Inteligentny system wykrywania oprogramowania złośliwego

Autorzy:
- Kacper Kurek
- Kacper Skrzypek
- Oskar Wrona

Opiekun pracy:
- dr inż. Mirosław Płaza

Plan prezentacji

1. Cel pracy
2. Motywacja
3. Metodyka przeprowadzonych badań
4. Proponowana metoda
5. Wyniki badań
6. Prezentacja działania
7. Podsumowanie

Cel pracy

Celem przedstawianego do oceny systemu jest wykrywanie złośliwego oprogramowania typu koń trojański, ransomware, spyware, adware, backdoor, rootkit oraz podobnych. System opracowano z wykorzystaniem algorytmów sztucznej inteligencji (AI) i metod uczenia maszynowego (ML).

Motywacja

1. Potencjał AI/ML w cyberbezpieczeństwie

1. AI i ML mają ogromny potencjał w poprawie skuteczności procesów zapewniających bezpieczeństwo.
2. Możliwość szybszego wykrywania i neutralizowania zagrożeń w dynamicznym środowisku cybernetycznym.

2. Motywacja do wdrożenia AI/ML

1. Analiza raportów bezpieczeństwa danych wskazuje na rosnące obszary podatności.
2. Konieczność skutecznej neutralizacji potencjalnych zagrożeń.

3. Aktualne trendy i możliwości

1. Trendy wskazują na duży potencjał zastosowań metod AI/ML w cyberbezpieczeństwie.
2. Narzędzia AI, takie jak TRAM (Threat Report ATT&CK Mapper), automatyzują procesy i analizy danych, identyfikując TTP (Tactics, Techniques, Procedures) z dużą precyzją.

4. Znaczenie automatyzacji

1. Raport CERT Orange podkreśla, że bez skutecznej automatyzacji kontrola zmieniających się zagrożeń byłaby bardzo trudna, a często niemożliwa.
2. Wykorzystanie AI zwiększa skuteczność i przygotowanie na przyszłe wyzwania w cyberbezpieczeństwie.

5. Przykłady narzędzi AI

1. TRAM (Threat Report ATT&CK Mapper) pomaga w automatyzacji procesów i analiz danych.
2. AI identyfikuje TTP (Tactics, Techniques, Procedures) z niezwykłą precyzją.

6. Przyszłość cyberbezpieczeństwa z AI

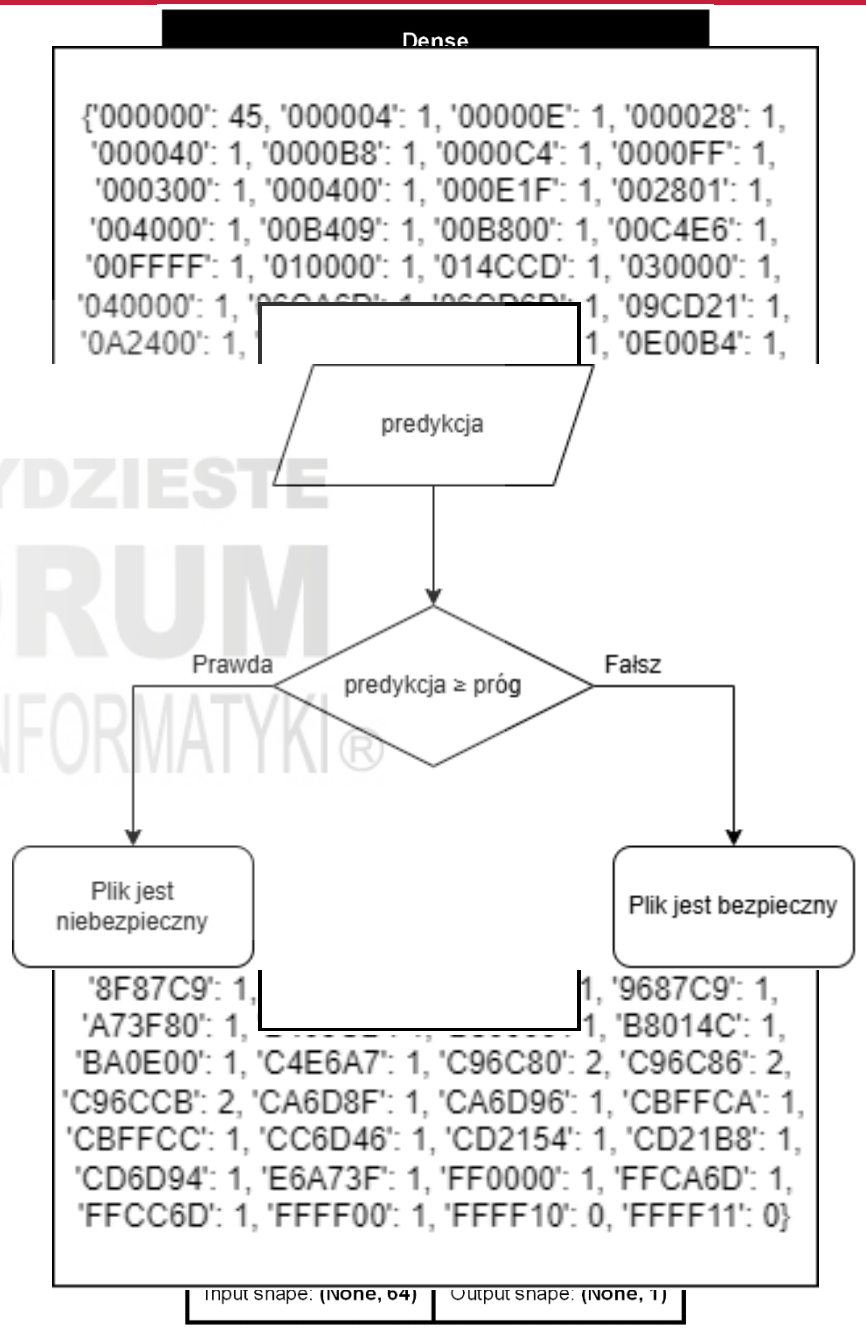
1. AI zwiększa skuteczność i gotowość na przyszłe wyzwania w obszarze cyberbezpieczeństwa.
2. Przedstawiany system będzie mógł być stosowany do skutecznej neutralizacji zagrożeń.

Opracowanie systemu obejmowało następujące etapy prac:

1. Zgromadzenie oraz przetworzenie dużych zbiorów danych
2. Określenie deskryptorów charakteryzujących cechy oprogramowania złośliwego
3. Opracowanie algorytmów oraz przeprowadzenie procesów uczenia modelu odpowiedzialnego za analizę oprogramowania złośliwego
4. Określenie sposobu ewaluacji systemu za pomocą stosownych metryk

Proponowana metoda

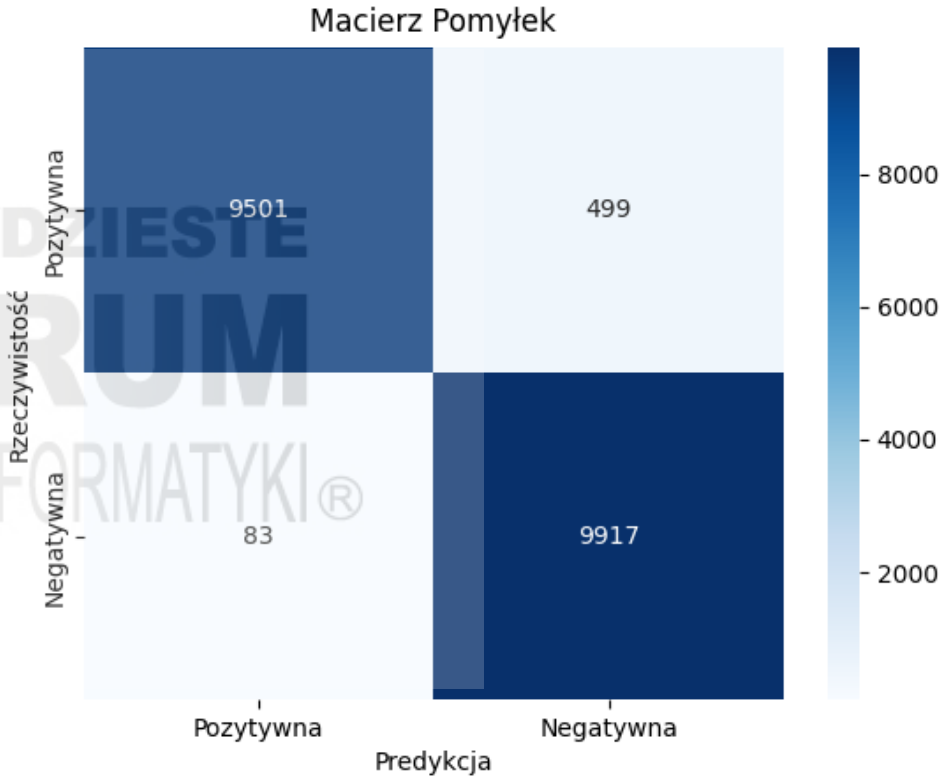
1. Wprowadzenie pliku
2. Przetworzenie pliku
   1. Odczyt pliku w postaci szesnastkowej
   2. Ekstrakcja trigramów
   3. Zliczenie wystąpień trigramów
   4. Redukcja wymiaru danych wejściowych średnią arytmetyczną
3. Klasyfikacja i analiza
   1. Predykcja wyniku z użyciem wytrenowanego modelu
   2. Binaryzacja wyniku predykcji wartością progową
4. Podjęcie decyzji
   1. Usuń plik
   2. Zachowaj plik



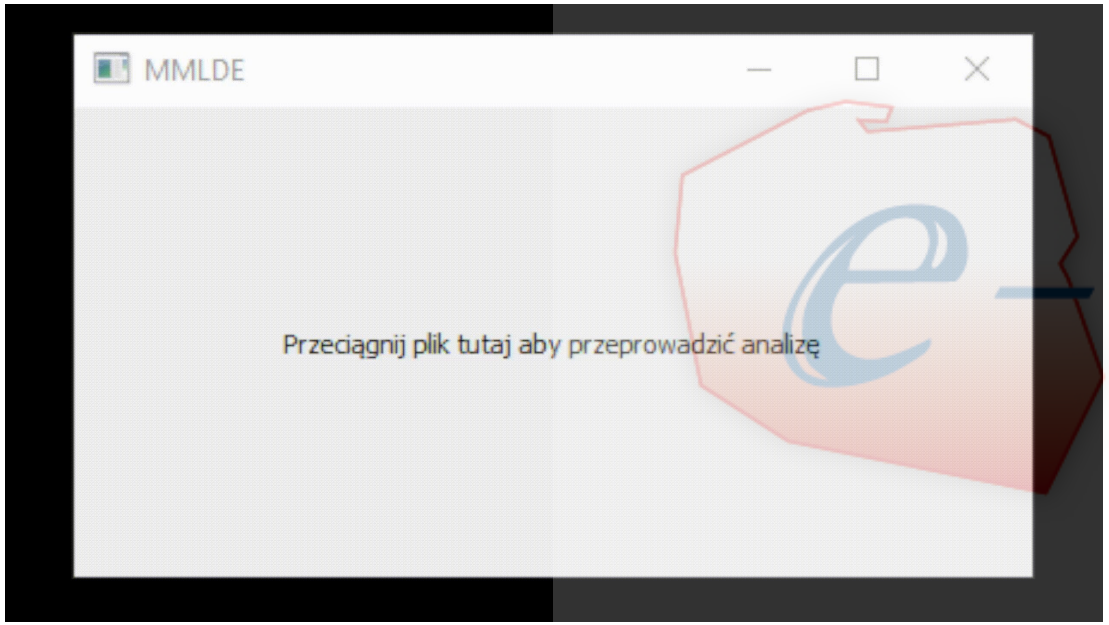
Wyniki badań

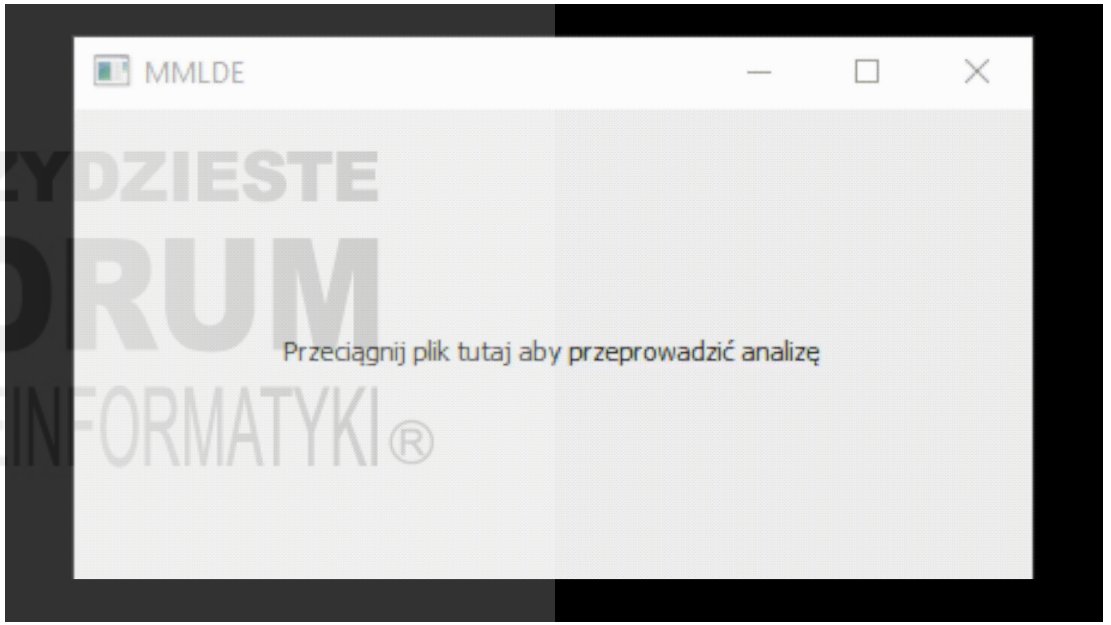
- Próbka testowa: 40 000 próbek
  - 10 000 plików nieszkodliwych 32 bit
  - 10 000 plików nieszkodliwych 64 bit
  - 10 000 plików szkodliwych 32 bit
  - 10 000 plików szkodliwych 64 bit

Iteracja	Accuracy	Precision	Recall	F_1
1	0.971	0.988	0.954	0.97
2	0.97	0.991	0.95	0.97
3	0.97	0.988	0.952	0.97
4	0.97	0.987	0.953	0.97
5	0.97	0.988	0.956	0.972
σ	0.0004	0.0014	0.002	0.0008



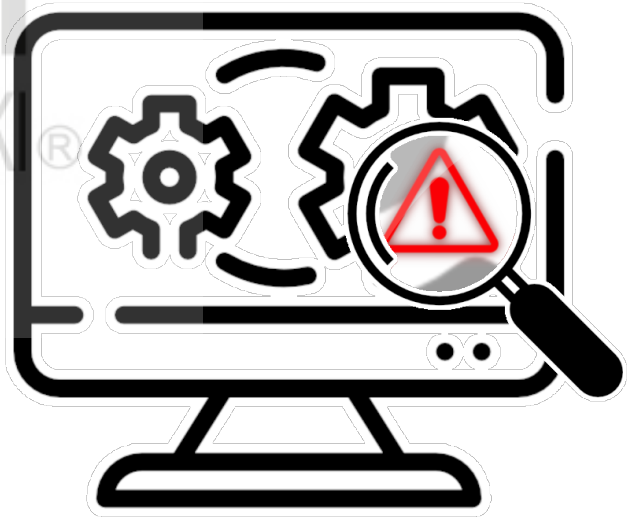
Prezentacja działania - LIVE





Podsumowanie

1. Zastosowanie AI/ML w ochronie systemów
    1. Praktyczne wykorzystanie algorytmów AI/ML do ochrony systemów i urządzeń przed cyberzagrożeniami.
    2. Model skanowania plików *.exe jako skuteczna metoda weryfikacji bezpieczeństwa systemów operacyjnych.
2. Dalsze prace i implementacja
    1. Planowana implementacja modelu spełniająca wymagania pracy w czasie rzeczywistym.
    2. Konieczność optymalizacji czasów przetwarzania plików dla lepszej efektywności.
3. Potencjalne kierunku rozwoju
    1. Zmiana sposobu generowania danych, np. przejście z trigramów na bigramy.
    2. Optymalizacja algorytmu tworzącego tensor z danymi.



Bibliografia

1. ENISA, "Raport ENISA Threat Landscape 2022". Dostęp: 15.07.2024 [Online]. Dostępne: https://nsarchive.gwu.edu/sites/default/files/documents/rmsj2i-vzrmv/2022-10-00-EU-ENISA-Threat-Landscape-2022.pdf
2. CERT Orange Polska, "Raport_CERT_Orange_Polska_2023". Dostęp: 15.07.2024. [Online]. Dostępne: https://cert.orange.pl/wp-content/uploads/2024/04/Raport_CERT_Orange_Polska_2023.pdf
3. Kaspersky, "Machine Learning for Malware Detection". Dostęp: 23.07.2024. [Online]. Dostępne: https://media.kaspersky.com/en/enterprise-security/Kaspersky-Lab-Whitepaper-Machine-Learning.pdf
4. M. Lester, „PE Malware Machine Learning Dataset.". Dostęp: 06.07.2024. [Online]. Dostępne: https://practicalsecurityanalytics.com/pe-malware-machine-learning-dataset
5. CERT.PL, *„Malware Database"*. Dostęp: 06.07.2024. [Online]. Dostępne: https://mwdb.cert.pl/
6. H. M i S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations", Int. J. Data Mining & Knowl. Manage. Process, t. 5, nr 2, s. 01–11, mar 2015. Dostęp: 23 lip 2024. [Online]. Dostępny: https://doi.org/10.5121/ijdkp.2015.5201
7. Virustotal, „virustotal.com". Dostęp: 18.07.2024. [Online]. Dostępne: https://virustotal.com