

GLPK Case Study 4 - 60016 Operations Research

We consider a public dataset¹ concerning Italian wines, called `wines.dat`. The dataset gives chemical analyses for a set of wines grown in the same region in Italy, but obtained from two different cultivars. For each cultivar, the analysis reports concentrations of alcohol, magnesium, flavanoids, wine color, protein concentration, and several other features. The dataset describes N_1 wines from the first cultivar and N_2 wines from the second cultivar. The question is whether it is sufficient to measure these chemical features to tell if a given wine comes from a cultivar or from the other.

Write a GMPL program that solves this problem for `wines.dat`. Solve it with `glpsol` and discuss what procedure has been used by the solver to find the answer.

¹The dataset comes from a public repository: <https://archive.ics.uci.edu/ml/datasets/Wine>.