Sid Singaravel

617-751-9572 | singaravel.s@northeastern.edu | linkedin.com/in/siddharthans/ | siddharthansingaravel.github.io/

EDUCATION

Northeastern University — Boston, MA
M.S. in Data Analytics, GPA: 3.95 — *Aug 2022 – Dec 2024*
Coursework: Neural Networks, Machine Learning, Database Analytics, Statistical Methods, Algorithms, MLOps

Anna University — Chennai, India
B.S in Electrical Engineering — *Jul 2014 – May 2018*

WORK EXPERIENCE

Data Science Co-op — Jul 2023 – Dec 2023
Fidelity Investments — *Boston, MA*
- Engineered XG-boosted trees to optimize retention-revenue trade-offs, generating $2 Million in revenue
- Built propensity score matching models to quantify impact of engagement initiatives on customer acquisition costs
- Implemented SMOTE-based data generation with focal loss adaptation, improving rare-class performance by 20%
- Architected workflows for Snowflake migration using dbt for model development, and standardized KPIs in Alation for cross-team data consistency (Spot Award: Alation Top Data User)

Senior Data Scientist — Apr 2021 – Jun 2022
ZoomRx, Ferma.ai — *Chennai, India*
- Implemented Bayesian A/B testing for iterative content optimization, attaining 15% uplift in click-through rates
- Applied DBSCAN and hierarchical clustering techniques to band physicians by behavioral trends, facilitating targeted outreach saving $300,000 in marketing spend (Spot Award: Innovation Champion)
- Built SQL-Python GUI product for periodic benchmarking aggregating 10 Million+ salesforce records
- Deployed RNN-based LSTMs (90% F1) for medical literature categorizing, reducing manual annotation efforts
- Created company specific market research statistical guidelines and led EDA workshops for 15 Analysts

Data Scientist — Jun 2018 – Mar 2021
ZoomRx — *Chennai, India*
- Led quantitative research studies for Fortune 500 pharmaceutical firms providing data-backed causal inference
- Coded data quality tests for HIPAA compliance – ensuring 100% accuracy in detecting data shifts, adverse events
- Deployed LEGO-like customizable exploratory data analysis blocks using PivotTable.js for non-tech stakeholders
- Designed multi-sheet Tableau dashboards in production, utilizing BigQuery and Alteryx to automate refreshes

Data Science Intern — Apr 2018 – May 2018
Institute of Remote Sensing — *Chennai, India*
- Devised spatiotemporal models to identify at-risk flood zones in coastal Chennai

Research Intern — Dec 2017 – Mar 2018
Nidec Control — *Chennai, India*
- Designed peak tracing algorithms inspired by whale foraging behavior (WOA-DE), reducing device latency by 35%

ACADEMIC PROJECTS

Sensitive Data Detection | *github.com/SiddharthanSingaravel/PII-Data*
- Architected BERT-based NLP models for sensitive content detection in student writing (92% recall)
- Implemented MLOps pipeline using Airflow, dbt, and DVC, enabling continuous model improvement

MetFaces: Diffusion Image Synthesis | *siddharthansingaravel.github.io/diffusion*
- Engineered Latent Diffusion img2img models for 128x128 image synthesis on data curated by NY Met Museum
- Optimized model training using CUDA GPUs, Hyperopt, and MLflow, achieving FID scores under 30

TECHNICAL SKILLS

Languages: Python, C++, SQL (Postgres), R
ML Framework: PyTorch, TensorFlow, XGBoost, LangChain, scikit-learn, MLflow, Hyperopt, TensorBoard, CUDA
Tools: Git, Tableau, Docker, Excel, Google Cloud Platform, Amazon Web Services, Snowflake, Airflow, dbt, DVC