

“Ma’lumotlar tahlili” fanidan yakuniy nazorat savollari

“Bilish” darajasiga mansub savollar:

1. Ma’lumotlar tahlili (Data Analysis) fanining maqsad va vazifasi nimadan iborat?
2. Ma’lumotlar ilmi (Data Science) va Ma’lumotlar tahlili (Data Analysis) o’rtasidagi farqlarni izohlab bering.
3. Mashinali o’qitish (Machine Learning)ga ta’rif bering.
4. Strukturalangan va Strukturalanmagan ma’lumotlar tushunchalariga izoh bering.
5. O’lchov darajalari va ularning turlari(Nominal, Ordinal, Interval, Ratio).
6. Uzluksiz(Continuous) va Diskret ma’lumotlar tushunchalariga ta’rif bering.
7. O’rtacha, Mediana va Moda (Mode) tushunchalariga izoh bering.
8. Dispersiya va Standart og‘ish (Standard Deviation) tushunchalariga ta’rif bering.
9. Kovariatsiya va Korrelyatsiya tushunchalarini ta’riflang.
10. Normal taqsimot (Normal Distribution) tushunchasini izohlab bering.
11. Ma’lumotlar tahlilida Shovqin (Noise) nimani anglatishini izohlang.
12. Anomal qiymat(Outlier) tushunchasiga ta’rif bering.
13. Tushirib qoldirilgan qiymat turlari (MCAR, MAR va MNAR)ni izohlab bering.
14. Label Encoding va One-Hot Encoding kodlash turlariga ta’rif bering.
15. Min-Max masshtablash usuli nima uchun qo‘llanilishi izohlang.
16. Chiziqli regressiya (Linear Regression) tushunchasiga ta’rif bering.
17. Gradiyent pastlash(Gradient Descent) algoritmiga ta’rif bering.
18. KNN algoritmidagi “K” qiymatini nimani anglatishini izohlang.
19. Qarorlar daraxti (Decision Tree) algoritmini izohlab bering.
20. CRISP-DM metodologiyasining bosqichlarini sanab o‘ting.

“Tushunish” darajasiga mansub savollar:

21. Data analysis jarayoni qanday bosqichlardan iborat?
22. Nominal va Ordinal o’lchov darajalarini o’rtasidagi farq nimadan iboratligini misollar bilan tushuntiring?
23. Interval va Ratsional o’lchov darajalari orasidagi asosiy farqni tushuntiring.
24. O’rtacha va Mediana qiymatlari qachon farq qilishi mumkinligini tushuntiring.
25. Standart og‘ish katta bo‘lsa ma’lumotlar haqida qanday xulosa chiqarish mumkin?
26. Musbat va manfiy og‘ish (Skewness) nimani anglatadi? Misol keltiring.
27. Tekis taqsimot (Uniform Distribution) va Normal taqsimot (Normal Distribution) o’rtasidagi farqni tushuntiring.
28. Nuqtali (Scatter plot) diagrammadan nimalarni tushunib olish mumkin?
29. Z-score yordamida anomal qiymatlarni aniqlash jarayonini tushuntiring.
30. IQR usuli bilan anomal qiymatlarni aniqlashning mantig‘ini tushuntiring.
31. Tushib qolgan qiymatlarni o’rtacha qiymat ortqali to‘ldirishning kamchiliklarini tushuntiring.
32. One-Hot Encoding nima uchun zarur? Misol orqali tushuntiring.

33. Z-score va Min-Max usullari qachon qo'llaniladi?
34. MSE va RMSE metirkalari model sifatini qanday baholaydi?
35. Gradiyent pastlash algoritmidagi o'qitish darajasi (Learning rate) kichik yoki katta qilib tanlab yuborilsa nima bo'ladi?
36. K-means klasterizatsiya algoritmi qanday ishlashini tushuntiring.
37. Overfitting va Underfitting tushunchalarini izohlang.
38. R^2 metirkasi nimani anglatishini misollar orqali tushuntiring?
39. Chalkashlik matritsasi (Confusion matrix) tushunchasi va u qachon ishlatilishini tushuntiring?
40. SEMMA metodologiyasidagi har bir bosqichining mazmunini tushuntiring.

“Amalda qo'llash” darajasiga mansub savollar

41. Quyidagi ma'lumotlar uchun O'rtacha qiymat, Median va Moda ni hisoblang:
[4, 7, 7, 9, 12, 15, 7]
42. Berilgan ma'lumotlar to'plami uchun Standart og'ishni hisoblang: [2, 4, 4, 4, 5, 5, 7, 9]
43. Nuqtalar: (1,2),(2,3),(6,7) berilgan. Yangi (3,4) nuqta uchun Evklid va Manhattan usullari orqali eng yaqin nuqtani hisoblang.
44. Quyidagi Data set uchun Varinace(Tarqalish) ni hisoblang. Dataset: 5,7,9,11
45. Min-Max normallashtirish usuli yordamida quyidagi ma'lumotlarni 0-1 oralig'iga o'tkazing. Ma'lumot: [10, 20, 30, 40, 50]
46. Mahsulot kategoriyalari (A, B, C) uchun One-Hot Encoding jadvalini tuzing.
47. IQR usulidan foydalanib quyidagi Data Set uchun anomal(outlier) qiymatlarni hisoblang. Data Set: 3, 5, 7, 11, 6, 4, 8, 58
48. IQR usulini qo'llab, [1, 2, 3, 4, 5, 100] ma'lumotlaridagi anomal(outlier) qiymatlarni aniqlang.
49. Quyidagi Data Setni Min-Max funksiyasi orqali normallashtiring.
Data Set : 5, 7, 8, 10, 100
50. Berilgan chalkashlik matritsasi asosida Precision, Recall va Accuracyni hisoblang: TP=50, FP=10, FN=5, TN=35
51. Quyidagi data set uchun o'rtacha qiymat, variance(tarqalish) va SD(standart og'ish) ko'rsatkichlarini hisoblang. Dataset: 2, 4, 6, 8, 10
52. Quyidagi data set uchun IQR usuli asosida Median, Q1 va Q3 qiymatlarini hisoblang. Dataset: 3, 5, 7, 9, 11
53. Quyidagi Data set uchun Kovariatsiya(Covariance) qiymatini hisoblang.
Data Set: X: 1,2,3
 Y: 2,4,6
54. Quyidagi Data set uchun Korrelyatsiya(Correlation) koeffitsiyentini hisoblang
Data Set: X: 1,2,3
 Y: 2,4,6
55. Regressiya usuli orqali tushib qolgan qiymatlarni to'ldirish jarayonini misollar bilan ko'rsating.

56. Gradient pastlash (Gradient Descent) algoritmini qadamma-qadam misollar orqali ko'rsating.
57. Quyidagi bashoratlangan va haqiqiy qiymatlar asosida MSE(O'rtacha kvadrat xatolik) ni hisoblang. Bunda $y: 5,7,9$ $\hat{y}: 6,6,10$
58. CRISP-DM jarayonini haqiqiy biznes muammosiga qo'llab ko'rsating.
59. Quyidagi kategorik ma'lumotlarni One-hot encoding orqali sonli ko'rinishga o'tkazing. Kategorik ma'lumot: Oq, qizil, sariq
60. Quyidagi bashoratlangan va haqiqiy qiymatlar asosida RMSE(O'rtacha kvadrat xatolik) ni hisoblang. Bunda $y: 5,7,9,10$ $\hat{y}: 6,6,10,10$

“Tahlil” darajasiga mansub savollar:

61. Ma'lumotlar tahlili (Data Analysis) va Mashinali o'qitish (ML) o'rtasidagi o'zaro bog'liqlikni tahlil qiling.
62. Normal taqsimot va og'gan taqsimot ma'lumotlariga qanday statistik usullar qo'llanilishini tahlil qiling.
63. O'rtacha qiymat va Medianning qaysi holda yaxshiroq markaziy tendensiya ko'rsatkichi ekanligini tahlil qiling.
64. Korelyatsiya ko'rsatkichi nimani anglatishini misollar orqali tahlil qiling.
65. Gistogramma (Histogram) va ustunli diagramma (Bar Chart) o'rtasidagi tuzilmaviy farqlarni tahlil qiling.
66. Z-score va IQR usullarining qaysi holatlarda ishonchsiz ekanligini tahlil qiling.
67. Tushirib qoldirilgan qiymat turlari(MCAR, MAR va MNAR) ga qarab har xil to'ldirish usullarining ta'sirchanligini tahlil qiling.
68. Label Encoding va One-Hot Encoding qaysi algoritmlarda muammo keltirib chiqarishi mumkin?
69. Min-Max mastablash vs Z-score usullaridan qaysi biri anomal qiymatlar (Outlier) mavjudligida samarali bo'lishini tahlil qiling.
70. MSE va RMSE ko'rsatkichlarini taqqoslab, ularning turli kontekstlardagi afzalliklarini tahlil qiling.
71. Chiziqli regressiya va Polynomial regressiya o'rtasidagi model murakkabligini tahlil qiling.
72. KNN algoritmidagi K qiymatining kichik va katta bo'lishi modelga qanday ta'sir qilishini tahlil qiling.
73. Qarorlar daraxti (Decision Tree) da bo'linish kriteriyasi (Gini vs Entropy) ta'sirini tahlil qiling.
74. Tasodifiy o'rmon (Random Forest) algoritmining Qarorlar daraxti (Decision Tree) algoritmiga nisbatan qanday ustunliklari borligini tahlil qiling.
75. K-means algoritmidagi K soni noto'g'ri tanlansa qanday oqibatlariga olib kelishini tahlil qiling.
76. Ensemble Learning usullari individual modellardan qanday farq qilishini tahlil qiling.
77. Overfitting sabablari va uning model aniqligiga ta'sirini tahlil qiling.

78. Ikkilik(Binary) va Ko'p sinfli (Multiclass Classification) sinflashtirish o'rtasidagi farqlarni tahlil qiling.

79. CRISP-DM va SEMMA metodologiyalarini qiyosiy tahlil qiling.

80. Joriy qilish (Deployment) bosqichida model samaradorligi pasayib ketishining mumkin bo'lgan sabablarini tahlil qiling.

“Sintez va xulosa” darajasiga mansub savollar

81. Ma'lumotlar to'plamiga qarab qaysi vizualizatsiya usuli mos ekanligini qanday baholaysiz?

82. Ma'lumotlarni qayta ishlash bosqichida ma'lumotlarni to'ldirish usullarining qaysi biri ishonchligini qanday baholaysiz?

83. Model baholashda Accuracy yoki F-score metrikalarining muhimlik darajasini misollar asosida baholang.

84. KNN va Qarorlar daraxti (Decision Tree) algoritmlaridan birini tanlash uchun qanday mezonlardan foydalanish mumkin?

85. Tasodifiy o'rmon(Random Forest) va Qarorlar daraxti (Decision Tree) algoritmlari orasida qaysi biri ma'lumotlarga mosligini qanday mezonlar asosida baholaysiz?

86. Regressiya modelida $RMSE=5$ va $RMSE=15$ bo'lgan ikki modeldan qaysinisi yuqori natija berishini baholang va sababi nimadalgini izohlang.

87. K-means va DBSCAN klasterlash algoritmlari orasida geometrik shakldagi ma'lumotlar uchun qaysinisi yaxshiroq?

88. Gradiyent pastlash (Gradient Descent) algoritmidagi o'rganish darajasi(Learning rate) qiymatini tanlashda nimalarga e'tibor berish kerak?

89. Overfitting muammosiga ega modelni deployment qilish to'g'rimi? Amaliy misollar asosida baholang.

90. Neyron tarmoqlari uchun o'qitiladigan ma'lumotlar ustida Min-Max masshtablash vs Standartlashtirish usullaridan qaysi birini qo'llagan afzal?

91. CRISP-DM metodologiyasi haqiqiy loyiha uchun qanchalik mos keladi? Afzallik va kamchiliklarini baholang.

92. Ikkilik(Binary) sinflashtirish modelida Precision va Recall metrikalari o'rtasida kelib chiqishi mumkin bo'lgan ziddiyatni baholang.

93. Ma'lumotlar sifatini baholashda qanday mezonlar eng muhim hisoblanadi?

94. Raqamli va kategoriyaviy ma'lumotlar uchun bir xil statistik usullarni qo'llab bo'lmasligining sababini izohlang.

95. Xususiyatlar o'rtasidagi kuchli bog'lanish model natijalariga qanday ta'sir qilishini bayon qiling.

96. Real loyihalarda ma'lumotlarni tozalash nima uchun eng ko'p vaqt olishini asoslab bering.

97. Berilgan modelning train data setda $accuracy=0.99$ va test data setda $accuracy=0.65$ qiymatlar bo'lsa, modelni qanday baholash mumkin?

98. Kategorik ma'lumotlarni kodlash (Data Encoding) usullarini ma'lumotlar hajmi va model turi nuqtai nazaridan baholang.

99. Regressiya modelini yaxshilashning beshta samarali usulini sanab bering.
100. SEMMA metodologiyasi CRISP-DM ga nisbatan qachon qulay bo'ladi?

Yakuniy nazorat imtihoni biletida 5 ta savol (har bir darajaga mansub bittadan) beriladi. Har bir savolning javobi 10 ballik mezonda baholanadi. Jami 50 ball.

**Raqamli texnologiyalar konvergensiya
kafedrasi katta o'qituvchisi**

X.Toliyev

**Raqamli texnologiyalar konvergensiya
kafedrasi katta o'qituvchisi**

N.Karimov