Group J GHC Tasks

(to be presented on Nov. 9, 2023)

**Support Vector Machines**

a) The soft margin support vector machine solves the following optimization problem:



What does the second term minimize (be precise!)? What is the purpose of C? How many examples are misclassified in the figure below! Add arrows to all examples in the figure below, whose ξi values are positive---the length of the arrow should correspond with the value to the respective ξi !What is the relationship between the value of ξi and example i being misclassified?



b) SVMs have been successfully used in conjunction with Kernels. How does this approach exactly work? Why do you believe it has been often successful for many datasets, achieving very high accuracies?

Group K GHC Tasks

(to be presented on Nov. 9, 2023)

a) In which steps does the APRIORI algorithm take advantage of the APRIORI property?

b) Assume the APRIORI algorithm identified the following 9 4-item sets that satisfy a user given support threshold: **abcd, acde, acdf, acdg, adfg, bcde, bcdf, bcdg, cdef, cdeg;** what initial candidate 5-itemsets are created by the APRIORI algorithm; which of those survive subset pruning?

c) How does APRIORI algorithm create rules after frequent item sets have been computed.

d) Assume an association rule *if smoke then cancer* has a confidence of 86% and a high lift of 2.1. What does this tell you about the relationship of smoking and cancer?