Studies on Topic Models of Document Corpora

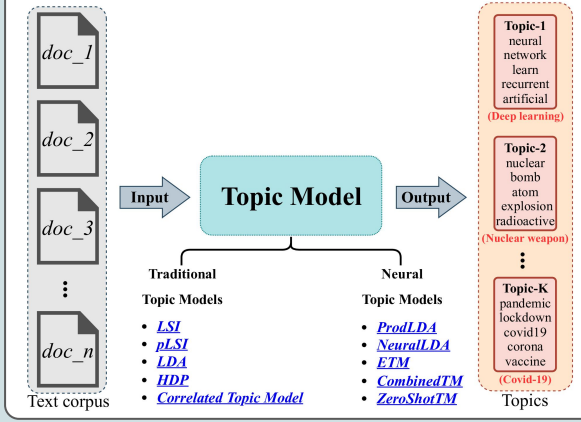
IACS, Kolkata

Suman Adhya

Supervised by: Dr. Debarshi Kumar Sanyal

acm India Council

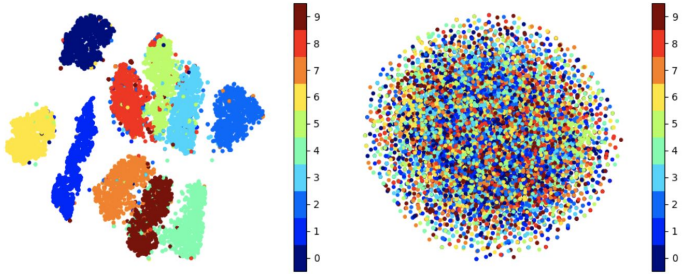
Introduction to Topic Modeling



Key Objectives of This Doctoral Work

- **Enhancing Negative Sampling for NTMs** – *IEEE Transactions on AI, ICON 2022.*
- **Incorporating Document Graph Representations in Topic Modeling** – *NAACL 2024.*
- **Knowledge Distillation for NTM Compression** – *ECIR 2023.*
- **Impact of Dropout Rates on NTM Performance** – *EACL 2023.*
- **LLM-Based Dynamic Topic Modeling** – *Ongoing.*
- **Addressing Posterior Collapse in VAEs** – *Ongoing.*

Ongoing Project: Mitigating Posterior Collapse Issue in VAEs



Posterior collapse occurs when the posterior distribution matches the prior distribution, leading to a loss of information about the latent structure of the input data.

$$\mathcal{O}_{\text{ELBO}}(\theta,\phi) := \mathbb{E}_{p_{\mathcal{D}}(\mathbf{x})}\left[\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z})]\right] - \mathcal{I}(\mathbf{x},\mathbf{z}) - \text{KL}(q_\phi(\mathbf{z})\|p(\mathbf{z}))$$

$$\therefore\uparrow \textbf{ELBO} \Rightarrow\downarrow \textbf{M.I. between } \mathbf{x} \textbf{ and } \mathbf{z}$$