Politechnika Śląska

KARTA MIKROWARSZTATU

Nazwa mikrowarsztatu: ***Bezpieczeństwo systemów przemysłowych wspomagane sztuczną inteligencją – aktualne wymagania i rekomendacje związane z aktami prawnymi (NIS2, CRA, AI Act)***

Nazwa Wydziału: Wydział Górnictwa, Inżynierii Bezpieczeństwa i Automatyki Przemysłowej

Prowadzący: dr hab. inż. Anna Manowska, prof. PŚ, dr hab. inż. Artur Kozłowski, prof. PŚ

Skrócony opis mikrowarsztatu (treści kształcenia):

Mikrowarsztat ma na celu prezentację oceny wpływu systemów na organizację, oceny narzędzi ML/AI, rozumienia zależności łańcuchów dostaw, a także umiejętności raportowania z zastosowaniem standardów CSIRT. Oczekiwane kompetencje obejmują dogłębne zrozumienie zasad zarządzania ryzykiem. W szczególności zakres tematyczny dotyczy możliwości wykorzystania generatywnej sztucznej inteligencji w analizie i zwiększaniu poziomu bezpieczeństwa systemów przemysłowych (ICS/SCADA, IIoT). Uczestnicy pracują na danych takich jak logi systemowe, alerty bezpieczeństwa i opisy incydentów. W ramach zajęć wykorzystują modele językowe do klasyfikacji zdarzeń, identyfikacji anomalii oraz wspomagania decyzji w zakresie reagowania na incydenty. Analizowane są również scenariusze ataków na systemy przemysłowe oraz ryzyka i ograniczenia wynikające z użycia AI (np. błędne rekomendacje, manipulacja danymi wejściowymi). Zajęcia obejmują również pracę praktyczną z narzędziami AI oraz ocenę możliwości ich zastosowania w środowisku przemysłowym.

Opis mikrowarsztatu:

Wykład (5 h)

1. **Architektura systemów przemysłowych i identyfikacja zagrożenia (ICS/SCADA/IIoT)** (1,5 h)
   - komponenty systemów
   - powierzchnia ataku
   - specyfika systemów OT
2. **Zastosowanie generatywnej AI w cyberbezpieczeństwie** (1,5 h)
   - analiza logów i incydentów
   - klasyfikacja zdarzeń
   - wspomaganie decyzji
3. **Ograniczenia możliwości stosowania modeli AI w systemach bezpieczeństwa** (1 h)
   - błędne rekomendacje
   - manipulacja wejściem (prompt/data poisoning)
   - ryzyko operacyjne
4. **Aspekty prawne i regulacyjne stosowania AI w systemach przemysłowych** (1 h)
   - AI Act (klasy ryzyka)
   - odpowiedzialność za decyzje systemów AI
   - bezpieczeństwo danych i zgodność (compliance)

Ćwiczenia (5 h)

1. **Analiza incydentów w systemach przemysłowych** (1,5 h)
   - klasyfikacja zdarzeń
   - identyfikacja przyczyn
   - ocena skutków
2. **Wykorzystanie generatywnej AI do analizy zdarzeń** (1,5 h)
   - praca z modelami językowymi
   - interpretacja wyników
   - identyfikacja błędów
3. **Ocena ryzyka i zgodności** (1 h)
   - przypisanie poziomu ryzyka (AI Act – uproszczone)
   - analiza scenariuszy
4. **Projekt koncepcji wdrożenia AI w przedsiębiorstwie** (1 h)
   - możliwość zastosowania AI
   - identyfikacja danych wejściowych
   - podstawowe wymagania bezpieczeństwa

Laboratorium (5 h)

1. **Analiza danych bezpieczeństwa (logi, alerty)** (1,5 h)
   - identyfikacja anomalii
   - przygotowanie danych
2. **Zastosowanie modeli generatywnych w analizie incydentów** (1,5 h)
   - klasyfikacja i opis zdarzeń

KRAJOWY PLAN ODBUDOWY
Rzeczpospolita Polska
Sfinansowane przez Unię Europejską NextGenerationEU

- wspomaganie decyzji

3. **Testowanie ograniczeń AI** (1 h)
   - błędne dane
   - manipulacja wynikami
   - ocena wiarygodności
4. **Mini-projekt: scenariusz wdrożenia AI w systemie przemysłowym** (1 h)
   - wybór obszaru (np. monitoring, detekcja incydentów)
   - identyfikacja ryzyk
   - elementy zgodności i bezpieczeństwa

Liczba godzin zajęć z bezpośrednim udziałem prowadzącego i studentów:	15
Liczba godzin przeznaczonych na pracę własną studenta:	15
Całkowita liczba godzin:	30
Liczba punktów ECTS:	1
Forma zaliczenia:	Raport z realizacji i obrona mini-projektu (scenariusz wdrożenia AI w systemie przemysłowym) oraz zaliczenie praktycznych zadań laboratoryjnych (protokół z oceny ryzyka i analizy incydentów).

Literatura:

1. Stalling W., Network Security Essentials, Pearson
2. NIST SP 800-82 – Guide to Industrial Control Systems Security
3. Chandola V. et al., *Anomaly Detection: A Survey*, ACM Computing Surveys
4. Materiały własne prowadzącego (studia przypadków ICS/SCADA)

Efekty uczenia się

Wiedza
Student zna i rozumie:

1. Architekturę i specyfikę systemów przemysłowych (ICS/SCADA/OT) oraz wybrane fakty i złożone zależności dotyczące zastosowania modeli generatywnej sztucznej inteligencji do analizy cyberzagrożeń
2. Podstawowe prawne, etyczne i operacyjne uwarunkowania wykorzystania AI w środowisku przemysłowym, ze szczególnym uwzględnieniem zapisów AI Act, klasyfikacji ryzyka oraz zagadnień zgodności (compliance).

Umiejętności
Student potrafi:

1. Analizować logi i incydenty bezpieczeństwa w systemach przemysłowych, stosować narzędzia AI do klasyfikacji zdarzeń oraz interpretować wyniki ich działania.
2. Komunikować wyniki analiz incydentów oraz rekomendacje dotyczące bezpieczeństwa systemów przemysłowych z użyciem terminologii specjalistycznej.
3. Współpracować w zespole przy analizie incydentów oraz realizacji mini-projektu wdrożeniowego AI w środowisku przemysłowym.
4. Samodzielnie rozwijać wiedzę i umiejętności w zakresie nowych zagrożeń cybernetycznych oraz zastosowań AI w bezpieczeństwie.

Kompetencje społeczne
Student jest gotów do:

1. Krytycznej oceny wyników działania systemów AI w kontekście bezpieczeństwa oraz weryfikacji ich poprawności.
2. Współdziałania w zakresie analizy ryzyk i incydentów oraz odpowiedzialnego podejścia do bezpieczeństwa systemów przemysłowych.
3. Przestrzegania zasad etyki zawodowej w zakresie analizy danych, bezpieczeństwa informacji oraz stosowania AI w środowiskach krytycznych.

Metody i kryteria oceniania:

Metody oceniania:

- ocena raportu z realizacji mini-projektu (scenariusz wdrożenia AI w systemie przemysłowym),
- ocena wykonania zadań laboratoryjnych (analiza logów, klasyfikacja incydentów, ocena ryzyka),
- weryfikacja wiedzy i umiejętności w trakcie zajęć (dyskusja, interpretacja wyników, praca z narzędziami AI).

Kryteria oceniania:

- **poprawność analizy incydentów i danych bezpieczeństwa** (identyfikacja zagrożeń, spójność wniosków),
- **umiejętność wykorzystania narzędzi AI** (adekwatność zastosowanych metod, interpretacja wyników),
- **ocena wiarygodności i ograniczeń AI** (identyfikacja błędów, krytyczna analiza wyników),
- **jakość opracowanego mini-projektu wdrożeniowego** (logika rozwiązania, uwzględnienie aspektów bezpieczeństwa i regulacji),
- **poprawność i przejrzystość prezentacji wyników** (struktura raportu, terminologia specjalistyczna).

Warunki zaliczenia:

- uzyskanie pozytywnej oceny z raportu końcowego,
- zaliczenie zadań laboratoryjnych,