

```
#Import the required libraries
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.impute import SimpleImputer

#Read the dataset
train = pd.read_csv('train.csv')
test = pd.read_csv('test.csv')

#Prepare the training data
X = train.drop(['SalePrice', 'Id'], axis=1)
Y = train['SalePrice']

# Handle missing values and non-numeric columns
X = pd.get_dummies(X)
test_processed = pd.get_dummies(test.drop('Id', axis=1))

# Align train and test sets
X, test_processed = X.align(test_processed, join='left', axis=1, fill_value=0)

# Fill any remaining missing values
imputer = SimpleImputer(strategy='mean')
X = imputer.fit_transform(X)
test_processed = imputer.transform(test_processed)

# 4. Train model
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X, Y)
```



```
RandomForestRegressor
RandomForestRegressor(random_state=42)
```

```
#Make predictions
predictions = model.predict(test_processed)

# 6. Prepare submission
import os

# Create the directory if it doesn't exist
os.makedirs('/mnt/data', exist_ok=True)

submission = pd.DataFrame({
    'Id': test['Id'],
    'SalePrice': predictions
})
```

```
})  
submission.to_csv("/mnt/data/submission.csv", index=False)
```

```
from google.colab import files  
files.download('/mnt/data/submission.csv')
```

