




# Harshita Ghate

Python Data Scientist, Tech Mahindra ([www.Techmahindra.com](http://www.Techmahindra.com))

 Hinjawadi- Pune, Maharashtra

 +91 91113-40181

02-may-1998



[harshita.ghate242@gmail.com](mailto:harshita.ghate242@gmail.com)



## Objective

Desirous for the role of Software Developer to design complex requirements of clients and customers into hi-tech and reliable software. Excellent knowledge of software design and development especially data structures and product-oriented designs and Overall 3+ years of work experience as a software developer.

To find a Position in a cutting-edge Organization that will utilize either or all of my **Python /Data Science/Machine Learning /Artificial Intelligence /Django REST API Web Development** Experience and skills.

## Technologies

Python | Data Science | Machine Learning | Django | PySpark | MongoDB | NumPY | Pandas | SQL | Unix | DevOps | GIT | Data Warehousing | Power BI | Tableau.

## Education

- BE ( Information Technology)
- University- Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal.

## Awards

- Recognized as a top performer at **Tech Mahindra** with the Performance Star award for consistently delivering exceptional results.
- Received appreciation for E2E Delivery from Client **Elisa**

## Certifications

Python Data Science

## Hobbies

- Traveling
- Sketching
- Playing Badminton
- Listening music

## Professional Experience

**Python Data Scientist - Tech Mahindra** ([www.Techmahindra.com](http://www.Techmahindra.com))

2021 Jan - Present

Actively contributing as a **Python Data Scientist** at Intelligent Metrics, collaborating with esteemed client **Payment Express** in Australia to drive advanced data-driven solutions.

### CORE TECHNOLOGY SPECIALIZATION:

**Python | Django REST API | Data Science | Machine Learning | PySpark | Data warehouse /ETL | Business Intelligence | SQL | Unix | AWS-Cloud | MongoDB.**

### VERTICAL /DOMAIN/SYSTEM SPECIALIZATION

**Payment Gateway | Supply Chain | Telecom OSS – Provisioning & Activation**

### Tools and Associated Technologies Used:

- ❖ Important Core Package worked with Django Rest Framework | Django-cors-headers | Django-import-export | Django-filter | Sentry-sdk | Django Storages | Django-extensions
- ❖ Core Data science Libraries Used : Matplotlib | NumPy | Pandas | Scikit Learn | Tensor flow | Seaborn | Keras | Scipy
- ❖ Web Framework involvement: Django REST
- ❖ UI Application Programming Language - HTML, CSS, JavaScript, JQuery and Bootstrap.
- ❖ Web Scraping Library: BeautifulSoup 4
- ❖ Unstructured Database Used – Mongo DB
- ❖ BI Tool used - Power BI, Parsing XML - **XML Schema**

**Process worked with:** DevOps + Agile + V- Model

**Familiar with Software Programming Language**

Python | Java | C++ | JavaScript | CSS | HTML | Bootstrap | JQuery

## PROFESSIONAL WORK EXPERIENCE

- Highly experienced Data Scientist with **3+ years'** experience in **Data Extraction, Data Modelling, Data Wrangling, Statistical Modeling, Data Mining, Machine Learning, and Data Visualization.**
- Expertise in transforming business resources and requirements into manageable **data formats and analytical models, designing** algorithms, building models, and developing data mining and reporting solutions that scale across a massive volume of structured and unstructured data.
- Worked in the core development team of 20+ members, for making a data pipeline related to data quality and data mesh.
- Proficient in managing the entire data science project life cycle and actively involved in all the phases of the project life cycle including **data acquisition, data cleaning, data engineering, features scaling, features engineering, statistical modeling,** testing and validation, and data visualization.
- Proficient in Machine Learning algorithms and **Predictive Modeling including Regression Models, Decision Trees, Random Forests, Sentiment Analysis, Naïve Bayes Classifiers, SVM, and Ensemble Models.**
- Collaborated with data engineers and operation team to implement ETL process, wrote and optimized SQL queries to perform data extraction to fit the analytical requirements.
- Explored and analyzed the customer-specific features by using Spark SQL.
- Performed univariate and multivariate analysis of the data to identify any underlying pattern in the data and associations between the variables.
- Performed data imputation using the Scikit-learn package in Python.
- Worked on data cleaning and ensured data quality, consistency, and integrity using Pandas and NumPy.
- Used SSIS to create ETL packages to Validate, **Extract, Transform, and Load data into the Data Warehouse and Data Mart.**
- Developed and implemented predictive models using machine learning algorithms such as linear regression, classification, multivariate regression, Naive Bayes, Random Forests, K-means clustering, KNN, PCA, and regularization for data analysis.
- Wrote complex Spark SQL queries for data analysis to meet business requirements.
- Perform **data manipulation, data preparation, normalization, and predictive modeling. Improve efficiency** and accuracy by evaluating models in Python
- Familiar with various **ML algorithms such as 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.
- Used various types of Core Data Science/ML Libraries /Packages such as PySpark, Pandas, Numpy, Seaborn, Matplotlib, Scikit Learn, Keras, Tensorflow, Scipy, Statsmodels, Plotly
- This project was focused on customer segmentation based on machine learning and statistical modeling efforts including building predictive models and generating data products to support customer segmentation
- Designed the front-end and back-end of the application utilizing Python on Django Web Framework.
- Responsible for creating website functionality with JavaScript, HTML, and CSS.
- For the development of the user interface of the website using HTML, CSS, JavaScript, and AJAX
- Used various types of Django libraries such as **Django-rest-framework ,Django-cors-headers , Django-debug-toolbar , Django-extensions ,Sentry-sdk , Django-allauth , Django-filter , Django-import-export**
- Experience in developing views and templates with Python and Django's view controller and templating language to create a user-friendly website interface.
- Experience in the required XML Schema documents and implemented the framework for parsing XML documents.
- Modify the existing Python/Django modules to deliver certain formats of data.
- Written Python scripts to parse **JSON documents and load the data in the database.**
- Performed data analysis using Pandas as API to convert data into tabular format
- Used Google APIs and create **visualizations such as pie charts, and donut charts and displayed them in the web application**
- For the development of the web applications utilized CSS and Bootstrap.
- To share the information across the applications used pickle/unpickle in Python.
- Utilized Python libraries like Numpy and matplotlib for generating graphical reports.
- Build SQL queries for performing various **CRUD operations like create, update, read, and delete.**
- Experienced with GIT version control and deployed the project to Heroku.

## TECHNICAL SKILLS

- ❖ Core Libraries worked: Matplotlib | NumPy | Pandas | Scikit Learn | Seaborn
- ❖ Process worked with: DevOps + Agile + V-Model
- ❖ Machine Learning: Logistics Regression, Linear Regression, K-means
- ❖ Analytic/Predictive modeling Tools: Microsoft Visual Studio, Jupyter Notebook, Anaconda, Google Collab
- ❖ Visualization Tools: Power BI, Python – Matplotlib
- ❖ ETL: Pyetl
- ❖ Programming Language - Python, SQL, Java
- ❖ Tools: SQL Developer
- ❖ Web Designing :- HTML, CSS, JavaScript
- ❖ Web server :- WSGI, Apache.
- ❖ Database Skills :- MySQL, Oracle, SSIS.
- ❖ Versioning Tools :- GIT
- ❖ Build Tools :- PyBuilder
- ❖ RDBMS :- Oracle 10 g
- ❖ Unstructured database :- Mongo DB
- ❖ Cloud Computing :- AWS EC2, S3
- ❖ Continuous Integration Tool :- Jenkins
- ❖ Data Modeling Tools :- Oracle Designer,
- ❖ Framework Software :- Django, Flask
- ❖ Defect Reporting Tool :- HPALM
- ❖ Deployment Tool :- MobaXterm
- ❖ Web Service :- REST –JSON, DRF (Django Rest Framework)
- ❖ Development Tools :- PyCharm, Notepad ++
- ❖ OS and Networking :- Windows 7, Windows 10, Windows 11, Ubuntu/Linux 18.2
- ❖ Certified Python Programme Professional:–CPPP.

## WORK EXPERIENCE DETAILS

Currently working as a **Python Data Scientist** with Tech Mahindra-<https://www.techmahindra.com/>,Pune since January 2021 to present.

## PROJECTS DETAILS

### [Project Sequence 1](#)

- **Project Name** : Intelligent Metrics
- **Vertical** : Payment Analytics
- **Client** : Payment Express, Australia
- **Technology & Tools** : Python, PySpark, Spark SQL, Plotly, Dash, Post Man, AWS, Docker, ETL, ML, Statistical Analysis
- **Role** : Python Data Scientist

### **Detail Project Overview and Workflow:**

Payment analytics software tracks online payments for e-commerce or subscription-based businesses. This type of software consolidates payment data from multiple sources to monitor customer payments. Payment analytics software is used mostly by accounting professionals to manage revenues and by sales managers to track the efficiency of their sales strategies. Payment analytics solutions can be delivered as a standalone solution or as a feature of payment gateway software. Subscription management software and e-commerce platforms also include

features for payment analytics.

#### Advantages and Core Functionalities:

Analyze Conversion Rates, where discover which currencies, price points, and locations have lower-than-normal payment conversion rates to hone your market strategy. Also Slice & Dice Sales Intelligence.

Enhanced sales strategies begin here. Track sales by currency, product, region, payment type, and more also we Reconcile in a Snap.

Close the books with ease. All payment data flows in and out of our integrated system for simplified reconciliation.

Eliminate Bad Transactions. High decline rates are bad for business. Cleanse Reports reveal where fraud, chargebacks, and refunds are hampering your revenue growth.

#### Task Handled:

- Part of the innovation team of 20+ members to explore new technology.
- Implemented Data Exploration to analyze patterns and select features using Python SciPy.
- Evaluated business requirements and prepared detailed specifications that followed project guidelines required to develop programs.
- Communicated and presented default customer profiles along with reports using Python and Tableau, analytical results, and strategic implications to senior management for strategic decision 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%
- Performed Data Enrichment jobs to deal with missing values, normalize data, and select features.
- Created reports and dashboards, by Power BI, to explain and communicate data insights, significant features, model scores, and performance of the new Recommendation system to both technical and business teams.
- Used Git 2. x for version control with the Data Engineer team and Data Scientists colleagues, used Agile methodology and SCRUM process for project development.
- Running SQL scripts, creating indexes, stored procedures for data mapping, and maintaining quality.
- Prepared Scripts in Python and Shell for Automation of administration tasks.
- 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.
- Worked on changes to open stack and AWS to accommodate large-scale data center deployment.
- Built models using Statistical techniques like Bayesian HMM and Machine Learning classification models like Linear Regression, Logistic Regression, and Random Forest.
- A highly immersive Data Science program involving Data Manipulation & Visualization, Web Scraping, Machine Learning, Python programming, SQL, GIT, Unix Commands, MongoDB, and Hadoop.
- Used pandas, NumPy, seaborn, Scipy, matplotlib, and scikit-learn, in Python for developing various machine learning algorithms.

#### Project Sequence 2

- **Project Name** : Maintenance Intelligent Grid
- **Client** : Elisa Telecom, Finland
- **System/Component** : Telecom, Supply Chain
- **Technology** : Python, Django, MySQL, GIT, Rest – Web service, Pandas, ETL, Statistical Analysis
- **Roles** : Python Data analyst / Python Development

#### Detail Project Overview and Workflow:

Grid analytics is a solution for end users to check the availability of the product in the Warehouse. Market Basket Analysis is a technique that identifies the strength of association between a pair/group of products that are purchased together. In simpler words, it's based on the concept that if a customer buys a product, they are more or less likely to buy another related product. Market Basket Analysis also uses machine learning/deep learning algorithms, like product recommendation engines.

SERP (Search Engine Result Page) analytics help users measure how customers interact with the search results on user SERP, which gives insight into how customer search engines are functioning. This gives users important data points about where to improve the search experience.

Association rule learning is a machine learning method that is used for discovering associations between products. It does not consider the order of the purchase of products. Product Analytics will show users the most popular products and availability along with a detailed insight into how often each product is purchased and how much revenue it generates.

Users can also see the key terms shoppers are searching and which product they end up buying, so users can update the product naming to better fit for customers.

#### **Task Handled:**

- Collaborated with a team of 10+ members to analyze system solutions for compliance.
- Involved in requirement analysis, design, estimation, and testing of the assigned tasks in open stack with BA.
- Used Git 2.x for version control with Data Engineer team and Data Scientists colleagues,
- Used Agile methodology and SCRUM process for project developing.
- Interpreting data, and analyzing results using statistical techniques.
- Developing and implementing data analyses, data collection systems, and other strategies that optimize statistical efficiency and quality.
- Design Python script concerning ensure functionalities meet customer requirements.
- Extensive experience in using Python packages such as NumPy, SciPy, Pandas, Beautiful Soap, Pickle, and OS.
- Created user interface for a website using Python, HTML5, CSS, JSON, and jQuery. Used CSS bootstrap framework for developing web applications.
- Acquiring data from primary or secondary data sources and maintaining databases.
- Work with stakeholders to determine how to use business data for valuable business solutions.
- Search for ways to get new data sources and assess their accuracy.
- Browse and analyze enterprise databases to simplify and improve product development, marketing techniques, and business processes.
- Create custom data models and algorithms.
- 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.
- Use predictive models to improve customer experience, ad targeting, revenue generation, and more.
- Develop the organization's test model quality and unit testing.
- Coordinate with various technical/functional teams to implement models and monitor results.
- Develop processes, techniques, and tools to analyze and monitor model performance while ensuring data accuracy.

#### **Project Sequence 3**

- **Project Name** : Industrial 4G LTE Routers
- **Client** : Elisa Telecom, Finland
- **System/Component** : Telecom OSS Provisioning& Activation CRM
- **Technology** : Python, Django, Oracle, GIT, SVN, Rest – Web service, Pandas, ETL, Statistical Analysis
- **Roles** : Python Data analyst / Python Development

#### **Detail Project Overview and Workflow:**

3X36 Industrial 4G Router uses a high-speed 4G module, and strong wireless communication with major networks, including TD-LTE, FDD-LTE, TD-SCDMA, WCDMA, EVDO, CDMA1X, GPRS/EDGE.

Using high-performance industrial-grade wireless modules, industrial grade high stability high precision components, -35 ~ +75 ° 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 comprehensive security features, they are well-equipped to secure the network edge. These high-performance, versatile LAN switches can be positioned.

At the edge of mid- to large-sized converged enterprise networks. At the aggregation layer, In a small enterprise network core, In the data center for GigE server connectivity and SDN applications. The Alcatel-Lucent Omni Switch 6560 Stackable Multi-Gigabit Ethernet LAN value switch family is an industry-leading campus access solution for enterprise networks. With multi-gigabit ports for high-speed IEEE 802.11 ac devices, 10 GigE uplinks, and 20 GigE stacking, the Omni Switch 6560 is the right solution for your next-generation network.

Offering a design optimized for flexibility and scalability as well as low power consumption, the Omni Switch 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:**

- Experience in developing entire frontend and backend modules using Python on Django Web Framework.
- Experience in working at various phases of projects such as analysis, design, development, and testing.
- Using the Django Framework model, implemented MVC architecture and developed web applications with superb interfaces.
- Created user interface of website using Python, HTML5, CSS, JSON and JQuery. Used CSS bootstrap framework for developing web applications.
- Extensive experience in using Python packages such as NumPy, SciPy, Pandas, Pickle, and OS.
- Preparing Low-level Design of Applications, to take part in software development activities.
- Conduct software analysis, programming, Unit and White box testing, and debugging.
- Identifying production and non-production application issues, ensuring designs comply with specifications, and Transforming requirements into stipulations.
- Managing Python application development, Developing, testing, implementing, and maintaining application software.
- Validated already developed Python reports. Fixed the identified bugs and redeployed the same.
- Develop technical designs for application development.
- Develop application code for Python programs, involved in client interaction to sort out the Requirement issue.
- Worked and extracted data from various database sources like Oracle, SQL Server, and DB2, regularly accessing the JIRA tool and other internal issue trackers for the Project development.
- Create the Report and share it with top management.

## PERSONAL DETAILS

- ❖ Father's name: Satish Ghatе
- ❖ Mother's name: Shalini Ghatе
- ❖ Permanent Address: Kanadiya Road, Indore, Madhya Pradesh
- ❖ DOB: 02-May-1998
- ❖ Marital Status: Unmarried
- ❖ Gender: Female
- ❖ Language: English, Hindi & Marathi
- ❖ Passport Availability: Yes

**Declaration: -**

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

Regards,  
Harshita Ghatе