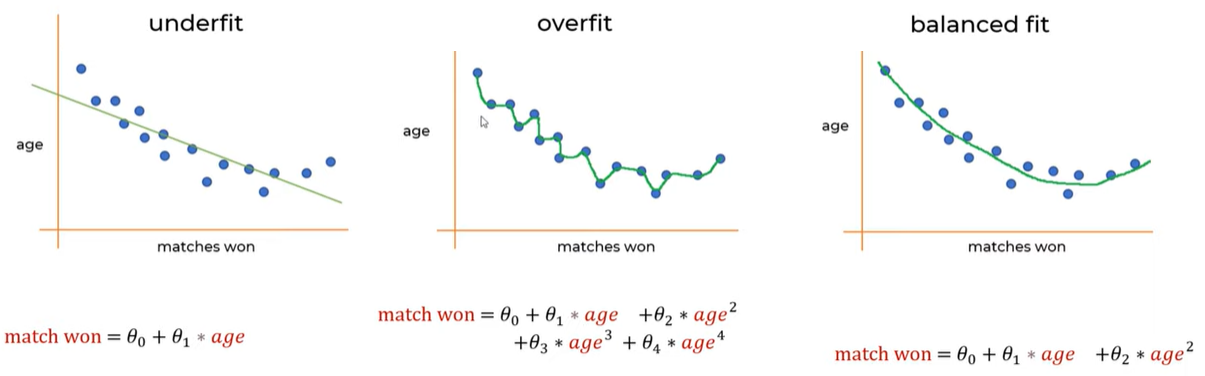
L1, L2 Lasso Ridge elastic

Regularizers

2 techiniques to deal with overfitting and underfitting



import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

import seaborn as sns

import warnings

warnings.filterwarnings('ignore')

dataset = pd.read\_csv('MELBOURNE\_HOUSE\_PRICES\_LESS.csv')

dataset.head(5)

dataset.columns

dataset = dataset.head(10000)

dataset.nunique()

dataset.isna().sum()

#if empty fill with zero

cols\_to\_fill\_zero=['Propertycount','Distance','Bedroom2','Bathroom','Car']

dataset[cols\_to\_fill\_zero]= dataset[cols\_to\_fill\_zero].fillna(0)

dataset.isna().sum()

dataset['Landsize']     = dataset['Landsize'].fillna(dataset.Landsize.mean())

dataset['BuildingArea'] = dataset['BuildingArea'].fillna(dataset.BuildingArea.mean())

#drop if there is null

dataset.dropna(inplace=True)

dataset['Landsize']     = dataset['Landsize'].fillna(dataset.Landsize.mean())

dataset['BuildingArea'] = dataset['BuildingArea'].fillna(dataset.BuildingArea.mean())

#drop if there is null

dataset.dropna(inplace=True)

dataset =pd.get\_dummies(dataset, drop\_first=True)

dataset.head()

dataset.dropna(inplace=True)

X = dataset.drop('Price',axis=1)

y = dataset['Price']

from sklearn.model\_selection import train\_test\_split

train\_X, test\_X, train\_y, test\_y = train\_test\_split(X, y, test\_size=0.3, random\_state=2)

from sklearn.model\_selection import train\_test\_split

train\_X, test\_X, train\_y, test\_y = train\_test\_split(X, y, test\_size=0.3, random\_state=2)

from sklearn.linear\_model import LinearRegression

reg = LinearRegression().fit(train\_X, train\_y)

reg.score(test\_X, test\_y)

reg.score(train\_X, train\_y)

from sklearn import linear\_model

lasso\_reg = linear\_model.Lasso(alpha=50, max\_iter=100, tol=0.1)

lasso\_reg.fit(train\_X, train\_y)

lasso\_reg.score(test\_X, test\_y)

lasso\_reg.score(train\_X, train\_y)

from sklearn import linear\_model

ridge\_reg = linear\_model.Ridge(alpha=50, max\_iter=100, tol=0.1)

ridge\_reg.fit(train\_X, train\_y)

ridge\_reg.score(test\_X, test\_y)

ridge\_reg.score(train\_X, train\_y)

LASSO

from sklearn import linear\_model

lasso\_reg = linear\_model.Lasso(alpha=50, max\_iter=100, tol=0.1)

lasso\_reg.fit(train\_X, train\_y)

lasso\_reg.score(test\_X, test\_y)

lasso\_reg.score(train\_X, train\_y)

RIDGE

from sklearn import linear\_model

ridge\_reg = linear\_model.Ridge(alpha=50, max\_iter=100, tol=0.1)

ridge\_reg.fit(train\_X, train\_y)

ridge\_reg.score(test\_X, test\_y)

ridge\_reg.score(train\_X, train\_y)