

k Test 2 Data Mining 2014

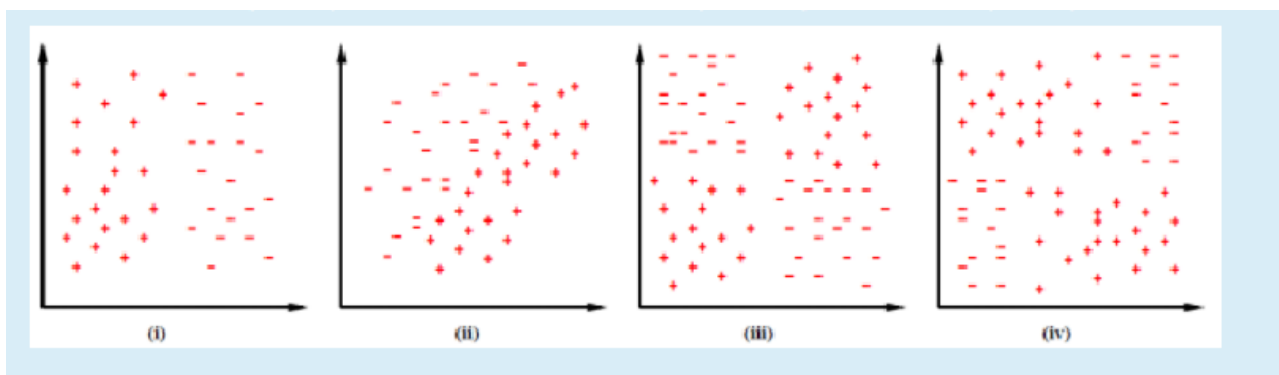
1. Какъв е моделът на следните неща (локален или глобален):
 индукция на дърво на решения - глобален
 naïve bayes - глобален
 kNN - локален
 невронна мрежа - глобален
2. Кой е коренът на дървото при дадена **(примерна)** таблица:
 - a. F1
 - b. F2
 - c. F3 - output би трябвало да е класът
 - d.

| F1 | F2 | F3 | Output |
|----|----|----|--------|
| a | b | c | + |
| c | a | c | + |
| a | a | c | - |
| c | c | a | - |
| a | b | b | - |

- e. Output
3. Кой атрибут ще сложите за корен на дървото?
 - a. F1
 - b. F2
 - c. F3
 - d. Output

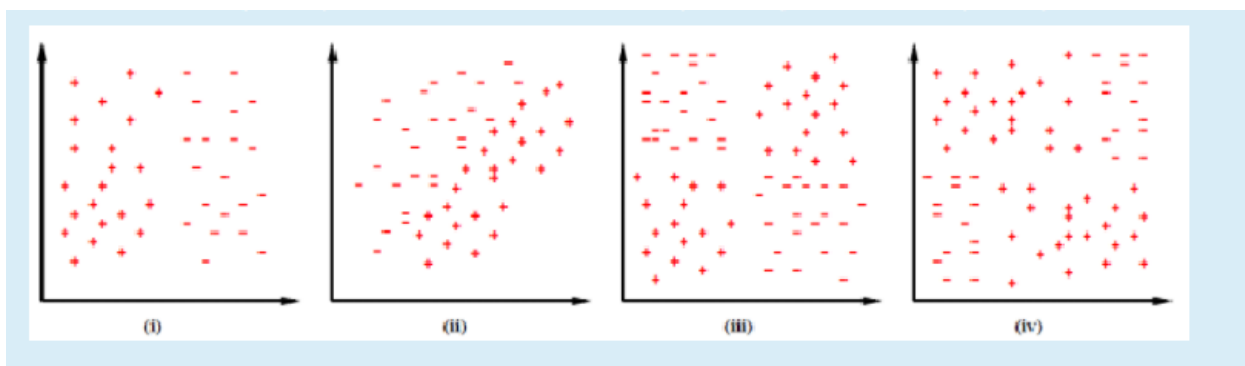
| F1 | F2 | F3 | Output |
|----|----|----|--------|
| a | a | a | + |
| c | b | c | + |
| c | a | c | + |
| a | a | b | - |
| c | b | a | - |
| c | b | b | - |

4. Може ли персептронът да реализира XOR?
 - a. да
 - b. не
5. Кое твърдение се отнася за ансамбловото учене (Ensemble learning)?
 - a) използват се различни алгоритми и се прави гласуване
 - б) използват се различни извадки от тестови данни и се прави гласуване
 - в) използват се к члена и се гласува
6. Кой от следните алгоритми спадат към йерархичните модели?
 - a. kNN
 - b. Agglomerative learning
7. Кой от следните алгоритми спадат към ученето с учител (supervised learning)?
 - a. ID3
 - b. DBSCAN
 - c. kNN
 - d. kMeans
 - e. C4.5
8. Кой от следните твърдения се отнасят за асоциативните правила?
 - a. минимален support
 - b. минимален confidence
9. Кой от следните не са верни за асоциативното учене?
 - a) резултатът от косинуса на ъгъла между данните
10. Кой от следните алгоритми за обучение се явява мързелив
 - a. Индукция на дърво на решенията
 - b. учене основано на примери
 - c. линейна регресия
 - d. k-means
11. Посочете правилното твърдение за обратното разпространение на грешката (Error Backpropagation)
 - a. използва се при обучението на еднослойна невронна мрежа (perception)
 - b. използва се при обучението на еднослойна и многослойна невронна мрежа
 - c. не се използва при обучението на невронни мрежи
 - d. използва се при обучението на многослойна невронна мрежа
12. Кое от следните обучаващи се множества не може да бъде разделено на дърво на решенията на 2 нива



- a. (i)
- b. (ii)
- c. (iii)
- d. (iv)

13. Кое от следните обучаващи се множества може да бъде правилно разделено от персептрона?



Отбележете едно или повече

- a. (i)
- b. (ii)
- c. (iii)
- d. (iv)

14. При ученето на дърво на решенията, по кой принцип избираме една от множеството допустими хипотези

- a. на Фишер

- b. на Окам
- c. на Тюринг
- d. на Бейс

15. При ученето на дърво на решенията, по кой принцип избираме една от множеството съвместими хипотези

- a. на Марков
- b. на Окам
- c. на Тюринг
- d. на Бейс

16. Каква е основната цел на кастренето (pruning) на дърво на решенията

- a. Да направи дървото по-компактно
- b. Да повишим точността върху обучаващото множество
- c. Да намалим пространството от хипотези, в които търсим
- d. Да предотврати прекомерно нагаждане към обучаващите примери (over-fitting)

17. Изводът чрез изброяване (inference by enumeration) се основава на:

- a. условните автомарности между автомарните събития
- b. пълното съвместно разпределение на автомарните събития
- c. доказателства, получени от автомарните събития
- d. единствено от автомарните събития

18. Ансамблово учене (Ensemble learning) е:

- a. от множество обучаващи се примера се избира представителна извадка (k на брой примера), така наречения парламент, който после се използва за обучение
- b. Учение на няколко класификатора, като се използва един и същи учещ алгоритъм, но с различни тегла на примерите в обучаващото множество. После тези класификатори вземат решение с гласуване.
- c. Учение на няколко класификатора, използвайки различни алгоритми за учене, които вземат решение с гласуване

- d. Намираме k най-близките съседни примери от обучаващото множество, които вземат решение с гласуване

19. Какъв модел учи всеки един от следните алгоритми (глобален (g) или локален (l))

- i. Учене основано на примери (kNN)
 - ii. Наивен Бейсов класификатор
- b. 1 l, 2 l
 - c. 1 l, 2 g
 - d. 1 g, 2 l
 - e. 1 g, 2 g

20. Какъв модел учи всеки един от следните алгоритми (глобален (g) или локален (l))

- i. Невронна мрежа
 - ii. Индукция на дърво на решенията
- b. 1 l, 2 l
 - c. 1 l, 2 g
 - d. 1 g, 2 l
 - e. 1 g, 2 g

21. Кой от изброените клъстерни алгоритми изгражда йерархичен модел отдолу нагоре

- a. агломеративен
- b. нито един от изброените
- c. йерархичен k-means
- d. k-means ++

22. Правилото (от началното състояние към целта (Progressive)) и обратно (от целта към началното състояние (Regressive)) търсене се използват само за строене изцяло (totally) наредени планове

- a. истина
- b. лъжа

23. Причинно-следствена връзка (casual link) при частично наредените планове се означава

$$A - p \rightarrow B$$

и означава:

- a. Между A и B трябва да има друго действие C, което да дава резултат p
- b. A трябва да се изпълни преди B
- c. Изпълнението на B трябва да започне непосредствено след изпълнението на A
- d. Не може да има действие C между A и B, което да отменя p

24. Алгоритъмът k-means е:

- a. лаком алгоритъм
- b. оптимален алгоритъм
- c. алгоритъм използващ търсене в дълбочина с възврат
- d. глобално търсещ алгоритъм

25. Алгоритъмът k-means е:

- a. локално търсещ алгоритъм
- b. оптимален алгоритъм
- c. алгоритъм използващ търсене в дълбочина с възврат
- d. глобално търсещ алгоритъм

26. Кое от следните множества от примери съдържа шум

- a. (i)
- b. (ii)
- c. (iii)

(i)

| X_1 | X_2 | Y |
|-------|-------|-----|
| 1 | 1 | + |
| 4 | 2 | - |
| 4 | 5 | - |
| 5 | 5 | + |

(ii)

| X_1 | X_2 | Y |
|-------|-------|-----|
| 1 | 1 | + |
| 5 | 5 | - |
| 4 | 5 | - |
| 5 | 5 | + |

(iii)

| X_1 | X_2 | Y |
|-------|-------|-----|
| 1 | 1 | + |
| 4 | 2 | - |
| 4 | 5 | + |
| 5 | 5 | + |

27. Какво е частично нареден план?

<http://fmi.wikidot.com/ai05> ctrl+f частично наредени планове

28. Кое от следните множества от примери е линейно разделяемо

- a. (i)
- b. (ii)
- c. (iii)

| | | | |
|-----|-------|-------|-----|
| (i) | X_1 | X_2 | Y |
| | 1 | 1 | + |
| | 4 | 2 | - |
| | 4 | 5 | - |
| | 5 | 5 | + |

| | | | |
|------|-------|-------|-----|
| (ii) | X_1 | X_2 | Y |
| | 1 | 1 | + |
| | 5 | 5 | - |
| | 4 | 5 | - |
| | 5 | 5 | + |

| | | | |
|-------|-------|-------|-----|
| (iii) | X_1 | X_2 | Y |
| | 1 | 1 | + |
| | 4 | 2 | - |
| | 4 | 5 | + |
| | 5 | 5 | + |

29. Кое от изброените клъстерни алгоритми изгражда йерархичен модел отгоре надолу

- a. k-means++
- b. йерархичен k-means
- c. агломеративен
- d. нито един от изброените

30. PCA (Principal component analysis)

- a. Решава класификационна задача
- b. Намалява размерността на пространството от примери
- c. Решава регресионна задача
- d. Свежда примерите до представителна извадка с по-малък размер

31. Ученето основано на примери (instance based learning) (kNN) може да бъде определено като

- i. мързеливо учене
- ii. учене по аналогия
- b. И двете твърдения са верни
- c. Твърдение 1 е вярно, а 2 е невярно
- d. Твърдение 2 е вярно, а 1 невярно

е. И двете твърдения не са верни

32. За k-means са дадени 2 твърдения:

i. трябва да се направят няколко произволни рестартирания

ii. строи йерархичен класификатор

b. И двете твърдения са верни

c. Твърдение 1 е вярно, а 2 е невярно

d. Твърдение 2 е вярно, а 1 невярно

е. И двете твърдения не са верни

33. Дадена е следната контекстно-свободна граматика, където X^* означава 0 или повече срещания - Това не е за HAC - това е от NLP

$S \rightarrow NP VP$
 $S \rightarrow first\ S\ then\ S$
 $NP \rightarrow Determiner\ Modifier\ Noun \mid Pronoun \mid ProperNoun$
 $Determiner \rightarrow a \mid the \mid every$
 $Pronoun \rightarrow she \mid he \mid it \mid him \mid her$
 $Modifier \rightarrow Adjective^* \mid Noun^*$
 $Adjective \rightarrow red \mid violet \mid fragrant$
 $Noun \rightarrow rose \mid dahlia \mid violet$

$VP \rightarrow Verb\ NP$
 $VP \rightarrow IntransitiveVerb$
 $VP \rightarrow Copula\ Adjective$
 $Verb \rightarrow smelled \mid watered \mid was$
 $IntransitiveVerb \rightarrow smelled \mid rose$
 $Copula \rightarrow was \mid seemed \mid smelled$
 $ProperNoun \rightarrow Spike$

Изречението the red red rose rose rose може да бъде генерирано от дадената граматика а.
Посочете дали твърдението е грешно или не

a. Истина

b. Лъжа

34. Дадена е следната контекстно-свободна граматика, където X^* означава 0 или повече срещания Това не е за HAC - това е от NLP

$S \rightarrow NP VP$
 $S \rightarrow first\ S\ then\ S$
 $NP \rightarrow Determiner\ Modifier\ Noun \mid Pronoun \mid ProperNoun$
 $Determiner \rightarrow a \mid the \mid every$
 $Pronoun \rightarrow she \mid he \mid it \mid him \mid her$
 $Modifier \rightarrow Adjective^* \mid Noun^*$
 $Adjective \rightarrow red \mid violet \mid fragrant$
 $Noun \rightarrow rose \mid dahlia \mid violet$

$VP \rightarrow Verb\ NP$
 $VP \rightarrow IntransitiveVerb$
 $VP \rightarrow Copula\ Adjective$
 $Verb \rightarrow smelled \mid watered \mid was$
 $IntransitiveVerb \rightarrow smelled \mid rose$
 $Copula \rightarrow was \mid seemed \mid smelled$
 $ProperNoun \rightarrow Spike$

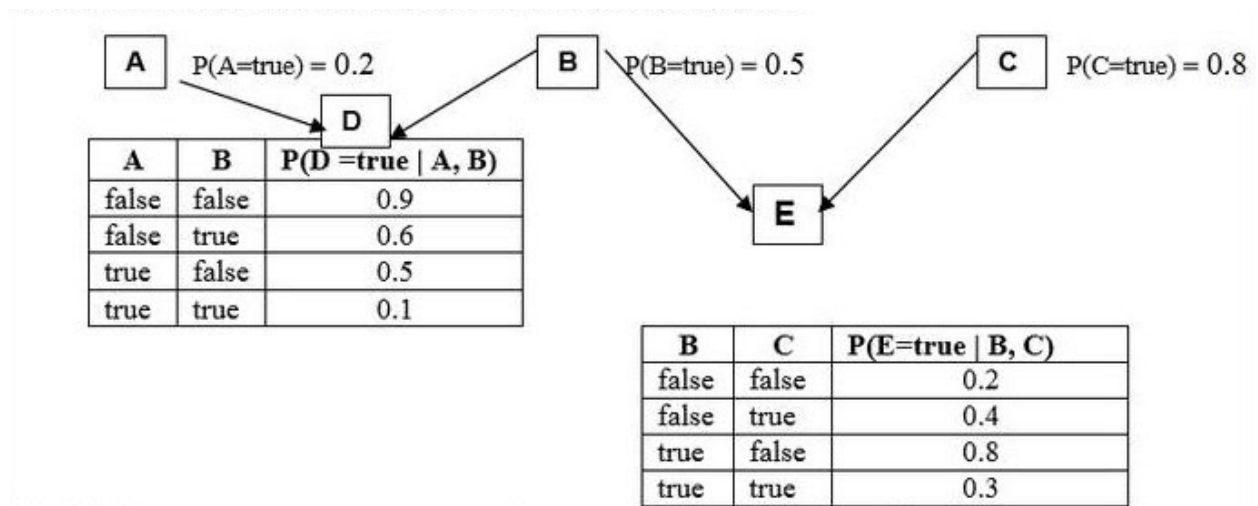
Изречението she was a violet violet violet може да бъде генерирано от дадената граматика.
Посочете дали твърдението е грешно или не

- a. Истина
- b. Лъжа

35. Посочете кои алгоритми спадат към учене без учител (Unsupervised learning)

- a. kMeans
- b. C4.5
- c. kNN
- d. EM-алгоритъм
- e. DBSCAN
- f. ID3

36. Дадена е Бейсова мрежа, където A–E са булеви променливи



Каква е вероятността всичките булеви променливи да са true

- a. 0.0024
- b. 0.00144
- c. 0.03
- d. 0.08

37. Ако Температура е условно независима от Главоболие при дадена диагноза Грип, то вярно или невярно е всяко едно от твърденията

- i. $P(\text{Главоболие}, \text{Температура} \mid \text{Грип}) = P(\text{Главоболие} \mid \text{Грип})P(\text{Температура} \mid \text{Грип})$
- ii. $P(\text{Температура} \mid \text{Главоболие}, \text{Грип}) = P(\text{Температура} \mid \text{Грип})$

- b. вярно, вярно
- c. вярно, невярно
- d. невярно, вярно
- e. невярно, невярно

38. Ако Температура е условно независима от Главоболие при дадена диагноза Грип, то вярно или невярно е всяко едно от твърденията

- i. $P(\text{Главоболие} \mid \text{Температура}, \text{Грип}) = P(\text{Главоболие} \mid \text{Грип})$

- ii. $P(\text{Грип} \mid \text{Главоболие, Температура}) = P(\text{Грип})P(\text{Главоболие} \mid \text{Грип})P(\text{Температура} \mid \text{Грип}) / P(\text{Температура})P(\text{Грип})$
- b. вярно, вярно
- c. **вярно, невярно**
- d. невярно, вярно
- e. невярно, невярно
39. Ако Температура е условно независима от Главоболие при дадена диагноза Грип, то вярно или невярно е всяко едно от твърденията
- i. $P(\text{Температура} \mid \text{Главоболие, Грип}) = P(\text{Температура} \mid \text{Грип})P(\text{Температура} \mid \text{Главоболие})$
- ii. $P(\text{Температура, Главоболие} \mid \text{Грип}) = P(T \mid \text{Грип})P(\text{Температура} \mid \text{Главоболие})$
- b. вярно, вярно
- c. вярно, невярно
- d. невярно, вярно
- e. **невярно, невярно**
40. Индукция на дървото ID3 учи локален модел
- a. вярно
- b. **невярно**
41. Невронната мрежа учи локален модел
- a. истина
- b. **лъжа**
42. При обучение на логистична регресия се използва метод на градиентно изкачване
- a. **истина**
- b. лъжа
43. Може ли “изключващо или” хог да се реализира с невронна мрежа с едно скрито ниво?
- a. **да**
- b. не
44. Кой от изброените клъстерни алгоритми изгражда нейерархичен модел?
- a. Агломеративен
- b. bisecting k-means
- c. нито един от изброените
- d. **k-means**
45. Правото (от началното състояние към целта (Progressive)) и Обратното (от целта към началното състояние (Regressive)) търсене се използват само за строене на изцяло (totally) наредени планове.
- a. **Истина**
- b. Лъжа

46. Логистичната регресия предполага условна независимост на атрибутите/характеристиките

- a. Истина
- b. Лъжа

47. Дадени са съвместните вероятности

Positive

| | circle | square |
|------|--------|--------|
| red | 0.2 | 0.02 |
| blue | 0.02 | 0.01 |

Negative

| | circle | square |
|------|--------|--------|
| red | 0.05 | 0.3 |
| blue | 0.2 | 0.2 |

Пресметнете

$P(\text{positive} \mid \text{red} \wedge \text{circle})$

- a. 0.25
- b. 0.80
- c. 0.57
- d. 0.20

48. За k-means алгоритъма са дадени следните 2 твърдения:

- i. Трябва да се направят няколко произволни рестартирания

ii. Строи йерархичен класификатор

b. Твърдение 1 е вярно, твърдение 2 невярно

c. Двете твърдения са неверни

d. Твърдение 1 е невярно, твърдение 2 вярно

e. Твърдения 1 и 2 са верни