

DATA SCIENTIST DATA SCIENTIST DATA SCIENTIST - CERNER Kansas City, KS Around 7+ years of experience in IT field with 3+ years as Data Scientist with strong technical and business experience, and communication skills to drive high-impact business outcomes through data-driven innovations and decisions . Expertise in Statistical analysis, Predictive modeling, Text mining, Supervised learning, Unsupervised Learning, and Reinforcement learning. Expert in developing NLP models for Topic extraction, Sentiment Analysis Strong mathematical background in Linear algebra, Probability, Statistics, Differentiation and Integration Having good experience in Amazon Web Services (AWS). Expertise in transforming business requirements into analytical models, designing algorithms, building models, developing data mining and reporting solutions that scale across a massive volume of Structured and unstructured data Proficient in Data Mining Methods, Factor Analysis, ANOVA, Hypothetical testing, normal distribution and other advanced Statistical modeling both Linear and nonlinear (LOGISTIC REGRESSION, LINEAR REGRESSION, NA VE BAYES, DECISION TREES, RANDOM FOREST, NEURAL NETWORKS, SVM, CLUSTERING, KNN) Experience with Deep learning techniques such as Convolutional Neural Networks, Recurrent Neural Networks by using Keras , TensorFlow and OpenCV Worked on several python packages like NumPy, Pandas, Scikit Learn, Matplotlib, Beautiful Soup, Pickle, SciPy, Python, PyTables etc. Proficient in implementing Dimensionality Reduction Techniques like Principal Component Analysis, t-Stochastics Neighborhood Embedding (t-SNE), and Linear Discriminant Analysis (LDA) Experience in implementing data analysis with various analytic tools, such as Anaconda 4.0, Jupyter Notebook 4.X, RStudio, RShiney. Expertise in all aspects of Software Development Lifecycle (SDLC) from requirement analysis, Design, Development Coding, Testing, Implementation, and Maintenance Hands on advanced SQL experience summarizing, transforming, segmenting, joining datasets. Experience in visualization tools like Tableau 9.X, 10.X for creating dashboards Experience in using GIT Version Control System Authorized to work in the US for any employer Work Experience DATA SCIENTIST CERNER - Kansas City, MO January 2018 to Present RESPONSIBILITIES: Participated in all phases of Machine Learning and Data Mining; data collection, data cleaning, developing models, validation, visualization Used Pandas,

NumPy, seaborn, scipy, matplotlib, scikit-learn, nltk in python for developing various machine learning algorithms. Implemented Bagging and Boosting to enhance the model performance. Leveraged disparate data sources that provide deep customer insight including online transactional data, web data, payment, orders history and marketing campaigns exposure data. Implemented the end-to-end platform for performing user behavior analytics using unsupervised machine learning. Implemented Classification using supervised algorithms like Logistic Regression, Decision trees, KNN, Naive Bayes. Data transformation from various resources, data organization, features extraction from raw and stored. Designed and implemented system architecture for AmazonEC2 based cloud -hosted solution for the client. Identified outliers and inconsistencies in data by conducting exploratory data analysis (EDA) using python NumPy and Seaborn to see the insights of data and validate each feature. Prepare scripts to ensure proper data access, manipulation and reporting functions with R programming. Involved in converting HIVE/SQL queries into Spark transformations using Spark RDDs, and Python Developed Ad-hoc Queries for moving data from HDFS to Hive and analyzing the data using HIVE QL Validated models using cross-validation and loss function to measure model performance. Created Confusion Matrix and ROC. Expert in performing Text Mining and Text classification in NLP by using TfidfVectorizer. Developed NLP models for Topic Extraction, Sentiment Analysis. Addressed overfitting by implementing of the algorithm regularization methods like L2 and L1. Performed price sensitivity and variation analysis across different marketing channels and conducted exploratory data analysis on variables such as lifetime value and profit score. Built data pipelines, implemented code modularization involving package creation and co-developed REST API's using Flask for production deployment. Performed data discovery and build a stream that automatically retrieves data from multitude of sources (SQL databases, external data such as social network data, user reviews) to generate KPI's using Tableau. ENVIRONMENT: Anaconda, Azure, Python, R Studio, R Shiny, Jupyter Notebook, VS code, Spyder,AWS, Oracle, SSMS, Unix, Tableau, HDFS, SPARK, IMPALA, HIVE, Hue. DATA SCIENTIST/DATA ANALYTICS SPECIALIST EQUIFAX, ST. LOUIS January 2016 to December 2017 RESPONSIBILITIES: Analyzed the data using various machine learning

algorithms whether to extend/not credit limit to an existing applicant and to approve/not new credit line to a new applicant will likely result in profit or loss based on various circumstances like credit history, utilization rate, income, age, location, hard enquiries & number of delinquencies. Extracted terabytes of structured and unstructured data by using SQL queries and performed data mining tasks including handling missing data, data wrangling, feature scaling, outlier analysis in python by importing pandas. Conducted data investigation, discovery & mapping tools to scan every single data record. Worked on predictive and what-if analysis using Python from HDFS and successfully loaded files to HDFS from Teradata and loaded from HDFS to HIVE. Performed data analysis, data validation, data cleansing, and data verification to identify data mismatch using Relational Data modeling (3NF) and Dimensional Data Modeling. Performed exploratory data analysis on all the features to understand feature importance and analyzed the behavior of features by using different statistical approaches. Studied the feature distribution with the help of Probability Density Function, Cumulative Distribution Function, Percentiles, Quantiles to draw some insights. Developed automated model training, testing & deployment via machine learning continuous delivery pipelines. Built decision tree model from the set of training data using the information entropy and the attribute with the highest normalized information gain is chosen to make the decision of credit approval. Created S3 buckets and managing policies for S3 buckets and Utilized S3 bucket and Glacier for Archival storage and backup on AWS. Experienced in designing and deploying AWS Solutions using EC2, S3, EBS, an Elastic Load Balancer (ELB) and auto scaling groups. Able to create scripts for system administration and AWS using languages such as BASH and Python. Used ML algorithms logistic regression, support vector machine, k nearest neighbors, Na ve Bayes, bagging, boosting, ensemble learning to analyze the data based on the features selected for data-driven decisions. Performed text analysis on the reviews of the products using NLP techniques like Bag of Words, Term Frequency-Inverse Document Frequency, Word2vec, Average Word2vec with help of NLTK, Beautiful soup libraries. Used machine learning algorithms to forecast the company's short-term and long-term growth in terms of revenue, number of customers, various costs, stock changes etcetera. Used Classified instances, Relative Operating

Characteristic curve (ROC) and Confusion Matrix to find the accuracy of the models built. Acquired knowledge on designing, iterating and fine-tuning neural network model's architecture for runtime efficiency to achieve optimal performance. Visualized results in python using Matplotlib, Seaborn libraries of Scikit-learn and used Tableau to create the interactive dashboards to present results for team members, management and clients ENVIRONMENT: Anaconda, Python, R Studio, Jupyter Notebook, VS Code, Spyder, PyCharm, SSMS, Unix, Tableau, Jira, HDFS, SPARK, IMPALA, HIVE, Hue. PYTHON DEVELOPER MAVIN SOLUTION - Hyderabad, Telangana July 2012 to August 2015 INDIA RESPONSIBILITIES: Involved in the design and development of different web-based applications based on client's requirements. Designed use case diagrams, class diagrams, sequence diagrams and state diagrams. Learned new technical skills as required for the system like Django, Flask Frameworks and Model-View-Controller (MVC) design pattern. Developed applications using Flask ( Python frameworks). Designed email marketing campaigns and created responsive web forms that saved data into a database using Python/ Django Framework. Developed Python scripts to read from Excel files, generate XML configuration files and for generating IP access frequency lists in different data logs. Deployed web applications to Google App Engine. Learnt to deploy projects using Jenkins. Utilized Pandas - Python library for analyzing data and data structures. Managed large datasets using Pandas data frames and SQLite. Performed front-end development for web initiatives to ensure usability, using HTML and CSS and enhanced quality, feel, and usability of consumer-facing website. Tested all completed work to ensure proper and error free functionality. Collaborated with a team of instructors and programmers to develop the curriculum and guidelines for workshops to teach the logic of programming. Created and ran custom SQL queries, stored procedures and created an application to store client phone calls and emails that were routed to various developers. Performed data profiling and analysis, applied various data cleansing rules, designed data standards, architecture and designed the relational models. Maintained metadata (data definitions of table structures) and version controlling for the data model. Environment: Python, Django, Flask, SQLite, SSMS, Google App Engine, Jenkins, Pandas, HTML, CSS PYTHON DEVELOPER

SUTHERLAND - Hyderabad, Telangana May 2011 to June 2012 INDIA RESPONSIBILITIES

Actively involved in interacting with front end users, project lead and business analyst to gather user requirements and online system specifications. Followed Agile Methodologies to manage full life-cycle development of the project. Designed and developed communication between client and server using Secured Web services. Written backend programming in Python and used the Django Framework to develop the application. Participated in entire lifecycle of the projects including Design, Development, and Deployment, Testing and Implementation and support. Implemented user interface guidelines and standards throughout the development and maintenance of the website using the HTML5, CSS3, JavaScript Developed views and templates with Python and Django's view controller and templating language to create a user-friendly website interface. Developed RESTful services using Django. Developed and tested many features for dashboard using Python, CSS, JavaScript. Used JavaScript and XML to update a portion of a webpage. Successfully migrated the Django database from SQLite3 to PostgreSQL with complete data integrity. Worked on Jenkins continuous integration tool for deployment of project. Created custom T-SQL procedures to read data from flat files to dump to SQL Server database using SQL Server import and export data wizard. Developed user defined functions based on the requirements and used various built-in functions. Handled errors using Exception handling (try, catch) extensively for the ease of debugging and displaying the error messages in the application. Developed batch scripts for scheduling data migration scripts. Created clustered, non-clustered indexes and indexed views to optimize the queries performance. Coordinated with onsite folks and mentored the offshore team Worked PL/SQL in Oracle database for writing queries, functions, stored procedures and triggers. ENVIRONMENT: Python, Django, JavaScript, HTML, CSS, XML, MYSQL, T-SQL, SSMS, MS- Excel, MS-word, T-SQL, Windows Server. TOOLS & TECHNOLOGIES Languages & Packages Python, SQL, TensorFlow, PySpark, Numpy, Pandas, Keras, NLTK, Caffe Languages & Packages Python, SQL, Numpy, Pandas, Scikit-Learn, Matplotlib TensorFlow, Keras, NLTK, Tableau, MySQL. Databases Hadoop Ecosystem SQL Server. HDFS, Map Reduce, HIVE, IMPALA, Spark (PySpark) and Kafka Mathematical Matrix

operations, Differentiation, Integration, Probability, Statistics, Linear Algebra, Geometry Machine Learning Algorithms Logistic Regression, Linear Regression, K Means Clustering Algorithm, Decision Trees, Support Vector Machines, Na ve Bayes, Hierarchical Clustering. Deep Learning Techniques Artificial Neural Networks, Convolutional Neural Networks, Multi-layer Perceptron's, Recurrent Neural Networks, LSTM, Back Propagation, Chain rule, Choosing Activation Functions, Drop Out, Optimization algorithms. User Interfaces HTML, CSS, Java Script, XML. Version control Tools Git Tools Spyder, Jupyter Notebook, VS Code, R Studio, SQL Server, Oracle, Anconda, R Shiney, Tableau, Plotly. Operating Systems Windows, Linux Methodologies Agile, Scrum Education Master's Skills Linux, Business Intelligence, SQL, Statistical analysis, NLP, AWS, Python, HIVE/SQL

Name: Jerry Aguirre

Email: dbaker@example.org

Phone: +1-539-244-8292x3704