

E0 334 - Deep Learning for NLP

Assignment 4

Validation set results due by 15th Nov, 09:59 PM

Note: Use the following link for submitting your results of Assignment 4.

TBA

Problem:

In this assignment, you will study different topic modeling techniques designed using LLMs on the following datasets:

1. Dataset for Topic Modeling for Research Articles (available at <https://www.kaggle.com/datasets/blessondensil294/topic-modeling-for-research-articles>)
2. Trip Advisor Dataset (available at <https://www.kaggle.com/datasets/arnabchaki/tripadvisor-reviews>)

The references for this assignment include:

1. M. Groothendorst, BERTopic, 2022. Available at <https://arxiv.org/pdf/2203.05794>
2. Pham et al, TopicGPT, 2024. Available at <https://arxiv.org/pdf/2311.01449v2>
3. Wang et al, Prompting Large Language Models for Topic Modeling, 2023. Available at <https://arxiv.org/pdf/2312.09693v1>

Note: Evaluation metrics and assignment submission links will be provided later.