

Deep Khut

+1 (240) 926-7303 | deepk2@umbc.edu
github.com/Deepkhut | [linkedin.com/in/deepkhut](https://www.linkedin.com/in/deepkhut)

Education

University of Maryland-Baltimore County, MD, USA

August 2023 – May 2025*

Master of Professional Studies in Data Science

3.78/4.0

Relevant Coursework: Data Analysis and Machine Learning, Big Data Processing, Data Management, NLP, Deep Learning
Nirma University, India

July 2019 – May 2023

Bachelor of Technology in Instrumentation and Control Technology

8.2/10

Skills

Languages: Python, JavaScript, C, C++, HTML/CSS, Matlab

Databases: MySQL, PostgreSQL, MongoDB

Libraries: NumPy, Pandas, OpenCV, Matplotlib, Seaborn, Nltk, Sickit-learn

Frameworks: Keras, TensorFlow, PyTorch, Bootstrap, Apache Spark

Tools & Technologies: Git, AWS, GCP

Experience

Jekson Vision Pvt. Ltd. | Deep Learning Intern

January 2023 – May 2023

- Implemented data augmentation techniques to enhance the efficiency of medical capsule and tablet inspection systems by 70%.
- Engineered a semi-automatic data annotation tool using YOLO, ResNet, and U-Net; reduced image annotation time by 40%, allowing the data science team to process 5,000 images weekly with enhanced accuracy.
- Constructed a data flow pipeline that automated the annotation process for segmentation and detection models; eliminated redundant steps, saving an estimated 15 hours per week in manual labor across the data team.
- Added a bulk upload feature which reduced the manual work of adding products into a database.

Tata Chemicals Ltd. | Automation Intern

May 2022 – July 2022

- Utilized Python to collect and analyze large production data sets, identifying key trends and patterns that led to a 50% reduction in production errors.
- Created a real-time interactive dashboard in Tableau that visualized production KPIs, enhancing data accessibility for 20+ team members and enabling data-driven decision-making that improved production efficiency by 15%.
- Spearheaded the integration of MongoDB Atlas, enabling live data access for a centralized dashboard; this innovation resulted in a 50% reduction in data retrieval time for critical reporting tasks.
- Compiled and presented detailed analytical reports to senior management, identifying 3 major inefficiencies; recommendations implemented resulted in a 15% boost in productivity across the customer service division.

Projects

Employee Attrition Prediction

January 2024 – April 2024

Tech Stack: Python, scikit-learn, Pandas, NumPy, Matplotlib

- Trained a machine learning classification model that achieved 88% test accuracy by conducting comprehensive Exploratory Data Analysis (EDA) and feature engineering using the scikit-learn library.
- Performed hyper-parameter tuning with cross-validation (up to 5 folds), improving model performance to 92%.

Exploratory Data Analysis (EDA) on Diabetic Patients Data

September 2023 – November 2023

Tech Stack: Python, Pandas, NumPy, Seaborn, Matplotlib

- Conducted EDA on a dataset of 70,000 diabetic patient records to clean, analyze, and visualize data.
- Mapped relationships between age and metabolic health metrics through a rigorous analysis of 60,000+ records; findings have been adopted in ongoing research projects aimed at improving diabetes management strategies.
- Analyzed patient demographics, blood-work, and lifestyle variables to evaluate risk factors, providing insights into variances in diabetes prevalence among different populations.

Traffic Sign Recognition for Autonomous Vehicles using Deep Learning

January 2021 – January 2022

Tech Stack: Python, TensorFlow, Keras, OpenCV

- Implemented a Convolutional Neural Network (CNN) architecture incorporating convolutional, pooling, dropout, flattening, and dense layers for traffic sign recognition, achieving 97% test accuracy.
- Leveraged batch normalization techniques to boost accuracy up to 99
- Explored and applied pre-trained architectures such as AlexNet, LeNet, and DenseNet to enhance model performance.

Publication

A Novel Approach to Tackle and Predict Absentism of Students using Deep Learning and Data Analysis