

Optimizer

Fixed Learning Rate Momentum Gradient

SGDP (2021)

SGD (1971)

LION (2023)

(a)

Adaptive Learning Rate Momentum Gradient

Adam (2015)

AdaMax (2015)

AdamP (2021)

AdamW (2019)

LAMB (2020)

Nadam (2018)

RAdam (2020)

Adan (2023)

(b)

Estimated Learning Rate Momentum Gradient

AdaBelief (2019)

AdaBound (2019)

AdaFactor (2018)

LARS (2018)

NovoGrad (2020)

Sophia (2023)

(c)

Adaptive Learning Rate Gradient

AdaGrad (2011)

AdaDelta (2012)

RMSProp (2012)

(d)