Data: high dimensional data $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subseteq \mathbb{R}^D$, Target Perplexity P, Number of iterations T
Result: low dimensional data $\mathbf{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_N\} \subseteq \mathbb{R}^d$ such that $d \ll D$ (usually d = 2 or 3)
subtract mean $\forall i = 0..N \quad \mathbf{x}_i = \mathbf{x}_i - \bar{\mathbf{x}}$;
recale data $\forall i = 0..N$, $k = 0..D \quad (\mathbf{x}_i)_k = (\mathbf{x}_i)_k \,/\, \max_{i',k'} (\mathbf{x}_{i'})_{k'}$;
compute squared eucledian distances $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|^2$;
for $i = 0..N$ **do**
    initialize σ_i;
    repeat
        $\forall j = 0..N \quad p_{j|i} = \frac{\exp(-d_{ij}/2\sigma_i^2)}{\sum_{k \neq i} \exp(-d_{ik}/2\sigma_i^2)}$;
        $\mathrm{per}_i = 2^{\sum_j p_{j|i} \log_2 p_{j|i}}$;
    until $per_i = P$;
end
symmetrize $p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N}$;
sample inital $\mathbf{Y}^{(0)} = \{\mathbf{y}_1^{(0)}, \ldots, \mathbf{y}_N^{(0)}\}$ from gaussian distribution;
for $t = 1..T$ **do**
    compute $\forall i, j = 0..N \quad q_{ij} = \frac{(1 + \|\mathbf{y}_i^{(t-1)} - \mathbf{y}_j^{(t-1)}\|^2)^{-1}}{\sum_{k \neq l} (1 + \|\mathbf{y}_k^{(t-1)} - \mathbf{y}_l^{(t-1)}\|^2)^{-1}}$;
    compute $\forall i = 0..N \quad \frac{\partial C}{\partial \mathbf{y}_i^{(t-1)}} = 4 \sum_j (p_{ij} - q_{ij})(\mathbf{y}_i^{(t-1)} - \mathbf{y}_j^{(t-1)})(1 + \|\mathbf{y}_i^{(t-1)} - \mathbf{y}_j^{(t-1)}\|^2)^{-1}$;
    update $\forall i = 0..N \quad \mathbf{y}_i^{(t)} = \mathbf{y}_i^{(t-1)} + \eta \frac{\partial C}{\partial \mathbf{y}_i^{(t-1)}} + \alpha(t)(\mathbf{y}_i^{(t-1)} - \mathbf{y}_i^{(t-2)})$;
    subtract mean $\forall i = 0..N \quad \mathbf{y}_i^{(t-1)} = \mathbf{y}_i^{(t-1)} - \bar{\mathbf{y}}^{(t-1)}$;
end

Algorithm 1: t-SNE Algorithm