**מגישים:**

איתמר ריינמן, ת"ז 326935285

בן אליאב, ת"ז 330053893

אייל אל אני, ת"ז 213950496

שיטות חיזוי בפינטק

חיזוי מחירים של דגנים וסוגי יבול

**תיאור העבודה**

הנכס אותו אנחנו מעוניינים לבדוק במחקר שלנו הוא מחירי חוזים עתידיים של מוצרי יסוד, בהתמקדות על תבואה (תירס, חיטה, שיבולת שועל) ופולי סויה. כל המוצרים נסחרו בבורסת הנגזרות של שיקגו (Chicago Board of Trade, CBOT), הנמצאת בתוך בורסת הסחורות של שיקגו (CME Group). בורסת הנגזרות של שיקגו היא בורסה שעוסקת במכירת סחורות חקלאיות, חוזים עתידיים של מגוון סחורות (לא רק סחורות חקלאיות,כגון זהב, כסף, אג״ח של האוצר האמריקאי, אנרגיה ועוד). השוק בחוזים עתידיים מהסוג הזה מאופיין על ידי High Risk – High Reward ומאוד הפכפך.

הסחורות עצמן נמצאות תחת הפיקוח של ה-CBOT, האחראים למדידת האיכות של המוצרים וקביעת הסטנדרטים לסחר. המידע המדויק של כל אחד מהמוצרים והמידות שנסחרות בכל חוזה מופיעים בטבלה למטה:

טבלה 1 – פירוט המוצרים הנסחרים. נשים לב שיחידות המידה bushels הן יחידות מידה של נפח השקולות ל-8 גלון.

|  |  |  |  |
| --- | --- | --- | --- |
| שם מוצר | קוד מוצר | זן | כמות בחוזה |
| Chicago Corn (Maize) | ZC | No. 2 Yellow Corn | 5000 bushels |
| Chicago Oats | ZO | No. 2 Oats | 5000 bushels |
| Chicago Soybean | ZS | No. 2 Yellow | 5000 bushels |
| Chicago Wheat | ZW | No. 2 Soft Red Winter | 5000 bushels |

לפי מאמרים קודמים העוסקים בשוק החוזים העתידיים של סחורות מהסוג הזה [1,2,3], לא מקובל להשתמש בפיצ׳רים חיצוניים למען חיזוי המחיר בעתיד. בנוסף, לפי דוח של Economist Intelligence (EIU) [4], מחירי ה-commodities בשנת 2024/2025 יושפעו בעיקר מ-אל ניניו, הנסיגה של רוסיה מהסכם סחר עם אוקראינה, רגולציות על סחר ירוק וחוקים מחמירים על ייצוא ממדינות מייצאות כגון הודו. הדברים הנ״ל לאו דווקא מתבטאים בתור פיצ׳רים שנוכל להכניס למודל ומתארים מקרה הרבה יותר ״מאקרו״. לכן בחרנו בחיזוי להתמקד בחיזוי מחירים לטווח הקצר, איפה שיש פחות השפעה על תהליכים עולמיים ויותר דמיון בין המחירים שחוזים לבין המחירים הנצפים בעבר.

במודל שלנו שנפרט בהמשך בחרנו להשתמש רק בפיצ׳ר חיצוני אחד – ריבית פריים באמריקה. בחרנו להוסיף את הפיצ׳ר של הפריים מפני שבדרך כלל כאשר הריבית עולה, מחירי ההלוואה עולים, מה שמאוד משפיע על ההחלטות של משקיעים להשקיע במניות / חוזים עתידיים על פני השקעות באג״ח. נמצא בעבר קשר סיבתי המקשר בין מחירי הריבית לבין שינויים במחירי מניות בארצות הברית [5], כמובן יחס הפוך מהרציונל שמתואר למעלה. חוץ מהפריים אשר בדרך כלל נשאר קבוע לאורך תקופה ארוכה, בחרנו להוסיף פיצ׳ר של הפרש ימים בין כל שני ימי מסחר מפני שמתוך 365 ימים בשנה, הבורסה פתוחה רק ב-252 ימים. הבדל בין הפסקה של יומיים בין ימים לשני ימי עבודה רצופים יכול לבוא לידי ביטוי בשינוי במחירים עד היום הבא.

הנתונים הכלכליים עצמם נאספו משני מקורות, AHDB ו-Yahoo Finance. AHDB (Agriculture and Horticulture Development Board) היא רשות החקלאות בממלכה המאוחדת. היא מספקת תמיכה דרך מו״פ ומידע שוק על מנת לסייע לחקלאים ולתעשייה הרחבה לייעל את התהליכים שלהם. המידע מ-AHDB נובע מנתונים רשמיים שפורסמו על ידי CME. השתמשנו גם בממשק של Yahoo Finance ב-Python על מנת לגשת ליותר מידע, ובשביל לחשב מדדים כגון סיכון Sharpe שמסתמך גם על נתוני מניה חסרת סיכון אחרת. בכל הנתונים, בחרנו להתייחס למחיר ה-Close של כל יום עסקים בתור המחיר של אותו היום, מפני שזה המידע שהתלכד עם המידע שנמצא בטבלה של AHDB. המטבע המשומש בסחר של הסחורות הוא הדולר האמריקאי.

הנתונים מ-AHDB מספקים נתונים יומיים על כל החוזים העתידיים בתאריך מסוים (החל מ-1.1.2018) עם תאריך תפוגה עד מספר שנים הפרש. לדוגמה, עבור השני לינואר, 2018, יש מידע על חוזים עם תאריך תפוגה החל ממרץ 2018 עד דצמבר 2020. תאריכי התפוגה של החוזים תמיד בחודשים מרץ, מאי, יולי, ספטמבר ודצמבר. זה אומר שלכל תאריך ולכל סחורה, יש מידע על לכל הפחות חמישה סוגי חוזים בשנה, ומינימום עשרה בסך הכל (אם ישנם חוזים לעוד שנתיים). הנתונים מ-Yahoo Finance גם יומיים לכל יום שהבורסה פתוחה, אך הם יותר ארוכי טווח (מתחילים בשנת 2000). מהצד השני, ניתן לגשת רק למחירי ה-Nearby של כל אחד מהחוזים. כלומר, יש מידע רק על החוזים עם תאריך התפוגה הקרוב ביותר (מסומן עם קוד =F[קוד מוצר]). מצורף בתחתית הקובץ קישור למקור הנתונים ב-AHDB. הגישה לנתונים מ-Yahoo Finance מתבססת על הספרייה ה-Open Source, yfinance של Python המשתמשת ב-API של Yahoo Finance.

המטרה הכללית של העבודה היא להבין יותר לעומק את שוק הסחר בחוזים עתידיים של מוצרי מזון בסיסיים. המטרה הסופית היא להשתמש במה שלמדנו על מנת לבנות תכנית השקעה בשוק, תוך התחשבות בטרנדים במחירי החוזים ומדדי סיכון (Sharpe). נשתמש בשיטות סטטיסטיות ושיטות בלמידת מכונה בשביל לחזות את מחירי החוזים ואת רמת הסיכון לתקופה הקרובה. כל הקוד מופיע ב- <https://github.com/ireinman/fintech/tree/master> .

המודלים העיקריים שהכנו:

1. מודל רגרסיה לינארית החוזה את מחיר החוזה עם תפוגה בספטמבר של שנה הבאה כתלות במחירי החוזים עם תפוגה בחודשים מוקדמים יותר. הקוד מופיע ב regression.ipynb בתיקייה.

|  |  |
| --- | --- |
| המשתנה התלוי | המשתנים הבלתי תלויים |
| מחיר ZC בספטמבר שנה הבאה | מחירי ZC במרץ, מאי, יוני שנה הבאה |
| מחיר ZO בספטמבר שנה הבאה | מחירי ZO במרץ, מאי, יוני שנה הבאה |
| מחיר ZS בספטמבר שנה הבאה | מחירי ZS במרץ, מאי, יוני שנה הבאה |
| מחיר ZW בספטמבר שנה הבאה | מחירי ZW במרץ, מאי, יוני שנה הבאה |

הבדיקה הזאת מודדת את הקורלציה בין מחירי החוזים השונים באותה נקודת זמן. המדדים הסטטיסטים המעניינים אותנו כאן הם RMSE (Root Mean Squared Error) ו-R^2 (Pearson Correlation). ביצענו מבחן רגרסיה נוסף המודד את R^2, RMSE במקרה שבו המשתנים הבלתי תלויים זה מחירי כל שאר הסחורות בתאריכי תפוגה אחרים, במקום רק מחירי אותה סחורה.

1. מודל SARIMA לתחזית סדרתית בשביל לחזות מחיר חוזה עתידי רק על בסיס מחירי העבר שלו. נמצא בקובץ sarima\_updated.ipynb בתיקייה. התנסינו בהרבה מאוד שילובים של היפרפרמטרים למודל ב-Order וב-Seasonal Order, כאשר ניסינו לתפוס את המורכבות והחזרתיות של המודל אך לא באמצעות יותר מידי פרמטרים. המשתנה התלוי אותו חוזים הוא רק מחיר החוזה לכל מוצר בחודש ספטמבר של שנה הבאה, כלומר הערך ב 2.1.2019 יהיה המחיר של חוזה לתאריך תפוגה בספטמבר 2020. המדדים הסטטיסטיים המעניינים אותנו כאן הם AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) ו-Log-likelihood בשביל למדוד את טיב המודל. זהו מעין מודל Baseline שמודד כמה ניתן ללמוד על החוזה ללא אף מידע חיצוני.
2. מודל למידה עמוקה (LSTM) לתחזית סדרתית של וקטור המחירים לכל סחורה, לחוזים עם תפוגה מרץ, מאי, יוני, ספטמבר בשנה הבאה. המשתנים הבלתי תלויים הם מחירי כל החוזים של כל הסחורות מהימים שלפני. במילים אחרות, מנסים לחזות את המחיר של חוזה לספטמבר 2024 שנמכר ב-2.1.23 על בסיס כל המידע הקיים ברשותנו מאותו החוזה ומחוזים אחרים, ובאמצעות ריבית הפריים הנמצאת ברשותנו. המודל אומן עם פונקציית הפסד MSE וגם התוצאות של המודל נמדדו באמצעות MSE.
3. מודל למידה עמוקה (LSTM) לתחזית של מדד הסיכון לשנה הבאה של השקעה בחוזה ה-Nearby לכל אחד מארבע הסחורות. מדד הסיכון שהשתמשנו בו הוא ה-Sharpe Ratio [6]. על מנת לחשב את ה Risk Free Rate הנמצא בחישוב המדד, השתמשנו במניה ^IRX (US 3 Month Treasury Bill) כמקובל [7]. חישבנו את ה-Sharpe Value לתקופת רבעון על פני במקום לחשב את הערך השנתי מפני שמניית ה-Nearby בעלת תאריך תפוגה יותר קצר ולכן נרצה לדד מה יהיו הרווחים שלה בטווח יותר קצר. בשביל שנוכל למדוד במידה יותר רחבה, באמצעות yfinance הבאנו מידע מחוזים של סחורות אחרות מתחומים שונים (מתכת, אנרגיה ועוד). המידע על כל החוזים שהשתמשנו בהם נמצא בתוך commodities\_tickers.json בתיקייה, והמודל של הלמידה נמצא בקובץ sharpe\_index.ipynb. השתמשנו בפונקציית הפסד MSE והתוצאות נמדדו גם באמצעות MSE.

מקורות

1. Ohyver, M., & Pudjihastuti, H. (2018). Arima model for forecasting the price of medium quality rice to anticipate price fluctuations. Procedia Computer Science, 135, 707-711
2. Kumar, R. (2021). Predicting wheat futures prices in India. Asia-Pacific Financial Markets, 28(1), 121-140
3. Ouyang, H., Wei, X., & Wu, Q. (2019). Agricultural commodity futures prices prediction via long-and short-term time series network. Journal of Applied Economics, 22(1), 468-483.
4. EIU. (2023). Commodities outlook 2024. [*https://www.eiu.com/n/campaigns/commodities-outlook-2024/*](https://www.eiu.com/n/campaigns/commodities-outlook-2024/)
5. Hashemzadeh, N., & Taylor, P. (1988). Stock prices, money supply, and interest rates: the question of causality. *Applied economics*, *20*(12), 1603-1611.
6. Sharpe, W. F. (1966). "Mutual Fund Performance". *Journal of Business. 39 (S1): 119–138.*
7. Maverick, J. (2024). What the Sharpe Ratio Means for Investors. *Investopedia* [*https://www.investopedia.com/ask/answers/010815/what-good-sharpe-ratio.asp*](https://www.investopedia.com/ask/answers/010815/what-good-sharpe-ratio.asp)

*האתר ממנו השגנו את הנתונים:* <https://ahdb.org.uk/cereals-oilseeds/futures-prices>