Variational Imbalanced Regression:

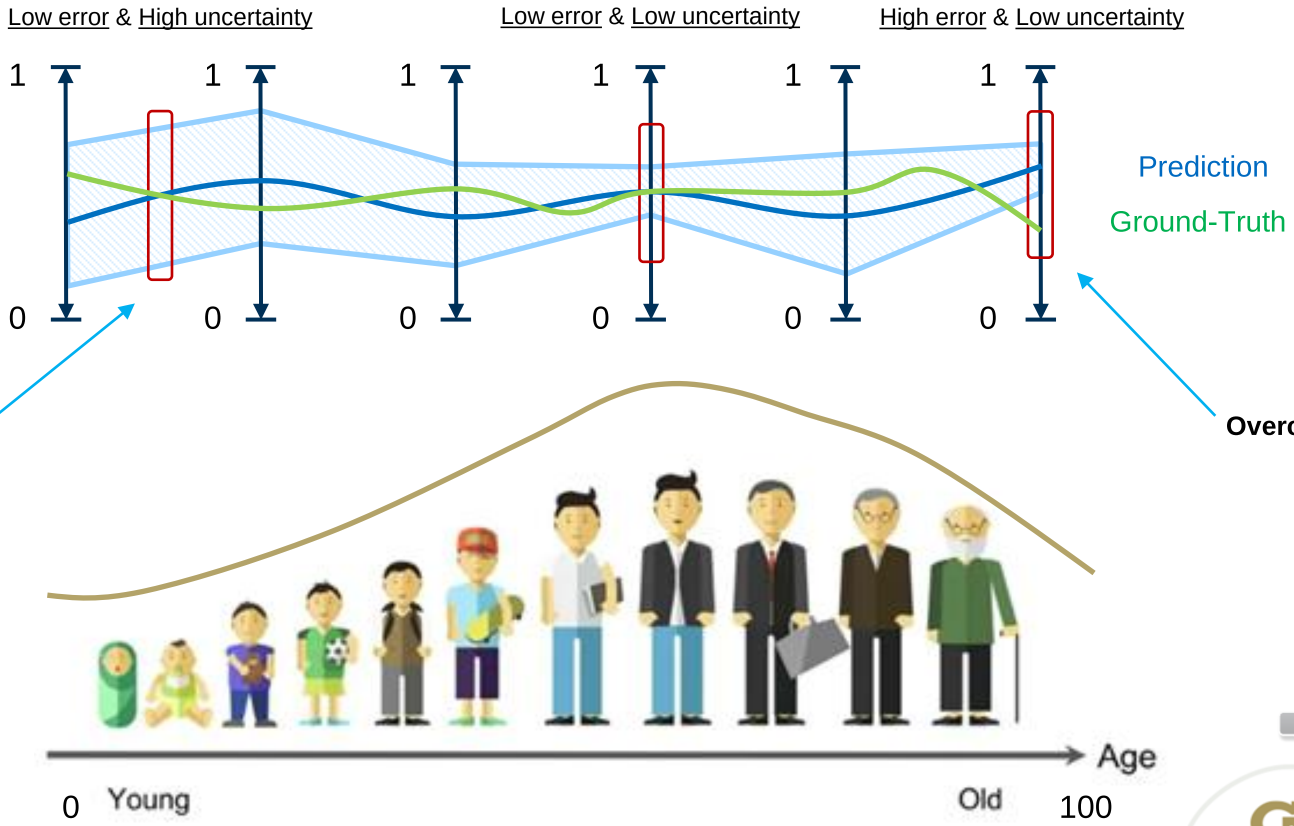
Fair Uncertainty Quantification via Probabilistic Smoothing

Ziyan Wang, Hao Wang

Motivation

Uncertainty Quantification

Underconfident



Overconfident

Age Distribution

Georgia Tech

Method (Feature Space): Constructing $q(z_i|\{x_i\}_{i=1}^{N})$

I.I.D. vs. N.I.D. (Neighboring and Identically Distributed)

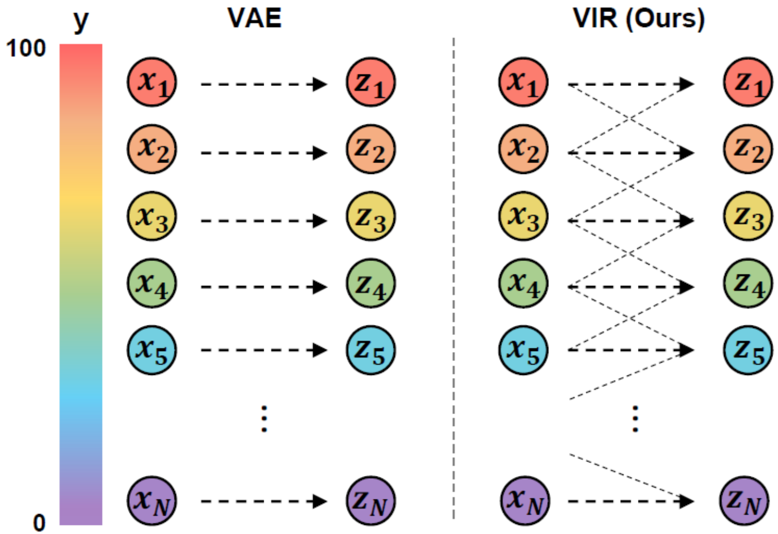
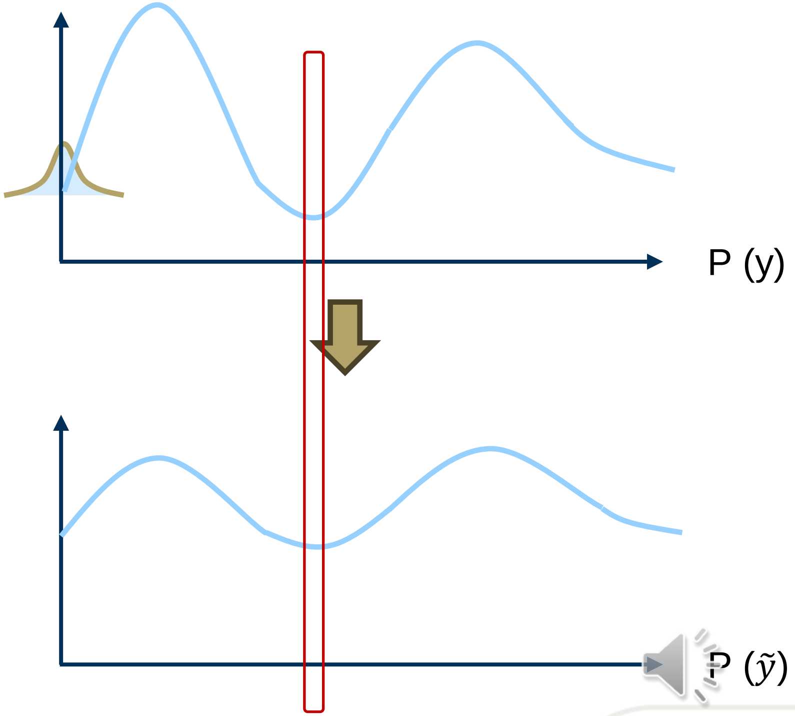
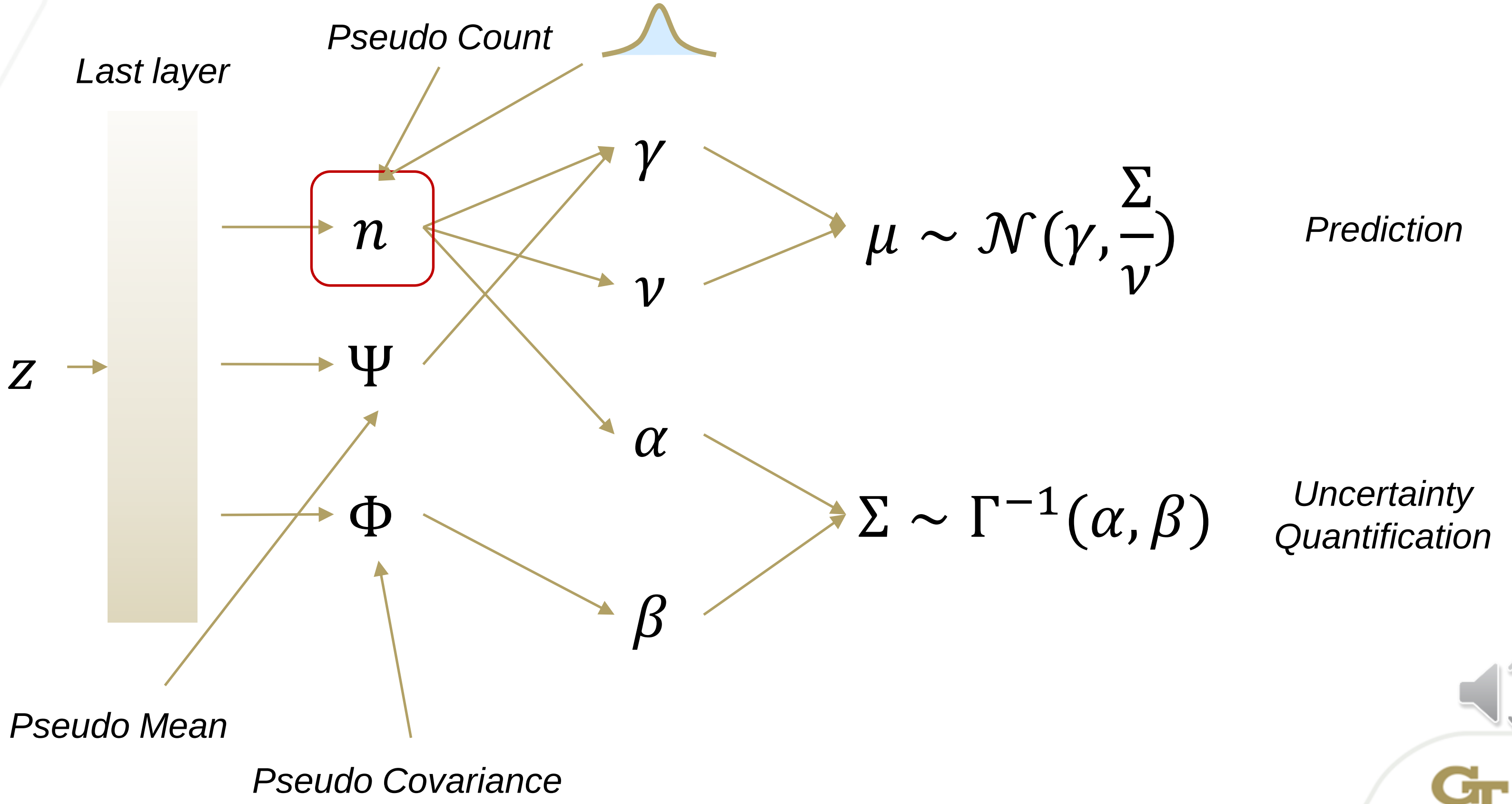


Figure 1: Comparing inference networks of typical VAE [25] and our VIR. In VAE (**left**), a data point's latent representation (i.e., **z**) is affected only by itself, while in VIR (**right**), neighbors participate to modulate the final representation.



Method (Label Space): Constructing $p(y_i|z_i)$



Theorem: Generalization Bound

Theorem 4.1 (Generalization Bound of VIR). *In imbalanced regression with bins $\mathcal{B}$, for any finite hypothesis space of predictions $\mathcal{H} = \{\widehat{Y}_1, \dots, \widehat{Y}_{\mathcal{H}}\}$, the transductive prediction error of the empirical risk minimizer $\widehat{Y}^{ERM}$ using the* $\boxed{\textit{VIR estimator with estimated propensities } \widetilde{P}\ (P_i > 0)}$ *and given training observations O from $\mathcal{Y}$ with independent Bernoulli propensities P, is bounded by:*

$$R(\widehat{Y}^{ERM}) \leq \widehat{R}_{\mathrm{VIR}}(\widehat{Y}^{ERM}|\widetilde{P}) + \underbrace{\frac{\Delta}{|\mathcal{B}|}\sum_{i=1}^{|\mathcal{B}|}\left|1 - \boxed{\frac{P_i}{\widetilde{P}_i}}\right|}_{\textit{Bias Term}} + \underbrace{\frac{\Delta}{|\mathcal{B}|}\sqrt{\frac{\log(2|\mathcal{H}|/\eta)}{2}}\sqrt{\sum_{i=1}^{|\mathcal{B}|}\boxed{\frac{1}{\widetilde{P}_i^2}}}}_{\textit{Variance Term}}. \quad (7)$$

- Increase Bias;
- **Significantly** reduce Variance in the imbalanced setting;
- Lower Generalization Error.

Results

Metrics	MSE ↓				MAE ↓				GM ↓			
Shot	All	Many	Medium	Few	All	Many	Medium	Few	All	Many	Medium	Few
VANILLA [49]	101.60	78.40	138.52	253.74	7.77	6.62	9.55	13.67	5.05	4.23	7.01	10.75
VAE [25]	99.85	78.86	130.59	223.09	7.63	6.58	9.21	13.45	4.86	4.11	6.61	10.24
DEEP ENSEMBLE [27]	100.94	79.3	129.95	249.18	7.73	6.62	9.37	13.90	4.87	4.37	6.50	11.35
INFER NOISE [29]	119.46	95[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.95	6.58	10.86
SMOTER [40]	114.34	93[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.65	5.69	8.49
SMOGN [5]	117.29	10[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.9	6.19	8.44
SQINV [49]	105.14	87[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.57	5.73	7.77
DER [1]	106.77	91[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.59	6.43	10.49
LDS [49]	102.22	83[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.39	5.8	7.45
FDS [49]	101.67	86[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.41	5.97	6.29
LDS + FDS [49]	99.46	84[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.36	5.45	6.79
RANKSIM [13]	83.51	71[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.13	5.37	6.89
LDS + FDS + DER [1]	112.62	94[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.75	6.74	7.68
VIR (OURS)	**81.76**±0.10	**70.61**[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**4.07**±0.02	**5.05**±0.03	**6.23**±0.05
OURS VS. VANILLA	+19.84	+7.79	+47.05	+111.38	+0.78	+0.23	+2.08	+4.16	+0.64	+0.16	+1.96	+4.52
OURS VS. INFER NOISE	+37.70	+24.41	+58.37	+123.93	+1.54	+1.23	+2.26	+4.31	+1.16	+0.88	+1.53	+4.63
OURS VS. DER	+25.01	+20.68	+30.96	+67.33	+1.10	+0.92	+1.52	+3.15	+0.78	+0.52	+1.38	+4.26
OURS VS. LDS + FDS	+17.70	+13.49	+20.73	+66.91	+0.56	+0.62	+0.77	+1.28	+0.31	+0.29	+0.40	+0.56
OURS VS. RANKSIM	+1.75	+1.38	+7.67	+6.69	+0.03	+0.10	+0.37	+0.17	+0.12	+0.06	+0.32	+0.66

Metrics	NLL ↓				AUSE ↓			
Shot	All	Many	Med	Few	All	Many	Med	Few
DEEP ENS. [27]	5.311	4.031	6.726	8.523	0.541	0.626	0.466	0.483
INFER NOISE [29]	4.616	4.413	4.866	5.842	0.465	0.458	0.457	0.496
DER [1]	3.918	3.741	3.919	4.432	0.523	0.464	0.449	0.486
LDS + FDS + DER [1]	3.787	3.689	3.912	4.234	0.451	0.460	0.399	0.565
VIR (OURS)	**3.703**	**3.598**	**3.805**	**4.196**	**0.434**	**0.456**	**0.324**	**0.414**
OURS VS. DER	+0.215	+0.143	+0.114	+0.236	+0.089	+0.008	+0.125	+0.072
OURS VS. LDS + FDS + DER	+0.084	+0.091	+0.107	+0.038	+0.017	+0.004	+0.075	+0.151

Georgia Tech

Conclusion

- **Identify the problem** of probabilistic deep imbalanced regression.
- **Propose VIR** to simultaneously cover two desiderata: balanced accuracy and uncertainty quantification.
- Show that VIR can achieve **better generalization bound**.

Code will be soon available at

https: //github.com/Wang-ML-Lab/variational-imbalanced-regression.

Paper

Code

Georgia Tech