SANJIB KHETAN

H: 8598850497 | khetan.sanjib@gmail.com | GitHub: https://github.com/sanjib-khetan

PROFESSIONAL SUMMARY

An Enthusiastic professional, Developer, Researcher preferably as AI Engineer having experience and knowledge in Deep Learning, Machine Learning algorithms and translating those algorithms and the research work into POCs, commercially viable products and services.

SKILLS:

Programming Language:

Python, Data Structure and Algorithm, NoSQL

Techniques:

Machine Learning, Deep Learning, Computer Vision, NLP, Transformer, Feature Engineering, Random Forest, LSTM, RNN, CNN, Boosting, Bagging, Ensemble Learning, Transfer Learning, Object Classification, Object Detection, GAN, SVM, Text Summarisation, Auto-encoder, Classification, Data Cleaning, Regression, OCR, Clustering, Anomaly Detection, Dimensionality Reduction, Sentiment Analysis, Framework Designing and Development, Product Pipeline design

Libraries:

TensorFlow, Pytorch, Scikit-Learn, NLTK, Spacy, Flask, PIL, MongoDB, Numpy, Pandas

WORK HISTORY:

PREZENT.AI | *Senior Data Scientist | May 2023 - Present*

Search and Recommendation Engine for the entire Product(Platform)

Solr, Hybrid Search, BLISS++, LLM, Semantic Search, LLM Finetuning.

Built the Entire Hybrid Search Engine Module for the Product having both Keyword Search and Embedding Search using LLM Embedding.
Trained the Sentence-Transformer Models on our own data which gave better and Expected results.
Implemented BLISS++ approach for the Search which reduces the Indexing efforts and heavy cost of Semantic Search in real time.

Entity Recognition System For Sanitizer using LLM and Retraining a Transformer Model

LLM, Transformer Model, Amazon Claude

Developed and Train a Custom NER model leveraging Transformer Architectures on our own data to detect various Entities Relevant to the Project.
Additionally, We have also used the Amazon Claude Model for Entity Recognition tasks, using prompt engineering techniques yielding favourable outcomes, which provides lesser Hallucinations.

TREDENCE ANALYTICS | *Senior Data Scientist | Sep 2021 - May 2023*

Business Insights Generation Tool For E-Commerce Data.

Time-Series Analysis, Anomaly Detection, Clustering, K-Means, Z-Score, DBSCAN, NLG.

Built and Deploy a Tool which Generates all the Business Insights combining the Sales, Traffic and Customers Data. Built different set of Algorithmic Trees which uses different ML methods and Business Rules to generate all the Insights which replaces an Analyst's work. Also used DBSCAN, I-FOREST, Z-SCORE, K-Means Clustering and other methods for Insight Generation.

INNOMINDS SOFTWARE | *Data Scientist | Apr 2021 - Aug 2021*

Recommendation System to find Similar Question-Answer Pair. (Semantic Search and Image Search)
NLP, Flask, Sentence Embedding, FAISS, MongoDB, ResNet50, Clustering, BERT

Built the Entire Search API for Question to Question search, which takes the question and finds the similar question from the database to map it with the corresponding Learning Objective.
Designed entire structure and implemented the pipeline for Search Functionalities by taking the entire mapped data and build FAISS model on top the embedding vector to search. Created the API for this project and deployed it.

SONY INDIA SOFTWARE CENTRE | *Data Scientist | (Payroll of JGD Tech) | Aug* 2019 - Jan 2021

Correcting AI (Improving detection accuracy using Representer Function)
Explainable AI (XAI), PyTorch, YOLOV3, Representer Points

Improved the accuracy of an **Object Detection model** with the help of **Representer Function Algorithm**. Built an algorithm on top of **YOLO algorithm and Representer function Algorithm** which can improve the correctness of the detection model, which includes modifying the YOLO V3 loss function and YOLO V3 optimisation function with its implementation.
Train and Built different object detection model for different dataset with implementation of complete YOLOV3 architecture, training, optimisation process with Representer function algorithm.

WIPRO RESEARCH | *Data Scientist* | Wipro AI Research Team | Dec 2016 - Aug 2019

Explainable AI - Build Tool to Explain AI decision
TensorFlow, Torch, Keras, LRP, Python

Built a **Deep Learning Library** to generate heatmap (from **LRP**) independent of any architecture, any framework supports **TensorFlow**, **PyTorch**, **Keras** which includes implementation of all the deep learning layer from scratch for ex: **CNN**, **FC**, **ReLU** or **Pooling** layers.
Train and Built different Image and Text Classification models for the Tool.
Complete Designing and Development of LRP algorithms for different deep learning network.

PAPERS AND PATENTS:

Published "**Single window platform for Explainable AI**" paper with my team member in ADCOM 2018 conference

Filed Patent in WIPRO for 1."correction of OCR".
2." Detection and correction of misclassification with Explainable AI" in Wipro PATENT CARNIVAL with my team member.

ACHIEVEMENTS:

Secured:

Wipro **PRODIGY AWARD** for **best debut performer** in Machine Learning for the year 2017-2018.
All India 1st rank in Coding Assessment (**DIGITHON**), conducted by Wipro Digital among 3000 employees through Hackerearth.

EDUCATION:

B. TECH | *Computer Science and Engineering* | *V S S U T BURLA | ODISHA | 2016*
12th Science *| JAWAHAR NAVODAYA VIDYALAYA NARLA | ODISHA | 2011*