|  |  |
| --- | --- |
| Data Analyst/ Date Scientist  sutirtha mukherjee | +91 7864831687 | sutirthapg@gmail.com  https://www.linkedin.com/in/sutirtha-mukherjee-55a624a1/  Kolkata, 700032 |

# Education

phd in theoretical physics– Jadavpur University and IACS, Kolkata, India 2016

M.Sc. in physics– The University of Burdwan, West Bengal, India 2009

# Skills

|  |  |  |
| --- | --- | --- |
| **Python**  Pandas, NumPy, matPlotLib, Seaborn, beutifulsoup2  **Machine Learning and Deep Learning**  scikit-learn, Keras | **SQL**  MySQL, SQL Server, PostgreSQL,  SQLite  **Excel**  Pivot Table, Pivot Chart, VLOOKUP | **Power Bi**  DAX, Power Query, Visualization  **C and FORTRAN programming**  Matrix diagonalization, Monte  Carlo, MPI-OMP parallelization |

# Projects

credit card fraud detection with machine learning –ml Internship Project **–**CodSoft, Kolkata October 2023

* Developed Python-based machine learning (ML) models to detect credit card fraud in a dataset of 1.3 million transactions.
* Conducted data cleaning, exploratory data analysis (EDA), and feature engineering using **Pandas**, **Matplotlib**, and **Seaborn**.
* Implemented ML algorithms, including Logistic Regression, Random Forests, and XGBoost, to create predictive models classifying transactions as fraudulent or legitimate.

customer churn prediction – ml Internship Project **–** CodSoft, Kolkata November 2023

* Developed machine learning models to predict customer churn in commercial banks.
* Conducted data cleaning, transformation, and EDA to extract insights and prepare data for predictive modeling.
* Compared various ML algorithms, including SVM, Random Forest, Boosting to construct a robust churn prediction model.

**Spam SMS Detection** – ML Internship Project **–** CodSoft, Kolkata November 2023

* Implemented Python and Natural Language Processing (NLP) to identify spam messages in a dataset of 5500 SMS messages.
* Applied text processing techniques, including TF-IDF, to convert textual data into numerical vectors suitable for ML.
* Utilized ML algorithms, specifically Logistic Regression and Support Vector Classifier, achieving a 96% accuracy and f1-score.
* Optimized model performance through EDA, feature engineering, and fine-tuning hyper-parameters.

Customer Sales Analysis – Personal Project **–** Philadelphia, PA November 2019

* Utilized **SQL** to extract data from 8 different related tables from customer sales databases using **JOIN** and **VIEW**
* Transformed and filtered data by using aggregating and filtering function to improve reporting process
* Loaded and visualized data with **Python** to identify key business intelligences that can improve sales performance

# Work Experience

post doctoral researcher **–** Korea Institute for Advanced Study **–** Seoul, South KoreaAugust 2016-August 2022

* Developed computational codes to investigate the physics of intriguing materials.
* Published multiple research articles in highly impactful scientific journals

Machine learning internship – CodSoft – Kolkata October 2023-November 2023

* Developed predictive models across four distinct machine learning projects.

# Data-Related Certifications

* IBM Data Science Professional Certificate. July, 2023
* Microsoft Power BI Desktop for Business Intelligence December, 2023
* Advanced SQL: MySQL for Ecommerce & Web Analytics December, 2023