SUMANSHU ARORA

Senior Data Scientist (Manager)

Phone: (+61) 481 707 280
DOB: 08 November 1989
Email: arora.sumanshu@gmail.com
LinkedIn:
https://linkedin.com/in/sumanshuarora

PROFILE

Data Science professional having over 10 years of experience in the data industry. A highly organized, motivated & passionate individual who thrives to solve complex business problems using his skills, knowledge and positive attitude with a proven record of building relationships, managing projects from concept to completion and decision-making ability in the most challenging situations. Always keen to learn, share feedback and seek feedback to grow and contribute to team's and organization's growth

EDUCATION

Bachelor of Technology
Information Technology
Lovely Professional University (India)

Graduated, May 2011
GPA 7.68

SKILL HIGHLIGHTS

- Team Management
- Coaching & Mentoring
- Project Solutioning
- DevOps

- Data Science
- Machine Learning
- Deep Learning
- Data Analytics

- Programming
- MLOps
- Natural Language Processing
- Data Visualization

TECHNICAL SKILLS

- Python
- SAS
- Django

- PowerBI
- Tableau
- QlikSense

- SQL
- NoSQL
- Teradata

MACHINE LEARNING FRAMEWORKS

- Pytorch
- Scikit-learn

- Tensorflow
- Statsmodels

- Keras
- Pycaret

OTHER SKILLS

- Google Cloud Platform
- AWS
- Openshift

- MLFlow
- REST
- Kafka

- Docker
- Github
- Jenkins

WORK EXPERIENCE

July 2019 – To Date
Senior Data Scientist (Manager)
ANZ Bank, Sydney

Machine Learning automation professional in the Institutional banking domain empowering internal operational customers by making their processes faster, accurate and cost efficient with help of statistical modeling and machine learning techniques.

- **Machine Learning Framework** – Machine learning framework based on micro service architecture paradigm wherein any new internal or external service can be added easily and gets compared against all others helping bank to only pay for the best services and customers to always get best results based on historical accuracy
- **Entity Extraction from unstructured documents** – Deep learning NER model framework in Pytorch which is capable of learning domain specific entities. Model developed is a hybrid model which uses capabilities of CNN, RNN and CRF to produce high quality predictions. Model evaluation metrics showed over 80% F1 score on two different domain problems
- **Cloud Introduction** – Leading the initiatives to introduce GCP and AWS in the team and also benchmark their performance against internal solutions
- **Text Templating Solution** In-house structured form parsing solution in Python for scenarios where labeled data is hard to get. The solution requires template setup with just one annotated image of scanned form and becomes capable of predicting values for on any new form instances
- **Email Classification –** A NLP based project using RNN model to categorize emails that are received on a common mailbox into 14 different classifications. These emails were indexed manually by a group of colleagues for further action before I built the model. The model achieved over 90% accuracy enabling business to implement automatic processing of emails based on calibration analysis of model results

July 2018 – July 2019
Analyst - Data Scientist, Data Analyst & Campaign
Westpac Bank, Sydney

An in-house consultant who supported various Data Science, Customer Analytics and Marketing teams to deliver different short-term projects. In just short span of time, I was been able to add value to the wider department through sound knowledge of various ML techniques, analytical skills, logical thinking and attention to detail. My responsibilities in this role included building targeting models, developing digital marketing campaigns, product performance and strategy analytics etc. Below are some key projects that I had undertaken in this role –

- **Behavior Adoption Prediction** – Created a logistic regression model using Python to predict probability of customers to adopt a certain behavior and targeted customers in middle deciles through a campaign to reward them for adopting and maintaining the behavior
- **Leads forecasting** – Created an ARIMA time series forecasting model to forecast the volume of leads to be generated over next 6 months in order to help CRM team efficiently manage resourcing and capacity issues

February 2018 – July 2018
Sr. Analyst – Insights & Analytics
Virgin Australia, Sydney

Data analytics to accelerate the business growth for loyalty program of Virgin Australia by analyzing the performance of existing partners along with tier profiling for potential partners. The role also involved deep dive analysis to answer ad-hoc business problems with help of data using SAS, dashboard creation for various partners using Tableau and stakeholder management to ensure continued growth with each partner.

August 2014 – January 2018
Manager - Remediation Strategy & Customer Analytics
Royal Bank of Scotland
Gurgaon, Haryana

Worked in various teams including Collection Strategy, Customer Experience & Remediation Strategy within my tenure of 3.5 years with the bank as Data Science & Data Analytics professional. These roles helped me enhance my business knowledge of different banking verticals. I exhibited multiple skills during my tenure including Classification Model Development, NLP, Data Analytics, Project Management, Mentoring, People Management etc.

Some highlights in terms of projects undertaken are -

- **Churn/Attrition Propensity Model**: Built logistic regression model using SAS to predict customer's propensity to leave the bank within 6 months after remediation
- **Complaints Categorization:** With help of NLP and Naive Bayes classification model, created a supervised classification model to categorize complaints in 10 custom categories which in turn helped bank to understand major pain points and route them to respective departments for actions like training, process changes etc.
- **Recoveries Customer Segmentation:** Built an unsupervised segmentation model to segment customers in Recoveries phase with respect to their I&E, Loan Value, Vulnerability, Account Balance etc. in order to understand the different bucket and help strategy team with tailored recovery approach
- **Mortgage Deals Analysis Kept Rate:** Trend analysis to find out the impact of Festive Season on UK Debt Collection World

June 2011 – April 2014
Functional System Analyst – HR Consulting
AON Hewitt
Gurgaon, Haryana

Data analysis, MI production & Insights Generation in retirement a.k.a Defined Benefits domain wherein I used to help the employees of our clients to effectively manage their retirement funds

Responsibilities –

- Data Analytics
- Develop monthly/weekly/daily reports for the retirement data of the employees
- Work closely with Testing team to provide right data for testing
- Tracking and fixing issues with the code in SAS jobs used by configuration team