Puspha Raj Pandeya

puspand@siue.edu · https://www.puspharaj.com.np · (636) 378-8382 · St Peters, Missouri
LinkedIn · GitHub

PROFESSIONAL SUMMARY

Experienced Data Engineer and Machine Learning enthusiast with a strong background in Computer Engineering. Skilled in statistical analysis, machine learning model development, and data-driven solutions, with proficiency in Python and DBMS. Passionate about leveraging ML and data-driven insights to tackle scientific problems. Always eager to explore innovative solutions that drive meaningful impact.

EDUCATION

Masters in Computer Science   Ongoing
Southern Illinois University Edwardsville

Bachelors in Computer Engineering   Sep 2017 - Aug 2022, GPA 3.03
Nepal College of Information Technology

SKILLS

Software Development : Python · Pandas · SQL · MongoDB · Git · Docker · Flask · Fast API · HTML · CSS

Data : Data Engineering · Machine Learning · Computer Vision · NLP · ETL · ELK · Big Data · Airflow · Kafka

General Tools : Microsoft Word · Excel · PowerPoint · PowerBI · DBeaver · MySQL Workbench · AWS

Soft Skills : Time Management · Organizational Skills · Interpersonal Communication · Multitasking

PROFESSIONAL EXPERIENCE

Cedar Gate Technologies   **Jun 2023 - August 2024**
Data Engineer

- Analyzed US healthcare data to identify patient cost distributions and the legitimacy of their insurance claims.
- Cleaned, transformed, aggregated, and mapped raw patient data to a company-wide schema.
- Standardized fields like relationship codes, insurance plans, etc., and conducted member information matching across data sources, while utilizing statistical techniques to interpret and analyze member demography and member months data.

Dekods INC.   **Jan 2023 - Feb 2023**
Freelance Data Analytics Consultant

- Cleaned and processed 100,000 record files for a UNICEF initiative to upgrade various infrastructures in Nepal's rural communities.
- Exploratory data analysis was carried out using Pandas and Python, and data cleaning and transformation scripts were created.
- Created and implemented a PowerBI dashboard for the UNICEF project team to display important metrics.
- Used DAX and best practices in visualization to deliver insights into rural families.

Eydean INC.   **Nov 2021 - Jan 2023**
Associate Machine Learning Engineer and Python Developer

- Developed sales forecasting models using time series analysis techniques like ARIMA and SARIMA.
- Created a recommendation system using content and collaborative filtering methods using techniques like cosine similarities and Non-negative matrix factorization.

- Developed a chatbot for an e-commerce client using NLP and neural networks.
- Contributed to the development of a Logistic Management System aimed at improving product delivery by finding the shortest path and using resources efficiently. Dijkstra algorithm was used to find the shortest path, and a navigation system was built using open street map.
- Designed and implemented a data pipeline using the ELK stack and Python for creating a sales analysis dashboard in Kibana.
- Designed and implemented efficient ETL processes to integrate data from multiple databases into comprehensive data warehouses, enabling faster and more accurate data analysis and visualizations.
- Designed, built, and containerized scalable backend systems for multiple projects using Flask, FastAPI, and MongoDB, resulting in efficient and easily maintainable applications.
- Migrated data from SQL to NoSQL databases using SQL and Pandas for denormalization and cleaning.

Technocolabs Softwares   **Oct 2021 - Nov 2021**

Machine Learning Intern

- Worked on a project called 'Using-NLP-Techniques-to-Predict-Song-Skips-on-Spotify-based-on-Acoustic-Data'.
- Lead a team of 6 people on this project to ensure efficient completion of the project.
- Created a sequential model for skip prediction using LSTM and Light GBT algorithms.

CERTIFICATIONS

AWS Academy Graduate - AWS Academy Data Analytics
Amazon Web Services

Python for Everybody Specialization
Coursera

Capstone: Retrieving, Processing, and VisualizingData with Python
Coursera

Python Data Structures
Coursera

UsingDatabases with Python
Coursera

PROJECTS

Brain Tumor Segmentation

Sentiment Analysis for Business Reviews

Cardiovascular Disease Detection