MANIDEEP ANNARAPU

+1 (940) 220-0208 ⬦ manideepannarapu@my.unt.edu ⬦ [Linkedin](#) ⬦ [Portfolio Website](#)

EDUCATION

Master of Science (M.S.), Computer Science (GPA: 4.0)   **Jan 2023 - May 2024**
University of North Texas, Denton, TX
Coursework: Big Data and Data Science, Machine Learning, Fundamentals of Database, Analysis of Computer Algorithms, Software Engineering, Information Retrieval and Web Search, Natural Language Processing, Data Analytics

Bachelor of Technology (B.Tech), Computer Science   **Jun 2018 - Aug 2022**
St. Martin's Engineering College, Hyderabad, India

TECHNICAL SKILLS

Languages and Libraries: Python, SQL, C++, Java, HTML, CSS, Pandas, Matplotlib, Scikit-Learn, TensorFlow, Keras, PyTorch, NLTK, NumPy, Seaborn
Data Engineering: Hadoop, Hive, Apache Spark, PySpark, HDFS, ETL Pipeline, Informatica, Data Manipulation
Tools: Tableau, Power BI, Microsoft Excel, SQL Server, Linux, Git, Cloud Computing (AWS), CI/CD
Other Skills: Prompt Engineering, Vector Databases, Computer Vision, Data Preprocessing, Model Evaluation

EXPERIENCE

Teaching Assistant, University of North Texas, Denton, TX   **Aug 2023 - May 2024**
- Assisted in the instruction of CSCE-4110 Data Structures and Algorithms, helping over 100 students and building strong relationships with students and faculty to ensure a collaborative learning environment.
- Took on responsibilities such as grading, providing feedback, and designing assignments to enhance student learning.

Summer AI/CS Research, University of North Texas, Denton, TX   **May 2023 - Jul 2023**
- Implemented dimensionality reduction techniques, reducing data dimensions by 50% while maintaining 95% of the original information by generating neural receptive fields and visualizing them from the sensory data.
- Explored various computer vision techniques and image processing using Python libraries.

Data Scientist, Onlane Solutions Pvt Ltd, Hyderabad, India   **Jul 2022 - Dec 2022**
- Developed predictive models to analyze customer data and provide actionable insights for marketing strategies, resulting in a 20% increase in customer retention.
- Architected intricate ETL workflows to streamline data integration from multiple sources; leveraged Excel, SQL, and Power BI to deliver real-time project performance reports, enhancing decision-making and boosting project efficiency by 40%.
- Presented data-driven insights to non-technical stakeholders, aiding in better business decision-making.

SDET, Value Labs, Hyderabad, India   **Mar 2022 - Jun 2022**
- Automated testing processes using the TestNG framework and Selenium, reducing manual testing efforts by 70%. Contributed to the design and execution of test plans and test cases.

PROJECTS

BCG Data Science Job Simulation on Forage   **May 2024 - Jun 2024**
- Completed a customer churn analysis simulation, demonstrating advanced data analytics skills, identifying essential client data through requirements gathering and outlining a strategic approach.
- Conducted efficient data analysis and data transformation using Python. Employed data visualization techniques for trend interpretation.
- Optimized a random forest model, achieving an 85% accuracy rate in prediction.

Diabetic Retinopathy Detection   **May 2024**
- Developed an advanced image processing system to detect diabetic retinopathy from retinal images using deep learning techniques. Implemented convolutional neural networks (CNNs) with TensorFlow to classify the severity of diabetic retinopathy, achieving a classification accuracy of 93%.
- Conducted extensive data preprocessing and augmentation to handle imbalanced datasets and improve model performance. Used Matplotlib and other visualization tools to display detection results and highlight areas of concern in retinal images.

Hotel Booking Demand Forecasting and Analysis May 2024
- Developed multiple predictive models to analyze hotel booking behaviors and trends, employing regression and ANOVA techniques, leading to actionable insights on customer preferences and booking dynamics.
- Leveraged Excel for preliminary data analysis and hypothesis testing to validate assumptions and format data before deep statistical analysis, enhancing the robustness of pricing strategies and customer segmentation.
- Utilized Excel's advanced analytical tools to perform initial ANOVA tests. Improved inventory management by analyzing booking data and forecasting demand.
- Utilized multiple classification and regression models from Scikit-Learn and developed predictive models to forecast hotel bookings.

Prediction and Global Analysis of Mental Health — Python, CNN, TensorFlow, Jupyter Feb 2024
- Developed a comprehensive predictive model for analyzing mental health indicators using deep learning neural networks (CNNs), achieving an accuracy of 92% by analyzing global trends in mental health.
- Integrated deep learning methodologies, providing insights into factors influencing mental health on a global scale. Used strong Python coding skills to preprocess and clean large datasets, ensuring data precision, efficiency, uniformity, and dependability for model training, resulting in a 50% reduction in data processing time.
- Demonstrated familiarity with Machine Learning models, particularly TensorFlow.

Forecasting Customer Buying Products Using Big Data and Data Science Nov 2023
- Analyzed and presented customer buying habits to enhance customer experience by providing recommendations. Employed efficient ETL pipelines, storing datasets in Hadoop HDFS and using Hive for analysis, demonstrating distributed and large-scale data processing capabilities.
- Used PySpark for analytics and integrated Spark with HDFS, achieving a 25% improvement in query performance. Applied Random Forest for classification, resulting in a 15% increase in model accuracy.
- Explored financial markets to understand consumer behavior and applied data-driven strategies for improving customer engagement.

Content-Based Image Retrieval Jul 2023
- Designed and developed an ML system that takes an image as input and displays similar images, achieving a retrieval accuracy of 88%.
- Used the VGG16 model from the TensorFlow library to extract features from the input and dataset. Matched input and dataset images using cosine similarity and visualized output using Matplotlib, reducing retrieval time by 20%.

Web Application Development for Furniture Company Database Jul 2023
- Worked on the end-to-end development of a web application for a furniture company database using ASP.NET Core, enhancing the efficiency and user experience of data management.
- Utilized the ASP.NET framework to create an efficient and user-friendly platform that streamlined data management and enhanced customer experience. Used Visual Studio Code and SQL Server Management Studio to develop and manage the application. Implemented web services to enable seamless data communication.

Interactive Sales Dashboard for Istanbul Shopping Malls Jul 2023
- Developed interactive sales dashboards and visualizations using Tableau, providing comprehensive insights into revenue generated by different shopping malls in Istanbul. Utilized SQL for data analysis.
- Enabled stakeholders to make data-driven decisions, leading to a 20% increase in sales performance through targeted marketing strategies.

Med-Bot Feb 2023 - Mar 2023
- Built a chatbot using Django and Python that utilizes natural language processing with a pretrained Generative AI model BERT - Large Language Model (LLM), achieving a response accuracy of 90%. Deployed the application on GitHub, demonstrating strong proficiency in machine learning, web development, text classification, and version control systems.