**מבוא**

שנת 2020 נפתחה עם מגפת הקורונה, שיצרה משבר עולמי בכל תחומי החיים, ובפרט בשוק ההון שהתרסק בצורה היסטורית וחסרת תקדים [1].

מאז ומתמיד **משברים בשוק ההון נחשבו להזדמנות** – משקיעים נוהגים לנצל משברים על מנת לקנות ניירות ערך במחירים אטרקטיביים, ולהרוויח את העלייה המגולמת בהתאוששות מהמשבר.

הדרך לניצול הפוטנציאל הנ"ל רצופת מכשולים, ולהלן העיקריים שבהם :

1. קשה מאוד **לחזות** אילו חברות ישרדו את המשבר. כך למשל, לאחר משבר הסאב-פריים בשנת 2008 חברות רבות ירדו מנכסיהן ופשטו רגל [2].
2. קשה מאוד **לתזמן את הכניסה לשוק –** זוהי בעצם נקודת המינימום של נייר הערך המועמד לקנייה, המאפשרת למשקיע את התזמון המושלם להיכנס לשוק במחיר אטרקטיבי.

שיטה נפוצה שהתפתחה במרוצת השנים להתמודדות עם הבעיה הראשונה היא **פיזור**. המשקיע הממוצע יחלק את הונו בין אפשרויות השקעה שונות, כדי להיות חסין מירידות ערך מאסיביות, ולייצר רשת ביטחון הגונה.

אחת הדרכים להשיג את הפיזור הנ"ל היא מכשיר פיננסי בשם **תעודת סל**. זוהי תעודה המונפקת על ידי גוף פיננסי (בנק או בית השקעות), ומאפשרת למשקיע לעקוב אחר נכס בסיס כגון **מדדי מניות** (מדדים המהווים ממוצע משוקלל של מספר רחב של מניות).

ההתמודדות עם **הבעיה השנייה היא שאלת מיליון הדולר של עולם ההשקעות,** ורבים מאמינים שבלתי אפשרי לחזות את התנהגות השוק. במרוצת השנים התפתחו שתי אסכולות מקבילות סביב השאלה הזו : Timing the markets VS Time in the markets [3] (השקעה לטווח הארוך אל מול תזמון השוק).

בפרויקט זה ננסה לאתגר אלגוריתמים שונים של בינה מלאכותית ולמידה עמוקה עם היכולת לתזמן את השוק בזמן משבר, ולמקסם את הרווח של המשקיעים = "קנה בזול ומכור ביוקר".

**תיאור הפתרון המוצע לבעיה**

תמצית הפתרון :

ב-30 השנים האחרונות שוק ההון התפתח במספר מובנים :

* ידע משברים רבים, מהם ניתן ללמוד על אופן ההתאוששות של השוק.
* נוספו אינדיקטורים רבים שנועדו לצוד תופעות שונות בשוק במטרה להבין לאן פניו. האינדיקטורים הללו משמשים משקיעים רבים ככללי אצבע בתכנון ההשקעות שלהם.

ברצוננו, בהינתן נייר ערך מסוים, ויום מסוים בעיצומו של משבר, לתת חיזוי לגבי מידת הכדאיות של כניסה לשוק. יום שפוטנציאל הרווח בו הוא גדול יקבל ציון קרוב ל-1, ויום בעל פוטנציאל רווח מינימלי יקבל ציון קרוב ל-0.

לשם כך, נשתמש במשברים קודמים כדוגמאות אימון לאלגוריתם הלמידה שלנו. כל יום מסחר בעיצומו של משבר יהווה דוגמא.

* **התכונות** של יום המסחר הם אינדיקטורים ממקורות שונים בהם האלגוריתם יתחשב.
* **התיוג** של יום המסחר הוא פוטנציאל הרווח שלו (על פי חישוב שיפורט בהמשך).

תיאור שלבי המערכת :

בשלב זה נתאר בפרוטרוט את החלקים השונים במערכת, וניגע באתגרים המרכזיים בכל אחד מהם.

1. **Data Extractor**

חלק זה אמון על הגדרת "משבר" על סמך patterns של ביצועים. הוא מקבל נתונים של ביצועים היסטוריים של נייר ערך, וקובע תקופות בהיסטוריה בהם נייר ערך זה היה במשבר.

ההגדרה למשבר היא תקופה בזמן שעונה על הקריטריונים הבאים :

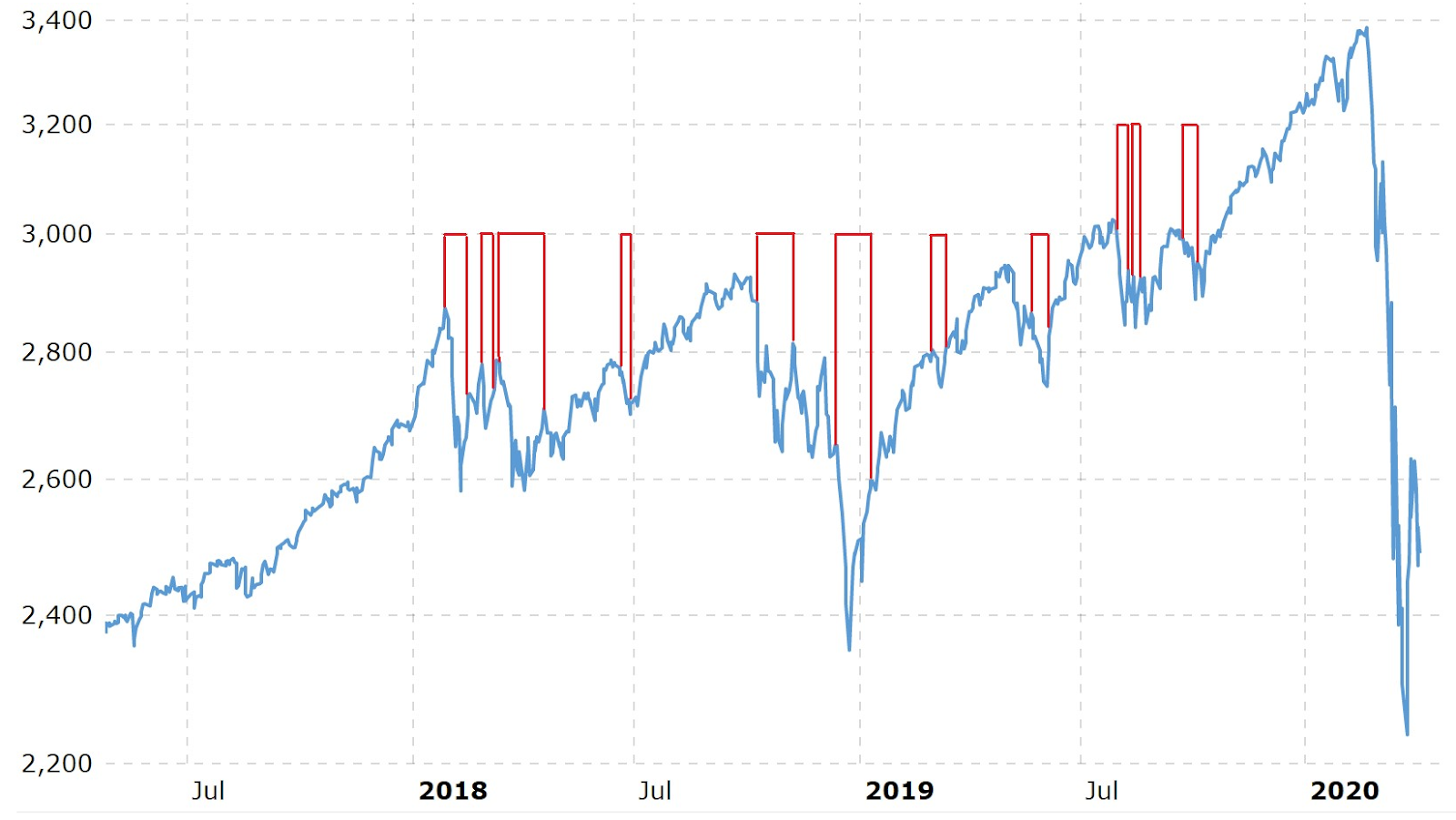
* + אורך מינימלי : לא כל יום של ירידות בשווקים ייחשב כמשבר. נרצה להגדיר סף מינימום (במונחים של ימי מסחר) שמצביע על משבר משמעותי מספיק.
  + פגיעת מינימום : לא כל רצף של ימי ירידה מעיד על משבר. נרצה שהירידה תהיה חדה מספיק (או מתונה וארוכה), ולכן נרצה להגדיר סף למחיר בשיא המשבר (נקודת המינימום ביחס למחיר המקורי).

סיום המשבר ייקבע ליום הראשון בו נייר הערך חזר להיות קרוב למחיר שממנו התחיל המשבר.

הערות :

* + האורך המינימלי, פגיעת המינימום, ורמת הקירבה למחיר המקורי (סוף המשבר) הם היפר-פרמטרים של המודל, וייקבעו בהתאם למדידות על סמך ניסויים שנבצע בהמשך.
  + קיימים משברים בהם נצפו ימי עליות בודדים אשר לא מעידים על סוף המשבר (לדוגמה, כתוצאה מהזרמה ממשלתית של כספים). כדי להתמודד עם הבעיה, נגדיר היפר-פרמטר נוסף שמאפשר ימי תיקון.

להלן ביצועי האלגוריתם על משברים בשלושת השנים האחרונות (בכחול – ביצועי האמת של מדד S&P500 בבורסה האמריקאית, באדום – תקופות אותן האלגוריתם הגדיר כמשבר).



1. **Data Combiner**

חלק זה אמון על איסוף אינדיקטורים ותכונות מימי המסחר שבתוך תקופות המשבר.

הגדרות :

* + Nasdaq, S&P500 : מדדים מובילים בעולם המניות האמריקאי. סחירים בבורסה באמצעות תעודות סל או קרנות מחקות.
  + אינפלציה : נתון המתאר את עליית המחירים במשק. נמדדת באמצעות מדד המחירים לצרכן, הבודק את השינוי במחירו של סל מוצרים קבוע (עלייה של מחירו גוררת עלייה באינפלציה).
  + Earning Ratio : היחס בין רווחי החברה לשווי המניה. זוהי דרך נפוצה לחשב רווח למניה.
  + Dividend Rate : מספר הדולרים שמשקיע מקבל מהחברה בה השקיע לכל מניה שקנה, כתוצאה מרווחי החברה.
  + Dividend Yield : היחס בין ה-Dividend Rate של החברה לבין מחיר המניה שלה.
  + Price-Earnings Ratio : היחס בין שווי השוק של חברה לבין הרווח הנקי השנתי שלה. במדדים, מבצעים ממוצע משוקלל של היחס לפי החברות המשתתפות במדד.
  + CAPE Ratio : מתאר את שווי השוק של מדד ה-S&P500 חלקי הרווח הממוצע, מתואם לאינפלציה, ב-10 השנים האחרונות.
  + מדד הפחד (VIX) : מדד התנודתיות של בורסת האופציות בשיקאגו (CBOE), הנועד להתחקות אחר אלמנטים פסיכולוגיים של משקיעים בשוק.
  + APA : תאגיד נפט אמריקאי. משקיעים רבים רואים בו כמכשיר פיננסי שעוזר במעקב אחר מחיר הנפט בעולם. זה עשוי להיות אינדיקטור טוב לתקופות משבר כי משקיעים רבים עוברים להשקיע בסחורות בתקופות כנ"ל.
  + GLD : קרן סל שמתיימרת לעקוב אחר מחיר הזהב בעולם. זה עשוי להיות אינדיקטור טוב לתקופות משבר כי משקיעים רבים עוברים להשקיע בסחורות בתקופות כנ"ל.
  + TNX : תשואת אג"ח ממשלתי אמריקאי ל-10 שנים. ככלל, אג"חים ממשלתיים נחשבים לאפיק השקעה סולידי ובטוח, ואופן המסחר בהם קשור קשר הדוק לתקופות משבר.

תכונות בתדירות יומיות :

1. מחיר סגירה של S&P500 (Close) : מחיר המניה בתום יום מסחר.
2. מחיר גבוה של S&P500 (High) : העסקה היקרה ביותר (לפי מחיר מניה בודדת) שהתבצעה באותו יום.
3. מחיר נמוך של S&P500 (Low) : העסקה הזולה ביותר (לפי מחיר מניה בודדת) שהתבצעה באותו יום.
4. מחזור מסחר (Volume) : שווי כולל של העסקאות עבור מניה זו שבוצעו באותו יום.
5. מחיר סגירה של nasdaq (Nas-Close).
6. מחיר גבוה של nasdaq (Nas-High).
7. מחיר נמוך של nasdaq (Nas-Low).
8. מחזור מסחר של nasdaq (Nas-Volume).
9. מחיר סגירה של מדד הפחד (Vix-Close).
10. מחיר גבוה של מדד הפחד (Vix-High).
11. מחיר נמוך של מדד הפחד (Vix-Low).
12. מחיר סגירה של APA (Apa-Close).
13. מחיר גבוה של APA (Apa-High).
14. מחיר נמוך של APA (Apa-Low).
15. מחזור מסחר של APA (Apa-Volume).
16. מחיר סגירה של GLD (Gld-Close).
17. מחיר גבוה של GLD (Gld-High).
18. מחיר נמוך של GLD (Gld-Low).
19. מחזור מסחר של GLD (Gld-Volume).
20. מחיר סגירה של TNX (Tnx-Close).
21. מחיר גבוה של TNX (Tnx-High).
22. מחיר נמוך של TNX (Tnx-Low).
23. מחזור מסחר של TNX (Tnx-Volume).

תכונות בתדירות חודשית :

1. אינפלציה.
2. Earning Ratio.
3. Dividend Rate.
4. Dividend Yield.
5. Price-Earning Ratio.
6. CAPE Ratio.

הערות :

1. כדי שנוכל להשתמש במסווג בזמן אמת, נרצה שהתכונות יתבססו על העבר בלבד. לכן בעת סיווג של יום מסחר, המסווג לא ייקח בחשבון תכונות של יום המסחר הנוכחי, אלא של קודמיו.
2. מאותה סיבה, התכונות החודשיות הן במבט לאחור. כך למשל, ביום מסחר של חודש מרץ, נוכל להסתכל על הנתונים החודשיים של ינואר-פברואר.
3. כל תכונה מפוצלת בפועל למספר תכונות, על מנת שנוכל לייצר זיכרון למודל על סמך תקופות זמן משמעותית יותר. כך למשל, נרצה לקחת בחשבון תכונות יומיות של חמישה ימים אחורה, ולא של יום בודד. האורך המדויק הוא היפר פרמטר של המודל.
4. **Features Converter**

כפי שראינו בחלק הקודם, חלק מהתכונות הן בתדירות יומית וחלקן בתדירות חודשית. חלק זה אמון על הגדרת ערכים יומיים שמהווים אינדיקטורים טובים לערכים החודשיים, כדי שנוכל להשתמש בערכים החודשיים במסווג.

Todo – להחליט איך עושים את זה, וספציפית גם להתייחס לזה שחלק מהנתונים הם סוף חודש וחלק מהתונים הם תחילת חודש, ואיך התמודדנו עם זה.

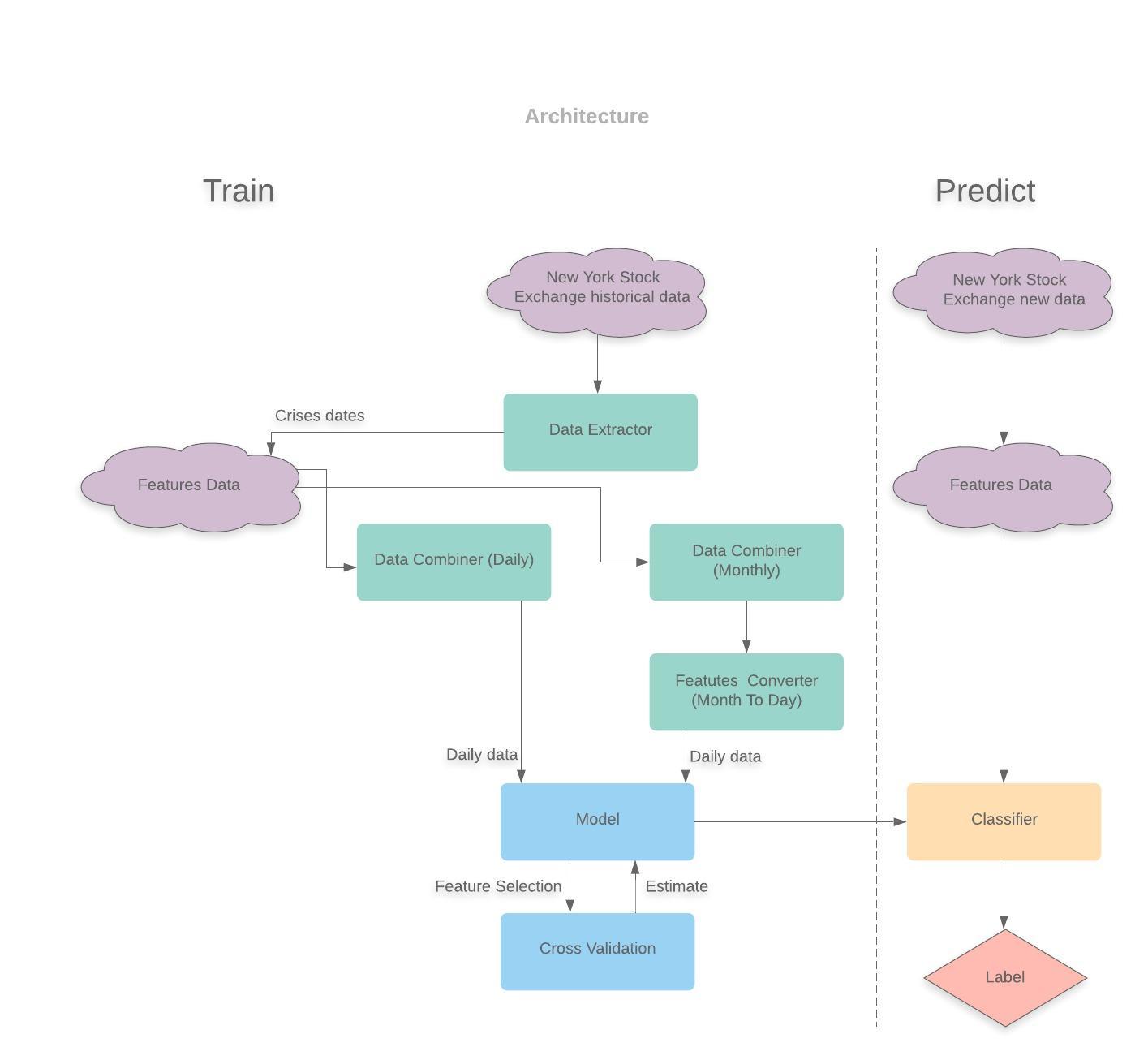
1. Model

TODO – איך מחושב פוטנציאל הרווח, תיוג של ימי מסחר וכו'...

1. Cross Validation

TODO - הערכה על סמך ביצועים ולקיחה של תת קבוצה אידיאלית של תכונות לסיווג.

מבט-על :



**ביבילוגרפיה**

[1] <https://www.maariv.co.il/business/world/Article-760391>

[2] <https://he.wikipedia.org/wiki/%D7%94%D7%9E%D7%A9%D7%91%D7%A8_%D7%94%D7%9B%D7%9C%D7%9B%D7%9C%D7%99_%D7%94%D7%A2%D7%95%D7%9C%D7%9E%D7%99_(2008)>

[3] <https://zoefin.com/learn/market-timing-vs-time-in-the-market/>