

OBJECTIVE

Seeking a position as an analyst utilizing knowledge of mathematics, statistics and technical analysis skills.

SUMMARY OF SKILLS

Strong fundamentals on Statistical analysis/model-building/data mining/experimental prediction
Practical data analytics and Visualization/Presentation Skills with big data experience
SAS, R, MYSQL, Hive, Pig, PYTHON, MATLAB, Linux and Microsoft Excel

WORK EXPERIENCE

DataTek Inc.

Statistical Analyst

Fairfax, VA

November 2012 - present

- Big Data Analytics for healthcare using Hive/Hadoop, Pig.
- Collaborate with engineering on creating and executing edit-check, data cleaning, data formatting and data enriching using SAS/SQL.
- Conduct and maintain exploratory analysis (regression analysis, generalized linear model, multivariable analysis, survival analysis and categorical analysis, etc.) as well as predictive and decision models using SAS/R.
- Generate data-driven reports, data visualization (Excel / SAS).
- Develop SAS applications (macros, programs, etc.) to support clinical trials from Phase I to III for new drug application, and ensure full compliance with FDA regulatory and guidelines.

Cross Pixel Inc.

Data analyst

New York, NY

August 2011 – November 2012

- Hive scripts to aggregate data into queryable hive tables and CSV segment file for deliverables.
- Syndicate reports to business owners and engineers, and visualize data for easy, insightful consumption by executives (Tableau).
- Categorize users based on scoring system and find the relations between different data sources.

EDUCATION AND PROJECTS

State University of New York at Buffalo, NY

Master of Science in Statistics

Jan 2012

- Project of pairwise-comparison between three multivariate logistic regression models to predict intracranial aneurysm rupture status based on independently significant flow and geometrical parameters with receiving operating curve (ROC) and area under the curve (AUC) (software used: R 2.14.0).

Master of Science in Mathematics

June 2010

- Project of simulating blood flow in human cardiovascular system using numerical methods, cooperated with Mathematics Department of Brown University. (software used: Python v2.7.2)
- Project of approximating Discontinuous Functions and incorporated Stochastic Parameters by adding classical filter. Results showed that Gibbs phenomena are efficiently reduced after small number steps. Computing time is greatly reduced and more accurate approximation could be gained. (software used: Matlab 7.8)

CERTIFICATIONS

SAS Base/Advanced/Business Programming for SAS 9 Certification