מערכת זיהוי פייק ניוז בטיקטוק

<u>דוח אפיון - פרויקט גמר תשפ״ה</u>

24.11.2024

<u>שם הארגון המבקש:</u>

IMT School for Advanced Studies Lucca, Italy

שם מנחה:

מר נדב וולך

<u>: מגישים</u>

יונתן כהן - 207906777 גל סבן - 313133084



תוכן עניינים:

מטרת הפרויקט	3-4
ספרות	5-11
מפרט	12.
הצעת פתרונות /הגדרת דרישות	12-13
סיכום	14
לוח זמנים של הפרויקט (גאנט)	15-17.
רשימת מקורות	18

1. מטרת הפרויקט

בשנים האחרונות, התופעה של פייק ניוז ברשתות חברתיות הפכה לאחד הנושאים המדוברים ביותר. קיימים כיום מחקרים רבים וכלים שונים המתמקדים בהתמודדות עם הבעיה, אך עדיין קיים צורך משמעותי במנגנון אפקטיבי שיזהה פייק ניוז ומשתמשים בעייתיים המפיצים אותם. המערכת המוצעת תתבסס על עיבוד ותמלול תוכן וידאו (בהתאם לפורמט בטיקטוק), יחד עם ניתוח שפה טבעית (NLP) וכלי למידת מכונה (ML). כמו כן, היא תשתמש בפרמטרים שמספקת פלטפורמת טיקטוק לזיהוי משתמשים בעלי פוטנציאל בעייתי ומקורות מידע המפיצים תוכן מטעה או שגוי.

מטרת הפרויקט היא לפתח מערכת לזיהוי תוכן פייק ניוז ומשתמשים שמפיצים אותו בטיקטוק. המערכת תתבסס על עיבוד ותמלול וידאו ושימוש בעיבוד שפה טבעית (NLP) .

NLP-היא עיבוד שפה טבעית בתחום מדעי המחשב ובינה מלאכותית שמתמקד באינטראקציה בין מחשבים לבין שפות טבעיות .

נעזר גם בלמידת מכונה כדי לאפשר למערכות מחשב להבין ולעבד את השפה האנושית בצורה שתאפשר להם לבצע משימות מסויימות בעזרת שימוש במודלים מתמטיים וסטטיסטיים.

1.1התוצר הסופי של הפרויקט:

המערכת תוכל לעבד ולהבין תוכן וידאו , לזהות דפוסים ושפה שעלולים להעיד על פייק ניוז , לספק ניתוח ודירוג לתוכן משתמשים.

היא תכלול ממשק שבו ניתן להזין סרטונים מטיקטוק ולקבל ניתוח על רמת האמינות של התוכן והסיכוי שהמשתמש מפיץ מידע שגוי .

המערכת תשלב שלוש טכנולוגיות עיקריות:

עיבוד ותמלול וידאו: ניתוח הסרטון יתחיל בהמרת תוכן וידאו לטקסט. ניתן להשתמש בכלים כמו Google Speech-to-Text, המאפשרים זיהוי דיבור ומספקים תמלול מדויק של אודיו.

עיבוד שפה טבעית (NLP): לאחר תמלול הסרטון, המערכת תעבור על התוכן הטקסטואלי במטרה לזהות דפוסי שפה שיכולים להעיד על מידע שקרי, מניפולטיבי או חדשות חשודות. לדוגמה, ניתוח של טקסטים לפי קטגוריות כמו סנטימנט (רגש חיובי/שלילי), זיהוי סגנונות רטוריים בעייתיים (כמו אזהרות או טענות קונספירטיביות), והשוואה מול מקורות מידע אמינים.

למידת מכונה: על מנת לשפר את הדיוק, המערכת יכולה לכלול מודלים למידת מכונה. אלה יוכלו להכשיר את המערכת לזהות משתמשים עם דפוסים בעייתיים, דוגמת פרופילים שמעלים תוכן בעייתי או משתמשים בתכנים מסוימים שוב ושוב.

1.2נחיצות ודרישות:

עם הצמיחה המהירה של טיקטוק כמדיה חברתית מובילה, קיים גידול בכמות הפייק ניוז והמידע המטעה בפלטפורמה.

אנשים רבים מקבלים מידע מהפלטפורמה, ולעיתים מתקשים להבחין בין אמת לשקר. המערכת תסייע ביצירת סביבה בטוחה יותר, שבה משתמשים יכולים לצפות בתכנים עם רמת אמינות גבוהה.

הכלי יכול לסייע גם למפעילי פלטפורמות, מגופים ממשלתיים ועד ארגונים חברתיים, בזיהוי והסרה של תכנים מסוכנים.

2. רקע וסקירת ספרות:

בשנים האחרונות, תופעת הפייק ניוז (Fake News) הפכה לבעיה גלובלית משמעותית, עם השפעה רחבה על תחומים שונים בחיי היומיום, כולל פוליטיקה, בריאות הציבור, וכלכלה. הפצת חדשות שקריות מתבצעת בעיקר דרך הרשתות החברתיות, שהפכו למקור מרכזי לצריכת מידע בעידן הדיגיטלי. פלטפורמות כמו Facebook, Twitter, TikTok ו-WhatsApp משמשות ככלי רב עוצמה לשיתוף מידע, אך בה בעת מאפשרות הפצה מהירה של תוכן מטעה ומסולף.

הפייק ניוז אינו רק בעיה של מידע כוזב – הוא משפיע ישירות על עמדות הציבור ועל יכולתם לקבל החלטות מושכלות. לדוגמה, במהלך מערכות בחירות, חדשות כוזבות על מועמדים או נושאים פוליטיים מסוימים עלולות לשנות את דפוסי ההצבעה. בתחום הבריאות, מידע כוזב על חיסונים או טיפולים רפואיים תרם לאי-אמון במערכות הבריאות, ולעיתים אף גרם לנזק ממשי לבריאות הציבור.

יתר על כן, הפייק ניוז מערער את יציבות הממשלות ואת אמון האזרחים במוסדות ציבוריים. בזירה הפוליטית, לדוגמה, נעשה שימוש שיטתי בפייק ניוז כדי לקדם אג'נדות מסוימות או לפגוע ביריבים פוליטיים. בעידן של מהירות שיתוף מידע, קשה מאוד לעקוב אחרי מקורות שקריים ולהפריכם בזמן, מה שמוביל לנזק מיידי ולתפוצה רחבה של תוכן מסולף.

מול אתגר זה, נוצר צורך בכלים טכנולוגיים מתקדמים, כמו מנועי זיהוי אוטומטיים ומודלים של בינה מלאכותית, כדי להתמודד עם הבעיה ולהפריך את המידע הכוזב במהירות. לצד הטכנולוגיה, ישנה חשיבות רבה בחינוך הציבור לקרוא ביקורתית ולהבחין בין מידע אמין לשקרי.

ביצועים/ידע בתחום ופתרונות קיימים :

-Claim Buster.1

הוא מערכת אוטומטית לאימות עובדות, המשלבת למידת מכונה, עיבוד שפה טבעית ושאילתות למסדי נתונים. המערכת נועדה לסייע בזיהוי טענות הדורשות בדיקה והצלבה עם מידע אמיתי במטרה להילחם במידע שקרי (פייק ניוז). היא תומכת בתמלול וידאו על מנת לזהות טענות אמיתיות או שקריות בדיונים חיים, נאומים ותוכניות טלוויזיה.

:איך זה עובד אלגוריתמית

- Claim Monitor.1- עוקב אחרי מקורות כמו רשתות חברתיות, תוכניות טלוויזיה ואתרים, ומספק טקסטים לבחינה.
- עבור וידאו, Claim Monitor משתמש בתמלול אוטומטי (כגון תמלול כתוביות של תכניות טלוויזיה או דיונים פוליטיים) כדי לייצר טקסט שניתן לנתח אותו.
- Claim Spotter.2- נותן ציון לכל משפט, המצביע על הסבירות שהוא מכיל טענה עובדתית שדורשת -בדיקה.
 - Claim Matcher.3- משווה טענות חדשות עם טענות שכבר נבדקו על ידי אנשי מקצוע (כגון -Claim Matcher.3).
 - שולח שאלות למאגרי מידע חיצוניים כדי לאמת את נכונות הטענות. -Claim Checker.4
 - -Fact-check Reporter.5 מציג את תוצאות האימות למשתמשים, דרך אתרי אינטרנט, טוויטר או API.

[1]

Azure Text Analytics-I Azure Video Indexer.2

Azure Video Indexer •

כלי לניתוח תוכן וידאו המפיק אוטומטית שמע מהוידאו, מבצע תמלול דיבור לטקסט (Speech-to-Text) וזיהוי ישויות כמו שמות, מקומות ותאריכים. מתאים לשליפת מידע טקסטואלי מנתוני וידאו מורכבים.

<u>: אלגוריתם</u>

- שליפת שמע: כלי כמו FFmpeg או Azure Media Services מפרידים את האודיו מהוידאו.
 - Deep Neural Networks תמלול דיבור לטקסט: המערכת משתמשת במודלים של
 Whisper ובמודלים טרנדיים כמו (DNNs)
 - זיהוי ישויות: מבוצע באמצעות Named Entity Recognition (NER) על הטקסט המופק.

Azure Text Analytics •

שירות מתקדם לעיבוד טקסט שמבצע ניתוח רגשות (Sentiment Analysis), זיהוי ישויות (עיבוד טקסט שמבצע ניתוח רגשות (Language Detection), ומיצוי ביטויים (Key Phrase Extraction). מתאים לזיהוי דפוסי שפה ולזיהוי טענות שעשויות להיות פייק ניוז.

: אלגוריתם

- עיבוד ראשוני: הטקסט מפורק למילים (Tokenization) ומומר לייצוג מספרי (Embeddings).
- ניתוח סמנטי: מודלים כמו Transformers מנתחים הקשרים בין מילים ומבצעים סיווג רגשות, זיהוי ישויות, והפקת תובנות.
 - פלט מנותח: תוצאה מוחזרת כ-JSON עם ביטויים מרכזיים, ישויות, ורגשות.

ספריות ושפות תכנות<u> רלוונטיות:</u>

שפות תכנות

- 1. **Python** השפה המובילה ב-NLP.
- .2 לשילוב ביישומי ווב. JavaScript/Node.is
 - .Azure SDK לשימוש אינטגרלי עם **C#** .3
 - . Java לפיתוח יישומים רחבים.

Python ספריות

- תמלול) Azure Speech Services לעבודה עם **azure-cognitive services** .1 דיבור).
 - .Azure Text Analytics לניתוח טקסט באמצעות azure-ai-text analytics .2
 - .3 **FFmpeg** לעיבוד וידאו ושליפת שמע.
 - . אישית מותאם אישית אישית spaCy .4
 - .NLP להרצת מודלים מתקדמים של Hugging Face Transformers .5
 - .6 או PyTorch לפיתוח מודלים מותאמים לאימות פייק ניוז.
 - .7 pandas לעיבוד וניתוח נתונים.

<u>JavaScript ספריות</u>

- .Azure-ב לניתוח טקסט ב @azure/ai-text-analytics .1
- Azure לתמלול דיבור באמצעות **@azure/cognitiveservices-speech-sdk** .2 .Speech
 - .3 לטיפול בקבצי וידאו ושמע. Fluent-ffmpeg

[2]

:(speak AI) דבר בינה מלאכותית . 3

כלי מתקדם מבוסס בינה מלאכותית לתמלול, ניתוח וניהול נתוני אודיו ווידאו. הכלי מאפשר להעלות הקלטות או להקליט ישירות באפליקציה, ליצור מכשירי אודיו וידאו מותאמים אישית, ולנתח מידע חשוב מתוך שיחות, ראיונות וסרטונים.

בין היתר, המערכת מספקת תמלול אוטומטי, זיהוי מילות מפתח, נושאים, ומגמות סנטימנט, לצד הפקת דוחות ניתוח מתקדמים. Speak מסייע לארגונים לנהל נתונים, לשתף תובנות ולבנות מאגרי מדיה אישיים הניתנים לשיתוף.

<u>תכונות עיקריות:</u>

- זיהוי ישויות ושמות.
- חיפוש מעמיק במידע. •
- .API אינטגרציות עם ממשקי
 - . ניהול מדיה יעיל.
- הפקת דוחות ולכידת אודיו וידאו בקנה מידה.

כדי לעבוד עם Speak Al באמצעות Python, ניתן להשתמש בספריות ובכלים הבאים:

- בקשות (Requests): ספרייה זו מאפשרת שליחת בקשות HTTP, מה שמאפשר
 אינטראקציה עם ממשקי ה-API של API.
- pandas: ספרייה לעיבוד וניתוח נתונים, המאפשרת ניהול נתונים המתקבלים מ-Speak Al.
 - numpy: ספרייה לחישובים מתמטיים, שימושית לעיבוד נתונים מספריים.
 - שודול מובנה בפייתון לעבודה עם נתוני JSON, הפורמט הנפוץ להחזרת נתונים מ- json:
 API.
- logging: מודול מובנה לניהול לוגים, המסייע במעקב אחר פעולות ותקלות במהלך העבודה עם ה-API.

בנוסף, Speak AI מספקת תבנית קוד פתוח ב-GitHub בשם "Speak-python-boilerplate", המאפשרת יצירת מיקרו-שירותים בסביבת Python. תבנית זו כוללת קבצים ותבניות מוגדרות מראש ליצירת נקודות קצה של Flask API ופריסה באמצעות Docker , שימוש בתבנית זו יכול להקל על תהליך הפיתוח והאינטגרציה עם Speak AI.

[3], [4]

. כלים טכנולוגיים לווידוא עובדות (Fact-Checking Tools):

כדי להתמודד עם התפשטות של פייק ניוז ברשתות החברתיות ובאינטרנט, פותחו כלים וטכנולוגיות המאפשרים לאמת את המידע בצורה אוטומטית וממוקדת:

● Google Fact Check Tools: פלטפורמת גוגל המאפשרת חיפושים על חדשות שקריות:
 [5] או הונאות ברשת, ומספקת גישה למידע מהימן ממסדי נתונים של בדיקות עובדות.

מה הכלי עושה?

- 1. חיפוש טענות שנבדקו: מאפשר לחפש טענות ולראות אם נבדקו על ידי ארגוני בדיקת עובדות כמו FactCheck.org או Snopes.
 - 2. הצגת מקורות אמינים: מציג את מקור הטענה, גוף בדיקת העובדות, והמסקנה (למשל, "נכון", "שקרי").
 - 3. שימוש בתקן **ClaimReview:** הכלי מתבסס על נתונים שסומנו על ידי ארגוני בדיקה כטענות שנבדקו.

אלגוריתם:

- משתמש באלגוריתמי חיפוש של גוגל ובמאגרי מידע מרכזיים.
- מציג תוצאות מתויגות מבדיקות עובדות שנעשו על ידי גופים מוסמכים.
 - . קל לשימוש מזינים טענה או נושא ומקבלים תוצאות רלוונטיות. ●
- TinEye: מנוע חיפוש הפוך לזיהוי מקור של תמונות, ובכך מאפשר לזהות אם התמונה
 שוכתבה או שוחררה בהקשר שקרי. [6]

מה הכלי עושה?

- 1. זיהוי מקור התמונה:
- עוזר למצוא את המקור המקורי של תמונה שהועלתה.
 - 2. גילוי שימושים חוזרים:
- מזהה אם התמונה פורסמה שוב באתרים אחרים, כולל שינויים שנעשו בה.
 - חשיפת הקשר שקרי:
 בוחן אם התמונה נערכה, הוצאה מהקשרה, או שוכתבה למטרות מטעות.

אלגוריתם:

Matching Algorithm: •

TinEye משתמש באלגוריתמים להשוואת חתימות דיגיטליות של תמונות, כולל תבניות צבעים, טקסטורות ומאפיינים ויזואליים.

- מאגר נתונים: הכלי מבצע חיפוש במסדי נתונים עצומים של תמונות ברחבי
 האינטרנט.
- זיהוי שינויים:מנתח וריאציות של אותה תמונה (שינוי גודל, עריכה, חיתוך).

של את אמינותם של :Whois כלי לזיהוי מידע על דומיינים ומעקב אחר אתרים, המסייע לוודא את אמינותם של :אתרים

● FactCheck.org: מאגרי בדיקות עובדות שנבדקים ומעדכנים באופן שוטף: Snopes ו- PactCheck.org: חדשות ומידע חשוד, המסייעים למנוע את התפשטות פייק ניוז.

המערכות הצפויות לצמוח:

התחום נמצא בתהליכי פיתוח מתקדמים, וישנן מערכות שמבוססות על חיבור בין טכנולוגיות מתקדמות ואלגוריתמים למערכות אנושיות.

לדוגמה, מאגרי מידע גלובליים המנטרים את הפצת פייק ניוז באמצעות מנועי חיפוש חכמים, מערכות ניתוח שפה טבעית ואלגוריתמים של למידת מכונה, שמטרתן להעלות את המודעות לפייק ניוז ולספק כלים לבדיקת אמיתות המידע.

שימוש במערכת:

המערכת מספקת כלים לאוטומטיזציה של תהליך בדיקת העובדות ומאפשרת זיהוי ומענה מיידי של טענות שקריות או מטעות, כולל טענות שנאמרות בוידאו. היא מאפשרת לתמלל את הדיבור ולבצע את האימות במהירות, גם עבור תוכן חי, כמו דיונים פוליטיים, נאומים וראיונות.

: מפרט 3

דרישות פונקציונליות:

- המערכת תתממשק לאתר אינטרנט מסוים אותו ניצור לטובת הלקוח.
- המערכת צריכה לאפשר למשתמש לבצע התחברות עם שם משתמש וסיסמה.
- על הלקוח להכניס סרטון וידאו מהרשת החברתית טיקטוק למערכת, לעיבוד הנתונים.
 - לאחר הכנסת הסרטון הלקוח יקבל הודעה שהסרטון נקלט בהצלחה.
 - . במידה והוידאו לא נקלט תוצג הודעת שגיאה.
- במידה והוידאו נקלט בהצלחה המערכת תציג הודעה על מידע מהימן, והלקוח יקבל פירוטאם המידע שהתקבל באמצעות תמלול הוידאו שהתבצעה במערכת הוא אמת או שקר.
 - הלקוח יוכל לצפות בנתונים סטטיסטים אודות המידע על סמך הסתברויות שונות.

דרישות לא פונקציונאליות:

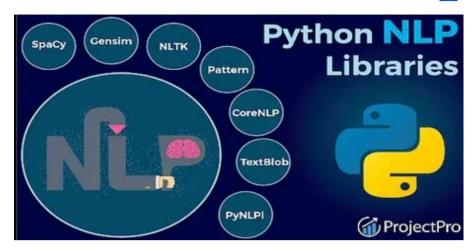
- לטובת בניית האתר נבנה קובץ html וניצור ממשק ידידותי ללקוח , בנוסף נשתמש בעיצוב css . סביבת עבודה תהיה
 - בניית צד שרת לשמירת מאגרי המידע וכתיבת הקוד.
- ◆ לאחר הכנסת הסרטון למערכת הוא יתקבל באמצעות פייתון , נחלץ את האודיו מהוידאו . על מנת לעבד את קבצי הוידאו נשתמש בספרייה VideoFileClip (ספרייה של פייתון).
 - : לאחר חילוץ קובץ אודיו נשתמש בספריות תמלול , אחת מהן נקראת . SpeechRecognition
 - . ספרייה זו היא מהירה ויעילה, למשל לסווג טקסט ועוד.
- בתום תהליך התמלול כל המילים והמשפטים שנקלטו באודיו הופכים למחרוזת אחת או יותר
 - לאחר שנעשה התמלול טקסט , נשתמש באחד הספריות NLP לנתח ולעבד את הטקסט. spaCy : ספרייה שימושית בתחום זה
- נאמן מודל ML ובו נשתמש בטקסט שקיבלנו על מנת לסווג נתונים שקריים / אמיתיים. לשם כך נבצע אימון של המודל ולבסוף נעשה הערכה סטטיסטית.

: ספריות לדוגמא

import pandas as pd

- from sklearn.feature_extraction.text import TfidfVectorizer
 - from sklearn.model_selection import train_test_split •
 - from sklearn.ensemble import RandomForestClassifier •
- from sklearn.metrics import classification report, accuracy score

[9]



<u>לסיכום :</u>

- . נבצע תמלול וידאו בזמן אמת באמצעות כלי מתקדם שניצור להמרת דיבור וידאו
- ננתח שפה טבעית NLP לזיהוי דפוסים בעייתיים בטקסט, כולל זיהוי סנטימנט ומניפולציות רטוריות.
 - למידת מכונה **ML** לזיהוי דפוסים של משתמשים בעייתיים, כולל שימוש ב-Clustering . Classification
- .FactCheck.org-ו ClaimBuster ו-FactCheck.org.
 - ננטר בזמן אמת של סרטונים וטענות ממאגרי טיקטוק על מנת להתריע ולסמן תוכן בעייתי.
 - ניצור ממשק משתמש ידידותי שמציג את הדירוגים של אמינות התוכן.

מערכת כזו תוכל לשפר באופן ניכר את זיהוי הפייק ניוז, במיוחד בפלטפורמות דינמיות כמו טיקטוק, ולהפוך את הרשתות החברתיות למקום יותר בטוח ומבוקר. הלקוח יוכל בקלות לאתר האם המידע אמין ותהיה לו אפשרות להשתמש במידע ככלי מה שמביא יתרון גדול למשל ללקוחות/אנשים מסוימים שרוצים למשל לבדוק את האמינות של המידע שהם צורכים והאם זה רלוונטי להם.

דרישות כלים טכנולוגיים בסיסיים לבניית הפרויקט:

- על מנת ליצור עבודה נוחה יותר במהלך ביצוע python + oop למידת שפת הפרויקט.
 - למידת שפת מכונה ML
 - למידת כלים הקשורים לפלטפורמת הרשת טיקטוק.
 - למידה כיצד ניתן לעבד ולתמלל וידאו.
 - . בסיומו Git-hub בסיומו •



[8] תמונה המציגה את הממשקים השייכים לשימוש בשפה טבעית.

4. לוח זמנים פרויקט (גאנט):

משימות	פירוט המשימה	תאריך התחלה משוער	מספר ימים נדרשים	תאריך סיום משוער
דוח אפיון	כתיבת סקריה מפורטת על דרישות הפרויקט , רקע וסקירת ספרות , קביעת פגישה עם המרצה ותיאום ציפיות	07/11/2024	17	24/11/2024
	המו צה ווניאום ציפיות הפרויקט			
תכנון ועיצוב המערכת:	הבנת דרישות המערכת.	25/11/2024	20	15/12/2024
בניית ממשק לקוח (Frontend):	יצירת קבצי HTML עם מבנה דפי האתר. הוספת עיצוב ותמיכה באמצעות CSS. יצירת ממשק ידידותי ואינטראקטיבי בסביבת העבודה VSCode בנוסף פגישה עם המנחה	16/12/2024	17	02/01/2025
פיתוח צד שרת (Backend):	תכנון מבנה מאגרי המידע. בניית API לניהול התקשורת בין הלקוח לשרת. חיבור המאגר לממשק באמצעות כתיבת קוד צד שרת.	03/01/2025	31	03/02/2025
עיבוד וידאו:	הכנסת סרטונים למערכת. חילוץ אודיו מהווידאו באמצעות ספריית VideoFileClip.	04/02/2025	25	01/03/2025
תמלול אודיו לטקסט:	שימוש בספריית SpeechRecognitio n לתמלול האודיו לטקסט. בדיקות לוודא שהתמלול מדויק ומהיר.	02/03/2025	27	29/03/2025
עיבוד טקסט באמצעות NLP:	עיבוד טקסט בעזרת ספריית spaCy. ביצוע ניתוח תחבירי, זיהוי ישויות, וסיווג תוכן.	30/03/2025	21	20/04/2025

בניית מודל Machine Learning:	תכנון המודל. בחירת אלגוריתם מתאים לסיווג טקסט (Fake/True).	21/04/2025	24	15/05/2025
אימון המודל:	שימוש בנתונים מהתמלול לעיבוד ואימון המודל. שיפור ביצועי המודל .	16/05/2025	16	01/06/2025
הערכת המודל וסטטיסטיקות:	ביצוע הערכה סטטיסטית לתוצאות (מדדים כמו דיוק, רגישות). כתיבת דוח מסכם על ביצועי המודל(אופציונלי).	02/06/2025	43	15/07/2025
בדיקות, אינטגרציה והשקה:	בדיקת המערכת כולה Frontend) ו- Backend). בדיקות קצה לקצה. שיפור חוויית המשתמש והשקת הפרויקט.	16/07/2025	77	01/10/2025

Gantt Chart 07/11/2024 27/12/2024 15/02/2025 06/04/2025 26/05/2025 15/07/2025 03/09/2025 23/10/2025 דוח אפיון תכנון ועיצוב המערכת: בניית ממשק לקוח (Frontend): פיתוח צד שרת (Backend): :עיבוד וידאו :תמלול אודיו לטקסט :NLP עיבוד טקסט באמצעות בניית מודל Machine Learning: :אימון המודל הערכת המודל וסטטיסטיקות: בדיקות, אינטגרציה והשקה: בדיקות, בניית מודל עיבוד טקסט בניית ממשק הערכת המודל וסטטיסטיקות: תמלול אודיו פיתוח צד שרת תכנון ועיצוב Machine אינטגרציה אימון המודל: באמצעות עיבוד וידאו: לקוח דוח אפיון (Backend): לטקסט: המערכת: והשקה: Learning: NLP: (Frontend): 03/01/2025 תאריך התחלה משוער 16/07/2025 02/06/2025 16/05/2025 21/04/2025 30/03/2025 02/03/2025 04/02/2025 16/12/2024 25/11/2024 07/11/2024

21

27

25

31

17

20

17

מספר ימים נדרשים

77

43

16

24

<u>. רשימת מקורות :</u>

- [1] ClaimBuster •
- [2] Azure Video Indexer
 - [3] Speak Al •
 - [4] GitHubLibraries •
- [5] Google Fact Check
 - [6]<u>TinEye</u> •
 - [7] picture1 •
 - [8] picture2 •
 - [9] picture3 •