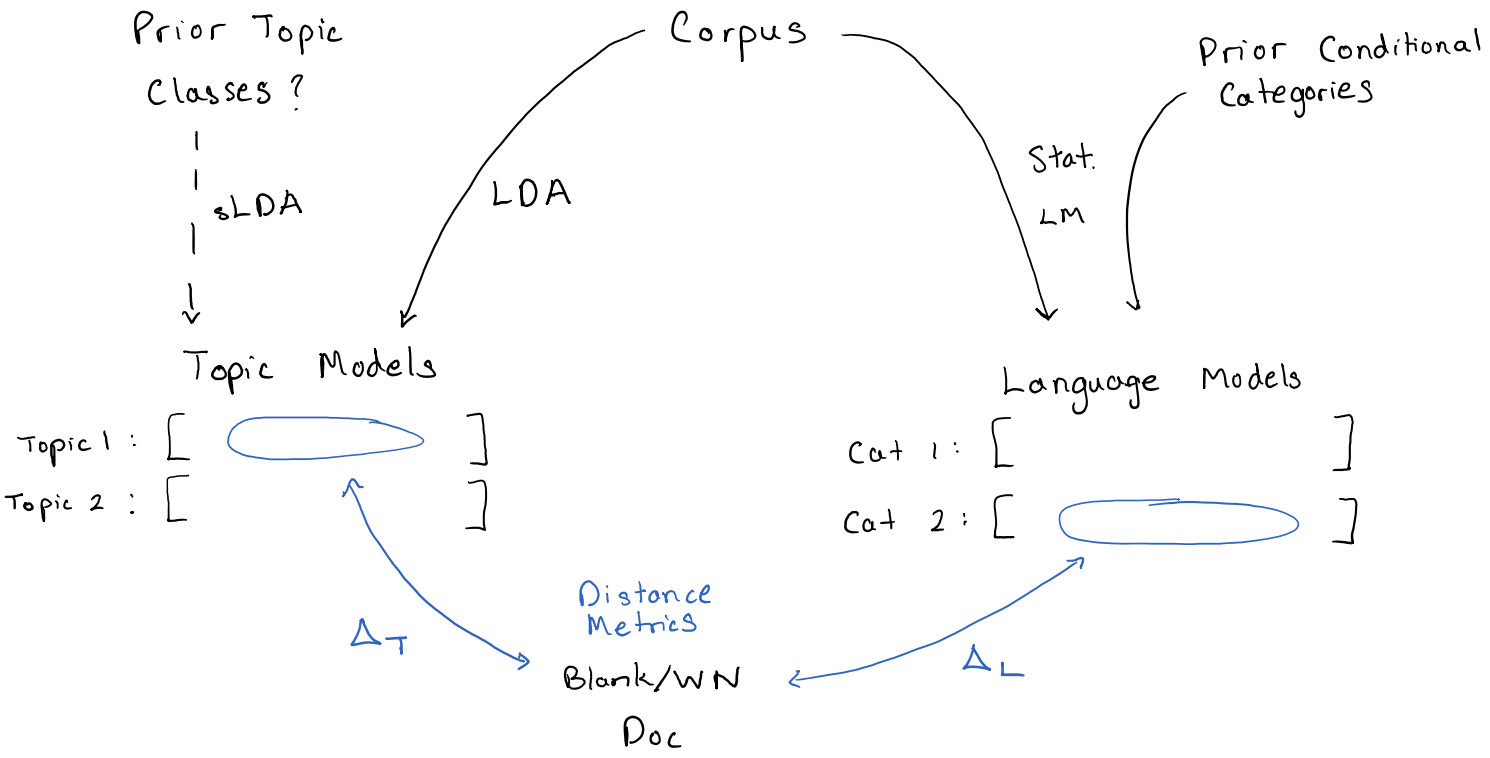
Project Planning

Sunday, April 1, 2018 7:09 PM



Distance Metrics:

$D = [d_1, d_2, \ldots, d_N]$

$T_i = [t_{i_1}, t_{i_2}, \ldots, t_{iK}]$

Word Net-weighted likelihood:

word-level: d_j, T_i

if d_j in T_i : $T_i(d_j) = p(d_j|T_i)$

else: find $\underset{t_{ik}}{\text{argmin}}\ WN(d_j, t_{ik}) = t_j \rightarrow \dfrac{T_i(t_j)}{WN(d_j, t_j)} = p(d_j|T_i)$

$$\frac{T_i(\underset{t_{ik}}{\text{argmin}}\ WN(d_j, t_{ik}))}{1 + WN(d_j, t_j)} = \tilde{p}(d_j|T_i)$$

Document-level: D, T_i

$$p(T_i|D) = \prod_D \tilde{p}(d_j|T_i)$$

For target topic T^*, maximize $p(T^*|D) = \prod_D \tilde{p}(d_j|T^*)$

Language Model:

Maximize likelihood $\underset{D}{\operatorname{argmax}}\ p(D \mid L_i)$ given language L_i