Program

Tenth Annual Winter Workshop
Bayesian Model Selection and Objective Methods

Department of Statistics
University of Florida
January 11-12,2008

Contents

Sponsors

National Science Foundation, InfoTech, the Graduate School, and the Department of Statistics of the University of Florida.

Organizing Committee

Hani Doss
George Casella
Malay Ghosh

Invited Speakers

Merli se Clyde, Duke University
Ming-Hui Chen, University of Connecticut
David Draper, University of California, Santa Cruz
Dean Foster, University of Pennsylvania
Ed George, University of Pennsylvania
Michael Jordan,University of California, Berkeley
David Madigan, Columbia University
Glen Meeden, University of Minnesota
Adrian Raftery, University of Washington
Marina Vannucci, Rice University
Yuhong Yang, University of Minnesota

Acknowledgements

The organizers thank the Department of Statistics staff, Robyn Crawford, Tina Greenly, Allison Pipkin, Carol Rozear, and Marilyn Saddler for their tremendous efforts in helping to set up this meeting and make it run smoothly.

Conference Schedule

Thursday, January 10, 2008

6:00-8:00 pm reception at Keene Faculty Center, Dauer Hall

All workshop sessions meet in the Rion Ballroom, 2nd floor, J. Wayne Reitz Union

Friday, January 11, 2008

8:00am Continental Breakfast

8:45am Welcome by:
Hani Doss, Conference Chair
Dr. Joseph Glover, Interim Dean, College of Liberal Arts & Sciences

9:00am **Session 1:** ***Regression Trees***
Chair: Hani Doss, University of Florida

Speakers: Ed George, University of Pennsylvania
Pre-Modeling Via BART

10:00-10:30am **Break and Conference photo at JWRU North side on the Colonnade steps**

10:45-12:45pm **Session 2:** ***Priors on Infinite-Dimensional Spaces***
Chair: George Casella, University of Florida

Speakers: Michael Jordan, University of California, Berkeley
Hierarchical Nonparametric Bayes

Merlise Clyde, Duke University
Towards Objective Priors and Nonparametric Regression and Classification

12:45-2:15pm **Lunch** (Gator Comer Dining Center)

2:15-3:15pm **Session 3:** ***Objective Priors***
Chair: Trevor Park, University of Florida

Speakers: Glen Meeden, University of Minnesota
A Noninformative Bayesian Approach to Finite Population Sampling Using Auxiliary Variables

Ming-Hui Chen, University of Connecticut
Objective Bayesian Variable Selection for Binomial Regression Models with Jeffreys's Prior

3:30-5:30pm **Poster Session:** Rion Ballroom, 2nd floor, J. Wayne Reitz Union

All workshop sessions meet in the Rion Ballroom, 2nd floor, J. Wayne Reitz Union

Saturday, January 12, 2008

8:00-8:30 Continental Breakfast

8:30-10:30am **Session 4:** ***Machine Learning***
Chair: Xueli Liu, University of Florida

Speakers: David Madigan, Columbia University
High-Dimensional Bayesian Classifiers

Dean Foster, University of Pennsylvania
On the Intrinsic Dimensionality of Multi-View Regression

10:30-11:00am **Break**

11:00-1:00pm **Session 5:** ***Model Selection***
Chair: Malay Ghosh, University of Florida

Speakers: Yuhong Yang, University of Minnesota
On Improving Traditional Model Selection Methods

Marina Vannucci, Rice University
Bayesian Methods for Variable Selection

1:00-2:3 0pm **Lunch** - free time

2:30-4:30pm **Session 6:** ***Model Uncertainty***
Chair: Arthur Berg, University of Florida

Speakers: Adrian Raftery, University of Washington
Online Prediction Under Model Uncertainty via Dynamic Model Averaging

David Draper, University of California, Santa Cruz
Bayesian model specification: What price model uncertainty?

5:00-8:30pm. **Pranzo Italiano:**
Hosted by Anne and George Casella
2245 NW 24th Ave.
Gainesville, FL 32605

Invited Talks

Objective Bayesian Variable Selection for Binomial Regression Models with Jeffreys's Prior

Ming-Hui Chen, University of Connecticut

We study several theoretical properties of Jeffreys's prior for binomial regression models with a focus on its applications to variable selection problems. We show that Jeffreys's prior is symmetric and unimodal about 0 and always has lighter tails than a ***t*** distribution and heavier tails than a normal distribution for a class of binomial regression models. We also develop an efficient importance sampling algorithm for calculating the prior and posterior normalizing constants based on Jeffreys's prior. Moreover, we show that the prior and posterior normalizing constants under Jeffreys's prior are scale invariant in the covariates. A closed form for Jeffreys's prior is obtained for saturated binomial regression models with binary covariates. Detailed simulation studies are presented to demonstrate its properties and performance in variable selection contexts and a real dataset is also analyzed to further illustrate the proposed methodology. This is a joint work with Joseph G. Ibrahim and Sungduk Kim.

Towards Objective Priors and Nonparametric Regression and Classification

Merlise Clyde, Duke University

In the univariate normal means hypothesis testing problem, Jeffreys recommended a Cauchy prior distribution to ensure consistency of Bayes factors under several situations. In the multiple regression setting, Zellner and Siow suggested multivariate Cauchy prior distributions obtained as a scale mixture of normal g-priors. Alternatively, independent Cauchy prior distributions are attractive, particularly in nonparametric regression problems where the number of potential predictors may greatly exceed the number of observations. In this talk we discuss the role of symmetric alpha-stable, in particular the Cauchy, process priors as a means to specifying prior distributions on infinite dimensional function spaces in nonparametric regression and classification problems. We show how the alpha-stable process priors may be represented as the limit of independent scale mixtures of normal priors and provide a generalization of the improper priors used in Tipping's Relevance Vector Machines in the framework of kernel regression. We discuss feature selection (variable selection) in the context of multivariate kernels. Finally, we present simulated and real data to illustrate the model performance.

Bayesian model specification: What price model uncertainty?

David Draper, University of California, Santa Cruz

(1) I will argue that, in problems of realistic complexity, for Bayesians to achieve the twin goals of coherence and calibration it's necessary to pay a price for model uncertainty;

(2) I will suggest two ways of paying this price:

(a) Bayesian nonparametric modeling, in which the price is built in automatically, by (in effect) adopting weaker prior assumptions on the space S of all possible models than those implicit in parametric modeling, or

(b) a Bayesian version of cross-validation, in which a portion of the data is set aside solely for assessing the calibration of the overall modeling process (which will typically involve a data-driven search through S); and

(3) I will present results on how much of the data needs to be set aside in strategy (b) to make it equivalent to strategy (a), and argue that this is a good way to quantify the price of model uncertainty.

On the Intrinsic Dimensionality of Multi-View Regression

Dean Foster, University of Pennsylvania

In the multi-view regression problem, we have a regression problem where the input variable can be partitioned into two different views, where it is assumed that either view of the example would be sufficient for learning — this is essentially the co-training assumption for the regression problem. For example, the task might be to identify a person, and the two views might be a video stream of the person and an audio stream of the person.

We show how Canonical Correlation Analysis, CCA, (related to PCA for two random variables) implies a ridge regression algorithm, where we can characterize the intrinsic dimensionality of this regression problem by the correlation of the two views. An interesting aspect of our analysis is that the norm used by the ridge regression algorithm is derived from the CCA — no norm or Hilbert space is assumed apriori (unlike in kernel methods).

Pre-Modeling Via BART

Edward I. George, University of Pennsylvania

Consider the canonical regression setup where one wants to learn about the relationship between ***y,*** a variable of interest, and x_b . . ., ***x_p, p*** potential predictor variables. Although one may ultimately want to build a parametric model to describe and summarize this relationship, preliminary analysis via flexible nonparametric models may provide useful guidance. For this purpose, we propose BART (Bayesian Additive Regression Trees), a flexible nonparametric ensemble Bayes approach for estimating *f(pc*i, . . ., ***x_p) =*** *E(Y* | xi, . . ., ***x_p),*** for obtaining predictive regions for future***y,*** for describing the marginal effects of subsets of xi, . . ., x_p and for model-free variable selection. Essentially, BART approximates *f* by a Bayesian "sum-of-trees" model where fitting and inference are accomplished via an iterative backfitting MCMC algorithm. By using a large number of trees, which yields a redundant basis for***f*** BART is seen to be remarkably effective at finding highly nonlinear relationships hidden within a large number of irrelevant potential predictors. BART also provides an omnibus test: the absence of any relationship between y and any subset of xi, . . ., x_p is indicated when BART posterior intervals for *f* reveal no signal. (This is joint work with Hugh Chipman and Robert McCulloch).

Hierarchical Nonparametric Bayes

Michael Jordan, University of California, Berkeley

Much statistical inference is concerned with controlling some form of tradeoff between flexibility and variability. In Bayesian inference, such control is often exerted via hierarchies—stochastic relationships among prior distributions. Nonparametric Bayesian statisticians work with priors that are general stochastic processes (e.g., distributions on spaces of continuous functions, spaces of monotone functions, or general measure spaces). Thus flexibility is emphasized and it is of particular importance to exert hierarchical control. In this talk I discuss Bayesian hierarchical modeling in the setting of two particularly interesting stochastic processes: the Dirichlet process and the beta process. These processes are discrete with probability one, and have interesting relationships to various random combinatorial objects. They yield models with open-ended numbers of "clusters" and models with open-ended numbers of "features," respectively. I discuss Bayesian modeling based on hierarchical Dirichlet process priors and hierarchical beta process priors, and present applications of these models to problems in bioinformatics and computational vision.

(Joint work with Yee Whye Teh and Romain Thibaux.)

High-Dimensional Bayesian Classifiers

David Madigan, Columbia University

Supervised learning applications in text categorization, authorship attribution, hospital profiling, and many other areas frequently involve training data with more predictors than examples. Regularized logistic models often prove useful in such applications and I will present some experimental results. A Bayesian interpretation of regularization offers advantages. In applications with small numbers of training examples, incorporation of external knowledge via informative priors proves highly effective. Sequential learning algorithms also emerge naturally in the Bayesian approach. Finally I will discuss some recent ideas concerning structured supervised learning problems and connections with social network models.

A Noninformative Bayesian Approach to Finite Population Sampling Using Auxiliary Variables

Glen Meeden, University of Minnesota

In finite population sampling prior information is often available in the form of partial knowledge about an auxiliary variable, for example its mean may be known. In such cases, the ratio estimator and the regression estimator are often used for estimating the population mean of the characteristic of interest. The Polya posterior has been developed as a non-informative Bayesian approach to survey sampling. It is appropriate when little or no prior information about the population is available. Here we show that it can be extended to incorporate types of partial prior information about auxiliary variables. We will see that it typically yields procedures with good frequentist properties even in some problems where standard frequentist methods are difficult to apply. Moreover one does not need to select a model which explicitly relates the characteristic of interest to the auxiliary variables.

Online Prediction Under Model Uncertainty via Dynamic Model Averaging

Adrian Raftery, University of Washington

We consider the problem of real-time prediction when it is uncertain what the best prediction model is. We develop a method called Dynamic Model Averaging (DMA) in which a state space model for the parameters of each model is combined with a Markov chain model for the correct model, allowing the correct model to vary over time. The state space and Markov chain models are both specified parsimoniously in terms of forgetting. The method is applied to predicting the output of a cold rolling mill, where the output is measured with a time delay. When only a small number of physically-based models were considered and one was clearly best, the method quickly converged to the best model, and the cost of model uncertainty was small. When model uncertainty and the number of models considered were large, our method ensured that the penalty for model uncertainty was small. This is joint work with Miroslav Karny, Josek Andrysek and Pavel Ettler.

Bayesian Methods for Variable Selection

Marina Vannucci, Rice University

In this talk I will address methods for Bayesian variable selection for high-dimensional data. I will start from the simple linear regression model and then extend methods to probit models for classification and to clustering settings. I will also consider models for survival data. I will show examples from genomics, in particular DNA microarray studies. The analysis of the high-dimensional data generated by such studies often challenges standard statistical methods. I will also assess performances on simulated data. Models and algorithms are quite flexible and allow us to incorporate additional information, such as data substructure and/or knowledge on gene functions.

On Improving Traditional Model Selection Methods

Yuhong Yang, University of Minnesota

In recent years, new methods, including those based on model combination, have been proposed to improve over traditional model selection methods from various perspectives. In this talk, we will share some thoughts/results on both positive and negative points of such approaches, contrasting model selection consistency, pointwise and uniform regression estimation. Results on adaptive and localized model selection/combination will be presented as well.

Poster Abstracts

A note on Bayes factors in semiparametric regression problems

Taeryon Choi, University of Maryland, Baltimore County

In this paper, we consider a Bayesian hypothesis testing problem when we have a parametric null model against a semiparametric alternative model. In particular, we investigate Bayes factors of a semiparametric regression model when the unknown regression function consists of parametric and nonparametric components. The nonparametric component is represented by an orthogonal basis where the coefficients are assumed to follow normal distributions with zero means and suitable variance structures. We identify the analytic form of Bayes factors and examine the asymptotic behavior of Bayes factors under the parametric null model as well as under the semiparametric alternative. Specifically, under appropriate conditions on the covariance structure of the regression function, we show that the Bayes factor is consistent, i.e. converges to infinity under the parametric null model while converges to zero almost surely under the semiparametric alternative as the sample size increases.

Monotonìe Regression via Variable Selection

S. McKay Curtis, North Carolina State University

In many areas of applied statistics, a researcher has substantive prior information that dictates a specific shape for a regression function but not a specific parametric form. Examples of these applications can be found in diverse areas such as economics, ecology, actuarial science, astronomy, and more. Methodological developments of shape-restricted inference start with the early works of Hildreth (1954) and Brunk (1955) and continue with the more recent developments of Mammen et al. (2001), Hall and Huang (2001), and Dette et al. (2006).

The variable selection problem consists of determining a method to select the best subset of predictors in a linear regression. Variable selection techniques also have a long history in the literature with early work by Gorman and Toman (1966) and recent work by Tibshirani (1996), Breiman (1996), George and Foster (2000).

In this paper, present a link between these two seemingly disparate areas of research. We begin with a monotonic regression model based on Bernstein polynomials. Under a simple reparametrization, we show that fitting this model is equivalent to the variable selection problem in linear models. We obtain model fits via the LASSO and demonstrate our method on several simulated data sets and the radio carbon data set analyzed by Hall and Huang (2001).

(Joint work with Sujit K. Ghosh.)

Incorporating Cost in Bayesian Variable Selection, with application to cost-effective measurement of quality of health care

D. Fouskakis, National Technical University of Athens, Greece

In the field of quality of health care measurement, patient sickness at admission is traditionally assessed by using logistic regression of mortality within 30 days of admission on a fairly large number of sickness indicators (perhaps on the order of 100) to construct a sickness scale, employing classical variable selection methods to find an "optimal" subset of 10-20 indicators. Such "benefit-only" methods ignore the considerable differences among the sickness indicators in cost of data collection, an issue that is crucial when admission sickness is used to drive programs that attempt to identify substandard hospitals by comparing observed and expected mortality rates (given admission sickness). When both data-collection cost and accuracy of prediction of 30-day mortality are considered, a large variable-selection problem arises in which costly variables that do not predict well enough should be omitted from the final scale.

We propose a prior setup which accounts for the cost of each variable and results in a set of posterior model probabilities which correspond to a generalized cost-modified version of BIC. We use reversible-jump Markov chain Monte Carlo (MCMC) methods to search the model space and check the stability of our findings with two variants of the MCMC model composition (MC^3) algorithm. Initially, we reduce our model space by dropping variables with low marginal posterior probabilities and we then estimate posterior model probabilities in the reduced space. Our cost-benefit approach results in a set of models with a noticeable reduction in cost and dimensionality, and only a minor decrease in predictive performance, when compared with models arising from the standard benefit-only analysis.

Additionally to the above, the practical relevance of the selected variable subsets is ensured, by enforcing an overall limit on the total data collection cost of each subset: the search is conducted only among models whose cost does not exceed this budgetary restriction. Trying to implement usual model search algorithms, will frequently fail if the actual best model is outside the imposed cost limit and when collinear predictors with high predictive ability are present. The reason for this failure is the existence of multiple modes with movement paths that are forbidden due to the cost restriction. Therefore, a population based trans-dimensional reversible-jump Markov chain Monte Carlo algorithm (population RJMCMC) is developed, where ideas from the population-based MCMC and simulated tempering algorithms are combined. Comparing the proposed technique with the simple RJMCMC we notice that population RJMCMC algorithm moves successfully and more efficiently between distinct neighborhoods of "good" models and achieves convergence faster. Our results are phrased in the language of health policy but apply with equal force to other quality assessment settings with dichotomous outcomes, such as the examination of drop-out rates in education, the study of retention rates in the workplace and the creation of cost-effective credit scores in business.

(Joint work with I. Ntzoufras and D. Draper)

Measuring Liquidity Costs in Agricultural Futures Markets: Conventional and Bayesian Approaches

Julieta Frank, University of Illinois

Estimation of liquidity costs in agricultural futures markets is challenging because bid-ask spreads are usually not observed. Spread estimators that use transaction data are available, but little agreement exists on their relative accuracy and performance. We evaluate four conventional and a recently proposed Bayesian estimators using simulated data based on Roll's standard liquidity cost model. The Bayesian estimator tracks Roll's model relatively well except when the level of noise in the market is large. We derive two improved Bayesian estimators that seem to have a higher performance even under high levels of noise which is common in agricultural futures markets. We also compute liquidity costs using data for hogs and cattle futures contracts trading on the Chicago Mercantile Exchange. The results obtained for market data are in line with the findings using simulated data.

(Joint work with Philip Garcia.)

Recentering Residuals In Bootstrapping Regression Models

Mihai C. Giurcanu, University of Louisiana at Lafayette

The purpose of this research is two-fold. First, we provide conditions for the consistency of the bootstrap distribution of OLS estimators in regression through the origin models. This problem has not been studied yet in the statistical literature, and an ad-hoc recentering procedure was usually employed. Second, we prove that under "regularity conditions", the bootstrap distribution of the two stage least squares (2SLS) estimators in instrumental variable (IV) regression models is weakly consistent without recentering residuals before resampling. As in the regression through the origin case, recentering is often employed, without verifying its necessity. These theoretical results give "rules" for when recentering residuals before resampling is necessary for these regression models.

Bayesian Variable Selection under Heredity Constraints

Woncheol Jang, University of Georgia

We propose a variable selection method for statistical models with high order interactions. The main challenge in this variable selection is to incorporate the effect heredity (Chipman, Hamda and Wu, 1997); higher order interaction can exist only if at least one of its parent effects exists. Using modern variable selection procedures such as LASSO may result in the violation of the effect heredity principle. We present a relatively simple Bayesian hierarchical variable selection procedure while still achieving the effect heredity principle. Examples in biomedical and engineering experiments are presented.

Bayesian Variable Selection Using Adaptive Powered Correlation Priors

Arun Krishna, North Carolina State University

The problem of selecting the correct subset of predictors within a linear model has received much attention in recent literature. Within the Bayesian framework, one of the popular choices among conjugate priors has been the Zellner's g-prior which is based on the inverse of empirical covariance matrix of the predictors. However Zellner's prior implicitly puts larger prior variance in the principal component directions with smaller eigenvalues, thus putting less information in the directions that are underdetermined by the data, particularly when the predictors are highly correlated. An extension of the Zeller's prior is proposed in this article which allow for a power parameter on the empirical covariance of the predictors. The power parameter helps control the degree to which correlated predictors are smoothed towards or away from one another. In addition, instead of using uniform prior on the model space, the empirical covariance of predictors is used to obtain suitable priors for model space consisting of all subsets of predictors variables. In this manner, the power parameter also helps to determine whether models containing highly collinear predictors are preferred or avoided. The proposed power parameter can be chosen via an empirical Bayes method which leads to a data adaptive choice of prior. Model selection is done via a stochastic search algorithm in cases where full enumeration of the models is not feasible. Simulation studies and real data examples have been presented to show how the power parameter is well determined from the degree of cross-correlation within predictors. The superior performance of the proposed method as compared to the standard use of Zellner's prior is also illustrated.

(Joint work with Sugit. K. Ghosh, Howard Bondell.)

Multiple Comparison Procedures for Long Memory Processes: Applications to US Stock Volatilities

Jaechoul Lee, Boise State University

Means of several United State stock price volatilities are compared to evaluate company's risk of stock investing. A fractionally integrated autoregressive moving average time series model is fitted to adequately take into the long memory present in the volatility in stock prices. As a simple and efficient method of mean comparisons, we modify typical uncorrelation-based multiple comparison procedures by adopting the equivalent sample size ideas. Performance of those proposed methods were examined by Monte Carlo simulations. A normal-inducing logarithmic transformation is employed to daily volatilities of several United State companies. High/low volatile companies were identified.

(Joint work with Kyungduk Ko, and Jason Arnold.)

Variable Selection in Multivariate Models with Block Structures

Dongyu Lin, University of Pennsylvania

Many variable selection algorithms include variable ranking as a principle because of its scalability and efficiency in data mining. Most of the ranking methods, however, are derived from an independence assumption of the observations, which is not necessarily true in reality, and may cause some misleading results by reason of the misspecified models. But sometimes observations can be grouped into blocks, each block of which is independent of the other blocks. We will look into two approaches to analyze such data: the sandwich covariance matrix estimator, which works well for linear models; and a method of Tukey, which is easier and possibly more general. Here we assume the design matrix is stochastic and blocks independently follow the same distribution, under which we provide extended interpretations of the two approaches and also the latter method is more robust in the sense that it is less sensitive to the block outliers.

A new Bayesian variable selection under the linear regression model

Yuzo Maruyama, University of Tokyo

In the normal linear regression model, the new Bayesian variable selection criterion is derived. Because Bayesian criteria are based on marginal density expressed by multiple integral, the numerical technique like MCMC or approximation of marginal density is extensively used. In this presentation, I will show that the special variant of Zellner's g-prior produces a simple closed form of marginal density from analytical calculation. Our criterion is not only consistent for model selection but also applicable for many regressors case ($n < p$).

(Joint work with Prof. Edward George, University of Pennsylvania.)

Bayesian estimation and testing in the normal mean problem

Marianna Pensky, University of Central Florida

We consider a problem of recovering a high dimensional vector $\boldsymbol{\mu}$ observed in white noise, where the unknown vector $\boldsymbol{\mu}$ is assumed to be sparse. The objective of the paper is to develop a Bayesian formalism which gives rise to a family of ***lo***-type penalties. The penalties are associated with various choices of the prior distributions $\boldsymbol{\pi_\eta}$ $(\cdot)$ on the number of nonzero entries of $\boldsymbol{\mu}$ and hence are easy to interpret. The resulting Bayesian estimators lead to a general thresholding rule which accommodates many of the known thresholding and model selection procedures as its particular cases corresponding to specific choices of $\boldsymbol{\pi_{\eta}}$ $(\cdot)$. Furthermore, they achieve optimality in a rather general setting under very mild conditions on the prior. We also specify the class of priors π_{η} $(\blacksquare)$ for which the resulting estimator is adaptive for a wide range of sparse sequences and consider several examples of such priors.

(Joint work with Felix Abramovich, Tel Aviv University and Vadim Grinshtein, The Open University of Israel)

Bayesian Kernel Regression and Classification

Zhi Ouyang, Duke University

We propose a general Bayesian framework for both nonparametric kernel regression and classification where the unknown mean function is represented as a weighted sum of kernel functions. We introduce alpha-stable Levy random fields to construct a prior on the unknown mean function, which lead to a specification of a joint prior distribution for the number of kernels, kernel regression coefficients, kernel centers, and kernel shape parameters. We show that the alpha-stable prior on the kernel regression coefficients may be approximated by t distributions, which is implemented by a Gamma prior distribution on the normal precision. A reversible-jump Markov chain Monte Carlo algorithm is developed to make posterior inference on the unknown mean function. In this algorithm, the regression coefficients are integrated out in calculating the likelihood, and the remaining coefficient precisions have much less correlation with the unknown mean function, which greatly improves the mixing of the Markov chain. For binary classification using a probit link, we augment the model with latent normal variables, hence the same method for Gaussian noise applies in the classification problem. We illustrate the approach on several simulated and real data sets.

(Joint work with Merlise A Clyde, Robert L Wolpert)

A Transformation-invariant Monotone Smoothing of Receiver Operating Characteristic Curves

Liansheng Tang, George Mason University

When a new diagnostic test is developed, it is of interest to evaluate its accuracy in distinguishing diseased subjects from non-diseased subjects. Receiver operating characteristic (ROC) curves serve as a popular evaluation tool. In this article we propose a monotone spline approach for estimating a single ROC curve. Unlike most of current ROC smoothing methods, our method ensures important inherent properties of underlying ROC curves which include monotonicity and transformation invariance. We compared the finite sample performance of the newly proposed ROC method with other ROC smoothing methods in large-scale simulation studies. We illustrated our method through two real life examples.

Bayesian Synthesis

Qingzhao Yu, Louisiana State University

Bayesian model averaging enables one to combine the disparate predictions of a number of models in a coherent fashion, leading to superior predictive performance. The improvement in performance arises from averaging models that make different predictions. In this work, we tap into perhaps the biggest driver of different predictions — different analysts — in order to gain the full benefits of model averaging. In a standard implementation of our method, several data analysts work independently on portions of a data set, eliciting separate models which are eventually updated and combined through Bayesian synthesis. The methodology helps to alleviate concerns about the sizeable gap between the foundational underpinnings of the Bayesian paradigm and the practice of Bayesian statistics.

We provide theoretical results that characterize general conditions under which data-splitting results in improved estimation which, in turn, carries over to improved prediction. These results suggest general principles of good modeling practice. In experimental work we show that the method has predictive performance superior to that of many automatic modeling techniques, including AIC, BIC, Smoothing Splines, CART, Bagged CART, Bayes CART, BMA, BART and LARS. Compared to competing modeling methods, the data-splitting approach 1) exhibits superior predictive performance for real data sets and simulations; 2) makes more efficient use of human knowledge; 3) selects sparser models with better explanatory ability and 4) avoids multiple uses of the data in the Bayesian framework.

First Name	Last Name	Organization
Alan	Agresti	University of Florida
Ayad	Ali	University of Florida
Howard	Bonded	NC State University
Jim	Booth	Cornell University
Eugenia	Buta	University of Florida
Ming-Hui	Chen	University of Connecticut
Jing	Cheng	University of Florida
Taeryon	Choi	University of Maryland
Mary	Christman	University of Florida
Merli se	Clyde	Duke University
Robyn	Crawford	University of Florida
Steven	Curtis	North Carolina State University
Shibasish	Dasgupta	University of Florida
Justin	Davis	University of Central Florida
Robert	Dorazio	University of Florida
Hani	Doss	University of Florida
David	Draper	University of California
Dean	Foster	University of Pennsylvania
Dimitris	Fouskakis	University of Athens
Julieta	Frank	University of Illinois at Urbana-Champaign
Claudio	Fuentes	University of Florida
Cyndi	Garvan	University of Florida
Jeremy	Gaskins	University of Florida
Ed	George	University of Pennsylvania
Mihai	Giurcanu	University of Louisiana
Meixi	Guo	University of Florida
Trung	Ha	University of Florida
Jim	Hobert	University of Florida
Woncheol	Jang	University of Georgia
Yongsung	Joo	University of Florida
Michael	Jordan	University of California
Goeran	Kauermann	University Bielefeld
Arun	Krishna	North Carolina State University
Paul	Kubilis	University of Florida
Jaechoul	Lee	Boise State University
Keunbaik	Lee	LSU-Health Science Center

First Name	Last Name	Organization
Zhen	Li	University of Florida
Dongyu	Lin	University of Pennsylvania
Xueli	Liu	University of Florida
Bo	Long	SUNY at Binghamton
Edgard	Maboudou - Tchao	University of Central Florida
David	Madigan	Columbia University
Yuzo	Maruyama	University of Tokyo
Glen	Meeden	University of Minnesota
Curtis	Miller	University of Florida
Nitai	Mukhopadhyay	Virginia Commonwealth University
Zhi	Ouyang	Duke University
Tezcan	Ozrazgat Basi anti	University of Florida
Trevor	Park	University of Florida
Marianna	Pensky	University of Central Florida
Allison	Pipkin	University of Florida
Adrian	Raferty	University of Washington
Jane	Ritho	College of Pharmacy
Vivekananda	Roy	University of Florida
Ananya	Roy	University of Nebraska-Lincoln
Carol	Rozear	University of Florida
Clyde	Schoolfield	University of Florida
Jihyun	Song	University of Florida
Doug	Sparks	University of Florida
Rebecca	Steorts	University of Florida
Lianshen	Tang	George Mason University
Marina	Vannucci	Rice University
Yanpin	Wang	University of Florida
Xiaoyin	Wang	Towson University
Song	Wu	University of Florida
Mark	Yang	University of Florida
Jie	Yang	University of Florida
Yuhong	Yang	University of Minnesota
John	Yap	University of Florida
Linda	Young	University of Florida
Qingzhao	Yu	Louisiana State University
Martine	Zandjanakou	University of Togo

Map

George and Anne Casella
2245 NW 24th Ave.
Gainesville, FL 32605