

1 Data 탐색 2 Feature engineering 3 Origin\_modeling 4 Modeling 5 질의응답

details 위기의 예측

카드 발급사들의 무분별한 신용카드 남발로 인하여 고객들은 상환 능력에 관계없이 신용카드를 과도하게 사용하여 신용카드 빚이 누적되는 결과를 초래. 잘 발달된 금융 시스템에서 위험 예측은 위기 관리보다 선행되어야 하는 과제.







프루젝트 각자 생각하는 시간을 가진 이후 서로의 의견을 토대로 변수를 수정 및 만들고 모델링을 진행하였다. 단계

process



keypoint 프로젝트 핵심

범주화 통합 IMIT\_BAL AGE EDUCATION

3 새로운 Feature

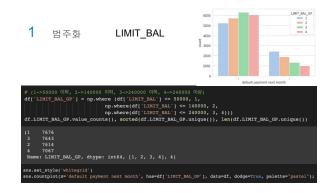
delay LIMIT\_USE\_AVG DELAY/LIMIT\_BAL

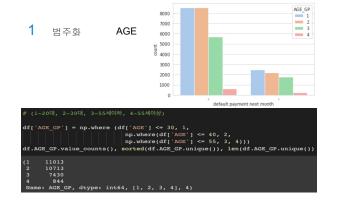
keypoint 프로젝트 핵심

범주화

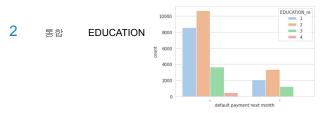
LIMIT\_BAL

AGE

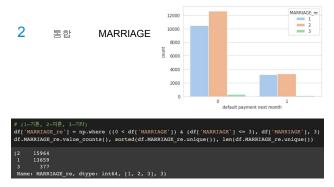








```
if('EDUCATION_re'] = np.where ((0 < df('EDUCATION')) & (df('EDUCATION') <= 3), df('EDUCATION', 4));</pre>
if.EDUCATION_re.value_counts(), sorted(df.EDUCATION_re.unique()), len(df.EDUCATION_re.unique())
      14030
10585
4917
468
:: EDUCATION_re, dtype: int64, [1, 2, 3, 4], 4)
```



```
새로운 Feature
keypoint
프로젝트
                       - DELAY
핵심
                      - LIMIT_USE_AVG
                       - DELAY/LIMIT BAL
```

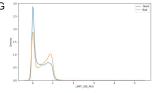
### 3 새로운 Feature DELAY

```
df['delay'].describe()
```



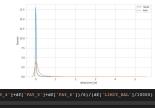
3 새로운 Feature LIMIT\_USE\_AVG

# 3 새로운 Feature LIMIT\_USE\_AVG



```
sns.kdeplot(df[df['default payment next month']==0]['LIMIT_USE_AVG'], label='Good')
sns.kdeplot(df[df['default payment next month']==1]['LIMIT_USE_AVG'], label='Bad')
plt.legend()
```

## 3 새로운 Feature DELAY/LIMIT BAL



plt.figure(figsize=(10.6))
sns.kdeplot(df(df'default payment next month')==0]('delay/Limit\_bal'), label='Good')
sns.kdeplot(df(df'default payment next month')==1]('delay/Limit\_bal'), label='Bad')
plt.xlabel('delay/Limit\_bal')



keypoint

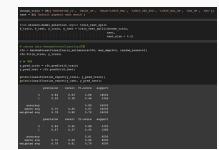
Origin\_modeling

daBoostClassifi

keypoint

Modeling

# RandomForestClassifier



# RandomForestClassifier

Modeling Modeling

