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 years of 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 Implementinga Microsoft Azure Al Solution
- Designing and Implementinga Data Science Solution on Azure
- Data bricks Certified ML Associate
- Data Science Professional from lnnomatics ResearchLabs, Hyderabad (IBM Certified)
- Data Science Specializationfrom E&ICT lIT ROORKEE

CONTACT

9557287321

manujjoshi52@gmail.com Haldwani, Uttrakhand

https://github.com/manujjoshi

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

EDUCATION

Uttar Pradesh Technical University

2010 - 2014	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, MachineLearning, Time series, API development, Microsoft Cognitive Services,MLFlow and MLOps, Docker, Kubernetes, Linux, API Testing and Cloud Computing.

Oct 2021-Jan 2022

Data Science Intern

lnnomatics Research Labs

Worked on Machine Learning and Deep Learning Projects, developed API and tested the same.
Created an end-to-end Plant Disease Detector Internship Project in a team.

Mar 2021-Apr 2021

Data Science & IOT Intern

The Sparks Foundation

Worked as a Data Science and IOT Intern in which Idealt with 8 ML business use cases for which I developed solutions and also built IOT use case for Covid mask detector.

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: Computer Vision, Natural Language Processing, Classical MachineLearning, Clustering, Time series, Azure Cognitive Services, Flask/Fast API, SQL, SQLAlchemy, Pyspark, Databricks, SuperResolution, MLFlow, MLOps, CI/CD and ADF, Docker, Kubernetes, API Testing, Postman

INTERNSHIPS/TRAINING

Innomatics Research Labs:

Project: Plant Disease Detector

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

Project: URL Shortened

- Used Flask for the back end along withHTML and CSS for the front endand 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, andextracted bike and car data to do dataanalysis to suggest the vehicle to the customer as per his budget and specifications.
- Scrapped the website with the helpBeautifulSoup and presentedthe demonstration to the team.

The Sparks Foundation:

Project: Data Analysis and Machine

- 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 onwhich Ihad to finetune all the 3 models according to business objectives.
- Created pipeline and hosted it with the help of Fast API.

Project: Faulty solar panel detection

- Used Detectron2 model to do the Image Segmentation of Faulty solar panels through Solarpanel images.
- Tagged the Solar panel images to train the Detectron2 model for faulty solar panel detection.

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 anddetect rooftops and landarea in satellite images.
- Created pipeline and hosted it with thehelp ofFast API.

Project: Intermittent time series forecasting

- Trained an intermittent time series models and created a CI/CD pipeline and deployed on AKS and also performed Data Drift Monitoring.
- Created a CI/CD pipeline on Azure using Docker and Kubernetes for this usecase.

Project: Open AI data pipeline

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

RESEARCH

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