

EDUCATION

- **University of Southern California** Los Angeles, CA
Master of Science in Data Science; GPA: 3.90 Sep. 2013 – Dec. 2014
- **Anna University** Chennai, India
Bachelor of Engineering in Computer Science; GPA: 3.80 (9.50/10.0) Aug. 2009 – July. 2013

EXPERIENCE

- **Pinterest** San Francisco, CA
Data Engineer Aug. 2018 – Present
 - **Recommendation Systems:** Improved pin recommendation algorithms using machine learning, increasing user engagement by 20% and enhancing content discovery across platforms
 - **Data Pipelines:** Built and optimized streaming pipelines for user data using Kafka, processing millions of events daily to support real-time analytics and reduce latency by 30%
 - **Dashboards:** Created and maintained analytics dashboards for ad performance, integrating with internal tools to provide actionable insights and improve campaign efficiency
- **Dropbox** San Francisco, CA
Software Engineer Jan. 2015 – Jul. 2018
 - **Analytics Tools:** Developed advanced analytics tools for user behavior tracking, leveraging machine learning to improve file sharing insights and increase user retention by 15%
 - **ETL Pipelines:** Built and optimized data pipelines for file metadata using Airflow, ensuring seamless data integration and reducing processing errors by 30%
 - **APIs:** Designed and deployed REST APIs for real-time data access, enabling third-party integrations and improving data accessibility for analytics teams
 - **Automation:** Automated data validation and processing workflows, leveraging Python and Jenkins to save 25 hours/week and streamline operations
- **Open Source Contributions** Remote
Contributor 2014 – 2017
 - **ML Libraries:** Contributed to LightGBM for faster model training, adding features for large-scale datasets used in production environments
 - **Visualization:** Developed Plotly components for data visualization, enabling interactive charting for over 400 open-source projects
 - **Blogging:** Authored over 50 technical blog posts on data pipelines, machine learning, and scalable systems, widely read in the tech community
- **Capgemini** Federica
Chicago, IL Summer 2013 *Data Analyst*
 - **Risk Models:** Built and validated risk assessment models for finance clients, achieving high accuracy and enabling data-driven investment strategies
 - **Reporting:** Developed and deployed Tableau dashboards for client reporting, automating data visualization and improving reporting efficiency

PROJECTS

- **Recommendation Engine:** Built a pin recommendation system
- **Visualization Tool:** Created dashboards for user analytics
- **Forecasting Model:** Designed a user engagement prediction system
- **ETL Pipeline:** Developed a scalable data integration framework

PROGRAMMING SKILLS

- **Languages:** Python, Java, SQL, JavaScript, Scala
- **Technologies:** AWS, Kafka, Tableau, Hadoop, Spark