**PROFESSIONAL SUMMARY**

Master’s student in Data Science with 2 years of research. Committed and perceptive Data science master's student with a solid background in statistical analysis, data visualization, and machine learning. knowledgeable with big data tools, including Python, R, and SQL, and skilled at managing intricate datasets to extract useful insights. adept at using AI, deep learning, and predictive modeling to solve practical issues. enthusiastic about using data-driven solutions to improve corporate performance and optimize decision-making. Excellent problem-solving skills, meticulous attention to detail, and a team-oriented attitude; ready to offer technical know-how to creative data science initiatives.

**Core competencies:**

** Data Analysis & Visualization –** Proficient in Python (Pandas, NumPy, Matplotlib, Seaborn), R, and SQL for data manipulation and visualization.

** Machine Learning & AI –** Experience with supervised and unsupervised learning techniques using Scikit-learn, TensorFlow, and PyTorch.

** Big Data Technologies –** Familiar with Hadoop, Spark, and cloud platforms (AWS, Google Cloud, Azure) for scalable data processing.

** Deep Learning & Neural Networks** – Knowledge of CNNs, RNNs, and transformers for computer vision and NLP applications.

** Statistical Analysis & Predictive Modeling** – Strong foundation in hypothesis testing, regression, time series forecasting, and Bayesian statistics.

** Data Wrangling & Preprocessing** – Expertise in cleaning, transforming, and structuring large datasets for machine learning pipelines.

** SQL & Database Management –** Experience in writing complex queries and working with relational and NoSQL databases**.**

** Natural Language Processing (NLP) –** Hands-on experience with text analytics, sentiment analysis, and topic modeling**.**

** Software Development & Version Control –** Familiarity with Git, Docker, and Agile methodologies for efficient collaboration.

** Problem-Solving & Critical Thinking** – Ability to translate business problems into data-driven solutions with a strong analytical approach.

**EDUCATION**

**University of New Haven ▪ West Haven, CT [**May] [2025] (expected 2025)

*Master of Science in Data Science*

* **Coursework:**  
  Finishing demanding courses in machine learning, deep learning, data mining, and artificial intelligence, along with practical experience developing predictive models and putting sophisticated algorithms into practice.
* Learned how to use statistical analysis, probability, and hypothesis testing to make data-driven decisions by applying these ideas to actual datasets.
* Gained expertise with cloud computing apps on AWS and Google Cloud, as well as big data technologies like Hadoop and Spark.
* Improved knowledge of SQL and database administration, working with both structured and unstructured data to facilitate effective processing and querying. a deeper understanding of computer vision and natural language processing (NLP), with the ability to apply deep learning methods to text and picture software.
* **Tools:**  
  ▪ Python ▪ R ▪ SQL ▪ TensorFlow ▪ Java ▪ Scala ▪ Bash/Shell scripting ▪ Power BI ▪ Excel ▪ Microsoft Azure ▪ Databricks ▪Git & GitHub/GitLab

**Mallareddy Engineering College For Women ▪ Hyderabad, Telangana** [Jun 2018 – May 2022]

*Bachelors in Computer Science*

* **Coursework:**
* Completed comprehensive coursework in Data Structures and Algorithms, Object-Oriented Programming, and Database Management Systems, gaining a strong foundation in problem-solving and software development.
* Acquired proficiency in Operating Systems, Computer Networks, and Distributed Systems, understanding system architecture, networking protocols, and cloud computing concepts.
* Developed hands-on experience in Software Engineering and Agile Development, working on real-world projects using industry best practices.
* Strengthened programming skills through courses in Python, Java, C++, and SQL, implementing efficient solutions for various computational problems.
* Additionally, explored Cybersecurity and Cryptography, learning about encryption, authentication, and secure coding principles. The coursework provided a well-rounded understanding of computer science principles, enabling the ability to design and develop scalable, efficient, and secure software solutions.

**DATA & AI PROJECTS AND PORTOFOLIO**

**University of New Haven ▪ West Haven, CT**  [May] [2025]–Present

*Master of Science in Data Science*

* [Summarize key project descriptions and your role in bullet points; include relevant KPIs to quantify your accomplishments, such as accuracy, speed, or efficiency improvements.]
* [Ensure that each project summary is balanced relative to the others (i.e., similar length)]
* **Project 1** – **Sentiment Analysis and Key Entity Detection Approach for Online Financial Texts**
* Created and put into practice a Sentiment Analysis and Key Entity Detection Approach for online financial texts, using Natural Language Processing (NLP) methods to glean insightful information from social media, financial news, and reports.
* Classified attitudes (positive, negative, and neutral) and identified important financial entities such as businesses, stock symbols, and economic indicators using machine learning models (Logistic Regression, SVM, and LSTMs) and deep learning frameworks (BERT and transformers).
* NLTK and spaCy were used in Named Entity Recognition (NER) methodologies to improve the accuracy of entity detection. enhanced text representation and contextual comprehension by the integration of sentiment lexicons, word embeddings, and TF-IDF.
* Deployed models on cloud platforms like AWS and Google Cloud for scalability after processing massive amounts of financial data with big data technologies like Hadoop and Spark.

**Skills:**

* **Entity Detection:** Proficiency in Named Entity Recognition (NER) utilizing libraries such as spaCy, NLTK, and transformers (BERT, RoBERTa) to extract important financial entities (e.g., company names, stock symbols, financial indicators).
* **Text Preprocessing & Feature Engineering:** Proficient in TF-IDF/word embeddings (Word2Vec, GloVe, FastText), tokenization, stemming, lemmatization, and stopword removal for financial text analysis.
* **Experience using supervised and unsupervised learning methods** : (such as Random Forest, SVM, LSTMs, and Transformers) for entity recognition and sentiment classification is a prerequisite for machine learning and deep learning.
* **Financial Data Processing:** Practical expertise with API-based financial text extraction from news articles, reports, and social media, as well as online scraping (BeautifulSoup, Scrapy).
* **Big Data and Cloud Computing:** Experienced with leveraging Google Cloud NLP APIs for scalable sentiment and entities, Apache Spark (PySpark), and AWS (S3, Lambda) to handle large-scale financial text datasets

**.**

* **Project 2**  – **Object Detection with PyTorch: Dataset, Methodology, and Evaluation**
* Using PyTorch, an object identification model was created and put into use to accurately recognize and categorize things in photos.
* Made use of the COCO dataset, a sizable collection of labeled photos for keypoint, segmentation, and object identification. To improve model generalization, the dataset was preprocessed utilizing picture augmentation methods as scaling, flipping, and normalization.
* YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and Faster R-CNN architectures were used to construct a deep learning-based object identification model. improved performance on the chosen dataset by utilizing Torchvision's pretrained models and refining them through transfer learning.
* Batch normalization, learning rate scheduling, and GPU acceleration were used to optimize model training for effective convergence.
* The model's performance was assessed using common metrics for object detection, such as Precision-Recall curves, Mean Average Precision (mAP), and Intersection over Union (IoU).
* To increase the accuracy of detection, significant hyperparameter tweaking and data augmentation were carried out. To properly comprehend findings, Matplotlib and OpenCV were used to visualize model predictions using bounding boxes and confidence scores.
* The project showcased the capacity to work with huge datasets, optimize model architectures, and assess performance for practical object identification applications. It also showed significant knowledge in deep learning, computer vision, and PyTorch.
* **Skills:**
* **Programming & Frameworks:**
* **Python** – Primary language for deep learning model development
* **PyTorch** – Deep learning framework for building and training object detection models
* **Torchvision** – Pre-trained models, transforms, and datasets for object detection tasks
* **OpenCV** – Image preprocessing, annotation, and augmentation
* **Object Detection Techniques:**
* Region-Based Convolutional Neural Networks (R-CNN, Fast R-CNN, Faster R-CNN, Mask R-CNN)
* Single Shot MultiBox Detector (SSD)
* You Only Look Once (YOLO)
* Feature Pyramid Networks (FPN)
* Transformers for Object Detection (DETR)
* **Dataset Handling & Preprocessing:**
* **COCO (Common Objects in Context)** & **Pascal VOC** – Standard datasets for training and evaluation
* **Custom Dataset Creation** – Collecting, labeling, and preparing images for training
* **Data Augmentation** – Using Albumentations, torchvision transforms, and OpenCV to enhance dataset diversity
* **Bounding Box Annotation** – Labeling tools like LabelImg and Roboflow
* **Model Training & Optimization:**
* **Transfer Learning** – Fine-tuning pre-trained models (ResNet, EfficientNet, MobileNet)
* **Backpropagation & Optimization Algorithms** – SGD, Adam, and learning rate scheduling
* **Batch Normalization & Dropout** – Improving model generalization and performance
* **Gradient Clipping** – Preventing exploding gradients in deep networks
* **Evaluation Metrics & Model Validation:**
* **Mean Average Precision (mAP)** – Primary metric for object detection performance
* **Intersection over Union (IoU)** – Measuring overlap between predicted and ground truth bounding boxes
* **Precision, Recall, F1-Score** – Assessing classification performance of detected objects
* **Confusion Matrix & ROC Curves** – Evaluating object detection accuracy
* **Deployment & Real-Time Inference:**
* **ONNX & TorchScript** – Converting PyTorch models for optimized inference
* **Flask & FastAPI** – Building REST APIs for deploying object detection models
* **Edge Deployment** – Running models on Raspberry Pi, NVIDIA Jetson, and TensorRT for real-time detection
* **Streamlit & Gradio** – Creating interactive web apps for showcasing object detection results
* **Capstone Project: [Title]** – Brief description of project, including methods and tools used (Python, R, SQL, etc.).

**[PROJECT TITLE] ▪ [City], [State]** [Month] [Year]– [Month] [Year]

*[Your Role], [School / Program]*

* [Summarize project key accomplishments in bullet points; include any relevant KPIs to quantify your accomplishments]
* [Ensure each of the project summary you include on your resume is balanced relative to the others (i.e., same length)]

**[PROJECT TITLE] ▪ [City], [State]** [Month] [Year]– [Month] [Year]

*[Your Role], [School / Program]*

* [Summarize project key accomplishments in bullet points; include any relevant KPIs to quantify your accomplishments]
* [Ensure each of the project summary you include on your resume is balanced relative to the others (i.e., same length)]

**PROFESSIONAL Portfolio ▪ City, United States**

* **Professional Portfolio** | [City, Country]  
  Data & AI Projects | [Group / Division]  
  ▪ Access my professional portfolio: [Add link here].

**{Example Professional Portfolio}**

**Website Link:**[**https://kshreya2k.github.io/My\_Portfolio/**](https://kshreya2k.github.io/My_Portfolio/)

**Github Link:**[**https://github.com/kshreya2k/Portfolio**](https://github.com/kshreya2k/Portfolio)

**Website Link** : <https://bronyayang.github.io/personal_website/>

**PUBLICATIONS**

* **Citation for the paper in APA format:** ▪ Author(s). (Year). *Title of the paper*. Publisher.
* **If you have specific author names, include them at the beginning:**  
  ▪ For example: Smith, J., & Doe, A. (Year). *Title of the paper*. Publisher.
* **Ensure to include the exact author names, year, title, and publisher for a complete and accurate citation.**

{Here’s the correct citation for the paper in APA format: For Example:

* Smith, J., & Doe, A. (2019). *Data Driven Decision Making in Academic Processes*. IEEE.
* Make sure to include the exact author names for a complete citation.}
* Smith, J., & Doe, A. (2019). *Data Driven Decision Making in Academic Processes*. IEEE.
* Smith, J., & Doe, A. (2019). *Data Driven Decision Making in Academic Processes*. IEEE.

**PROFESSIONAL CERTIFICATIONS & MEMBERSHIPS**

**[****Professional Certificate 1]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]

**[Professional Certificate 2]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]

**[Professional Certificate 3]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]

**[Professional Certificate 4]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]

**LANGUAGES**

**[Language 1]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]

**[Language 2]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]

**[Language 3]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]

**[Language 4]**  *[Certifying institution / professional association]* [Month] [Year]–[Month] [Year]