

Python Developer Python Developer Python Developer - Infor, Inc New York, NY Over 10 years of IT industry experience encompassing in Machine Learning, Datamining with large datasets of Structured and Unstructured data, Data Acquisition, Data Validation, Predictive modeling, DataVisualization. Extensive experience in Text Analytics, developing different statistical machine learning, Datamining solutions to various business problems and generating data visualizations using R, Python, and Tableau. Experience object-oriented programming (OOP) concepts using Python, C++ and PHP. Experience on advanced SAS programming techniques, such as PROC SQL (JOIN/ UNION), PROC APPEND, PROC DATASETS, and PROC TRANSPOSE. Integration Architect & Data Scientist experience in Analytics, Bigdata, BPM, SOA, ETL and Cloud technologies. Highly skilled in using visualization tools like Tableau, ggplot2, and d3.js for creating dashboards. Tagging of experience in foundational machine learning models and concepts: regression, random forest, boosting, GBM, NNs, HMMs, CRFs, MRFs, deep learning. Proficiency in understanding statistical and other tools/languages - R, Python, C, C++, Java, SQL, UNIX, QlikView data visualization tool and Anaplan forecasting tool. Proficient in the Integration of various data sources with multiple relational databases like Oracle/, MS SQL Server, DB2, Teradata and Flat Files into the staging area, ODS, Data Warehouse and DataMart. Experience in Extracting data for creating Value Added Datasets using Python, R, SAS, Azure and SQL to analyze the behavior to target a specific set of customers to obtain hidden insights within the data to effectively implement the project Objectives. Worked with NoSQL Database including HBase, Cassandra and MongoDB. Extensively worked on statistical analysis tools and adept at writing code in Advanced Excel, R, MATLAB, Python. Implemented deep learning models and numerical Computation with the help of data flow graphs using Tensor Flow Machine Learning. Worked with complex applications such as R, Stata, Scala, Perl, Linear, SAS, and SPSS to develop a neural network, cluster analysis. Experienced the full software lifecycle in SDLC, Agile and Scrum methodologies. Experience in designing stunning visualizations using Tableau software and publishing and presenting dashboards, Storyline on web and desktop platforms. Hands on experience in implementing LDA, Naive Bayes and skilled in RandomForest, DecisionTrees, Linear

and LogisticRegression, SVM, Clustering, neural networks, Principle Component Analysis and good knowledge of Recommender Systems. Experienced with machine learning algorithms such as logistic regression, random forest, XP boost, KNN, SVM, neural network, linear regression, lasso regression and k-means. Strong experience in Software Development Life Cycle (SDLC) including Requirements Analysis, Design Specification and Testing as per Cycle in both Waterfall and Agile methodologies. Skilled in using dplyr and pandas in R and python for performing Exploratory data analysis. Experience working with data modeling tools like Erwin, Power Designer and ERStudio.

Experience with data analytics, data reporting, Ad-hoc reporting, Graphs, Scales, PivotTables and OLAP reporting. Highly skilled in using visualization tools like Tableau, ggplot2, and d3.js for creating dashboards. Worked and extracted data from various database sources like Oracle, SQL Server, DB2, and Teradata.

Proficient knowledge of statistics, mathematics, machine learning, recommendation algorithms and analytics with an excellent understanding of business operations and analytics tools for effective analysis of data. Authorized to work in the US for any employer

Work Experience Python Developer Infor, Inc - New York, NY July 2017 to Present Description: Myntra is an Indian fashion e-commerce marketplace company headquartered in Bengaluru, Karnataka, India. The focus of the company is the online retailing of branded apparel. The main goal of the project is to develop an application to generate the tickets for the orders of the customers.

Responsibilities: Worked on the project from gathering requirements to developing the entire application. Worked on AnacondaPython Environment. Created, activated and programmed in Anaconda environment. Wrote programs for performance calculations using NumPy and SQLAlchemy. Wrote python routines to log into the websites and fetch data for selected options.

Used python modules of Urllib, urllib2, Requests for web crawling. Extensive experience in Text Analytics, developing different Statistical Machine Learning, Data mining solutions to various business problems and generating data visualizations using R, Python and Tableau. Used with other packages such as BeautifulSoup for data parsing. Involved in development of Web Services using SOAP for sending and getting data from the external interface in the XML format. Used with other packages such as BeautifulSoup for data parsing Worked on development of SQL and

stored procedures on MYSQL. Analyzed the code completely and have reduced the code redundancy to the optimal level. Design and build a text classification application using different text classification models. Used Jira for defect tracking and project management. Worked on writing and as well as read data from CSV and excel file formats. Involved in Sprintplanning sessions and participated in the daily AgileSCRUM meetings. Conducted every day scrum as part of the SCRUM Master role. Developed the project in Linux environment. Worked on resulting reports of the application. Performed QA testing on the application. Held meetings with client and worked for the entire project with limited help from the client. Environment: Python, Anaconda, Spyder (IDE), Windows 7, Teradata, Requests, urllib, urllib2, Beautiful Soup, Tableau, python libraries such as NumPy, SQLAlchemy, MySQL Data Scientist/Machine Learning Engineer AT &T - Dallas, TX February 2015 to July 2017 Description: AT&T Inc. provides telecommunications and digital entertainment services. The company operates through four segments: Business Solutions, Entertainment Group, Consumer Mobility, and International. Responsibilities: Used Pandas, NumPy, Seaborn, SciPy, Matplotlib, Scikit-learn, and NLTK in Python for developing various machine learning algorithms. Worked on different data formats such as JSON, XML and performed machine learning algorithms in Python. Setup storage and data analysis tools in Amazon Web Services cloud computing infrastructure. Implemented end-to-end systems for Data Analytics, Data Automation and integrated with custom visualization tools using R, Mahout, Hadoop and MongoDB. Utilized Spark, Scala, Hadoop, HBase, Cassandra, MongoDB, Kafka, Spark Streaming, MLLib, Python, a broad variety of machine learning methods including classifications, dimensionality reduction etc. and utilized the engine to increase user lifetime by 45% and triple user conversations for target categories. Used Spark Data frames, Spark-SQL, Spark MLLib extensively and developing and designing POC's using Scala, Spark SQL and MLLib libraries. Used Data Quality Validation techniques to validate Critical Data Elements (CDE) and identified various anomalies. Participated in all phases of Data-Mining, Data-collection, Data-Cleaning, Developing-Models, Validation, Visualization and Performed Gap Analysis. Data Manipulation and Aggregation from different source using Nexus, Toad, Business Objects and SmartView.

Implemented Agile Methodology for building an internal application. Focus on integration overlap and Informatica newer commitment to MDM with the acquisition of Identity Systems. Good knowledge of Hadoop Architecture and various components such as HDFS, Job Tracker, Task Tracker, Name node, Data node, Secondary Name node, and MapReduce concepts. Programmed a utility in Python that used multiple packages (SciPy, NumPy, Pandas) Implemented Classification using supervised algorithms like Logistic Regression, Decision trees, KNN, Naive Bayes. Updated Python scripts to match training data with our database stored in AWS Cloud Search, so that we would be able to assign each document a response label for further classification. Created SQL tables with referential integrity and developed queries using SQL, SQL PLUS and PL/SQL. Interaction with Business Analyst, SMEs and other Data Architects to understand Business needs and functionality for various project solutions Identifying and executing process improvements, hands-on in various technologies such as Oracle, Informatica, and BusinessObjects. Environment: AWS, Informatica, Python, HDFS, ODS, OLTP, Oracle 10g, Hive, OLAP, DB2, Metadata, MS Excel, Mainframes MS Vision, Map-Reduce, Rational Rose, SQL, and MongoDB. Sr. Python developer UnitedHealth Group - Minneapolis, MN November 2013 to February 2015 Description: Cisco is an American multinational technology that develops, manufactures, and sells networking hardware, telecommunications equipment, and other high-technology services and products. The main goal of the project is to report IT, security, mobility problems faced by the customers and suggest solutions to them. We developed a website for reporting issues and finding solutions to that problem easily. Responsibilities: Analyzed the requirements and designed the flow of task using flow charts and designed flow between pages of the UI. Documented the design solutions and created stories for client requirements. Written REST services using python and Apollo (internal to Cisco). Written Python Scripts to establish continuous workflows from different teams providing data. Written unit and integration tests in python to test the code. Implemented LDAP authentication to authenticate and authorize the Customers using python Rest Services. Generated client certificates in both pem and pfx formats using M2Crypto python module. Used SQLite3 database for caching the client requests. Written

LDAP search filters for both single level and multilevel. Complete UI development using AngularJS, CSS and HTML5. Dashboards with quick filters, parameters and sets to handle views more efficiently. Performed user validations on client side as well as server side. Improved code reuse and performance by making effective use of various design patterns. Efficient delivered code based on principles of Test Driven Development (TDD) and continuous integration to keep in line with Agile Software Methodology principles. Participated with QA to develop test plans from high-level design documentation. Used Rally for Agile software management. Primary contact for all issues in both development and production environments. Implemented the Long-term fix for incidents that are happened in production environment by finding the root cause. Completed the cisco white belt in security, which helps to develop the applications in a secured manner to protect from the threats. Environment: Python 2.7, Cassandra, MySQL, LDAP, Git, Linux, Windows, JSON, JQuery, HTML, XML, CSS, REST, Rally, Bootstrap, JavaScript, Angular JS, Agile, Bitbucket, Py Unit, PyCharm, Microsoft SQL server management studio, DataStax DevCenter, Apache Directory Studio, Ansible, Jenkins, Matplotlib, MOCK, Beautiful Soup, PyTest Machine Learning Engineer/Data Scientist Regions Bank - Birmingham, AL August 2011 to November 2013 Description: Regions Financial Corporation is a bank and financial services company headquartered in the Regions Center in Birmingham, Alabama. The company provides retail and commercial banking, trust, securities brokerage, mortgage and insurance products and services. Responsibilities: Worked with several R packages including knit, dplyr, Spark,R, Causal Infer, Space-Time. Coded R functions to interface with Caffee Deep Learning Framework. Used Pandas, NumPy, Seaborn, SciPy, Matplotlib, Scikit-learn, and NLTK in Python for developing various machine learning algorithms. Installed and used Caffee Deep Learning Framework Worked on different data formats such as JSON, XML and performed machine learning algorithms in Python. Setup storage and data analysis tools in Amazon Web Services cloud computing infrastructure. Implemented end-to-end systems for Data Analytics, Data Automation and integrated with custom visualization tools using R, Mahout, Hadoop and MongoDB. Worked as Data Architects and IT Architects to understand the movement of data and its storage and

ERStudio9.7. Utilized Spark, Scala, Hadoop, HBase, Cassandra, MongoDB, Kafka, Spark Streaming, MLLib, Python, a broad variety of machine learning methods including classifications, regressions, dimensionally reduction etc. and Utilized the engine to increase user lifetime by 45% and triple user conversations for target categories. Used Spark Data frames, Spark-SQL, Spark MLLib extensively and developing and designing POC's using Scala, Spark SQL and MLLib libraries.

Used Data Quality Validation techniques to validate Critical Data Elements (CDE) and identified various anomalies. Extensively worked on Data Modeling tools Erwin Data Modeler to design the Data Models. Developed various Qlik-View Data Models by extracting and using the data from various sources files, DB2, Excel, Flat Files and Bigdata. Participated in all phases of Data-Mining, Data-collection, Data-Cleaning, Developing-Models, Validation, Visualization and Performed Gap Analysis. Data Manipulation and Aggregation from a different source using Nexus, Toad, BusinessObjects, Power BI and SmartView. Implemented Agile Methodology for building an internal application. Focus on integration overlap and Informatica newer commitment to MDM with the acquisition of Identity Systems. Good knowledge of Hadoop Architecture and various components such as HDFS, Job Tracker, Task Tracker, Name Node, Data Node, Secondary Name Node, and MapReduce concepts. As Architect delivered various complex OLAP Databases/Cubes, Scorecards, Dashboards and Reports. Programmed a utility in Python that used multiple packages (SciPy, NumPy, Pandas). Implemented Classification using supervised algorithms like Logistic Regression, Decision Trees, KNN, Naive Bayes. Designed both 3NF data models for ODS, OLTP systems and Dimensional Data Models using Star and Snowflake Schemas.

Updated Python scripts to match training data with our database stored in AWS Cloud Search, so that we would be able to assign each document a response label for further classification. Created SQL tables with referential integrity and developed queries using SQL, SQLPLUS and PL/SQL. Designed and developed Use Case, Activity Diagrams, Sequence Diagrams, OOD (Object oriented Design) using UML and Visio. Interaction with Business Analyst, SMEs, and other Data Architects to understand Business needs and functionality for various project solutions Identifying and executing process improvements, hands-on in various technologies such as Oracle, Informatica,

and BusinessObjects. Environment: AWS, R, Informatica, Python, HDFS, ODS, OLTP, Oracle 10g, Hive, OLAP, DB2, Metadata, MS Excel, Mainframes MS Vision, Map-Reduce, Rational Rose, SQL, and MongoDB. Data Scientist Cisco - San Jose, CA May 2009 to August 2011 Description: UnitedHealth Group Inc. is an American for-profit managed health care company based in Minnetonka, Minnesota. It is sixth in the United States on the Fortune 500. UnitedHealth Group offers health care products and insurance services. UnitedHealth Group is the largest healthcare company in the world by revenue \$184 billion in 2016. UnitedHealth Group subsidiaries companies together serve approximately 115 million individuals in 2016. Responsibilities: Provided the architectural leadership in shaping strategic, business technology projects, with an emphasis on application architecture. Utilized domain knowledge and application portfolio knowledge to play a key role in defining the future state of large, business technology programs. Participated in all phases of data mining, data collection, data cleaning, developing models, validation, and visualization and performed Gap analysis. Developed MapReduce/Spark Python modules for machine learning & predictive analytics in Hadoop on AWS. Implemented a Python-based distributed random forest via Python streaming. Created ecosystem models (e.g. conceptual, logical, physical, canonical) that are required for supporting services within the enterprise data architecture (conceptual data model for defining the major subject areas used, ecosystem logical model for defining standard business meaning for entities and fields, and an ecosystem canonical model for defining the standard messages and formats to be used in data integration services throughout the ecosystem). Used Pandas, NumPy, seaborn, SciPy, Matplotlib, Scikit-learn, NLTK in Python for developing various machine learning algorithms and utilized machine learning algorithms such as linear regression, multivariate regression, naive Bayes, Random Forests, K-means, & KNN for data analysis. Conducted studies, rapid plots and using advance data mining and statistical modelling techniques to build a solution that optimize the quality and performance of data. Demonstrated experience in design and implementation of Statistical models, Predictive models, enterprise data model, metadata solution and data life cycle management in both RDBMS, Big Data environments. Analyzed large data sets apply machine learning techniques and develop

predictive models, statistical models and developing and enhancing statistical models by leveraging best-in-class modeling techniques. Worked on database design, relational integrity constraints, OLAP, OLTP, Cubes and Normalization (3NF) and De-normalization of the database. Developed MapReduce/Spark Python modules for machine learning & predictive analytics in Hadoop on AWS. Worked on customer segmentation using an unsupervised learning technique - clustering. Worked with various Teradata tools and utilities like Teradata Viewpoint, Multi Load, ARC, Teradata Administrator, BTEQ and other Teradata Utilities. Utilized Spark, Scala, Hadoop, HBase, Kafka, Spark Streaming, MLlib, Python, a broad variety of machine learning methods including classifications, regressions, dimensionally reduction etc. Developed LINUX Shell scripts by using NZSQL/NZLOAD utilities to load data from flat files to Netezza database. Designed and implemented system architecture for Amazon EC2 based cloud-hosted solution for the client. Tested Complex ETL Mappings and Sessions based on business user requirements and business rules to load data from source flat files and RDBMS tables to target tables. Environment: Erwin r9.6, Python, SQL, Oracle 12c, Netezza, SQL Server, Informatica, Java, SSRS, PL/SQL, T-SQL, Tableau, MLlib, regression, Cluster analysis, Scala NLP, Spark, Kafka, MongoDB, logistic regression, Hadoop, Hive, Teradata, random forest, OLAP, Azure, MariaDB, SAP CRM, HDFS, ODS, NLTK, SVM, JSON, Tableau, XML, Cassandra, MapReduce, AWS. Education Bachelor's Skills Analysis services, Application development, Hadoop, Hbase, Hdfs, Javascript, MySQL, Android, ASP, PHP

Name: Devin Wells

Email: lcruz@example.com

Phone: 001-241-792-9240x52952