DEEPAK PATHAK

AI Researcher & Data Scientist

deepakpathak0892@gmail.com deepakpathak0892 Kaiserslautern, Germany
dpkpathak.github.io Deepak Pathak

EXPERIENCE

AI Researcher
DFKI GmbH

Oct 2021 – Ongoing Kaiserslautern, Germany

- Developed crop yield prediction models by integrating multi-modal data (e.g., weather, soil properties, digital elevation maps).
- Led the design and implementation of preprocessing pipelines for raw combine harvester data.
- Built an interactive dashboard for visualizing and exploring crop field data using Dash, enabling stakeholders to gain actionable insights.
- Established a containerized MLflow server for large-scale experiment tracking, improving reproducibility and collaboration across teams.
- Designed machine learning models for predicting satellite collision risk and conducted feature importance analysis using SHAP.
- Containerized machine learning models for deployment to research partners.

Machine Learning Research Intern
Miele & Cie. KG

Dec 2020 – Aug 2021 Gütersloh, Germany

- Explored deep metric learning and contrastive learning methods for fine-grained image classification.
- Developed models based on state-of-the-art unsupervised and self-supervised deep learning techniques to enhance supervised model performance.
- Investigated methods for storing lower-dimensional representations of images using generative models.
- Utilized PyTorch and PyTorch Lightning for model development and Azure Databricks clusters for distributed training.

System Engineer - Application Developer
IBM

Aug 2015 — Mar 2019 Bangalore, India

- Implemented business processes using Oracle BPM Suite within a Service-Oriented Architecture (SOA) framework.
- Developed Java EE modules for distributed applications in the telecom sector.
- Created an automation solution using Spring Boot and Twilio API to flag high-priority production incidents.

Student Research Assistant
virtUOS, Universität Osnabrück

June 2020 — Dec 2020 Osnabrück, Germany

- Contributed to the development of SIDDATA, a digital study assistant, by managing the backend and integrating deep neural network-based recommender systems.

ABOUT

I am passionate about applying machine learning to real-world challenges, with a strong interest in developing software and designing robust backend systems. I enjoy creating practical solutions that bridge data science and technology.

SKILLS

Python | Macine Learning & Deep Learning | Crop Modelling | Time series | Multi-modal Learning | Pytorch | Data Visualization | Docker | Kubernetes

Hard-working | Analytical | Critical Thinking | Persistent

RECOGNITIONS

Eminence & Excellence "Spark" Award
For excellent contribution to telecom project at IBM

Manager's Choice Award
For development and automation activities at IBM

LANGUAGES

English
German

EDUCATION

M.Sc. in Cognitive Science
Universität Osnabrück

Apr 2019 – Aug 2021 Osnabrück, Germany

Bachelor of Technology (B.Tech) in Electronics Engineering
HBTI (Harcourt Butler Technological Institute)

Aug 2011 - June 2015 Kanpur, India

Working Student - Software Developer

Aitech Concept UG

Oct 2019 – Nov 2020 Wallenhorst, Germany

- Implemented object detection models using TensorFlow for tracking orders in restaurant settings.
- Developed deployable applications in Python (Django) for real-time object detection using surveillance camera feeds.

PUBLICATIONS

Journal Articles

- F. Mena, **D. Pathak**, H. Najjar, *et al.*, "Adaptive fusion of multi-modal remote sensing data for optimal sub-field crop yield prediction," *Remote Sensing of Environment*, vol. 318, p. 114 547, 2025, ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2024.114547.

Conference Proceedings

- M. Miranda, **D. Pathak**, M. Nuske, and A. Dengel, "Multi-modal fusion methods with local neighborhood information for crop yield prediction at field and subfield levels," in *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, 2024, pp. 4307–4311. DOI: 10.1109/IGARSS53475.2024.10640993.
- **D. Pathak**, M. Miranda, F. Mena, *et al.*, "Predicting crop yield with machine learning: An extensive analysis of input modalities and models on a field and sub-field level," in *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 2023, pp. 2767–2770. DOI: 10.1109/IGARSS52108.2023.10282318.
- C. Sanchez, **D. Pathak**, M. Miranda, *et al.*, "Influence of data cleaning techniques on sub-field yield predictions," in *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 2023, pp. 4852–4855. DOI: 10.1109/IGARSS52108.2023.10282955.

CERTIFICATION

Oracle

- Oracle Certified Associate, Java SE 7 Programmer, 2017
- Oracle PL/SQL Developer Certified Associate, 2017

IBM

- Data Science Foundations - Level 1
- Data Science Foundations - Level 2 (V2)
- Python for Data Science
- IBM Cloud Essentials
- IBM Agile Explorer

HackerRank

- Problem Solving (Basic) Certificate
- Python (Intermediate) Certificate