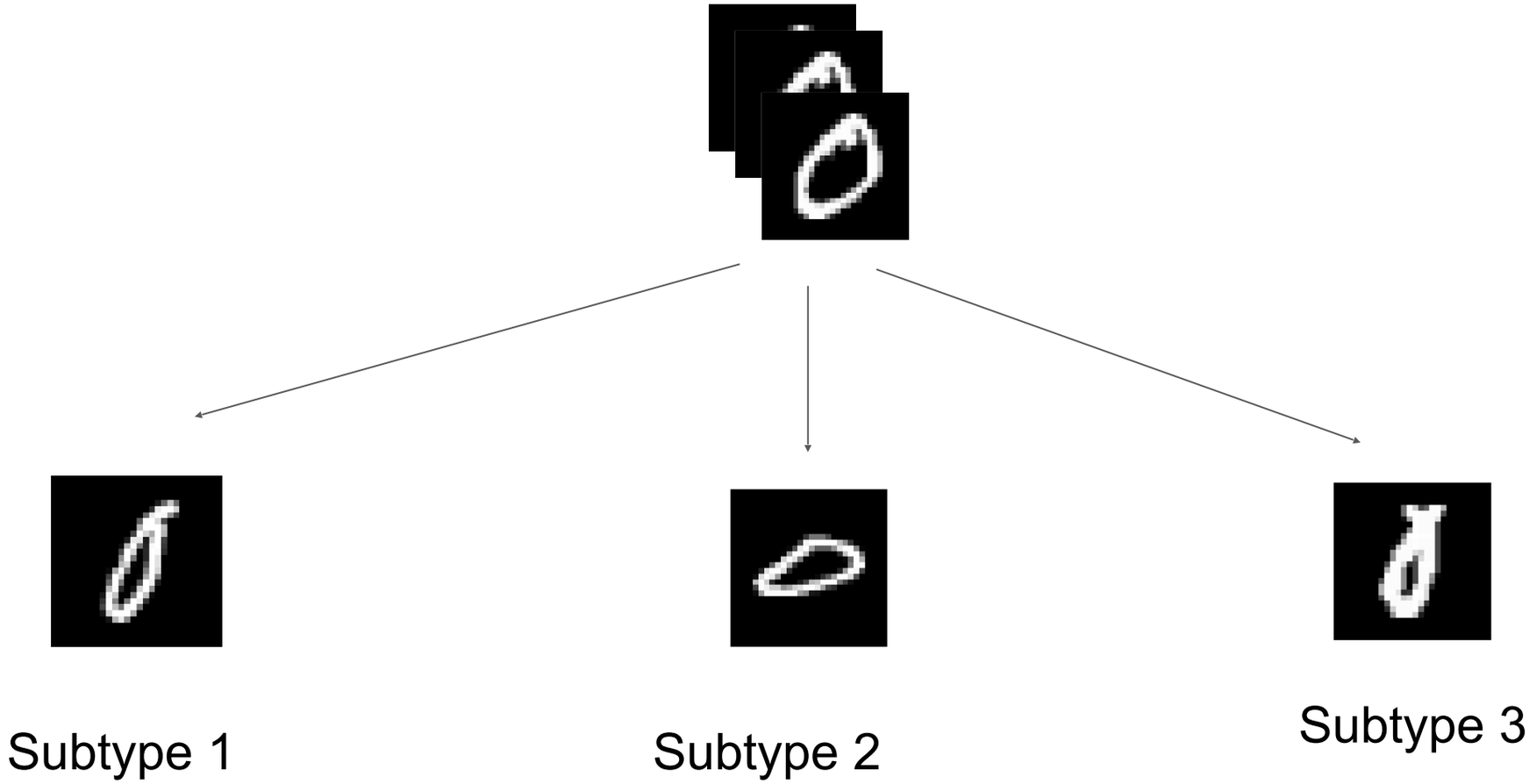
OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

INF

FAKULTÄT FÜR
INFORMATIK

Categorizing Classification Errors

Introduction



Why does it matter ?

- Understanding and developing correct Dataset structure
- Automation of Error detection categorization
- Effect of the dataset on Classification
- Explore the reasoning behind misclassification

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG | INF | FAKULTÄT FÜR INFORMATIK

Project Objective

- Dissect classification error subtypes
- Finding an efficient approach to categorize similar misclassified examples
- Implementing a way to visualize error subtypes

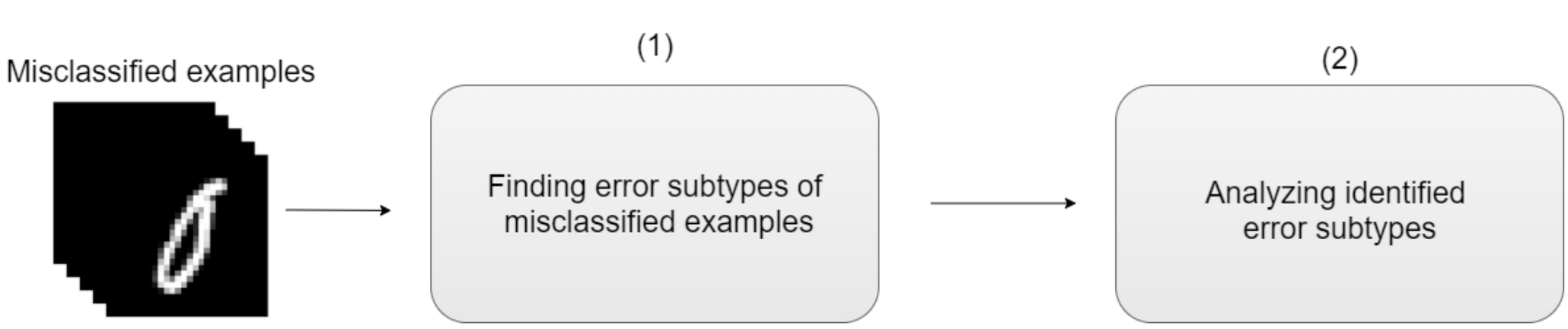
Background

- Understanding error behavior:
  - GradCAM
  - Layer-wise Relevance Propagation
  - Saliency Maps

- Grouping techniques:
  - Clustering approaches
  - Calculating similarities of feature vectors

Method: Concept

2 Stages:



Method: Baseline approach

- Stage 1 (Identifying error subtypes):
  - Applying a simple clustering approach (e.g. K-means clustering)
- Stage 2 (Understanding the error subtype):
  - Take the average of all examples in a cluster group

- In such way
  - we can categorize the error subtypes and define why they are wrong

Method: Baseline approach

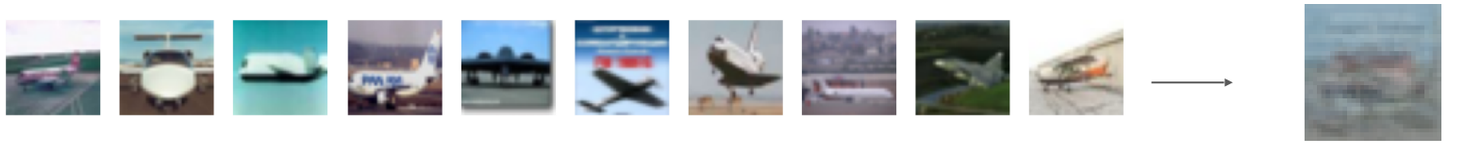
- Unfortunately such baseline approach does not work on all datasets (e.g. CIFAR10)
  - (1) Simple clustering approaches (e.g. K-means) will not easily group such images
  - (2) Averaging over all the examples in the group will result in a noisy image

- Example (MNIST):

Cluster example

Average of cluster

- Example (CIFAR10):



Cluster example

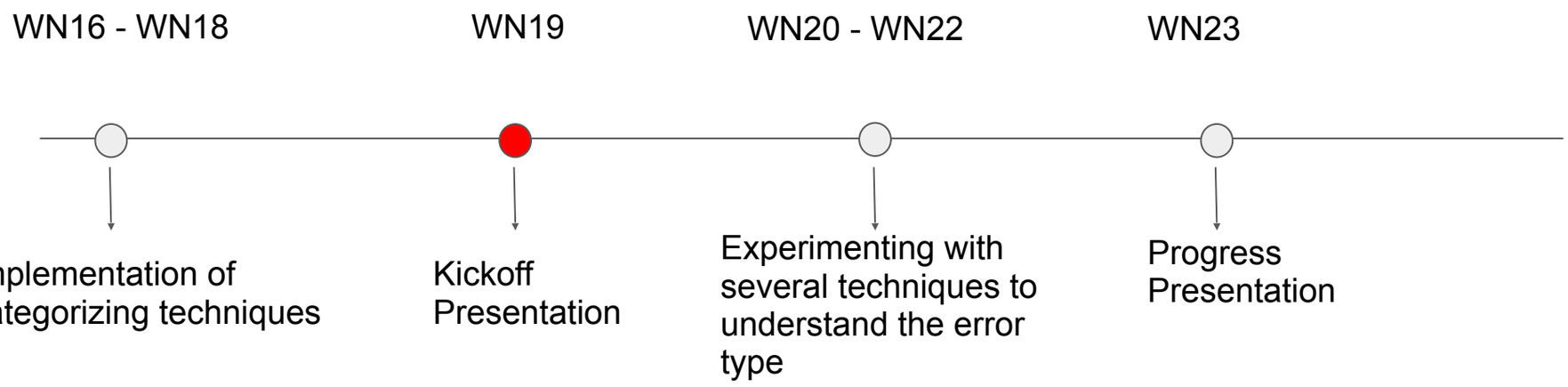
Average of cluster

Method: Approach

- Improve the grouping approach
  - Embeddings
  - Autoencoders

- Improve how to represent the error subtype
  - Hierarchical clustering representations
  - Applying clustering again on the groups
  - Saliency maps

Plan: Time plan



Plan: Work distribution

- Stage 1:
  - Embeddings (Khalifa)
  - Autoencoders (Hatim, Mohamed)

- Stage 2:
  - Hierarchical clustering representations (Khalifa)
  - Applying clustering again on the groups (Hatim)
  - Saliency maps (Mohamed)

Thank you for your attention