

Sukanya M Mathapati

Bengalore, Karnataka

+91-7259084297

sukanyamathapati2002@gmail.com

LinkedIn

GitHub

PROFILE SUMMARY

Dynamic and Results-Driven Data Scientist with a strong foundation in AI/ML, specializing in sentiment analysis, neural networks, and Large Language Models (LLMs). Proficient in Python, C, R, and full-stack web development, with expertise in data analysis, visualization, and cloud computing (AWS, Docker). Demonstrated leadership in spearheading innovative AI projects, driving significant improvements in client acquisition, product sales, and agricultural outcomes. Recognized for excellence in project execution and Agile management, with a proven track record of delivering impactful solutions in high-stakes environments. Strong communicator and effective leader with a commitment to advancing the field of AI through continuous learning and practical application.

EDUCATION

Basaveshwar Engineering College, Bagalkot

Bachelor of Engineering - Computer science - CGPA- 8.4

2020 – 2024

Bagalkot, India

Alvas PU College, Moodbidri,

PU - Science - Percentage - 92.17%

2018 – 2020

Moodbidri, India

COURSEWORK

- Data Structures & Algorithms
- Network Security
- Artificial Intelligence and Machine learning
- Web Development
- Operating Systems
- Database Management System (DBMS)
- OOPS Concept
- Data Science

PROJECTS

F.I.O.N.A | LLM's, GenAI, Transformers, AWS, OpenCV, NLP

June 2024

- Friendly Intelligently Organized Neural Assistant, is An advanced AI Language model developed to understand human emotions and provide appropriate response to user same as Human brain.
- Implemented Advanced Reinforcement learning, BERT & GPT-3 architecture to generate Human like language with remarkable accuracy of 94.00% .
- Utilized more then 10M dataset from custom data of openAI and Hugging face, more then 8M online websites for real time data through web scraping. with 15M training parameters.(Storage in AWS buckets)
- Achieved National Level first prize, State Level first prize and special recognition from ISRO, AICTE.

GenAI for Automated Software Development and Debugging | GPT-2, Docker, EC2

2023

- Developed a system that can assist developers in generating, debugging, and refactoring code, potentially reducing development time and minimizing human errors in software development processes.
- Fine-tuned GPT-2 model with a training accuracy of 92% on a dataset of 50,000 code snippets, debugging logs, and refactored code samples.
- Utilized a custom dataset created from public repositories and CodeSearchNet dataset, consisting of 50,000+ labeled examples.
- Employed transformer-based models (GPT-2) for code generation and debugging tasks, with additional logic for code refactoring. Enhanced with custom algorithms for code simplification, dead code removal, and function extraction, optimizing software quality and maintainability beyond traditional automation methods.

Object Detection for Blind People | YOLO-V3, OpenCV, AWS S3, Tacotron 2, GAN

2022-23

- Designed and implemented a real-time object detection system using YOLOv5, integrated with edge devices for on-the-go usage by blind individuals.
- Utilized Tacotron 2 with HiFi-GAN for high-quality, natural-sounding Text-to-Speech, enabling auditory feedback that accurately describes the surrounding environment to users.
- Fine-tuned YOLOv5 and Tacotron 2 models for efficiency, ensuring that both object detection and TTS components operate seamlessly and provide timely feedback.
- Designed a robust feedback loop by integrating auditory descriptions with contextual information, significantly improving the navigation and interaction experience for users.

EXPERIENCE

GeeksforGeeks

Jan 2023 – June 2023

Data Science and Machine Learning Intern

Remote, Bangalore

- Assisted in building real-time machine learning models using TensorFlow and scikit-learn, enhancing live data stream analysis with a 25% increase in accuracy through the application of advanced algorithms such as XGBoost and Random Forest.
- Created and optimized data pipelines capable of processing real-time data streams with a throughput increase of 40%, significantly reducing data processing time and increasing the scalability of data workflows.
- Engaged in live projects focusing on real-time data analytics and predictive modeling, including tasks such as feature engineering and hyperparameter tuning, which led to a 15% improvement in the performance of recommendation systems
- Improved advancements in real-time data processing and model accuracy contributed to a 10% increase in platform revenue by enhancing user satisfaction and engagement. The work also supported GeeksforGeeks' goal of providing cutting-edge educational resources.

Brainovision Solutions India Pvt.Ltd

June 2023 – Sept 2023

Data Science and AI Intern

Hyderabad, India

- Performed comprehensive exploratory data analysis (EDA) using Jupyter Notebooks and Pandas, uncovering critical insights and trends. Utilized statistical methods and machine learning techniques to improve business forecasting accuracy by 25%
- Supported research initiatives by implementing cutting-edge algorithms such as deep learning networks and natural language processing models. Used PyTorch and Keras to advance AI solutions, contributing to the company's innovation goals
- The contributions led to a notable enhancement in the company's AI capabilities, resulting in a 10% increase in client satisfaction and a 12% growth in project delivery speed. The work also supported the company's goal of staying at the forefront of technology and innovation in the industry.

TECHNICAL SKILLS

Languages: Python, Java, C, JavaScript, R, HTML5, CSS3

Tools: Power BI, Tableau, Docker container, Github

Frameworks: Flask, Django

Database: SQL, NoSQL, MongoDB, Cassandra

Artificial Intelligence & Machine Learning: LLM's, TensorFlow, scikit-learn, LightGBM, NLTK, spaCy, Transformers, GAN's, Langchain, Edge AI, Seaborn, Pytorch.

Cloud: AWS S3, AWS EC2, Google GCE

Data Science: Data Visualisation, Data Analytics, statistical Analysis, ROI analysis, Feature Engineering, Predictive Modeling

Bigdata: Apache Spark, Hadoop

Soft skills: Strong interpersonal skills, Scrum Master, Agile Project Management, Effective Leader, Continuous learning, Adaptable, Team player

CERTIFICATIONS

- Certified Analytics Professional (CAP)
- Cloudera Certified Professional Data Engineer (CCP Data Engineer)
- TensorFlow Developer Associate Certificate
- AWS Certified Data Analytics – Specialty
- Google Data Analytics professional
- Google Professional Data Engineer

ACHIEVEMENTS

- National First prize in Project Competition
- Special Recognition from ISRO, AICTE at State Level
- Approved Patent on AI Brain Technology(Language model)