Machine Learning Documentation

Dataset – Numeric info

#	Column	Non-Null Count Dtype
RangeIndex: 13580 entries, 0 to 13579 Data columns (total 21 columns):		
00	Suburb	13580 non-null object
01	Address	13580 non-null object
02	Rooms	13580 non-null int64
03	Type	13580 non-null object
04	Price	13580 non-null float64
05	Method	13580 non-null object
06	SellerG	13580 non-null object
07	Date	13580 non-null object
08	Distance	13580 non-null float64
09	Postcode	13580 non-null float64
10	Bedroom2	13580 non-null float64
11	Bathroom	13580 non-null float64
12	Car	13518 non-null float64
13	Landsize	13580 non-null float64
14	BuildingArea	7130 non-null float64
15	YearBuilt	8205 non-null float64
16	CouncilArea	12211 non-null object
17	Lattitude	13580 non-null float64
18	Longtitude	13580 non-null float64
19	Regionname	13580 non-null object
20	Propertycount	13580 non-null float64
dtypes: float64(12), int64(1), object(8) memory usage: 2.2+ MB		

Dataset name : Melbourne Housing

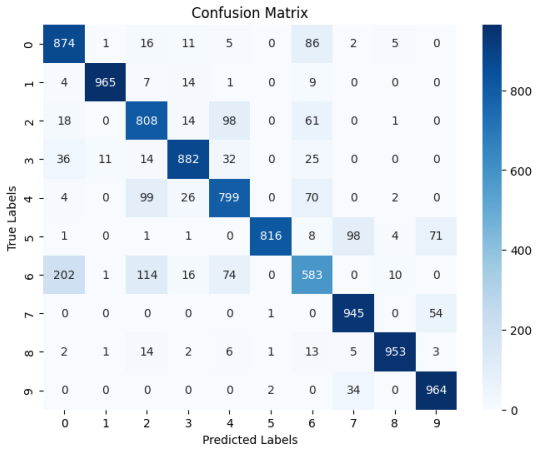
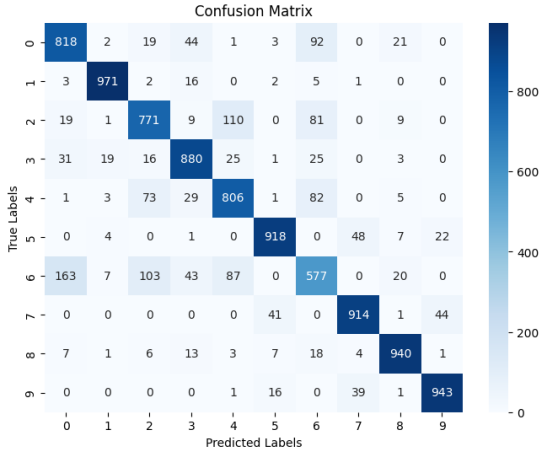
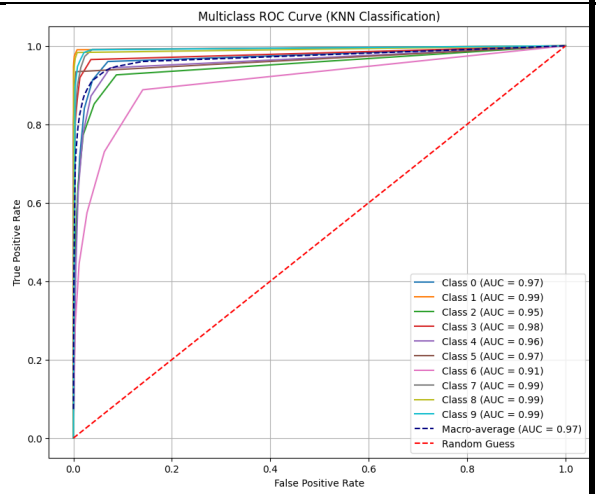
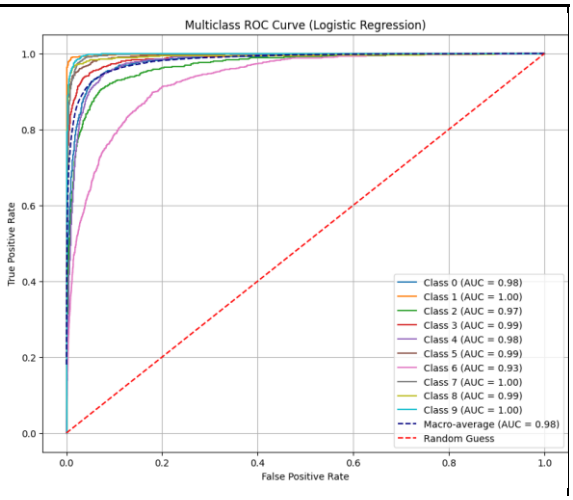
Training data : 10814

Testing data : 2704

Numeric Regression Comparison

	KNN	Logistic Regression
MAE	142,499.1804997887	172,953.1034871587
MSE	38,828,957,242.78912	48,856,264,964.51134
R^2	0.7999329109588743	0.7482669789004716

Image Classification Comparison

	KNN	Logistic Regression
Accuracy	0.8589	0.8538
Loss	0.20878811613789466	0.4028998395555939
Confusion Matrix		
Precision	0.8613943216647832	0.8517173505898177
Recall	0.8589	0.8538
ROC | AUC Graph		

Dataset name : Fashion MNIST

Number of Classes : 10

Dimensions (Columns without the label) : 28 x 28 → 784

Number of Rows : (70,000 total) → 60,000 for training | 10,000 for testing