SHIBASHISH NAYAK

Phone no.: +91-9337922962

Email id: shibashishnayak97@gmail.com

Github: shibashish97 (Shibashish Nayak) (github.com)

LinkedIn: (14) Shibashish Nayak | LinkedIn

Work Experience: Total 1.2 years experience as a Data Science Intern(1 year in Etlhive and 2 months in Cloud BC labs)

KEY SKILLS

- Data Science
- Python
- Data Visualization
- Machine Learning

FRAMEWORKS/LIBRARIES

- Pandas
- Sklearn
- Matplotlib
- Seaborn
- Tableau
- Flask

ALGORITHIMS

- Linear Regression
- Logistic Regression

PROFESSIONAL SUMMARY

Actively seeking job opportunities as Data Analyst, Data Science and Machine Learning Engineer who takes pride in building models that translate data points into business insights.

TECHNICAL SUMMARY

- Effectively perform **Exploratory data analysis** (EDA) and **data visualization techniques** on categorical and continuous variables.
- **Handle missing data values and outliers** within a data set by using various data pre-processing techniques.
- Insightful knowledge of concepts of **ensemble techniques,** aware about their advantages and disadvantages.
- Built **various regression and classification models** based upon the target variable and compareaccuracy of different models for the same problem.
- Know to **deal with collinearity** associated within a model and **improve accuracy.**
- Knowledge about **deployment of ML** models in **AWS EC2 Instance**

- Decision Trees
- Ensemble Techniques
- KNN
- Clustering
- Apriori
- Time Series Forecasting

DEPLOYMENT

- AWS EC2
- AWS S3
- AWS Redshift

EDUCATION

- CENTURION UNIVERSITY OF TECHNOLOGY AND MANAGEMENT
  B.E. (Mechanical Engineering) – (2014-2018)
  (8.9/10.00) CGPA

- V.N COLLEGE
  XII - 2014(60.5/100.00) PERCENTAGE

PROJECTS

- **3 WAY MATCH(BETWEEN PO, INVOICE AND GOODS RECEIPT)**
  shibashish97/project12: 3 way match (github.com)
  Goal: Using this model, company can automate 3 way match between Purchase Order , Invoice and Goods receipt.
  - First image preprocessing was done , so that we get better accuracy when we extract text from image.
  - Then we have use Easy OCR to extract text from image.
  - Then we have used Cosine Similarity to match the text between the 3 documents.

- **LOAN APPROVAL PREDICATION ANALYSIS**
  shibashish97/project3: Loan Status Prediction Using ML (github.com)
  shibashish97/Project4: Loan Status Prediction Using DL (github.com)
  Goal: Predict if the loan applicant is eligible for loan using both ML and DL.
  - Loan prediction is a classification problem as Loan Status column is categorical (Y, N).
  - Performed all before training the model that isdata profiling, data preprocessing and exploratorydata analysis.
  - Used Logistic Regressor, Decision Tree, Random forest and K-nearest Neighbors and Artifitial Neural Network to compare between the models and to obtain better accuracy.

- SARASWATI SHISHU VIDYA MANDIR

  X -2012

  (83.00/100.00)

  PERCENTAGE

WORK EXPERIENCE

- WORKED AS DATA SCIENCE INTERN AT ETLHIVE FOR ONE YEAR

  (Sep 21 to Sep 22)

- **HOUSING SALE PRICE PREDICTION**

  shibashish97/Project2: Housing Sale Price Prediction (github.com)

  Goal: To predict House sale price using different predictors.

  - Performed data profiling, data preprocessing and exploratory data analysis on the dataset.
  - Create Backward elimination OLS model and create prediction on test data.
  - Make final prediction by Linear Regression Method.

- **RESTAURANT REVIEW CLASSIFICATION**

  shibashish97/project6: Restaurant Review Classification Using ANN (github.com)

  Goal: Using Artifitial Neural Network predict Restaurant Review(Liked/Disliked).

  - Data profiling and missing data treatment and preprocessing of text data.
  - CountVectorize the preprocessed data and convert to array and predict accuracy on the basis of ANN.
  - Then do review classification using user input.

- **TIME-SERIES ANALYSIS ON AIRPASSENGER DATA.**

  shibashish97/Project5: Time Series Analysis (github.com)

  Goal: Forecasting number of passengers for airlines From 1949 to 1960 for every month.

  - Training has been done with an AR model.

  Techniques like rolling mean and Ad-fuller were used to convert the available time series into a stationary time series