

EDUCATION

Master of Science in Computer Science (3.75 GPA) <i>University of Central Florida</i> <i>Relevant Coursework:</i> Data Structures and Algorithms, Computer Architecture, Data Mining, Machine Learning, Natural Language Processing	May 2024 Orlando, USA
Bachelor of Engineering in Electronics and Communication (7.39 CGPA) <i>Visvesvaraya Technological University</i>	Jun 2019 Bangalore, India

SKILLS

Languages	: Python, SQL, C, R
Cloud/Database	: AWS, Oracle, IBM Db2, Sybase, Hive
Frameworks	: NumPy, Informatica, Tableau, Git, Apache Airflow, Power BI, Hadoop, UNIX, PyCharm, Autosys, VS Code, GitHub, JIRA

PROFESSIONAL EXPERIENCE

Graduate Research Assistant – <i>University of Central Florida</i>	<i>Aug 2023-present</i>
<ul style="list-style-type: none">Conducting in-depth analysis of students’ demographic data using Physics Online Learning Modules (OLM) for the National Science Foundation (NSF), employing advanced R analytics and visualization techniques.Investigating the intricate relationship between UCF students’ characteristics, leveraging statistics to derive actionable insights.	
Data Engineering Intern – <i>City of Charlotte</i>	<i>Jun 2023-Aug 2023</i>
<ul style="list-style-type: none">Collaborated with stakeholders and business analysts to gather requirements and understand the company's data needs.Streamlined the data cleaning process for unstructured crime dataset by developing automated scripts in Python, reducing manual effort by 50%.Planned and executed data-driven solutions by deploying data pipelines in Airflow and building scalable database on AWS, resulting in a 40% increase in operational output.Developed timelines for individual projects to deliver work products in sprints in Agile framework. Leveraged git and Jira for deployment, tracking, and cataloguing work to communicate progress to direct leadership.	
Graduate Research Assistant (Data Analytics) – <i>University of Central Florida</i>	<i>Jan 2023-May 2023</i>
<ul style="list-style-type: none">Demonstrated exceptional mastery in utilizing R packages to efficiently process and analyze extensive CDC Medicaid datasets, providing invaluable insights into factors influencing periodontal health outcomes among U.S. adults.Applied statistical methodologies and successfully interpreted results to create informative regression tables, making intricate information accessible and comprehensible.Expertly conveyed key findings, facilitating informed decision-making in CDC research and policy formulation.	
Associate Data Engineer – <i>Capgemini</i>	<i>Apr 2020-Aug 2022</i>
<ul style="list-style-type: none">Participated in full data engineering cycle, requirements gathering across cross-functional teams, initial data analysis, data modeling, unit and integration testing, deployment, monitoring, and maintenance of applications in Data Governance for a Fortune 100 Global banking firm.Designed, developed, and organized ETL workflows to map data from large-scale data structures to generate insights and address business needs.Achieved a 90% enhancement in SQL query performance through optimization and created dashboards to help developers track their progress and prioritize tasks which increased the work efficiency.Developed scripts to run and streamline data fetched by complex queries and built a dashboard to monitor the hygiene of Sybase database and organize potential risks to handle. Received appreciation for developing an efficient risk monitoring portal that reduced production issues by 25%.Designed and scheduled data pipelines in Autosys, reducing processing time by 30% and ensuring data availability for timely decision-making.	
Data Engineer – <i>Capgemini</i>	<i>Dec 2019-Mar 2020</i>
<ul style="list-style-type: none">Strategically designed and implemented robust data models (Snowflake, Star schema) within Oracle database, tailored to the unique demands of Zomato's online food delivery data in Bangalore, India.Spearheaded ETL pipeline development utilizing Talend, seamlessly harmonizing diverse data sources.Employed Power BI for dynamic visualization and report creation, empowering data-informed decision-making, leading to a noteworthy 30% enhancement in client business operations.	

PROJECTS

Smart Summarizer for CNN News Articles	<i>Apr 2023</i>
<ul style="list-style-type: none">Developed a summarization model for complex 2.5L+ CNN/Dailymail news articles.Designed and implemented abstractive (BART, LSTM) and extractive techniques (BERT, TextRank) for comprehensive and accurate content summaries.	
BRFSS Early Diabetes Prediction	<i>Dec 2022</i>
<ul style="list-style-type: none">Prepared comprehensive analysis to identify crucial variables contributing to the onset of Diabetes among US citizens.Employed exploratory data analysis, imputation methods, regression techniques and hyperparameter tuning to boost accuracy up to 98%.	
Credit Card Fraud Detection	<i>Nov 2022</i>
<ul style="list-style-type: none">Led and implemented an end-to-end ML project pipeline to classify fraud bank transactions from 2,84,000+ imbalanced dataset.Leveraged Apache Airflow to automate the pre-processing phase of the project, enhancing efficiency and responsiveness.	

PUBLICATIONS AND CONFERENCES

BRFSS Data Analysis : Early Diabetes Prediction	<i>In-Progress</i>
Attending Grace Hopper Celebration 2023, Orlando, FL	<i>09/2023</i>
2023 KPMG Women in Advisory Summit, Orlando, FL	<i>08/2023</i>

CERTIFICATIONS

Google Data Analytics Professional Certificate, Google, present Oracle Database SQL Certified Associate, Oracle IBM Data Science Specialization, IBM	
--	--

CO-CURRICULARS

Math Tutor , Government School – Bangalore, India	<i>2022</i>
Fellow, Girls Who Code , Rewriting the Code	<i>2023</i>