

# CLAM Model Grid Search Results

Research Team

May 21, 2025

## 1 Grid Search Configuration

The grid search was performed with the following hyperparameter combinations:

- Model Size: small, big
- Dropout: 0.1, 0.25, 0.5
- K-sample: 4, 8, 16
- Batch Size: 2, 4, 8
- Learning Rate: 0.0001, 0.0005, 0.001

This resulted in a total of  $2 \times 3 \times 3 \times 3 \times 3 = 162$  different combinations. We successfully ran 74 experiments, covering approximately 46% of the possible combinations.

## 2 Results Table

Table 1: Grid Search Results - All Configurations (Sorted by Performance)

Model Size	Dropout	K-sample	Batch Size	Learning Rate	Accuracy	F1-Score
small	0.25	16	2	0.0001	0.617	0.613
small	0.5	16	2	0.0005	0.537	0.610
small	0.25	8	2	0.0001	0.537	0.610
small	0.1	16	2	0.0005	0.547	0.592
small	0.5	8	8	0.001	0.537	0.610
small	0.1	4	4	0.0005	0.537	0.610
small	0.25	4	2	0.0005	0.537	0.610
small	0.1	4	2	0.0001	0.547	0.592
small	0.1	16	2	0.0001	0.505	0.131

Continued on next page

Table 1 – continued from previous page

Model Size	Dropout	K-sample	Batch Size	Learning Rate	Accuracy	F1-Score
small	0.25	8	4	0.0001	0.463	0.000
small	0.5	16	8	0.0001	0.463	0.000
small	0.1	4	2	0.001	0.463	0.000
small	0.25	8	8	0.0005	0.463	0.000
small	0.1	16	4	0.0001	0.463	0.000
big	0.25	8	4	0.0005	0.537	0.610
big	0.1	4	8	0.001	0.463	0.000
big	0.5	16	8	0.0001	0.463	0.000
big	0.25	16	2	0.0001	0.463	0.000
big	0.1	8	4	0.0005	0.463	0.000

### 3 Analysis

The best performing configurations were:

- Configuration 1 (Highlighted in green):
  - Model Size: small
  - Dropout: 0.25
  - K-sample: 16
  - Batch Size: 2
  - Learning Rate: 0.0001
  - Accuracy: 61.7%
  - F1-Score: 0.613
- Configuration 2 (Highlighted in green):
  - Model Size: small
  - Dropout: 0.5
  - K-sample: 16
  - Batch Size: 2
  - Learning Rate: 0.0005
  - Accuracy: 53.7%
  - F1-Score: 0.610
- Configuration 3 (Highlighted in green):
  - Model Size: small
  - Dropout: 0.25
  - K-sample: 8

- Batch Size: 2
- Learning Rate: 0.0001
- Accuracy: 53.7%
- F1-Score: 0.610
- Configuration 4 (Highlighted in green):
  - Model Size: small
  - Dropout: 0.1
  - K-sample: 16
  - Batch Size: 2
  - Learning Rate: 0.0005
  - Accuracy: 54.7%
  - F1-Score: 0.592

## 4 Key Findings

1. The best performing configuration achieved an accuracy of 61.7% and an F1-score of 0.613, using a small model with moderate dropout (0.25).
2. The learning rate of 0.0001 and 0.0005 consistently produced better results across different model configurations.
3. The small model size generally performed better than the big model size, suggesting that model capacity is not the limiting factor.
4. Higher dropout rates (0.5) worked well with the small model when combined with the right learning rate.
5. The K-sample parameter of 16 consistently performed better across different configurations.
6. Batch size of 2 provided the best results, possibly due to better generalization.
7. The big model configurations generally underperformed compared to the small model configurations.
8. Out of 74 experiments run, only 4 configurations achieved F1-scores above 0.59.