Politechnika Krakowska im. Tadeusza Kościuszki

Karta przedmiotu

obowiązuje studentów rozpoczynających studia w roku akademickim 2023/2024

Wydział Informatyki i Telekomunikacji

Kierunek studiów: Informatyka | Profil: Ogólnoakademicki

Forma sudiów: niestacjonarne | Kod kierunku: I

Stopień studiów: II

Specjalności: Cyberbezpieczeństwo

1 Informacje o przedmiocie

Nazwa przedmiotu	Przetwarzanie i ochrona danych typu Big Data
Nazwa przedmiotu w języku angielskim	Processing and protection of Big Data sets
Kod przedmiotu	WIiT I oIIN D8 23/24
Kategoria przedmiotu	Przedmioty specjalnościowe
Liczba punktów ECTS	4.00
Semestry	2

2 Rodzaj zajęć, liczba godzin w planie studiów

Semestr	Wykład	Ćwiczenia	Laboratorium	Laboratorium komputerowe	Seminarium	Projekt
2	18	0	0	0	0	9

3 Cele przedmiotu

Cel 1 Zapoznanie studentów z podstawowymi pojęciami i metodami bezpiecznego przetwarzania i ochrony danych typu Big Data.

Cel 2 Zapoznanie studentów z podstawowymi algorytmami zabezpieczania danych Big Data.

4 Wymagania wstępne w zakresie wiedzy, umiejętności i innych kompetencji

1 Podstawowe informacje dotyczące przetwarzania danych i systemów bazodanowych.

2 Podstawy bezpieczeństwa w systemach informatycznych

5 Efekty kształcenia

EK1 Wiedza Znajomość podstawowych pojęć z zakresu bezpieczeństwa danych w systemach Big Data.

EK2 Umiejętności Umiejętność rozpoznawania zagrożeń w systemach Big Data i anomalii w zgromadzonych danych.

EK3 Umiejętności Umiejętność implementacji algorytmów metod wykrywania ataków i anomalii w systemach Big Data.

EK4 Kompetencje społeczne Świadomość znaczenia nowoczesnych algorytmów wykrywania anomalii w dużych zbiorach danych i potrzeby ochrony i bezpiecznego przetwarzania danych w systemach Big Data.

6 Treści programowe

Projekt		
Lp	Tematyka zajęć Opis szczegółowy bloków tematycznych	Liczba godzin
P1	Definicja problemu - atak DDoS, wykrywanie anomalii w ruchu, etc. Podział na grupy projektowe, opracowanie planu pracy i harmonogramu wykonania zadań.	2
P2	Definicja i implementacja metody (metod) rozwiązania postawionego problemu.	3
P3	Wykonanie eksperymentów i sprawozdania.	3
P4	Prezentacja projektów.	1

Wykład		
Lp	Tematyka zajęć Opis szczegółowy bloków tematycznych	Liczba godzin
W1	Charakterystyka zbiorów danych typu Big Data - model 4V. Praktyczne aspekty bezpiecznego przetwarzania i przechowywania dużych zbiorów danych, łańcuch odpowiedzialności za bezpieczeństwo w systemach Big Data.	2
W2	Bazy danych NoSQL. Modele danych i ich bezpieczeństwo w bazach NoSQL.	2
W3	Wprowadzenie do platformy Hadoop. Model MapReduce. Systemy plików i przepływ danych w Hadoop. Projekt systemu HDFS. Manager klastra YARN.	4
W4	Wykrywanie anomalii w zbiorach danych.	4

Wykład		
Lp	Tematyka zajęć Opis szczegółowy bloków tematycznych	Liczba godzin
W5	Wprowadzenie do platformy Apache Spark, rozproszone kolekcje obiektów RDD (Resilient Distributed Dataset), operacje na zbiorach danych w oparciu o język zapytań Spark SQL. Bezpieczeństwo w systemach Apache Spark.	4
W6	Automatyzacja strategii obronnych w systemach typu Big Data.	2

7 Narzędzia dydaktyczne

N1 Wykłady

N2 Dyskusja, konsultacje

N3 Prezentacje multimedialne

N4 Ćwiczenia projektowe

N5 Praca zdalna - wykorzystanie MS TEAMS i/lub DELTA

8 Obciążenie pracą studenta

Forma aktywności	Średnia liczba godzin na zrealizowanie aktywności
Godziny kontaktowe z nauczycielem akademickim, w tym:	
Godziny wynikające z planu studiów	27
Konsultacje przedmiotowe	10
Egzaminy i zaliczenia w sesji	0
Godziny bez udziału nauczyciela akademickiego wynikające z nakładu pracy studenta, w tym:	
Przygotowanie się do zajęć, w tym studiowanie zalecanej literatury	18
Opracowanie wyników	25
Przygotowanie raportu, projektu, prezentacji, dyskusji	20
Sumaryczna liczba godzin dla przedmiotu wynikająca z całego nakładu pracy studenta	**100**
Sumaryczna liczba punktów ECTS dla przedmiotu	4.00

9 Sposoby oceny

Nie przeprowadza się testu wprowadzającego.

Ocena formująca

F1 Sprawozdania z wykonanych projektów.

F2 Aktywność na wykładach.

Ocena podsumowująca

P1 Średnia ważona ocen formujących (90%, 10%).

Warunki zaliczenia przedmiotu

W1 Pozytywna ocena z projektów - zaliczenie wszystkich elementów.

W2 Brak nieusprawiedliwionych nieobecności na obowiązkowych formach zajęć

Ocena aktywności bez udziału nauczyciela

B1 Ocena pracy indywidualnej studenta nad wykonaniem projektu - ocena raportów z postępów pracy wysyłanych prowadzącym zajęcia i prezentowanych podczas zajęć (również z wykorzystywaniem metod nauczania zdalnego).

Kryteria oceny

Efekt kształcenia 1	
Na ocenę 2.0	Student nie spełnia wymagań na ocenę 3.0.
Na ocenę 3.0	Student zna ogólne zasady dotyczące zagrożeń bezpieczeństwa oraz zabezpieczania oraz potrafi scharakteryzować środowiska Big Data w kontekście bezpieczeństwa danych. Student potrafi wymienić po jednym przykładzie z ww zagadnień.
Na ocenę 3.5	Student zna ogólne zasady dotyczące zagrożeń bezpieczeństwa oraz zabezpieczania oraz potrafi scharakteryzować środowiska Big Data w kontekście bezpieczeństwa danych, zarządzania użytkownikami oraz obsługi baz typu NoSQL Student potrafi podać różne przykłady dla każdego wymienionego zagadnienia.
Na ocenę 4.0	Student zna szczegółowo zasady dotyczące zagrożeń bezpieczeństwa oraz zabezpieczania oraz potrafi scharakteryzować środowiska Big Data w kontekście bezpieczeństwa danych, zarządzania użytkownikami oraz obsługi baz typu NoSQL Student potrafi podać różne przykłady dla każdego wymienionego zagadnienia.
Na ocenę 4.5	Student zna ogólne zasady dotyczące zagrożeń bezpieczeństwa oraz zabezpieczania oraz potrafi scharakteryzować środowiska Big Data w kontekście bezpieczeństwa danych, zarządzania użytkownikami oraz obsługi baz typu NoSQL Student potrafi podać różne przykłady dla każdego wymienionego zagadnienia. Potra podać oraz rozumie matematyczne podstawy wybranych metod zabezpieczania danych typu Big Data.

Na ocenę 5.0	Student zna ogólne zasady dotyczące zagrożeń bezpieczeństwa oraz zabezpieczania oraz potrafi scharakteryzować środowiska Big Data w kontekście bezpieczeństwa danych, zarządzania użytkownikami oraz obsługi baz typu NoSQL Student potrafi podać różne przykłady dla każdego wymienionego zagadnienia. Potra podać oraz rozumie matematyczne podstawy wybranych metod zabezpieczania danych typu Big Data. Student potrafi dobrać metody zabezpieczania ww środowisk do konkretnych przykładów przetwarzania danych typu Big Data.
Efekt kształcenia 2	
Na ocenę 2.0	Student nie spełnia wymagań na ocenę 3.0.
Na ocenę 3.0	Student posiada podstawową wiedzę o wykrywaniu anomalii w dużych zbiorach danych.
Na ocenę 3.5	Student zna podstawowe zagrożenia w systemach typu Big Data związane z przechowywaniem danych, zna metody rozpoznawania anomalii w tych danych.
Na ocenę 4.0	Student ma dość szeroką wiedzę na temat zagrożeń w systemach typu Big Data związanych z przechowywaniem danych, zna metody rozpoznawania anomalii w tych danych.
Na ocenę 4.5	Student ma szeroką wiedzę na temat zagrożeń w systemach typu Big Data, zna metody bezpiecznego przechowywania danych w tych systemach i metody rozpoznawania anomalii w dużych zbiorach danych strumieniowych i nieustruktyzowanych.
Na ocenę 5.0	Student doskonale orientuje się w zagadnieniach na temat zagrożeń w systemach typu Big Data, zna metody bezpiecznego przechowywania danych w tych systemach i metody rozpoznawania anomalii w dużych zbiorach danych strumieniowych i nieustruktyzowanych.
Efekt kształcenia 3	
Na ocenę 2.0	Student nie spełnia wymagań na ocenę 3.0.
Na ocenę 3.0	Student potrafi zaimplementować wskazane algorytmy detekcji ataków i anomalii w systemach Big Data oraz przeprowadzić proste eksperymenty z wykorzystaniem zaimplementowanych algorytmów.
Na ocenę 3.5	Student potrafi zaimplementować wskazane algorytmy detekcji ataków i anomalii w systemach Big Data oraz przeprowadzić proste eksperymenty z wykorzystaniem zaimplementowanych algorytmów, potrafi zinterpretować otrzymane wyniki.
Na ocenę 4.0	Student potrafi zaimplementować wskazane algorytmy detekcji ataków i anomalii w systemach Big Data oraz przeprowadzić złożone testy jednostkowe wykonanego oprogramowania oraz eksperymenty numeryczne na dostarczonych danych testowych, potrafi zinterpretować wyniki tych testów.
Na ocenę 4.5	Student potrafi zaimplementować wskazane algorytmy detekcji ataków i anomalii w systemach Big Data oraz przeprowadzić złożone testy jednostkowe wykonanego oprogramowania oraz eksperymenty numeryczne na dostarczonych danych testowych, potrafi zinterpretować wyniki tych testów za pomocą prostej analizy statystycznej.

Na ocenę 5.0	Student potrafi zaimplementować wskazane algorytmy detekcji ataków i anomalii w systemach Big Data oraz przeprowadzić wszechstronną analizę eksperymentalną i dokonać zaawansowanej analizy statystycznej otrzymanych wyników.
Efekt kształcenia 4	
Na ocenę 2.0	Student nie spełnia wymagań na ocenę 3.0.
Na ocenę 3.0	Student zna przynajmniej 1 przykład praktyczny metody wykrywania anomalii lub ataków DDoS w systemach typu Big Data.
Na ocenę 3.5	Student zna kilka przykładów metod wykrywania anomalii lub ataków DDoS w systemach typu Big Data.
Na ocenę 4.0	Student zna kilka przykładów metody wykrywania anomalii lub ataków DDoS w systemach typu Big Data., potrafi je przedstawić w prosty sposób.
Na ocenę 4.5	Student orientuje się w praktycznych zastosowaniach metod wykrywania anomalii lub ataków DDoS w systemach typu Big Data i potrafi wytłumaczyć zalety ich stosowania w praktyce.
Na ocenę 5.0	Student doskonale orientuje się w praktycznych zastosowaniach metod wykrywania anomalii lub ataków DDoS w systemach typu Big Data i potrafi wytłumaczyć zalety ich stosowania w praktyce na przykładach flagowych praktycznych projektów.

10 Macierz realizacji przedmiotu

Efekt kształcenia	Odniesienie danego efektu do szczegółowych efektów zdefiniowanych dla programu	Cele przedmiotu	Treści programowe	Narzędzia dydaktyczne	Sposoby oceny
EK1	I2_W03 I2_W04 I2_W05 I2_W08 I2_U11 I2_K02	Cel 1 Cel 2	W1	N1 N2 N3 N5	F1
EK2	I2_W06 I2_W08 I2_U01b I2_U02b I2_U03b I2_U05 I2_U07 I2_U08 I2_U12 I2_K01 I2_K02	Cel 1 Cel 2	P1 P2 W1 W2 W3 W4 W5 W6	N1 N2 N3 N4 N5	F1 F2 P1

Efekt kształcenia	Odniesienie danego efektu do szczegółowych efektów zdefiniowanych dla programu	Cele przedmiotu	Treści programowe	Narzędzia dydaktyczne	Sposoby oceny
EK3	I2_W01 I2_W03 I2_U02b I2_U03b I2_U07 I2_U11 I2_U12 I2_K02	Cel 1 Cel 2	P1 P2 P3 P4 W1 W2 W3 W4	N1 N2 N3 N4 N5	F1 F2 P1
EK4	I2_U03b I2_U04b I2_U05 I2_U06 I2_U07 I2_U08 I2_U10 I2_U11 I2_K01 I2_K02 I2_K03 I2_K04	Cel 1 Cel 2	P1 P2 P3 P4 W1 W2 W3 W4 W5 W6	N1 N2 N3 N4 N5	F1 F2 P1

11 Wykaz literatury

Literatura podstawowa

[1] Thomas Erl, Wajid Khattak, andPaulBuhler — *BigData Fundamentals Concepts, Drivers & Techniques*, , 2017, Pearson

Literatura uzupełniająca

[1] Hang-Jung Hsieh, Ting-Yuan Chan — *DETECTION DDOS ATTACKS BASED ON NEURAL-NETWORK USING APACHE SPARK*, -, 0,

12 Informacje o nauczycielach akademickich

Osoba odpowiedzialna za kartę

dr hab. prof.PK Joanna Kołodziej (kontakt: joanna.kolodziej@pk.edu.pl)

Osoby prowadzące przedmiot

1 Dr hab. prof. PK Joanna Kołodziej (kontakt: joanna.kolodziej@pk.edu.pl)

2 Mgr inż. Piotr Biskup (kontakt: piotr.biskup@pk.edu.pl)

13 Zatwierdzenie karty przedmiotu do realizacji

(miejscowość, data) (odpowiedzialny za przedmiot) (dziekan)

Przyjmuję do realizacji (data i podpisy osób prowadzących przedmiot)

.....

.....