

Akshat Dubey

Mobile: +49-1627917434
Email: dubeyakshat07@gmail.com
linkedin.com/in/akshat0007
github.com/dubeyakshat07

EDUCATION

PhD in Computer Science (Funded by the Robert Koch Institute) <i>Freie Universität Berlin</i>	August 2023 - Current <i>Berlin, Germany</i>
Integrated M.Sc. in Mathematics and Computing <i>Birla Institute of Technology - Mesra</i>	July 2018 - May 2023 <i>Ranchi, India</i>

EXPERIENCE

Teaching Assistant <i>Freie Universität</i>	April 2024 - Current <i>Berlin, Germany</i>
<ul style="list-style-type: none">• Winter Semester 2024/25 : Assisting Dr. habil. Georges Hattab with seminar-based lectures on Visualization for AI Explainability. The students have to present a research paper on XAI & it's application and present the insights on the methodology including results.• Summer Semester 2024: Assisted Dr. habil. Georges Hattab with Advanced Data Visualization for Artificial Intelligence (AI), Natural Language Processing (NLP), and Explainable Artificial Intelligence (XAI) course.• Providing hands-on training and experience to enrolled students.	
Research Associate <i>Robert Koch Institute: Center for Artificial Intelligence (Federal Ministry of Health)</i>	August 2023 - Current <i>Berlin, Germany</i>
<ul style="list-style-type: none">• Conducted advanced research in domain-centric Explainable Artificial Intelligence (XAI), enabling the adoption of AI in healthcare through trustworthy, regulatory-compliant solutions.• Developed a storytelling XAI framework utilizing knowledge distillation to exploit relationships between multimodal data (text and images), achieving enhanced domain-centric explainability.• Engineered Human-AI-Interfaces leveraging information visualization to improve transparency, trust, and decision-making in healthcare applications.• Conceptualized the "Nested Model for AI Design and Validation", facilitating the integration of regulatory-compliant workflows in healthcare AI systems.• Collaborated with multidisciplinary teams to drive innovative solutions for public health challenges and published high-impact research on XAI frameworks.	
Deep Learning Engineer <i>QpiVolta Technologies</i>	January 2023 - June 2023 <i>Bangalore, India</i>
<ul style="list-style-type: none">• Developed a Docker-based execution environment for machine learning workflows.• Successfully deployed Flyte using Rancher and Helm Chart, and integrated Flyte, a Python-based Machine Learning Workflow framework, which automatically manages Kubernetes clusters, into our workflow management system. Streamlined and automated various machine learning tasks, improving overall productivity.• Developed deep learning models for carbon capture using material structure information. Enhanced carbon capture efficiency predictions, contributing to sustainable energy solutions.• Trained MOFTransformer for predicting physical properties associated with metal-organic frameworks (MOFs). Achieved high accuracy in predicting properties such as CO₂ uptake and CO₂/N₂ selectivity.	
AI Scientist Intern <i>QpiVolta Technologies</i>	July 2022 – December 2022 <i>Bangalore, India</i>
<ul style="list-style-type: none">• Trained BERT (Bidirectional Encoder Representations from Transformers) for predicting reaction yields. Improved reaction yield predictions, aiding in reaction optimization and synthesis planning.• Successfully trained BART (Bidirectional and Auto-Regressive Transformer) on chemical smiles to predict various sequence-based tasks, especially product prediction from given reactants.• Enabled robust atom mapping on valid reaction SMILES (Simplified Molecular Input Line Entry System), using Attention, the atom-mapping information was learned by an ALBERT model trained in an unsupervised fashion on a large dataset of chemical reactions.	
Research Assistant <i>Ryerson University</i>	September 2021 – July 2022 <i>Toronto, Canada</i>

- Worked under Dr. Pawel Pralat (**Ryerson University, Canada**) and Dr. Bogumił Kamiński (**SGH Warsaw School of Economics, Poland**)
- Created **text embeddings** using different embedding techniques such as Bag of Words, TF-IDF, Word2Vec, and BERT embedding.
- Created graph embeddings using different embedding techniques such as **Node2Vec, Struc2Vec, and RolX**.
- Analyzed the correlation between graph embeddings and word embeddings.
- Performed in-depth analysis of different machine learning and deep learning models for detecting bots in a graph network using **graph embeddings, node features, and text features**.
- This project concluded with the acceptance of a research paper titled "**Detecting Bots in Social-Networks Using Node and Structural Embeddings**" at the **International Conference on Data Science, Technology and Applications - 2022, Lisbon, Portugal**.

Data Science Intern

March 2021 – September 2021

Labellerr, Automated AI and Data Annotation SAAS platform

Chandigarh, India

- Worked with Noom and Perceptly.
- My work primarily at Labellerr involved working with image data and building deep learning models for different tasks such as image classification, image segmentation, object detection, and masking.
- Performed the analysis and annotation of the image data.
- Implemented **transfer learning** by fine-tuning different architectures such as **ResNet, InceptionV3, EfficientNet** for image classification tasks.
- Implemented **U-Net** CNN-based deep learning architecture from scratch and trained on a variety of biomedical image data for the task of segmentation.
- Trained **Vision Transformer** on custom image data (Stock Keeping Units) provided by one of our clients for the task of object detection.
- Trained **Google AutoML** Object Detection API and deployed it.
- Deployed the deep learning models on **Google Cloud Platform (GCP)**

PUBLICATIONS

AI Readiness in Healthcare Domain through Storytelling XAI

First Workshop on Explainable Artificial Intelligence in Medical Domain at ECAI-2024

July 2024

A Nested Model for AI Design and Validation

Cell Press iScience Journal

October 2024

Pri-GP: Prior-Aware Distributed Gaussian Process Regression

The 23rd International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)

May 2024

Detecting bots in social-networks using node and structural embeddings

Journal of Big Data

July 2023

Supervised Contrastive Learning Approach for Facial Expression Recognition

Journal of Multimedia Tools and Applications

March 2023

MASTER'S THESIS

Supervised Contrastive Learning Approach for Segmentation of Gastro-Intestinal Tract MRI Images Supported by Explainable AI

- The University of Wisconsin-Madison GI Tract MRI Image Segmentation Dataset was used for this project.
- **Stratified group fold data sampling** method was used during the preparation of the data.
- Employed data augmentation to combat imbalanced data problem.
- Justified the use of **supervised contrastive training method** for the training of U-Net architecture.
- In the case of GI tract MRI segmentation, the contrastive training helped the **U-Net model to learn to differentiate between similar regions of the GI tract that may have subtle differences**, such as different parts of the colon or small intestine.
- This project resulted in more **accurate and robust segmentation** of the GI tract, which in turn improved the accuracy of medical diagnoses and treatments.

- The problem of imbalanced data was also solved by contrastive learning.
- The DICE Loss was used for the evaluation. The DICE Loss score was 0.78
- Used **GRAD-Cam** for the **interpretation and explainability** which are critical in healthcare, as they help to ensure that healthcare providers can trust the decision-making process of the AI system and make **informed decisions**.

TECHNICAL SKILLS

Languages: C, Python, SQL (MySQL)

Skills: Data Analysis, Data Visualization, Machine Learning, Deep Learning, Natural Language Processing, Time Series Analysis, Statistics, Dashboards

Data Analysis Tools: Python, Microsoft Excel

Machine Learning Libraries: sklearn, XGBoost, PyCaret

Deep Learning Libraries: Tensorflow, PyTorch, Hugging Face

Hands-on Experience: Docker, Kubernetes, FastAPI, Flyte (ML Orchestration Platform), Flask, Git, Heroku, Google Cloud Platform (GCP), Data Version Control (DVC), PySpark

ACHIEVEMENTS

Best Paper Award Certificate

11th International Conference on Data Science, Technology and Applications (DATA), Lisbon, Portugal *July 2022*

- This award certificate was awarded to me and my team at Ryerson University for our work on the research problem "Detecting Bots in Social-Networks Using Node and Structural Embeddings"

National Winner

EXL EQ 2022

April 2022

- National data science and analytics competition organized by EXL Analytics in India. Cash prize of 2000 USD

CO-CURRICULAR ACTIVITIES

President

Society for Data Science, BIT-Mesra

Ranchi, Jharkhand

Jul. 2021 – Present

Volunteer

National Service Scheme (Ministry of Youth Affairs and Sports), BIT-Mesra

Ranchi, Jharkhand

July 2018 – July 2020