Similarity-Aware Deep Attentive Model for Clickbait Detection

GRU’s:

GRU is a simplified version of Long Short-Term Memory (LSTM), which is an improved version of recurrent neural networks (RNN’s). RNN’s are unable to learn long term dependencies as effectively.

Recurrent neural networks (RNN) consider past information or memory for classification, but RNN’s face an issue called vanishing gradient problem and are unable to learn long-term dependencies. GRU is a simplified version of Long-Short term memory as it has a fewer number of gates (gates are neural network layers that determine which information should be kept or forgotten). GRU also has lesser matrix and tensor operation compared to LSTM making it’s training faster.

GRU works using 2 gates: update gate and reset gate. The update gate is responsible for determining what information about a new entry is retained or forgotten. In contrast, the reset gate determines how much the past data will be forgotten.

Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.35

He, T., Droppo, J.: Exploiting lstm structure in deep neural networks for speech recognition. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 5445–5449. IEEE (2016)