

Vaidik Shah

+919265202360 | vaidik1103@gmail.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

Graduate Computer Science student with hands-on experience in developing machine learning models using Python and Scikit-learn. Achieved high prediction accuracy in sales and solar insolation forecasting through robust data processing and optimized pipelines. Brings strong technical skills, a commitment to data integrity, and a keen interest in applying AI/ML in healthcare research settings.

EDUCATION

California State University, Long Beach

Aug 2025 - Aug 2027

Master of Science, Computer Science

Gujarat Technological University

Sep 2020 - Jun 2024

Bachelor of Science, Computer Science

• **GPA:** 9.4/10.0

PROFESSIONAL EXPERIENCE

Technolee

Jul 2024 - Present

ML Engineer

Ahmedabad, Gujarat

- Developed a logistic regression model to predict sales trends for 20+ products, achieving ~80% prediction accuracy while adhering to robust data handling practices.
- Conducted data cleaning, feature selection, and established a training pipeline using Python and scikit-learn, aligning with best practices for custom ML model development.
- Constructed and integrated a backend API to serve real-time predictions within a dashboard, ensuring seamless deployment and user accessibility.
- Leveraged Python, scikit-learn, Flask, and Pandas to deploy machine learning solutions, facilitating efficient data processing and model monitoring.

Indian Space Research Organization (ISRO)

Jan 2024 - Apr 2024

AI/ML Associate Engineer

Ahmedabad, Gujarat

- Engineered scalable machine learning models using Python, scikit-learn, and XGBoost to forecast solar insolation, achieving over 92% real-world prediction accuracy.
- Optimized model performance through feature engineering and hyperparameter tuning, reducing forecasting errors by 25% and improving computational efficiency by 20%.
- Streamlined a backend pipeline to preprocess large satellite datasets and automate model inference, enhancing data processing speed and accuracy.
- Designed and presented interactive data visualizations with Matplotlib and Seaborn to effectively communicate insights and forecasting trends to scientific stakeholders.

SELECTED PROJECTS

Academic Video Summarizer

Feb 2025

- Built a pipeline that uses Retrieval-Augmented Generation (RAG) to generate context-rich summaries of academic videos by integrating external reading materials
- Enhanced LLM performance through few-shot learning and prompt engineering, addressing knowledge gaps in pre-trained models and improving summary accuracy
- Conducted a user study confirming the system's effectiveness, outperforming traditional summarization methods in relevance and clarity for educational use

Solar Insolation Forecasting

Mar 2024

Jodhpur and Rewa

- Developed a high-accuracy forecasting model (>92%) to predict solar insolation in Jodhpur and Rewa, improving solar energy production efficiency by 15%
- Enabled better alignment of energy output with peak sunlight hours, resulting in a 20% boost in solar power generation efficiency
- Analyzed 10+ years of historical meteorological data and applied feature engineering to reduce forecasting errors by 25%
- Built scalable ML pipelines for data preprocessing, model training, and evaluation, enhancing reproducibility and deployment-readiness

TECHNICAL SKILLS

- **Languages:** Python, Java, JavaScript, TypeScript, C++
- **Frameworks & Libraries:** Node.js, Express.js, React
- **AI/ML Tools:** OpenAI API, Hugging Face, LangChain, RAG, prompt engineering, TensorFlow, PyTorch, LLM integration, Machine Learning, Artificial Intelligence, Scikit-learn, Model Monitoring
- **Developer Tools & Platforms:** Git, VSCode, Docker, Kubernetes, AWS, CI/CD, Agile
- **Education & Fields:** Computer Science, Data Science