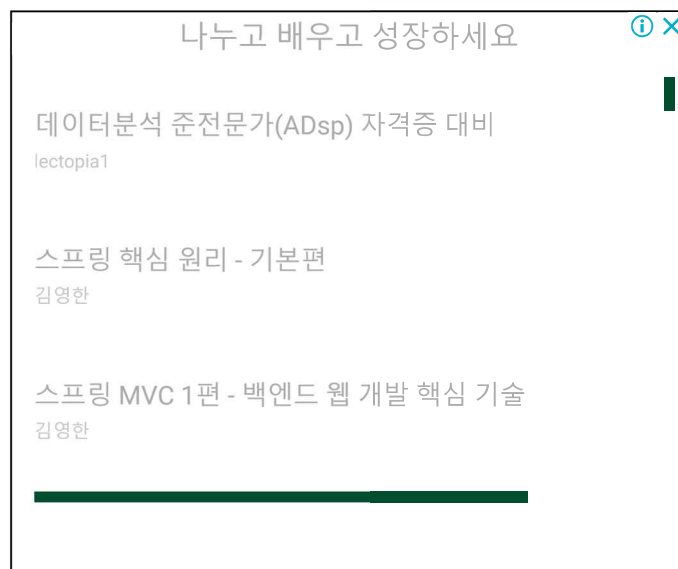


R/데이터분석자격증(ADP)

데이터분석전문가(ADP) 실기시험 기출문제 조사 - 2. 데이터 마이닝 & 기계 학습

👤 분석하는다니엘 2020. 11. 12.



이번 포스트는 데이터 분석 전문가(ADP) 실기시험을 준비하는 분들에게 도움이 되고자 확인된 내용을 공유하는 글입니다.

아래 기출 문제들의 분류는 제가 임의로 정리하여 오분류가 있을 수 있습니다. ^^

- 최종 업데이트 일자 : 2022-02-21

데이터 마이닝 & 기계 학습 (Data Mining & Machine Learning)

[1회, 2014-06-24]

- 고객 세분화 : 세분화 변수 생성 및 선정, 군집분석, 집단별 특성 분석 및 인사이트 도출

[2회, 2014-10-18]





- 로또 번호 연관분석

←

Google 광고

이 광고 그만 보기

이 광고가 표시된 이유 ⓘ

[6회, 2016-04-30]

- sales데이터 분석 : 고객/제품별 처리, 피어슨상관분석, 특정 유저와 유사한 15명 도출, 특정 유저당 5개 상품 추천

[7회, 2016-10-29]

- 홈쇼핑 구매데이터 분석 : 구매내역 데이터 제공, EDA, F/E, Classification

[8회, 2017-04-29]

- 백화점 사용패턴 분석 (하단 데이터 예시)

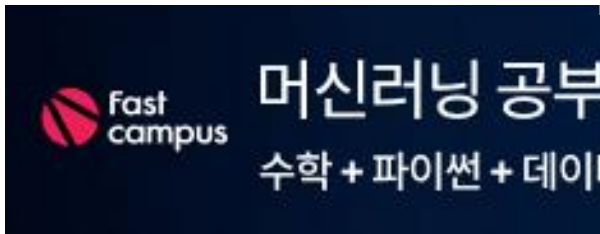
- Date(0000-00-00 00:00 형태)
- customer id
- product (한글 백화점 상품명)
- price (가격)

1. 파생 변수 생성 및 EDA
2. 군집분석 및 적절한 집단 갯수 설정
3. 세그멘테이션 별 의미 파악 (집단 라벨링)
4. 세그멘테이션 별 마케팅 인사이트 도출

[11회, 2018-10-27]

^

^



- 고객 구매데이터(transaction data) 분석
- 1. 이후 1개월 간 실제 고객이 구매할 것으로 예정되는 지점 추천
- 2. 5개 지점 추천 후 하나라도 맞으면 맞는 것으로 간주
- 3. 적중률 66.7% 이상인 경우만 채점, 그 이하인 경우 0점으로 간주

[13회, 2019-06-29]

- 마케팅 데이터 분석
- 1. 시각화 및 전처리를 통해 세분화 작업에 필요한 변수 추출 및 파생변수 생성
- 2. 군집분석을 통한 적절한 군집 수 파악 및 군집 정의
- 3. 마케팅 인사이트 도출
- 4. 문항별로 "시각화 필수" 조건이 있었음

[14회, 2019-09-07]

- California 주택 가격 예측 : EDA, NA처리, F/E, 집 가격 예측/검증

[15회, 2019-12-14]

- 제조 생산 데이터 분석
- 1. 데이터 탐색 : EDA
- 2. 데이터 전처리 : 변수 선택(VIF), 파생변수 생성, 데이터 분할(train/validation/test(20%))
- 3. 로지스틱 분석 : 분류1 을 판단 하는 모델 생성 (종속변수는 총 7개 분류, 분류1 외의 값은 0으로 치환), confusion matrix 해석
- 4. 로지스틱 분석 외 3개 이상 분류 모델 생성 및 결과 해석 : SVM 필수 포함, Precision/Sensitivity 결과 출력
- 5. 위 모델 중 최고 모델을 선택하여 최적의 군집 개수를 선택하고 클러스터링 수행 : F-1 Score 출력



11월 20일 11월 21일

패스트캠퍼스

자세

[17회, 2020-06-21]

- 집값 예측 문제
- 1. EDA
- 2. 모델 생성
 1. 데이터 분할
 2. 교호작용을 고려한 다중 선형 회귀 수행
- 3. 3가지 분류 모델 생성 및 비교, 좋은 모델 선택

[18회, 2020-09-19]

- 고객 등급 예측 모형
- 1. EDA & 결측값 채우기
- 2. 파생변수 3개 생성 & 이유 작성
- 3. Train-Test 분할(7:3) / SOM 군집분석 / 정오분류표
- 4. 분류분석 4가지

[19회, 2020-12-13]

- 고객 이탈 예측
 - 20개의 x와 1개의 y를 가지는 데이터로 y는 0과 1을 가짐
- 1. EDA 및 전처리
- 2. 분류모델 3개 적용 및 Confusion Matrix 출력
- 3. 3개 모델 앙상블 모형 생성 후 예측값 저장

[21회, 2021-06-05]





(school, sex, paid, activities, famrel, freetime, goout, dalc, walc, health, absences, 타겟변수 : grade)

- 참고링크 : <https://archive.ics.uci.edu/ml/datasets/student+performance>

1. 시각화 포함 탐색적 자료분석
2. 결측치 식별하고 결측치를 예측하는 두 가지 방법 정도를 쓰고, 선택한 이유를 설명
3. 범주형 변수 인코딩이 필요한 경우를 식별하고, 변환을 적용하시오. 선택한 이유를 설명.
4. 데이터 분할 방법을 2가지 쓰고 적절한 데이터 분할을 적용. 선택한 이유 설명.
5. svm, xgboost, randomforest 3개의 알고리즘 공통점을 쓰고 이 예측 분석에 적합한 알고리즘인지 설명.
6. 3모델 모두 모델링 해보고 가장 적합한 알고리즘 선택하고 이유 설명. 한계점 설명하고 보완 가능한 부분 설명. 현업에서 사용시 주의할 점 등에 대해 기술.

[22회, 2021-09-12]

- 머신러닝 (data: Pima Indian Diabetes)
1. 탐색적 데이터 분석: 결측치 확인, 히스토그램/박스플롯/페어플롯, 타겟변수 분포 그래프의 불균형 확인, 변수 전체의 상관관계, 이상치 처리 방안 제시, 위의 전처리 단계에서 얻은 함수 분석 시 고려사항 작성
 2. 클래스 불균형 처리: 오버샘플링, 언더샘플링 과정 설명하고 결과 작성, 둘 중 선택하고 그 이유 설명
 3. 모델링: 최소 3개 이상 알고리즘 제시하고 정확도 측면의 모델 1개와 속도 측면의 모델 1개를 구현, 모델 비교하고 결과 설명, 속도 개선을 위한 차원 축소 설명하고 수행, 성능과 속도 비교하여 결과 작성

[23회, 2021-12-18]

- 독립변수는 4개 (온도, 습도, 빛, CO2), 종속변수는 객실사용 여부(0, 1), 관측치 290개의 분류문제

1. 문항1

1. EDA 과정을 하고 결과를 제시하여라.
2. 결측치를 찾고 대안법을 제시하여라.
3. 데이터 질을 향상 시키기 위한 방법을 제시하여라.(구현 필요X)

2. 문항2

1. 데이터 불균형 식별하고 불균형을 처리할 오버샘플링 기법 선택하여 2개 기법 선정



2. 같은 데이터와 비교하여 오버샘플링이 아닌 경우에 대해 설명

평생소장 | 딥러닝

알고리즘 구현, 실습 프로젝트
정복하는 딥러닝

패스트캠퍼스

자세

참고 자료

분류 (Classification)

- 도요타 중고차 가격 예측, XGBoost
- R을 사용한 베이지 분류/예측 모델 (Naive Bayes classification in R)
- R을 사용한 의사결정나무 분석 (Decision tree analysis in R)
- 도요타 중고차 가격 예측, XGBoost

군집 분석(Clustering)

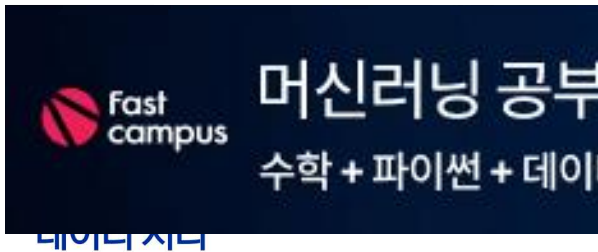
- R을 사용한 K-means 군집분석 (K-means clustering in R)
- 마케팅 - RFM 및 응용분야

연관 분석(Association Analysis)

- R을 사용한 연관성 분석 (association rules in R)
- 장바구니 분석을 통한 쇼핑 트렌드 예측-1
- 장바구니 분석을 통한 쇼핑 트렌드 예측-2
- 장바구니 분석을 통한 쇼핑 트렌드 예측-3

추천(Recommendation)





- 클래스 불균형을 다루는 방법 (Class Imbalance)
- 클래스 불균형(Class imbalance)
- 성능지표 계산을 위한 Confusion Matrix
- R 데이터마이닝 dplyr

ADP 시험 응시자 분들 모두 화이팅 입니다!!! ^^

라군 인 테라스 2차

세계가 주목하는 해양레저도시, 시화 중심에서 펼쳐지는 라군라이프!
<http://hillstatelagoon2.com>

4

구독하기

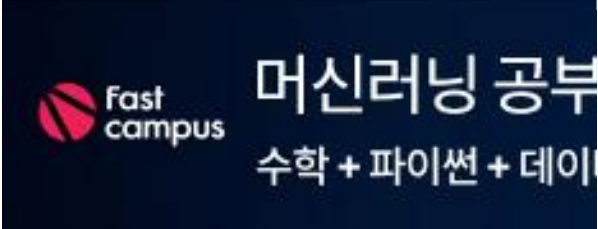
관련 글

- ▶ 데이터분석전문가(ADP) 19회 실기시험 준비 (updated:2020.12.09)
- ▶ 데이터분석전문가(ADP) 실기시험 기출문제 조사 - 3. 텍스트 마이닝
- ▶ 데이터분석전문가(ADP) 실기시험 기출문제 조사 - 2. 데이터 마이닝 & 기계 학습
- ▶ 데이터분석전문가(ADP) 실기시험 기출문제 조사 - 1. 통계 분석
- ▶ 데이터분석전문가(ADP) 19회 실기시험 준비
- ▶ 데이터분석전문가(ADP) 17회 실기시험 후기

태그

#ADP, #ADsP, #데이터분석, #데이터분석자격증, #데이터분석전문가, #빅데이터, #시험후기, #자격증

추천 글



댓글 0

이름	비밀번호
여러분의 소중한 댓글을 입력해주세요.	

☐ 비밀글

등록