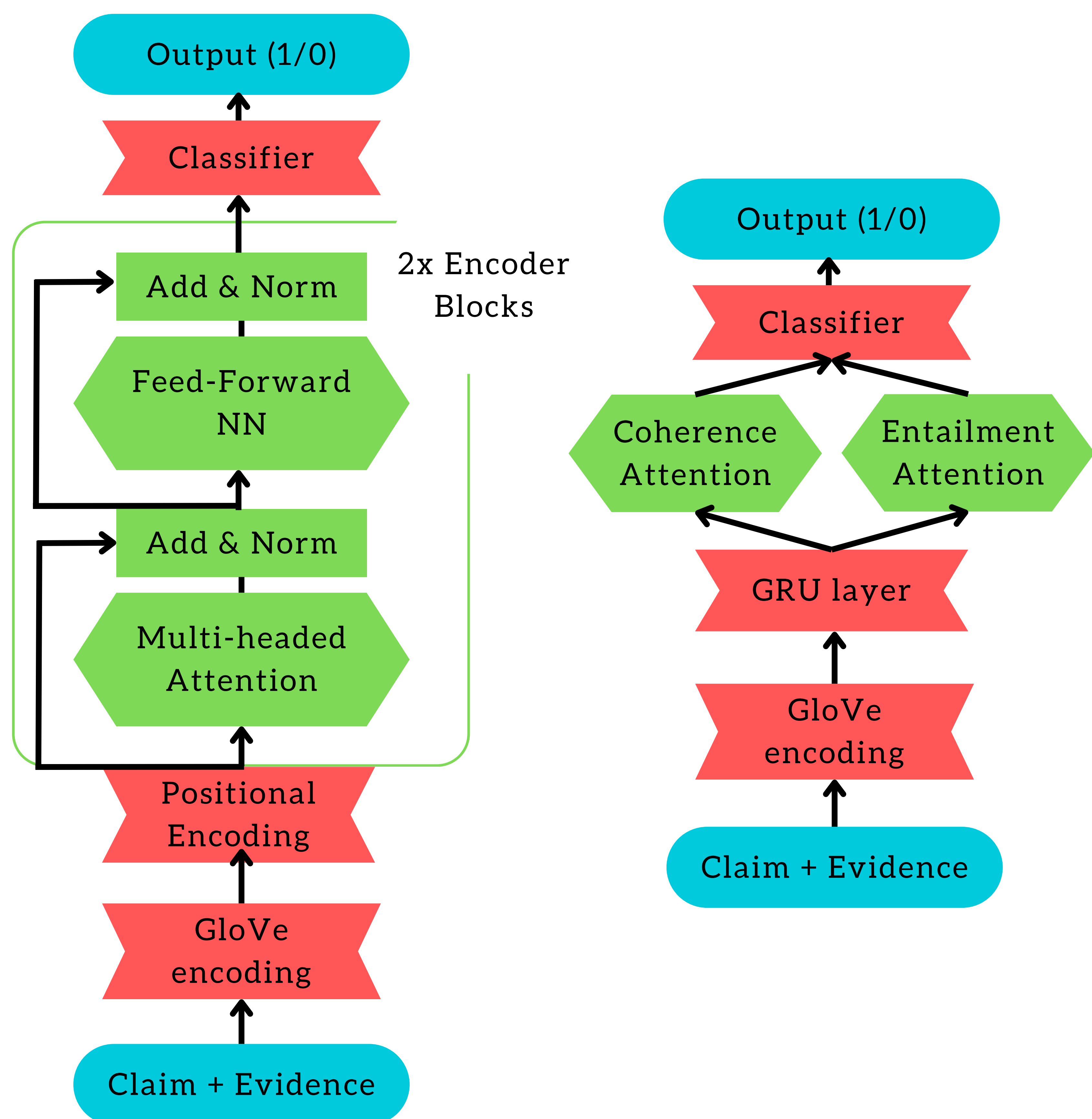


GROUP 70 - EVIDENCE DETECTION

By Chong Zhe Khang & Daniel Dobzinski

THE MODELS



APPROACH

We have implemented two approaches, each of which, given a claim and a piece of evidence, determines whether the evidence is relevant to the claim. One approach uses a transformer model, and the other employs an attention network.

INSPIRATIONS

δ -GClip

Also known as Delta-Gradient Clipping is an optimization algorithm which mitigates the exploding gradient problem as well as placing a lower bound on the update step to prevent updates becoming 0 if gradient of the loss function with respect to a parameter is tending towards ∞ .

Attention-based Sequence Processing Model

The attention network approach is inspired by "Sentence-level Evidence Embedding for Claim Verification with Hierarchical Attention Networks" paper published by Singapore Management University. The paper presents a model where each sentence of evidence is combined with the claim and processed using Coherence Attention. The resulting outputs are then combined using Entailment Attention.

WHY GLOVE?

GloVe is a specific type of word embedding developed by researchers at Stanford. It is designed to leverage both global statistics and local context in its training process. GloVe comes with several pre-trained models using various corpora and dimensions. For instance, models trained on Wikipedia and Gigaword are popular for general NLP tasks. These pre-trained vectors can be downloaded and used directly, saving significant time and computational resources.

TRAINING PROCESS

