Siddhant Rawat

Email-id: sidrawat26july@gmail.com

Mobile No.: 6398839371,

https://github.com/sidrawat22

ACADEMIC DETAILS

Year	Degree/Exam	Institute	GPA/Marks(%)
Sep, 2020 - Jun, 2024	B.TECH in Computer Science	Graphic Era University Dehradun	8.5/10
2019	C.B.S.E	Doon International School, Dehradun	82.00 %
2017	C.B.S.E	Doon International School, Dehradun	10/10

Experience

• Tata Consultancy Services (July 2024 - Present)

Applied advanced SQL and Python (Pandas, Matplotlib) for statistical modeling and risk pattern analysis on large datasets (500K+ rows). Developed automated variable creation scripts and data validation models, reducing manual efforts by 40% while improving model accuracy. Created predictive analytics dashboards using statistical techniques to identify business anomalies and enhance decision-making in production environments. Collaborated cross-functionally to implement model enhancements that directly improved business outcomes and operational efficiency.

Certifications

• Internshala (Oct 2022 - Nov 2022)

Exploratory Data Analysis And Machine Learning, Internshala Trainings, Online

• Microsoft Azure Fundamentals (Dec 2021)

Proving knowledge of cloud computing concepts, models, and services, such as public, private, and hybrid cloud, in addition to infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS).

• AWS Certified Solutions Architect - Associate (May 2023)

AWS Certified Solutions Architect - Associate showcases knowledge and skills in AWS technology, across a wide range of AWS services.

• TensorFlow Developer Certificate (September 2023)

TensorFlow showcases the knowledge and skills to solve deep learning and ML problems.

PROJECTS

• Text Detection(Human vs AI) (Nov, 2023 - Dec, 2023)

To create a machine detection model which can differentiate human generated text from machine generated text. The model is also capable of classifying different types of text generating models like ChatGPT, Dolly-v2, DaVinci, Bloom and Cohere. To achieve this TF-IDF vectorizer was used. It's based on the idea that words that appear more often in a document are more relevant to the document. It was also found that Extra Trees classifier gave the most satisfying results (94% accuracy).

• Emotion Detection (Jun, 2023 - July, 2023)

To create a machine detection model which can detect human emotions and classify them. The aim of this project is to design a network that is capable of classifying the emotions, and then use Opencv for the detection of the faces and then pass it for prediction.

• Hate Speech Detection (Dec,2022-Jan,2023)

Developed a classification model to detect hate speech in tweets posted on X (formerly Twitter). Performed extensive exploratory data analysis including frequency distributions, feature correlation matrices, and TF-IDF vectorization to extract meaningful patterns from text data. Evaluated model performance using statistical metrics such as precision, recall, F1-score, and confusion matrix. Compared Decision Tree (87.55%)

accuracy), Logistic Regression (89.68%), and Support Vector Machine (90.05%) classifiers to identify the most effective model for deployment.

• Credit Risk Prediction (Jan, 2021 - Feb, 2021)

Credit Risk Assessment Variable Engineering (Jan 2021 - Feb 2021) Developed predictive credit risk models using statistical and econometric techniques to assess loan default probability. Engineered custom risk variables through feature creation and selection, conducting multi-collinearity analysis and variable importance ranking. Implemented Support Vector Machine with advanced model validation techniques (ROC-AUC, confusion matrix) achieving 83.33% accuracy across imbalanced datasets, directly supporting credit decision-making processes.

TECHNICAL SKILLS

- **Programming Languages** Python(proficient), C, C++(basic).
- Databases & Querying DBMS, MySQL, ETL concepts.
- Tools and Frameworks AWS (Amazon Web Services), Tableau, Power BI, MS Excel, Git, Jupyter, VS Code.
- Machine Learning and Analysis TensorFlow, Keras, scikit-learn, NumPy, Pandas, Matplotlib, Seaborn, Classification models, Predictive analytics, Model validation, Performance optimization.
- Other Skills Data Cleaning, Data Visualization, Feature Engineering, Dashboarding, Econometric modeling, Statistical analysis, Variable engineering, Feature selection.