Helwan University

Faculty of Computers and Artificial Intelligence

Machine learning AI330

Fall 2023-2024 First Semester.

Machine learning Document

**Team number:70**

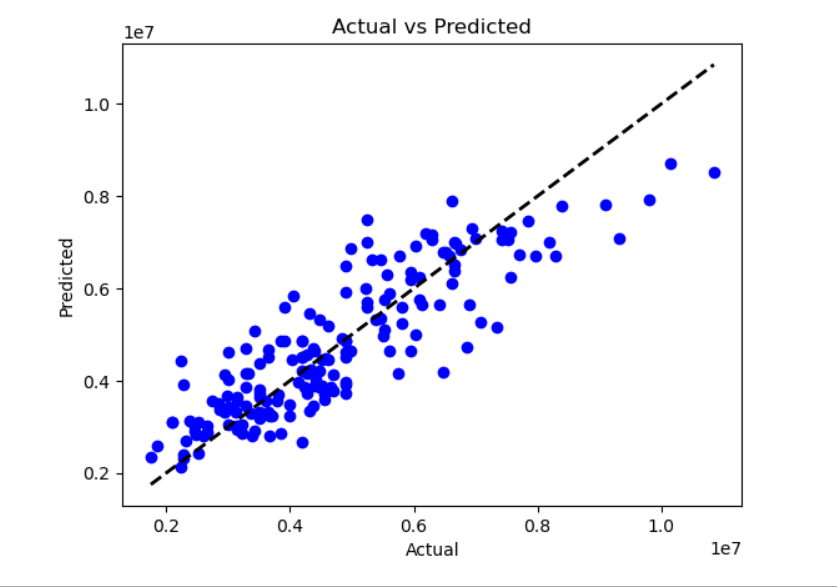
**1**

**General information**

|  |  |  |
| --- | --- | --- |
| *1* | Model name: | *Linear regression* |
| *2* | Name of dataset: | Hedonic prices and the demand for clean air |
| *3* | Total number of samples: | *545* |
| *4* | Size for each image (in case of image ): | *---* |
| *5* | Number of samples used in training | *365* |
| *6* | Number of samples used in *validation* | *73* |
| *7* | Number of samples used in *testing* | *180* |

**Implementation details**

|  |  |  |
| --- | --- | --- |
| *1* | *How many features were extracted:* | *10 features were selected out of 12 features* |
| *2* | *Names of feature extraction:* | *'area' , 'bedrooms' ,'bathrooms' , 'stories' , 'mainroad' ,'guestroom' , 'basement' ,'airconditioning' , 'parking' ,'prefarea'* |
| *3* | *Dimension of resulted features:* | *( 545 ,10 )* |



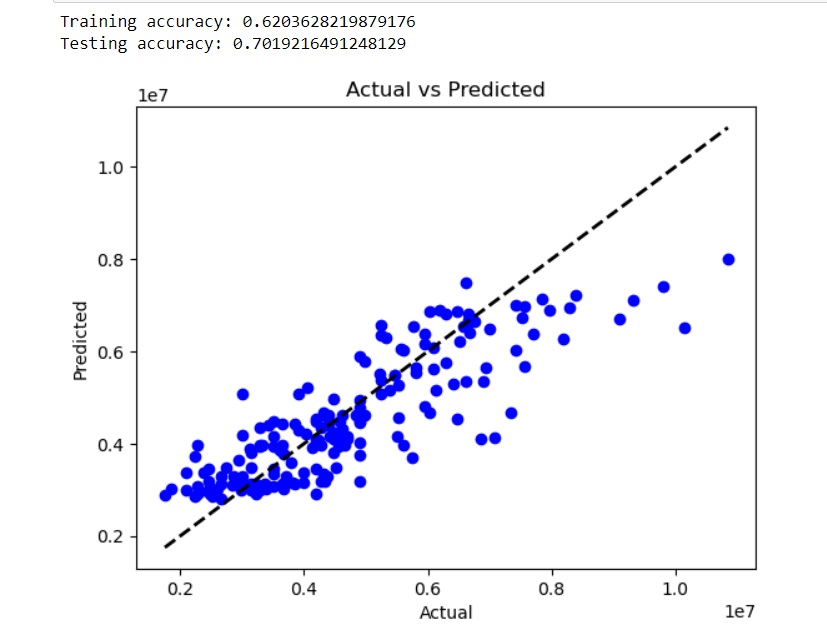
**2**

**General information**

|  |  |  |
| --- | --- | --- |
| *1* | Model name: | *K-Nearest Neighbors* |
| *2* | Name of dataset: | Hedonic prices and the demand for clean air |
| *3* | Number of classes and the labels: | *12* |
| *4* | Total number of samples: | *545* |
| *5* | Size for each image (in case of image ): | *----* |
| *6* | Number of samples used in training | *365* |
| *7* | Number of samples used in *validation* | *73* |
| *8* | Number of samples used in *testing* | *180* |

**Implementation details**

|  |  |  |
| --- | --- | --- |
| *1* | *How many features were extracted:* | *10 features were selected out of 12 features* |
| *2* | *Names of feature extraction:* | *'area' , 'bedrooms' ,'bathrooms' , 'stories' , 'mainroad' ,'guestroom' , 'basement' ,'airconditioning' , 'parking' ,'prefarea'* |
| *3* | *Dimension of resulted features:* | *( 545 ,10 )* |



3

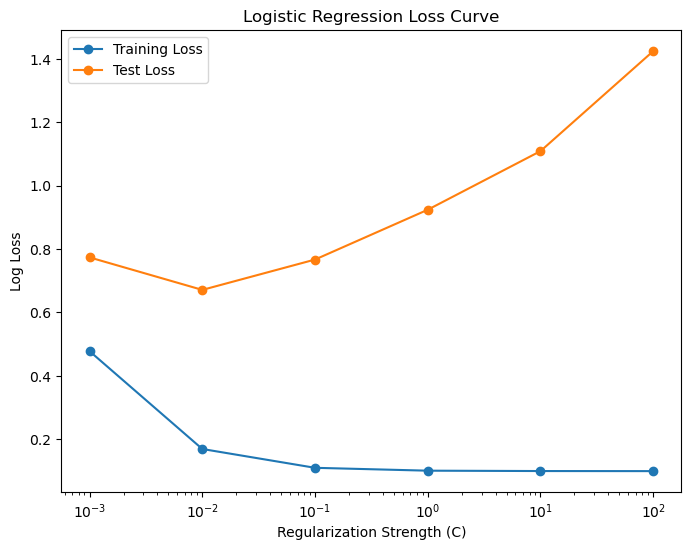
**General information**

|  |  |  |
| --- | --- | --- |
| *1* | Model name: | *Logistic regression* |
| *2* | Name of dataset: | Tomato detection |
| *3* | Number of classes and the labels: | *4* |
| *4* | Total number of samples: | *1300* |
| *5* | Size for each image (in case of image ): | *640x640 pixel 47.6 KB approximately* |
| *6* | Number of samples used in training | *700* |
| *7* | Number of samples used in *validation* | *300* |
| *8* | Number of samples used in *testing* | *300* |

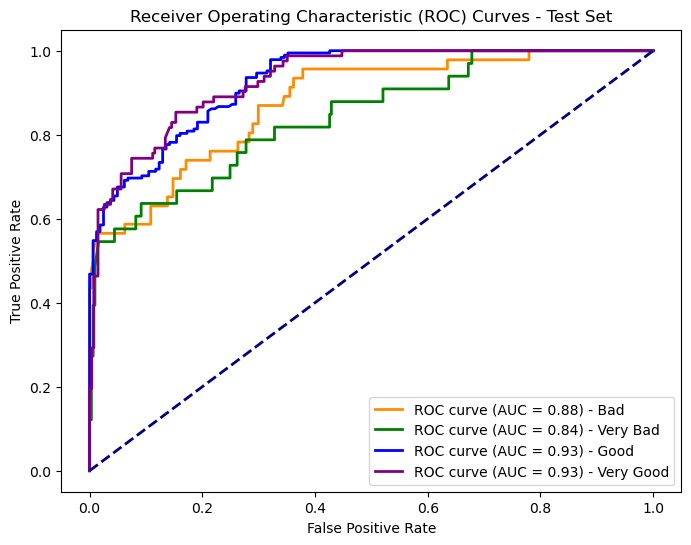
**Implementation details**

|  |  |  |
| --- | --- | --- |
| *1* | *Feature extraction description:* | *HOG* |
| *2* | *How many features were extracted:* | *2* |
| *3* | *Dimension of resulted features:* | *(714, 2)* |

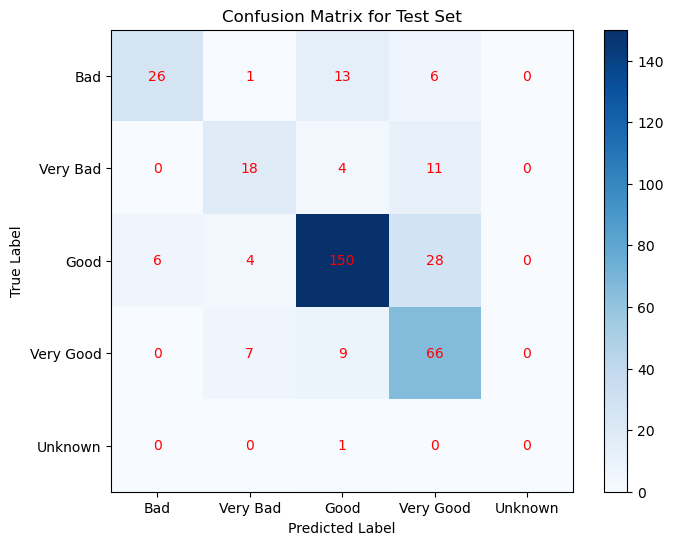
Loss curve:



ROC:



Confusion Matrix:



4

**General information**

|  |  |  |
| --- | --- | --- |
| *1* | Model name: | *K-means* |
| *2* | Name of dataset: | Tomato detection |
| *3* | Number of classes and the labels: | *4* |
| *4* | Total number of samples: | *1300* |
| *5* | Size for each image (in case of image ): | *640x640 pixel 47.6 KB approximately* |
| *6* | Number of samples used in training | *700* |
| *7* | Number of samples used in *validation* | *300* |
| *8* | Number of samples used in *testing* | *300* |

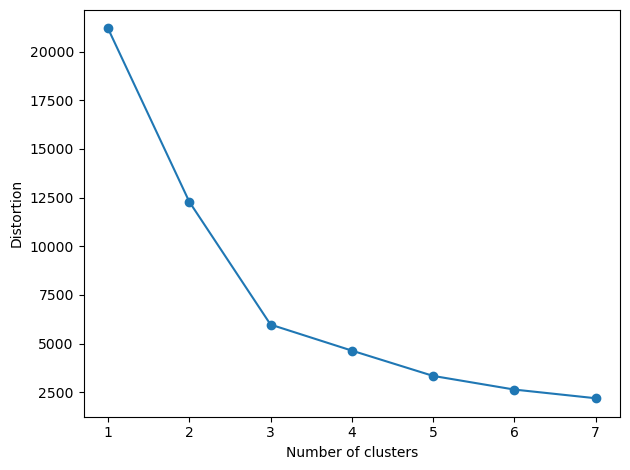
**Implementation details**

|  |  |  |
| --- | --- | --- |
| *1* | *Feature extraction description:* | *HOG* |
| *2* | *How many features were extracted:* | *2* |
| *3* | *Dimension of resulted features:* | *(714, 2)* |

Cross-validation (if used) :

|  |  |  |
| --- | --- | --- |
| *1* | Number of folds: | *5* |
| *3* | *Hyperparameters:* | *Number of clusters :3*  *Initial =10* |

Number of clusters with different distortion



K-means Clustering :

