Prakrit Sinha

Minneapolis, MN ▪ sinha239@umn.edu ▪ (469)-631-3903
▪ https://www.linkedin.com/prakritsinha ▪ https://github.com/prakrit14

EDUCATION

Master of Science, Data Science — **University of Minnesota, Twin-Cities**
GPA: 3.26 / 4.00 — *May 2025*
Coursework: Theory of Statistics, Applied Regression Analysis, Introduction to Data Mining, Principles of Database Systems, Computer Vision, Advanced Machine Learning, Optimization for Machine Learning

Bachelor of Technology, Electronics and Communication — **SRM University**
GPA: 3.5 / 4.00 — *May 2022*
Relevant Coursework: Linear Algebra, Calculus, Probability and Stochastic Processes, Applied Machine Learning, Scientific Python

PROJECT EXPERIENCE

Machine Learning Modeling for Heat Exchangers, Daikin Applied Americas — *September 2024 – December 2024*
- Currently working on an industrial capstone project along with a team of industrial and mechanical engineers in determining which variables are most influential in predicting heat transfer and pressure drops in brazed plate heat exchangers in HVAC systems.

Thermo-Crime Analysis for Minneapolis, Principles of Database Systems — *September 2023 – December 2023*
- Utilized **SQL**, **Python**, and **R** to gather, clean, and integrate data from various sources, investigating the correlations between weather temperature and crime rates.
- Employed **regression analysis** and **ANOVA** testing to build explanatory models and conduct **hypothesis testing**, presenting findings through meaningful **visualizations** and reaching a conclusion crime has a direct positive dependency on temperature.

Enhancing vision-language models for intelligent trash sorting, Computer Vision — *January 2024 – April 2024*
- Collaborated with an interdisciplinary team from data science and robotics to **fine-tune BLIP2** vision-language model on the Trashnet dataset to generate captions and feed it into **zero-shot model OWL-ViT** to detect and classify trash objects.
- Successfully attained **0.6 IoU similarity** in ground truth bounding boxes and **80% trash classification accuracy** on the TACO test dataset.

Fake News Detection Through Advanced ML Models, Machine Learning:Analysis and Methods — *January 2024 – April 2024*
- Worked with several state-of-the-art **LLM models** including BERT, RoBERTa, DistilBERT, LSTM and CNN to detect circulating fake news.
- Fine-tuned the language model **BERT** on the LIAR dataset to classify fake news. **Attained 60% and 35% accuracy** on binary and multi-label type classification respectively.

PROFESSIONAL EXPERIENCE

***Data Science Intern**, Highradius Corporation* — *May 2021 – August 2022*
- Utilized **Python** to process and **clean raw data** and execute the Synthetic Minority Over-sampling Technique (**SMOTE**) to tackle class imbalance within a customer dispute validity prediction model.
- Applied **supervised learning** techniques, resulting in a substantial **15% enhancement in classification accuracy**. **Proficiency in Python** was demonstrated through **effective preprocessing and feature engineering**, underscoring a dedication to leveraging innovative machine learning strategies for improved model performance in complex data scenarios.
- Enforced **ensemble techniques like boosting and bagging** for **anomaly detection** among customers to enhance the payment date prediction model to flag up to **5% more defaulters**.

LEADERSHIP & ACTIVITIES

***Instructor,** Highradius Corporation* — *April 2022 – July 2022*
- Trained 120 interns on machine learning fundamentals, achieving a remarkable **95% comprehension rate** and fostering a **90% satisfaction rate** among participants. Facilitated a conducive learning environment through adept instruction and guidance.
- Attained the highest conversion rate to full-time interns among 12 cohorts, exhibiting a notable **25% enhancement** compared to predecessors. Demonstrated the effectiveness of teaching and support strategies in promoting successful transitions for interns.

***Customer and Receiving Specialist,** UMN Reuse Program* — *September 2023 – May 2024*
- Processed customer purchases and assisted with loading logistics while providing courteous customer service.
- Organized and priced over 2000 lbs. of incoming material weekly, resulting in a **10% boost in overall sales**.

SKILLS, TOOLS, AND PACKAGES

- **Programming & Frameworks:** Python, SQL, R, Scikit-learn, PyTorch, TensorFlow, Keras, PySpark, NLTK, OpenCV
- **Data Mining and Machine learning:** association rules, anomaly detection, regression models, SVM, CART, ensemble models (random forest, XGBoost, AdaBoost), deep learning (MLP, CNN, RNN, Transformers), Unsupervised learning (KMeans, DBSCAN)
- **Data Management and Technologies:** MySQL, AWS, data preprocessing (wrangling, cleaning, merging, sampling, aggregation), Database modeling (ER model, indexing, transaction management), data warehousing
- **Statistical Analysis and Experimentation:** statistical data modeling, regression analysis, ANOVA testing, experiment design (A/B testing), visualization
- **Tools and Software:** MS Power BI, Tableau, MS Excel, MS PowerPivot, MS PowerPoint, MS Word, JupyterNB, Docker, Git