CHAITHANYA SAI KARNE

609-977-4367 ckarne@seas.upenn.edu linkedin.com/in/chaithanyasaikarne github.com/ckarneupenn

EDUCATION

University of Pennsylvania, Philadelphia, PA **Aug 2021 – May 2023**
Master of Science in Engineering in Data Science *CGPA: 3.75/4.0*

- **Relevant Coursework:** Machine Learning, Statistics for Data Science, Deep Learning for Data Science, Programming (Python and Java), Big Data Analytics, Database Design, Computational Linguistics

Manipal Institute of Technology, Manipal, India **Jul 2016 – Jul 2020**
Bachelor of Technology in Mechatronics Engineering; Minor Specialization in Data Science *CGPA: 4.0/4.0*

TECHNICAL SKILLS

Programming Languages: Python, Java, R, Apache Spark, JavaScript
Database Technologies: SQL, MySQL, Oracle, NoSQL (MongoDB, Neo4j), Redshift, Hadoop, Hive, Google BigQuery, AWS RDS
Machine Learning Libraries: PyTorch, Keras, Scikit-Learn, OpenCV, Matplotlib, Pandas, Numpy
Data Science/Miscellaneous: Data Mining, Data Wrangling, NodeJS, ReactJS, Git, Tableau, ETL, Excel, AWS - S3, Athena, SageMaker

PROFESSIONAL EXPERIENCE

Graduate Data Management Assistant **Jan 2023 – Present**
The Wharton School, University of Pennsylvania *Philadelphia, PA*

- Managed & formulated research summaries for 8 of the 'Wharton Research' data sets using Pandas library for EDA in Python

Graduate Teaching Assistant **Sep 2022 – Present**
University of Pennsylvania, for the 'Programming Languages and Techniques (CIT590)' course *Philadelphia, PA*

- Collaborated with a team of 9 assistants for designing 5+ Python & JAVA learning sessions to guide 100+ students
- Conducted 2 office hours, 1 recitation, responded to 10 student questions and graded 5 assignments weekly

Graduate Technical Assistant **Sep 2022 – Dec 2022**
Wharton Customer Analytics - Analytics Accelerator program, The Wharton School *Philadelphia, PA*

- Associated with Lowe's Companies, Inc., to identify 10+ best offline (BOPUIS) products which were found to cost below $40, for Lowes' backroom space, using Python based clustering and regression

Graduate Research Assistant **May 2022 – Aug 2022**
ESG & Political Risk Research Lab, The Wharton School, University of Pennsylvania *Philadelphia, PA*

- Developed 20+ Python scripts & 30+ AWS Athena SQL queries for projects aimed at examining the influence of businesses on social landscape by analyzing online news articles (GDELT data) of 260+ countries
- Analyzed over 100 million rows/articles using AWS S3 & SageMaker, devised 2 funded data pipelines by mapping location names in articles' text to PRIO GRID cells, ran BERT model to visualize articles' sentiments

Business Analyst Intern **Jan 2020 – Jul 2020**
MiQ Digital India (P) Ltd *Bangalore, India*

- Assessed the impact of 10+ global advertising campaigns & formulated recommendations for digital spend optimization
- Automated 5+ analytical workflows by streamlined Data Engineering, Predictive Modeling which helped generate higher revenue
- Explored customer density analysis & web scraping for 2 top tier clients utilizing R, Python, SQL, Redshift, Hive & Excel

ACADEMIC PROJECTS

Image Outpainting using Generative Adversarial Networks (GANs) | *Python, Scikit-Learn, PyTorch, CNN* **Feb 2023 – Present**

- Implement GANs using context encoders & single image generative models (CNNs) as baselines to predict the region beyond the borders of an image
- Try to build advanced models with local and global discriminator, & residual blocks to beat state of the art MSE score of 0.78

Weather Application for Planning Holidays | *SQL, AWS, DataGrip, NodeJS, ReactJS* **Sep 2022 – Dec 2022**

- Designed a web application using NodeJS and ReactJS that queries a MySQL database hosted on AWS, containing 1.6 million rows of weather data, to assist users to plan future holidays
- Optimized database queries (SQL) using caching, indexing and efficient projections to reduce query run time by 90%

Visual Question Answering | *Python, Scikit-Learn, Pandas, PyTorch, NLP* **Sep 2022 – Dec 2022**

- Built a Bottom Up Top Down visual question answering model with modified FasterRCNN weights for bottom up layer and attention based LSTM model for top down approach resulting in a VQA score of 0.77

DataCo Global Supply Chain Analysis & Late Delivery Prediction | *Python, Pandas, Seaborn, SQL* **Mar 2022 – May 2022**

- Performed Exploratory Data Analysis (EDA) and generated insightful temporal and geospatial visualizations for the DataCo Global supply chain data
- Achieved 75-85% accuracy using Random Forest, Logistic Regression & KNN models, while predicting 'late delivery' risk of orders

PUBLICATIONS

'KCSNBShiny' & 'KCSKNNShiny' R Packages | *R, R Shiny*

- Published and deployed 2 R packages (July 2019) onto CRAN (Comprehensive R Archive Network), applicable for predictive analysis, employing Naive Bayes & k-nearest neighbors algorithms to predict any variable in categorical and numeric data