Deep Learning

ניתוח סנטימנט ביטקוין

רועי ויונטה, מור סיגמן

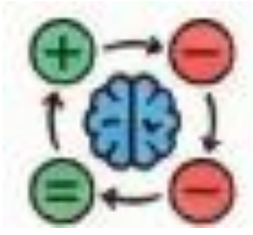
סמסטר א', תשפ"ו

עידן טוביס

מכון טכנולוגי חולון
Holon Institute of Technology

הקדמה

חיזוי מחיר הביטקוין

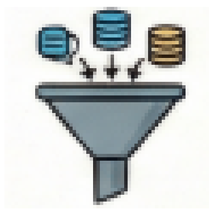


ניתוח סנטימנט

Twitter/X

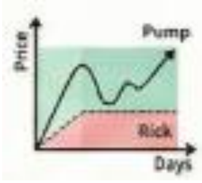
פיתוח מערכת מבוססת Deep Learning לניתוח סנטימנט מרשתות חברתיות (X/Twitter) ככלי לזיהוי מוקדם של מצבי סיכון או הזדמנות בשוק הביטקוין.

שיטה ומודלים



איסוף והבנת הנתונים

66,000 ציוצים ונתוני מחיר Yahoo



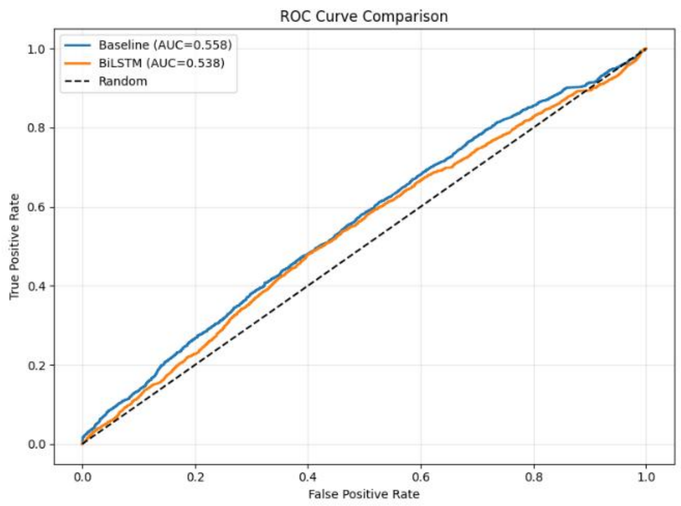
חיזוי r30

הגדרת r30 תשואה בטווח של 30 יום מרגע פרסום הציוץ

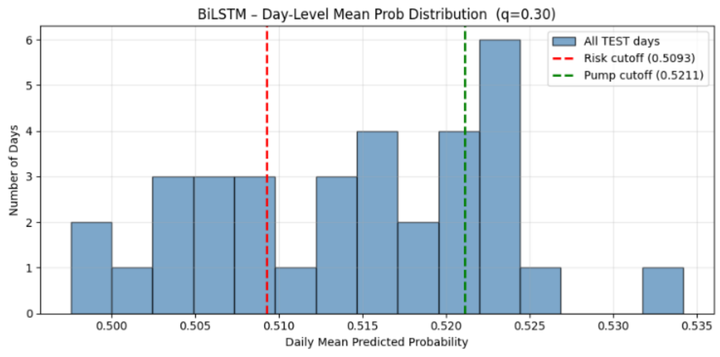
חלוקה טמפורלית

למניעת דליפת מידע לפי זמן כדי להבטיח אמינות בחיזוי.

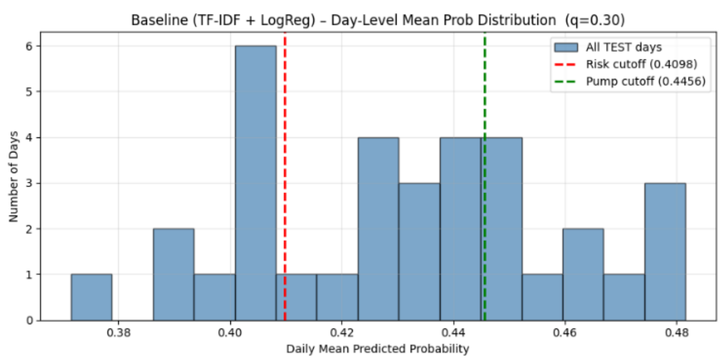
תוצאות



ברמת הציוץ הבודד, הרעש רב מידי. קשה להבחין בין מודלים באמצעות מדדים יבשים.



Metric	Baseline (TF-IDF)	BiLSTM (Deep Learning)
Spread (Pump-Risk)	32.55%	12.16%
p-value	0.0002 (מובהק)	0.1185 (לא מובהק)



Event Study - "שור שוויון" צבירת סנטימנט לרמת ימים, חשיפת סינגל מובהק במודל ה Baseline בלבד.

בדיקת יציבות

הרצת Pipeline על דאטה גדול: 330,000 ציוצים (1500 ביום) תוך שמירה על מתודולוגיה זהה.

הגדלת הדאטה שיפרה יציבות, אבל לא שינתה את המסקנה המרכזית — הסיגנל מהטקסט קצר מוגבל ולכן נדרש מודל/דאטה עשיר יותר כדי לראות שיפור משמעותי.

מסקנות

מניעת דליפת מידע (Data Leakage)

הסיגנל בטקסט חלש ברמת ציוץ, אבל מתגלה ברמת יום

השוואה הוגנת Baseline חזקה מוכיחה יושרה מדעית

בדיקת יציבות על דאטה גדול מחזקת שהממצאים לא מקריים

סיכום:

בנינו מערכת Deep Learning עם מתודולוגיה נקייה נגד דליפת מידע, וגילינו שהסיגנל בציוצים חלש ברמת ציוץ בודד אבל מתגלה יותר טוב ברמת ימים דרך — Event Study והבדיקה על K 330 ציוצים הראתה שהתמונה יציבה והמסקנות לא מקריות