



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده مهندسی کامپیوتر

پروژه کارشناسی
گرایش هوش مصنوعی

پیش‌بینی نوسان قیمت در بازار رمزارزها با استفاده از
یادگیری عمیق

پایان‌نامه

نگارش
سید علیرضا بختیاری

استاد راهنما
دکتر مریم امیرمزلقانی

تیر ۱۴۰۰

به نام خدا

صفحه فرم ارزیابی و تصویب پایان نامه- فرم تأیید اعضای کمیته دفاع

در این صفحه فرم دفاع یا تأیید و تصویب پایان نامه موسوم به فرم کمیته دفاع- موجود در پرونده آموزشی- را قرار دهید.

نکات مهم:

- نگارش پایان نامه/رساله باید به **زبان فارسی** و بر اساس آخرین نسخه دستورالعمل و راهنمای تدوین پایان نامه های دانشگاه صنعتی امیرکبیر باشد.(دستورالعمل و راهنمای حاضر)
- رنگ جلد پایان نامه/رساله چاپی کارشناسی، کارشناسی ارشد و دکترا باید به ترتیب مشکی، طوسی و سفید رنگ باشد.
- چاپ و صحافی پایان نامه/رساله بصورت **پشت و رو(دورو)** بلامانع است و انجام آن توصیه می شود.

به نام خدا



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

تعهدنامه اصالت اثر

تاریخ: تیر ۱۴۰۰

اینجانب سید علیرضا بختیاری متعهد می‌شوم که مطالب مندرج در این پایان‌نامه حاصل کار پژوهشی اینجانب تحت نظارت و راهنمایی اساتید دانشگاه صنعتی امیرکبیر بوده و به دستاوردهای دیگران که در این پژوهش از آن‌ها استفاده شده است مطابق مقررات و روال متعارف ارجاع و در فهرست منابع و مآخذ ذکر گردیده است. این پایان‌نامه قبلاً برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نگردیده است. در صورت اثبات تخلف در هر زمان، مدرک تحصیلی صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پیگیری قانونی خواهد داشت.

کلیه نتایج و حقوق حاصل از این پایان‌نامه متعلق به دانشگاه صنعتی امیرکبیر می‌باشد. هرگونه استفاده از نتایج علمی و عملی، واگذاری اطلاعات به دیگران یا چاپ و تکثیر، نسخه‌برداری، ترجمه و اقتباس از این پایان‌نامه بدون موافقت کتبی دانشگاه صنعتی امیرکبیر ممنوع است. نقل مطالب با ذکر مآخذ بلامانع است.

سید علیرضا بختیاری

امضا

پاس‌گزاری

اکنون که مراحل پژوهش، تدوین و نگارش پایان نامه به پایان رسیده است، از مادر و پدر عزیزتر از جانم متشکرم که درطول زندگی و دوران تحصیل همراه و مشوقم بوده‌اند و با ایثار و از خودگذشتگی و تحمل زحمات، مرا در این راه یاری نمودند. از سرکار خانم دکتر مزلقانی که به عنوان استاد راهنما با سعه صدر در مسیر این پژوهش، همواره راهنما و راهگشای اینجانب بوده‌اند تقدیر و تشکر می‌نمایم.

سیدعلیرضا، تنیاری

تیر ۱۴۰۰

چکیده

هدف از انجام این پروژه پیش‌بینی نوسان قیمت در بازار رمزارزها با استفاده از اطلاعات دفتر سفارشات و ویژگی‌های استخراج شده از آن و شبکه‌های عصبی بازگشتی است. پیش‌بینی نوسان در بازارهای مالی به طور کلی از اهمیت زیادی برخوردار است. به صورت سنتی تنها از اطلاعات و تاریخچه‌ی قیمت برای محاسبه‌ی نوسان در مدل‌های آماری استفاده شده‌است. در حالی که استفاده از اطلاعات موجود در دفتر سفارشات که شامل تمامی سفارشات موجود و در جریان یک بازار است می‌تواند دقت این مدل‌ها را افزایش دهد. اما به دلیل حجم زیاد دفتر سفارشات، در گذشته کمتر از آن در جهت پیش‌بینی نوسان استفاده شده‌است.

در این پروژه با واکشی و پیش‌پردازش اطلاعات دفتر سفارشات و سپس با استفاده از مدل‌های یادگیری شبکه‌ی عصبی بازگشتی، سعی در مدل کردن نوسان در بازارهای رمزارزها را داریم. نتایج به دست آمده نشان‌دهنده‌ی برتری مدل‌های یادگیری شبکه‌ی عصبی بازگشتی در مقایسه با دیگر معماری‌های شبکه‌های عصبی در پیش‌بینی نوسان است.

واژه‌های کلیدی:

پیش‌بینی سری‌های زمانی، نوسان قیمت، شبکه‌های عصبی، یادگیری عمیق، شبکه‌های عصبی بازگشتی، دفتر ثبت سفارشات

فهرست مطالب

عنوان

صفحه

۱	مقدمه، تعریف مساله، راه حل پیشنهادی	۱
۱-۱	مقدمه	۲
۲-۱	تعریف مساله	۳
۱-۲-۱	رمزارزها	۳
۲-۲-۱	دفتر سفارشات	۴
۳-۲-۱	نوسان قیمت در رمزارزها	۴
۳-۱	راه حل پیشنهادی	۵
۴-۱	بخش بندی پایان نامه	۶
۲	پیش زمینه و کارهای مرتبط	۷
۱-۲	مقدمه	۸
۲-۲	نوسان	۸
۱-۲-۲	تعریف نوسان	۸
۳-۲	مدل های نوسانی	۹
۱-۳-۲	مدل های گارچ	۱۰
۲-۳-۲	دفتر سفارشات	۱۳
۳-۳-۲	جمع بندی	۱۵
۳	روش پیشنهادی	۱۶
۱-۳	مقدمه	۱۷
۲-۳	استخراج ویژگی و سری زمانی از دفتر سفارشات	۱۷
۱-۲-۳	دفتر سفارشات تجمعی	۱۸
۲-۲-۳	استخراج ویژگی های دیگر از دفتر سفارشات	۱۹
۳-۲-۳	استاندارد سازی ویژگی ها	۲۱
۳-۳	یادگیری	۲۱
۱-۳-۳	شبکه های عصبی بازگشتی	۲۱
۲-۳-۳	جی آریو	۲۴

۲۵	۳-۳-۳ تابع خطا
۲۸	۴-۳ جمع بندی
۲۹	۴ پیاده سازی و کتابخانه ها
۳۰	۱-۴ مقدمه
۳۰	۲-۴ واکشی و پیش پردازش دفتر ثبت سفارشات
۳۰	۱-۲-۴ واکشی دفتر ثبت سفارشات
۳۳	۳-۴ مدل و یادگیری
۳۴	۴-۴ جمع بندی
۳۵	منابع و مراجع
۳۶	پیوست

شکل	فهرست اشکال	صفحه
۱-۱	روند کلی استخراج و نحوه کار رمزارزها (؟) .	۳
۲-۱	نمونه‌ای از دفتر سفارشات مربوط به بیتکوین .	۴
۱-۲	نمودار ساختار دفتر سفارشات به همراه برخی ویژگی‌های مربوط (؟) .	۱۳
۱-۳	نحوه جداسازی و ذخیره‌ی دفتر سفارشات .	۱۷
۲-۳	اندیس گذاری دفتر سفارشات برای ایجاد دفتر سفارشات تجمعی .	۱۸
۳-۳	نمونه‌ای از دفتر سفارشات معمولی و دفتر سفارشات تجمعی متناظر (؟) .	۱۹
۴-۳	شمای کلی شبکه‌های عصبی بازگشتی .	۲۲
۵-۳	نمایی از روابط ۷-۳ و ۸-۳ که نحوه محاسبات انجام شده درون هر سلول شبکه‌ی عصبی بازگشتی را نشان می‌دهند .	۲۲
۶-۳	ساختار داخل هر سلول از شبکه‌ی عصبی جی‌آریو .	۲۴
۷-۳	نمودار تابع میانگین مربعات خطا بر اساس میزان خطا یا ϵ .	۲۶
۸-۳	نمودار میانگین قدرمطلق خطا بر اساس میزان خطا یا ϵ .	۲۶
۱-۴	اطلاعات ورودی به ازای هر لحظه‌ی t در طول زمان .	۳۲
۲-۴	اطلاعات ورودی به ازای هر لحظه‌ی t در طول زمان .	۳۳
۳-۴	وارد شدن اطلاعات ورودی به جی‌آریو برای تخمین نوسان .	۳۳

صفحه

فهرست جداول

جدول

فصل اول

مقدمه، تعریف مساله، راه حل پیشنهادی

۱-۱ مقدمه

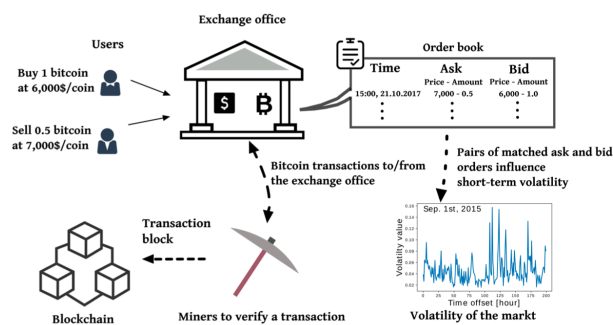
با ظهور کامپیوتر و علم رایانه بازارهای مالی جهان به سمت الکترونیکی شدن پیش رفتند. امروزه تقریباً تمام بازارهای مالی و حتی گاهی غیرمالی در بستر کامپیوتری و هوشمند فعال هستند تا جایی که بازارهایی مانند نزدک^۱ یا بازارهای مربوط به رمزارزها^۲ تنها در این بسترها قابل دسترسی می‌باشند و هیچ مکان فیزیکی‌ای برای مبادله درون این بازارها وجود ندارد.

تعداد زیادی از فعالان اقتصادی در این بازارها فعالیت روزانه خود را دنبال می‌کنند که محور اصلی فعالیت آن‌ها پیش‌بینی ویژگی‌های مختلف بازار مانند قیمت دارایی‌ها، میزان ریسک سرمایه‌گذاری در یک دارایی، تغییرات و نوسان تغییرات قیمت است. اما همان‌طور که انتظار می‌رود این زمینه نیز از پیشرفت هوش مصنوعی بی‌نصیب نمانده‌است و دستخوش تغییرات زیادی شده است. یکی از این تغییرات به وجود آمدن عوامل خودکار خرید و فروش است. عوامل هوشمندی که با تکیه بر ابزارها و روش‌های از پیش تعیین شده اقدام به خرید و فروش در این بازارها می‌کنند. پیش‌بینی سری‌های زمانی نقشی کلیدی در فعالیت تمامی این عوامل دارد، که باعث شده است مطالعات گسترده‌ای بر روی سری‌های زمانی مالی صورت بگیرد.

از ویژگی‌هایی که همواره مورد توجه فعالان این بازارها بوده است محاسبه و پیش‌بینی میزان نوسان قیمت و نه خود آن در بازارهای مالی در بازه‌های زمانی مشخص می‌باشد. داشتن تخمینی از نوسان آینده‌ی بازار و همین‌طور میزان تغییرات قیمت در ساعات آینده می‌تواند استراتژی‌های خرید و فروش هر عاملی، خواه هوشمند و خواه انسانی، را به کلی دستخوش تغییر کند. علاوه بر آن، نوسان یک دارایی از اهمیت بسیاری در جهت محاسبه‌ی سود و امکان واگذاری اوراق بهادار مشتقه‌ی آن برخوردار است. از این رو تحقیقات زیادی در مورد تخمین میزان نوسانات بازارهای مالی انجام شده است که هر یک سعی کرده‌اند از منابع داده‌ی گوناگون، از تغییرات قیمت تا کامنت‌های شبکه‌های اجتماعی استفاده کنند.

¹NASDAQ

²Cryptocurrency



شکل ۱-۱: روند کلی استخراج و نحوه کار رمزارزها (؟).

۲-۱ تعریف مساله

۱-۲-۱ رمزارزها

ارز دیجیتال یا رمزارز نوعی پول مجازی است که از فناوری رمزنگاری استفاده می‌کند و معمولاً به صورت غیرمتمرکز اداره می‌شود. رمزارزها میتوانند مانند سایر ارزهای بدون پشتوانه برای مبادله، انجام تراکنش، خرید آنلاین و ... مورد استفاده قرار بگیرند.

بیتکوین^۳ اولین پیاده سازی موجود از این رمزارزهاست که هم اکنون از بیشترین ارزش بازار در میان تمامی رمزارزها برخوردار است (؟). بیتکوین بدون مدیریت مرکزی عمل می‌کند و از طریق اینترنت و تکنولوژی شبکه‌ی بلوکی^۴ خود پرداخت‌های بین کاربران را ذخیره و تایید می‌کند. هر پرداخت جدید به وسیله‌ی بیت کوین باید به اطلاع شبکه‌ی موجود رسانده شود و یک گزارش از آن در زنجیره‌ی بلوکی بیتکوین ذخیره شود. برای تایید و ذخیره شدن در زنجیره‌ی بلوکی اصلی، تراکنش‌ها باید توسط رایانه‌هایی که وظیفه‌ی تایید تراکنش‌ها را دارند تایید شوند. این رایانه‌ها استخراج‌کننده^۵ نامیده می‌شوند و در ازای عملیات تایید مقداری بیتکوین به عنوان پاداش دریافت می‌کنند.

علاوه بر انتقال مستقیم بیتکوین در زنجیره‌ی بلوکی، مبادله‌ی بیتکوین و دیگر رمزارزها در صرافی‌های آنلاین نیز صورت می‌گیرد. دفتر سفارشات^۶ داده ساختاریست که در این صرافی‌ها نگهداری می‌شود و شامل سفارشات خرید و فروش است (۱).

³Bitcoin

⁴Blockchain

⁵Miner

⁶Order book

SUM	TOTAL	SIZE (BTC)	BID (USD)	ASK (USD)	SIZE (BTC)	TOTAL	SUM
12657.1972	12657.1972	2.498	5067.26900000	5071.50000000	2.500	12678.7500	12678.7500
14181.7535	1524.5563	0.301	5064.67100000	5073.28600000	0.301	1527.1495	14205.8995
46750.4048	32568.6514	6.433	5062.84300000	5073.43400000	3.765	19103.6771	33309.5766
59403.0298	12652.6250	2.500	5061.05000000	5073.43700000	0.010	51.3620	33360.9386
84703.9298	25300.9000	5.000	5060.18000000	5073.78000000	7.500	38053.3500	71414.2886
122653.8548	37949.9250	7.500	5059.99000000	5074.06900000	1.500	7611.1035	79025.3921
122664.3068	10.4520	0.002	5059.28600000	5074.96500000	2.988	15164.8536	94190.2457
131337.3008	8672.9940	1.715	5057.75900000	5075.40000000	5.000	25377.0000	119567.2457
257637.3258	126300.0250	25.000	5052.00100000	5077.00000000	1.258	6387.5962	125954.8419
257676.9463	39.6205	0.008	5052.00000000	5079.14300000	1.500	7618.7145	133573.5564
273.518 BTC			1305280.554 USD	1705516.579 USD			307.754 BTC

شکل ۱-۲: نمونه‌ای از دفتر سفارشات مربوط به بیتکوین.

۱-۲-۲ دفتر سفارشات

دفتر سفارشات داده ساختاریست که نگهدارنده‌ی وضعیت فعلی بازار و سفارشات خریداران و فروشندگان است. هر سفارش خرید^۷ شامل دو عدد حجم و قیمت است، که بیان کننده‌ی تقاضای خرید به مقدار حجم مشخص شده از آن دارایی در قیمتی کمتر یا مساوی با قیمت مشخص شده‌اند. در مقابل هر تقاضای فروش^۸ نیز نشان دهنده‌ی تقاضای فروش به مقدار حجم مشخص شده از آن دارایی در قیمتی بیشتر و یا مساوی با قیمت مشخص شده است. معامله و جابه‌جایی دارایی تنها هنگامی صورت می‌گیرد که قیمت پیشنهادی خریدار بیشتر و یا مساوی با قیمت پیشنهادی فروشنده باشد، در چنین شرایطی به میزان حجم مشخص شده از دارایی فروشنده به دارایی خریدار انتقال صورت می‌گیرد و هر دو سفارش از دفتر سفارشات خارج می‌شوند.

۱-۲-۳ نوسان قیمت در رمزارزها

به دلیل استقبال گسترده از رمزارزها در سال‌های گذشته، مطالعات زیادی بر روی ویژگی‌های آماری بازار رمزارزهای مختلف انجام شده است. یکی از زمینه‌های اصلی این مطالعات نوسان قیمت در این بازارها بوده‌است. بر خلاف افزایش اندازه بازار رمزارزها، این بازارها همچنان دارای نوسان بسیار زیادی نسبت به بازارهای هم اندازه هستند که محل بررسی بوده است (؟). برخی این میزان از نوسان را به دلیل عدم وجود روش دقیقی برای ارزش‌گذاری این ارزها که مورد پذیرش همگان باشد می‌دانند. این مسئله موجب شده است تا مطالعات مختلفی به بررسی این ویژگی این بازارها بپردازند.

^۷bid

^۸ask

۳-۱ راه حل پیشنهادی

در این پروژه، سعی داریم تا با استفاده از اطلاعات موجود در دفتر سفارشات یک مدل نوسانی بر پایه یادگیری عمیق ارائه بدهیم. مطالعات مختلفی نشان دهنده اهمیت و تاثیر دفتر سفارشات بر متغیرهای مختلف بازارها بوده‌اند. در مقایسه با دیگر منابع داده‌ی موجود، دفتر سفارشات حجیم‌تر است و شامل اطلاعات جزئی‌تری از وضعیت فعلی بازار است. این اطلاعات که شامل سفارشات انجام نشده هستند تا حدودی خبر از آینده‌ی نزدیک بازار می‌دهند، و از این جهت در مقدار و جهت تغییرات بازار که همان نوسان است تاثیر قابل توجهی دارند(؟) و(؟).

از طرفی شبکه‌های عصبی در سال‌های گذشته در بسیاری از زمینه‌ها به صورت گسترده به کار گرفته شده‌اند و نتایج بسیار امیدوار کننده‌ای ارائه داده‌اند. با این وجود، شبکه‌های عصبی معمولی در استفاده از دنباله‌ی طولانی‌ای از داده‌های ورودی و یادگیری در چنین شرایطی مشکلات مختلفی دارند و کارا نیستند(؟). در این میان، شبکه‌های عصبی بازگشتی^۹، به کمک طراحی خود و استفاده از دنباله‌ای از اطلاعات ورودی برای پیش‌بینی اطلاعات آینده، عملکرد بهتری در این زمینه ارائه کرده‌اند. هرچند که این مدل‌ها نیز در یافتن و استفاده از اطلاعات بلندمدت در دنباله‌ی طولانی از اطلاعات ناتوانند. شبکه‌های بازگشتی معمولی در مواجهه با دنباله‌ی طولانی از اطلاعات، به ویژه در مسائل سری‌های زمانی، با مشکل ناپدید شدن گرادیان^{۱۰} مواجه می‌شوند(؟) و(؟). به طور دقیق‌تر گرادیان بعد از طی کردن مسیر طولانی مقدار بسیار کمی اتخاذ می‌کند که باعث مشکل در روند یادگیری می‌شود. در چنین شرایطی مدل تنها حافظه‌ی کوتاه مدت دارد و توانایی استفاده از روابط طولانی‌مدت تر را نخواهد داشت.

برای رفع این مشکل شبکه‌های حافظه طولانی کوتاه‌مدت^{۱۱} ارائه شده‌اند(؟). این شبکه‌ها با استفاده از متغیرها و ارتباطات مخصوصی که برای حافظه در نظر گرفته‌اند، در طول انتشار معکوس^{۱۲} مقادیر گرادیان را تا حدودی حفظ می‌کنند و از ناپدید شدن آن جلوگیری می‌کنند. در این پروژه تصمیم داریم با استفاده از گونه‌ی دیگری از این ساختارهای اصلاح شده به نام جی‌آریو^{۱۳} و دنباله‌ای از اطلاعات استخراج شده از دفتر سفارشات، اقدام به پیش‌بینی سری زمانی نوسان در این بازارها کنیم.

⁹Recurrent Neural Network

¹⁰Vanishing Gradient Problem

¹¹Long short-term memory

¹²Back Propagation

¹³Gated recurrent unit

۴-۱ بخش‌بندی پایان‌نامه

در بخش بعدی، کارهای مشابه و مربوط به پروژه را بررسی می‌کنیم و مورد مطالعه قرار می‌دهیم. در میان کارهای مطالعه شده بعضاً عمیق‌تر شده و به توضیح مدل‌های احتمالاتی برای پیش‌بینی نوسان می‌پردازیم. در بخش روش پیشنهادی، نحوه‌ی استخراج ویژگی‌های مورد استفاده از دفتر سفارشات را توضیح خواهیم داد، و سپس به ارائه‌ی ساختار مدل پیشنهادی و نحوه‌ی آموزش خواهیم پرداخت. سپس در بخش پیاده‌سازی جزییات نحوه‌ی پیاده‌سازی، کتابخانه‌های استفاده شده و نحوه‌ی کار سیستم را توضیح خواهیم داد. بعد از آن در بخش ارزیابی به معرفی منابع داده‌ی استفاده شده و مقایسه‌ی روش‌های مختلف خواهیم پرداخت. در بخش پایانی نیز به بیان پیشنهادات برای کارهای آینده و جمع‌بندی نهایی خواهیم پرداخت.

فصل دوم

پیش‌زمینه و کارهای مرتبط

۱-۲ مقدمه

در این بخش، ابتدا به تبیین مفهوم نوسان و بیان مدل ریاضی آن از دیدگاه‌های متفاوت می‌پردازیم. سپس سعی می‌کنیم تا برخی از کارهای مرتبط در این حوزه را معرفی کنیم. از راهکارهای متفاوتی برای پیش‌بینی سری‌های زمانی مختلف و همچنین نوسان همان سری‌های زمانی استفاده شده است که در ادامه به مطالعه‌ی برخی از آنها خواهیم پرداخت.

هرچند به صورت سنتی اکثر مدل‌های مربوط به محاسبه‌ی نوسان آماری هستند، فعالیت‌های تحقیقاتی متعددی در این زمینه در سالهای گذشته انجام شده است که دامنه وسیعی از ابزارهای حوزه هوش مصنوعی را به کار گرفته‌اند. دامنه مجموعه داده‌ی مورد استفاده در این کارها نیز گوناگون است، درحالی که برخی تنها از اطلاعات قیمت و حجم مورد مبادله استفاده می‌کنند، برخی دیگر پا را فراتر گذاشته و از اطلاعات دفتر سفارشات نیز استفاده می‌کنند. برخی دیگر با استفاده از داده‌های با فرکانس بسیار بالا و در حد چند میلی‌ثانیه سعی در تخمین وضعیت بازار در بازه‌ی زمانی کوتاهی در آینده را با دقتی بالاتر از حد معمول دارند که در معاملات با فرکانس بالا^۱ کاربرد دارد. اما تلاش اصلی همواره در جهت به دست آوردن تخمینی از وضعیت بازار در چند دقیقه‌ی آینده در یک بازار مالی بوده است. مدل‌های یادگیری عمیق و به ویژه مدل‌های شبکه‌ی عصبی بازگشتی نیز در سال‌های گذشته در زمینه‌ی پیش‌بینی سری‌های زمانی بیشتر به کار گرفته شدند و حتی برخی مطالعات از پیش‌تازی این روش‌ها در برخی کاربردها خبر می‌دهند.

۲-۲ نوسان

در این بخش ابتدا نوسان را به صورت دقیق تعریف می‌کنیم و سپس به مطالعه‌ی تعدادی از مدل‌های نوسانی معروف و نحوه‌ی عملکرد آنها می‌پردازیم.

۱-۲-۲ تعریف نوسان

ابتدا به معرفی نمادهای مورد استفاده برای تعریف نوسان می‌پردازیم. P_t را قیمت یک دارایی در لحظه‌ی t در نظر می‌گیریم. بازده^۲ در لحظه‌ی t به معنای سود یا ضرریست که دارایی مورد نظر در آن لحظه

¹High-frequency trading

²Return

تجربه می‌کند. بر پایه‌ی P_t بازده در لحظه‌ی t به صورت زیر تعریف می‌شود:

$$r_t = \ln(P_t) - \ln(P_{t-1}) \quad (۱-۲)$$

بر این اساس، میانگین بازده در یک پنجره‌ی N تایی به صورت زیر تعریف می‌شود:

$$\bar{r}_t = \frac{\sum_{i=0}^{N-1} r_{t-i}}{N} \quad (۲-۲)$$

به همین شکل واریانس تغییرات قیمت برای پنجره‌ای به طول N ، که به عنوان نوسان شناخته می‌شود، به صورت زیر محاسبه می‌شود:

$$v_t = \sqrt{\frac{\sum_{i=0}^{N-1} (r_{t-i} - \bar{r})^2}{N}} \quad (۳-۲)$$

طبق رابطه‌ی بالا می‌توان سری زمانی نوسان، v_t را در طول زمان تشکیل داد.

۳-۲ مدل‌های نوسانی

در ادامه به معرفی ویژگی‌های مدل‌های نوسانی مختلف و بررسی نحوه‌ی کارکرد برخی از آن‌ها می‌پردازیم. مطالعات پیشین به استفاده و بررسی گسترده از مدل‌های بر پایه‌ی مدل گارچ^۳ برای پیش‌بینی نوسان پرداخته‌اند (؟) و (؟). برخی دیگر از روش‌های مبتنی بر مدل‌های احتمالاتی گرافی برای استفاده از اطلاعات موجود در دفتر سفارشات استفاده کرده‌اند (؟).

³GARCH

۱-۳-۲ مدل‌های گارچ

برای درک نحوه‌ی کارکرد مدل‌های گارچ ابتدا باید از گونه‌ی ساده‌تری از مدل‌ها به نام آریما^۴ که خود شامل سه بخش کاهنده‌ی خودکار^۵، میانگین متحرک^۶ و بخش تجمعی^۷ هستند شروع کنیم.

مدل‌های کاهنده‌ی خودکار

$AR(p)$ یک مدل کاهنده‌ی خودکار از مرتبه‌ی p است که از پس‌افت‌های^۸ سری زمانی اصلی به عنوان متغیرهای پیش‌بینی همان سری زمانی در آینده استفاده می‌کند. مقدار p بیان‌کننده‌ی مرتبه‌ی مدل است. به عنوان مثال، $AR(1)$ از پس‌افت اول سری زمانی، یعنی x_{t-1} ، برای پیش‌بینی x_t استفاده می‌کند. بنابراین می‌توان یک مدل کاهنده‌ی خودکار از مرتبه‌ی p را به صورت زیر تعریف کرد(؟):

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \epsilon_t \quad (۴-۲)$$

که ϵ_t در آن نویز سفید^۹ است. در واقع ϵ هر بار بعد از مشاهده‌ی مقدار اصلی سری زمانی به صورت اختلاف پیش‌بینی مدل و مقدار اصلی محاسبه می‌شود. ϵ در حالت ایده‌آل نویز سفید است، چرا که در این صورت می‌توانیم مطمئن باشیم که مدل تمامی الگوهای موجود در سری زمانی را مدل کرده است و مقدار باقیمانده^{۱۰} از سری زمانی کاملاً تصادفی و غیرقابل مدل کردن است.

مدل‌های میانگین متحرک

این مدل‌ها از جهاتی به مدل‌های کاهنده‌ی خودکار شبیه هستند، اما به جای استفاده از پس‌افت‌های قبلی سری زمانی، از خطاها (ϵ) برای پیش‌بینی استفاده می‌کنند. مدل $MA(q)$ یک مدل میانگین متحرک از مرتبه‌ی q است که از تعداد q خطای قبلی برای پیش‌بینی سری زمانی استفاده می‌کند، که به صورت

^۴Arima^۵Auto Regressive^۶Moving Average^۷Integrated^۸Lag^۹White Noise^{۱۰}Residual Value

زیر تعریف می‌شود:

$$x_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} + \mu \quad (5-2)$$

که μ بیانگر میانگین ثابت است.

مدل‌های آریما

این مدل‌های که به نوعی ترکیب دو مدل قبلی هستند، از تمامی ویژگی‌های موجود در آن‌ها بهره می‌برند. این مدل‌ها ابتدا متغیر جدید y_t را به صورت زیر تعریف می‌کنند:

$$y_t = x_t - x_{t-1} \quad (6-2)$$

علت تعریف y_t به این صورت از بین بردن رشد خطی یک سری زمانی یا روند سیر^{۱۱} آن است. بدیهی است که در صورت توانایی پیش‌بینی سری زمانی y می‌توان به راحتی سری زمانی x را با داشتن یک مقدار اولیه به دست آورد. مدل آریما سری زمانی y_t را به صورت زیر مدل می‌کند:

$$y_t = \phi_0 + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (7-2)$$

که مانند مدل‌های قبلی مقادیر p و q مرتبه‌ی مدل آریما را مشخص می‌کنند. p مرتبه‌ی قسمت کاهنده‌ی خودکار و q مرتبه‌ی قسمت میانگین متحرک است. مدل پارامتر دیگری به نام i نیز دارد که عموماً مقدار ۱ دارد اما بسته به کاربرد ممکن است تغییر کند و مقداری بیشتر از ۱ بگیرد. i در مدل آریما مشخص می‌کند که عمل تفاضل‌گیری از سیگنال اولیه چندبار انجام شود. به عنوان مثال، رابطه‌ی ۶-۲ نشان‌دهنده‌ی یکبار تفاضل‌گیری از سیگنال اولیه است. در صورتی که i مقداری بیشتر از یک داشته باشد، باید عمل تفاضل‌گیری برای y_t نیز تکرار شود تا سری زمانی جدید تشکیل شود و سپس از رابطه‌ی ۷-۲ برای مدل کردن آن استفاده شود.

¹¹Trend

مدل‌های آرچ

مدل‌های بر پایه‌ی آرچ^{۱۲} با تمرکز بر پیش‌بینی نوسان به وجود آمده‌اند. این مدل میزان نوسان هر نقطه‌ی زمانی را با استفاده از خطای پیش‌بینی در نقاط قبل از آن یعنی ϵ ها مدل می‌کند. این مدل نیز مانند مدل‌های پیشین برای پایه‌های مختلف تعریف می‌شود. تعریف یک مدل آرچ به ازای پایه‌ی r برابرست با:

$$\epsilon_t = w_t \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \dots + \alpha_r \epsilon_{t-r}^2} \quad (۸-۲)$$

که در حالت ایده‌آل، w_t در آن باید نویز سفید باشد که بعد از مشاهده‌ی مقدار اصلی نوسان دنباله‌ی هدف قابل محاسبه می‌باشد.

مدل‌های گارچ

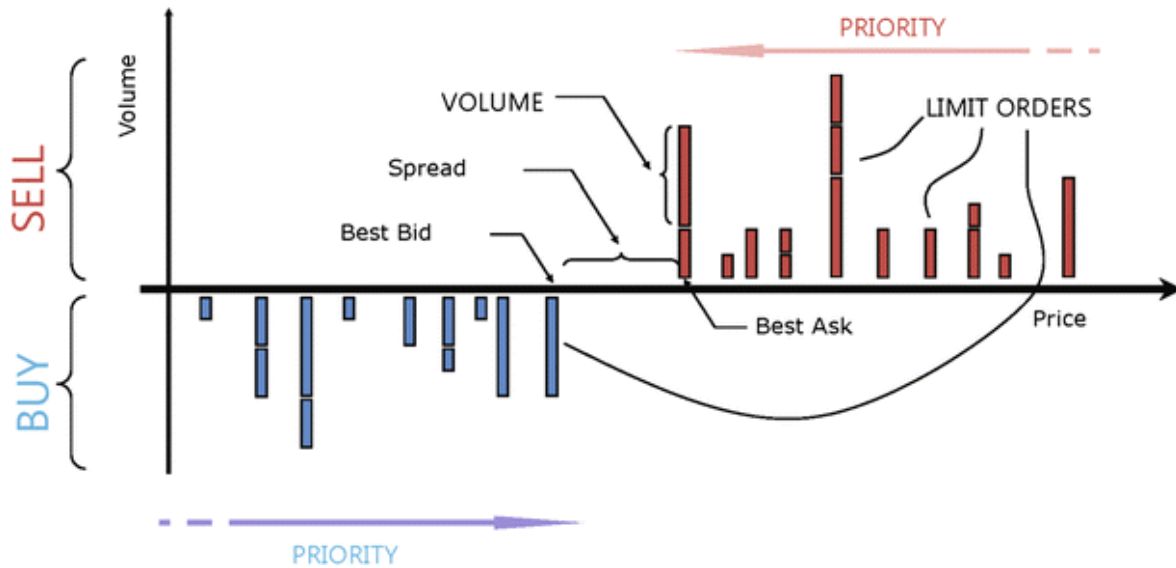
نحوه‌ی بدست آمدن مدل گارچ از روی مدل آرچ بسیار شبیه به روند تبدیل مدل کاهشی خودکار به آریماست. در تبدیل مدل کاهشی خودکار به آریما شاهد بودیم که با اضافه کردن یک مجموعه متغیر جدید که نشان‌دهنده‌ی خطای پیش‌بینی‌های گذشته بودند، مدل کامل‌تری ساخته شد که می‌توانست از خطاهای قبلی خود در جهت بهبود پیش‌بینی‌های بعدی استفاده کند. در تبدیل آرچ به گارچ نیز با اضافه کردن پس‌افت‌های خود سری زمانی نوسان به مدل قبلی، می‌توانیم دنباله‌های بیشتری را مدل کنیم. رابطه‌ی گارچ برای پایه‌های p و q به صورت زیر است:

$$\epsilon_t = w_t \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_p \epsilon_{t-p}^2 + \beta_1 \delta_{t-1}^2 + \dots + \beta_q \delta_{t-q}^2} \quad (۹-۲)$$

در تعریف بالا δ_t به معنای نوسان در لحظه‌ی t است که به صورت زیر قابل محاسبه است:

$$\delta_t = \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \dots + \alpha_q \epsilon_{t-q}^2} \quad (۱۰-۲)$$

¹²Autoregressive conditional heteroskedasticity



شکل ۲-۱: نمودار ساختار دفتر سفارشات به همراه برخی ویژگی‌های مربوط (؟).

با استفاده از ۲-۹ و ۲-۱۰ می‌توان رابطه‌ی گارچ را به صورت زیر بازنویسی کرد:

$$\epsilon_t = w_t \sqrt{\delta_t + \beta_1 \delta_{t-1}^2 + \dots + \beta_q \delta_{t-q}^2} \quad (۱۱-۲)$$

که مانند مدل‌های قبلی w_t نویز سفید است.

۲-۳-۲ دفتر سفارشات

در این بخش به مطالعه‌ی بخشی از مطالعات انجام شده بر روی اطلاعات دفتر سفارشات خواهیم پرداخت. ابتدا به معرفی ساختار دفتر سفارشات خواهیم پرداخت و سپس مشاهده خواهیم کرد که این داده ساختار چگونه اطلاعات مربوط به وضعیت بازار و سفارشات را در خود جای می‌دهد. در ادامه به بررسی چند نمونه از مطالعات انجام شده بر روی دفتر سفارشات خواهیم پرداخت و با اهمیت این داده ساختار در مطالعه‌ی بازارهای مالی آشنا خواهیم شد.

ساختار دفتر سفارشات

به طور خلاصه دفتر سفارشات شامل تمامی سفارش‌های خرید و فروش در جریان در یک بازار مالی و مربوط به یک دارایی مشخص است. این دفتر شامل سفارشات به صورت یک دوتایی مرتب شامل قیمت و حجم متناظر است. تمام سفارش‌هایی که دارای قیمت یکسانی باشند با یکدیگر تجمیع می‌شوند و به

صورت یک دوتایی مرتب با همان قیمت و مجموع حجم همه‌ی آن‌ها در دفتر ثبت می‌شوند. این دوتایی‌های مرتب را می‌توان بر اساس قیمت مرتب کرد تا بتوان ترتیبی برای نحوه‌ی جور کردن آن‌ها در جریان فعالیت بازار در نظر گرفت. در صورت مرتب‌سازی این دوتایی‌ها بر اساس قیمت، مانند شکل ۱-۲، شاهد جداسازی سفارش‌های مربوط به خرید و سفارش‌های مربوط به فروش خواهیم بود. علت این امر آن است که سفارش‌های خرید همواره قیمت کمتری نسبت به قیمت سفارش‌های موجود برای فروش دارند، زیرا که اگر چنین نبود سفارش خرید به محض ایجاد با یک سفارش فروش متناظر می‌شد و معامله انجام می‌شد و هر دو سفارش از دفتر سفارشات خارج می‌شدند.

سپس در حالی که دوتایی‌ها بر اساس قیمت مرتب‌سازی شده‌اند می‌توان یک ترتیب برای نحوه‌ی جورسازی و اجرای آن‌ها در نظر گرفت. سفارش‌های فروش به ترتیب از قیمت کم به زیاد دارای اولویت زیاد به کم هستند چرا که خریداران همواره تمایل به خرید یک دارایی با کمترین قیمت ممکن را دارند و در نتیجه اگر قیمت پیشنهادیشان بیشتر از چند سفارش فروش مختلف باشد، اقدام به خرید از ارزان‌ترین آن‌ها می‌کنند. برعکس همین مسئله برای سفارش‌های خرید برقرار است. بدین ترتیب سفارشی که بیشترین میزان قیمت پیشنهادی را داشته باشد از بالاترین اولویت برخوردار است و در صورتی که سفارش فروش جدیدی با قیمتی کمتر از چند سفارش خرید موجود در بازار ایجاد شود، به طور خودکار با گران‌ترین سفارش خرید جور شده و اجرا می‌شود.

از طرف دیگر ویژگی‌های دیگری نیز در دفتر سفارشات وجود دارد که مورد علاقه‌ی تحلیلگران و فعالین این حوزه بوده است. از مهم‌ترین آن‌ها می‌توان به شکاف قیمت^{۱۳} که در شکل ۱-۲ نیز مشاهده می‌شود اشاره کرد. شکاف قیمت بیانگر اختلاف موجود بین ارزان‌ترین پیشنهاد فروش و گران‌ترین پیشنهاد خرید است. به طور معمول افزایش میزان شکاف قیمتی خبر از اختلاف میان خریداران و فروشندگان می‌دهد که می‌تواند باعث افزایش یا کاهش شدید قیمت در آینده‌ی نزدیک بشود که در هر دو حالت با افزایش نوسان همراه خواهد بود.

در مطالعه‌ی نشان داده شده است که با استفاده از ویژگی‌های مختلف دفتر سفارشات می‌توان نوسان را با دقت بیشتری نسبت به روش‌هایی که از این اطلاعات استفاده نمی‌کنند تخمین زد(؟). این مقاله با استفاده از یک مدل احتمالاتی گرافی و اطلاعات دفتر سفارشات، موفق شده است که نوسان را با دقت بیشتری نسبت به مدل‌های گارچ پیش‌بینی کند. همچنین نکته‌ی دیگری که در این مقاله حائز اهمیت است، مقایسه‌ی نتایج نهایی به دست آمده توسط این مدل با مدل‌های ساده‌ی شبکه‌ی عصبی است که نشان دهنده‌ی برتری شبکه‌های عصبی در برخی بازه‌های زمانی نیز هست.

¹³Spread

۳-۳-۲ جمع‌بندی

در این فصل با برخی روش‌های آماری برای پیش‌بینی سری‌های زمانی و نوسان آن‌ها آشنا شدیم. سپس به بررسی و معرفی دقیق‌تر دفتر سفارشات پرداختیم.

فصل سوم

روش پیشنهادی



شکل ۳-۱: نحوه‌ی جداسازی و ذخیره‌ی دفتر سفارشات

۱-۳ مقدمه

در بخش قبلی با برخی پیشنیازهای مربوط به مسئله اصلی و برخی از مطالعات انجام شده در استفاده از دفتر سفارشات آشنا شدیم. در این بخش قصد داریم تا روش پیشنهادی برای حل مسئله را تشریح کنیم. ابتدا به نحوه‌ی ذخیره‌سازی و استخراج ویژگی‌های مختلف از دفتر سفارشات می‌پردازیم، سپس به سراغ تشکیل سری زمانی نوسان قیمت و ایجاد نمونه‌های برچسب دار می‌رویم. در ادامه با مدل‌های شبکه‌ی عصبی بازگشتی و به ویژه جی‌آریو آشنا می‌شویم و با استفاده از آن‌ها مدل نوسانی خود را تشکیل می‌دهیم.

۲-۳ استخراج ویژگی و سری زمانی از دفتر سفارشات

همانطور که در فصل قبل اشاره کردیم دفتر سفارشات را می‌توان به صورت دنباله‌ای از دوتایی‌های مرتب شامل قیمت و حجم متناظر در نظر گرفت. با جدا سازی دو بخش خرید و فروش و همچنین جداسازی دو بخش قیمت و حجم برای هر دفتر ثبت سفارش در طول زمان به یک آرایه‌ی سه بعدی که شامل تمامی اطلاعات دفتر سفارشات اولیه است دست پیدا می‌کنیم (شکل ۳-۱). در این شکل آرایه‌ی سه بعدی مد نظر با ابعاد $2 * 2 * Depth$ نمایش داده شده است که $Depth$ یا عمق در آن به معنای تعداد سفارش‌های موجود در هر سمت دفتر سفارشات است. در این پروژه عمق دفتر سفارشات در هر دو سمت خرید و فروش را برابر در نظر می‌گیریم تا نمایش بهتری از داده‌ها داشته باشیم. هر دندانه‌ی^۱ قیمتی به یک قیمت مشخص برای دارایی مورد مطالعه اشاره می‌کند. فواصل این دندانه‌ها توسط صرافی مربوطه

¹Tick

Vb_D	...	Vb_3	Vb_2	Vb_1	Va_1	Va_2	Va_3	...	Va_D
Pb_D	...	Pb_3	Pb_2	Pb_1	Pa_1	Pa_2	Pa_3	...	Pa_D

شکل ۳-۲: اندیس گذاری دفتر سفارشات برای ایجاد دفتر سفارشات تجمعی

تعیین می شود و گاهی نسبت ثابتی با آخرین قیمت معامله شده دارد، اما عمدتاً این فواصل به صورت ثابت و پیش فرض برابر مقدار کمی مانند ۱۰ سنت در نظر گرفته می شوند. علت این امر ایجاد انعطاف بیشتر برای خریداران و فروشندگان در سفارشات است.

۱-۲-۳ دفتر سفارشات تجمعی

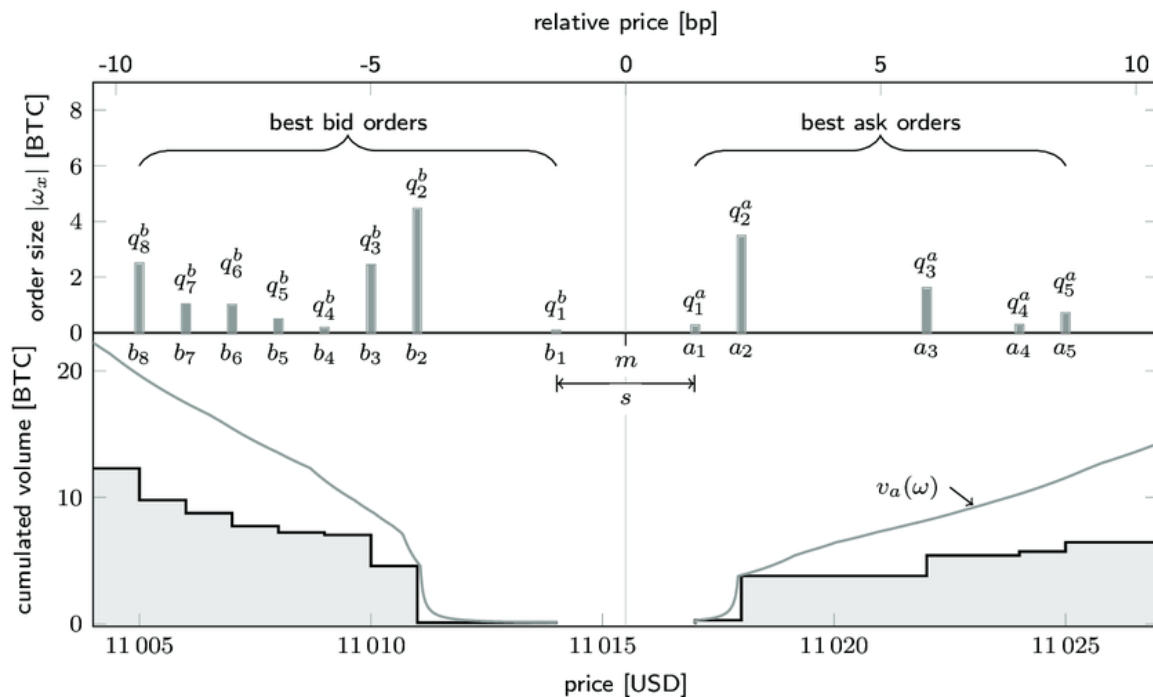
در مرحله ی بعدی دفتر سفارشات تجمعی را تعریف می کنیم. هدف اصلی از تعریف دفتر سفارشات تجمعی دستیابی به نمایش بهتری از عرضه و تقاضا نسبت به دفتر سفارشات عادی است (؟). حجم های تجمعی را بر اساس حجم های اولیه (شکل ۳-۲) تعریف می کنیم:

$$\begin{aligned}\bar{V}b_i &= \sum_{j=1}^i Vb_j \\ \bar{V}a_i &= \sum_{j=1}^i Va_j\end{aligned}\quad (1-3)$$

که Vb_i و $\bar{V}b_i$ در آن به ترتیب بیانگر حجم خرید در دندانه ی i ام خرید و حجم تجمعی خرید در دندانه ی i ام هستند. به همین شکل، Va_i و $\bar{V}a_i$ در آن به ترتیب بیانگر حجم فروش در دندانه ی i ام فروش و حجم تجمعی فروش در دندانه ی i ام هستند.

با تشکیل این دنباله های جدید، می توان تفسیر جدیدی برای دفتر سفارشات در نظر گرفت. در این نمایش جدید \bar{V} ها نشان دهنده ی حداکثر میزان حجم موجود برای خرید (یا فروش) در صورت تمایل به پرداخت (یا دریافت) قیمت متناظر با هر دندانه هستند. بدین شکل هر دندانه ی سفارشی معنای کامل تری نسبت به حالت قبلی پیدا می کند چرا که محل قرارگیری سفارش های خریداران و فروشندگان در دفتر سفارشات در این نمایش جدید گنجانده شده است.

به طور مثال، قرار دادن یک سفارش خرید با حجم بسیار زیاد در نزدیکی بالاترین قیمت پیشنهادی



شکل ۳-۳: نمونه‌ای از دفتر سفارشات معمولی و دفتر سفارشات تجمعی متناظر (؟)

خرید تمامی پیشنهاد‌های با قیمت کمتر را تحت تاثیر قرار می‌دهد و از احتمال اجرا شدن آن‌ها می‌کاهد. اینگونه سفارشات با حجم بالا که به عنوان دیوار قیمتی^۲ شناخته می‌شوند از اهمیت بسیار زیادی برخوردار هستند چرا که وجود یا عدم وجودشان باعث تفسیر متفاوتی از سفارش‌های بعد از آنها می‌شوند.

۳-۲-۲ استخراج ویژگی‌های دیگر از دفتر سفارشات

در این مرحله به استخراج برخی دیگر از ویژگی‌های دفتر سفارشات که در برخی از مطالعات انجام‌شده بر روی دفتر سفارشات از آن‌ها استفاده شده است می‌پردازیم.

قیمت میانی

قیمت میانی برای هر دفتر سفارشات به صورت میانگین بیشترین قیمت خرید و کمترین قیمت فروش در نظر گرفته می‌شود. این مقدار را به صورت زیر می‌توان تعریف کرد:

$$MidPrice = \frac{(Pb_1 + Pa_1)}{2} \quad (۲-۳)$$

^۲Price Wall

قیمت میانی می‌تواند تخمین خوبی از قیمتی که معامله‌ی بعدی در آن انجام می‌شود به دست دهد و از این جهت حائز اهمیت است.

شکاف قیمت

شکاف قیمت^۳ اختلاف بین بالاترین قیمت خرید و پایین‌ترین قیمت فروش است. که به صورت زیر محاسبه می‌شود:

$$PriceSpread = Pa_1 - Pb_1 \quad (3-3)$$

تغییر قیمت یا بازده

مقدار تغییر قیمت که به آن بازده نیز گفته می‌شود، معمولاً به صورت لگاریتمی در مطالعات استفاده می‌شود. این مقدار با استفاده از دو قیمت میانی متوالی قابل محاسبه است که به صورت مقابل نشان داده می‌شود:

$$PriceChange_{t+1} = \ln\left(\frac{MidPrice_{t+1}}{MidPrice_t}\right) = \ln(MidPrice_{t+1}) - \ln(MidPrice_t) \quad (4-3)$$

شکاف قیمتی وزن دار

این مقدار بیانگر اختلاف موجود میان مجموع تقاضای خرید و فروش است. برای محاسبه‌ی این مقدار به طور معمول از ۱۰٪ اول دفتر سفارشات استفاده می‌کنیم تا یک میانگین وزن دار از مجموع تقاضاهای خرید و فروش موجود در این بازه بسازیم. این مقدار را به صورت مقابل تعریف می‌کنیم:

$$WheightedSpread = \frac{1}{n} \left(\sum_{i=1}^n Pa_i * Va_i - \sum_{i=1}^n Pb_i * Vb_i \right) \quad (5-3)$$

$$n = Depth/10$$

³Price Spread

نوسان

نوسان قیمت را نیز طبق فرمول ۳-۲ برای پنجره‌های مختلف استخراج می‌کنیم و به مجموعه داده ورودی برای هر نقطه اضافه می‌کنیم.

۳-۲-۳ استانداردسازی ویژگی‌ها

در ادامه برای یادگیری بهتر لازم است تا ویژگی‌های استخراج‌شده را از نظر آماری استاندارد کنیم. یک ویژگی استاندارد دارای میانگین صفر و واریانس یک است، و برای استانداردسازی هر ویژگی باید ابتدا میانگین مقدار آن و انحراف معیار آن را در طول زمان را به دست بیاوریم و سپس طبق رابطه‌ی زیر ویژگی را استاندارد کنیم:

$$X_{standard} = \frac{X - \bar{X}}{\delta} \quad (۳-۶)$$

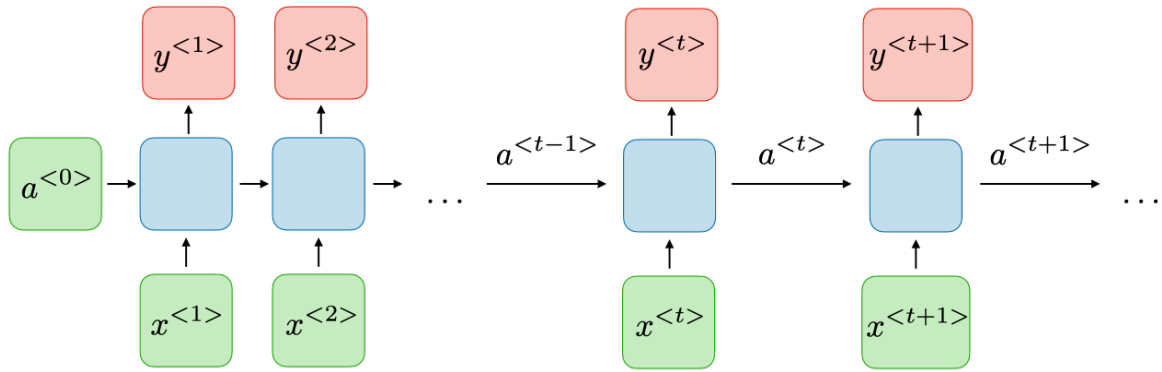
که X بیانگر مقدار ویژگی در هر نقطه، \bar{X} میانگین ویژگی X در طول زمان و δ برابر انحراف معیار ویژگی X است.

۳-۳ یادگیری

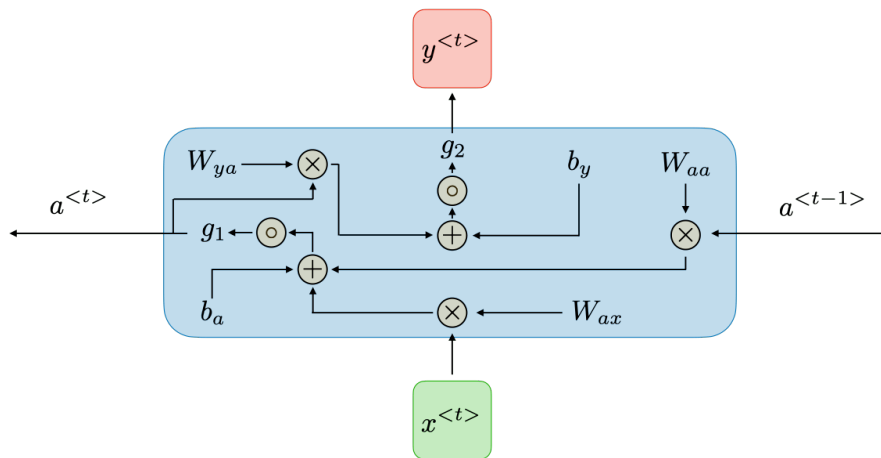
در این پروژه قصد داریم تا با استفاده از مدل‌های شبکه‌ی عصبی بازگشتی اقدام به یادگیری توزیع نوسان بر حسب ویژگی‌های استخراج‌شده کنیم. در این راستا، ابتدا شبکه‌های بازگشتی عمیق ساده را معرفی می‌کنیم و مشکلات آن‌ها را بررسی می‌کنیم. سپس به ساختار مدل جی‌آریو می‌پردازیم و راه حل این مدل برای مشکلات گفته شده را بررسی می‌کنیم. در انتها نحوه‌ی یادگیری بر اساس نمونه‌ها توسط این مدل را تشریح می‌کنیم.

۱-۳-۳ شبکه‌های عصبی بازگشتی

شبکه‌های عصبی بازگشتی برای یادگیری داده‌هایی که طبیعت دنباله‌ای دارند طراحی شده‌اند (؟). این شبکه‌ها به دلیل ساختار دنباله‌ای در زمینه‌های مختلفی مانند پردازش زبان طبیعی و همچنین پیش‌بینی سری‌های زمانی به صورت گسترده استفاده شده‌اند. این شبکه‌ها یک دنباله‌ی ورودی از داده‌ها را می‌گیرند و دنباله‌ای دیگر به عنوان خروجی تولید می‌کنند.



شکل ۳-۴: شمای کلی شبکه‌های عصبی بازگشتی



شکل ۳-۵: نمایی از روابط ۷-۳ و ۸-۳ که نحوه‌ی محاسبات انجام شده درون هر سلول شبکه‌ی عصبی بازگشتی را نشان می‌دهند.

برای هر نقطه‌ی زمانی t ، تابع فعال‌سازی $\alpha^{<t>}$ و تابع خروجی $y^{<t>}$ به صورت زیر محاسبه می‌شوند:

$$\alpha^{<t>} = g_1(W_{aa}\alpha^{<t-1>} + W_{ax}x^{<t>} + b_a) \quad (۷-۳)$$

$$y^{<t>} = g_2(W_{ya}\alpha^{<t>} + b_y) \quad (۸-۳)$$

که ماتریس‌های W_{ax} و W_{aa} ، W_{ya} ، b_y ، b_a در آن ضرایبی هستند که توسط هر سلول به اشتراک گذاشته می‌شود. از طرفی g_1 و g_2 نیز توابع فعال‌سازی^۴ هستند که در ادامه با آن‌ها خواهیم پرداخت.

^۴Activation Function

استفاده از شبکه‌های عصبی بازگشتی مزایا و معایب متفاوتی دارد که در ادامه به آن‌ها اشاره می‌کنیم.

معایب شبکه‌های بازگشتی

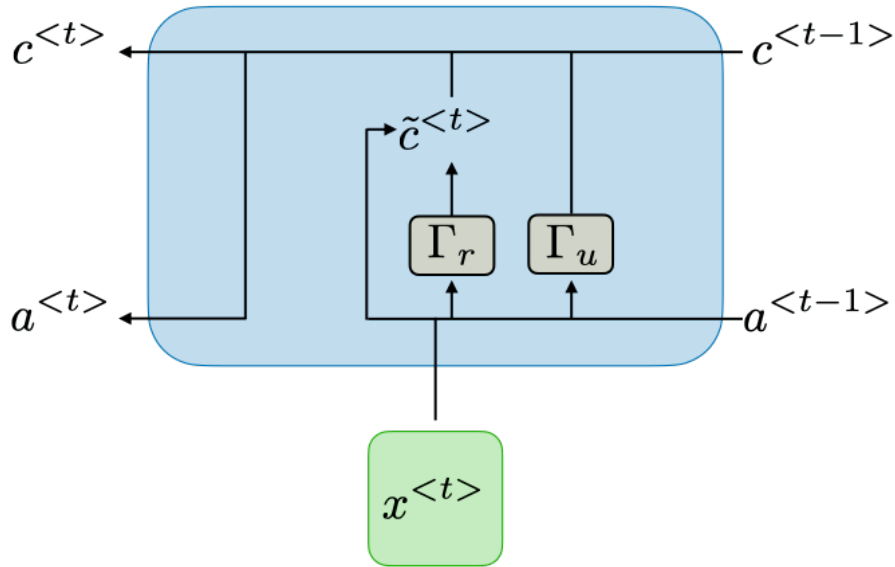
همانطور که پیش‌تر گفته شد، شبکه‌های عصبی بازگشتی از مشکل از بین رفتن گرادیان رنج می‌برند (؟). این مسئله از تضعیف مقدار گرادیان در انتقال بین سلول‌ها ناشی می‌شود. هنگامی که فاصله‌ی نقطه‌ای که خطا در آن محاسبه شده با یک سلول زیاد می‌شود، گرادیان خطا نسبت به متغیرهای آن سلول مقدار بسیار کمی می‌گیرند که مانع از انجام یادگیری توسط مدل می‌شود. این مسئله باعث می‌شود تا ارتباطات بلندمدت در دنباله‌ی ورودی توسط مدل یاد گرفته نشود. در ادامه خواهیم دید که شبکه‌ی جی‌آریو چگونه با استفاده از حافظه‌ی داخلی این مشکل را برطرف می‌کند.

از طرف دیگر، به دلیل انتقال یک طرفه‌ی اطلاعات در این شبکه‌ها، امکان انتقال خطا از سلول‌های جلویی به عقبی وجود ندارد و به همین دلیل ورودی‌های بعدی در وضعیت فعلی متغیرهای سلول تأثیری نخواهند داشت. هرچند این مسئله در پیش‌بینی سری‌های زمانی خیلی مطرح نیست چرا که اغلب بر اساس دنباله‌ای از داده‌های ثبت شده در یک بازه‌ی زمانی به پیش‌بینی نقطه‌های بعدی می‌پردازیم. اما این مشکل در پردازش زبان طبیعی بسیار حائز اهمیت است که موجب پیدایش شبکه‌ها بر پایه توجه^۵ شده‌است.

مزایای شبکه‌های بازگشتی

این شبکه‌ها به خاطر به اشتراک‌گذاری وزن‌ها در طول زمان، نسبت به افزایش طول ورودی حساسیتی ندارند و با افزایش طول ورودی تعداد پارامترهای آن‌ها افزایش نمی‌یابد. این ویژگی این شبکه‌ها را برای استفاده در مسائل پیش‌بینی سری‌های زمانی که گاهی دارای ورودی بسیار بزرگ هستند مناسب می‌کند. از طرفی، با به اشتراک‌گذاری وزن‌ها بین سلول‌ها، به نوعی دانش موجود برای نقاط مختلف زمانی به اشتراک گذاشته می‌شود، که باعث به وجود آمدن مدل‌های کاراتری می‌شود. اطلاعات و خروجی توابع فعال‌سازی سلول‌های قبلی نیز در هر سلول به کار گرفته می‌شوند که ویژگی بسیار مناسبی در پیش‌بینی سری‌های زمانیست.

^۵Attention



شکل ۳-۶: ساختار داخل هر سلول از شبکه‌ی عصبی جی‌آریو.

۳-۳-۲ جی‌آریو

جی‌آریو گونه‌ای از شبکه‌های عصبی بازگشتیست که برای حل مسئله‌ی گردایان از بین رونده و مدل‌سازی روابط طولانی مدت در داده‌های ورودی به وجود آمده است. این شبکه‌ها با استفاده از ساختاری مشابه به حافظه اقدام به حفظ اختیاری اطلاعات سلول‌های گذشته و انتقال آن‌ها به سلول‌های بعدی می‌کنند. برای متوجه شدن بهتر نحوه‌ی کارکرد این مدل از شبکه‌های عصبی روابط ریاضی آن‌ها را تشریح می‌کنیم:

$$\tilde{c}^{<t>} = \tanh(W_c[\Gamma_r * \alpha^{t-1}, x^{<t>}] + b_c) \quad (9-3)$$

$$\tilde{c}^{<t>} = \tanh(W_c[\Gamma_r * \alpha^{t-1}, x^{<t>}] + b_c) \quad (10-3)$$

$$a^{<t>} = c^{<t>} \quad (11-3)$$

که Γ در آن مشخص کننده دروازه‌هاست که هرکدام هدف مشخصی دارند. رابطه ریاضی برای هر دروازه به صورت زیر است:

$$\Gamma = \sigma(Wx^{<t>} + Ua^{<t-1>} + b) \quad (12-3)$$

که U ، b و W ضرایب متناظر با هر دروازه هستند و σ تابع سیگموئید است. که به صورت زیر تعریف می‌شود:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3-13)$$

در هر سلول جی‌آریو Γ_r مشخص کننده میزان ارتباط اطلاعات ورودی جدید در زمان t است و به آن دروازه‌ی ربط گفته می‌شود. این دروازه مشخص می‌کند که تا چه میزان اطلاعات ورودی جدید وارد حافظه‌ی بلندمدت یا همان متغیر \bar{c} بشود. از طرف دیگر، Γ_u مشخص می‌کند که اطلاعاتی که از قبل درون حافظه وجود داشته اند تا چه میزان باید فراموش شوند و چه میزان از آن‌ها باید حفظ شود و ازین جهت به آن دروازه‌ی به‌روزرسانی گفته می‌شود.

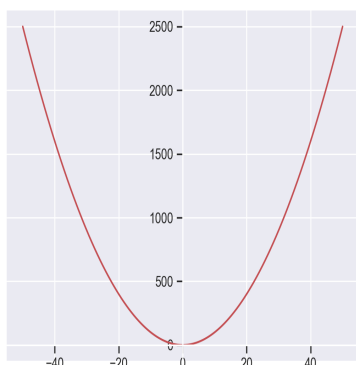
۳-۳-۳ تابع خطا

معیارهای مختلفی برای گزارش خطا در پیش‌بینی سری‌های زمانی وجود دارد. تعریف تابع خطا در مطالعات بر روی سری زمانی اهمیت بسیار زیادی دارد، چرا که توابع خطای مختلف نتیجه‌های بسیار متفاوتی در مدل نهایی به وجود می‌آورند. به طور مثال، توابع خطایی که با داده‌های پرت^۶ مانند باقی داده‌ها رفتار می‌کنند نتیجه‌ی قابل قبولی در مدل نهایی ارائه نمی‌دهند، چرا که این داده‌ها در سری‌های زمانی عموماً غیر قابل پیش‌بینی هستند و به همین خاطر می‌توانند خطای قابل توجهی را به مدل‌ها تحمیل کنند. از طرف دیگر، هدف از به کار گیری مدل نوسان داشتن تخمین خوبی از نوسان در طول زمان به صورت عمومیست و نه فقط بازه‌های خاصی که پیش‌بینی نوسان در آن‌ها بسیار سخت است، به همین خاطر سعی در طراحی و استفاده از تابع خطا، جهت‌دهی یادگیری به سمتی است که در تعداد زیادی از نقاط تخمین خوبی از دنباله‌ی هدف توسط مدل حاصل شود. از همین جهت انواع مختلفی از توابع خطا در این حوزه به کار گرفته شده‌اند که در ادامه به معرفی تعدادی از آن‌ها می‌پردازیم.

⁶Outlier

میانگین مربعات خطا

این معیار خطا که به نوعی از نرم دوم خطا استفاده می‌کند، با جمع زدن و میانگین‌گیری از مربعات خطا، معیاری از دقت عملکرد سیستم گزارش می‌دهد که رابطه‌ی آن به صورت زیر است:



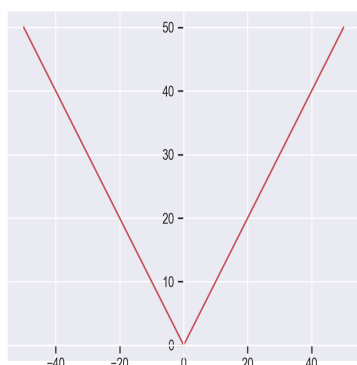
$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (۱۴-۳)$$

شکل ۳-۷: نمودار تابع میانگین مربعات خطا بر اساس میزان خطا یا ϵ

این تابع خطا نسبت به داده‌های پرت بسیار حساس بوده و ازین جهت جای بهبود دارد.

میانگین قدرمطلق خطا

این معیار خطا که به نوعی از نرم اول خطا استفاده می‌کند، با جمع زدن از قدر مطلق خطا، معیاری از دقت عملکرد سیستم گزارش می‌دهد که رابطه‌ی آن به صورت زیر است:



$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (۱۵-۳)$$

شکل ۳-۸: نمودار میانگین قدرمطلق خطا بر اساس میزان خطا یا ϵ

میانگین قدر مطلق خطا نسبت به تمامی نقاط موجود در مجموعه داده رفتار یکسانی دارد. علت این امر آن است که شیب نمودار خطا به ازای تمامی مقادیر ϵ ثابت است و مانند میانگین مجموع مربعات خطا برای داده‌ها پرت افزایش نمیابد. ازین جهت می‌توان گفت که این تابع خطا نسبت به داده‌ها پرت حساسیت کمتری دارد.

میانگین نسبی قدرمطلق خطا

میانگین نسبی قدرمطلق خطا^۷ مانند میانگین قدرمطلق خطا از نرم اول خطا برای محاسبه‌ی هزینه‌ی نهایی استفاده می‌کند، اما در مقابل این خطا را به مقدار ورودی اصلی تقسیم می‌کند که باعث می‌شود نوعی عادی سازی^۸ صورت بگیرد. رابطه‌ی این تابع خطا به صورت زیر است:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{\hat{Y}_i} \right| \quad (۱۶-۳)$$

میانگین نسبی قدرمطلق خطا تابع خطای مناسبی برای سری‌های زمانی است. سری‌های زمانی اغلب با افزایش و کاهش‌های شدید در طول زمان مواجه هستند. در مقابل پیش‌بینی‌های انجام‌شده بر روی هر سری زمانی با خطاهایی همراه هست که با مقدار سری زمانی در آن نقطه نسبت مستقیمی دارند. به بیان دیگر با افزایش مقدار یک سری زمانی در طول زمان (به عنوان مثال افزایش نوسان بازار) پیش‌بینی آن با همان میزان خطای مطلق قبلی بسیار دشوارتر می‌شود و در نتیجه میزان خطای مدل نیز افزایش پیدا می‌کند. این مسئله باعث می‌شود تا مقادیر بزرگتر سری زمانی خطای بیشتری نسبی به مقادیر کوچک‌تر به مدل تحمیل کنند.

برای حل این مشکل تابع میانگین نسبی قدرمطلق خطا پیشنهاد شده است که با تقسیم مقدار مطلق خطا به مقدار سری زمانی در آن نقطه معیار بهتری از عملکرد سیستم نسبت به توابع دیگر به دست می‌دهد. از نگاهی دیگر این تابع خطا نسبت به افزایش و کاهش سری‌زمانی حساس نیست که باعث می‌شود خطای یکنواختی در طول سری زمانی ارائه کند.

^۷Mean Absolute Percentage Error

^۸Normalization

۴-۳ جمع‌بندی

در این فصل با نحوه‌ی استخراج ویژگی‌ها مختلف از دفتر سفارشات آشنا شدیم و سپس به معرفی شبکه‌های عصبی بازگشتی و ساختار آن‌ها پرداختیم. در انتها نیز توابع خطای مورد استفاده در مطالعه‌ی سری‌های زمانی را معرفی کردیم. در فصل بعدی روش و جزییات پیاده سازی را تشریح می‌کنیم.

فصل چهارم

پیاده‌سازی و کتابخانه‌ها

۴-۱ مقدمه

در این قسمت قصد داریم بخش پیاده‌سازی و کتابخانه‌های استفاده شده را توضیح دهیم. پروژه با زبان برنامه‌نویسی پایتون پیاده‌سازی شده است. با توجه به معماری مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی، برای پردازش با سرعت مناسب بهتر بود از پردازنده‌های گرافیکی^۱ استفاده کنیم. به همین منظور، تمام کدهای پروژه بر روی سرویس کولب^۲، که به صورت رایگان امکان استفاده از پردازنده‌های گرافیکی به کاربران می‌دهد، اجرا شده است. در ادامه این بخش به تشریح قسمت‌های مختلف پیاده‌سازی و کتابخانه‌های استفاده شده می‌پردازیم.

۴-۲ واکشی و پیش‌پردازش دفتر ثبت سفارشات

در این قسمت ابتدا نحوه‌ی واکشی اطلاعات دفتر ثبت سفارشات از صرافی آنلاین بایننس^۳ به وسیله کتابخانه‌ی سی‌سی‌ایکس‌تی^۴ را تشریح می‌کنیم. سپس به نحوه‌ی ذخیره‌سازی سری‌های زمانی و ساخت مجموعه داده برچسب خورده برای یادگیری می‌پردازیم. در انتها نیز پیاده‌سازی مدل جی‌آریو و یادگیری مدل و تابع خطا را تشریح می‌کنیم.

۴-۲-۱ واکشی دفتر ثبت سفارشات

هر رمز ارز در هر صرافی دارای دفتر سفارشات مخصوص به خود است. این دفاتر شامل سفارش‌های از جنس آن رمز ارز در آن صرافی هستند که بنا به سیاست هر صرافی می‌توان تاریخچه‌ی آن‌ها را درخواست کرد و مورد مطالعه قرار داد. در اینجا ما از داده‌های دفتر ثبت سفارشات صرافی بایننس که هم‌اکنون بزرگترین صرافی رمزارزها در دنیا است استفاده می‌کنیم. صرافی بایننس اجازه‌ی واکشی اطلاعات مربوط به دفتر سفارشات رمزارزهای گوناگونی که در آن مبادله می‌شوند را به کاربران خود می‌دهد که ما در این میان به مطالعه‌ی پنج رمزارز بزرگ‌تر می‌پردازیم.

¹Graphics Processing Unit

²Google Colaboratory

³Binance

⁴ccxt

رمزارزهای مورد مطالعه:

- بیتکوین^۵
- اتریوم^۶
- بایننس کوین^۷
- کاردانو^۸
- دوچ کوین^۹

اطلاعات مربوط به دفتر سفارشات این رمزارزها در صرافی بایننس و در بازه‌ی زمانی ابتدای ۲۰۲۰ تا ابتدای ۲۰۲۱ در این پروژه به کار گرفته شده است. برای واکنشی اطلاعات مربوط به این رمزارزها از کتابخانه‌ی سی‌سی‌ایکس‌تی^{۱۰} که یک کتابخانه‌ی مخصوص برای مطالعه و خرید و فروش رمزارزهاست استفاده شده است.

اطلاعات مربوط به هر رمزارز به صورت مجموعه‌ای از دفترهای سفارش در طول زمان و با فاصله‌ی زمانی یک دقیقه‌ای جمع‌آوری شده است این دفترها که هر کدام موجودیتی سه‌بعدی هستند (شکل ۳-۱) در طول زمان کنار یکدیگر قرار می‌گیرند و در نهایت یک آرایه‌ی چهاربعدی تشکیل می‌دهند. اما همانطور که پیش‌تر گفته شد این دفترها ابتدا به دفترهای سفارش تجمعی تبدیل می‌شوند و سپس قیمت میانی از آن‌ها استخراج می‌شود. به همین دلیل در ادامه، دیگر نیازی به دندان‌های قیمت نخواهد بود چرا که فاصله‌ی آن‌ها توسط صرافی برابر مقدار ثابتی از مقدار میانی در نظر گرفته می‌شود و به همین دلیل نیازی به ذخیره‌سازی و وارد کردن این دندان‌های با طول ثابت به مدل شبکه‌ی عصبی وجود نخواهد داشت. بر این اساس، با حذف دندان‌های قیمت و ذخیره‌سازی حجم‌های متناظر مجموعه‌داده‌ی ورودی را تشکیل می‌دهیم.

از طرف دیگر ویژگی‌هایی که در قسمت روش حل مسئله به آن‌ها اشاره کردیم را به مجموعه‌داده اضافه می‌کنیم که شامل قیمت میانی، شکاف قیمت، تغییر قیمت (بازده)، شکاف قیمت وزن دار و نوسان با

⁵Bitcoin

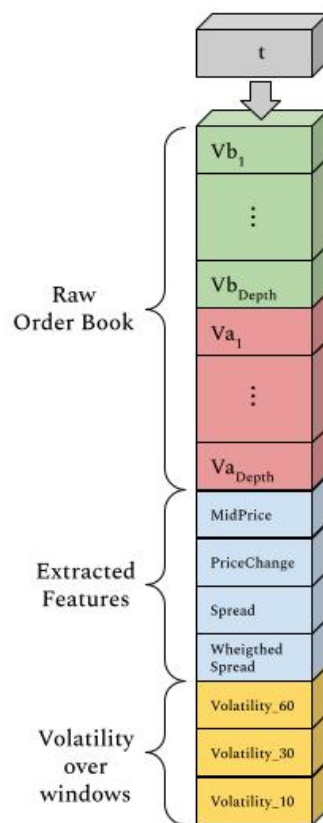
⁶Ethereum

⁷Binance Coin

⁸Cardano

⁹Doge Coin

¹⁰<https://github.com/ccxt/ccxt>



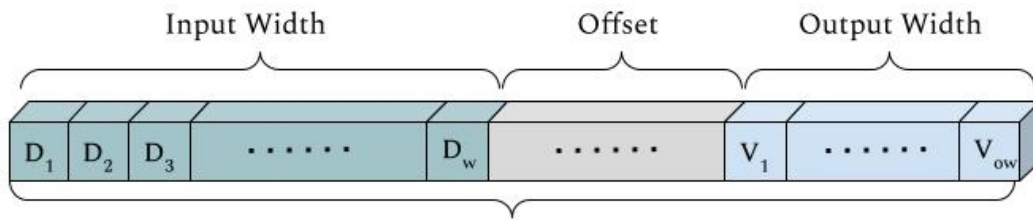
شکل ۴-۱: اطلاعات ورودی به ازای هر لحظه‌ی t در طول زمان

دامنه‌های متفاوت است.

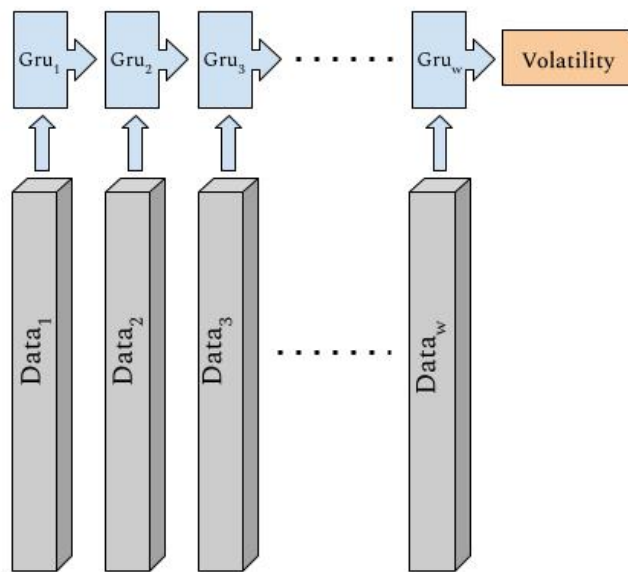
بدین ترتیب در هر نقطه از زمان مجموعه‌ای از اطلاعات جمع‌آوری و محاسبه می‌شود که می‌تواند برای یادگیری استفاده شود. در ادامه می‌توانیم با کنار هم قرار دادن دنباله‌ای از ستون‌هایی از این جنس به یک یک ماتریس دست پیدا کنیم که هر ستون آن می‌تواند به یکی از واحدهای مدل جی‌آریو به عنوان ورودی داده شود.

هر دنباله‌ی ورودی که خود شامل تعدادی از ستون‌های شکل ۴-۱ است، می‌تواند توسط یک عدد که میزان نوسان در دقیقه‌ی آینده و یا پنج دقیقه‌ی آینده بعد از آن برچسب بخورد و برای آموزش به مدل داده شود. بنابراین نحوه‌ی برچسب‌زنی مجموعه‌ی نهایی به سه متغیر طول پنجره‌ی ورودی، حاشیه^{۱۱} و طول پنجره‌ی پیش‌بینی بستگی خواهد داشت (شکل ۴-۲).

¹¹Offset



شکل ۴-۲: اطلاعات ورودی به ازای هر لحظه‌ی t در طول زمان



شکل ۴-۳: وارد شدن اطلاعات ورودی به جی‌آریو برای تخمین نوسان

به این صورت مجموعه‌داده نهایی به صورت مجموعه‌ای دنباله‌های ورودی با طول مشخص و برابر W و خروجی متناظر با نوسان در بازه‌ای در آینده با طول OW خواهد بود. طبیعی است که این مقادیر بسته به هدف استفاده از مدل و توانایی پردازشی سیستم تعیین می‌شوند.

۴-۳ مدل و یادگیری

برای پیاده‌سازی مدل جی‌آریو از کتابخانه‌ی تانسورفلو^{۱۲} استفاده کرده‌ایم. این کتابخانه که مخصوص طراحی و آموزش شبکه‌های عصبی است، شامل پیاده‌سازی انواع مختلفی از شبکه‌های عصبیست که باعث شده در سال‌های گذشته به طور گسترده به کار گرفته شود.

¹²Tensorflow

اطلاعات بدست آمده در مرحله‌ی قبل مانند شکل ۴-۳ به عنوان ورودی به جی‌آریو داده می‌شوند و در نهایت به وسیله‌ی خطای محاسبه شده آموزش داده می‌شود.

۴-۴ جمع‌بندی

در این فصل کتابخانه‌های استفاده شده را معرفی کردیم و به بیان جزئیات ساخت مجموعه داده مورد استفاده پرداختیم. در ادامه به پیاده‌سازی شبکه‌ی عصبی جی‌آریو و نحوه‌ی انجام عمل یادگیری به وسیله‌ی داده‌های برچسب خورده پرداختیم.

منابع و مراجع

- [1] Næs, Randi and Skjeltorp, Johannes A. Order book characteristics and the volume–volatility relation: Empirical evidence from a limit order market. *Journal of Financial Markets*, 9(4):408–432, 2006.

English Summary

Nowadays, machine learning models are widely used in varied applications and issues. One of the important challenges for these models, especially deep ones, is adversarial learning and adversarial examples which are generated especially for machine learning models. Natural Language Processing is also an area in which adversarial learning and adversarial examples should be considered as a major concern. In this project, our purpose is to devise a defense system against text adversarial examples generated to fool NLP models. Masked language modeling, one of the phases in pretraining of deep BERT, is used to design an approach to identify adversarial examples in text. We also use pre-trained word embeddings (GLOVE) to find the correct form of those words identified as adversarial. To evaluate our method, the performance of our defense system is measured on three transformer-based models and three english text classification datasets. The results show that our method outperforms semi-character neural networks, another novel defensive method against text adversarial examples, in all experiments.