**Ranjeet Kumar**– + 91 7384710470 Email- kumarranjeet632@gmail.com

Associate Data Scientist

Netcracker Technology.

***Technical Qualification Headlines:-***

Currently working as a **’Associate Data Scientist** ’ in Netcracker Technologywith client Screwfix.

**Technical qualification :** Strong work experience with technology **Data Science/Machine learning /Data Analytics /BigData /Business Intelligence.**

**Software progrmming Language Work exp :**  **Python,SQL**

**Core Libraries worked with : Pandas, Numpy, Matplotlib, Seaborn, Scikit Learn, Tensorflow, Keras, Scipy, Spacy, Ploty,NLTK, PySpark**

**Machine learning Algorithm Work exp : Linear Regression, Logistic Regression, KNN, Naïve Bayes, Decision Tree, Random forest, k-means, Support Vector Machine, XG Boost,** [**Artificial neural networks**](https://en.wikipedia.org/wiki/Artificial_neural_network)**-MLP, NLP,stemming, lemmatization, Bag of words**

**Machine learning & Deep learning Algorithm Strong Understanding : RNN, LSTM, CNN, Time Serise, Encoder and Decoder, [Boosting](https://en.wikipedia.org/wiki/Ensemble_learning" \l "Boosting), Word2Vec, Computer Vision.**

**Tools & other Framework Work exp :**

**Python - Spark | BI Tool used - Tableau | Frame work involvement : Django rest |Web scrapping Libraries : Beautiful Soup 4 |Unstructured DB – Mongo DB |structured DB – MySQL**

Software progrmming Language Understanding : Java | c++ | Java script | CSS | html |Boot Strap | J query

Process worked with : DevOps, Agile, Vmodel

Business Implementation Programming Experience : **ETL – Python , Machine learning –Python ,Business Intelligence**

Strong area : **Python , Data Analytics , Machine Learning Algorithm ,Artificial Intelligence, SQL , Unix**

Domain /Component Type worked with : **Pyment and E- Commerce and Telecom**

***Objective:-***

As **Data scientist / Web Developer** is a challenging carrier 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 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.

***Experience Summary:-***

**Data Science / ML & Data Analyst :-**

* 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.
* 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 **pandas, NumPy, Seaborn, SciPy, Matplotlib, Ploty, 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,NLP.
* 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 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.
* 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.

***Technical Skills:-***

* Version Control:Git, GitHub
* Devops Tool : Docker
* Cloud Computing: AWS EC2, S3
* IDE s/ Development Tools: PyCharm
* Operating Systems: Windows, Linux
* Protocols: TCP/IP, HTTP/HTTPS, SOAP
* Deployment Tools: Heroku, Jenkins
* Web Service : REST –JSON, DRF (Django Rest Framework)

***Educational Qualification:-***

* B.Tech(ECE) – Dr. A.P.J. Abdul Kalam Technical University, formerly Uttar Pradesh Technical University (UPTU).

**Working Zone – Organization:-**

Currently working as **Associate** **Data Scientist** With Netcracker Technology,Punesince may 2018 to till date.

**Awards:-**

* Performance of Star performer award Received from Netcracker for good performances .
* Appreciation for E2E Delivery Received from Client Screwfix .

**Project**

**Project – 1**

* **Project Name : Intelligent Metrics**
* **Vertical :** E- commerce , Payments
* **Client : Screwfix,United Kingdom**
* **Technology & Tool :** Python, Pyspark, Spark SQL, Plotly, Dash , Docker,ML ,Statistical Analysis
* **Roles:** Data scientist

**Detail Project Overview and Workflow :**

**Screw fix** Intelligent Metrics 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 under performing, as well as what you can do to optimize sales. Product matrices 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 Repository collection ,Customer Behavior Analytics ,Marketing Performance analysis.
* Conversion Optimization ,Product analytics.
* **Task Handled :**
* Implemented Data Exploration to analyze patterns and to select features using Python SciPy.
* Build customer segmentation and product categorization using of NLTK, lemmatization and TF-IDF.
* 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.
* Created HBase tables to store various data formats of data coming from different portfolios,Worked on improving performance of existing Pig and Hive Queries.
* Structuring the Data Marts to store and organize the customer's data.
* 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 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.

**Project - 2**

* **Project Name :** Industrial LTE Routers
* **Vertical :** Telecom OSS – Provisioning & Activation [Inventory Management ]
* **Client :** Elisa Telecom , Finland
* **Technology & Tool :** Python ,Django REST , Oracle ,GIT , Rest – Web service ,Pandas ,Statistical Analysis
* **Role –** Data scientist

**Detail Project Overview and Workflow :**

Using high-performance industrial-grade wireless modules, industrial grade high stability high precision components, -35 ~ +75 º C extremely wide temperature design, easily adapt to the high-temperature cold working environment for unmanned systems to provide reliable group network with stable communication

These compact, high-density, advanced Gigabit and multi Gigabit Ethernet (GigE) platforms are designed for the most demanding of converged networks. Offering high-speed flexible uplinks, 200G stacking, industry leading 95W PoE, and high density 10G multi-gigabit ports ready for Wi-Fi 6, these platforms are the right choice for the next generation of enterprise switching networks.

Offering a design optimized for flexibility and scalability as well as low power consumption, the OmniSwitch 6560 is an outstanding edge solution. It uses the field-proven Alcatel-Lucent Operating System (AOS) to deliver highly available, secure, self-protective, easily managed and eco-friendly networks.

* **Task Handled:**
* Understand the FRS document and analyse the requirement,Analyzed system requirements specifications and also in client interaction during requirements specifications.
* 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 re-branding 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 life-cycle 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 - 3**

* **Project Name :** Profitability and expense analytics**.**
* **Vertical :** E- commerce , Payments
* **Client : Screwfix,United Kingdom**
* **Technology & Tool :** Python, Pyspark, Spark SQL, Plotly, Dash, Post Man Microsoft Azure, Docker,ML ,Statistical Analysis
* **Roles:** Data scientist

**Detail Project Overview and Workflow :**

Businesses need to clearly understand the importance of collecting and analyzing data. Business needs the right analysis helps endures extract the relevant information that suits your payment environment. From there, user can use that information operationally to improve the way your system performs. user need to know customer payment environment—that way user can get the historical data, and tie it back into the real time transaction flow and reporting.

* 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.
* Implemented Data Exploration to analyze patterns and to select features using Python SciPy.
* Created HBase tables to store various data formats of data coming from different portfolios.
* Coordinate and manage data analytics activities with stakeholders
* Perform multivariate analysis, predictive modeling, market basket analysis using sophisticated statistical techniques
* Interpret & translate analytic output into insights
* Determine the source of the data, coordinate extraction and acquisition.

**Project - 4**

* **Project Name :** Profitability and delay analytics**.**
* **Vertical :** Flight delay Analysis and Prediction.
* **Technology & Tool :** Python, Plotly, Docker,ML ,Statistical Analysis**,** seaborn, matplotlib, Data cleaning, delay analysis, data preprocessing.
* **Roles:**  Data scientist

**Detail Project Overview and Workflow :**

Air travel is becoming increasing complex with multiple variables impacting the same. The flight delay is one of such variables impacts carrier, Airport and passenger and may result in significant commercial loss or reputation loss to all the stakeholders and thus huge cost on the economy. Thus, prediction of delay is crucial not only from view point of customer from time management perspective and carrier for retention of customer faith but also from Airport point of view for managing the traffic more efficiently to optimize the number of arriving flights by appropriate adjustment of schedules. The contribution of weather conditions has been identified to be the very important contributor of these delays.

**Personal Details:-**

Father’s Name : Late Shivyogi Yadav

Mother’s Name : Mrs. shiro devi

Gender : Male

Languages Known : English, Hindi.

Date of Birth : 18-Aug-1993

Date:-