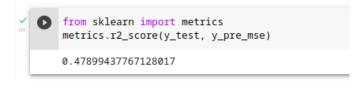
به جز فیچرهای نامبرده و اجاره تمام فیچرهای دیگر را دراپ می کنیم.

Preprocessing را که شامل حذف کردن دیتای پرت اسکیل کردن و .. است انجام می دهیم. همچنین heating را که شامل مدل های زیادی است را ساده تر میکنیم تا با get_dummies ستون های زیادی به دیتافریم اضافه نشه د

سيس أن را از حالت كتگوريكال خارج كرده و نومريكال مي كنيم.

در ادامه باید مدل رگرسیون خطی با میانگین مربع خطاها را پیاده سازی کنیم. این مدل باید توابع فیت و پردیکت را داشته باشد. هر کدام را با توجه به الگوریتم ریاضی آن پیاده سازی میکنیم.

در آخر آنرا را با دیتایی که در دست داریم اموزش می دهیم. سپس r2_score آنرا محاسبه می کنیم.



درباره r2_score:

بهترین امتیاز ممکن 1.0 است و می تواند منفی باشد (زیرا مدل می تواند خودسرانه بدتر باشد). یک مدل ثابت که همیشه مقدار مورد انتظار y را پیش بینی می کند، بدون توجه به ویژگی های ورودی، امتیاز 0.0 را دریافت می کند.

2. برای مقایسه مدل پیاده سازی شده با linear regression از پکیج sklearn مدلی از آن میسازیم و آموزش میدهیم.

میبینیم مدلی که از linear regression با mse ساختیم از مدل sklearn بهتر آموزش داده شده است.

3. برای انجام این بخش نیاز به پیش پردازش های زیادی است که در پروژه قبلی انجام شده و از انها استفاده میکنیم.

```
▼ a) Ridge

{x}
    (320] ridge = Ridge(alpha=1)
            ridge.fit(X_train, y_train)
y_pre_ri = ridge.predict(X_test)
    / [321] metrics.r2_score(y_test, y_pre_ri)
            0.5535044428178639
      b) Lasso
    \frac{\checkmark}{2s} [322] lasso = Lasso(alpha=0.1)
            lasso.fit(X_train, y_train)
            y_pre_la = lasso.predict(X_test)
<>
    [323] metrics.r2_score(y_test, y_pre_la)
\equiv
            0.5537851596018275
>_
```

مدل های lasso و ridge را نیز با استفاده از دیتای پیش پردازش شده اموزش می دهیم. میبینیم که lasso از بین ۴ مدل آموزش داده شده, بهتر عمل کرده است.