

Sahith Kumar Singari

Data Scientist (m/f/d) Machine Learning

Personal Info

Phone

+49 15215715704

E-mail

sahithkumar1999@gmail.com

Skills

Team leadership

Teamwork: Proven ability to collaborate effectively in a team

Programming Languages: Python (Proficient)

Machine Learning: Familiarity with ML concepts and algorithms

Data Manipulation: Experience in collecting, cleaning, and preparing data

Analytical Skills: Strong analytical and problem-solving abilities

Fast Learner: Quick adaptation to fast-paced AI research and development

Version Control: Git

A/B Testing: Familiarity

SQL

Links

LinkedIn

www.linkedin.com/in/sahithkumar

As a motivated and aspiring Machine Learning enthusiast currently pursuing a Master's degree in Information Technology at Frankfurt University of Applied Science, I bring a strong foundation in machine learning, proficiency in Python, and a passion for leveraging data-driven insights to solve complex problems. With a solid educational background complemented by certifications in natural language processing and artificial intelligence, I am eager to embark on a career in data science. My dedication to continuous learning, coupled with hands-on experience in developing machine learning models, positions me to contribute effectively as an Entry-Level Data Scientist in the InsurTech sector. I am excited to apply my skills, collaborate with cross-functional teams, and drive impactful solutions that align with the goals of the organization while staying at the forefront of advancements in the field of Data Science

Education

Master's

Frankfurt University of Applied Science / 09/2022
Frankfurt am Main

Master of Science in Information Technology

Courses

Build Chatbots with Python

Code academy / 08/2023

Getting Started with Natural Language Processing

Codeacadamy / 08/2023

Artificial Intelligence Foundations: Machine Learning

LinkedIn Learning / 05/2023 - 07/2023

Introduction to Conversational AI

LinkedIn Learning / 08/2023

Machine Learning with Python: Foundations

LinkedIn Learning / 08/2023

Projects

Handwriting Recognition using K-Means

08/2023

Codeacadamy

- Description: This project focuses on Handwriting Recognition using K-Means clustering, inspired by real-world applications like the U.S. Postal Service, ATMs, Evernote, and Expensify. It demonstrates the power of machine learning in understanding handwritten content, such as addresses, bank checks, task lists, and receipts.

- The project leverages scikit-learn, an open-source machine learning library in Python, and K-Means clustering to analyze and cluster images of handwritten digits. Python's data science ecosystem, including NumPy and Matplotlib, is employed for data preprocessing and visualization.

- Results: The project showcases AI/NLP skills by successfully applying K-Means clustering to group handwritten digits. It demonstrates the potential for automating handwritten content recognition, opening doors to applications in various domains like postal services, finance, note-taking, and expense management.

Languages

English

German

Scalar Encoder with Buckets

11/2022 - 03/2023

Frankfurt University of Applied Science

- Description: This project involves the implementation of a Scalar Encoder with Buckets in C#. It showcases AI/NLP skills by creating a specialized encoder that can discretize continuous numerical data into categorical "buckets." Such encoders are valuable in various applications, including natural language processing and machine learning, for feature engineering and data preprocessing tasks.

- Technologies Used:

C#: The project is developed using the C# programming language.

.NET Core: Utilizing .NET Core for cross-platform compatibility.

Visual Studio: An integrated development environment (IDE) for C# development.

Math Libraries: Leveraging mathematical libraries for data manipulation and encoding algorithms.

- Results: The project successfully implements a Scalar Encoder with Buckets in C#. It provides a valuable tool for data scientists and machine learning practitioners to preprocess continuous data efficiently. By discretizing numerical features into buckets, it simplifies data analysis and enhances the performance of AI/NLP models, ultimately contributing to more accurate and efficient natural language processing and machine learning applications.

Conversational AI Bot with Retrieval-Based Learning

Skills Demonstrated:

- Proficiency in Python for data processing and natural language understanding.
- Application of NLP (Natural Language Processing) techniques, including tokenization and lemmatization.
- Implementation of TF-IDF vectorization and cosine similarity for context-based response generation.
- Chatbot development and user interaction design.

Key Achievements:

- Successfully developed a functional Conversational AI Bot capable of understanding and generating responses based on user input.
- Implemented text preprocessing techniques to enhance the quality of input data and improve the bot's comprehension.
- Utilized TF-IDF vectorization and cosine similarity to identify relevant responses from a corpus of text.
- Ensured user engagement with greetings and appropriate responses.
- Created an interactive and user-friendly chat flow for seamless communication.

Shape Matching using Dynamic Time Warping (DTW) Algorithm

Skills Demonstrated:

- Proficiency in Python and relevant libraries for image processing, including OpenCV and numpy.
- Implementation of the DTW algorithm for shape matching.
- Graphical User Interface (GUI) development using tkinter.
- Data visualization and results storage in Excel format.

Key Achievements:

- Developed a user-friendly GUI to facilitate easy interaction with the shape matching application.
- Successfully implemented the DTW algorithm for shape matching, allowing comparisons between a given image and multiple template sequences.
- Visualized matching results to enhance understanding and analysis.
- Efficiently saved matched images and matching data for future reference.

