

# Sourav Mandal *Data Analyst*

📍 Rajbagan Chakdaha Nadia West Bengal,  
741222 Chakdaha, India

☎ 08618241725

🇮🇳 Indian

in [www.linkedin.com/in/sourav-mandal-390064210](https://www.linkedin.com/in/sourav-mandal-390064210)

🔗 <https://sites.google.com/view/sourav9827>

✉ [mandal.srv1998@gmail.com](mailto:mandal.srv1998@gmail.com)

📅 27/11/1998

💍 Unmarried

🌐 <https://github.com/Sourav9827>



## 🧠 SKILLS

### Python

Pandas, Numpy, Matplotlib, Seaborn, Flask, Plotly, Dash, Sklearn etc.

### Tableau

Data Visualization, Analytics, Dashboards etc.

### MS-Office

Excel, Power point, Word, Access.

### HTML/CSS

Flask and django api frameworks

### Power BI

DAX, Data Visualization, Dashboarding etc.

### Databases

SQL and Mongo DB.

### Team Leader

I lead a group of interns under me, monitor their work and as a team we solve various task in my present company.

### R Software

ggplot2, dplyr, plotly kniter, data.table, mlr3 etc.

## 📁 PROFESSIONAL EXPERIENCE

2017 – present  
Chakdaha, India

### Teacher(Mathematics), Self Employed

I am a private tutor in mathematics and teach mathematics to students of class XI and above, and science and mathematics to students from class VIII to XII.

02/2022 – 11/2022  
Work from home, India

### Mentor of NH Mapping, Highway Delite

Monitor the work of interns under me, give training and work on the backend maintenance of the website.

08/2021 – 09/2021  
Work from home, India

### Data Analyst Intern, The Sparks Foundation

## 📁 PROJECTS

07/2022 – present

### 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

12/2022 – 01/2023

### BBC News Category Classification,

Conducted a natural language processing (NLP) analysis of BBC news articles to classify them into different categories

- Used NLP techniques to classify News Category as Entertainment, Political, Tech or Sports
- Utilized Python libraries such as Pandas, scikit-learn, and NLTK to clean and analyze the data
- Trained and evaluated a Deep learning model (BERT) on a dataset of over 1400 News Article
- Achieved an accuracy of 97.96% on the test set

10/2022 – 11/2022

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

- Used NLP techniques to classify tweets as positive or negative
- Trained a machine learning model on a dataset of over 10,000 tweets
- Achieved an accuracy of 95% on the test set

06/2022 – 07/2022

### California Housing Price Estimation,

Developed a machine learning model to predict housing prices in California using data from the California Census

- Utilized Python libraries such as NumPy, Pandas, and scikit-learn to preprocess and analyze the data
- Trained and evaluated the model using cross-validation and various metrics such as mean squared error and R squared.
- Created a machine learning model that achieved an R squared value of 0.87 on the test set, demonstrating strong predictive power

- Successfully implemented ML Ops practices, allowing the model to be trained and deployed automatically, reducing the time and effort required to maintain the model.

06/2022 – 06/2022

#### Algerian Forest fire prediction,

*Created a machine learning model that can accurately predict the likelihood of future forest fires in Algeria* [↗](#)

- Conducted a data analysis of past forest fire occurrences in Algeria, including meteorological and vegetation data
- Developed a machine learning model to predict the likelihood of future forest fires using the data
- Utilized the Pandas library to preprocess the data and the scikit-learn library to train and evaluate the model
- Achieved an accuracy of 98.76% on the test dataset

06/2022 – 06/2022

#### 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

03/2022 – 04/2022

#### 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.

03/2022 – 04/2022

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

[↗](#)

- Developed a Python script to scrape product reviews from Flipkart and images from website (<https://github.com/Sourav9827/Image-Scraper.git>)
- Developed a Python script to scrape images from google search. (<https://github.com/Sourav9827/ReviewScrapperFlipkart.git>)
- Utilized the BeautifulSoup 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

## EDUCATION

09/2021 – present  
Santipur, India

**Bachelor in Education (B.Ed), W.B.U.T.T.E.PA**

02/2021 – 02/2022  
Chakdaha, India

**Diploma in Information Technology, Nehru Yuva Computer Shiksha Kendra**

08/2019 – 08/2021  
Kalyani, India

**M.Sc Mathematics(Pure), University Of Kalyani**

07/2016 – 07/2019  
Kalyani, India

**B.Sc (Hons.) Mathematics, University Of Kalyani**

11/2021 – present

**Full Stack Data Science program, ineuron.ai** [↗](#)  
*Data Science and Analytics, Machine Learning, Deep Learning 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)**

## 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**  
Nadia, 21/04/2022

## COURSES

09/2020 – 11/2022	<b>Introduction to R Software, NPTEL</b>
07/2021 – 10/2021	<b>Learning Analytics Tools, NPTEL</b>
07/2021 – 10/2021	<b>Probability for Computer Science, NPTEL</b>
07/2021 – 09/2021	<b>Introduction to Machine Learning, NPTEL</b>
04/2021 – 05/2021	<b>Introduction to Machine Learning, Duke University(Coursera)</b>
01/2021 – 04/2021	<b>Data Analytics with Python, NPTEL</b>
01/2021 – 03/2021	<b>Data Science for Engineers, NPTEL</b>
09/2020 – 11/2020	<b>Programming, Data Structures And Algorithms Using Python, NPTEL</b>
09/2020 – 11/2020	<b>Introduction to R Software, NPTEL</b>
09/2020 – 10/2020	<b>Python for Data Science, NPTEL</b>