MIPS and related problems in ML

Manohar, Saketh

March 31, 2018

Extreme Classification application: Let $\mathcal{X} \subset \mathbb{R}^n$ (n is in millions) and $\mathcal{Y}$ be an extremely large discrete set. ($c \equiv |\mathcal{Y}|$ is in millions). You are given a initial labeled set of examples $(x_i, y_i), i = 1, \ldots, m$, using which model parameters $w_1, \ldots, w_c$ for each class are learned using some standard algorithm. Note that this is a multi-labeled classification. Hence y is a vector representing all relevant labels.

You are additionally given a set of unlabeled examples (again millions), $z_1, \ldots, z_u$. Let $t \in \mathbb{N}$ be a user defined (given) parameter. At beginning of iteration k, let the model parameters be $w_1^{(k)}, \ldots, w_c^{(k)}$. In iteration k do the following:

1. For each z_i:
   (a) Among the values $\langle w_1^{(k)}, z_i \rangle, \ldots, \langle w_c^{(k)}, z_i \rangle$, find the class/index with the t^{th} highest value of the inner-product. Let this class be p.
   (b) Among the values $\langle w_1^{(k)}, z_i \rangle, \ldots, \langle w_c^{(k)}, z_i \rangle$, find the class/index with the $(t+1)^{th}$ highest value of the inner-product. Let this class be q.
   (c) Let $\Delta_i \equiv \langle w_p^{(k)}, z_i \rangle - \langle w_q^{(k)}, z_i \rangle$. This intuitively represents the 'gap' at the t^{th} position.
2. Find the i with least Δ_i. Let this be z_*.
3. Ask the oracle the label for this z_*. Include this example and retrain, thus updating the model parameters.