

Aadish Chopra

aadishchopra16@gmail.com | 3128689603 | Redmond, WA | [Linkedin](#) | [Github](#) | [Kaggle](#)

SUMMARY

Experience of 5+ years as a data scientist, skilled in NLP & a strong background in statistics and computer programming.

EDUCATION

Master of Science in Management Information Systems	Dec 2017
University of Illinois at Chicago – Liautaud Graduate School of Business	
Bachelor of Technology in Electronics and Comm. - Rajasthan Technical University	May 2015

TECHNICAL SKILLS

Machine Learning: Regression, Random Forest, BERT, xgboost, SVM, RNN, Bayesian Networks, scikit-learn, PyTorch, TensorFlow, NLP, Transformers, Encoder-Decoder

Languages: Python, R, SQL, Bash, Linux Shell Scripting, VBA, PowerShell

Tools: SQL Developer, Oracle DV, Tableau, Power BI, Docker, Excel

Big Data: PySpark, Spark-SQL, Oracle Cloud, Azure, ETL

EXPERIENCE

Machine Learning Engineer: NIELSEN (for LucidaTech) June 2022 – January 2023

- Text classifications on Item descriptions for various consumers using BERT. See project section for more details
- Developed highly available distributed framework using Python, Spark clusters and Redis Cache for request routing (queue management) for various consumers to run predictive models. Using Microsoft Azure ecosystem (Kubernetes) for hosting docker containers
- Aided the support team for triaging monitoring issues by publishing and creating runbooks for the L2 support team
- Designed a schema for handling the requests and associated attributes in PostgreSQL

Data Scientist / Member Technical Staff (In Machine Learning): ORACLE Jan 2020- Mar 2022

- Developed a regression model to optimize Inventory Levels based on Lot Size, historical use, Item Type and other factors
- Created Bayesian Networks for Order Management module to inform business decisions like "How long should we have holds on Order?" and to predict Order Fulfillment
- Data Modelling and developed facts and dimensions (OLAP) for functional areas such as Inventory Management, Inventory Balances and Valuation, Sales Orders, Cost Accounting, Account Receivables, Shipment, NetSuite ERP
- Built data pipelines and did ETL using Spark. Transformed code base from ODI to spark which improved Performance by 6 hours
- Developed scripts for automating processes such as entity relationship diagrams, null handling, p-sql, ddl, table scans

Data Analytics Consultant: ZURICH INSURANCE (for DISYS) Nov 2018 – Dec 2019

- Development and automation of Claims Analysis Tool using PowerShell and SharePoint REST API. RECAT Tool allows clients to get a quick summary of their claims. Improved utilization of resources of upto 2 FTE and saved \$24,000 on server purchase.
- Developed tree based predictive models for grading real estate properties and for low premium machinery and crop insurance
- Built dashboards like Utilization Management and Demand Modelling using SSRS, FlexDashboard (R) and PowerBI.
- Created a document management system using .NET framework and RCurl. Reduced report building time by 300%

Data Scientist (INTERN): LOYOLA MEDICAL CENTER June 2017- Aug 2017

- Worked on Ordinary Differential Equations, stepwise regression for Hepatitis B and Hepatitis D virus, explaining the pharmacokinetics of drugs and virus interactions
- Applied Biphasic Models to fit viral kinetics to predict their treatment outcomes, time to cure and analyze accuracy

PROJECTS

Text Classification - NIELSEN (for LucidaTech)

- Formation of Item categories using Item Descriptors. Model accuracy was 87%
- Resulted in accurate forming of Item categories which helped the downstream teams (e.g., campaign team for promotions/advertising)
- Pre trained distil BERT uncased model was used and fine tuned for multiclass outputs