Sourav Mandal *Python Developer*

Vill.:Rajbagan, P.O.: Chakdaha, Dist.: Nadia, State: West Bengal, Pin: 741222 mandal.srv1998@gmail.com
+91 7908790565 27/11/1998 Indian sourav-mandal-390064210 Sourav9827 Portfolio

Skills

Python (*Pandas, Numpy, Matplotlib, Seaborn, Flask, Plotly, Dash, Sklearn etc.*)

Team Leader (*I lead a group of interns under me, monitor their work and as a team we solve various task in Highway Delite.*)

Data Analytics (*SQL, Tableau, PowerBI, Snowflake, Pentaho etc.*)

Deep learning (*ANN, CNN and RNN*)

NLP (*NLTK, Machine Learning, RNN, LSTM, GRU, Self-Attention model, Transforms, Chatbot etc.*)

Machine Learning (*Linear and Logistic Regression, SVM, Xgboost, Random Forest, Naive Bayes etc.*)

Databases (*SQL and Mongo DB.*)

Big Data (*PySpark, Hadoop, Kafka and Spark*)

Computer Science (*Python, SQL, Machine Learning and Data Science*)

MLOps (*Github Action, CI/CD pipelines, Deployment Techniques In AWS,AZURE, Dockers And Kubernetes, MLflow, DVC.*)

MS-Office (*Advanced Excel, Power point, Word, Access.*)

Cloud Services (*AWS, Azure, Heroku, Render etc.*)

Teaching (*Data Science, Mathematics and Computer Science*)

HTML/CSS (*Flask and django api frameworks*)

R Software (*ggplot2, dplyr, plotly kniter, data.table, mlr3 etc.*)

Certificates

Data Analytics Consulting Virtual Internship(KPMG)
(*Data Quality Assessment Data Insights Data Insights and Presentation May 8th, 2021*)

Data@ANZ Program (*Exploratory Data Analysis Predictive Analytics May 13th, 2021*)

Full Stack Data Science (*ineuron.ai*)

Python for Data Science and Machine Learning Bootcamp
(*udemy*)

Introduction to Machine Learning (*Coursera*)

Data Science foundations (*Great Learning*)

Education

Full Stack Data Science program, *ineuron.ai*
03/2022 – 03/2023

Data Science and Analytics, Machine Learning, Deep Learning and Big data basics.

Diploma in Information Technology,
Nehru Yuva Computer Shiksha Kendra
02/2021 – 02/2022 | Chakdaha, India

Typing, MS Office, Advanced Excel, HTML, CSS and Basic C Programming.

Bachelor in Education (B.Ed), *W.B.U.T.T.E.P.A*
09/2021 – 09/2023 | Santipur, India

On-site School Teaching Internship, Seminars, Micro-teaching, Presentations, Models, Learning Designs etc.

Professional Experience

Subject Matter Expert, *Vertocity*
03/2023 – present | Remote, India

Training students on Data Science and Analytics
SQL, Advanced Excel, Python, Statistics, Machine Learning, Power BI, Time series analysis, Deep Learning and NLP.

Data Analytics intern, *ineuron.ai*
05/2023 – 07/2023 | Remote, India

- **Project Title:** NBA Data Analytics Project
- **Tools:** Python, SQL and Power BI
- **Final Output:** A presentation of the findings, including visualizations and recommendations.
- **Roles and Responsibilities:**
  - Collected(Webscrapping) and cleaned NBA player data.
  - Developed and executed data analysis queries.
  - Created visualizations of the data.
  - Presented the findings and documentations to stakeholders.

Data Science Intern, *ineuron.ai*
12/2022 – 02/2023 | Remote, India

- **Project Title**: News Article sorting
- **Objective:** To classify news articles into predefined categories like Sports, Technology, Entertainment, etc. using deep learning methods.
- **Tools:** Python, Google's BERT model, Azure App Services, Flask, HTML/CSS , Pandas, NumPy, Scikit-learn, TensorFlow, and Keras.
- **Final Output:** Achieved 97.96% accuracy on test set, deployed model on Azure App Services using Flask web app for predictions
- **Roles & responsibilities:**
  - Pre-processing the data to make it suitable for the model
  - Fine-tuning the BERT model for News Category Classification
  - Developing the Flask application for the model
  - Deploying the model on Azure app services
  - Testing and debugging the application
  - Documentation of the project

Teacher(Mathematics and CS), *Self Employed*
2017 – present | Chakdaha, India

- Freelancing Teacher of Mathematics and Computer Science
- Creating Presentations and Learning designs for everyday classes.
- Good understanding of Concepts and Application in real life situations.
- Noticeable improvements in Student's Academic results.

Associate, *Highway Delite*
02/2022 – 11/2022 | Remote, India

- Building database for highway related data.
- Train and monitor the work of new interns.
- Research and create documents for tourism data
- Verify and upload the highway information in the backend of the website

Data Analyst Intern, *The Sparks Foundation*
08/2021 – 09/2021 | Remote, India

- Prediction using Supervised ML
- Prediction using Unsupervised ML
- Exploratory Data Analysis - Retail
- Exploratory Data Analysis - Terrorism
- Prediction using Decision Tree Algorithm

M.Sc Mathematics(Pure), *University Of Kalyani*
08/2019 – 08/2021 | Kalyani, India

- Specialization in Complex Analysis and Differential Geometry.
- Advanced Research Project on Wave Equations.

B.Sc (Hons.) Mathematics, *University Of Kalyani*
07/2016 – 07/2019 | Kalyani, India

Courses

Associate Financial Analysis, *invact.com*
06/2023 – 12/2023

TensorFlow Developer Certificate: Zero to Mastery, *Udemy*
08/2023 – 12/2023

Prompt Engineering, *deeplearning.ai*
04/2023 – 04/2023

Introduction to R Software, *NPTEL*
09/2020 – 11/2022

Learning Analytics Tools, *NPTEL*
07/2021 – 10/2021

Probability for Computer Science, *NPTEL*
07/2021 – 10/2021

Introduction to Machine Learning, *NPTEL*
07/2021 – 09/2021

Data Analytics with Python, *NPTEL*
01/2021 – 04/2021

Data Science for Engineers, *NPTEL*
01/2021 – 03/2021

Programming, Data Structures And Algorithms Using Python, *NPTEL*
09/2020 – 11/2020

Introduction to R Software, *NPTEL*
09/2020 – 11/2020

Python for Data Science, *NPTEL*
09/2020 – 10/2020

Languages

English • Hindi • Bengali

Declaration

I hereby declare that all the information furnished above is correct to the best of my belief. I am responsible for the authenticity of all the information.

Sourav Mandal
Chakdaha, 31/12/2023

Projects

Adult Census Income Prediction project

Objective: To develop a Python-based machine learning model for predicting adult income levels. It is currently in the deployment phase.
Tools: Python, ML algorithms, FastAPI app, railway.ai for hosting, Circle CI for CI/CD.
Methodology: Using Python libraries like scikit-learn and pandas, the project preprocesses data, engineers features, and trains the ML model on labeled adult census data. Model evaluation achieved an 83% accuracy on the test dataset.
Results: The ML model achieved 83% accuracy, demonstrating its effectiveness in predicting adult income levels.

Stores Sales Prediction, *Creating a machine learning model that can accurately predict sales for Big Mart Stores*

- Developed a machine learning model to predict sales for Big Mart Stores using data on various features such as location, product type, and promotional activities.
- Utilized Python libraries such as Pandas and scikit-learn to clean and analyze the data, and trained and evaluated the model using a combination of supervised and unsupervised learning techniques.
- Implementing a continuous integration and deployment (CI/CD) pipeline to automate the model training and deployment process.

Earthquake Analysis, *Created Earthquake Analysis Dashboard, predicted risk & intensity, improved disaster response.*

- **Objective**: Analyze earthquake data and create a predictive model for earthquake risk and intensity.
- **Tools**: SparkR, Power BI.
- **Methodology**: Created an ETL pipeline using SparkR to ingest and clean data. Built a machine learning model to predict earthquake risk and intensity. Visualized the results using Power BI.
- **Results**: Provided insights into the earthquake patterns, hotspots, and trends. Predicted the countries at risk for earthquakes and their intensity levels.
- **Impact**: Improved disaster preparedness, risk mitigation, and emergency response strategies for the affected regions.

Tweets Sentiment Analysis,
Data analysis and machine learning using Python and NLTK

- **Objective:** Analyze the sentiment of tweets and classify them as positive, negative, or neutral using machine learning
- **Tools:** Python, Scikit-Learn, SciPy, NLTK, Pandas, Heroku, Falsk, HTML/CSS.
- **Final Output:** Achieved 85% accuracy on test set, deployed model on Heroku as a web app for predictions
- **Roles & responsibilities:** Sole contributor, responsible for all aspects of the project including data collection, preprocessing, model training, and deployment. Managed the entire project from start to finish.

Adidas Sales Dashboard, *Created a sales dashboard using Excel, Power Query, and DAX to analyze sales trends, improve forecasting, and marketing strategies for Adidas.*

- **Objective:** Create a sales dashboard for Adidas to visualize and analyze sales data.
- **Tools:** Excel, Power Query, DAX.
- **Methodology:** Extracted and transformed data from multiple sources, created calculated columns and measures, and designed interactive visuals.
- **Results:** Provided insights into sales trends, product performance, and customer behavior, leading to data-driven decisions.
- **Impact:** Improved sales forecasting, inventory management, and marketing strategies for Adidas.

Mushroom Classification, *Develop an accurate and reliable model for classifying mushrooms.*
09/2022 – 12/2022

- **Objective:** Classify mushrooms into edible or poisonous categories using machine learning
- **Tools:** Python, scikit-learn, pandas, SciPy, NumPy, Flask, Heroku
- **Data**: Mushroom dataset with over 8,000 records and 22 features
- **Preprocessing**: Data cleaning, encoding categorical variables, splitting into training and test sets

- **Model training:** Comparison of various algorithms and selection of best performing algorithm for classification
- **Final output:** Achieved 99% accuracy on test set, deployed on Heroku as a Flask web app for predictions
- **Roles & responsibilities:** Sole contributor, responsible for data collection, preprocessing, model training, deployment, and project management

Algerian Forest fire prediction, *The model could be used to identify areas that are at high risk of forest fires.*

- **Objective:** Develop a machine learning model that can predict the occurrence of forest fires in Algeria.
- **Tools:** Python, Scikit-learn, Jupyter notebook, Pandas, Matplotlib, seaborn, Flask, heroku, HTML/CSS.
- **Data:** Historical data on forest fires in Algeria from 2000 to 2019.
- **Final Output:**
  - A machine learning model that can predict the occurrence of forest fires in Algeria with an accuracy of 98.76%.
  - A Flask app that integrates the model and allows users to input real-time data and get predictions.
- **Roles & Responsibilities:**
  - Sole contributor, responsible for all aspects of the project.
  - Successfully collected, preprocessed, and trained the model.
  - Deployed the model on Heroku and made it available to users.

Puppy adoption site, *Designed and developed a web application using the Flask framework in Python to allow users to browse and adopt puppies*

- Created a full-featured web application that allows users to browse and adopt puppies, including a form for submitting adoption applications
- Implemented a form for users to submit applications to adopt a puppy, including validation and error handling
- Stored the application data in a MySQL database and used Flask-SQLAlchemy to query and manipulate the data
- Utilized Jinja templates to create dynamic web pages and integrated Bootstrap for a responsive design
- Successfully deployed the app on Heroku, making it accessible to users around the world.

Review and Image Scrappers,
Flipkart Review Scrapper and google image scrapping

03/2022 – 04/2022

- Developed a Python script to scrape product reviews from Flipkart and images from website (https://github.com/Sourav9827/Image-Scrapper.git)
- Developed a Python script to scrape images from google search. (https://github.com/Sourav9827/ReviewScrapperFlipkart.git)
- Utilized the Beautiful Soup library to parse HTML and extract relevant data
- Stored the data in a sqlite database and used the Pillow library to process and resize the images

HR Attrition Dashboard, *Excel-based HR dashboard improved employee retention and engagement.*

- **Objective**: Create a dashboard to analyze employee attrition and provide insights for HR decisions.
- Tools: Excel, Power Query, DAX.
- **Methodology**: Extracted and transformed employee data from multiple sources, created calculated columns and measures, and designed interactive visuals.
- **Results**: Provided insights into attrition rates, reasons, and patterns, leading to data-driven decisions.
- **Impact**: Improved employee retention, engagement, and satisfaction for the company.

Adventure Works Reports, *Analyzed sales data for Adventure Works Cycles using Power BI*

- Created a comprehensive dashboard that allowed the sales team to track their performance and identify opportunities for improvement
- Created a dashboard to track key performance indicators such as sales, revenue, profit, and returns
- Utilized Power BI features such as filters, slicers, and visualizations to present the data in an interactive and easy-to-understand format

Insurance Fraud Detection, *Develop and implement a fraud detection system for an insurance company.*

- **Problem Statement/Objective:** To develop a machine learning model that can accurately detect fraud in insurance claims to minimize losses and improve claim processing efficiency.
- **Tools:** Python, Jupyter Notebook, Scikit-learn, Pandas, NumPy, Flask, Heroku, HTML/CSS.
- **Final Output:**
  - A machine learning model with accuracy of 77.2% in detecting fraudulent claims that can predict the likelihood of a claim being fraudulent based on various features such as claim amount, policy holder history, and claim type.
  - The model was deployed on Heroku as a web application that can be used to submit insurance claims and receive an instant fraud prediction.
- **Roles & Responsibilities:** Sole developer and project manager. Conducted data analysis and preprocessing, trained and evaluated machine learning models, and deployed the final model on Heroku.
  -

Customer Details Dashboard, *Excel-based customer dashboard provided insights and improved business performance.*

- **Objective:** Create a dashboard to analyze customer data and provide insights for business decisions.
- **Tools:** Excel, Power Query, DAX.
- **Methodology:** Extracted and transformed customer data from multiple sources, created calculated columns and measures, and designed interactive visuals.
- **Results:** Provided insights into customer demographics, behaviors, and preferences, leading to data-driven decisions.
- **Impact:** Improved customer retention, acquisition, and satisfaction for the business.

Case Study SQL - Danny's Diner,
Analyse Danny's Restaurant's Database

- **Objective:** Analyzing Danny's Diner's data using SQL to help improve customer experience and make informed business decisions.
- **Tools:** SQL, Google Sheets (datasets provided)
- **Methodology:** Write SQL queries to answer specific questions about customer spending, visiting patterns, menu preferences, loyalty program, and points calculation. Create basic data tables for easy data inspection without SQL joins.
- **Results:** Insights generated on customer spending, visiting patterns, and menu preferences. Evaluation of loyalty program effectiveness. Calculation of points earned by customers.
- **Impact:** Enable Danny to understand customer behavior, enhance personalized experiences, expand the loyalty program, and optimize menu offerings based on customer preferences.

Calefornia House Price Estimation, *Predictive Modeling for California Housing Prices: Harnessing Data for Accurate Price Projections.*

- **Objective:** Develop a predictive model for California house prices using machine learning.
- **Tools:** Employed Python, Pandas, Scikit-Learn, XGBoost for data analysis and modeling, Github actions and Heroku.
- **Methodology:**
  - Implemented regression models, feature engineering, and hyperparameter tuning.
  - Led data pipeline creation & model deployment for real-world application.
- **Results and Impact:** Achieved a robust model with 92% accuracy, aiding in informed real estate decisions and enhancing market understanding.