Mathematical Foundations of Computing
©American Institute of Mathematical Sciences
Volume **3**, Number **4**, November **2020**
doi:10.3934/mfc.2020026
pp. **i–ii**

PREFACE OF THE SPECIAL ISSUE ON ANALYSIS IN DATA SCIENCE: METHODS AND APPLICATIONS

Data analytics is playing vital roles in many branches of physical and social sciences. In the past two decades, scientific research has been employing more and more data-driven methodologies. The rapid advances of computing hardware and the mathematical analysis of data science have been providing a solid foundation for a broad scope of successful applications, which in turn have been stimulating and inspiring the development of applied and computational mathematics. This special issue is devoted to the discussion of some recent developments and cutting edge topics on the methods and applications of data science and analytics.

The special issue opens with the paper "Sketch-based Image Retrieval via CAT Loss with Elastic Net Regularization", where the authors Jia Cai, Guanglong Xu, and Zhensheng Hu study an interesting problem of pairing a sketch (which is a fast drawing with lines lacking details) of an object with its real picture. To solve the problem, the authors adopt the Sketch-a-Net CNN with soft attention module to construct the feature map, where five convolutional layers with ReLU unit, together with dropout and max-pooling are used. A combination of the absolute distance loss and the triplet loss is used on the feature space with the elastic net regularization to train the feature map. Comprehensive experiments are provided to show that the algorithm proposed is competitive with the state-of-the-art methods in the literature.

The next paper "Inpainting via Sparse Recovery with Directional Constraints" by Xuemei Chen and Julia Dobrosotskaya, considers the image inpainting problem. The images are modeled by matrices and the painted parts are modeled by missing values to recover. A new framework of regularization is introduced with not only the total variation penalty, but also the adaptive directional constraints which are constructed with filters chosen based on the properties of the domain with missing values. The implementation of the proposed algorithm is provided in detail with the alternating direction method of multiplier (ADMM). Sufficient conditions are derived for the uniqueness of the solution of the proposed minimization scheme.

The third paper by Qiang Fu, Yanlong Zhang, Yushu Zhu, and Ting Li on "Network Centralities, Demographic Disparities, and Voluntary Participation" considers the data set of a nationally representative random-digit-dial (RDD) survey of individuals aged 22-65 in the United States. From the perspective of network centrality, the authors explore the demographic features of the civic networks with a focus on race and gender. Several important findings include the greater network centrality of civic networks among women and whites than those among men and non-whites, the importance of religious organizations in civic networks compared to ethnic/civil-rights organizations and labor unions, and so on.

The fourth paper "Modeling Interactive Components by Coordinate Kernel Polynomial Models" by Xin Guo, Lexin Li, and Qiang Wu, studies a coordinate kernel polynomial-based scheme not only for learning the regression function but also for variable selection and interactive component identification. The model assigns a

reproducing kernel for each dimension of the input space, summing them up with adjustable weights and raise the sum to a power to be the reproducing kernel for regression learning. Variable selection is achieved by putting ℓ_1 penalty on the weights for kernel. Interactive component identification is implemented by expanding the power of the kernel and studying the corresponding product of weights. Generalization error bounds are derived with the help of Rademacher chaos complexity.

The paper by Ying Lin and Qi Ye on "Support Vector Machine Classifiers by Non-Euclidean Margins" studies support vector machine (SVM) with non-Euclidean margins. The motivations include better exploiting more geometrical structures, sparsity, and the application of the well known reproducing kernel Banach space to SVM. The equivalence between soft margin and hard margin SVM in Euclidean space is generalized to SVM with non-Euclidean margins. The application of the reproducing kernel Banach space to SVM is studied. The existence and uniqueness of the models studied are derived. Solid experimental studies are provided with both artificial and real data.

The final paper of the issue, "AIMS: Average Information Matrix Splitting" by Shengxin Zhu, Tongxiang Gu, and Xingping Liu, considers the linear mixed models that are widely used in statistics and social sciences. The authors prove that the average of the Fisher information matrix and the observed information matrix can be decomposed into two terms, where the first term $\mathcal{I}_A$ enjoys a light computational complexity, and the second term is a random matrix with mean zero. Therefore, it makes sense to approximate the Fisher information matrix by the observed information matrix and $\mathcal{I}_A$.

We would like to use this opportunity to thank the authors for submitting their latest work for publication in this special issue, and thank the anonymous referees for their valuable review comments.

Guest Editors:
Xin Guo
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hong Kong, China, and
School of Mathematics and Physics
The University of Queensland
Brisbane, QLD 4072, Australia
xin.guo@uq.edu.au
Lei Shi
School of Mathematical Sciences
Fudan University
Shanghai, China
leishi@fudan.edu.cn