OpenMax with Clustering for Open-Set Classification

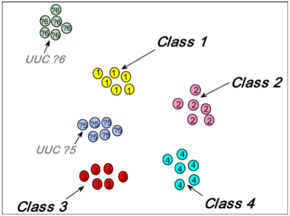
Dennys Huber

February 5, 2024

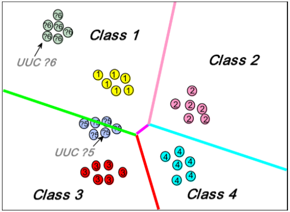
University of Zurich

Table of contents

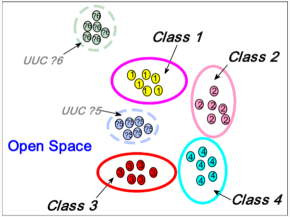
Open-Set



(a) Distribution of the original dataset



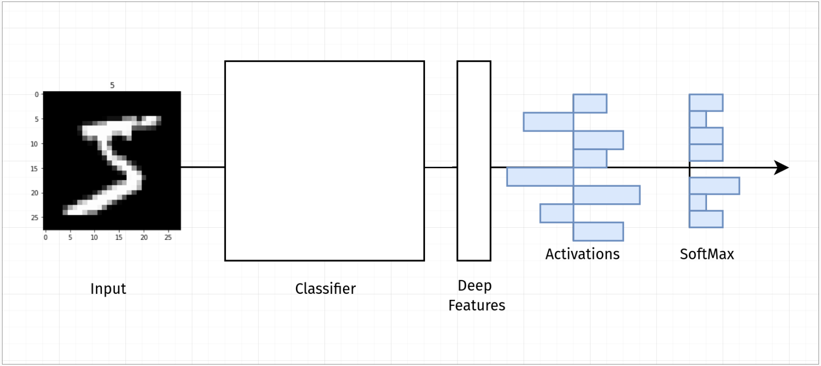
(b) Closed-set classification problem



(c) Open-set classification problem

OpenMax

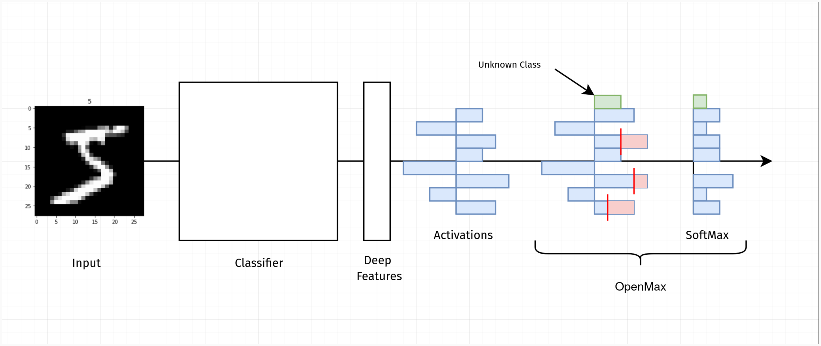
softmax



Introducing OpenMax

- Extension of SoftMax during testing
- Distance based
- Deep features
- Extreme Value Theory
- Heuristic

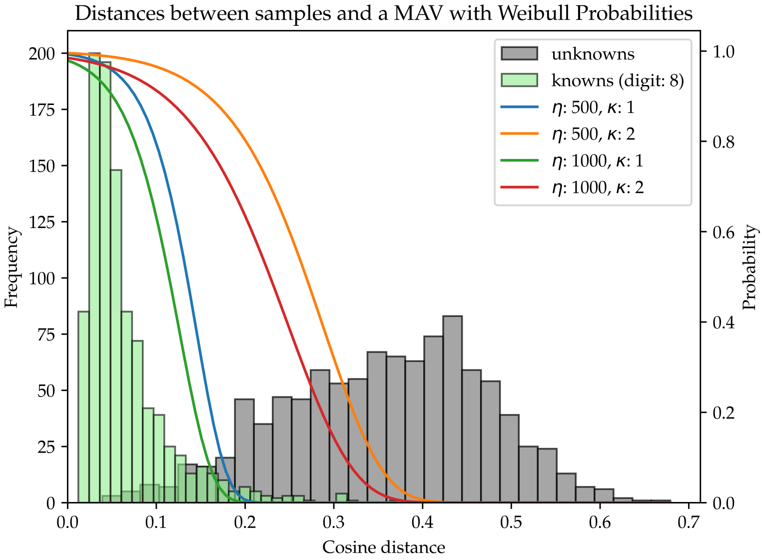
OpenMax Overview



Calibrating OpenMax

- Training
- Class representation
- Mean Activation Vector μ
- Features from penultimate layer ϕ
- Correctly classified samples
- Weibull Distribution
- Tail size η & Distance Multiplier κ

Building a Weibull Distribution



OpenMax in Detail

For a testing sample:

1. Sort and select α classes (logit value)
2. Weibull Probability ω based on distance μ_i and ϕ
3. $\hat{z} = z \circ \omega$
4. $\hat{z}_{N+1} = \sum_i z_i(1 - \omega_i)$
5. $\mathrm{softmax}(\hat{z})$

Approach

Improving OpenMax

- Handling negative logits
  - Value Shift
  - Adjust Probabilities
- Introducing weight factors
  - $\varphi_N = \frac{1}{N-1}$
  - $\varphi_\omega = \frac{1}{\sum_i (1-\omega_i)}$
- Removing alpha parameter

Research Questions 1

RQ1: Can OpenMax performance be enhanced by accounting for negative activation values?

Open-Set Classification Rate

Correct Classification Rate (CCR)

- Known samples
- Threshold θ

$$\mathrm{CCR}(\theta) = \frac{|\{k_c \,|\, \operatorname{argmax}_{1 \leq n \leq N} y_{c,n} = \tau_c \wedge y_{c,n} \geq \theta\}|}{|K|} \tag{1}$$

False Positive Rate (FPR)

- Unknown samples
- Threshold θ

$$\mathrm{FPR}(\theta) = \frac{|\{u_c \,|\, \operatorname{argmax}_{1 \leq n \leq N} y_{c,n} \geq \theta\}|}{|U|} \tag{2}$$

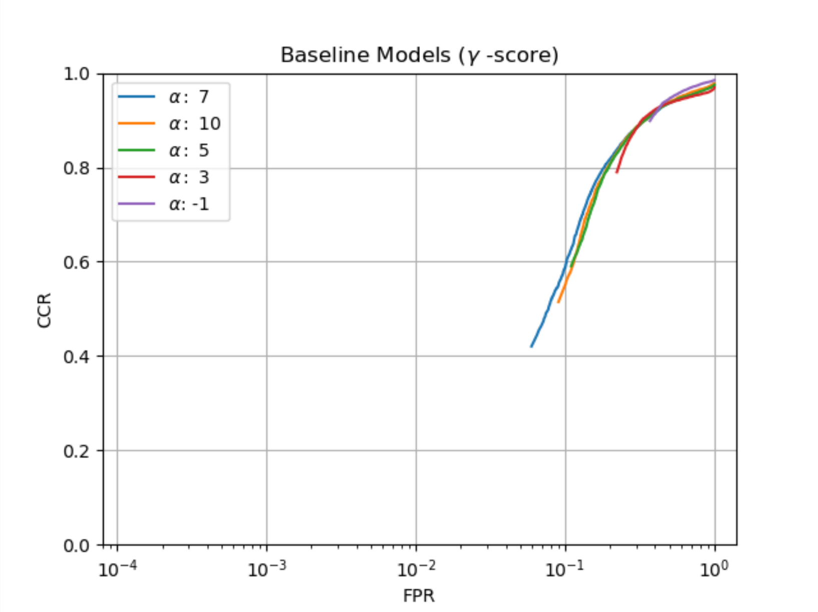
Defining Scoring System

- Σ-Score - Thresholding
- γ-Score $\gamma = \frac{(\gamma^+ + \gamma^-)}{2}$

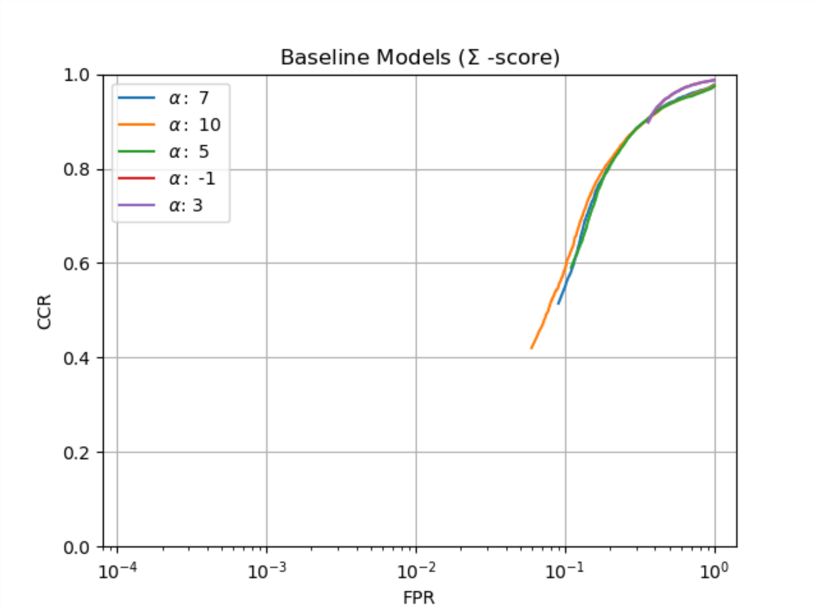
$$\gamma^+ = \frac{1}{|K|} \sum_{c=1}^{|K|} y_{\tau_c} \tag{3}$$

$$\gamma^- = \frac{1}{|U|} \sum_{c=1}^{|U|} \left(1 - \operatorname{argmax}_{1 \leq n \leq N} y_{c,n}\right) \tag{4}$$

Baseline (γ)

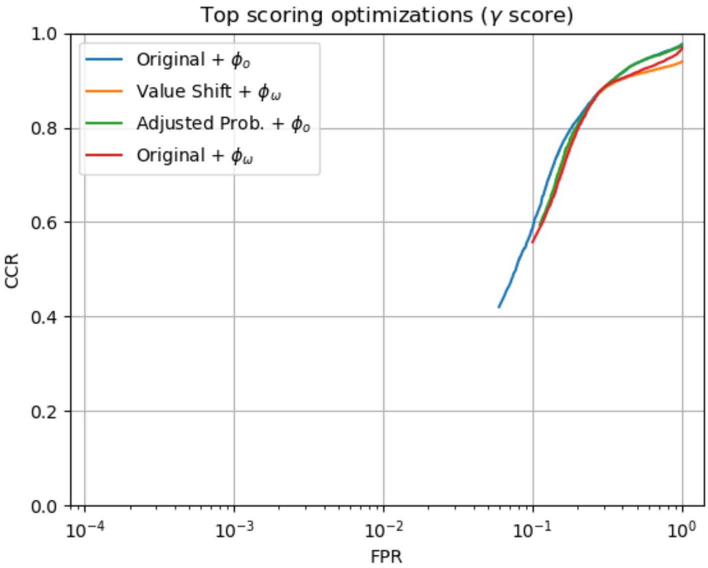


Baseline (Σ)

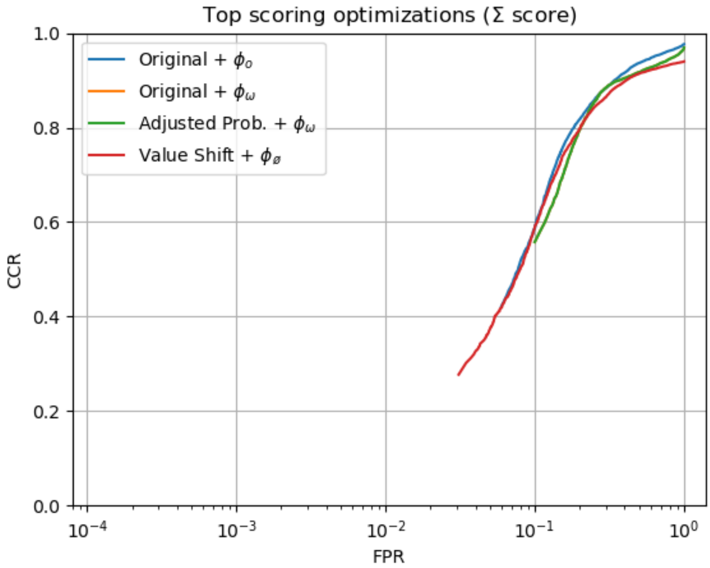


Baseline Models (Σ -score)
1.0
0.8
0.6
0.4
0.2
0.0
CCR
α: 7
α: 10
α: 5
α: -1
α: 3
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
FPR

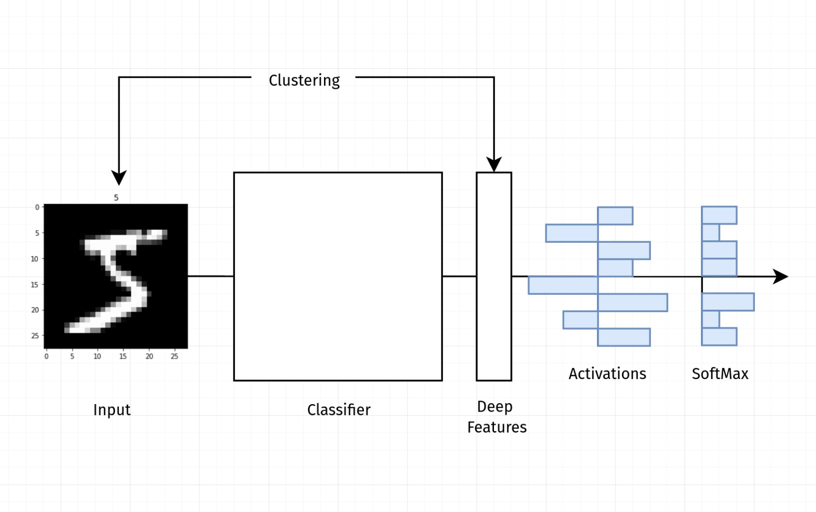
Results RQ1 (γ)



Results RQ1 (Σ)



Where to apply the clustering?



Introducing Clustering to OpenMax

Input Clustering:

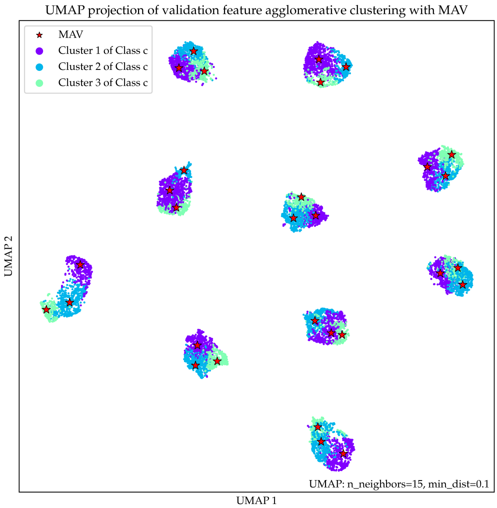
- Input data
- Per class
- Each cluster a class

Features Clustering:

- Training Features
- Validation Features
- Per class

Combination of both types

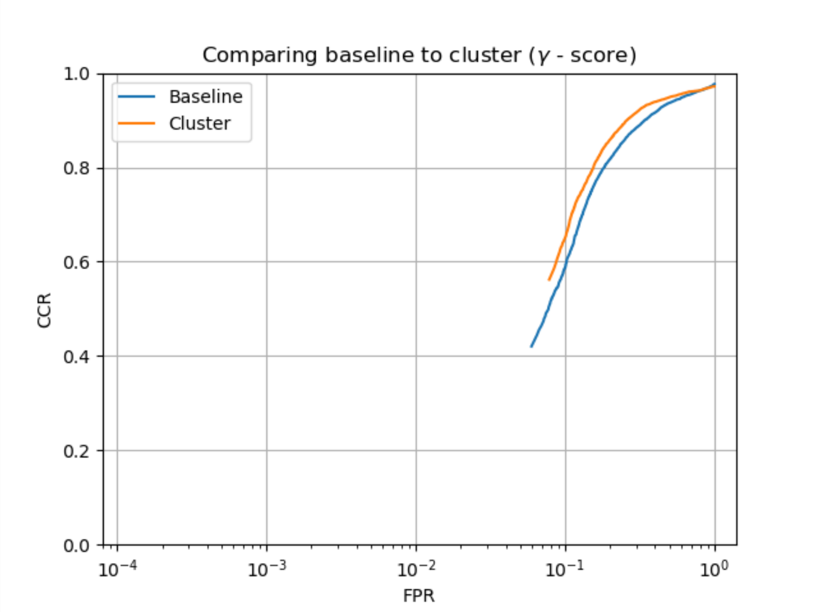
Visualizing Clustering



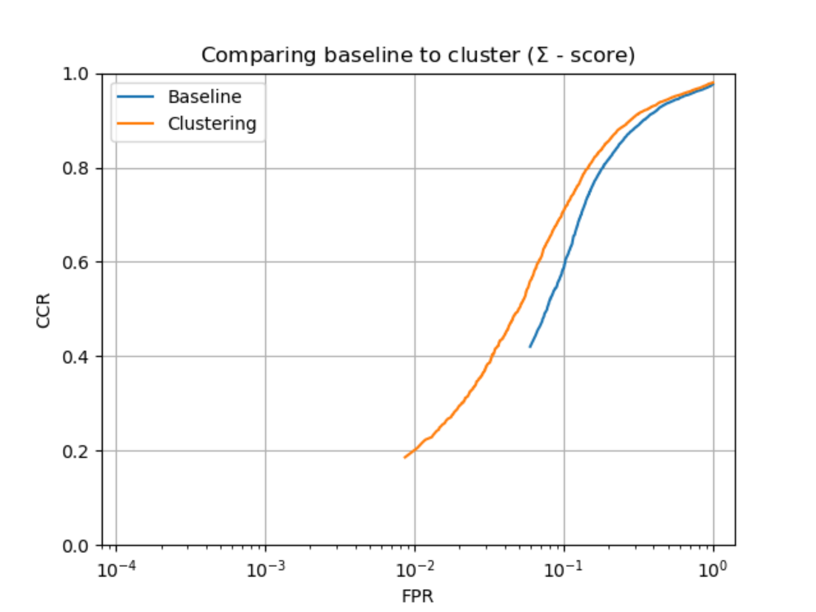
Research Questions 2

RQ2: Can clustering improve OpenMax's performance?

Results RQ2 (γ)



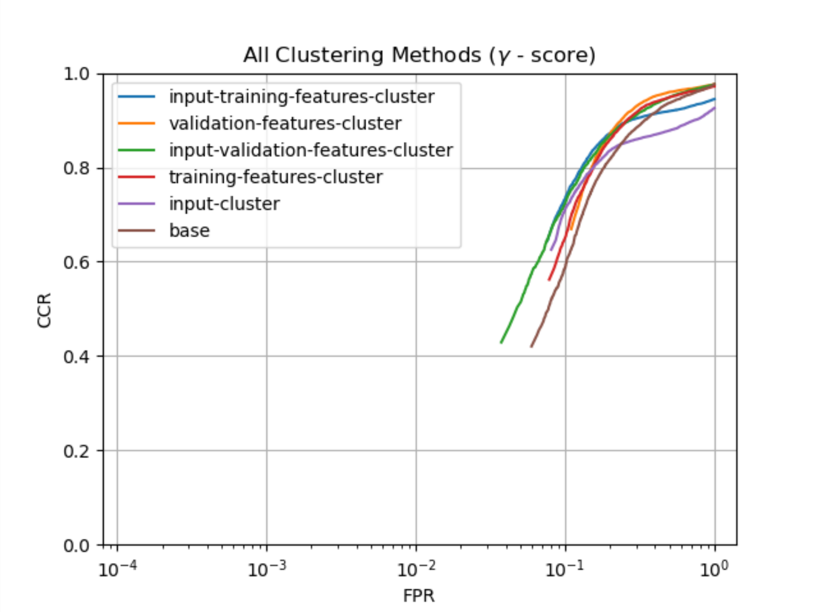
Results RQ2 (Σ)



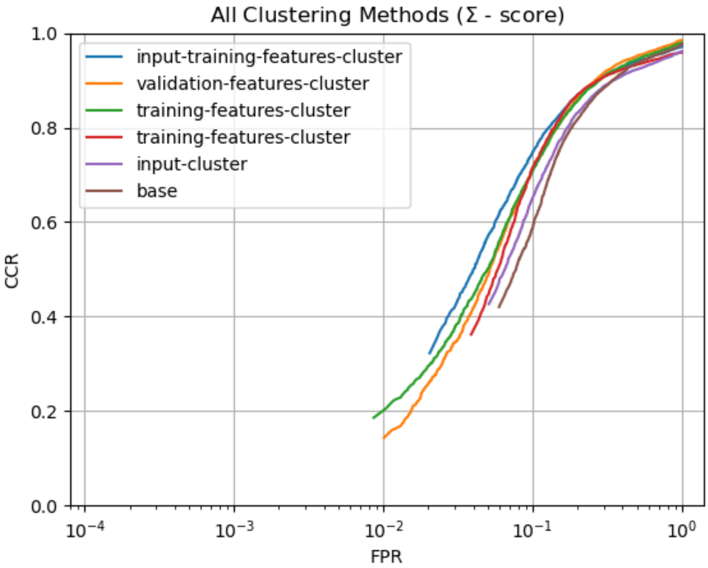
Research Questions 3

RQ3: If clustering improves OpenMax, which clustering type is optimal, and by using which parameters?

Results RQ3 (γ)



Results RQ3 (Σ)



Discussion

Limitations & Future Work

- Datasets
- Other Clusetring Algorithms
- Per clusters parameters
- Costum Loss function

Questions?