Ishan Dutta

MACHINE LEARNING INTERN

CONTACT

📞 9560227020

✉️ ishandutta92@gmail.com

📍 Delhi NCR- India- 201014

Linked Profile Link-
linkedin.com/in/ishandutta0301

Github
github.com/Dermisfit0301

Kaggle
kaggle.com/dermisfit

EDUCATION

2023
Data analyst masters & Data science from Simplilearn, Remote

2021
BBA From business intelligence SCMS NOIDA, Noida, UP

SKILLS

- Programming Languages: Python, R
- Data Analysis: Data preprocessing, feature engineering, exploratory data analysis (EDA), statistical analysis
- Data Visualization: Tableau, Power BI
- Machine Learning: Model building, MLOps, classification algorithms, clustering algorithms, Random Forest Regressor
- Web Development: Streamlit, Flask
- Cloud Services: AWS (deployment)
- Languages: Oracle SQL
- Business Intelligence Tools: Tableau, Microsoft Power BI

PROFILE

I am a proactive and results-driven Data Analyst with hands-on experience in machine learning and data analysis. With a passion for uncovering insights from complex datasets and a drive to contribute to innovative projects, I have a proven track record of utilizing advanced techniques to transform raw data into valuable insights. My strong analytical skills, proficiency in various programming languages, and commitment to teamwork make me a valuable asset to any data-driven project.

EXPERIENCE

Machine Learning Intern
Innomatics Research labs, Remote India |
March 2023 - June 2023

Roles & Responsibilities

- Collaborated with the team to develop interactive web applications using Streamlit and Flask frameworks in Python, enhancing user experience and engagement.
- Applied Python programming to preprocess and manipulate raw data, ensuring its quality and suitability for machine learning tasks. Implemented effective feature engineering techniques to improve model performance.
- Utilized Python libraries and frameworks to create machine learning models that effectively solved specific business problems, contributing to the development of innovative solutions.
- Conducted thorough data analysis, extracting meaningful insights and patterns from datasets. Translated these insights into actionable recommendations and presented them in the form of clear and concise reports.
- Effectively communicated findings and analysis results to stakeholders, fostering a clear understanding of the implications and potential applications of the data-driven insights.

- Implemented MLOps (Machine Learning Operations) practices to streamline the deployment and management of machine learning models. Ensured seamless integration of models into production environments.
- Utilized AWS (Amazon Web Services) to deploy machine learning applications on the cloud, improving accessibility, scalability, and reliability of the applications.
- Employed MS Excel to meticulously clean and format data, eliminating inconsistencies and ensuring data integrity for accurate analysis and modeling.

Data Analyst Intern
The Sparks Foundation, Remote- Singapore | January 2023 March 2023
Roles & Responsibilities

- I had the opportunity to utilize Python and R programming languages to perform in-depth data analysis, predictive modeling, and data visualization.
- Throughout the month of March 2023, as a remote intern based in Singapore, I was immersed in conducting exploratory data analysis and employing statistical techniques to uncover meaningful patterns within the datasets I worked with as part of my role at The Sparks Foundation.
- In January 2023, my role as a Data Analyst Intern with The Sparks Foundation allowed me to apply my skills in Python and R programming to dissect and analyze data, ultimately contributing to the development of valuable predictive models while working remotely.
- As the calendar turned to March 2023, my remote position at The Sparks Foundation had me engaged in the intricate task of creating interactive and informative dashboards using tools like Tableau and Power BI, ensuring effective communication of data-driven insights to key stakeholders.
- From January to March 2023, my tenure as a Data Analyst Intern at The Sparks Foundation exposed me to the dynamic realm of data analysis, where I utilized Python and R to extract insights, conducted comprehensive exploratory analyses, and developed compelling interactive dashboards for stakeholders, all while operating from Singapore.

Operations Intern MPP
MPP New Delhi, India | December 2021 - April 2022
Roles & Responsibilities

- Tracked product flow from pick-up to delivery via logistics systems.
- Handling various transactions and working with the accounts department.
- Procured equipment and handled their transportation to the site.
- Made use of MS Excel for optimization and scheduling purposes.
- Rearranged shipping schedules to combine loads, and maximize vehicle usage.

PROJECTS

Resume Keyword Analysis and Automation
- Developed a web scraping tool using Selenium to automatically parse job descriptions from LinkedIn.
- Utilized NLP techniques to extract and analyse keywords from job descriptions.
- Employed clustering algorithms from the sci-kit-learn package to group keywords based on frequency.

2.Plant Health Detection Application
- Created an application using computer vision to detect plant diseases based on leaf images.
- Utilized Python and OpenCV for image processing, enhancement, and segmentation.
- Developed machine learning models to classify and identify plant diseases.

Laptop Price Estimation

- Built a machine learning model using the Random Forest Regressor algorithm to estimate laptop prices.
- Utilized features such as drive type, operating system, RAM, and warranty type to predict laptop prices.
- Developed a web application using Streamlit framework to provide an easy-to-use interface for price estimation.

Pubspook - Pub Finder Web Application

- Developed a web application using Streamlit framework to help users find pubs near their location in the UK.
- Integrated features such as filtering pubs by local authority and town, displaying pub counts by regions, and searching for pubs based on user coordinates.
- Implemented the Euclidean distance concept to calculate the distance between user coordinates and pub locations and displayed the top 5 pubs within walking range of the user based on the shortest distance.

Buri Aatma Classification Application

- Developed a web application using the Streamlit framework to classify evil spirits into their respective categories for effective invocation of mantras.
- Made use of classification models like Logistic regression, decision tree classifier and naive bayes to develop the model.
- Leveraged Streamlit (dashboard development), Pickle (dumping model), and MLflow (to select the best performing model) to simplify and streamline the model development and deployment process.

CERTIFICATIONS

- Data analysis using Excel from Macquarie University via Coursera . IBM Data Science with R, IBM, Jul 2021
- Data analyst toolbox from John Hopkins University via Coursera

LANGUAGES

- Bengali: First Language
- English: Advanced
- German: Beginner
- Hindi: Advanced
- Urdu: Intermediate

I do hereby declare that the particulars of information and facts stated here in above are true, correct and complete to the best of my knowledge and belief.

Ishan Dutta