דו"ח התקדמות – פרויקט גמר  
מערכת לזיהוי פישיניג

**PhishGuard AI**

איל רונן

גל ארז

מנחה: ד"ר אריק פארן

**מטרת הפרויקט**

בשנים האחרונות, התקפות פישינג הפכו לאחד האיומים המשמעותיים ביותר על אבטחת המידע והפרטיות ברשת. תוקפים משתמשים בשיטות מתוחכמות ומתפתחות כדי להטעות קורבנות ולגנוב מהם מידע רגיש, כגון פרטים פיננסיים וזהויות. למרות המודעות הגוברת לנושא, רבים עדיין נופלים קורבן להתקפות אלו, מה שמדגיש את הצורך במנגנונים חדשניים ויעילים יותר לזיהוי ומניעת פישינג.

מטרת הפרויקט PhishGuard AI היא ליצור מערכת מתקדמת, המשלבת טכניקות של בינה מלאכותית ועיבוד שפה טבעית לזיהוי והתמודדות עם איומי פישינג. המערכת תבצע ניתוח של תוכן מיילים והודעות טקסט ותיישם אלגוריתמי למידת מכונה לזיהוי תבניות ומאפיינים חשודים האופייניים להתקפות פישינג.

התוצר הסופי של הפרויקט יהיה כלי אוטומטי ויעיל, המסוגל לסרוק כמויות גדולות של טקסט, להתריע על ניסיונות פישינג פוטנציאליים ולספק המלצות פעולה למשתמשים. בכך, PhishGuard AI תסייע בהפחתה משמעותית של הסיכון לגניבת מידע ונזקים פיננסיים, תוך שיפור ההגנה הכוללת על הפרטיות והביטחון של משתמשים במרחב הדיגיטלי.

**רקע וסקירת ספרות:**

בשנים האחרונות, נזקי הפישינג הפכו למשמעותיים וניכרים יותר מאי פעם.

לפי ,Anti-Phishing Working Group (APWG)ברבעון הרביעי של 2023 נצפו מעל מיליון התקפות פישינג ברחבי העולם. סך הכל בשנת 2023 היו כמעט חמישה מיליון מקרי פישינג, מספר שיא של כל הזמנים. התקפות כנגד פלטפורמות מדיה חברתית הגיעו ל-42.8% מכלל התקפות הפישינג בסוף 2023. מספר מתקפות המבוססות על העברות בנקאיות (BEC) עלה ב-24% ברבעון הרביעי לעומת הרבעון הקודם והסכום הממוצע לניסיון הוא כ-56,195 דולר [1].

דו"ח ה-FBI Internet Crime Complaint Center (IC3) מצביע על כך שבשנת 2021 לבדה, הנזק הכלכלי הנובע מהתקפות פישינג עמד על כ-6.9 מיליארד דולר בארה"ב, גידול של כ-35% ביחס לשנת 2020 [2].

פישינג אינו פוגע רק באנשים פרטיים, אלא גם בחברות וארגונים ברחבי העולם. סקר של חברת הסייבר Proofpoint מצא כי בשנת 2021, כ-74% מהארגונים ברחבי העולם חוו התקפת פישינג מוצלחת, כאשר העלות הממוצעת לארגון עמדה על כ-14.8 מיליון דולר, הכוללת אובדן פרודוקטיביות, קנסות רגולטוריים ופגיעה במוניטין [3].

דו"ח של IBM העריך את העלות הממוצעת של זליגת נתונים הנובעת מהתקפות פישינג בכ-4.96 מיליון דולר בשנת 2022 [4].

נתונים אלה מדגישים את הצורך הדחוף בפתרונות מתקדמים להתמודדות עם איום הפישינג הגובר.

PhishGuard AI מנסה לתת מענה לבעיה זו על ידי פיתוח כלים חדשניים המבוססים על בינה מלאכותית ולמידת מכונה, שיסייעו לזהות ולבלום התקפות פישינג בזמן אמת ולצמצם את הנזקים הפיננסיים והתדמיתיים עבור המשתמשים.

**מפרט ותכנון מפורט של מרכיבי הפרויקט:**

כדי לממש את מערכת PhishGuard AI לזיהוי והתמודדות עם התקפות פישינג בזמן אמת, אנו מתכננים לשלב טכניקות מתקדמות של עיבוד שפה טבעית (NLP) ולמידת מכונה. המערכת תנתח את תוכן המיילים, הודעות טקסט ותיישם אלגוריתמים חכמים לזיהוי תבניות ומאפיינים חשודים האופייניים להתקפות פישינג.

להלן השלבים והאבני דרך העיקריים במימוש המערכת:

1. בניית מאגר נתונים מקיף:

- איסוף דוגמאות מגוונות של הודעות פישינג והודעות אמיתיות.

- סידור, תיוג וניקוי הנתונים כדי להכינם לשימוש באלגוריתמי הלמידה.

- חלוקה לסטים של נתוני אימון, בדיקה ואימות.

2. פיתוח מנגנון לעיבוד שפה טבעית:

- בחירת ספריות NLP מתאימות בשפת Python.

- יישום האלגוריתמים וניתוח תחבירי וסמנטי של טקסט.

- חילוץ מאפיינים לשוניים רלוונטיים, כמו מילות מפתח, דפוסי משפט ורגש.

3. יצירת מודל למידת מכונה לסיווג התקפות פישינג:

- התנסות באלגוריתמים שונים כולל עצי החלטה, SVM, רשתות נוירונים וכו'.

- אימון מודלים על סט הנתונים המוכן ובדיקת הביצועים שלהם.

- כיול פרמטרים והשוואת מודלים לבחירת הגישה המיטבית לסיווג פישינג.

4. בניית ממשק משתמש אינטואיטיבי:

- עיצוב ממשק נוח וידידותי למשתמש.

- הצגה ויזואלית של התוצאות.

- הוספת מדריכי משתמש להסברת תהליך השימוש במערכת.

5. אינטגרציה, בדיקות מקיפות ופיילוט:

- שילוב כלל הרכיבים לכדי מערכת עובדת ותקינה.

- תיקון באגים, שיפורי ביצועים והתייעלות בקוד ובשימושיות.

**גאנט:**

ינואר - (דוח איפיון): הערכת ותיזמון היקף הפרויקט, היעדים והדרישות המפורטות.

פברואר - (לימוד החומרים החדשים הנדרשים לפרויקט): רכישת ידע תיאורטי ומיומנויות טכניות הכרחיות בעולם למידת המכונה ביניהן NLP.

מרץ - (תכנון המערכת): פיתוח ארכיטקטורת המערכת, החלטה כיצד אנו רוצים שתעבוד ובאיזה אופן לאחר הבנת הרקע והתיאוריה.

אפריל - (הפקת נתונים ותיוגם): התחלת איסוף נתונים, ייצור נתונים בעזרת מערכות שונות, תיוג ועיבוד ראשוני לפיתוח מערכת.

אמצע אפריל עד תחילת מאי - חופשת פסח

מאי - (השוואה ולמידה של אלגוריתמים): איסוף נתונים, תיוג ועיבוד הדאטה. בחינת אלגוריתמים ראשוניים, השוואה בין סוגים שונים של אלגוריתמים, מדידת הביצועים שלהם ובחירת האלגוריתמים בהם נשתמש.

יוני – (התחלת פיתוח המערכת): פיתוח סביבה ראשונית ותחילת סיווג המידע בעזרת האלגוריתמים הנבחרים.

יולי – (פיתוח UI ובדיקות): פיתוח ויצירת ממשק משתמש UI והתחלת בדיקות המערכת.

אוגוסט - (סיום פיתוח המערכת): סיום פיתוח מערכת עם שילוב תכונות מלא של האלגוריתמים שנבחרו. עריכת בדיקות מקיפות מול דרישות, חידוד ושיפור הפרויקט על סמך משוב ודיוק התוצאות הרצויות עד להגשה סופית.

תמונה שמכילה טקסט, קו, מספר, גופן

התיאור נוצר באופן אוטומטי

**מקורות:**

1. Anti-Phishing Working Group (APWG) Phishing Activity Trends

<https://docs.apwg.org/reports/apwg_trends_report_q4_2023.pdf>

2. FBI Internet Crime Report 2021. <https://www.ic3.gov/Media/PDF/AnnualReport/2021_IC3Report.pdf>

3. Proofpoint State of the Phish Report 2022. <https://www.proofpoint.com/us/resources/threat-reports/state-of-phish>

4. IBM Cost of a Data Breach Report 2022. <https://www.ibm.com/reports/data-breach>