

Karneeshwar Sendilkumar Vijaya

+1 (469) 922-7009 | karneeshwarsv@gmail.com | [linkedin.com/in/karneeshwar](https://www.linkedin.com/in/karneeshwar) | github.com/karneeshwar | [Website/Portfolio](#)

EDUCATION

The University of Texas at Dallas

Dallas, TX

Masters of Science in Computer Science

August 2021 - May 2023

- Courses: Data Structures and Algorithms, Big-Data Analytics, Artificial Intelligence, Machine Learning

National Institute of Technology, Tiruchirappalli

Trichy, India

Bachelor of Technology

August 2013 - June 2017

- Courses: Basics of Programming, Probability & Statistics, Calculus, Operations Research, Rapid Prototyping

RELEVANT EXPERIENCE

Copart, Inc

Dallas, TX

Data Scientist Intern

September 2020 - Present

- Performed Descriptive/Predictive Time-Series Analysis on the data gathered through the different business processes, used Python Data Science toolkit to access database, develop pipelines for time-series prediction and draw inferences on the result
- Developed APIs using Flask to process a variety of new data based on user requests as an additional feature to an existing vehicle damage score prediction model

Renault Nissan Technology Business Centre India (RNTBCI)

Chennai, India

Engineer

July 2020 - April 2021

- Developed scripts in Python to automate post-processing of several simulated crash test results in Meta-post, reduced lead time of projects by 40%
- Saved 1.5M Euros in material and tool cost by developing a linear regression model based on the previous vehicle performance data to predict the design parameters for new vehicle projects in the same platform

TECHNICAL SKILLS

Languages: C, C++, Java, Python, R, PySpark, HTML, CSS, Javascript, MATLAB

SQL: MySQL

NoSQL: MongoDB, HBase, Cassandra

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 | *PySpark, 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

ACTIVITIES

- Head of Events at Prodigy'2017 (National symposium of Production Engineering Department, NIT-Trichy)
- Event manager at Pragyan'2016 (Annual international techno-management festival of NIT-Trichy))
- Manager at Festember Social Responsibility (Community service department of the inter-college cultural fest)