Abhishek Tiwari

New Delhi, India | +918130034421 | Tiw.abhishek04@gmail.com
LinkedIn | Portfolio | GitHub

PROFILE: Data Scientist

I am a freelance data scientist, Currently, developing an AI-powered Website, that lets users share images using face recognition. I have a background in mechanical engineering and groundwater management, I used data science to predict missing metrological data in my master's thesis. Later took formal training in Machine learning, Deep learning, Data analysis, Web Development, and Cloud Infrastructure. Now, I am looking to apply and enhance my data-driver problem-solving skills to tackle challenges in important projects.

SKILLS

Data structures and algorithms, Data Visualization, Big Data/Database.

Programming languages
- Python
- SQL
- Bash/Shell

IDE-Code Editors
- Anaconda-Jupyter
- Pycharm
- VS Code

Web Development Frameworks:
- Flask/Django
- Streamlit

Cloud & DevOps:
- AWS infrastructure
- Docker

Deep Learning
- Deep Neural Network
- Recurrent Neural Network
- Convolutional Neural Network
- Tensorflow
- Pytorch

Data Analysis
- Tableau
- Power Bi
- Matplotlib/ Seaborn
- Pandas/ Numpy
- Google Analytics
- Pyspark

Machine Learning
- Scikit-Learn
- XGBoost
- Arima/Arma
- Natural Language Toolkit

Management methodology tools
- Jira
- Scrum
- Agile

Language and Other Skills
- English-C1/C2
- German-B1/B2
- MS. Office Suite
- CCNA Certified

PROJECTS

Milestone Project at "DataScientest":
- **Project Focus**: Weather prediction in Australia, using 10 years of data to forecast rain, temperature, and wind speed.
- **Objective**: Predict rain for the next day, with applications in agriculture and logistics.
- **Challenges**: Missing data and class imbalance (more dry days than rainy days).
- **Solutions**:
  - Used KNN imputation for missing values.
  - Applied SMOTE for oversampling to balance the dataset.
- **Modeling Approach**: Tested multiple machine learning models, including Random Forest and XGBoost.
- **Results**:
  - Random Forest and XGBoost delivered high accuracy and recall in predicting rain.
  - XGBoost with SMOTE oversampling: Accuracy Train Set: 0.917, Accuracy Test Set: 0.862.
- **Skills Demonstrated**: Data preprocessing, feature engineering, and machine learning application for real-world predictions.

Berlin Rent Prediction Model Using Machine Learning (GitHub):
- A fully functional website hosted on an AWS ec2 server that predicts rent for different apartment types in several locations in Berlin (Predictor: XGBoost).
- Train Accuracy: 0.927, Test Accuracy: 0.872.

A Footballer's Image classifier (GitHub):
- A fully functional website that uses photos to classify football players (Classifier: SVM), Accuracy: 0.887.

A Microsoft Power-BI and SQL Sales Analysis Project (GitHub):
- A Microsoft Power-BI, Dashboard showing information regarding sales of A company over four years (2017-2020) consists of 5 slides named Overview, Yearly, Market, Customers, and Product Analysis.

WORK EXPERIENCE

Remote Freelancer Data Scientist July 2024- Present
Lens Painters, New Delhi, India.

- Leveraged face detection & recognition technology to create a website that shares images using AI-based facial recognition techniques.
- Improved detection accuracy by optimizing face detection models and enhancing real-time data pipelines.

Foundry Production Employee (Gravity Casting) January 2023 – January 2024
Bharat Forge Aluminiumtechnik Gmbh., Brand-Erbisdorf, Germany.

- Quality control of the aluminum casting (NDT analysis of alloying elements - visually and with EPOCH 650 defect detector) for detection of porosity and holes.
- Spectrochemical analysis of aluminum samples for precise quantification of the element composition.
- Control of the process flow by Kuka-robot programming.

Working Student Position as a Production Assistant June 2020- November 2020 &
Volkswagen Fahrzeugwerk, Zwickau, Germany. June 2021- November 2021

- Assembly of the engine and the outer body for VW cars.

Amazon Shipping/Warehouse Associate October 2018- May 2019
Amazon, Winsen-Luhe, Germany.

- Picking, packing, and shipping customer orders.

Project and Execution Function Engineer July 2015- July 2017
Global Water Separation Systems Pvt Ltd, New Delhi, India.

- Planning and installation of industrial water separation systems.
- Data acquisition, management, and analysis.
- Developed positive working relationships with stakeholders to coordinate work activities effectively.

EDUCATION

Diploma in Data Science May 2024
DataScientest, Puteaux, France.

- **Machine Learning Models:** Gained expertise in regression, classification, and clustering using algorithms like Logistic Regression, Random Forest, and K-Means.
- **Deep Learning Fundamentals:** Developed skills in neural networks, convolutional networks, and transfer learning using Keras and TensorFlow.
- **Data Engineering:** Proficient in SQL, API integration, and distributed data processing with PySpark.
- **MLOps:** Hands-on experience with Docker, Streamlit, and deploying machine learning models in cloud environments (AWS).
- **Time Series and Anomaly Detection:** Worked on advanced time series forecasting (ARIMA) and anomaly detection methods (KNN, LOF).

M. Sc., Groundwater Management June 2023
Technische Universität Bergakademie, Freiberg, Saxony, Germany.

- Master's thesis: Hydrochemistry of Lake Zwenkau in Leipzig and its future aspects. Analysis of the meteorological data of the district of Leipzig and chemical analysis of Lake Zwenkau.
- Modules in Management: Project Management, Resource Management, Corporate Sustainability & Risk Management, Human Resources & Organization Management.
- Modules in Modeling: Geoscientific Communication (GIS), Hydrogeological Flow, and Transport Modeling of Water (FeFLow).
- Modules in Environmental Sciences: Environmental Management & Policies and Hydrogeology.

Bachelor of Technology, Mechanical Engineering June 2015
Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India.

- Bachelor's thesis: Principles of magnetic levitation and their applications.