MANPREET KAUR
DATA ANALYST

0411 048 215
Melbourne, Victoria

manpreetkaur.data@gmail.com
www.linkedin.com/in/manpreet-kaur-aus/
https://github.com/ManpreetKaur-Aus

Originally trained as a physicist, my research career provided me with great skills, such as problem-solving, analytical, and logical thinking to tackle difficult questions. After the parenting break, I wanted to work in a different field where I could leverage my analytical skills. Data Science seemed a great way to use my science and technical skills. Working in the industry helped me gain skills, but was not limited to consulting, data science, Coding, and stakeholder management. Above all, I learned that my favorite part was data visualization. I love the power of visualizing data, communicating the solutions, and the story, that the data tells us; that drives me. So finally, I decided to upskill my data analytics skills, which empowered me with the suitable tools and expertise to take up the next data challenge.

SKILLS

- **Programming :** Python, SQL
- **Tools/IDE :** PyCharm, PySpyder, Jupyter Notebook, Excel, Google Colab
- **Databases :** Postgresql
- **Visualization :** Tableau, PowerBi
- **Machine Learning :** Linear & Logistic Regression, SVM, Decision Tree, Random Forest, Xgboost, anomaly detection, KNN, KMeans, Naïve Bayes, PCA, Hyperparameter tuning
- **Libraries:** NumPy, Pandas, Matplotlib, Scikit-learn, Seaborn, Plotly
- **Web/API:** Flask
- **Cloud :** Amazon Web Service (AWS-S3, EC2, Elasticbeantalk, rekognization)

ARTICLES

- Concept of Various Probability Distributions in Statistics
- Hypothesis testing: A complete description

PROJECTS

Capstone Project -National Pollutant Inventory
2022

General Assembly

- Utilized PostgresSQL query and python to gain insights about substance emission in states/cities, their reduction techniques, and transfer reporting.
- Cleaning and further visualizing the SQL query results by using python libraries
- Validated the accuracy of NPI Facility Emissions Estimations and forecasting using a machine learning model
- Created visualizations and dashboards using Tableau

Data Sprint Project -Wind turbine industry in the U.S.
2022

General Assembly

- Performed analysis to study the different parameters of the wind turbine industry
- Transformed data into useful information to help support business decisions
- Conducted descriptive statistical analyses in Excel and Python
- Created visualizations and dashboards using Tableau/Power BI
- Reviewed the team's findings, insights, and recommended actions
- Utilized the data analytics workflow within an agile development framework

Fault Detection In Wafers Based On Sensor Data

This project aims to build a machine learning model that predicts the quality of wafers (a thin piece of semiconductor material) based on the inputs from various sensors. It helps in deciding whether a wafer needs to be replaced or not.

- Created binary classification application using Flask and python to predict whether the wafer is working or not
- Performed explanatory data analysis (EDA) and built custom models for clustering data using KMeans and classifying using Random Forest and Xgboost
- Deployed the application using AWS (ElasticBeanstalk)

CERTIFICATIONS

- Tableau 10 A-Z: Hands-On Tableau Training For Data Science from Udemy.com
- The Complete SQL Bootcamp from Udemy.com
- Python for Data Science and Machine Learning Bootcamp from Udemy.com
- Complete Python Bootcamp: Go from zero to hero in Python 3 from Udemy.com

EDUCATION

General Assembly
2022
Data analytics immersive Course

Guru Nanak Dev University, India
Ph.D. (Physics)

INTERESTS

In my own time, I love to make creative creations using all-natural resources and colors, which brings something into reality from my imagination.

PROFESSIONAL EXPERIENCE

Machine Learning Associate Engineer
June 2021 - Dec 2021
DIUS

- QGIS for the geospatial analysis of the target area of interest and understanding patterns in the geospatial data
- Integration of Python GIS APIs with Machine Learning libraries, such as scikit-learn in Jupyter Notebook
- Applied the machine learning algorithms for prediction, clustering, classification, and visualization of geospatial data
- Used the bespoke software tools for prospective targeting, which resulted in 85% of accuracy in multiple target AOI
- Worked closely with the business stakeholders, including client and Senior management, on various occasions, which included requirement gathering, presenting solutions, proposals, and POCs

Machine learning Intern
Jan 2021 – May 2021
Neuron Intelligence AI

- The objective of this project is to predict the bankruptcy of different companies.
- Engaged in all project life-cycle stages, including requirements gathering, data collection, data transformation/processing, building validating models, and Performing explanatory data analysis (EDA).

Visiting Scientist at CSIRO
May 2019 – Sep 2019
CSIRO Manufacturing, Clayton, VIC.

- Data analysis of drug delivery system and formation of a hydrogel, polymer synthesis, Preparation, and characterization.
- Data cleaning, visualization, and data interpretation of data related to biomedical materials and compression studies of different materials.

Assistant Professor
2012-2013, 2015 2016
Department of Applied Sciences (Physics), ABES Institute of Technology, India.
Department of Applied Sciences (Physics), GTBIT, New Delhi, India.

VOLUNTEER EXPERIENCE

- Worked as a visiting scientist at CSIRO, Clayton, VIC.
- Associated with STEM Professionals in Schools Program by CSIRO Education and Outreach