

Federated Learning Client Participation Rates and Zero-Shot VQA-v2 Accuracy in Multimodal Models

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: How does varying the client participation rate in federated learning impact the zero-shot accuracy of multimodal vision-language models on the VQA-v2 benchmark. It is well understood that client-master communication can be a primary bottleneck in Federated Learning. In this work, we address this issue with a novel client subsampling scheme, where we restrict the number of clients allowed to communicate their updates back to the master. 9 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Optimal Client Sampling for Federated Learning. Research question: How does varying the client participation rate in federated learning impact the zero-shot accuracy of multimodal vision-language models on the VQA-v2 benchmark?.

2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.8/10.

3 Results

8 papers retrieved. 9 claims extracted; 0 independently verified. Quality review score: 3.8/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The Federated EMNIST (FEMNIST) image dataset is a well-balanced dataset with data of similar quality on each client.	×	0.06
The unchanged EMNIST validation set consists of 40,832 images.	×	0.02
Dataset 1 has 32,906 training images in total.	×	0.01
10% of clients in Dataset 1 hold 82% of the training images.	×	0.02
Dataset 2 has 29,906 training images in total.	×	0.01
20% of clients in Dataset 2 hold 90% of the training images.	×	0.02
Dataset 3 has 27,599 training images in total.	×	0.01
50% of clients in Dataset 3 hold 98% of the training images.	×	0.02
The code and datasets for the experiments are available at https://github.com/SamuelHorvath/FL-optimal-client-sampling .	×	0.08

References

- <http://arxiv.org/abs/2506.02887v2>
- <http://arxiv.org/abs/2010.13723v3>
- <http://arxiv.org/abs/2509.20877v1>