



# Dr. Priyanka Chaudhary

priyanka.chaudhary18@gmail.com

Indian | B EU/EFTA Permit

<https://www.linkedin.com/in/pchaudha/>

Pfäffikon SZ, 8808 | Mobile: +41 78 241 1868



## DATA SCIENTIST

Researcher with 10+ years across Machine Learning, Computer Vision including a PhD from ETH Zürich. My core strength is building end to end vision systems from problem definition and dataset design to model development, evaluation, and clear communication of results to technical and non-technical partners. I have published on flood forecasting with transformers, uncertainty quantification using deep ensembles and on multi-task ranking with weak supervision approach. I bring strong engineering fundamentals in Python, SQL, and production minded ML systems (CI/CD, containerization, model versioning, drift monitoring).

## CORE SKILLS

**Languages:** Python, C/C++, SQL

**ML Methods:** Machine Learning, Natural Language Understanding, NLP, Transformers, CNN, Segmentation, Detection, Uncertainty, Ensembles

**MLOps / ML Systems:** Docker, Kubernetes, CI/CD, Training & evaluation pipelines, Deployment, Monitoring (drift, prediction churn), Model versioning

**Frameworks & Tools:** PyTorch, TensorFlow, Keras

**Engineering:** Git, Linux, Data Structures & Algorithms

## WORK EXPERIENCE

### Data Scientist

ETH Zürich, Switzerland

Apr 2024 – Sep 2025

- Co-developed a global biodiversity monitoring pipeline on Google Earth Engine, integrating 85 global map layers to compute the SEED index on a ~1 km grid (30 arc seconds).
- Implemented key pipeline components and features, scaling global runs across 846 ecoregions with typical end-to-end runtime of ~13 hours per full execution.
- Supported benchmarking and validation by linking pixels to minimally modified reference areas ( $HMI < 0.05$ ) to create consistent baselines across regions.

### Data Scientist Intern

Swiss Re Group, Switzerland

Sep 2021 – Feb 2022

- Owned delivery of an unsupervised claims triaging ML product for UK claims handlers to support payment decisions, iterating with stakeholders to map their workflow, define requirements and validate outputs against real decision patterns.
- Built, tested, documented and deployed a toolkit of ML utilities on the internal platform to improve reuse and delivery speed for subsequent ML initiatives.

### Student researcher

Osram GmbH, Germany

Apr 2016 – Oct 2017

- Research and development of cross platform GUI using WxPython.
- Translation of MATLAB code into C++.

### Software Development Engineer

Samsung R&D Institute, India

Jun 2013 – Jun 2015

- Debugged production issues via logs and crash traces, performed root cause analysis, delivered fixes.
- Owned C/C++ file system API and integration work for core operations (mount, unmount, read, write, seek, copy, unlink) with a focus on correctness and reliability under device constraints.
- Collaborated with cross functional teams to integrate changes across product variants and ensure stable behaviour in production builds.

### Other Projects

- GCP MLOps: Vertex AI training + sweeps (PyTorch ResNet18 CIFAR10), experiment tracking to GCS; reached 76.74% test accuracy at 10 epochs on NVIDIA T4.
- Cloud Run inference: FastAPI service loading model from GCS, versioned health endpoint, packaged in Docker and stored in Artifact Registry.
- CI/CD: GitHub tag → Cloud Build → staging deploy + smoke test → digest based promotion to prod, with optional canary traffic split (1%).

## RESEARCH EXPERIENCE

### PhD Candidate

ETH Zurich, Switzerland

Nov 2018 – Nov 2023



# Dr. Priyanka Chaudhary

priyanka.chaudhary18@gmail.com

Indian | B EU/EFTA Permit

<https://www.linkedin.com/in/pchaudha/>

Pfäffikon SZ, 8808 | Mobile: +41 78 241 1868



## Deep Learning for Urban Flood Depth Estimation

- Best Paper Award at ISPRS Geospatial Week 2019 in Semantic Scene Analysis and 3D Reconstruction from Images and Images Sequences track.
- Built and evaluated an image-based flood depth estimation pipeline using multi-task ranking approach on 7000 unique flood images (5-fold CV) and achieving <10 cm MAE on test data. Used GPU-accelerated training to enable fast iteration and reproducible benchmarking.
- Built strict generalization benchmarks for flood forecasting (10 catchments test holdout, 18 validation, plus 3 rainfall events) for 2-hour-ahead water-depth maps, including scalable training (256x256 patches, 14-channel inputs) and uncertainty via deep ensembles (best-case MAE 21 cm for >1 m depths at 1 m resolution).

Tools: Python, PyTorch, Keras, Matplotlib

## Data Mining Lab

TUM, Germany

Apr 2016 – Sep 2016

- Built an end-to-end pipeline to predict article popularity from the Mashable Online News Popularity dataset (39,645 articles, 61 attributes).
- Detected data integrity issues in the original dataset (151 pages with non-standard structure causing missing share extraction) and rebuilt a clean dataset (40 parsed attributes) and parsed 17 additional features from raw pages.
- Trained and benchmarked tree ensembles, achieving 0.6887 accuracy with Random Forest and comparing selection strategies including Recursive Feature Selection (RFE) and custom greedy selection. Also experimented with XGBoost as a boosted tree alternative.

Tools: Python, R

## Natural Language Understanding

ETH Zurich, Switzerland

Feb 2019 – Aug 2019

- Implemented a simple LSTM Language model to perform various experiments.
- Built Story Cloze baselines using sentiment consistency, reaching 57.2% accuracy (TextBlob) and 62.2% (VADER). Repurposed RACE to generate 263,597 training samples and fine-tuned a Transformer model, achieving 75% test accuracy.

Tools: Python, Tensorflow

---

## EDUCATION

### ETH Zürich

Nov 2018 – Nov 2023

PhD Candidate, Scientific Assistant in *Machine Learning and Computer Vision*

### Technical University Munich (TUM)

Oct 2015 – Sep 2018

*Computer Science, 1.5 (1.0 being the highest possible grade), Passed with distinction*

### Delhi Technological University

Aug 2009 – Jun 2013

*Software Engineering, 75.27, First Division with distinction*

---

## PUBLICATIONS

- 2024 - Flood Water Depth Prediction with Convolutional Temporal Attention Networks at *Water Journal*.
  - 2022 - Flood Uncertainty Estimation Using Deep Ensembles at *Water Journal*.
  - 2020 - Water level prediction from social media images with a multi-tasking approach at *ISPRS Journal of Photogrammetry and Remote Sensing*.
  - 2019 - Flood-Water Level Estimation from Social Media Images at *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.
- 

## LANGUAGE SKILLS

English: Native

Hindi: Native

German: Intermediate