|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SUMMARY** | | | | | |
| * Statistician by training (BS, MS, PhD). 15+ years of experience of handling, cleaning data and creating data summaries * 7+ years of teaching and research * 2 years in health care industry analyzing large administrative health data * 5+ years of predictive modeling with machine learning algorithms (Lasso, CART, Gradient Boosting, glmnet) * Expert R programmer (15+ years) * Experienced SAS programmer (5+ years) * 2+ years of Python (NumPy, Pandas, Scikit-learn) * 3+ years of SQL (SAS, MySQL, Oracle) * 1+ year of Tableau, Microsoft Azure ML | | | | | |
| **SPEACILITIES** | | | | | |
| * Machine Learning and Predictive modeling * Oracle SQL * LASSO, adaptive LASSO, SCAD, and regularization methods for high dimensional data * Supervised and unsupervised learning for feature extraction and dimension reduction, K-Mean, ANN * Multivariate data analysis: Multivariate multiple regression, Logistic regression, PCA, Factor Analysis (FA), Discriminant Analysis (DA) and classification, and cluster analysis, multilevel and mixed modeling, GAM * Design and Analysis of experiments * Advanced programming, optimization, and Monte Carlo simulation with R * Data analysis with SAS; SAS/SQL, GLM, MIXED, GLIMMIX, RSREG, FACTEX and Multiple comparison procedures * Cloudera Hadoop and MapReduce plus familiarity with Spark, Pig, Hive (coursework on Big Data from Coursera)   **Current Research:**   * Improving the accuracy of predictive models via James-Stein-type shrinkage and LASSO-type regularization methods and some hybrid estimators connecting these two types of estimators.   **Teaching (2016):**   * Applied Multivariate Analysis (Graduate level. Using SAS/R) * Introduction to Statistical Computing with SAS (Grad level. SAS) * Introduction to SQL (MySQL and Oracle)   **Taught:**   * Applied Multivariate Analysis (Grad; mostly PhD students) * Multivariate Statistical Analysis (theoretical treatment, grad level) * Advanced Statistical Programming (Grad level: SAS, R) * Theory of Probability and Mathematical Statistics * Survival Analysis (K-M estimator, Censored data analysis, Cox proportional hazards Regression) * Computer Simulation and Modeling | | | | | |
| **EDUCATION / TECHNICAL COURSEWORK** | | | | | |
| **Machine Learning Engineer,** *Enrolled**in Nanodegree program at**Udacity* | | | | *Feb 2016 – Present* | |
| * Foundations on Artificial Intelligence, Machine learning, and Data Science * Applying Machine Learning algorithms such as Gradient Booting, Random Forest, and Ensemble techniques for building and predictive models | | | | | |
| **Introduction to Big Data,** *Course completed at Coursera* | | *Oct 12 – Nov 9, 2015* | | | |
| * Introduction to Hadoop, MapReduce, Hive, Pig, and Spark * Getting data in and out of HDFS, running basic MapReduce jobs with Cloudera Hadoop distribution | | | | | |
| **PhD in Statistics,** *University of Windsor, Windsor, Ontario, Canada* | | *Sep 2007 – Jun 2012* | | | |
| * Dissertation topic: Shrinkage and penalty estimation with linear and partially linear regression models * Published three peer reviewed articles in statistical and computational journals | | | | | |
| **EMPLOYMENT** | | | | | |
| **Assistant Professor of Statistics,** *University of Northern Colorado, Colorado, USA* | | | *Aug 2013 – Present* | | |
| * Cutting edge research in predictive modeling with tools such as Lasso, Elastic Net, and Regression regularization * Taught programming with R, and SAS (basic and advanced programming, R functions, SAS macro, SAS/SQL, ODS) * Taught multivariate analysis (Logistic regression, Segmentation, Cluster Analysis, and dimensionality reduction with PCA, SVD, and Factor Analysis). Sample work: <https://github.com/raheems/computer-programming> | | | | | |
| **Assistant Professor of Statistics,** *University of Wisconsin-Green Bay, WI, USA* | | | *Aug 2012 – July 2013* | | |
| * Taught undergraduate courses on Mathematical Statistics, Intro to Statistics * Member of admissions committee in the Environmental Science and Policy graduate program * Served on masters level thesis and project committees and provided statistical consultation and design of experiments and analysis of data * Published research on “Shrinkage and absolute penalty estimation in linear regression models”  <http://dx.doi.org/10.1002/wics.1232> | | | | | |
| **Data Analyst,** *Windsor-Essex County Health Unit, Windsor, Ontario, Canada* | | | *Jan 2011 – Aug 2012* | | |
| * Created technical report while working with epidemiologists * Managed and analyzed large administrative health datasets and created data summaries for stakeholders * Prepared internal reports by analyzing complex survey data (Canadian Community Health Survey, CCHS) * Organized and participated meetings with stakeholders and regional partners   Supervised and monitored performance of a research assistant   * Sample work on data analysis and programming with Stata software: <https://github.com/raheems/DataAnalysis/tree/master/Stata/HealthReport> | | | | | |
| **Statistical Consultant and PhD Student,** *University of Windsor, Ontario, Canada* | | | *Sep 2007 – Jun 2012* | | |
| * Completed coursework on Bayesian statistics, Experimental Design, Statistical Inference * Wrote an R package on shrinkage estimation in linear models. See codes here:  <https://github.com/raheems/r-packages> * As a consultant at the Center for Statistical Consulting, University of Windsor, provided consultations on the following projects:   + ***Tilbury District Family Health Team, Ontario, Canada*** Designed code-plan and prepared data entry form Cleaned and analyzed data using SAS Produced summary results through cross tables and graphics   + ***Canadian National Survey on Child Restraint Us****e* Designed efficient data entry forms, Suggested analysis plan   + **Hospital Patients’ Falls data, Windsor, ON, Canada** Cleaned and prepared raw data for analysis, Suggested analysis plans. * Conducted methodological research and published peer reviewed articles in computational journals (see below for a list of publications) | | | | | |
| **Research Assistant,** *University of Waterloo, Waterloo, Ontario Canada* | | | *Sep 2005 - Aug 2007* | | |
| * Worked with a professor to perform literature review on survival analysis using the Gamma frailty model * Participated in collaborative research in statistical genetics * Worked in a team of researchers to manage and analyze statistical genetics data that participated in the Genetic Analysis Workshop 15. The title of our project was "Utilizing SNPs to feature the candidate genes: HLA-DRB1, CTL4, PADI4 and PTPN22 in three Rheumatoid Arthritis studies: CANADA, NARAC and UK, (jointly with Hania Wormald and others) at Genetic Analysis Workshop 15 (GAW15), November 11-15, 2006, Florida, USA" | | | | | |
| **MS Student of Statistics,** *McMaster University, Hamilton, Ontario, Canada* | | | *Sep 2003 – May 2005* | | |
| * Completed coursework on Mathematical statistics, Multivariate analysis, Statistical Inference * Performed large scale simulation on estimating parameters of a logistic regression model. * Title of project work: Inferential methods for bivariate logistic models  <https://dl.dropboxusercontent.com/u/29485908/permalink/macproject.pdf> | | | | | |
| **Statistical Consultant,** *International Center for Diarrheal Disease Research (ICDDR’B), Bangladesh* | | | | | *May – Aug 2005* |
| * At the Population Science division, provided consultation on two research projects involving fitting logistic regression model to survey data * Cleaned the data and prepared the data for analysis * Created data summaries through tables, and graphs * Fitted logistic regression to predict binary outcome variable | | | | | |
| **Lecturer of Applied Statistics,** *Institute of Statistical Research and Training (ISRT), University of Dhaka, Bangladesh.* | | | *Jan 2001 – Aug 2003* | | |
| * Taught statistics at undergraduate level * Taught computer programming with S-Plus, R, and FORTRAN * Performed collaborative research involving analysis of Health Survey Data, and publishing research findings in peer reviewed journals * Supervised two MSc students and guided them in their thesis works | | | | | |
| **TECHNICAL SKILLS** | | | | | |
| **Statistical Analyst (15+ years)**   * Linear and nonlinear models, LASSO, SCAD * Optimization * Classification and Regression Trees (**CART**) * Random Forest, and ensemble methods * Multivariate statistical analysis, ANOVA, MANOVA * Logistic regression * Dimensionality reduction with PCA, and SVD * Later factor identification with Factor Analysis (FA) * Segmentation with Discriminant Analysis (DA) and Cluster analysis * **Multilevel modeling**, mixed modeling * Generalized Additive Modeling (GAM) * **Design and Analysis of Experiments** (DOE): factorial, split-plot, repeated measures and crossover designs, response surface methods * Computer Aided Designs | **Statistical/Computer Programming (1-15 years)**   * Expert **R** programmer including package development (15+ years) * Extensive Monte Carlo simulation with R (15+) * **SAS**, SAS/SQL, GLM, MIXED, GLIMMIX, RSREG (5+) * **Python**: NumPy, pandas, IPython (working knowledge) (1+) * Familiar with **Tableau** for data visualization * 4+ years of Virtual Private Server management with **Debian**, CentOS, and Slackware Linux distro * Knowledge of **C# and JAVA with selenium** for software testing and software quality assurance * Knowledge of **SDLC** and Agile development process   **Relational Database (2-5 years)**   * Oracle SQL and Oracle SQL Developer (2+ years) * SAS/SQL (5+ years) | | | | |
| **SCHOLARLY PUBLICATIONS** | | | | | |
| 1. Sanku Dey and **Enayetur Raheem** (2016, accepted). A multilevel multinomial logistic regression model for identifying risk factors of anemia in children aged 6-59 months in northeastern states of India. <http://www.tandfonline.com/doi/abs/10.1080/23311835.2016.1159798> 2. S. Ejaz Ahmed, and **E. Raheem**, (2012). Shrinkage and absolute penalty estimation in linear models, *WIREs Computational Statistics*, 4(6):541-553. <http://dx.doi.org/10.1002/wics.1232> 3. **Enayetur Raheem**, S. Ejaz Ahmed, and Kjell A. Doksum (2012). Absolute penalty and shrinkage estimation in partially linear models, *Computational Statistics and Data Analysis*, 56(4):874-891 <http://dx.doi.org/10.1016/j.csda.2011.09> 4. **Enayetur Raheem**, and S. Ejaz Ahmed, (2011). Positive-shrinkage and pretest estimation in linear regression: A Monte Carlo study with applications. *Journal of the Iranian Statistical Society*. 10(2):267-289 <http://arxiv.org/abs/1109.2527> 5. S. Ejaz Ahmed, **Enayetur Raheem** and Shakhawat Hossain (2010). Absolute Penalty Estimator in International Encyclopedia of Statistical Science, Lovric, Miodrag (Ed.), Springer. 6. A. K. Md. Ehsanes Saleh and **Enayetur Raheem** (2015, Submitted). Improved LASSO. <http://arxiv.org/abs/1503.05160> 7. A. K. Md. Ehsanes Saleh and **Enayetur Raheem** (2015, Submitted). Penalty, Shrinkage, and Preliminary Test Estimators under Full Model Hypothesis. <http://arxiv.org/abs/1503.06910> 8. A.K.Md. E. Saleh, M. Arashi, and **E. Raheem** (In progress). Shrinkage estimation in sparse regression models. | | | | | |
| **SCHOLARLY PRESENTATIONS AND TALKS** | | | | | |
| 1. **Enayetur Raheem** and Niloofar Ramezani. Getting your photo “explored”on Flickr-A predictive model using photo meta-data. Presented at the Joint Statistical meetings (JSM), Aug 8-13, Seattle, WA, USA 2. **Enayetur Raheem** and Hasinur Rahaman Khan (2015). Identifying factors of anemia in children – a multilevel multinomial logistic regression model using SAS PROC GLIMMIX. Accepted for presentation at the Western Users of SAS Software Educational Forum and Conference, September 9-11, 2015, San Diego, California, USA 3. Expository talk on predictive modeling. University of Northern Colorado, February 11, 2015 4. Lasso and its cousins. University of Northern Colorado, January 28, 2015 5. A Unified Approach to Shrinkage Estimation in Linear Regression Models (With S. E. Ahmed), Poster presentation at the Join Statistical Meetings, Boston, MA, Aug 2-7, 2014 6. Invited Judge at the 7th Annual Probability and Statistics Day, Department of Mathematics and Statistics, University of Maryland, Baltimore County, April 26-27, 2013, Baltimore, MD, USA. 7. Erie St. Clair and South West Local Health Integration Networks’ Falls Prevention Best Practices Collaborative, London, Ontario, April 11, 2012. 8. Absolute Penalty and Shrinkage Estimation in Linear Models, Poster at the Joint Statistical Meetings, Miami Beach, Florida, Jul 30 - Aug 4, 2011. 9. Absolute Penalty and B-spline-based Shrinkage Estimation in Partially Linear Models, International Workshop on Perspectives on High-dimensional Data Analysis, Fields Institute, Toronto, Ontario, Canada, June 9-11, 2011 10. “shrink”: An R Package for Shrinkage Estimation in Linear Regression Models, colloquium at the Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario, Canada, November 18, 2010. 11. Shrinkage M-Estimation in Partially Linear Models, presented at the 38th Annual Meeting of the Statistical Society of Canada, May 23 to 26, 2010 at Universit´e Laval, Quebec City, Canada. 12. Shrinkage Versus Lasso in Partially Linear Models presented at 2009 Canadian Mathematical Society Annual Meeting, December 5-7, Windsor, ON, Canada. 13. Obstructive sleep apnea case study. Statistical Society of Canada (SSC), Annual Meeting, 2006, London, ON, Canada. 14. Attended Southern Ontario Statistical Graduate Students’ Seminar Days (SOSGSSD) held at McMaster University, Hamilton, ON, Canada. 15. On inferential methods for bivariate logistic distribution. Institute of Statistical Research and Training, University of Dhaka, Bangladesh, 2005 16. On tests of pseudorandom numbers. Institute of Statistical Research and Training, University of Dhaka, Bangladesh, 2005 17. Fifth International Symposium on Optimization and Statistics. December 28–30, Aligarh Muslim University, Aligarh, India, 2002 | | | | | |
| **PRODUCTIVITY SOFTWARE** | | | | | |
| * MS Office (Word, Excel, PowerPoint) | | | | | |