

Computers and Artificially intelligence

AI330_ Machine Learning Projects_Fall2023

ID	Name	Grade
20211050	ياسين محمد يس نصر	
20210192	انس صالح موسى الصادق	
20210191	اندرو عاطف شكر الله	
20210190	اندرو شريف شكري عزيز	
20210956	منار مصطفى محمود مكرم	
20211002	ندى محمد فكري احمد عيسى	

General Information on dataset:

Numerical datasets	Image datasets
Name of data sets=House Prices - Advanced Regression Techniques	Name of image= <i>Fruits 360 Dataset</i> Link : https://www.kaggle.com/moltean/fruits
Total number of sample: 1858 rows 81 columns	Size=1GB
Training=1460 rows Testing=1460 rows Validation= 1858 rows used Total number of samples=2919 rows & 81 columns	Used 5 classes Labels: <ol style="list-style-type: none">1. Apple red yellow2. Avocado3. Banana4. Blueberry5. watermelon
	Training=67692 images Testing=22688 images Validation=2346 images in 5 classes

Implementation details:

1. feature extraction for numerical:
you use 70 column only

0	LotFrontage
1	LotArea
2	OverallQual
3	YearBuilt
4	YearRemodAdd
5	MasVnrArea
6	BsmtFinSF1
7	BsmtFinSF2
8	BsmtUnfSF
9	TotalBsmtSF
10	1stFlrSF
11	2ndFlrSF
12	LowQualFinSF
13	GrLivArea
14	BsmtFullBath
15	BsmtHalfBath
16	FullBath
17	HalfBath
18	KitchenAbvGr
19	TotRmsAbvGrd
20	Fireplaces
21	GarageYrBlt
22	GarageCars
23	GarageArea
24	WoodDeckSF
25	OpenPorchSF
26	EnclosedPorch
27	3SsnPorch
28	ScreenPorch
29	PoolArea
30	MiscVal
31	SalePrice
32	MSZoning_RL
33	MSZoning_RM
34	LotShape_Reg
35	Neighborhood_NAMES
36	Neighborhood_NoRidge
37	Neighborhood_NridgHt
38	Neighborhood_OldTown
39	HouseStyle_2Story

40	RoofStyle_Gable
41	RoofStyle_Hip
42	Exterior1st_VinylSd
43	Exterior2nd_VinylSd
44	MasVnrType_BrkFace
45	MasVnrType_Stone
46	ExterQual_Gd
47	ExterQual_TA
48	Foundation_CBlock
49	Foundation_PConc
50	BsmtQual_Gd
51	BsmtQual_TA
52	BsmtExposure_Gd
53	BsmtExposure_No
54	BsmtFinType1_GLQ
55	HeatingQC_TA
56	CentralAir_Y
57	Electrical_SBrkr
58	KitchenQual_Gd
59	KitchenQual_TA
60	FireplaceQu_Gd
61	GarageType_Attchd
62	GarageType_BuiltIn
63	GarageType_Detchd
64	GarageFinish_RFn
65	GarageFinish_Unf
66	PavedDrive_Y
67	SaleType_New
68	SaleType_WD
69	SaleCondition_Partial

In image:

Using HOG feature extraction

2. Yes, use cross validation, number of fold=5, ratio of training/validation=80/20