|  |  |  |
| --- | --- | --- |
| Ritika Shrivastava  Python | Data Science | Machine Learning | Django –REST | ETL/BI |Cloud AWS | Java Script | |  |
| Machine Learning Engineer,IBM - https://www.ibm.com/-Pune   |  |  | | --- | --- | |  | | |  | | |  | **CAREER OBJECTIVE** | |  | | | A challenging carrier as **Data scientist / Web Developer** where **Python Machine Learning /AI/Data Intelligence** skills can be effectively used and upgraded.  **Data Scientist** with strong math 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.  Looking for a challenging career in the field of IT-Software especially **Data scientist/ML/AI +Python Programming** where my Strong a sql and unix knowledge and experience in Programming Concepts and Methodologies in Software Development are shared and with Development encouraged. | | |  | | |  | **PROFESSIONAL EXPERIENCE** | |  | |   Currently working as a ’Python Data Scientist ’ for Voice and Message Routing Optimization with client Aussie Broadband,Australia  **Technical qualification Headlines :**  **Experience Summary - Technology ,Tools and platform**  **Experience with Core Data Science/ML Libraries /Packages :**  PySpark | Pandas| Numpy | Seaborn | Matplotlib |Scikit Learn | Keras **|** Tensorflow | Scipy |Statsmodels | Plotly  **Experience with Core Django REST Libraries /Packages:**  Django-rest-framework |Django-cors-headers | Django-debug-toolbar | Django-extensions |Sentry-sdk | Django-allauth |Django-filter | Django-import-export |  **Major Tools , Associated Technologies & Environment**   * Unstructured Database Used – Mongo DB * UI Application Programming Language - HTML, CSS, JavaScript, JQuery and Bootstrap. * Web Scraping Library : Beautiful Soup 4 * Web Framework involvement : Django REST * BI Tool used - Tableau * Parsing XML - **xml Schema**   **Experience with Machine Learning Algorithm**  Linear Regression | Logistic Regression | KNN |K-Means |Naïve Bayes |Mean, Median, Mode |Decision Tree |Random Forest  Support Vector Machine |Principal Component Analysis |NLP |XG Boost  **Process worked with :** DevOps + Agile + V- Model  **Software Programming Language Experience:**  Python |Java | C++ |Java Script |CSS |HTML |Boot Strap |J Query  **Business Implementation Programming Experience:**  Machine learning – Python | ETL – Python | Business Intelligence  **Strong Experience Areas**  Python | SQL | Unix Data Analytics –ETL |Machine Learning Algorithm | Django Web service-REST  **EXPERIENCE SUMMARY**  **Data Science / ETL/ ML & Django Rest Task as a Full stack Developer**   * 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 Hadoop, Hive,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 and Factor Analysis, testing and validation using 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 Hadoop, Spark, Hive, Pig, PySpark, Spark SQL,PySpark** Hands on experience in implementing LDA, Naive Bayes and skilled in Random Forests, Decision Trees, **Linear and Logistic Regression**, SVM, Clustering, neural networks, 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. * 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. * 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, PivotTables 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. * Extensive experience in Data Visualization including producing tables, graphs, listings using various procedures and tools such as Tableau.   **Django REST Task :**   * Good experience in developing web applications implementing **Model View Control architecture using Django REST application frameworks.** * Strong expertise in development of **web based applications using Python,HTML, XML, KML, CSS, DHTML, 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,** * Experience in working with **Python ORM Libraries including Django ORM, SQLAlchemy.** * Excellent Interpersonal and communication skills, efficient time management and organization skills, ability to handle multiple tasks and work well in a team environment.   **TECHNICAL 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, My SQL, MongoDB * IDE s/ Development Tools: PyCharm * 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)   **WOKING ZONE – ORGANIZATION**  Currently working as **Data Scientist** With IBM , <https://www.ibm.com/>Punesince Oct 2018 to till date.  **PROJECTS DETAILS**    **Project Sequence 1**   * **Project Name :** Voice and Message Routing Optimization * **Vertical :** Telecom OSS –Provisioning & Activation [Inventory Management ] * **Client** Aussie Broadband, Australia * **Technology & Tool :** Python ,Django REST , Oracle ,GIT , Rest – Web service ,Pandas ,etl,Statistical Analysis * **Role –** Python Data scientist   **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 2.7Tbps 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 2.7Tbps 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 analyse the requirement,Analyzed system requirements 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. * Writing backend programming in Python,Used JavaScript and XML to update a portion of a webpage. * Develop 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. * 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. * Updated the client admin tool to have the new features like support for internalization, support for customer service and etc. * Implemented responsive vector maps and charts on web pages using the data from MongoDB . * Collaborate with Product Management and User Experience experts regarding product definition, schedule, scope and project-related decisions, Manage, collaborate and coordinate 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,Work 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. * Documented the design solutions and created stories for client requirements.   **Project Sequence 2**   * **Project Name : Product Analytics &Metrics Visualization** * **Vertical :** E- commerce , Payments * **Client : Mohawk Industries, USA** * **Technology & Tool :** Python, Pyspark, Spark SQL, Plotly, Dash, Post Man Microsoft Azure, , Docker,ETL,ML ,Statistical Analysis * **Roles:** Python Data scientist   **Detail Project Overview and Workflow :**  Product Analytics & Metrics Visualization is the process of gathering data from all areas that have an impact on user online store and using this information to understand the trends and the shift in consumers' 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 price-wise is a very important data point user need to consider if user want to compete in the e-commerce space. Product analysis allows user to understand which products and product lists are performing, why they’re performing or underperforming, as well as what you can do to optimize sales.Product matrics su as Number of products, Number of active products, Gross Margin,Number of Reviews,Product detail views,Product add to carts,Product removes from cart, Product conversion rate are vital element for business .  Important Modules :   * Sales Analytics ,Data Reprository collection ,Customer Behaviour Analytics ,Marketing Performance analysis * Conversion Optimization ,Product analytics * **Task Handled :** * 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. * Designed an A/B experiment for testing the business performance of the new recommendation system. * Supported MapReduce Programs running on the cluster. * 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 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. * Performed Data Enrichment jobs to deal missing value, to normalize data, and to select features. * Developed multiple MapReduce jobs in java for data cleaning and pre-processing,Analyzed the partitioned and bucketed data and compute various metrics for reporting. * Extracted data from Twitter using Java and Twitter API. Parsed JSON formatted twitter data and uploaded to database. * Developed Hive queries for analysis, and exported the result set from Hive to MySQL using Sqoop after processing the data. * 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, 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 colleagues,Used Agile methodology and SCRUM process for project developing. * KT with the client to understand their various Data Management systems and understanding the data,Creating meta-data and data dictionary for the future data use/ data refresh of the same client. * 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 like Bayesian HMM and Machine Learning classification models like XG Boost, SVM, Random Forest. * 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 3**   * **Project Name : Virtual Tesrminal** * **Vertical :** E- commerce , Payments * **Client : Mohawk Industries, USA** * **Technology & Tool :** Python, Pyspark, Spark SQL, Plotly, Dash, Post Man Microsoft Azure, , Docker,ETL,ML ,Statistical Analysis * **Roles:** Python Data scientist   **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 Ecommerce 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:**   * 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. * 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. * Created HBase tables to store various data formats of data coming from different portfolios. * Provide summary statistics of key performance metrics and other measures deemed significant business units * Coordinate and manage data analytics activities with stakeholders * Perform multivariate analysis, predictive modelling, cluster, market basket analysis using sophisticated statistical techniques * Interpret & translate analytic output into insights * Determine the source of the data, coordinate extraction and acquisition * Work 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 * Assess data quality, identify gaps in the data and eliminate irrelevant data   **Personal Details:**   * Fathers name :- Satish Kumar Shrivastava * Mother name :- Maniprabha Shrivastava * Current address :- Radhavallabh nivas, katthi mohhala ward no.23,singhpur, shahdol(m.p) 484001. * Date of birth:- 17/08/1992 * Martial status :- Single * Language:- Hindi , English   Location :  Pune :  Regards:  Ritika Shrivastava | |  |  | | --- | --- | |  | | |  | | |  | | |  | **CONTACT** | |  | | | **Phone**  **+**91- 9713261086 | | | **Email** | | | ritikashrivastava17@gmail.com | | |  | | |  | **EDUCATION** | |  | | | *\* B.E - Computer Science Engineering-RKDF Institute of Science and technology.*  *\* University :- Rajiv Gandhi university, Bhopal(Madhya Pradesh* | | |  | | |  | **RELEVANT SKILLS** | |  | | |  | |   Python/ETL/Big Data /Business Intelligence /Analytical Language using PySpark / Django REST/AWS-Cloud/Mongo DB/ Unix  **AWARDS**   * Received Performance Star performer award from IBMfor good performances . * Received appreciation for E2E Delivery from Client **Mohawk Industries, USA** |  |