NEHA KOLAMBE

+1 (720) 641-1565 neha.kolambe@colorado.edu linkedin.com/in/nehakolambe/ github.com/nehakolambe neha.kolambe

EDUCATION

University of Colorado, Boulder — **August 2024 – May 2026**
Master of Science in Computer Science, ***GPA: 4.0/4.0*** — *Boulder, USA*

Sardar Patel Institute of Technology — **August 2018 – May 2022**
Bachelor of Technology in Information Technology, ***GPA: 9.07/10*** — *Mumbai, India*

TECHNICAL SKILLS

Languages: Python, Java, C++, SQL, Bash
Frameworks and Libraries: Hive, JDBC, Spark, REST APIs, Junit, Cucumber, Flask, Docker, Spring, Spring Boot, Django, Tkinter, Swing, AWT, TensorFlow, Pandas, NumPy, Data Visualization, Pytorch
Databases: MySQL, Oracle, MSSQL, PostgreSQL
Tools and Technologies: Maven, Gradle, Git, YAML, JSON, Terraform, IntelliJ, Eclipse, Microsoft Visual Studio, PyCharm, Mockito, Openshift Fabric, Bitbucket, WordPress, Tableau, Helm, Jenkins, Excel, Jupyter
Cloud and Distributed Computing: Kubernetes, GCP, Google Cloud Compute, Google Cloud Storage (GCS), Google Kubernetes Engine (GKE), Firestore, Cloud Run Functions, RabbitMQ, Apache Airflow, Hadoop
Operating Systems: Linux, macOS, Windows

EXPERIENCE

Splunk, a CISCO compnay | *Security Data Scientist Intern* — **June 2025 – September 2025**

- Designed and deployed an AI-based **Smart Alert system** to reduce manual review of large volumes of alerts, built on ∼**400k security alerts** and **100+ engineered features** to automate classification and streamline analyst workload.
- Developed an interactive **UI** integrated with the **ML pipeline**, enabling one-click application of AI-suggested actions or manual overrides, ensuring flexibility and analyst trust.
- Project selected to be presented at **.conf**, Splunk's flagship global **security and data conference**, recognizing its innovation and **impact**.

Deutsche Bank | *Technology Senior Analyst* — **July 2022 – July 2024**

- Built a fault-tolerant **data ingestion pipeline** in **Java** and **Spring Boot** to consolidate data from diverse sources into **Oracle**, improving reliability by **35%** and streamlining **NII calculations**.
- Designed and implemented an **automated testing framework** for onboarding processes, reducing manual QA by **50%** and accelerating **release cycles**.
- Enhanced **CI/CD pipelines** using **TeamCity** and deployed **containerized services** on **OpenShift**, reducing deployment effort by **30%**.
- Delivered **reconciliation features** through tight collaboration with stakeholders, ensuring **100% on-time delivery** and accurate **financial reporting**.

PROJECTS

SemEval-2025 Task 8 — **April 2025**
Collaborated on a research project for **SemEval 2025 Task 8**, developing question answering systems over tabular data using **LLaMA 3 models**. Implemented and evaluated three **LLM-based strategies**: **baseline full table input**, **retrieval augmented generation (RAG)**, and **two phase prompting with column filtering**. Leveraged **FAISS indexing** and **prompt engineering** to optimize performance across diverse question types. Achieved up to **49.7%** accuracy, surpassing prior benchmarks and demonstrating improvements in answer quality and column selection.

VoxOff: AI-Powered Karaoke Web App - Link — **April 2025**
Built a full-stack karaoke web application with real-time vocal separation, forced lyrics alignment, and a synchronized lyrics player. Designed a hybrid architecture combining local preprocessing with **GKE-based microservices** for scalable audio and lyrics processing. Integrated **YouTube** and **Genius APIs**, and deployed infrastructure using **RabbitMQ**, **Terraform**, **Firestore**, and **Google Cloud Storage**. Implemented **Google OAuth authentication** for account-based history and personalized recommendations. Achieved **2–3 minute processing per song** and showcased at **CU Boulder's ATLAS Expo**, processing **100+ songs without failure**.

NarrateNow: Chapter-wise audiobook generation service - Link — **December 2024**
Developed a Python-based scalable application that converts EPUB books into chapter-wise audiobooks using the **Google Text-to-Speech API**. Designed a **Kubernetes-based microservice architecture**, employing **RabbitMQ** and **Redis** for task orchestration and real-time tracking, while integrating **Google Cloud Storage** for secure data handling. The service processed an 11-chapter book (150 pages) in under **4 minutes** and demonstrated the ability to handle **50+ simultaneous requests** without errors, emphasizing reliability and scalability.

MoodyTunes: A DL-based Music Mood Classifier - Link — **December 2024**
Developed a **deep learning-based music classification system** using **Python**, **TensorFlow**, and **Keras** to categorize songs into emotional moods as **happy**, **sad**, **calm**, and **energetic**. Integrated enriched datasets from **Spotify (278k songs)** and **Last.fm**, achieving **90.67% accuracy** by leveraging **MFCCs**, **spectral contrast**, and **user tags**. Optimized neural networks such as **ANN**, **RCNN**, and **LSTM**, resulting in an average performance boost of **4%** across all models.

PUBLICATIONS

Detection of Mental Disorder on Social Platform - Link — **February 2022**
Engineered an **LSTM-based prediction model** with **97%** reliability to identify mental disorders based on consumer behavior on social platforms. The platform offered **comparative analysis** of data over specific periods, aiding users in assessing potential conditions. A **custom survey tool** collected behavioral evidence, providing **actionable insights** into trends, while **advanced machine learning techniques** ensured the model's **robustness** and **reliability**.