



Zestaw zagadnień egzaminacyjnych

dla kierunku *Inżynieria danych*

obowiązuje od roku akad. 2021/2022

1. Istota i funkcje zarządzania. Wybrane metody zarządzania.
2. Zasady programowania w języku Python – reguły indentacji, typy obiektów wbudowanych.
3. Jednostki pamięci komputerowej. Rodzaje danych komputerowych możliwych do przetwarzania we współczesnych systemach komputerowych. Reprezentacja liczby rzeczywistej (zmiennoprzecinkowa) i liczby całkowitej ze znakiem i bez znaku w pamięci komputera.
4. Budowa, zadania i zasada działania systemu operacyjnego.
5. Paradygmat programowania. Przykładowe paradygmaty. Koncepcja programowania sterowanego zdarzeniami.
6. Metodyka programowania RAD (Rapid Application Development). Przykładowe narzędzia i biblioteki RAD.
7. Krótka charakterystyka programowania orientowanego obiektowo. Pojęcie klasy i jej składowych. Definicja obiektu. Enkapsulacja, dziedziczenie, polimorfizm.
8. Proste i złożone typy danych w językach programowania (w tym wskaźniki, referencje, dynamiczne struktury danych).
9. Rodzaje błędów w programie komputerowym. Pojęcie wyjątku, obsługa wyjątków.
10. Definicja interfejsu użytkownika. Rodzaje interfejsów użytkownika.
11. Organizacja dostępu do baz danych w aplikacjach; przykładowe biblioteki i technologie.
12. Etapy powstawania programu komputerowego. Pojęcia: algorytm oraz implementacja algorytmu. Rola podprogramów w implementacji algorytmów.
13. Algorytmy zachłanne i dynamiczne, ich zastosowanie. Przykłady zastosowań.
14. Algorytmy rekurencyjne i iteracyjne. Przykłady rekurencyjnych i iteracyjnych algorytmów optymalnych i nieoptymalnych.
15. Algorytmy numeryczne. Uzasadnienie konieczności obliczeń z użyciem algorytmów numerycznych. Przykłady takich algorytmów.
16. Relacyjna baza danych, pojęcia podstawowe, proces projektowania bazy danych.
17. Postaci normalne tabeli, proces normalizacji, logika powiązań między tabelami w bazie danych.
18. Indeksy w bazach danych; definicja pojęcia, cel stosowania.
19. Język poleceń SQL. Klasyfikacja poleceń języka SQL (DDL, DQL, DML, TCL, DCL). Przykłady poleceń z każdej klasyfikacji. Odpowiedniość w QBE.



20. Operacje selekcji, projekcji, sortowania, agregacji danych oraz pola obliczeniowe w relacyjnej bazie danych. Przykłady realizacji takich operacji w SQL oraz QBE.
21. Rodzaje złączeń tabel w bazie danych oraz ich implementacja w języku SQL (co najmniej trzy przykłady).
22. System zarządzania bazą danych. Zasady działania takich systemów w trybie desktop oraz klient-serwer.
23. Problemy występujące we współbieżnym dostępie do danych w bazach danych. Pojęcie transakcji. Zbiór właściwości gwarantujących poprawne przetwarzanie transakcji (ACID).
24. Pojęcie hurtowni danych. Cechy hurtowni danych. Ładowanie danych do hurtowni danych – proces ETL.
25. Wielowymiarowe modele danych. Modele OLAP i OLTP. Zapytania analityczne na kostkach wielowymiarowych.
26. Uczenie nadzorowane i nienadzorowane; cechy charakterystyczne. Przykładowe metody dla każdego rodzaju uczenia. Przykłady zastosowań.
27. Sztuczne sieci neuronowe, architektura, proces uczenia. Przykłady zastosowań.
28. Specyfika uczenia się ze wzmocnieniem – cechy charakterystyczne. Przykłady zastosowań.
29. Etapy przedsięwzięcia projektowego.
30. Diagnozowanie jakości danych. Weryfikacja, czyszczenie i wzbogacanie danych.
31. Poziomy ustrukturalnienia danych. Transformacja danych jako element przygotowania danych do analiz (kodowanie, kategoryzacja, normalizacja, agregacja).
32. Etapy postępowania w procesie integracji danych w systemach informatycznych.
33. Architektura aplikacji w systemie operacyjnym Android i najważniejsze pliki składowe takiej aplikacji. /Architektura sprzętowa i programowa urządzeń sieciowych CISCO. (jedno z pytań, w zależności od wyboru przedmiotu w toku studiów)
34. Podstawowe komponenty i cykl życia aplikacji w systemie operacyjnym Android. / Podstawowe topologie fizyczne i logiczne sieci komputerowych. Protokoły TCP/IP, schematy adresowania. (jedno z pytań, w zależności od wyboru przedmiotu w toku studiów)
35. Wstępna eksploracja i graficzna prezentacja danych ilościowych i jakościowych.
36. Wizualizacja zależności występujących między parami zmiennych różnego typu (I-I, J-J, I-J).
37. Pojęcie szeregu czasowego. Metody identyfikacji składowych szeregu czasowego. Przykłady metod prognozowania szeregów czasowych z trendem oraz szeregów czasowych z trendem i wahaniami sezonowymi.



38. Rodzaje instrukcji w programie komputerowym. Klasyfikacja i znaczenie operatorów. Przykłady.
39. Przetwarzanie danych w arkuszu kalkulacyjnym Excel. Adresowanie względne i bezwzględne, moduł Model danych, tabele przestawne.
40. Kopia zapasowa jako element bezpieczeństwa danych. Strategie tworzenia kopii bezpieczeństwa, ich wady i zalety.
41. Jednokierunkowe funkcje hashujące – podstawowe własności i zastosowanie.
42. Ataki internetowe i sposoby obrony przed takimi atakami.
43. Modele sieci komputerowych. Zadania routera i przełącznika w sieci komputerowej. Usługi sieciowe DHCP i DNS.
44. Język znaczników na przykładzie HTML lub XML, przykłady zastosowań. Struktura dokumentu HTML i jego formatowanie za pomocą języka arkuszy stylów (CSS).
45. Wzorce architektoniczne w aplikacjach webowych. Wielowarstwowa struktura aplikacji internetowej współpracującej z bazą danych.
46. Programowanie liniowe i nieliniowe. Wykorzystanie modułu Solver (MS Excel) do rozwiązywania problemów badań operacyjnych. Przykłady zastosowań.
47. Struktura bilansu w systemie informacyjnym rachunkowości.
48. Różnica pomiędzy grafiką wektorową, a grafiką rastrową. Przykładowe formaty grafiki wektorowej i rastrowej.
49. Atrybuty definiujące komputerowy model koloru. Charakterystyka modeli kolorów: RGB, HSB i CMYK.
50. Język programowania VBA w aplikacjach MS Office, makra języka. Możliwości zastosowania.
51. Zmienne losowe (dyskretne i ciągłe), ich rozkłady i charakterystyki. Rozkład normalny i Centralne Twierdzenie Graniczne. Zastosowania.
52. Testy parametryczne i nieparametryczne. Wnioskowanie na podstawie przedziału ufności.
53. Modele danych panelowych, istota, przykłady zastosowań.

Specjalność: Analityka danych i modelowanie

1. Zadania języka programowania wbudowanego w środowisko aplikacji analitycznej (4GL w SAS). Przykłady wykorzystania myśliwości takiego języka w analityce danych.
2. Zagadnienie Text Mining. Możliwości zastosowań. Podstawowe zasady i techniki eksploracji danych tekstowych.
3. Wielokrotne modele regresyjne (prognostyczne i klasyfikacyjne), dobór zmiennych objaśniających do modelu. Weryfikacja merytoryczna i statystyczna w modelowaniu regresyjnym.



4. Decyzyjne drzewa regresyjne i klasyfikacyjne. Obszary zastosowań.
5. Koncepcja analizy asocjacji. Struktura, ocena i interpretacja reguł asocjacyjnych.
6. Analiza skupień; metody grupowania i podziału danych. Przykłady zastosowań.

Specjalność: Inżynieria zasobów danych

1. Specjalne tryby dopasowania argumentów funkcji w języku programowania Python.
2. Obsługa plików binarnych i tekstowych w programie w języku Python. Współpraca z bazami danych.
3. Modelowanie wiedzy w sieci semantycznej.
4. Semantyczne bazy danych. Zarządzanie wiedzą w semantycznej bazie danych.
5. Rozproszony system plików na przykładzie HDFS. Algorytm MapReduce.
6. Istota Big Data. Przykładowe zastosowania.