Abhishek Suman

abhisuman.2512@gmail.com 9304353539 Bhubaneshwar, Orissa
linkedin.com/in/abhishek-suman-8113b3183

abhisindhu2512 · GitHub

Professional Summary

Microsoft Certified Data Scientist with 2+ years of experience in executing data-driven solutions to increase internal data processing efficiency, accuracy, and utility. Experienced at creating data regression models, using predictive data modeling, and analyzing data mining algorithms to deliver insights and implement action-oriented solutions to complex business problems.

Skills:

- Python, SQL, R
- Power BI, Tableau
- Machine Learning, Deep Learning, Artificial Intelligence, Pyspark, Azure Machine Learning, MLops
- Azure, AWS
- Advance Excel, PowerPoint, DBMS, Hadoop, Databricks
- Data Processing, Data Mining, Statistical Modeling, and Analysis, Data Visualization
- Tensorflow, Pytorch, Keras, Sklearn, NLTK, Spacy, Regex, Pandas, Matplotlib, Statsmodels, Numpy

Experience:

Analyst- Data Science

StratLytics Consulting Private Limited — April 2021 – Present

- Working on Automation COE projects where my responsibility is to develop a machine learning model and provide services on the MLops side.
- Using web scraping techniques and API to extract data from online sources.
- Using different data mining techniques to extract insightful information from structured and Unstructured data.
- Worked on Credit Risk Scorecard Project where build a Logistic Regression to develop a scorecard.
- Build a Hybrid (LDA + Rule-Based) model and did Text mining to extract information for the Email Routing NLP Projects.
- Worked on a Demand and Forecasting project where used four different models – ARIMA, LSTM, AutoML (Azure), and Simple Model (Rolling 7-day Moving average) model to forecast demand of SKUs(Products) as per different locations(Warehouse location) for the next 104 weeks.

Education:

Maulana Abul Kalam Azad University of Technology, West Bengal formerly WBUT

Master of Science - MS, Data Science and Analytics — 2019 - 2021

Rajiv Gandhi Prodyogiki Vishwavidyalaya

Bachelor of Engineering - BE, Mechanical Engineering — 2012 - 2016

Licenses & certifications

Microsoft Certified: Azure Data Scientist Associate - Microsoft

Link:- https://www.credly.com/badges/f64699ad-11d9-4b17-8fa2-7dfb55f75426/public_url

Microsoft Certified: Azure AI Fundamentals - Microsoft

Link:- https://www.credly.com/badges/4aeaec6e-abbd-4680-a502-6515988a930e/public_url

Projects:

1. **Finding Credit Card Defaults using Machine Learning and Neural Network**

   Data consisted of borrower-specific information like gender, credit limit, education, age, and payment behavioural information like past 5 months' billing amount, payment amount, payment status, and target variable of next month's default status. Used K means and Density-Based Scanning (DBSCAN) to segment the portfolio. Model accuracy (10-Fold Cross Validation) increased from 82% to 88% for Logistic Regression, and 83% to 87.8% for Random Forest but remain unchanged at 99.32% for the Sequential Neural Network model. Project Link:- https://github.com/abhisindhu2512/Machine-Learning-and-AI-Projects/blob/main/Credit_Card_Project.ipynb

2. **Auto Text Generation using Pytorch**

   The Project task is to create a model that takes a sequence of words and then completes a whole sentence by autogenerating some text. In this project, I used Pytorch to generate the model. The dataset used in this project consists of a list of sentences. For feeding the text in the model it needs to be prepared to involve the following task cleaning the text mean removing punctuation, symbols, etc from the text following tokenization, padding to make the sequence equal length after that converting the text into an integer.
   Link:- Machine-Learning-and-AI-Projects/Auto_Text_Generation (1).ipynb at main · abhisindhu2512/Machine-Learning-and-AI-Projects · GitHub

3. **Twitter Sentiment Analysis using NLP**

   This project aimed to survey people thinking about Mr. Narendra Modi based on tweets on Twitter. For this purpose, I have collected 100 tweets about Narendra Modi from Twitter using Twitter API in my local MySQL database. Then do some cleaning on the text and then text classification using NLTK library to classify the text into strongly positive, positive, negative, and strongly negative. Then used the dashboard to see my analysis result.
   Link:- Machine-Learning-and-AI-Projects/Twitter Sentiment_Analysis_Project.ipynb at main · abhisindhu2512/Machine-Learning-and-AI-Projects · GitHub