MRINAL WALIA

Toronto, Ontario | 905-781-2945 | walia8@uwindsor.ca
linkedin.com/in/mrinal-walia-b0981b158 | github.com/abhiwalia15 | mrinalwalia.netlify.app

TECHNICAL SKILLS

Programming Languages: Python(primary), R, HTML, JavaScript, ChatGPT Prompt Engineering
Tools and Technologies: GCP, Git, Bitbucket, Docker, SonarQube, Jira, PyCharm, AWS, Azure
Data Science: Data Visualization, Data Analysis, Data Exploratio, Gathering & Cleaning, Performance Metrics
Machine Learning: TensorFlow, PyTorch, Big Data, OpenCV, NLP, Deep Learning, Neural Networks, LSTM
Frameworks/Tools: Apache Spark, Hadoop, VertexAI, Jupyter Lab, NumPy, Pandas, Matplotlib, MS Tools, Canva
Databases: SQL, Single Store DB, MongoDB, Big Query, Cloud SQL
Other: Airflow, PyTest, Unit Test, Pylint, Coverage, Sphinx, GitHub Action, Azure DevOps, PowerBI, Looker, Unix

EDUCATION

Master of Applied Computing (*Specialised in Artificial Intelligence*)   Jan 2022 – Sept 2023
University of Windsor, Windsor, Ontario
Bachelor of Engineering in Computer Science and Engineering   Aug 2017 – Jul 2021
Dayananda Sagar College of Engineering**,** Bengaluru, India

PROFESSIONAL EXPERIENCE

Data Scientist | COOP, Loblaws Companies Limited**,** Toronto, Canada   Jan 2023 – Aug 2023
Promo Optimization Engine *(Python, GCP, BigQuery, VMs, Jupyter Lab, BackTesting)*
- Developed and executed an enhanced **backtesting methodology** for wholesale store groups, resulting in an additional **$35M** in yearly revenue (based on yearly revenue in 2022).
- Pioneered the revamp of **rating metrics** for the **backtesting pipeline** of Promo Optimization Tool, introducing more interpretable and comprehensive metrics to **evaluate model performance**.
- Performed **data analysis** to identify competitors to Loblaws Divisions based on price differences for various products and integrated the feature to the **front-end tool** for Business to access the competition pricings.
- Led the development of a comprehensive **newsletter team**, resulting in a **30% increase** in engagement.

Data Scientist | **Technical Writer,** OpenMined, Google, UK   May 2022 – Dec 2022
PySyft *(rST, Markdown, Canva, Figma, Git, Slack, Notion, Python, SQL, Jupyter Notebook)*
- Executed **privacy-preserving machine learning** and **federated learning algorithms** from **PySyft** library on sensitive and **real private data.**
- Built a one-click end-to-end **testing CI/CD pipeline** using GitHub actions.
- Created comprehensive **Jupyter Notebooks** guides and tutorials (**Pull Request**: #6633, #6669, #6759, https://github.com/OpenMined/PySyft/pulls/abhiwalia15) to facilitate knowledge sharing and collaboration within the **open-source data science** community.

Junior Data Scientist | TESSACT**,** Bengaluru, India   May 2020 – Apr 2022
Tech *(Python, OpenCV, AWS [DynamoDB, S3, SageMaker], PySpark, Jira, Git, TensorFlow, NLP)*
- Improved **RMSE** and **LogLoss** metric of a **machine learning model** for a FinTech project by **1.35%** by tactfully framing business problems and communicating challenges and results to **key partners**.
- Implemented end-to-end **data solutions** using **cloud** platforms, applying advanced analytical methods to manipulate and analyze large, complex datasets for actionable insights.
- Collaborated with cross-functional teams, including senior Data Scientists, to develop and implement **scalable deployments scripts** using **SQL query** processing methods in production environments for **CI/CD.**
- Designed and integrated various **ETL processes** using cutting-edge technologies such as **AWS S3, Single Store, Lambda, Kafka, Cloud Watch, EC2,** and **API Gateway** for efficient data processing and management.
- Consolidated **data scraping strategies**, encompassing dynamic scraping and API extraction, to prepare datasets.
- Developed and administered a robust **Data Quality Validation Framework**, establishing rules to ensure comprehensive **data quality** checks and processes.

ACADEMIC PROJECTS

Stocks Analysis & Prediction *(Python, R, StreamLit, LSTM, APIs, NoSQL, TensorFlow, Git)*
GitHub: 36 Stars & 30 Forks (https://github.com/abhiwalia15/AI-for-Finance-Stocks-real-time-analysis-)
- Designed and coded an interactive dashboard for predicting stocks prices (98% accuracy) for next 30 days.

Learn Python for DS and ML *(Pandas, NumPy, Jupyter Notebook, Plotly, Python)*
GitHub: 35 Stars & 14 Forks (https://github.com/abhiwalia15/Python-for-Data-Science-and-Machine-Learning-Bootcamp)
- Developed an extensive Python program covering data visualization, NumPy, pandas, web scraping, and machine learning with SciKit Learn.

Facial Emotion Recognition *(Python, TensorFlow, Keras, RestAPIs, NLP, MongoDB, OpenCV, Git).*
GitHub: 6 Stars & 3 Forks (https://github.com/abhiwalia15/Facial-Expression-Recognition)
- Achieved 94% accuracy by designing a neural network architecture to recognize 7 types of emotions in real-time.

Detecting Postures Deformities *(Pose Estimation, Python, Deep Learning, OpenCV, Git, MatPlotLib, Seaborn)*
GitHub: (https://github.com/abhiwalia15/Sitting-Posture-Recognition)
- Presented at the *MATJOURNALS, 2021;* awarded '*Best Paper Presentation*' in the Machine Learning category.
- Elaborated on deep learning algorithms like BlazePose, DeepPose, OpenPose, etc. used for pose estimation.

COVID-19 Cases Prediction *(Python, Machine Learning, Data Science, Streamlit, Tensorflow.js)*
GitHub: (https://github.com/abhiwalia15/COVID-19)
- Conducted time series analysis on COVID-19 data, identifying trends in affected cases, deaths, and recoveries.
- Created data visualizations and dashboards to share valuable insights with the data science department.

IPL Team Prediction *(Python, Pandas, seaborn, plotly-dash, cufflinks, numpy, jupyter notebook, MS Excel)*
GitHub: (https://github.com/imhardikj/Team-Predictions)
- Developed an extensive Python program to find the best performing players in cricket sports busing the dataset and technologies like data visualization, NumPy, pandas, web scraping, and machine learning with SciKit Learn.

ACHIEVEMENTS

- **Data Science Writer–** Professional technical writer in Data Science domain (6+ years) holding 300+ articles with a background working for multiple magazines, publications, start-ups, open-source organizations, etc.
- **Google Season of Docs`22–** Aced a competition by Google to work with one of its open-source organization on a fully-funded project with a Data Science library named PySyft.
- **GitHub Excellence–** Recognized among India's top 100 contributors in 2020, with 3300+ commits across 30+ repositories with 50+ stars in personal projects.
- **Smart India Hackathon–** Received a silver medal (out of 50 teams) for presenting the project as tech lead from AI & ML Innovation Lab@DSCE at the National level.

CERTIFICATIONS

- Microsoft Certified: **Azure** Fundamentals
- Side Hustle Strategies for **Data Science and Analytics Experts**, LinkedIn Learning
- Software Development Life Cycle (**SDLC**), LinkedIn Learning
- Planning and Releasing Software with **JIRA** (2018), LinkedIn Learning

- Git Started with **GitHub**, Git
- **Machine Learning** with **Python**, Coursera-**IBM** (9 Courses)
- **Data Science** Professional Certificiate, Coursera-**IBM** (9 Courses)
- Feature Engineering with **PySpark**, DataCamp