Explaining the spatial pattern of the debris flow and flood hazard in High Mountain Asia

Varvara Bazilova, Tjalling de Haas, Walter Immerzeel

- Estimated probabilities of the catchment being debris flow vs flood dominated
- Area, Melton ratio, mean/median slope and some permafrost conditions are important
- Climate does not add a lot of new information to the classification

I. Data:
- Target values (Y):
  - manually identify the catchments, based on the shape of the alluvial fan ("target value")
  - build a dataset with the characteristics of these catchments (features) for each catchment

- Features (X):
  - morphometric (resemble the shape of the catchments)
  - climatic (describe the temperature and precipitation regime) - from ERA5-Land
  - glaciers (RGI v6.0)
  - permafrost (Obu et al., 2018)

II. Methods:
- Build the model:
  - Catboost (gradient boosted trees)
- Evaluate the model:
  - random guess accuracy: 75.7%
  - accuracy: 91%, 92%
  - area under ROC curve: 0.91, 0.92

III. Results:
- Feature importance. How did model make this prediction?
  - small areas, small Melton ratio (relief*area^{0.5}) make it more likely to be classified as "debris flow" (1)
  - high mean and median slope bring the model result closer to the "flood" class (0)
  - low fraction of sporadic permafrost drags the model towards "flood" (0) class

IV. Next steps:
- Expand the dataset to cover more areas
- Add vegetation coverage as predictor
- Apply the model beyond the training dataset
- Explore the regional differences and spatial patterns
- Assess the climate change impact ("force" the model with climate scenarios)