Karneeshwar Sendilkumar Vijaya

+1 (469) 922-7009 | karneeshwarsv@gmail.com | linkedin.com/in/karneeshwar
github.com/karneeshwar | karneeshwar.github.io/portfolio/

Education

The University of Texas at Dallas — Dallas, TX
Masters of Science in Computer Science — *August 2021 - May 2023*

National Institute of Technology, Tiruchirappalli — Trichy, India
Bachelor of Technology in Engineering — *August 2013 - June 2017*

Relevant Experience

Copart, Inc — Dallas, TX
Software Engineer Intern (Data Platform) — *September 2022 - Present*

- Performed Descriptive/Predictive Time-Series Analysis on the data gathered through the different business processes using Python packages to draw inferences. Developed pipelines to perform Time-Series predictions for a variety of granularities and optimized them to handle large amounts of data
- Developed APIs to perform CRUD operations on a variety of new data based on user requests as an additional feature to be processed in an existing vehicle damage score prediction model
- Built a scheduling system using RabbitMQ, Celery, and Cron beats that can perform computation every night for the data gathered by the end of a business day eliminating manual processing of data. Facilitated other teams to use these results directly for their projects which helped in a 20% reduction in their lead time

Renault Nissan Technology Business Centre India (RNTBCI) — Chennai, India
Software Developer — *May 2019 - April 2021*

- Implemented API end-points using Java Spring framework to allow engineering team users to directly retrieve vehicle performance data (read-only) increasing the availability of essential data to required teams
- Created scripts in Databricks using Scala to completely automate ETL operations on semi-structured incoming vehicle data through Azure DataLake for Analyzing and Reporting tasks

Automation Engineer — *July 2017 - April 2019*

- Developed scripts in Python to automate the post-processing/reporting tasks in several engineering teams, reducing the lead time of projects by 40%

Technical Skills

Languages : Java, Python, PySpark, HTML, CSS, Javascript, C++, MATLAB, R
DB/Cloud : SQL, MongoDB, Cassandra, AWS (Lambda, EC2, S3), GCP Bigquery, Azure (Databricks, DataLake)
Frameworks : Spring/Spring boot, Flask, React, Bootstrap
Tools : Maven, Docker, Jenkins, Postman, Git, Linux

Projects

Twitter Streaming Sentiment Analysis, UT-Dallas | *PySpark, Kafka, Kibana, Docker* — *July 2022*

- Developed a structured streaming application to perform sentiment analysis on filtered tweets based on keywords
- Twitter API was used to retrieve data and streamed it via Kafka. The tweets were classified as positive or negative using pipeline text classification and data were filtered for the given keywords, for instance: covid19, coronavirus
- Logstash, Elasticsearch, and Kibana were used to store, visualize and analyze the polarity of filtered tweets

Plot Summary Based Search Engine for Movies, UT-Dallas | *Scala, SparkSQL, Databricks* — *June 2022*

- Built a search engine in Databricks to list top 10 movies which are closely related to the user's search input based on plot summaries using the Carnegie Mellon University's Movie Summary Corpus with over 42000 movie summaries
- The user search input could be either a single term or a query of multiple terms. MapReduce was used to compute TF-IDF and cosine similarity for single-term searches and multi-term queries respectively

DoorDash Database System Design, UT-Dallas | *SQL* — *March 2022 - May 2022*

- Developed an Entity-Relationship diagram to represent all the entities and their relation in the DoorDash system
- Mapped the ER diagram to a Relational Model by following the database normalization rules
- Created tables and implemented 3 triggers and stored procedures using PL/SQL like door-dasher age check, monthly pay stub computation for door-dasher, and overall bill calculation for customer