

NAME

pctest – Classify sequences using a trained classifier.

SYNOPSIS

pctest [*options*] *serialized-classifier.txt input.txt*

DESCRIPTION

Read a trained classifier and a set of sequences and apply the classifier to each sequence of the set. A confusion matrix and total classification loss is given for the input set.

PARAMETERS

--confusion-matrix <file>

Write a confusion matrix for the sequence set to the given file.

--dump-classifier <file>

Write a textual description of the serialized classifier.

INPUT FILE FORMAT

The input file consists of n lines, each containing a filename and a class label, separated by any whitespace character. The filename corresponds to a sequence file. If only two classes are considered, the class labels are "1" and "-1", for the positive and negative class, respectively. If more classes are considered, the first class is labeled "0", and the following classes "1", "2", etc.

Each sequence file consists of p lines, each line representing one sequence element. Each element contains a number of items (positive integer numbers), separated by whitespace characters. The items within each element must be ordered and contain no duplicate item.

BUGS

No bugs known, if you find any, please send a bug report to me. I will try to fix it.

AUTHOR

Sebastian Nowozin <sebastian.nowozin@tuebingen.mpg.de>

SEE ALSO

pboost(1), **pspan**(1)