KRITIK SETH

New York, NY | (551)-344-6726 | kritikseth@nyu.edu | linkedin.com/in/kritikseth | github.com/kritikseth | kritikseth.com

EDUCATION

New York University, Center for Data Science   **New York, USA**
Master of Science in Data Science **(GPA: 3.62/4.00)**   **May 2024**
- Relevant Courses: Machine Learning, Big Data, Computational Cognitive Modelling (Reinforcement Learning), Probability & Statistics, Natural Language Processing, Optimization and Computational Linear Algebra.

NMIMS University, MPSTME   **Mumbai, India**
Bachelor of Technology in Data Science **(GPA: 3.88/4.00)**   **May 2022**
- Relevant Courses: Programming in C++, Machine Learning, DL, CV, NLP, Financial Engineering & Risk Management, Financial Institutions & Markets, Business Visualization, Cloud Computing, Statistical Methods, AI, Applied Mathematics.

TECHNICAL SKILLS

- Programming Languages & Frameworks: Python, SQL, R, C, C++, MATLAB, PL-SQL, Hadoop (Map-Reduce)
- Tools: Git, Tableau, PowerBI, Snowflake, Airflow, Spark, PyTorch, TensorFlow, scikit-learn, NumPy, Pandas, Langchain, AWS.

RELEVANT EXPERIENCE

Logitix – Data Scientist (Florida, US)   **June 2023 – Present**
- Trained a machine learning model to predict ticket tiers with 94% accuracy, securing lucrative partnerships with multiple prestigious sports venues and directly *generating $100K in revenue* through ticket sales.
- Streamlined categorization by performing unsupervised machine learning on ticket sales data, implementing BIRCH and K-Means clustering algorithms to create 5 tiers hence enabling efficient analysis.
- Collaborated with the strategy analyst team to enhance clustering algorithms, boosting model accuracy and reliability, and developed a business solutions dashboard to convey technical insights to non-technical stakeholders through data storytelling.

Persistent Systems – Data Science Intern (Mumbai, IN)   **Jan – April 2022**
- Accelerated manual classification of cells in histopathological images, resulting in *80% increase in efficiency*, by building Image Segmentation Models to detect and count different types of cells.
- Enhanced accuracy by 15% and expedited preprocessing with *40% increase in speed to 3 seconds* by streamlining pipeline to incorporate Deep Learning model for keyword extraction on text, post speech-to-text conversion.

AkzoNobel – Data Science Intern (Mumbai, IN)   **Sept 2021 – Mar 2022**
- *Improved accuracy of model by 20%*, as measured by its ability to classify colors based on reflection values, by implementing ensemble of Random Forest and Light Gradient Boosting Models using Scikit-Learn.
- Simplified color recipe-generating process by building Machine Learning models to generate color recipes using solid colors.

Kenmark ITAN – Junior Data Science Associate (Mumbai, IN)   **April – July 2020**
- Led development of text-cleaning pipeline, *reducing processing time by 40% to 7 seconds* and expediting integration of data.
- Implemented a baseline recommendation system using sentiment analysis for a client's social media application, leading to an *increase in user retention time by 3 minutes* as validated through *A/B testing*.
- Conducted and facilitated knowledge transfer by hosting a *tutoring session for 11 full-time staff members*.

Sapio Analytics – Data Analyst Intern (Mumbai, IN)   **April – June 2020**
- *Maximized supply chain efficiency* of COVID-19 vaccine deliveries by spearheading the development of a collaborative dashboard (Tableau & Dash), leveraging AWS to extract key metrics. Presented it to Andhra Pradesh government leaders.
- Analyzed historical data and market trends to predict need of essential supplies at hyper-granular level in India (ad hoc queries).
- Managed SQL database (over 40 tables with 100,000 rows) for COVID-19 Project, integrated by mobile and web applications.

SELECTED PROJECTS

Backtesting Financial Analysts' Future Predictions (Open AI, LangChain, OpenCV, AI)   **May – June 2023**
- Utilized LLMs and Langchain for backtesting, extracting key information and timelines for outcome detection.

Moving Target Interception - Multi-Agent Reinforcement Learning (MARL) (Python, Numpy, OpenCV)   **Mar – May 2023**
- Engineered an innovative MARL framework, training agents to make co-ordinated decisions to capture an evasive thief.

Music Recommendation System (Spark, Dask, Python, Hadoop)   **Mar – April 2023**
- Developed *collaborative filtering* based music recommendation system on large-scale interactions data (50GB+), achieving 3 fold improvement in mean average precision over baseline.

Analyzing Optimal Video Game Playing Conditions (TensorFlow, sklearn, Scipy, statsmodels, LightGBM)   **Nov – Dec 2022**
- Collaborated with a cross-functional team to execute a Kolmogorov-Smirnov *statistical test*, validating Moore's Law.
- Trained a neural network model with *2x improvement* in predicting FPS compared to traditional ML approaches.

Multi-Disease Detection using Retinal Fundus Images (PyTorch, TensorFlow, CNN, OpenCV)   **Aug – Oct 2021**
- Achieved 0.93 weighted F1 Score in identifying 45 diseases through training an ensemble Convolutional Neural Network.

Swachhdata - *50,000 downloads* (Regex, Git, PyPI, NLTK, OpenCV, Gensim, NumPy, and Pandas)   **May – Aug 2021**
- Programmed *3,000+ lines* to develop Swachhdata library, delivering modular preprocessing tools for data, text, and images.

Wherebnb (Python, Flask, TensorFlow, Scikit-Learn, HTML-CSS, and Tableau, Cloud)   **Aug – Oct 2020**
- Built an Airbnb clone leveraging Deep Learning for precise price and popularity predictions of real listings.
- Implemented state-of-the-art *text-generation* RNNs to analyze listings and provide hosts with tailored title recommendations.