산불 진화준비 단계 분석

1. 04\_허건호.csv, 05\_김시형.ipynb, 6\_박지원.ipynb 사용 (& 데이터 추가 재구조화 및 제거 진행)

2. 분석에 사용한 feature

1차 전체 feature

'화재발생일시', '화재발생\_연', '화재발생\_월', '화재발생\_일', '주말여부', '화재발생시간대', '시군구명', '읍면동명', '발화지점', '날씨', '온도', '습도', '시간단위풍속', '특보내용', '발화원인종합', '소방서명', '서센터명', '소방지역대명', '출동소요시간', '현장소방서거리', '현장안전센터거리', '현장소방지역대거리', '전체인력수합계', '화재진압시간', '동원장비수', '인명피해수', '재산피해금액', '피해면적'

위의 feature 중 분석에 필요한 feature 선별

선별 feature 확인

'화재발생\_월', '화재발생\_일', '주말여부', '화재발생시간대', '시군구명', '발화지점', '날씨', '온도', '습도', '시간단위풍속', '특보내용', '발화원인종합', '출동소요시간', '현장소방서거리', '현장안전센터거리', '전체인력수합계', '화재진압시간', '동원장비수', '피해면적'

2차 feature 재구조화

a. '화재발생\_월' 을 계절로 변환

(한국민족문화대백과사전 참고 “일반적으로 3·4·5월을 봄, 6·7·8월을 여름, 9·10·11월을 가을, 12·1·2월을 겨울이라고 한다.” )

b. '화재발생\_일’ 을 월초, 월말, 월중순으로 변환

10일 단위로 끊어서 구분

분석에 사용된 feature

'한달내의 시기', '계절', '주말여부', '화재발생시간대', '시군구명', '발화지점', '날씨', '온도', '습도', '시간단위풍속', '특보내용', '발화원인종합', '출동소요시간', '현장소방서거리', '현장안전센터거리', '전체인력수합계', '화재진압시간', '동원장비수', '피해면적'

3. '한달내의 시기', '계절', ‘화재발생시간대’, '시군구명' , '발화지점', '날씨', '발화원인종합', '특보내용' 에 대한 원 핫 인코딩 진행

4. DT 코드 동일

5. 총 3번의 분석 진행

1. 모든 피처를 전부 사용한 결과 유의미한 결과가 나오지 않음.   
     
   R² Score: -21.404874523513673  
   RMSE on training set: 4.429457305695217e-18  
   RMSE on test set: 101.62059362533478

<GridSearchCV>  
R² Score: -21.350820024800406

RMSE on training set: 3.145731055703614

RMSE on test set: 101.49793354508876

최적의 파라미터 선정의 문제인지 확인후 모델 시각화 해본 결과 ‘전체인력수합계’ 로만 분류가 진행됨. 해당 피처가 모델 형성에 문제를 야기하는지 파악위해 해당 피처 및 관련된 피처에 대한 조작 진행

1. 화재의 피해 규모에 따라 전체인력수합계가 결정되는 부분이라는 생각 (장비 수도 이와 동일) 해당 feature drop  
     
   R² Score: 0.1477133021850603  
   RMSE on training set: 4.506279594042046e-18  
   RMSE on test set: 19.81998514406888  
     
   <GridSearchCV>  
   R² Score: 0.2293723804424408  
   RMSE on training set: 33.48306418333812

RMSE on test set: 18.846588780068767  
  
일정 부분 모델 결과의 큰 향상이 있음

1. 추가로 feature에 대한 조작 진행 (위와 비슷한 맥락에서 화재진압시간 처리)

<GridSearchCV>  
R² Score: -0.007693133541981201  
RMSE on training set: 33.7606668888528

RMSE on test set: 21.5513569048797  
  
유의미한 결과 도출하지 못함

결론

<GridSearchCV>  
R² Score: 0.2293723804424408  
RMSE on training set: 33.48306418333812

RMSE on test set: 18.846588780068767  
텍스트, 도표, 평면도, 폰트이(가) 표시된 사진

자동 생성된 설명