ניתוח וניטור מידע כוזב ברשת החברתית טוויטר במהלך משבר הקורונה



מציגים: יקיר אטיאס ומנגשה אטלאיי מנחה: ד״ר אבי יוסיפוף מחלקה: מערכות מידע

רקע ומצב נוכחי:

שנה: תשפ״ד

בעידן הרשתות החברתיות, מידע כוזב יכול להתפשט ברשת בשניות ובמקרי חירום ליצור נזקים בלתי הפיכים.

מוקד הפרויקט שלנו הוא מידע שקרי לגבי תיאוריית הקונספירציה ה-5G ברשת החברתית טוויטר במהלך תקופת הקורונה.

הגדרת הבעיה:

ריבוי תיאוריות הקונספירציה בטוויטר עשוי להקשות על אנשים להבחין בין מידע מדויק ולא מדויק, מה שעלול להוביל לבלבול וחוסר אמון במקורות מידע רשמיים.

בנוסף, להפצה של תיאוריות קונספירציה בטוויטר בזמן חירום יכולה להיות גם השפעות חברתיות ופסיכולוגיות שליליות.

מטרות הפרוייקט:

מטרת המחקר היא להבין את עומק הדיון והפצה של מידע כוזב בנושא ה-5G במדינות אירופה וישראל.

מחקר זה מתאר תוכנית לניתוח ולניטור מידע כוזב ותדירותו, בעניין תאוריית ה-5G בטוויטר, בתקופת הקורונה.







<u>יישום מערכת:</u> פיתוח מערכת מבוססת ענן להצגת ממצאי המחקר. המערכת בעלת ממשק ידידותי למשתמש היכולה לשמש קובעי מדיניות, חוקרים, עיתונאים וכלל הציבור.

שיטות ועיבוד נתונים:

הפרויקט שלנו משלב שיטות חישוביות וסטטיסטיות שונות כדי לנתח ולפרש נתוני טוויטר הקשורים לתיאוריית ה-5G, ולהשוות את השיח בין שפות שונות באירופה.

להלן הצעדים והכלים העיקריים בהם השתמשנו:

- ניתוח טקסט: ביצענו ניתוח טקסט כדי לזהות את המילים והביטויים הנפוצים ביותר הקשורים ל5G.
- ,NLTK **ניתוח סנטימנט (רגש):** באמצעות כלי ויידר מספריית קבענו את הרגש (חיובי, שלילי או ניטרלי) של כל ציוץ.
- ניתוח סדרות זמן: השתמשנו במודל ARIMA כדי לנתח את נפח הציוצים לאורך זמן. זה עזר לנו להבין את המגמות והדפוסים בדיון בכל תיאוריית קונספירציה.
- עיבוד שפה: השתמשנו ב-BERT, מודל עיבוד שפה מתוחכם, כדי לסווג את הסנטימנט של ציוצים.
- ניתוח מתאם: חישבנו את המתאם בין נפחי הציוצים של שפות שונות.

ניתוח הנתונים

הגדרת יעדי מחקר וניסוח

הגדרת יעדים ברורים למחקר

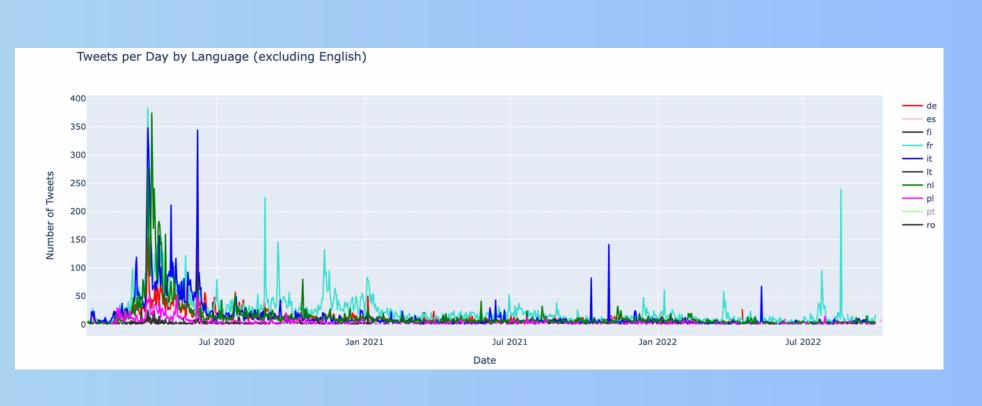
וניסוח שאילתות לאחזור ציוצים

הקשורים ל-5G מטוויטר בשפות

<u>שאילתות:</u>

שונות.

כמות הציוצים בכל יום בתקופת המגיפה:

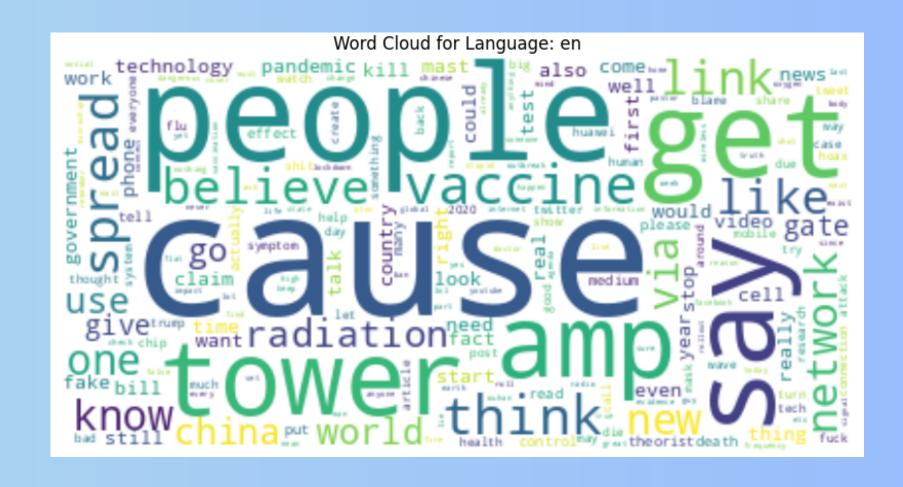


<u>איסוף נתונים והכנה:</u>

לאחסן אותם במסד נתונים ייעודי

ולנקות את הנתונים כדי להתכונן

ענן המילים המדוברות ביותר באנגלית:



התפלגות עמדות האוכלוסייה לגבי ה-50 בשפה האנגלית (BERT MODEL):



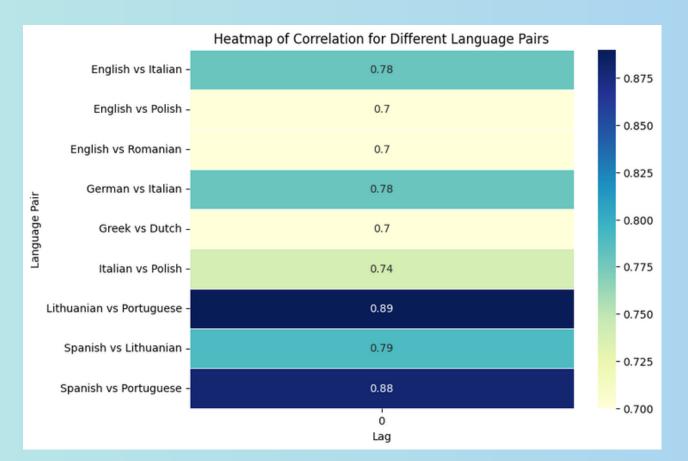
הקורלציות הגבוהות ביותר בין השפות בתקופה אפס:

ניתוח רשת וטקסט. בניית מודלי

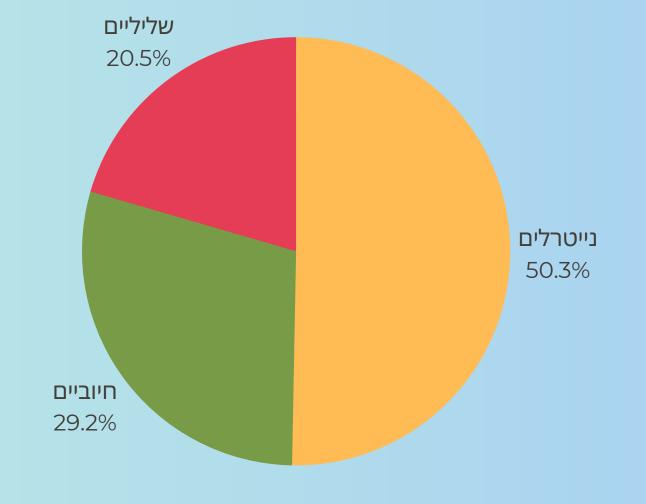
התורמים להפצת מידע מוטעה

חיזוי לזיהוי גורמי מפתח

בזמן חירום.



התפלגות רגישות הציוצים בכל השפות:



תוצאות ומסקנות:

גילינו קורלציה אוטומטית חיובית בין שפות, כלומר, עוצמת הציוצים הדנים בתיאוריית ה-5G נוטה להיות מושפעת מהדיונים של היום הקודם.

פעולה עתידית פוטנציאלית: פיתוח מערכת ניטור בזמן אמת שמתריעה לארגוני בריאות הציבור ולפלטפורמות המדיה כאשר מתרחשת עלייה חדה בדיונים בתיאוריות הקונספירציה. זה יאפשר התערבות בזמן.

גיוון בין שפות: ראינו שתרבות הדיון ותכיפות השיח משתנים בין שפה לשפה.

בנוסף, חלק מהשפות (לדוגמא- הולנדית) הציגו תגובות מאוחרות, מצביע על כך ששפות מסוימות עשויות לעכב תגובות אסטרטגיות לאירועים או דיונים גלובליים.

פעולה פוטנציאלית: פיתוח אסטרטגיות התערבות מותאמות המתחשבות בהבדלים הלשוניים והתרבותיים. זה יכול להיות כרוך בשיתוף פעולה עם גורמים משפיעים או ארגונים מקומיים כדי לספק מידע מדויק ורלוונטי מבחינה תרבותית.

סיווג רגשות עם BERT: על ידי שימוש ב- BERT לסיווג רגשות כלפי הקונספירציה, השגנו הבנה עמוקה ומדוייקת יותר של עמדת הציבור בזמן אמת.

פעולה פוטנציאלית: ניטור <u>קבוע</u> של רגשות כלפי תיאוריות קונספירציה יכול לעזור להעריך את האפקטיביות של מסעות פרסום לציבור ולהתאים אסטרטגיות בהתאם, ובכך ליצור מסרים יעילים יותר לבריאות הציבור.











iii plotly Dash