Sounak Mukherjee

✉ m.sounak2000@gmail.com | 🅻 mSounak | ☎ +91-6290480560 | ⌨ github.com/mSounak | k kaggle/redhunter20

EXPERIENCE

Pratilipi ⬈
May 2022 - present

Data Scientist   Bangalore

- Performed multiple experiment on *Collaborative Filtering* model and *Content Based* models to improve total reads by **50%** and also **2x** the subscription purchases.
- Written data ingestion scripts for AWS Personalise. And deployed new campaigns, which helped to achieve daily reads of **150k**.
- Modified two algorithms for recommender systems and improved total reads by **20%**.
- Built a Bi-LSTM model from scratch to predict the gender of a person from their name and achieve a F1-score of **93%**.
- Wrote ETL pipeline to get user meta data from 3rd party API.
- Built and maintain the Machine learning infrastructure using **MetaFlow** and **AWS Batch**.
- Wrote **terraform** module to automate **AWS ECS** service deployment.

Bobble.ai ⬈
February 2022 - May 2022

AI Intern   Remote

- Wrote pipelines to process terabytes of conversational data.
- Performed *intent recognition* on *hindi macronics* corpus.
- Developed a model to find the sentiment of brands in the conversational texts.
- Trained *FastText* on large datasets to create sentence embedding and word embedding.
- Performed clustering on large conversational data.

PROJECTS

Targeted-sentiment-analysis ⬈
December 2021

Implemented Transformation Net for targeted sentiment analysis

- Built a custom model to get the polarity of the target in the sentence.
- Then created an API for the backend using FastAPI, containerized using docker and deployed on a kubernetes cluster.

Machine-Translation ⬈
November 2021

Trained a Transformer from scratch and used for machine translation

- Encoder-Decoder architecture to produce an input sequence representation with applied positional embedding to increase token position awareness in sentence. Decoder connects it all together and generates required translation.
- It translates English sentences to German have plans to expand its support for Indian language. Created the front-end in streamlit and deployed the application in EC2 instance.

TECHNICAL SKILLS

- **Languages:** Python, C/C++, Dart, HTML/CSS
- **Libraries:** Scikit-Learn, Matplotlib, Pandas, Numpy, Seaborn, Spacy
- **Framework:** TensorFlow, Keras, PyTorch, Flutter
- **Others:** Linux, Git, AWS, Streamlit, Docker, Kubernetes, FastAPI, Terraform, MetaFlow, MLFlow, Weight Biases

COURSEWORK

- **Academics:** NLP, Machine Learning, DMDW, Data Structures, Operating Systems, Database Management Systems, Web Technology
- **Others:** Coursera Deep Learning Specialization by Andrew Ng., GANs Specialization by DeepLearning.ai

EDUCATION

Kalinga Institute of Industrial Technology
Bhubaneswar

B.Tech Computer Science and Engineering, **CGPA: 9.15**   2019–2023