Chaithanya Sai Karne

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

Education

University of Pennsylvania **Aug 2021 – May 2023**
Master of Science in Engineering in Data Science *CGPA: 3.78/4.0*
- **Relevant Coursework:** Applied Machine Learning, Statistics for Data Science, Programming Languages and Techniques (Python and Java), Big Data Analytics, Databases (Database Design), Computational Linguistics

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

Technical Skills

Languages: Python, Java, R, MATLAB, JavaScript
Database Technologies: SQL, MySQL, Oracle, MongoDB, Hive, DataGrip, Redshift, AWS RDS
Technologies/Frameworks/Others: PyTorch, Keras, Scikit-Learn, Apache Spark, Data Mining, Data Wrangling, NodeJS, ReactJS, Git, GitHub, Matplotlib, Seaborn, Eclipse, LaTeX, Tableau, ETL, AWS - S3, Athena, SageMaker

Experience

Graduate Technical Assistant **Sep 2022 – Present**
Wharton Customer Analytics - Analytics Accelerator program, The Wharton School *Philadelphia*
- Associated with Lowe's Companies, Inc., to identify the products a given store should keep a limited backroom stock of using Python based Machine Learning models
- Collaborated with Align Technology to optimize the product marketing of the product 'Itero' & maximize sales

Graduate Teaching Assistant **Sep 2022 – Present**
University of Pennsylvania, for the 'Programming Languages and Techniques (CIT590)' course *Philadelphia*
- Guided 100+ students by designing 3+ advanced Python & JAVA learning sessions, preparing, and grading exams

Graduate Research Assistant **May 2022 – Aug 2022**
ESG & Political Risk Research Lab, The Wharton School, University of Pennsylvania *Philadelphia*
- Developed 20+ complex Python scripts & 30+ AWS Athena SQL queries for projects aimed at examining the influence of businesses on the social landscape by analyzing online news articles (GDELT data lake) of 260+ countries
- Analyzed over 100 million rows/articles using AWS S3, Athena SageMaker, devised 2 funded data pipelines by connecting location names in the articles' text to PRIO GRID geo cells, ran BERT model to visualize news articles' sentiments

Business Analyst **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
- Created 5+ automated analytical workflows by streamlined Data Engineering, Predictive Modeling for efficient ad spend
- Implemented customer clustering & classification, web analytics utilizing R, Python, SQL, Hive & Advanced Excel

Academic Projects

Zomato Hotels Complete Data Analytics | *Python, sklearn, Matplotlib, Seaborn, Folium* **Oct 2022 – Nov 2022**
- Analyzed all Zomato Indian restaurants: Data Cleaning for Restaurant overview analysis, top 5 budget restaurants analysis, Geographical analysis of restaurants using Folium
- Achieved accuracy of 79%-84% while predicting whether a restaurant is good or bad using Random Forest, Logistic Regression, K-Nearest Neighbors and xgboost

Fake News Predictor | *Python, sklearn, NLP, NLTK, Vectorization* **Oct 2022 – Nov 2022**
- Developed a Python script with Bag of Words to predict fake news using Multinomial Naive Bayes with 90% accuracy

Supply Chain Analysis and Late Delivery Predictor | *Python, Google Cloud Console* **Mar 2022 – May 2022**
- Performed Exploratory Data Visualization to visualize the Supply Chain Statistics of the company DataCo Global
- Achieved accuracy of 70% and 60% while predicting the late delivery risk of orders using Categorical and Gaussian Naive Bayes models. Then used Random Forest to predict the late delivery risk and boosted the accuracy to 85%

Masked Face Recognition | *Python (CNN, OpenCV, Keras)* **Oct 2021 – Dec 2021**
- Coded a Python program to detect people's faces & names with masks properly & improperly worn with 65% accuracy

Publications

- **"KCSNBShiny"** and **"KCSKNNShiny"** R Packages: Developed using RStudio (R Shiny) applicable for predictive analysis, employing Naive Bayes k-nearest neighbors algorithms to predict any variable in categorical and numeric data, published and deployed onto CRAN (Comprehensive R Archive Network) in July 2019)