

SANKALP SINGH GAHARWAR

704-232-8126 | sankalpgaharwar@gmail.com | linkedin.com/in/sankalpgaharwar
github.com/sankalpg10 | public.tableau.com/app/profile/sankalp.gaharwar1646

EDUCATION

North Carolina State University| Raleigh, NC| Aug 2019 – May 2021

Master of Computer Science| CGPA - 4.0/4.0

Relevant Coursework: Software Engineering, Design and Analysis of Algorithms, Human-Computer Interaction, Deep Learning, Neural Networks, Machine Learning and Data Analysis, Privacy in Digital Systems, Internet Protocols, Applied Statistics, Distributed Systems

Guru Gobind Singh Indraprastha University| New Delhi, India| Aug 2014 – Aug 2018

Bachelor of Technology in Computer Science

Relevant Coursework: Machine Learning, Data Structures, Operating Systems, Database Management Systems, Web Engineering

SKILLS

Programming Languages: Java, Python, C++, GoLang

Web-Development Technologies: Flask, Django, JavaScript, Angular, Node.JS, ASP.Net, React, jQuery, Bootstrap, HTML5, Web API

Databases: MySQL, SQL, MongoDB, NoSQL, PostgreSQL, Amazon S3

Machine Learning and Data Analytics: Large-Language Models (LLM), Prompt Engineering, ChatGPT, GPT-4, Scikit-Learn, SciPy, NumPy, Pandas, TensorFlow, PyTorch, PySpark, Keras, Matplotlib, Regression, Classification, AdaBoost, XgBoost, KNN

Tools and Frameworks: Micro Services, AWS Lambda, AWS SageMaker, SOAP UI, SOAP Web Services, REST Web Services, Spring, Maven, JUNIT, Git, Linux, Docker, MVC, REST API, Hadoop, Tableau, ECL, Hive, DataSwarm, Microsoft Azure, DataBricks, AWS EC2, JIRA, YAML, Jenkins, Amplitude, Datadog, Big Query, Kubernetes, MS Power BI

WORK EXPERIENCE

Lexis Nexis Legal Solutions | Software Engineer| Raleigh, NC| September 2021 – Present

- Worked on building a **AWS Cloud-based Search-Engine** solution for Legal Analytics, using tools such as **Python-based Natural Language Processing NLP libraries, Web API, AWS Cloud Services, ECL libraries** and **MySQL**, to leverage relevant information for enhancing customer experience, and reduced customer-based error report frequency by **35%**.
- Led the development and deployment of a multi-regional Legal Drafts Assistance Tool, achieving **80%** improvement on user error-report feedback benchmarks, by optimizing the Jurisdiction Recognition Service, through **LLM-driven techniques**, achieving a **98% recall rate**, and minimizing false-negatives by integrating CRF with negative sampling. Fine-tuned **miniLLM-Cross-Encoder** on **AWS SageMaker** for enhancing **precision score results by 11%**.
- Improved the re-ranking performance of grounding materials in terms of being aligned with user requirements, by developing a **hybrid ML re-ranker using BERT, BM25, ElasticSearch**. Implemented **AWS Lambda** micro-service solutions for **Advanced Filtering and Boosting** algorithms to boost NDCG scores for text recommendations by **17%**.
- Contributed to developing a high performance computer cluster system that uses **Big Data analytics** and **Python-based machine learning libraries** on **AWS EC2 instances** to classify legal content for automating Data Mining tasks.
- Employed Human-Computer-Interaction-based prototyping methods to improve existing **ReactJS** UIs for performing CRUD operations on metrics. Reduced average query times for customers by **91%** across a cross-regional user base.

NC Department of Transport| Software Developer| Raleigh, NC| Dec 2020 – August 2021

- Built a distributed client-server application in **Java**, using containerized Micro Services on **Docker**, that combines effective Data Analytics for **Natural Language Processing** with Data Visualization capabilities through **Tableau**, for providing intuitive web dashboards. Deployed the application on **AWS EC2** instance.
- Formulated the tool as a Web Application employing **REST API Web Services** using **Spring Boot, SOAP Web Services** with **SOAP UI**, and **Maven** Configuration with support for unit testing through **JUNIT**.
- Devised an end-to-end framework to establish Single Source of Truth for 3200 IDC metrics spread over 500+ **Tableau** dashboards
- Conceptualized the system architecture, assigned user stories, and defined workflows for dedicated role-based dashboards, with customized views, functionality and access rights.

NTT Data Services Pvt. Ltd.| Software Development Senior Associate| Noida, India| Jun 2018 – Jul 2019

- Designed and programmed high performance, cloud-based web applications based on the **Smart UI architecture**, through **MVC-based** models on **Django, Flask**, and **ASP.Net-MVC5, Micro-services, Web API, RESTful API**, and **Entity framework**, for accomplishing a diverse set of tasks ranging from database access to distributed deployment over **Azure Cloud Infrastructure**.
- Built data ingestion tool for importing tables from **Oracle, MYSQL, Sybase** databases to **Hadoop** using **Apache Oozie** and **Flask** for dynamic job orchestration and processing of **10M+** rows data using **Airflow, PySpark, Azure Databricks & SQL**

Tata Consultancy Services Ltd. | Software Engineering Intern | Gurugram, India | Jun 2017 – Aug 2017

- Engineered interactive server-side web applications for real-time vehicular tracking using **ReactJS**, **AngularJS**, **Relay**, **GraphQL**, **NodeJS**, **MongoDB**, and **Redux**.
- Implemented the backend aspects of the application for real-time data access by employing **Node.JS** with **MongoDB**.
- Constructed Nodes/Edges Graphical (NEG) Model using **YAML & GraphQL** to connect multiple entities with 100K+ data points