

Optimization techniques for Kernel Logistic Regression on large-scale datasets: A comparative study

José Ángel Martín-Baos¹, Ricardo García-Ródenas¹, Luis Rodríguez-Benitez²

¹Department of Mathematics, University of Castilla-La Mancha, Spain, ²Department of Information and System Technologies, University of Castilla-La Mancha, Spain.

Abstract

In recent years, machine learning techniques have been increasingly applied to modelling the decision-making processes of individuals. One technique that has shown good results in the literature for modelling complex behaviours is the Kernel Logistic Regression (KLR). However, standard KLR implementations have a time complexity of $\mathcal{O}(n^3)$, which is not feasible for large datasets. To overcome this limitation, one of the proposed alternatives is to approximate the kernel matrix using the Nyström method. The aim of this work is to evaluate the Nyström KLR model on large-scale datasets and to study, at the experimental level, which of the optimisation techniques that allow training this model is the most efficient. As results, the authors show that the Nyström method efficiently computes the objective function and its gradient, enabling the training of KLR models with up to 10^5 parameters. Then, it is evaluated the performance of several optimisation methods, including gradient descend, Momentum, Adam, and L-BFGS-B. It can be concluded that L-BFGS-B is the most efficient method for training the Nyström KLR model. However, given enough computational time and proper hyperparameter tuning, the Adam method can also yield good results.

Keywords: *Random Utility Model; Machine Learning; Kernel Logistic Regression; Nyström method; Numerical Optimisation.*
