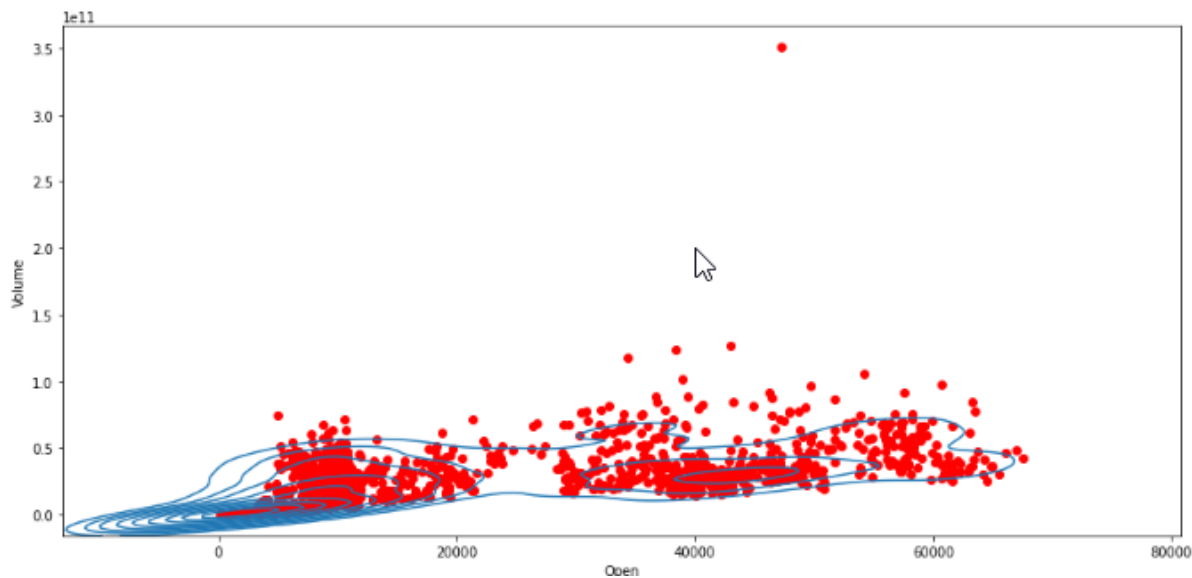
نمودار زیر همبستگی بین ستون هاس مختلف دیتاست را به ما نشان میدهد:



این جدول در هر سطر به ترتیب تعداد داده های موجود ، میانگین ، انحراف معیار ، کمترین مقدار، چارک اول، چارک دوم، چارک سوم و بیشترین مقدار هر ستون را در خود گنجانیده است جدول زیر توصیف آماری کاملی از تمام ستون ها به نمایش میگذارد:

	Open	High	Low	Close	Adj Close	Volume
count	2852.000000	2852.000000	2852.000000	2852.000000	2852.000000	2.852000e+03
mean	12438.445471	12762.284354	12074.918764	12444.167710	12444.167710	1.543296e+10
std	16581.359304	17009.706219	16084.298465	16576.975235	16576.975235	1.991359e+10
min	176.897003	211.731003	171.509995	178.102997	178.102997	5.914570e+06
25%	635.306519	641.894257	625.045227	636.637512	636.637512	8.639670e+07
50%	6566.449951	6672.479980	6452.399902	6572.459961	6572.459961	5.853965e+09
75%	11529.242188	11816.213135	11286.059326	11545.715820	11545.715820	2.627581e+10
max	67549.734375	68789.625000	66382.062500	67566.828125	67566.828125	3.509679e+11

شکل ۲-۶: توصیف آماری داده های تحقیق



شکل ۲-۷: نمودار اسکتر پلات بین ویژگی حجم و قیمت باز شده

و اسکتر پلات بین قیمت بسته شده و قیمت باز شده بیت کوین بررسی شده است. و هر گونه تحلیل آماری که شناخت از نوع داده را به محقق داده باشد در این کد استفاده شده است. دیده گاه اماريست و نوع رسم نمودار ممکن است متفاوت باشد. در این تحقیق بیشتر از نمودار اسکتر بین ویژگی ها برای شناخت از ویژگی ها استفاده شده است.

۳- معماری شبکه عصبی مصنوعی:

در ابتدا به برخی از تعاریف تحقیق و ادبیات تحقیق خواهیم پرداخت.

تعاریف تحقیق

هوش مصنوعی؛ هوش مصنوعی که گاهی اوقات هوش ماشینی نیز نامیده می شود شبیه سازی فرایندهای هوش طبیعی^۹ توسط ماشین ها به ویژه سیستم های رایانه ای است. به عبارت دیگر، هوش مصنوعی به سامانه هایی گفته می شود که می توانند واکنش هایی مشابه رفتارهای هوشمند انسانی از جمله، درک شرایط پیچیده، شبیه سازی فرایندهای تفکری و شیوه های استدلالی انسانی و پاسخ موفق به آن ها، یادگیری و توانایی کسب دانش و استدلال برای حل مسائل را داشته باشند، بطور خلاصه هوش مصنوعی را دانش ساخت و طراحی عامل هوشمند تعریف کرده اند. این علم کاربرد های فراوانی در علوم رایانه، علوم مهندسی، تجارت، پزشکی و بسیاری از علوم دیگر دارد. بعنوان مثال: در پزشکی تجزیه و تحلیل صدا قلب، ربات های پرستار، ارائه مشاوره و پیش بینی احتمال مرگ بیمار برای هر روش جراحی..... در امور مالی و تجارت تجزیه و تحلیل بازار های مالی، پیش بینی قیمت سهام ها، معاملات الگوریتمی، مدیریت دارای و... از کاربرد های هوش مصنوعی در این علوم هستند. هوش مصنوعی، موضوعی بسیار گسترده است که شاخه های متعددی دارد. شاخه های هوش مصنوعی عبارتند از: یادگیری ماشینی، شبکه های عصبی^{۱۱} سیستم های خبره^{۱۲}، پردازش زبان طبیعی^{۱۳}، تشخیص گفتار^{۱۴} و بینایی ماشین^{۱۵} و- رباتیک^{۱۶} و منطق فازی است.

یادگیری ماشین: یادگیری ماشینی شاخه ای از هوش مصنوعی^۸ و علوم کامپیوتر^۹ است که بر استفاده از داده ها و الگوریتم ها برای تقلید از روشی که انسان ها یاد می گیرند تمرکز دارد و به تدریج دقت آن را بهبود می بخشد. یادگیری ماشین به عنوان بخشی از هوش مصنوعی در نظر گرفته می شود که به مطالعه الگوریتم های کامپیوتری می پردازد که می تواند به طور خودکار از طریق تجربه و با استفاده از داده ها بهبود یابد. الگوریتم های یادگیری ماشین مدلی را بر اساس داده های نمونه می سازند که به داده های آموزشی معروف است تا پیش بینی ها یا تصمیم گیری ها را بدون برنامه ریزی صریح انجام دهند. الگوریتم های یادگیری ماشین در کاربردهای متنوعی مانند پزشکی، فیلتر کردن ایمیل، تشخیص گفتار و بینایی کامپیوتری استفاده می شوند. انواع الگوریتم های یادگیری

^۸Artificial Intelligence

^۹Natural Intelligence

^{۱۰}machine learning

^{۱۱}Neural network

^{۱۲}Expert Systems

^{۱۳}Natural language processing

^{۱۴}speech recognition

^{۱۵}Machine vision

^{۱۶}Robotic

^{۱۷}Fuzzy logic

^{۱۸}Artificial intelligence

^{۱۹}computer science

ماشین عبارت اند از: درخت تصمیم،^{۲۴}لاجستیک رگرسیون،^{۲۵}ماشین بردار پشتیبان،^{۲۶}لاسته بند بیز،^{۲۷}نزدیک ترین همسایه^{۲۸}... می توان اشاره کرد.

یادگیری عمیق^{۲۹} یادگیری عمیق که در زبان فارسی به یادگیری ژرف نیز ترجمه شده است بخشی از خانواده یادگیری ماشین می باشد که بر روش های تمرکز دارد که مبتنی بر الگوریتم های شبکه عصبی مصنوعی هستند. این الگوریتم ها تلاش اند که مغز انسان را شبیه سازی کنند. به طور خلاصه در یادگیری عمیق شبکه های عصبی مصنوعی و الگوریتم های مشابه مغز بشر از مجموعه های عظیم داده مهارت های مورد نظر را فرا می گیرند. همانطور که ما از طریق تجربه چیزهای جدید یاد می گیریم الگوریتم یادگیری عمیق نیز با هر بار تکرار یک کار مهارت خود را نسبت به دفعات قبلی بهبود می بخشد. دلیل استفاده از عبارت یادگیری عمیق این است که شبکه های عصبی لایه های مختلف یا عمیقی دارند که یادگیری را ممکن می سازد.

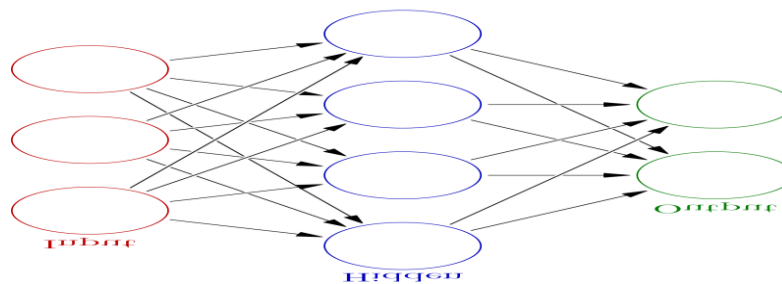
داده کاوی^{۳۰}: داده کاوی فرایندی برای تبدیل داده های خام به اطلاعات مفید می باشد، داده کاوی فرایند استخراج و کشف الگوها در مجموعه داده های بزرگ است که شامل روش هایی در محل تلاقی یادگیری ماشین ، آمار و سیستم های پایگاه داده است. به عبارت دیگر داده کاوی یک زیرشاخه بین رشته ای علوم کامپیوتر و آمار با هدف کلی استخراج اطلاعات (با روشهای هوشمند) از مجموعه داده و تبدیل اطلاعات به یک ساختار قابل درک برای استفاده بیشتر است.

الگوریتم های معاملاتی: استفاده از برنامه های کامپیوتری برای ورود به سفارش های معاملاتی بدون دخالت انسان. این الگوریتم ها که می توانند بیش از یکی باشند، برای انجام معاملات بررسی های لازم را از جنبه های گوناگونی مانند زمان بندی، قیمت و حجم روی سفارشات و بازار انجام داده و تصمیم می گیرند. این امر کمک می کند تا بازار سرمایه به روشی اصولی تر و به دور از دخالت احساسات انسانی پیش رود که یکی از نتایج آن بالارفتن نقدینگی در بازار است.

اکنون وقت آن است تا به صورت مختصر به شرح برخی از انواع شبکه های عصبی و کاربرد آنها بپردازیم :

^{۲۴}Decision tree
^{۲۵}Logistic Regression
^{۲۶}Support vector machine
^{۲۷}Naive Bayes Classifiers
^{۲۸}KNN
^{۲۹}Deep learning
^{۳۰}Artificial neural network
^{۳۱}Data mining

شبکه های عصبی مصنوعی که معمولاً به سادگی شبکه های عصبی نامیده می شوند، سیستم های محاسباتی هستند که از شبکه های عصبی بیولوژیکی الهام گرفته شده اند که مغز حیوانات را تشکیل می دهند. یک شبکه عصبی مصنوعی مبتنی بر مجموعه ای از واحدها یا گره های متصل به هم به نام نورون های مصنوعی است که نورون ها را در یک مغز بیولوژیکی مدل سازی می کنند. هر اتصال، مانند سیناپس های یک مغز بیولوژیکی، می تواند سیگنالی را به نورون های دیگر منتقل کند. یک نورون مصنوعی سیگنال ها را دریافت می کند و سپس آنها را پردازش می کند و می تواند به نورون های متصل به آن سیگنال دهد. "سیگنال" در یک اتصال یک عدد واقعی است و خروجی هر نورون با مقداری تابع غیرخطی از مجموع ورودی های آن محاسبه می شود. اتصالات لبه نامیده می شوند. نورون ها و لبه ها معمولاً وزنی دارند که با ادامه یادگیری تنظیم می شود. وزن باعث افزایش یا کاهش قدرت سیگنال در یک اتصال می شود. نورون ها ممکن است آستانه ای داشته باشند که سیگنال تنها در صورتی ارسال شود که سیگنال مجموع از آن آستانه عبور کند.



شکل ۱-۳ شماتیکی از یک شبکه عصبی ساده

نحوه یادگیری شبکه های عصبی مصنوعی:

شبکه های عصبی با پردازش مثال ها، که هر کدام حاوی یک «ورودی» و «نتیجه» شناخته شده هستند، یاد می گیرند (یا آموزش می بینند) و ارتباط های احتمالی را بین این دو تشکیل می دهند که در ساختار داده های خود شبکه ذخیره می شوند. آموزش یک شبکه عصبی از یک مثال مشخص معمولاً با تعیین تفاوت بین خروجی پردازش

شده شبکه و خروجی هدف انجام می شود. این تفاوت همان خطا است. سپس شبکه وزن های خود را بر اساس یک قانون یادگیری و با استفاده از این مقدار خطا تنظیم می کند، تنظیمات پیاپی باعث می شود که شبکه خروجی را تولید کند که به طور فزاینده مشابه خروجی هدف باشد پس از تعداد کافی از این تنظیمات، آموزش را می توان بر اساس معیارهای خاصی خاتمه داد.

کاربرد شبکه های عصبی مصنوعی:

شبکه های عصبی مصنوعی به دلیل توانایی خود در بازتولید و مدل سازی فرآیندهای غیرخطی، در بسیاری از رشته ها کاربرد پیدا کرده اند. از جمله این کاربرد ها میتوان به موارد زیر اشاره کرد:

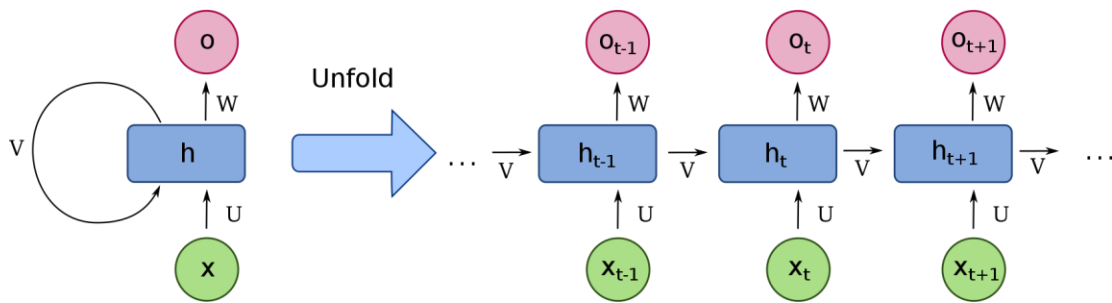
حوزه های کاربردی شامل شناسایی و کنترل سیستم (کنترل وسیله نقلیه، پیش بینی مسیر، کنترل فرآیند، مدیریت منابع طبیعی)، شیمی کوانتومی، تشخیص الگو (سیستم های رادار، شناسایی چهره، طبقه بندی سیگنال، بازسازی سه بعدی، تشخیص اشیا و موارد دیگر)، تجزیه و تحلیل داده های حسگر است و....

شبکه های عصبی بازگشتی^{۲۹}

شبکه عصبی بازگشتی یک کلاس از شبکه های عصبی مصنوعی است که در آن اتصالات بین گره ها یک گراف جهت دار یا غیر جهت دار را در امتداد یک دنباله زمانی تشکیل می دهند. این مسئله به آن اجازه می دهد تا رفتار پویا زمانی را نشان دهند. شبکه های عصبی بازگشتی که از شبکه های عصبی پیشخور^{۳۰} مشتق شده اند، می توانند از حالت داخلی (حافظه) خود برای پردازش توالی با طول متغیر ورودی ها استفاده کنند. شبکه های عصبی مکرر از نظر تئوری تورینگ کامل هستند و می توانند برنامه های دلخواه را برای پردازش توالی دلخواه ورودی ها اجرا کنند.

^{۲۹}recurrent neural network

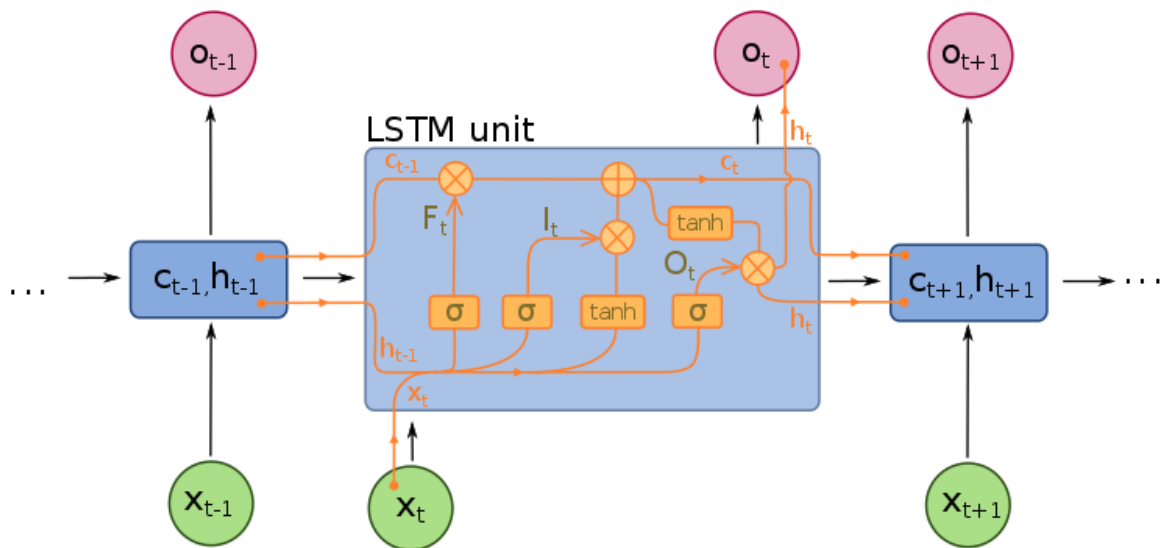
^{۳۰}feedforward neural networks



شکل ۳-۲ نحوه عملکرد یک شبکه عصبی بازگشتی را به نمایش میگذارد

در ادامه به معرفی مختصری از رایج ترین ، قدرتمن ترین و پرکاربرد ترین شبکه عصبی بازگشتی حال حاضر میپردازیم که در حال حاضر در ابعاد گسترده از آن استفاده می شود:

حافظه کوتاه مدت بلند مدت^۲!



در شکل ۳-۳ شماتیکی از یک شبکه حافظه کوتاه مدت بلند مدت نمایش داده شده است

حافظه کوتاه مدت بلند مدت یک سیستم یادگیری عمیق است که از مشکل ناپدید شدن گرادیان جلوگیری می کند. حافظه کوتاه مدت بلند مدت معمولاً با گیت های تکراری به نام «دروازه های فراموش»^۳ تقویت می شود،

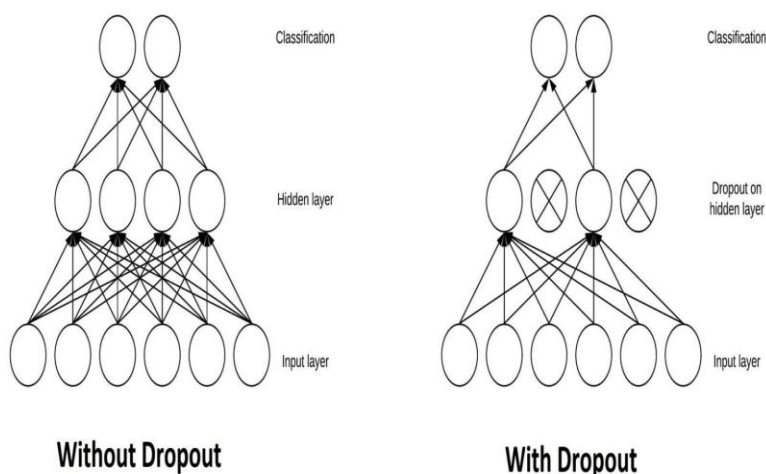
^۲Long short-term memory

^۳forget gate

حافظه کوتاه مدت بلند مدت از ناپدید شدن یا منفجر شدن خطاهای پس از انتشار جلوگیری می‌کند. این شبکه ها می‌تواند وظایفی را بیاموزند که به خاطراتی از رویدادهایی نیاز دارند که هزاران یا حتی میلیون‌ها گام زمانی گسسته پیشتر اتفاق افتاده‌اند.

شرح لایه های مختلف در شبکه های عصبی:

دراپ اوت^۳: در هر گام از آموزش به صورت تصدفی برخی از نورن ها را غیرفعال می کند این مسئله باعث آموزش مستقل از نورون شبکه می شود و همچنین از بیش برازش (اورفیت) شدن شبکه جلوگیری می کند.

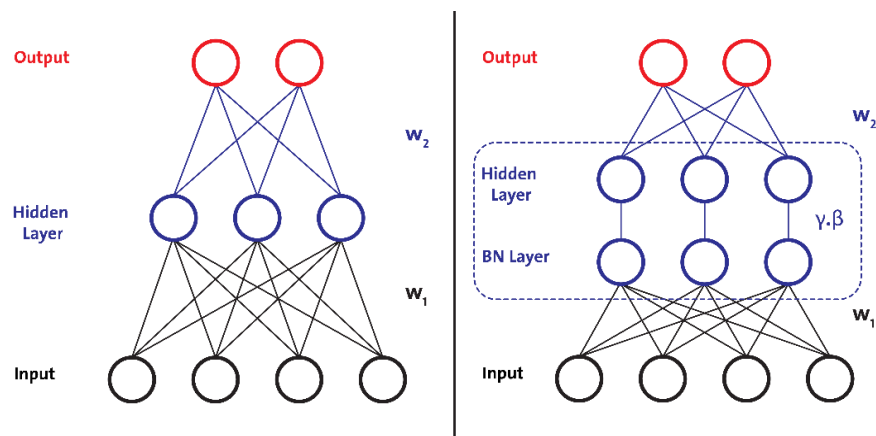


شکل ۳-۴ نحوه عملکرد لایه dropout را بیان می کند

بتج نرمالیز: بعد از اعمال وزن ها و قبل از تابع فعال سازی به صورت یک لایه در شبکه قرار میگیرد و داده ها را نرمالیز می کند. باعث می شود سرعت یادگیری افزایش پیدا کند.

^۳Dropout

^۴Batch normalize



شکل ۳-۵ جایگاه و نحوه قرارگیری لایه batch normalize در شبکه را نشان میدهد

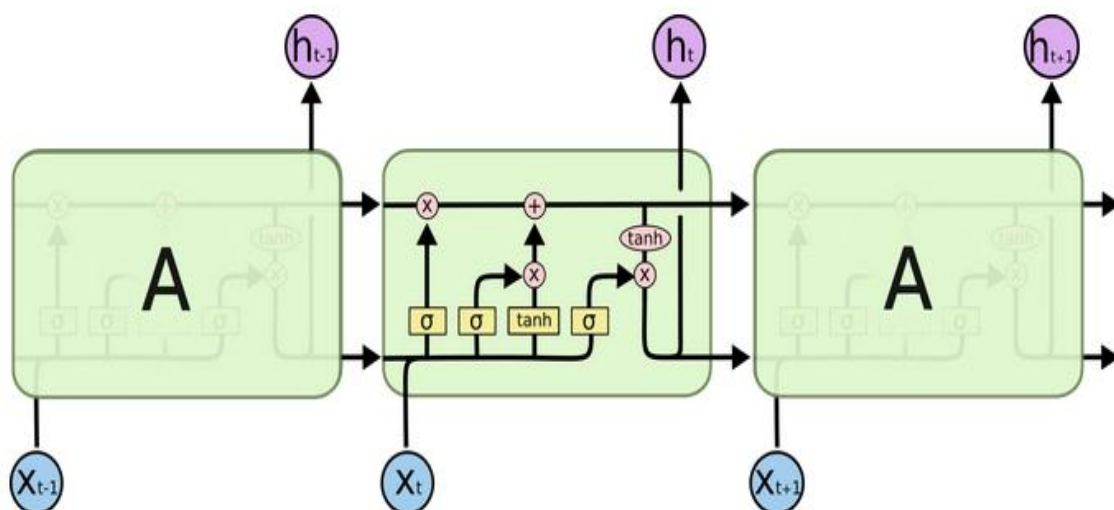
لایه دنس (متراکم)^۴ در هر شبکه عصبی، یک لایه متراکم به لایه ای گفته می شود که عمیقاً به لایه قبلی خود متصل است، به این معنی که نورون های لایه به هر نورون لایه قبلی خود متصل هستند. این لایه پرکاربردترین لایه در شبکه های عصبی مصنوعی است. نورون لایه متراکم در یک مدل از هر نورون لایه قبلی خود خروجی دریافت می کند، جایی که نورون های لایه متراکم ضرب ماتریس-بردار را انجام می دهند.

در زیر فرمول کلی ضرب ماتریس-بردار آورده شده:

$$\begin{aligned}
 \mathbf{Ax} &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
 &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} .
 \end{aligned}$$

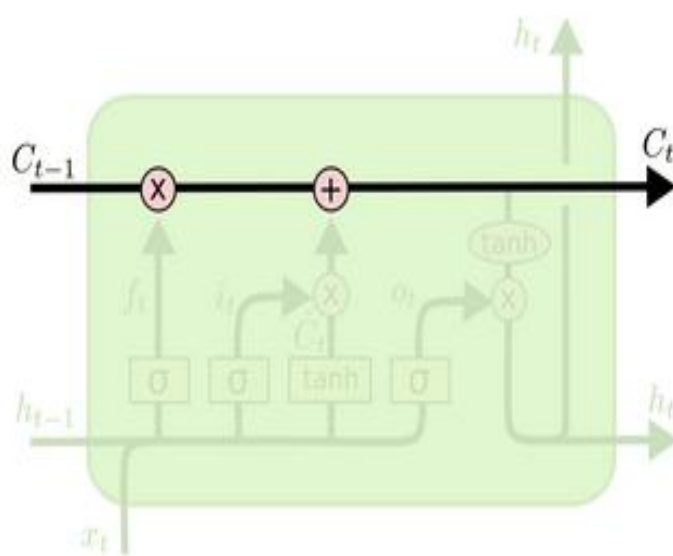
که در آن A یک ماتریس $(m \times n)$ x یک ماتریس $1 \times n$ است. مقادیر درون ماتریس پارامترهای آموزش دیده لایه های قبلی هستند و همچنین می توانند توسط پس انتشار به روز شوند. پس انتشار متداول ترین الگوریتم مورد استفاده برای آموزش شبکه های عصبی پیشخور است. به طور کلی، پس انتشار در یک شبکه عصبی، گرادیان تابع هزینه را با توجه به وزن های شبکه برای ورودی یا خروجی واحد محاسبه می کند.

لایه شبکه های حافظه کوتاه مدت : شبکه های حافظه کوتاه مدت به طور گسترده ای برای حل کارهای متوالی مختلف مورد استفاده قرار گرفته اند. بیایید دریابیم که این شبکه ها چگونه کار می کنند. این شبکه ها ساختار کمی پیچیده تر نسبت به شبکه های بازگشتی ساده دارند در ادامه به شرح کامل این شبکه می پردازیم:



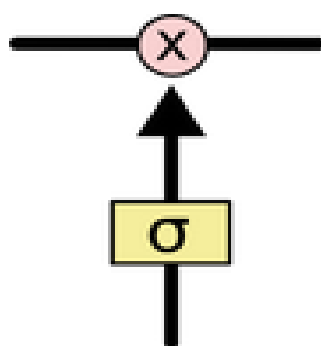
در شکل ۳-۶ ساختار یک نورن در شبکه ال اس تی ام نمایش داده شده است.

کلید ال اس تی ام ها وضعیت سلولی است، خط افقی که از بالای نمودار عبور می کند.



شکل ۳-۷ سلول حالت در شبکه lstm را نشان میدهد

سلول حالت به نوعی شبیه تسمه نقاله است. این به طور مستقیم در کل زنجیره اجرا می شود، تنها با برخی فعل و انفعالات خطی جزئی. بسیار آسان است که اطلاعات بدون تغییر در امتداد آن جریان یابد. اس تی ام توانایی حذف یا اضافه کردن اطلاعاتی را به وسیله که با دقت توسط ساختارهایی به نام گیت تنظیم می شود به سلول حالت را دارد، گیت ها راهی هستند که به صورت اختیاری اطلاعات را از خود عبور می دهند. آنها از یک لایه شبکه عصبی با تابع فعالسازی سیگموئید و یک عملیات ضرب نقطه ای تشکیل شده اند.



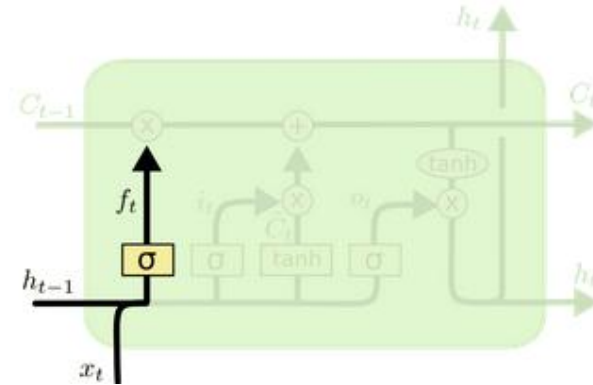
شکل ۳-۸ لایه سیگموئید در شبکه lstm را نشان میدهد

لایه سیگموئید اعدادی بین صفر و یک را خروجی می دهد و توضیح می دهد که چه مقدار از هر جزء باید عبور کند. مقدار صفر به معنای "نگذارید هیچ چیز عبور کند"، در حالی که مقدار یک به معنای "اجازه دهید همه چیز از بین برود!" می باشد.

یک LSTM دارای سه مورد از این دروازه ها برای محافظت و کنترل وضعیت سلول است.

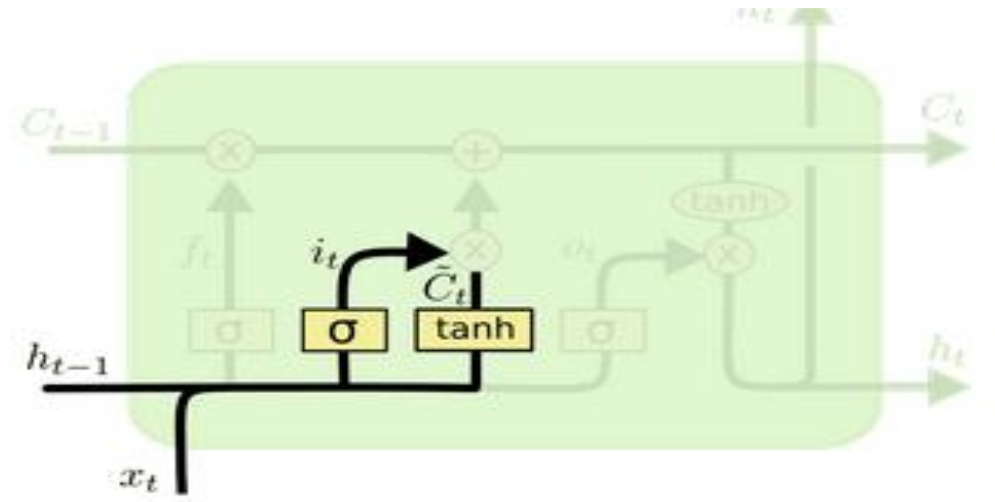
نحوه کار ال اس تی ام رو قدم به قدم بررسی کنیم:

اولین قدم در LSTM ما این است که تصمیم بگیریم چه اطلاعاتی را از سلول حالت دور بریزد. این تصمیم توسط یک لایه سیگموئید به نام "دروازه فراموش" گرفته می شود. به h_{t-1} و x_t نگاه می کند و برای هر عدد در سلول حالت C_{t-1} عددی بین ۰ و ۱ خروجی می دهد. عدد ۱ نشان دهنده "کاملاً این را حفظ کنید" در حالی که ۰ نشان دهنده "به طور کامل حذف شود".



$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f)$$

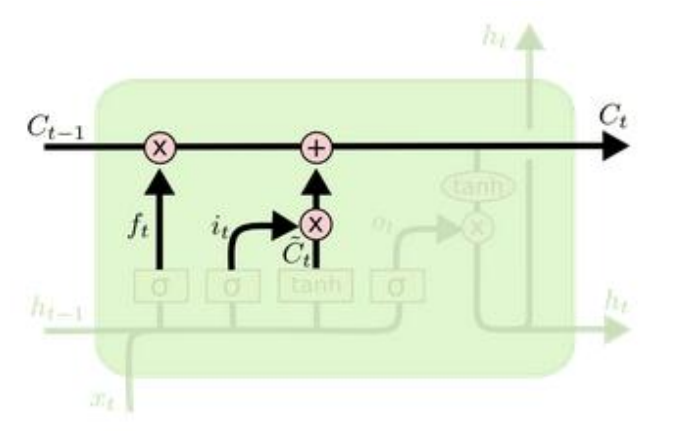
گام بعدی این است که تصمیم بگیریم چه اطلاعات جدیدی را در سلول حالت ذخیره کنیم. این گام دو بخش دارد. ابتدا، یک لایه سیگموئید به نام "لایه دروازه ورودی" تصمیم می گیرد که کدام مقادیر را به روز کنیم. سپس، یک لایه \tanh یک بردار از مقادیر کاندید جدید، \tilde{C}_t ایجاد می کند که می تواند به سلول حالت اضافه شود. در مرحله بعدی، این دو را با هم ترکیب می کنیم تا یک به روزرسانی برای سلول حالت ایجاد کنیم.



$$i_t = \sigma (W_i \cdot [h_{t-1}, x_t] + b_i)$$

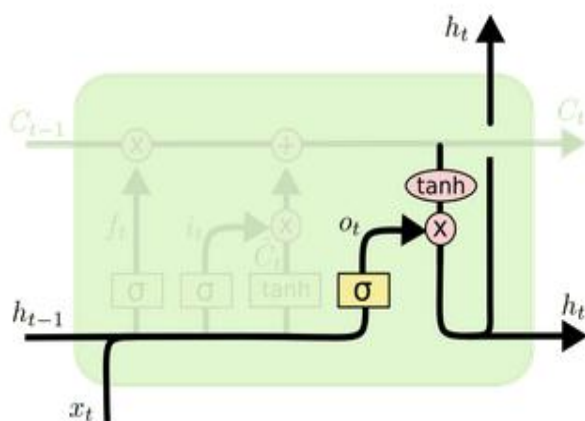
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

اکنون زمان آن فرا رسیده است که سلول حالت قدیمی، C_{t-1} ، به سلول حالت جدید C_t به روزرسانی شود. برای اینکار ابتدا سلول حالت را در f_t ضرب میکنیم با این کار مواردی را که میخواستیم فراموش کنیم از سلول حالت حذف میکنیم و سپس عبارت را با $(i_t * C_t)$ جمع می کنیم اکنون سلول حالت جدید بدست آمده است.



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

در نهایت، ما باید تصمیم بگیریم که چه چیزی را خروجی خواهیم داشت. این خروجی بر اساس سلول حالت ما خواهد بود، اما یک نسخه فیلتر شده خواهد بود. ابتدا یک لایه سیگموئید اجرا می کنیم که تصمیم می گیرد چه قسمت هایی از سلول حالت را خروجی بگیریم. سپس، سلول حالت را از طریق \tanh قرار می دهیم (برای فشار دادن مقادیر بین -1 و 1) و آن را در خروجی دروازه سیگموئید ضرب می کنیم، به طوری که فقط قسمت هایی را که تصمیم گرفتیم خروجی بگیریم.



$$o_t = \sigma (W_o [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh (C_t)$$

-معماری مدل شبکه عصبی ال اس تی ام

مدل LSTM یک نوع از خانواده ی شبکه های عصبی بازگشتی^۴ است. شبکه های عصبی بازگشتی تلاش می کنند تا زمان یا یک رفتاری که به ترتیب وابسته است را مدل بکنند. و ترفند این شبکه به همین منظور آن است که خروجی لایه ی شبکه عصبی را در هر قدم به عنوان ورودی همان لایه ی شبکه عصبی برای استفاده در قدم بعدی، داده شود.

مدل شبکه های عصبی بازگشتی و مدل GRU بیشترین کاربرد را در مدل های به اصطلاح state-of-the-art را امروزه دارند. از این مدل ها برای تشخیص گفتار^{۳۷}، تولید گفتار^{۳۸} و تولید متن (Text Generation) کاربرد دارد. به عنوان مثال در ایجاد زیر نویس برای یک فیلم می شود از آن استفاده کرد.

شبکه عصبی بازگشتی یک کلاس از شبکه های عصبی مصنوعی است که در آن اتصالات بین گره ها یک گراف جهت دار یا غیر جهت دار را در امتداد یک دنباله زمانی تشکیل می دهند. این مسئله به آن اجازه می دهد تا رفتار پویا زمانی را نشان دهند. شبکه های عصبی بازگشتی که از شبکه های عصبی پیشخور^{۳۹} مشتق شده اند، می توانند از حالت داخلی (حافظه) خود برای پردازش توالی با طول متغیر ورودی ها استفاده کنند.

ایجاد مدل شبکه های عصبی بازگشتی برای حل مشکلاتی بود که مدل های اولیه RNN داشتند که به یک مورد اشاره می شود :

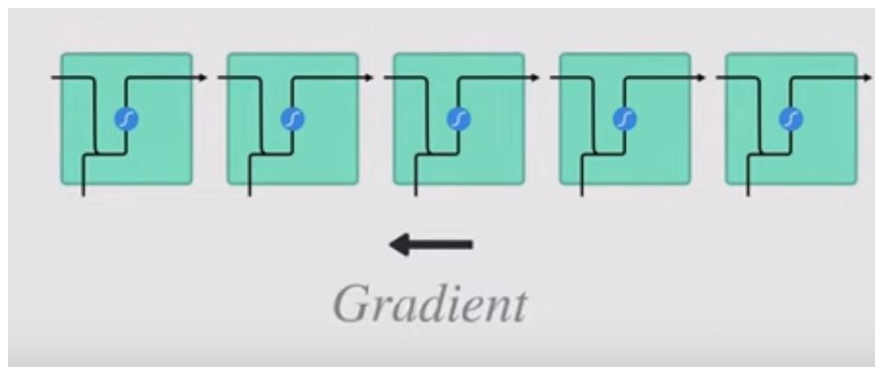
مشکل به صفر رسیدن این مشکل به این دلیل ایجاد می شود که زمانی که مدل در فاز یادگیری است و عمل backpropation انجام می شود تا لایه ها به ترتیب از آخرین لایه تا اولین لایه ها آموزش ببینند. رفته رفته مقدار gradient کاهش پیدا می کند. با کاهش مقدار gradient، لایه های اولیه نمی توانند یادگیری موثری نسبت به لایه های اولیه داشته باشند. برای حل این مشکل از مدل LSTM استفاده شده است.

^۴RNN

Speech Recognition

Speech Synthesis

^{۳۶}feedforward neural networks



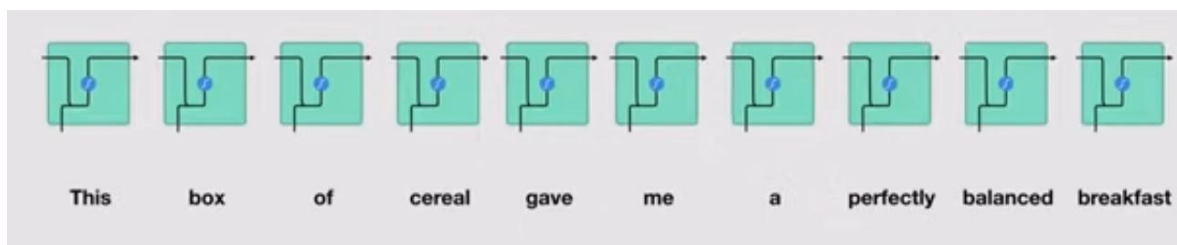
از نتایج منفی به صفر رسیدن مقدار **gradient**، این است که مدل دنباله های کوچکی را می تواند یاد بگیرد و مقدار نخواهد بود با متن هایی که شامل مقادیر زیادی ورودی کاراکتر به عنوان یک دنباله هستند، کار کند.

نحوه ی کار شبکه های عصبی بازگشتی:

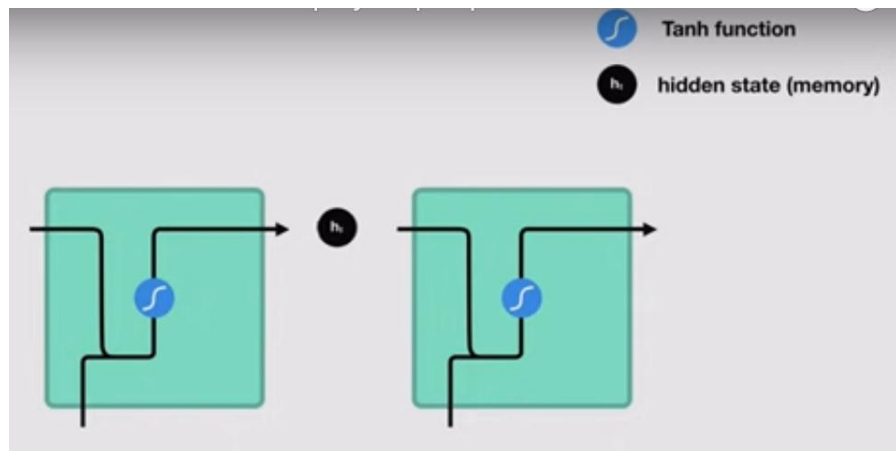
فرض کنید که می خواهیم بر مبنای متنی که کاربری در مورد محصولی نوشته است بفهمیم که آیا کاربر از خرید این کالا راضی است یا خیر؟

کاری که شبکه های عصبی بازگشتی انجام می دهد مانند کاری است که انسان انجام می دهد. انسان با به خاطر سپردن کلمه های کلیدی، مانند: این کالا خیلی خوب بود، قطعا دوباره آن را خواهیم خرید و کلمه های کلیدی دیگر و حذف موارد اضافی که اطلاعاتی در مورد میزان رضایت مشتری ندارند؛ به ذهن می سپارد که آیا کاربر از خرید محصول راضی است یا خیر.

حال برای آنکه بفهمیم معماری شبکه های عصبی بازگشتی چگونه کار می کند اول یک مدل ساده از RNN را بررسی میکنیم.

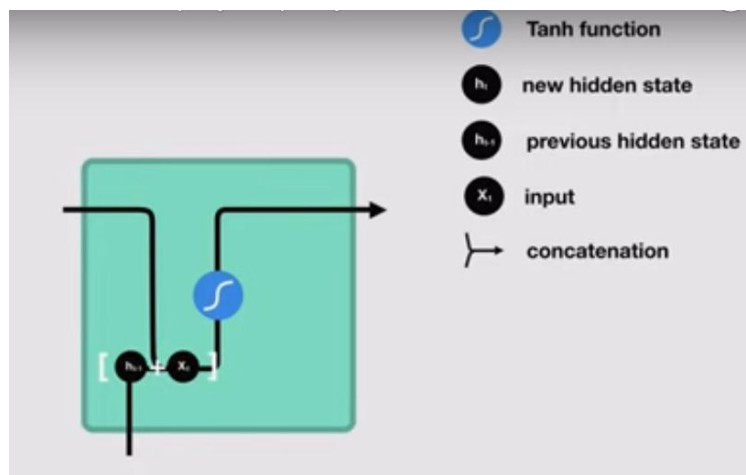


ابتدا سلول های RNN هر کدام یکی از کلمات جمله را به عنوان ورودی دریافت می کنند. این نکته باید ذکر شود که هر کلمه به صورت یک وکتور عددی به عنوان ورودی داده می شود.



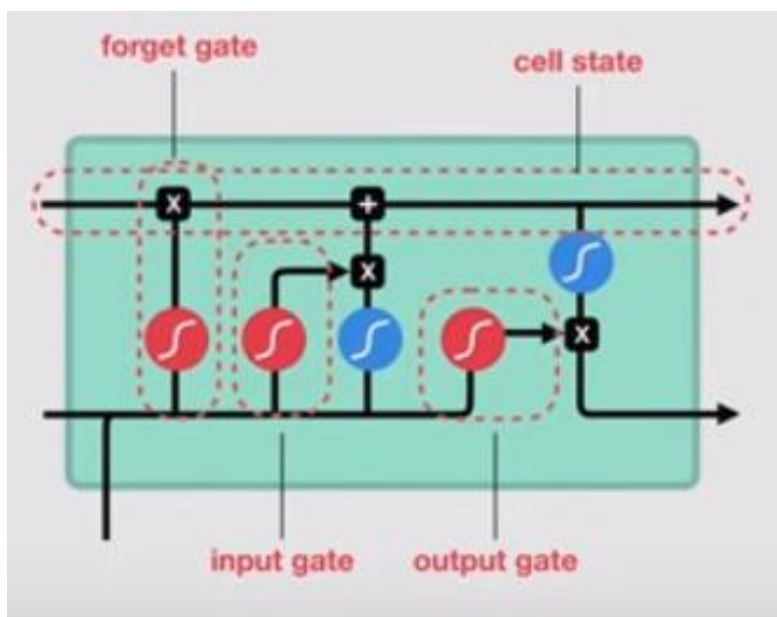
خروجی هر سلول RNN به عنوان وضعیت پنهان (Hidden State) به سلول بعدی داده می شود. و شامل اطلاعاتی است از کلمه های قبلی که توسط مدل دیده شده است. در اینجا کلمه هایی که تا رسیدن به این سلول خوانده شده است به عنوان یک وضعیت پنهان استفاده می شود.

اینکه چگونه وضعیت پنهان محاسبه می شود. به شکل زیر دقت شود :



ابتدا داده ی ورودی مثلا کلمه ی بارانی به ورودی به عنوان x داده می شود سپس این کلمه با وضعیت پنهان قبلی که از دیده شدن کلمه های اینجا هوا ترکیب شده و یک وکتور جدید می سازد. سپس این وکتور بعد از عبور از تابع تانژانت هایپربولیک مقداری بین -1 و 1 پیدا می کند. سپس این مقدار به عنوان وضعیت پنهان جدید به سلول دیگر RNN داده می شود.

حال برای درک مدل LSTM به شکل زیر نگاه شود:



سلول شبکه های عصبی بازگشتی از ۴ بخش مهم تشکیل شده است:

۱. گیت فراموشی
۲. گیت ورودی
۳. گیت خروجی
۴. وضعیت پنهانی

که در ادامه به توضیح هر گیت می پردازیم :

گیت فراموشی^۴:

گیت فراموشی به عنوان ورودی خود، ورودی یک کلمه مثل باران را با وضعیت پنهان از کلمه های قبلی دیده شده ترکیب می کند. سپس این ترکیب را به تابع سیگموئید خود می دهد. تابع سیگموئید ورودی خود را به اعدادی بین صفر و یک نگاشت می دهد. عدد صفر یعنی من این اطلاعات را می خواهم فراموشی کنم و عدد یک یعنی می خواهم این اطلاعات حذف شود. در ادامه مشاهده می شود که خروجی گیت فراموشی در وضعیت

سلول قبلی ضرب می شود. این همان ضربی است که صفر خروجی گیت فراموشی باعث حذف اطلاعات سلول وضعیت قبلی و ضرب شدن یک باعث ماندن آن اطلاعات می شود.

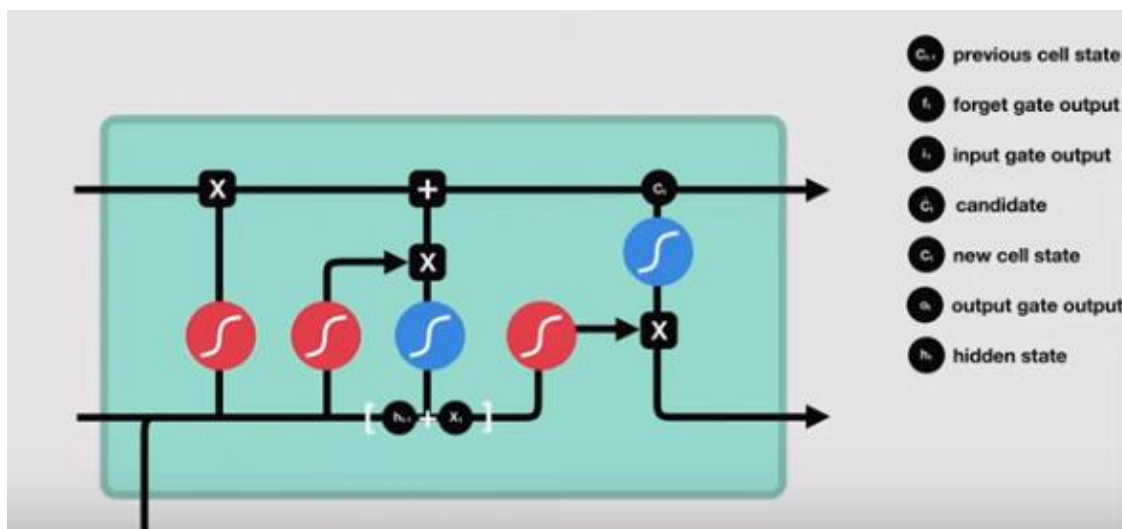
گیت ورودی:^{۴۲}

گیت ورودی شامل دو ورودی است به توابع آبی و قرمز رنگ در شکل بالا. ورودی به تابع سیگموید مثل گیت فراموشی اطلاعاتی که باید حذف شوند را مشخص می کند. تابع آبی رنگ، تابع تانژانت هایپربولیک است که باعث میشود ورودی بین ۱ و ۱- قرار بگیرد. این تابع باعث می شود تا یادگیری به درستی انجام شود. سپس خروجی هر دو تابع قرمز و آبی در هم ضرب می شود تا فقط اطلاعات مهم بماند. این اطلاعات که حالا اطلاعات مهم ترکیب شده از ورودی و خلاصه ای از اطلاعات وضعیت مخفی کلمه های قبلی متن است با وضعیت سلول قبلی ترکیب می شود.

گیت خروجی:

وظیفه ی گیت خروجی؛ محاسبه ی وضعیت پنهان جدید در خروجی خود است. که این کار را از حذف اطلاعات بی ارزش ترکیب ورودی و وضعیت پنهانی قبلی با استفاده از تابع سیگموید و ضرب کردن آن در وضعیت سلول قبلی بدست می آورد.

نتیجه ی همکاری این گیت ها با یکدیگر باعث می شود تا تنها اطلاعات مهم باقی بمانند و سلول وضعیت قبلی بتواند مانند حافظه اطلاعاتی که از خواندن یک متن بدست آورده را در خود ذخیره کند.



ما در این مثال از یک شبکه عصبی LSTM ساده که شامل ۱۶ واحد (unit) لایه ی LSTM می شود استفاده کرده ایم. تعداد سلول های LSTM در هر واحد یا یونیت استفاده کردیم به صورت خودکار از روی ابعاد ورودی داده به مدل مشخص می شود. چون در مثال ما پنجره ی زمانی ۲۴ است در هر یک از ۱۶ واحد، ۲۴ سلول LSTM وجود دارد.

در نهایت ما یک لایه ی از نوع dense layer به منظور انجام عمل پیش بینی به معماری خود اضافه می کنیم. که شامل یک واحد است و این واحد یک پیش بینی را تولید می کند که می توان آن را با یک برجسب مقایسه کرد.

ما همچنین نیاز داریم که مقدار return_sequence را برابر True قرار دهیم. کاری که این پارامتر میکند آن است که سلول های موجود در هر واحد یا یونیت که آن ها نیز در لایه ی LSTM هستند. همه ی خروجی های خود را به یونیت های دیگر بدهند.

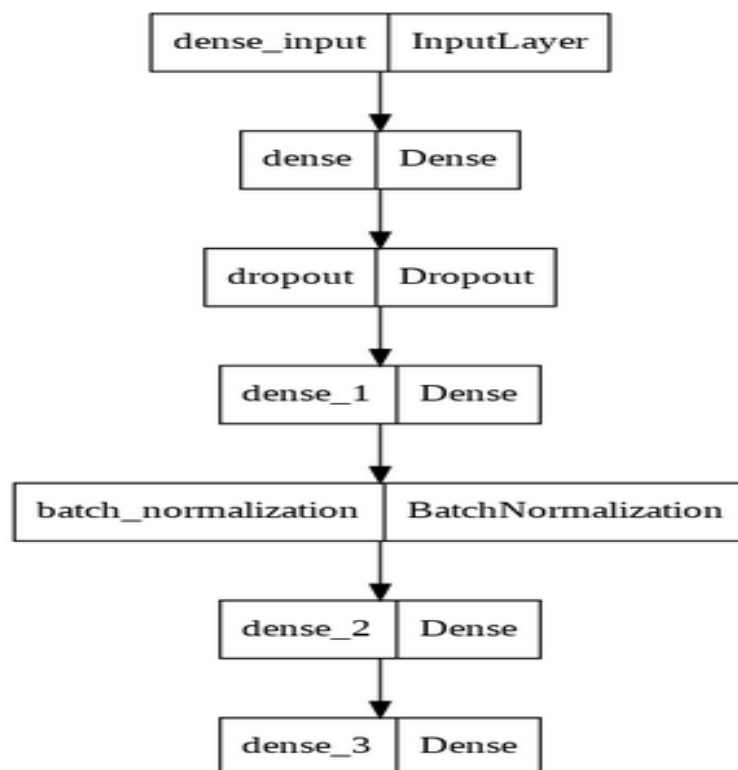
تابع loss، تابع بهینه ساز و متریک های مورد استفاده در مدل : تابع loss ما MSE یا Mean Squared Error است که به معنی به توان ۲ رساندن اختلاف بین پیش بینی مدل و برجسب است.

برای تابع بهینه ساز adam استفاده میکنیم.

متریک ما برای اندازه گیری اختلاف بین پیش بینی و برجسب absolute error یا همان تفاضل ریاضی بین پیش بینی و برجسب می باشد.

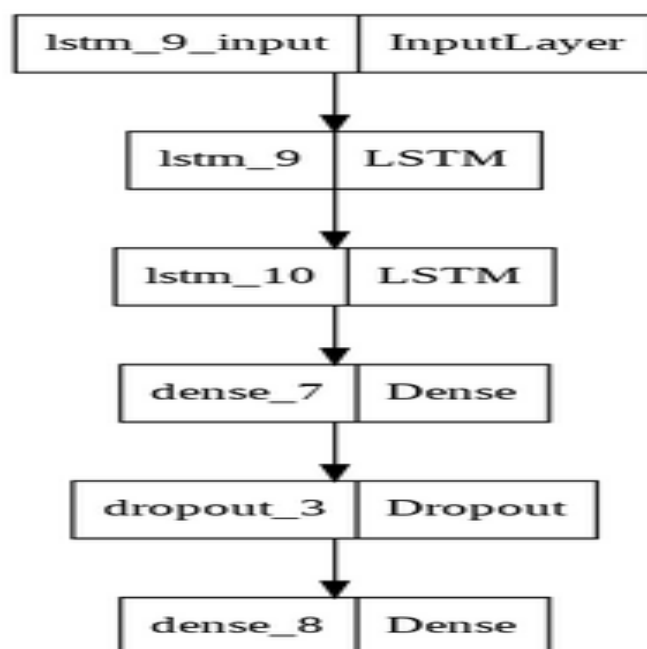
در این پژوهش دو شبکه عصبی با کمک کتابخانه کراس برای پیش بینی قیمت بست شدن بیت کوین در هر روز آموزش داده و سپس دقت آنها را مورد ارزیابی قرار داده ایم شبکه عصبی ابتدا یک شبکه عصبی neural network(ann) و پس از آن از شبکه عصبی LSTM استفاده شده است. که از ۱۰۰ دوره نسل آموزش یا epochs را داریم و با نمودار history روند تابع loss را بررسی کرده ایم که این روند باید نزولی باشد. که ساختار هر کدام از این شبکه ها در ادامه آورده شده است:

شبکه اول: شامل ۵ لایه متراکم(dense) می باشد و ۱ لایه دارپ اوت (dropout)، یک لایه بتچ نرمالیز (batche normailze) می باشد.



شکل ۳-۸ معماری شبکه اول را نشان میدهد.

شبکه دوم: سه لایه ال اس تی ام (lstm) و یک لایه دارپ اوت (dropout) و دو لایه متراکم (dense) را شامل می شود.



شکل ۳-۹ معماری شبکه دوم را نشان میدهد.

آدام^{۴۳}

یک الگوریتم بهینه‌سازی است که می‌تواند به جای روش کلاسیک گرادیان نزولی تصادفی برای به‌روزرسانی وزن‌های شبکه بر اساس داده‌های آموزشی استفاده شود. آدام توسط Diederik Kingma از OpenAI و Jimmy Ba از دانشگاه تورنتو در مقاله (پوستر) ICLR ۲۰۱۵ خود با عنوان "Adam: A Method for Stochastic Optimization" ارائه شد، مزایای جذاب استفاده از Adam در مسائل بهینه‌سازی غیر محدب توسط توسعه دهندگان آن به شرح زیر بیان شده است:

۱. پیاده‌سازی ساده
۲. کارآمد برای محاسبات
۳. به حافظ کمی نیاز دارد
۴. برای اهداف غیر ثابت مناسب است.
۵. فرایامترها تفسیر بصری دارند و معمولاً نیاز به تنظیم کمی دارند

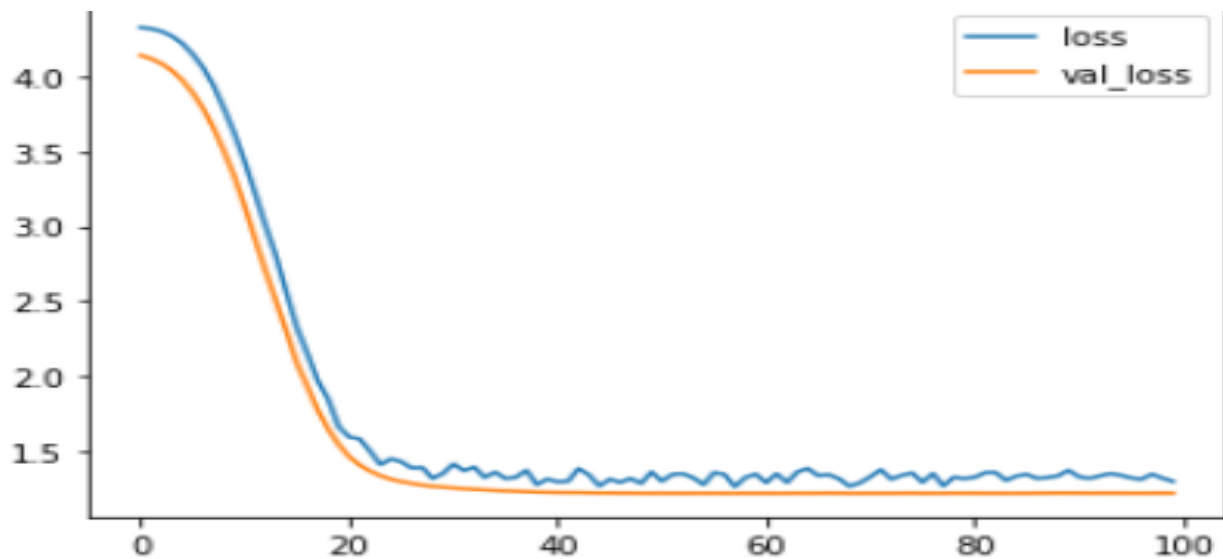
^{۴۳}adam

خطای میانگین مربعات^{۴۴}

در ریاضیات و آمار، خطای میانگین مربعات روشی برای برآورد میزان خطاست که در واقع تفاوت بین مقادیر تخمینی و آنچه تخمین زده شده، است.

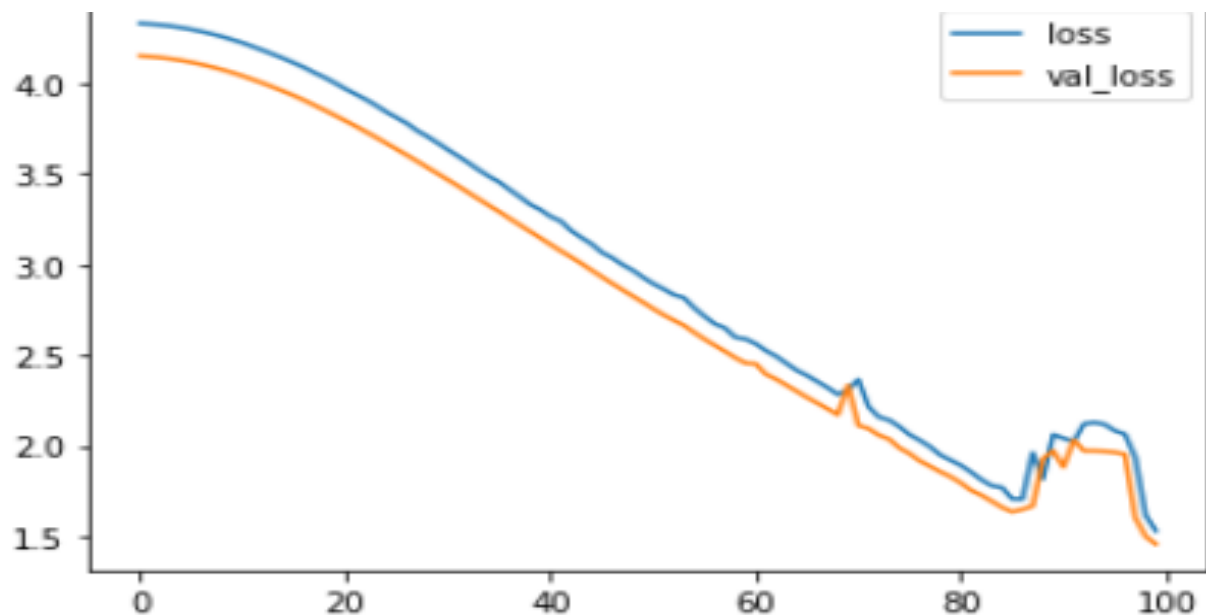
$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

ما در این تحقیق از تابع هزینه MSE و تابع بهینه سازی adam برای هر دو شبکه عصبی که توسعه داده ایم استفاده کرده ایم سپس هر شبکه در ۱۰۰ اپیاک آموزش دیده که در ادامه نمودار روند تغییر مقدار تابع هزینه برای هر دو شبکه نمایش داده شده:



شکل ۳-۱۰ نمودار روند تغییر مقدار تابع هزینه برای شبکه اول می باشد

^{۴۴}Mean squared error



شکل ۳-۱۱ نمودار روند تغییر مقدار تابع هزینه برای شبکه دوم می باشد

۴-نتیجه گیری:

پیش بینی زمانی یک روش خیلی قدرتمند ولی با محدودیت هایی نیز همراه است. همیشه گذشته اطلاعات کافی برای پیشبینی آینده به ما نمی دهد.

در ادامه به شرح معیار های ارزیابی که در این پژوهش مورد استفاده قرار داده ایم می پردازیم و مقدار هر کدام را این معیار ها را برای هر شبکه به تفصیل بیان خواهیم کرد:

R2_score:

یکی از معیارهای ارزیابی عملکرد برای مدل های یادگیری ماشین مبتنی بر رگرسیون است. به ضریب تعیین نیز معروف است. این معیار با اندازه گیری میزان واریانس در پیش بینی های توضیح داده شده توسط مجموعه داده کار می کند. به بیان ساده، تفاوت بین نمونه های موجود در مجموعه داده و پیش بینی های انجام شده توسط مدل است.

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

:Mean absolute error

در آمار، میانگین خطای مطلق معیاری از خطاهای بین مشاهدات زوجی است که یک پدیده را بیان می کنند.

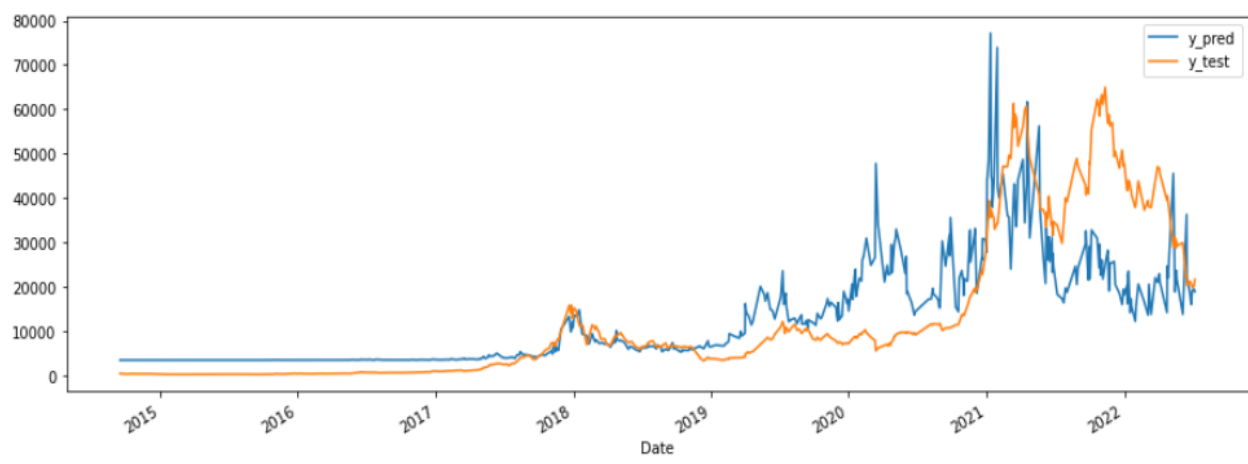
$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}$$

و معیار mean squared error مقدار هر معیار برای شبکه اول به شرح زیر می باشد:

R2 score : 0.543

Mean squared error : 1.861

Mean absolute error : 7175.270



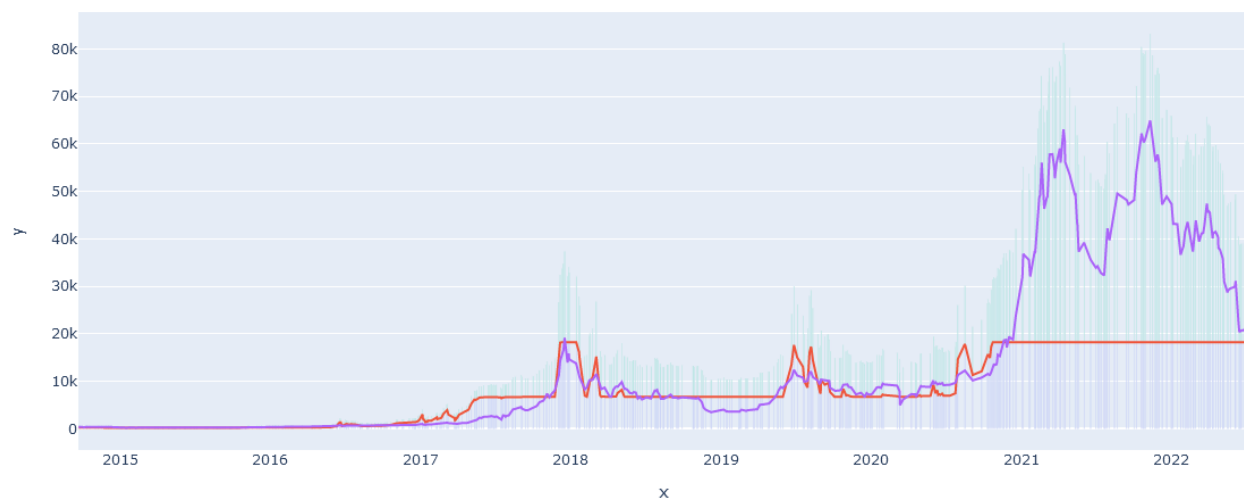
شکل ۴-۱ اختلاف بین مقدار پیش بینی شده توسط شبکه اول از مقادیر واقعی را نشان میدهد (خط نارنجی نشان دهنده مقدار مورد انتظار و خط آبی مقداری که مدل پیش بینی کرده است)

مقدار هر معیار برای شبکه دوم:

R2 score : 0.458

Mean squared error : 0.312

Mean absolute error : 5973.618



شکل ۴-۲ اختلاف بین مقدار پیش بینی شده توسط شبکه دوم از مقادیر واقعی را نشان می‌دهد (خط بنفش نشان دهنده مقدار مورد انتظار و خط قرمز مقداری که مدل پیش بینی کرده است)

پیشنهادهای آتی تحقیق:

محققان می‌توانند در آینده از مدل سری‌های زمانی آریما برای پیش‌بینی استفاده کنند. روند آن را با پیش‌بینی تحقیق حاضر مقایسه کنند.

- [1] Dimson, E., & Mussavian, M. (1998). A brief history of market efficiency. *European financial management*, 4(1), 91-103.
- [2] F. E. T. Burton and S. N. Shah, "Efficient market hypothesis," *CMT Level I 2017: An Introduction to Technical Analysis*, 2017.
- [3] Zhao, Z., Rao, R., Tu, S., & Shi, J. (2017, November). Time-weighted LSTM model with redefined labeling for stock trend prediction. In *2017 IEEE 29th international conference on tools with artificial intelligence (ICTAI)* (pp. 1210-1217). IEEE
- [4] Hardesty, L. (2017). Explained: neural networks. *MIT News*, 14.
- [5] Dupond, Samuel (2019). "A thorough review on the current advance of neural network structures". *Annual Reviews in Control*. 14: 200–230.
- [6] Abiodun, Oludare Isaac; Jantan, Aman; Omolara, Abiodun Esther; Dada, Kemi Victoria; Mohamed, Nachaat Abdelatif; Arshad, Humaira (2018-11-01). "State-of-the-art in artificial neural network applications: A survey". *Heliyon*. 4 (11): e00938.
- [7] Tealab, Ahmed (2018-12-01). "Time series forecasting using artificial neural networks methodologies: A systematic review". *Future Computing and Informatics Journal*. 3 (2): 334–340.
- [8] Hytyniemi, Heikki (1996). "Turing machines are recurrent neural networks". *Proceedings of STeP '96/Publications of the Finnish Artificial Intelligence Society*: 13–24.

[9] Gers, Felix A.; Schraudolph, Nicol N.; Schmidhuber, Jürgen (2002). "Learning Precise Timing with LSTM Recurrent Networks" (PDF). *Journal of Machine Learning Research*. 3: 115–143. Retrieved 2017-06-13.

[10] Hochreiter, Sepp (1991), *Untersuchungen zu dynamischen neuronalen Netzen*, Diploma thesis, Institut f. Informatik, Technische Univ. Munich, Advisor Jürgen Schmidhuber

[11] Schmidhuber, Jürgen (January 2015). "Deep Learning in Neural Networks: An Overview". *Neural Networks*. 61: 85–117.