**Studia II stopnia Kierunek : Informatyka specjalność: Analityka danych/Data Science**



**Zagadnienia na egzamin magisterski**

1. Omówić jakie wyróżnia się etapy w przetwarzaniu i analizie danych w uczeniu maszynowym i krótko je scharakteryzować.

2. Omówić po co się stosuje walidację krzyżową w uczeniu maszynowym oraz jakie metody walidacji wyróżniamy.

3. Opisać kryteria oceny wyników w uczeniu maszynowym. Proszę omówić metody oceniana skuteczności predykcji klasyfikatorów.

4. Opisać metody stosowane do wstępnego przetwarzania danych? Dlaczego ten etap jest ważny w uczeniu maszynowym?

5. Omówić algorytm k-najbliższych sąsiadów (KNN).

6. Opisać metodę maszyny wektorów nośnych (SVM).

7. Opisać etap normalizacji, po co się go stosuje w uczeniu maszynowym?

8. Omówić klasyfikator lasu losowego.

9. Omówić regresję wieloraką i regresją grzbietową.

10. Omówić typy, metody treningu, strukturę i zasadę działania sztucznych sieci neuronowych.

11. Omówić metody optymalizacji hiper-parametrów klasyfikatorów, w tym algorytmy genetyczne.

12. Klasy i metody abstrakcyjne - podać przykłady praktycznego zastosowania.

13. Pojęcie hermetyzacji oraz metody praktycznej realizacji w językach obiektowo zorientowanych.

14. Omówić mechanizm polimorfizmu.

15. Omówić zasadnicze cechy mechanizmu dziedziczenia.

16. Omówić zasadnicze elementy modelu MVC (Model-View-Controler).

17. Omów krótko środowisko OpenMP.

18. Omów krótko środowisko MPI.

19. Omów taksonomię Flynna.

20. Wymień typy sieci głębokich.

21. Wyjaśnij pojęcia: uczenie nadzorowane, uczenie nienadzorowane, uczenie ze wzmacnianiem.

22. Wymień przykłady baz danych typu No-Sql.

23. Czym jest Hadoop ?

24. Opisz cechy Haara.

25. Wymień i scharakteryzuj krótko przykładowy algorytm binaryzacji obrazu.

26. Wymień 3 przykładowe testy do sprawdzania równości wariancji.

27. Scharakteryzuj metodę Anova.

28. Omów test Kruskala - Wallisa.

29. Scharakteryzuj technologię EJB.

30. Wyjaśnij pojęcie ścieżki krytycznej w teorii zarządzania projektami.

31. Omów metodykę SCRUM.

32. Omów podstawowe charakterystyki opisujące szereg czasowy.

33. Wyjaśnij skrót ETL.

34. Wymień 3 frameworki wykorzystywane w rozwiązywaniu problemów z dziedziny uczenia maszynowego.

35. Omów mechanizm dependency injection.

36. Wyjaśnij skrót REST.

37. Wyjaśnij pojęcia: bagging i boosting w kontekście uczenia maszynowego.

38. Co to jest boxplot ?

39. Omów Prawo Moore’a.

40. Co to jest chaos deterministyczny ?

41. Przedstawić i omówić budowę perceptronu. Gdzie przechowywana jest wiedza?

42. Omówić algorytm regresji liniowej. Czy jest to przykład uczenia nadzorowanego, czy nienadzorowanego? Dlaczego?

43. Na czym polega strategia uczenia one-versus-all? W przypadku jakich problemów jest ona stosowana?

44. Przedstawić metodę gradientu prostego. Jakie są jej założenia? W jakim celu się ją stosuje?

45. Wymienić przynajmniej 5 wzorców projektowych GoF. Wybrać i scharakteryzować 2 z nich. Jakie problemy dzięki nim są rozwiązywane?

46. Czym różni się proces od wątku? Wyjaśnić oba pojęcia i podać główne różnice.

47. Różnice pomiędzy relacyjną i nierelacyjną bazą danych. Przykłady zastosowań obu typów baz danych.

48. Opisać algorytm analizy głównych składowych (PCA). Na czym polega i kiedy się go stosuje?

49. Konteneryzacja a wirtualizacja. Opisz różnice i podaj przykłady technologii.

50. Proszę wyjaśnić na czym polega uczenie zespołowe?