MANUJ KUMAR JOSHI

ASSOCIATE DATA SCIENTIST

Associate Data Scientist, result oriented, dedicated and can work on own initiative and can deliver on time with a high level of integrity and flexibility in a collaborative environment utilizing my creativity and technical skills.
Want to secure a position as Data Scientist in a company where I can utilize my 1.10 yearsof experience in Data Science to make a positive contribution to the organization.

EXPERTISE

- Critical Thinking
- Leadership
- Research based approach
- Communication
- Problem Solving

LANGUAGE

- English
- Hindi

CERTIFICATES

- Microsoft Azure Al Fundamentals
- Microsoft Azure Fundamentals
- Designing and Implementing a Microsoft Azure Al Solution
- Designing and Implementing a Data Science Solution on Azure
- Databricks Certified ML Associate
- Databricks Certified Generative AI
- Data Science Professional from lnnomatics Research Labs, Hyderabad (IBM Certified)
- Data Science Specialization from E&ICT llT ROORKEE

CONTACT

9557287321

manujjoshi52@gmail.com

Haldwani, Uttrakhand

https://github.com/manujjoshi

https://www.linkedin.com/in/manujjoshi/

EDUCATION

2010 - 2014 — Uttar Pradesh Technical University
Bachelor of Technology in Computer Science (64%)

2010 — CBSE XII (65%)

2008 — CBSE X (68%)

WORK EXPERI ENCE

April 2022 -Present

Associate Data Scientist
Celebal Technologies

Worked on various technologies like Computer vision, NLP, Databricks, Machine Learning, Time series, API development, Microsoft Cognitive Services, MLFlow and MLOps, Docker, Kubernetes, Linux, API Testing and Cloud Computing.

Apr 2020-Mar 2022

Data Analyst
iTrac Technology

Worked on Numpy, Pandas, web scraping,
Data Analysis and Visualization and SQL for analyzing data.

Mar 2018-Mar 2020

Developer
iTrac Technology

Worked on writing scalable, optimized codes.

Apr 2016-Feb 2018

Software Trainer
Skill Bird

Worked as a Computer Science Faculty in NGO

SKILLS

- Programming: Python
- Operating System: Windows and Ubuntu
- Data Visualization: Tableau and Power BI
- Code Management: Git and GitHub
- Tools IDE: Jupiter and vs. code
- Cloud: Azure
- Data Science: Machine Learning, Computer Vision, Natural Language Processing, Classical Machine Learning, Clustering, Time series, Azure Cognitive Services, Flask/Fast API, SQL, SQLAlchemy, Pyspark, Databricks, Super Resolution, MLFlow, MLOps, CI/CD , DevOps, ADF, Docker, Kubernetes, API Testing, Postman, Prompt Engineering, Generative AI and LLMs

INTERNSHIPS/TRAINING

lnnomatics Research Labs:

Project: Plant Disease Detector

- Used CNN to build the model and tried three pre-trained models, VGG16, VGG19and lnceptionV3 in a team of 5 members and deployed it.

Project: URL Shortener

- Used Flask for the back end along with HTML and CSS for the front end and SQLAlchemy for the database purpose.
- Through that shortened URL the user will be redirected to the original URL and the same shortened URL can also be used in the future for re-direction.

Project: Web Scraping

- Scrapped an automobile website, and extracted bike and car data to do data analysis to suggest the vehicle to the customer as per his budget and specifications.
- Scrapped the website with the help of BeautifulSoup and presented the demonstration to the team.

The Sparks Foundation:

Project: Data Analysis and ML

- Worked on multiple datasets to do Descriptive Analysis, Diagnostic Analytics, Predictive analysis, and Prescriptive Analysis.
- Used the patterns to predict the response with the help of Regression, Classification and Clustering Algorithms.

Project: COVID-19 Mask Detector

- COVID-19 Mask Detector which detect the mask of the person in video and photo.

HOBBIES AND INTEREST

- Traveling
- Photography
- Writing
- Reading

PROJECTS

Celebal Technologies:

Project: Customer Analytics

- Developed models using NLP for Text Summarization, Sentiment Analysis and NER using BERT and Pytorch.
- Created Transcriptions for the Project on which I had to finetune all the 3 models according to business objectives.
- Created pipeline and hosted it with the help of Fast API.

Project: Short coming of Solar panels

- Used SOTA model to do the Image Segmentation of solar panels through Solar panel images.

Project: Solar Potential

- Detect the potential surface Area for the solar panel installation via satellite images.
- Used Real ESRGAN model to super resolute satellite images from 320px to1280px.
- Used MMDetection model to segment and detect roof tops and land area in satellite images.
- Created pipeline and hosted it with the help of Fast API.

Project: Q&A pipeline

- Created an Open AI pipeline which extracts data from files and answers questions from text present in file.

RESEARCH

- Early diagnosis of MS disease: The vision transformer model is proposed to classify between a healthy brain and an unhealthy brain.
- Condition Monitoring of Hydraulic rig: We will predict the conditionof a hydraulic rig, based on the sensor data provided by the 17 sensors. The main objective of this work is to determine the sensors and featureswhich are more effective in detecting a given type of fault.
- Telecom sector use case to suggest a good recharge plan to the customer.
- Thermal power loss detection on solar panels.
- Time series to forecast appropriate electricity generation.