THE UNIVERSITY OF BRITISH COLUMBIA

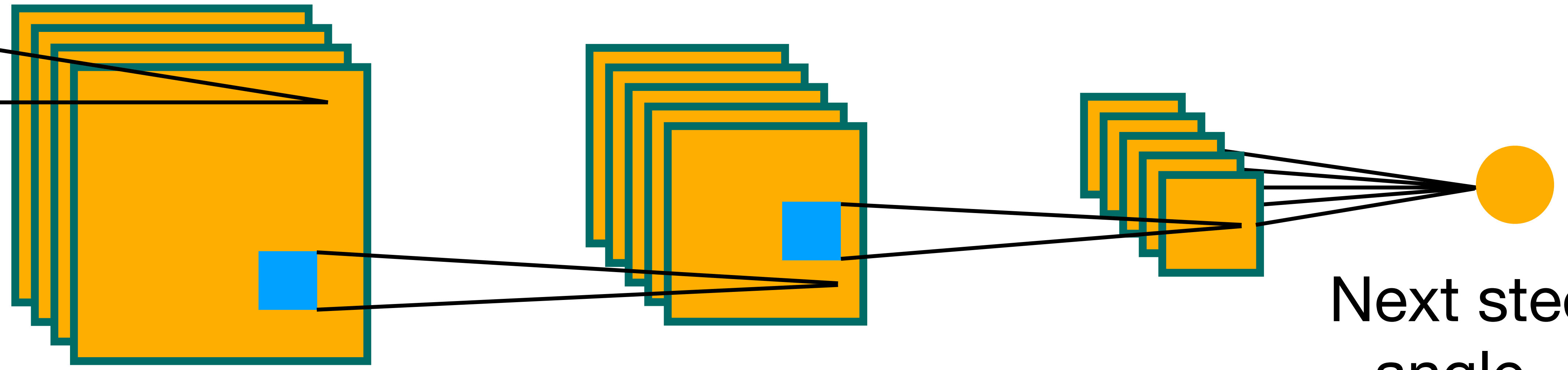
ICLR International Conference On Learning Representations

Learning Label Encodings for Deep Regression

Deval Shah, Tor M. Aamodt

https://github.com/ubc-aamodt-group/RLEL_regression

Deep Regression

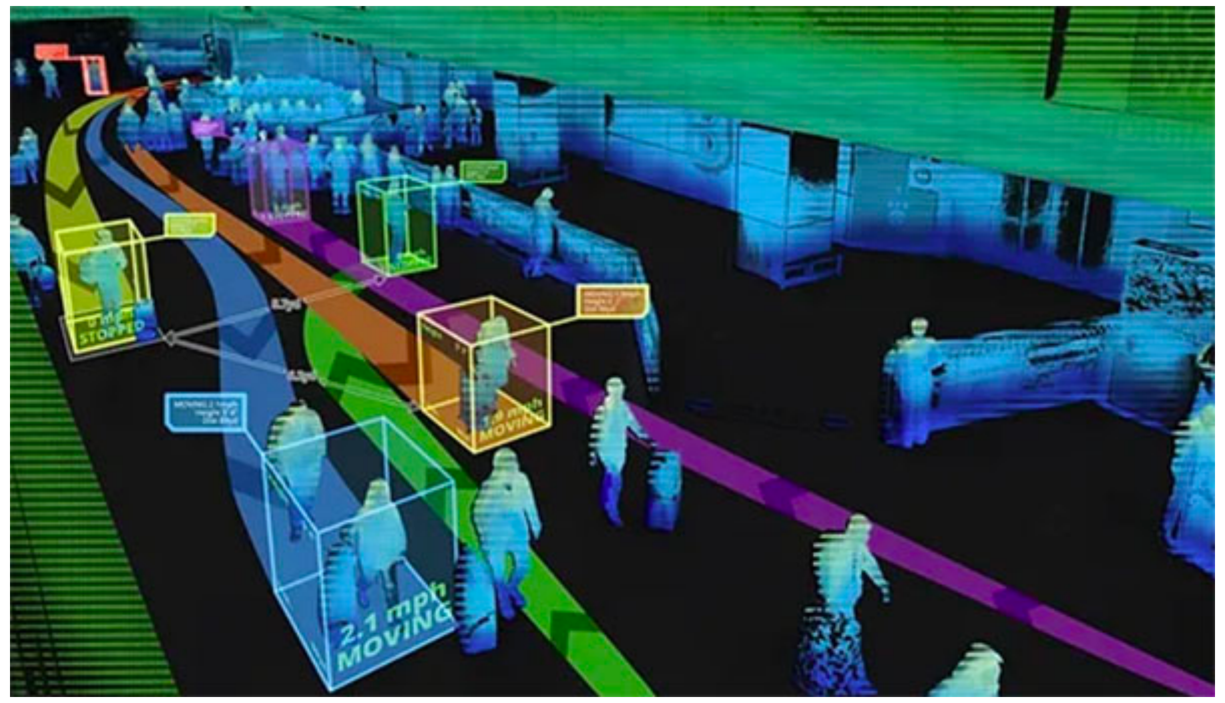


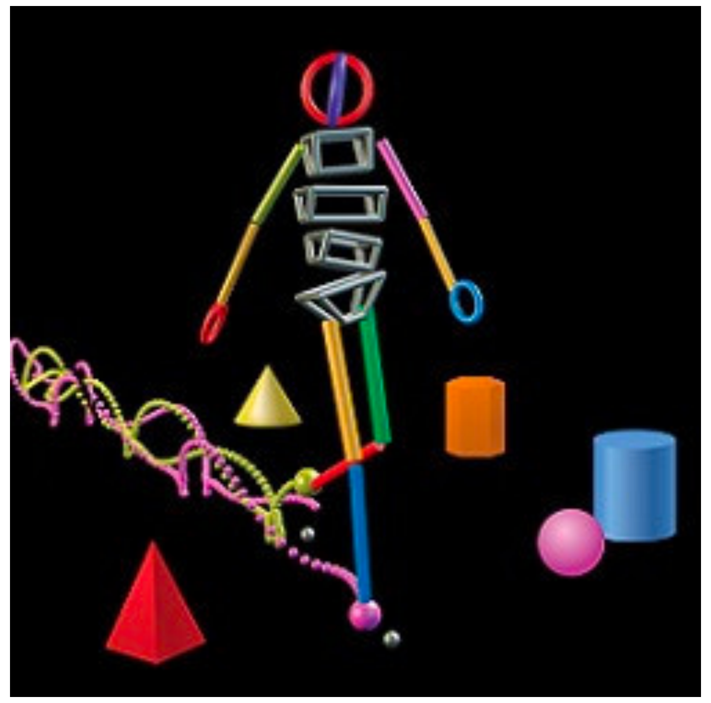
Direct regression:

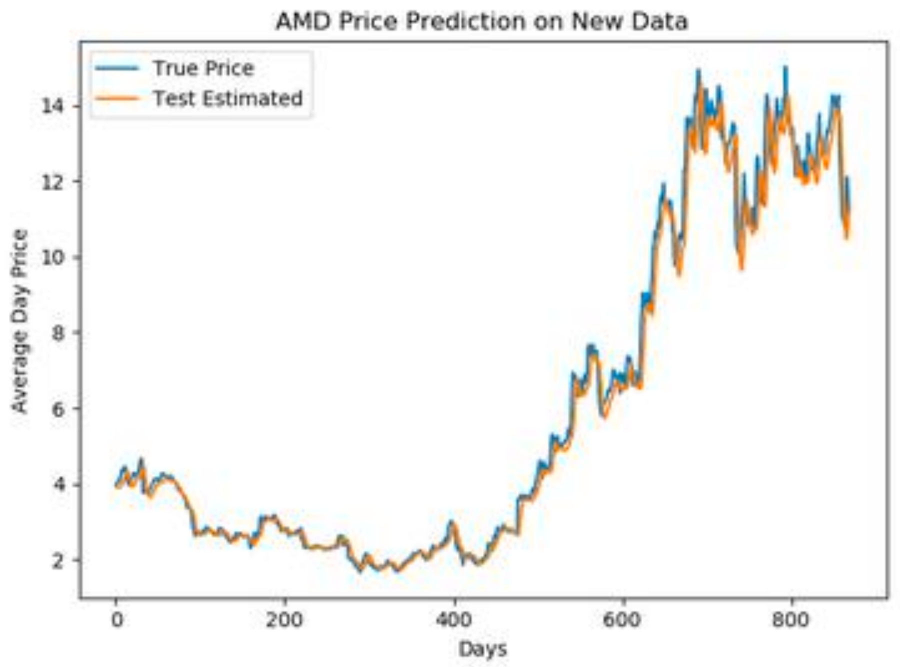
Loss = (Prediction-Target)^2

Or

Loss = |Prediction-Target|







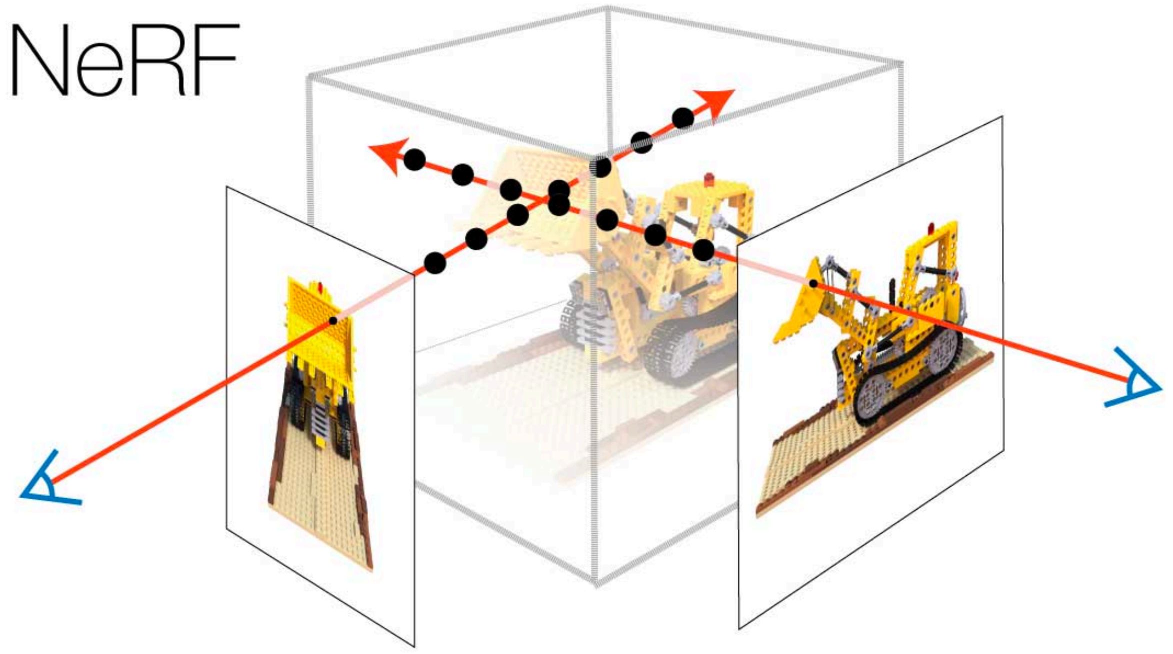


Image sources:

[1] https://developer.nvidia.com/blog/researching-and-developing-an-autonomous-vehicle-lane-following-system/

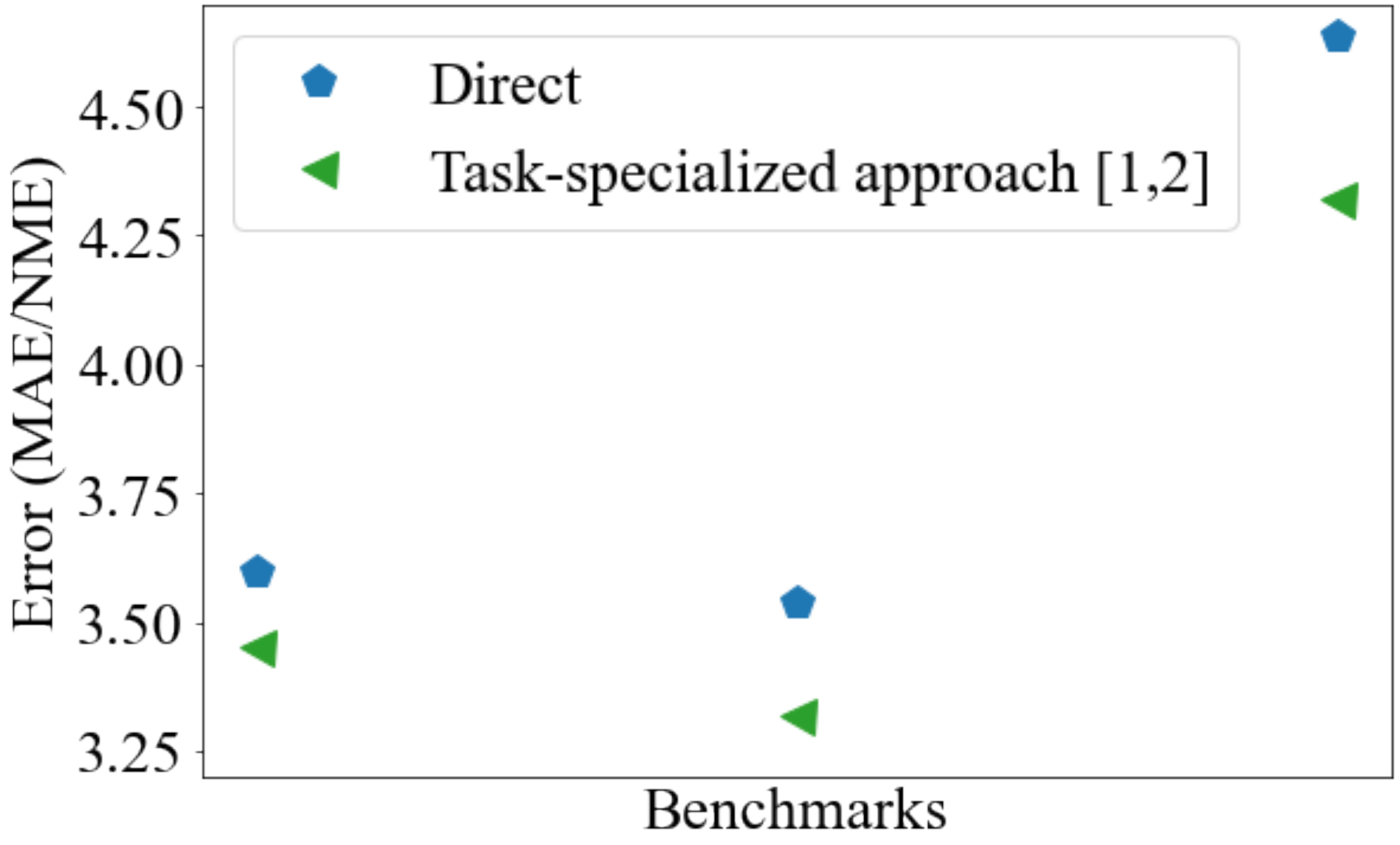
[2] https://www.robotics247.com/article/seoul_robotics_marks_growth_democratizes_3d_perception_ces_2022

[3] O. Salzman, "Sampling-based robot motion planning," *Communications of the ACM, 2019*

[4] https://paperswithcode.com/task/time-series-forecasting

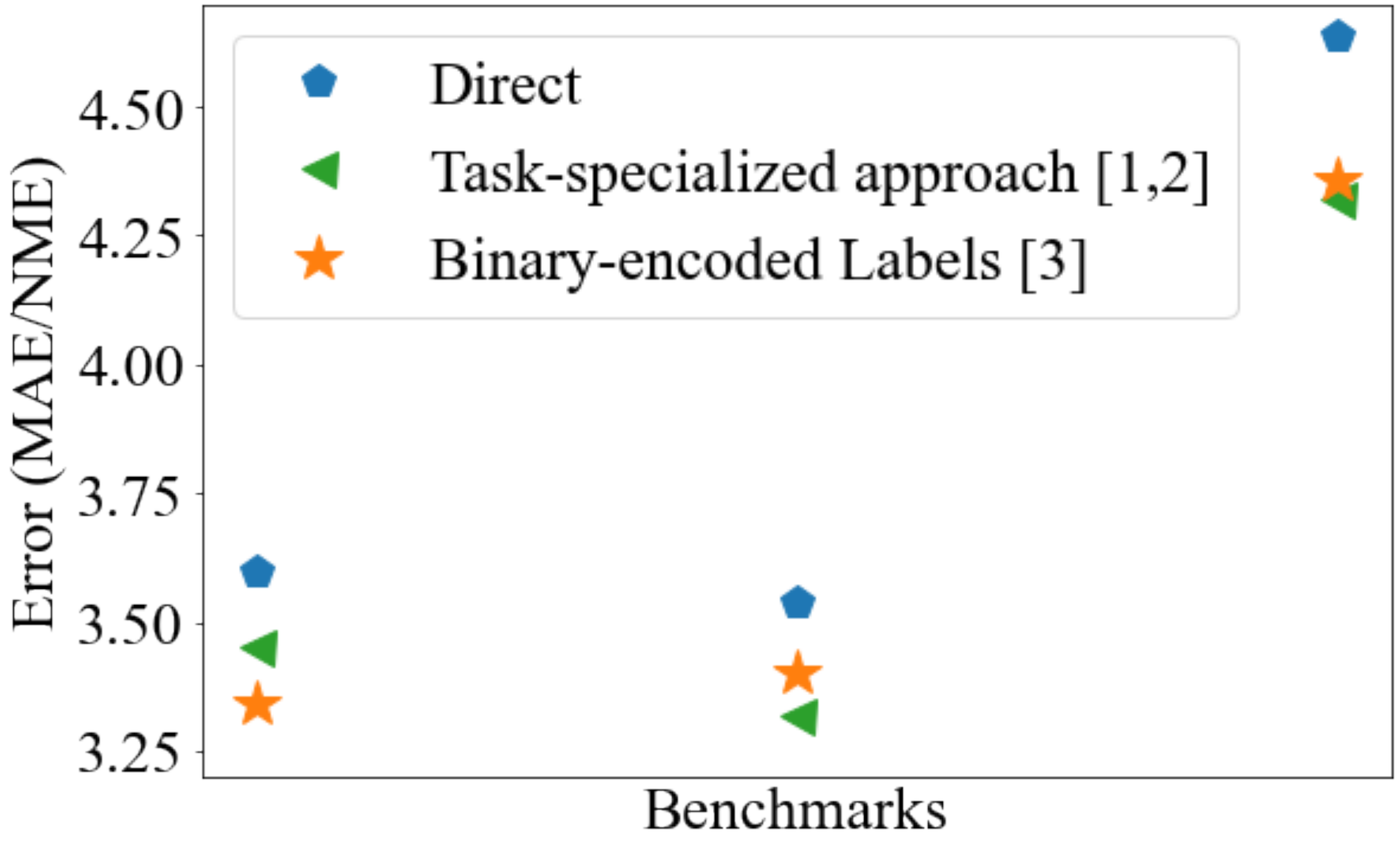
[5] Mildenhall et al., NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, ECCV 2020

Deep Regression



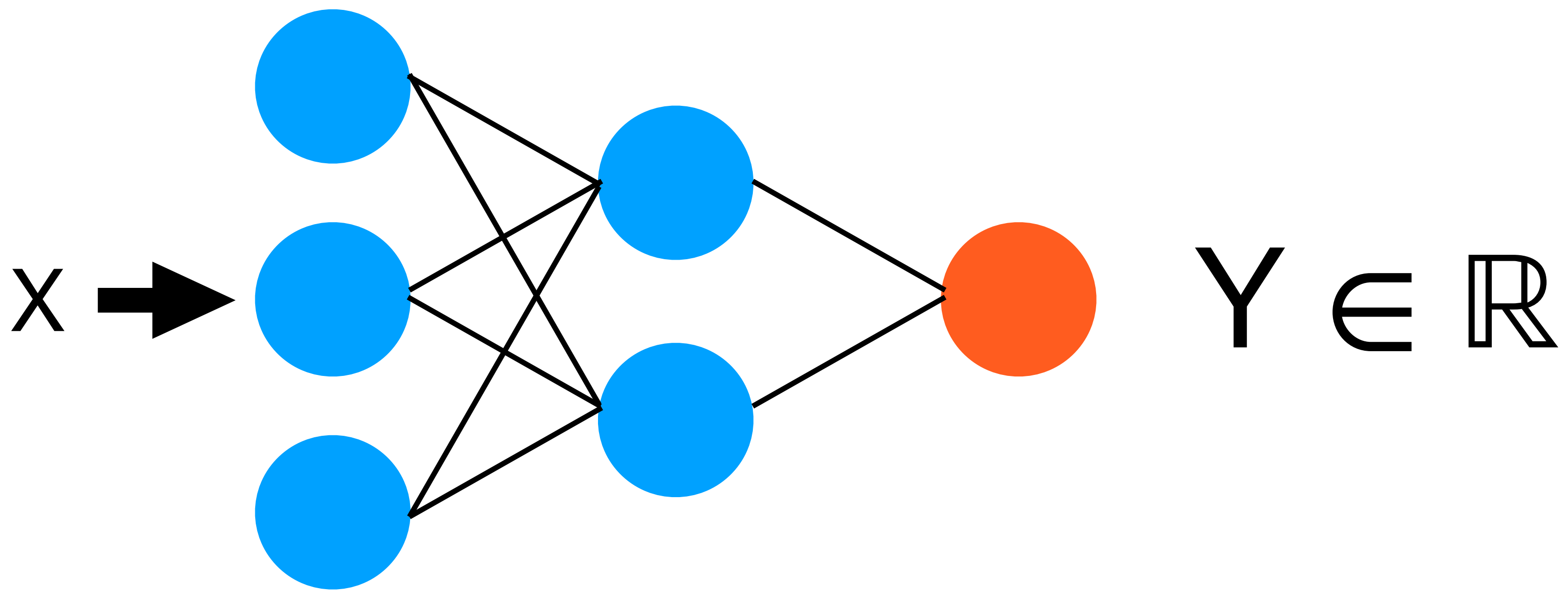
[1] Wang et al., Deep high-resolution representation learning for visual recognition, PAMI 2020
[2] Xu et al., Anchorface: An anchor-based facial landmark detector across large poses , *ArXiv*, abs/2007.03221, 2020

Deep Regression

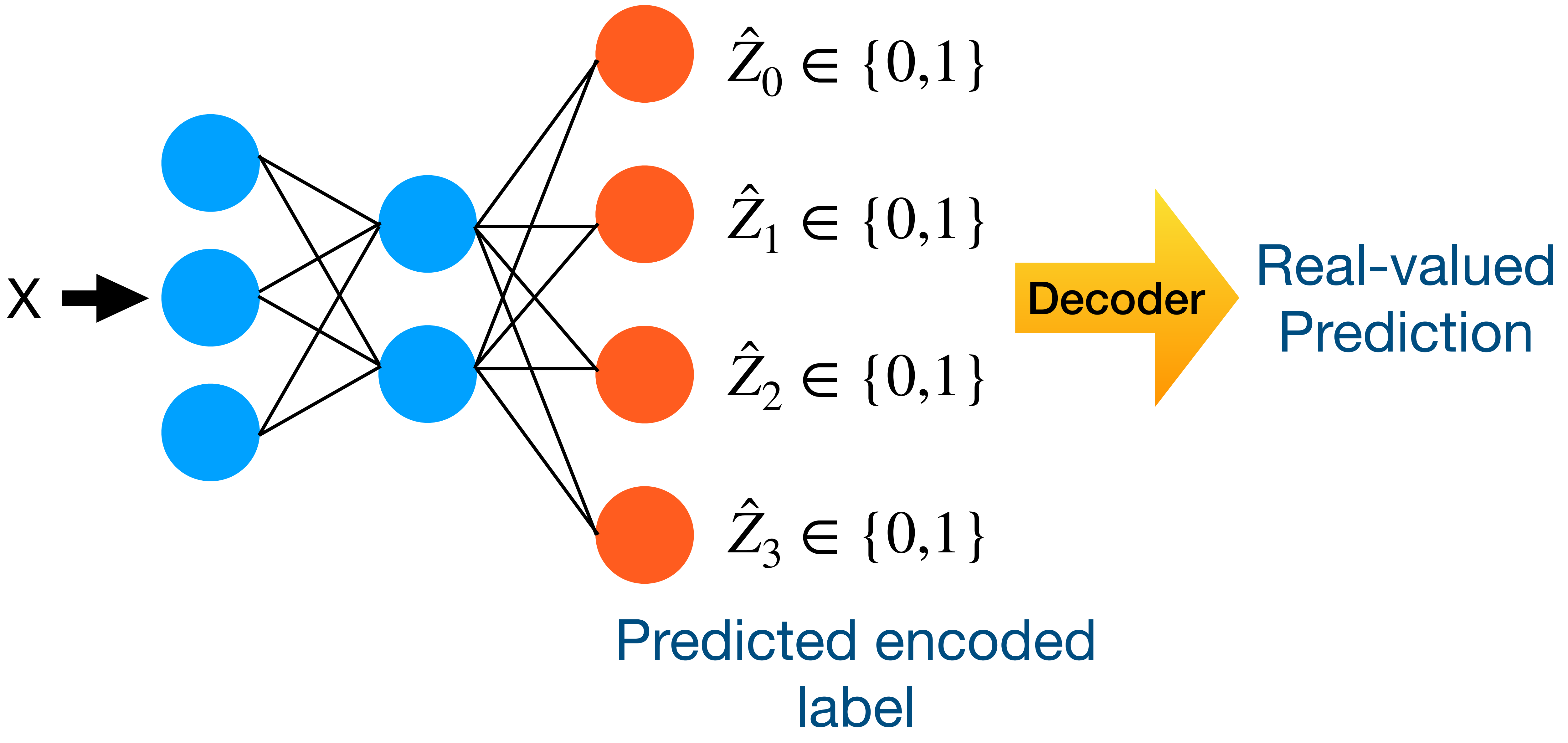


[1] Wang et al., Deep high-resolution representation learning for visual recognition, PAMI 2020
[2] Xu et al., Anchorface: An anchor-based facial landmark detector across large poses , *ArXiv*, abs/2007.03221, 2020
[3] Shah et al., Label Encoding for Regression Networks, ICLR 2022

Regression by Binary Classification



Regression by Binary Classification

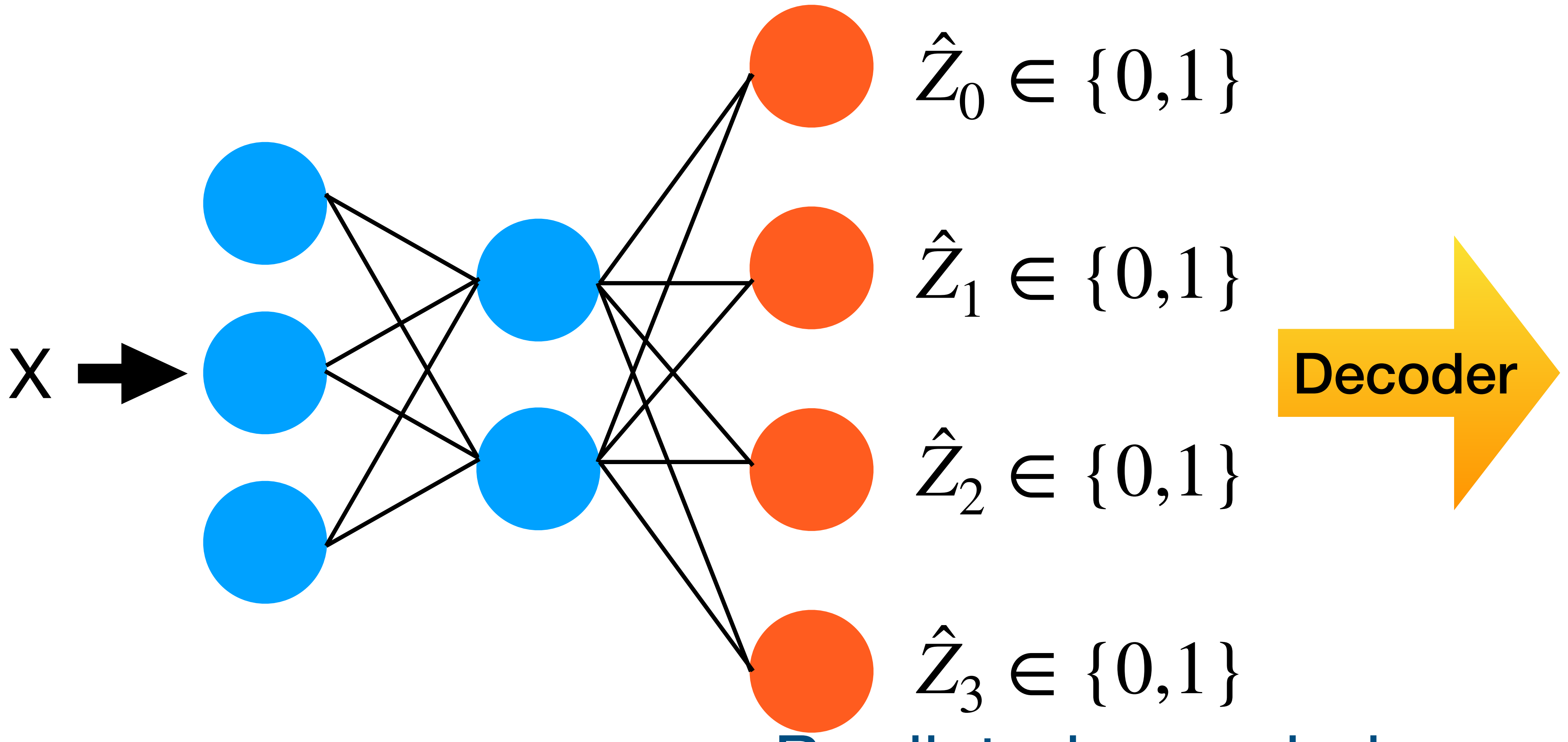


	Z_0	Z_1	Z_2	Z_3
Y $\in (1,2]$	1	0	0	0
Y $\in (2,3]$	1	1	0	0
Y $\in (3,4]$	0	1	1	1
Y $\in (4,5]$	0	0	1	1

Target encoded label

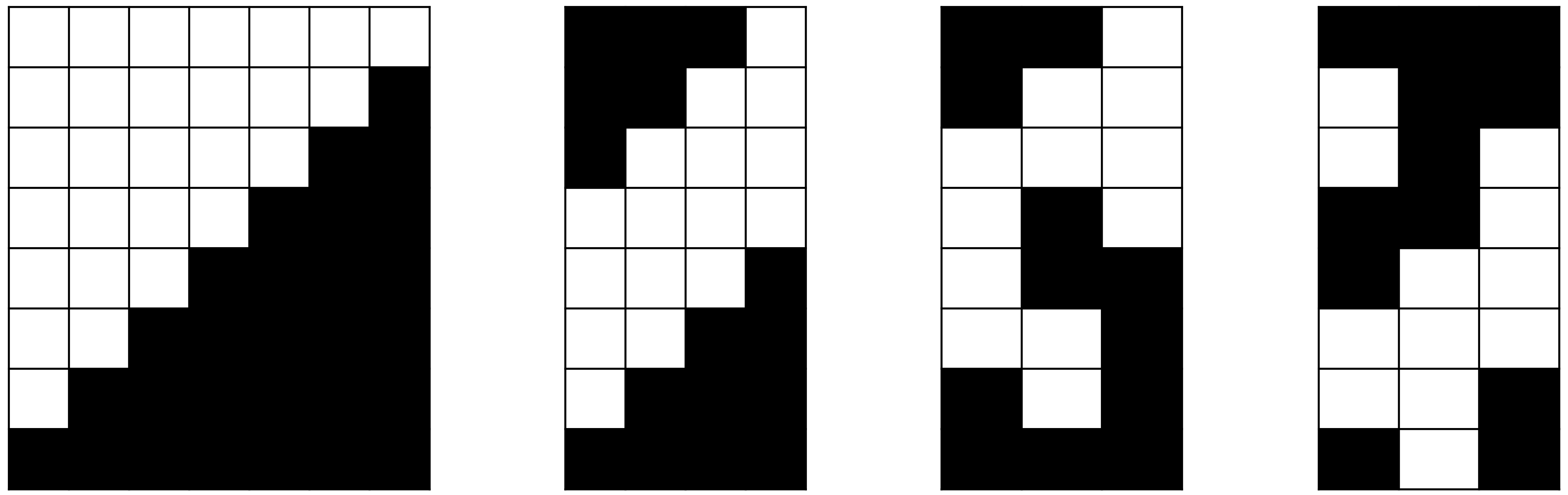
[1] Deval Shah, Zi Yu Xue, and Tor M. Aamodt. Label encoding for regression networks. ICLR 2022
[2] Li & Lin, Ordinal regression by extended binary classification, NeurIPs 2006
[3] T. G. Dietterich & G. Bakiri. Solving Multiclass Learning Problems via Error-Correcting Output Codes, JAIR 1995
[4] Song et al, Error-Correcting Output Codes with Ensemble Diversity for Robust Learning in Neural Networks. *AAAI*, 2021
[5] Niu et al., Ordinal regression with multiple output CNN for age estimation, CVPR 2016

Regression by Binary Classification



Real-valued Prediction

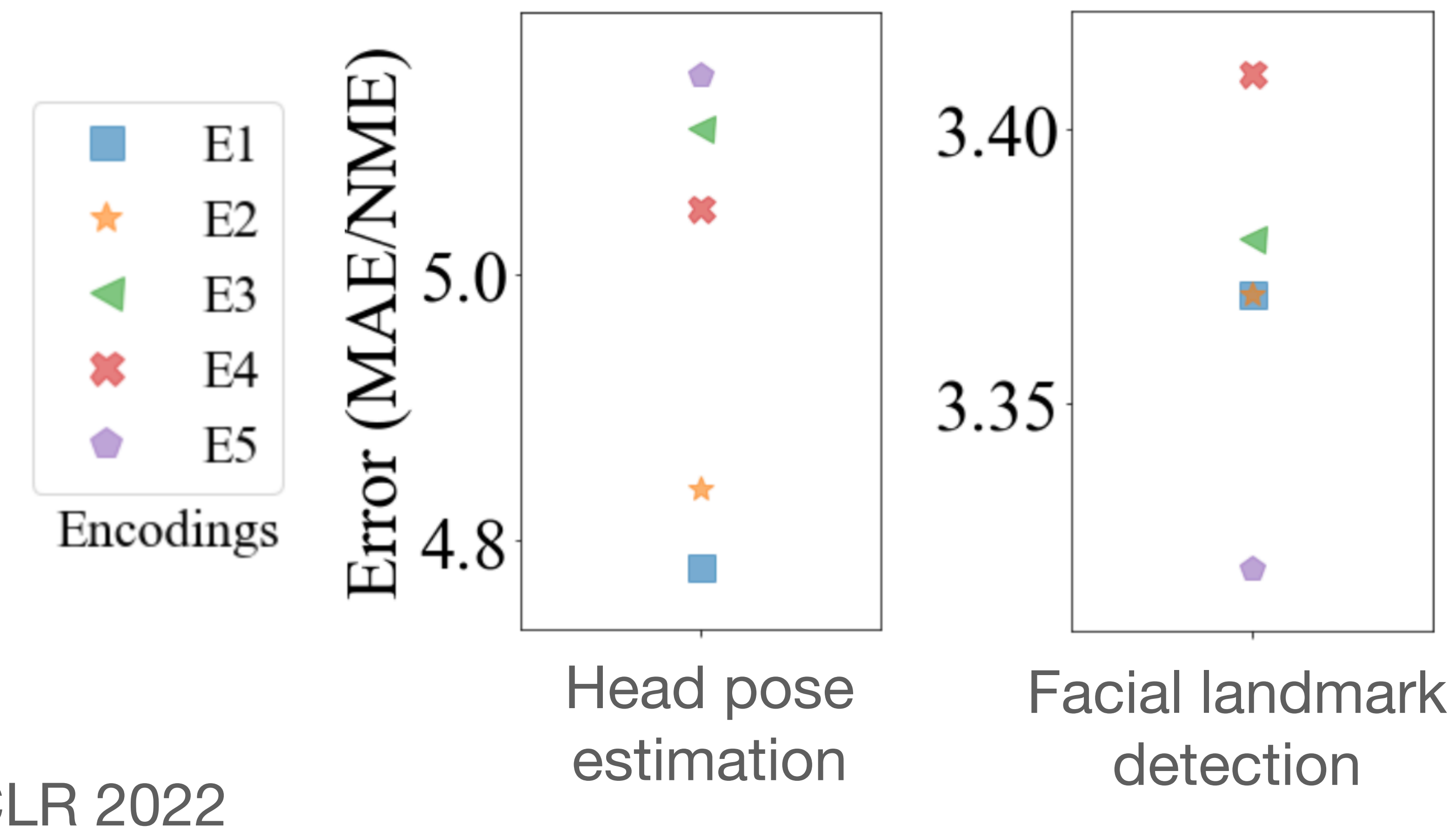
Predicted encoded label



Large design space

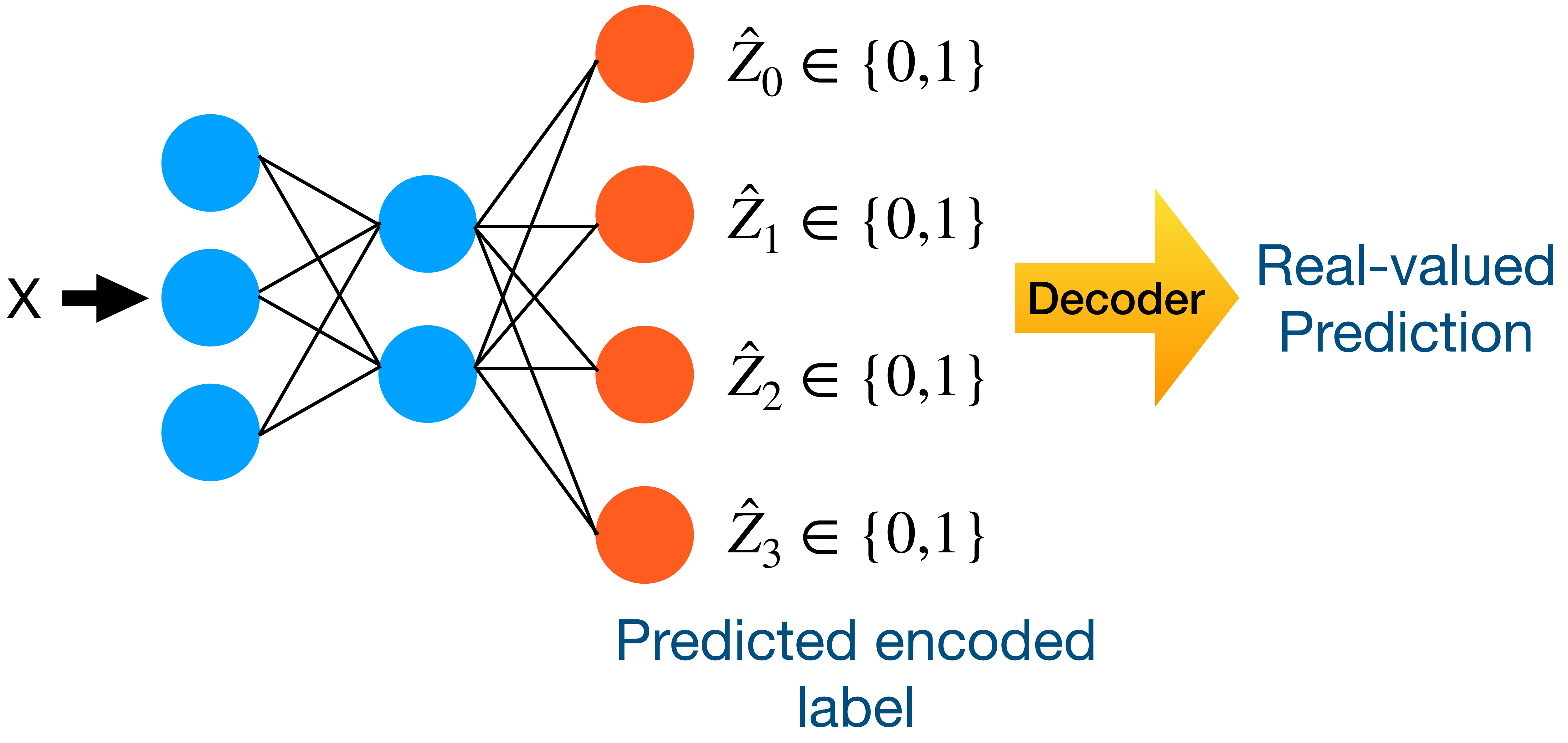
	Z_0	Z_1	Z_2	Z_3
Y $\in$ (1,2]	?	?	?	?
Y $\in$ (2,3]	?	?	?	?
Y $\in$ (3,4]	?	?	?	?
Y $\in$ (4,5]	?	?	?	?

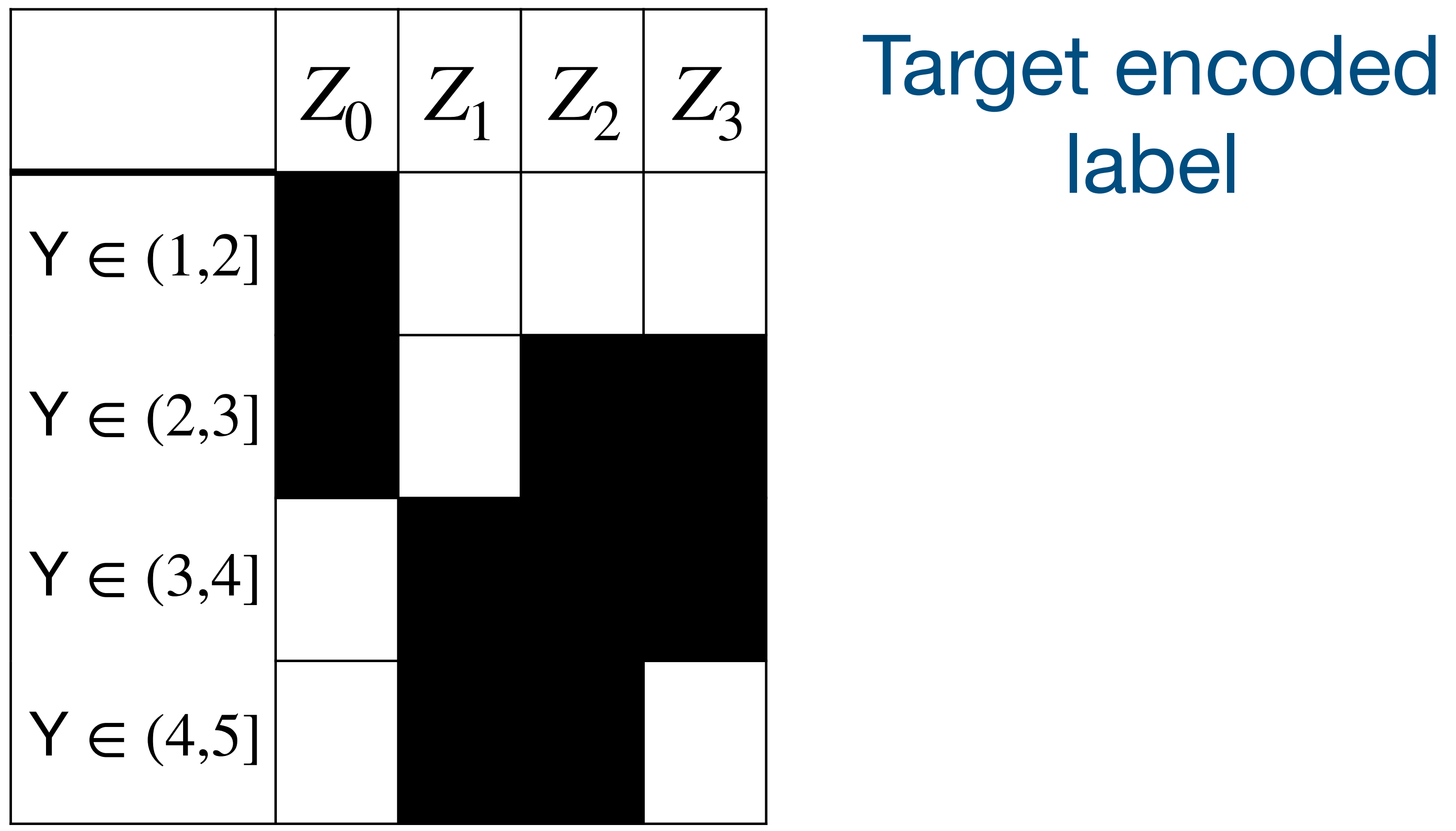
Target encoded label



[1] Deval Shah, Zi Yu Xue, and Tor M. Aamodt. Label encoding for regression networks. ICLR 2022
[2] Li & Lin, Ordinal regression by extended binary classification, NeurIPs 2006

Regression by Binary Classification





Regularized Label Encoding Learning

Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

Challenges: Learning Label Encoding

Problem dependent design choice and discrete design space

Vast design space

M Bits

N Values

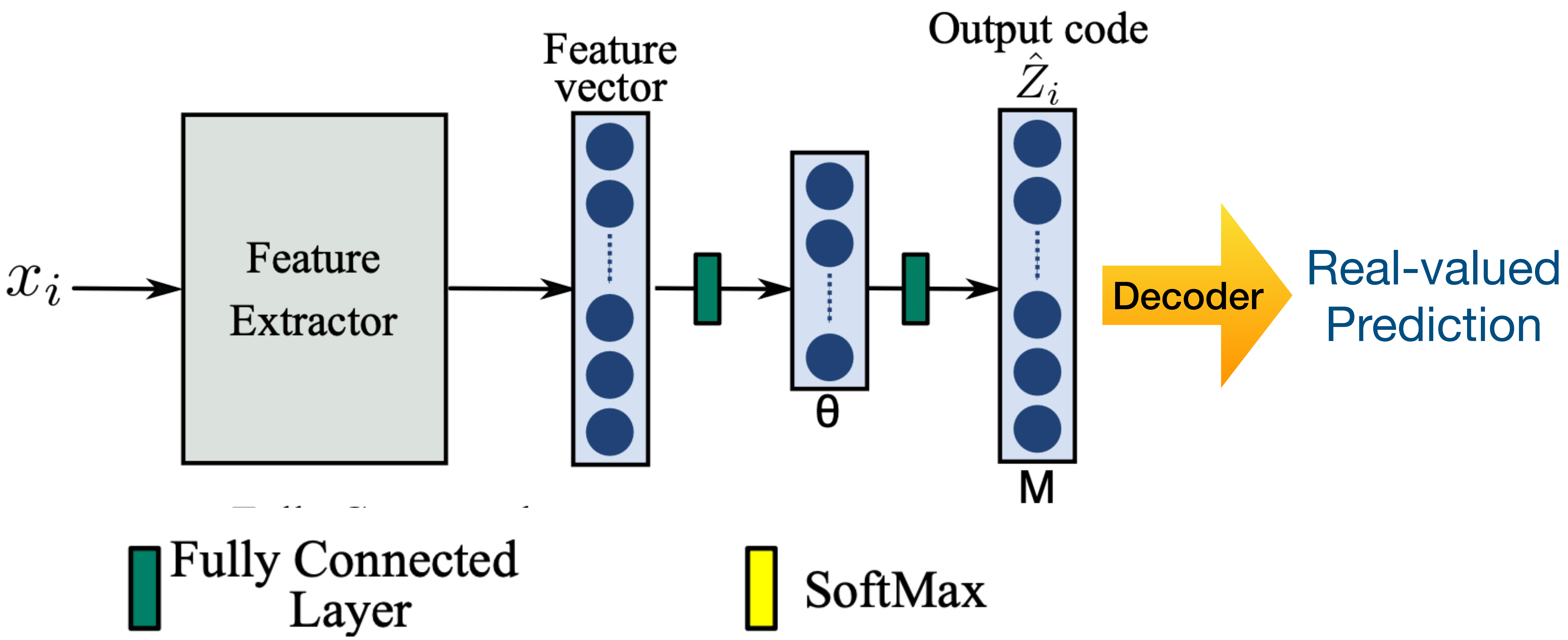
$$\frac{2^M!}{(2^M - N)!}$$

10^{89}

Regularized Label Encoding Learning

- Use continuous search space of label encodings and co-optimize label encodings and network parameters
- Constrain the search space using regularizers

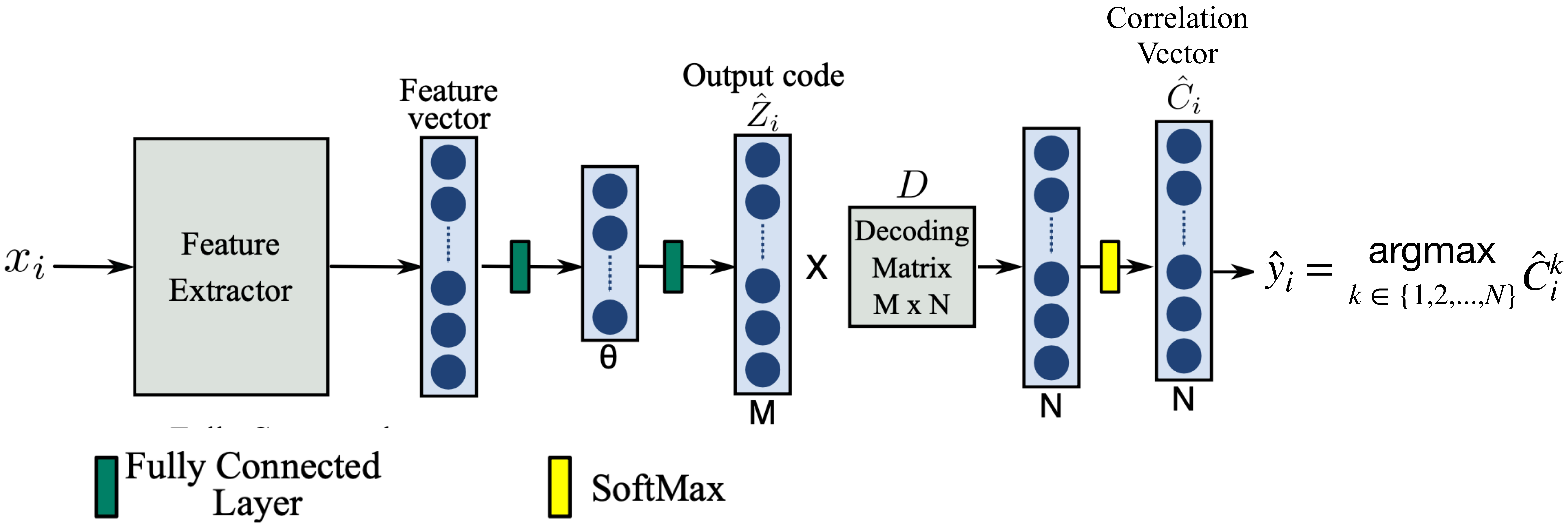
Regression by Binary Classification



Loss = Binary_cross_entropy (Output code, Target code)

Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

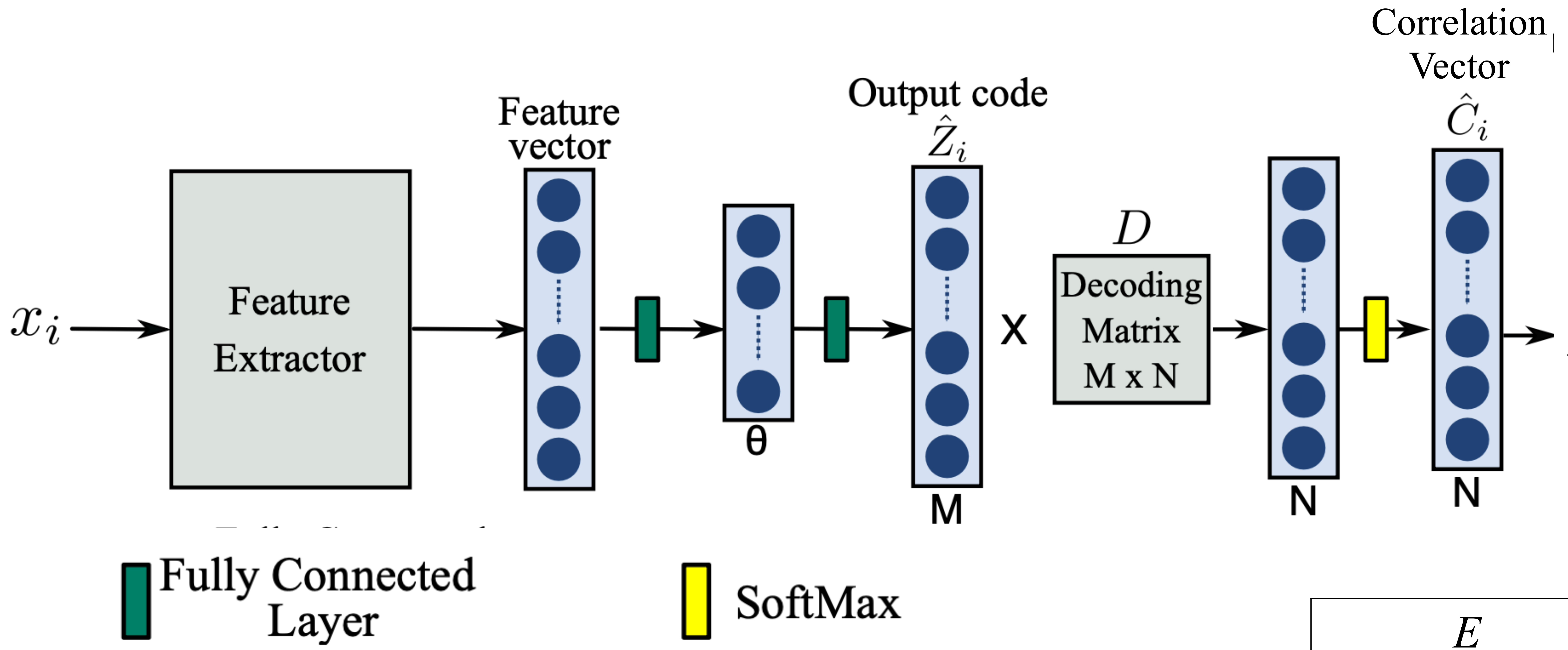
Label Encoding Learning



Loss = Cross entropy (Decoded output, Target label)

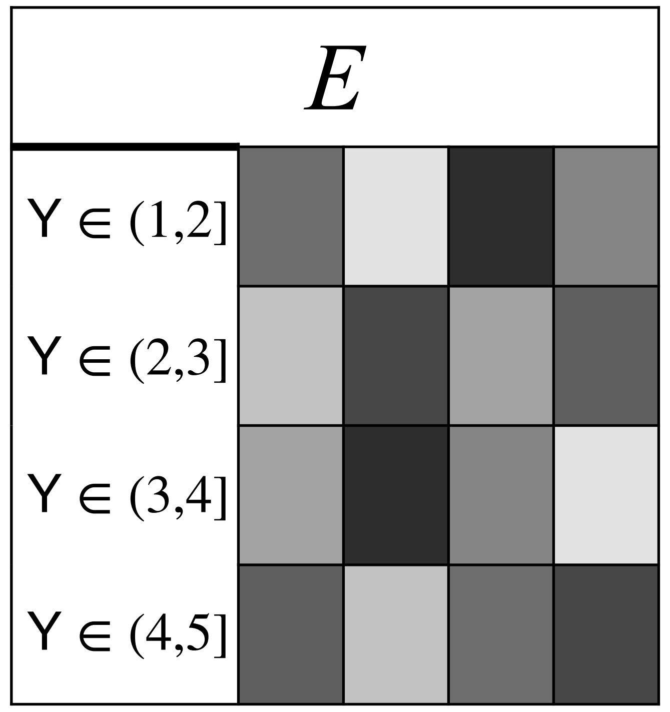
Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

Label Encoding Learning



$$\hat{y}_i = \underset{k \in \{1,2,\ldots,N\}}{\text{argmax}} \hat{C}_i^k$$

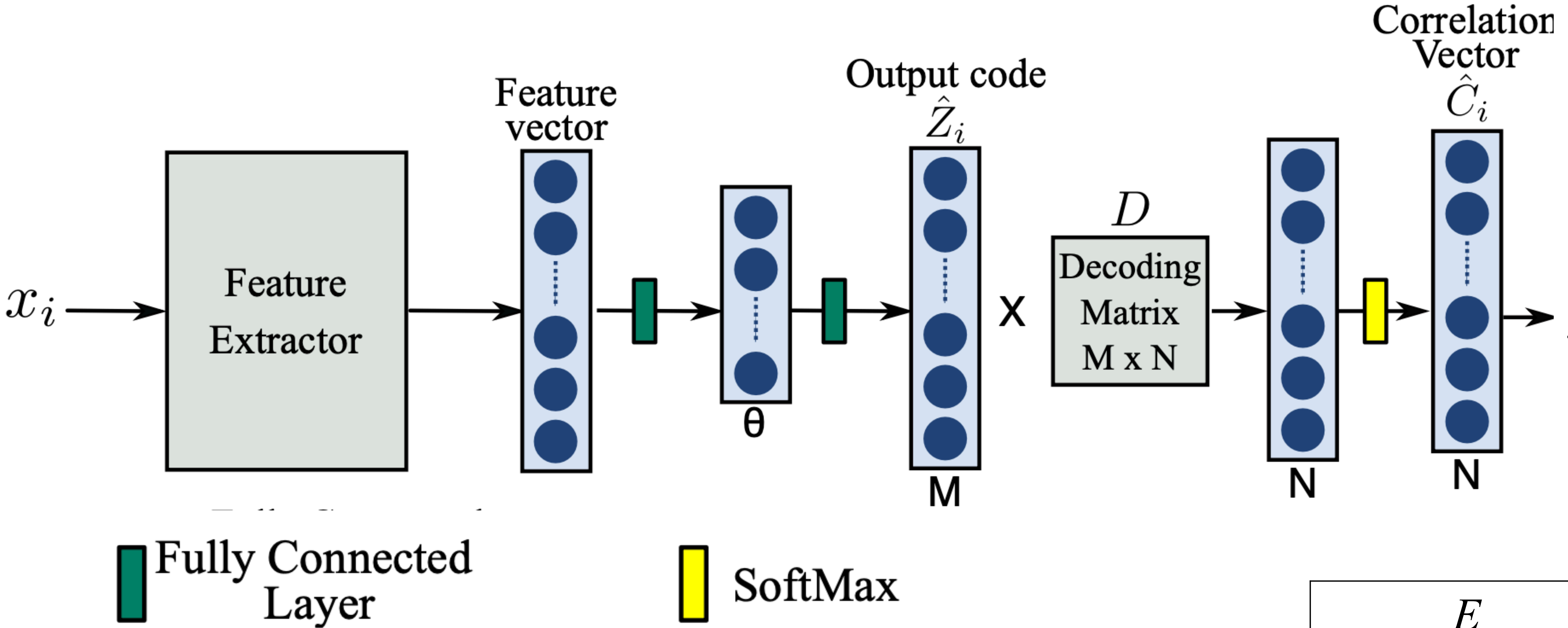
Loss = Cross entropy (Decoded output, Target label)



$$E_{k,:} = \frac{1}{|\mathbb{S}_k|} \sum_{i \subset \mathbb{S}_k} \hat{Z}_i$$

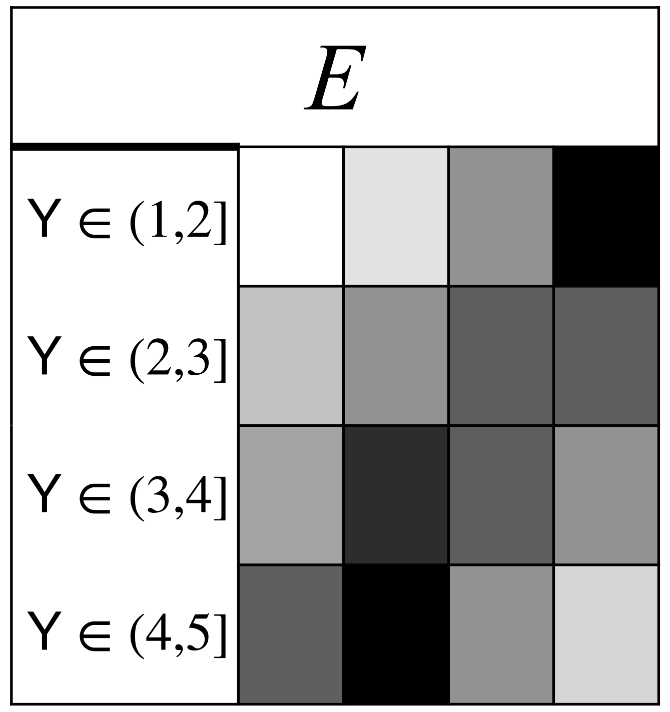
Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

Regularized Label Encoding Learning



$$\hat{y}_i = \underset{k \in \{1,2,\ldots,N\}}{\text{argmax}} \hat{C}_i^k$$

Loss = Cross entropy (Decoded output, Target label)
+ Regularizer (E)

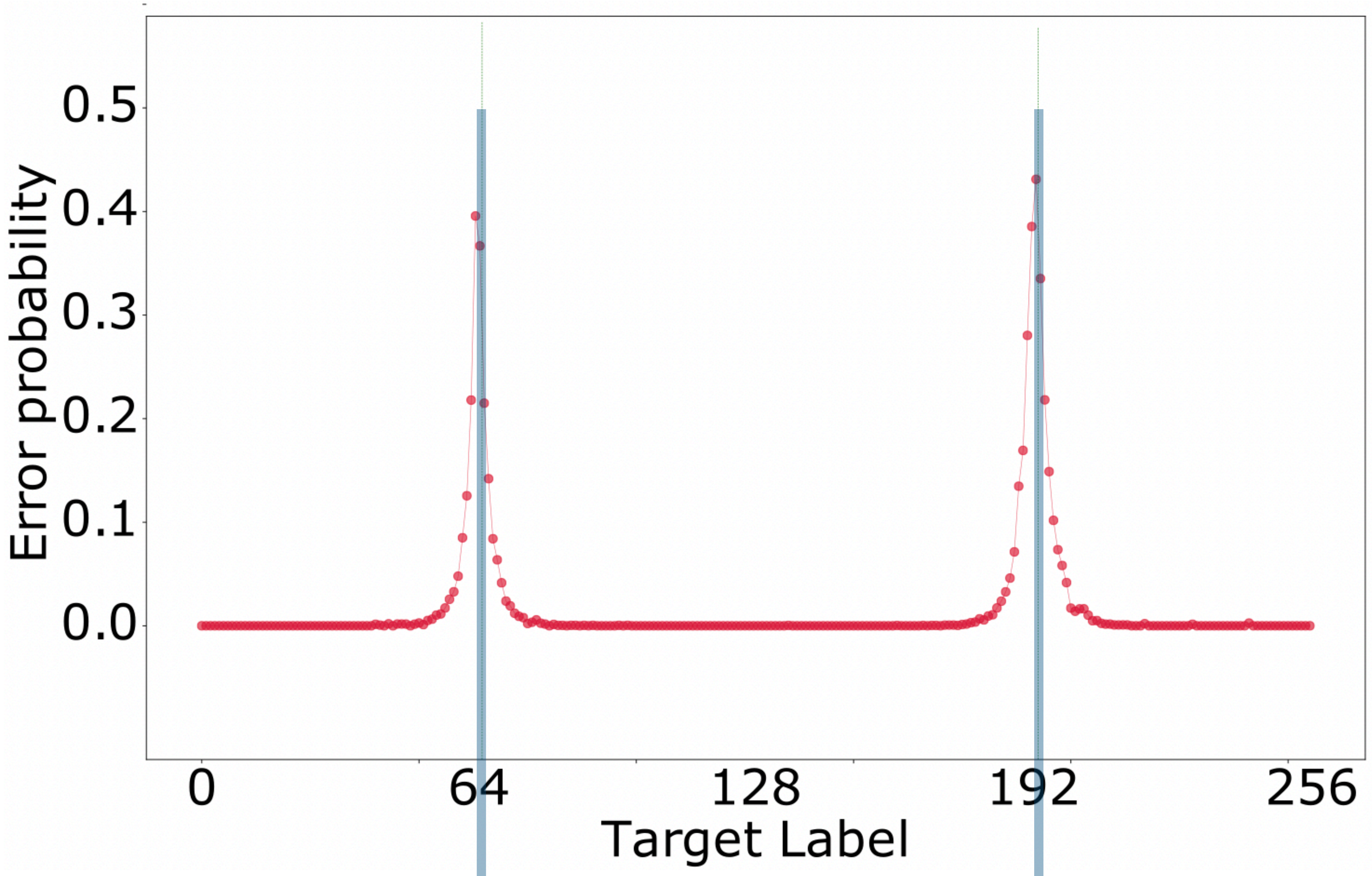


$$E_{n,:} = \frac{1}{|\mathbb{S}_n|} \sum_{i \subset \mathbb{S}_n} \hat{Z}_i$$

Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

Regularized Label Encoding Learning

Design Properties of label encoding

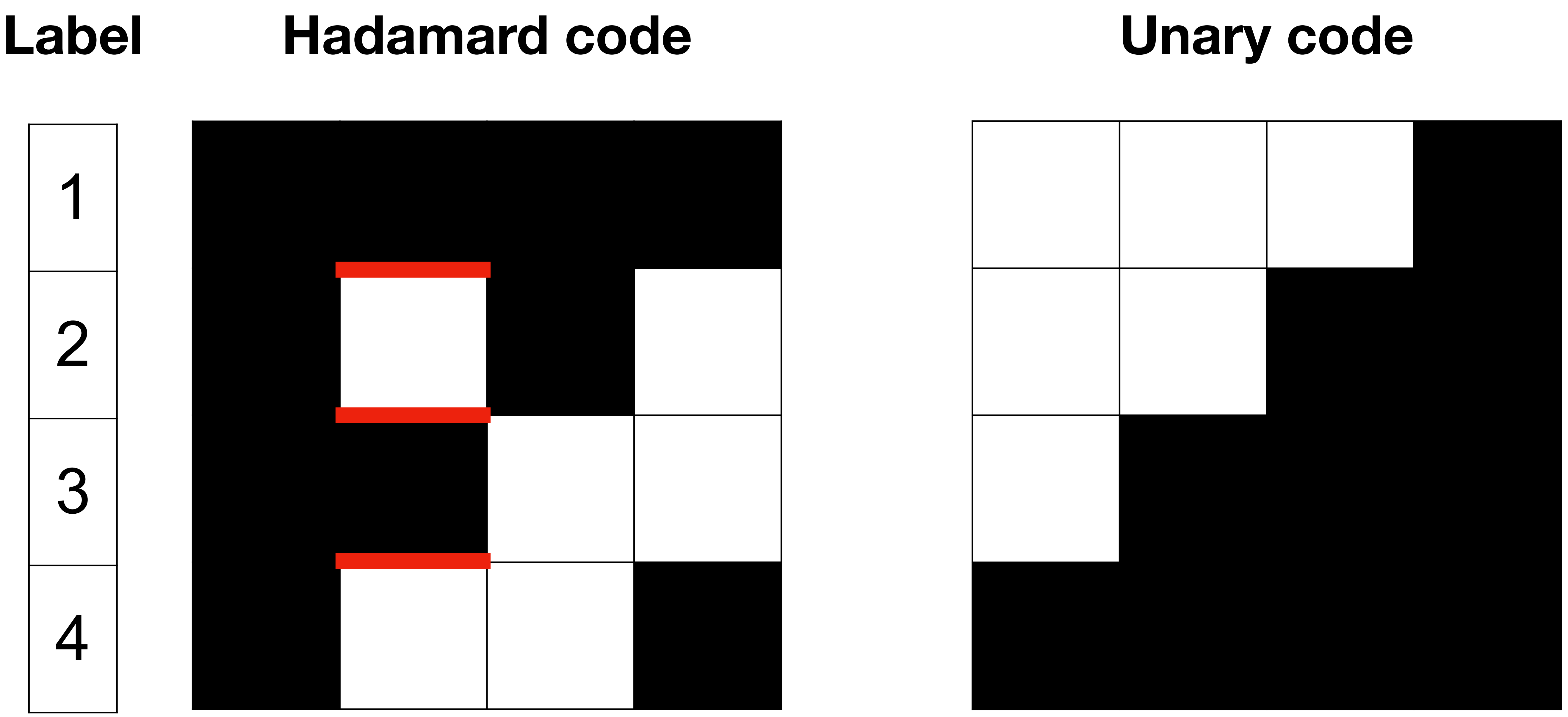


Target label	0	…	63	64	65	…	…	190	191	192	…	255
Classifier’s target	0	0	0	1	1	1	1	1	1	0	0	0

Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

Regularized Label Encoding Learning

Design Properties of label encoding



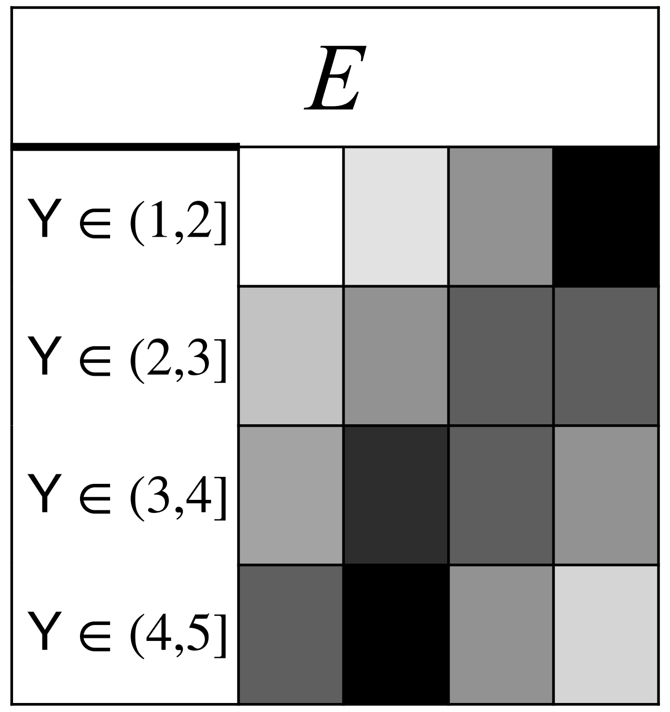
Hamming Distance	2	1	Error-correction capability
Classification Error	~50%	~2%	Bit-error probability

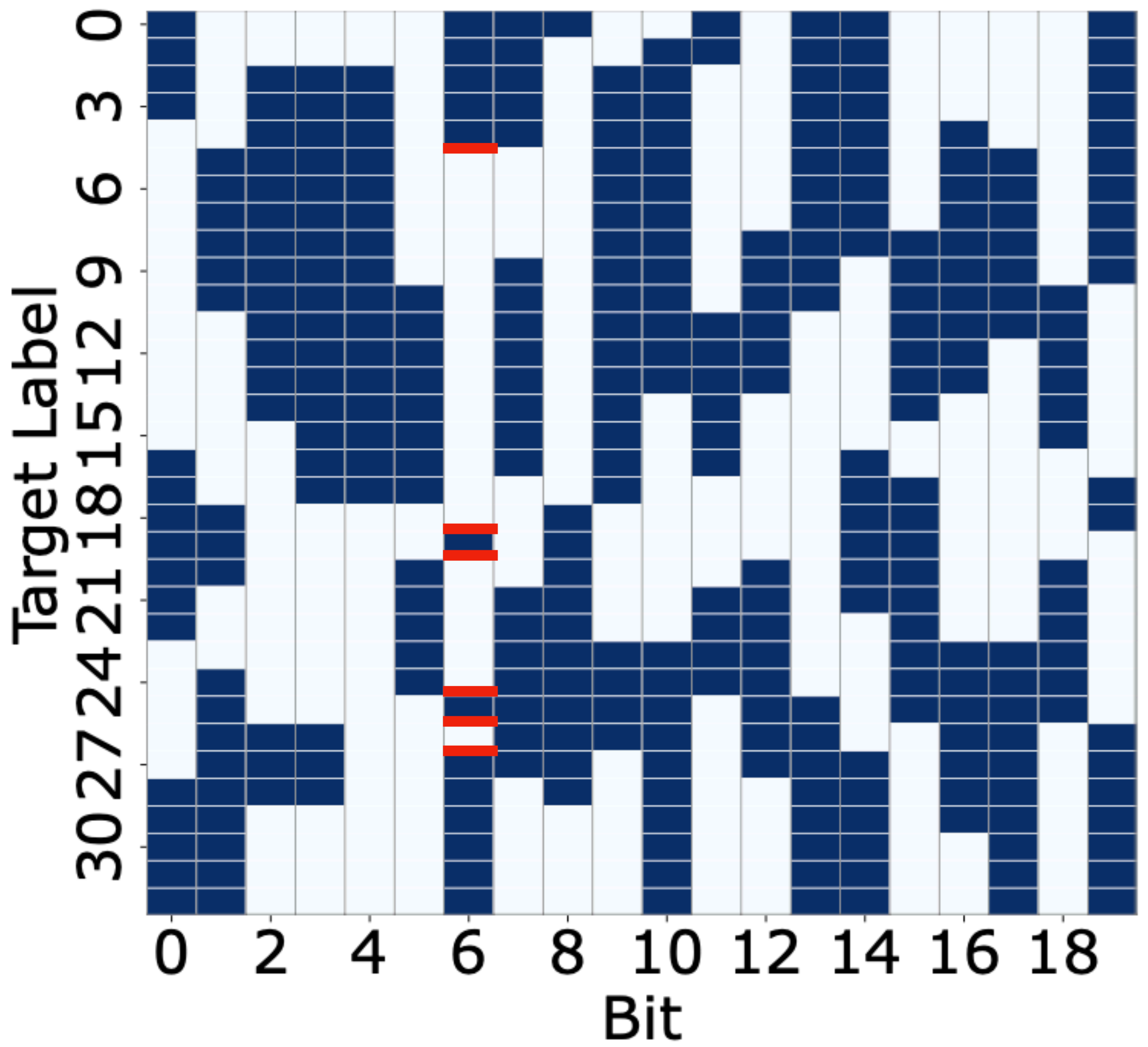
Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

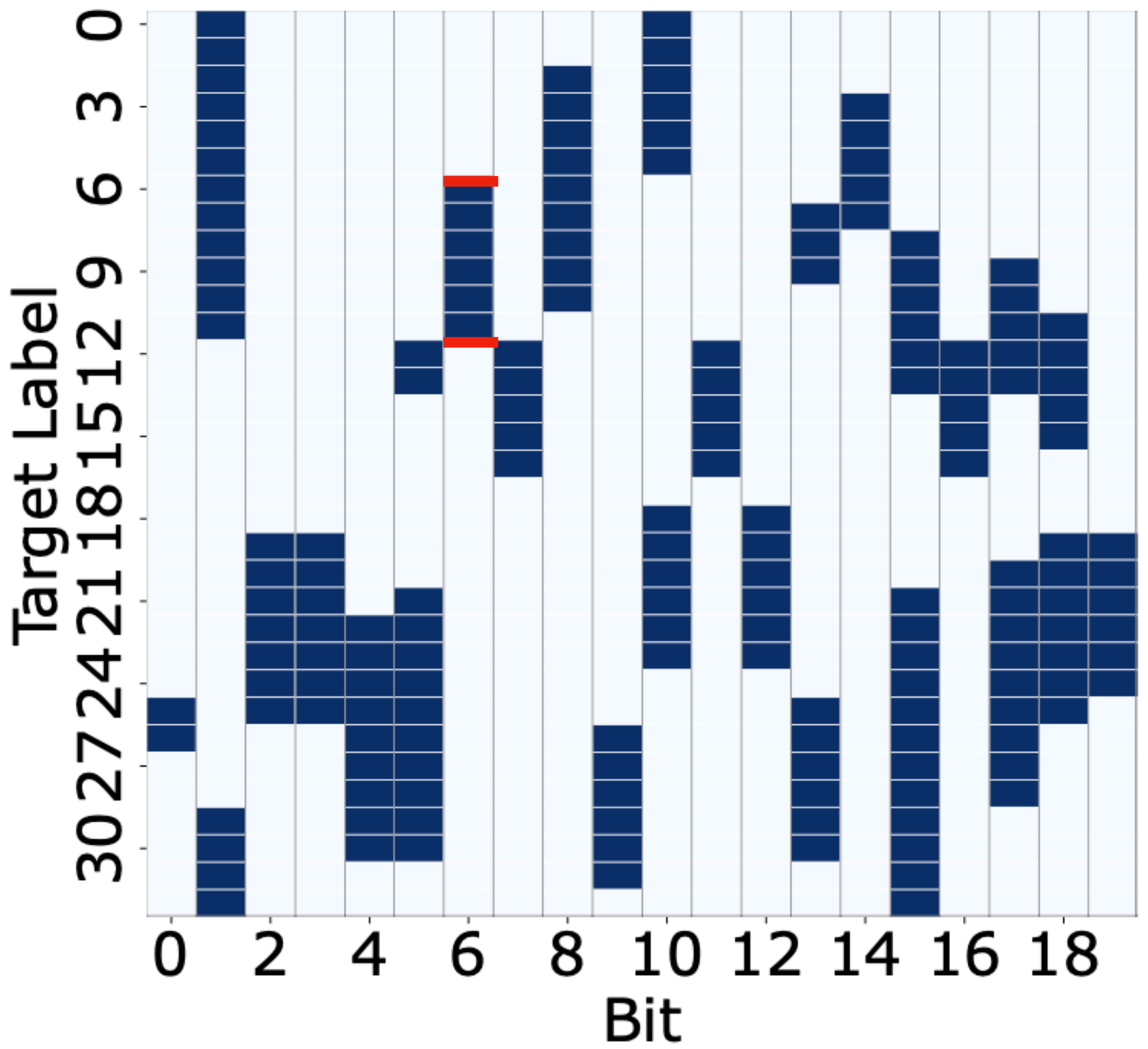
Regularized Label Encoding Learning

- Regularizing number of bit transitions

$$\sum_{j=1}^{M} \sum_{i=1}^{N-1} |E_{i,j} - E_{i+1,j}|$$







Regression Error	Without regularizer: 4.13	With regularizer: 3.73
Average #bit transition	Without regularizer: 3.9	With regularizer: 2.5

Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

Regularized Label Encoding Learning

Design Properties of label encoding

| Predicted value - Target value |

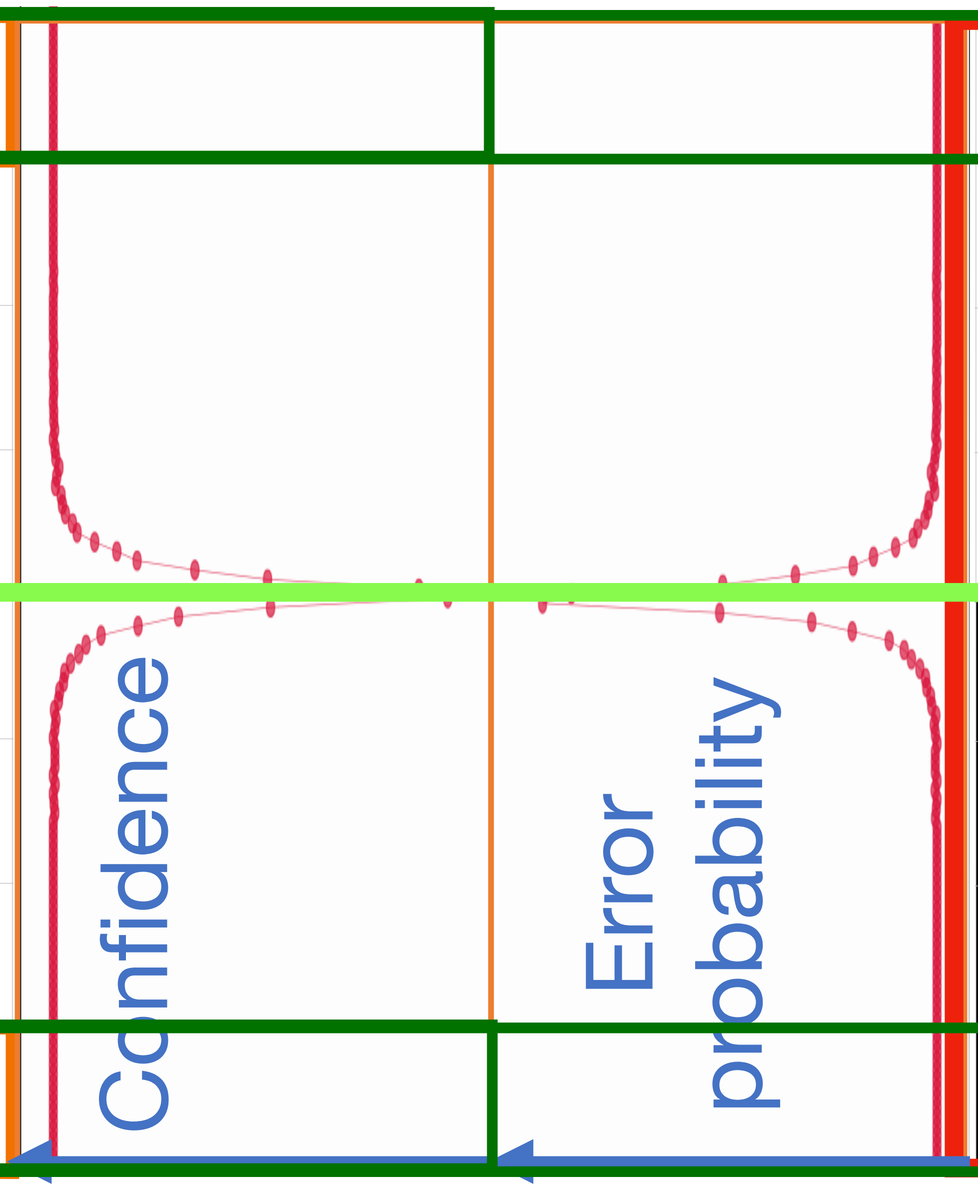
Label	Unary Code							
1	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	1	1
3	0	0	0	0	0	1	1	1
4	0	0	0	0	1	1	1	1
5	0	0	0	1	1	1	1	1
6	0	0	1	1	1	1	1	1
7	0	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1

2

7

Hamming Distance

Label	Johnson Code			
1	0	0	0	0
2	0	0	0	1
3	0	0	1	1
4	0	1	1	1
5	1	1	1	1
6	1	1	1	0
7	1	1	0	0
8	1	0	0	0

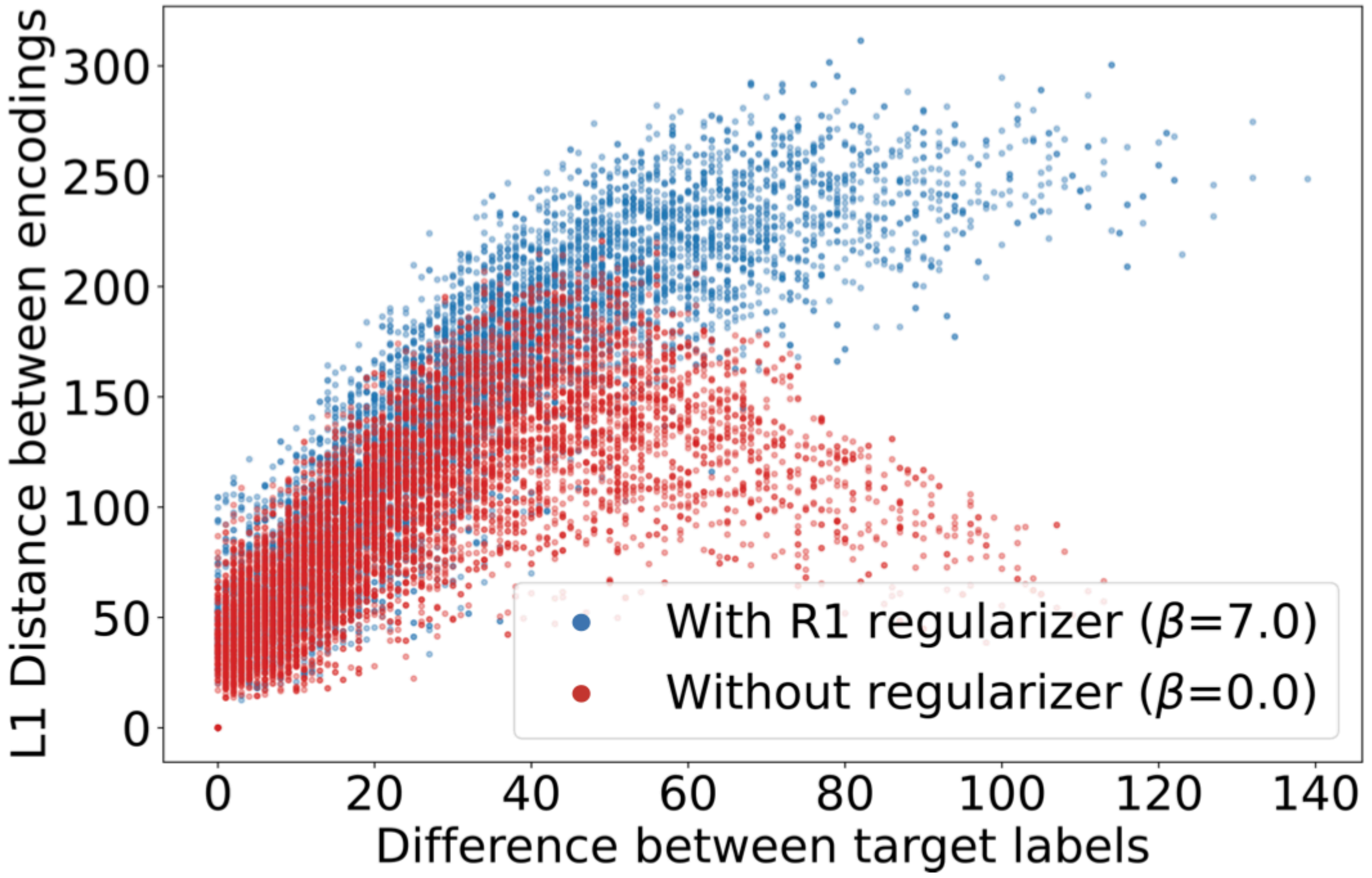


Deval Shah and Tor M. Aamodt, Learning Label Encodings for Deep Regression

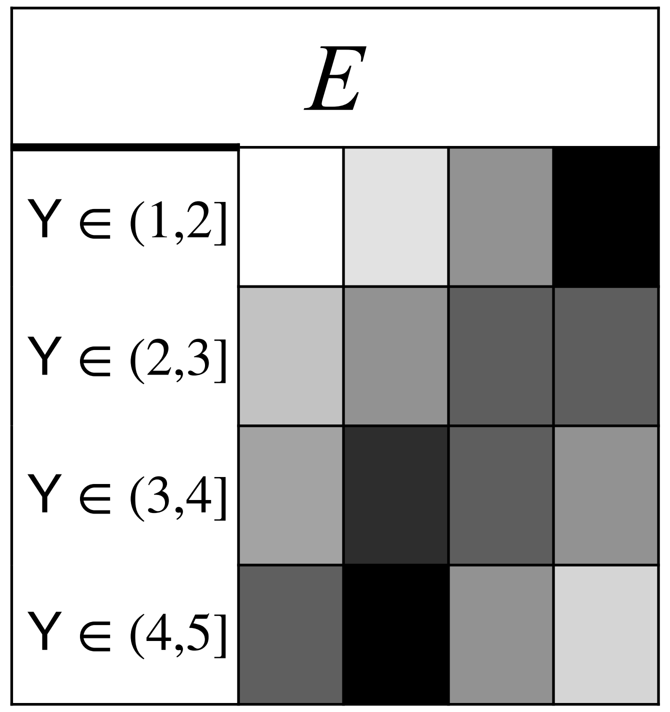
Regularized Label Encoding Learning

- Regularizing the distance between label encodings

$$||E_{i,:} - E_{j,:}||_1 \propto |i - j|$$



Regression Error:
Without regularization: 4.98
With regularization: 4.71

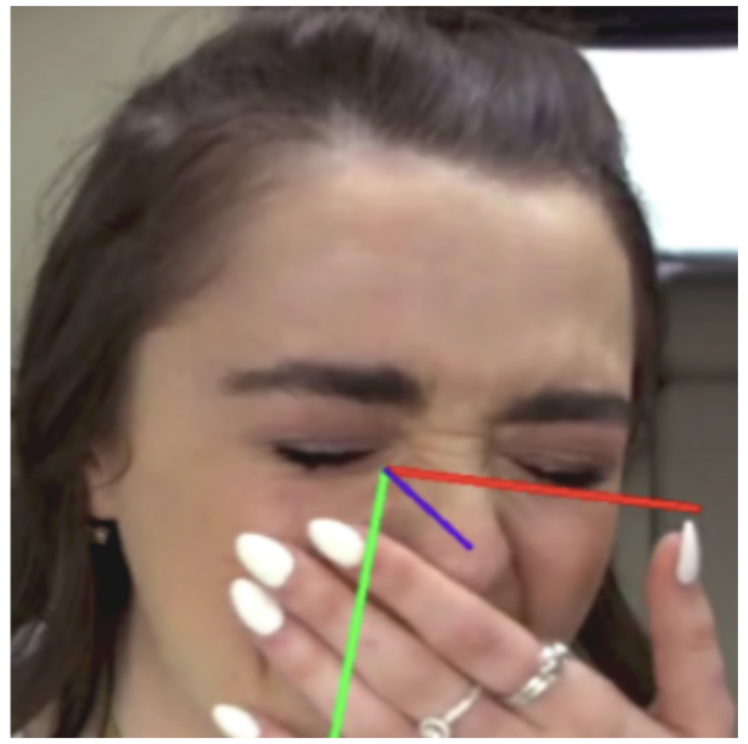


$R(E)$

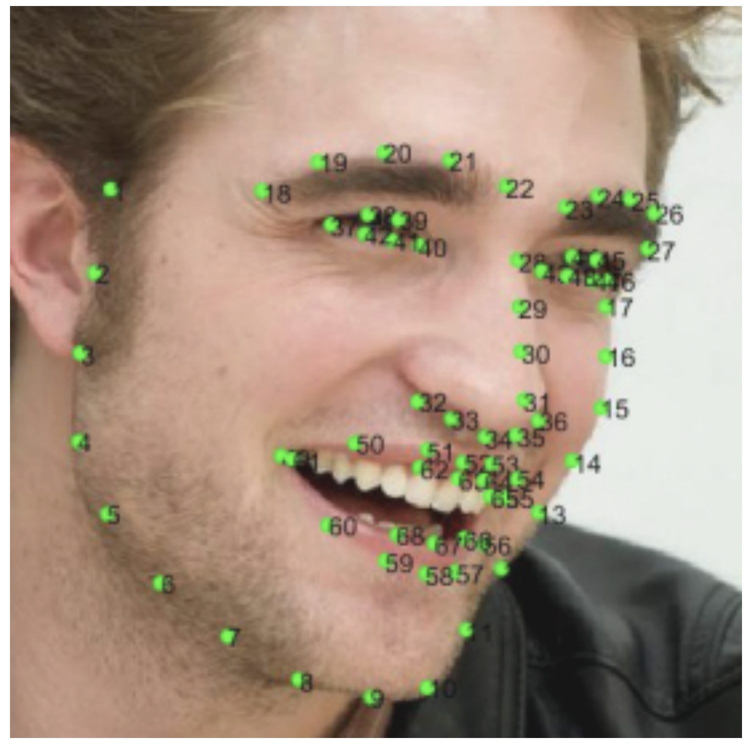
$$E_{n,:} = \frac{1}{|\mathbb{S}_n|} \sum_{i \subset \mathbb{S}_n} \hat{Z}_i$$

$R(\hat{Z}_i)$

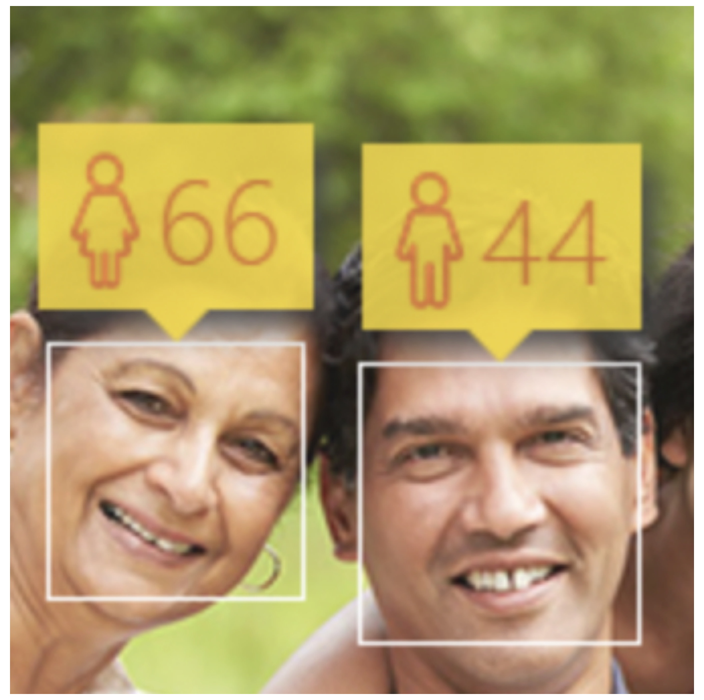
Evaluation



Head pose estimation



Facial landmark detection



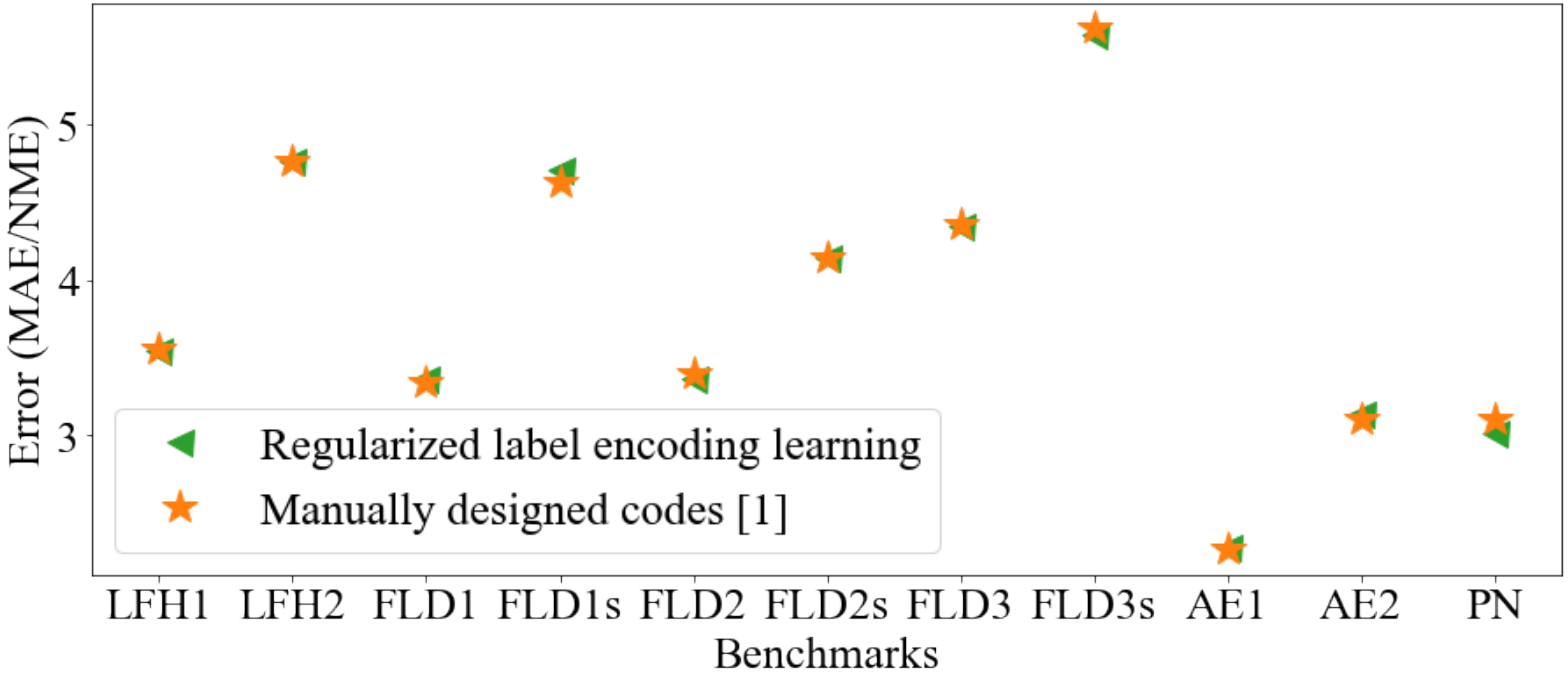
Age estimation

End-to-end autonomous driving

Image sources:
[1] https://github.com/natanielruiz/deep-head-pose
[2] D. Merget, M. Rock and G. Rigoll, "Robust Facial Landmark Detection via a Fully-Convolutional Local-Global Context Network," *CVPR* 2018
[3] https://techxplore.com/news/2015-05-microsoft-age-estimate-tool-unleashed-real-time.html
[4] https://www.idtechex.com/en/research-article/how-close-are-we-to-autonomous-cars/19191

Evaluation

Comparison of Label Encoding Design Approaches

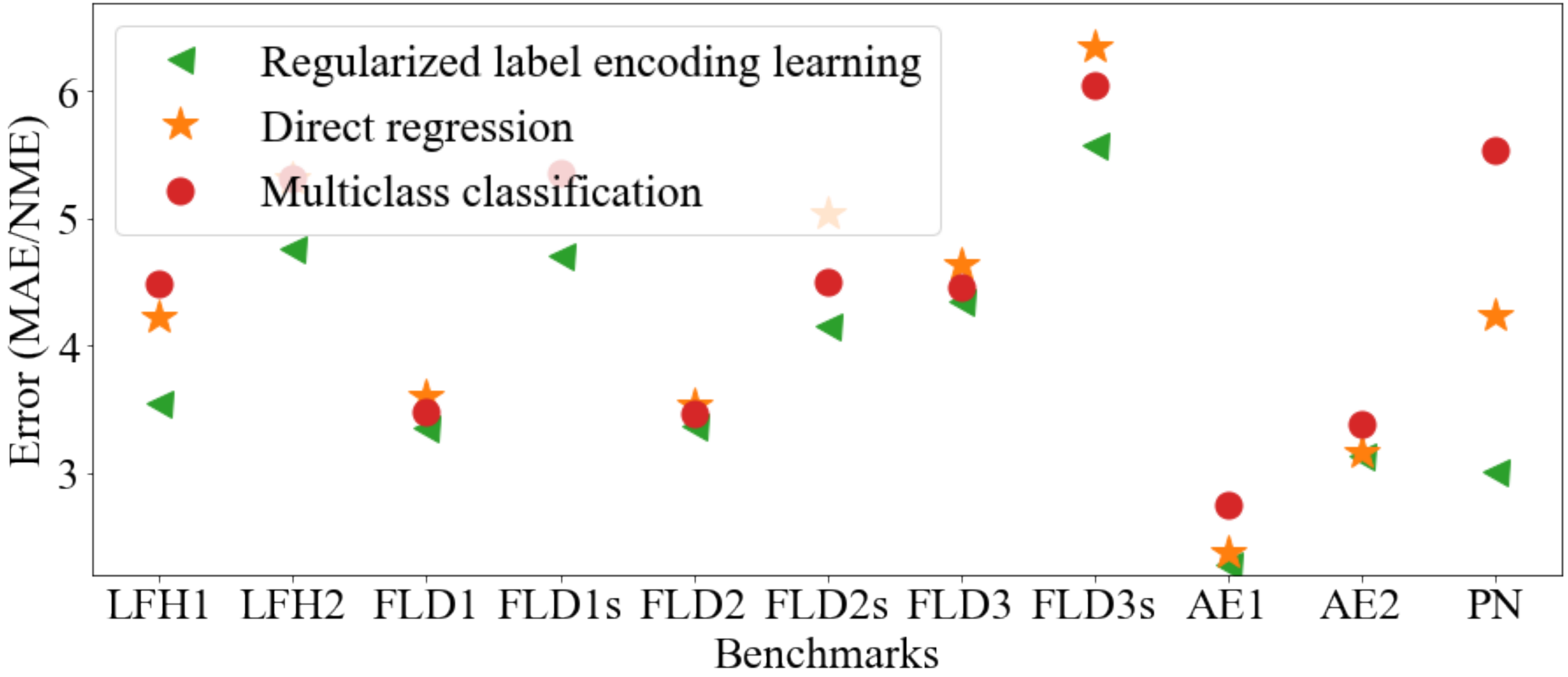


Outperforms simulated annealing and autoencoder based approaches

[1] Shah et al., Label Encoding for Regression Networks, ICLR 2022

Evaluation

Comparison with Regression Approaches



12.7% and 8.3% lower error than direct regression and multiclass classification, respectively

Check our paper for more information!

https://github.com/ubc-aamodt-group/RLEL_regression