

核心思想为：分区间进行逻辑回归

$$f(x) = \sum_{i=1}^m \frac{e^{\mu_i \cdot x}}{\sum_{j=1}^m e^{\mu_j \cdot x}} \cdot \frac{1}{1 + e^{-w_i \cdot x}}$$

所以搭建模型时只需要考虑两个部分：一部分是softmax，一部分是LR，前向传播时，输入分别经过两部门后进行相乘得到最终输出