



# Pranay Bhatkar

Data Scientist – IBM – <https://www.ibm.com/>–

## My Contact

✉ [Pranaysbhatkar@gmail.com](mailto:Pranaysbhatkar@gmail.com) ☎  
+919730157253  
📍 Senapati Bapat road  
Gokhale Nagar Pune, Maharashtra  
Pin code: 411016

## Education

- University: Amravati University Graduation
- University: Mumbai
- Master

## Skill

- Regression, classification, and cluster analysis
- Python, SQL, Pyspark
- Tableau, Plotly, Seaborn, Matplotlib
- Scikit Learn, Keras, Tensorflow AWS,
- Ec2, S3
- Statistical Analysis

## More about me

- Father name : Sanjay Bhatkar Mother
- name: Mangla Karmankar
- Current address: District: Senapati Bapat Road  
Gokhale Nagar Pune, Maharashtra  
Pin code: 411016
- Marital status: Unmarried
- Gender: Female
- Language: Hindi, Marathi, English

## Career Objective

A challenging carrier as Data Scientist / Web Developer where my Python Machine Learning /AI/Data Intelligence skills can be effectively used and upgraded. 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. Looking for a challenging career in the field of IT-Software Industry especially for roles such as 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 my all-rounder development is encouraged.

## Experience

Currently working with client Worldpay, United Kingdom Data Scientist/ML Engineer for Virtual Terminal - — ml implementation on E-commerce payments system

**Associated Technology :** AWS, Docker, Power BI, ETL(Data Migration), Linux, IoT, Android

**Work Experience with clients:** Aussie Broadband Australia

**System/Vertical Worked with :** Payment analytics | Telecom OSS – Provisioning and Activation | E-commerce, Payments

## 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 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

Strong Experience Areas : Python | SQL | Unix Data Analytics –ETL |Machine Learning Algorithm | Django Web service-REST

Technical Qualification Headlines

Deep Learning | Python Data Science |Big Data | Machine Learning | Django REST | Mongo DB |PySpark| ETL / Data warehousing | Business Intelligence | Unix |

## Professional work Experience

### Data Science/Machine learning/ETL/BL 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 Hadoop, Hive and 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 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. 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. 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.
- 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

- |                            |  |
|----------------------------|--|
| • Frameworks:              | Scikit-learn, Tensorflow, PyTorch                                      |
| • Programming Languages:   | Python, Pyspark, SQL   |
| • Data manipulation:       | Pandas, Numpy  |
| • Version Control:         | Git, GitHub.   |
| • Cloud Computing:         | AWS EC2, S3, Azure   |
| • Analytic Tools:          | PySpark.   |
| • Databases:               | Oracle 10g, My SQL, MongoDB  |
| • Visualization Tools:     | Tableau, Matplotlib, Seaborn   |
| • Web Scraping Library:    | Beautiful Soup 4   |
| • IDEs/ Development Tools: | PyCharm, Anaconda – Jupyter Nb, Atom, Spyder, Microsoft Visual Studio. |
| • Operating Systems:       | Windows, Red Hat Linux.  |
| • Deployment Tools:        | Heroku, Jenkins.   |

### Working zone Organization

Currently working as **Python Data Scientist** With IBM, <https://www.ibm.com/> Pune since August 2020 to till date

## Projects

### Projects Sequence 1

Project Name: Virtual Terminal  
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 integrates 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 Handled:

- 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.
- 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.
- Coordinated and managed Data Analytics activities with stakeholders
- 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, and coordinate extraction and acquisition.
- 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.
- Assess data quality, identify gaps in the data and eliminate irrelevant data.

## Projects Sequence 2

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  
Roles: 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 the 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, Analyse system requirements specifications and also in client interaction during requirements specifications.
- Worked on the 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.
- Writing backend programming in Python, Performed Unit testing, Integration Testing, GUI and web application testing using Rspec.
- Worked on Jenkins continuous integration tool for deployment of the project.
- Worked on changes to open stack and AWS to accommodate large-scale data centre deployment.
- Worked on Jenkins continuous integration tool for deployment of the project, worked on changes to open stack and AWS to accommodate large-scale data centre 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 new features like support for internalization, support for customer service, 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 clipboard to have the new features as per the client's requirements, Used many regular expressions 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 a team of developers on Python applications for RISK management.
- Taken part in the entire lifecycle of the projects including Design, Development, 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.

**Declaration: -**

I hereby declare that the above-mentioned information is correct to the best of my knowledge

Date:

Location:

Regards,

Pranay Sanjay Bhatkar

