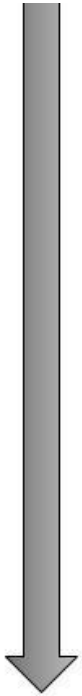
Mini Aria and A Billion Block Diagrams

Other mini ideas before meeting The One

- Case Citation Count based Concept Graph
- Solr Recall Experiments
- Identifying Legal Issue

Preprocessing Text



CLEAN UP

CONVERT TO WORDS

CREATE DICTIONARIES

IND2WORD
WORD2IND

PRETRAINED EMBEDDINGS

IND2VEC

CLEAN UP

CONVERT TO WORDS

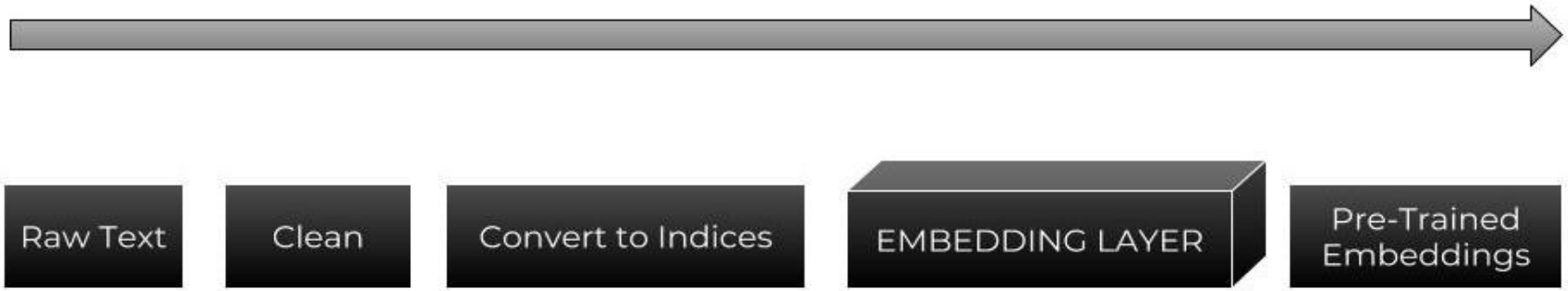
WORD2IND

Terminology

- **Sequence Length:** Length of the input sequence / sentence
- **Batch Size:** Number of sequences bundled for one forward pass
- **Forward Pass:** Sending the input through the model
- **Backward Pass:** Calculate gradients of Loss function wrt model parameters
- **Gradient Descent / Step:** Update model parameters in backward pass
- **Optimizer:** Algorithm that does the step, e.g SGD, Adam, etc.
- **Epoch/ Iteration:** Do a step for training data
- **Layer:** Abstraction of a particular deep learning functionality
- **Model:** Architecture created by stacking layers
- **Loss Function:** Math equation which helps train the model
    - Penalizes the model when it gets something wrong
    - Minimal value when model gets it right

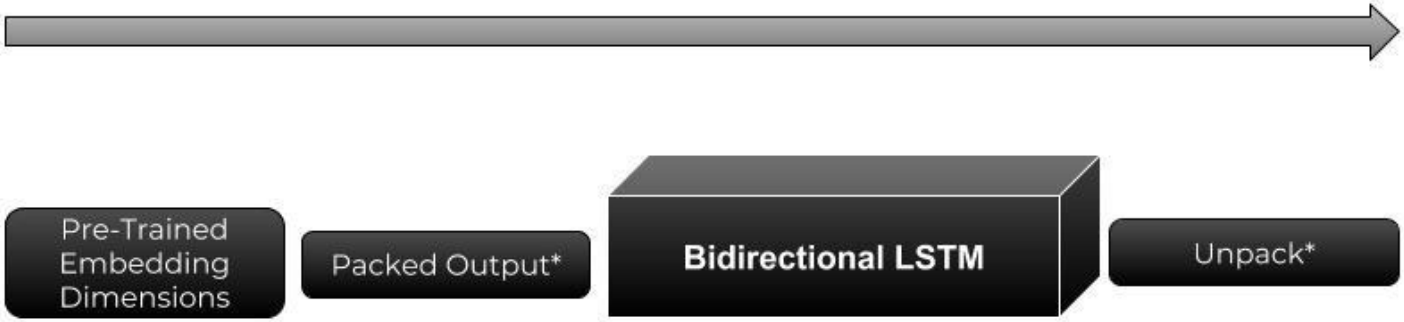
Layers in the Model

Embedding Layer



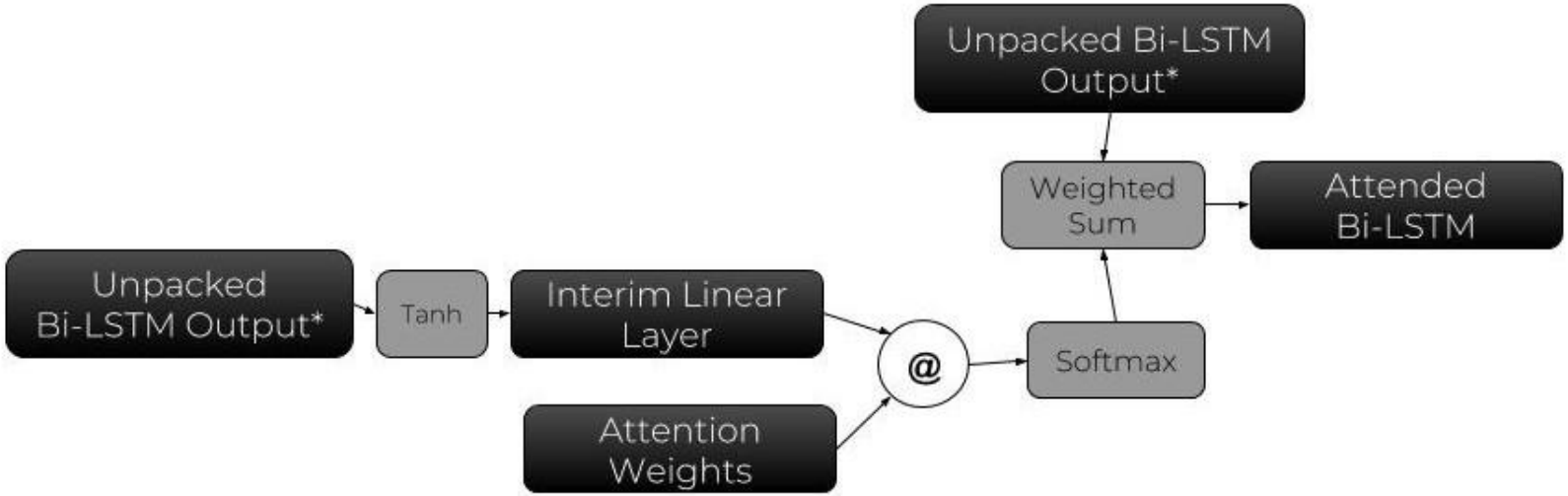
- Input: Indexed Text - Output: Pre-Trained or Random Embeddings
- Input Dimensions: seq_len * batch_size
- Output Dimensions: seq_len * batch_size * Pre-Trained_Embedding_Size
- Layer Dimensions: Vocabulary_Size * Pre-Trained_Embedding_Size

LSTM



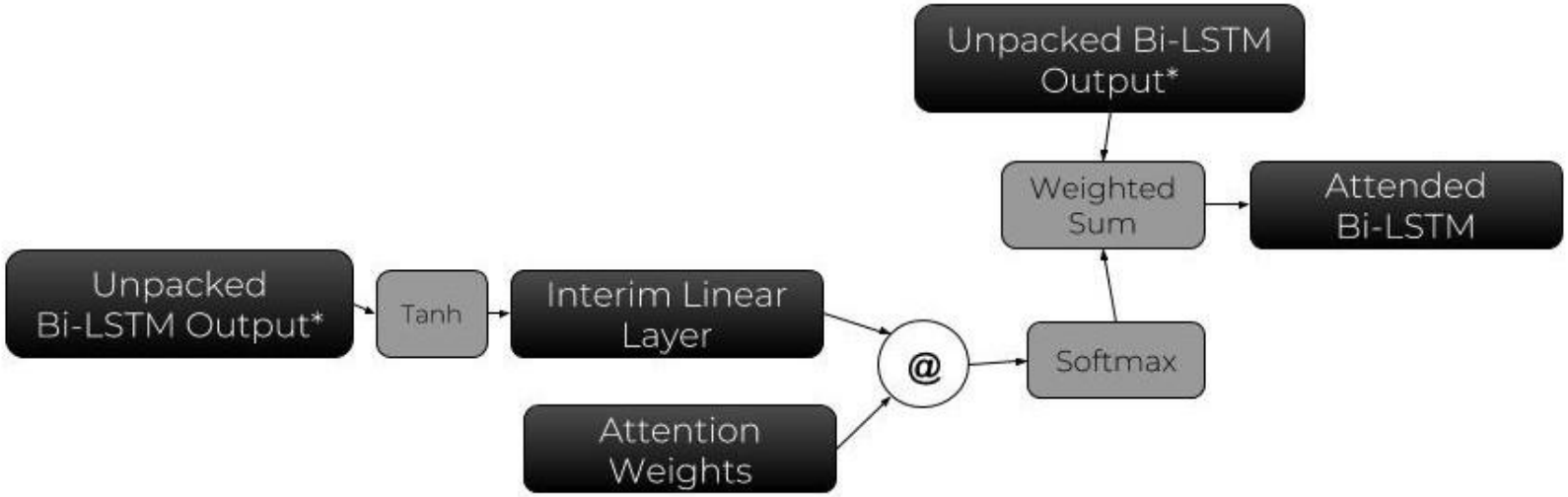
- LSTM_Dimensions: (num_LSTM_layers * num_hidden_units * num_directions)
- Input: Embeddings - Output: Hidden states
- Input Dimensions: seq_len * batch_size * PreTrained_Embedding_Size
- Output Dimensions: seq_len * batch_size * LSTM_Dimension
- Layer Dimensions: PreTrained_Embedding_Size * LSTM_Dimension

Attention Model



- Input: BiLSTM Hidden Units
- Input Dimensions: seq_len * batch_size * LSTM_Dimensions
- Output: Weighted or Attended BiLSTM Output
- Output Dimensions: seq_len * batch_size * LSTM_Dimension

Attention Model

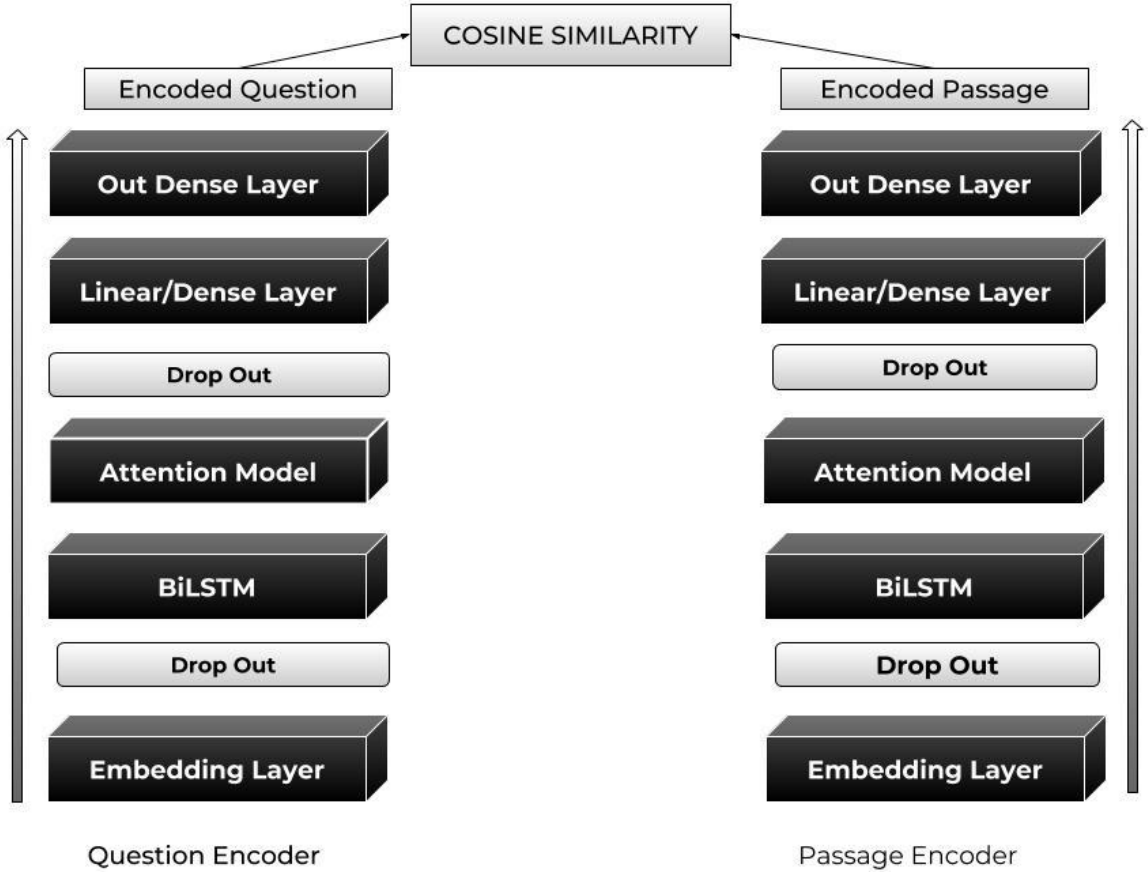


- Interim Linear Layer: LSTM_Dimension * Linear_Units_of_User_Choice(L1)
- Interim Linear Layer Output: seq_len * batch_size * L1
- Attention Weights (Vector): L1

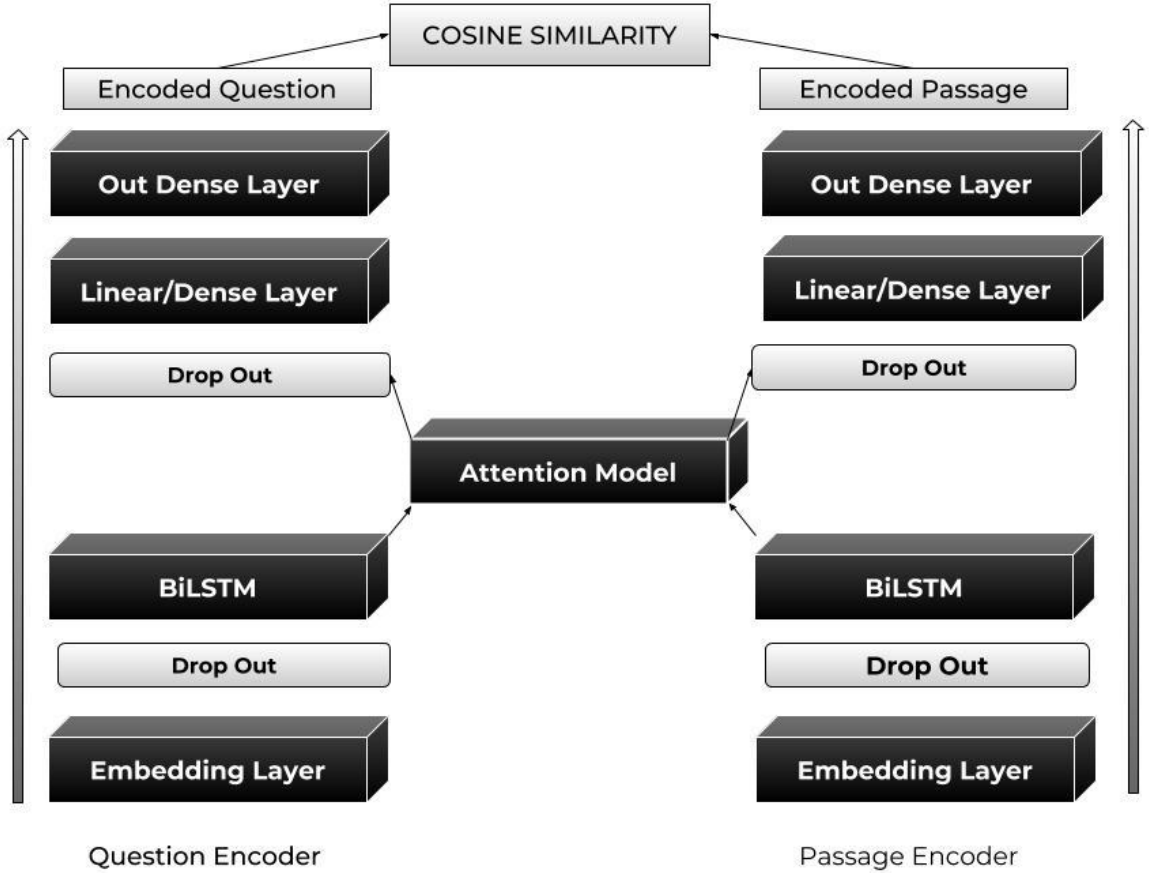
@ operation: Input: seq_len * batch_size * L1, L1 - Output: seq_len * batch_size

Models

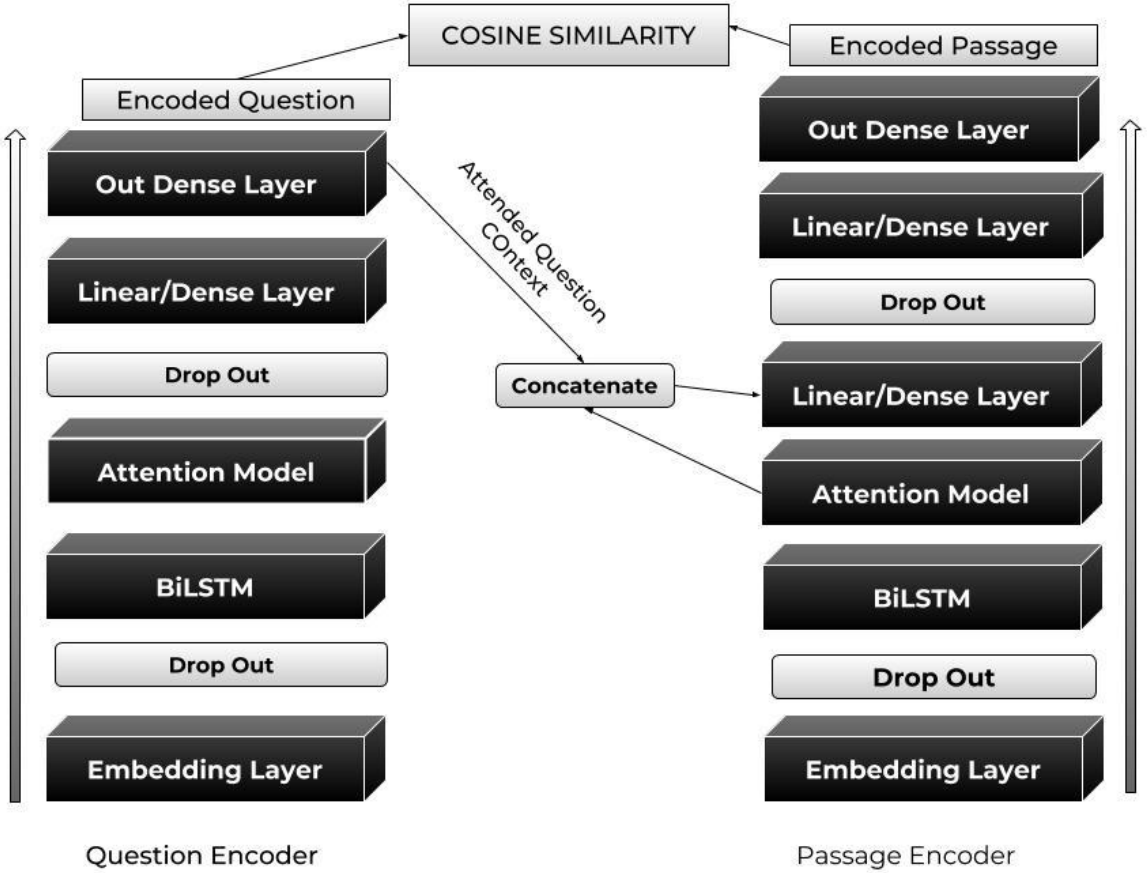
Siamese Network



Common Attention Variant

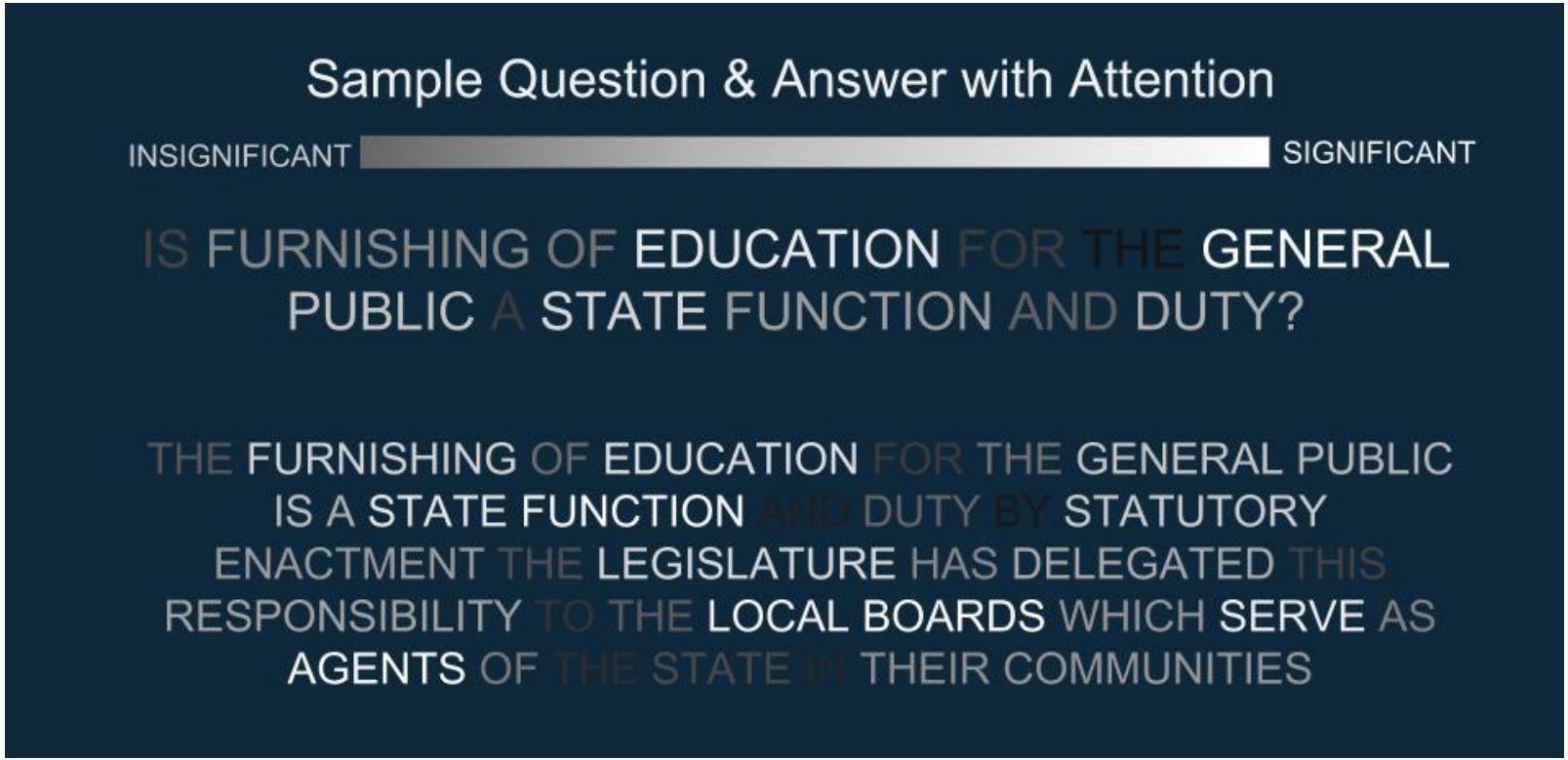


Context Attention Variant



Results

Attention


Sample Question & Answer with Attention
INSIGNIFICANT
SIGNIFICANT
IS FURNISHING OF EDUCATION FOR THE GENERAL PUBLIC A STATE FUNCTION AND DUTY?
THE FURNISHING OF EDUCATION FOR THE GENERAL PUBLIC IS A STATE FUNCTION AND DUTY BY STATUTORY ENACTMENT THE LEGISLATURE HAS DELEGATED THIS RESPONSIBILITY TO THE LOCAL BOARDS WHICH SERVE AS AGENTS OF THE STATE IN THEIR COMMUNITIES

nDCG (No Solr)

@N	nDCG
Individual Attention Model	
2	0.87
3	0.91
Common Attention Model	
2	0.87
3	0.90
Context Attention Model	
2	0.56
3	0.62

Latent Semantic Representation For Ranking

Manasa Bharadwaj

Thanks to Elias, Paul, Phileas, Hao-Wei for technical support. Antonia and Simon for being awesome denmates.