Variable Learning Rate

- Idea: speed up the convergence by increasing the learning rate on flat surfaces and decreasing it when the slope increases.
- Update the weights.
- Calculate the squared error (over the entire training set).
- If the error increases by more than a predefined % $\theta$ (1-5%):
  - Weight update is discarded.
  - Learning rate is decreased by some factor ($1 > \alpha > 0$) [$\alpha = 5\%$ typically].
  - Momentum is set to 0.
- If the error increases by less than $\theta$:
  - Weight update is accepted.
  - Learning rate is unchanged.
  - If momentum has been set to 0, it is reset to its original value.
- If the error decreases:
  - Weight update is accepted.
  - Learning rate is increased by some factor $\beta > 1$.
  - If momentum has been set to 0, it is reset to its original value.