

Shashwat Siddhant

Master of Data Science & Analytics
Royal Holloway, University of London

Phone: +44 7745589611
Email: shashwat007@gmail.com
Github: github.com/Siddhant08
LinkedIn: uk.linkedin.com/in/shashwat007

Personal Summary

Data Scientist with focus on analytics and development. With professional experience in Python for development and data science, Database management using SQL and relational databases, Automation using Selenium, Siddhant is proficient in working with data sets of all sizes, building Machine Learning models and drawing insights to add value to business.

Education

2019-20	Royal Holloway, University of London Master of Data Science & Analytics
2013-17	Manipal Institute of Technology, Manipal Bachelor of Technology, Computer Science

Work Experience

May'21-Present	Data Scientist – Vave MGA at Canopus Managing Agents Ltd, London, United Kingdom <ul style="list-style-type: none">End-to-end automated ETL of claims bordereau data from multiple Product Partners using Python and MS SQL Server.Developed application on Qlik Sense for tracking claims progression, portfolio analysis and catastrophe impact using dashboards with intuitive visualizations and KPI's.Built a catastrophe classification model using TF-IDF algorithm for vectorization of text and Random Forest classifier to tag insurance claims based on their loss descriptions. Achieved an F-score of 0.91 and an overall accuracy of 0.92.Worked on insurance risks data to compute the company's propensity to bind it using a data set with 50+ variables and 10M records through Generalized Linear Modelling. Competitor analysis done by using their next best premium bids for a risk.Built a machine learning regressor using Decision Trees in Pyspark on Databricks to predict loss ratios on claims using flood insurance data of 60M records to identify pockets of areas where the company's business has potential to grow. Insights used for altering the pricing algorithm for policies.Linked policies to claims using Fuzzy logic technique in Python on various fields from claims bordereau and exposures data to track the policy portfolio across different risk locations. Achieved ~98.5% match.Worked with Third Party Administrators to set up a process to for reconciling claims and risks data. Automated the data comparison and report generation process to aid the TPA and Vave in reconciliation.
Jun'20 – Sep'20	Data Scientist (Part-time) – In Diverse Company, London, United Kingdom <ul style="list-style-type: none">Developed a virtual assistant for the users on the company's learning platform in the form of a chat bot that identifies the intent behind a user's query about the platform and gives an appropriate response. The chat bot model was built as an Artificial Neural Network using Tensorflow in Python.
Aug'17 – Jun'19	Data Scientist – SG Analytics, Pune, India <ul style="list-style-type: none">Developed web crawlers built in Python using Selenium and BeautifulSoup to scrape social media data.Created dashboards for data reports using Tableau and Amazon Redshift database.ETL of social media platforms data using Python, MS-Excel, and Amazon Redshift DB.Sentiment Analysis of Instagram comments and tweets to understand the audience's reaction to newly launched shows using the libraries NLTK and SpaCy in Python. Achieved a recall score of 0.82.Developed an algorithm in Python to identify flow of rap artists by studying the lyrics for the biggest media and entertainment firm in the US.Built a classification model using Random Forests to analyze and predict the behavior of a wafer for fabrication for a firm dealing with semiconductor devices. Optimized the production process by bringing down the costs involved in testing phase by 24%. Successfully communicated with the client and led a team of 3 people for this project to achieve the targets before the deadline.
Dec'16 – May'17	Data Analyst Intern – FanZoo Sports Pvt Ltd, New Delhi, India <ul style="list-style-type: none">Auto pick feature for fantasy gaming– Greedy algorithm(Knapsack) implementation in Python using sports data set. The feature was used by 61% of the customer base.Twitter AI Bot – Using Twitter API's and Microsoft NLP Bot framework to automate marketing of product which increased the website traffic by 250%.Fantasy Player Projections – Prediction model using regression analysis to predict a player's performance. Achieved an r2 value of 0.86 on the model.Social Media Marketing – Content development and managing social media accounts.

Jun'16 - Jul'16

Summer Research Intern, Indian Institute of Science, Bangalore, India

- Set up a transmission system using **Python** between **IoT device**(R-pi model 3b) and server system with database connection.
- Analysis and calibration of sensor data using **Artificial Neural Network** (ANN) for a breath analyzer for diseases.

Technical Skills

Languages	Python, C++, C, MATLAB, R, PySpark
Data Science	Pandas, Scikit-learn, Scipy, Tensorflow, Matplotlib, Selenium, Qlik Sense, Tableau, NLTK, SpaCy, MS-Excel, MS-SQL
Operating Systems	Android, Linux, Windows

Projects

- Worked on churn prediction model using **Logistic Regression**, **Random Forest** and **customer segmentation** for personalized marketing on a customer data set for a telco company under Dr. Carol Hargreaves, National University Singapore.
- **Sentiment Analysis** – Tweet fetching based on keywords and sentiment analysis of text in English, Dutch, French, Spanish, and German – as an international freelance consultant for a client in the Netherlands.
- **Credit risk analysis** of bank customers for loan approval prediction and achieved a recall score of 0.89 using **Random Forest** classifier.
- Movie Suggestion Application for Windows using C# and IMDB API as part of Code Fun Do Hackathon organized by Microsoft.
- Online ticket booking system using ASP.NET for UI and C# for backend.
- Privacy Preserving Data Mining using Python, **Apriori algorithm** and MD5.

Publications and Certifications

- Cyber security and its ten domains on Coursera by University System of Georgia.
- Co-authored a paper with fellow intern and senior Professor on '**Graphical analysis of Real-Time Data from IoT**' published in International Journal of Innovation and Scientific Research (IJISR).
- Co-authored a paper on '**Security in NFC**', published in IJTRE Volume 3, Issue 2, October 2015, ISSN(Online): 2347-4718

Extra-curriculars and Personal Skills

- Participated in *Code.Fun.Do Hackathon* organized by Microsoft on college campus. **Worked in a team** and **communicated** with all members successfully to complete the project in the 24-hour deadline.
- **Presented and explained** a self-written research paper in front of experienced PhD professors and participants during TechTatva'15, MIT, Manipal Technical Fest.
- Secured first place in Revels '16 Men's Cricket Tournament as a part of the college department team.
- Workshops organizer for Revels'15, MIT Manipal Cultural Fest. Booked halls for workshops, **negotiated, and coordinated** with the people invited to give workshops, arranged for travel tickets, and stay.
- **Finisher** in London Summer 10k race at Regent's Park in June, 2022.