

7-dars uchun REPORT

7-darsda biz **Encodingni** davom ettirdik va so'zngida **For Loopga** ham to'hdaldik. Bunda **Frequency Encoding**, **Target Encoding** va **Ordinal Encodinglar**.

- **Frequency Encoding** bu takrorlanishalr sonini umumiy elementlar soniga nisbati bilan almashtirishdir masalan datasetda dog degan element 10 marta uchragan va umumiy dataset 1000 tadan iborat uni $10/1000 = 0.01$ tashkil etadi.

- **Target Encoding** esa bu output yoki target qiymatining o'rta arifmetigiga ko'ra encoding qilishidir. Masalan; dog 10 marta kelgan va uning most target qiymatini esa 20ga teng shunda $10/20 = 0.5$ tashkil etadi.

- **Ordinal Encoding** esa bu ketma-ketlikka ko'ra sonlar bilan almsahtirishdir. Masalan kiyim o'lchamlarida Ordinal Encoding bilan kattadan kichikka yoki kichikdan kattaga qarab haraktlanadi. Bunda kichik->o'rta-> katta taritib mavjud bo'lsa.

Ikkinci bo'llimda **For Loopga** to'ltaldik. O'rtada bir qator **Data Typeslar** (**int**, **float**, **string**, **object**, **boolean**), **Operations** (arifmetik va mantiqiy amallar), **Conditions** (**if**, **elif**, **else**) kodlarni boshqarish va fitirlash uchun larni qaytardik.

For Loop deb jarayonlarni avtomatlashtirish uchun ishlatalinadigan kodlar bloki.

```
# 1 dan boshlab to 11 gacha sonlar ichida faqat juft
sonlarni chiqaring

for i in range(2,11):
    if i%2==0:
        print(i)
```

range(start, stop, step) shu shaklda ham qilsa bo'ladi if shart operatorisiz ham

```
for i in range(2, 11, 2):
    print(i)
```

Yana biz For Loop + Preprocessing va For Loop + Encodingga ham to'hdaldik.
Bunda

```
# hamma columnlarni nomini chiqarib beradi
```

```
df.columns
```

Boolean orqali True yoki False qiymatlar bilan bironta tushib qolgan qiymat mavjudmi deya tekshiramiz

```
df.isnull().any()
```

```
# demak for loop orqali datasetdagi barcha missing
valuelarni ichiga kirib to'ldirib chiqish uchun
```

```
for col in df.columns:
    if df[col].isnull().any():
        if df[col].dtype == 'object':
            df[col].fillna(df[col].mode()[0],
                           inplace=True)
        else:
```

```
        df[col].fillna(df[col].mean(),  
                     inplace=True)
```

Shu tartibda **For Loop + Preprocessing** bajargan holda barcha tushib qolgan qiymatlarni o'z data typeda to'ldirib chiqdik.

For Loop + Encoding

```
# Birinchi navbatda biz skleardan LabelEncoderni import  
qilib olishimiz kerak  
from sklearn.preprocessing import LabelEncoder  
  
encoder = LabelEncoder()  
encoder  
  
# endi esa biz encoding + for loopni ko'rib chiqamiz  
  
for col in df.columns:  
    if df[col].dtype=='object':  
        if df[col].nunique() <= 5:  
            dummies = pd.get_dummies(df[col],  
                                     prefix=col, dtype=int)  
            df = pd.concat([df.drop(columns=col),  
                           dummies], axis=1)  
        else:  
            df[col] = encoder.fit_transform(df[col])
```