**מערכת לסיווג הודעות פישיניג – PhishGuard AI**

**דוח איפיון – פרויקט גמר תשפ"ד**

**מנחה:**

ד"ר אריק פארן

**מגישים:**

**תמונה שמכילה ציור, שרטוט, אומנות קליפיפם, איור

התיאור נוצר באופן אוטומטי**

תוכן עניינים

[רקע 3](#_Toc156750448)

[מוטיבציה 3](#_Toc156750449)

[מטרת הפרויקט 3](#_Toc156750450)

[תוצאות ומטרות מבוקשות 3](#_Toc156750451)

[מפרט 4](#_Toc156750452)

[פתרונות קיימים כיום בשוק: 5](#_Toc156750453)

[גאנט 6](#_Toc156750454)

[סקירה ספרותית 7](#_Toc156750455)

# 

# רקע

בשנים האחרונות, התקדמות טכנולוגית והתפתחות בתחום האינטרנט והתקשורת הדיגיטלית יצרו סביבה פורה להתפתחות תופעות הונאה דיגיטליות, בראשן הפישינג. ניסיונות הפישינג מתבצעים בדרכים שונות, כאשר המטרה המרכזית היא להטעות אנשים ולשכנעם לחשוף מידע אישי או פיננסי. חשיבות הפרויקט מתבטאת בכך שהוא פונה לזיהוי ומניעת ניסיונות אלו, תוך שימוש בטכנולוגיות מתקדמות בתחום הבינה המלאכותית ועיבוד שפה טבעית.

# מוטיבציה

בעידן הדיגיטלי, אחד האתגרים המרכזיים בתחום אבטחת המידע הוא התמודדות עם תופעת הפישינג. הפרויקט נובע מהצורך להתמודד עם אתגרים אלו והגנה על משתמשים מפני ניסיונות הונאה דיגיטליים מתוחכמים.

# מטרת הפרויקט

המערכת שאנו מפתחים מיועדת לזיהוי ולמניעת ניסיונות פישינג, למען שיפור אבטחת המידע והגנה על פרטיות המשתמשים. במרכז הפרויקט נפתח מערכת המבוססת על למידת מכונה, עם מטרה לזהות תבניות טקסט ומאפיינים הייחודיים לניסיונות פישינג.

המערכת תבצע אנליזה מעמיקה של מבנה ההודעה, כולל ניתוח שפה ושימוש במילות מפתח, וכן תבצע בדיקות של קישורים ותכנים חשודים. תהליך זה יכלול שימוש באלגוריתמים של למידת מכונה, המאפשרים לזהות ולהבחין בין הודעות תקינות לבין ניסיונות הפישינג.

# תוצאות ומטרות מבוקשות

1. איתור וניתוח של אלמנטים מרמזים:הפרויקט יתמקד באיתור וניתוח של אלמנטים כגון שפת פנייה, סימנים לינגוויסטיים שכיחים בהודעות פישינג, ואסטרטגיות פסיכולוגיות שמתקיפים משתמשים בהם כדי להטעות את הקורבן. דגש יינתן על זיהוי טקסטים וקישורים חשודים, כדי לזהות ככל האפשר את ניסיונות הפישינג.
2. דיוק בזיהוי ניסיונות פישינג: נשאף להשיג דיוק גבוה בזיהוי ניסיונות הפישינג, על מנת למזער את הסיכון לחשיפת מידע רגיש. המערכת תעבוד בשימוש אלגוריתמים שיכולים לזהות את התבניות של מתקפות הפישינג על מנת לסווגן כראוי.
3. מהירות ניתוח ותגובה: אנו מקווים לפתח מערכת שמסוגלת לנתח ולזהות ניסיונות פישינג במהירות, כך שההגנה תהיה יעילה ומיידית.
4. מדדים להערכת הצלחת הפרויקט: הפרויקט יערך לפי מדדים כגון אחוז הזיהוי המדויק של ניסיונות פישינג והיכולת להתמודד עם תרחישים חדשים ולא ידועים מראש.

# מפרט

**הגדרת דרישות והצעת פתרונות:**

דרישות המערכת:

ניתוח טקסט: המערכת תתוכנן לנתח טקסטים של הודעות דוא"ל והודעות טקסט, כדי לזהות אלמנטים של נסיונות פישינג.

ממשק משתמש: פיתוח ממשק נוח ואינטואיטיבי למשתמשים שירוץ בצורה לוקאלית על מחשבים, המאפשר ניתוח והצגה ברורה של הנתונים.

למידת מכונה: שימוש באלגוריתמי למידת מכונה כדי לנתח ולזהות תבניות טקסט של פישינג.

אלגוריתמים:

* מציאת אלגוריתמים ספציפיים לזיהוי פישינג, המתבססים על תכונות שפה וסימנים לינגוויסטיים, והתאמתם למערכת שלנו.
* אלגוריתמים שמבחינים בין הודעות תקינות לבין ניסיונות פישינג, באמצעות למידת מכונה.

# פתרונות קיימים כיום בשוק:

פתרונות מבוססי למידת מכונה ובינה מלאכותית

* **Anti-Phishing Working Group (APWG):**

ארגון בינלאומי המאחד גורמים מתחומים שונים למאבק בפשעי סייבר.

* **Netcraft:**

חברה המציעה שירותים נגד פישינג וכלי סרגל דפדפן לזיהוי אתרי פישינג.

* **PhishLabs:**

חברת ביון נגד איומי סייבר המספקת שירותי זיהוי פישינג.

* **Microsoft's SmartScreen & Google Safe Browsing:**

כלים המשולבים בדפדפנים לסינון תוכן פישינג.

פתרונות מבוססי עיבוד שפה טבעית (NLP)

* **Microsoft's SmartScreen & Google Safe Browsing:**
* משתמשים ב-NLP לזיהוי אתרים ותכנים מזויפים באינטרנט.

פתרונות מבוססי זיהוי תבניות

* **OpenDNS:**

שירות סינון רשת מבוסס ענן המשתמש בזיהוי תבניות ואלגוריתמים לחסימת אתרי פישינג.

* **Symantec's Anti-Phishing Toolbar:**

סרגל כלים לדפדפן המזהה וחוסם אתרי פישינג.

טכנולוגיות אחרונות בשימוש:

* **AI-driven phishing detection:**

פיתוח מודלים מתקדמים לניתוח תוכן דוא"ל והתנהגות משתמשים.

* **Real-time phishing detection:**

כלים מתקדמים לזיהוי וחסימת פישינג בזמן אמת.

* **Blockchain-based solutions:**

פתרונות המבוססים על בלוקצ'יין ליצירת מערכות אימייל מבוזרות ובטוחות.

# גאנט

תמונה שמכילה טקסט, קו, מספר, גופן

התיאור נוצר באופן אוטומטי

**פירוט:**

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): 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.

# סקירה ספרותית

**1. Anti-Phishing Training Effectiveness**:

* **Source**: ["Don’t click: towards an effective anti-phishing training"](https://hcis-journal.springeropen.com/articles/10.1186/s13673-020-00237-7) from Human-centric Computing and Information Sciences, SpringerOpen.
* **Summary**: This article (SpringerOpen, 2020) delves into the effectiveness of anti-phishing training programs. It underscores the necessity of educating users about phishing tactics and offers a comparative analysis of various training methods, emphasizing their role in mitigating phishing risks.

**2. Comprehensive Survey of Phishing Attacks**:

* **Source**: ["A Comprehensive Survey of Phishing Attacks: Characteristics, Detection, and Prevention"](https://www.mdpi.com/2227-7080/11/6/161) from MDPI.
* **Summary**: Providing a thorough survey on phishing attacks (MDPI, 2023), this paper outlines the key characteristics of phishing, detection techniques, and preventive measures. It’s a valuable resource for understanding the intricate details of phishing strategies and the methodologies employed in their detection and prevention.

**3. Current Trends in Phishing**:

* **Source**: [Cybersecurity Magazine.](https://cybersecurity-magazine.com/?s=phishing)
* **Summary**: Cybersecurity Magazine offers a wealth of articles and reports on contemporary phishing trends. It covers the use of IoT devices in phishing, the emergence of ransomware, and deepfake technology, providing a nuanced understanding of modern phishing strategies and their technological evolution.

**4. Integration of NLP in Cybersecurity**:

* **Source**: [The Cyber Express.](https://thecyberexpress.com/?s=phishing)
* **Summary**: The integration of Natural Language Processing (NLP) in cybersecurity, as discussed in The Cyber Express, illustrates the potential of NLP in enhancing threat communication and response. This perspective is crucial for understanding how advanced technologies like AI and NLP can be leveraged in the fight against phishing.