

Yash Soni

Maryland, United States | +1 (667) 320-5799 | yashnitinsoni.17@gmail.com |
www.linkedin.com/in/yashsoni176 | <https://github.com/yashnitinsoni> | [Portfolio](#)

EDUCATION

University of Maryland, Baltimore County | Master's in Information Systems (*GPA: 3.89*) *Jan 2024 – Dec 2025*
Relevant Courses: Data Visualization, Data Mining, Advanced Database Projects, Data Analysis for Cybersecurity, Statistical Learning for Data Analysis, Decision Support Systems
Vishwakarma Institute of Technology | Bachelor of Technology in Computer Science *Aug 2018 – Jun 2022*

WORK EXPERIENCE

Software Engineer, Harman India, Pune, India *Jan 2022 – Dec 2023*

- Integrated REST API data into data pipelines and backend systems using Python, enabling real-time analytics pipelines and improving access to structured data for dashboards and reporting.
- Optimized SQL queries, improving database performance by 30% and accelerating analytics workflows.
- Developed and automated data processing workflows in Python to clean, transform, and validate datasets, ensuring 100% accuracy, improving consistency across reports, and reducing manual reporting effort for analysts.
- Implemented data validation and anomaly detection using SQL constraints and Python scripts, cutting reporting errors by 40%, enhancing data reliability, and ensuring high-quality information for business dashboards.
- Collaborated with a team of 5+ engineers and analysts to optimize pipeline efficiency by 25%, enhancing scalability, automation, processing speed, and enabling faster decision-making through improved analytics.
- Worked in an Agile environment, managing tasks with Jira and maintaining code quality through Git/GitHub version control, performing reviews, and ensuring smooth collaboration across technical teams.

LEADERSHIP & INVOLVEMENT

Software Developer, The Robotics Forum - VIT, Pune, India *Jan 2021 – Dec 2021*

- Taught over 50 students Data Structures and Machine Learning by explaining topics, solving problems, and building projects to apply the concepts.

PROJECTS

AI Nutrition Analyzer (Python, TensorFlow, Keras, FastAPI, Streamlit)

- Developed a deep learning model achieving 90%+ accuracy in classifying food images for nutritional analysis.
- Trained on the Food-101 dataset with 101,000 images across 101 categories to ensure robust food recognition.
- Built a FastAPI backend to process images and return calorie and macro data in real-time.
- Created a user-friendly Streamlit interface for uploading images and displaying instant nutrition info.

Online Proctor System (Python, OpenCV, TensorFlow, YOLOv4)

- Engineered an AI-based exam monitoring system to detect cheating with over 90% accuracy.
- Merged YOLOv4 and facial landmark detection to identify faces, eyes, mouths, and mobile devices.
- Applied real-time detection of suspicious behaviors to support secure and uninterrupted remote exams.
- Tracked eye gaze and movement to flag suspicious behavior during remote assessments.

Customer Churn Prediction (Python, Scikit-learn, Pandas, NumPy, Matplotlib, Joblib)

- Designed a machine learning model with 81.76% accuracy to predict customer churn for telecom data.
- Performed data cleaning, encoding, and scaling on 7,000+ customer records for better model input.
- Identified key churn indicators like contract type, monthly charges, and customer tenure.
- Boosted F1-score by 12% by optimizing feature selection and reducing false positives.

SKILLS

Programming: Python, SQL, C++, Kotlin, JavaScript
Data Analytics & Visualization: Tableau, Power BI, Pandas, NumPy, Matplotlib, Seaborn, Excel
Machine Learning & Statistical Analysis: Scikit-learn, TensorFlow, Statistical Modeling, Feature Engineering
Data Engineering: Apache Spark, ETL (Extract, Transform, Load), Data Pipelines, Data Warehousing
Databases: MySQL, PostgreSQL, SQLite, MongoDB, Firebase
Cloud & DevOps: AWS (S3, EC2, Lambda, Redshift), Docker, GitHub Actions
Certifications: [AWS Certified Cloud Practitioner](#), [HackerRank SQL\(Advanced\)](#)