

HIMALAYA MANDAL

Vill.: Pathardihi P.O.: Hatirampur Dist.: Bankura Pin: 722121

Phone No.: +918436776454

Email: himalayamandalcse@gmail.com

LinkedIn: www.linkedin.com/in/himalaya-mandal

Xing: www.xing.com/profile/Himalaya_Mandal/cv



PYTHON | ML | DJANGO | REST-API | PYSPARK | KAFKA | AIRFLOW | WEB SRAPING | NLP | AWS | SQL | DOCKER | POSTGRESQL | GIT

EXPERIENCE

August, 2022 TO Present

SENIOR SOFTWARE ENGINEER, CALSOFT INC.

JOB ROLE: Understanding the requirements for new features and deliver within the time committed. Also, leading a data engineering team which is part of CALSOFT R&D team.

October, 2020 TO June 2021

SENIOR SOFTWARE ENGINEER, HOMZHUB ADVISOR

JOB ROLE: Leading a backend development team for developing mobile and web application. Design a scraper which scrapes data from real estate website, build model and recommend user with real-estate property, news etc. etc.

February, 2020 TO August 2020

NLP Developer, WSE Games

JOB ROLE: Building a recommendation-based system which will recommend other post, blogs, forum jobs etc. As per the behavior of the user on the platform.

June, 2019 TO January, 2020

PYTHON DEVELOPER, ECLIPSE TECHOCONSULTING PVT.LTD

JOB ROLE: Responsible to work with client to understand the requirement and implement it within deadline.

September, 2018 TO April, 2019

PYTHON ANALYST, INDIAN CYBER SECURITY SOLUTION

JOB ROLE: Placed in training and development team, I was responsible to provide training on python and Machine Learning and also develop VAPT tools for internal VAPT team.

SKILLS:

Programing Language:

Python, SQL

Databases:

RDBMS: MySQL, PostgreSQL. NOSQL: MongoDB, Redis

Cloud:

AWS, Azure

Web Framework:

Django, Rest-API

Data Science:

Machine Learning, Data Engineering, NLP, Recommendation System, model building, Collaborative Filtering, BERT, TF-IDF, word2vec, Classification, clustering, svm, PySpark, spacy, Scikit Learn, Pandas, Matplotlib, Seaborn.

Data Pipeline & Scheduler:

Kafka, Airflow, Celery

Containerization:

Docker

Scraping & Automation:

Requests, BeautifulSoup4, Selenium,

Version Control & Other Tools:

GIT, Bash language, Jupyter Notebook, Debian

EDUCATION

2005-2011

SECONDARY, MASANJHAR ADIBASHI HIGH SCHOOL

PERCENTAGE: 71%

Board: WBSE

2011-2014

HIGER SECONDARY, INDPUR GOENKA HIGH SCHOOL

SPECALIZATION: SCIENCE

Percentage: 63%

Board: WBCHSE

2014-2018

B. TECH, GURU NANAK INSTITUTE OF TECHNOLOGY

SPECALIZATION: COMPUTER SCIENCE AND ENGINEERING

CGPA: 7.10

Board: WBUT

PROJECT

1. DATA ENGINEERING PIPELINE TO HANDLE LARGE AMMOUT DATA

Client: Calsoft Inc

Description: Our project focuses on building a highly efficient data engineering pipeline capable of processing vast amounts of data from various data sources in minimal time. To achieve this, we leverage Kafka for data streaming, distributing the workload across multiple nodes. Airflow is employed to monitor the data in Kafka nodes and initiate data engineering workflows designed using PySpark.

Our primary goal is to achieve remarkable data processing speed, aiming to handle 80,000 rows of data, with each row containing an extensive 10,000 columns. By implementing this robust pipeline, we can process large-scale datasets rapidly and seamlessly, enabling timely data insights and analysis.

Technology Used: PySpark, Apache Airflow, Apache Kafka, aws, docker, PostgreSQL

2. APSTRA OPERATING SYSTEM

Client: Calsoft Inc

Description: This project involves developing a comprehensive data center network management operating system. The system aims to efficiently connect and oversee numerous network switches and routers.

Users can access the user-friendly Apstra dashboard, which serves as a central control hub for all network configurations. Additionally, the platform enables seamless uploading and deployment of Interface maps and device profiles across multiple network devices.

Docker is utilized to facilitate container management and deployment, enhancing the system's scalability and flexibility.

Technology Used: Python, docker, git, Jenkins, Concierge

3. HOMZHUB BACKEND & RECOMANDATION

Client: Homzhub Advisor

Description: My primary role in this project was to develop the backend using Django and rest-api, incorporating a recommendation system and web scraper. The platform utilizes user behavior to provide personalized recommendations for real-estate properties, news, blogs, loans, and more. To gather real estate-related data, I implemented web scraping techniques using libraries such as BS4, requests, and Selenium.

For data processing and handling the large volume of information, I utilized Kafka to stream the data into multiple dynamic dags. Within these dags, PySpark was employed for data filtering and extraction. The processed data was then stored in MongoDB.

Next, I focused on the model processing phase, where the data was vectorized and classified. Convolutional Neural Network (CNN) was chosen as the classification model, enabling efficient and accurate classification of the data.

Overall, the project involved creating a robust backend infrastructure, implementing recommendation features based on user behavior, and utilizing web scraping and CNN for data processing and classification

Technology Used:

Python, Bash, Django, rest-api, pyspark, spacy, bert, Machine learning, nlp, Airflow, kafka, aws(ec2, s3, rds, sqs, sns), azure(databricks), NoSQL (MongoDB, Redis) & SQL(PostgreSQL) database, docker, git, jira

4. RECOMANDATION API FOR WENATURALISTS

Client: WSE GAMES

Description: Building a recommendation system based of user data on WeNaturalists platform (Mobile and web). As per user behavior on the platform, user will get recommended by other user's post, blogs, forms etc. etc. Collaborative filtering has been used for recommendation system and Django with rest-api used for user interaction.

Technology Used:

Python, Django, rest-api, pandas, numpy, spacy, nlp, celery, aws(ec2, s3, rds, sqs), NoSQL (MongoDB) database, git, jira

5. STARTWRITE INDIA

Client: Eclipse Techno consulting solution pvt. Ltd.

Description: It's a desktop application developed using Tkinter. This Software will help children to learn English alphabets in different font.

Technology Used: Python, tkinter, matplotlib

6. WEBSITE VULNERABILITY SCANNER

Client: Indian Cyber Security Solution

Description: A website/mobile application which can scan for vulnerability of another website or network endpoint followed by OWASP Top 10 (2019).

Technology Used: Python, Django, rest-api, NoSQL:(MongoDB), SQL(MySQL), request, SQLmap, Nmap

RESEARCH WORK:

1. ICMP smurf attack detection using neural network.

Git: https://github.com/xhimalaya/ICMP_smurf_attack_detection/

2. Bezier curve algorithms

Git: <https://github.com/xhimalaya/brezier>

Certification:

- Big Data Analysis from CDAC Kolkata.
- Certified by Microsoft on database fundamentals.
- Certified by IIT Mumbai on C Programming.

Hobbies:

Photography, Hacking, Riding