



**JUNIOR DATA ENGINEER
DEEP ROOT ANALYTICS
ARLINGTON, VA**

Company Overview

Deep Root is an audience-powered analytics company that: identifies a client's ideal audiences, analyzes how the audiences consume media for optimal advertising, integrates the audiences into platforms for linear, addressable and digital ad delivery, and measures how the audiences are being impacted by paid and earned media.

Our data and technology stack are reflective of our company mission to bring a data-driven approach to television advertising. We work with multiple datasets that qualify as "big data" following the "Four V's" criteria: they are Varied in their layouts, they are large, Voluminous datasets, they have Velocity in that they are frequently updated, and they have Veracity in that they represent the observed behavior of television viewers, advertisers and content. As such, our technology stack relies heavily on cloud services, data pipelines and data science libraries.

Job Description

Deep Root Analytics is looking for a Junior Data Engineer to work in our engineering department. We are seeking someone with a technical background who has experience in shuttling data from one place to another, preferably in a professional ETL environment. The areas of responsibility primarily cover building out new pipeline components and maintaining existing ones for a complex technology stack that spans a variety of languages and frameworks.

The engineering team is responsible for data health and quality in every step of the pipeline process, from initial ingestion to deployment and visualization. As a result, debugging can require a deep dive into several interfacing pieces of software, and on any given day a team member can expect to work on multiple components that perform very different functions. The ideal candidate must be comfortable moving between systems, and importantly, they must be curious and hard-working, as we expect them to tackle these problems using their own experience and independent research abilities.

Specifically, the Junior Data Engineer will be tasked with the following:

- Manage, modify, and maintain our proprietary software responsible for data storage and transformation of data from a wide variety of sources and delivery methods.
- Design and build new components that scale to efficiently ingest, normalize, and process data from a growing number of different sources.
- Run distributed computing jobs using Spark and Elastic MapReduce to prepare and transform terabytes of time-series and event data for modeling.
- Integrate external APIs into current products and utilize their data to streamline and add value to current offerings.
- Assist and service the Data Science department through tasks such as building datasets, reducing dimensionality, and training and testing models.



- Assist data analysts on product buildout and ad hoc analysis requests from the client service team by ensuring they have access to the data they need.

Qualifications & Skills

- Bachelor's degree in Mathematics, Computer Science, or related field.
- 2+ years of experience using Python 3 to leverage its strong data science libraries, including Scikit-Learn and Pandas.
- Strong in at least one other language; experience shell scripting, especially bash.
- Proficient with different flavors of SQL, especially PostgreSQL, including a knowledge of under-the-hood concepts like indexing and analysis of query plans.
- Experience extracting data from, and pushing data to, a variety of sources including relational and non-relational databases, RESTful APIs, flat files, FTP servers, and distributed file systems.
- Excellent communication skills, especially when explaining difficult technical concepts to people in non-technical roles.
- Strong analytical skills, especially when working with multiple large datasets.
- Experience with "XaaS" cloud services — we are an AWS shop but will consider candidates with similar experience on other cloud platforms.

Preferred Qualifications

- Proficiency in JavaScript used in both front-end (React) and back-end (node.js) contexts, especially for middleware APIs.
- Experience displaying data with Tableau, d3.js, or other data visualization tools.
- Experience working with decision trees, clustering algorithms, and other supervised and unsupervised machine learning techniques.
- A basic understanding of the mathematics underlying models and transformations used commonly in data science today.

Characteristics of a Successful Candidate

- We are looking for smart, driven, likable and curious people.
- You need to be highly adaptable to a rapidly growing and changing business environment.
- You are a quick learner and problem solver.
- You keep on top of latest developments in field.
- You have an ability and eagerness to constantly learn and teach others.
- You are an exceptional problem-solver and have a self-starter attitude.

Application:

- Please email a resume, sample work product, and three references to info@deeprootanalytics.com with the subject line: DATA ENGINEER. Please be prepared to explain and walk through code samples.
- Applicants should also be prepared to answer detailed technical questions and be comfortable whiteboarding problems in a language of their choosing.