Supplementary Material for:
Neural Topic Model with Attention for Supervised Learning

1 Detailed model inference

Starting from Equation (8), we can perform the reparameterization trick as below:

$$\begin{aligned}
\log p_{\Theta,\Psi}(l,\boldsymbol{d}) &= \log \int_{\boldsymbol{t}} \frac{p_\Theta(\boldsymbol{t})}{q_\Phi(\boldsymbol{t}|\boldsymbol{d})} q_\Phi(\boldsymbol{t}|\boldsymbol{d}) p_\Psi(l|\boldsymbol{t}) p_\Theta(\boldsymbol{d}|\boldsymbol{t}) d\boldsymbol{t} \\
&= \log \mathbb{E}_{q_\Phi(\boldsymbol{t}|\boldsymbol{d})}\big[\frac{p_\Theta(\boldsymbol{t})}{q_\Phi(\boldsymbol{t}|\boldsymbol{d})} p_\Psi(l|\boldsymbol{t}) p_\Theta(\boldsymbol{d}|\boldsymbol{t})\big] \\
&= \mathbb{E}_{q_\Phi(\boldsymbol{t}|\boldsymbol{d})}\big[\log p_\Theta(\boldsymbol{d}|\boldsymbol{t}) - \log \frac{q_\Phi(\boldsymbol{t}|\boldsymbol{d})}{p_\Theta(\boldsymbol{t})} + \log p_\Psi(l|\boldsymbol{t})\big] \\
&= \mathbb{E}_{q_\Phi(\boldsymbol{t}|\boldsymbol{d})}\big[\log p_\Theta(\boldsymbol{d},\boldsymbol{t}) - \log q_\Phi(\boldsymbol{t}|\boldsymbol{d}) + \log p_\Psi(l|\boldsymbol{t})\big] \\
&\quad + \mathrm{KL}(q_\Phi(\boldsymbol{t}|\boldsymbol{d})||p_\Theta(\boldsymbol{t}|\boldsymbol{d})) \qquad (1)
\end{aligned}$$

Where Ψ represents all the parameter from the RNN attention model. Θ are the generative parameters and Φ are the variational parameters. μ_0, σ_0 are omitted because they are constants. Since $\beta \subset \Theta$, β is omitted too.

Since KL-divergence is always non-negative, we construct the variational objective function, also called the evidence lower bound (ELBO) of $\log p_{\Theta,\Psi}(l,\boldsymbol{d})$ as below:

$$\begin{aligned}
\mathcal{L} =& \mathbb{E}_{q_\Phi(\boldsymbol{t}|\boldsymbol{d})}\big[\log p_\Theta(\boldsymbol{d},\boldsymbol{t}) - \log q_\Phi(\boldsymbol{t}|\boldsymbol{d}) + \log p_\Psi(l|\boldsymbol{t})\big] \\
=& \mathbb{E}_{q_\Phi(\boldsymbol{t}|\boldsymbol{d})}\big[\log p_\Theta(\boldsymbol{d}|\boldsymbol{t}) + \log p_\Psi(l|\boldsymbol{t})\big] \\
& - \mathrm{KL}(q_\Phi(\boldsymbol{t}|\boldsymbol{d})||p(\boldsymbol{t})) \qquad (2)
\end{aligned}$$