Suyash **Sathe**

Melbourne, VIC 3000

☐ (+61) 470 372 165 | ✉ suyash0106@outlook.com | ⌨ Suyash0106 | in SuyashLinked0106

Summary

A data-savvy IT professional with a strong programming background in Python, and over 2-years of experience in data science project work. Highly proficient in data skills including data wrangling, modelling, visualization and machine learning demonstrated by specialization in advanced data analytics from Monash University. Acquired comprehensive statistical and problem-solving capabilities working as a Mathematics Consultant for Cambridge University Press. Currently associated with Telstra as a Python Developer with strong focus on data processing and Robotic Process Automation technology.

Skills

Programming	Python, R, MySQL, Java (J2EE), HTML, CSS, Selenium, RPA
Data Visualization	Tableau, Power BI, Python, R visualization packages, Microsoft Excel
Data Analytics	Machine Learning, Data wrangling, Statistical Modelling, NLP, Classification, Regression, Clustering, Feature Engineering
Business Forecasting	Time-series analysis using ETS, ARIMA, Fourier, Multiple Regression models
Big Data Processing	Hadoop, MapReduce, Spark
Others	Git, Shell Scripting
Cloud Platforms	AWS, Azure, Google Cloud Platform

Work Experience

Synaptron/Telstra   *Melbourne, Australia*

PYTHON DEVELOPER   *Apr. 2020 - Present*

- Involved in development of BOH automation framework to process customer orders and automate other business processes using Python
- Configure API endpoints for data transfer between business applications and write business logic to transform and load the data into company's database
- Test, debug and fix code blocks in the application program and perform root cause analysis of automation failures
- Perform daily data queries to compile and validate customer orders along with monitoring and remediating failed orders
- Identify anomalies and missing information in the source data as per business requirements and report it to the source system
- Work closely with the SMEs to identify opportunities to automate manual tasks using Robotic Process Automation technology
- Responsible for development and testing of back-end modules and deliver complete solution within strict timelines

Find Sports   *Melbourne, Australia*

PROGRAMMER ANALYST   *Nov. 2019 - Apr. 2020*

- Designed and implemented data processing scripts using Python to maintain data accuracy and completeness
- Developed a Machine Learning model to classify product listings into the respective categories and bypass manual product categorization
- Analyzed product listings on marketplaces and improved e-store metrics
- Consolidated and maintained IT systems in the company while maintaining a smooth flow between the IT and the operations team
- Managed day-to-day software vendor relationship (e.g. troubleshooting software, communicating feedback)
- Collaborated with Find Sports' overseas team while delegating routine tasks and integrated solutions

Cambridge University Press   *Melbourne, Australia*

MATHEMATICS CONSULTANT   *Nov. 2018 - Dec. 2019*

- Created advanced mathematical and statistical content for publishing at Cambridge University Press
- Prototyped descriptive statistical content using R and Python to demonstrate the process of exploratory data analysis
- Designed fundamental random sampling, simulation and hypothesis testing solutions for the associated modules using R
- Developed material for inferential and predictive statistical analysis
- Evaluated analysis results to ensure content integrity and accuracy, and presented the results to the publishers

Accenture Solutions Pvt Ltd   *Pune, India*

APPLICATION DEVELOPMENT ANALYST   *Nov. 2016 - Jan. 2018*

- Designed, developed and tested new software applications for the banking and insurance industry to enhance user experience and access
- Implemented new change requests (CR) and established compatibility with third-party software products by designing and developing modules for modification and integration of services
- Orchestrated new Web Services (WS) to meet client requirements for functionality, scalability and performance
- Involved with Cloud solutions, Java projects, DevOps and performance testing activities, developing framework of the project's multichannel architecture
- Created a multi-node cluster of Linux machines on AWS and deployed Hadoop framework for distributed processing and storage of datasets

Education

Monash University *Melbourne, Australia*

M.S. in Data Science *Feb. 2018 - Nov. 2019*

- Excelled in machine learning and advanced data analytics coursework with a WAM of 71.3
- Recipient of Information Technology International Merit Scholarship
- Achieved High Distinction (HD) in Python Programming, Data Wrangling, Big Data Processing and Data Visualization

Savitribai Phule Pune University *Pune, India*

B.E. in Computer Engineering *Aug. 2012 - Jun. 2016*

- Graduated with a WAM of 71.0
- Published thesis on "Deploying and researching Hadoop algorithms on virtual machines and analyzing log files"

Projects

Domain Generating Algorithms (DGA) Detection using ML *Freelance*

Technology Stack: Python Pandas, Flask, scikit-learn, Jupyter Notebook *Mar. 2020*

- Implemented and delivered an end-to-end machine learning application to detect domain names generated from DGAs with 92% accuracy
- Developed 3 different ML models with Python and scikit-learn and analyzed their performances to select the best model for prediction
- Built a multi-class classification model to further predict the domain subclass
- Integrated the ML models with a web application using Python Flask and built a GUI for predicting malicious emails

Product Categorization *Find Sports*

Technology Stack: Python Pandas, Scikit-learn, Jupyter Notebook *Feb. 2020*

- Pre-processed the raw text and created a sparse matrix representation of the textual information
- Implemented Gaussian Mixture Models for classifying the dataset of text documents into their natural categories
- Used expectation maximization (EM) algorithm to model the classifier and determine its optimal model parameters
- Performed principal component analysis (PCA) on product category clusters and deployed the model to predict product categories

Predicting the Critical Temperature of a Superconductor *Monash University*

Technology Stack: R, R-Studio *Oct. 2019*

- Performed multivariate data analysis on the superconductor data containing 21,623 records and 170 features
- Applied various statistical techniques and feature engineering algorithms using R to obtain 20 most significant features for predicting the critical temperature
- Successfully implemented 3 different models for predicting the critical temperature including linear regression, random forest and gradient boosted model (GBM)

Delivery Fare Prediction Model *Monash University*

Technology Stack: Python, Scikit-learn, Tableau, Jupyter Notebook *Dec. 2018*

- Transformed and integrated datasets to make it suitable for applying graphical and non-graphical EDA methods
- Performed data cleansing using Python and Tableau to fix discrepancies in the dataset
- Implemented a multivariate linear regression model using Python to predict the delivery fare of a package

Resume Preprocessor *Monash University*

Technology Stack: Python, Scikit-learn, PyCharm *Jul. 2018*

- Implemented text parser using regular expressions to extract data from raw text files and stored it into XML and JSON files
- Analyzed the data and converted it into numerical representations suitable for recommender systems and information retrieval algorithms
- Built sparse representations for the resumes including word tokenization and vocabulary generation

Certifications

| 2019 | **GCP Big Data and Machine Learning Fundamentals,** Google Cloud | *Melbourne* |
| 2019 | **Pearson Test of English (PTE),** Score: 88/90. PTE Academic Technical Advisory Group | *Melbourne* |

Volunteer Work

| 2019 | **Volunteered as a Python tutor for post-graduate programming bootcamp,** Monash University | *Melbourne* |
| 2018 | **Mentored a group of 10 post-graduates at Peer Mentor Program,** Monash University | *Melbourne* |

Reference

Available upon request