DEVARATA OZA

New Brunswick, New Jersey, USA, 08901 ☏ 732-522-6179 ✉ do309@rutgers.edu

in https://www.linkedin.com/in/devarata/ ☐ https://github.com/devarata ⊕ https://devarata.vercel.app/

EDUCATION

Rutgers University, New Brunswick Sep 2022 – May 2024
Master of Science in Computer Science *GPA: 3.8/4.0*

Thadomal Shahani Engineering College, Mumbai Aug 2015 – May 2019
Bachelor of Engineering in Computer Engineering *CGPA: 8.34/10*

TECHNICAL SKILLS

Languages & Databases: Python, C++, JavaScript, Java, Shell, SQL, Apache Hive, Oracle, MySQL, MongoDB

Big Data & Cloud: PySpark, Apache Spark, SparkSQL, Apache Hadoop, HDFS, MapReduce, AWS, Docker

ML & Analysis: Spark MLlib, TensorFlow, Pytorch, Pandas, NumPy, Matplotlib, scikit-learn, NLTK, spaCy, OpenCV

Web Development & DevOps: ReactJS, BeautifulSoup (bs4), Requests, Git, Airflow, CI/CD

WORK EXPERIENCE

Center for Advanced Biotechnology and Medicine, Rutgers University Jun 2023 – Sep 2023
Research Assistant (Machine Learning Engineer) *New Brunswick, NJ*
Project: Mice Monitoring System | **Skillset:** *PyTorch, Deep Learning, Computer Vision, ReactJS, Raspberry Pi, SQL*
- **Created scalable cost-saving mice cage monitoring system** with 160 cages per rack using **YOLOv8 model** trained on a custom dataset, comprising over **10000 annotated images** from video recordings and web scraping.
- **Designed Pi-based night vision system** for mice monitoring, processing video every 10s with **optimized storage**.
- **Developed 3 ML models** (mice counting, injury detection) and **mice disease detection** (Alopecia, Dermatitis, Tumors, Nasal discharge), while designing a **web interface using ReactJS** integrated with **Tecniplast API**.

Deutsche Bank Jul 2019 – Jul 2022
Senior Analyst (Data Scientist) *Pune, India*
Project: Rate Card and Matter-Invoice | **Skillset:** *PySpark MLlib, sklearn, NLP, Spacy, NLTK, matplotlib, Hive, SQL, HQL*
- Designed **machine learning-based multiclass classification system** for **automated invoice classification**, processing **approx. 4 million invoices in minutes**.
- **Optimized hyperparameters** to elevate **XGBoost Classifier** accuracy from **92.7%** to **97.8%**.
- **Boosted classification** by **20%** through **feature engineering** and **spaCy's Named Entity Recognition**.
- Employed **PySpark, Apache Hadoop,** and **Apache Hive** for Big Data processing of **5 GB daily** in Parquet format.
- Developed **Error Logging System** for Spark jobs and data pipelines, **enhancing troubleshooting by 50%**.
- **Led migration** of 1 TB+ legal historical data from **Oracle to Big Data** for enhanced processing.
- **Designed 200+ SQL queries** for data consolidation across data marts, adhering to business logic.

PERSONAL PROJECTS

Portfolio Website | *Next.js, TypeScript, Tailwind CSS, React Context API, Framer Motion* Dec 2023 - Jan 2024
- Developed responsive **Next.js web app**, incorporating **TypeScript** for type safety, **Tailwind CSS** for UI design.
- Integrated **React Context API** for efficient **state management** and Framer Motion for advanced interactive animations and smooth scrolling effects.

Walk-these-ways | *PyTorch, Isaac Gym Simulator, Reinforcement Learning, Actor-Critic Policy* Sep 2023 - Dec 2023
- Developed neural network for **robotic dog** in Isaac Gym, focusing on efficient **locomotion** & collision-free navigation.
- Implemented **Actor-Critic Proximal Policy (PPO)** for **precise, collision-free movement** in Isaac Gym, combining **reinforcement learning** with **integrated obstacle avoidance** for **adaptive navigation**.

BuyMe - Online Auction System | *HTML, JavaScript, Java, JDBC, SQL, MySQL, Tomcat* Mar 2023 - Apr 2023
- Designed eBay-like relational database, efficient user interaction and **data modeling** for enhanced system performance
- Crafted **user accounts, login system, bidding mechanisms, and advanced search functionalities**, along with administrative tools for robust auction management.

Music Recommender System | *Annoy, Tensorflow, Python, Docker* Jul 2021 - Aug 2021
- **Developed recommendation system** with **Annoy**, achieving a cosine similarity of **0.85** among top 10 neighbors.
- **Processed 13,000 audio examples**, generating **128D embeddings per sec** with Max Audio Embedding Generator.
- Implemented Annoy with **100 trees** for fast music recommendations, achieving an **average response time of 50 ms**.

ACHIEVEMENTS & AWARDS

- Recognized as department's best-performing employee, receiving **Deutsche Bank Recognition** Award for H2 2020.
- Awarded 1st place at TSEC Project-Expo'19 for **Smart Waste Management** , Internet of Things (IoT) Project