Python / Big Data Developer Python / Big Data Developer Python / Big Data Developer - VALIC Group, AIG AIG Global Data Services: Currently working in AIG Global Data Services as Data Science- Python developer. Worked extensively on Big Data analytical models developed in Python. Worked on regression models, Random forest algorithm, Regular expressions andparallel processing. Worked on various Python data structures including list, dictionaries, comprehensions, data-frames, vectors. Experience of working in live production environment and Unix platforms. Capable of working efficiently in aggressive, time bound schedules and high stress environment. Excellent team player and strong individual ownership. Experience in building a Regression and Classification models using machine learning techniques. Implemented many codes in python to automate the intermediate process while building the models Active participate on kaggle Tata Consultancy Service: In Tata Consultancy Service worked as Assistant System Engineer-Traineefor one year Worked on Web Development frameworkSpring, Hibernet Work Experience Python / Big Data Developer VALIC Group, AIG - New York, NY April 2016 to Present Technologies: Python: os, sys, time, ConfigParser, subprocess, shlex Big Data Technologies: PySpark, Sqoop DB: Oracle Responsibilities: Involved in model development and optimization. Implemented Autosys for automatic frequency base run of model. Implemented wrapper in Python which handles entire execution flow of model. Used Pythonsub-process module to call the PySpark job, SFTP and Oracle store procedure. Developed a scoring calculation algorithmin PySpark on AIG's Hadoop cluster. Used Sqoop to fetch the data from Oracle database and also send it back. Implemented file parser to read the data from configuration files. Automate the whole process in a single run using shell scripting. Responsible in restructuring and finalizing architecture for the overall process. Logistic regression used for scoring probability of surrendering an account. Extensively used python list, dictionaries, comprehensions and generators for various data manipulations. Python /Big Data Developer AIG Internal Tool Building - Bangalore, Karnataka January 2016 to August 2016 Technologies: Python: os, sys, re, time, csv, numpy, math, subprocess Big Data Technology: PySpark Responsibilities: Worked extensively in the complete end to end development of the tasks involved right from the Ingestion till the end to get a summary

files. Implemented PySpark applicationto create schema file from the data set Perform profiling task on categorical features to identify the junk value, count null and empty values, pattern of different categories etc. Calculation of statistics on numeric features such as mean, median, percentile, skewness etc. using PySpark Automate the whole process in a single run using shell scripting. Implemented a code to generate report files from the Profiling task. Iterators and foreach used vastly to make code efficient. Used Python sub-process module to invoke Spark job via python wrapper. Developed Python text analytics using re (regular expressions) to find pattern and generate the schema file. Education M-TECH in Computer Science International Institute of Information Technology - Bangalore, Karnataka 2016 B-TECH in Electronics Gujarat Technological University 2012 Additional Information Technical Skills Python: Libraries used -os, csv, re, numpy, pandas, sklearn, collections, sys, PDFMiner, Scrapy, Beautifulsoup, subprocess, zipfile, ConfigParser, time Hadoop: Spark, Sqoop, Hive DB and Data Formats: Oracle PL/SQL, MySQL, SQL Server, SQLite, JSON Tools: Eclipse, Android Studio, Tortoise SVN, AncondaPython

Name: Anthony Stephens

Email: barbara68@example.net

Phone: 3527081320