

# SANTHOSHKUMAR CHANDRASEKAR

chandrasedkar.sant@northeastern.edu | LinkedIn | +1 (857) 230-8285 | Boston, MA

## Education

### NORTHEASTERN UNIVERSITY

Master of Science in Data Analytics Engineering · GPA 4.0/4.0

Boston, MA  
Sep 2024 – Dec 2026

Relevant Coursework: Data Management, Computation and Visualization, Data Mining, MLOps, Deep Learning, Applied GenAI

### COLLEGE OF ENGINEERING, ANNA UNIVERSITY

Bachelor of Engineering in Electronics & Communication

Chennai, India  
Aug 2017 – Aug 2021

## Work Experience

### Data Analyst Co-op | FIDELITY INVESTMENTS | Boston, MA

Jul 2025 – Dec 2025

- Laid the analytics foundation for a Health & Welfare offering that was new to the Customer Knowledge and Strategic Insights (CKSI) team, building the **first standardized reporting framework and metrics** to support lead generation and growth of the H&W business.
- Built a single **Snowflake** source of truth from siloed legacy systems (**1.3M+** employees, **3M+** enrollments), enforcing accuracy, completeness, and consistency checks to guarantee data integrity.
- Designed **Tableau** Factbook dashboards spanning four core dimensions (Employees, Plan Sponsors, Enrollments, and Engagement), applying **correlation, key driver, and trend analysis** to surface drivers and leading indicators across **50+** plan sponsors and **12M+** annual engagements.
- Partnered cross-functionally with Sales, Marketing, Health, and senior leadership, translating insights and dashboards to each team's lens to drive plan-sponsor prospecting, employee outreach, and evolution of the H&W offering.

### Data Analyst | BARCLAYS | Chennai, India

Dec 2022 – Feb 2024

- Served as **Product Owner** for the Mortgage portfolio, partnering with data scientists and campaign managers to align insights with business goals.
- Segmented customers using transaction and engagement data to target high-value mortgage prospects, lifting campaign response rates by **40%** and reducing wasted marketing spend by **5%**.
- Built and executed **50+** marketing campaigns on **Teradata** and CRM platforms and built **Tableau** and **Excel** dashboards to analyze product performance, which improved user retention by **15%**.
- Replaced manual reporting with automated **SQL/Python** pipelines, cutting turnaround by **70%** and freeing analysts for higher-value work.

### Software Engineer | LTIMINDTREE (formerly MINDTREE) | Chennai, India

Aug 2021 – Oct 2022

- Designed and optimized **ETL** pipelines using **Informatica**, **SSIS**, and **Azure Data Factory** to integrate multi-source data for an insurance client, improving data-flow efficiency by **75%**.
- Deployed and operated production ETL workflows through **CI/CD** and **Git**, monitoring scheduled jobs and resolving failures to ensure reliable, on-time data delivery.
- Built **SSAS OLAP cubes** and dimensional data models to power analytical reporting, reducing data redundancy by **30%**.
- Developed **20+** interactive **Power BI** and **Tableau** dashboards for client stakeholders to monitor key business metrics and support data-driven decisions.

## Academic Projects

### MOMENTO: AI-Powered Reader Compatibility Platform (IE7374 MLOps) [GitHub] [Demo]

Jan 2026 – Apr 2026

- Originated and owned the product end to end, from concept and data to machine-learning design, for a private reading application that matches readers by how they *Think* and *Feel* about book passages, and built its front-end in **JavaScript/JSX** with **Firestore Auth**.
- Enforced production-readiness through a **three-metric release gate** backed by an offline evaluation harness on a 450-interpretation synthetic benchmark and a **PII/injection/profanity/spam data-quality layer** that auto-triggered retraining on metric regression.
- Built the **agent-based** compatibility engine (Gemini 2.5 Flash) with a **Bradley-Terry** ranking model, and engineered the MLOps pipeline with **Airflow**, **TFDV**, **DVC**, and **MLflow** on **GCP**. Placed **2nd of 30 teams** at a Google-sponsored MLOps Expo.

### THE EYE: Explainable Object Detection for 73 Everyday Objects (IE7615 Deep Learning) [Live demo] [GitHub]

Sep 2026 – Oct 2026

- Engineered a human-in-the-loop labeling pipeline that turned **6,942** unlabeled photos into a verified detection dataset, combining OWLv2 text and image prompts with a self-trained YOLO by majority vote and routing disagreements to review (**98.9%** box accuracy).
- Designed THE EYE's core explainability layer, which labels every prediction as correct, misplaced, misnamed, or missed with its root cause and compares five models side by side, then deployed it publicly on **Oracle Cloud** with HTTPS, auto-restart, and per-user rate limiting.
- Enforced research-grade rigor: every model choice made on validation by pre-registered, SHA-256-sealed rules, all results reported with bootstrap confidence intervals, and a five-model ensemble rejected on evidence despite a **6.9-point** gain.

## Technical Skills

- Languages:** Python (pandas, NumPy, scikit-learn, PyTorch), SQL, R, DAX
- BI & Analytics:** Tableau, Power BI, Excel, KPI tracking, product analytics, marketing analytics, customer segmentation, A/B & hypothesis testing, bootstrap statistics, root cause analysis
- Data Engineering & Databases:** ETL (Informatica, SSIS, Azure Data Factory), PySpark, data modeling, schema validation, data governance, SQL Server, Snowflake, Teradata, PostgreSQL, BigQuery, MongoDB
- Machine Learning & Deep Learning:** Random Forest, XGBoost, K-Means, ARIMA, CNNs, computer vision, object detection (YOLO), transfer learning, open-vocabulary detection (OWLv2), segmentation (SAM), model explainability, feature engineering, predictive and statistical modeling
- Generative & Agentic AI:** LLMs, LLM evaluation, RAG, LangChain, multi-agent systems, prompt engineering, NLP
- MLOps & Cloud:** Apache Airflow, MLflow, DVC, TFDV, Docker, Kubernetes, CI/CD (GitHub Actions), Linux server deployment (Caddy, systemd), GCP, AWS, Oracle Cloud