

## EDUCATION

---

- **Stanford University** Stanford, CA  
*Master of Science in Computer Science; GPA: 3.95* *Sep. 2013 – Dec. 2014*
- **Indian Institute of Technology** Delhi, India  
*Bachelor of Technology in Computer Science; GPA: 3.70 (9.25/10.0)* *Aug. 2009 – July. 2013*

## EXPERIENCE

---

- **Microsoft** Redmond, WA  
*Software Engineer* *Nov. 2017 – Present*
  - **Azure ML:** Developed and optimized scalable machine learning pipelines for Azure, integrating distributed training systems to handle terabyte-scale datasets, achieving a 25% improvement in model deployment efficiency and reducing training time by 20%
  - **Data Processing:** Designed and implemented high-throughput data ingestion pipelines using Apache Spark, processing millions of records daily, resulting in a 35% reduction in processing latency and improved data availability for analytics teams
  - **Visualization Tools:** Created and deployed internal BI dashboards with real-time analytics for cross-team collaboration, incorporating interactive features that enhanced decision-making and improved team efficiency by 30%
- **Amazon** Seattle, WA  
*Software Engineer* *Feb. 2015 – Oct. 2017*
  - **Recommendation Engine:** Led the development of a personalized recommendation system for e-commerce platforms, utilizing collaborative filtering and deep learning models, increasing sales conversions by 15% and improving user engagement metrics
  - **Data Pipeline:** Designed and deployed ETL pipelines for real-time inventory updates using AWS Glue, ensuring seamless data flow across distributed systems and reducing data sync errors by 40%
  - **Automation Tools:** Developed and maintained automation scripts for testing and deployment processes, leveraging Python and Jenkins, which reduced manual testing time by 60% and streamlined CI/CD workflows
  - **User Analytics:** Built a comprehensive user behavior tracking system integrated with Redshift, enabling real-time insights into customer interactions and supporting data-driven product improvements
- **Open Source Contributions** Remote  
*Contributor* *2014 – 2017*
  - **ML Libraries:** Contributed advanced features to scikit-learn, including new model evaluation metrics and optimization algorithms, adopted by thousands of developers for machine learning workflows
  - **Web Tools:** Developed and enhanced Plotly components for interactive data visualizations, enabling dynamic charting capabilities used in over 500 open-source projects
  - **Tech Writing:** Authored over 40 detailed blog posts on cloud computing, machine learning pipelines, and scalable architectures, gaining a wide readership in the data science community
- **Quantiphi** Boston, MA  
*Data Engineer* *Summer 2013*
  - **AI Models:** Developed and fine-tuned NLP models for text classification in client-facing projects, achieving a 90% accuracy rate and enabling real-time sentiment analysis for business applications
  - **Data Platform:** Built a scalable data ingestion system for real-time analytics, integrating multiple data sources and reducing data processing bottlenecks by 25%

## PROJECTS

---

- **AI Chatbot:** Built an NLP-based chatbot for customer support, deployed on AWS
- **Data Dashboard:** Created interactive dashboards for sales analytics using D3.js
- **Prediction System:** Designed a demand forecasting model for retail using public datasets
- **ETL Framework:** Developed a scalable ETL framework for multi-source data integration

## PROGRAMMING SKILLS

---

- **Languages:** Python, Java, JavaScript, SQL, Go    **Technologies:** AWS, Kubernetes, React, Spark, Snowflake