SANCHARIKA DEBNATH

Data Scientist

Passionate Data Scientist adept at emerging tech for innovative solutions, with data science and back-end skills.

+91 (947) 723 0267 - sancharikadebnath@gmail.com - linkedin.com/in/sancharika-debnath - github.com/sancharika

WORK EXPERIENCE

AI ML Engineer

Giggr Technology, Karnataka, India *JUNE 2023 - FEBRUARY 2024*

- Developed an agile event system integrating facial recognition for attendance, object identification via camera data, and GPS. Achieved **87.4%** accuracy using **OpenCV, YOLO, Deep-Face, Geo location API, and JavaScript.**
- Collaborated with team to develop Neo4j architecture, integrating it collaboratively for data mining concepts, Data and Analytics using technologies such as **Neo4j, Graphical Database, REST API, and GCP, GitHub Actions.**
- Delivered a custom public API via Elasticsearch for 1.3M+ Indian education institutes, using web scraping. Technology Used: **Amazon Web Services, Elastic-Search, Shell Scripting**
- Designed a Generative AI stack Bot based on business requirements, advanced scoring, Firebase integration, boosting interaction by **90%**. Technologies Used: **Dialogflow, Firebase, Open AI, NLP, GPT, Text-to-Speech**

Back-end Developer

Leapon, Remote *FEBRUARY 2023 - JUNE 2023*

- Developed a customized back-end and Django REST API for a website with 50+ advisors, integrating Unit Test and Bug fixing for feature optimization via CI/CD tools. Used **Python - Django, SQL, AWS, Agile Scrum, GIT, JIRA**.

Data Science Intern

Maersk Global Service Centres, Karnataka, India *JUNE 2022 - FEBRUARY 2023*

- Developed office seat pre-booking web app, integrating Spring Boot & Next.js, boosting user booking efficiency by **80%** in production. Technologies Used: **Spring-Boot, Next.js, Anchor UI, PostgreSQL**
- Implementation of models to cluster ports using clustering & estimation methods, enhancing analytical capabilities, with **Gaussian Mixture** for Cost estimation. Utilized **Pyspark, Databricks, Microsoft Azure.**
- Approach to predict container turn time and Forecasting of Attachment-ratio in distributions logistics, considering quantitative metrics with **FLAML** as the best AutoML Utilized **ETL, Databricks and Azure ML.**

Data Science Winter Intern

HighRadius Corporation, Orissa, India *JANUARY 2022 - APRIL 2022*

- Created a regression invoice prediction system using machine learning models, and Statistical concepts to optimize financial operations. Utilized **XGBoost** with **76.15%** accuracy and frameworks including Keras.

TECHNICAL SKILLS

Programming Languages: Python, C++, R, Java, JavaScript
Technical Skills: NLP, Machine Learning and Data Mining, Deep Learning, Computer Vision, Data Analysis, GenAI
Libraries: Pandas, NumPy, SkImage, TensorFlow, PyTorch, scikit-learn, Keras, OpenCV, NLTK, spaCy, Transformers
Tools: AWS / Azure / GCP, Neo4j Graph Database, Cassandra, Pinecone, SQL / query languages, Tableau
Frameworks: Keras, TensorFlow, PyTorch, LangChain, Llama, Django, PySpark, GitHub, Databrick, Azure ML

PROJECTS

- **Career Enchanter**: Developed Generative AI pipeline for job hunting with resume reviews, personalized recommendations, cover letter generation, job suggestions, interview preparation and ATS score calculation, using Natural Language Processing (NLP): **BERT Transformer, MLops, Gemini-Pro, Large Language Models**.
- **LLM IPO Analyzer**:Build Generative AI/ML model for detailed investment statistical analysis in startup IPO. Used prompt engineering, vector databases, and deployment. **Llama, HuggingFace, Langchain, Pinecone**
- **Furniture Classification**: Developed a furniture classifier using **deep learning models and Transfer Learning models**, employing neural networks like **ResNet-50** on cloud platform, **AWS SageMaker** for improved accuracy and efficiency in image classification. Achieved **83.16%** accuracy. Utilized **Transfer Learning, ResNet-50**

PUBLICATIONS

- **Hyperspectral Image (HSI) Compression and Classification**: experimentation on a research framework for HSI using CNN bottleneck AutoEncoder to exploit **Hyper Spectral Net** for classification. Achieved **0.998** accuracy. Used **TensorFlow, Autoencoder, Hybrid SN**. Conducted **Jan'22 to Mar'22**, published in **IJCISIM**.

EDUCATION

Kalinga Institute Of Industrial Technology **Bhubaneswar, Orissa, India**
Bachelor of Technology, Information Technology, *CGPA:* 8.83 *JUNE 2019 - JULY 2023*