Shubh Mehta

Linkedin: shubh9mehta
Github: github.com/shubh9mehta
Email: shubh9mehta@gmail.com
Mobile: +1(812) 803-4508

EDUCATION

- **Indiana University Bloomington** — Bloomington, Indiana, USA
  Master of Science in Data Science; ***GPA: 3.14/4*** — *August 2023 - Present*
  Relevant Coursework: *Data Mining; Introduction to Statistics; Elements of Artificial Intelligence; Computer Vision; Management, Access and Use of Big and Complex Data*

- **SRM Institue of Science and Technology** — Chennai, India
  Bachelor of Technology - Computer Science Engineering; ***CGPA: 8.81/10*** — *June 2019 - May 2023*
  Courses: *Data Mining and Analytics, Digital Image Processing, Database Management Systems, Advance Calculus and Complex Analysis, Design and Analysis Of Algorithms, Artificial Intelligence*

SKILLS SUMMARY

- **Languages**: Python, R, C++, C, SQL, HTML, CSS
- **Frameworks**: Scikit-learn, TensorFlow, Pandas, PyTorch, Matplotlib, Numpy, Seaborn, SciPy
- **Softwares**: Tableau, Ms Excel, GIT, MySQL, SQLite, SAS Business Intelligence, Microsoft Power BI
- **Soft Skills**: Interpersonal Communication, Flexibility in Dynamic Environments, Collaborative Partnership,
  . Ethical Judgment, Strategic Problem Resolution

EXPERIENCE

- **TeknoBiz Solutions Pvt Ltd.** — Remote
  Data Analyst Intern — *May 2022 - Sep 2022*
  - Facilitated comprehensive data analysis to identify trends and patterns, enhancing sales revenue predictions by 20%.
  - Improved machine learning workflow, contributing to a 15% increase in model accuracy for predictive analytics projects.

PROJECTS

- **Customer Churn Analysis in Telecom Sector**
  Conducted an extensive data-driven analysis to understand customer churn dynamics of a telecom company.
  - Leveraged Python, Pandas, NumPy, Matplotlib, Seaborn, and Plotly for data cleansing and visualization, achieving a 30% boost in extracting actionable insights.
  - Employed Logistic Regression, Random Forest, and Gradient Boosting for predictive modeling, enhancing data accuracy by 12% and customer retention by 25% through optimized strategies and engagement tactics.
  - Managed a dataset with 7,043 entries and 21 variables, facilitating strategic decision-making and contributing to targeted retention efforts with measurable success.

- **Customer Personality Analysis and Segmentation for Tailored Marketing Strategies**
  Project analyzing customer data to customize products and target marketing, improving engagement and satisfaction.
  - Leveraged Jupyter Notebook for data preprocessing and feature engineering on a 2240-entry dataset, and applied K-means clustering for customer segmentation by behaviors and demographics, enhancing marketing precision.
  - Implemented data cleansing, including handling missing values in a 2240-record dataset, achieving a silhouette score of 54, indicating strong cluster formation for improved customer targeting and satisfaction.

- **Post Covid-19 Analysis on Healthcare**
  Analysis of COVID-19's impact on healthcare using machine learning for predictive accuracy.
  - Led groundbreaking analysis of COVID-19's effect on global healthcare using Gradient Boosting, XGBoost, and LightGBM for predictive insights; simultaneously devised a Python and Pandas data cleansing process that reduced data discrepancies by 25%, boosting data integrity for analytics.
  - Crafted and optimized analytical models including Multi-Layer Perceptron and LightGBM, achieving an R-squared of up to 0.9643 and significantly lowering MSE, enhancing forecast accuracy and decision-making quality.

- **Sales Data Analysis and Visualization for Computer Hardware Company**
  Constructed Tableau dashboards to visualize sales data, identifying trends for driving strategic decisions.
  - Developed two Tableau dashboards; the first analyzed sales performance across regions with advanced visualizations for strategic insights, while the second offered an in-depth profit analysis post-sales, merging financial metrics for enriched fiscal understanding.

- **Time Series Analysis of Cryptocurriencies Using LSTM**
  Enhanced Ethereum price prediction using LSTM, improving investment returns by 20% through deep learning.
  - Built LSTM models in Python and TensorFlow, improving Ethereum price forecast accuracy by 15%, and enhanced financial prediction in cryptocurrency through feature-engineered time series analysis. Collaborated on integrating AI with trading strategies, increasing simulated portfolio returns by 20%.

ONLINE COURSES

- **Machine Learning**   Coursera
  Stanford University

- **Data Science**   Internshala
  Internshala

- **Database Foundation**   Oracle Academy
  Oracle

- **Building Web Applications in PHP**   Coursera
  University of Michigan

COLLEGE VOLUNTEER EXPERIENCE

- **Committee Member at Techno Management Fest-AARUSH**   SRMIST
  Connected with Icons of various fields to be our guest and take workshops for the students.

- **Event Organizing Committee of C.Tech Club**   SRMIST
  Organized events and directed workshops reaching over 5000 students.