Performative Power of Federated Learning

Andrew Isaacson

November 2025

Abstract

This paper introduces Participant Selection (ParS), a targeting strategy that enables Federated Learning to exploit user heterogeneity for engagement optimization. By allocating model intensity proportionally to client sensitivity rather than uniformly, FL can focus on the most responsive subpopulations. While FL is typically viewed as a privacy-constrained approximation of centralized training, we present a counter-intuitive result: FL structurally outperforms centralized approaches in the Performative Prediction setting. We prove analytically that the FL-to-centralized power ratio equals $\sqrt{E[\varepsilon^2]}/E[\varepsilon] \geq 1$, with the gap increasing as user heterogeneity grows. We validate these findings on synthetic data and real Reddit community data, demonstrating up to 13% improvement in Performative Power over centralized baselines.

1 Introduction

Performative Power measures a platform's ability to shift user behavior distributions through deployed models. In recommendation systems, algorithms influence engagement patterns by optimizing interactions with content. A social media platform seeking to maximize engagement must consider not just what content to recommend, but how different users respond to recommendations [2].

1.1 Performative Power

Performative Power quantifies a model's capacity to influence users and shift data distributions. Following Hardt et al. [2], we define:

$$P := \sup_{f \in F} \frac{1}{|U|} \sum_{u \in U} \mathbb{E}\left[\operatorname{dist}(z(u), z_f(u))\right] \tag{1}$$

where U is the set of all users, F is the set of available actions (e.g., recommendation policies), and $\operatorname{dist}(z(u), z_f(u))$ measures the behavioral shift for user u after action f is taken. This captures the maximum distributional shift achievable across the user population.

Traditional recommendation systems assume static user behavior distributions. Performative Prediction [2] recognizes that deployed models reshape the distributions they operate on. When a platform deploys model θ, users shift their behavior, creating a feedback loop between model and data.

1.2 Federated Learning

Federated Learning (FL) was originally developed for privacy preservation [1], enabling model training without centralizing raw user data. Local models are trained at the client level, then aggregated to update a global model. This approach satisfies data protection requirements such as GDPR while still enabling machine learning at scale [1].

This paper investigates whether FL provides advantages beyond privacy for maximizing Performative Power. We present a counter-intuitive result: FL does not merely approximate centralized training under privacy constraints—it structurally exceeds centralized performance when users exhibit heterogeneous sensitivities to recommendations.

1.3 Federated Learning in Performative Prediction

Jin et al. [3] introduced Performative Federated Learning with the P-FedAvg algorithm, which converges to a Performative Stable solution at rate $O(1/T)$. Algorithm 1 shows the full-participation version.

Algorithm 1: Performative FedAvg (P-FedAvg) [3]

Input: Clients $i \in \{1, \ldots, N\}$, weights $\{p_i\}$, learning rates $\{\eta_t\}$, aggregation interval E.
Output: Model sequence $\{\theta^t\}$.

Initialize: Server broadcasts initial model $\theta_i^0 = \theta^0$ to all clients.
for $t = 0, \ldots, T-1$ **do**
    for *each client* $i \in \{1, \ldots, N\}$ *in parallel* **do**
        Sample data $Z_i^{t+1} \sim D_i(\theta_i^t)$;
        Compute local update: $w_i^{t+1} = \theta_i^t - \eta_t \nabla \ell(\theta_i^t; Z_i^{t+1})$;
    end
    if $t + 1 \in \mathcal{I}_E$ **then**
        $\theta^{t+1} = \sum_{j=1}^{N} p_j w_j^{t+1}$;
        Broadcast: $\theta_i^{t+1} \leftarrow \theta^{t+1}$ for all i;
    else
        $\theta_i^{t+1} \leftarrow w_i^{t+1}$ for all i;
    end
end

Jin et al. proved convergence under both full and partial participation. Our work confirms these findings and extends the analysis to Performative Power comparison.

1.4 Participant Selection under Budget Constraints

We introduce *Participant Selection* (ParS), where the server allocates model intensity proportionally to client sensitivity. Under an L_2 budget constraint $\sum_i \alpha_i \theta_i^2 \leq B^2$, the optimal FL targeting strategy is:

$$\theta_i^* = \frac{\varepsilon_i \cdot B}{\sqrt{E[\varepsilon^2]}} \tag{2}$$

This allocates greater influence to high-sensitivity clients while respecting the total resource budget. Centralized deployment, by contrast, must use uniform θ across all clients. The full derivation appears in Appendix **??**.

Our main theoretical contribution shows that under equivalent resource constraints, the FL-to-centralized Performative Power ratio is:

$$\frac{P_{\text{FL}}}{P_{\text{central}}} = \frac{\sqrt{E[\varepsilon^2]}}{E[\varepsilon]} = \sqrt{1 + \frac{\text{Var}(\varepsilon)}{E[\varepsilon]^2}} \geq 1 \tag{3}$$

The inequality follows from Cauchy-Schwarz, with equality only when $\text{Var}(\varepsilon) = 0$ (all clients identical). FL's advantage grows with user heterogeneity.

2 Methodology

2.1 Distribution Model

We adopt a linear Gaussian response model. For user u, the behavior distribution under deployed model θ is:

$$D_u(\theta) = \mathcal{N}(\mu_u + \varepsilon_u \cdot \theta, \sigma^2) \tag{4}$$

where μ_u is user u's baseline behavior (without model influence), $\varepsilon_u \in (0, 1)$ is the user's sensitivity, and σ^2 is noise variance.

This satisfies the Wasserstein sensitivity assumption from Jin et al. [3]:

$$W_1(D_u(\theta), D_u(\theta')) = |\varepsilon_u| \cdot |\theta - \theta'| \tag{5}$$

Higher ε_u indicates a more responsive user whose behavior shifts more readily with recommendations.

2.2 System Architecture

Our implementation contains three levels:

Users: Individual users u with baseline μ_u and sensitivity ε_u. Behavior is sampled from $D_u(\theta)$ when model θ is deployed.

Clients: Groups of users (e.g., subreddits) indexed by $i \in \{1, \ldots, N\}$. Each client i has:

- User set U_i with aggregate sensitivity $\varepsilon_i = \frac{1}{|U_i|} \sum_{u \in U_i} \varepsilon_u$

- Population weight $\alpha_i = |U_i| / \sum_j |U_j|$, the fraction of total users belonging to client i

Server: Coordinates P-FedAvg training and computes population statistics $E[\varepsilon]$, $E[\varepsilon^2]$, and $\mathrm{Var}(\varepsilon)$.

2.3 P-FedAvg Implementation

We implement P-FedAvg with the following specifications:

- **Loss Function:** Squared loss $\ell(\theta; z) = \frac{1}{2}(\theta - z)^2$
- **Learning Rate:** Decaying schedule $\eta_t = \beta/(t + \gamma)$ with $\beta = 2$, $\gamma = 10$, satisfying convergence conditions in [3]
- **Aggregation:** Weighted averaging $\theta^{t+1} = \sum_i p_i \theta_i^t$ where $p_i = \alpha_i$
- **Local Steps:** $E = 5$ SGD steps per round
- **Batch Size:** $B = 4$ samples per local step

The Performative Stable solution for our Gaussian model is:

$$\theta^{PS} = \frac{\mu}{1 - \varepsilon} \tag{6}$$

where $\mu = E[\mu_u]$ and $\varepsilon = E[\varepsilon_u]$.

3 Experimental Setup

3.1 Synthetic Data

We generate synthetic populations to validate theoretical results under controlled conditions:

- $N = 10$ clients with 50 users each (500 total users)
- Client sensitivities $\varepsilon_i \sim \mathcal{N}(0.5, 0.3)$, clipped to $\varepsilon_i > 0.1$
- User sensitivities $\varepsilon_u \sim \mathcal{N}(\varepsilon_i, 0.05)$, adding within-client heterogeneity
- Baselines $\mu_u \sim \mathcal{N}(5.0, 0.5)$
- Training rounds $T = 200$
- Noise $\sigma = 1.0$
- Budget $B = 1.0$

Ground-truth sensitivities allow exact verification of theoretical predictions.

3.2 Reddit Data

For realistic validation, we derive sensitivities from Reddit engagement metrics across 50 subreddits with 1,000 users per subreddit (50,000 total users).

Sensitivity Derivation: We hypothesize that communities with high engagement variance and strong user reactions are more responsive to algorithmic recommendations. Each metric is normalized to $[0, 1]$ via min-max scaling:

$$\mathrm{norm}(x; \ell, h) = \ell + (h - \ell) \cdot \frac{x - \min(x)}{\max(x) - \min(x) + 10^{-9}} \tag{7}$$

The combined sensitivity score weights four engagement signals:

$$\varepsilon_i^{\mathrm{raw}} = 0.4 \cdot \tilde{v}_i + 0.3 \cdot \tilde{e}_i + 0.2 \cdot \tilde{a}_i + 0.1 \cdot \tilde{c}_i \tag{8}$$

where:

- $\tilde{v}_i$ (**Score variance**): Variance in post scores within the subreddit. High variance indicates unpredictable content performance, suggesting users react strongly to algorithmic curation.
- $\tilde{e}_i$ (**Votes per comment**): Ratio of total votes to comments. High values indicate active engagement rather than passive consumption.
- $\tilde{a}_i$ (**Comments per author**): Average comments per unique author. Repeat engagement suggests susceptibility to recommendation-driven behavior loops.
- $\tilde{c}_i$ (**Controversy ratio**): $\frac{\text{mean\_downs}}{\text{mean\_ups}+1}$. Polarized communities with high downvote rates may have more reactive users.

The final sensitivity is scaled and clipped:

$$\varepsilon_i = \mathrm{clip}(0.1 + 0.8 \cdot \varepsilon_i^{\mathrm{raw}},\ 0.05,\ 0.95) \tag{9}$$

4 Results

4.1 Synthetic Data

Figure 1 shows P-FedAvg convergence on synthetic data. Full participation converges fastest, confirming Jin et al.'s theoretical results [3]. Both partial participation schemes (I and II) also converge, with slightly higher variance.
FL achieves approximately 9% higher Performative Power than centralized training. The FL algorithm allocates resources more effectively across clients by targeting high-sensitivity users, compared to the uniform distribution required by centralized deployment.

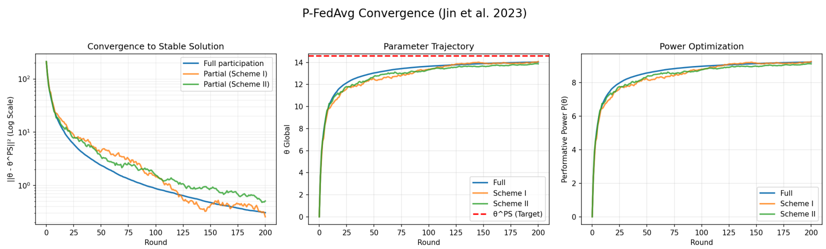

Figure 1: P-FedAvg convergence on synthetic data. All participation modes converge to θ^{PS}, with full participation achieving fastest convergence.

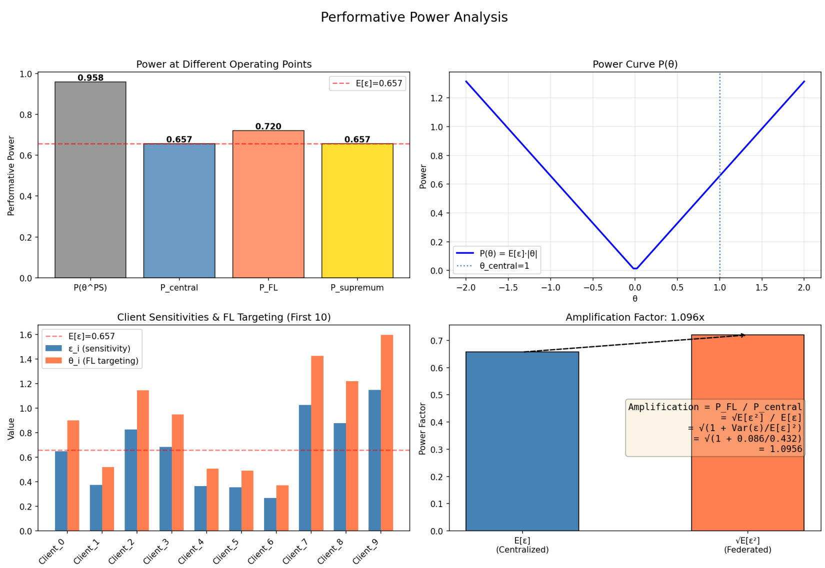

Figure 2: Power analysis on synthetic data. FL outperforms centralized by exploiting sensitivity heterogeneity.

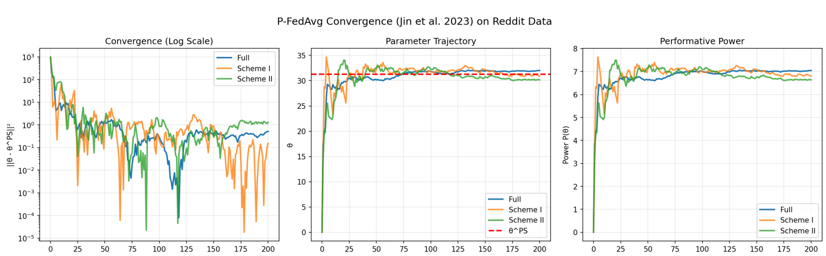

Figure 3: P-FedAvg convergence on Reddit data (50 subreddits, 50,000 users).

4.2 Reddit Data

Figure 3 demonstrates P-FedAvg convergence on Reddit data. Despite greater heterogeneity than synthetic data, the algorithm achieves stable convergence. The Reddit dataset exhibits greater heterogeneity ($\mathrm{Var}(\varepsilon) = 0.087$) than synthetic data ($\mathrm{Var}(\varepsilon) = 0.018$), resulting in larger FL advantage. This confirms our theoretical prediction: amplification grows with $\mathrm{Var}(\varepsilon)$.

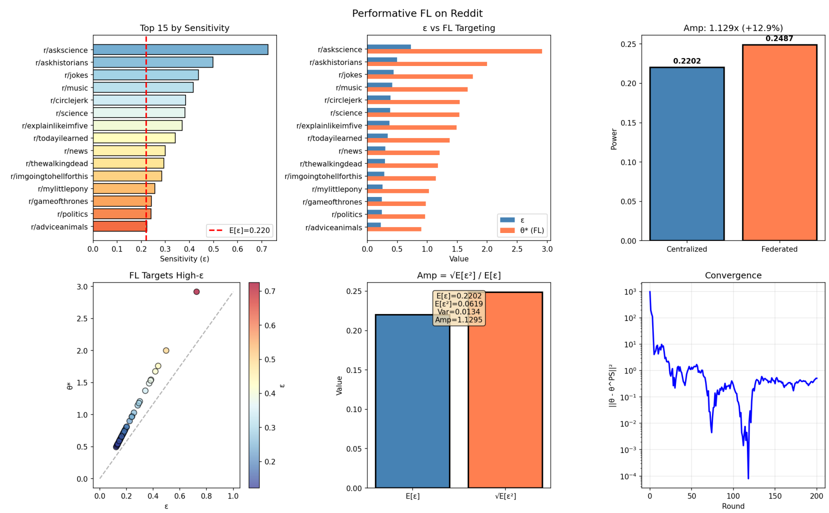

Figure 4: Performative Power comparison on Reddit data. FL achieves 1.13× the power of centralized training.

4.3 Discussion

Our results demonstrate three key findings:

1. FL structurally outperforms centralized training. This advantage stems not from privacy constraints but from FL's ability to learn heterogeneous client sensitivities and target accordingly.

2. Amplification increases with heterogeneity. Reddit's diverse communities yield larger FL advantage (12.9%) than synthetic data (9.2%), confirming the theoretical relationship between $\mathrm{Var}(\varepsilon)$ and amplification.

3. ParS enables principled resource allocation. High-sensitivity communities receive proportionally stronger model influence ($\theta_i^* \propto \varepsilon_i$), while low-sensitivity communities receive less, optimizing total Performative Power under fixed budget.

5 Conclusion

We have shown that Federated Learning can outperform centralized training for maximizing Performative Power. The FL-to-centralized power ratio equals $\sqrt{E[\varepsilon^2]}/E[\varepsilon] \geq 1$, growing with user heterogeneity. Our Participant Selection (ParS) strategy allocates model intensity proportionally to client sensitivity: $\theta_i^* = \varepsilon_i B/\sqrt{E[\varepsilon^2]}$.

Experiments on synthetic data and Reddit communities validate these findings, demonstrating up to 12.9% improvement in Performative Power for FL over centralized baselines.

5.1 Implications

Social media platforms can leverage FL not merely for privacy compliance but as a superior optimization strategy. The ability to learn and exploit user heterogeneity provides a structural advantage over uniform deployment. Platforms with more diverse user bases benefit most from this approach.

5.2 Ethical Considerations

While we frame Performative Power in terms of engagement optimization, the same techniques could target users in potentially harmful ways. Platforms deploying these methods should consider the ethical implications of sensitivity-based targeting, particularly for vulnerable populations.

5.3 Limitations and Future Work

Linear model assumption. Our analysis uses a linear Gaussian response model. Real user behavior may exhibit nonlinear responses, saturation effects, or threshold dynamics.

Sensitivity estimation. We assume sensitivities can be accurately estimated. In practice, this requires sufficient interaction data and may be noisy.

Multi-objective optimization. Real platforms optimize multiple metrics (engagement, revenue, retention). Extending ParS to multi-objective settings is a natural direction for future work.

Code Availability: `https://github.com/xkabot/PowerFL`

A Derivation of Optimal FL Targeting

A.1 Problem Setup

We seek to maximize total performative power under an L_2 budget constraint. Let α_i denote the weight (population fraction) of client i, and ε_i denote client i's sensitivity.

Centralized: Must deploy uniform θ to all clients.

$$\max_{\theta} \sum_i \alpha_i \varepsilon_i |\theta| \quad \text{s.t.} \quad \theta^2 \leq B^2 \tag{10}$$

Since $\sum_i \alpha_i = 1$, this simplifies to $\max_\theta E[\varepsilon] \cdot |\theta|$ subject to $|\theta| \leq B$. The solution is $\theta^* = B$, yielding:

$$P_{\text{central}} = E[\varepsilon] \cdot B \tag{11}$$

Federated: Can deploy different θ_i per client.

$$\max_{\theta_1,\ldots,\theta_N} \sum_i \alpha_i \varepsilon_i \theta_i \quad \text{s.t.} \quad \sum_i \alpha_i \theta_i^2 \leq B^2 \tag{12}$$

(We assume $\theta_i \geq 0$ without loss of generality since power depends on $|\theta_i|$.)

A.2 Lagrangian Solution

Form the Lagrangian with multiplier $\lambda \geq 0$:

$$\mathcal{L}(\theta, \lambda) = \sum_i \alpha_i \varepsilon_i \theta_i - \lambda \left(\sum_i \alpha_i \theta_i^2 - B^2 \right) \tag{13}$$

KKT Conditions:

1. *Stationarity:* Taking partial derivatives with respect to θ_i:

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = \alpha_i \varepsilon_i - 2\lambda \alpha_i \theta_i = 0 \tag{14}$$

Solving for θ_i:

$$\theta_i = \frac{\varepsilon_i}{2\lambda} \tag{15}$$

2. *Complementary Slackness:* $\lambda \left(\sum_i \alpha_i \theta_i^2 - B^2 \right) = 0$
3. *Primal Feasibility:* $\sum_i \alpha_i \theta_i^2 \leq B^2$

A.3 Solving for λ

Assume the constraint $\lambda > 0$. Substituting $\theta_i = \varepsilon_i / 2\lambda$ into the constraint:

$$\sum_i \alpha_i \left(\frac{\varepsilon_i}{2\lambda} \right)^2 = B^2 \tag{16}$$

$$\frac{1}{4\lambda^2} \sum_i \alpha_i \varepsilon_i^2 = B^2 \tag{17}$$

$$\frac{1}{4\lambda^2} E[\varepsilon^2] = B^2 \tag{18}$$

$$\lambda = \frac{\sqrt{E[\varepsilon^2]}}{2B} \tag{19}$$

A.4 Optimal Solution

Substituting λ back into $\theta_i = \varepsilon_i / 2\lambda$:

$$\theta_i^* = \frac{\varepsilon_i}{2 \cdot \frac{\sqrt{E[\varepsilon^2]}}{2B}} = \frac{\varepsilon_i \cdot B}{\sqrt{E[\varepsilon^2]}} \tag{20}$$

This is the *Participant Selection* (ParS) strategy: allocate model intensity proportionally to client sensitivity ε_i.

A.5 FL Performative Power

The resulting performative power is:

$$\begin{aligned} P_{\text{FL}} &= \sum_i \alpha_i \varepsilon_i \theta_i^* && (21) \\ &= \sum_i \alpha_i \varepsilon_i \cdot \frac{\varepsilon_i \cdot B}{\sqrt{E[\varepsilon^2]}} && (22) \\ &= \frac{B}{\sqrt{E[\varepsilon^2]}} \sum_i \alpha_i \varepsilon_i^2 && (23) \\ &= \frac{B \cdot E[\varepsilon^2]}{\sqrt{E[\varepsilon^2]}} && (24) \\ &= B \cdot \sqrt{E[\varepsilon^2]} && (25) \end{aligned}$$

A.6 Amplification Factor

The ratio of FL to centralized power is:

$$\text{Amplification} = \frac{P_{\text{FL}}}{P_{\text{central}}} = \frac{\sqrt{E[\varepsilon^2]}}{E[\varepsilon]} \tag{26}$$

Using the variance decomposition $E[\varepsilon^2] = \text{Var}(\varepsilon) + E[\varepsilon]^2$:

$$\text{Amplification} = \sqrt{\frac{E[\varepsilon^2]}{E[\varepsilon]^2}} = \sqrt{1 + \frac{\text{Var}(\varepsilon)}{E[\varepsilon]^2}} \tag{27}$$

A.7 Bound via Cauchy-Schwarz

Claim: Amplification ≥ 1, with equality iff $\text{Var}(\varepsilon) = 0$.

Proof: By the Cauchy-Schwarz inequality:

$$E[\varepsilon]^2 = \left(\sum_i \alpha_i \varepsilon_i\right)^2 \leq \left(\sum_i \alpha_i\right)\left(\sum_i \alpha_i \varepsilon_i^2\right) = 1 \cdot E[\varepsilon^2] = E[\varepsilon^2] \tag{28}$$

Therefore:

$$\text{Amplification} = \frac{\sqrt{E[\varepsilon^2]}}{E[\varepsilon]} \geq \frac{\sqrt{E[\varepsilon]^2}}{E[\varepsilon]} = 1 \tag{29}$$

Equality holds iff $\varepsilon_i = c$ for all i (i.e., all clients have identical sensitivity), which implies $\mathrm{Var}(\varepsilon) = 0$. □

References

[1] *Federated learning: Overview, strategies, applications, tools and future directions.* National Library of Medicine. DOI: 10.1016/j.heliyon.2024.e38137. 2024.

[2] Moritz Hardt, Meena Jagadeesan, and Celestine Mendler-Dünner. *Performative Power**. arXiv. arXiv:203.17232v2. 2022. URL: `https://arxiv.org/pdf/2203.17232`.

[3] Kun Jin et al. *Performative Federated Learning:A Solution to Model-Dependent and Heterogeneous Distribution Shifts.* arXiv preprint. arXiv:2305.05090. 2023. URL: `https://arxiv.org/pdf/2305.05090`.