

Let's talk about  
~~incentives.~~  
trust



Institute of Data Science

**GLOW.AI**



## **Paid with Models:**

Optimal Contract Design for Collaborative Machine Learning

Bingchen Wang\*, Zhaoxuan Wu, Fusheng Liu, Bryan Kian Hsiang Low



# Toward Trustable Model-centric Sharing for Collaborative Machine Learning

## ● Motivation

Models as Rewards

Optimal Contract

Results





# Toward Trustable Model-centric Sharing for Collaborative Machine Learning

## ● Motivation

Models as Rewards

Optimal Contract

Results

 The Guardian

### US stops sharing intelligence on Russia with Ukraine

US no longer providing information about targets in Russia, in latest blow after halting military aid.

1 week ago



 BBC

### Trump vows to quit Paris climate pact and 'drill, baby, drill'

President Donald Trump has once again vowed to withdraw the US from the Paris climate agreement, the world's most important effort to tackle rising...

1 month ago

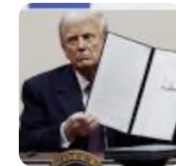


 Reuters

### Trump revokes Biden executive order on addressing AI risks

Biden's order required developers of AI systems that pose risks to U.S. national security, the economy, public health or safety to share the...

1 month ago



How are these related to **Collaborative Machine Learning**?



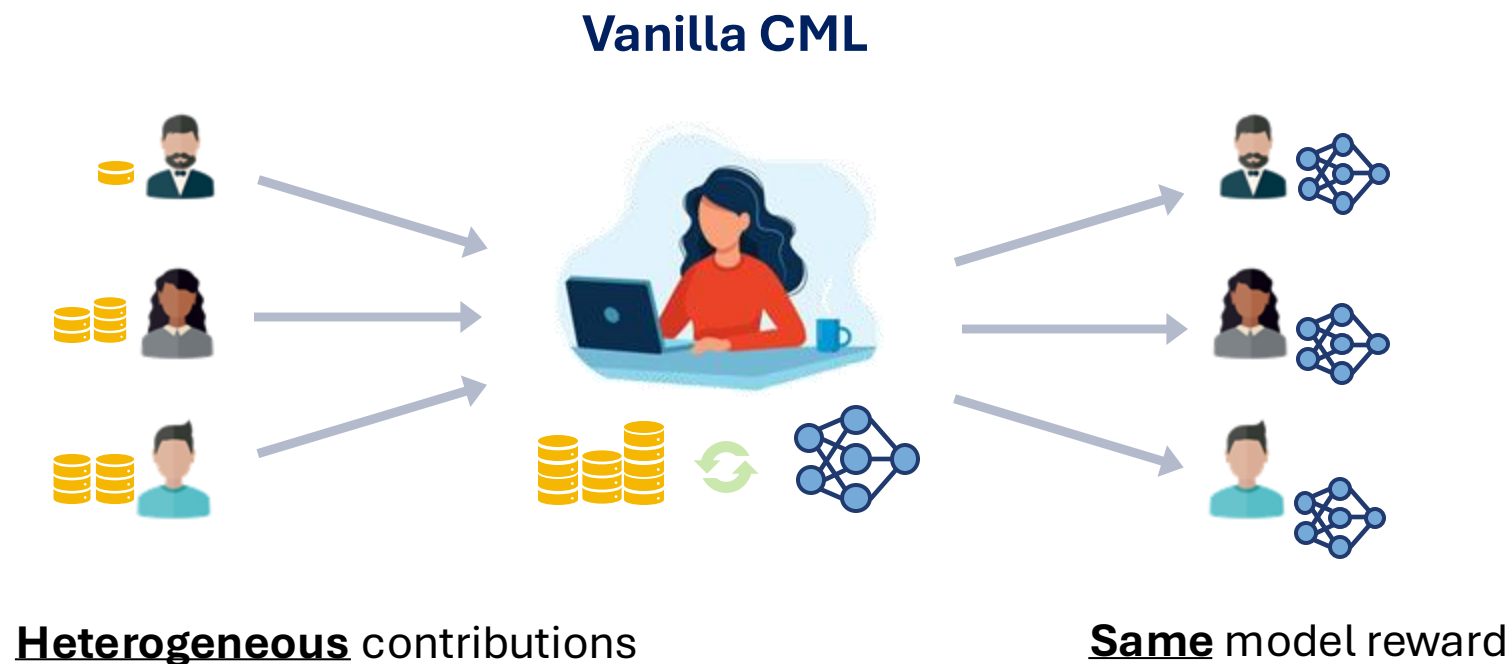
# Toward Trustable Model-centric Sharing for Collaborative Machine Learning

## ● Motivation

Models as Rewards

Optimal Contract

Results







# Toward Trustable Model-centric Sharing for Collaborative Machine Learning

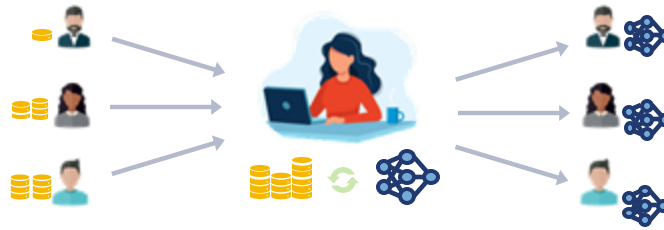
## ● Motivation

Models as Rewards

Optimal Contract

Results

## Vanilla CML



Cost is **private**, but benefit is **shared**.

## Collaboration failures in the wild

The Guardian

### US stops sharing intelligence on Russia with Ukraine

US no longer providing information about targets in Russia, in latest blow after halting military aid.

1 week ago



BBC

### Trump vows to quit Paris climate pact and 'drill, baby, drill'

President Donald Trump has once again vowed to withdraw the US from the Paris climate agreement, the world's most important effort to tackle rising...

1 month ago



Reuters

### Trump revokes Biden executive order on addressing AI risks

Biden's order required developers of AI systems that pose risks to U.S. national security, the economy, public health or safety to share the...

1 month ago



If they fail to incorporate the **positive externalities** of their actions, then...



# Toward Trustable Model-centric Sharing for Collaborative Machine Learning

## ● Motivation

Models as Rewards

Optimal Contract

Results





# Catastrophic Freeriding for CML

## ● Motivation

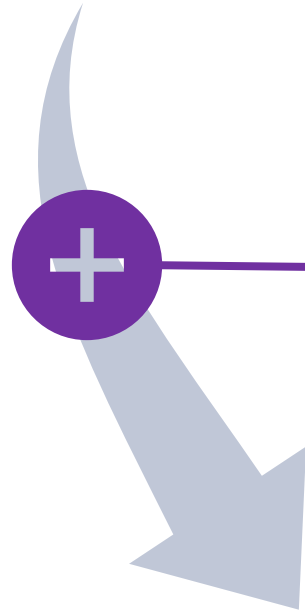
Models as Rewards

Optimal Contract

Results



Conflict of interests



Karimireddy, Guo and Jordan (2022)



Parties have **different contribution costs.**

Catastrophic  
freeriding





# We cannot talk about trust without talking about money.

## ● Motivation

Models as Rewards

Optimal Contract

Results

 Reuters


Trump welcomes Ukraine's willingness to sign minerals deal,  
talk peace

NEW YORK, March 4 (Reuters) - U.S. President Donald Trump said on Tuesday he appreciated Ukrainian President Volodymyr Zelenskyy's...

1 week ago



Signing deals or contracts is a practical way to restore trust in a collaborative setting

 The Kyiv Independent

US to 'immediately' resume security assistance, intelligence sharing to Ukraine

US Secretary of State Marco Rubio (L) and US National Security Advisor Mike Waltz (R) speak with the media following meetings with a Ukrainian delegation in...

2 days ago







# Building trust with model rewards ~~money~~

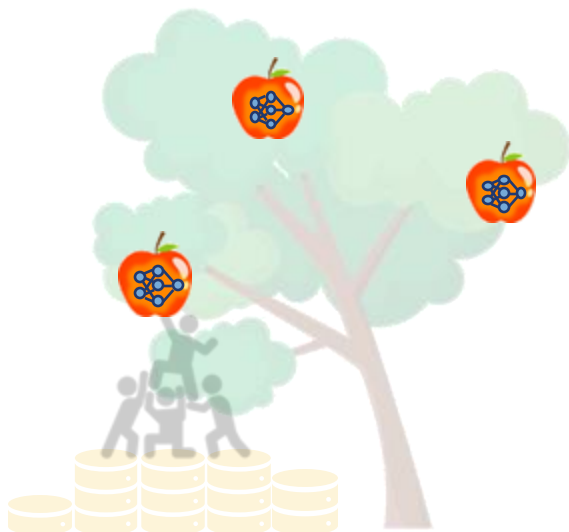
Model rewards are different from money:

Motivation

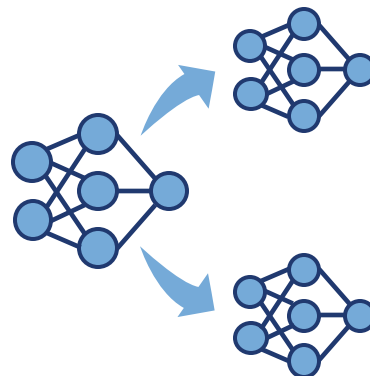
## ● Models as Rewards

Optimal Contract

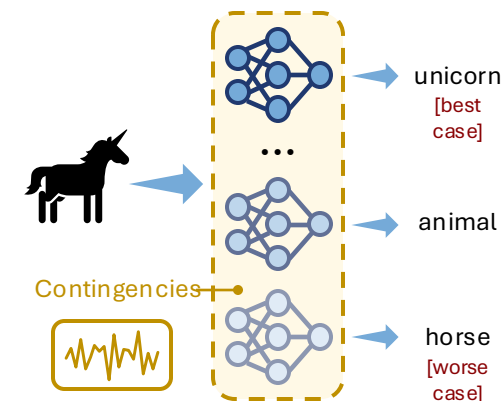
Results



(1) Non-rivalrous.



(2) Stochastic ex-ante.



Assigning **deterministic model rewards** risks **insolvency** or **under-utilization** of the budget.



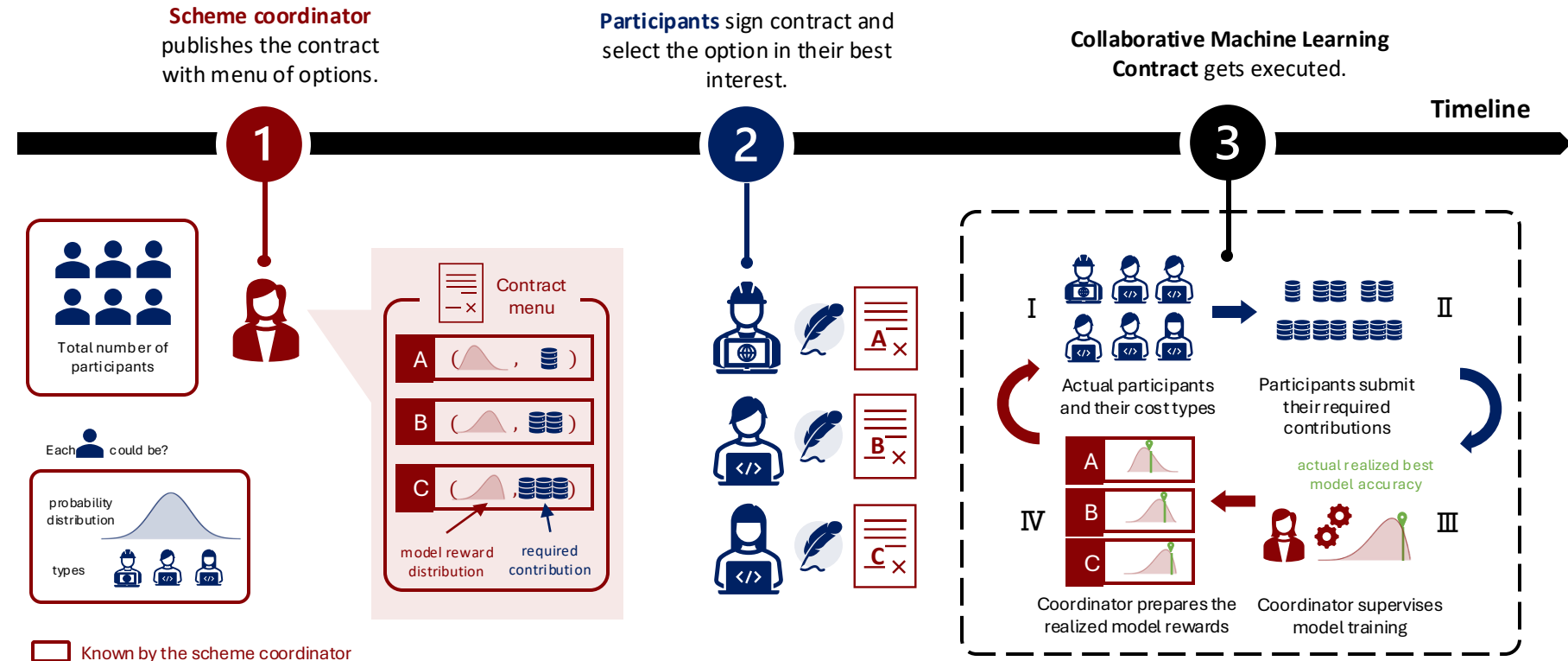
# CML with Optimal Contract Design

## Motivation

## Models as Rewards

## Optimal Contract

## Results





# Constrained Optimization

Motivation

Models as Rewards

## ● Optimal Contract

Results



$$\max_{(\mathbf{r}_i, m_i)_{i=1}^I} \mathbb{E}_{n \sim \text{Multi}(N, p)} \left[ a \left( \sum_{i=1}^I n_i m_i \right) \right]$$



$$\text{s. t. } \begin{cases} \mathbb{E}_{n_i \geq 1} [v(r_i)] - c_i m_i \geq f_i, \forall i \\ \mathbb{E}_{n_i \geq 1} [v(r_i)] - c_i m_i \geq \mathbb{E}_{n_j \geq 1} [v(r_j)] - c_j m_j, \forall i, j \\ \|r(n)\|_{\infty} \leq a(\sum_{i=1}^I n_i m_i), \forall n \in \text{Multi}(N, p) \end{cases}$$



Individual Rationality



Incentive Compatibility



Budget Constraint

Hard to solve directly.





# First-moment Problem

Motivation

Models as Rewards

● **Optimal Contract**

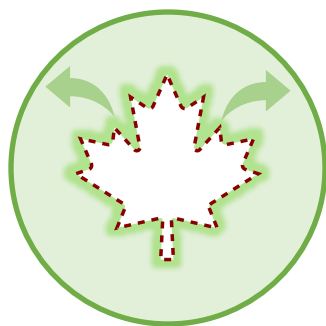
Results



original problem



map solution back to  
the original problem



relax the constraint



optimize in the  
new domain

$$\begin{aligned} & \max_{(t_i, m_i)_{i=1}^I} \mathbb{E}_{n \sim \text{Multi}(N, p)} \left[ a \left( \sum_{i=1}^I n_i m_i \right) \right] \\ \text{s. t. } & \begin{cases} t_i - c_i m_i \geq f_i, \forall i \\ t_i - c_i m_i \geq t_j - c_j m_j, \forall i, j \\ t_i \leq \mathbb{E}_{n_i \geq 1} \left[ v \left( a \left( \sum_{i=1}^I n_i m_i \right) \right) \right], \forall i \end{cases} \end{aligned}$$



Fewer variables



Fewer constraints



Convex





# Experiment Results

Motivation

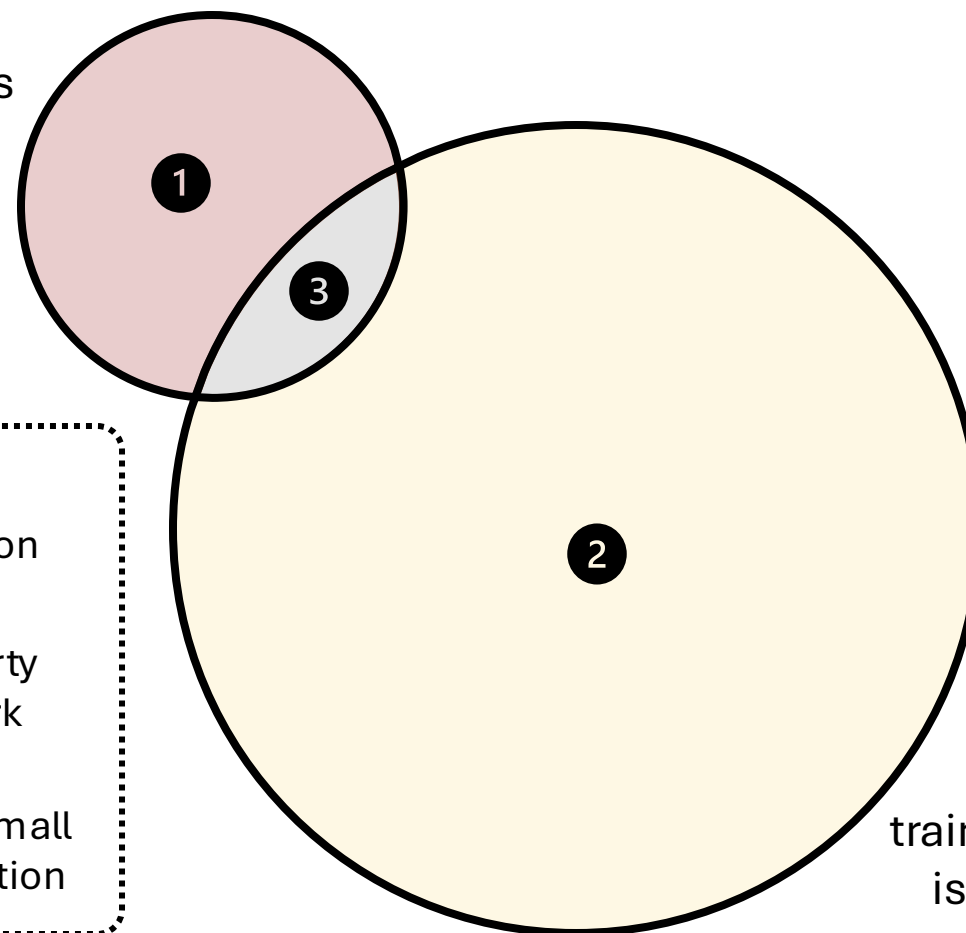
Models as Rewards

Optimal Contract

## ● Results



training a model is  
manageable  
 $f_i > 0$



1  Big-firm  
cooperation

2  Small-party  
teamwork

3  Big-and-small  
collaboration

$f_i = 0$   
training a model  
is too costly





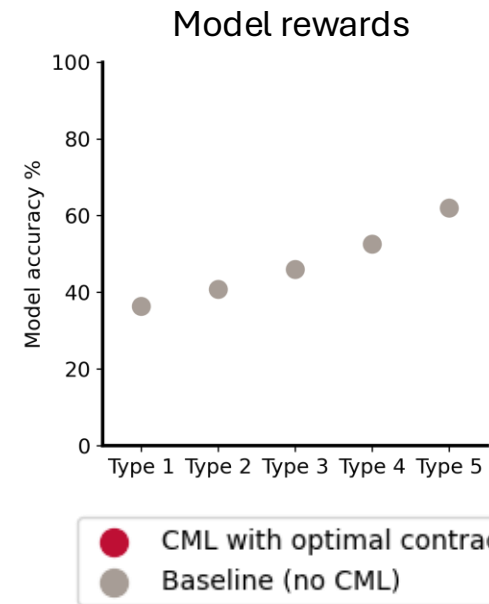
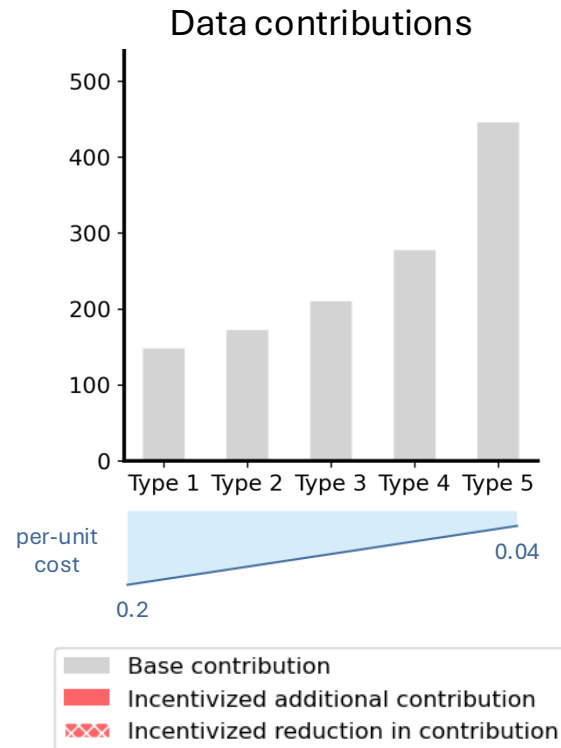
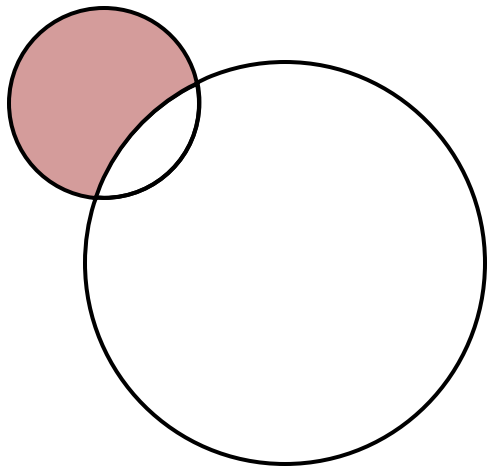
# Experiment Results (Big-firm cooperation)

Motivation

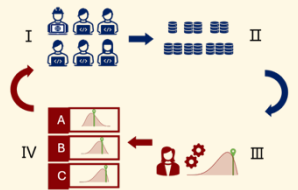
Models as Rewards

Optimal Contract

## ● Results



CML + OCD



OFF



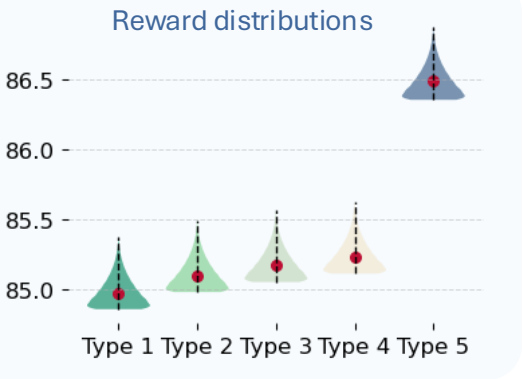
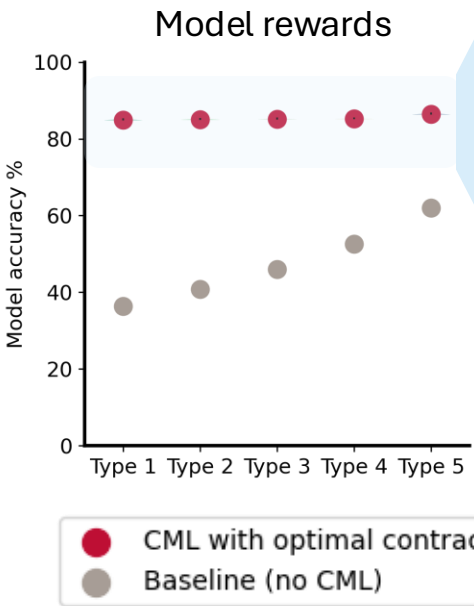
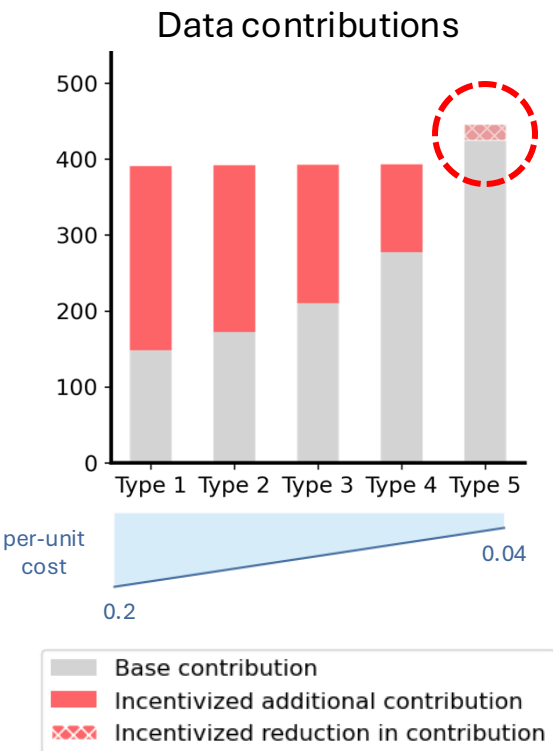
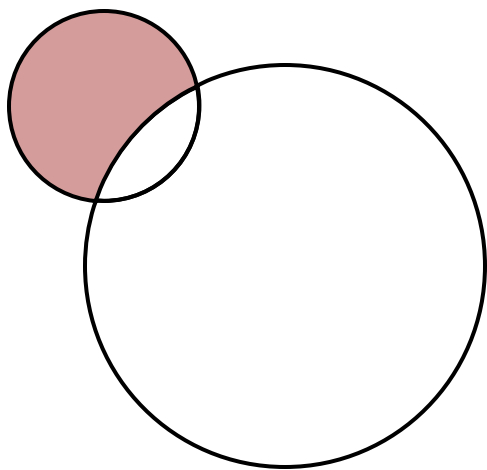
# Experiment Results (Big-firm cooperation)

Motivation

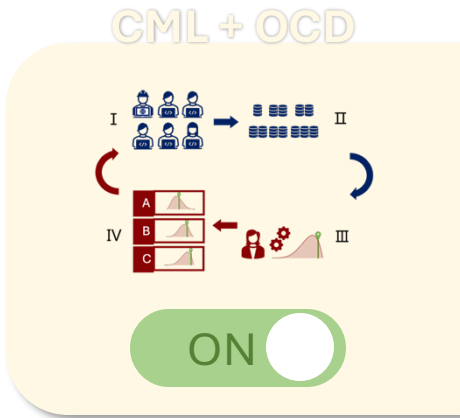
Models as Rewards

Optimal Contract

● **Results**



A party can be incentivized to **contribute less**.





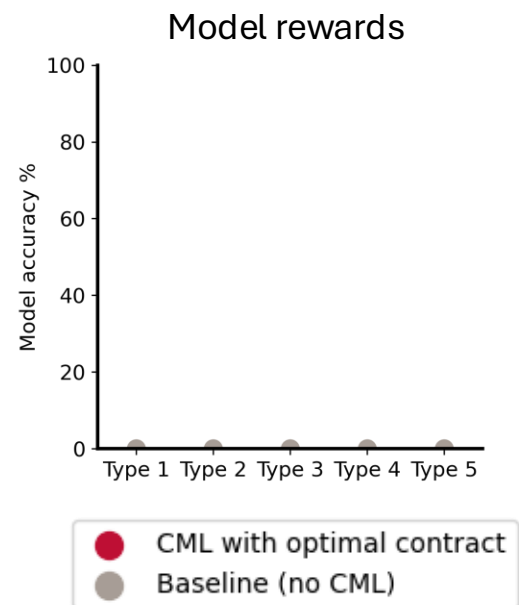
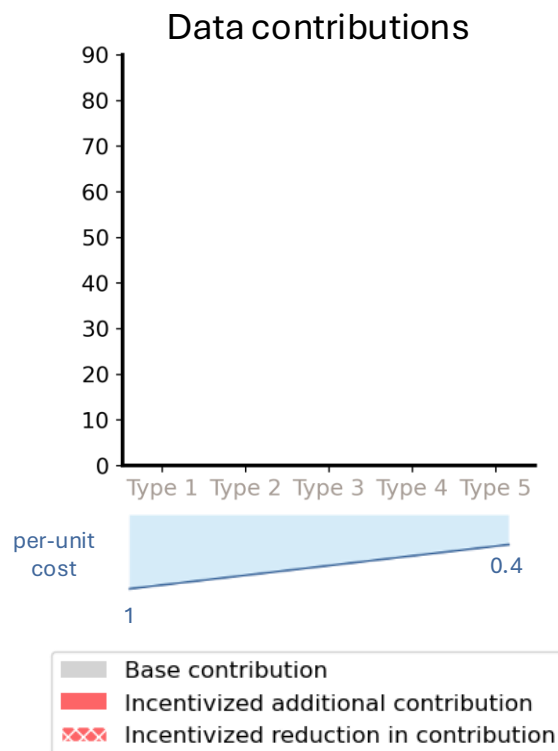
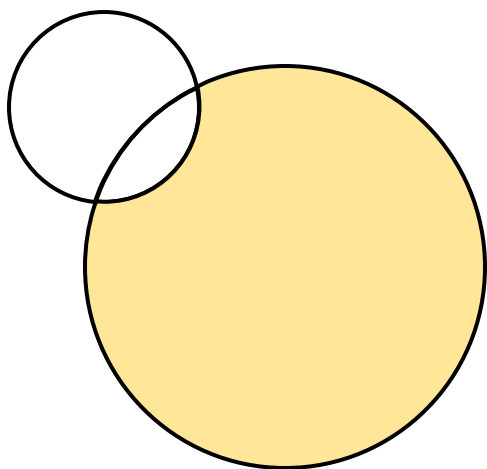
# Experiment Results (Small-party teamwork)

Motivation

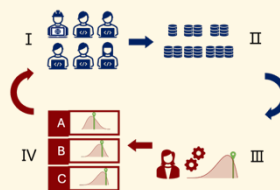
Models as Rewards

Optimal Contract

## ● Results



CML + OCD



OFF



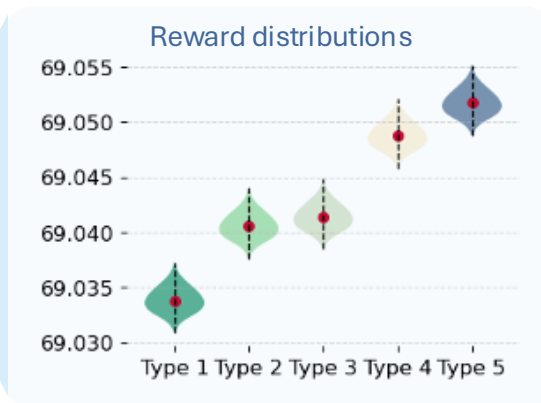
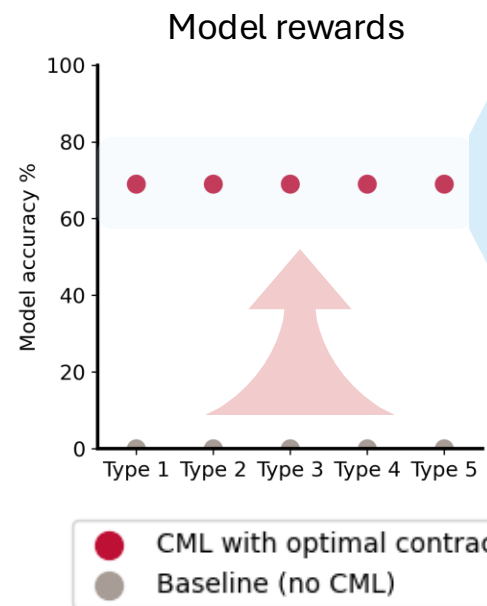
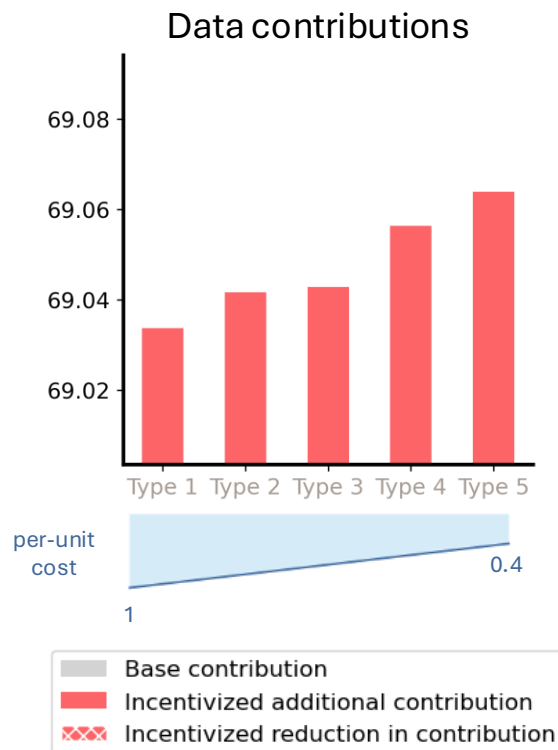
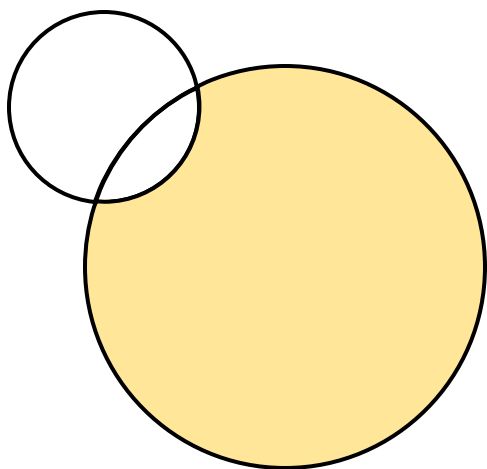
# Experiment Results (Small-party teamwork)

Motivation

Models as Rewards

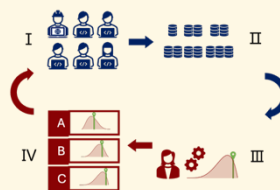
Optimal Contract

## ● Results



Parties successfully surmount model training barrier.

CML + OCD



ON



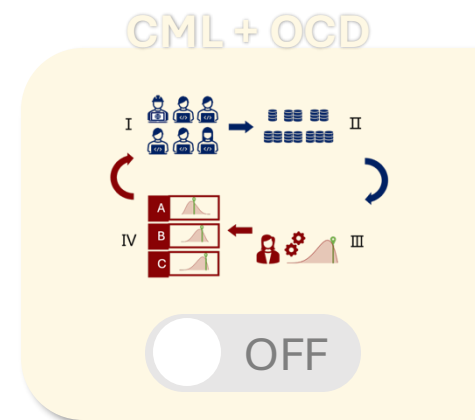
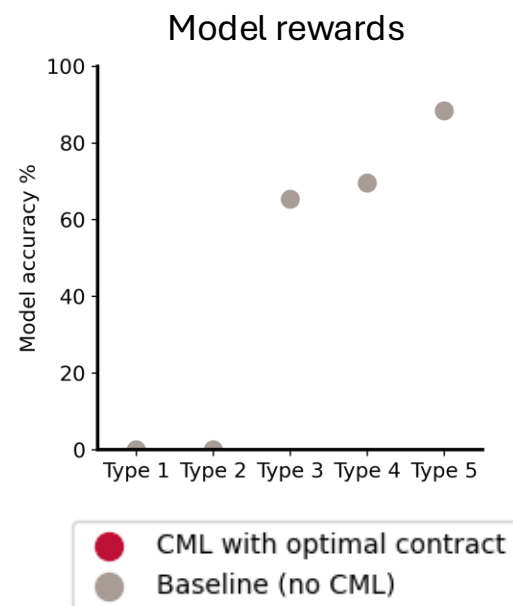
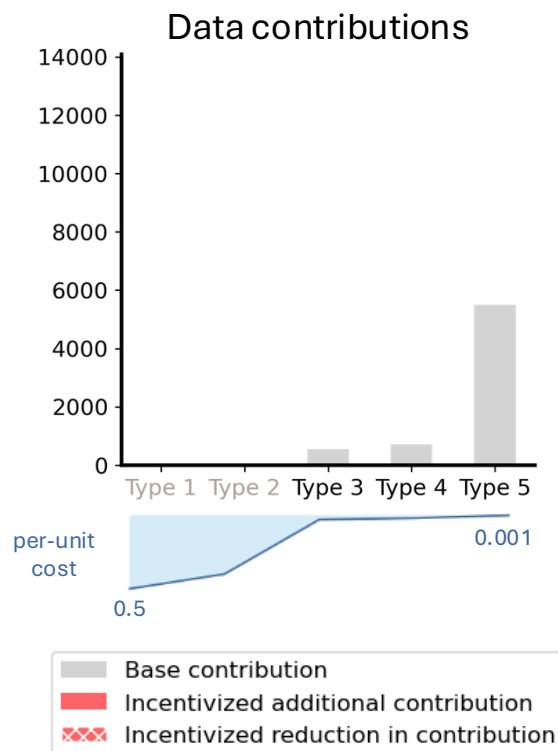
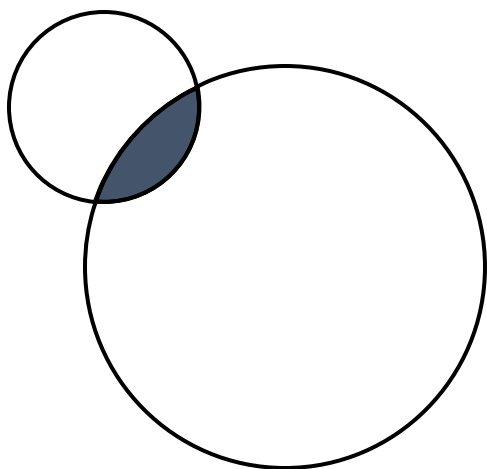
# Experiment Results (Big-and-small collaboration)

Motivation

Models as Rewards

Optimal Contract

## ● Results







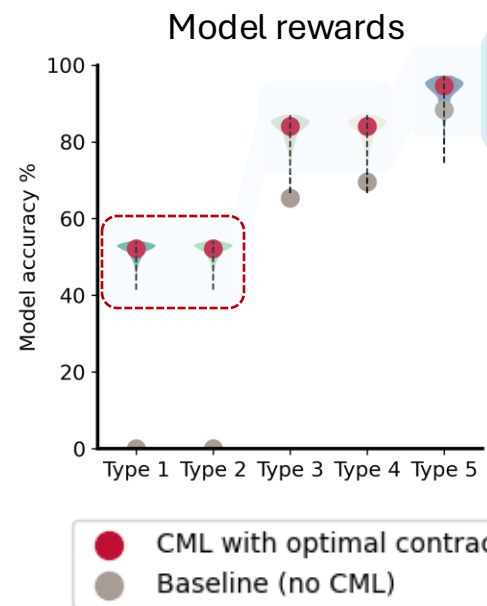
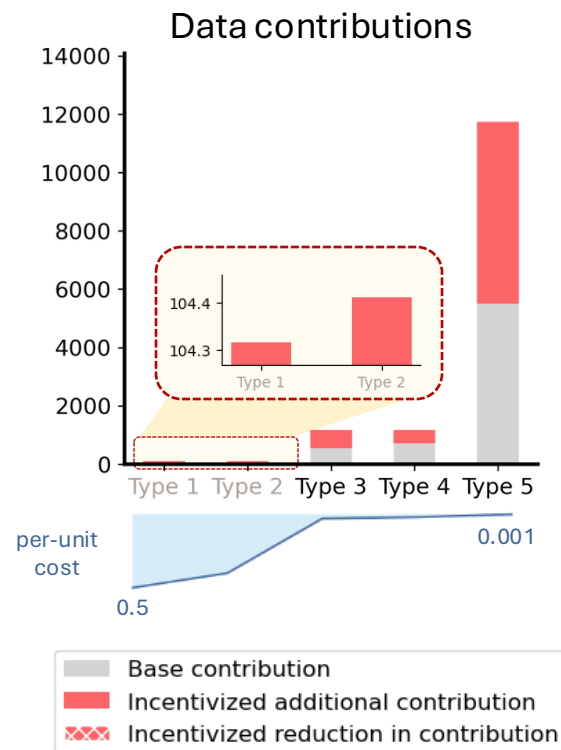
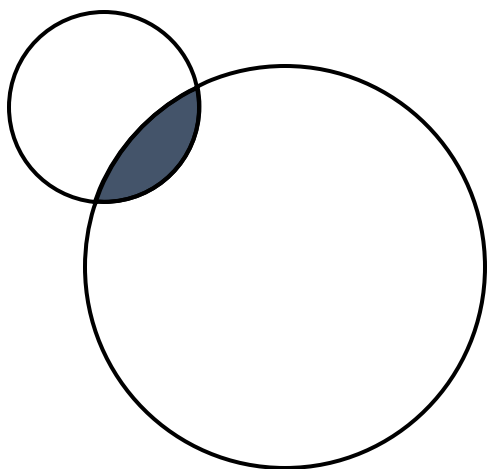
# Experiment Results (Big-and-small collaboration)

Motivation

Models as Rewards

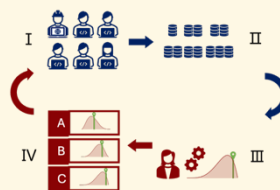
Optimal Contract

## ● Results



Small parties can **still gain** from collaboration.

CML + OCD



ON



Motivation

Models as Rewards

Optimal Contract

## ● Results

# Conclusion

"You do not rise to the level of your goals. You fall to the level of your systems."  
— James Clear, *Atomic Habits*.

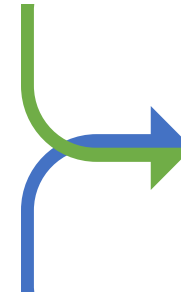
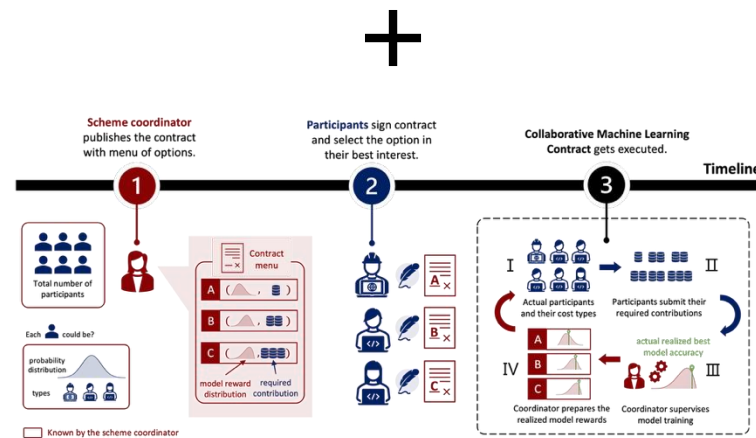


A promising crowdsourcing paradigm to democratize AI?

Full paper



Code





# Toward Trustable Model-centric Sharing for Collaborative Machine Learning

Motivation

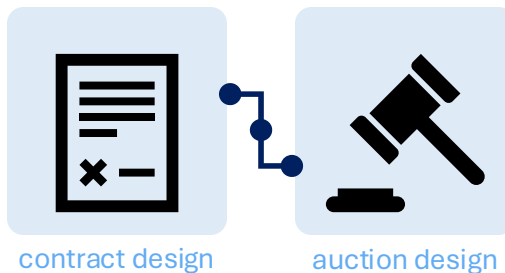
Models as Rewards

Optimal Contract

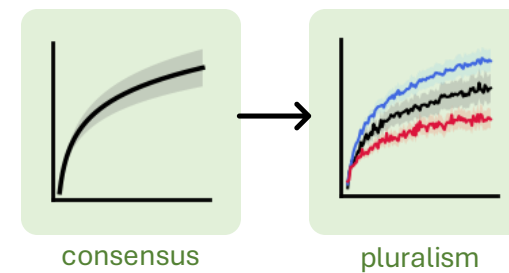
Results

## ● Limitations & Future Work

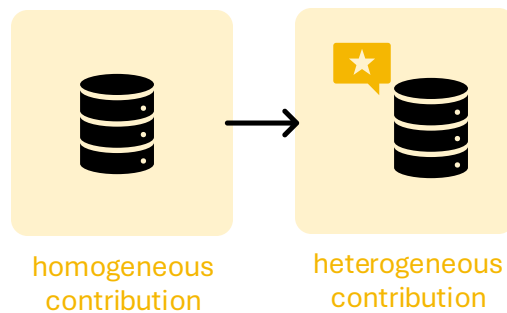
Empirical justification for behavioral assumptions



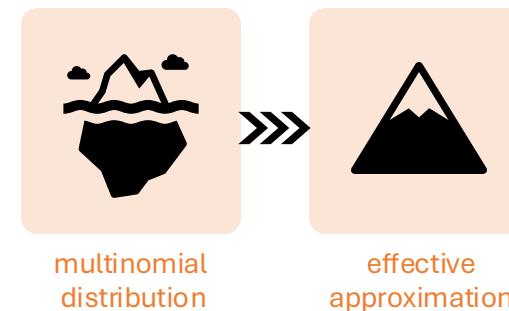
Prior-training accuracy function & valuation function



Incorporating quality in the contribution measure



Distributional assumption & combinatorial challenge



Full paper

Code

