

Survival analysis: finding relevant epistatic SNP pairs using Model-Based Multifactor Dimensionality Reduction

Francois Van Lishout^{1,2}, Celine Vens³, Victor Urrea⁴, M. Luz Calle⁴, Louis Wehenkel^{1,2}, Kristel Van Steen^{1,2}

¹*Montefiore Institute, University of Liege, 4000 Liege, Belgium*

²*GIGA-Research, University of Liege, 4000 Liege, Belgium*

³*Computer Science Department, K.U. Leuven, 3000 Leuven, Belgium*

⁴*Departement of Systems Biology, University of Vic, 08500 Vic, Spain*

Correspondence: f.vanlishout@ulg.ac.be

Key words: Model-Based Multifactor Dimensionality Reduction, Epistasis, Survival analysis

Analyzing the combined effects of genes (and/or environmental factors) on the development of complex diseases is quite challenging, both from the statistical and computational perspective, even using a relatively small number of genetic and non-genetic exposures. Several data-mining methods have been proposed for interaction analysis, among them, the Multifactor Dimensionality Reduction Method (MDR). Model-Based Multifactor Dimensionality Reduction (MB-MDR), a relatively new dimensionality reduction technique, is able to unify the best of both nonparametric and parametric worlds, and has proven its utility in a variety of theoretical and practical settings. Until now, MB-MDR software has only accommodated traits that are measured on a binary or interval scale. Time-to-event data could therefore not be analyzed with the MB-MDR methodology. MB-MDR-3.0.0 overcomes this shortcoming of earlier versions. We show the added value of MB-MDR for censored traits by comparing the implemented strategies with more classical methods such as those based on a parametric regression paradigm. The simulation results are supplemented with an application to real-life data.