

# Pràctica 2 APC

## Classificació

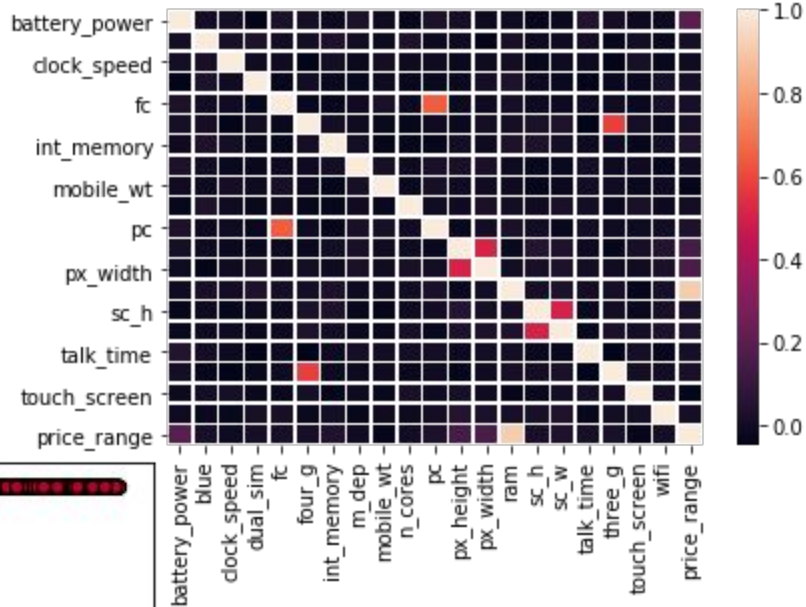
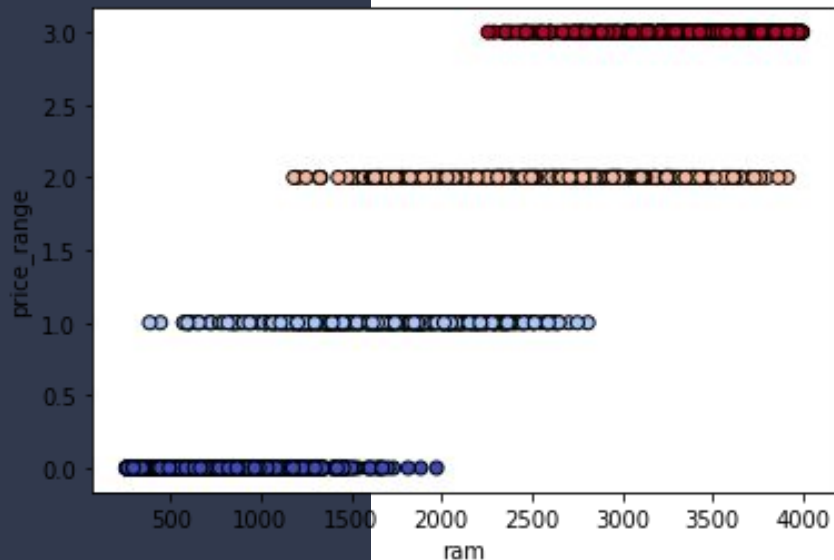
Àlex Correa Orri 1564967  
Júlia Pumares Benaiges 1566252

# Apartat B: IRIS dataset

Mètode	0.5	0.7	0.8
SVM	0.85	0.76	0.77
KNeighbors	0.67	0.73	0.7
Decision Tree	0.63	0.8	0.7
Naive Bayes	0.85	0.78	0.73
Lineal	0.73	0.64	0.61
Logistic	0.67	0.78	0.5
Adaboost	0.76	0.38	0.7
Bagging	0.69	0.67	0.67
Perceptron	0.37	0.51	0.5

# 1. EDA

- Classificar el preu d'un mòbil en 4 grups de rangs de preu.
- 20 atributs
  - 6 binaris
  - 14 numèrics
- 500 mostres d'entrenament per classe



## 2. Preprocessing

- Reduir la dimensió agafant els atributs més correlacionats
- Estandaritzar



### 3. Model Selection

Méthode	0.5	0.7	0.8
SVM	0.91	0.93	0.94
KNeighbors	0.9	0.91	0.91
Logistic	0.85	0.84	0.84
Decision Tree	0.74	0.75	0.75
Perceptron	0.57	0.72	0.69
Bagging	0.72	0.71	0.33
Adaboost	0.84	0.7	0.78

## 4. Cross Validation

SVM:

```
Scores      : [0.93  0.9225 0.9225 0.9225 0.9225]
Scores Mean : 0.924
Scores std  : 0.00300000000000000248
```

KNeighborsClassifier:

```
Scores      : [0.905 0.91  0.9275 0.905 0.9075]
Scores Mean : 0.910999999999999999
Scores std  : 0.008455767262643871
```

LogisticRegression:

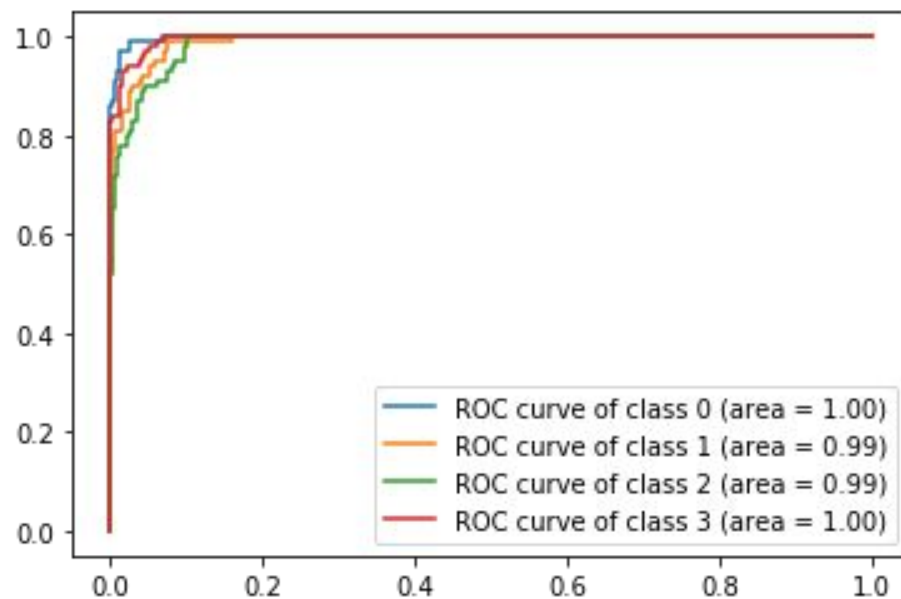
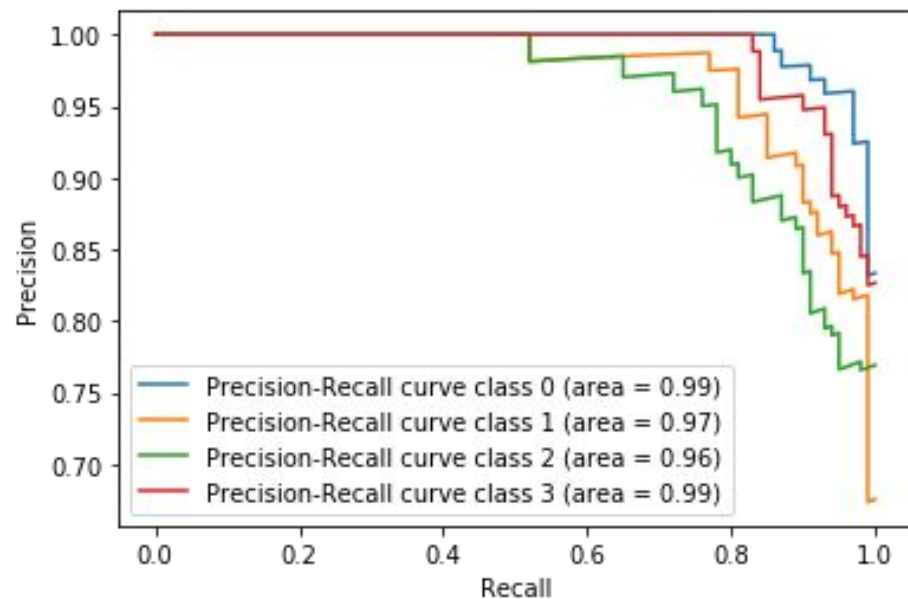
```
Scores      : [0.845 0.8325 0.855 0.845 0.82 ]
Scores Mean : 0.839499999999999999
Scores std  : 0.012083045973594577
```

BaggingClassifier:

```
Scores      : [0.6175 0.6525 0.7075 0.7225 0.5025]
Scores Mean : 0.640500000000000001
Scores std  : 0.07865112840894277
```

# 5. Metrics Analysis SVM

Correct classification SVM:				0.94
	precision	recall	f1-score	support
class 0	0.94	0.98	0.96	100
class 1	0.95	0.91	0.93	100
class 2	0.92	0.93	0.93	100
class 3	0.96	0.95	0.95	100
accuracy			0.94	400
macro avg	0.94	0.94	0.94	400
weighted avg	0.94	0.94	0.94	400

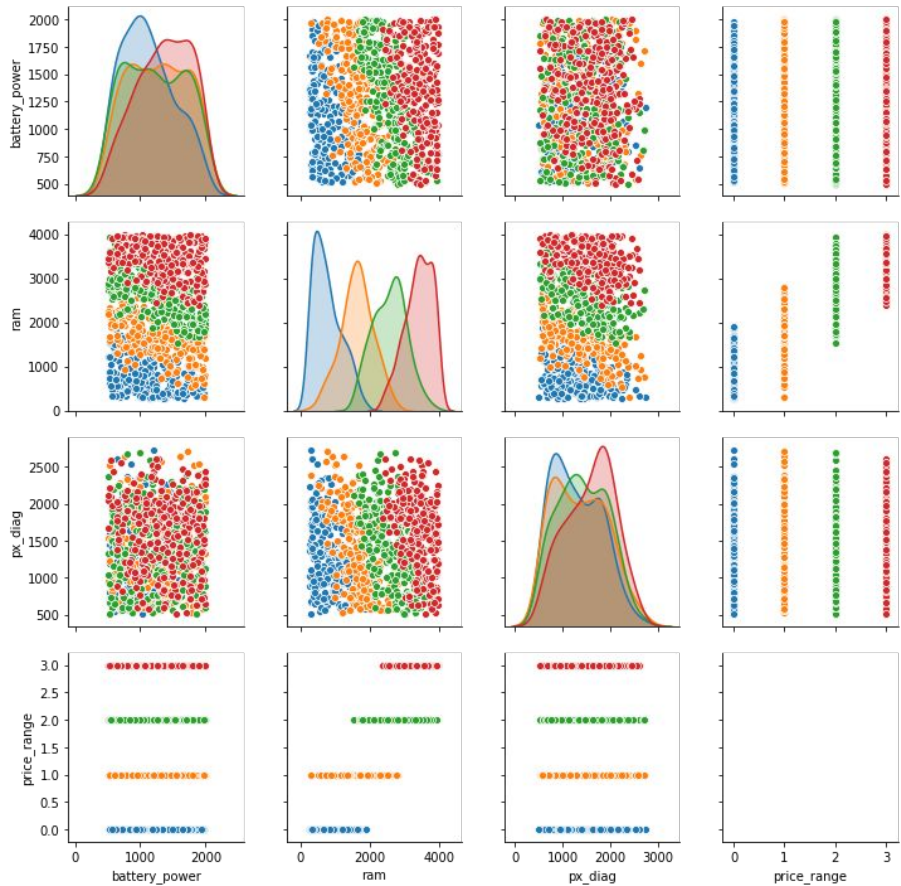


# 6. Hyperparameter Search

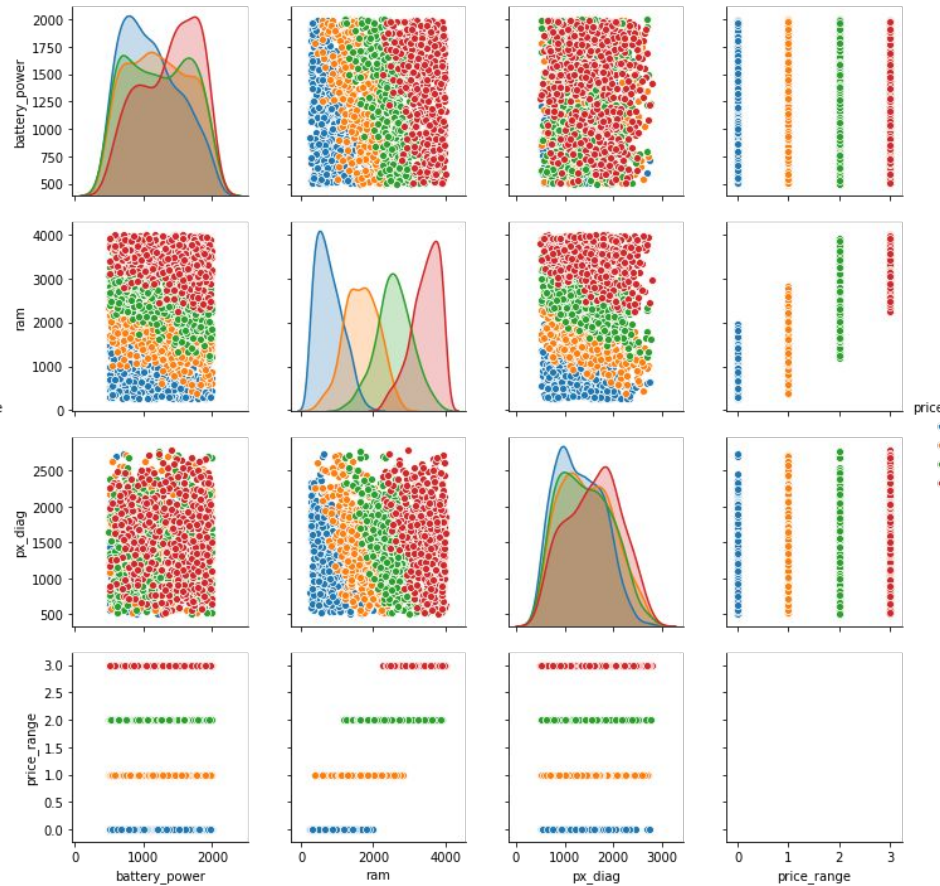
SVM:

- $C = 35$
- $\gamma = 0.1$
- score = 0.9425





TRAIN



TEST

# Conclusions

Els mètodes que ens han funcionat millor han sigut el SVM i el KNeighbors