Intermediate Data Scientist / Statistician

CareerBuilder.com is looking for a highly experienced statistician / data scientist with a curious mind, skilled at improving product and process through statistics, data mining, and visualization. We are looking for someone to join our team to work on groundbreaking R&D projects, to leverage massive structured, unstructured, transactional and real-time data sets from a variety of sources and analyze customer usage patterns in order to make actionable recommendations using statistics, data mining, machine learning, business understanding and common sense.

If you are a statistics, machine learning or data mining whiz (a.k.a. data scientist) who is equally at home discussing insights with business owners, customers, researchers or developers, and is willing to roll up your sleeves and do the hard work – we want to hear from you! You will be working on CareerBuilder.com’s Client Health Index, an early warning system designed to attribute root cause analysis for product performance opportunities. Additionally, you will be spearheading the charge to take CareerBuilder.com’s data analysis efforts into a new frontier, as we expand upon the insights we’ve already learned. Your work will be observed by C-Levels, senior executives and other leaders looking to discover drivers of performance across CareerBuilder.com’s most important product lines.

Additional Responsibilities:

- Interact with internal and potentially external clients to understand their requirements for predictive analytics applications.
- Understand the business goals and initiatives of the project and combine business modeling skills with outstanding data analysis.
- Leverage predictive modeling techniques to generate reliable forecasts based on historical data sets from multiple sources, accounting for missing data imputation when necessary.
- Cooperate with the Business Intelligence team to design and execute replicable data acquisition and utilization processes. Acquiring, analyzing, cleaning and structuring data is an essential part of your daily responsibilities.
- Collaborate with IT, BI, Sales, Customer Operations and Business Development teams to provide integration and implementation of the algorithms you have developed in the production environment.
- Create new statistical and visualization methods for highlighting critical variables that effect product performance, and for demonstrating the effect of client changes over time.
- Incorporate Bayesian techniques to discover previously unknown relationships.
Intermediate Data Scientist / Statistician

In order to solve these challenges you should be able to leverage off-the-shelf or open-source technologies as well as in-house engineering, and feel comfortable with big data solutions, applications and infrastructure (Hadoop, MapReduce, Mahout, Hive/Pig with UDF):

Requirements:

- Advanced degree in Statistics or Math; Publications in high quality cited scientific journals and conferences / patents are a definite plus
- 3+ years of industrial data-mining / analytics experience including applied techniques in data mining, machine learning, or graph mining using R, SPSS, Python, Matlab, or Java.
- Strong statistical Software Experience: R, SAS JMP, Design Expert, Minitab, Log and StatExact, Simca P(PCA and PLS), Crystal Ball, @Risk, WinBUGS,
- Extensive experience with SQL, Hadoop, Hive, Pig, MapReduce: Write UDFs to analyze unstructured data
- Strong collaboration skills and an excellent ability to understand and formulate problems and clearly communicate insights as a result of your extensive technical and business background.
- Experience with BaysiaLab a plus
- Experience turning ideas into actionable insights. Able to persuade stakeholders and champion effective techniques through product development. We are looking for an eager learner and you should be able to communicate how you have demonstrated success at telling the story hidden within the data.