

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Statystyka i analiza danych		Kod 1010514351010511921
Kierunek studiów Informatyka	Profil kształcenia (ogólnoakademicki, praktyczny) ogólnoakademicki	Rok / Semestr 3 / 5
Ścieżka obieralności/specjalność -	Przedmiot oferowany w języku: polski	Kurs (obligatoryjny/obieralny) obligatoryjny
Stopień studiów: I stopień	Forma studiów (stacjonarna/niestacjonarna) niestacjonarna	
Godziny Wykłady: 16 Ćwiczenia: - Laboratoria: 16 Projekty/seminaria: -		Liczba punktów 4
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) podstawowy		(ogólnouczelniany, z innego kierunku) z danego kierunku
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki nauki techniczne		Podział ECTS (liczba i %) 4 100%
Odpowiedzialny za przedmiot / wykładowca:		
prof. dr hab. inż. Jerzy Stefanowski email: Jerzy.Stefanowski@cs.put.poznan.pl tel. (0-61) 665-2903 Instytut Informatyki ul. Piotrowo 2, 60-965 Poznań		
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z matematyki, w szczególności rachunku prawdopodobieństwa (pojęcia jak: prawdopodobieństwo dyskretne i ciągłe; wartości oczekiwane i wariancja; rozkłady zmiennych losowych - zwłaszcza dwupunktowy i normalny; dystrybuanta standaryzowanego rozkładu normalnego, centralne twierdzenie Lindberga - Levy'ego; podstawy estymacji punktowej i przedziałowej).
2	Umiejętności:	Powinien posiadać umiejętność rozwiązywania podstawowych zadań związanych z powyższymi pojęciami probabilistycznymi, rozwiązywania układów równań liniowych oraz wykonywania obliczeń z wykorzystaniem macierzy. Konieczna jest też podstawowa sprawność praktycznej obsługi arkuszy kalkulacyjnych (wraz z wykorzystaniem dostępnych funkcji oraz definiowania własnych formuł przetwarzania danych) oraz umiejętność pozyskiwania informacji ze wskazanych źródeł literaturowych.
3	Kompetencje społeczne	Student musi również rozumieć konieczność samokształcenia i poszerzania swoich umiejętności. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza i kreatywność.
Cel przedmiotu:		
Celem przedmiotu jest praktyczne wprowadzenie w zagadnienia statystycznej analizy danych i jej zastosowań w informatyce i naukach technicznych. W szczególności zamierza się:		
1. Przekazać studentom postawy wiedzy ze statystyki obejmujące: wstępną analizę danych (w zakresie graficznego przetwarzania danych i miar statystyki opisowej); przejście od modelu probabilistycznego do wnioskowania statystycznego; testowanie hipotez; analizę regresji oraz analizę zależności zmiennych jakościowych. 2. Rozwijać u studentów umiejętności rozwiązywania prostych problemów z zakresu planowania eksperymentów, pozyskania danych łącznie z wykorzystaniem technik losowego wyboru prób, obliczania podstawowych miar statystycznych, oceny wiarygodności wartości tych miar, doboru właściwych testów statystycznych dla wyjaśnienia zjawisk reprezentowanych przez dane oraz interpretacji ich wyników. Ponadto zwraca się uwagę na poprawne opracowywanie wyników badania statystycznego oraz stosowania tabelarycznej i graficznej prezentacji danych. 3. Kształtować umiejętności wykorzystania oprogramowania statystycznego oraz algorytmizacji przetwarzania danych eksperymentalnych.		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza:		

1. ma wiedzę w zakresie statystyki matematycznej potrzebną do zrozumienia wybranych działów informatyki oraz innych nauk technicznych i eksperymentalnych - [K_W3, KW8]
2. zna podstawowe pojęcia i metody statystyki opisowej, prezentacji danych, wnioskowania statystycznego oraz analizy regresji, które są przydatne do rozwiązywania prostych zadań informatycznych z zakresu analizy algorytmów i ich oceny eksperymentalnej, oceny efektywności systemów informatycznych, sztucznej inteligencji - [K_W8]
3. zna podstawowe pojęcia i metody statystyki opisowej, prezentacji danych, wnioskowania statystycznego oraz analizy regresji, które są przydatne do rozwiązywania prostych zadań informatycznych z zakresu baz danych, inżynierii oprogramowania, wspomagania decyzji oraz systemów pomiarowych - [K_W8]
Umiejętności:
1. potrafi wykorzystać metody statystyki opisowej i wnioskowania statystycznego (głównie testowania hipotez) do rozwiązywania zadań informatycznych - [K_U8]
2. potrafi wyznaczyć charakterystyki liczbowe jednowymiarowych danych pomiarowych - [K_U7]
3. potrafi wykonać weryfikacje hipotez dotyczących parametrów i własności danych - [K_U8]
4. potrafi dokonać analizy prostego problemu analizy danych, dokonać wyboru narzędzi pozyskania danych i ich statystycznej analizy oraz przedstawić uzyskane wyniki - [K_U7]
Kompetencje społeczne:
1. rozumie potrzeby poszerzania kompetencji w zakresie analizy danych i poszukiwania jej zastosowań w informatyce i pokrewnych dziedzinach - [K_K1]

Sposoby sprawdzenia efektów kształcenia

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

- a) w zakresie wykładów:
 - na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach;
- b) w zakresie ćwiczeń:
 - na podstawie oceny bieżącego postępu realizacji zadań,

Ocena podsumowująca:

Sprawdzanie założonych efektów kształcenia realizowane jest przez:

- ocenę przygotowania studenta do poszczególnych zajęć laboratoryjnych (sprawdzian wejściowy),
- ocenianie ciągle, na każdym zajęciu (odpowiedzi ustne) - premiowanie przyrostu umiejętności posługiwania się poznanymi metodami i ich wykorzystania w dostępnym oprogramowaniu (ocena wykonania analizy problemu lub zadania przedstawionego w instrukcji do wybranego ćwiczenia),
- ocenę raportu z analizy problemów (przygotowywanego częściowo w trakcie zajęć, a częściowo po ich zakończeniu) oraz zadań domowych (weryfikujących dodatkowo umiejętności praktyczne lub wymagających samodzielnego zebrania dodatkowych danych i ich analizy),
- ocenę wiedzy i umiejętności związanych z realizacją zadań laboratoryjnych poprzez 2 kolokwia w semestrze (obejmującego także problemy obliczeniowe),
- ocenę wiedzy z zakresu wykładów poprzez kolokwium w formie testu wielokrotnego wyboru oraz rozwiązywania kilku prostych zadań (tego typu test obejmuje na ogół około 10 pytań i zadań)

Uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:

- wskazanie nowych obszarów zastosowań praktycznych zagadnień będących przedmiotem ćwiczeń,
- zaproponowanie pozyskania nietypowych, nowych zestawów danych dla powyższego ćwiczenia,
- efektywność zastosowania zdobytej wiedzy podczas rozwiązywania zadanego problemu w instrukcji ćwiczenia,
- uwagi związane z udoskonaleniem materiałów dydaktycznych lub wskazywanie trudności percepcyjnych studentów umożliwiające bieżące doskonalenie procesu dydaktycznego.

Elementy powyższej oceny są dokumentowane w postaci prac pisemnych studenta, raportów oraz arkuszy ocen jego aktywności na zajęciach.

Treści programowe

<p>Program wykładu obejmuje następujące zagadnienia:</p> <ol style="list-style-type: none"> 1. Zadania analizy danych. Mierzenie zmiennych i rodzaje skal pomiarowych. Tworzenie szeregów rozdzielczych oraz histogramów. Wstępna graficzna analiza danych. 2. Miary statystyki opisowej (wartości przeciętne i rozproszenia). 3. Pojęcia rachunku prawdopodobieństwa przydatne we wnioskowaniu statystycznym (Pojęcia związane z rozkładem normalnym i dwupunktowym, estymatory punktowe). 4. Podstawowe metody doboru prób losowych. 5. Wprowadzanie do testowania hipotez (testy dla wartości średniej dla jednej próby). 6. Ocena ryzyka błędów popełnianych w testowaniu hipotez oraz test t-studenta dla wartości średniej w małej próbie. 7. Inne testy parametryczne (dla dwóch zbiorowości, wariancji). 8. Badanie zależności między zmiennymi liczbowymi, miary współzależności, współczynniki kowariancji i korelacji. 9. Regresja liniowa. Rola obserwacji "nietypowych" w analizie regresji. Ocena wiarygodności modelu regresji. 10. Analiza danych jakościowych. Test chi-2 - wykorzystanie do testowania zgodności rozkładów. 11. Badania zależności dwóch zmiennych nominalnych. Miary siły związku. 12. Wybrane testy nieparametryczne (test znaków oraz Wilcoxon) oraz korelacja rangowa. 13. Zasady planowania i wykonywania prostego badania statystycznego. 14. Elementy graficznej prezentacji danych i wyników analizy. Przegląd wybranych przykładów oprogramowania. 15. Kolokwium ? test zaliczeniowy. <p>Podczas laboratorium studenci rozwiązują zadania albo większe problemy dotyczące przykładów stosowania metod statystycznych w analizie danych pochodzących z problemów informatycznych, innych technicznych, marketingowych lub zarządzania. Zakres ćwiczeń jest powiązany w przebiegiem wykładu (w przypadku niektórych ćwiczeń wymaga samodzielnego uzupełnienia wiedzy w oparciu o wskazane podręczniki).</p> <p>W szczególności przebieg ćwiczeń obejmuje:</p> <p>Tabelaryczna i graficzna prezentacja danych (1 ćwiczenie); Wprowadzanie do pakietów statystycznych oraz miary statystyki opisowej (2 ćwiczenia); Pobierane prób oraz estymacja punktowa w modelu normalnych i dwu-punktowym (1 ćwiczenie); Testowania hipotez ? parametry rozkładu dwu-punktowego (1 ćwiczenie) i średnia w rozkładzie normalnym (1 ćwiczenie); Inne testy parametryczne - t-Studenta, porównywanie dwóch populacji, ocena wariancji (2 ćwiczenia); Analiza korelacji i regresja liniowa (2 ćwiczenia); test chi-2 (1 ćwiczenie), testy nieparametryczne (1 ćwiczenie); studium przypadku badania statystycznego (1 ćwiczenie) oraz spotkania związane z realizacją kolokwium i zaliczenia całego laboratorium.</p> <p>W ramach ćwiczeń laboratoryjnych studenci poznają także i uczą się stosowania wybranego oprogramowania (klasyczne arkusze kalkulacyjne jak Ms Excel oraz specjalizowane oprogramowanie ? pakiet R).</p> <p>Do każdego z ćwiczeń przygotowane są instrukcje ich wykonania oraz niezbędne zestawy gotowych danych.</p> <p>Cześć wymienionych wyżej treści programowych realizowana jest w ramach pracy własnej studenta.</p> <p>Metody dydaktyczne:</p> <ol style="list-style-type: none"> 1. Wykład: prezentacja słowna ilustrowana przykładami podawanymi na tablicy, rozwiązywanie zadań, pokaz multimedialny wcześniej przygotowanej prezentacji. 2. Ćwiczenia laboratoryjne obejmujące: rozwiązywanie zadań obliczeniowych, ćwiczenia praktyczne analizy danych, wykonywanie eksperymentów ? pobierania prób losowych, dyskusja, pokazy multimedialne, studium przypadków (samodzielne zebrania danych do wybranego studium i ich analiza oraz przedstawienie wniosków wraz z oceną ich wiarygodności) 3. Instrukcje do przebiegu ćwiczeń, zastawy danych oraz tzw. slajdy prezentacyjne są udostępniane studentom stopniowo z odpowiednim wyprzedzeniem jako pliki dostępne na specjalnym serwerze dydaktycznym. 	
<p>Literatura podstawowa:</p> <ol style="list-style-type: none"> 1. Statystyka dla studentów kierunków technicznych i przyrodniczych, J. Koronacki, J. Mielniczuk, WNT, Warszawa, 2006 (I wydanie 2004) 2. Statystyka w zarządzaniu, A. Aczel, PWN, Warszawa, 2006 3. Statystyka praktyczna, W. Starzyńska, PWN, Warszawa, 2006 	
<p>Literatura uzupełniająca:</p> <ol style="list-style-type: none"> 1. Przystępny kurs statystyki. Tom 1, A. Stanisław, Statsoft, Kraków, 1997 2. Przewodnik po pakiecie R. P. Biecek, Oficyna Wyd. GiS. Wrocław 2008 3. Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Część II: Statystyka matematyczna, W. Kryszczyński, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski, Warszawa 1997 	
<p>Bilans nakładu pracy przeciętnego studenta</p>	
<p>Czynność</p>	<p>Czas (godz.)</p>

1. udział w zajęciach laboratoryjnych / ćwiczeniach :	16
2. przygotowanie do ćwiczeń laboratoryjnych:	12
3. dokończenie (w ramach pracy własnej) raportów z ćwiczeń laboratoryjnych oraz wykonanie tzw. zadań domowych:	12
	2
4. udział w konsultacjach związanych z realizacją procesu kształcenia, w szczególności ćwiczeń laboratoryjnych / projektu	20
	16
5. przygotowanie do sprawdzianów / kolokwium	20
6. udział w wykładach	
7. zapoznanie się ze wskazaną literaturą / materiałami dydaktycznymi (10 stron tekstu naukowego = 1 godz.), 200 stron	
Obciążenie pracą studenta	
forma aktywności	godzin
ECTS	
Łączny nakład pracy	98
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	34
Zajęcia o charakterze praktycznym	40
	4
	1
	2