

# Otázky ke zkoušce

## Matematické metody analýzy dat / Matematické metody pro analýzu dat

1. Co je multimodální data a multimodální systém.
2. Co jsou spojitá, diskrétní a sčítací data.
3. Co jsou shluky.
4. Jak je formulována úloha shlukování dat.
5. Jak je formulována úloha klasifikace dat.
6. Čím se liší úlohy shlukování a klasifikace dat.
7. Napište a vysvětlete mnohorozměrný statický normální model pro spojitá data ve tvaru rovnice.
8. Jak je definován šum mnohorozměrného statického normálního modelu pro spojitá data.
9. Napište a vysvětlete mnohorozměrný statický normální model pro spojitá data ve tvaru hustoty pravděpodobnosti.
10. Napište a vysvětlete kategorický model pro diskrétní data ve tvaru tabulky.
11. Napište a vysvětlete kategorický model pro diskrétní data ve tvaru pravděpodobnostní funkce.
12. Napište a vysvětlete Poissonův model pro sčítací data.
13. Jaké předpoklady má Poissonův model pro sčítací data.
14. Napište a vysvětlete Bayesovo pravidlo.
15. Jak lze Bayesovo pravidlo využít pro odhad parametrů modelu.
16. Jak se přepočítává statistika při odhadu mnohorozměrného statického normálního modelu.
17. Jak se určí dimenze počáteční informační matice při odhadu mnohorozměrného statického normálního modelu.
18. Jak se určí bodové odhady parametrů mnohorozměrného statického normálního modelu.
19. Jak funguje algoritmus Bayesovského odhadu mnohorozměrného statického normálního modelu.
20. Jak se přepočítává statistika při odhadu kategorického modelu.
21. Jak se určí dimenze počáteční statistiky při odhadu kategorického modelu.
22. Jak se určí bodové odhady parametrů kategorického modelu.
23. Jak funguje algoritmus Bayesovského odhadu kategorického modelu.
24. Jak se přepočítává statistika při odhadu Poissonova modelu.
25. Jak se určí bodové odhady parametrů Poissonova modelu.
26. Jak funguje algoritmus Bayesovského odhadu Poissonova modelu.
27. Co je model směsi distribucí.
28. Co jsou komponenty modelu směsi distribucí obecně.
29. Co je ukazovátka modelu směsi distribucí.
30. Co jsou normální komponenty modelu směsi distribucí.
31. Jak generujeme v Matlabu multimodální spojitá data z modelu směsi normálních komponent.
32. Co ovlivňuje vzdálenosti mezi shluky spojitých normálních dat.

33. Jaký je základní princip rekurzivního Bayesovského odhadu modelu směsi distribucí obecně.
34. Co je inicializace algoritmu rekurzivního Bayesovského odhadu modelu směsi distribucí obecně.
35. Jak inicializujeme počet normálních komponent pro algoritmus rekurzivního Bayesovského odhadu modelu směsi distribucí.
36. Jak inicializujeme počáteční statistiky normálních komponent pro algoritmus rekurzivního Bayesovského odhadu modelu směsi distribucí.
37. Jak se určí vzdálenosti dat od jednotlivých normálních komponent (proximity) při rekurzivním Bayesovském odhadu modelu směsi distribucí.
38. Jak se určí váhy komponent při rekurzivním Bayesovském odhadu modelu směsi distribucí.
39. Jak se určí bodový odhad ukazovátka při rekurzivním Bayesovském odhadu modelu směsi distribucí.
40. Jak probíhá update statistik normálních komponent s váženými daty při rekurzivním Bayesovském odhadu modelu směsi distribucí.
41. Co jsou kategorické komponenty modelu směsi distribucí.
42. Jak generujeme v Matlabu multimodální diskrétní data z modelu směsi kategorických komponent.
43. Jak se určí vzdálenosti dat od jednotlivých kategorických komponent (proximity) při rekurzivním Bayesovském odhadu modelu směsi distribucí.
44. Jak probíhá update statistik kategorických komponent s váženými daty při rekurzivním Bayesovském odhadu modelu směsi distribucí.
45. Co jsou Poissonovy komponenty modelu směsi distribucí.
46. Jak generujeme v Matlabu multimodální sčítací data z modelu směsi Poissonových komponent.
47. Co ovlivňuje vzdálenosti mezi shluky sčítacích Poissonových dat.
48. Jak inicializujeme počet Poissonových komponent pro algoritmus rekurzivního Bayesovského odhadu modelu směsi distribucí.
49. Jak inicializujeme počáteční statistiky Poissonových komponent pro algoritmus rekurzivního Bayesovského odhadu modelu směsi distribucí.
50. Jak se určí vzdálenosti dat od jednotlivých Poissonových komponent (proximity) při rekurzivním Bayesovském odhadu modelu směsi distribucí.
51. Jak probíhá update statistik Poissonových komponent s váženými daty při rekurzivním Bayesovském odhadu modelu směsi distribucí.
52. Co je Poissonova regrese.
53. Jak se odhadují parametry Poissonovy regrese pomocí linearizace a metody nejmenších čtverců.
54. Jak se odhadují parametry Poissonovy regrese pomocí metody maximální věrohodnosti.
55. Jak se určí predikce dat pomocí Poissonovy regrese.
56. Co jsou trénovací a testovací data.
57. Vysvětlete princip naivního Bayesova klasifikátoru.
58. Jak funguje klasifikace dat pomocí metody naivního Bayese.
59. Co je logistická regrese.
60. Jak se odhadují parametry logistické regrese pomocí metody maximální věrohodnosti.
61. Jak funguje klasifikace dat pomocí logistické regrese.

62. Jak se určí přesnost klasifikace.
63. Co je chybová matice.
64. Vysvětlete princip algoritmu k-means pro shlukování dat.
65. Jak se určí vzdálenosti mezi datovými body při shlukování dat pomocí k-means.
66. Vysvětlete princip algoritmu k-medoids pro shlukování dat.
67. Vysvětlete princip algoritmu fuzzy shlukování c-means.
68. Vysvětlete princip algoritmu DBSCAN pro shlukování dat na základě hustoty.
69. Vysvětlete princip hierarchického shlukování.
70. Jak funguje klasifikace dat pomocí algoritmu k-nejbližších sousedů.
71. Jak funguje klasifikace dat s rozhodovacími stromy.
72. Jak se určí nejlepší rozhodovací strom pomocí Giniho indexu nebo entropie.
73. Co jsou náhodné lesy.
74. Jak funguje klasifikace dat pomocí podpůrných vektorových strojů.
75. Vysvětlete základní princip klasifikace dat pomocí neuronových sítí.