
Bagian 1: Konsep Dasar Support

1. Apa yang diukur oleh nilai support dalam association rule mining?
Jawaban: **c. Seberapa sering itemset muncul dalam seluruh transaksi**
2. Jika itemset {A,B} muncul dalam 15 dari 100 transaksi, berapa support-nya?
Jawaban: **d. 15%**
3. Minimum support digunakan untuk:
Jawaban: **c. Menentukan batas bawah frekuensi itemset agar dianggap penting**

Bagian 2: Konsep Confidence

4. Confidence dari rule $A \rightarrow B$ mengukur:
Jawaban: d. Probabilitas kemunculan B jika A sudah muncul
5. Jika support {A,B} = 25% dan support {A} = 50%, maka confidence rule $A \rightarrow B$ adalah:
Jawaban: a. 50%
6. Suatu rule dikatakan kuat jika:
Jawaban: c. Memiliki confidence tinggi dan support memadai

Bagian 3: Hubungan Support dan Confidence

7. Mengapa suatu rule dengan confidence tinggi tapi support rendah bisa menyesatkan?
Jawaban: d. Karena mungkin hanya kebetulan pada sedikit transaksi
8. Jika support {X} = 60% dan support {X,Y} = 30%, maka:
Jawaban: c. Confidence $X \rightarrow Y$ adalah 50%
9. Manakah yang TIDAK benar tentang support dan confidence?
Jawaban: c. Confidence selalu lebih besar dari support
10. Dalam market basket analysis, jika support {susu,tisu} = 8% dan confidence susu \rightarrow tisu = 70%, artinya:
Jawaban: d. 70% pembeli susu juga beli tisu

Bagian 4: Aplikasi Praktis

11. Jika minimum support = 5% dan minimum confidence = 50%, rule mana yang LAYAK?
Jawaban: a. Hanya Rule 3
 12. Mengapa confidence saja tidak cukup untuk mengevaluasi kualitas rule?
Jawaban: a. Karena bisa menyesatkan jika item consequent sudah sangat umum
-

Bagian 5: Itemset Lattice dalam Association Analysis

13. Apa yang dimaksud dengan itemset lattice dalam association analysis?
Jawaban: a. Struktur data yang menyimpan semua kemungkinan kombinasi itemset dari suatu dataset
 14. Dalam itemset lattice, itemset {A, B} merupakan superset dari itemset apa?
Jawaban: c. {A} dan {B}
 15. Jika diketahui itemset {Milk, Bread} memiliki support 30%, dan {Milk} memiliki support 50%, bagaimana hubungan keduanya dalam lattice?
Jawaban: a. {Milk} adalah subset dari {Milk, Bread}
 16. Mengapa itemset lattice penting dalam association rule mining?
Jawaban: a. Mempermudah penghitungan frequent itemset dengan prinsip antimonotonicity
 17. Jika {A, B} adalah frequent itemset, mana yang pasti juga frequent berdasarkan properti lattice?
Jawaban: b. {A}
 18. Apa yang terjadi pada itemset lattice jika minimum support dinaikkan?
Jawaban: c. Jumlah frequent itemset berkurang
 19. Dalam konteks lattice, apa yang dimaksud dengan closed itemset?
Jawaban: a. Itemset yang tidak memiliki superset dengan support yang sama
 20. Jika {Diaper, Beer} adalah itemset dengan support tertinggi, apa implikasinya dalam lattice?
Jawaban: a. Semua subsetnya juga memiliki support tinggi
-

Bagian 6: Algoritma Apriori

21. Apa tujuan utama dari algoritma Apriori dalam association analysis?
Jawaban: b. Menemukan semua frequent itemset yang memenuhi minimum support
 22. Tahap pertama dalam algoritma Apriori adalah...
Jawaban: b. Membangkitkan candidate itemset berukuran 1 (1-itemset)
 23. Prinsip dasar Apriori yang memungkinkan optimasi pencarian frequent itemset adalah...
Jawaban: a. Jika sebuah itemset tidak frequent, semua supersetsnya juga tidak frequent
 24. Mengapa algoritma Apriori menggunakan pendekatan generate-and-test?
Jawaban: a. Untuk mengurangi jumlah candidate itemset yang perlu diperiksa
-

Bagian 7: Sifat Antimonotone

25. Apa yang dimaksud dengan sifat antimonotone dalam association analysis?
Jawaban: a. Jika sebuah itemset tidak memenuhi minimum support, semua supersetsnya juga tidak memenuhi
26. Contoh penerapan sifat antimonotone adalah...
Jawaban: a. Jika {A} tidak frequent, maka {A,B} tidak perlu diperiksa

27. Bagaimana sifat antimonotone membantu optimasi algoritma Apriori?
Jawaban: a. Dengan memangkas candidate itemset yang tidak mungkin frequent
28. Jika $\{A, B\}$ memiliki support 20% dan minimum support adalah 25%, apa yang dapat disimpulkan tentang $\{A, B, C\}$?
Jawaban: a. $\{A, B, C\}$ pasti memiliki support $\leq 20\%$
-

Bagian 8: Konsep Lift

29. Apa yang diukur oleh lift dalam association rule mining?
Jawaban: d. Seberapa kuat hubungan antara antecedent dan consequent dibanding kemunculan acak
30. Jika $\text{lift}(A \rightarrow B) = 1$, artinya:
Jawaban: c. A dan B independen (tidak ada hubungan)
31. Nilai lift yang dianggap menunjukkan hubungan positif yang kuat adalah:
Jawaban: b. >1
-

Bagian 9: Leverage

32. Leverage mengukur:
Jawaban: b. Selisih antara frekuensi gabungan aktual dan yang diharapkan jika independen
33. Jika $\text{support}\{A\}=0.4$, $\text{support}\{B\}=0.3$, dan $\text{support}\{A,B\}=0.25$, berapa $\text{leverage}(A \rightarrow B)$?
Jawaban: a. 0.05
34. Leverage bernilai negatif ketika:
Jawaban: d. Itemset muncul lebih jarang dari yang diharapkan
-

Bagian 10: Conviction

35. Conviction digunakan untuk mengukur:
Jawaban: d. Seberapa sering rule memprediksi salah dibanding asumsi independensi
36. Jika $\text{conviction}(A \rightarrow B) = 1.5$, interpretasinya adalah:
Jawaban: c. Rule 1.5 kali lebih baik dari prediksi acak
37. Conviction bernilai ∞ ketika:
Jawaban: d. Confidence = 100%
-

Bagian 11: Perbandingan Metrik

38. Manakah yang sensitif terhadap kemunculan item consequent yang sudah umum?
Jawaban: a. Lift dan conviction
39. Untuk rule yang sama, jika:
- Lift = 2.5

- Leverage = 0.15
 - Conviction = 1.8
Interpretasinya adalah:
Jawaban: d. Hubungan positif kuat ($lift > 1$, $leverage > 0$, $conviction > 1$)
40. Jika lift tinggi tapi leverage rendah, kemungkinan karena:
Jawaban: d. Support consequent sangat kecil
-

Bagian 12: Frequent Itemset Generation

41. Apa tujuan utama dari Frequent Itemset Generation?
Jawaban: a. Menemukan semua itemset yang memenuhi minimum support
42. Algoritma manakah yang menggunakan pendekatan "generate-and-test" untuk Frequent Itemset Generation?
Jawaban: c. Apriori
43. Prinsip apa yang digunakan Apriori untuk mengurangi pencarian itemset?
Jawaban: d. Sifat antimonotone (jika itemset tidak frequent, supersetnya juga tidak)
-

Bagian 13: Rule Generation

44. Setelah mendapatkan frequent itemset, Rule Generation bertujuan untuk:
Jawaban: b. Membuat aturan asosiasi yang memenuhi minimum confidence
45. Bagaimana confidence rule $A \rightarrow B$ dihitung?
Jawaban: c. $\text{Support}(A \cup B) / \text{Support}(A)$
46. Untuk itemset $\{A, B, C\}$, berapa banyak candidate rule yang mungkin dihasilkan?
Jawaban: a. 6 ($A \rightarrow BC$, $B \rightarrow AC$, $C \rightarrow AB$, $AB \rightarrow C$, $AC \rightarrow B$, $BC \rightarrow A$)
-

Bagian 14: Proses Gabungan

47. Urutan yang benar dalam association rule mining adalah:
Jawaban: d. Frequent Itemset Generation \rightarrow Rule Generation
48. Mengapa Frequent Itemset Generation dilakukan terlebih dahulu sebelum Rule Generation?
Jawaban: a. Untuk mengurangi jumlah candidate rule yang harus dihitung confidence-nya
49. Jika itemset $\{X, Y, Z\}$ tidak frequent, apa implikasinya terhadap rule generation?
Jawaban: b. Tidak ada rule dari itemset ini yang akan dipertimbangkan
-

Bagian 15: Optimasi Proses

50. Teknik pruning dalam rule generation bertujuan untuk:
Jawaban: d. Menghilangkan rule yang tidak memenuhi minimum confidence

51. Apa keuntungan menggunakan metrik tambahan seperti lift setelah rule generation?

Jawaban: d. Mengidentifikasi rule yang benar-benar menarik (tidak sepele)

52. Dalam FP-Growth, frequent itemset dihasilkan dengan:

Jawaban: c. Membangun FP-tree dan mining conditional pattern base

Berikut adalah lanjutan soal multiple choice lengkap beserta jawaban final untuk Bagian Strategi Merging ($F_{k-1} \times F_1$ dan Apriori Merging Strategy) dalam Association Rule Mining, digabung langsung ke daftar sebelumnya — menjadikannya total 62 soal:

Bagian 16: $F_{k-1} \times F_1$ Merging Strategy

53. Apa yang dimaksud dengan $F_{k-1} \times F_1$ Merging Strategy dalam pembangkitan candidate itemset?

Jawaban: a. Menggabungkan semua frequent (k-1)-itemset dengan semua frequent 1-itemset

54. Kelemahan utama dari $F_{k-1} \times F_1$ Merging Strategy adalah:

Jawaban: c. Menghasilkan terlalu banyak candidate itemset yang sebenarnya tidak perlu

55. Dalam $F_{k-1} \times F_1$, jika $F_2 = \{\{A,B\}, \{A,C\}, \{B,C\}\}$ dan $F_1 = \{\{A\}, \{B\}, \{C\}\}$, candidate itemset berukuran 3 yang dihasilkan adalah:

Jawaban: b. $\{\{A,B,C\}\}$

Bagian 17: Apriori Merging Strategy

56. Apriori Merging Strategy berbeda dengan $F_{k-1} \times F_1$ karena:

Jawaban: d. Hanya menggabungkan (k-1)-itemset yang memiliki (k-2) item pertama sama

57. Keuntungan utama Apriori Merging Strategy dibanding $F_{k-1} \times F_1$ adalah:

Jawaban: c. Menghasilkan lebih sedikit candidate itemset yang tidak perlu

58. Dalam Apriori Merging Strategy, jika $L_2 = \{\{A,B\}, \{A,C\}, \{A,D\}, \{B,C\}, \{B,D\}\}$, candidate itemset berukuran 3 yang dihasilkan adalah:

Jawaban: b. $\{\{A,B,C\}, \{A,B,D\}\}$

Bagian 18: Perbandingan Kedua Strategi

59. Mengapa Apriori Merging Strategy dianggap lebih efisien daripada $F_{k-1} \times F_1$?

Jawaban: d. Karena menghasilkan lebih sedikit candidate itemset yang tidak perlu diuji

60. Dalam konteks optimasi algoritma, tahap apa yang tetap diperlukan baik pada $F_{k-1} \times F_1$ maupun Apriori Merging Strategy?

Jawaban: d. Pruning untuk menghilangkan candidate itemset yang subsetnya tidak frequent

61. Jika kita memiliki frequent itemset $\{A,B\}$, $\{A,C\}$, dan $\{B,C\}$, berapa banyak candidate itemset berukuran 3 yang dihasilkan oleh:

$F_{k-1} \times F_1$: ...

Apriori: ...

Jawaban: a. (a) 6, (b) 1

62. Pernyataan mana yang benar tentang kedua strategi?

Jawaban: b. Apriori Merging Strategy lebih efisien tetapi lebih kompleks implementasinya

Berikut adalah lanjutan soal multiple choice lengkap beserta jawaban final untuk Bagian Representasi Biner dan Market Basket Data dalam *Association Rule Mining*, disusun terstruktur dan digabung langsung ke daftar sebelumnya (nomor dimulai dari 63):

Bagian 19: Konsep Dasar Representasi Biner

63. Apa tujuan utama penggunaan representasi biner dalam market basket analysis?

Jawaban: d. Mengodekan keberadaan item dalam transaksi dengan 0 dan 1

64. Dalam matriks biner, apa arti nilai 1 pada sel (i,j) ?

Jawaban: c. Item j muncul dalam transaksi i

65. Keuntungan utama representasi biner adalah:

Jawaban: d. Efisiensi komputasi untuk sparse data

Bagian 20: Market Basket Data

66. Market basket data secara alami cocok direpresentasikan sebagai:

Jawaban: d. Matriks biner berukuran (transaksi \times item)

67. Contoh market basket data adalah:

Jawaban: b. Daftar belanjaan di supermarket

68. Dalam konteks retail, analisis market basket data membantu:

Jawaban: d. Menemukan pola belanja pelanggan

Bagian 21: Transformasi Data

69. Bagaimana merepresentasikan transaksi {Roti, Susu, Telur} dalam bentuk biner untuk item [Roti, Susu, Telur, Sabun]?

Jawaban: c. [1, 1, 1, 0]

70. Jika ada 1000 item unik dan rata-rata transaksi mengandung 10 item, berapa sparsity matriks biner-nya?

Jawaban: d. ~99%

Bagian 22: Aplikasi Praktis

71. Mengapa representasi biner tidak cocok untuk data yang menyimpan kuantitas item?

Jawaban: d. Karena hanya mencatat keberadaan bukan jumlah

72. Teknik alternatif jika perlu menyimpan kuantitas item adalah:

Jawaban: d. Menggunakan matriks integer

Berikut adalah lanjutan soal multiple choice lengkap beserta jawaban final untuk Bagian Candidate Generation dan Pruning dalam *Association Rule Mining*, dengan struktur konsisten dan melanjutkan penomoran sebelumnya (mulai dari nomor 73):

Bagian 23: Candidate Generation

73. Apa tujuan dari Candidate Generation dalam algoritma Apriori?

Jawaban: b. Membangkitkan itemset potensial yang akan diuji untuk memenuhi minimum support

74. Bagaimana cara algoritma Apriori membangkitkan candidate itemset berukuran k?

Jawaban: a. Menggabungkan itemset berukuran (k-1) yang memiliki (k-2) item yang sama

75. Misalkan kita memiliki itemset frequent {A,B} dan {A,C}, candidate itemset berukuran 3 yang mungkin dibangkitkan adalah...

Jawaban: a. {A, B, C}

Bagian 24: Pruning

76. Apa tujuan dari Pruning dalam algoritma Apriori?

Jawaban: a. Menghilangkan candidate itemset yang tidak mungkin frequent berdasarkan subsetnya

77. Jika {A, B} tidak frequent, apa yang terjadi pada {A, B, C} setelah Pruning?

Jawaban: a. {A, B, C} dihapus dari candidate itemset karena salah satu subsetnya tidak frequent

78. Pruning memanfaatkan sifat apa untuk mengoptimasi proses pencarian frequent itemset?

Jawaban: a. Sifat antimonotone

Bagian 25: Gabungan Candidate Generation dan Pruning

79. Setelah Candidate Generation, mengapa Pruning diperlukan?

Jawaban: a. Untuk mengurangi jumlah candidate itemset yang harus dihitung support-nya

80. Manakah pernyataan yang benar tentang Candidate Generation dan Pruning?

Jawaban: a. Candidate Generation memperluas itemset, sedangkan Pruning mempersempitnya berdasarkan sifat antimonotone

Berikut adalah lanjutan soal multiple choice lengkap dengan jawaban final untuk Bagian Konsep Dasar, Kandidat Itemset, dan Aplikasi Apriori, melanjutkan penomoran sebelumnya (mulai dari nomor 81) dengan struktur yang konsisten:

Bagian 26: Konsep Dasar

81. Itemset yang memenuhi minimum support disebut:

Jawaban: d. Frequent itemset

82. Itemset {A,B} memiliki support 20% ($\text{min_sup}=25\%$). Status itemset ini adalah:

Jawaban: d. Infrequent

Bagian 27: Kandidat Itemset

83. Itemset yang belum diuji support-nya disebut:

Jawaban: c. Candidate itemset

84. Mengapa {A,B,C} bisa menjadi not candidate itemset?

Jawaban: d. Karena subsetnya {A,B} infrequent

Bagian 28: Aplikasi Algoritma

85. Dalam algoritma Apriori, itemset mana yang TIDAK akan dibangkitkan?

Jawaban: d. Not candidate itemset

86. Jika {A} frequent dan {B} frequent, tapi {A,B} belum diuji, status {A,B} adalah:

Jawaban: b. Candidate itemset

Bagian 29: Analisis Kasus

Gunakan data berikut ($\text{min_sup}=30\%$):

- $\text{Support}\{A\}=40\%$, $\text{Support}\{B\}=50\%$, $\text{Support}\{C\}=20\%$
- $\text{Support}\{A,B\}=25\%$, $\text{Support}\{B,C\}=15\%$

87. Status itemset {C}:

Jawaban: d. Infrequent

88. Status itemset {A,B,C}:

Jawaban: d. Not candidate

89. Itemset mana yang pasti di-prune oleh Apriori?

Jawaban: a. {B,C}

Bagian 30: Prinsip Dasar

90. Pernyataan yang BENAR tentang hubungan ketiga jenis itemset:

Jawaban: b. Not candidate \rightarrow Tidak perlu diuji \rightarrow Bisa jadi infrequent

91. Itemset yang tidak memenuhi min_sup tapi tetap di-generate sebagai candidate:

Jawaban: c. Tidak ada (prinsip antimonotone mencegah ini)

92. Jika $\{A,B\}$ frequent dan $\{A,C\}$ frequent, maka $\{A,B,C\}$ adalah:

Jawaban:

d. Candidate itemset