Summary: We offer a variational Bayes approximation to the multinomial probit model for basis expansion and kernel combination. Our model is well-founded within a hierarchical Bayesian framework and is able to instructively combine available sources of information for multinomial classification. The proposed framework enables informative integration of possibly heterogeneous sources in a multitude of ways, from the simple summation of feature expansions to weighted product of kernels, and it is shown to match and in certain cases outperform the well-known ensemble learning approaches of combining individual classifiers. At the same time the approximation reduces considerably the CPU time and resources required with respect to both the ensemble learning methods and the full Markov chain Monte Carlo, Metropolis-Hastings within Gibbs solution of our model. We present our proposed framework together with extensive experimental studies on synthetic and benchmark datasets and also for the first time report a comparison between summation and product of individual kernels as possible different methods for constructing the composite kernel matrix.

Keywords: variational Bayes approximation; multiclass classification; kernel combination; hierarchical Bayes; Bayesian inference; ensemble learning; multi-modal modelling; information integration