Omkar Kulkarni

Zebrastraat - 39, Gent, Belgium - 9000
E-mail: kulkarnios@gmail.com
Phone: 0032 - 485061870

Skills:

Statistical Programming	: R, Python, SAS
Clustered	: Apache Spark
Bayesian Modelling	: JAGS, Quantile regression, Bayesian Linear regression
Predictive Modelling	: OLS and logistic regression, LDA, KNN, Cross validation, Bootstrap, Ridge and Lasso, PCR and PLS regressions. Classification techniques.
Non Linear Modelling	: GAMs, Decision Trees, Random Forest, SVMs

Education

Advanced Masters in Statistical Data Analysis

Sep 2015 – Jan 2017

UGent, Ghent, Belgium

Thesis - ongoing: We (with an IoT company) plan to extract information from data stream in real time using machine learning algorithms. Tools in use shall be Python, Spark, Spark Streaming and R. For Sensor fusion Kalman filter is designed and Gauss-Newton method is used to minimize the cost function.

Courses completed:

- Principles of Statistical Data Analysis
- Analysis of Continuous data
- Analysis of Categorical Data
- Statistical Computing (R, Shiny and SAS)
- Statistical Inference
- Bayesian Statistics
- Analysis of High Dimensional Data (MDS, penalized regression, Large Scale Hypothesis Testing)
- Data mining and Big Data Science
- Big Data Science (Python, Pandas, Spark, D3)

Extra Courses: Algorithms and Programming (Python)

Bachelor of Engineering in Electronics and Communication

2006 – 2010

Visvesvaraya Technological University (VTU), Belgaum, India

Academic Projects:

Data Mining : Predicted mathematics scores of Portuguese students with 30+ predictors. During the process compared various prediction techniques such as OLS regression, Principal component regression, Ridge and Lasso techniques, Decision trees and Random Forests etc. Identified the significant factors using Zero Inflated (Poisson) models.

Bayesian Analysis: Comparison of Bayesian T-test and ANOVA to frequentist t-test and ANOVA on dataset of Belgian Falcons. Compared performance of Bayesian quantile regression with OLS regression. Analyzed convergence for MCMC on Logistic & Probity regression. Used R and JAGS for it.

High Dimensional: The goal was to predict at what age do infants develop type 1 diabetes, we had information on examination of the infant gut microbiome. We had 777 observations and 2239 Operational Taxonomic Units. Used Multiple Correspondence analysis and PCR for data visualization and compared models made using Principal component regression, Lasso and Ridge regression all three based on cross validation setups to arrive at optimum model.

Experience

CloudLeaf Inc Belgium
Intern Oct 2016 – present

- Pursuing my master's thesis.

Ghent University Belgium
Teaching Assistant Nov 2016 – Dec 2016

- Create R scripts to explain predictive modelling for continuous data cases

Nihilent Technologies India
Test Analyst Nov 2014–Jul 2015

- Created Business Process Maps for various processes in an Agile setup.
- Streamlined project analysis for Aero Card project along with requirement gathering.
- Part of testing team for MediSwitch – Universal Health Records project – South Africa.

CAREER PURSUIT India
Co-Founder Apl 2013–May 2015

- Co-Founded a start up with 2 others. We provided training services to educational institutions and corporate organizations.
- My major role was to come up with marketing strategy and increase our visibility in Pune.

INFOSYS India
Systems Engineer Sep 2010–Aug 2012

- Was part of product development unit of *Finacle* - Universal **Banking** Solution.
- Worked on Oracle for Unix-AIX and Windows-SQL Database platforms and application deployment on IBM-WAS (UNIX and Windows OS).
- Part of the team to deliver *Finacle* E-Banking product to banks like Equity Bank Kenya, RCBC Bank, RBCZ Bank, ICICI Bank.

VAYAVYA LABS LTD India
Intern Jan 2009–Mar 2009

- Worked on Bus Functional Modelling and implemented Wishbone and AMBA protocols in HDL-Verilog.

References available upon request