MANI KOHLI

Data Scientist

kohlimanis@gmail.com
+1 587-201-6885
https://www.linkedin.com/in/manikohli
https://mani-kohli.github.io/

SUMMARY

A motivated Data Scientist with unique industrial and educational experience in data science/analytics, software engineering and healthcare/medicine. Driven towards an ultimate goal of enhancing insights from data in society.

EDUCATION

2020 - 10 -

Certificate in Advanced Statistics
The Institute for Statistics Education | Arlington, VA

2018 - 09 -
2019 - 06

Master of Data Science (MDS)
University of British Columbia | Vancouver, BC
GPA 3.70

2017 - 07 -
2018 - 04

Certificate in Data Analytics, Big Data, and Predictive Analytics
Ryerson University | Toronto, ON
GPA 4.13

2013 - 01 -
2016 - 09

Medical Program
Caribbean Medical University | Curacao / Chicago, IL
GPA: 3.55 | Honors Grant Award (x2)

2005 - 09 -
2009 - 09

Medical Program
Poznan University of Medical Sciences | Poznan, Poland

2000 - 01 -
2002 - 04

Master of Science, Software Engineering
University of St. Thomas | St. Paul, MN
GPA 3.94 | President's Grant Award

1994 - 09 -
1997 - 09

BSc Major Psychology, Minor Mathematical Sciences
University of Alberta | Edmonton, AB

TECHNICAL SKILLS

Data Science: Descriptive Statistics & Probability, Communication & Argumentation, Data Wrangling, Algorithms & Data Structures, Data Visualization, Statistical Inference & Computation (incl. Bayesian), Database & Data Retrieval, Supervised Machine Learning, Workflows, Unsupervised Machine Learning, Feature & Model Selection, Collaborative Software Development, Spatial & Temporal Models (Time Series), Privacy Ethics & Security, Web & Cloud Computing, Experimentation & Causal Inference, Advanced Machine Learning (Natural Language Processing (NLP)).

Programming: R, Python, Java, VB, Regular Expression, JSON, SPSS, SAS

Big Data Analytics: Introduction: Big Data, Data Analytics: Basic Methods, Data Analytics: Advanced Methods, Big Data Tools (Hadoop, MapReduce, Hive, Pig, Spark, Hortonworks), Data Organization for Data Analysts (RDB & NoSQL).

Database Systems: RDBMS (SQL, PostgreSQL, SQLite, MySQL), NoSQL (MongoDB)

Visualization packages and Tools: ggplot2, Ploty, Matplotlib, Seaborn, Shiny, Tableau

Reporting and Integrated Environments: Jupyter Notebook, RMarkdown, Ioslides, RStudio, Linux, Docker

Cloud Platforms and ML Software: AWS, Azure, Google Collab, IBM Watson, Weka

Software Engineering: Software engineering processes/models, System analysis/design, Object Oriented (OO) Analysis & Design, OO Patterns, OO Database design, RDBMS design, Multimedia, SQA, Project Management, Technical writing

DATA SCIENCE PROJECTS (SELECTED)

Portfolio (GitHub): https://mani-kohli.github.io/

2020
- **Facial Expression Recognition:** Python, CNN, Deep Learning, Machine Learning, Flask
- **Fake News Categorization:** Python, NLP, LSTM, Deep Learning, AI, WordCloud

2019
- **Order Rate Prediction:** Python, R, PostgresSQL, Logistic Regression, Tableau, Pipeline, Docker
- **Image Enhancement:** Python, R, Testing/Test Coverage, Build
- **Mental Health in Tech Jobs:** Shiny, R

2018
- **Income Prediction:** R, Decision Trees, Workflow, Scripts

EXPERIENCE

2019 - 04 - 2019 - 06

Data Scientist
FreshPrep | Vancouver, BC
UBC MDS Capstone Project: Order Rate Prediction (Team collaboration with 4 members)
- **Exploratory data analysis** and **data wrangling** in R and PostgresSQL.
- **Engineered features** and applied supervised learning techniques to create a predictive model using **Logistic Regression** and **Random Forrest** methods in R and Python.
- Final **model** achieved with a 4.61 **Mean Absolute Percentage Error (MAPE).**
- Developed an interactive **Tableau** dashboard providing **insights for marketing strategies** based on past orders and model **predicted** future orders (for up to 3 weeks in advance).
- Packaged project into a self-contained **Docker** product.

2018 - 01 - 2018 - 04

Data Scientist
Ryerson University | Toronto, ON
Ryerson Certificate Capstone Project: Fire Fighter Injury Analysis
- **Exploratory data analysis** and **data wrangling** in R.
- **Feature Analysis** with **Random Forest** supervised method in R with **ROC** accuracy score of 97%.

2016 - 09 - 2017 - 02

Data Scientist Mentee
Ardaman Singh (Mentor), Principal, Axtria | New Jersey, NY / India
- Review, research and discussion of credit financial information of a leading financial institution.
- Researched potential areas for leveraging national citizens' database (Nepal, Asian Development Bank project) to improve citizen services and transparency.

2003 - 04 - 2005 - 09

Systems Analyst
Boston Scientific Corporation | Maple Grove, MN
- Contributed to the concurrent **strategic development, integration and maintenance** of multiple corporate clinical and supportive business software systems supporting FDA clinical data requirements within **dynamic deadlines.**
- Provided corporate global **education, training and solutions** for clinical software systems.
- **Lead groups** (2-3 coworkers) in corporate file, application and directory standardization project.

2002 - 04 - 2003 - 04

Systems Analyst
Arran Technology Inc. | Minneapolis, MN
- Contributed to the **design, development and maintenance** of various enterprise software and web application systems in small team environments within **rigid deadlines**.
- Successfully **project managed** by working closely with clients to provide data, statistical and database requirements of consumer website and web applications for clientele business analysis.

2001 - 09 - 2002 - 04

Software Engineer
Arran Technology Inc. | Minneapolis, MN
UST MSc Software Engineering Project: Website Marketing Analysis Tools
- Analyzed client data requirements for real estate website, **created** SQL **stored procedures**, integrated **Online Analytical Processing (OLAP) and statistical reporting tools** which **increased client productivity.** Provided complete lifecycle documentation.

2001 - 04 - 2002 - 04

Junior Software Engineer
e.Intelligence Inc. | Minneapolis, MN
- **Co-developed** client configuration components for corporate Predictive Statistical Analysis and Modeling Enterprise Software, eliminating the need for technical knowledge. Provided **testing/SQA.**