Ribhay Singh

singhribhay@gmail.com | 925-315-0834 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

Arizona State University | Tempe, AZ   **December 2025**
Bachelor of Science in Data Science (Computer Science Track)
- Awarded **Dean's List** for academic excellence in Fall 2024 and Spring 2025

INDUSTRY EXPERIENCE

Initialyze - San Francisco Bay Area
Software Engineer Intern (Data Engineering/Analytics)   **May 2024 – August 2024**
- Led a Data ingestion project to acquire and integrate data from internal web services into Adobe Experience Platform using Dataflow UI and REST APIs, improving data integration **throughput by 20%**.
- Built scalable ETL pipelines in Databricks notebooks, reducing data processing time and improving validation accuracy for distributed datasets.
- Developed SQL profiling and validation quersing using MySQL and AEP Query Service across **35 tables,** enhancing reporting accuracy and ensuring data integrity.
- Contributed to schema creation and data governance processes within AEP to support scalable data solutions.

PERSONAL PROJECTS

FitPro - Fitness Analytics Platform | Python, PostgreSQL, scikit-learn, Databricks, Tableau, Streamlit
- Built a full-stack fitness analytics platform using machine learning that predicts calorie burn with **80% R^2 score**, analyzes workout trends and forecasts goal achievement using XGBoost and Random Forest models.
- Engineered a PostgreSQL data pipeline with Databricks for real-time data ingestion, transformation and goal tracking.
- Developed interactive data visualization dashboards in Streamlit and Tableau to support data analytics, visualize workout trends and personalize calorie burn insights, validated through user testing across multiple workout types and goal scenarios.

EchoMind AI - Mental Health Journaling Tool | Python, PostgreSQL, Streamlit, FastAPI , GPT-3.5
- Built a mental health journaling tool that summarizes and analyzes user entries using an LLM powered by GPT-3.5, achieving **90% sentiment alignment** during internal validation.
- Designed and deployed a FastAPI REST API backend with SQLite to process **100+ journal entries**.
- Implemented an intuitive Streamlit frontend for real-time journaling and instant feedback.

Real Estate Price Prediction Model | Python, scikit-learn, pandas, Flask, HTML/CSS
- Developed a predictive model using multiple linear regression to estimate housing prices, achieving **85% R^2 score**.
- Cleaned and analyzed a Kaggle housing dataset, applying EDA and feature engineering to identify key factors.
- Deployed a Flask web application with HTML/CSS for live user inputs and real-time price predictions.

World Happiness Report | R (dplyr, ggplot2), Python (pandas, Numpy, Matplotlib)
- Conducted data mining and multivariate analysis on the 2019 World Happiness Report across **150 countries** to quantify the influence of GDP, life expectancy, and social support on happiness scores.
- Performed regression, correlation, and feature analysis to identify key socio-economic drivers.
- Automated Python pipelines for EDA, model validation, and result verification.

TECHNICAL SKILLS

Languages: Python, R, Scala, SQL, Java, C++, HTML/CSS/JavaScript
Libraries & Frameworks: scikit-learn, PyTorch, TensorFlow, pandas, NumPy, FastAPI, Flask, Streamlit
ML & Data: EDA, Feature Engineering, Data Wrangling, Predictive Modeling, NLP, Model Evaluation, A/B Testing, Statistics
Tools: Tableau, Power BI, Azure, VS Code, Databricks, Spark, Git/Github, Adobe Experience Manager, Adobe Experience Platform
Certifications: Complete Data Science Bootcamp, Complete SQL Bootcamp, HTML/CSS/ Javascript/React - Online Certification Course, Adobe Experience Manager Masterclass, Databricks Fundamentals Accreditation