# תקציר מנהלים:

תיאור הפרויקט:

מטרת הפרויקט היא תכנון ופיתוח מערכת משולבת למסחר בשוק ההון, כלומר לניהול תיק מסחר בשוק ההון עבור השוק הפרטי. המערכת תשאף להניב יחס רווח לסיכון גבוה ככל האפשר, יחס המכונה גם כמדד שארפ. מסחר באופן הזה יאפשר סחר בצורה רציונלית ללא השפעה רגשית ויחסוך לסוחרים זמן רב של עבודת מסחר ידנית וספקולטיבית.

בפרויקט זה נבנה מערכת מסחר חצי-אוטומטית. כלומר, האלגוריתם יקבל את הנתונים והמידע מהשוק ומהמדיה החברתית, ינתח אותם בצורה אוטומטית ויספק הוראות מסחר ומידע שנועד לעזור לסוחרים לקבל החלטות, אך בניגוד למערכת מסחר אוטומטית מלאה, הוא לא יבצע את פקודות המסחר עבור הסוחרים בצורה אוטומטית. מערכת מסוג זה מאפשרת לנו להגיע לתוצאות טובות בצורה נוחה ומהירה, ויחד עם זאת לספק לסוחרים שליטה מלאה על הפעולות המתבצעות בחשבון שלהם. המערכת משלבת בין ניתוח טכני וסטטיסטי של הנתונים בצורה אנליטית לבין הניסיון, הידע והאינטואיציה האנושית.

מערכת הפרויקט מבוססת למידה עמוקה והיא מורכבת משלושה חלקים מרכזיים: איסוף מידע, עיבוד מידע ושערוך מחיר מניות להשקעה.

איסוף נתונים: קבלת נתונים עדכניים בזמן אמת בנוסף לנתונים היסטוריים על מניות הנסחרות בבורסה. קבלת מידע, כגון מועדי דיווח לבורסה, חדשות רלוונטיות על החברות הנסחרות בבורסה כגון דוחות רבועניים, עסקאות בעלי עניין וכו' , בצורה אוטומטית.

לאחר בחינת מספר חלופות לתקשורת הנתונים הוחלט השימוש בזחלן, וכתיבת הקוד בשפת Python. שפת הקוד Python מאפשרת בניית אלגוריתמים למסחר בשוק ההון, יש לה ספריות זמינות באינטרנט לציבור הרחב והיא מספקת סביבת עבודה לפיתוח האלגוריתם המלא.

מדד ההצלחה לחלק זה של הפרויקט יהיה מידת זמינות הנתונים של מניות, דוחות רבעוניים ועסקאות בעלי עניין. היקף הזמינות שמאפשר השימוש בזחלן מראה על עמידה במדד זה.

עיבוד מידע: חלק זה מתמקד בחילוץ רשימת ציוצים מטוויטר, עיבודם, והערכתם באמצעות מודל עיבוד מקדמים, זאת במטרה לקבלת פלט כקובץ אקסל המכיל את שם המניה, סמלה, הציוץ, שם המשתמש, שם המצייץ וערך השפעת הציוץ על מחיר המניה. פלט זה נועד להיות כהכנה לשימוש בשלב השערוך.

### שערוך:

חלק זה מתמקד בחילוץ רשימה קצרה של מניות בעלות יחס רווח לסיכון גבוה. ישנם שלוש שיטות עליהן ניתן לבסס את תהליך הסינון ומיצוי המידע: ניתוח טכני של גרף המניות (Technical Analysis), ניתוח פונדמנטלי של הדוחות הרבעוניים של כל מנייה המפורסמים לציבור בכל שלושה חודשים (Fundamental Analysis) וניתוח סנטימנטלי של המדיה והרשתות החברתיות במטרה לצבור מידע רלוונטי על המניות הנסחרות בבורסה ועל מצב השוק הכללי (Sentimental Analysis). בפרויקט זה נתמקד בשיטת הניתוח הטכני. כאשר נבצע ניסיון להשתמש בשילוב ניתוח טכני ופונדמנטלי, ובשילוב ניתוח טכני וסנטימנטלי.

ניתוח טכני מתבסס על התיאוריה לפיה דפוסי עבר של המחיר והמגמה של מנייה כלשהי הנסחרת בבורסה, יכולים לעזור להעריך את ההתנהגות העתידית של אותה מנייה. החלק העיקרי של הניתוח הטכני מתרכז סביב השימוש בלמידה עמוקה וחילוץ מאפיינים באופן אוטומטי. אל אף שישנם מדדים רבים הקיימים בניתוח הטכני בעולם הפיננסי, אנו ננסה בעבודה זו למטב את חילוף המאפיינים באמצעות למידה עמוקה בתקווה שכך נקבל תוצאות טובות יותר ממדדים מוכנים הנמצאים כיום בשימוש בשוק ההון ונגישים לכולם.

בעבודה זו השתמשנו בשלושה מודלים, הראשון LSTM המבוססת על רשת נוירונים חוזרת אשר בפועל מתבססת על פונקציית רגרסיה, כאשר ההערכה היא שבעקבות מידע מהעבר נוכל לשערך את ערך מחיר הסגירה של המניה ביום הבא. בנוסף למערכת זו, השתמשנו גם במודל השני MLP התבססנו על בינה מלאכותית לחיזוי לינארי של ערך מחיר הסגירה של מניה מסוימת בהתבסס על מספר כניסות, לטובת הבדיקות נוסה המודל שלוש פעמים עם מידע שונה כקלט, פעם אחת עם מידע טכני בלבד כקלט, פעם נוספת עם מידע טכני ועסקאות בעלי עניין כקלט ופעם השלישת עם מידע טכני וערך סנטינמטלי כקלט. לאחר שני מודלים אלו LSTM וMLP נוסה שימוש במדד שנקרא MACD (Moving Average Convergence Divergence). מדד זה הוא מדד מבוסס מומנטום (momentum) שמטרתו לזהות נקודות בגרף המנייה שבהן המגמה העתידית תהיה חיובית. בחרנו להשתמש ב-MACD לאחר סקירת מספר חלופות ובדיקות רבות שהראו כי זהו המדד המתאים ביותר לפרויקט שלנו.

מדדי ההצלחה לחלק זה יהיו מציאת מניה עם מדד שארפ (מדד יחס רווח לסיכון) של לפחות 0.5, וסטיית תקן שלא עולה על 10%.

החלק האחרון בפרויקט הוא סימולטור אשר מקבל כל יום פרדיקציה (ניבוי) עבור כל מנייה, בוחר את האחת הטובה ביותר לפי מיון רגיל ומחליט האם הוא נכנס אליה בהינתן סף כניסה מסוים מידע של 5 שנים על אותה המניה וערך של חיזוי על קבוצת האימות בלפחות 60%. הסימולטור בכל עסקה משקיע סכום לפי העדפת המשתמש.

בסוף הסימולטור מחשבים את מדד השארפ הכולל לכל תקופת המסחר, את כמות הכסף שהרווחנו ואת סטיית התקן בגרף הכסף, כמו גם ערכים נוספים שאפשרים לנו להעריך את הביצועים כגון טעות ריבועית ממוצעת ועוד.

התוצר: התוצר הסופי של פרויקט הוא אלגוריתם מבוסס למידה עמוקה המיועד למסחר בשוק ההון. האלגוריתם מנתח נתוני עבר של מניות ומאמן מודל שמטרתו לחזות את ההתנהגות העתידית של אותן מניות. התוצר הסופי מכיל סימולטור שמבצע Backtesting ובודק מה היו ביצועי האלגוריתם אילו היינו מריצים אותו על נתוני עבר (בנספחים, מידול, מודל שלוש).

לפי הרצת הסימולטור האלגוריתם הגיע לתוצאות טובות בהרבה מהיעדים שנקבעו ב-SOW ועמד בכל המדדים שהוגדרו. אם זאת חשוב לציין כי האלגוריתם אינו הורץ בפועל במסחר ישיר בבורסה ולכן אין נתונים אמיתיים על איכות האלגוריתם בזמן אמת.

סיכום מסקנות והצעה לעבודת המשך: הפרויקט הוכיח כי ניתן להשתמש באינדיקטורים המשמשים סוחרים ידניים באמצעות למידת מכונה כדי לסחור בבורסה באופן אוטומטי ומוצלח.

פיתוח הפרויקט לא כולל בתוכו את הפונקציות האחריות לניהול תיק וביצוע פעולות מול הבורסה, העבודה שנעשתה היא בעיקרה עבודת מחקר לבחירת סוג המודל המתאים ביותר ובדיקת האלגוריתם.

ניתן לשפר ולפתח את הפרויקט על ידי מימוש סורק אמיתי אשר עובר על כל המניות בבורסה כדי למצוא בכל יום מספר מצומצם של המניות הכדאיות ביותר למסחר, להוסיף לניתוח הטכני גם ניתוח פונדמנטלי מעמיק לאחר איסוף המידע וניתוח סנטימנטלי ממספר מקורות, כמו כן

ניתן לממש הפונקציות הרלוונטיות אשר יהיו אחראיות למסחר מול הבורסה, ניהול סיכוני התיק וטיוב פוזיציות ובכך להשלים את החלק החסר בפרויקט זה.

# Executive Summary:

#### the project's description:

The purpose of the project is to design and develop an integrated system for trading in the capital market, ie to manage a trading portfolio in the capital market for the private market. The system will strive to yield a profit-to-risk ratio as high as possible, a ratio also known as the Sharpe Index. Trading in this way will allow trading in a rational way without emotional impact and will save traders a lot of time of manual and speculative trading work.

A semi-automated trading system was built in this project. That is, the algorithm will receive the data and information from the market and social media, analyze it automatically and provide trading instructions and information designed to help traders make decisions, but unlike a fully automated trading system, it will not execute trading orders for traders automatically. This type of system allows us to achieve good results in a convenient and fast way, while at the same time providing traders with complete control over the operations performed on their account. The system combines technical and statistical analysis of the data in an analytical way with human experience, knowledge and intuition.

The project system is based on deep learning and consists of three main parts: information gathering, information processing and stock price revaluation for investment.

Data collection: Receiving up-to-date data in real time in addition to historical data on shares traded on the stock exchange. Receipt of information, such as reporting dates to the stock exchange, relevant news about the companies traded on the stock exchange, such as quarterly reports, stakeholder transactions, etc., automatically.

After examining several alternatives to data communication, it was decided to use a crawler, and write the code in Python. The Python code language enables the construction of algorithms for capital market trading, has libraries available online to the general public and provides a work environment for developing the full algorithm. The measure of success for this part of the project will be the degree of availability of shares, quarterly reports and stakeholder transactions. The extent of availability that allows the use of the crawler shows compliance with this index.

Information processing: This section focuses on extracting a list of tweets from Twitter, processing them, and evaluating them using a coefficient processing model, in order to obtain output as an Excel file containing the stock name, symbol, tweet, username, tweeter name and the effect of the tweet on stock price. This output is intended to be in preparation for use in the revaluation phase.

predictions: This section focuses three methods on which the process of filtering and extracting information can be based: technical analysis of the stock chart (Technical Analysis), fundamental analysis of the quarterly reports of each stock published to the public every three months (Fundamental Analysis) and sentimental analysis of media and social networks. Stocks traded on the stock exchange and on the general market situation (Sentimental Analysis). In this project we will focus on the method of technical analysis. When we make an attempt to use a combination of technical and fundamental analysis, and a combination of technical and sentimental analysis. Technical analysis is based on the theory that past patterns of the price and trend of any stock traded on the stock exchange, can help assess the future behavior of that stock. The main part of the technical analysis revolves around the use of deep learning and feature extraction automatically. Although there are many metrics that exist in the technical analysis in the financial world, we will try in this work to optimize the exchange of characteristics through deep learning in the hope that we will get better results from ready-made indices that are currently used in the capital market and accessible to all.

Success indices for this part will be finding a stock with a Sharp index (profit-to-risk ratio index) of at least 0.5, and a standard deviation not exceeding 10%.

The last part of the project is a simulator that receives a daily prediction for each stock, chooses the best one by regular sorting and decides whether it enters it given a certain entry threshold. . The simulator in each transaction invests an amount according to the user's preference. At the end of the simulator we calculate the overall Sharpe index for the entire trading period, the amount of money we earned and the standard deviation in the money graph, as well as other values ​​that allow us to evaluate performance such as average quadratic error and more.

Product: The final product of a project is a deep learning algorithm designed for trading in the capital market. The algorithm analyzes past stock data and trains a model that aims to predict the future behavior of those stocks. The final product contains a simulator that performs backtesting and checks what the algorithm's performance would have been if we had run it on past data (in appendices, modeling, model three). By running the simulator the algorithm achieved much better results than the targets set in the SOW and met all the defined indices. If so it is important to note that the algorithm is not actually run in direct trading on the stock exchange and therefore there is no real data on the quality of the algorithm in real time. Summary of Conclusions and Suggestion for Continued Work: The project has demonstrated that indicators used by manual traders through machine learning can be used to trade the stock market automatically and successfully. The development of the project does not include the functions responsible for portfolio management and execution of operations with the stock exchange,The work done is mainly research work to select the most appropriate model type and test the algorithm. The project can be improved and developed by realizing a real scanner that goes through all the stocks on the stock exchange to find every day a limited number of the most profitable stocks to trade, add to the technical analysis also an in-depth fundamental analysis after gathering the information and sentimental analysis from several sources.The relevant functions that will be responsible for trading with the stock exchange, portfolio risk management and position improvement can be realized, thus completing the missing part of this project.