



Maresh

Python Data Science | Machine Learning | Django REST API | PySpark | Mongo DB | SQL | ETL / Dwh | Business Intelligence | Unix

Objective

Looking for a challenging role as **(Python Developer)** in a good organization to utilize my **Python Machine Learning / Django REST / Data Intelligence** skills for the growth of the organization as well as to enhance my knowledge about emerging trends in the IT sector. In the career of 3.3 years of experience acquired a good knowledge about Data Science and Web Application Development field. Data Scientist with strong Statistics and Mathematics background and 3+ years of experience using Predictive Modeling, Data Processing, and Data Mining Algorithms to solve challenging business problems. Involved in Python Open-Source Community and passionate about Deep Reinforcement Learning. A self-motivated IT professional seeks for a thriving, exciting, and growing career in the field of IT-Software Industry especially for roles such as Data Scientist/ML/AI/Python Developer where my profound Data Science / Web Application Development along with Python Programming, SQL and UNIX knowledge and experience in Programming Concepts and Methodologies in Software Development are shared and my all-rounder development is encouraged.

Technical Qualification Headlines

Currently working as a **'Python-Data Scientist'** for 'Voice and Message Routing Optimization' with client Aussie Broadband, Australia at IBM Pune. Strong work experience with Technologies such Python Data Science, Machine Learning, Django REST API, PySpark, SQL, MongoDB, UNIX, ETL Dwh, Business Intelligence.

Core Libraries Worked With:

- Pandas
- Numpy
- Seaborn, Matplotlib, Scipy
- Scikit Learn
- Keras, Tensorflow
- Python ETL – Spark
- BI Tool used - Tableau
- Framework involvement : Django REST
- Web Scraping Library: BeautifulSoup 4
- Unstructured Database – Mongo DB

Machine Learning Algorithm Work Experience:

- Linear Regression
- Logistic Regression
- KNN
- K-Means
- Naïve Bayes, XG Boost
- Decision Tree
- Random Forest
- Support Vector Machine
- Principal Component Analysis
- NLP

Process worked with: DevOps + Agile

Software Programming Language Experience:



- Python
- Java
- C++
- Java Script
- CSS
- HTML
- Boot Strap
- J Query

Business Implementation Programming Experience:

- Machine learning – Python
- Django REST API
- ETL – Python
- Business Intelligence

Strong Areas

- Python
- SQL,
- Unix
- Django Web service-REST
- Machine Learning Algorithm
- Data Analytics –ETL

Professional Work Experience

Data Science / Machine Learning & Django Rest / ETL Task:

- Highly efficient Data Scientist/Data Analyst with 3+ years of experience in Data Analysis, Machine Learning, Data mining with large data sets of Structured and Unstructured data, Data Acquisition, Data Validation, Predictive modeling, Data Visualization, Web Scraping.
- 3+ years of Experience in design, development, testing and implementation of various stand - alone and client-server architecture based enterprise application software in Python on different domains .
- Experience in statistical programming languages with Python including Big Data technologies like Pyspark
- Proficient in managing entire data science project life cycle and actively involved in all the phases of project life cycle including Data Acquisition, Data Cleaning, Data Engineering, Features Scaling, Features Engineering, Statistical Modeling (Decision Trees, Regression Models, Clustering), Dimensionality Reduction using Principal Component Analysis (PCA) and Factor Analysis, Testing and Validation using AUC/ROC plot, K - fold Cross-Validation and Data Visualization.
- Experience and deep understanding of Statistical modeling, Multivariate Analysis, Model Testing, Problem Analysis, Model Comparison, and Validation.
- 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.
- Skilled in performing Data Parsing, Data Manipulation and Data Preparation with methods including describe data contents, compute descriptive Statistics of data, Regex, split and combine, remap, merge, subset, reindex, melt and reshape.
- Experience in using various Libraries and Packages in python-like NLP, Pandas, NumPy, Seaborn, SciPy, Matplotlib, Sci-kit-learn, Beautiful Soup, Json, CSV, Highly skilled in using visualization tools like Tableau for creating dashboards.
- Extensive experience in Text Analytics, generating data visualizations using Python and creating dashboards using tools like Tableau.

- Hands on experience with big data tools like PySpark, Spark SQL. Hands on experience in implementing LRA, Naive Bayes and skilled in Random Forests, Decision Trees, Linear and Logistic Regression, SVM, Clustering, Principle Component Analysis.
- Good Knowledge in Proof of Concepts (PoC's), gap analysis and gathered necessary data for analysis from different sources, prepared data for data exploration using data munging.
- Good industry knowledge, analytical & problem-solving skills and ability to work well within a team as well as an individual.
- Experience in designing stunning visualizations using Tableau software and publishing and presenting dashboards, Storyline on web and desktop platforms.
- Experience and Technical proficiency in Designing, Data Modeling Online Applications, Solution Lead for Architecting Data Warehouse/Business Intelligence Applications.
- Experience with Data Analytics, Data Reporting, Ad-hoc Reporting, Graphs, Scales, Pivot Tables and OLAP reporting.
- Highly skilled in using Hadoop (pig and Hive) for basic analysis and extraction of data in the infrastructure to provide data summarization.
- Worked and extracted data from various database sources like Oracle, SQL Server, DB2, regularly accessing JIRA tool and other internal issue trackers for the Project development.
- Highly creative, innovative, committed, intellectually curious, business savvy with good communication and interpersonal skills.

Django REST Task:

- Good experience in developing Web Applications implementing **Model-View-Control Architecture using Django REST Application Framework.**
- Strong expertise in development of **Web-Based Applications using Python, HTML, CSS, JavaScript, JSON and JQuery.**
- Experience in working with **Python ORM Libraries including Django ORM, SQLAlchemy.**
- Good knowledge of **Web Services with protocols SOAP, REST.**
- Experience in working with **continuous deployment using Heroku and Jenkins.**
- Good experience in working with **Amazon Web Services like EC2, Virtual private clouds (VPCs), Storage models (EBS, S3, instance storage), Elastic Load Balancers (ELBs).**
- Experience in using various version control systems like **Git, and GitHub.** Having good knowledge in using **NoSQL databases like MongoDB .Proficient in writing SQL Queries.**
- Excellent Interpersonal and communication skills, efficient time management and organization skills, ability to handle multiple tasks and work well in a team environment.

Key Skills

- Frameworks: Django, CSS , Bootstrap.
- Web Technologies: HTML, CSS, Java Script, JQuery, JSON, XML.
- Programming Languages: Python, HTML, CSS, SQL.
- Version Control: Git, GitHub.
- Cloud Computing: AWS EC2, S3.

- Analytic Tools: PySpark.
- Application servers: Apache Tomcat, Nginix, JBoss, WebLogic, WebSphere.
- Databases: Oracle 10g, MySQL, MongoDB.
- IDEs/Development Tools: PyCharm, Anaconda – Jupyter Nb, Atom, Spyder, Microsoft Visual Studio.
- Operating Systems: Windows, Red hat Linux.
- Protocols: TCP/IP, HTTP/HTTPS, SOAP.
- Deployment Tools: Heroku, Jenkins.
- Issue Trackers: HPALM 12.5.
- Web Service: REST–JSON, DRF (Django Rest Framework)

Educational Details

Masters in Technology, (M. Tech - CAD/CAM & Automation) from VJTI (Govt. Autonomous), Mumbai, Mumbai University)

Achievements

- Received “Star Performer Award” from IBM for good performance.
- Received appreciation for E2E Delivery from client Mohawk Industries, USA.

Projects

Project Sequence 1

- **Project Name** : **Product Analytics & Metrics Visualization**
- **Vertical** : E-commerce, Payments
- **Client** : Mohawk Industries, USA
- **Technology & Tool** : Python, Data Science, Machine Learning, PySpark, Spark SQL, ETL, Statistical Analysis
- **Roles** : Python Data Scientist

Detail Project Overview and Workflow:

Product Analytics and Metrics Visualization is the process of collecting, gathering data from all areas that have an influence on user online store and using this information to understand the trends and the shift in consumer's behavior to make data-driven decisions that will drive more online sales. Price and stock information about business competitors and where Business are positioned against them pricewise is a very important data point user need to consider if user want to sustain and compete in the E-Commerce sector. Product Analysis allows user to understand which products and product lists are well performing, why they're performing or underperforming, as well as what we can do to optimize and enhance sales. Product Metrics includes number of active products, number of products, gross margin, product detail views, number of reviews, product added to different carts, product removes from cart, product conversion rate etc; which are vital elements for business.

Important Modules: Data Repository Collection, Customer Behavior Analytics, Sales Analytics, Marketing Performance analysis, Product analytics, Conversion Optimization.

Task Handled:

- Implemented Data Exploration to analyze patterns and to select features using Python and SciPy.
- Built Factor Analysis and Cluster Analysis Models using Python and SciPy to classify customers into different Target Groups.
- Evaluated business requirements and prepared detailed specifications that follow project guidelines required to develop written programs.
- Participated in Data Acquisition with Data Engineer team to extract historical and real-time data by using PySpark.
- Communicated and presented default customers profiles along with reports using Python and Tableau, analytical results and strategic implications, to senior management for strategic decision making. Developed scripts in Python to automate the customer query addressable system using python which decreased the time for solving the query of the customer by 45%. Collaborated with other functional teams across the Risk and Non-Risk Groups to use standard methodologies and ensure a positive customer experience throughout the customer journey.
- Supported MapReduce Programs running on the cluster.
- Performed Data Enrichment jobs to deal missing value, to normalize data, and to select features.
- 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, Bing Ads, AdWords, AdSense, Criteo, Smartly, SurveyMonkey, and Mailchimp) in order to complete business & marketing analysis and assessment.
- Used Git 2.x for version control with Data Engineer team and Data Scientists colleague. Used Agile methodology and SCRUM process for project developing.
- Structuring the Data Marts to store and organize the customer's data.
- Running SQL scripts, creating indexes, stored procedures for data analysis, Data Lineage methodology for Data Mapping and maintaining data quality.
- Prepared Scripts in Python and Shell for Automation of administration tasks.
- Maintained PL/SQL objects like packages, triggers, procedures etc. Mapping flow of trade cycle data from source to target and documenting the same.
- Performing QA on the data extracted, transformed and exported to excel.
- Participated in all phases of Data Mining; Data Collection, Data Cleaning, Developing Models, Validation, Visualization and performed Gap Analysis.
- Extracted data from HDFS and prepared data for Exploratory Analysis using data Munging.
- Built models using Statistical techniques and Machine Learning Classification Models like XG Boost, SVM, Random Forest. etc.
- A highly immersive Data Science program involving Data Manipulation & Visualization, Web Scraping, Machine Learning, Python programming, SQL, GIT, Unix Commands, NoSQL, MongoDB, Hadoop.
- Used Pandas, Numpy, Seaborn, Scipy, Matplotlib, Scikit-learn, 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.

Project Sequence 2

- **Project Name** : **Virtual Terminal**
- **Vertical** : E- Commerce , Payments
- **Client** : Mohawk Industries, USA
- **Technology & Tool** : Python, Django REST-API, PyTest, Postman, ETL, UNIX
- **Roles:** : Python Developer

Detail Project Overview and Workflow:

This Payment Gateway Solution integrate all your payments along with banking, expense management & accounting - all in ONE online bank account. This Digital Ecommerce Payment Gateway is to approve the transaction process between merchant and customer. It plays a vital role in the online transaction process and authorizes transactions between merchants and customers. Every detail of a transaction can be potentially reported on. From the Virtual Terminal or user which initiated the transaction, to a simplified summary of all transactions – Mohawk Industries has Business Intelligence to streamline existing processes. Reports can be integrated with several accounting packages, or customized if necessary, making reconciliation and financial reporting a breeze. Available in real time and in a variety of formats: CSV, PDF and in XML - Reports can be securely accessed via Pay line, or for corporate clients this solution can automate this service. There is no limit to the size of report which may be generated, and all historical merchant data is readily accessible. This payment gateway is a merchant service provided by a E-commerce application service provider that authorizes credit card or direct payments processing for e-businesses, online retailers, bricks and clicks, or traditional brick and mortar. The payment gateway may be provided by a bank to its customers but can be provided by a specialized financial service provider as a separate service, such as a payment service provider. This payment gateway facilitates a payment transaction by the transfer of information between a payment portal (such as a website, mobile phone or interactive voice response service) and the front-end processor or acquiring bank.

Task Handled:

- Understand the FRS document and analyze the requirement, specifications and also in client interaction during requirements specifications.
- Worked on backend of the application using Python, Django-REST API, JSON and JQuery.
- Involved in analysis and design of the application features, designed and developed communication between client and server using Secured Web services.
- Created UI using JavaScript and HTML5/CSS, Developed and tested many features for dashboard using Python, Java, Bootstrap, CSS, JavaScript and JQuery.
- Wrote backend programming in Python, used JavaScript and XML to update a portion of a webpage.
- Developed consumer based features and applications using Python and Django in Test Driven Development and Pair Based Programming.
- Performed Unit testing, Integration Testing, GUI and Web Application testing using Rspec.
- Implemented user interface guidelines and standards throughout the development and maintenance of the website using the HTML, CSS, JavaScript and JQuery.
- Worked on Jenkins continuous integration tool for deployment of project, worked on changes to open stack and AWS to accommodate large-scale data center deployment.

- Used advanced packages like Mock, patch and beautiful soup (b4) to perform Unit Testing. Worked on translation of web pages to different languages as per client requirements.
- Implemented responsive vector maps and charts on web pages using the data from MongoDB.
- Collaborated with Product Management and User Experience experts regarding product definition, schedule, scope and project-related decisions. Also managed, collaborated and coordinated the work of an offshore development team.
- Worked on rebranding the existing web pages to clients according to the type of deployment.
- Worked on updating the existing clip board to have the new features as per the client requirements. Used many Regular Expressions in order to match the pattern with the existing one.
- Skilled in using collections in Python for manipulating and looping through different user defined objects. Worked with team of developers on Python applications for RISK management.
- Designed an A/B Experiment for testing the business performance of the new recommendation system.
- Taken part in entire lifecycle of the projects including design, development, and deployment, testing and implementation and support. Improved code reuse and performance by making effective use of various design patterns

Project Sequence 3

- **Project Name** : Customer Segmentation and Product Development
- **Client** : Aussie Broadband, Australia.
- **Vertical** : Telecom OSS Provisioning and Activation –Inventory Management
- **Technology & Tool** : Python, Data Science, Statistical Analysis, EDA, Machine Learning Algorithms
- **Role** : Python Data Scientist.

- **Detail Project Overview and Workflow:**

Telecom industry uses data science for different purposes-Maximizing Profit, Planning Efficient Business and Market Strategies, Data Visualization, Performing Data Transmission, etc. With the help of predictive analytics to the core the industry can understand their clients well, customer segmentation is one such application of the Telecom industry, if the telecom organization can divide the market and then focus on the content then success is not far. Strategic planning and careful planning continued throughout the product development life cycle will improve a loyal customer base, by using the insights collected from data analysis and predictive analytics targeted product development is feasible. Real time data is used to gather real time issues solve them and give responses. Real time analytics collates information related to location, customer profiles usage, traffic and network. Customer sentimental analysis assesses the negative and positive feedback of a consumer for a product. When the collected information is analyzed properly, main insights newest trends are revealed. A real time solution to the issues experienced by the clients is also possible. Text analysis techniques also enable consumer sentiment analysis.

Task Handled:

- Evaluated business requirements and prepared detailed specifications that follow project guidelines required to develop written programs.
- Implemented Data Exploration to analyze patterns and to select features using Python and SciPy.
- Built Factor Analysis and Cluster Analysis models using Python SciPy to classify customers into different target groups.

- Implemented Data Exploration to analyze patterns and to select features using Python and SciPy.
- Performed Multivariate Analysis, Predictive Modelling, Cluster, Market Basket Analysis using sophisticated statistical techniques.
- Interpreted & translated analytic output into insights.
- Determined the source of the data, coordinate extraction and acquisition.
- Assess data quality, identify gaps in the data and eliminate irrelevant data.
- Implemented Data Exploration to analyze patterns and to select features using Python and SciPy.
- Built Factor Analysis and Cluster Analysis Models using Python and SciPy to classify customers into different Target Groups.
- Performed Data Enrichment jobs to deal missing value, to normalize data, and to select features.
- Supported MapReduce Programs running on the cluster.
- Participated in Data Acquisition with Data Engineer team to extract historical and real-time data by using Hadoop-MapReduce and HDFS.
- Communicated and presented default customers profiles along with reports using Python and Tableau, analytical results and strategic implications, to senior management for strategic decision making. Developed scripts in Python to automate the customer query addressable system using python which decreased the time for solving the query of the customer by 45%. Collaborated with other functional teams across the Risk and Non-Risk Groups to use standard methodologies and ensure a positive customer experience throughout the customer journey.
- Developed multiple MapReduce jobs for Data Cleaning and Pre-processing, analyzed the partitioned and bucketed data and compute various metrics for reporting.
- Created HBase tables to store various data formats of data coming from different portfolios.
- Created reports and dashboards, by using 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, Bing Ads, AdWords, AdSense, Criteo, Smartly, SurveyMonkey, and Mailchimp) in order to complete business & marketing analysis and assessment.
- Used Git 2.x for version control with Data Engineer team and Data Scientists colleague. Used Agile methodology and SCRUM process for project developing.
- Structuring the Data Marts to store and organize the customer's data.
- Running SQL scripts, creating indexes, stored procedures for data analysis, Data Lineage methodology for Data Mapping and maintaining data quality.
- Worked with brands and business units to ensure that all consumer and behavioral data is stored in a Centralized Enterprise Repository using Amazon Web Services, Redshift and C3.
- Participated in all phases of Data Mining; Data Collection, Data Cleaning, Developing Models, Validation, Visualization and performed Gap Analysis.
- Extracted data from HDFS and prepared data for Exploratory Analysis using data Munging.
- Built models using Statistical techniques like Bayesian HMM and Machine Learning Classification Models like XG Boost, SVM, Random Forest. etc.
- A highly immersive Data Science program involving Data Manipulation & Visualization, Web Scraping, Machine Learning, Python programming, SQL, GIT, Unix Commands, NoSQL, MongoDB, Hadoop.

- Used Pandas, Numpy, Seaborn, Scipy, Matplotlib, Scikit-learn, 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.
- Created HBase Tables to store various data formats of data coming from different portfolios.
- Provided summary statistics of key performance metrics and other measures deemed significant business units.

Project Sequence 4

- **Project Name** : **Voice and Message Routing Optimization**
- **Client** : Aussie Broadband, Australia.
- **Vertical** : Telecom OSS Provisioning and Activation –Inventory Management
- **Technology & Tool** : Python, Django REST, Oracle ,GIT , Rest – Web Service ,Pandas, ETL, Statistical Analysis.

- **Role** : Python Developer

- **Detail Project Overview and Workflow:**

It is a high performance modular and fully redundant aggregation router, designed to enable high quality network service delivery for RAN and fixed/mobile converged metro aggregation networks. In its category, it sets a new benchmark for port density by scaling up to 144x10G and 24x100G interfaces and offering up to 2.7 Tbps switching capacity in a space efficient 5RU chassis with front access for all field replaceable units allowing an overall lower OPEX. It supports VPN services over IP/MPLS networks, service provider SDN, service exposure using NETCONF/YANG, extensive quality of service and precise synchronization features. The Router 6274 has strong security features such as IPSec and vendor software authentication for ubiquitous deployment. With 2.7 Tbps of switching capacity, the Router 6274 delivers performance needed to fully support LTE, LTE Advanced, 5G, Fixed Mobile Convergence and Enterprise applications. The Router 6274 is part of the Ericsson Router 6000 Series, a radio integrated and subscriber aware IP transport family of products. The Router 6000 offers a range of high-performance routers with resiliency features and form factors optimized for the various needs of metro and backhaul networks. This equipment is an advanced 4G/5G access router and pre-aggregation router with 100Gb forwarding capacity.

Task Handled:

- Understand the FRS document and analyze the requirement, specifications and also in client interaction during requirements specifications.
- Designed the front end of the application using Python, HTML, CSS, AJAX, JSON and JQuery. Worked on backend of the application, mainly using Active Records.
- Involved in analysis and design of the application features, designed and developed communication between client and server using Secured Web services.
- Created UI using JavaScript and HTML5/CSS, Developed and tested many features for dashboard using Python, Java, Bootstrap, CSS, JavaScript and JQuery.
- Wrote backend programming in Python, used JavaScript and XML to update a portion of a webpage.
- Developed consumer based features and applications using Python and Django in Test Driven Development and Pair Based Programming.
- Performed Unit testing, Integration Testing, GUI and Web Application testing using Rspec.



- Implemented user interface guidelines and standards throughout the development and maintenance of the website using the HTML, CSS, JavaScript and JQuery.
- Worked on Jenkins continuous integration tool for deployment of project, worked on changes to open stack and AWS to accommodate large-scale data center deployment.
- Used advanced packages like Mock, patch and beautiful soup (b4) to perform Unit Testing. Worked on translation of web pages to different languages as per client requirements.
- Implemented responsive vector maps and charts on web pages using the data from MongoDB.
- Worked on rebranding the existing web pages to clients according to the type of deployment.
- Worked on updating the existing clip board to have the new features as per the client requirements. Used many Regular Expressions in order to match the pattern with the existing one.
- Skilled in using collections in Python for manipulating and looping through different user defined objects. Worked with team of developers on Python applications for RISK management.
- Taken part in entire lifecycle of the projects including design, development, and deployment, testing and implementation and support. Improved code reuse and performance by making effective use of various design patterns