ICLR

北京大学 健康医疗大数据国家研究院
NATIONAL INSTITUTE OF HEALTH DATA SCIENCE AT PEKING UNIVERSITY

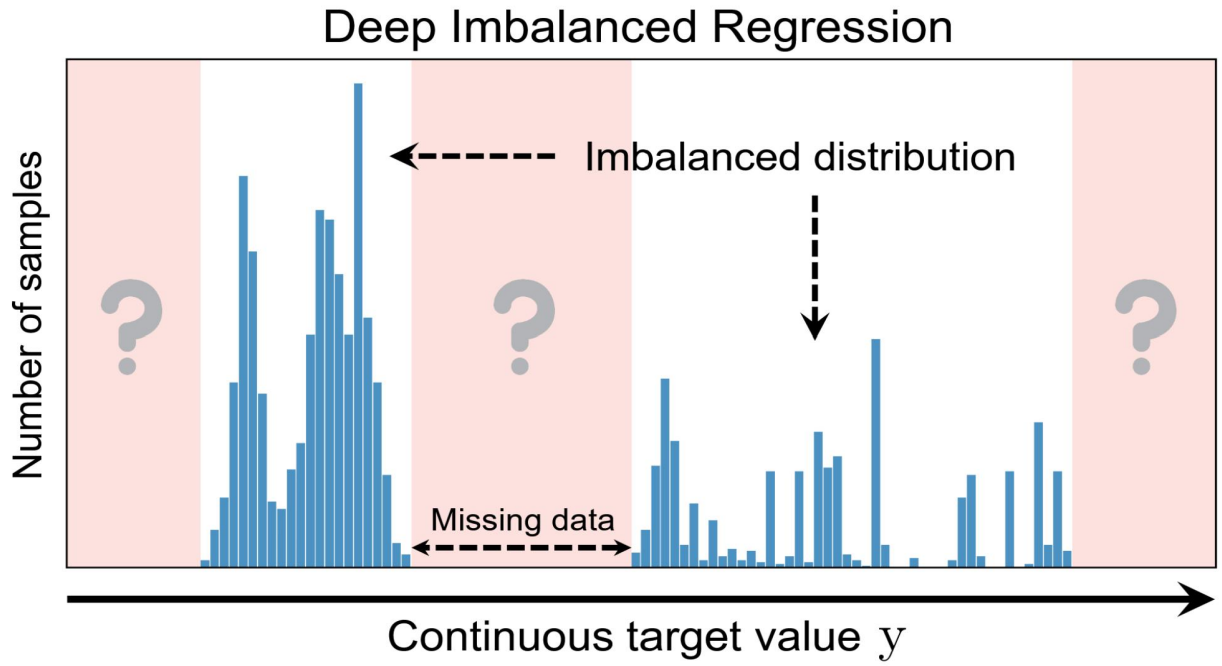
Dist Loss: Enhancing Regression in Few-Shot Region Through Distribution Distance Constraint

Guangkun Nie, Gongzheng Tang, Shenda Hong

Peking University

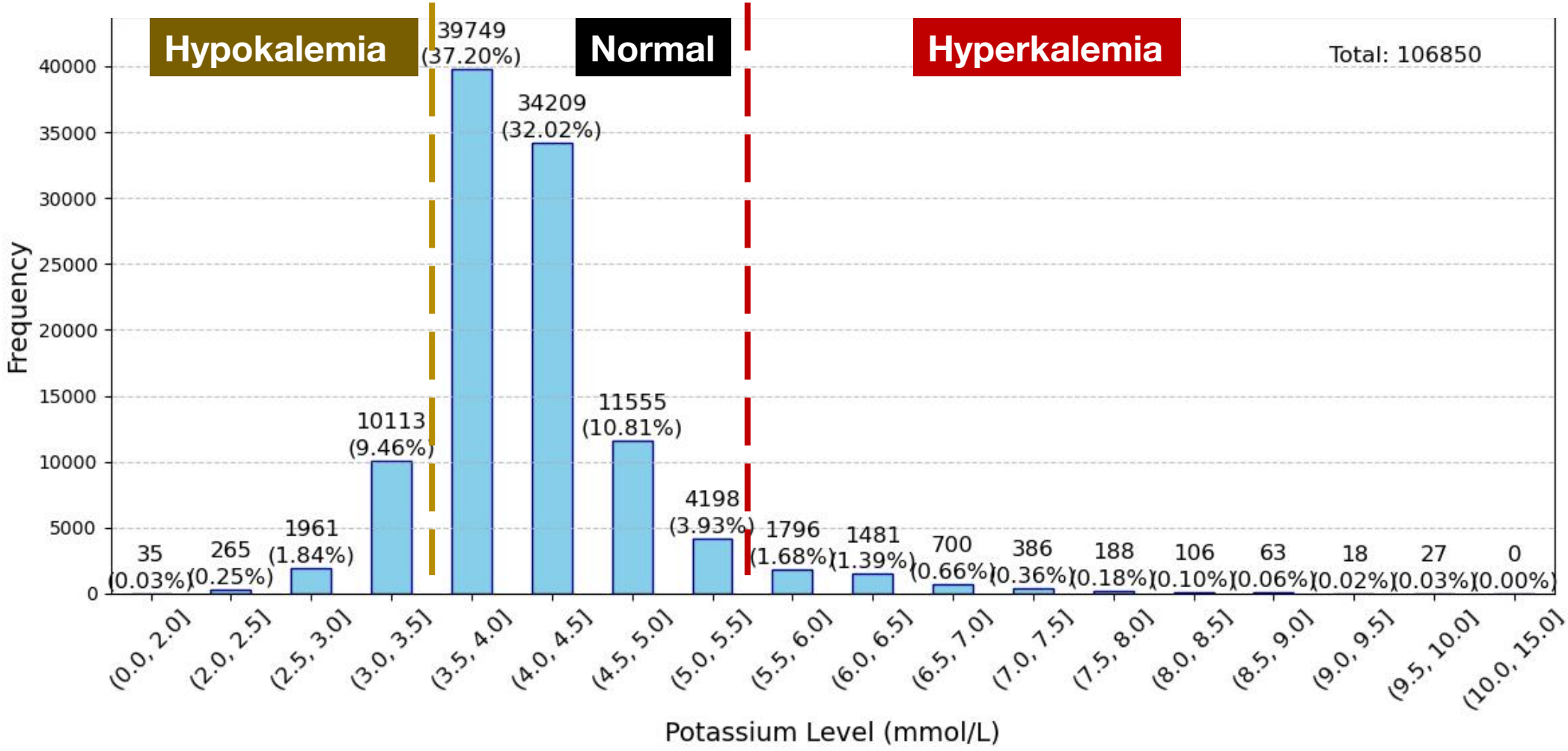
Definition of deep imbalanced regression (DIR)

Learning from imbalanced data with continuous targets, tackling potential missing data in certain regions, and generalizing to the entire target range.



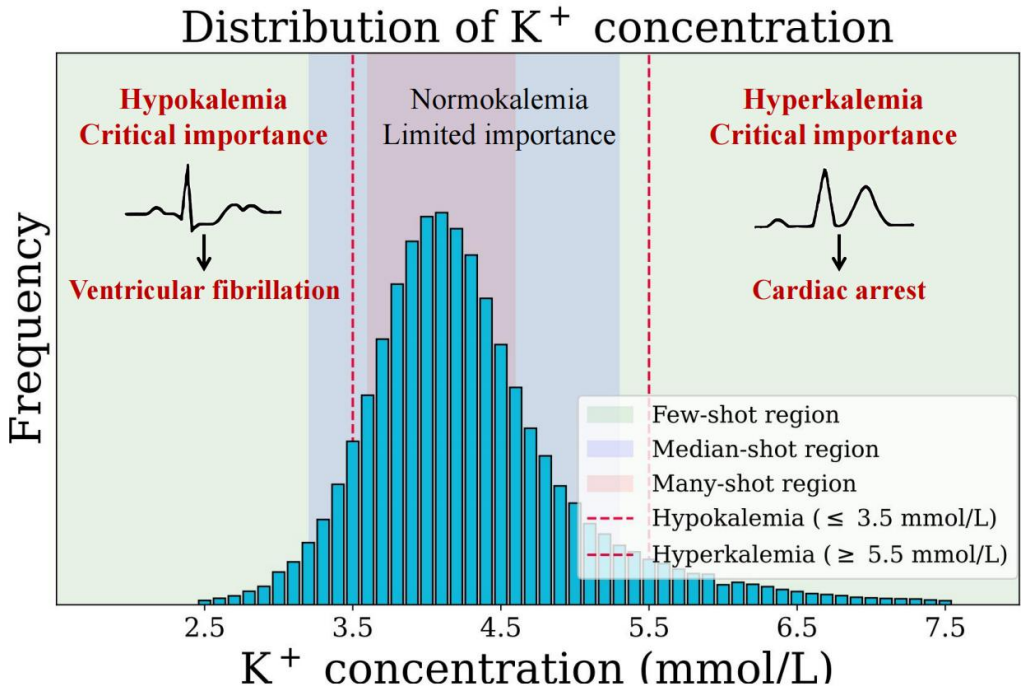
Yang, Yuzhe, et al. "Delving into deep imbalanced regression." International conference on machine learning. PMLR, 2021.

Real-world case study: DIR of blood potassium (K+) levels



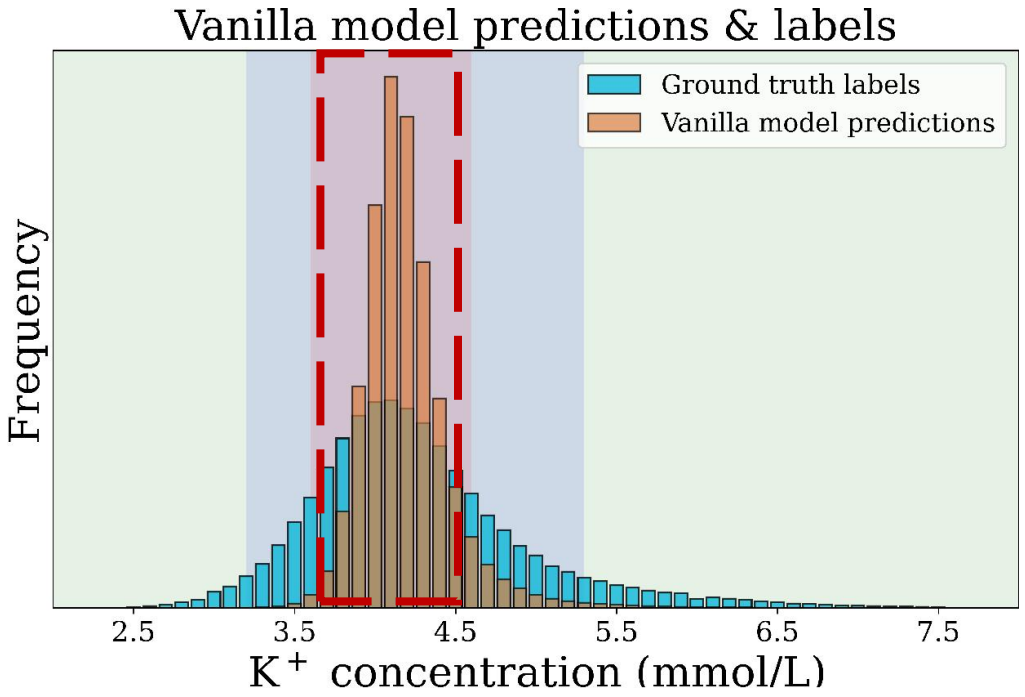
Motivation of Dist Loss

Significant distribution discrepancy between the labels and predictions



L1 Loss

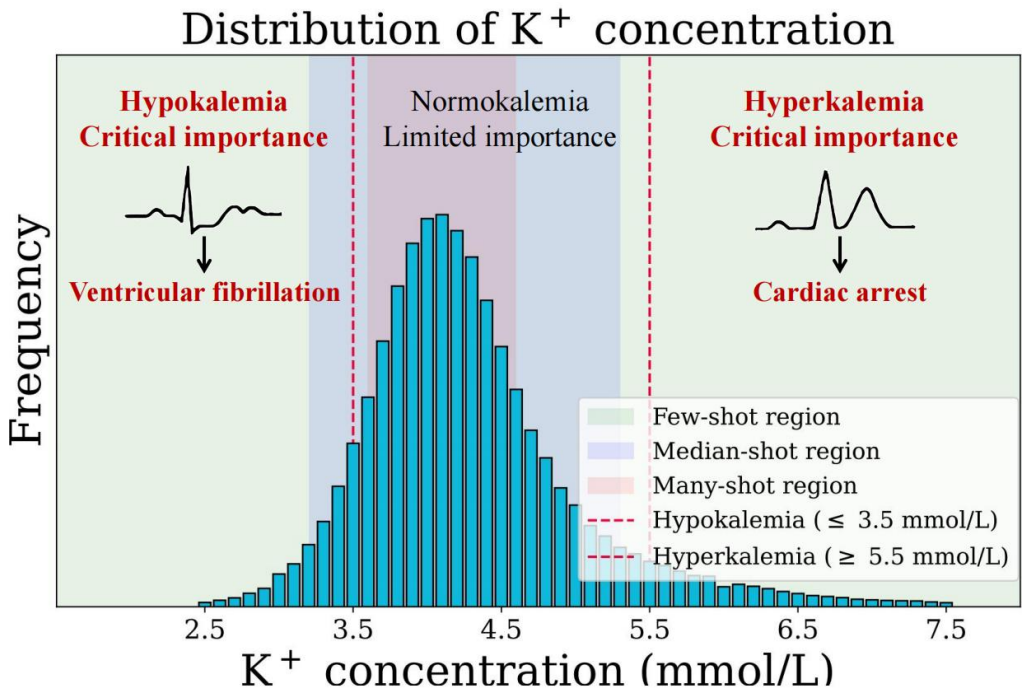
$$loss(x, y) = \frac{1}{n}\sum_{i=1}^{n}|y_i - f(x_i)|$$



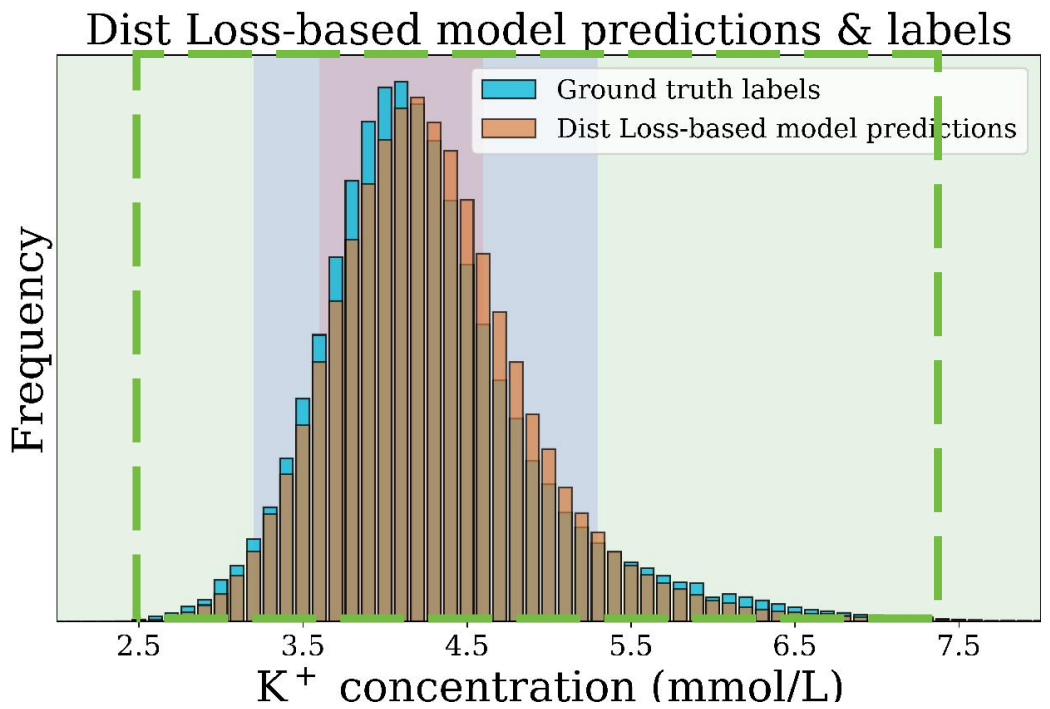
Mitigating such distribution descrepancy between the labels and predictions

Performance of Dist Loss

A much smaller distribution discrepancy between the labels and predictions



Dist Loss



Implementation of Dist Loss

Dist Loss = Sample-level Loss + Distribution-level loss

Sample-level loss: precise sample-level prediction

Distribution-level loss: regularizes model outputs to prevent few-shot samples (edges) from shifting toward many-shot regions (center)

Implementation of Dist Loss

Dist Loss = Sample-level Loss + Distribution-level loss

Sample-level loss: precise sample-level prediction

Can be any loss designed to reduce sample-level prediction error, such as L1 and L2 loss.

Implementation of Dist Loss

Dist Loss = Sample-level Loss + Distribution-level loss

Distribution-level loss: regularizes model outputs to prevent few-shot samples (edges) from shifting toward many-shot regions (center)

How to achieve this? KL divergence or other methods?

Implementation of Dist Loss

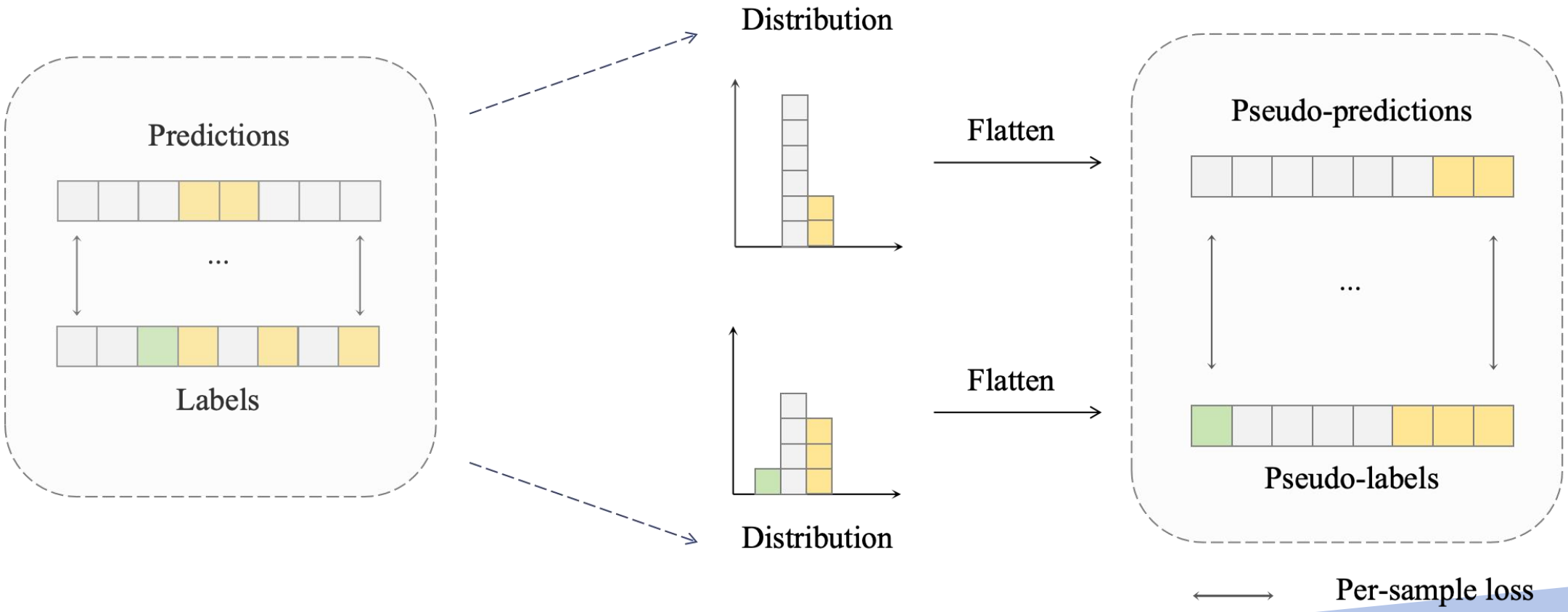
Dist Loss = Sample-level Loss + Distribution-level loss

No, not differentiable in regression

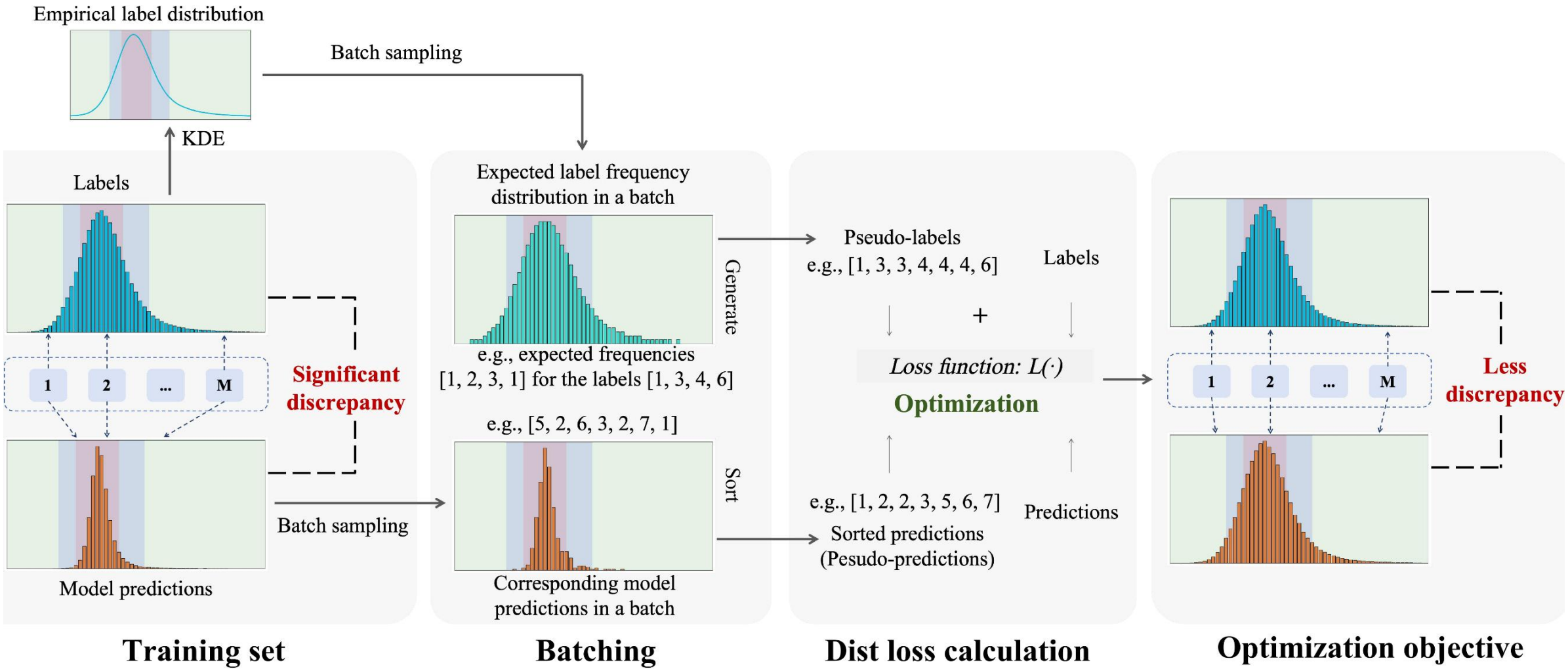
How to achieve this? KL divergence or other methods?

Implementation of Dist Loss

How to design a differentiable distribution-level loss?



Implementation of Dist Loss



Implementation of Dist Loss

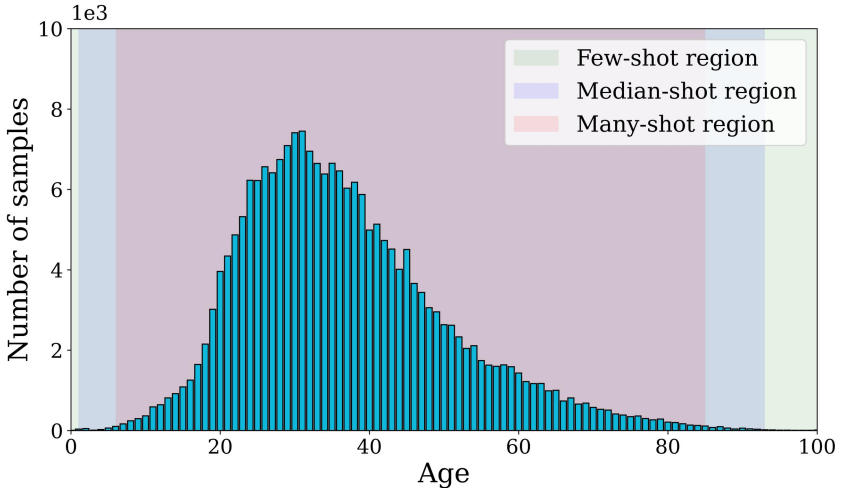
Dist Loss = Sample-level Loss + Distribution-level loss

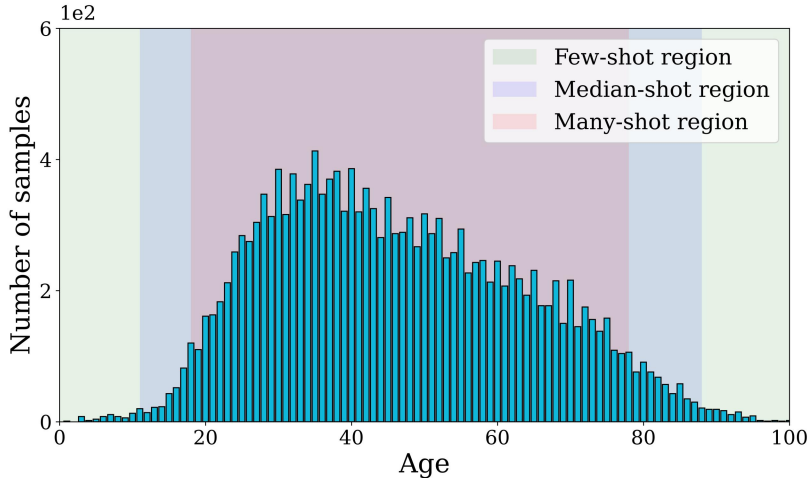
What does Dist Loss do exactly

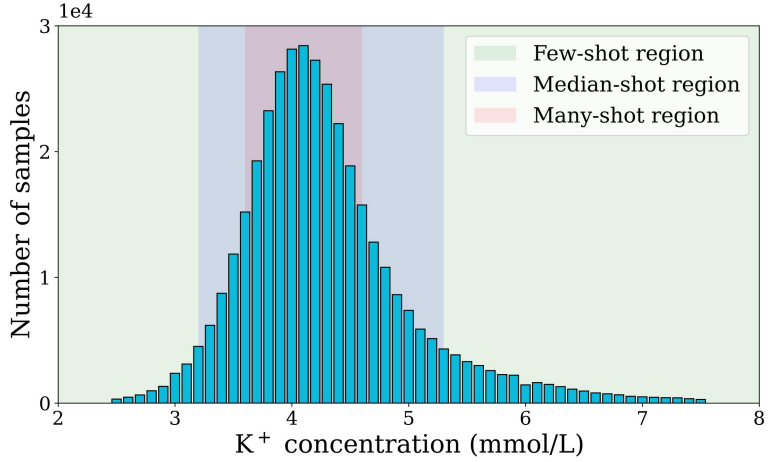
Simultaneously optimizing sample-level prediction error and distribution-level discrepancy.

Benchmarks

Both validation and test sets are balanced







IMDB-WIKI-DIR

- 191,509 training samples
- 11,022 validation samples
- 11,022 testing samples

AgeDB-DIR

- 12,208 training samples
- 2,140 validation samples
- 2,140 testing samples

ECG-K-DIR

- 365,549 training samples
- 5,098 validation samples
- 5,098 testing samples

Baselines

- LDS[1]: Label Distribution Smoothing
- FDS[1]: Feature Distribution Smoothing
- RankSim[2]: Ranking Similarity
- ConR[3]: Contrastive Regularizer
- Balanced MSE[4]: Balanced Mean Squared Error

[1]. Yang, Y., et al. (2021). Delving into deep imbalanced regression. International Conference on Machine Learning (ICML), PMLR.
[2]. Gong, Y., et al. (2022). RankSim: Ranking similarity regularization for deep imbalanced regression. International Conference on Machine Learning (ICML), PMLR.
[3]. Keramati, M., et al. (2024). ConR: Contrastive regularizer for deep imbalanced regression. International Conference on Learning Representations (ICLR).
[4]. Ren, J., et al. (2022). Balanced MSE for imbalanced visual regression. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

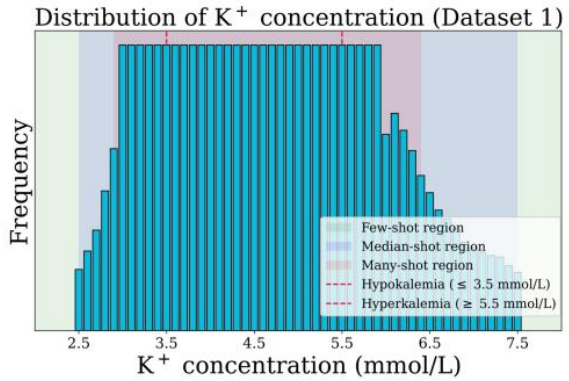
Results

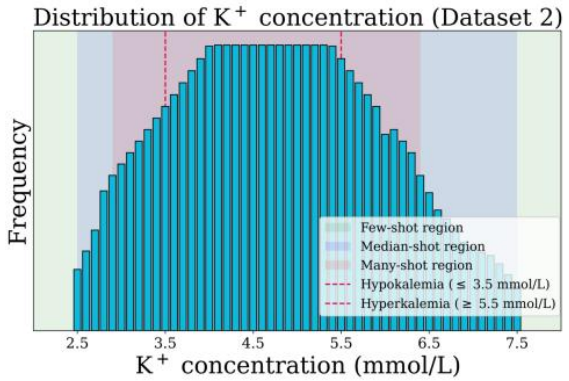
	MAE			GM		
	IMDB-WIKI-DIR	AgeDB-DIR	ECG-K-DIR	IMDB-WIKI-DIR	AgeDB-DIR	ECG-K-DIR
Vanilla	26.930	12.894	1.771	21.254	9.789	1.578
+ LDS	22.753	11.279	1.510	12.803	7.846	1.190
+ FDS	24.908	11.161	1.737	14.361	7.361	1.529
+ Ranksim	25.999	12.569	1.791	19.690	9.495	1.600
+ ConR	25.408	12.623	1.756	17.022	8.787	1.556
+ Balanced MSE	23.542	9.613	1.417	**12.603**	6.248	1.046
+ **Dist Loss (Ours)**	**22.550**	**9.122**	**1.329**	14.288	**5.453**	**0.978**
Ours vs. Vanilla	**+ 4.380**	**+ 3.772**	**+ 0.442**	**+ 6.966**	**+ 4.336**	**+ 0.600**
Ours vs. LDS	**+ 0.203**	**+ 2.157**	**+ 0.181**	**- 1.485**	**+ 2.393**	**+ 0.212**
Ours vs. FDS	**+ 2.358**	**+ 2.039**	**+ 0.408**	**+ 0.073**	**+ 1.908**	**+ 0.551**
Ours vs. Ranksim	**+ 3.449**	**+ 3.447**	**+ 0.462**	**+ 5.402**	**+ 4.042**	**+ 0.622**
Ours vs. ConR	**+ 2.858**	**+ 3.501**	**+ 0.427**	**+ 2.734**	**+ 3.334**	**+ 0.578**
Ours vs. Balanced MSE	**+ 0.992**	**+ 0.491**	**+ 0.088**	**- 1.685**	**+ 0.795**	**+ 0.068**

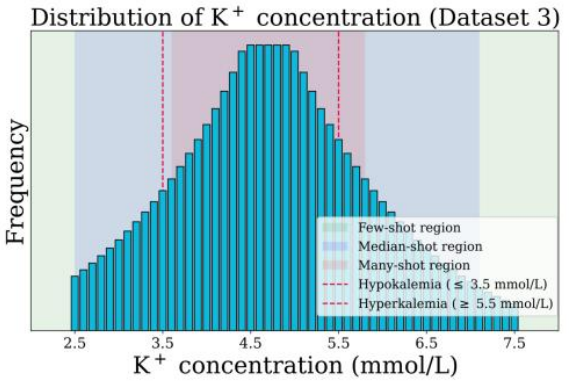
Results

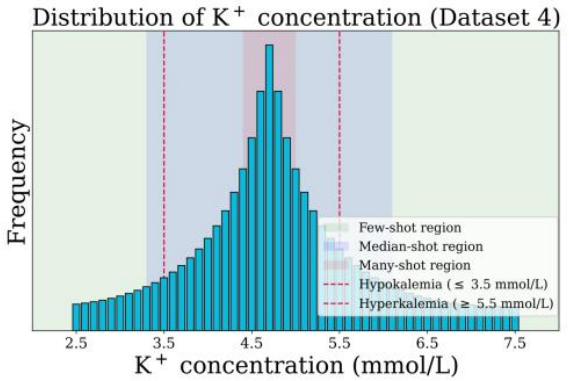
	MAE			GM		
	IMDB-WIKI-DIR	AgeDB-DIR	ECG-K-DIR	IMDB-WIKI-DIR	AgeDB-DIR	ECG-K-DIR
+ LDS	22.753	11.279	1.510	**12.803**	7.846	1.190
+ LDS + **Dist Loss**	**22.331**	**10.437**	**1.325**	13.021	**7.051**	**0.957**
+ FDS	24.908	11.161	1.737	**14.361**	7.361	1.529
+ FDS + **Dist Loss**	**24.112**	**10.444**	**1.428**	14.929	**6.696**	**1.099**
+ Ranksim	25.999	12.569	1.791	19.690	9.495	1.600
+ Ranksim + **Dist Loss**	**23.772**	**12.102**	**1.325**	**15.422**	**8.515**	**0.970**
+ ConR	25.408	12.623	1.756	17.022	**8.787**	1.556
+ ConR + **Dist Loss**	**22.700**	**12.303**	**1.336**	**14.713**	9.123	**0.987**
+ Balanced MSE	23.542	9.613	1.417	**12.603**	6.248	1.046
+ Balanced MSE + **Dist Loss**	**22.597**	**9.110**	**1.357**	14.238	**5.585**	**0.996**

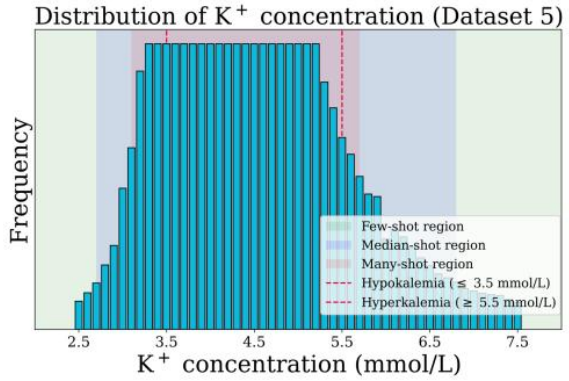
Results

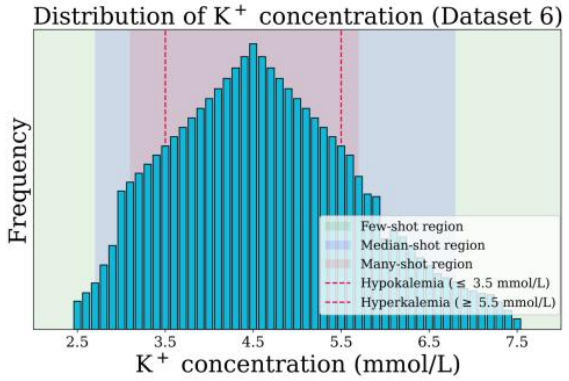

Distribution of K+ concentration (Dataset 1)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)

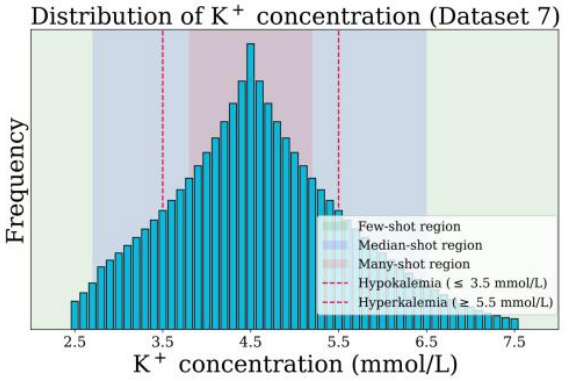

Distribution of K+ concentration (Dataset 2)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)

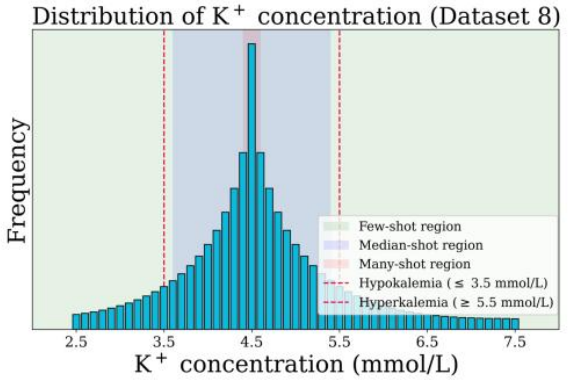

Distribution of K+ concentration (Dataset 3)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)


Distribution of K+ concentration (Dataset 4)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)


Distribution of K+ concentration (Dataset 5)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)


Distribution of K+ concentration (Dataset 6)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)


Distribution of K+ concentration (Dataset 7)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)


Distribution of K+ concentration (Dataset 8)
Frequency
Few-shot region
Median-shot region
Many-shot region
Hypokalemia (≤ 3.5 mmol/L)
Hyperkalemia (≥ 5.5 mmol/L)
2.5
3.5
4.5
5.5
6.5
7.5
K+ concentration (mmol/L)

Results

	MAE							
Methods	Dataset 0	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7
Vanilla	2.701	2.676	2.658	1.979	2.679	2.647	2.624	1.888
LDS	2.684	2.703	2.642	1.962	2.672	2.507	2.644	1.901
FDS	**1.865**	2.368	2.191	**1.790**	2.223	2.625	1.908	1.665
Ranksim	2.470	2.327	2.273	1.831	2.314	2.192	2.258	1.725
ConR	2.461	2.343	2.308	1.828	2.193	2.274	2.255	1.742
Balanced MSE	1.997	1.984	1.981	1.831	1.906	1.863	1.815	1.708
Dist Loss	1.955	**1.873**	**1.963**	1.822	**1.852**	**1.803**	**1.730**	**1.638**

ICLR

北京大学 健康医疗大数据国家研究院
NATIONAL INSTITUTE OF HEALTH DATA SCIENCE AT PEKING UNIVERSITY

Thank you

Guangkun Nie, Gongzheng Tang, Shenda Hong

Peking University

https://github.com/Ngk03/DIR-Dist-Loss

https://arxiv.org/pdf/2411.15216