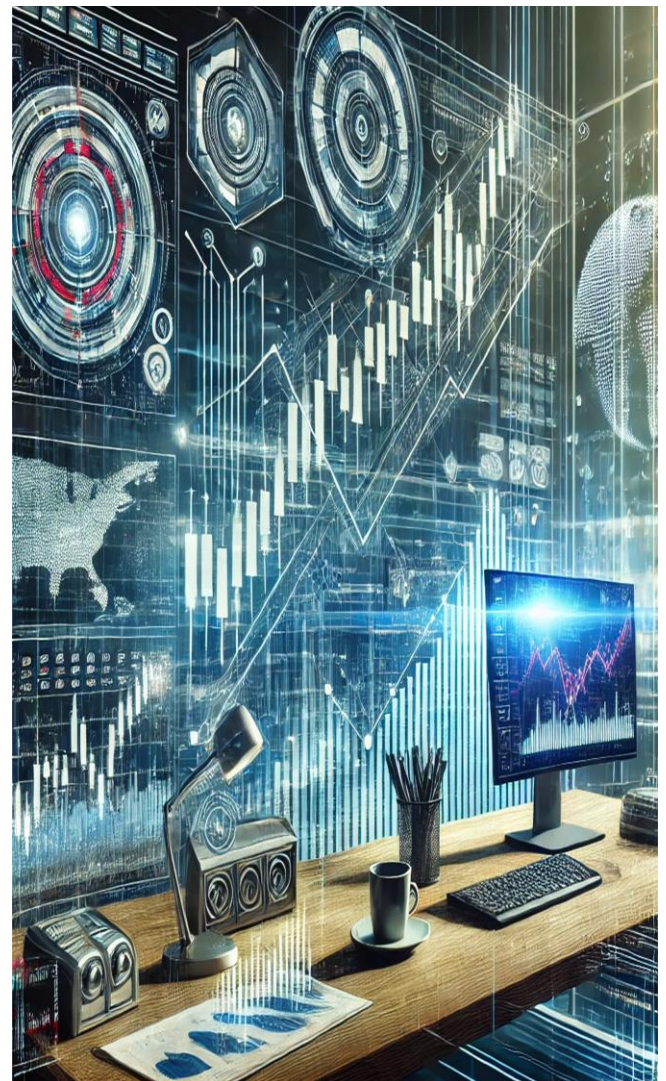


תקציר

פרויקט הגמר עוסק בפיתוח מודל חיזוי משולב לחיזוי מגמות במדד S&P 500 תוך שילוב של ניתוח טכני עם ניתוח סנטימנט של חדשות כלכליות. מטרת העל של הפרויקט הייתה לסייע למשקיעים פרטיים ולמנהלי השקעות לקבל החלטות מושכלות, להפחית שגיאות ולהגביר את דיוק התחזיות. הפרויקט כלל תהליך ביקורת ולמידה מתודולוגי, על פי עקרונות CRISP-DM, והוביל לתובנות פרקטיות וחשיבה מחודשת שתסייע בהובלת פרויקטים עתידיים בתחום הדאטה והפיננסים.



דוח הערכת המודלים

השקעות בסיכון נמוך זה אצלנו-NexTrade

מגישים – אריאל קריב ומשי בר



תוכן עניינים

2	הערכת התוצאות	1.
8	תהליך ביקורת ולמידה	2.

1. הערכת התוצאות

במהלך שלבי הפיתוח וההרצה של המודלים השונים, נבחנה ההתאמה של תוצרי הפרויקט ליעדים העסקיים שהוגדרו כבר בשלבים הראשונים במסמך ההבנה העסקית. כזכור, מטרת העל של הפרויקט הייתה לסייע למשקיעים פרטיים ולמנהלי השקעות לקבל החלטות השקעה מבוססות, תוך חיזוי רמת הסיכון במדד S&P 500 על בסיס תנודות שוק וחדשות כלכליות בזמן אמת.

מענה ליעדים העסקיים

המודל הסופי שבחרנו (GRU (Gated Recurrent Unit הוכיח יכולת לספק תחזית שבועית (ולא רק יומית כפי שתוכנן בתחילה), תוך שמירה על רמת דיוק גבוהה ואינטגרציה של ניתוח טכני עם רגשות חדשותיים. בכך, ענה המודל בפועל על ההגדרה המקורית של מטרות העסק:

- תחזית ברורה אם לקנות /להחזיק /למכור, בהתאם למגמה שזוהתה.
- אינטגרציה של מדדים טכניים עם תובנות מרגשות חדשות כלכליות, תוך הפקת תחזית משולבת.
- שיפור משמעותי בזמן תגובה ובמהימנות התחזיות, כאשר תוצאות מתקבלות בתוך פחות מ-30 שניות, ועומדות ביעדי זמינות ואמינות.

בהירות התוצאות והצגתן

בהרצות שביצענו למודל, GRU התקבל פלט מסודר שכלל את תחזית מחירי הסגירה לחמשת ימי המסחר הבאים, יחד עם המלצה לפעולה המבוססת על כיוון המגמה שזוהה. הפלט כולל את ערכי התחזיות באופן מספרי, כאשר לכל יום מוצג המחיר החזוי, וכן סיווג מילולי פשוט שמבהיר את המשמעות עבור המשקיע – לדוגמה:

- "מגמת עלייה – שקול קנייה"
- "מגמה שלילית – שקול מכירה"
- "מגמה ניטרלית – החזק"

בנוסף לתחזיות, המערכת מציגה גם את ערכי הקלט שנשמרו- כגון אינדיקטורים טכניים וציוני הסנטימנט שהוזנו למודל- כדי לאפשר למנתח הנתונים להבין על בסיס אילו ערכים התקבלו ההמלצות.

נכון לשלב זה, הפלט מתקבל ישירות מתוך קוד המודל כטבלה או מערך נתונים, ומיועד לשימוש אנליטי או לצרכי בדיקה. אנו מתכננים בהמשך הדרך לממשק את תוצרי המודל לדשבורד חזותי, על מנת להקל על השימוש בו על ידי משתמשים שאינם טכניים, אך בשלב הנוכחי הדגש הוא על דיוק התחזית ופשטות הצגת הפלט הטקסטואלי והמילולי.

תובנות ייחודיות שעלו במהלך התהליך

אחת התובנות הבולטות והמשמעותיות שזיהינו במהלך ההרצות המרובות של המודלים הייתה היכולת של מודל GRU לשלב באופן ישיר גם את משתני הסנטימנט שנגזרו מהחדשות הכלכליות, מבלי להזדקק למודל שפה מתקדם כמו BERT או RoBERTa שפעל בנפרד.

בשלבם הראשוניים של הפרויקט, תכננו לבנות מודל חיזוי כפול LSTM- לניתוח טכני ו- BERT לניתוח רגשות חדשותיים, אך בפועל- לאחר תהליך ניסוי וטעייה והכשרת הנתונים ברמה גבוהה, גילינו שניתן להזין את משתני הסנטימנט כגון- "overall sentiment label", "sentiment score" ישירות לתוך שכבות הקלט של ה-GRU. הדבר אפשר לנו לייעל משמעותית את הביצועים של המערכת, לקצר זמני חישוב, להפחית תלות במשאבי GPU ולשמור על מבנה תחזיתי חכם ואחיד.

זהו ממצא ייחודי בפרויקט שלנו, מכיוון שהוא מוכיח שניתוח חדשות כלכליות יכול להיות מוטמע ישירות כחלק ממודל תחזית שוק טכני, תוך שמירה על רמת ביצועים גבוהה.

מענה לשאלות העסקיות שהוגדרו

1. הערכת רמת הסיכון בהשקעה ב- S&P 500

המערכת שבנינו מספקת תחזית מבוססת שילוב בין שני מקורות מידע:

- אינדיקטורים טכניים - מדדים מתמטיים המבוססים על נתוני מחירי העבר של המדד. לדוגמה:

- RSI - מדד שמודד האם השוק נמצא במצב של קנייה יתרה ($RSI > 70$) או מכירה יתרה ($RSI < 30$), מה שמעיד לעיתים על פוטנציאל היפוך מגמה.

- MACD - מדד שבוחן את ההפרש בין שני ממוצעים נעים (קצר וארוך טווח), ומסייע לזהות מומנטום בשוק.

- ניתוח סנטימנט מחדשות כלכליות - ניתוח טקסטואלי המזהה את "הטון" של כתבות כלכליות כלפי השוק. לדוגמה:

- כתבה שמסווגת כ- Bullish מרמזת על ציפייה לעלייה בשוק, לעומת כתבה Bearish שמעידה על חשש מירידת ערך.

שילוב של ניתוח טכני עם רגשות החדשות מאפשר זיהוי מצבים של סיכון גבוה או נמוך, וכתוצאה מכך - המלצה מושכלת לפעולה.

2. מצבים בהם כדאי למשקיע לשקול השקעה או מכירה

המצבים שזוהו כמשמעותיים להחלטת השקעה הם כאשר מתקיימת התאמה בין אינדיקטורים טכניים לסנטימנט החדשות, למשל:

- RSI נמוך מאוד (מתחת ל-30) המעיד על שוק שנמצא במכירת יתר, יחד עם חדשות חיוביות (Bullish), עשויים לרמז שזהו זמן מתאים לרכוש.

- אם MACD מציג מגמה שלילית וחדשות Bearish, זהו סימן חזק לכך שייתכן שכדאי למכור.

- במקרה של חוסר התאמה למשל MACD חיובי אך סנטימנט שלילי, תופיע המלצה ל-"המתן" או "החזק" עד שהמגמה תתבהר.

3. חדשות שמשפיעות באופן חזק על המדד וכיצד הן נכללו במודל

ניתוח הנתונים הראה כי חדשות העוסקות במדדים כלכליים מרכזיים הן המשפיעות ביותר על תנודות מדד S&P 500, לדוגמה:

- CPI (Consumer Price Index) - מדד מחירי הצרכן, מדד האינפלציה העיקרי בארה"ב.
- Fed Interest Rate Decision - החלטת הריבית של הבנק הפדרלי, בעלת השפעה עצומה על שוק המניות.
- Unemployment Rate - שיעור האבטלה, משפיע על תחזיות הצמיחה וביקושים.

חדשות מסוג זה נבחרו והוזנו לתוך המודל באופן מבוקר, כאשר לכל כתבה חושב - sentiment score - ציון מספרי בין -1 ל-1, שמתאר את הטון הכללי שלה. תכונה זו הוזנה ישירות לרשת הנוירונים במודל כחלק מתהליך קבלת ההחלטות שלה.

דירוג המודלים לפי התאמה ליעדים

סטטוס	תיאור ביצועים	שימוש עיקרי	מודל
מאושר (סופי)	עמד בכל הקריטריונים: מהירות, דיוק, שילוב נתונים, קלות שימוש	תחזית שבועית על מדד S&P 500 כולל סנטימנט חדשות	GRU
נדחה	איטי יותר, מדויק פחות, לא מאפשר שילוב ישיר עם חדשות	תחזית יומית על סמך מדדים בלבד	LSTM
נדחה	קל משקל אך לא הוסיף ערך מובהק לפרויקט בהשוואה לשילוב בתוך GRU	ניתוח רגשות בטקסט כלכלי	DistilBERT
נדחו	כבדים חישובית, חיזוי מדויק אך איטי ומורכב מדי	ניתוח חדשות NLP מתקדם	RoBERTa / BERT

שאלות חדשות שעלו מתוך ההרצות

תוך כדי עבודה עם המודלים עלו מספר שאלות חדשות, שחלקן עשויות להוות בסיס לפרויקטים עתידיים:

- האם ניתן להרחיב את מערכת התחזית גם לשווקים נוספים, כמו NASDAQ או מדדים אירופיים?
- האם ניתן לשלב תחזיות גולשים או רגשות של קהל ממשתמשים ברשתות חברתיות - Twitter, Reddit - כדי להוסיף נדבך חדש של סנטימנט?
- האם ניתן להוסיף שכבת הסבר אוטומטית שתסביר למשתמש מדוע הומלצה פעולה מסוימת? - לדוגמה, "בגלל RSI נמוך ו-CPI חיובי".
- שאלות אלו מדגישות שהפרויקט הצליח לא רק לספק פתרון, אלא גם לפתוח פתח לחשיבה עסקית ודאטאית עמוקה יותר.

התמודדות עם אתגרים ושינויים תוך שמירה על המטרה

לאורך הדרך בוצעו התאמות רבות, אך כולן נשמרו במסגרת מטרת-העל:

- מעבר לעבודה בסביבת Google Colab לצורך ביצועים גבוהים.
- ביטול הצורך בשני מודלים לטובת איחוד ל-GRU הכולל גם את משתני הסנטימנט.
- מעבר מתחזית יומית לתחזית שבועית- תוצאה ישירה מהבנת צרכי המשתמש הסופי.

כל שינוי שנעשה נבדק אל מול מטרות הפרויקט המקוריות, תוך הקפדה שלא לפגוע ברלוונטיות העסקית של הכלי שפותח. להפך, שינויים אלה אפשרו לנו לשפר את חוויית המשתמש, לייעל את תהליך קבלת ההחלטות, ולהפוך את המערכת ליישומית יותר.

לסיכום, הפרויקט עמד במלואו ביעדים העסקיים שהוגדרו בשלב הראשון:

- סיפק כלי ניתוח שמבוסס על תחזיות עדכניות ואמינות.
- הציג חיזוי ברור והמלצה מובהקת לפעולה.
- עמד בזמני תגובה קצרים והציג תוצאה תוך פחות מ-30 שניות.
- שיפור ניהול הסיכונים- הבחנה בין מטרה תיאורטית ליישום מעשי:

אחת ממטרות העל שהוגדרו בתחילת הפרויקט הייתה לסייע בהפחתת שגיאות בהחלטות השקעה ובכך לצמצם הפסדים – יעד שאפתני שמוסגר כהפחתה של לפחות 10%. בפועל, אמנם הצלחנו לבנות מודל המספק תחזיות ברמת דיוק גבוהה, אך אין ביכולתנו במסגרת פרויקט גמר אקדמי לבדוק בצורה מלאה את ההשפעה הישירה של המודל על שיפור ניהול הסיכונים או הפחתת ההפסדים של משקיעים אמיתיים.

הסיבה לכך היא כפולה:

1. היעדר אינטגרציה מעשית בעולם האמיתי- המודל לא הופעל על תיקי השקעות חיים של קרן אמיתית, ולכן לא התאפשר מעקב אחרי החלטות השקעה אמיתיות שבוצעו על פיו.

2. חסמים רגולטוריים ובירוקרטיים- הכנסה של מודל כזה לשימוש ממשי בקרן השקעות דורשת רגולציה, אישורים והתנסות הדרגתית בשוק- תהליך שמעבר להיקף פרויקט גמר אקדמי.

לפיכך, ניתן לומר שהמודל עומד בדרישות הביצועיות והטכניות שהגדרנו (כולל רמת דיוק, זמן תגובה, בהירות פלט), אך טרם נבחן במבחן המציאות של שוק אמיתי לצורך הערכת ההשפעה הפיננסית הישירה שלו.

בכך, ענה המודל על קריטריוני ההצלחה שנקבעו מראש והוכיח את עצמו ככלי פרקטי המספק ערך אמיתי למשקיעים.

2. תהליך ביקורת ולמידה

תהליך פיתוח המודל בפרויקט זה אינו הסתכם בבנייה טכנית של מערכת חיזוי, אלא כלל בתוכו למידה עמוקה והשתכללות מקצועית, אשר יהוו בסיס לפרויקטים עתידיים שנבצע. בהתאם לרוח מתודולוגיית CRISP-DM, הקדשנו זמן משמעותי לרפלקציה שיטתית- בחינה של הצלחות, טעויות, ודפוסי פעולה שחזרו על עצמם. מסקנות אלו מאפשרות לנו לגבש עקרונות פעולה ברורים לעבודה עתידית בתחום מדעי הנתונים.

הגדרה מדויקת של צרכי המודל מראש

מהתנסותנו למדנו כי אחד הגורמים המשמעותיים להצלחת תהליך המידול הוא הגדרה מדויקת של הדרישות מהמודל כבר בתחילת הדרך- לא רק מבחינת הנתונים הקיימים, אלא בעיקר מבחינת הציפיות מהתוצאה. בפרויקטים הבאים נדע לנסח מראש אילו מאפיינים חשובים לנו יותר: האם אנו זקוקים למודל מהיר או מדויק? מודל שמייצר פלט מספרי או קטגוריאלי? האם המידע הנדרש הוא כמותי, טקסטואלי, או היברידי?

בפרויקט הנוכחי, רק לאחר הרצות מרובות והשוואה בין LSTM, GRU, BERT ו-DistilBERT התחדדה ההבנה אילו מודלים משרתים טוב יותר את צרכינו- למשל, שה- GRU אפשר שילוב ישיר של עמודות סנטימנט והיה קל משקל יותר. בפעם הבאה, נוכל לחסוך זמן על ידי ביצוע מיפוי תכונות המודל הדרוש לנו, ולאתר מראש מודלים קיימים המתאימים לבעיה הנתונה.

בקרה על איכות הנתונים ותגובה לעמודות חסרות ערך

תובנה חשובה נוספת נוגעת ליכולת לזהות עמודות שעלולות להיראות משמעותיות- אך בפועל אינן תורמות למודל או אף פוגעות בו. לדוגמה, בטבלת החדשות הכלכליות מצאנו עמודות כמו ticker_sentiment שהכילו שיעור גבוה מאוד של ערכים חסרים או אפסיים.

במקום להפעיל את המודל עם רעש מיותר, פיתחנו עמודות חלופיות מבוססות ריכוז רגשות יומיים וממוצעי סנטימנט כוללים (overall_sentiment_score), אשר אפשרו למודל להבין טוב יותר מהם המאפיינים המשפיעים באמת על כיוון השוק. השימוש בגישה זו שיפר את הדיוק והוריד את השפעתם של משתנים לא עקביים. מכאן ואילך, ניישם כלים אנליטיים מוקדמים כדי להעריך אילו עמודות מספקות ערך מובהק למידול, ואילו דורשות הסרה, שינוי או החלפה.

גמישות מחשבית ופתיחות למודלים חדשים

הפרויקט גם חידד עבורנו את הצורך בפתיחות מחשבית מול מודלים שאינם בהכרח מוכרים או ראשוניים שנבחרו. בתחילת הדרך נטינו לבחור בפתרונות כמו LSTM או BERT שנראו לנו מתקדמים ומבוססים. אך עם התקדמות הפרויקט, למדנו לזהות את היתרונות של מודלים "מתחרים" כמו GRU או DistilBERT גם אם לא הכרנו אותם לעומק לפני כן.

למשל, גילינו שה-GRU מאפשר ניתוח של נתוני זמן משולבים עם עמודות סנטימנט - ובכך אפשר לנו לוותר לחלוטין על שימוש במודל שפה נפרדת. ההחלטה לאחד תהליכים תחת מודל אחד הייתה תוצאה של קריאת פלטים, ניתוח ביקורתי של הביצועים, והבנה מה באמת עובד - ולא מה תוכן מראש.

המעבר לתחזית שבועית (במקום יומית בלבד), למשל, גם הוא לא נבע מהתכנון המקורי אלא מהתבוננות מושכלת על תועלת הפלט למשתמש הקצה. תובנה זו מוכיחה לנו כי נדרשת גמישות אמיתית במהלך כל שלבי העבודה - תכנון, הרצה וניתוח.

איזון בין ניסויים לאיטרציות יתר

נקודה חשובה שלמדנו קשורה לניהול קצב ההרצות. כחוקרים מתחילים, חוונו תחושת חוסר ביטחון בהרצות ראשוניות שהניבו תוצאות לא מספקות. אך עם הזמן הבנו כי אין טעם לרוץ לאינספור איטרציות ללא תכלית - ושדווקא איטרציה אחת עם תכנון מדויק עשויה להניב תובנה טובה יותר מעשר הרצות לא ממוקדות.

למדנו לזהות מתי נכון לעצור, לנתח תוצאה, ולחשוב מחדש לפני ריצה נוספת. שמירה על איזון בין בדיקה, המתנה, ולמידה מניבה תהליך מדויק, חסכוני, ובעל ערך אמיתי.

סיכום התהליך

העבודה על פרויקט הגמר שלנו הייתה בית ספר מעשי לעקרונות מפתח בעבודה דאטה-מדעית נכונה. מתוך ההתנסות נבנו לנו חמישה עקרונות להמשך הדרך:

1. הגדרת צרכים מוקדמת וברורה של המודל הרצוי.
 2. הבנה עמוקה של איכות וערך העמודות- לא רק קיומן.
 3. נכונות לבחון מודלים שאינם מוכרים ולבחור את הטוב ביותר בפועל, לא על פי מוניטין.
 4. גמישות מחשבתית להגיב לפלטים ולשנות כיוון אם הנתונים מצדיקים זאת.
 5. שליטה בקצב ובכמות האיטרציות- לחפש איכות ולא כמות.
- בזכות פרויקט זה, נרכשו כלים ותובנות שיילוו אותנו בהמשך דרכנו, ויבטיחו שכל פרויקט עתידי יתנהל בצורה מושכלת, יצירתית ויעילה יותר.