Deepak Shrestha

409-454-1698 | dshrestha4@lamar.edu | linkedin.com/in/iamdeepak42 | github.com/iamdeepak42 | iamdeepak42.github.io

Education

Lamar University Jan. 2022 – Dec. 2023
Master of Science in Computer Science *Beaumont, TX*
Coursework: Computer Vision, Data Science and Big Data Analysis, Statistical Inference, Analaysis of Algorithms, Cloud Computing, Database Design, Software Engineering

Experience

Graduate Research Assistant Jan 2023 – Present
Lamar University partnered with Texas A&M Agrilife Research *Beaumont, TX*
- Collaborated on a computer vision project for automated rice plant counting and rice plot segmentation
- Developed a VGGNet-based CNN model, achieving 93% accuracy in rice plant identification
- Currently enhancing project usability in precision agriculture through improved plot segmentation methods

Software Engineer Jan 2020 – Dec 2021
Parsedom *Kathmandu, Nepal*
- Designed a data extraction framework using Python and NodeJS, extracting data from 200+ websites
- Analyzed and visualized data for 50+ clients using pandas and seaborn, deriving insights
- Engineered Flask-based RESTful APIs, enhancing data exchange efficiency and system's data handling capacity

Data Analyst June 2019 – Dec 2019
Deerwalk Inc. *Lexington, MA*
- Performed detailed analysis of Healthcare Effectiveness Data and Information Set (HEDIS) measures, generating actionable insights to influence healthcare decisions.
- Collaborated with clients on managing data intake, clarifying and meeting data requirements for over 20+ projects
- Leveraged SQL for data analysis and mining, resolving 100+ data-related issues to optimize processes

Projects

Automated Rice Seedling Counting | *Pytorch, CNN, openCV, scipy, matplotlib*
- Developed a deep learning model for precise rice seedling counting from drone imagery
- Designed the model to provide two types of outputs: locations of rice seedlings and their counting confidence scores
- Employed Euclidean and Cross-Entropy losses to optimize point regression and proposal classification
- The model achieved a 93% success rate in accurately identifying rice plants

Image-Based PM2.5 Pollution Analysis | *Matlab, Autoencoder, Image Processing*
- Developed a lightweight image-based method for PM2.5 concentration estimation
- Utilized a 3-layer autoencoder with saturation and super-pixel techniques for image feature extraction
- Validated the model on a 1,460-photo dataset, achieving superior results compared to state-of-the-art methods, with SoftMax and Gaussian regression for PM2.5 estimation

Disaster Response Pipeline | *Python, Natural Language Processing, Flask, scikit-learn, nltk, pandas, sqlite*
- Developed a web application utilizing NLP techniques to classify emergency messages into distinct categories
- Implemented an ETL pipeline, facilitating efficient data reading, cleaning, and storage in a SQLite database
- Built a machine learning pipeline using Random Forest to classify 36 distinct categories
- Deployed a Flask-based web app that utilizes a pre-trained model (stored as a pickle file) to generate and display predictive categories based on user-input text

Technical Skills

Languages: Python, SQL (Postgres), JavaScript, Matlab, R, C++, Java
Frameworks: Node.js, Flask, Scrapy, Apify, Sklearn, OpenCV, PyTorch
Developer Tools: Git, Google Cloud Platform, VS Code,PyCharm, PyCharm, Jupyter Notebook, Amazon Redshift
Libraries: pandas, NumPy, Matplotlib, Scipy, ggplot, seaborn, nltk

Certifications

Mathematics for Machine Learning and Data Science | *Coursera* July 2023
Data Scientist Nanodegree | *Udacity* April 2020