Problem/Data	Recommended Optimizers	Reasoning
Large Datasets	SGD, Momentum, Adam	Efficient for handling large volumes of data
Sparse Datasets	Adagrad, RMSprop, Adadelta	Adaptive learning rates work well with sparse features
Non-Stationary Data	RMSprop, Adadelta, Adam	Robust to changing data distributions
Complex Error Surfaces	NAG, Adam	Better handling of non-convex optimization landscapes
Deep Learning	Adam (often default)	Popular choice due to good performance and ease of use
Simple Models/Starting Point	SGD	Easy to implement and understand, good baseline