Algorithm 1 Stochastic Gradient Descent (SGD) with Constant Step Size and Stopping Criterion

Require: Objective function $f(x)$, gradient $\nabla f_I(x)$ for mini-batch I of size s, step size $\eta > 0$, initial point x_0, tolerance $\epsilon > 0$ maximum iterations $max - iter$

Initialize $x \leftarrow x_0$
Initialize $s \leftarrow s_0$
Initialize $\eta \leftarrow \eta_0$
Initialize maximum iterations $max - iter = 10000$
Initialize $\epsilon = 10^{-6}$
for $k = 1$ to $max - iter$ **do**
  Sample a mini-batch I_k of size s uniformly at random from the set $\{1, \cdots, n\}$
  Compute stochastic gradient $g_k \leftarrow \nabla f_{I_k}(x)$
  Update parameters: $x \leftarrow x - \eta g_k$
  if $\|g_k\|_2 < \epsilon$ **then**
    break
  end if
end for
return x