

## **Credit Scoring - business process automation**

### **Example exam questions**

1. What is Gini?
2. What is WoE?
3. What is ROC?
4. What is CAP?
5. What is the idea and purpose of methods of calibrations?
6. How is the process of variable preselection?
7. What are the stability statistics of variables?
8. What are the statistics of prediction of a variable?
9. What methods do we have for assessing models?
10. How do we create segments / categories of continuous variables? (Categorization of a continuous variable, binning)
11. How do we create random samples (training, validation) for modeling?
12. What criteria identify a good model?
13. What is a bad rate?
14. What is risk?
15. What is KL?
16. What is IV?
17. What is logit?
18. What is model monitoring?
19. What is Reject Inference?
20. List the steps of model building for an acceptance process that includes Reject Inference.
21. How is the default event defined?
22. What is the decision engine/scoring engine/decision system?
23. What is ABT?
24. List examples of ABT variables.
25. What does it mean that you must not build the model on data taken from a future, how do we define a future?
26. What is collinearity and how is it measured?
27. Give examples of ABT variables that can have a Gini predictive power of about 90%?
28. Can the model have a Gini value of 90%?
29. Why are variable corrections have to be made in the final model for stability reasons?
30. What is stability in modeling, why is it important?
31. Why is the report of Gini's evolution over time being created?
32. Why do we create variable reports, what components of the report are important?
33. What is variable clustering? What does it give?
34. What are the stepwise methods for selecting variables?
35. What is the Score method for selecting variables in logistic regression?
36. What procedure do we calculate collinearity measures?
37. What procedure counts KS statistics?

38. What procedure calculates Gini statistics?
39. How do we calculate partial scores?
40. Which variable in the scorecard is the most important?
41. What benchmarks do we use for variable preselection?
42. What benchmarks do we use to assess models?
43. What conditions should be met by good categorization of a continuous variable, segmentation/binning?
44. How do we calculate the profitability of the credit acceptance process?
45. What is the profit curve?
46. How do we determine the cut-off for the model?
47. What is the expected loss and how is it determined?
48. What is PD and how is it determined?
49. How can be estimated the risk, PD for rejected applications?
50. How does the predictive power of the model affect the profit of the credit acceptance process?
51. How do we create segments / groups for a qualitative variable, binning?
52. What are the Entropy and Gini Index statistics for?
53. Why do we use two sets, training and validation?
54. Why do we define the indeterminate status in the default definition?
55. What is the difference between acceptance/application and behavioral modeling?
56. What is the difference between the definition of default in acceptance/application and behavioral usage?
57. To what group does he assign the scoring code if it is a value not provided for in inequalities of all groups, otherwise what is the default value of partial score?
58. What pros and cons do we have from using more models on segments instead of one big common model?
59. Why are variable interactions sometimes created in the model?
60. What is the difference between the interaction of variables used in the model with the same variables used separately?
61. What is variable interaction?
62. How to choose the observation period in the definition of default event, should it be long or short?
63. Can the power of the model as determined by Gini statistics change significantly after implementation?
64. Please provide an example of an acceptance strategy, some rule that does not have to be associated with scoring models?
65. What is the business model for cheap acquisition - expensive cross selling?
66. Why should you use response models in acceptance strategies?
67. Response definition of event.
68. What is response rate?
69. Please provide a simple example of the CLTV model for the business model: cheap installment loan and expensive cash loan?
70. What is CLTV?
71. Please provide an example of segmentation of the acceptance/application portfolio (loan applications) which has business sense.
72. What is the difference between application and behavioral variables?

73. Please provide typical Gini values for acceptance/application and behavioral models
74. Please provide typical values of the Gini, KS, IV statistics for variables.
75. What does the change of the sign in estimated beta by the WoE approach in logistic regression?
76. What is the difference between Dummy and WoE approaches in scorecard building?
77. What function is used for calibration from score into modeled probability of event?
78. What are Gains and Lift statistics?
79. What are the statistics of Gini, Gains, Lift? When do we use them? Does model that have a large Gini value also have a high value of Lift or Gains?
80. What shape of CAP or ROC is more desirable for models used for acceptance / rejecting applications? Do we focus on a part of the curve or all parts are important?
81. What is Shapley method? What reports can we do?
82. Please present pros and cons of usage of risk scorecard models contra AI models in credit risk acceptance process.