Data Valuation using Reinforcement Learning

Jinsung Yoon, Sercan O. Arik, Tomas Pfister

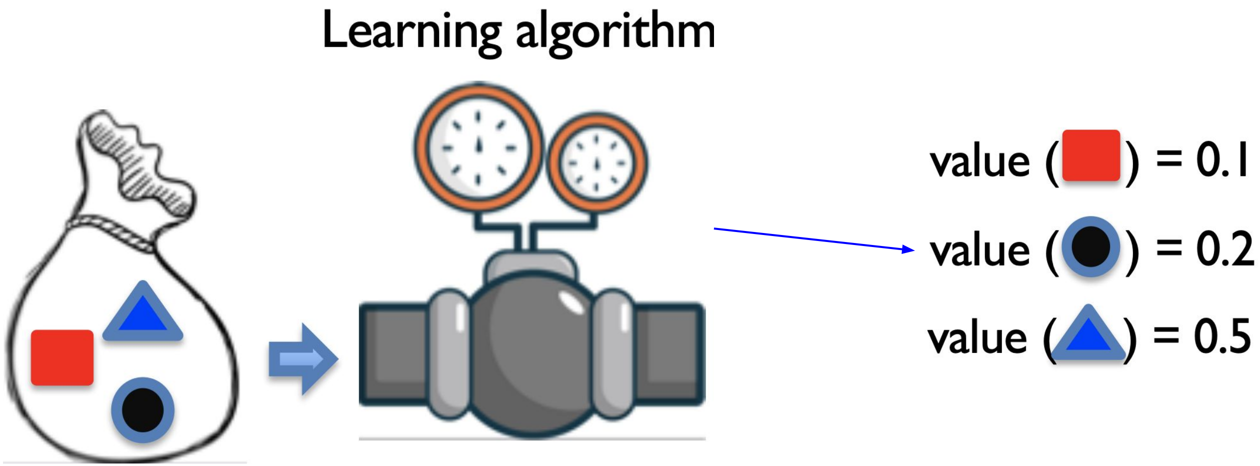
Google Cloud AI

Google Cloud

2020 International Conference on Machine Learning (ICML 2020)

Problem Definition

- What is data valuation?
  - How much does each data contribute to the trained model



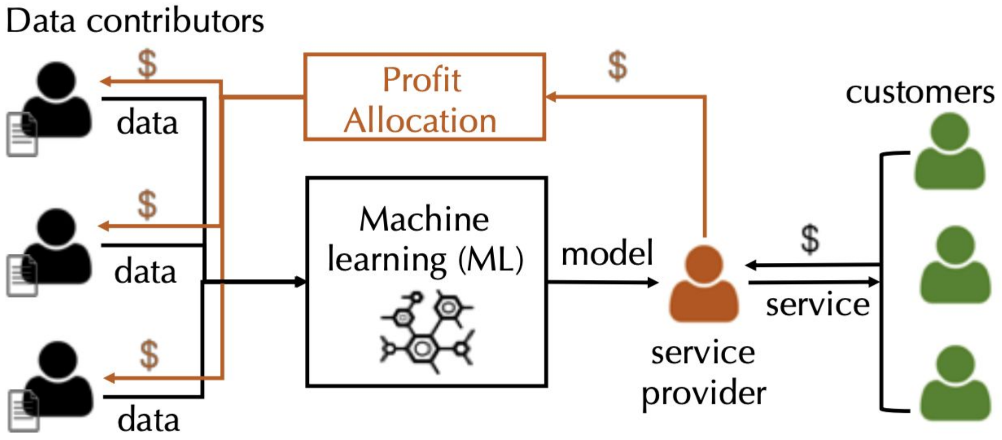
Google Cloud

Amirata Ghorbani, James Y. Zou, ***Data Shapley: Equitable Valuation of Data for Machine Learning***, *ICML*, 2019

Objective & Use-cases

- Learn in reliable way

- **Data valuation**
  - Fair valuation for the labelers and data provider
  - Insights about the dataset



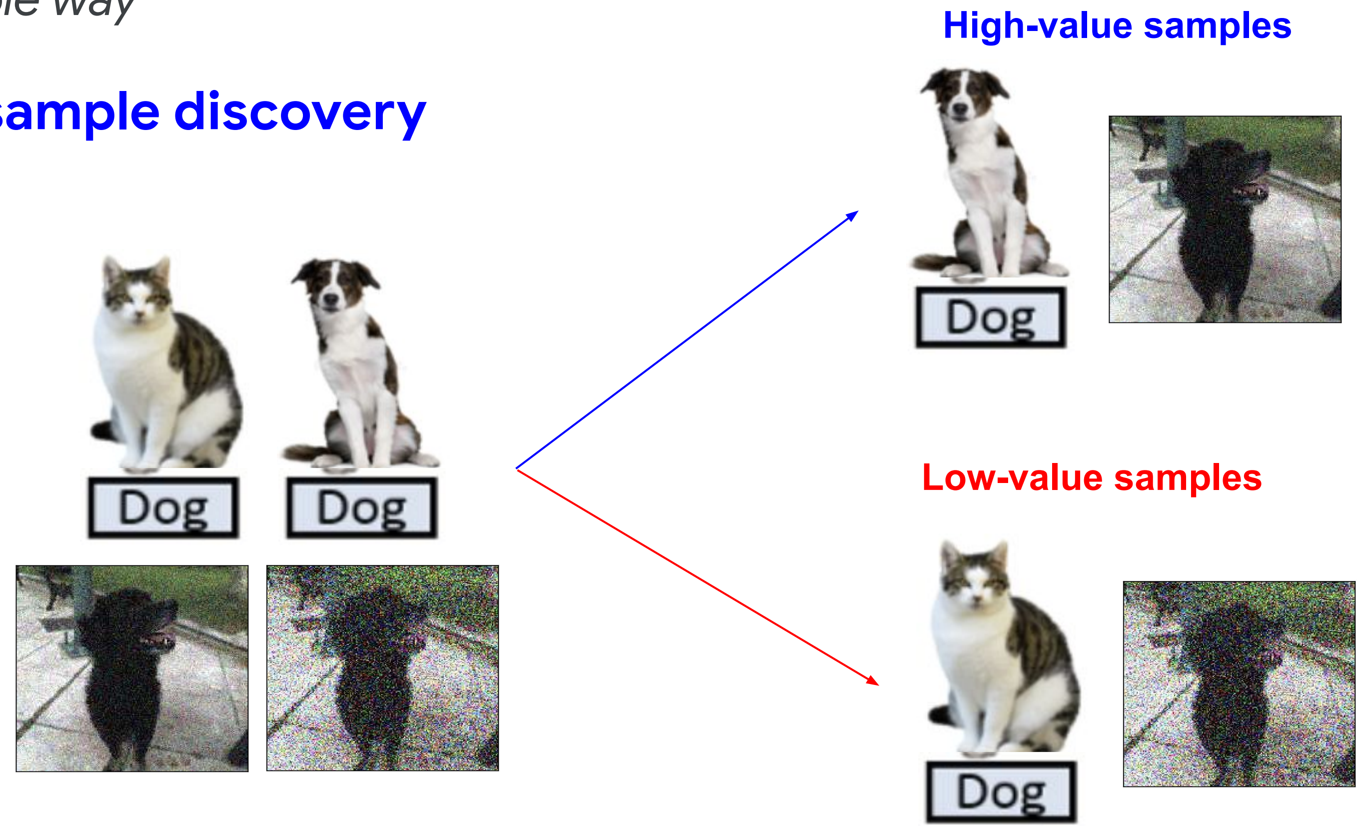
Google Cloud

Ruoxi Jia et al., ***Towards Efficient Data Valuation Based on the Shapley Value***, *AISTATS*, 2019

Objective & Use-cases

- *Learn in reliable way*

- **Corrupted sample discovery**

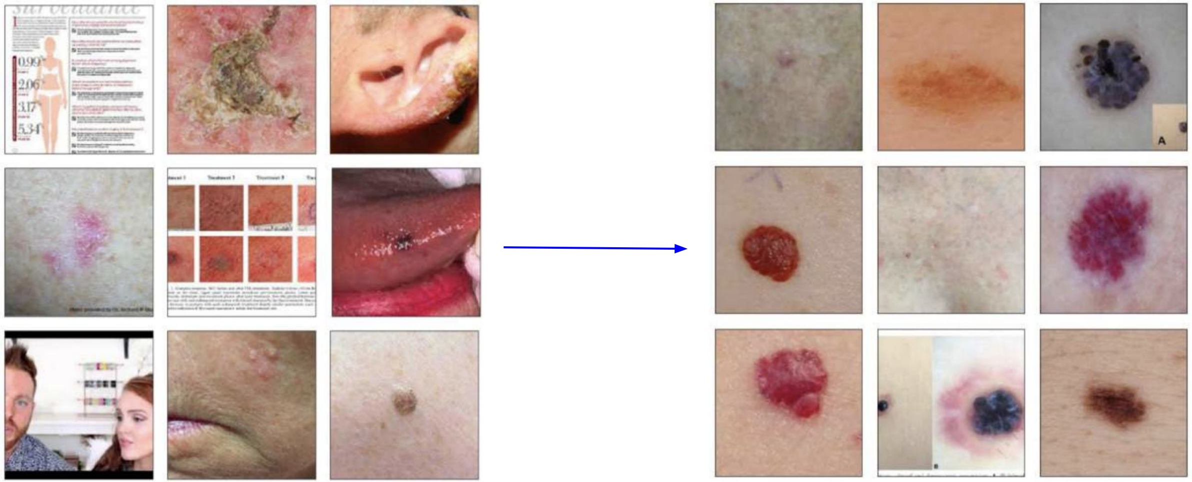


Google Cloud

Objective & Use-cases

- *Learn in reliable way*

- **Robust learning with noisy (or cheaply-acquired) datasets**
  - Augmented learning



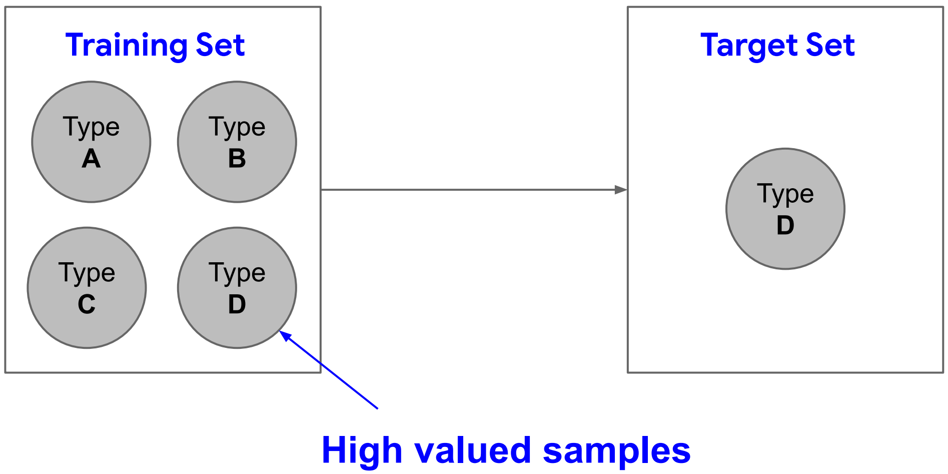
Cheaply-acquired samples **High valued samples**

Amirata Ghorbani, James Y. Zou, ***Data Shapley: Equitable Valuation of Data for Machine Learning***, *ICML*, 2019

Objective & Use-cases

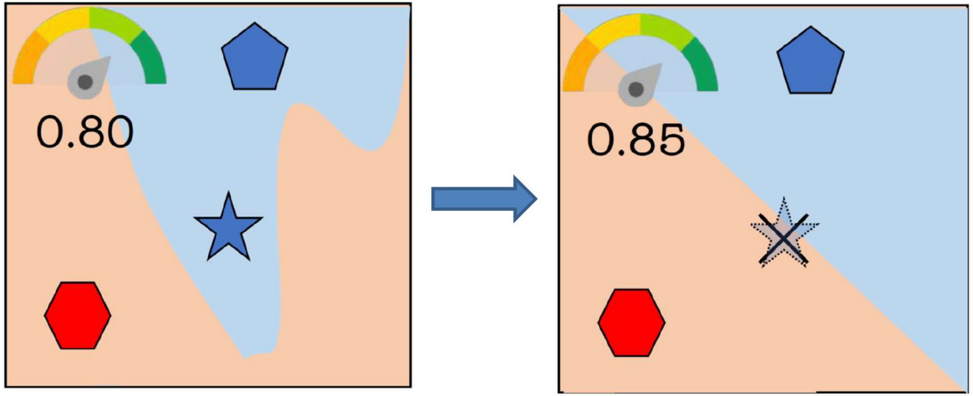
- *Learn in reliable way*

- **Domain adaptation**
  - Assigns higher values on the samples from the target distribution



Google Cloud

Related works - Leave-one-out



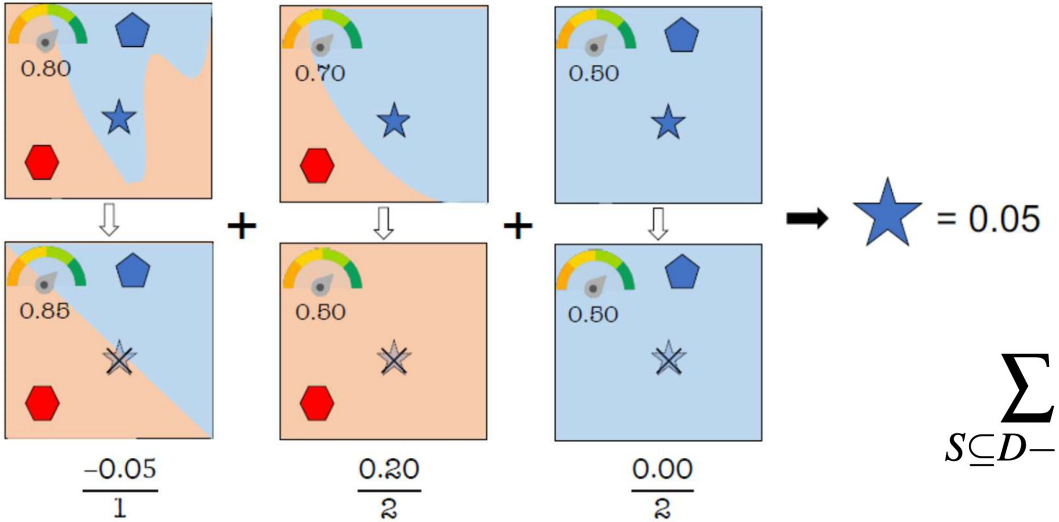
$$V(D) - V(D - \{i\})$$

- Not reasonable when there are two similar training samples.

Google Cloud

Amirata Ghorbani, James Y. Zou, ***Data Shapley: Equitable Valuation of Data for Machine Learning***, *ICML*, 2019

Related works - Data Shapley



$$\sum_{S \subseteq D - \{i\}} \frac{V(S \cup \{i\}) - V(S)}{\binom{n-1}{|S|}}$$

- **Computational complexity is exponential** with the number of samples.

Google Cloud

Amirata Ghorbani, James Y. Zou, ***Data Shapley: Equitable Valuation of Data for Machine Learning***, *ICML*, 2019

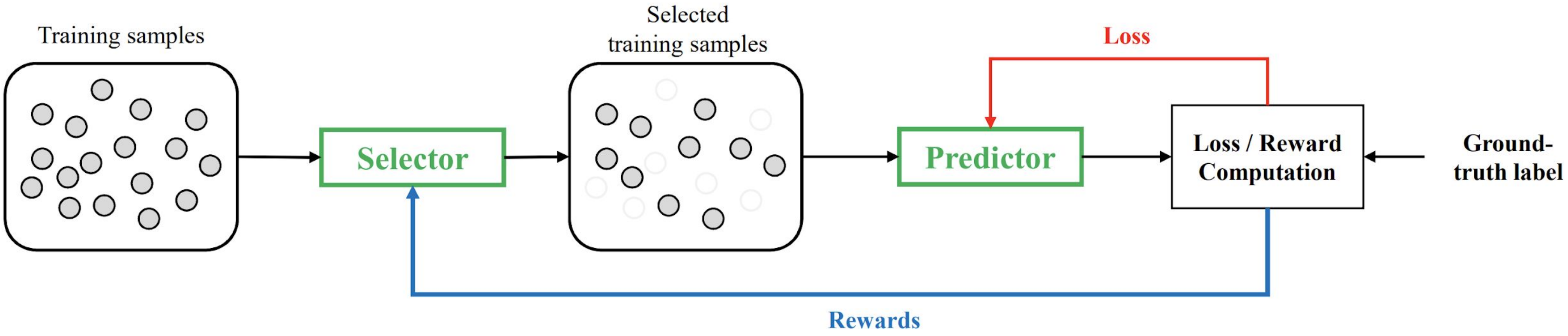
Challenges & Motivation

- **The search space is extremely large.**
  - Impossible to explore the entire space.

- **Training processes can be non-differentiable**
  - Selection operation (i.e. sampler block) is non-differentiable.
  - Performance metrics can be non-differentiable (accuracy, AUC).
  - End-to-end back-propagation may not be possible.

- **Reinforcement learning** is an efficient way to explore large search space and to handle non-differentiable process.

High-level figure for DVRL



- Jointly train **selector** and **predictor** in an end-to-end way.

Google Cloud

Problem formulation

To minimize the validation loss

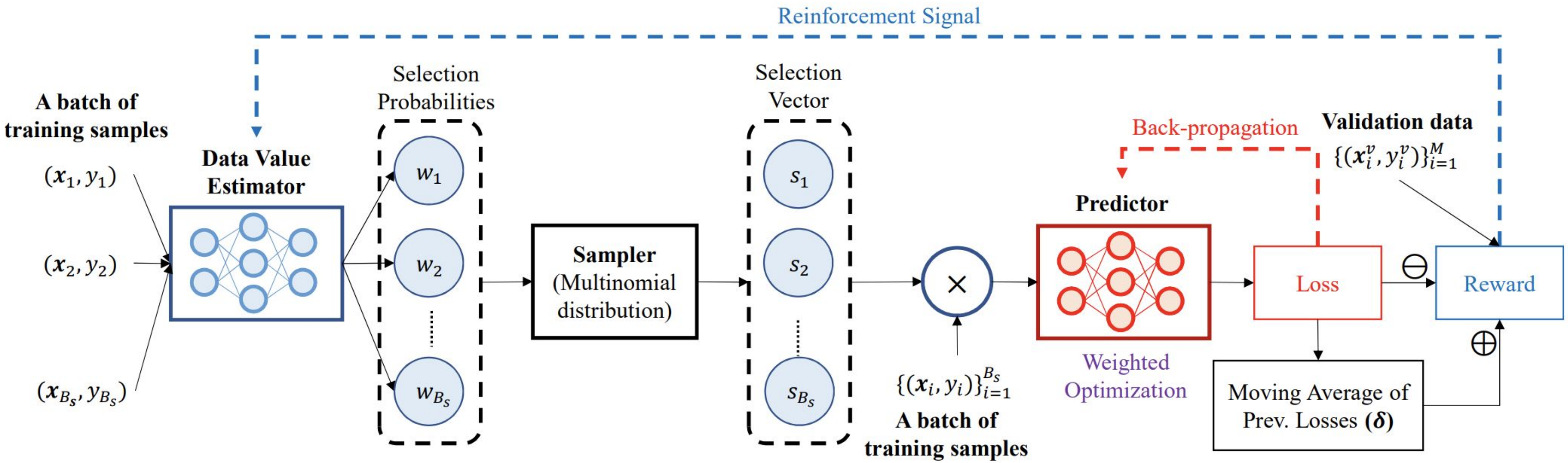
$$\min_{h_\phi} \quad \boxed{\mathbb{E}_{(\mathbf{x}^v, y^v)\sim P^t}\left[\mathcal{L}_h(f_\theta(\mathbf{x}^v), y^v)\right]}$$
$$\text{s.t.} \quad f_\theta = \arg\min_{\hat{f}\in\mathcal{F}} \boxed{\mathbb{E}_{(\mathbf{x},y)\sim P}\left[h_\phi(\mathbf{x}, y)\cdot\mathcal{L}_f(\hat{f}(\mathbf{x}), y)\right]}$$

Weighted optimization for predictor

- Components
  - Training set: $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N \sim \mathcal{P}$
  - Validation set: $\mathcal{D}^v = \{(\mathbf{x}_k^v, y_k^v)\}_{k=1}^L \sim \mathcal{P}^t$
  - Predictor model: $f_\theta : \mathcal{X} \rightarrow \mathcal{Y}$
  - Data valuation model: $h_\phi : \mathcal{X} \cdot \mathcal{Y} \rightarrow [0, 1]$

Google Cloud

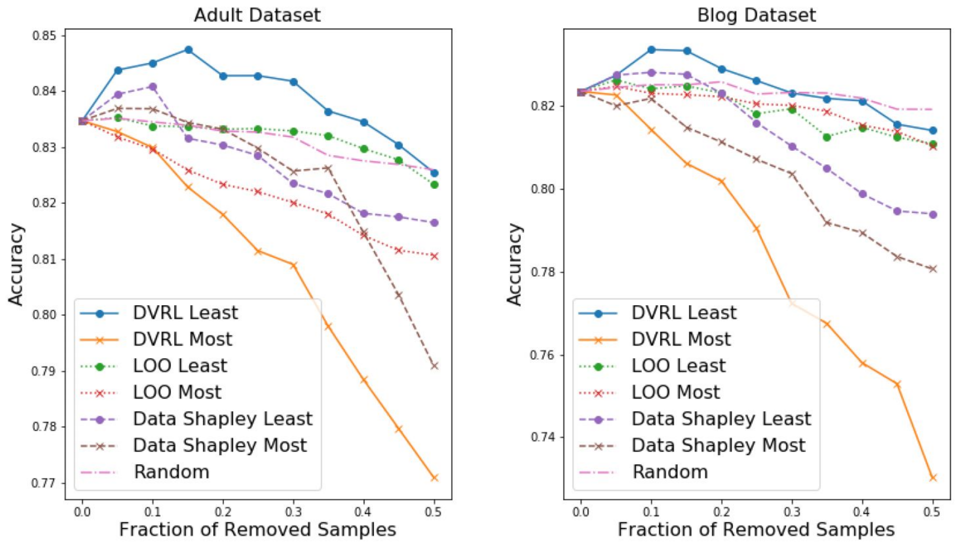
Block diagram



Experiments - How to **quantitatively** evaluate the data valuation?

- Remove high / low valued samples
- Corrupted sample discovery
- Robust learning with noisy data
- Domain adaptation

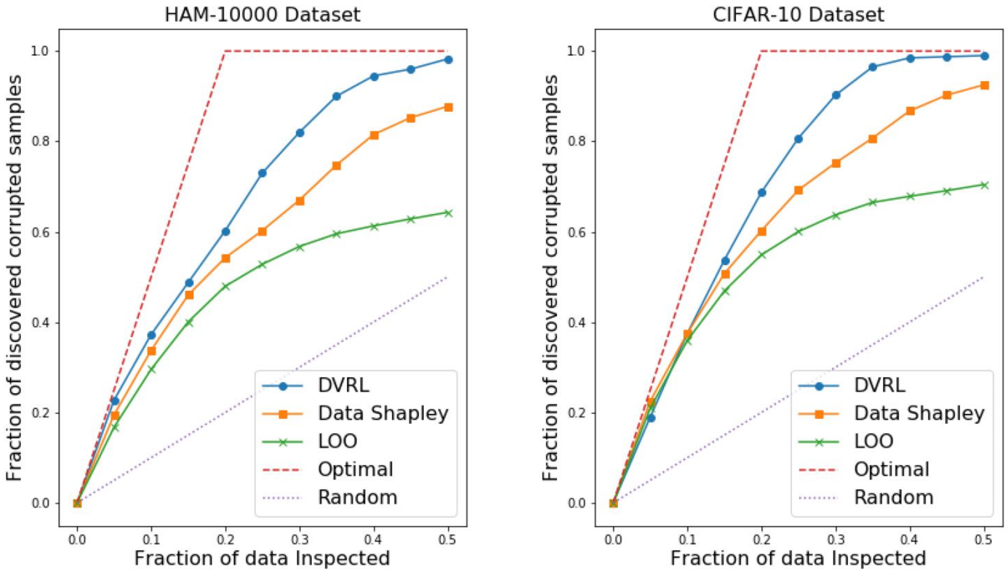
Results - Remove high / low valued samples



- **Standard supervised learning setting** (train, validation, test datasets come from the same distribution)
- Remove **high** valued samples: **Fastest** performance degradation
- Remove **low** valued samples: **Slowest** performance degradation

Google Cloud

Results - Corrupted sample discovery



- Corrupted sample setting (20% of label noise)
- **Highest True Positive Rate (TPR)** for corrupted sample discovery

Google Cloud

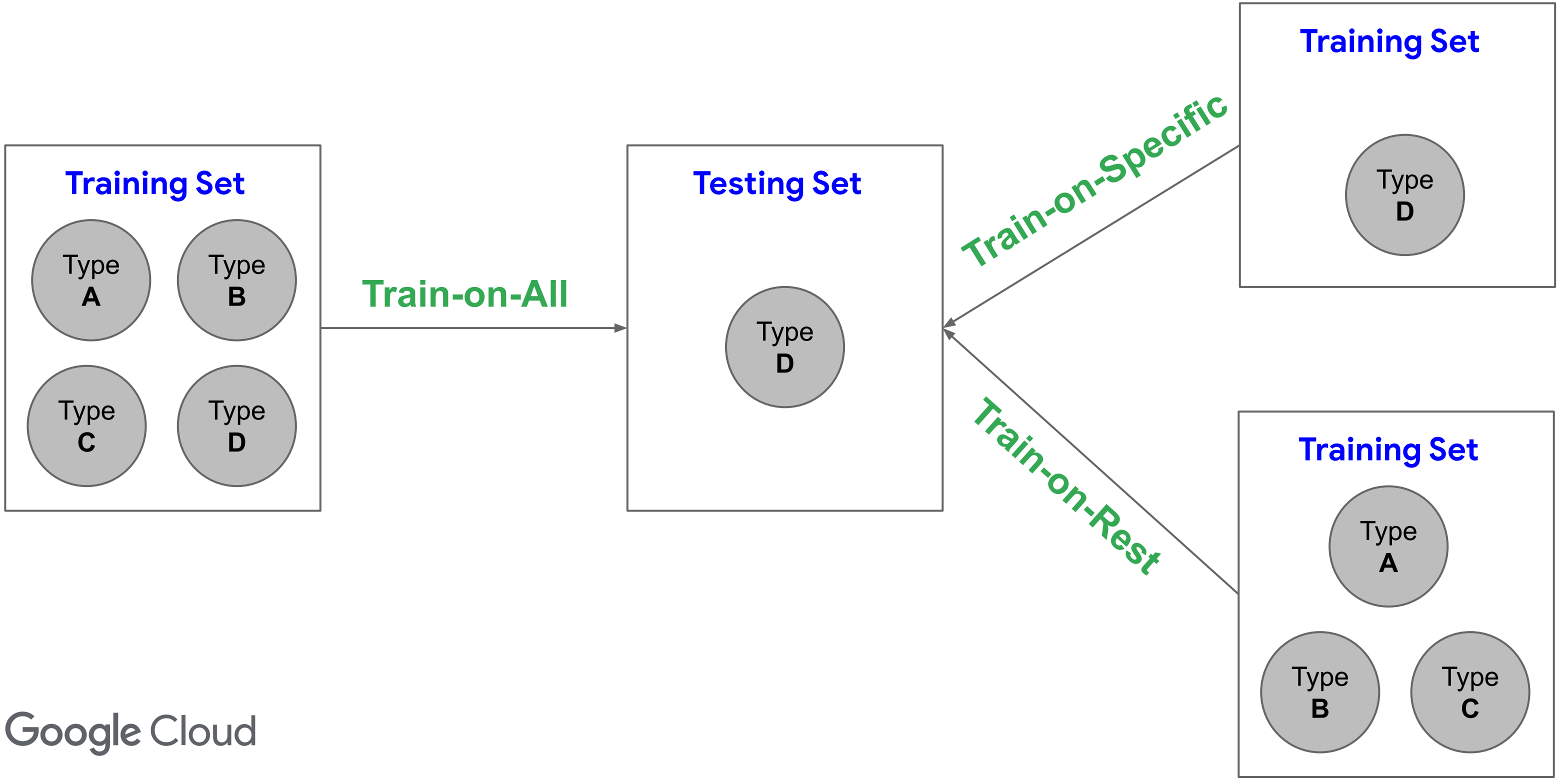
Results - Robust learning with noisy labels (40%)

Noise (predictor model)	**Uniform** (WideResNet-28-10)		**Background** (ResNet-32)	
Datasets	CIFAR-10	CIFAR-100	CIFAR-10	CIFAR-100
Validation Set Only	46.64 ± 3.90	9.94 ± 0.82	15.90 ± 3.32	8.06 ± 0.76
Baseline	67.97 ± 0.62	50.66 ± 0.24	59.54 ± 2.16	37.82 ± 0.69
Baseline + Fine-tuning	78.66 ± 0.44	54.52 ± 0.40	82.82 ± 0.93	54.23 ± 1.75
MentorNet + Fine-tuning	78.00	59.00	-	-
Learning to Reweight	86.92 ± 0.19	61.34 ± 2.06	86.73 ± 0.48	59.30 ± 0.60
DVRL	$\mathbf{89.02 \pm 0.27}$	$\mathbf{66.56 \pm 1.27}$	$\mathbf{88.07 \pm 0.35}$	$\mathbf{60.77 \pm 0.57}$
Clean Only (60% Data)	94.08 ± 0.23	74.55 ± 0.53	90.66 ± 0.27	63.50 ± 0.33
Zero Noise	95.78 ± 0.21	78.32 ± 0.45	92.68 ± 0.22	68.12 ± 0.21

- Proves **scalability** of DVRL in terms of **complex models** (WideResNet-28-10 and ResNet-32) and **large datasets** (CIFAR)
- **State-of-the-art** robust learning performance

Mengye Ren et al., ***Learning to Reweight Examples for Robust Deep Learning***, *ICML*, 2018

Results - Domain adaptation on Retail dataset



Google Cloud

Results - Domain adaptation on Retail dataset

Predictor Model (Metric: RMSPE)	Store Type	*Train on All*		*Train on Rest*		*Train on Specific*	
		Baseline	DVRL	*Baseline*	DVRL	*Baseline*	DVRL
XGBoost	A	0.1736	**0.1594**	0.2369	**0.2109**	0.1454	**0.1430**
	B	0.1996	**0.1422**	0.7716	**0.3607**	0.0880	**0.0824**
	C	0.1839	**0.1502**	0.2083	**0.1551**	0.1186	**0.1170**
	D	0.1504	**0.1441**	0.1922	**0.1535**	0.1349	**0.1221**
Neural Networks	A	0.1531	**0.1428**	0.3124	**0.2014**	0.1181	**0.1066**
	B	0.1529	**0.0979**	0.8072	**0.5461**	0.0683	**0.0682**
	C	0.1620	**0.1437**	0.2153	**0.1804**	0.0682	**0.0677**
	D	0.1459	**0.1295**	0.2625	**0.1624**	0.0759	**0.0708**

- **Significant gain** on Train on Rest setting (**largest domain mismatch**)
- **Reasonable gain** on Train on All setting (**most common setting**)
- **Marginal gain** on Train on Specific setting (**no domain mismatch**)

Results - Domain adaptation in other domains

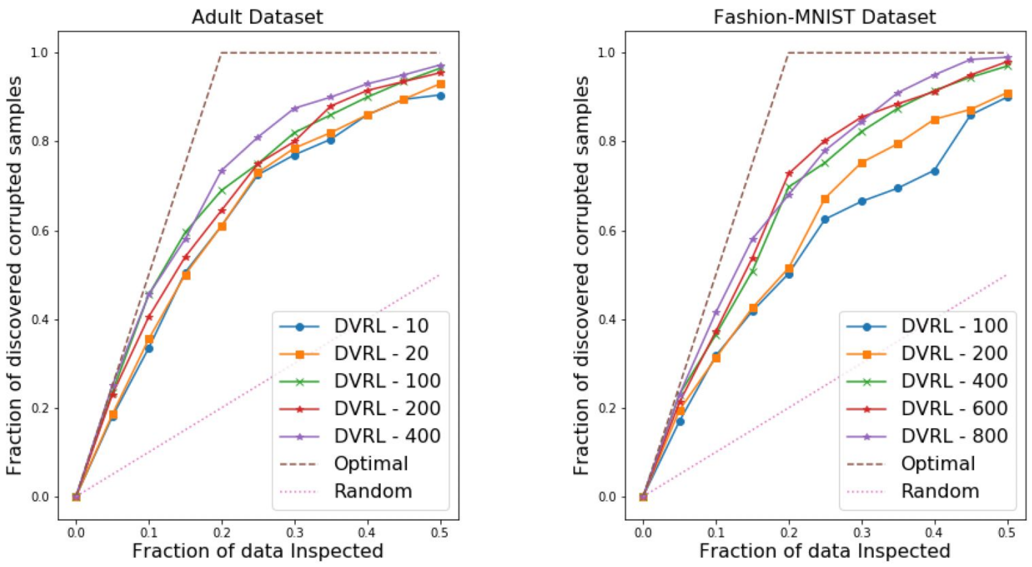
Source	Target	Task	*Baseline*	Data Shapley	**DVRL**
Google	HAM10000	Skin Lesion Classification	.296	.378	**.448**
MNIST	USPS	Digit Recognition	.308	.391	**.472**
Email	SMS	Spam Detection	.684	.864	**.903**

- **Main source of gain:**
    - DVRL **jointly optimizes** the data valuator and corresponding predictor model

Google Cloud

Amirata Ghorbani, James Y. Zou, ***Data Shapley: Equitable Valuation of Data for Machine Learning***, *ICML*, 2019

Discussion: How many validation samples are needed?



- **A small number of validation samples** are enough for DVRL training.
- Reasonable performances even with **10 validation samples** on Adult data.

Google Cloud

Codebase of DVRL

DVRL - Github:
https://github.com/google-research/google-research/tree/master/dvrl

DVRL- AI-Hub:
https://aihub.cloud.google.com/u/0/p/products%2Fcb6b588c-1582-4868-a944-dc70ebe61a36