Shashank Satish Damodaran

(847) 262-6438 | shashank-satish-d@u.northwestern.edu | in/shashank-satish | shawshank-cs.github.io

EDUCATION

Northwestern University — **Evanston, IL**

*Master of Science (**M.S.**) in Computer Science* | **GPA: 3.83/4** — **(expected) Mar 2021**

Relevant Coursework: Distributed Systems, Databases, Operating Systems, Intensive Program Design, Machine Learning, NLP, Data Science, Social Network Analysis, Statistical Language Modelling, Intelligent Systems (Practicum)

Labs/Groups: Database Systems and Security Research Lab, Northwestern Network & Neighborhood Initiative (N3)

Rajiv Gandhi Technological Institute (RGPV) — **Bhopal, India**

*Bachelor of Engineering (**B.E. HONS**) in Computer Science & Engineering* | **CGPA: 8.63/10** — **Aug 2015 - Jun 2019**

Relevant Coursework: Algorithms, Data Structures, Cloud Computing, Object Oriented Design, Software Development Life Cycle

RELEVANT EXPERIENCE

Quantitative Data Analyst - *Institute for Policy Research, Northwestern University* — **Evanston, IL | Dec 2020 - Present**

- Modeled co-evolution of *Chicago Police Department*'s official assignment network data ***(18MM rows)*** with 22 years' of police misconduct data ***(CPDP: 43 tables)*** via *Goldfish* library in R to analyze tie evolution in *Dynamic Network Actor Models (DyNAMs)*.
- Defined and Calibrated network boundaries after extensive cleaning and detailed exploratory analysis in GCP's BigQuery engine.
- Identified individual, organizational and neighborhood factors responsible for burgeoning misconduct through police networks.

Research Assistant (Databases) - *McCormick School of Engineering, Northwestern University* — **Evanston, IL | Jan 2020 - Present**

- Contributed to *Prof. Jennie Rogers'* **NSF-funded** data federation project : ***VaultDB*** to securely query distributed hospital data.
- Generated ***aggregate operator*** functionality in C++ for the query execution engine in ***both encrypted and plaintext*** settings.
- Developed a robust test harness for secure *multi-party computation* for testing operators across ***TPC-H*** benchmark queries.
- Lead build-pipeline development in ***Docker***, aiding successful *deployment at 5* major participant hospitals *in Chicagoland* region.
- Enumerated query plans and integrated operator reordering in *Apache Calcite* to enhance *query optimizer's* performance by ***37%***.

Backend Engineering Intern - *Practicum LLC* — **St. Louis, MO (Remote) | Jun – Aug 2020**

- Developed a learning-cum-trading platform for novice quant traders returning an average market edge of ***5.76%*** over live trade.
- Designed database schema to handle data from YahooFinance API through endpoints in *Flask*, backtested up to five years' worth daily/hourly trade data, and translated financially proven strategies like *Statistical Arbitrage* and *Mean Reversion* to Python.

SELECT PROJECTS

RAFT protocol for Consensus and Leader Election — **Apr - Jun 2020**

- Integrated a reliable, fault-tolerant, thoroughly tested emulation of RAFT to achieve consensus-based leader election in GoLang.
- Significantly improved upon a simple primary-backup KV store, built atop a highly distributed implementation of Map-Reduce.

Social Network Mapping of Bali Bombing Terrorists — **May 2020**

- Scrutinized Bali bombers' network (between '02-'05) in R to statistically model the formation and dismantling of terrorist cells.
- Predicted missing ties, unidentified recruiters via ERGM & fine-tuned autologistic actor model to delineate 'terrorist' tendencies.
- Optimized SIENA for longitudinal analysis, identified relational shifts, and spotted potential (crucial) canaries in the network.

Health-centric U.S. restaurant trend analysis — **Feb - Mar 2020**

- Analyzed ***84,962*** US restaurants from Yelp (state-wise), classifying them on regional cuisine preferences and health metrics.
- Implemented end-to-end Data science pipeline in Python, and enhanced performance of Machine Learning models by ***8.745%.***
- Developed Tableau dashboard with an elaborate time-series predictive analysis to pilot re-opening of restaurants post COVID.

High availability Web server hosting through Fail-over Clustering — **Jun - Jul 2018**

- Managed data-center virtualization on VMWare Workstation cluster formation on MS Server-2008 R2 using SAN.
- Strategically deployed and tested fail-over cluster (with Quorum) for maintaining cloud-based services and web apps.

TECHNICAL SKILLS

Languages:	Java (preferred), Python (working proficiency), C++ (low-level dev), GoLang, R
Databases:	PostgreSQL, MySQL, Oracle 10g, MongoDB, SQL Server, CockroachDB, BigQuery
Machine Learning:	KNN, SVM, CNN, clustering, random forest, decision trees, sklearn, text mining, pandas
Web Development:	HTML/CSS, Javascript, Flask, API, REST, Postman, JSON, XML, MVC Model, SDLC
Tools:	Git, LINUX, VIM, Docker, MS Server R2, Hyper-V, Jupyter, Tableau, RStudio, AWS, GCP, Clion, IntelliJ

PUBLICATIONS

- Vishal Shrivastava, **Shashank Satish Damodaran** & Megha Kamble (2020) *Adalward: a deep-learning framework for multi-class malicious web page detection*, **Journal of Cyber Security Technology**, 4:3, 153-195, January 2020.
    - Presented a ***5-layer*** deep learning framework to identify and classify malicious websites as spam/phishing or spyware.
    - Hypothesized and verified a set of supervised models; achieving an accuracy of ***99.76% on 1MM+ web URLs***.

CERTIFICATIONS / ACTIVITIES

- **DBMS** (IIT-Kharagpur), **Databases for Data Science** (IBM), **Cloud Computing Associate** (CRISP), **Getting started with CockroachDB**
- ***Finalist*** (6/463) - ***Smart India Hackathon*** (software edition), *2019*; ***Winner*** - Wireless Robotics (IIT-Kanpur Zonal Round) - *2016*