Saurabh Kumar Srivastava

Varanasi, Uttar Pradesh 221206

+916388126556 srisaurabh2001@gmail.com LeetCode GitHub LinkedIn

Internship

Data Science Intern **Jan 2024 - Sep 2024**

iNeuron.ai

- Built a sales prediction model with **92% accuracy** using logistic regression, analysing 10,000+ product prices across 3 categories.
- Cleaned and pre-processed 6 months of sales datasets, improving data quality from using Python, **Power bi** and **Excel**.
- Conducted EDA on **$2M worth** of sales data, identifying key pricing trends using Pandas, Seaborn and Matplotlib.
- Made the dashboard on power bi on actionable features and outcome.

Projects

Diamond Price Prediction | *Python3, Scikit-learn, Flask* **Sep 2024 - Nov 2024**

- Designed and implemented a **supervised machine learning model** to predict diamond prices based on key attributes such as carat, cut, colour, and clarity.
- Developed and fine-tuned models using **Linear Regression, Decision Trees,** and **Random Forest**, achieving a **95% R^2 score**, ensuring high prediction accuracy.
- Engineered and optimized features using **Pandas and NumPy**, improving model performance and reducing error variance.
- Built a **Flask-based API** for real-time price estimation and integrated **MySQL** for storing prediction history and insights.
- Applied software engineering best practices, including modular code design, version control (Git), and API integration for scalability.
- **GitHub Repository Link**: github.com/saurabhsrivastava2001/DiamondPricePrediction

Digit Classification *| Python3, Keras, CNN, TensorFlow, Bash, Scikit-learn* **May 2024 - Aug 2024**

- Built a CNN model using Keras and TensorFlow, achieving **98.5% accuracy** on the 10,000-image test set.
- Processed 70,000+ images by applying normalization and one-hot encoding for optimized training.
- Designed a 4-layer architecture with 32+64 convolutional filters, ReLU activation, and **dropout regularization** (0.25-0.5) to enhance generalization.
- Trained for 10+ epochs using Adam optimizer, improving efficiency, and reducing overfitting.
- **GitHub Repository Link**: github.com/saurabhsrivastava2001/MNIST_digit_clssification

Certificates

Full Stack Data Science Master's Program *(Ineuron.ai)* **15 May'23 – 16 Sept'24**

Python basics *(HackerRank)* **4th Oct 2024**

SQL basic, intermediate *(HackerRank)* **5th Oct 2024**

Technical Skills

- **Languages:** Python, SQL, Java, C++
- **Libraries:** Pandas, NumPy, Matplotlib, Seaborn, scikit-learn, TensorFlow, Keras, NLTK, OpenCV, spaCy
- **Web Development:** Flask, Django, HTML/CSS, Bootstrap, REST APIs
- **Machine Learning:** Regression, Random Forest, SVM, KNN, XGBoost, LightGBM, Ensemble Methods
- **Deep Learning:** Neural Networks, CNN, RNN, LSTM, GRU, Transformers, Transfer Learning, Time-Series analysis
- **Computer Vision:** Object Detection, Image Segmentation, Feature Extraction
- **NLP:** Text Classification, Sentiment Analysis, Named Entity Recognition
- **Data Analysis:** EDA, Statistical Analysis, Power BI, A/B Testing, Time Series Analysis
- **Data Formats:** JSON, XML, YAML, CSV, Parquet, Avro
- **Databases & Tools:** MySQL, MongoDB, PostgreSQL, Git, Jupyter Notebook, MLflow

Education

Lovely Professional University Punjab **2024 – Current**

Master of Computer Applications; CGPA: 8.8 *Jalandhar, Punjab*

Udai Pratap College **2019 – 2022**

Bachelor of Science – Mathematics; CGPA:7.05 *Varanasi, Uttar Pradesh*

National Inter college **2018 – 2019**

Intermediate CGPA – 8.01 *Varanasi, Uttar Pradesh*