

Sourav Mandal *Data Scientist*

📍 Vill.:Rajbagan, P.O.: Chakdaha, Dist.: Nadia, State: West Bengal, Pin: 741222 ✉ mandal.srv1998@gmail.com

☎ +91 7908790565 📅 27/11/1998 🇮🇳 Indian in sourav-mandal-390064210 🔄 Sourav9827 📁 Portfolio



👤 Profile

As a highly motivated and skilled data scientist with a postgraduate degree in Mathematics, I possess a solid understanding of various data science concepts and tools, including Python, SQL, Machine Learning, Deep Learning, NLP, and PowerBI. With hands-on experience in multiple data science projects and internships, I have developed a strong problem-solving and analytical mindset. Additionally, my past experience in teaching mathematics and mentoring a team of interns for a backend operation internship showcases my excellent communication and leadership skills. As a result, I am confident in my ability to leverage my skills and experience to contribute to any data-driven organization's success.

🧠 Skills

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

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

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

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

Databases (SQL and Mongo DB.)

Big Data (PySpark, Hadoop, Kafka and Spark)

Deep learning (ANN, CNN and RNN)

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

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.)

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

Teaching (Data Science, Mathematics and Computer Science)

📄 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.

💼 Professional Experience

Data Science SME, Vertocity

03/2023 – present | Remote, India

Creating Dashboards, Generating Insights and Automation.

Providing Training 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 (Webscraping) 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

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.

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

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.

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.

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.
-

California 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.