Vishwani Sati

vishwani.sati@gmail.com | (+1) 872 806-6879 | 1021 Garnett Pl, Evanston, IL 60201 | linkedin.com/in/vishwanisati

EDUCATION

Northwestern University, The Graduate School — Evanston, Illinois
Master of Science in Computer Science — December 2022 (Expected)
- Cumulative GPA: 3.667/4.0

Amity University, ASET — Noida, India
Bachelor of Technology in Computer Science and Engineering — July 2020
- Cumulative GPA: 8.29/10.0

SKILLS

- **Programming Languages:** Python, R, C++, SQL, MATLAB, JavaScript, React, HTML, CSS
- **Big Data & Machine Learning:** Spark, Hadoop, MongoDB, Python (scikit-learn, NumPy, pandas, Matplotlib)
- **Data Science & Miscellaneous Technologies:** A/B Testing, ETL, Data Science Pipeline (Cleaning, EDA, Visualization, Modeling, Interpretation), Statistics, Time Series, Excel, Git

PROFESSIONAL EXPERIENCE

Amazon.com, Inc. — East Palo Alto, California
Software Development Engineer Intern, AWS Aurora — June 2022 – September 2022
- Built auroracr (v1.0.0), a developer tool to execute dry run build commands for mentioned packages into custom version sets
- Developed a Git Commit Formatting Checker into auroracr to perform package-specific checks on developer's commit message
- Vended auroracr to the BuilderToolbox; wrote User, Developer, and Test documentation; presented results to senior leadership

VMware, Inc. — Bengaluru, India
Associate Technical Engineer, vSphere — September 2020 – June 2021
- Applied knowledge of networking, storage, multi-platform Operating Systems (Windows, Linux, UNIX) to maintain business crucial applications in the cloud computing virtualization platform of vSphere
- Diagnosed and solved the critical technical issues pertaining to vCenter Server, ESXi hosts and Virtual Machines
- Certifications: VMware vSphere 6.7 Certified Professional, VMware vSphere 6.7 Foundations

SLAC National Accelerator Laboratory, Stanford University — Stanford, California
Research Intern, Data Analytics — November 2019 – January 2020
- Analyzed PingER data with 823K observations to measure and compare the internet network performance of India and Pakistan
- Utilized Python (scikit-learn, pandas) to identify 'K' using Silhouette method and implement clustering on the data, which concluded faster RTT for India than Pakistan by 8%

BISITE Research Group, University of Salamanca — Salamanca, Spain
Undergraduate Researcher, Machine Learning — August 2019 – December 2019
- Designed a wearable PPE (with sensors) to collect real-time data regarding worker's physical condition and the environment
- Utilized Python to develop a random forest classifier, which detects anomalous (risk) situations with an accuracy of 94.53%
- Publication: Sanchez S, Lecumberri F, **Sati V** et al. (2020). "Edge Computing Driven Smart Personal Protective System Deployed on NVIDIA Jetson and Integrated with ROS". CCIS, vol 1233. Springer. *doi.org/10.1007/978-3-030-51999-5_32*

Wisma R&D Lab, University of Malaya — Kuala Lumpur, Malaysia
Research Intern, Information Retrieval — May 2018 – June 2018
- Analyzed TREC data for 6 different Web Tracks to compute the Pearson correlation between 22 evaluation metrics, and derived the metric pairs with mean correlation measure of greater than 0.8
- Utilized R to implement a linear regression model to predict high-cost evaluation metrics from low-cost evaluation metrics, which reduced the cost of evaluation by 21%

PROJECTS and LEADERSHIP

Graduate Teaching Assistant, Northwestern University — Evanston, Illinois
Distributed Systems — March 2022 – June 2022
- Led recitations, graded coursework, and advised 60+ students on basic principles behind distributed systems and main paradigms used to organize them

Introduction to Computer Systems — January 2022 – March 2022
- Coordinated with instructor and 5 other TAs to host office hours, answer 70+ students' queries on conceptual framework of modern computer system and system programming in C in the Unix environment

Automated Data Cleaning — Evanston, Illinois
Project Member, Data Science — January 2022 – March 2022
- Developed dRTE, a data tool which accurately rectifies errors in the data layer to achieve data completeness and consistency
- Designed and implemented data quality rules to identify 'dirty data' and clean raw data, automating data quality monitoring

Sentiment Analysis on User Reviews — Evanston, Illinois
Project Member, Data Science Pipeline and Machine Learning — September 2021 – December 2021
- Analyzed Yelp dataset with 4000K observations to perform sentiment analysis on reviews using Liu Bing's lexicon method
- Utilized Python and R to implement an SVM classifier to predict the star ratings of the reviews with an accuracy of 91.81%