**Kusheshwar**

**PROFESSIONAL SUMMARY:**

* A database professional with **5+ years** of progressive experience in **Data Analytics, Statistical Modeling, Visualization** and **Machine Learning**. Excellent capability in collaboration, quick learning and adaptation.
* Experience in **Data mining** with large datasets of Structured and Unstructured data, Data Acquisition, Data Validation, Predictive modeling, Data Visualization.
* Experience in integrating data, profiling, validating and data cleansing transformation and data visualization using **Python** and **R.**
* Experience in migration from heterogeneous sources including **Oracle** to **MS SQL Server**.
* Hands on experience in design, management and visualization of databases using **Oracle, MySQL, SQL Server**, and **MSBI**.
* Hands-on experience on **Python** and libraries like **NumPy, Pandas, Matplotlib, Seaborn, NLTK, Sci-Kit learn, SciPy**.
* Experience in extracted data from various database sources like **Oracle, MS SQL Server, DB2, MYSQL.**
* Strong programming expertise in **Python, Java, R, C++, C, Fortran** and strong in **Database SQL.**
* Good analytical and communication skills and ability to work independently with minimal supervision and perform as part of a team.
* Theoretical foundations and practical hands-on projects related to **(i) supervised learning (linear and logistic regression, boosted decision trees, Support Vector Machines, neural networks, NLP), (ii) unsupervised learning (clustering, dimensionality reduction, recommender systems), (iii) probability & statistics, experiment analysis, confidence intervals, A/B testing, (iv) algorithms and data structures**.
* Experience in dimensionality reduction using techniques like **PCA and LDA**.
* Intensive hands-on Boot camp on **Data Analytics** course spanning from Statistics to Programming including data engineering, data visualization, machine learning and programming in **R, SQL**.
* Experience in data analytics, predictive analysis like Classification, Regression, Recommender Systems.
* Experience in Descriptive Analysis Problems like Frequent **Pattern Mining, Clustering, Outlier Detection**.
* Worked on Machine Learning algorithms like Classification and Regression with **KNN Model, Decision Tree Model, Naïve Bayes Model, Logistic Regression, SVM Model and Latent Factor Model**.
* Expertise and knowledge in **TensorFlow** to do machine learning/deep learning package in **python**.
* Good Exposure on **SAS analytics**.
* Good Exposure in deep learning with Tensor flow in **python**.
* Good knowledge in **MSBI**, **SSIS, SSRS, SSAS,** **Tableau, Power BI** for interactive data visualizations.
* Experience in both **Waterfall** (**Software Development Life Cycle -SDLC) and Agile Project Environment.**
* Experience understanding the code written in different programming languages **Java, HTML, CSS, C#, Python, Java Script** anddatabases like **Oracle and SQL Server.**
* Experience in **SQL** queries to setup/modify test date in **Oracle database.**

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Languages** | Python, R, Java, C, C++, Fortran, and SQL |
| **Databases** | SQL, Databases MSSQL-Server, MYSQL, DB2, MS Access, NoSQL, MongoDB, Cassandra. |
| **Packages** | Pandas, NumPy, Seaborn, SciPy, Matplotlib, Sci-kit-learn, ggplot2, caret, dplyr, Rweka, gmodels, RCurl, C50, twitter, NLP, Reshape2, rjson, plyr, |
| **Web Technologies** | JDBC, HTML5, DHTML and XML, CSS3 &6, Web Services, WSDL |
| **Data Modelling Tools** | Erwin r 9.6, 9.5, 9.1, 8.x, ER/Studio, MS Visio, SAP Power designer |
| **Reporting Tools** | MS Office (Word/Excel/Power Point/ Visio), Tableau, Crystal reports XI, Business Intelligence, SSRS, Business Objects 5.x/ 6.x, Cognos7.0/6.0. |
| **ETL Tools** | Informatica Power Centre, SSIS. |
| **Version Control Tools** | SVM, GitHub |
| **BI Tools** | SSRS, Tableau |
| **Operating System** | Windows variants, Linux, Unix, |

**PROFESSIONAL EXPERIENCE:**

**Yum! Brands – Louisville, KY Jul 2018 – Present**

**Data Analyst**

**Responsibilities:**

This project is about developing an algorithm that accurately predicts the demand of products among multiple classes based on the historical sales data available on multiple products. Further, the aim was to improve the profit by maintaining the right stock of products whose demand is high while avoiding the loss of maintaining unnecessary products.

* Gathering business requirement from client and approach formulation and design methodology to match client requirements
* Requirement gathering and Data analysis of different types of data.
* **Extraction by developing a pipeline using Amazon Redshift to retrieve the data from S3**
* Using Pandas data frames for working with data and manipulating data.
* Performed Data Cleaning, features scaling, features engineering using R, Pandas and NumPy packages in python and build models using deep learning frameworks.
* Implemented end-to-end systems for Data Analytics, Data Automation and integrated with custom visualization tools using R and Mongo DB.
* Replacement of missing data and perform a proper EDA to understand the time series data
* Involved working on different databases like XML, Oracle and SQL of different platforms etc.
* Created SQL tables with referential integrity and developed queries using SQL, SQL PLUS and PL/SQL.
* Designed and developed Use Case, Activity Diagrams, Sequence Diagrams, OOD (Object oriented Design) using UML and Visio.
* Identifying and executing process improvements, hands-on in various technologies such as Oracle, Informatica, and BusinessObjects.
* Involved working in **Data science using Python 2.x/3.x on different data transformation** and validation techniques like Dimensionality reduction using Principal Component Analysis (PCA) and A/B testing, Factor Analysis, testing and validation using ROC plot, K- fold cross validation, statistical significance testing.
* Check the existence of trend and seasonality in the data
* Improvise the model until achieving best accuracy
* Generated visualizations using Tableau and R-Shiny to present the findings
* Also worked with several R packages including knitr, dplyr, SparkR, CausalInfer, spacetime.

**Environment:** Python 2.x/3.x, R, Linux, Spark, TensorFlow, Tableau SQL Server 2012, Microsoft Excel, MATLAB, SQL, Scikit-learn, Pandas, AWS(S3/Redshift), XML

**DOITT City of New York – NY Jan 2016 - Jun 2018**

**Data Analyst**

The New York City Department of Information Technology & Telecommunications (DoITT) provides for the sustained, efficient, and effective delivery of IT services, infrastructure, and telecommunications to enhance service delivery to the City's residents, businesses, and Visitors.      
  
As the City’s technology leader, we are responsible for maintaining the core IT infrastructure and systems that touch every aspect of city life from public safety to human services, from education to economic development crossing the full spectrum of governmental operations.  
DoITT.

**Responsibilities:**

* Involved in Analysis and Marketing Team to make business decision.
* Worked with large structured and unstructured datasets.
* Extracting the source data from Oracle tables, **MS SQL Server**, sequential files and excel sheets.
* Understand Datasets with Descriptive Statistics.
* Performed Data Cleaning, features scaling, features engineering using Python
* Replacement of missing data and perform a proper EDA to understand the time series data
* Involved working on different databases like XML, Oracle and SQL of different platforms etc.
* Design rich data Visualizations to communicate complex ideas to business leaders and executives.
* Extraction of data from different **API**s using **RESTful API** services.
* Documented methodology, data reports and machine learning model results and communicated with the project team.
* Setup storage and data analysis tools in Amazon Web services cloud computing infrastructure.
* Used **SVN** to commit the Changes into the main **EMM application trunk**.
* Understanding and implementation of **text mining concepts, graph processing** and **semi structured and unstructured data processing**.
* Analyzed the partitioned and bucketed data and compute various metrics for reporting.
* Involved in loading data from **MYSQL,** **RDBMS** and **web logs**.
* Extracted data from **Twitter** using **Java and Twitter API**. Parsed **JSON** formatted twitter data and uploaded to database.
* Launching **Amazon EC2 Cloud Instances** using **Amazon Images (Linux/ Ubuntu)** and Configuring launched instances with respect to specific applications to improve **robustness**.

**Environment**: Python, Pandas, Numpy, Matplotlib, MS Excel, SQL, Spark (MLlib, PySpark), Tableau, AWS, MongoDB, exploratory analysis, feature engineering, SVM, JSON, XML and MapReduce.

**Amnil Technologies – Nepal Apr 2014 - Oct 2015**

**Data Analyst**

**Responsibilities:**

* Analyze business information requirements and model class diagrams and/or conceptual domain models.
* Managed the project requirements, documents and use cases by IBM Rational Requisite Pro.
* Assisted in building an Integrated LogicalData Design, propose physical database design for building the DataMart.
* Gather & Review Customer Information Requirements for OLAP and building the data mart.
* Worked with SQL statements, import and export data, and generate reports in Teradata.
* Calculated and analyzed claims data for provider incentive and supplemental benefit analysis using Microsoft Access and Oracle SQL.
* Designed email marketing campaigns and also created responsive web forms that saved data into a database using Python/ Django Framework.
* Worked on Hadoop single node, Apache spark, Hive installations.
* Installation, Configuration, Integration, Tuning, Backup, Crash recovery, Upgrades, Patching,
* Monitoring System Performance, System and Network Security and Troubleshooting of Linux/Unix Servers.
* Implemented public segmentation using unsupervised machine learning algorithms by implementing the k-means algorithm using R.

**Environment**: SQL Server 2008, R2/2005 Enterprise, SSRS, SSIS, Crystal Reports, Windows Enterprise Server 2000, DTS, SQL Profiler, and Query Analyzer.

**EDUCATION:** Available upon request

**REFERENCES:** Available upon request