

# Exploring Deep Learning Models for Sentiment Classification in Customer Reviews

**Kazi Habibur Rahaman (20101559)**

## Proposed Tasks:

- Sentiment classification using CNN model [1] (Kim, 2014)
- Sentiment classification using RNN model (LSTM network) [2] (Wang, J., 2018)
- Sentiment classification using Transformer [3] ((Yang et al., 2019))

## Dataset:

- Training Data: IMDB Dataset of 50K Movie Reviews
  - <https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>
- Training Data: 515K Hotel Reviews Data in Europe
  - <https://www.kaggle.com/jiashenliu/515k-hotel-reviews-data-in-europe>
- Testset: Tripadvisor Bangladesh attraction spots reviews
  - <https://drive.google.com/file/d/16tRD77sLtztMetzO-VecMZHEh9EisSut/view?usp=sharing>

## Contribution:

- A comparison between CNN, RNN and Transformer
- Find out the best deep learning model depending on real life data, dataset size and performance.

## Real Life Applications:

- Find the best locations to visit depending on their sentiment classification counts
- Tune model for single location to evaluate the spot
- Not only location, the research paper will also reveal which model should be preferred for sentiment analysis

References:

Kim, Y. (2014, August 25). [1408.5882] Convolutional Neural Networks for Sentence Classification. arXiv. Retrieved December 5, 2023, from <https://arxiv.org/abs/1408.5882>

Wang, J., Liu, T., Luo, X., & Wang, L. (2018). An LSTM Approach to Short Text Sentiment Classification with Word Embeddings. Retrieved December 5, 2023, from <https://aclanthology.org/O18-1021/>

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Ruslan Salakhutdinov, R., & Le, Q. V. (2019, June 19). [1906.08237] *XLNet: Generalized Autoregressive Pretraining for Language Understanding*. arXiv. Retrieved December 5, 2023, from <https://arxiv.org/abs/1906.08237>