Aninda Kundu

✉ anindakundu1994@gmail.com · 📞 +91 9831846782 · ○ aninda1994 · in /aninda-kundu

PROFILE

I am a Data Analyst having rich 1.8+ years of experience in executing full life-cycle of development projects. Passionate about Data Science & AI, having good programming skills, well versed knowledge of making full end-to-end Data Science pipeline from scraping to deploying applications, having good analytical and problem solving skills. Hands on experience to solve problems like classification, Regression, Clustering.

SKILLS

Languages
Python, SQL

Tools
Jupyter Notebook, Pycharm IDE, Anaconda IDE, Spyder, Postman, Postgres, Perforce, GIT, Jira.

ML/DL Algorithms
Machine Learning Algorithms like Linear regression, Ridge & Lasso regression, Logistic regression, Random Forest, KNN, Ensemble and Boosting Techniques, Clustering Algorithms(K-Means, DB Scan), etc.
Deep Learning Algorithms like ANN, CNN, advanced RNN with LSTM, Transfer Learning Techniques like VGG-16, Resnet 50, Inception V2 etc.

Frameworks & Libraries
Scikit-learn, Pandas, Tensorflow, Keras, NumPy, NLTK, Spacy, Flask, Seaborn, Matplotlib, BeautifulSoup, Selenium.

Operating System
Linux (Ubuntu, Fedora), Windows

Database Technologies
PostgreSQL

PROFESSIONAL EXPERIENCE

Adamas University

Assistant professor

Feb 23 - Present

EXL Services 🔗

Business Analyst

March 2022–Feb 2023 Gurgaon, India

- Currently working as a Business Analyst in US Insurance Domain
- Working on Lawyers pricing project. Using Regression model to predict the loss and frequency of a policy.
- Involved in various data collection, data preprocessing and model building technique.
- Build a custom based generalized web scrapping tool using python BeautifulSoup and Selenium.

Ribbon Communication 🔗
Software Developer Technical Analyst

April 2021-March 2022 Bangalore, India

- Formerly worked as a Software Developer in Telecommunication Domain.
- Worked on agile methodology and jira tool. Followed full process of software Development life cycle.
- Involved in different kind of sanity testing procedure.
- Build an Anomaly detection tool using machine learning which can detect error lines from a log file that contains millions of lines.

Vaultedge Software Pvt LTD 🔗
Software Developer Intern

July 2018 – April 2019 Bangalore, India

- Formerly worked as a paid intern in Mortgage and legal domain.
- Extract relevant information from legal data using rule based technique and machine learning technique.

PROJECTS

A K-means based approach to detect copy number variation in a genome.
AIM : To build a clustering model that can detect the duplication or insertion of unwanted DNA sequence in a genome.

- Technologies Used – Machine learning algorithms includes K-means clustering algorithm, Bio informatics alignment algorithm (Smith-waterman algorithm), Python, sklearn, Pycharm IDE.
- Copy number variation (CNV) is a form of structural variation caused by duplications or deletions of a large DNA segment, which may have a vital impact on human health, causing many neurological diseases as well as cancer. The task is to detect copy number variation in genomic regions.

A Custom based web scrapping tool to extract attorney level information from different websites.
AIM : To build a generalized scrapping tool that can extract relevant information from different websites.

- Technologies Used : Python, BeautifulSoup, Selenium.
- For Lawyers pricing model we required attorney level information (exa: Attorney name, Designation, Address, State, Firm name, Area of practice, Revenue etc). To extract these information from different formatting based website is quite time consuming. With the help of our custom based scrapping tool we can extract relevant information from different web sites with just a single click and save it in csv format. It will save a lot of time and effort.

Anomaly detection from system generated log file.
AIM : During smoke test or full regression test of a feature or a product there we can find out that multiple things can be broken or we can face multiple issue regarding some feature. These all information contains in a log file. These log file contains millions lines of information, so for a developer it is really hard to debug the log file and find the cause of the error. We are automating these process using machine learning technique. Out of those millions lines our model will only highlight those errors which caused a failure.

- Technologies Used – Python, Machine learning, Log reduce, KNN, NLP for data preprocessing, Node red.
- Collected successful logs for model training where failure test case should not be present.
- Created a user interface where a user can put path of a log file and our tool will visualize that how many test cases executed successfully and how many test cases failed, also for failed test cases it will reduce the log file information and only show the lines which caused error.

A Machine learning based application which recommends best crop to grow, best fertilizer to use and diseases caught by crop.
AIM : Detect proper crop and fertilizer based on soil and weather condition using machine learing technique. Detect diseased crop using deep learning technique.

- Technologies used : Python, Sklearn, Tensorflow 2.x, Keras, Random forest classifier, XG boost classifier, CNN, Keras tuner, VGG-16, Resnet 50, Inception V3,
- In the crop recommendation application, the user can provide the soil data and the application will predict which crop the user should grow.
- For the fertilizer recommendation application, the user can provide the soil data and the type of crop they want to grow and the application will predict which type of fertilizer is good for that particular soil.
- For the last application the user can provide an image of a diseased plant leaf and the application will predict what diseases it is and will also give a little background about the

disease and suggestion to cure it.

- **Github Link:**
https://github.com/aninda1994/Harvestify

An end to end application to predict air quality index.
AIM : To build an application that can predict the air quality index level.

- Technologies used : Python, Web scrapping, BeautifulSoup, Linear regression, Ridge and lasso regression, Random forest regressor, XG boost regressor, ANN, Hyperparameter Optimization, Sklearn, Pandas, Model evaluation, Model deployment.
- Prediction of pm2.5 of Bengaluru by analyzing previous year data.

Github Link: https://github.com/aninda1994/Bengaluru-AQI

Deployment link: https://bengaluru-aqi-prediction.herokuapp.com/

COURSES

- **Deep learning Masters (Ineuron.ai)**
- **Coursera - Applied Machine learning in Python (University of Michigan)**
- **Coursera – Applied Text mining in Python (University of Michigan)**
- **Coursera – Introduction to Data Science in Python (University of Michigan)**
- **Udemy – Natural Language Processing with Python**

EDUCATION

- **Vellore Institute of technology (2017-2019, Chennai, India), M.tech Computer science & engineering with specialization in Big data Analytics and Data science (7.8 CGPA)**

- **Meghnad Saha Institute of Technology (2012-2016, Kolkata, India), B.tech Computer Science & engineering (7.34 CGPA)**

- **WBCHSE(70%), 2012, Kolkata, India.**

- **WBBSE(68%), 2010, Kolkata, India.**

LANGUAGE

English, Bengali, Hindi

INTERESTS

Playing Football, Badminton | Research and experimenting Machine learning and deep learning applications