

# חיזוי שערי מט"ח באמצעות

## RNN – LSTM



נתנאל שושן

בהנחיית ד"ר מיה הרמן

א2023

# תוכן עניינים

<b>3</b>	<b>1 מבוא</b>
3	1.1 המ撒ר בשוק ההון
3	1.1.1 רקע
3	1.1.2 סקירת סוגים מסחר שונים
4	1.1.3 אתגרים בתחום
5	1.2 סוגים ניתוחים והשפעתם על חיזוי השער
5	1.2.1 ניתוח טכני (Technical analysis)
5	1.2.2 ניתוח בסיסי (Fundamental analysis)
<b>6</b>	<b>2 מבוא לבינה מלאכותית</b>
6	2.1 מבוא
6	Artificial Neural Networks (ANNs)
8	2.2.1 סוג רשתות ושימושיהן בעולם האמיתי
13	Deep learning for timeseries
13	2.3.1 סקירת השימושים השונים של ניתוח סדרות זמן
<b>14</b>	<b>3 חיזוי מחيري מט"ח באמצעות רשתות LSTM</b>
14	3.1 מבוא
14	3.2 משתנים בעלי השפעה על חיזוי שעריו המט"ח
17	3.3 חיזוי מחيري מט"ח באמצעות רשת LSTM
17	3.3.1 מבוא
18	3.3.2 הגדרת המודל
28	3.3.3 השוואה בין מודל LSTM היברידי לבין מודל LSTM רגיל
29	3.3.4 תוצאות ומסקנות – הערצת ביצועים
<b>31</b>	<b>4 סיכום</b>
<b>32</b>	<b>5ביבליוגרפיה</b>

# 1 מבוא

## 1.1 המסחר בשוק ההון

### 1.1.1 רקע

המסחר במט"ח ובסחרות מתחיל בעבר הרחוק, עוד לפני הספירה, מתוך הצורך לשלם במטרות שונות עבור טובין מסווג – טקסטיל, מזון, החלפת כספים עבור אנשים שנסעו ממדינה למדינה (בכל אמצעי שהוא) ועוד. במהלך השנים, מותט מט"ח שהבנקים החזיקו עליה בהדרגה יחד איתה, ולו גם המסחר, במותט המשתתפים וההיצע. ביום, אין מקום פיזי אליו כלם מתכנסים, אלא רוב המסחר במטרות חוץ / סחרות ו- מנויות מתבצע בצורה אלקטרוניות. שוק המט"ח פועל 5/24 בשונה משוק המניות שם השעות משתנות בהתאם לשוק בו נסחרת המניה (NYSE/HKG/TLV).

### 1.1.2 סקירת סוגי מסחר שונים

**Commodities Trading** – המסחר בסחרות התחליל בוצרים כלליים כגון בקר, חיטה, תירס, פול, סוויה ועוד. בהמשך המסחר התרחב אל מוצרים קיימים בכל השוקים כגון: גז טבעי, נפט, קפה, דלק, פחם, סוכר, אלומיניום, אורז, בסף ועופרת וכיום קיימות Commodity Exchanges בהן ניתן לבצע מסחר בצורה ישירה או לבצע מסחר דרך ברוקרים שמקשור באופן עקיף עם בורסות אלו ( OTC ) ומאפשרים לקנות לשלוח בסחרות אלו. ( בהמשך, במידה והאפשרות OTC – Over The Counter קיימת, אציין את זה בסוף הפיסקה)

**Indices Trading** – "אינדקס" הוא שיטה למדידת ביצועים של קבוצת מנויות. דוגמה נפוצה תהיה S&P 500. קבוצה זו מכילה בתוכה את 500 החברות הגדולות ביותר בארה"ב שנסחרות בבורסות בארה"ב וערוכה מהוות ממד לביצועי המניות של אותן חברות. אינדקסים נחזרים דרך בורסה בה הם רשומים או באמצעות OTC.

**Stock Trading** – מניה היא נייר ערך שמקנה חלק בעלות בחברה מסוימת. המסחר במניות מתיחס לקנייה ומכירת מנויות ופונה יותר לsonian פעילים בשוק, שימושיים לטוח קצר. כמו כן, מניה נסחרת בבורסה בה החברה הונפקה. לדוגמה העניין, TSLA נסחרת ב-NASDAQ בעוד TEVA נסחרת בבורסת תל-אביב (TLV). שוק המניות ושוק האג"חים נכנסים תחת "שוק זהן".

**Derivative Trading** – שוק המעו"ף ( מכירים עתידיים ופיננסים ) הוא שוק של מכירים פיננסים שחסרים בבורסה, ונגדים מנכסים מסוימים. השוק מורכב מאופציות ( Options ) וחוזים עתידיים ( Futures ). נגזרת יכולה להיסחר דרך בורסה או באמצעות OTC.

**Forex Trading** – הנושא שבו עוסוק. המסחר במט"ח הוא התהילך של קניה / מכירה של מטבעות. שוק המט"ח ייחודי מכמה סיבות: הוא הגדל בעולם לפי ווליום, שעות הפעולות שלו ( 5/24 ), המשתנים המשפיעים על שעריו המטבעות, יכולת להשתמש במינוף ועוד אפשרויות מסחר שונות שעלייה נדבר בהמשך.

### 1.1.3 **אתגרים בתחום**

המטרה בכלל מסחר היא למקסם את ה *Return on investment* – ROI. חלק מהשאלות הנשאלות לפני השקעה יהיו: מהם הפרמטרים שמהווים אומדן?-האם כדאי להשקיע באותו המוצר / ביחסון / אג"ח / מטבע / ני"ע ? כל השקעה ברוחבה בסיכון מסוים. מהם הסיכון? כיצד ניתן סיכונים בצורה נכונה תוך שמירה על רוחניות בתיק? מהם הגורמים המשפיעים על ההשקעה?

ובכן, שאלות אלו מהוות חלק מזמן מבנק שאלות גדול ותנאים שונים שנובל להציב לפני השקעה. האתגרים בהם הסוחר עמד תלולים בסוג ההשקעה אותה הוא מעוניין לבצע. בכלל שוק מסחר ישנים גורמים המשפיעים על המסחר. לדוגמה: בשוק המניות – דוחות רבוניים של החברה, עמידה בתחזיות ובמות המשקיעים. לכל שוק ישנים גורמים בעלי השפעה על התנהלות שעריו המחייר. אתגר של הסוחר הוא לבצע ניתוח נכון של הנתונים וכותצאה מכך להבין את מצבו של השוק ולקבל החלטה לגבי אותו הנכס בו הוא רוצה להשקיע, כיצד ישקיע והואפן שבו תתנהל ההשקעה.

## 2.1 סוגים נייחים והשפעתם על חיזוי השער

### 1.2.1 ניתוח טכני – Technical Analysis

ניתוח טכני הוא תהליך שמאפשר לסוחר להזות הזדמנויות מסחר. הדרכ לזהות הזדמנויות מסחר יכולה להיות ע"י צירוף אינדיקטורים סטטיסטיים, תנודות מחיר, ווליום מסחר ועוד דרכים ושיטות נוספות עליהן נרחיב בחלק 3. התהליך יכול להתבצע ע"י הסוחר או מערכת אוטומטית שתחשב ותפיק Signal לקניה / מכירה ותבצע את המחסר ללא מגע אדם.

### 1.2.2 ניתוח בסיסי – Fundamental Analysis

ניתוח בסיסי, כניתוח טכני – מטרתו היא לסייע לסוחר אינדיקטיה לחיזוי הזדמנויות מסחר. בניתוח זה, הסוחר / מערכת מסתמכים על דוחות כגון – ריביות, תל"ג, אחוז אבטלה, CPI ועוד'. לכל אחד מdochות אלו ישנה תחזית. עמידה בתחזית או חריגה ממנה – יכולה לגרום לעליות או ירידות בהתאם.

לדוגמא: דוח CPI – Consumer Price Index U.S. דוח זה מודד את השינוי במחיר של טובים ביחס לצרכן. ביום, הוא מהווים מפתח למדידת שינויים בטרנדים של רקישות או אינפלציה.

נכיה והערך בדוח האחרון שפורסם היה 7.7%, וברגע ישנה תחזית של 7.3%.

תוצאה שקטנה מהתחזית תהווה סימן לירידה בערך הדולר (Bearish) ותוצאה גדולה מהתחזית תהווה סימן לעלייה בערך הדולר (Bullish).

קייםות שיטות מסחר המשלבות את שני סוגי הניתוחים. הניתוחים יכולים להתבצע באמצעות הסוחר או באמצעות מערכות אוטומטיות.

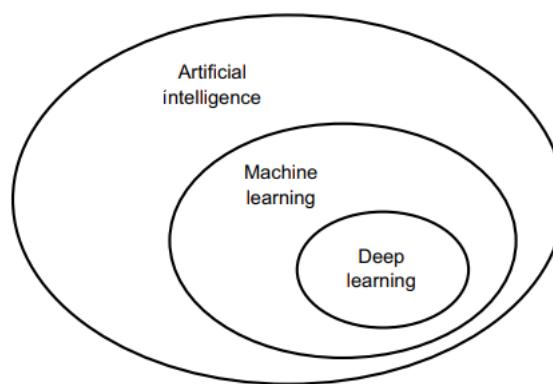
## 2 מבוא לבינה מלאכותית

### 2.1 מבוא

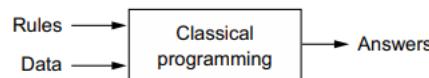
בפרק זה נסביר מהי בינה מלאכותית, ממה היא מורכבת, מהו המושג "רשות נירוגנים מלאכותית" והמילה AI שאנו שומעים הרבה בשנים האחרונות. נסקר בקצרה את שימושה בעולם האמיתי ולאחר מכן את ההגדרות נצלול אל הגדרת סוגים מיידת עמוקה והשימושים בהם.

### Artificial Neural Networks (ANNs) 2.2

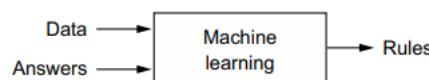
לפני שנתחיל בהגדרת המושג "למידה עמוקה", נענה על השאלה: מהי בינה מלאכותית? בינה מלאכותית ניתנת לתיאור באמצעות מרכיבות הדורשות חשיבה, שמתבצעות בדרך כלל על אדם. היא משלבת בתוכה תחומיים כגון למידה חישובית, למידה עמוקה וגישות נוספת שלא מצריכות מידע למידה<sup>[6]</sup> (Symbolic AI)



אחד הדריכים לגרום למחשב להיות שימושי היא בעזרת כתיבת סט חוקים שנקבע מראש. המחשב יקבל את הנתונים והחוקים ויפיק תשובה בהתאם, כפי שניתן להתרשם באירוע מטה, גישה זו היא בעצם גישת הפיתוח הקלסית שאנו מכירים.



לעומת זאת, התהילך של למידה חישובית (Machine Learning) שונה מגייסת הפיתוח הקלסית. גישה זו מורכבת מקלט התשובות והנתונים והיא תפיק את החוקים. לדוגמה: אם ירד גשם ביום מסויים. המערכת תקבל נתונים עבר ייחד עם התשובות של האם ירד גשם ועפ"י חוקים סטטיסטיים, היא תלמד להפיק תוצאות בהתאם.



קיימים מגוון רחב של סוגי למידה שונים, בהםם:

**למידה מפוקחת** – המודל מקבל בקלט דוגמאות של קלטיים ופלטאים תקינים, לומד את החוקיות ומהזיר תחזית על סט נתונים מבחר שיקבל בהמשך ובכך ניתן להעריך את רמת דיוק המודל. בקטגוריה זו נכללות בדרך כלל בעיות סיוג שמטרתן לחזות סוג מחלוקת, כגון זיהוי ספרות בכתב יד לפי תמונה, ובעיות גרסיה שמטרתן לחזות ערך נומרי, כגון מחירידירות לפי מאפיינים מסוימים.

**למידה לא מפוקחת** – המודל משתמש בדוגמאות שאין עליהם פליים נוכנים. ככלור על הסוכן ללמידה התבניות מובילו לשיסופק לו משוב. בקטגוריה זו נכללות בדרך כלל בעיות כמו בעיות הקבוצה (Clustering) שמטרתה לחלק מידע לקבוצות על סמך מאפיינים זהים בין הפריטים, הערכת צפיפות (Density Estimation) שמחפשת לקבוע את ההסתברות של קבלת קלטיים מסוימים, הצגת נתונים (Scatter plot) והктנתה מייד (Visualization) שמציגה נתונים בהתאם על ערכי משתנים (Dimensional Reduction) ש命ירה מרחבים רב-ממדים לבעיות עם מספר ממדים מצומצם יותר.

**למידת חיזוקים** – טכנית אימון שבה "הסוכן" פועל בסביבה ולומד באופן עצמאי את החוקיות באמצעות פידבק שמקבל. הסוכן ינסה למיקם את ערך ה-reward ולפי כן, ילמד את החוקיות או "סט הפעולות" הנדרש כדי לקבל את ה-reward הגבוה ביותר על הפעולה שביצע. בקטגוריה זו נכללות בעיות כגון משחק מחשב בהם הסוכן יכול לקבל פידבק של האם ניצח או לא ולפי כן ינסה ללמידה ולמיקם את הפעולות שלו על מנת שניצח. דוגמה נוספת יכולה להיות Google's AlphaGo.

בקשר של סינר זה, למידת חיזוקים מעורבת משלבי מרכיבים עיקרים, הסוכן והסביבה. במהלך אימון הסוכן,\*socen\* עובר דרך היסטוריים (Data Candles) אחד אחר השנוי ותמיד יבצע פעולה אחת בלבד: בניסת קניה, יצאה מקניה, כניסה מכירה, יצאה מכירה, נטרלי). במהלך תהליכי האימון, הסבירה תיעקב אחר פעולות וביצועי הסוכן ותתגמל אותו בהתאם בעוצמת פונקציית ה-reward.

ניתן להשתמש בכל אחת מסוגי הלמידה שציינו לעיל למטרת חיזוי שערי מט"ח. למידת חיזוקים נוספת שכבבה של הסתגלות ותגובה לשוק שימושיים וגרסאות לא יכולים.\* לעומת זאת, למסוגים וגרסאות ייש את היתרון של חיזוק מדויק ואמין.

\*socen RL שאומן בצורה לא נכונה יוכל למצוא "שיטות"- "טריקים" לרמות כדי למיקם את התגמול מובילו לניצח במסחר.

## 1.2.2.1 סוגים של רשתות ומערכות המבוססות על דוגמאות

1. **(Feedforward Neural Network) FNN:** זהו הסוג הבסיסי ביותר של רשת נירונים שבה הנתונים זורמים בכיוון אחד, משבבת קלט לשכבות פלט, מבלי לחזור אחרת. סוג זה משמש למשימות כגון סיווג תמונות, תרגום שפה וזיהוי דיבור.

2. **(Convolutional Neural Network) CNN:** סוג זה נועד לעבד נתונים בעלי טופולוגיה דמוית רשת, כגון תמונת. הוא משמש בדרך כלל ליזוי תמונות וניתוח וידאו.

3. **(Recurrent Neural Network) RNN:** סוג זה נועד לעבד נתונים עוקבים, כגון סדרות זמן או שפה טבעית. הוא יכול לטפל בקלט באורכים שונים ולשמור על מצב פנימי (state) שנייה. יהיה להשתמש בו כדי לעבד את הקלט הבא ברצף הנתונים.

4. **(Long Short-Term Memory) LSTM:** סוג זה הוא וריאציה של RNN שנועדה להתגבר על בעיית "השיפוע הנעלם" (Vanishing Gradient) \*. הוא משתמש בדרך כלל למשימות עיבוד שפה טבעית כגון מודול שפה, זיהוי דיבור, תרגום מכונה, יצירת טקסט, ניתוח סנטימנטים, חיזוי סדרות זמן, ניתוח וידאו, ועוד.

5. **(Gated Recurrent Unit) GRU:** סוג זה הוא וריאציה נוספת של RNN שנועדה להתגבר על הבעיה "השיפוע הנעלם". הוא משתמש בשערים כדי לשלוט בזרימת המידע והוא יעיל יותר מבחינה חישובית-M-LSTM.

6. **(Autoencoder) AE:** סוג זה נועד ללמידה ייצוג קומפקטי של נתונים הקלט, שנקרא קידוד. משמש בין היתר משימות כגון הפחמת מימד (Dimensionality reduction) וזיהוי חריגות (Anomaly Detection).

7. **(Generative Adversarial Networks) GAN:** סוג זה נועד ליצור נתונים חדשים שדומים למערך נתונים נתון. הוא מורכב משתי רשתות: generator ו-discriminator. ה-"מחולל" יוצר נתונים חדשים וה-"מאפיין" מנסה לבדוק בין הנתונים שנוצרו לבין הנתונים האמיתיים.

8. **(Variational Autoencoder) VAE:** סוג זה הוא וריאציה של מקודד אוטומטי (AE), המשמש במקודדים ובמפענחים הסתברותיים שמאפשרים לו ליצור דגימות חדשות שדומות לנוטני האימון. הוא משמש בדרך כלל למשימות כמו יצירת תמונות ועיבוד שפה טבעית.

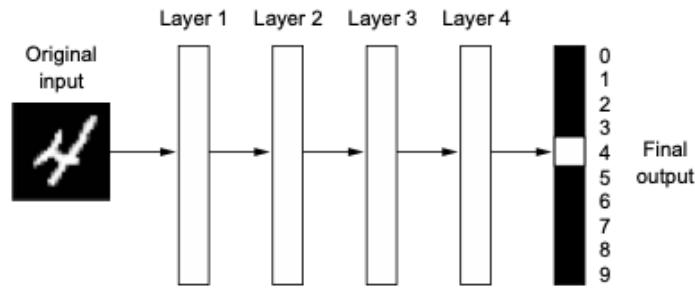
9. **Transformers:** סוג זה הוצג במאמר "תשומת לב היא כל מה שאית/ה צריך/ה"<sup>[4]</sup>. הוא מבוסס על מנגנון הקשב העצמי, נמצא היום בחזית המחקר ומשמש משימות עיבוד שפה טבעית כגון תרגום שפה, סיכום טקסט, מודול שפה (BERT,GPT-3), ניתוח סנטימנטים (מבוצע בין היתר בזיהוי תוכן פוגעני/אובדן/מסית ברשות החברתיות הגדולות), יצירת דיאלוג, ביצוע תמונה בזיהוי תוכן (Image Captioning), זיהוי דיבור ועוד.

\* Vanishing Gradient – בעיה השיפוע הנעלם מתרחשת בעת אימון רשותות עמוקות, שבהן שיפועי המשקלות ביחס לשגיאה הופכים קטנים מאוד. זה יכול לקרות כאשר פונקציית ההפעלה ( Activation function ) המשמשת בראשת אינה מתאימה היטב לרשותות נירונים עמוקות בגין RNN או כאשר ערכי המשקלות לא מאותחלות כמו שצරיך. בעיה זו עלולה לגרום לתהילך האימון להיות איטי מאוד, מכיוון שהשיפועים של המשקלות ביחס לשגיאה קטנים מדי לביצוע עדכונים משמעותיים למשקלות. השיפועים יכולים להיות כל כך קטנים עד כדי שהם נעלים וכך בעיה זו נקראת "הSHIPוע הנעלם". בעיה זו חמורה יותר ב- RNN ו- FNN כי השיפועים צריכים לעבור הפצה (propagation) דרך שכבות רבות. ניתן לטפל בעיה זו ע"י שימוש בפונקציות הפעלה שמתאימות יותר לרשותות אלו בגין ReLU או ע"י שימוש בטכניקות כגון Batch Normalization.

בסמינר זה נתעסק בכל הקשור ללמידה عمוקה "Deep Learning". תחום זה הוא תת-תחום ספציפי של למידת מכונה. מדובר בתפיסה חדשה של ייצוג למידה מנתונים ששמה דגש על לימוד שכבות עוקבות של ייצוגים בעלי משמעויות שהולכת וגוברת. המילה "עומקה" אינה מתייחסת לסוג של הבנה عمוקה יותר אלא מייצגת את הרעיון של שכבות עוקבות של ייצוגי נתונים. עומק המודל מייצג את כמות השכבות שתורמות למודל הנתונים.

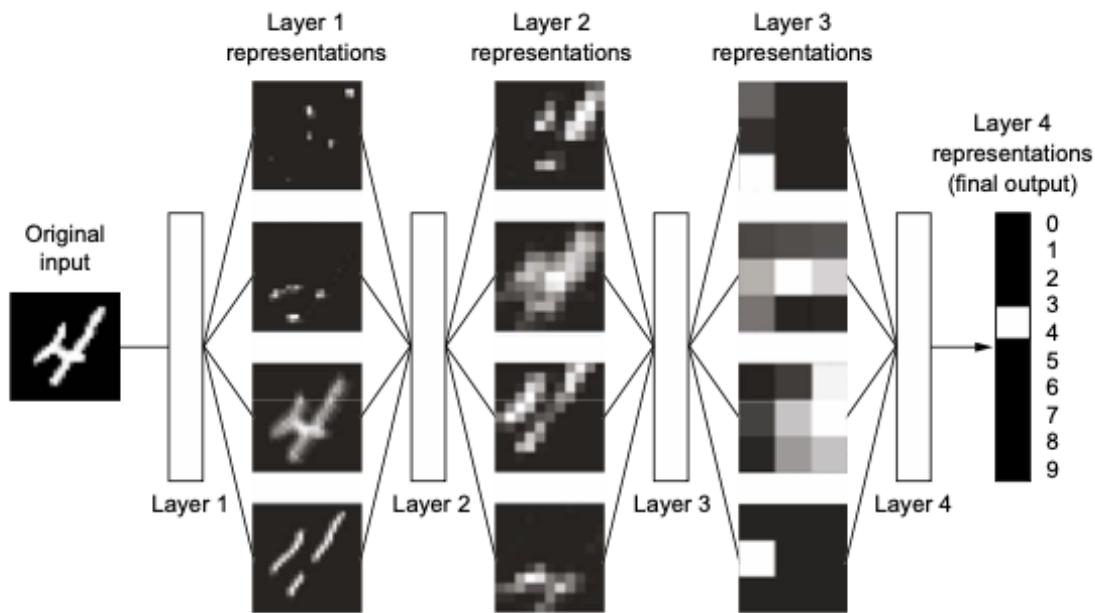
קיימים מספר שמות נוספים בתחום בגין:  
 ."Layered representations learning", "Hierarchical representations learning"  
 בגישה המודרנית קיימים עשרות או אפילו אלף שכבות עוקבות של ייצוגים וכולם נלמדים אוטומטית מחשיפה לנוטוני אימון. גישות אחרות של למידת מכונה נוטות להתמקד בלמידה שכבה אחת או שתיים בלבד של ייצוגים של הנתונים והן נקראות "shallow learning".

בלמידה عمוקה, ייצוג השכבות הללו נלמדים באמצעות מודלים שנקראים רשותות נירונים מלאכותיות (Artificial Neural Networks – ANNs) שモבנות זו על זו. המונח "רשתות נירונים מלאכותיות" מתייחס לננוירוביולוגיה, ולמרות שחלק מהמודלים המרכזיים בלמידה عمוקה פותחו בחלוקת ע"י קבלת השראה מההבנה שלנו של המוח, מודלים של למידה عمוקה אינם מודלים של המוח. אין הוכחה לכך שהמוח מישם משהו כמו מגנוני הלמידה שאנחנו משתמשים בלמידה عمוקה ובום, ניתן להתרשם מלא מעט מאמרם שמצינים של למידה عمוקה פועלת כמו המוח, ונוטה לגרום לנו לחשב כי התהום קשור לננוירוביולוגיה. בסופו של יום, במידה عمוקה היא מסגרת מתמטית למידת ייצוגים נתונים. איך נראים הייצוגים שנלמדו ע"י אלגוריתם ללמידה عمוקה? נציג זאת ע"י הסתכלות על דוגמא של רשת בעלת 4 שכבות שמטרתה היא לזהות את הספרה שבתמונה.<sup>[6]</sup>



איור 1 - רשת למידה عمוקה המשוגגת ספרות [6]

הרשות הופכת את התמונה הספרטית לייצוג שונים מהתמונה המקורי בצורה כזו שהמידע עבר מסטר "פילטרים" ויצא נקי.



איור 2 - יציג מודיע שמלמד ע"י מודל סיוג ספרות [6]

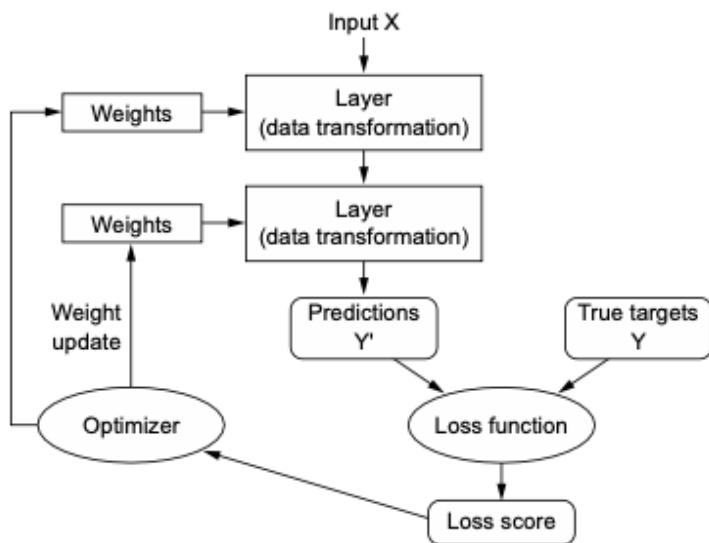
לסיבום, מבחינה טכנית: למידה عمוקה היא דרך רב-שלבית ללמידה ייצוג של נתונים. זה רעיון פשוט שמורכב ממנגנונים פשוטים ובקנה מידה גדול יחסית, יכול להראות כמו קسم.

בשלב זה, אנו יודעים שלמידת מכונה עוסקת במיפוי קלטיים (בגון תמונות) למטרות (בגון סיוג "חтол") דבר שנעשה על ידי התבוננות בדוגמאות רבות של קלט ומטרות. אנו יודעים שרשתות עצביות עמוקות

עשות את המיפוי של קלט למטרה באמצעות רצף عمוק של טרנספורמציות נתונים פשוטות (שכבות) וטרנספורמציות נתונים אלו נלמדות על ידי חסיפה לדוגמאות.

בעת נדבר על כיצד הלמידה זה מתרחשת, באופן קונקרטי. המפרט של מה ששבבה עושה לנוטנו הקלט שלו טמון במשקלים של השכבה, שבמהותם הם חבורת מספרים. במקרים טכניים, היינו אומרים שהטרנספורמציה המיושמת על ידי שכבה מותאמת לפי המשקלות שלה. (משקלות נקבעות לעיתים גם הפרמטרים של שכבה). בהקשר שלנו, למידה פירושה מציאת קבוצת ערכים עבור המשקלים של כל השכבות בראשת, כך שהרשת תפפה באופן נכון קלטים למטרות המשויות אליו. אבל זה העניין: רשות עצבית עמוקה יכולה לבצע עשרה מיליון ואף יותר פרמטרים. מציאת הערכים הנכונים עבור כלום עשויה להיראות ממשימה מרתקעה, במיוחד בתחום בו כרך שניינו הערך של פרמטר אחד ישפייע על ההתנהגות של כל האחרים!

כדי לשולט על הפלט של רשות ניורונים, אנו צריכים להיות מסוגלים למדוד עד כמה הפלט הזה רחוק ממה שציפינו. זהו התקן של פונקציית ההפסד של הרשות, הנקראת לעיתים גם objective function / cost function. פונקציית ההפסד לוקחת את התוצאות של הרשות והמטרה האמיתית (מה שאנו רצים שהרשת תוציא) ומחשבת ציון מרחק (distance score), ולובכת עד כמה הרשות הצלחה בדוגמה הספציפית זו. הטריק הבסיסי בלמידה עמוקה הוא להשתמש בניקוד זה כמשוב כדי שנוכל להתאים מעט את ערך המשקלות, בכיוון שוריד את ציון ההפסד עבור הדוגמה הנוכחיות. התאמה זו היא תפקido של ה- Optimizer, שמיישם את מה שנקרה אלגוריתם ההפצה לאחר (Backpropagation) [6] - האלגוריתם המרכזי בלמידה עמוקה.



איור 3 – רשות ניורונים עמוקה, השימוש בפונקציית ההפסד ושינוי ערך המשקלות בעזרת ה- Optimizer [6]

אלגוריתם ההתפשטות לאחר ( Backpropagation ) הוא תהליך המשמש לעדכון המשקלים של הרשת במהלך האימון. ה-optimizer מיישם אותו כדי למצוור את השגיאה בין הפלט החזו ( Prediction Y ) לבין הפלט האמיתי ( True targets Y ). הוא פועל על ידי חישוב שיפוע השגיאה ביחס לכל משקל ברשת. לאחר מכן, נעשה שימוש בשיפוע כדי לעדכן את המשקלות בהתאם להفور לשיפוע, וכך שהשגיאה תழער. תהליך זה חוזר על עצמו במהלך האיטרציות שנקבעו מראש שידועות כ-epochs, עד שהשגיאה מתחת לסף מסויים או שהביצועים של הרשת בתנאי האימון מפסיקים להשתפר. הוא מתייחל בחישוב השגיאה בין הפלט החזו לבין הפלט האמיתי של הרשת. לאחר מכן, הוא מפיץ את השגיאה לאחר דרך הרשת, החל משכבות הפלט ועובד אחורה לשכבות הקלט. בכלל שכבה, השגיאה משמשת לחישוב שיפוע השגיאה ביחס למשקלים של אותה שכבה. ברגע שהמשקלים מחושבים עבור כל השכבות, המשקלות מתעדכנות באמצעות האופטימיזר כגון Gradient Descent / ADAM.

## Deep Learning for Time Series 2.3

### 2.3.1 סקירת השימושים השונים של ניתוח סדרות זמן

סדרת זמן יכולה להיות לייצג נתונים המתקבלים באמצעות מדידות באינטראולים קבועים, כמו מחיר שער מט"ח, צרכית חשמל شمال שעתית, מכירות של חנות ועוד. סדרות זמן נמצאות כמעט בכל תחום, בין אם אנו משתמשים על תופעות טבע, מגז אויר, דפוסי פעילות של משתמשים באתר אינטרנט, עסקאות אשראי. תחום זה, בשונה מתחומים אחרים, דורש הבנה של טרנדים, נתונים מחזוריים ושינויים פתאומיים ויחד עם זה, ללא ספק שהמשמעות הנפוצה ביותר שקשורה לסדרות זמן תהיה חיזוי. מהו יקרה בעתיד? מהו שער המט"ח הבא? מה תהיה צרכית החשמל מספר שעות מראש על מנת שנוכל לצפות את הביקוש? מהו מג האויר כמה ימים מראש כדי שנוכל לתכנן את הטיול? ובקרה – **תחזית**. כמו כן, ישנו לא מעט דברים אחרים שניתן לעשות עם סדרות זמן כגון:

- סיווג – בהינתן סדרת זמן של מבקר באתר, סוג האם המבקר הוא בוט / אדם.
- זיהוי אירעים – זיהוי אירע ספציפי בסדרת זמן. מתבצע ע"י זיהוי "keyword" כגון "Hey Siri/Alexa/Google".
- זיהוי אנומליות – זיהוי דברים חריגים בסדרת זמן. בגין עסקאות אשראי / פעולות חריגות בחשבון / פעולות חריגה בראשת הארגונית או אפילו זיהוי דבר חריג בתהילה ייצור מוצר. חלק זה מתבצע באמצעות למידה לא-מודרבת מאוחר ואנו לא יודען את האנומליה שאויטה לנו מחפשים וכן אן אפשר לבצע אימון על אנומליות ספציפיות.
- תחבורה – חיזוי דפוסי תנעה, ליעל נתיבי תחבורה וניטות
- שירותים בריאות – ניבוי התפרצויות מחלות, ניטור סימנים חיוניים וייעול אספקת שירותי בריאות
- קלאלות – ניבוי יבול, ייעול מערכות השקיה וניטור בריאות היבול

אלן רק מספר דוגמאות, קיימות טכניקות רבות של למידה عمוקה הנמצאות בשימוש גם בתחוםים רבים אחרים והתחום מתפתח במהירות. עם עלייה בזמנים של נתונים וכוח חישוב, השימוש יעלה.

# 3 חיזוי מחירי מט"ח באמצעות LSTM

## 3.1 מבוא

חיזוי מחירי מט"ח (מטבע חזע) היא משימה מתוגרת בגל התנודתיות הגבוהה והאי-לינאריות של השוק. בשנים האחרונות, נעשה שימוש רב בטכניקות למידה عمוקה עבור חיזוי מחירי מט"ח בשל יכולת למדל קשרים מורכבים ולא לינאריים בנתונים. אחת הארכיטקטורות הנפוצות ביותר עבור חיזוי מחירי מט"ח היא LSTM.

רשתות (Recurrent Neural Network) RNN שתוכננו LSTM הם סוג של (Long Short-Term Memory) RNN שתוכננו במיוחד לטיפול בסדרות זמן. הן מבצעות זאת ע"י כך שהם מאפשרים לרשף לשומר על מצב זיכרון פנימי שנិיתן לעדכן לאורך זמן. מצב פנימי זה יכול לשומר תליות רכבות טווח בנתונים, דבר שהופך את ה-LSTMs למתאימים עבור חיזוי מחירי מט"ח.

בסמינר זה, אנו מצאים שימוש ב-LSTM כדי לחזות את המחיר העתידי של צמד מטבעות ספציפי, כגון USD/EUR או YJP/USD. תחילת נאוסף נתונים היסטוריים של צמד המטבעות ונכון אותם מראש (pre processing) כדי שייתאימו למודל. לאחר מכן נאמת ונבדוק את המודל באמצעות הנתונים שהכנו ונעריך את הביצועים שלו על ידי השוואת המחיר שהמודל חזה מול המחיר הנוכחי. בנוסף, נחקור גם את השימוש בארכיטקטורות שונות כגון LSTM - stacked bidirectional LSTM וטכניקות שונות כגון, נורמליזציה של נתונים (data normalization) ושינוי קנה מידת (feature scaling) לשיפור הביצועים של המודל.

## 3.2 משתנים בעלי השפעה על חיזוי שערי מט"ח

ישנם מספר משתנים שיכולים להשפיע על התוצאות של חיזוי המט"ח בبنיהם:

### 1. נתונים היסטוריים:

ಐוכות, זמינות, נקיות, דיקוק וכמות הנתונים ההיסטוריים המשמשים לאימון ובדיקה המודל יכולה להיות השפעה משמעותית על הביצועים. המטרה היא לקבל מידע כמה שיותר נקי ומדויק כדי למנוע הטויות (biases) בלשון מודול ובכך לחתוף בצורה טובה יותר את הדפוסים והמגמות בשוק.

### 2. עיבוד מקדים של הנתונים (Data preprocessing):

בהתיליך בبنית מודל, הוא בעל השפעה גדולה על הביצועים של המודל בגין:

- טיפול בנתונים חסרים: הסרת השורות המכילות ערכים חסרים / טיפול בערכים אלו בשיטה כלשהיא.
- ניקוי הנתונים: הסרת חריגים וחוסר עקביות מהנתונים.
- נורמליזציה: המרת הנתונים לסקלה שבין 0-1 ע"י שימוש ב $\text{ax-min} / \text{ax-max}$ .
- פיצול הנתונים: חלוקה לקבוצות אימון / מבחן.

על ידי עיבוד מקדים של הנתונים אנו יכולים להפוך את המודל לחזק יותר בעל יכולת התחדשות גבוהה יותר עם השוק. בנוסף, הדבר עשוי לעזור בהפחיתת הוצאות החישובית והגדלת הביצועים של המודל.

### 3. הנדסת תכונות (Feature engineering):

התהילך של הפיכת נתונים גולמיים לתכונות שיב轮廓 לשמש כקלט למודל. הנדסת תכונות היא שלב חשוב בתהילך בנית המודל ויכולת להיות לה השפעה משמעותית על ביצועיו. בהקשר שלנו, היא יכולה לכלול יצירת תכונות חדשות המבוססות על נתונים היסטוריים, כגון:

- אינדיקטורים טכניים: ממצאים נאים, מדד חזק יחסית (RSI) ורכזות בולינגר.
- מדדי טרנדים: ממנטום וקצב שינוי (ROC)
- עונתיות: דקה, שעה, יום, שבוע, חודש

כל אחת מתכונות אלו יכולה לעזור בהבנת דפוסים ומוגמות חשובות בשוק שאולי לא נראים בקלות בנתונים המקוריים. בין היתר נוכל לומר שהנדסת תכונות יכולה לבחור תת-קבוצה של התכונות המקוריות, על ידי שימוש בטכניקות כגון בחירת תכונות או הפקחת ממד, מה שיכל לעזור בהפחיתת מורכבות המודל והגדלת יכולת הפרשנות.

### 4. ארכיטקטורת הרשת / מודל (Model / Network architecture):

הבחירה בין LSTM, מספר השכבות, מספר הנוירונים, סוג פונקציית הפעלה, שימוש ב-  
stacked LSTM או LSTM bidirectional יכולה גם היא להשפיע לשיפור את ביצועי המודל.

### 5. כיוון הiperפרמטרים (Hyperparameters tuning):

התאמת ערכי פרמטרים מסוימים במודל כדי לעל את הביצועים שלו. הפרמטרים האלה לא נלמדים במהלך האימון, אלא נקבעים על ידי המשtamש. בינהו:

- מספר השכבות במודל.
- מספר הנוירונים בכל שכבה של המודל.
- קצת הלמידה: הקצב שבו המודל מעבדן את המשקولات במהלך האימון.
- גודל אצווה (batch size): מספר הדגימות בשימוש באיטרציה אחת אל אימון.
- מספר epochs: מספר הפעמים שהמודל מאומן על כל סט הנתונים.
- קצת הנשירה (Dropout rate): הקצב שבו נוירונים נושרים במהלך האימון כדי למנוע התאמת יתר.
- פונקציית הפעלה: הפונקציה המשמשת להציג א-lienאריות של המודל.

לפרמטרים האלה יכולה להיות השפעה גדולה על ביצועי המודל, ומציאת הערכים האופטימליים עבורם דורשת ניסוי וטעייה. ישנן טכניקות שונות כגון אלגוריתם גנטית, חיפוש רשת, חיפוש אקראי ואופטימיזציה

ביסיאנית בהן ניתן להיעזר כדי לבזין את הפרמטרים. על ידי ביצוען הפרמטרים, ניתן להפוך את המודל למדוק וחזק יותר, בעל יכולת התחווות עם המורכבות והתנודתיות של השוק.

## 6. גורמים חיצוניים:

אירועים שקיימים מחוץ לשוק עצמו יכולים להשפיע על השוק ועל מחירי המטבעות. אירועים אלה יכולים להיות:

- כלכליים: שינוי ריבית, דוחות תל"ג, אינפלציה ושיעור אבטלה.
- פוליטיים: בחירות, שינוי מדיניות ממשלתית ויחסים בינלאומיים.
- טבע: הוריקנים, רעידות אדמה ואירועי טבע אחרים.

כל אירוע הללו יכולה להיות השפעה משמעותית על השוק ועל מחירי המטבעות, והם לא תמיד ניתנים לחיזוי. לדוגמה, שינוי ריבית בלתי צפוי או תוצאות בחירות בלתי צפויות יכולה לגרום לשינוי פתאומי בשוק, מה שעלול להפוך את התחזיות של המודל לפחות אמינות. לכן, חשוב לקח בחשבון את הגורמים החיצוניים הללו בעת בניית המודל והערכת ביצועיו. אחת הדרכים לעשות זאת היא על ידי הכללת חדשות ונימוח סנטימנטים במודל, שיחד יכולים לעזור לתפום / "להבין" את ההשפעה של גורמים חיצוניים על השוק. בנוסף, חשוב שתהיה לנו במידות של נתונים איקוטיים המכbsים תנאי שוק שונים, לרבות תקופות של השפעה על גורמים חיצוניים, וזאת כדי שנוכל להבшир ולבדוק את המודל בצורה יעילה.

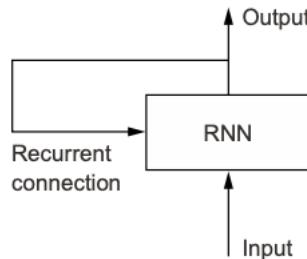
7. **תנודתיות (Volatility)**: התנודתיות של צמד מטבעות מתייחסת למידת השונות של שער החליפון של צמד מטבעות לאורך זמן. צמד מטבעות עם תנודתיות גבוהה יהיה מגבר רחב של תנודות בשער החליפון שלו, בעוד שצלAMD מטבעות עם תנודתיות נמוכה, יריה שער החליפון יציב יותר. במקרה שלנו, תנודתיות היא משתנה חשוב לשער חישוב מכיוון שהיא יכולה להשפיע על ביצוע המודל. אם המודל אינו חזק מספק כדי להתמודד עם תנודתיות גבוהה, יתכן שהוא לא יוכל לחזות במדויק שינויים פתאומיים בשוק וזה עלול להוביל לביצועים גרועים יותר של המודל והתחזיות לא אמינות. לכן, המודל צריך להיות מתוכנן להיות חזק מספק כדי להתמודד עם תנודתיות גבוהה ושינויים פתאומיים בשוק. ניתן להשיג זאת על ידי שימוש בטכניקות כגון נורמליזציה, שינוי קנה מידה וביצוען של הפרמטרים.

### 3.3 חיזוי מחרי מט"ח באמצעות רשת LSTM

#### 3.3.1 מבוא

מאפיין עיקרי של רוב הארכיטקטורות, הוא שאין להן זיכרון. כל קלט שמוצג להן, מעובד באופן עצמאי ללא מצב שנשמר בין הקלטים. כדי לעבד סדרת זמן עם רשותות אבלו (רשותות שאינן מוחזריות), נוצר רציף לרשות את כל הרצף של הסדרה בביטחון אחת – דבר שיحقق את הסדרה לבעל נקודת אחת. כפי שאנו מכירים, רשת שעובדת בצורה כזו היא תקרה feedforward neural network. לעומת זאת, ב-RNN החיבור בין נירונים הוא לא רק feedforward (הזנה קדימה), אלא קיימת הזנה חוזרת (feedback connection) בין נירונים. חיבורים אלו מאפשרים למידע לזרום ברשות לאורך זמן. לדוגמה: כאשר אנו קוראים את המשפט הזה, אנו קוראים ומעבדים מילה אחר מילה וכך שמירה על זיכרון מיילים שקראנו קודם וע"י כך, הדבר מאפשר לנו ייצוג של המשמעות שהמשפט מעביר.

הweeneyון המרכזי של RNNs הוא שהם יכולים לשמור על state, או זיכרון, שנייתן לעדכן ולהעביר משלב זמן אחד לאחר. ה-state מיוצג בדרך כלל על ידי וקטור הפעלה שהוא רגילים לקרוא לו המצב הנסתור (hidden state). בכל שלב זמן, הקלט הנוכחי וה המצב הנסתור הקודם משולבים כדי ליצור מצב חדש נסתור עדכני, שיימש בהמשך להפקת הפלט של הרשת. ל-RNN קיימת "בעיית השיפוע הנעלם" עליה דיברנו בפרק 2, LSTM שהוא וריאציה של RNN מתמודד על בעיה זו בעזרת שימוש בשערים (gates) עליו נדבר

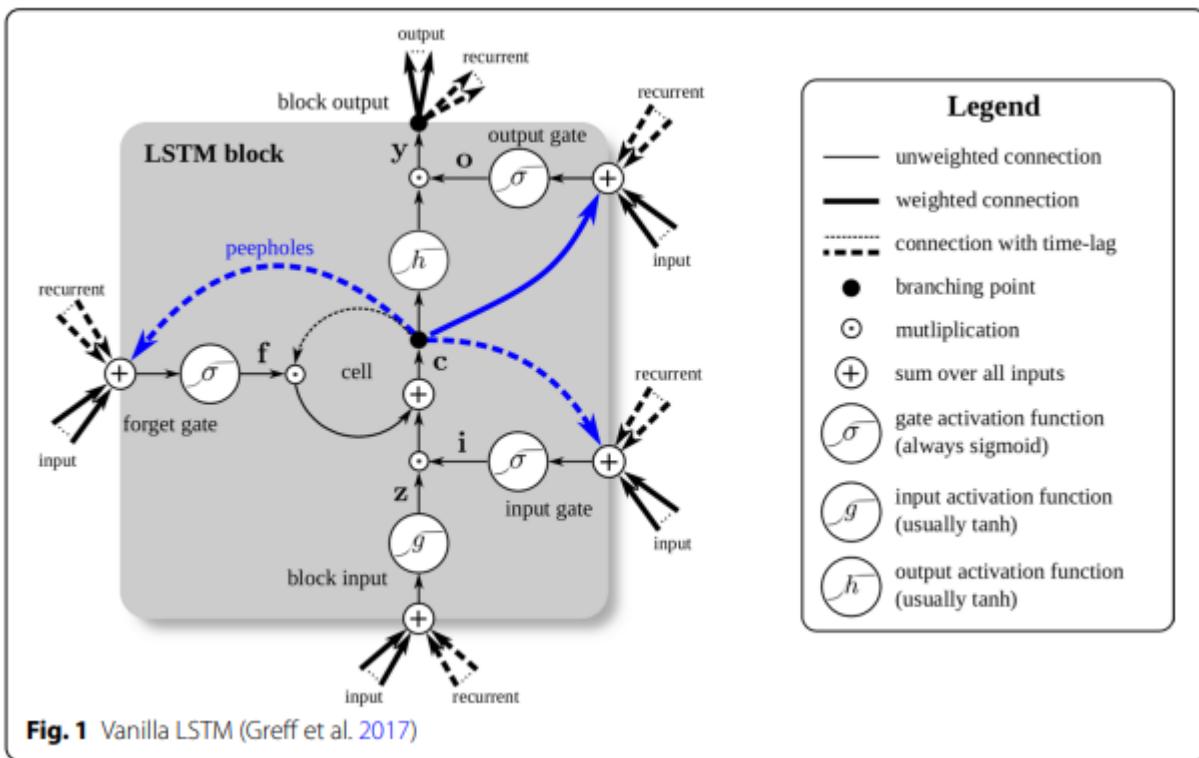


בסעיף הבא.

איור 4 - רשת RNN בעליות לולאה פנימית

### 3.3.2 הגדרת המודל

LSTM מורכב ממספר שערים, הוא כולל שער קלט, שער פלט ושער שכיחה. שערים אלו שולטים בזרימת המידע אל התא וממנו. שערים אלו מיושמים ע"י רשותות נירונים בעלות פונקציית אקטיבציה מסוג sigmoid / tanh אשר מייצרות ערך בין 0 ל-1 עברו כל אלמנט בקלט.



בלוק LSTM מסורתי. [1]

ניתן לסווג את המשקלים של רשת LSTM באופן הבא:

משקל קלט:  $W_z, W_i, W_f, W_o \in \mathbb{R}^{N \times M}$

משקלים חוזרים:  $R_z, R_i, R_f, R_o \in \mathbb{R}^{N \times N}$

משקל חור הצצה (המשקלים מהתאים הפנימיים של לשערי הכפל):  $p_i, p_f, p_o \in R^N$

משקל הטייה:  $b_z, b_i, b_f, b_o \in \mathbb{R}^N$

בasher  $z$  הוא בлок הקלט,  $i$  הוא שער הקלט,  $f$  הוא שער השכחה,  $o$  הוא שער הפלט,  $N$  הוא מספר בлокי LSTM ו- $M$  הוא מספר הקלטים. על ידי הצגה של  $x^t$  כווקטור קלט,  $y^t$  כווקטור פלט,  $c^t$  בתא בזמן  $t$ , חישוב ה-forward pass של המודל המוצע [1] תהיה באופן הבא:

אופן החישובים של השיערים מתרבעץ באופן הבא:

$$\bar{z}^t = W_z x^t + R_z y^{t-1} + b_z$$

$$z^t = g(\bar{z}^t)$$

$$\bar{i}^t = W_i x^t + R_i y^{t-1} + p_i \odot c^{t-1} + b_i$$

$$i^t = \sigma(\bar{i}^t)$$

$$\bar{f}^t = W_f x^t + R_f y^{t-1} + p_f \odot c^{t-1} + b_f$$

$$f^t = \sigma(\bar{f}^t)$$

$$c^t = z_t \odot i^t + c^{t-1} \odot f^t$$

$$\bar{o}^t = W_o x^t + R_o y^{t-1} + p_0 \odot c^t + b_o$$

$$o^t = \sigma(\bar{o}^t)$$

$$y^t = h(c^t) \odot o^t$$

בasher  $\sigma$  מציינת את פונקציית האקטיבציה הלוגיסטיבית (sigmoid),  $g$  ו- $h$  הן הפונקציות  $tanh$  ו- $\odot$  (מציאן את המכפלה בין שני וקטורים).

### **האינדיקטורים טכניים:**

אינדיקטורים טכניים הם חישובים מתמטיים המבוססים על המחיר, הנפק או הריבית הפתוחה של נייר ערך או סחורה. הם מספקים אותות (signals) לגבי מגמות בשוק ומשיעים לשוורדים לקבל החלטות מושכלות על ידי ניתוח נתוני השוק מהעבר. אינדיקטורים טכניים משמשים בניתוח טכני וניתן לישן אותם על גרפים של מנויות, סחרות, מטבעות ומניות פיננסיים אחרים. דוגמאות לאינדיקטורים טכניים כוללים ממוצעים נועים (MA), רצועות בולינגר, RSI, MACD ו-Stochastics. עלפי הfonקציונליות שלהם, ניתן לקבעם לשולש קטגוריות: מפרטים, מוביילים ותנודתיות.

אינדיקטורים בפיגור – הידועים גם כאינדיקטורים מגמה, עוקבים אחר פעולות העבר בינם:  
MA - Moving Average, Moving average convergence divergence (MACD)

אינדיקטורים מוביילים – הידועים גם כאינדיקטורים מבוססי מומנטום, מטרתם לחזות שינוי מגמת ROC – Rate – RSI-Relative Strength Index - Of Change.

מדדים מבוססי תנודתיות (volatility) – רמות התנודתיות במחיר. BB – Bollinger Bands – הוא המציג הנפוץ ביותר המבוסס על תנודתיות.

אינדיקטורים אלה אינם מהווים ערבוט לביצועים עתידיים, אך הם יכולים להיות שימושיים בסיווג לסוחרים ליהות הזדמנויות קניה או מכירה פוטנציאלית.

האינדיקטורים הטכניים שמסוקרים בסימן זה לפי [1] מתוארים להלן:

### Moving Average (MA)

ממוצע נع הוא אינדיקטור שעוקב אחר מחיר מגמה (או פיגור) שמחליק מחירים לפי חישוב הממוצע שלהם בתקופת זמן שהוגדרה מראש. חישוב הממוצע שלהם מסיע בסיכון רעים ובכך ניתן להזות לא רק את כיוון המגמה, אלא גם לקבע נקודות תמייה והתנגדות פוטנציאליות.

$$MA = \frac{P}{N}, N = \text{number of days}, P = \sum \text{closed prices in the last } N \text{ days.}$$

### Moving Average Convergence Divergence (MACD)

הוא אינדיקטור טכני המשמש במסחר כדי לזהות מגמות במחירים השוק. הוא מחושב באופן הבא:

$$MACD = EMA[12] - EMA[26]$$

$$signal = EMA_{MACD}[9]$$

$$hist = MACD - signal$$

ערך ה-MACD מחושב על ידי הפחתת EMA26 (ממוצע נע מעריבי) מ-EMA12. (המספרים מציננים את מספר הימים עליהם הממוצע מחושב).

לקבלת אותן (signals) עברו קניה או מכירה, יש לחשב ממוצע נע מעריבי בעל ערך 9, על תוצאה MACD. בולו, באשר קו ה-MACD חוצה מעלהקו האות, הוא נחשב לאות קניה (bullish) פוטנציאלי ובאשר הוא חוצה מתחת לנקו האות, הוא נחשב לאות מכירה (bearish) פוטנציאלי. MACD בשילוב עם אינדיקטורים טכניים אחרים עוזר בניתוח וקבלת החלטות מסחר מושכלות.

### Rate Of Change (ROC)

אינדיקטור מומנטום המודד את אחוז השינוי במחיר המטיב על פני תקופה מסוימת. הוא מחושב באופן הבא:

$$\text{rate of change} = \frac{\text{close}_{\text{today}} - \text{close}_{N \text{ days ago}}}{\text{close}_{N \text{ days ago}}}$$

ROC מחושב על ידי חלוקת הפרש בין השער הנוכחי לבין מחיר השער לפני  $N$  ימים, חלקו ממחיר השער לפני  $N$  ימים. הערך שמתתקבל מבוטא באחוז ומשורטט בשורה על קו ישר בגרף. ROC עולה, מצביע על בה שמחיר המניה עולה בקצב מהיר מ-בעבר, בעוד ש-ROC יורד מציין על ההפוך. אנו משתמש ב-ROC בשילוב עם אינדיקטורים אחרים כעזר לניתוח וקבלת החלטות מסחר מושכלות.

## Momentum

אינדיקטור טכני המודד את קצב שינוי המחיר של המטבע. אינדיקטור זה עוזר בזיהוי כיוון תנועות המחיר ואת העוצמה של טרנד. חישוב זה מבוצע על ידי השוואה של המחיר הנוכחי למחיר ממספר ימים קודמים.

## Relative Strength Index (RSI)

אינדיקטור המודד את עוצמת פועלות המחיר של המטבע. מחושב על ידי השוואת הרוחים הממוצעים של מטבע לבין ההפסדים הממוצעים של אותו מטבע – בפרק זמן מוגדר. מדד RSI מתבטא בערך בין 0 ל-100, כאשר ערכים גבוהים מצביעים על קניית יתר או מכירת יתר של אותו מטבע.

## Bollinger Bands (BB)

רצועות בולינגר משמשות למידדת התנודתיות (volatility) בשוק ולספק הגדרה יחסית של מחירים גבוהים ונמוכים. הוא מורכב משלושה קווים בגרף: MA במרכז, אחד מעליו ואחד מתחתלו שבמרחך 2 סטיות תקין מערכו. תנודתיות עולה / יורדת -> המרחק בין הפס העליון לתחתון גדל / תקען בהתאם שערים שmaguiim לקו העליון, מעידים על קניית יתר בעוד שערים שmaguiim לקו התחתון מעידים על מכירת יתר.

## Commodity Channel Index (CCI)

משמש בעיקר ליזוי ממצבים פוטנציאליים של קניית יתר או מכירת יתר בשוק. ה-CCI מודד את הסטייה של המחיר הנוכחי מהממוצע הסטטיסטי שלו, כאשר ערכים גבוהים מצביעים על תנאי קניית יתר וערכים נמוכים מצביעים על תנאי מכירת יתר.

## סט הנתונים

כפי שהוזכר בפרק 1, שיעורי הריבית והאינפלציה הם שני אינדיקטורים בסיסיים לחזקה של כלכלת המקרה של ריבית נמוכה, אנשים נוטים לקנות כדי לחזק את הכלכלת ובמקורה ההפון, הכלכלת הופכת שברירית.[1] כיוון, גרמניה וארה"ב הן שתיהן מהכלכלה החזקות בעולם. בנסיבות אלה, קיים קשר חזק

**Table 1 Macroeconomic data and the currency pair used in the data set**

Term	Explanation
<i>Close(EURUSD)</i>	Daily close value of euro/dollar currency pair
<i>Open(EURUSD)</i>	Daily open value of euro/dollar currency pair
<i>High(EURUSD)</i>	Daily high value of euro/dollar currency pair
<i>Low(EURUSD)</i>	Daily low value of euro/dollar currency pair
<i>Inflation Rate<sub>EU</sub></i>	Monthly inflation rate for the European area
<i>Inflation Rate<sub>USA</sub></i>	Monthly inflation rate for the US area
<i>Interest Rate<sub>GER</sub></i>	Monthly interest rate in Germany
<i>Interest Rate<sub>EU</sub></i>	Monthly interest rate in the European area
<i>FED Funds Rate</i>	Daily Fed funds rates
<i>Close(S&amp;P 500)</i>	Standard and Poor—daily close American stock market index
<i>Close(DAX)</i>	Daily close German stock index

איור 5 – גורמים מאקרו-כלכליים בסט הנתונים [1]

בין שוק המניות לערך המטבע שלהם. ל-DAX – מzd המניות הגרמני יש קשר חזק על שער היוו בעוד של-S & P 500 הוא מzd מניות אמריקאי שימושי על הדולר. שיעורי הריבית של הבנקים המרכזיים הם גם גורמים חשובים הקובעים את מחירי המטבעות ולכן, שיעורי הריבית שנקבעים על ידי הבנק המרכזי של אירופה או על ידי ה-fed (שאחראים באלה"ב) משפיעים באופן ישיר על היוו והدولר בהתאם.

ב-[1] מבוצע מחקר כפי שנראה בהמשך, המראה את ההשפעה של גורמים מאקרו-כלכליים על הערך של הצמד USD/EUR. הגורמים הנלקיים בחישוב מתוארים בטבלה מטה.

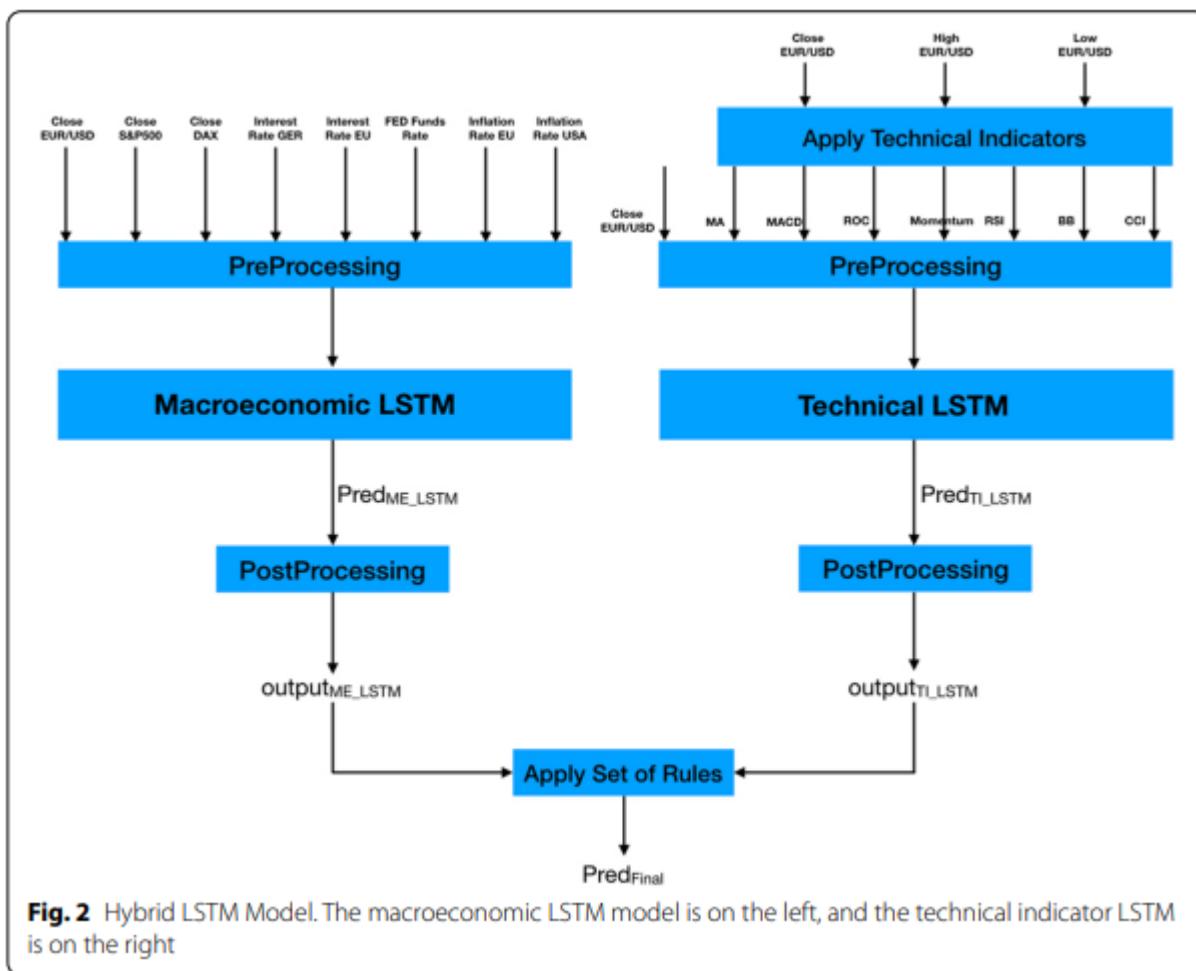
הנתונים שנלקחים בחשבון במסגרת מחקר זה כוללים שער פריחה וסירה של המטבעות, מדדים, אינפלציה, ריביות ונ נתונים כלכליים. הנתונים הם בטוחה התאריכים שבין 01/2013-01/2018.

## מודל LSTM היברידי המשתמש באינדיקטורים מאקרו-כלכליים ובאינדיקטורים טכניים

במודל המוצע, נבנה LSTM היברידי כדי לחזות את התנועה של הצמד EUR/USD המשתמש באינדיקטורים מאקרו-כלכליים ובאינדיקטורים טכניים. מודל זה מכיל שני LSTM תתי-מודלים שיימדו מנתונים שונים, קרי אחד לכל סוג, כפי שניתן להתרשם באIOR מטה.

המודל יעבוד באופן הבא:

1. עיבוד מקדים של הנתונים.
2. אימון ME-LSTM (macro-economic) ועיבוד הפלט שלו.
3. אימון TI-LSTM (Technical Indicators) ועיבוד הפלט שלו.
4. מימוש אסטרטגיות שונות על ידי שימוש בתוצאות של שני-LSTMs שלנו.



איור 6 - מודל LSTM היברידי המשולב משני LSTM בתתי-מודלים, אחד לכל סוג אינדיקטור [1]

כל אחד מתתי-המודלים נבדק באופן עצמאי ושילוב תוצאותיהם יחד גורר מודל LSTM .

### **מבנה מודל ME-LSTM**

מודל זה נועד בכדי לבדוק את ההשפעות המאקרו-כלכליות על מחירי השערים של הצמד USD/EUR. הדברים שנלקחו בחשבון, עליהם הסבירנו לעיל הם:

- ריביות של גרמניה והאיחוד האירופאי
- ריביות צפויות של ה-FED
- שער אינפלציה ב-EU וב-US
- שער סגירה של מדד S&P 500
- שער סגירה של מדד DAX

לאחר העיבוד המקדים, מודל זה התאמן על הцמד תוך כדי התחשב במשתנים לעיל על שער הסגירה של USD/EUR.

### **מבנה מודל TE-LSTM**

מודל זה נבנה בכדי לבדוק את ההשפעה של אינדיקטורים טכניים על שינוי המחיר של הцמד EUR/USD. האינדיקטורים הטכניים בהם בוצע שימוש הם:

- MA עם מחזור של 10
- MACD עם מחזור קצר של 12 וארוך של 26
- ROC עם מחזור של 2
- Momentum עם מחזור של 4
- RSI עם מחזור של 10
- BB עם מחזור של 20
- CCI עם מחזור של 20

לאחר העיבוד המקדים, מודל זה התאמן על הцמד תוך כדי התחשב באינדיקטורים לעיל על שער הסגירה של USD/EUR.

המודל המוצע ב[1] אינו מודל משולב של שני תת-המודלים יחד. השילוב יבוצע לפי מנגנון החלטה מבוסס בלאים שיפעל בעיבוד מאוחר.

## תהליכי האימון והסיווג

בידוע, האופן שבו המודל מאומן, שלב ההכנה המוקדם, ניקוי הנתונים, הגדרת הפרמטרים, מספר האיטרציות, פונקציות הפעלה, במתה השכבות והיפרפרמטרים מיציה של פרמטרים ועוד, הינם דברים שתלויים בחוקר. במקרה שלנו, קיימות 3 מחלקות סיווג והחלוקת להן תהיה עפ"י חציית סף מסוים שיוגדר מראש. הסיבה שהחוקרים ב-[1] חילקו מגעה מתוך מקום של הבנת הממצב וחלוקת לשולש פעולות אותן החוקר/הסוכן (עליו אפרט בסיכון) יוכל לבצע.

שלושת מחלקות אלו הן:

Class\_inc : המתייחסת לעלייה במחיר שמעל הסף.

Class\_dec : המתייחסת לירידה במחיר שמעל הסף – יכולה לשמש למיכוחות.

Class\_noact : המתייחסת לשינוי במחיר ש מתחת לסת – ככלمر שינויים לא משמעותיים.

האלגוריתם שהוצע לחישוב היסטוגרמה בפי שמצין מטה, נועד עבור המחלקה השלישית ובכך עוזר במניעה של בדיקת כל האפשרויות ובכך מצמצם את מרחב החיפוש. הרעיון הוא לקבוע את הגבול העליון של הסף בהתבסס על כיסוי של 85% מכל השינויים וכדי להשיג זאת, מתבצע ניתוח היסטוגrama על שער סגורה של הצמד USD/EUR כדי להחליט את התפלגות שינוי המחיר בימים עוקבים ע"י חלוקה למחלקות (bins), סידורו בסדר יורד וסכימה שלhn עד חציית סף של כ-85% או עד מספר המחלקות.

---

### Algorithm 1 Histogram Analysis.

```
1: procedure HISTOGRAM_ANALYSIS(close_diff[ ],number_of_bins)           ▷ close(t) - close(t - k)
2:   bin_counts[ ],bin_max_diff[ ]  $\leftarrow$  perform_histogram(close_diff[ ],number_of_bins)
3:    $\triangleright$  bin_counts, bin_max_diff are reverse sorted arrays according to bin_counts values
4:   temp_sum  $\leftarrow$  0
5:   sum_bin_counts  $\leftarrow$  sum(bin_counts[ ])
6:   i  $\leftarrow$  1
7:   while i  $\leq$  number_of_bins do
8:     temp_sum  $\leftarrow$  temp_sum + bin_counts[i]
9:     if temp_sum/sum_bin_counts  $>$  0.85 then
10:      break
11:      i  $\leftarrow$  i + 1
12:   threshold_upper_bound  $\leftarrow$  bin_max_diff[i]
13:   return threshold_upper_bound
```

---

איור 7 - ניתוח היסטוגרמה וקביעת ערך סף[1]

ערך הספַּף נקבע על פי מzd אנטרופיה. אנטרופיה נחשבת מzd לai-ודאות על קבוצת מצבים אפשרית. אנטרופיה ניתנת להגדרה על משטנה בדיז או רציף, ואני נמנע מלפרט עליה במסגרת סמיינר זה.

היא מחושבת באופן הבא:

$$Entropy(x) = \sum -p(x) \log_2 p(x)$$

בכדי לקבל התפלגות מואזנת, אנו מחשבים את האנטרופיה של התפלגות המחלקה (*bin*) באופן איטרטיבי לכל ערך ספַּף עד לקבלת הפרש מקסימלי. אבל כמו שציינו, אנו מחשבים את ספַּף זה באלגוריתם 1. מאחר ואנו מגבלים את כמות האיטרציות לפי ספַּף של 85% (threshold), נמצא ספַּף סופי, כך שים קסם את ערך האנטרופיה. האופן שבו נבצע זאת מפורט באלגוריתם 2 הנועד לחישוב ספַּף זה, כפי שנitin להתרשם מטה.

---

### Algorithm 2 Threshold Calculation.

---

```

1: procedure CALCULATE_THRESHOLD(close_diff[ ])
2:   threshold_upper_bound  $\leftarrow HISTOGRAM\_ANALYSIS(close\_diff[ ], 10)$  ▷ close(t) - close(t - k)
3:   temp_threshold  $\leftarrow 0$ 
4:   best_entropy  $= -\infty$ 
5:   while temp_threshold  $< threshold\_upper\_bound do
6:     labels  $\leftarrow \{0\}$ 
7:     indexes_incr  $\leftarrow indexOf(close\_diff[ ] > temp\_threshold)$ 
8:     indexes_decr  $\leftarrow indexOf(-close\_diff[ ] > temp\_threshold)$ 
9:     labels{indexes_incr}  $\leftarrow 2$ 
10:    labels{indexes_decr}  $\leftarrow 1$ 
11:    entropy  $\leftarrow calculate\_entropy(labels[ ])$ 
12:    if entropy  $> best\_entropy then
13:      best_entropy  $\leftarrow entropy$ 
14:      threshold  $\leftarrow temp\_threshold$ 
15:    temp_threshold  $\leftarrow temp\_threshold + 0.00001$ 
16:   return threshold$$ 
```

---

איור 8 - חישוב ספַּף בעל ערך אנטרופיה מקסימלי [1]

אלגוריתם 2 בא לעזור במצטומם מרחב החיפוש ע"י מציאת *threshold* בעל ערך אנטרופיה מקסימלי. לולאת *while* משמשת למציאת ספַּף שבין 0 לבין *threshold\_upper\_bound* עם הוספה של 0.00001 ובכך מצמצמת מרחב החיפוש. לכל ערך ספַּף, מספר העליות והירידות של ערך הספַּף נקבעות (מסומנות כ-2 עבור עליות, -1- עבור ירידות ו-0 עבור אין שינוי) ולאחר מכן, ערך האנטרופיה עבור אותה התפלגות מחושבת. בסוף לולאת *while*, נבחרת התפלגות (הספר) בעלת ערך האנטרופיה המקסימלי והוא תשמש את אותן המחלקות עליהם דיברנו לעיל.

## שלב העיבוד המאוחר – (Postprocessing)

המטרה של כל העיבוד עליון דיברנו לעיל הייתה לדעת כיצד נוכל לסוג למחקות. החיזויים המוצעים מחת-הרשתות ME-LSTM ו-TI-LSTM-[1] שולבו באמצעות הכללים הבאים:

- אם המודל חזה את המחלקה `class_noact` אז שההחלטה תהיה `class_noact` בלבד נטראלי.
- אם שני המודלים חזו את אותם משתני מטרה (`labels`) אז שההחלטה תהיה לפי אותו משתנה מטרה.
- אם שני המודלים חזו דברים שונים, נבחר את המטרה בעלת ההסתברות הגבוהה ביותר. אם ההסתברות שווה, אז שנבחר את החיזוי של LSTM-TI.

גישה זו מאפשרת לצמצם את בנות העסקאות לטובות חיזויים מדויקים.

## מדד ביצועים

마חר ואנו מדברים על שני סוגי וישנים שני מודלים אותם אנו צריכים לבדוק, הוצע ממדד חדש לממדית השיטה המוצעת. ממדד זה מודד את רמת הדיווק הייחסית במספר העליות והירידות במשתני המטרה. ככלומר ניתן להסתכל על ממדד זה באופן הבא: היחס שבין מספר הערות הרווחיות לבין מספר העסקאות הכוללות. הערכים הבאים הם הערכים שמחושבים במדד הביצועים.

- `True_inc`: the number of true predictions increases
- `True_dec`: the number of true predictions decreases
- `False_dec_noact`: the number of predictions of the no-action class decreases
- `False_inc_noact`: the number of predictions of the no-action class increases
- `False_inc_dec`: the number of predictions of the decrease class increases
- `False_dec_inc`: the number of predictions of the increase class decreases

*profitAccuracy*

$$= \frac{TrueDec + TrueInc}{FalseDecNoact + FalseIncNoact + TrueDec + FalseIncDec + FalseDecInce + TrueInc}$$

### 3.3.3 השוואת מודל LSTM היברידי לבין מודל LSTM וריגל:

נשאלת השאלה, מהן היתרונות בשימוש במודל היברידי? מדוע הדבר כדאי? בקצרה - יותר מידע = יותר דיק.

בעת בעבר להשוואה:

#### 1. גיון של נתונים:

- **היברידי** – ע"י שילוב של היסטוריות מחירים יחד עם אינדיקטורים טכניים ואינדיקטורים מאקרו-אקונומיים, המודל מקבל תמונה נרחבת של השוק. זהו דבר הכרחי לאחר ומחירי שערים בשוק המט"ח מושפעים מגוון רחב של גורמים ביניהם: אירועים כלכליים, התפתחויות פוליטיות וסנטימנטניים בשוק.
- **ריגל** – מסתמך בעיקר על היסטוריות מחירים. פחות מידע לגבי תנאי השוק, דברים העשוי להשפיע על דיק החיזוי.

#### 2. חסינות:

- **היברידי** – הכללה של תכונות נוספות לתהفور את המודל לחסן וחזק יותר לשינויים בשוק, וזאת בגלל שהוא יכול לתפוס מגמות ויחסים שאינם ברורים מנתוני המחיר בלבד.
- **ריגל** – מכיוון שהוא מסתמך רק על מחירים היסטוריים, הוא עשוי להיות רגיש יותר לרעש השוק ופחות חזק לשינויים.

#### 3. דיק:

- **היברידי** – בדרך כלל, מודל היברידי יטו להיות בלי דיק חיזוי טובים יותר מכיוון שהם יכולים ללבוד אוסף עשיר יותר של מידע ותבניות שונות. הם מסוגלים להבין יותר את הגורמים שימושיים את השינויים במחירים.
- **ריגל** – עשוי שלא לדיק באותה מידע ובמיוחד במלת שכזו, בה מחירים מושפעים מאוד מגורמים מאקרו-כלכליים.

#### 4. חשיבות וסיבוכיות:

- **היברידי** – יותר מורכבת וıntנסיבית מבחינה חשיבותית בגלל התכונות הנוספות. הדבר דרש בחירת תכונות (*feature engineering*) מדויקת מאוד והמנון כיסויות וטעה. "זריקת" כל האינדיקטורים האפשריים לא ישפרו את ביצועי המודל ولكن נדרש בחירה קפדנית של כמה תכונות ספציפיות לאמון, היישרו את ביצועי המודל.
- **ריגל** – פשוט יותר ופחות אינטנסיבי מבחינה חשיבותית. עם זאת, דבר זה עשוי לבוא על חשבון של דיק.

#### 5. הבנה של המודל ויכולת פירוש:

- **היברידי** – למורת שהסביר ופירש של *LSTM* הינה מטלה מאוד קשה, ככלומר חסרות פרשנות, הגישה ההיברידית יכולה לספק קצת תובנות לגבי התכונות שמניעות את התחזיות, ועשויה להציג הבנה מסוימת של מכנייקת השוק.
- **ריגל** – מספק פחות תובנות לגבי הגורמים שתרומים לתחזיות מכיוון שהוא לוקח בחשבון רק את המחיר ההיסטורי.

### 3.3.4 תוצאות ומסקנות

הטבלה מטה מסכמת את ביצועי המודלים בחיזוי יומם אחד קדימה. ראשית נתחיל בסקירת ביצועי המודלים כל אחד בנפרד ולאחר מכן נסקור את ביצועי המודל היברידי.

#### תוצאות עבור המודל מאקרו-כלכלי (ME-LSTM)

**Table 4 ME\_LSTM model: one-day-ahead result summary**

	Profit_accuracy (%)	# of transactions
Iterations=50	46.50	243/243
Iterations=100	55.00	60/243
Iterations=150	48.97	243/243
Iterations=200	52.27	52/243
Average	50.69	149.50/243

Macro-Economic LSTM: One day ahead predictions [1]

כפי שניתן לראות, קיימת שונות גבוהה במספר העסקאות בעוד שנות נמוכה יותר ברמת הדיוק.

#### תוצאות עבור המודל הטכני (TI-LSTM)

**Table 5 TI\_LSTM model: one-day-ahead result summary**

	Profit_accuracy	# of transactions
Iterations=50	50.26	189/243
Iterations=100	50.93	161/243
Iterations=150	53.11	177/243
Iterations=200	54.43	94/243
Average	52.18	155.25/243

Technical LSTM: One day ahead predictions [1]

כפי שניתן לראות, השונות קטנה יותר גם במספר העסקאות וקטנה אפילו יותר ברמת הדיוק.

## תוצאות עבור המודל היברידי (ME-TI-LSTM)

**Table 7 Hybrid model: one-day-ahead predictions**

Iterations		Hybrid model-modification based on ME_LSTM		Hybrid model-modification based on TI_LSTM	
ME_LSTM	TI_LSTM	Profit_accuracy (%)	# of transactions	Profit_accuracy (%)	# of transactions
50	50	70.80	137/243	70.80	137/243
50	100	73.50	117/243	74.36	117/243
50	150	69.60	125/243	77.60	125/243
50	200	81.63	52/243	82.35	52/243
100	50	78.13	32/243	65.63	32/243
100	100	69.23	26/243	65.38	26/243
100	150	70.59	34/243	70.59	34/243
100	200	73.17	46/243	75.00	46/243
150	50	76.56	128/243	78.13	128/243
150	100	72.64	106/243	78.30	106/243
150	150	73.17	123/243	80.49	123/243
150	200	100.00	8/243	100.00	8/243
200	50	80.00	20/243	76.47	20/243
200	100	80.77	28/243	80.77	28/243
200	150	84.00	26/243	82.61	26/243
200	200	83.33	28/243	85.71	28/243
Average		77.32	64.75/243	77.76	64.75/243

מודל היברידי: תוצאות חיזויים אחד קדימה [1]

בטבלה לעיל מסכמת את ערכי רמת הדיק ואות מספר העסקות עבור על מקרה במודל זהה. כפי שניתן לראות, בחלק מהמקרים בהם מספר האיטרציות גדול (200), במרבית העסקאות צנחה באופן משמעותי. רמת הדיק הממוצעת עבור 16 מקרים עומדת על 77.32% למודל היברידי המבוסס על אינדיקטוריים טכניים ו-77.76% למודל היברידי מבוסס אינדיקטוריים טכניים

## **4 סיכום**

בஸמינר זה התחלתי בסקירה של המסחר בשוק ההון, תוך כדי דיון על הגורמים בעלי השפעה על המסחר, האופן שבו הוא מתבצע ודריכים שונות לביצוע מסחר. משם עברתי לפרק 2 שם סקרה את המושג בינה מלאכותית, מהי המשמעות מאחורי מילה זו, סוג רשותות וארכיטקטורות שונות, אופני למידה שונים, בעיות בתחום וכיוצא מתמודדים איתן. הסברתי את השימושים השונים של הארכיטקטורות והרחבתי על השימושים ברשותות מחזירות. בפרק 3, התחלתי עם חזרה קצרה של מהי רשות ניירונים מחזירית, ומשם עברתי להסביר על LSTM. הצגתני מחקר שבו מבוצע השימוש במודל LSTM היברידי לצורך חיזוי שערי מطق והציג את ביצועיו.

אני עושה סמינר זה לאחר קורס המבוא לבינה מלאכותית והקורס ברירת מחדע שם רוב הנושאים נלמדו. אני רוצה להודות לך (מיה) על קורס מושלם שגרם לי להחליט מה אני רוצה לעשות בחיים.

נתנאל

## Bibliography

- [1] Yıldırım, D.C., Toroslu, I.H. & Fiore, U. Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. *Financ Innov* 7, 1 (2021). <https://doi.org/10.1186/s40854-020-00220-2>
- [2] Hu Z, Zhao Y, Khushi M. A Survey of Forex and Stock Price Prediction Using Deep Learning. *Applied System Innovation*. 2021; 4(1):9. <https://doi.org/10.3390/asi4010009>
- [3] OTC foreign exchange turnover in April 2022 by Bank for International Settlements. [https://www.bis.org/statistics/rpfx22\\_fx.htm](https://www.bis.org/statistics/rpfx22_fx.htm)
- [4] Attention is All You Need, [arXiv:1706.03762](https://arxiv.org/abs/1706.03762)
- [5] Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering, [arXiv:1707.07998](https://arxiv.org/abs/1707.07998)
- [6] Chollet,F (2017). Deep Learning with Python (2<sup>nd</sup> ed.). Manning Publication