

Przykładowe pytania egzaminacyjne z zajęć

Credit Scoring - automatyzacja procesu biznesowego

1. Co to jest Gini?
2. Co to jest WoE?
3. Co to jest ROC?
4. Co to jest CAP?
5. Metoda kalibracji, na czym polega?
6. Jak wygląda etap preselekcji zmiennych?
7. Jakie mamy statystyki stabilności zmiennych?
8. Jakie mamy statystyki predykcyjności zmiennych?
9. Jakie mamy metody oceny modeli?
10. Jak tworzymy segmenty / atrybuty zmiennych ciągłych? (Kategoryzacja zmiennej ciągłej, binning)
11. Jak tworzymy próby losowe (treningowy, walidacyjny) do modelowania?
12. Jakie kryteria decydują o dobrym modelu?
13. Co to jest bad rate?
14. Co to jest ryzyko?
15. Co to jest KL?
16. Co to jest IV?
17. Co to jest logit?
18. Na czym polega monitoring modelu?
19. Co to jest Reject Inference?
20. Wymień etapy budowy modelu akceptacyjnego uwzględniającego metody Reject Inference?
21. Jak definiuje się zdarzenie default?
22. Co to jest silnik / system decyzyjny?
23. Co to jest ABT?
24. Wymień przykładowe zmienne ABT?
25. Co to znaczy, że nie wolno budować modelu na danych z przyszłości?
26. Co to jest współliniowość i jak się ją mierzy?
27. Podaj przykłady zmiennych ABT, które mogą mieć moc predykcyjną Giniego około 90%?
28. Czy model może mieć wartość Giniego 90%?
29. Dlaczego wykonuje się poprawki zmiennych w finalnym modelu ze względu na stabilność?
30. Czym jest stabilność w modelowaniu, dlaczego jest ważna?
31. Dlaczego wykonuje się raport ewolucji Giniego w czasie?
32. Dlaczego tworzymy raporty zmiennych, jakie składowe raportu są ważne?
33. Na czym polega klasteryzacja zmiennych? Co daje?
34. Na czym polegają metody krokowe wyboru zmiennych?
35. Na czym polega metoda Score wyboru zmiennych w regresji logistycznej?

36. Jaką procedurą liczymy mierniki współliniowości?
37. Jaką procedurą liczy się statystyka KS?
38. Jaką procedurą liczy się statystykę Glniego?
39. Jak liczymy oceny cząstkowe?
40. Jaka zmienna w karcie skoringowej jest najważniejsza?
41. Jakie banchmarki stosujemy przy preselekcji zmiennych?
42. Jakie banchmarki stosujemy przy ocenie modeli?
43. Jakie warunki powinna spełniać dobra kategoryzacja zmiennej ciągłej, podział na segmenty / atrybuty, binning?
44. Jak liczymy opłacalność procesu akceptacji kredytowej?
45. Co to jest krzywa profit?
46. Jak wyznaczamy cut-off dla modelu?
47. Co to jest strata oczekiwana i jak się ją wyznacza?
48. Co to jest PD i jak się je wyznacza?
49. Jak można estymować ryzyko, PD dla wniosków odrzuconych?
50. Jak moc predykcyjna modelu wpływa na profit procesu akceptacji kredytowej?
51. Jak tworzymy segmenty / grupy dla zmiennej jakościowej, binning?
52. Do czego służą statystyka Entropi i Index Giniego?
53. Dlaczego używamy dwóch prób train i valid?
54. Dlaczego definiujemy status nieokreślonego w definicji default?
55. Czym się różni modelowanie aplikacyjne od behawioralnego?
56. Czym się różni definicja default aplikacyjna od behawioralnej?
57. Do jakiej grupy przydziela kod do skórowania, jeśli jest to wartość nieprzewidziana w nierównościach wszystkich grup, inaczej jaka jest wartość domyślna oceny cząstkowej?
58. Jakie plusy i minusy mamy z używania większej liczby modeli na segmentach zamiast jednego dużego?
59. Dlaczego czasami tworzy się interakcje zmiennych w modelu?
60. Czym różni się interakcja zmiennych od modelu z tymi samymi zmiennymi osobno?
61. Co to jest interakcja zmiennych?
62. Jak dobierać okres obserwacji w definicji zdarzenia default, czy powinien być długi czy krótki?
63. Czy moc modelu liczona statystyką Giniego może znacząco się zmienić po wdrożeniu?
64. Proszę podać przykład strategii akceptacji, jakiejś reguły, która nie musi być związana z modelem skoringowym?
65. Na czym polega model biznesowy tania akwizycja – droga sprzedaż krzyżowa?
66. Dlaczego w strategiach akceptacji warto używać modeli response?
67. Definicja zdarzenia response?
68. Co to jest response rate?
69. Proszę podać prosty przykład modelu CLTV dla modelu biznesowego tani kredyt ratalny i droga gotówka?
70. Co to jest CLTV?
71. Proszę podać przykład segmentacji portfela aplikacyjnego (wniosków kredytowych), która posiada sens biznesowy?
72. Czym się różni zmienna aplikacyjne od behawioralnych?
73. Proszę podać typowe wartości Giniego dla modeli aplikacyjnych, behawioralnych?
74. Proszę podać typowe wartości statystyk Gini, KS, IV dla zmiennej?

75. O czym świadczy zmiana znaku estymowanej bety metodą WoE w regresji logistycznej?
76. Czym różnią się metody liczenia ocen punktowych Dummy od WoE?
77. Jaką funkcją kalibruje się ocenę punktową do wartości prawdopodobieństwa modelowanego zdarzenia?
78. Co to są statystyki Gains i Lift?
79. Do czego służą statystyki Gini, Gains, Lift, kiedy ich używamy, czy każdy model, który ma dużą wartość Giniego ma też dużą wartość Lift lub Gains?
80. Jak kształt CAP lub ROC jest bardziej pożądany dla modeli stosowanych do akceptacji / odrzucania wniosków?
81. Co to jest metoda Shapley'a, jakie raporty można wykonać?
82. Proszę omówić zalety i wady stosowania karty skoringowej względem modeli AI w procesie akceptacji kredytowej.