Raghvendra Mall

Data Scientist

My objective is to obtain a deep understanding of sparsity in large scale machine learning and utilize model based kernel methods for big data learning. I aim to design advanced data-driven models for business analytics and decision support.

Biography

Nationality: Indian
Date of Birth: 05.03.1988
Sex: Male

Contact

✉ Kasteelpark Arenberg 10
Heverlee (Leuven) B-3001
Belgium
+32 484 287856
rmall@esat.kuleuven.be

Programming Skills

C\C++ ▲▲▲▲△
Python ▲▲▲△
C# ▲▲▲△
R ▲▲ △ △△
Java ▲▲ △ △△
Julia ▲ △ △ △ △

Tools & Platforms

Matlab ▲▲▲▲△
Weka ▲▲▲▲△
Gephi ▲▲▲ △ △
MySQL ▲▲▲ △ △
Hadoop ▲ △ △ △ △
Cvx ▲▲ △ △△
Lpsolve ▲▲ △ △△
SPSS ▲ △ △ △ △

Personal Skills

Communication ▲▲▲▲△
Learning Curve ▲▲▲▲▲
Problem Solving ▲▲▲▲△
Team Player ▲▲▲▲△
Quality Conscious ▲▲▲▲▲

Education

2012–2015 **Doctorate** Machine Learning
KU Leuven (Belgium)
Title: Sparsity in Large Scale Machine Learning using Kernel Methods
Advisor: Professor Johan Suykens (johan.suykens@esat.kuleuven.be)
Explored the role of sparsity in kernel methods for big data learning. Worked on classical machine learning problems like classification, regression and feature selection devising scalable sparse models using least squares support vector machines. Developed 2 sampling techniques for big data in order to obtain representative subset from data. Proposed several scalable flat and hierarchical clustering methodologies for complex networks, images and datasets using a primal-dual optimization framework (kernel spectral clustering). These algorithms can efficiently scale to 10^6-10^7 data points on a modern laptop.

2006–2011 **MS by Research + BTech** Computer Science (Dual Degree)
IIIT-Hyderabad, India C.G.P.A. 9.15/10
Title: PERturbed Frequent Itemset based Classification Techniques
Advisor: Prof. Vikram Pudi (vikram@iiit.ac.in)
Proposed novel perturbed frequent itemset based classification techniques using the Apriori framework. Proposed an effective methodology to classify real-attributed market-basket data.

Courses

- Phd: Introduction to R, Advanced R-Programming in R and Beyond, Graphics in R, Support Vector Machines-Methods and Applications, Machine Learning (Stanford University), Practical Machine Learning, Exploratory Data Analysis and Computing for Data Analysis (Johns Hopkins University), Social Network Analysis (University of Michigan), Networked Life (University of Pennsylvania) and Experiment for Improvement (McMaster University).
- MS by Research + BTech: Pattern Recognition, Artificial Intelligence, Data Warehousing and Data Mining, Web Data and Knowledge Management, Computer Programming, Algorithms, Data Structures, Theory of Computation, Linear Algebra, Numerical Analysis and Principles of Programming Languages.

Work Experience

2013–2013 **Research Associate**

Qatar Computing Research Institute (Doha), Qatar

Advisor: Prof. Halima Bensmail (hbensmail@qf.org.qa)

Contributions include:

- Developed primal-dual framework for feature extraction using LS-SVM.
- Applied the same for gene classification datasets.
- Published a paper in a peer-reviewed international conference.
- Extended sparse reductions of kernel spectral clustering to omics datasets.

2011–2012 **Research Intern**

Microsoft R & D Private Limited (Bangalore), India

Contribution include:

- Developed hierarchical clustering techniques to identify micro markets.
- Applied the same on large scale keyword-advertiser graphs.
- Devised methods to optimize the revenue for Bing search engine by efficiently ordering the mainline ads.

2010–2011 **Research Intern**

TALARIS Team, INRIA (LORIA), Nancy, France

Advisor:Prof. Jean-Charles Lamirel (jean-charles.lamirel@loria.fr)

Contribution include:

- Proposed several incremental growing neural gas clustering techniques.
- Applied the same for biological and scientometric text datasets.
- Task was to identify & track evolution of macro & micro topics.
- Used information retrieval quality metrics like macro, micro and cumulative precision and recall.
- Published multiple peer-reviewed international conference papers.

Top Publications

1. Mall R., Langone R., Suykens J.A.K., "Multilevel Hierarchical Kernel Spectral Clustering for Real-Life Large Scale Complex Networks", PLOS One, e99966, vol. 9, no. 6, Jun. 2014.
2. Mall R., Suykens J.A.K., "Very Sparse LSSVM Reductions for Large Scale Data", IEEE Transactions on Neural Networks and Learning Systems, in press.
3. Mall R., Langone R., Suykens J.A.K., "Kernel Spectral Clustering for Big Data Networks", Entropy, Special Issue: Big Data, vol. 15, no. 5, May 2013, pp. 1567-1586.
4. Mall R., Jumutc V., Langone R., Suykens J.A.K., "Representative Subsets For Big Data Learning using k-NN graphs", In Proc. of IEEE International Conference on Big Data (IEEE BigData 2014), Washington D.C., U.S.A., 2014, pp. 37-42.
5. Mall R., Langone R., Suykens J.A.K., "Self-Tuned Kernel Spectral Clustering for Large Scale Networks", in Proc. of the IEEE International Conference on Big Data (IEEE BigData 2013), Santa Clara, United States of America, Oct. 2013, pp. 385-393.
6. Full list of publications available at: `http://web.iiit.ac.in/~raghvendramall`

Professional Interests

Machine Learning, Big Data, Data and Graph Clustering, Feature Extraction, Bioinformatics, Time-series analysis, Predictive Analytics, Mathematical modelling, Optimization, Complex Networks, Climate and Fault Detection.