SHUBHAM MALIK

ML Data Scientist

shubhamalik2000@gmail.com | +91 9013584157 | New Delhi, India
[LinkedIn](#)|[GitHub](#)|[HackerRank](#)
Portfolio: https://shubhamalik2000.github.io/portfolio/

Data Scientist with an experience of about 9 months, and badge for 5-star HackerRank python coder, I aspire to build ML/DL models to foster data driven decisions. My current finding being studying Arctic Sea Ice shrinkage (current rate of 12.2%) and drying up of Aral Sea (89% lost in last few decades).

PROFESSIONAL EXPERIENCE

Data Scientist   **Sept 2023 – June 2024**
AI Variant | Bengaluru, India

Project 1: Resume Recognition | Python, Visualization, Natural Language Processing

- Build a resume recognition model using **NLP (natural language processing)** to help the recruiting team select the right candidate for the job role.
- Successfully increased resume parsing accuracy by 25% through iterative model tuning and validation.
- **Reduced the manual resume screening time by 50%** by deploying an efficient and scalable solution.

Project 2: Stock Price prediction Forecast | Python, Visualization, Deep Learning

- Led a of **group of 7** by setting example to work the stock price of Amazon.
- Continual optimization of models resulted in finally using a **LSTM model (deep learning model) based on Keras and TensorFlow**, which had the **RMSE value in the range of 4-5**.
- Helped stakeholders **reduce analysis time from 30 minutes to 2 minutes**, along with **increasing the profit by about 50%.**

Project 3: Case Study – NAMMA Yatri | SQL, exploratory data analysis

- Our case study, utilizing SQL queries on trip-related data, delves into user behaviour, driver performance, and payment trends within OLA App.
- Key insights cover cancellation rates, conversion rates, top-performing drivers, and popular locations, offering a comprehensive understanding of the app's operational dynamics.

TECHNOLOGICAL SKILLS

- **Data Science / Machine Learning / Deep Learning / AI:** Python, Data Visualization, Supervised and Unsupervised Machine Learning algorithms (regression, classification (logistic regression, SVM, k-NN, Decision tree, Random Forest), clustering (DBSACN, k-mean etc.)), ANN, CNN, EDA, NLP, Feature engineering, feature selection and extraction, LLM etc.
- **Mathematics for ML / DL:** Statistics and Probability, Algebra, Calculus, Linear Algebra, Advance statistical techniques

- **Python Packages and framework:** Scikit-learn, TensorFlow, Keras, Cloud Vision, API, NumPy, Pandas, SciPy, Beautiful Soup, PyTorch, Matplotlib, Seaborn
- **Programming languages:** Python, R (Basic)
- **Database:** MySQL
- **Visualization:** Tableau, PowerBI, Excel and Advance Excel for interactive dashboards
- **Cloud Development**: AWS and Azure
- **Big data**: Hadoop and Spark

EDUCATION

BSc Mathematics | University of Delhi (2018 - 2021) | Grand CGPA 8.716
Gold medalist during final year (10 CGPA)
Well-built base in Linear Algebra, Calculus, Pattern Analysis, Algorithms and statistics and probability

12th standard, PCM with computer science (C++) | CBSE (2018) | 92.6%
Laid the requisite foundation for mathematics and computer science

CERTIFICATIONS / PROFESSIONAL TRAINING

Data Scientist Certification | ExcelR (2024)
Data Analyst Basic Quiz | CareerNinja, LearnTube (2024)

Achievements

HackerRank Python 5-star coder | Link