Supriti Ghosh

Software Engineer · ML Engineer · Ex- NetApp Inc

Pittsburgh, PA 15217, USA

(+1) 515-735-5090 | supriti.ghosh.ju@gmail.com | supritighosh.github.io | supritighosh | supritighosh

Summary

I am looking for a full-time job opportunity in software engineering. I have worked as a Software Engineer II at NetApp Inc and worked in the Cloud Volumes Service (CVS) team and Hybrid Cloud Engineering (HCE) team to design and develop software features. I take on new challenges with my passion and design technical solutions. By improving the competencies and skills, a strategic emphasis enables engineering to be a part of the expansion of the company.

Professional Experience

NetApp Inc

Cranberry Township, PA 16066

Software Engineer II — *June 2022 - Jan 2023*

- **Cloud Volumes Service (CVS):**
  - Developed CVS in the Storage as a service (STaas) model for backup services in **Microsoft Azure, Google Cloud Platform (GCP)** utilizing **Kubernetes, Golang, and MySQL**.
  - Built solutions for scheduled and on-demand disaster recovery tasks for Azure which includes the lifecycle management of scheduled volume replication using CRR (Cross-Region Replication), scheduled backup to cloud, and on-demand backup, and restore from an object store in the cloud.
- **Hybrid Cloud Engineering (HCE):**
  - Designed and developed software features in HCE and build solutions that help customers in the E-Series solutions automation team with **Ansible collections**.
  - Developed infrastructure to automate the deployment of E-Series solutions, especially the SANtricity, Host, and BeeGFS collections which also includes working around developing and testing in high availability (HA) solutions.

Core Member at WIT NANE

- Worked with the team to brainstorm ideas and thoughts to help women who are interested to grow in technology.
- Led a session in YWIT 2022 event to inspire young women to continue to pursue their interest in technology education and careers and and presented web scraping in python which means how to process of gathering information using automatic method to obtain large amounts of data from websites where most of this data can be unstructured data in Python. .

University of South Dakota

Vermillion, SD 57069

Graduate Research Assistant, Biocomputational Biological Engineering Lab, Biomedical Engineer — *January 2022 - May 2022*

- **Software Engineer:**

  Web Applications: Worked in a team of 11 engineers to develop and maintain search-engine based biomedical system used by million users and more than thousands scientists and developed web applications using **JavaScript**, **ReactJS**, **SailsJS**, **GraphQL** & **Elastic Search** to analyze and process data. Ensured applications security and ability to interact with multiple APIs and databases.

  NLP Applications: Designed NLP applications using **Python** to automatically categorize data. Used effective text representations to transform natural language into useful features. Trained and developed NLP models and evaluated them for deployment.

Graduate Research Assistant, 2AI: Applied Artificial Intelligence Research Lab, Computer Science — *January 2021 - December 2021*

- **Machine Learning (ML) Engineer:** To screen Covid-19, used **Python**, **TensorFlow**, **Keras**, **scikit-learn**, **Convolutional Neural Networks**. For the purpose of data visualization, used **Matplotlib** & **Tableau**. [GitHub]

  DNN in Chest X-ray to screen Covid-19: Implemented neural network on the balanced dataset to prevent from possible bias. Collected 10k chest x-rays to implement binary classification (Covid-19 vs Non-Covid-19). Used the different pre-trained models of deep learning to compare the performance. Generated lung segmentation and heatmap to screen and localize interpretable abnormalities.

  CheXNet to screen Covid-19: Implemented CheXNet on a dataset of 4.6k to screen Covid-19 using chest X-ray images. CheXNet is initially designed for radiologist-level pneumonia detection in chest X-rays (CXRs). Implemented feature extraction technique using deep learning which showed success in identifying Covid-19. [Publication] [Presentation]

  Literature Review of Covid-19 Screening Algorithms using Chest X-rays: Reviewed 50 peer-reviewed papers and compared the results for screening chest X-rays. Also collected their datasets and compared performance with relation to the size and quality of data. [Publication]

Coursera

Remote

Mentor, DeepLearning.AI — *August 2021 - June 2022*

- Mentored & guided learners from all over the world with programming methods, code reviews & technical problems.
- Checked requirement satisfaction and helped engineers with solutions and new ideas by actively participating in Discussion Forums.

NAMI (National Alliance on Mental Illness) Central Iowa *Ames, IA 50010*

Intern, Data Analyst *March 2020 - November 2020*

- Interpretation of data to analyze results using statistical training. Also developed and implemented data collection systems.
- Optimized statistical efficiency and quality. Filtered and cleaned data by reviewing reports using **Python**.
- Developed web platform with **PHP** and worked with management to prioritize business and information needs.

Jahangirnagar University *Dhaka, Bangladesh*

Research Assistant, Institute of Information Technology *January 2016 - December 2017*

- **Project: Identify Fingerprint Using Minutiae Matching in Biometric Security System** · This research presents a new method of identifying fingerprints in biometric security systems because the fingerprint is one of the best examples of biometric security and it can identify the personal information and it is much more secure than any other biometric identification system.
- The experiment was performed in **Matlab** & **SQL**. The experimental result exhibits the performance of the proposed method. [Publication] [GitHub]

Education

University of South Dakota (USD) *Vermillion, SD 57069*

MS in Computer Science *January 2021 - May 2022*

- **CGPA:** 3.8/4.0, **Thesis:** Deep Features to Analyze Pulmonary Abnormalities in Chest X-rays due to Covid-19 [Dissertation] [Presentation] [GitHub], **Advisor:** Dr. KC Santosh
- **Courses:** IoT & Security · Mathematics for Machine Learning · Seminar · AI in Medical Imaging · Design and Analysis of Computer Algorithms · Information Storage and Retrieval · Machine Learning Fundamentals · Advanced Artificial Intelligence

Jahangirnagar University (JU) *Dhaka, Bangladesh*

MS in Information Technology *January 2016 - January 2018*

- **CGPA:** 3.71/4.0, **Project:** Identify and Recognize Person Using Iris Biometric Security System [GitHub], **Advisor:** Dr. Mohammad Abu Yousuf
- **Courses:** Advanced Database System · Probability & Stochastic Process · Advanced Digital Communication · Information Systems & Securities · Client Server Technologies & Cloud Computing · Advanced Mobile & Cellular Communication Technologies · Software Project Management & Quality Assurance · Geographical Information Systems · E-Commerce & E-Governance · Computer Vision & Image Processing

Jahangirnagar University (JU) *Dhaka, Bangladesh*

BS in Information Technology *December 2011 - December 2015*

- **CGPA:** 3.58/4.0, **Thesis:** Novel Method to Assess Motion Blur Kernel Parameters and Comparative Study of Restoration Techniques Using Different Image Layouts [Publication] [GitHub] **Advisor:** Dr. Mohammad Abu Yousuf
- **Courses:** Data Structures · Algorithm Analysis · Digital Logic Design · Object Oriented Programming · Operating System · Computer Architecture · Wireless & Mobile Communication · Computer Network Security · Database Management System · Differential & Integral Calculus · Complex Variable & Vector Algebra · Statistical & Probability Theory · Discrete Mathematics · Computational Mathematics · Simulation & Modeling · Electronic Devices & Circuits · Information System Analysis · Computer Network & Internet Technology · Signal & System · Software Engineering · Computer Graphics · Web Technologies · Microprocessor & Interfacing · Telecommunication System · Management Information System · Introduction to Bio-informatics · Artificial Intelligence & Neural Network · Parallel & Distributed System · Multimedia System & Application · Human Computer Interfacing · Digital Image Processing & Pattern Recognition · Mobile Application Development

Technical Skill

Platforms	Microsoft Azure, Google Cloud Platform (GCP), AWS Sagemaker
Programming Language	Python, Golang, C, C++, Java, C#
DevOps	CI/CD, GIT, Linux, Agile Methodologies, Scrum
Cloud Technologies	Kubernetes, Ansible, RESTful API
Web Technologies	HTML, CSS, PHP, JavaScript, ReactJS, SailsJS, GraphQL, Elastic Search
Database	MySQL, PostgreSql, Tableau, Power BI, Oracle, Microsoft SQL Server
Machine Learning	Scikit-learn, Keras, TensorFlow, PyTorch
Frameworks & IDEs	Visual Studio, Goland, Netbeans, .NET, Matlab
Other Skills	Latex, Adobe Photoshop, Shell, CISCO Packet Tracer

Projects

Neural Network *Python*

Course Project for Machine Learning Fundamentals Course [GitHub] *January 2022 - May 2022*

- Used 400 images in total (200—200 - balanced data) in this project. Trained Neural Network with three different activation functions: a) linear, b) Sigmoid, and c) Tanh. Used k (=5)-fold cross-validation and computed the confusion matrix, precision, recall, and F1-score. Reported which kernel performs better of all.

Covid-19 Recognition in CT Scans using Artificial Intelligence (AI) guided tools

Python

Course Project for Advanced Artificial Intelligence Course [GitHub] *January 2022 - May 2022*

- AI-guided algorithms have been utilized to screen CT scans for Covid-19 analysis. A total of 1, 810 CT scan datasets have been collected for this project where 1, 267 Covid-19 patients' and 543 healthy patients' CT Scans. The pre-trained models InceptionNet V3 and U-net have been used for training purposes. K-fold cross-validation has been used to verify a better model. Our main goal was to show which model can perform better to detect, localize and segment Covid-19 cases using CT scan images so that we can use one or two globalized models for Covid-19 analysis.

Document Clustering

Python, scikit-learn, Matplotlib

Course Project for Information Storage and Retrieval Course [GitHub] *August 2021 - December 2021*

- In this project, used Python to cluster documents. At first, fetched Wikipedia articles and then represented each article as a vector. Performed k-means clustering to cluster documents and then evaluated the result.

Fake Job Detection

Python, NLP

Data Science Project [GitHub] *May 2021 - August 2021*

- Used Python libraries to detect fake jobs. The dataset contains of 18k job descriptions where around 800 false job descriptions is also included. Used logistic regression model because it can be used when the dependent variable is binary and also the dataset has been used to train and classify suspicious job descriptions.

Sensor Data Analysis for Internet-of-Things

Python, esp32, Raspberry Pi 4

Course Project for IoT & Security Course [GitHub] *January 2021 - May 2021*

- In this project, collected temperature and humidity data for continuously two hours on five different days using esp32, DHT11, MQTT, Raspberry Pi 4, breadboard and saved all the data in google firebase. For data plotting, used python libraries for line plotting, box plotting of temperature and humidity data and also line grid plots for three different days of temperature and humidity data. Then data has been pre-processed and linear regression models have been used for data analysis of MSE, AIC and also to plot the actual and predicted values.

IoT Big Data Management

PostgreSQL, Python

Course Project for IoT & Security Course [GitHub] *January 2021 - May 2021*

- Used PostgreSQL to manage big datasets for Internet-of-Things in this project. Also used python to change the time format (UNIT time to dd-mm-yy) and after that, included 4000 rows of data in the PostgreSQL to store and process.

Identify and Recognize Person Using Iris Biometric Security System

Matlab, SQL

Masters Project [GitHub] *August 2016 - December 2017*

- Used automated iris recognition for personal identification to verify both uniqueness of the human iris and also its performance as a biometric based system. The performance of research was measured for stored database which is scored 0% each for False Reject Rate (FRR) and False Accept Rate (FAR) and consequently, iris recognition is shown to be a precise and reliable biometric technology.

Hotel Management System

MySQL

Course Project for Information Storage and Retrieval Course [GitHub] *August 2021 - December 2021*

- Designed a database for the hotel management system. In this project, created relations between customers, HR, services of the hotel etc. It would allow the hotel management to handle all hotel activities.

Final Result Processing System

C#, MySQL

Semester Project for Software Engineering [GitHub] *July 2014 - December 2014*

- In this project, developed an application software in C# entitled "Final Result Processing System" which is a desktop application where the teachers can insert students databases, calculate results and Grade Point Average. The teachers could log in and update the student databases and results. The mark would be calculated automatically and saved in this software for future use. The application is built through C#, MySQL and provides the flexibility to add, modify or recreate new results for students.

Publications

International Conference/Journal Papers *Google Scholar*

- KC Santosh, **Supriti Ghosh**, Debasmita GhoshRoy, "Deep Learning for Covid-19 Screening using Chest X-rays in 2020: A Systematic Review" *International Journal of Pattern Recognition & Artificial Intelligence (IJPRAI).* [IJPRAI]
- KC Santosh, **Supriti Ghosh**, "CheXNet for the Evidence of Covid-19 Screening using 2.3K Positive Chest X-rays" *The 4^{th} International Conference on Recent Trends in Image Processing & Pattern Recognition (RTIP2R), December 2021.* [RTIP2R] [Presentation]
- **Supriti Ghosh**, Mohammad Abu Yousuf, "Novel Method of Identifying Fingerprint Using Minutiae Matching in Biometric Security System" *International Journal of Advanced Engineering, Management and Science (IJAEMS), ISSN: 2454-1311, Vol-2, Issue-7.* [IJAEMS]
- Munira Akter Lata, **Supriti Ghosh**, Farjana Bobi, Mohammad Abu Yousuf, "Novel method to assess motion blur kernel parameters and comparative study of restoration techniques using different image layouts" *5^{th} International Conference on Informatics, Electronics and Vision* (ICIEV 2016), Dhaka, Bangladesh. [IEEE XPlore]

Presentation

2022	**YWIT 2022**, Presented in the YWIT 2022 at the NetApp Inc.	*C. Township, PA*
2022	**IdeaFest 2022**, Presented research in the IdeaFest 2022 at the University of South Dakota (USD).	*Vermillion, SD*
2021	**RTIP2R'2021**, Presented paper in the International Conference on Recent Trends in Image Processing & Pattern Recognition (RTIP2R).	*Msida, Malta*
2021	**IdeaFest 2021**, Presented research proposal in the IdeaFest 2021 at the University of South Dakota (USD).	*Vermillion, SD*
2021	**CSC 790 - Seminar**, Presented my research in the seminar lecture series of department of computer science at the University of South Dakota (USD).	*Vermillion, SD*

Certifications

Specialization Courses Offered on NetApp Inc

- Using SANtricity Software · NetApp · [Certification]

Specialization Courses Offered on Coursera

Google Data Analytics Professional Certificate · Google · [Certification]

- Foundations: Data, Data, Everywhere · [Certification]
- Ask Questions to Make Data-Driven Decisions · [Certification]
- Prepare Data for Exploration · [Certification]
- Process Data from Dirty to Clean · [Certification]
- Analyze Data to Answer Questions · [Certification]
- Share Data Through the Art of Visualization · [Certification]
- Data Analysis with R Programming · [Certification]
- Google Data Analytics Capstone: Complete a Case Study · [Certification]

PostgreSQL for Everybody · University of Michigan · [Continuing]

- Database Design and Basic SQL in PostgreSQL · [Certification]
- Intermediate PostgreSQL · [Certification]
- JSON and Natural Language Processing in PostgreSQL · [Continuing]

Applied Data Science with Python · University of Michigan · [GitHub] [Certification]

- Introduction to Data Science in Python · [GitHub] [Certification]
- Applied Plotting, Charting & Data Representation in Python · [GitHub] [Certification]
- Applied Machine Learning in Python · [GitHub] [Certification]
- Applied Text Mining in Python · [GitHub] [Certification]
- Applied Social Network Analysis in Python · [GitHub] [Certification]

Practical Data Science · DEEPLEARNING.AI & AMAZON WEB SERVICES · [GitHub] [Certification]

- Analyze Datasets and Train ML Models using AutoML · [GitHub] [Certification]
- Build, Train, and Deploy ML Pipelines using BERT · [GitHub] [Certification]
- Optimize ML Models and Deploy Human-in-the-Loop Pipelines [GitHub] [Certification]

Python 3 Programming · University of Michigan · [GitHub] [Certification]

- Python Basics · [GitHub] [Certification]
- Python Functions, Files, and Dictionaries · [GitHub] [Certification]
- Data Collection and Processing with Python · [GitHub] [Certification]
- Python Classes and Inheritance · [GitHub] [Certification]
- Python Project: pillow, tesseract, and opencv · [GitHub] [Certification]

Python for Everybody · University of Michigan · [GitHub] [Certification]

- Programming for Everybody (Getting Started with Python) · [GitHub] [Certification]
- Python Data Structures · [GitHub] [Certification]
- Using Python to Access Web Data · [GitHub] [Certification]
- Using Databases with Python · [GitHub] [Certification]
- Capstone: Retrieving, Processing, and Visualizing Data with Python · [Certification]

Courses Offered on Coursera

- **Machine Learning with Python** · IBM · [Certification]

Courses Offered on Google Digital Garage

- **The Fundamental of Digital Marketing** · Google · [Certification]

Symposium Offered on Jahangirnagar University

- **CCNA Voice Primer**
- **Demystifying Neural Networks In-Vivo: A Computational Approach**
- **Database Design and Implementation with Web Application**

Activities

Volunteer Opportunities

- **Core Member of WIT NANE**, NetApp Inc, Cranberry Township, PA 16066. [July 2022 - January 2023]
- **Session Lead at YWIT 2022**, NetApp Inc, Cranberry Township, PA 16066. [September 2022]
- **Journal Reviewer**, SN Computer Science, Remote [October 2022]
- **AI Ethics Mentor**, Teens in AI x Harvard x MIT Hackathon, Cambridge, MA. [March 2022]
- **Social Media Coordinator**, Association for Computing Machinery (ACM), University of South Dakota, Vermillion, SD 57069. [July 2021 - May 2022]
- **AI Ethics Mentor**, Teens in AI, San Francisco, CA. [March 2021]
- **Mentor**, YSS (Youth Standing Strong), Ames, IA 50010. [Sept 2019 - Oct 2020]
- **Program Coordinator (Professional Activities)**, IEEE Student Branch, Jahangirnagar University, Dhaka, Bangladesh. [May 2015 - May 2016]
- **Volunteer (Technical Activities)**, International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT). [May 2015]

Honors & Awards

2018	**University Merit Scholarship**, Top 10% in Graduate Level, Jahangirnagar University	*Dhaka, BD*
2016	**University Merit Scholarship**, Top 20% in Undergraduate Level, Jahangirnagar University	*Dhaka, BD*

Organizations

ACM	**Member**, Association for Computing Machinery (ACM)	*2020 - Present*
IEEE	**Student Member**, Institute of Electrical and Electronics Engineers (IEEE)	*2015 - 2016*

Research Experience

Data Science	Artificial Intelligence, Machine Learning, Natural Language Processing
Deep Neural Network	Convolutional Neural Network, Object Detection
Healthcare Data Analytics	Pattern Recognition, Computer Vision, Image Processing