MOHIT BHALKE

Associate Software Engineer

✉ mohitbhalke0712@gmail.com 📞 +91 8830580293 🌐 *https://github.com/mohitbhalke07* 📍 Pune, Maharashtra

Data-driven professional skilled in data analysis, data science, and machine learning. Proven ability to build predictive models, streamline workflows, and deliver actionable insights. Proficient in Python, SQL, and frameworks like Scikit-learn, LangChain, Flask and Django. Passionate about solving complex problems and driving data-informed decisions.

Experience

Associate Software Engineer | DeltaFx Technologies   *May 2023 - Present*

- Designed and implemented **interactive dashboards** using **Python (Pandas, Plotly, Dash**/Streamlit) to visualize key metrics for business performance and client trading behavior.
- Developed and deployed an **NLP-based chatbot** for automated customer support using **text preprocessing, intent classification, and NER**, reducing support response time by **40%**.
- Worked with historical client transaction data to identify patterns and behaviors, laying the groundwork for **predictive analytics**.
- Explored and prototyped **ML models for lead scoring and client segmentation** using **clustering and classification algorithms**, enhancing targeting accuracy for sales teams.
- Took initiative to **optimize existing codebase**, reducing code redundancy by **25%**, increasing modularity, and enhancing system maintainability

Data Analyst | Careervira   *July 2022 - July 2023*

- Collaborated with cross-functional teams (**marketing, product, and tech**) to define data requirements and deliver actionable insights for strategic decision-making.
- Performed **data extraction, cleaning, and analysis** on user behavior and course engagement data using **Python (Pandas, NumPy), SQL, and Excel** and Automated repetitive workflows using Python scripts and **Google AppScript**, cutting down recurring report generation time by **50%**
- Conducted **statistical analysis** and **A/B testing** to evaluate marketing strategies and feature effectiveness.
- Developed **predictive models** for **user segmentation** and **sentiment analysis**, boosting the efficiency of personalized marketing by up to **35%**.

Data Science Intern | YoShops   *April 2022 - June 2022*

- Collaborated with a diverse team of interns spanning various departments, contributing to projects involving the development of **machine learning models, web scraping techniques, and sentiment analysis.**
- Simultaneously participated in an intensive 8-week training program in **Data Science & Machine Learning**.

Skills

- **Programming Languages:** Python, PHP
- **Frameworks & APIs:** Django, Flask, CodeIgniter3, FastAPI
- **Databases & Cloud:** MySQL, MongoDB, ElasticSearch, Vector Database, AWS
- **Data Analysis & ETL:** Statistical Analysis, Predictive Analytics, Data Cleaning, Web Scraping, Automation, ETL, Excel
- **Machine Learning & Generative AI:** Supervised Learning, Unsupervised Learning, NLP, Predictive Modeling, BERT, LangChain, RAG, Hugging Face Transformers
- **Tools & Libraries:** Scikit-learn, Pandas, NumPy, Seaborn, Matplotlib, TensorFlow, Scrapy, Selenium, NLTK, Transformers, Keras, OpenAI API
- **Data Visualization:** Tableau, Matplotlib, Seaborn, Streamlit, Dash

Projects

- **Recommendation System**
  Stack: Python, Streamlit, HuggingFace, LangChain, Pinecone, OMDb API
  - Built a **GenAI**-powered **movie recommendation system** using **BERT-based sentence embeddings** and **Pinecone vector DB**, enabling semantically rich and personalized movie suggestions.
  - Transitioned from traditional NLP (**TF-IDF** + **cosine similarity**) to a modern GenAI stack (**HuggingFace** + **LangChain**), improving contextual relevance and scalability for large datasets.
  - Deployed an interactive **Streamlit** web app with poster previews, detailed IMDb metadata, and a **"You may also like"** section powered by **vector-based similarity retrieval**.

- **Commodity Price Predictor**
  Stack: Python, Streamlit, Scikit-learn, XGBoost, GridSearchCV, Joblib, Pandas
  - Designed and deployed a unified machine learning solution for predicting prices across **cars, mobile phones, and houses** using structured datasets and domain-specific **regression models**.
  - Built robust preprocessing pipelines with **OneHotEncoding**, **target encoding**, and **feature scaling**, enabling consistent handling of **categorical and numerical data** across commodity types.
  - Achieved strong predictive performance, achieving optimal R^2 **scores up to 91%**, with **MAE** and **RMSE metrics** tailored per domain for transparency and model effectiveness.
  - Improved model reliability and accuracy using **GridSearchCV**-based **hyperparameter tuning**, minimizing error margins across all three commodities.
  - Designed a modular **Streamlit** application featuring a dynamic UI that adapts input fields based on selected commodity, allowing seamless interaction and **real-time predictions**.

- **Fraud Detection System Using Anomaly Detection in E-commerce**
  - Designed and implemented a supervised ML system to identify fraudulent transactions within a simulated e-commerce setting (e.g., Flipkart seller dashboard).
  - Enabled real-time fraud detection by integrating a sync trigger and automated background processing for continuous prediction updates.
  - Assessed model performance using precision, recall, and F1-score, and fine-tuned hyperparameters to optimize the balance between false positives and negatives.
  - Built an interactive dashboard to visualize transaction trends, fraud predictions, and actionable insights through dynamic charts and filters.

Education

- **B-Tech Honor's | Data Science | Sanjivani College Of Engineering**

- **B-Tech | Mechanical Engineering | Sanjivani College Of Engineering**

- **HSC | Science | New English Jr. College**

- **SSC | St. Xavier's High School 2017**

Certification

- Applied Machine Learning in Python | Coursera
- Statistics for Data Science with Python | Coursera
- Fundamental Of Visualization with Power Bi | Coursera
- Deep Learning | NPTEL
- SQL For Data Science | Coursera
- AWS Academy Cloud Foundations | AWS
- Visual Analytics With Tableau | Coursera
- Python Programming | NPTEL

Language

- English
- Hindi
- Marathi