We would first imeplment one hot encoding to the features that are categorial

For example the feature:

"cut" - has labels fair, good, very good, permium, ideal

The label "fair" using one hot encoding becomes a feature vector such as:

[1, 0, 0, 0, 0, 0]

The label "good" using one hot encoding becomes a feature vector such as:

[0, 1, 0, 0, 0, 0]

We would then feed each of the features into

$$\frac{1}{n}\sum_{i=1}^{n}(\hat{y_i} - (w^T x_i + c))^2$$

where x_i composes of all the features for a particular diamond

For instance some of the features that would be put into the equation would be:

carat weight = 0.4

cut - [1, 0, 0, 0, 0]

z = 0.5

$\therefore x_1$ = [0.4, 1, 0, 0, 0, 0, 0.5 etc...]