قائمة الاسئلة

امتحان نهاية الفصل الدراسي الثاني - للعام الجامعي 1446 هـ - الموافق -2025/2024مـ-مركز الحاسوب :: أدوات تنقيب البيانات - (125)-

أ.علي عبدالله الشعيبي

1) ما هي المشكلة الشائعة في البيانات التي تحتاج إلى تنظيف؟
   1) + البيانات المفقودة
   2) - البيانات الصحيحة
   3) - البيانات المنظمة
   4) - البيانات الضخمة

2) أي مما يلي يُستخدم للتعامل مع القيم المفقودة في البيانات؟
   1) - تعويض القيم المفقودة
   2) - حذف الصفوف التي تحتوي على قيم مفقودة
   3) - استخدام التنبؤ بالقيم المفقودة
   4) + كل ما سبق

3) ما هي الخطوة الأولى في استخدام برنامج ويكا (Weka)؟
   1) + تحميل البيانات
   2) - تنظيف البيانات
   3) - تحليل البيانات
   4) - تصدير النتائج

4) أي نظام يستخدم تعلّم الآلة سيحتاجُ لثلاثة مكونات رئيسية وهي:
   1) - المتغيرات ، المصفوفات ، الدوال
   2) + البيانات ، الميزات ، الخوارزميات
   3) - التعلم الموجه ، التعلم غير الموجه ، التعلم المشترك
   4) - التنقيب ، الروبوتات ، البرامج

5) ما هي الخوارزمية المستخدمة في برنامج ويكا (Weka) لبناء أشجار القرار؟
   1) - Naive Bayesian
   2) + J48
   3) - Support Vector Machine
   4) - K-Nearest Neighbors

6) أي من تنسيقات الملفات التالية تعتبر التنسيق المفضل لبرنامج WEKA ؟
   1) - CSV
   2) - JSON
   3) - XML
   4) + ARFF

7) ما هي خوارزمية التصنيف المستخدمة لبناء أشجار القرار؟
   1) - K-Nearest Neighbors (KNN)
   2) - Naive Bayes
   3) + Decision Tree
   4) - Support Vector Machine (SVM)

8) لنفترض أنك بحاجة لاقتراض بعض المال عن طريق بطاقتك الائتمانية. كيف سيعرف البنك إذا كنت تستطيع فعلا أن تسدد هذا القرض أم لا
   1) + باستخدام خوارزميات التصنيف
   2) - باستخدام خوارزميات الانحدار
   3) - باستخدام البرمجة الكائنية
   4) - باستخدام خوارزميات التعليم المعزز

9) ما هو تطبيق تنقيب البيانات في مجال التعليم؟
   1) - تحليل سلوك العملاء
   2) - تحليل البيانات المالية
   3) - تشخيص الأمراض
   4) + تحليل أداء الطلاب

10) ما هي التقنية المستخدمة لتقليل الأبعاد في البيانات؟

1) - K-Means
2) - Decision Tree
3) - Association rule learning
4) + PCA (Principal Component Analysis)

11) تُقسِّم عملية ..... الكائنات على أساس ميِّزات غير معروفة
1) - الانحدار
2) - تقليل الأبعاد أو التعميم
3) - تعلم قواعد الارتباط
4) + التجميع

12) ما أهمية تنقيب البيانات في التجارة الإلكترونية؟
1) - تقليل كميات البيانات
2) + تحسين التوصيات للعملاء
3) - حذف المنتجات غير المرغوبة
4) - زيادة وقت تحميل الموقع

13) تعتمد على أخذ مجموعة من الخوارزميات ذات الفعالية العادية، وتجبرها على تصحيح أخطاء بعضها بعضًا
1) - طريقة الشبكات العصبية
2) + طريقة المجموعات Method Ensemble
3) - الواجهات Interface
4) - ليس مما سبق

14) ما هو تنقيب البيانات (Mining Data)؟
1) - عملية استخراج البيانات من قواعد البيانات
2) - عملية تخزين البيانات في قواعد بيانات
3) + عملية اكتشاف أنماط مفيدة من كميات كبيرة من البيانات
4) - عملية تحليل البيانات يدويًا

15) المصطلح المرادف لتنقيب البيانات هو
1) - Object Oriented Programming (OOP)
2) + Knowledge Discovery in Data (KDD)
3) - Supervised Learning (SL)
4) - Unsupervised Learning (UL)

16) ما هي الخطوة الأولى في عملية تنقيب البيانات؟
1) - تنظيف البيانات
2) + فهم المشكلة وتحديد الأهداف
3) - تطبيق خوارزميات التصنيف
4) - تصدير النتائج

17) يستخرج تنقيب البيانات معلومات مخفية من البيانات، لكنه غير قادر على تقييم قيمة هذه المعلومات
1) + صحيح
2) - خاطئ

18) تستهلك هذه المرحلة بشكل عام حوالي 80% الى 90% من وقت مشروع تنقيب البيانات
1) - مرحلة النشر
2) - مرحلة فهم البيانات
3) + مرحلة إعداد البيانات
4) - مرحلة النمذجة

19) في هذه المرحلة، يتم تطبيق خوارزميات تنقيب البيانات لبناء النموذج
1) - Problem definition
2) - Data understanding phase
3) - Data preparation phase
4) + Modeling phase

20) واحدة من تقنيات تنقيب البيانات هي:
1) - الطريقة المثلثية

2) - البرمجة الموجهة
3) - قانون الجذب
4) + النمذجة التنبؤية

21) من لغات البرمجة الشائعة في تنقيب البيانات:
1) - HTML
2) - C++
3) + Python
4) - Cobol

22) ما هي أداة تنقيب البيانات الشهيرة التي توفر واجهة رسومية؟
1) + Weka
2) - MySQL
3) - Photoshop
4) - Powerpoint

23) أي من التالي ليس من مراحل تنقيب البيانات؟
1) - جمع البيانات
2) + اختبار البرامج
3) - إتنظيف البيانات
4) - تحليل الأنماط

24) ما الفرق بين تنقيب البيانات وتحليل البيانات التقليدي؟
1) - التحليل التقليدي أكثر دقة
2) - التحليل التقليدي أسرع من تنقيب البيانات
3) + تنقيب البيانات يعتمد على التعلم الآلي واكتشاف الأنماط
4) - لا يوجد فرق

25) تشمل تطبيقات كشف الانحراف
1) - كشف المخالفات المرورية
2) - كشف مخالفات توزيع الغاز
3) - كشف التعامل بالعملات الاجنبية
4) + كشف الاحتيال في استخدام بطاقات الائتمان

26) ما نوع المشاكل التي يحلها تنقيب البيانات؟
1) - تصنيف البيانات
2) - التنبؤ بالقيم المستقبلية
3) - اكتشاف الأنماط المخفية
4) + جميع ما سبق

27) هي تقنية تنقيب البيانات تتضمن تحويل البيانات الخام إلى تنسيق مفهوم
1) - تقنية تحليل الارتباط
2) + المعالجة الأولية للبيانات
3) - النمذجة التنبؤة
4) - تقنية الجار الاقرب

28) من مراحل المعالجة الاولية للبيانات
1) + تكامل البيانات
2) - ضخامة البيانات
3) - جودة البيانات

29) يتم التعامل مع القيم المفقودة في مرحلة
1) - تقليل البيانات
2) + تنظيف البيانات
3) - تحويل البيانات
4) - دمج البيانات

30) أكثر طرق تحويل البيانات شيوعًا واستخدامًا هما:
1) - التطبيع (التسوية) Normalization

2) - التوحيد القياسي Standardization
3) - كلاهما غير صحيح
4) + كلاهما صحيح

31) باستخدام تقنيات ..... يمكن تمثيل مجموعة البيانات بطريقة مختصرة دون المساس فعليًا بسلامة البيانات الأصلية.
1) + تقليل البيانات
2) - فرز البيانات
3) - تجميع البيانات

32) تكون الفئات المستهدفة معروفة بالفعل، أي أن بيانات التدريب مصنفة بالفعل بإجابات فعلية.
1) - التصنيف المسبق
2) + التصنيف اللاحق
3) - التصنيف العددي
4) - التصنيف الاسمي

33) سمة الفئة التي تمثل مخرج لجميع السمات الأخرى تُعرف ب
1) - السمات المستقلة attributes independent
2) - السمة المحايدة attribute neutral
3) + السمة التابعة attribute dependent

34) تسمى السمات التي تحتوي على أرقام
1) - categorical
2) - nominal
3) + numerical
4) - ليس مما سبق

35) هو عملية من خطوتين. الخطوة الأولى هي تدريب النموذج والخطوة الثانية هي اختبار النموذج للتأكد من دقته
1) + التصنيف
2) - برمجة الكائنات
3) - مجموعة البيانات
4) - البحث

36) من خلال ..... يتعلم النظام وينشئ بعض القواعد للتنبؤ.
1) - تحليل بيانات الاختبار
2) + تحليل بيانات التدريب
3) - تحليل بيانات غير معروفة
4) - التعرف على نوع الملفات

37) في مصنف شجرة القرار ، يتم إجراء التنبؤات باستخدام
1) - حلقات تكرارية
2) - مصفوفات
3) + جمل شرطية متعددة
4) - دوال

38) تعتمد خوارزمية شجرة القرار على مفهوم
1) + كسب المعلومات
2) - التنقل بين الاشجار
3) - فهم كل شجرة
4) - كل ما سبق

39) إذا كان هناك ..... فهناك معلومات وإذا لم يكن هناك ..... فلا توجد معلومات
1) - تدفق
2) - انحياز
3) + عدم يقين
4) - ليس فيما سبق

40) هي متوسط كمية المعلومات المقدمة من مصدر البيانات
1) - الاحتمالية probability
2) - سمة التقسيم attribute split

3) - كسب المعلومات gain Information
4) + الإنتروبي entropy