

ROY RAJJYOTI

rajyotiroy2020@gmail.com

+91-8240781936

Expertise in executing full life-cycle development projects within time, budget & quality parameters, as per project management & best practice guidelines in Data Science, Machine Learning, NLP and Artificial intelligence.

Work Experience

Total Experience: 2 years and 8 months

Wipro Ltd.

October 2018 – Present

Project 1 – Data Scientist

- Recommend similar products & review similarities on an e-commerce website.
- Performed EDA to analyze the data set and features to be created from the existing data.
- Involved in Data Cleaning and performed data pre-processing such as Removing HTML tags, Removing Punctuations, Removing Stop words, expanding Contractions etc.
- Generated Vectors (BoW , tf-idf etc) from Text to determine the text based product similarity
- Removing near duplicates items.
- Performed Text Semantic Based feature Similarity(Content based Filtering)
- Identify the reviews given on products are duplicates or not
- Metric Used for accuracy : Log Loss
- Technical skills ML libs: Python- Scikit, NLTK, fuzzywuzzy
- Programming Language: Python

Project 2 – Data Scientist

- Identify whether a given piece of file/software is a malware.
- Performed Exploratory Data Analysis Class distribution
- Performed feature extraction from byte files
- Performed Multivariate analysis of features from byte files
- Performed feature extraction from asm files
- Minimized multi-class error
- Metrics Used for accuracy : Multi class log-loss and Confusion matrix
- Programming Language: Python
- Technical skills ML libs: Python -Scikit , NLTK, Random Forest, Xgboost.

Project 3- Project Engineer

- Privileged Access Management (PAM)
- User on-boarding, Account on-boarding.
- Privileged Session Manager Monitoring, control and isolation of privileged sessions.
- Creation of Password policy for managing the accounts with respect to Client Password policy.
- Securing and maintain privileged accounts into EPV (CyberArk Enterprise Password vault).
- Maintaining component servers.

Skills

- Data Science, Machine Learning and Deep Learning.
- Python, Keras ,Tensorflow
- Statistical analysis, Quantitative Analysis , Predictive Modelling
- Natural Language Processing (NLP)
- ML Libraries such as pandas, matplotlib, sckit learn ,PyTorch , NLTK/ SpaCy, Regular Expressions, etc
- Machine Learning algorithms such as Classification, Regression, Clustering.
- Understanding of various Machine Learning algorithms such as Logistic regression, Linear Regression, Support Vector Machines (SVM), Ensembles and Random Forest.
- Knowledge of Clustering Algorithms such as k-Means, K-Medoids and DBSCAN.
- Deep Learning Algorithms like Multi level Perceptron (MLP) , Data Mining , LSTM, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN).
- CyberArk

Certifications

- NLP with Python certification from LinkedIn learning
- Data Scientist Certification from Udemy
- CyberArk Trustee

Academic Qualifications

- Electrical Engineering (Aug 14 – Jun 18)
[Meghnad Saha Institute of Technology, Kolkata](#)
7.43 – CGPA