Ludwig-Maximilians-Universität München
Institut für Informatik
Prof. Dr. Thomas Seidl
Janina Sontheim, Maximilian Hünemörder

Knowledge Discovery and Data Mining I

WS 2019/20

Exercise 3: Classification Evaluation, m-fold Cross Validation, Naïve Bayes Classifier

Exercise 3-1 Evaluation of classifiers

Given a data set $D = \{o_1, \ldots, o_n\}$ with known class labels $C(o_i) \in \mathcal{C} = \{A, B, C\}$ of the objects. In order to evaluate the quality of a classifier K, each object $o_i \in D$ is additionally classified using K, yielding class label $K(o_i)$. The results are given in the table below.

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$C(o_i)$	A	B	A	C	C	B	A	A	A	B	B	C	C	C	B
$K(o_i)$	A	A	C	C	B	B	A	A	A	C	A	A	C	C	B

(a) Setup the confusion matrix.

(b) Compute the accuracy / classification error.

(c) For each class $i \in \mathcal{C}$ compute precision and recall.

(d) To get a complete measure for the quality of the classification with respect to a single class, the F_1-measure (the harmonic mean of precision and recall) is commonly used. It is defined as follows:

$$F_1(K, i) = \frac{2 \cdot \text{Recall}(K, i) \cdot \text{Precision}(K, i)}{\text{Recall}(K, i) + \text{Precision}(K, i)}$$

Compute the F_1-measure for all classes.

Note: "F_1-measure" may refer to the same formula but computed using a different precision and different recall in other applications. It is a specialization of F_β with equal weighting of precision and recall.

(e) So far, the F_1-measure is only defined for classes and not yet useful to get an overview of the overall performance of the classifiers. For this, one commonly takes the average over all classes using one of the following two approaches:

(i) **Micro Average F_1-Measure**: The values of TP, FP and FN are added up over all classes. Then precision, recall and F_1-measure are computed using these sums.

(ii) **Macro Average F_1-Measure**: Precision and recall are computed for each class individually, afterwards the average precision and average recall are used to compute the F_1-measure.

Compute the Micro- and Macro-Average F_1-measures for the example above. What do you observe?

Exercise 3-2 m-fold Cross Validation

Suppose, you have a 2-dimensional dataset consisting of 5 classes with 90 objects each, arranged as follows

$$\overbrace{x_0, \ldots, x_{89}}^{C(x)=0}, \overbrace{x_{90}, \ldots, x_{179}}^{C(x)=1}, \ldots, \overbrace{x_{360}, \ldots, x_{449}}^{C(x)=4},$$

and that the classes are linearly separable (i.e. can be separated using a hyperplane). Suppose further, that someone has produced a poor implementation of the m-fold cross validation procedure and applied it in combination with a multi-class linear classifier to obtain the following results:

m	2	3	5	6	10
accuracy	20%	40%	0%	100%	100%

What is the problem with the implementation of the m-fold cross validation? Describe and explain the result for each value of m in short and precise sentences. How could the implementation be improved?

Exercise 3-3 Naive Bayes

The skiing season is open. To reliably decide when to go skiing and when not, you could use a classifier such as Naive Bayes. The classifier will be trained with your observations from the last year. Your notes include the following attributes:

- The weather: The attribute `weather` can have the following three values: sunny, rainy and snow.
- The snow level: The attribute `snow level` can have the following two values: $\geq$50 (There are at least 50 cm of snow) and $<$50 (There are less than 50 cm of snow).

Assume you wanted to go skiing 8 times during the previous year. Here is the table with your decisions:

weather	sunny	rainy	rainy	snow	snow	sunny	snow	rainy
snow level	$<$50	$<$50	$\geq$50	$\geq$50	$<$50	$\geq$50	$\geq$50	$<$50
ski?	no	no	no	yes	no	yes	yes	yes

(a) Compute the *a priori* probabilities for both classes `ski = yes` and `ski = no` (on the training set)!

(b) Compute the conditional distributions for the two classes for each attribute.

(c) Decide for the following weather and snow conditions, whether to go skiing or not! Use the Naive Bayes classificator for finding the decision.

day	**weather**	**snow level**
A	sunny	$\geq$50
B	rainy	$<$50
C	snow	$<$50