KARNATI SAI PRASHANTH

@ karnatisaiprashanth.com ☎ +1 518-514-8868 in karnati-sai-prashanth-ab902b205 ⌂ /KS-prashanth
🌐 ks-prashanth.github.io/KSPrashanth.github.io/

ABOUT ME

I am a passionate tech enthusiast increasingly drawn to data science and artificial intelligence. My journey commenced out of fascination with AI's potential to revolutionize our world for the better. As time progressed, my focus pivoted towards utilizing data and AI methodologies to tackle intricate challenges. Equipped with a robust grasp of machine learning, statistics, and Natural Language Processing, I am dedicated to delving into the expansive horizons of AI and data science. My goal is to spearhead innovation and create meaningful solutions that make a substantial impact.

EDUCATION

Master's - Data Science
State University of New York at Albany

📅 2024 – Ongoing   📍 Albany, USA

B.Tech - Computer Science And Engineering (Artificial Intelligence)
Amrita Vishwa Vidyapeetham

📅 2020 – 2024   📍 Coimbatore, India

CGPA : 8.12

Higher Secondary Education
FIITJEE Junior College

📅 2018 – 2020   📍 Visakhapatnam, India

CGPA : 9.36

Secondary School
The Presidential School

📅 2018   📍 Visakhapatnam, India

Percentage : 86

TECHNICAL SKILLS

- **Programming:** Python, R, Matlab, c.
- **Packages:** NumPy, Pandas, TensorFlow, SK-learn.
- **Data Visualization:** Matplotlib, Seaborn, PowerBi, Tableau, Excel.
- **Mathematics:** Linear Algebra, Statistics, Calculus
- **Database:** MySQL, MongoDB
- **Cloud Platforms:** Azure, AWS

PROJECTS

Voice and Message Based Home Automation

📅 Jul 2022

I developed an IoT-based home automation system using a Raspberry Pi. This system includes voice recognition, a Telegram bot, and real-time data integration, all designed to provide a user-friendly experience.

Bridging the Language Gap: Empowering Hate Speech Detection with Pervasive Attention on Dravidian Languages

📅 Jul 2023

This project addresses hate speech detection in code-mixed Dravidian languages, specifically Tamil, using pervasive attention and the random kitchen sink (RKS) method, achieving an accuracy of 77.4

Parts of Speech Tagging in Telugu Using Generative Pre-Trained Models

📅 May 2024

Parts of Speech (POS) tagging is crucial in natural language processing. Major languages have advanced, but Telugu struggles due to limited annotated data. Researchers are using large language models (LLMs) for Telugu-specific POS tagging, achieving 98.01% accuracy with methods like Mixtral.

Matrix Factorization for Predicting the Movie Rating

📅 Aug 2022

Implemented Matrix Factorization and Collaborative Filtering for movie rating prediction, considering genre and favorite actors/actresses to enhance accuracy while handling missing ratings.

Data-Driven Student Achievement Analysis and Machine Learning Deployment in AWS with CI/CD Integration

📅 Jul 2023

Utilized Machine Learning on Portuguese school datasets encompassing student grades, demographics, social factors, and school-related features, then deployed the model on AWS with CI/CD Integration.

TECHNICAL INTERESTS

Machine Learning   Data Analysis

Natural Language Processing

Artificial Intelligence

PUBLICATIONS

Bridging Language Barriers: Exploring Hindi-to-English Speech-to-Speech Translation for Multilingual Communication

📅 01-08-2023

Accepted at 4th Congress of Intelligence Systems 2023
Developed a structured Hindi-to-English S2S translation model with STT, text translation, and TTS modules, achieving an average MOS of 3.8, addressing language barriers for effective multilingual communication in India.

Enhancing Telugu Part-of-Speech Tagging with Deep Sequential Models and Multilingual Embeddings

📅 03-12-2023

Accepted at ICON-23
The paper investigates deep sequential models for POS tagging in Telugu, a low-resource language. Using Universal dependencies dataset, it explores RNNs, LSTMs, GRUs, and stacked variants. Employing multilingual BERT embeddings, stacked LSTM achieves the highest F1 score of 0.8812, proving their effectiveness for Telugu POS tagging.

Findings of the shared task on hate and offensive language detection in telugu codemixed text (hold-telugu).

📅 14-03-2024

Accepted at Dravidian LangTech2024
This study analyzes submissions from multiple teams for the Hate and Offensive Language Detection in Telugu Codemixed Text (HOLD-Telugu) task, part of DravidianLangTech 2024. Teams were challenged to develop models to identify harmful content in Telugu codemixed social media text. The dataset, comprising annotated YouTube comments, was manually compiled. Out of 23 participating teams, results were evaluated based on macro F1-score to create the rank list.

Towards a More Inclusive Telugu Internet: Creating Corpus and Developing Models to Detect Hate Speech in Telugu Code-Mixed Social Media Text

📅 Under Review

This paper presents machine learning and deep learning models for detecting hate speech in Telugu-English code-mixed social media text data, addressing challenges in sentence complexity and low-resource language settings, achieving optimal performance with distilbert-based features.

Exploring Current Transformer-Based Models in Speech Processing Tasks: A Concise Review

📅 Under Review

This paper explores advancements in speech processing, focusing on transformer-based techniques like Conformer models, and provides a comprehensive comparison of AI-based methodologies to guide further research in the field.

CERTIFICATIONS

📅 Issued 2020 - 2023

- Data Science Virtual Experience Program - British Airways
- Microsoft Certified: Azure Fundamentals
- Microsoft Certified: Azure Data Fundamentals
- Microsoft Certified: Azure AI Fundamentals

VOLUNTEER CONTRIBUTIONS

LIVE-IN-LABS®

📅 Feb 2023 - Dec 2023   📍 Rampura, Haryana

Participated in the Live-in-Labs® Program in rural Mahendragarh district, Haryana, as part of a sustainability development initiative by Mata AmritAnandamayi Math and The United Nations. In a team of six, we underwent training in tools like Participatory Rural Appraisal and Human Centered Design to identify and address challenges faced by the community of Rampura. The collected data is being analyzed to construct a proposed developmental project.

LANGUAGES

- Telugu
- English
- Hindi
- Tamil

PERSONAL DETAILS

- **Date Of Birth:**
  16th March 2002
- **Hobbies:**
  Photography, Swimming, Playing Video Games

EXTRA - CURRICULAR ACTIVITIES

- Central Co-Ordinator for the events GOKULASTAMI and AMRITAOSTAVAM.
- Head of the Drones Team for the ANOKHA Technical Event.
- Elected as the Class and Cultural Representative for my branch.