Naga Sanka

Data Scientist

Contact

- Troy, MI
- (248) 345 8742
- nsanka@umich.edu
- linkedin.com/in/nagasanka
- github.com/nsanka
- nsanka.medium.com

Skills

Data Mining	Git/GitHub
Data Analytics	Statistics
Data Visualization	Matlab
Machine Learning	Docker
Python	Spark
MySQL/Postgress	BigQuery
TensorFlow	NLP
Hadoop	DNN

Honors-Awards

- ★ Received BRAVO award in Stellantis for the work to recognize employees providing exceptional service.
- ★ Received star performer award for two consecutive years in Pacific Mindware for exceptional work.
- ★ Received DAAD Scholarship for IIT-Masters Sandwich Program for 2003-2004

Certifications

- ★ From Data to Insights with Google Cloud Specialization – Coursera
- ★ Web Application Technologies and Django – Coursera
- ★ Introduction to Git and GitHub – Coursera

Data Scientist familiar with gathering, cleaning and organizing data for use by technical and non-technical personnel. Advanced understanding of statistical, algebraic and other analytical techniques. Highly organized, motivated and diligent with significant background in data analysis.

Projects

Spotify Playlist Recommender | *Nov 2021 – Dec 2021* | WebApp

- Designed, developed and deployed an end-to-end Spotify playlist recommender system using unsupervised learning clustering approach to discover music based on the user's listening preferences.

Predict Analyst Recommendations on Publicly Listed Companies | *Apr 2021 – May 2021* | Report

- Extracted S&P 500 information from IBES and CRSP Databases on WRDS using SQL queries. Performed feature engineering to understand the importance of various features and used PCA to reduce the dimensionality of feature set.
- Built classification models using Gaussian Naïve Bayes, Support Vector Machines and Radom Forest to classify the changes in analysts' recommendations on companies they follow and described the underlying structure of those recommendations.

Explore Various Parameters Which Affect Text Classifier Performance | Dec 2020

- Built Logistic Regression model using biography articles from Wikipedia to predict the nationality. Various options used to test the performance are dataset size, count vectorizer, TF-IDF vectorizer, stop word removal, hyper parameter tuning, number of features, lower case text.

Analyze Autonomous Vehicle Driving Data | *Apr 2020 – May 2020* | Report

- Performed exploration, preparation and visualization of autonomous driving datasets using Spark RDD and DataFrame to answer key questions related to vehicle pose and driving patterns. Used Folium Maps and Plotly for visualizations.

Exploratory Data Analysis of Covid-19 Dataset | Mar 2020 | Notebook

- Performed Exploratory Data Analysis (EDA) on the South Korean COVID-19 dataset to demonstrate the expertise in data manipulation, cleaning and visualization. Used various Python modules such as Pandas, PySpark, Matplotlib, Altair, Plotly and Folium.

Work Experience

Stellantis, Auburn Hills, MI   **Mar, 2014 – Present**
Advanced Senior Engineer

- Developed offline and online data augmentation methods to improve absolute vehicle pose estimation using deep neural networks (DNN).
- Hands on experience in installing, configuring and using Hadoop ecosystem components like Hadoop, MapReduce, HDFS, Hive and Spark.
- Implemented multi-objective optimization based on approximation using classification/clustering methods with Design of Experimental (DOE) data.

- Researched, designed and implemented scalable applications to reduce time, effort and cost for a first time right design.
- Developed pipelines for efficient modeling and custom Python modules for process automation which reduced analysis time from 3 weeks to 2 days.
- Developed databases, programs and processes for integration and implementation across powertrain virtual engineering department.

ESI NA Inc, Farmington, MI   **May, 2009 – Mar, 2014**
Sr. Software/Application Engineer

- Worked closely with other business analysts, development teams and infrastructure specialists to deliver high availability solutions for mission-critical applications.
- Reviewed C++/Python/Java code and worked with development and texting teams to fix problems, run test scenarios and prepare for deployment.
- Introduced methodologies and best practices to software development division to enhance product definition and application customization.

Pacific Mindware Engineering Pvt Ltd, India   **Sep, 2004 – May, 2009**
Applications Manager

- Led 4-member development team in delivery of high-end customized engineering applications to streamline process of Computation Fluid Dynamics based simulation technology.
- Designed, created, tuned physical database objects to manage automotive vehicle components.

Education

- **University of Michigan – School of Information**   **2019 – 2021**
  Masters, Applied Data Science
- **University of Stuttgart**   **2003 – 2004**
  MTech Thesis
- **Indian Institute of Technology Guwahati**   **2002 – 2004**
  MTech, Fluid and Thermal Engineering
- **Jawaharlal Nehru Technological University, Hyderabad**   **1998 – 2002**
  BTech, Mechanical Engineering