Train Perplexity

Perplexity (y-axis, log scale) vs Epochs (x-axis)

Legend:
- TopK, lr=0.001
- TopK_EF, lr=0.0001
- TopK_SCAM, lr=0.001
- ImpK_s_SCAM, lr=0.001