



KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Statystyka i analiza danych [S1Inf1>SAD]

Przedmiot

Kierunek studiów
Informatyka

Rok/Semestr
2/4

Studia w zakresie (specjalność)
–

Profil studiów
ogólnoakademicki

Poziom studiów
pierwszego stopnia

Język oferowanego przedmiotu
polski

Forma studiów
stacjonarne

Wymagalność
obligatoryjny

Liczba godzin

Wykład
32

Laboratorium
24

Inne
0

Ćwiczenia
0

Projekty/seminaria
0

Liczba punktów ECTS

4,00

Koordynatorzy

prof. dr hab. inż. Jerzy Stefanowski
jerzy.stefanowski@put.poznan.pl

Wykładowcy

Wymagania wstępne

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). 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 arkusza kalkulacyjnego (wraz z wykorzystaniem dostępnych funkcji oraz definiowania własnych formuł przetwarzania danych) oraz umiejętność pozyskiwania informacji ze wskazanych źródeł literaturowych. 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 podstawy 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.

Przedmiotowe efekty uczenia się

Wiedza:

1. ma wiedzę w zakresie statystyki matematycznej potrzebną do zrozumienia wybranych działów informatyki oraz innych nauk technicznych i eksperymentalnych (K1st_W4, K1st_W5)
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 (K1st_W7)

Umiejętności:

1. potrafi wyznaczyć charakterystyki liczbowe jednowymiarowych danych pomiarowych (K1st_U3)
2. potrafi wykorzystać metody statystyki opisowej i wnioskowania statystycznego (głównie testowania hipotez) do rozwiązywania zadań informatycznych (K1st_U4)
3. potrafi wykonać weryfikację hipotez dotyczących parametrów i własności danych (K1st_U4)
4. potrafi dokonać analizy prostego problemu analizy danych, dokonać wyboru narzędzi pozyskania danych i ich statystycznej analizy oraz przedstawić uzyskane wyniki (K1st_U4)

Kompetencje społeczne:

1. ma wiedzę w zakresie statystyki matematycznej potrzebną do zrozumienia wybranych działów informatyki oraz innych nauk technicznych i eksperymentalnych (K1st_W4, K1st_W5)
2. rozumie potrzeby poszerzania kompetencji w zakresie analizy danych i poszukiwania jej zastosowań w informatyce i pokrewnych dziedzinach (K1st_K1, K1st_K2)

Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

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 - realizowane z wykorzystaniem testów pisemnych dwa razy podczas semestru (testu wielokrotnego wyboru oraz rozwiązywania kilku prostych zadań (tego typu test obejmuje na ogół około 10 pytań i zadań));

b) w zakresie ćwiczeń: na podstawie oceny bieżącego postępu realizacji zadań oraz sprawdzianów,

Ocena podsumowująca: Sprawdzanie założonych efektów kształcenia realizowane jest przez:

- ocenę przygotowania studenta do poszczególnych zajęć laboratoryjnych (sprawdzian wejściowy)
- ocenianie ciągle, na każdych zajęciach (odpowiedzi ustne) – premiowanie przyrostu umiejętności posługiwania się poznanymi metodami i ich wykorzystania w dostępnym oprogramowaniu
- 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),

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

Treści programowe

Program wykładu obejmuje następujące zagadnienia:

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 korelacji.
9. Regresja liniowa. Rola obserwacji 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 Wilcoxa) 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.

Podczas laboratorium studenci rozwiązują zadania albo większe problemy dotyczące przykładów stosowania metod statystycznych w analizie danych pochodzących z problemów informatycznych, innych technicznych, marketingowych lub zarządczych. Zakres ćwiczeń jest powiązany w przebiegu wykładu (w przypadku niektórych ćwiczeń wymaga samodzielnego uzupełnienia wiedzy w oparciu o wskazane podręczniki).

W szczególności przebieg ćwiczeń obejmuje:

Tabelaryczna i graficzna prezentacja danych (1 ćwiczenie); Wprowadzanie do pakietów statystycznych oraz miary statystyki opisowej (2 ćwiczenia); Pobieranie prób oraz estymacja punktowa w modelu normalnym i dwu-punktowym (1 ćwiczenie); Testowanie 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.

W ramach ćwiczeń laboratoryjnych studenci poznają także i uczą się stosowania wybranego oprogramowania (np. specjalizowane oprogramowanie – pakiet R).

Tematyka zajęć

brak

Metody dydaktyczne

Wykład: prezentacja multimedialna, ilustrowana przykładami podawanymi na tablicy. Ćwiczenia laboratoryjne: prezentacja multimedialna prezentacja ilustrowana przykładami podawanymi na tablicy oraz wykonanie zadań podanych przez prowadzącego - ćwiczenia praktyczne.

Literatura

Podstawowa

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

Uzupełniająca

1. Przystępny kurs statystyki. Tom 1, A. Stanisław, Statsoft, Kraków, 1997
2. Przewodnik po pakiecie R. P. Biecek, Oficyna Wyd. GiS. Wrocław 2008
3. Statystyka praktyczna, W. Starzyńska, PWN, Warszawa, 2006

Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	100	4,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	58	2,00
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu)	42	2,00