Sr . Python Developer Sr .Python Developer Sr . Python Developer - Tiaa-cref ,Cary,NC Work Experience Sr . Python Developer Tiaa-cref - Cary, NC, US August 2017 to Present Responsibilities: Implemented Data Exploration to analyze patterns and to select features using Python SciPy. Built Factor Analysis and Cluster Analysis models using Python SciPy to classify customers into different target groups. Using R and Python withggplot2 packages performed an extensive graphical visualization of overall data, including customized graphical representation of revenue reports, specific item sales statistics and visualization. Evaluated business requirements and prepared detailed specifications that follow project guidelines required to develop written programs. Used Oozie workflow engine to run multiple Hive and Pig jobs. Participated in Data Acquisition with Data Engineer team to extract historical and real-time data by using python Library Pandas. Analyzed the partitioned and bucketed data and compute various metrics for reporting. Involved in loading data from RDBMS and web logs into SQL and NOSql Extracted data from Twitter using Java and Twitter API. Parsed JSON formatted twitter data and uploaded to database. Launching Amazon EC2 Cloud Instances using Amazon Images (Linux/ Ubuntu) and Configuring launched instances with respect to specific applications. Developed Hive queries for analysis, and exported the result set from Hive to MySQL using Sqoop after processing the data. Analyzed the data by performing Hive queries and running Pig scripts to study customer behavior. Created HBase tables to store various data formats of data coming from different portfolios. Worked on improving performance of existing Pig and Hive Queries. Created reports and dashboards, by using D3.js and Tableau 9.x, to explain and communicate data insights, significant features, models scores and performance of new recommendation system to both technical and business teams. Utilize SQL, Excel and several Marketing/Web Analytics tools (Google Analytics, AdWords) in order to complete business & marketing analysis and assessment. Used Git 2.x for version control with Data Engineer team and Data Scientists colleagues. Used Agile methodology and SCRUM process for project developing. Environment: Python 2.7, Django, HTML5/CSS, PostgreSQL, MS SQL Server 2013, MySQL, JavaScript, Jupyter Notebook, VIM, Pycharm, Shell Scripting, Angular.JS, JIRA. Python Developer Amneal - Long Island April 2015 to June 2017 Responsibilities

Perform Data Profiling to learn about behavior with various features such as traffic pattern, location, Date and Time etc. Extracted the data from hive tables by writing efficient Hive queries. Performed preliminary data analysis using descriptive statistics and handled anomalies such as removing duplicates and imputing missing values. Analyze Data and Performed Data Preparation by applying historical model on the data set in Tablue. Performed data cleaning and feature selection using MLlib package in PySpark and working with deep learning frameworks such as Caffe, Neon etc. Conducted a hybrid of Hierarchical and K-means Cluster Analysis using IBM SPSS and identified meaningful segments of customers through a discovery approach. Develop Spark/Scala, Python, R for regular expression (regex) project in the Hadoop/Hive environment with Linux/Windows for big data resources. Used clustering technique K-Means to identify outliers and to classify unlabeled data. Evaluate models using Cross Validation, Log loss function, ROC curves and used AUC for feature selection and elastic technologies like Elastic Search, Kibana etc. Work with NLTK library to NLP data processing and finding the patterns. Categorize comments into positive and negative clusters from different social networking sites using Sentiment Analysis and Text Analytics. Analyze traffic patterns by calculating autocorrelation with different time lags. Ensure that the model has low False Positive Rate and Text classification and sentiment analysis for unstructured and semi-structured data. Addressed over fitting by implementing of the algorithm regularization methods like L2 and L1. Use Principal Component Analysis in feature engineering to analyze high dimensional data. Create and design reports that will use gathered metrics to infer and draw logical conclusions of past and future behavior. Perform Multinomial Logistic Regression, Random forest, Decision Tree, SVM to classify package is going to deliver on time for the new route. Implemented different models like Logistic Regression, Random Forest and Gradient-Boost Trees to predict whether a given die will pass or fail the test. Perform data analysis by using Hive to retrieve the data from Sql to retrieve datafrom Oracle database and used ETL for data transformation. Perform Data Cleaning, features scaling, features engineering using pandas and numpy packages in python. Develop Map Reduce pipeline for feature extraction using Hive and Pig. Create Data Quality Scripts using SQL and Hive to validate successful data load and

quality of the data. Create various types of data visualizations using Python and Tableau. Communicate the results with operations team for taking best decisions.   Collect data needs and requirements by Interacting with the other departments.   Environment: Python 2.x, R, HDFS, Hadoop 2.3, Hive, Linux, Spark, IBM SPSS, Tableau Desktop, SQL Server 2012, Microsoft Excel, Matlab, Spark SQL, Pyspark. Python Developer Rite Aid - Richmond, VA, US September 2012 to March 2015 Responsibilities: Worked in comprehending and examining the client's business requirements.   Used Django frameworks and Python to build dynamic webpages.   Developed tools for monitoring and notification using Python.   Enhanced the application by using HTML and Java script for design and development.   Used data structures like directories, tuples, Object Oriented class based inheritance features for making complex algorithms of networks.   Created PHP/MySQL back-end for data entry from Flash and worked in tandem with the Flash developer to obtain the correct data through query string   Involved in designing database Model, API's, Views using python to build an interactive web based solution.   Generated Python Django Forms to record data of online users.   Implemented Data tables to add, delete, update and display patient records and policy information using PyQt.   Implemented a module to connect and view the status of an Apache Cassandra instance using python.   Developed MVC prototype replacement of current product with Django.   Improved the Data Security and generated report efficiently by caching and reusing data.   Created UI using JavaScript and HTML5/CSS3.   Managed datasets using Panda data frames and MYSQL. Queried the database queries using Python-MySQL connector and retrieved information using MySQLdb.   Recorded the online users' data using Python Django forms and implemented test case using Pytest.   Developed the application using the Test-driven methodology and designed the unit tests using Python Unit test framework.   Created web application prototype usingJQuery and Angular JS.   Deployed the project into Heroku using GIT version control system,   Maintained and Updated the application in accordance to the clientele's requirement   Environment: Python 3, Django 1.6, Tableau 8.2, Beautiful soup, HTML5, CSS/CSS3, Bootstrap, XML, JSON, JavaScript, JQuery, Angular JS, Backbone JS, Restful Web services, Apache spark, Linux, Git, Amazon s3, Jenkins, MySQL, Mongo DB, T-SQL, Eclipse. Education

Bachelor's Skills HTML5 (3 years), Java (3 years), Linux (6 years), Python (6 years), SQL (6 years)

Additional Information Technical Skills: Languages Java 8, Python, R Packages ggplot2, caret, dplyr, Rweka, gmodels, RCurl, C50, twitter, NLP, Reshape2, rjson, plyr, pandas, numPy, Seaborn, sciPy, matplot lib, sci-kit-learn, Beautiful Soup, Rpy2.   Web Technologies JDBC, HTML5, DHTML and XML, CSS3, Web Services, WSDL Data Modelling Tools Erwin r 9.6, 9.5, 9.1, 8.x, Rational Rose, ER/Studio, MS Visio, SAP Power designer Big Data Technologies Hadoop, Hive, HDFS, MapReduce, Pig, Kafka Databases SQL, Databases SQL-Server, My SQL, MS Access Reporting Tools   MS Office (Word/Excel/Power Point/ Visio), Tableau, Crystal reports XI, Business Intelligence, SSRS, Business Objects 5.x/ 6.x, Cognos7.0/6.0.   ETL Tools Informatica Power Centre, SSIS. Version Control Tools SVM, GitHub BI Tools Tableau, Tableau Server, Tableau Reader, Amazon Redshift Operating System Windows, Linux, Unix, Macintosh HD, Red Hat

Name: Chase Suarez

Email: awilson@example.org

Phone: 399-405-6055