

המכללה האקדמית גליל מערבי

בית הספר לניהול

חוג לניהול ,

חיזוי מחירי ביטקוין

מבוסס על מודלים של רשתות עצביות

מוגש על ידי:

איליה אלפר

ת.ז. 326897089

בהנחיית:

ד"ר שאמי לביב



1 מקדמה

במסגרת התואר הראשון בניהול התבקשתי לכתוב עבודה לסמינר. הבחירה שלי נפלה על נושא המטבעות המבוזרים, כי כולם מדברים על זה ורציתי להתעמק בנושא. התחלתי לחקור נושא זה. לצערי התאכזבתי ממה שמצאתי! השבע בן טיהוריה ואול האמתי: כול התענות של מטבעות מבוזרים לא מצדקים את העצמם! לדוגמה נכח את BITCOIN הו נוצר לטובת זה שאנשים לא היה תלועים ממשלות ושממשלה או איכולה לשלות בהם. בפועל זה לא ממש עובד כי מטבע תשלום הווה לא עובד ולא יכול לעבוד קח כי לדוגמה מעבר כספים איכול לקחת יומיים, כי עצמאות ממשלות זה גם לא עובד כי בורסות מחויבים לדאבך לממשל. בלוקצ'ין ואיכות הסביבה! צריחת חשמל של כריית מטבעות שבה ל 121,36 טרה-וואט שעות בשנה. לדוגמה בישראל צריחת חשמל שנתי 12554. בערך דומה אבל אנחנו משבים בן צריחה של מיליוני אנשים וצריחה של מכשיר ספקולטיבי שמשמש בנוסף לעסקאות לא חוקיות. אז למה צריך מטבעות מבוזרים? כודם כול זה נטו מכשיר ספקולטיבי למסחר בבורסות שונות. משחר בדברים לא חוקיים כמו שמים, שרותי חקרים, מסחר בנשק, ועד

במקביל לחקר במסגרת הלמודים נחשפתי לקורס של מבוא למדעי הנתונים ומאוד התחברתי לתחום זה והמשכתי ללמוד אתזה יותר לעומק מחוץ למסגרת הלמודים. זה תחום לא חדש אבל מחודש. תחום זה קיים משנות ה-80 אבל רק בשנות ה-00 היגיעו מחשבים שמסוגלים ליישם את ה"ריות מלפני 20 שנה. זה תחום ומכדם אותנו קדימה!

בעבודה זו אני רוצה להציג שילוב של שתי דברים שכיימים בעולם הזה בסחות יכולות עיבוד של מחשבים בימינו.

תוכן עניינים 2

1	מקדמה.....	2
2	תוכן עניינים.....	3
3	מבוא	4
4	סקירת ספרות.....	5
4.1	ביטקוין :	5
4.2	רשת נוירונים:	5
4.3	חיזוי סדרת זמן:.....	6
5	שיטת המחקר	6
5.1	אימון רשתות עצביות.....	6
5.1.1	שיטות לאימון רשתות עצביות	6
5.2	Adamאלגוריתם אופטימיזציה	7
5.3	LSTM - דגמי זיכרון לטווח קצר ארוך.....	8
5.4	שכבות רשת עצביות	9
5.4.1	שכבות קונבולוציוניות.....	9
5.5	שיטות להערכת איכות התחזית.....	9
6	ניסוח הבעיה.....	10
7	משתני המחקר ומדידתם	11
	במחקר נעשה שימוש בספריות הבאות:	12
8	דיון	16
9	ניתוח תוצאות חיזוי.....	17
10	סיכום.....	18
11	ביבליוגרפיה.....	19

בשנים האחרונות לא מעט מחקרים יצאו על המטבעות הדיגיטליים. תחום זה הינו תחום חדש המעלה הן שאלות אתיות והן שאלות מגוונות אחרות הדורשים הבנת מספר מושגים בסיסיים בנושא המטבעות הדיגיטליות. [1] Bohme, R., at all (2015) מסביר בצורה יסודית את אופן הפעולה של "מטבע מבוזר" כגון שימוש במטבע בקנייה, שימוש בכלכלה שחורה, ועוד.

ביטקוין מתוכנן כמערכת מזומנים עמית לעמית. [2] Baur, D., & Dimpfl, T. (2021) טוענים שכדי לעבוד כמטבע, עליו להיות יציב או להיות מגובה על ידי ממשלה. אפשר לרות שהתנודתיות של מחירי הביטקוין היא קיצונית וגבוהה כמעט פי 10 מהתנודתיות של שערי החליפין העיקריים. בגלל זה, Baek, C., & Elbeck, M. (2015) [3] מגיעים למסקנה שסוחרים עשוי להפיק תועלת משימוש במערכות מסחר לטווח קצר. חיזוי מחירים עתידיים עבור ביטקוין הוא עניין של דאגה ונושא מחלוקת עבור חוקרים. למרות זאת ניתוח מוכשר של מערכות סוציו-אקונומיות המאפשר לך ביעילות לארגן תהליכים שונים בתנאי השוק, אשר יבואו לידי ביטוי בעתיד בדינמיקה חיובית של התפתחות.

כיום, חקר שוק מטבעות הקריפטו מבטיח כיוון המחקר הכלכלי. מודלים מתקדמים של מסחר מאפשרים לחוקרים לחזות את השוק, אך לעתים קרובות יותר, בעת קבלת החלטות לגבי השקעות, רובן מונחות על ידי התרגול של מסחר באינטואיציה. בהתחשב בכך שלרוב המכריע של משתתפי השוק יש ידע בניתוח וכמובן, יכולה להיות הזדמנות לשחק ברווח בבורסה, אז הנוכחית המצב פרדוקסלי. יתרה מכך, בעת קבלת החלטה אינטואיטיבית, ולא על בהתבסס על ניתוח פרודוקטיבי, כפי שמראה בפועל, רוב הסוחרים נשארים במינוס, כלומר, הם מפסידים.

מחירי הביטקוין הם סדרת זמן מאוד תנודתית. מחירי הביטקוין מושפעים מגורמים שונים, חדשות, אירועים זמניים, סנטימנט המשקיעים והתערבות אנושית. חיזוי מחירי ביטקוין היא משימה מורכבת ומשמעותית עבור מוסדות פיננסיים ומשקיעים פרטיים. כדי להפחית ביעילות את סיכוני ההשקעה ולהשיג החזר יציב על ההשקעה, מדענים רבים הציגו מספר רב של מודלים של חיזוי. [4] Cohen, G., (2020) משתמש באופטימיזציה של נחיל חלקיקים ומוזהה שגם טכניקות Darvas Box וגם טכניקות רגרסיה לינארית יכולות לעזור לסוחרים לחזות את מגמות המחיר של הביטקוין. [5] Deng, J. et al (2019) בונים מודל תיאורטי לאבולוציה של הערך הפנימי של החוזים העתידיים.

בשנים האחרונות, חוקרים רבים יישמו אלגוריתמים שונים של למידת מכונה כדי לנתח ולחזות את מחירי הביטקוין, כמו רשתות עצביות, למידה מרובת ליבות, ניתוח רגרסיה שלב ולמידה עמוקה. Cocco, L., et al (2019) [6] השתמש באלגוריתם גנטי כדי לבחור את האסטרטגיות עם הביצועים הטובים ביותר מתוך קבוצה של כללי מסחר מוגדרים מראש. הכללים שהוגדרו מראש שולבו כדי ליצור אסטרטגיות באמצעות מוטציה אקראית. McNally, S. et al (2018) [7] השתמש ברשת עצבית חוזרת אופטימלית בייסאנית וב-LSTM כדי לחזות את כיוון מחיר הביטקוין בדולר ארה"ב. הם גם השתמשו במודל ARIMA כדי להשוות את שיטות הלמידה העמוקה. [8] Atsalakis, G. et al (2019) מתמקד בשיטת אינטליגנציה חישובית במיוחד בקר נזירופזי היברידי על מנת לחזות את שער החליפין של ביטקוין. מודל זה השתמש בגישה נזיר-מטושטש וברשתות עצביות מלאכותיות. [9] Lahmiri, S., & Bekiros, S. (2019) הטמיע אלגוריתמים של למידת מכונה כדי לחזות את שער החליפין של המחיר היומי של מטבעות קריפטוגרפיים בזמינות גבוהה כמו BTC, ריפל ומזומן דיגיטלי. הם יישמו RNN ו-GRNN (Generalized Regression Neural Network) כדי לקבל שיעור החיזוי המדויק של מטבעות קריפטוגרפיים בעלי נזילות גבוהה.

בדומה לחוכרים אחרים אני רוצה לעשות פיתוח ויישום תוכנה של אלגוריתם לעיבוד מקדים של נתונים וביצוע לאחר מכן של תחזית התנהגות של סדרות זמן פיננסיות.

4.1 ביטקוין:

ביטקוין הוא פרוטוקול תקשורת מקוון המאפשר שימוש במטבע וירטואלי, כולל תשלומים אלקטרוניים. מאז הקמתו בשנת 2009 על ידי קבוצה אנונימית של מפתחים [10](Nakamoto 2008), הכללים של ביטקוין עוצבו על ידי מהנדסים ללא השפעה ניכרת של עורכי דין או רגולטורים. במקום לאחסן עסקאות בכל שרת או קבוצת שרתים בודדים, ביטקוין בנוי על יומן עסקאות שמופץ על פני רשת של מחשבים משתתפים. הוא כולל מנגנונים לתגמול השתתפות כנה, לאתחל קבלה על ידי מאמצים מוקדמים, ולהישמר מפני ריכוזי כוח. העיצוב של הביטקוין מאפשר עסקאות בלתי הפיכות, נתיב קבוע של יצירת כסף לאורך זמן והיסטוריית עסקאות ציבורית. כל אחד יכול ליצור חשבון ביטקוין, ללא תשלום וללא כל הליך בדיקה מרוכז - או אפילו דרישה לספק שם אמיתי. באופן קולקטיבי, כללים אלה מניבים מערכת שמוכנת כגמישה יותר, פרטית יותר ופחות מתאמה לפיקוח רגולטורי מאשר צורות תשלום אחרות. ביטקוין מעניין כלכלנים כמטבע וירטואלי עם פוטנציאל לשבש מערכות תשלום קיימות ואולי אפילו מערכות מוניטריות. (Bohme, R., et al (2015)[1])

Biais, B., et al (2020)[11] בונים מודל של שיווי משקל רציונלי של ציפיות חופפות לדורות המתייחס בין המחיר של מטבע קריפטוגרפי ליסודותיו: עלויות העסקאות והיתרונות. המודל מראה כיצד יש לתמחר את היסודות הללו, ומדגיש את האינטראקציה בין המחירים העתידיים הצפויים לבין היסודות. המודל גם מראה שניתן להגדיל את תנודתיות מחירי שיווי המשקל על ידי תנודתיות חיצונית שאינה קשורה ליסודות. לאחר מכן, המכילים את משוואת תמחור שיווי המשקל, תוך הסתמכות על מערך נתונים שנאסף ביד של אירועים בסיסיים שגורמים לסוכנים לבצע עסקאות בביטקוין בקלות. באמצעות הנתונים הללו אנו בונים פרוקסי ליסודות הביטקוין: עלויות העסקאות והיתרונות שלו. חוקרים מראים כי יסודות אלה הם גורמים מכריעים של תשואות ביטקוין, ומספקים מדדים כמותיים של חשיבותם היחסית לאורך זמן. בהתאם למשמעות התיאורטית לפיה מחירי שיווי משקל יכולים להפגין תנודתיות חיצונית, הכיול שלנו גם מראה שחלק גדול מהשונות במחירים אינו מוסבר על ידי שינויים בבסיס. קבלו את החזרות הביטקוין הממומשות הגדולות מאוד, עלויות העסקאות המכילות שלנו והיתרונות גדולים מאוד, וללא ספק בלתי סבירים.

4.2 רשת נוירונים:

Lahmiri, S., & Bekiros, S. (2019)[9] החיל לראשונה את הרשתות הנוירוניות העמוקות (DNNs) על חיזוי מחירי מטבעות קריפטוגרפיים, הבהיר את יכולת הניבוי לטווח הקצר של מטבעות קריפטוגרפיים, ומצא שהדיוק הניבוי של רשתות עצביות לטווח קצר ארוך (LSTM) גבוה יותר מזה של רגרסיה כללית. רשתות עצביות LSTM [12](Hochreiter, S., & Schmidhuber, J. 1997) הוכחו כיעילות באופן משמעותי בחיזוי מחירי ביטקוין בשל יכולתן לזהות תלות ארוכת טווח ולאחסן מידע זמני לטווח ארוך וקצר. חוקרים השיגו תוצאות טובות יותר עם רשת LSTM, בעבודותיהם לחזות את מחירי הביטקוין ולהשוות אותם עם הביצועים של מודלים שונים [13](Uras, N., et al 2020), [14](Linardatos, P., & Kotsiantis, S. 2020). יתרה מכך, רשתות עצביות קונבולוציוניות (CNNs) יושמו גם לחיזוי שוק מטבעות קריפטוגרפיים. במחקר [15](Alonso-Monsalve, S. et al, 2020), חוקרים שילבו את CNN עם רשת עצבית LSTM לחיזוי מגמות שוק בתדר גבוה עבור מגוון מטבעות קריפטוגרפיים. הניתוח האמפירי שלהם מראה שהוספת שכבות קונבולוציוניות משפרת את ביצועי החיזוי וכי מבנה הרשת ההיברידי מספק את תוצאות החיזוי הטובות ביותר בניסוי. (Li, Y., & Dai, W. (2020) [16] מציעים מודל רשת עצבית היברידי המבוסס על רשתות עצביות של CNN ו-LSTM, ותוצאות ניסויים מראות שרשת עצבים היברידיים

CNNLSTM יכולה לשפר ביעילות את הדיוק של חיזוי ערך וחיזוי כיוון בהשוואה לרשת עצבית בעלת מבנה יחיד. מודל הרשת העצבית ההיברידית המשלב CNN עם LSTM שימש גם חוקרים רבים לחיזוי סדרות זמן של נתונים פיננסיים כגון חיזוי תנודתיות זהב [17](Vidal, A., & Kristjanpoller, W. 2020), מחירי מניות וכו'. הביצועים המצוינים שלו הוכחו גם בתחומים רבים אחרים (Tian, C. et al 2018), (Lei, W. et al 2017), [18][19][20].

4.3 חיזוי סדרת זמן:

חקר סדרות זמן תמיד היה תחום חשוב של למידת מכונה. על ידי בניית רשתות עצביות המבוססות על למידה עמוקה, אנו מסוגלים לחלץ ולנצל את המידע הנסתר המיוצג על ידי נתוני מטבעות דיגיטליים גולמיים של מטבעות דיגיטליים על מנת לבצע תחזיות מחיר מדויקות ויעילות [9](Lahmiri, S., & Bekiros, S. 2019). בשנים האחרונות, החוקרים התקדמו ללא הרף בתחום חיזוי סדרות זמן פיננסיות (Ariyo, A. et al 2014), (Cao, J. et al 2019), (Adhikari, R., & Agrawal, R. K. 2014), [21][22][23]. גישות מחקר מסורתיות רבות מתמקדות בלימוד דפוסים פנימיים, כגון אוטוקורלציה, בסדרת זמן מסוימת. עם זאת, עבור תרחישי מציאות, במיוחד במטבעות קריפטוגרפיים, שוקי מניות ותחומים פיננסיים אחרים, ברוב המקרים אנו עוסקים בסדרות זמן רב-משתניות. [24] Guo, Q. et al (2019) פיתח מודל סדרת זמן מבנית של למידה עמוקה לטיפול בתשומות של סדרות זמן רב-משתניות בקורלציה. המודל שלהם מסוגל למנף תלות בין מספר סדרות זמן מתואמות ולחלץ תכונות הבדלות משוקללות ללמידה טובה יותר של מגמות.

5 שיטת המחקר

בסעיף זה מתוארים הפרטים של השיטה המוצעת שלנו LSTM והבלוקים הבסיסיים.

5.1 אימון רשתות עצביות

אימון רשתות עצביות הוא תהליך שבו משקולות ופרמטרים אחרים מותאמים על ידי מודלים של הסביבה. סוג הלמידה נקבע לפי אופן התאמת הפרמטרים. האימון מתבצע עד ששגיאת הרשת העצבית מגיעה לאפס או מתקרבת לערך מקובל.

5.1.1 שיטות לאימון רשתות עצביות

ישנן ארבע פרידיגמות למידה:

1. למידה מפוקחת היא דרך שבה המערכת מקבלת תשובה נכונה ידועה מראש לכל אות כניסה. בתהליך, המשקולות מותאמות כדי למזער שגיאות. משמש בעיקר לבעיות סיווג ורגרסיה.
2. למידה ללא מורה מבוססת על זיהוי עצמי של תכונות ותלות, ומנתחת אותם. שיטת הלמידה הבלתי מפוקחת משמשת בעיקר באשכולות, זיהוי חריגות, בעיות אסוציאציות, וגם מקודדים אוטומטיים משתמשים בסוג זה של למידה.
3. צורה מעורבת מרמזת שיש כמות קטנה של נתונים עם תשובות ידועות וכמות גדולה של נתונים עם תשובות לא ידועות מראש. שיטה זו משמשת כאשר קשה להדגיש מאפיינים חשובים ולא ניתן לסמן את כל האובייקטים.
4. למידת חיזוק כוללת הכשרת גורם כלשהו בסביבה באמצעות מערכת תגמול. עבור הפעולה הנכונה, הסוכן מקבל פרס. מנקודת המבט של הקיברנטיקה, זה אחד מסוגי הניסוי הקיברנטי.

בבעיה זו ייעשה שימוש בלמידה מפוקחת לא קלאסית. בדרך כלל משתמשים ב-GAN ללמידה ללא פיקוח, במשימה זו פועלת רשת המפלה כמורה ועוזרת לבחור את הפרמטרים הטובים ביותר לרשת המפלה. הסימנים הם

הערכים הקודמים, והתשובה הנכונה היא ערך התקופה העתידית. זה ממזער את השגיאה באי-התאמה בין ערך היעד לערך שחזה המודל.

5.2 אלגוריתם אופטימיזציה ADAM

אופטימיזציה המבוססת על ירידה בשיפוע סטוכסטי (SGD) הפכה לנפוצה בתחומים רבים של מדע וטכנולוגיה, מאחר ובעיות רבות יכולות להיות מיוצגות כאופטימיזציה של פונקציה פרמטרית סקלרית כלשהי, שהאופטימיזציה שלה מבטיחה מציאת נקודת המינימום (או המקסימום). אם פונקציה ניתנת להפרדה ביחס לפרמטרים שלה, ירידת שיפוע היא טכניקת אופטימיזציה יעילה למדי, שכן לחישוב נגזרות חלקיות מסדר ראשון ביחס לפרמטרים של הפונקציה יש אותה מורכבות חישובית כמו פשוט לאמוד את הפונקציה.

פונקציות אובייקטיביות יכולות לרוב להכיל מקורות רעש, כל הפונקציות הללו דורשות שיטות אופטימיזציה סטוכסטיות יעילות. SGD הוכחה כטכניקת אופטימיזציה יעילה ויעילה שמילאה תפקיד מרכזי בסיפורי הצלחה רבים של למידת מכונה.

Adam (Adaptive Moment Estimate) - אלגוריתם אופטימיזציה שהוא הרחבה של השתלשלות השיפוע הסטוכסטית שהפכה לאחרונה לנפוצה יותר ביישומי למידה עמוקה בראייה ממוחשבת ועיבוד שפה טבעית.

השיטה Adam (Kingma, D. P., & Ba, J. 2014) [25] דורשת רק גרדיאנטים מסדר ראשון עם דרישות זיכרון נמוכות. השיטה מחשבת שיעורי למידה אדפטיבית אינדיבידואלית עבור פרמטרים שונים בהתבסס על הערכות של הרגע הראשון והשני של השיפועים. השיטה משלבת את היתרונות של שתי שיטות פופולריות לאחרונה: AdaGrad, שעובדת היטב עם שיפועים דלילים, ו-RMSProp, שעובדת היטב בהגדרות מקוונות ודינמיות.

חלק מהיתרונות של Adam הם שגדלי עדכון פרמטרים אינם משתנים לשינוי קנה מידה של גרדיאנט, גדלי הצעדים שלו מוגבלים בערך על ידי ההיפרפרמטר של גודל הצעד.

תן $f(\theta)$ להיות הפונקציה האובייקטיבית: פונקציה סקלרית סטוכסטית הניתנת להבדלה ביחס לפרמטר θ . המשימה היא למזער את הערך הצפוי של פונקציה זו, $E[f(\theta)]$ ביחס לפרמטר θ . באמצעות $f_1(\theta), \dots, f_T(\theta)$ אנו מצינים את המימושים של הפונקציה הסטוכסטית בשלבי הזמן הבאים $1, \dots, T$. סטוכסטיות יכולה לבוא מהערכת תת-דגימות אקראיות (מיני-אצט) של נקודות נתונים, או שהיא יכולה לנבוע מרעש פונקציונלי מהותי. עבור $g_t = \nabla \theta f_t(\theta)$ זהו השיפוע, כלומר. הווקטור של נגזרות חלקיות של f_t ביחס ל- θ , מוערך בשלב הזמן t .

האלגוריתם מעדכן את הממוצעים הנעים האקספוננציאליים של השיפוע (m_t) והשיפוע בריבוע (v_t), כאשר הפרמטרים ההיפרפרמטרים $\beta_1, \beta_2 \in [0, 1)$ שולטים בקצב הדעיכה האקספוננציאלי של הממוצעים הנעים הללו. הממוצעים הנעים עצמם הם הערכות של הרגע הראשון (ממוצע) והרגע הגולמי השני (שונות לא ממוקדת) של השיפוע. עם זאת, ממוצעים נעים אלו מאותחלים כאפסים (וקטורים של) אפסים, מה שמוביל לאומדני רגע המוטים לאפס, במיוחד בשלבי זמן ראשוניים ובמיוחד כאשר קצבי הדעיכה קטנים (כלומר, β קרוב ל-1). החדשות הטובות הן שניתן לנטרל את הטיית האתחול הזו בקלות, וכתוצאה מכך להערכות מותאמות הטיה m_t ו- v_t .

אחד המאפיינים העיקריים של אלגוריתם אדם הוא בחירה קפדנית של גדלי צעדים. בהנחה ש- $e = 0$, שלב

הרווח של הפרמטר האפקטיבי בשלב הזמן t הוא $\Delta t = \alpha \cdot \frac{\overline{m_t}}{\sqrt{v_t}}$. לגודל הצעד האפקטיבי יש שני גבולות

עליונים: $|\Delta t| \leq \alpha$ ו- $(1 - \beta_1) > \sqrt{1 - \beta_2}$ במקרה $(1 - \beta_1) > \sqrt{1 - \beta_2}$ ו- $|\Delta t| \leq \alpha$ אחרת. המקרה הראשון מתרחש רק במקרה החמור ביותר של דלילות: כאשר השיפוע היה אפס בכל שלבי הזמן מלבד שלב הזמן

הנוכחי. במקרים פחות דלילים, גודל הצעד האפקטיבי יהיה קטן יותר. עבור $\beta_1 = 1 - \beta_2$ יש לנו, $\frac{|mt|}{\sqrt{vt}} < 1$ אז $|\Delta t| < a$. בתרחישים נפוצים יותר, יש $\frac{\overline{mt}}{\sqrt{vt}} \approx \pm 1$ בגלל $|E[g]/\sqrt{E[g^2]}| \leq 1$.

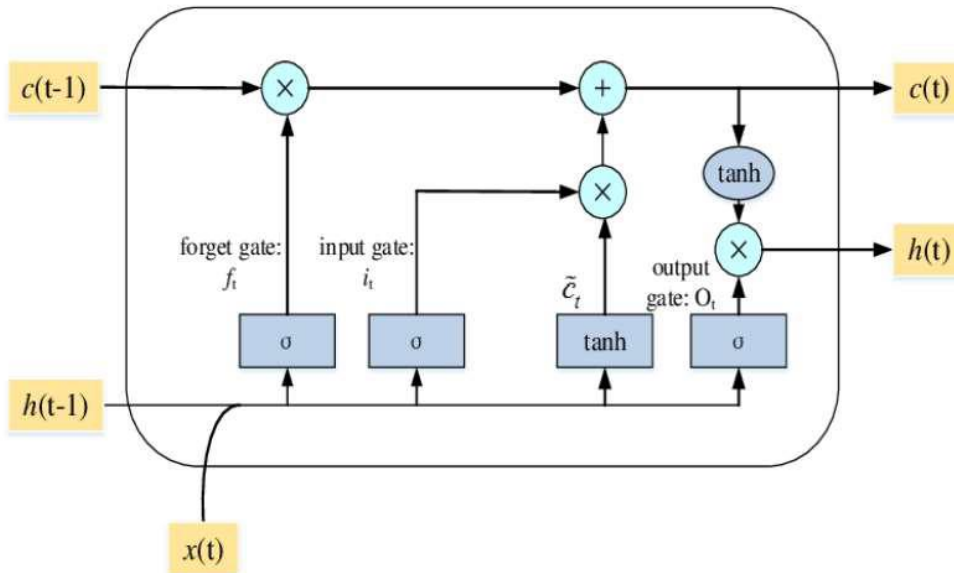
הגודל האפקטיבי של הצעדים הננקטים במרחב הפרמטרים בכל שלב זמן מוגבל בערך על ידי הגדרת גודל הצעד α , כלומר $|\Delta t| > \alpha$. ניתן להבין זאת כהקמת אזור ביטחון סביב הערך הנוכחי של הפרמטר, שמעבר לו אומדן השיפוע הנוכחי אינו מספק מידע מספיק. זה בדרך כלל מקל יחסית לדעת את הסולם הנכון α מראש. לדוגמה, עבור מודלים רבים של למידת מכונה, ידוע מראש כי אופטימיות טובות נמצאות, בסבירות גבוהה, באזור נתון כלשהו במרחב הפרמטרים; לדוגמה, אין זה נדיר שיש התפלגות מוקדמת על פרמטרים. מכיוון ש- α קובע (גבול עליון) לגודל השלבים במרחב הפרמטרים, ניתן להסיק את סדר הגודל הנכון של α , כך שניתן להגיע לאופטימום מ- θ_0 במספר כלשהו של איטרציות. עם קצת שימוש לרעה בטרמינולוגיה, בואו נקרא ליחס mt/\sqrt{vt} יחס האות לרעש (SNR). עם SNR קטן יותר, גודל הצעד האפקטיבי Δt יהיה קרוב יותר לאפס. זהו תכונה רצויה, מכיוון ש-SNR קטן יותר פירושו יותר אי ודאות לגבי האם הכיוון של mt הוא בכיוון של השיפוע האמיתי. לדוגמה, ערך ה-SNR בדרך כלל מתקרב ל-0 לקראת האופטימום, וכתוצאה מכך צעדי רווח אפקטיביים קטנים יותר של פרמטרים: סוג של חישול אוטומטי. גודל הצעד האפקטיבי Δt אינו תלוי גם בקנה המידה של ההדרגות; שינוי קנה מידה גרדיאנטים g עם פקטור c יקנה קנה מידה mt עם פקטור c ו- vt עם פקטור c^2 , אשר מבטלים: $\frac{c\overline{mt}}{\sqrt{c^2vt}} = \frac{\overline{mt}}{\sqrt{vt}}$.

בבעיה זו, אלגוריתם הלמידה של אדם משמש כשיטה בסיסית לאופטימיזציה של הפרמטרים של רשתות עצביות. החלטה זו נבעה מהעובדה ששיטה זו נמצאת בשימוש נרחב בעבודות אחרות והיא אלגוריתם מקובל לסוג זה של בעיות.

5.3 LSTM - דגמי זיכרון לטווח קצר ארוך

בין מודלים של סדרות זמן, בולטים מודלים של זיכרון לטווח קצר ארוך (LSTM), שיש להם את היכולת לחזות דינמיקה של השוק למספר שרירותי של שלבים.

התכונה הגדולה ביותר של רשתות עצביות חוזרות היא שאתה יכול לקבל מידע תלוי לטווח ארוך על הנתונים. LSTM נמצא בשימוש נרחב בתחומים רבים ועשה התקדמות משמעותית ביישומים רבים. מכיוון שה-LSTM יכול



לזכור מידע לטווח ארוך על הנתונים, העיצוב של ה-LSTM מונע את בעיית התלות לטווח ארוך. נכון לעכשיו, LSTM הוא מודל חיזוי של סדרות זמן פופולרי מאוד. מודל LSTM (תא) מורכב מ-5 רכיבים עיקריים. על ידי שימוש בממרכיבים אלה, נתונים ארוכי טווח וקצר טווח מודגמים. התא מוצג באיור 1

הבה נשקול את המכשיר של תא LSTM ביתר פירוט. כפי שניתן לראות באיור 1, הוא מורכב ממספר חלקים. מצב התא (ct) מייצג את הזיכרון הפנימי שלו. הוא מאחסן מידע לטווח קצר וארוך טווח. המצב הנסתר (ht) מספק מידע על מצב הפלט. מידע זה נקבע על ידי התאמה בין נתוני הקלט לבין המצב הנסתר הקודם. בנוסף, המצב הסמוי מאפשר לך לחלץ אחד מסוגי הזיכרון (לטווח קצר או ארוך), שני הסוגים, שיכולים להיות שימושיים בחיזוי הבא.

באמצעות שער הקלט (it), נקבעת כמות המידע המגיע מהקלט הנוכחי למצב התא. הפונקציה של מעקב אחר מידע על הקלט הנוכחי והמצב הקודם של התא הזורם למצב הנוכחי של התא מרפא בשער השכחה (ft). תפקידו של שער הפלט (ot) הוא לעקוב אחר המידע שעושה את המעבר מהמצב הנוכחי של התא למצב הנסתר. זה חשוב מכיוון שהוא מאפשר ל-LSTM לבחור סוג אחד או שניהם של זיכרונות: לטווח ארוך ולטווח קצר.

5.4 שכבות רשת עצביות

5.4.1 שכבות קונבולוציוניות

שכבת הקונבולוציה היא אבן הבניין העיקרית של רשת עצבית קונבולוציונית. שכבת הקונבולוציה כוללת פילטר משלה עבור כל ערוץ, שגרעין הקונבולוציה שלו מעבד את השכבה הקודמת פרגמנט אחר פרגמנט (מסמכים את תוצאות תוצר המטריצה עבור כל פרגמנט). המשקלים של גרעין הקונבולוציה (מטריקס קטן) אינם ידועים ונקבעים במהלך האימון. תכונה של השכבה הקונבולוציונית היא מספר קטן יחסית של פרמטרים שנקבעים במהלך האימון.

למרות ש-CNN פותחו באופן מסורתי עבור נתוני תמונה דו-ממדיים, ניתן להשתמש בהם למודל של בעיות חיזוי של סדרות זמן דו-ממדיות.

5.5 שיטות להערכת איכות התחזית

הבה נשים לב לעובדה שכאשר עושים תחזית, הם שוכחים לעתים קרובות מהדרכים לשמש בהערכת תוצאות. ובמציאות, אנחנו רואים את התחזית, אבל אין השוואה לעובדה. בנוסף, שגיאה נוספת מתרחשת לעתים קרובות למדי כאשר משתמשים בשני דגמים או יותר. יחד עם זאת, האפקטיביות והדיוק שלהם כלל לא ברורים. לכן, על מנת למנוע את החסרונות המפורטים, נעשה שימוש במדדי איכות בקביעה מספרית של ביצועים.

להדרכה והערכה השוואתית של ארכיטקטורות שונות של רשתות עצביות, משתמשים לרוב במדדים MSE (שגיאת ריבוע ממוצעת), MAPE (ממוצע שגיאת אחוז מוחלט) ו-MDA (דיוק כיווני ממוצע), מכיוון שהם מעריכים מאפיינים שונים של תחזיות מודל (שגיאה כוללת, שגיאות לגבי ערך היעד וקביעה נכונה של כיוון השינוי, בהתאמה)

כאשר אנו צריכים להתאים עקומה לנתונים שלנו, הדיוק של התאמה זו יאומד על ידי התוכנית לפי השגיאה הסטנדרטית (MSE). שגיאת הריבוע הממוצעת של השורש מחושבת על ידי הנוסחה:

$$MSE = \frac{1}{n} * \sum_i^n (y_i - \widetilde{y_i})^2$$

MSE - שגיאת RMS

n-מספר תצפיות

y_i -קואורדינטת תצפית בפועל

\tilde{y}_i — חזויה תצפית קואורדינטת

MAPE נועד לתת לדגם משמעות ויזואלית עוד יותר. מייצג ביטוי כממוצע טעות מוחלטת באחוזים. מבוטא באחוזים. מחושב לפי הנוסחה

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \tilde{y}_i}{y_i} \right|$$

MAPE -טעות מוחלטת ממוצעת באחוזים

n-מספר תצפיות

y_i -קואורדינטת תצפית בפועל

\tilde{y}_i — חזויה תצפית קואורדינטת

MAD -הסטייה המוחלטת הממוצעת מוגדרת כמנה של סכום השאריות מודולו של מספר התצפיות. מחושב לפי הנוסחה:

$$MAD = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i|$$

MAD-מתכוון לסטייה מוחלטת

n-מספר תצפיות

y_i -קואורדינטת תצפית בפועל

\tilde{y}_i — חזויה תצפית קואורדינטת

6 ניסוח הבעיה

הליקוי העיקרי שזוהה בעבודת המחקר בתחום חיזוי סדרות זמן פיננסיות הוא זה ברוב העבודות היעילות של מודלים מורכבים מוערכת בהשוואה למודלים בסיסיים מאוד. מודלים אטומיים, כגון ARIMA או LSTM, או דומים מאוד ארכיטקטורות. לכן, מחקר זה מוקדש להשוואה עקרונית סוגים שונים של דגמים. בנוסף, כמה מאמרים סותרים זה את זה, כמו כאשר משווים את אותן שיטות על נתונים שונים, מתקבלות תוצאות הפוכות תוצאות, אז בדוק את הממצאים האלה מול הביצועים בפועל.

המודלים והשכבות הבאים נבחרו למחקרים ניסיוניים:

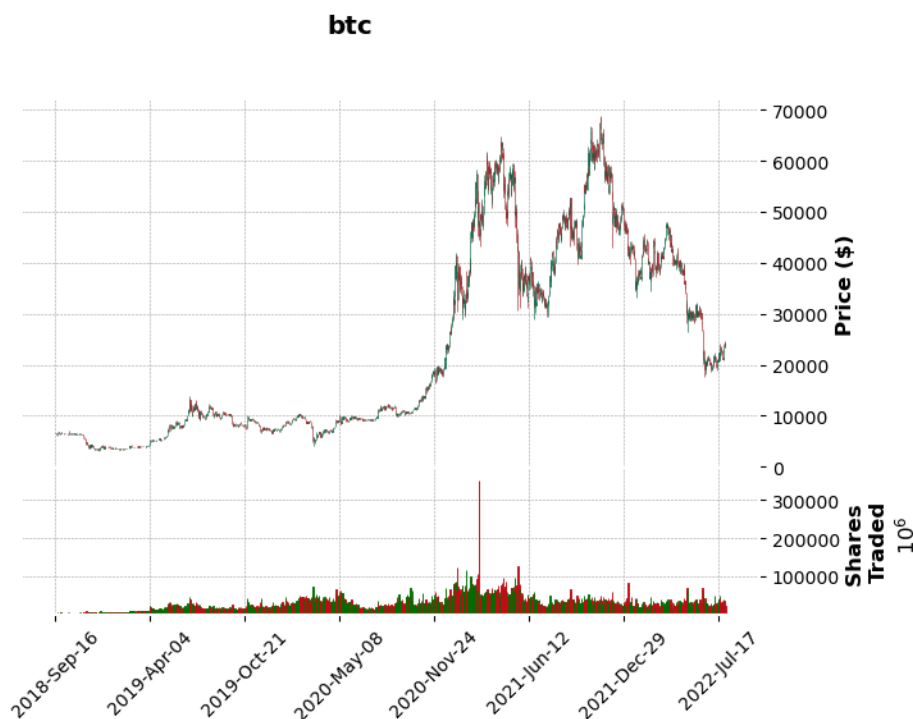
1. LSTM כרשת העצבית הבסיסית והנפוצה ביותר. LSTM נועד ליצור שכבות זיכרון לטווח קצר לטווח ארוך לאחסון מידע.
 2. Dense-שכבה מקבלת מידע מכל הצמתים של השכבה הקודמת. במילים אחרות, Dense משמש ליצירת שכבות נסתרות (מחוברות לחלוטין) של הרשת.
 3. שכבת Conv1D מחליקה את סדרת הזמן של הקלט, כך שכאשר משתמשים בשכבה זו, אין צורך להוסיף ערכי ממוצע נע או סטיית תקן נעים לפונקציות הקלט.
 4. LSTM + Conv1D + Dense - דגם היברידי המשלב את הקלאסי למידת מכונה LSTM, שכבה קונבולוציונית D1 ושכבה צפופה ואחריה להתמזג כדי לקבל תחזית יחידה.
- על מנת להעריך באופן מקיף את איכות התחזית של כל דגם, נעשה זאת המדד הבא מחושב: MSE (שגיאת ריבוע ממוצעת). השגיאה הממוצעת בריבוע מתאימה את המודל כדי למזער את הסכום הממוצע של שגיאות בריבוע.
- מחקר זה בוחן חיזוי צעד אחד קדימה למנוע בעיות הקשורות לשגיאה המצטברת של התקופה הקודמת עבור חיזוי מחוץ למערך האימונים.

7 משתני המחקר ומדידתם

על מנת לחקור באופן אמפירי את השפעת שיטות החיזוי המוצעות, מודלים של רשתות עצביות של ציטוטים יומיים של ביטקוין בנויים כמושא מחקר במאמר זה לצורך ניתוח והערכה שלאחר מכן של ערכי הניבוי של הסדרה. נתונים ראשוניים - סדרות זמן של מחזוריות יומית לתקופה מ 17/09/2018 עד 01/08/2022. מערך הנתונים נטען באמצעות ספריית yfinance. פיסת נתונים מוצגת באיור לנתונים המוצגים יש את הפרמטרים הבאים:

	Open	High	Low	Close	Volume	
Date						
2018-09-16	6536.680176	6544.330078	6460.100098	6517.180176	3273730000	Open - מחיר פתיחה יומי
2018-09-17	6514.060059	6540.209961	6257.520020	6281.200195	3910780000	Close - מחיר סגירה יומי
2018-09-18	6280.910156	6384.180176	6265.709961	6371.299805	4180090000	Low - מחיר מינימום יומי
2018-09-19	6371.850098	6448.459961	6208.339844	6398.540039	4431340000	High - מחיר מקסימלי יומי
2018-09-20	6398.850098	6529.259766	6395.950195	6519.669922	4348110000	Volume - נפח המסחר יומי
...	
2022-07-27	21244.169922	22986.529297	21070.806641	22930.548828	31758955233	Date - תאריך
2022-07-28	22933.640625	24110.470703	22722.265625	23843.886719	40212386158	
2022-07-29	23845.212891	24294.787109	23481.173828	23804.632812	35887249746	
2022-07-30	23796.818359	24572.580078	23580.507812	23656.207031	28148218301	
2022-07-31	23652.070312	24121.642578	23275.703125	23336.896484	23553591896	

האיור משרטט את ערכי מערך הנתונים עבור ביטקוין:



תהליך העיבוד המקדים של סדרת הזמן הראשונית, אימון הרשת העצבית וקבלת תוצאות חיזוי יודגם שלב אחר שלב עבור ביטקוין.

כדי ליישם מודל רשת עצבית לחיזוי עלות הביטקוין, נעשה שימוש בשפת התכנות Python Jupyter.

במחקר נעשה שימוש בספריות הבאות:

Scikit-learn – ספרייה זו מכילה יישומים מוכנים של אלגוריתמים רבים של למידת מכונה, כמו גם מדדי הערכת אלגוריתמים.

Keras – ספריית רשת עצבית המיועדת ליישום תפעולי של רשתות למידה עמוקה. הוא מכיל שכבות מוכנות, מסוגל לשלב אותן לרשת, לאמן וליישם תחזיות בעזרתו.

TensorFlow – המכונה וספריית הלמידה העמוקה מאפשרת לייעל את החישובים שנעשו במהלך האימון.

Matplotlib – מכיל סט גדול של כלים להמחשת נתונים בצורה של גרפים דו-ממדיים ותלת-ממדיים.

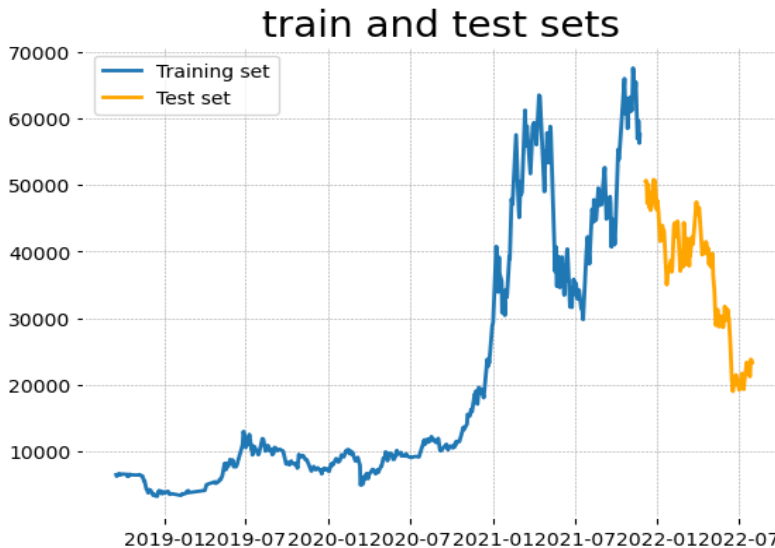
Pandas – ספרייה המיועדת לעיבוד וניתוח נתונים. ספריית pandas משמשת להצגת מידע טבלאי ויזואלית.

Numpy – ספרייה לתמיכה במערכי נתונים רב מימדיים גדולים.

מכיוון שאיכות התחזיות תלויה במידה רבה בנתוני הקלט, יש צורך גם לבחור את מערך הנתונים האופטימלי לאימון ובדיקה. מדגם ההכשרה לא צריך להיות פחות משנה על מנת לקחת בחשבון שינויים עונתיים במחירי המניות ולא צריך להיות גדול מדי כדי להימנע מהסבה מחדש של המודל וזמן מופרז ומשאבי חישוב.

בנוסף לגודל המדגם, למבנה הנתונים תפקיד משמעותי: חיזוי מחיר הסגירה של ביטקוין יכול להתבסס אך ורק על מחירי הסגירה של ימים קודמים, וכן על נתונים נוספים, כגון מחיר הפתיחה, המקסימום ומחירי הסגירה של הימים הקודמים. מינימום ליום, כמו גם נפח ליום.

לצורך פעולת הרשת כולה, יש צורך לנרמל את הנתונים המשמשים. עבור משימה זו, שיטת MinMaxScaler נקראה מספריית scikit-learn. עבור כל ערך במערך הנתונים, MinMaxScaler מוריד את הערך המינימלי ממערך הנתונים ומחלק בטווח - ההפרש בין המקסימום למינימום המקורי (ברירת המחדל היא 0 ל-1). השימוש במחלקה זו אינו מקרי - MinMaxScaler מאפשר לאובייקט "לזכור" את התכונות של הנתונים שבהם הוא ממוקם. זה לא משנה באופן משמעותי את המידע המוטבע בנתונים המקוריים.



לפני אימון הרשת העצבית, יש צורך גם לפצל את המדגם הראשוני לאימון ובדיקה. בפועל, תהליך פיצול המדגם מתבצע באופן אקראי; במקרה של חיזוי שיטה זו אינה מתאימה, שכן יש צורך בכך הרשת העצבית לא ידעה את הנתונים שעליה לחזות. לשם כך, המדגם מחולק במפורש לשני חלקים ביחס מסוים. בעבודה זו החלוקה מתרחשת ביחס של 6 ל-1. האיור מציג תרשים המציג את התפלגות מדגם המחיר עבור דוגמאות ההדרכה והמבחן.

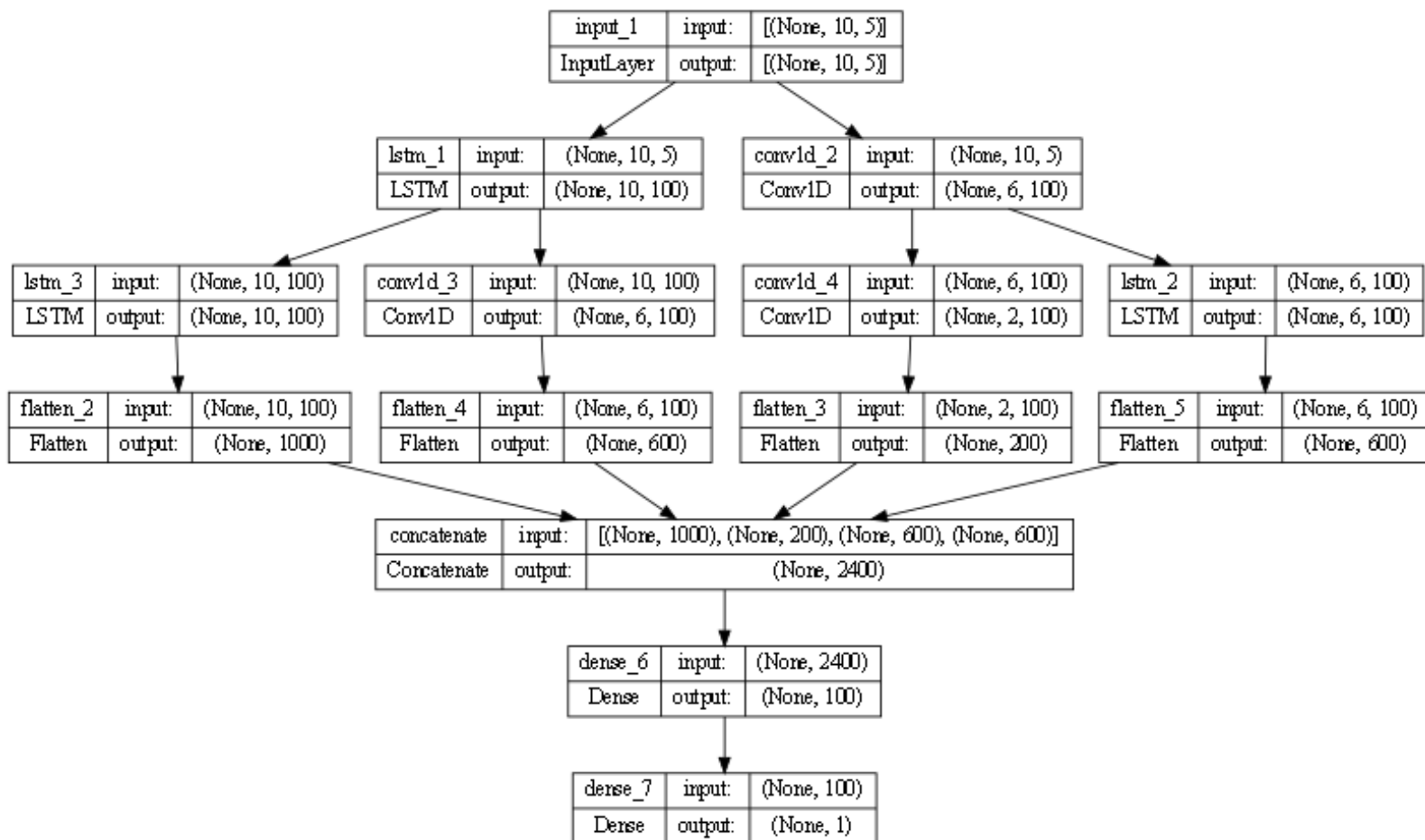
עבור סדרת הזמן הראשונית של מחירי הביטקוין, מודל LSTM נבחר בניסוי, שכבות Conv1D בהתחלה ושכבות צפופות בסוף, עם השגיאה הריבועית הממוצעת הקטנה ביותר בעת חיזוי ליום אחד. מודל זה יישמש לביצוע התחזית ליום הבא.

כדי לפתור את הבעיה, פותחה רשת עצבית LSTM-Conv1D-Dense, בעלת הטופולוגיה שתוארה לעיל. האלגוריתם של המודל מוצג באופן הבא: נתונים מוכנים ממערך הנתונים מוזנים לקלט של המודל. יתר על כן, הנתונים מעובדים ומועברים דרך השכבות LSTM, Conv1D, Flatten, Concatenate, Dense. כל אחת מהשכבות הללו מבצעת פעולות מתמטיות מסוימות, וכתוצאה מכך המשקולות של מודל זה משתנות. לאחר אימון הרשת על נתוני האימון, מסופקים נתוני בדיקה שהרשת העצבית עדיין לא ראתה. בדרך זו ניתן לדעת את דיוק המודל. כדי לפשט, זה הגיוני להביא את הפרמטרים של הרשת העצבית בטבלה

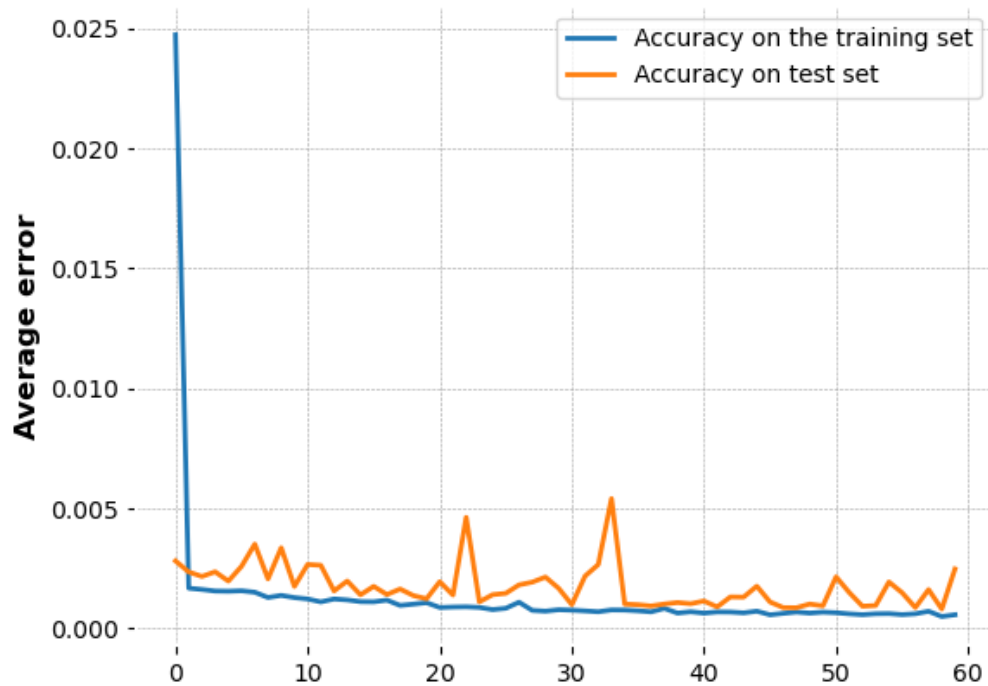
מספר שכבות LSTM	3
מספר השכבות המחוברות במלואן (Dense)	2
מספר שכבות חד ממדיות (Conv1D)	3
מספר שכבות טרנספורמציה (Flatten)	4
פונקציית הפעלה	linear, relu
גודל אצווה	5
מספר תקופות	60

אנו משתמשים בפונקציות ההפעלה הבאות:
 1. Unit Linear Rectified (relu) היא פונקציית ההפעלה הנפוצה ביותר עם למידה עמוקה. פונקציה זו מחזירה 0 אם היא לוקחת שלילי ארגומנט, במקרה של ארגומנט חיובי, הפונקציה מחזירה את המספר בעצמה.

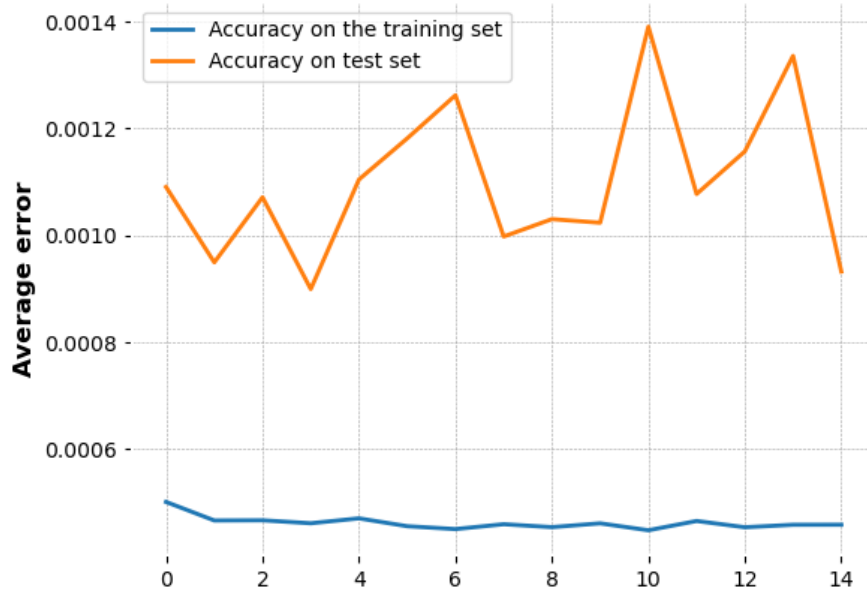
2. פונקציה לינארית (linear) היא קו ישר, כלומר התוצאה של פונקציית הפעלה זו פרופורציונלית לארגומנט שעבר. בניגוד הפונקציה הקודמת, היא מאפשרת לך לקבל מגוון של ערכי פלט, ולא רק בינארי 0 ו-1, אשר פותר את בעיית הסיווג עם מספר רב של מחלקות.



ב-60 עידינים, אתה יכול להבין למה המודל שלנו מסוגל, כלומר, הוא לומד ומגיע לשיאו, ולאחר מכן אין טעם לאמן אותו עוד יותר. דיוק מודל LSTM-Conv1D-Dense



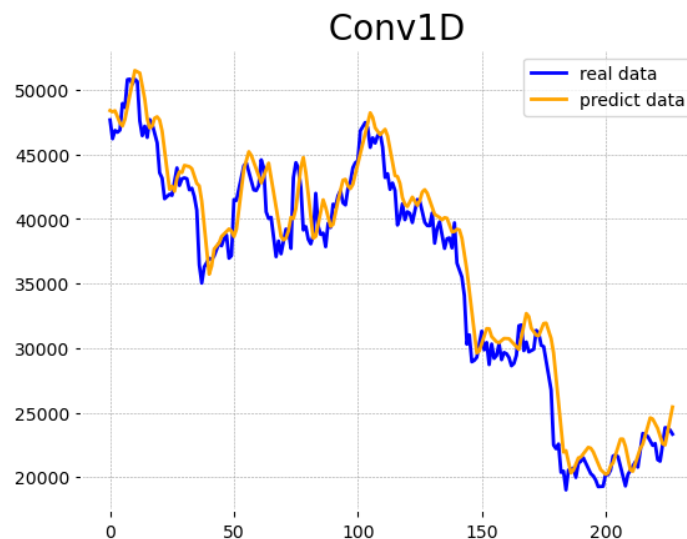
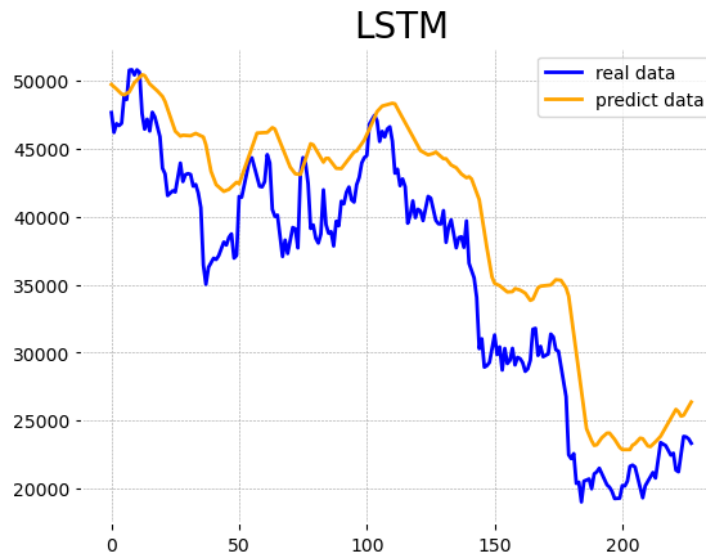
הגרף יורד בהתמדה, השגיאה פוחתת, ואז אתה עדיין יכול לאמן את הרשת. הכשרה נוספת התקיימה ב-15 עיידים, אבל איך ככלל בתכנות, הציון מתחיל מ-0. בוא נעריך את איכות האימון. ציר ה-x מייצג את עידן האימון. הדיוק על מערך ההדרכה ועל מערך האימות מוצג באיור.

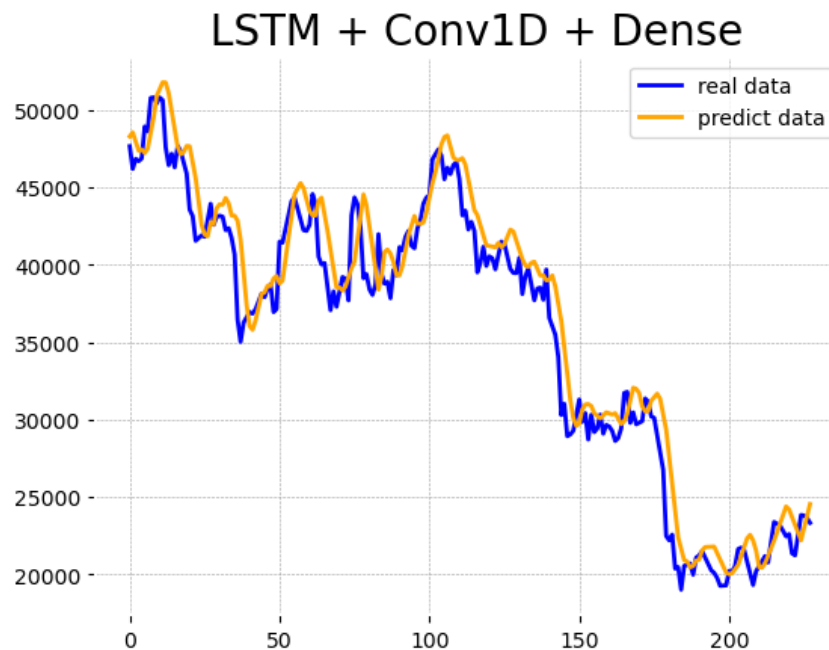


בדיקת המודל שפותח לחיזוי סדרות זמן של מחירי ביטקוין

ברגע שנוצר מודל LSTM-Conv1D-Dense מאומן היברידי, מתבצע חיזוי ישיר. כדי לבדוק את המודל שפותח, חיזנו את ההתנהגות של סדרות זמן של מחירי סגירת ביטקוין ליום אחד קדימה באמצעות מחירי הפתיחה, הגבוהים, הנמוכים ביותר והסגירה של 10 הימים הקודמים. עבור התחזית, נלקחים ערכי 10 הימים הראשונים ועליהם נבנית תחזית ליום ה-11, כלומר עוברים על הערכים בשלב של יום אחד, מנתחים את ה-10 האחרונים ימים ובונים תחזיות ליום ה-11, מהיום ה-2 עד ה-11 בונים תחזית ליום ה-12, מהיום ה-3 ועד היום ה-12 נבנית תחזית ליום ה-13, וכן הלאה.

הדמיה של נתונים ראשוניים וחזויים עלפי מודלים





9 ניתוח תוצאות חיזוי

לאחר ביצוע התחזית עבור כל חברה מהרשימה, מחושבים האינדיקטורים לאיכות התחזית שהתגשמה

	Conv1D	LSTM	LSTM-Dense-Conv1D
MSE	0.979	0.962	0.979
MAPE	4.719	18.252	4.719

בניתוח ההערכות המתקבלות של התחזיות, אנו יכולים להסיק שהיישום של הרשת העצבית של התפיסה הרב-שכבתית עבור בעיית החיזוי מוצלח. ברור שהמודל ההיברידי LSTM-Dense-Conv1D עדיף במידה רבה על ה-LSTM מבחינת איכות התחזיות שנבנו.

חיזוי מחירי הביטקוין ממלא תפקיד חשוב בשיקוף המגמות הכלליות של שוק מטבעות הקריפטו ויש לה ערך השקעה מעשי גדול. כך, כתוצאה מהעבודה, הושגה מטרת המחקר: בוצע ניתוח שיטות חיזוי, פותח מודל רשת עצבית היברידית.

כתוצאה מהניתוח התיאורטי בוצעה סקירה של השיטות הקיימות לחיזוי מחירי שוק הקריפטו. בנוסף, בוצע ניתוח של עקרונות הפעולה והתכונות של רשתות עצביות ישירות, חוזרות ונשנות, וכן קבוצות נבחרות של מודלים היברידיים ושיטות ספציפיות שנבחרו למחקר ניסיוני. הניתוח שבוצע איפשר לזהות מגמה של סיבוך של שיטות חיזוי ופופולריות של מחקרים על סוגים חדשים של מודלים היברידיים. כמו כן זוהה פער מחקרי, המורכב מהיעדר מחקרים המשווים סוגים שונים של מודלים היברידיים ורשתות עצביות בינם לבין עצמם.

במהלך העבודה, נוצר מודל היברידי LSTM-Conv1D-Dense לחיזוי מחירי סגירה, המיישם את הארכיטקטורה של רשת עצבית חוזרת הבנויה על אלמנטים של זיכרון לטווח קצר ארוך (LSTM), שכבה צפופה ואחד -שכבה קונבולוציונית ממדית. מודל זה נבחר בשל העובדה שמערך הנתונים של מחירי הסגירה של ביטקוין הוא סדרת זמן ויש צורך להתייחס אליו באופן קבוע ולקחת בחשבון את ההקשר לטווח ארוך. הרשת העצבית החוזרת לזיכרון לטווח קצר LSTM עושה עבודה טובה במשימה זו.

עלייה במספר שכבות LSTM הובילה לשיפור משמעותי בתחזית, המראה את יעילות המודל המוצע. עם זאת, הליך זה צריך לקחת בחשבון את הארכיטקטורה של ציוד המחשב, שכן עלייה נוספת במספר השכבות יכולה להוביל לתוצאות גרועות. הוכח שהוספת שכבת צפופה נוספת מחוברת במלואה לפני שכבת הפלט יכולה להפחית את שגיאת חיזוי הרשת, ובכך לשפר את ביצועי המודל.

מחקר זה יכול להיות שימושי הן עבור חברות ניהול ומחלקות השקעות, והן עבור משקיעים פרטיים. מחקר נוסף בתחום זה יכול להיות מוקדש הן לניתוח השוואתי של האפקטיביות של רשתות עצביות ושיטות חיזוי היברידיות שאינן מוזכרות בעבודה זו, והן ליצירה. של דגמים היברידיים חדשים.

1. Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of economic Perspectives*, 29(2), 213-38.
2. Baur, D. G., & Dimpfl, T. (2021). The volatility of Bitcoin and its role as a medium of exchange and a store of value. *Empirical Economics*, 61(5), 2663-2683.
3. Baek, C., & Elbeck, M. (2015). Bitcoins as an investment or speculative vehicle? A first look. *Applied Economics Letters*, 22(1), 30-34.
4. Cohen, G. (2020). Forecasting Bitcoin trends using algorithmic learning systems. *Entropy*, 22(8), 838.
5. Deng, J., Pan, H., Zhang, S., & Zou, B. (2021). Optimal bitcoin trading with inverse futures. *Annals of Operations Research*, 304(1), 139-163.
6. Cocco, L., Tonelli, R., & Marchesi, M. (2019). An agent-based artificial market model for studying the bitcoin trading. *IEEE Access*, 7, 42908-42920.
7. McNally, S., Roche, J., & Caton, S. (2018, March). Predicting the price of bitcoin using machine learning. In *2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP)* (pp. 339-343). IEEE.
8. Atsalakis, G. S., Atsalaki, I. G., Pasiouras, F., & Zopounidis, C. (2019). Bitcoin price forecasting with neuro-fuzzy techniques. *European Journal of Operational Research*, 276(2), 770-780.
9. Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons & Fractals*, 118, 35-40.
10. Nakamoto, S., & Bitcoin, A. (2008). A peer-to-peer electronic cash system. *Bitcoin*.—URL: <https://bitcoin.org/bitcoin.pdf>, 4, 2.
11. Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., & Menkveld, A. J. (2020). Equilibrium bitcoin pricing. *Available at SSRN 3261063*.
12. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
13. Uras, N., Marchesi, L., Marchesi, M., & Tonelli, R. Forecasting bitcoin closing price series using linear regression and neural networks models (2020). *arXiv preprint arXiv:2001.01127*.
14. Linardatos, P., & Kotsiantis, S. (2020). Bitcoin price prediction combining data and text mining. In *Advances in Integrations of Intelligent Methods* (pp. 49-63). Springer, Singapore.
15. Alonso-Monsalve, S., Suárez-Cetrulo, A. L., Cervantes, A., & Quintana, D. (2020). Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators. *Expert Systems with Applications*, 149, 113250.
16. Li, Y., & Dai, W. (2020). Bitcoin price forecasting method based on cnn-lstm hybrid neural network model. *The journal of engineering*, 2020(13), 344-347.
17. Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, 157, 113481.

18. Tian, C., Ma, J., Zhang, C., & Zhan, P. (2018). A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. *Energies*, 11(12), 3493.
19. Lei, W., Wang, X., Liu, M., Ilievski, I., He, X., & Kan, M-Y. (2017). SWIM: A Simple Word Interaction Model for Implicit Discourse Relation Recognition. In C. Sierra (Ed.), *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)* (pp. 4026-4032). International Joint Conferences on Artificial Intelligence.n.
20. Lei, W., Jin, X., Kan, M. Y., Ren, Z., He, X., & Yin, D. (2018, July). Sequicity: Simplifying task-oriented dialogue systems with single sequence-to-sequence architectures. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1437-1447).
21. Ariyo, A. A., Adewumi, A. O., & Ayo, C. K. (2014, March). Stock price prediction using the ARIMA model. In *2014 UKSim-AMSS 16th international conference on computer modelling and simulation* (pp. 106-112). IEEE.
22. Adhikari, R., & Agrawal, R. K. (2014). A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications*, 24(6), 1441-1449.
23. Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical mechanics and its applications*, 519, 127-139.
24. Guo, Q., Lei, S., Ye, Q., & Fang, Z. (2021, July). MRC-LSTM: A hybrid approach of multi-scale residual CNN and LSTM to predict bitcoin price. In *2021 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
25. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.