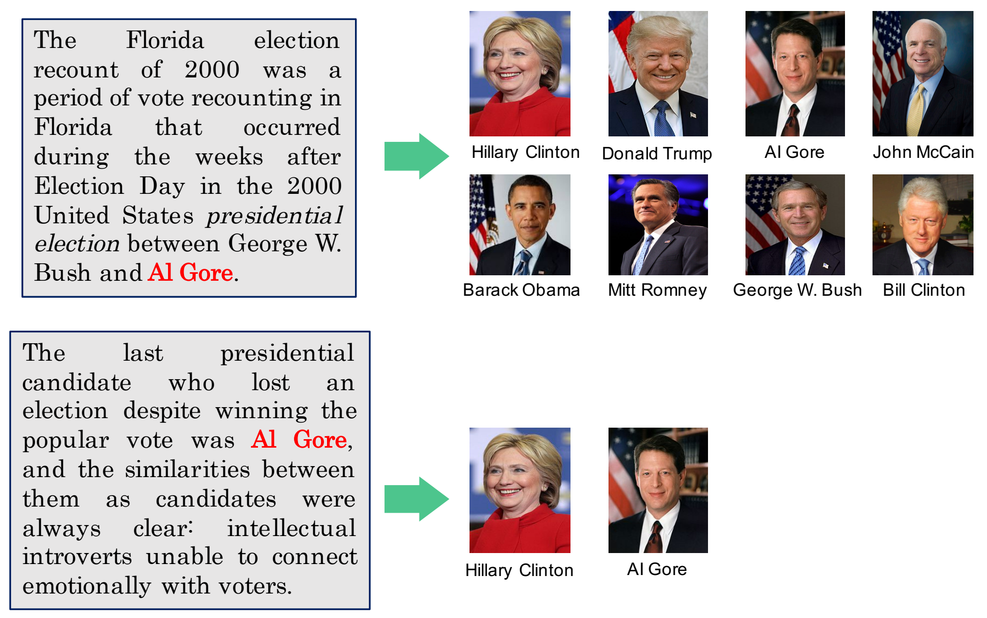
NEURAL MODEL FOR CONTEXTUAL NAMED ENTITY RETRIEVAL

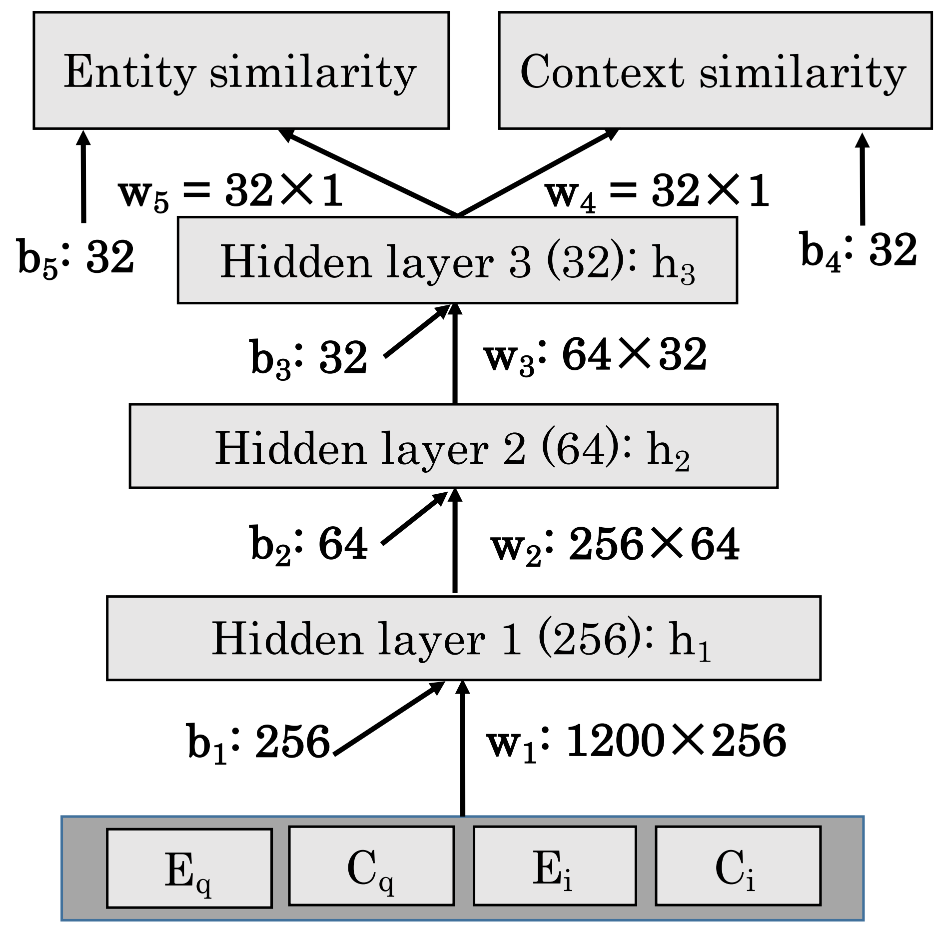
Sheikh Muhammad Sarwar, Anna Fariha

MOTIVATION

- Entity set expansion depends on both **exemplar entity** and **context**.
- **Explanation** of the retrieved entities helps understanding relevance.



ARCHITECTURE



Concatenated embedding vector

METHODOLOGY AND EVALUATION

- The dataset consists of a collection of sentences with topically annotated entities. Entities that are the answers to the same list question are grouped together. We randomly pick one query sentence and retrieve other sentences that contain entities similar to the entities of the query sentence.
- **MTL** takes into account both entity similarity context similarity.
- We applied **weak supervision** for sentence similarity (*Less accuracy may be desired!*). We trained a **Siamese LSTM** on **SNLI** (Stanford) dataset and used that network as a weak supervisor (79% accuracy on validation).
- We have used **Fasttext** Embedding (Facebook Research): Skipgram model, minimum word count = 5, dimension = 300, number of embedded words = 507,865.
- **Evaluation**: A sentence is relevant if it contains a similar or topically bound entity with respect to the query entity. We pick 10 sentences from 2000 candidate sentences and compute **recall@10**. We use **mean average precision** for measuring ranking quality.

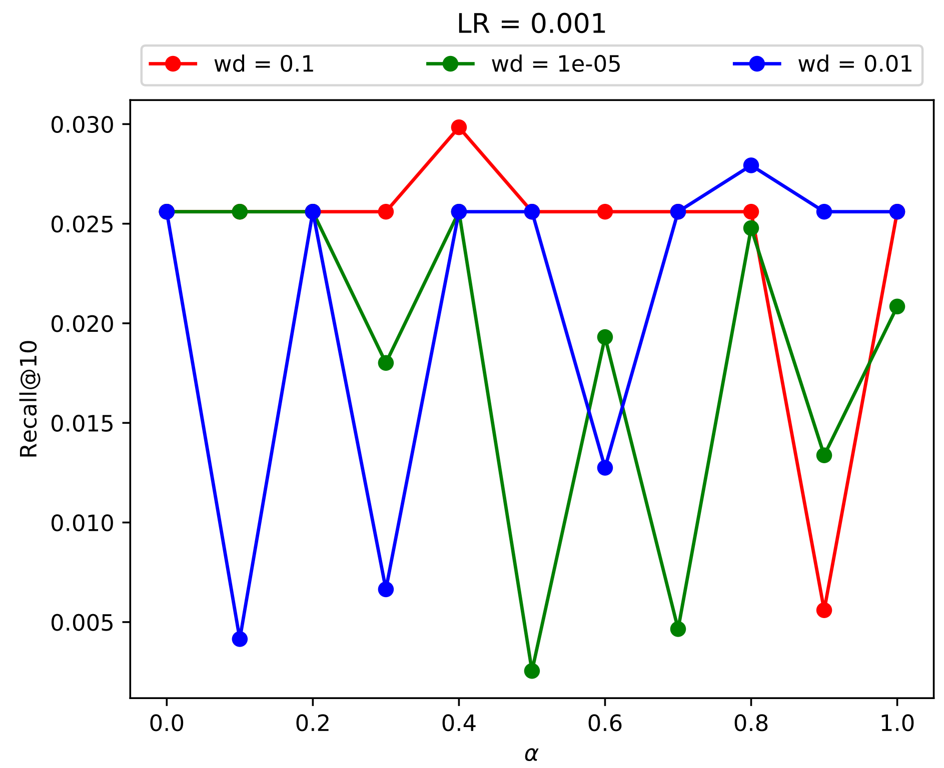
Ranking scheme: $E_i \prec E_j \leftrightarrow SimScore(E_q, E_i) > SimScore(E_q, E_j)$

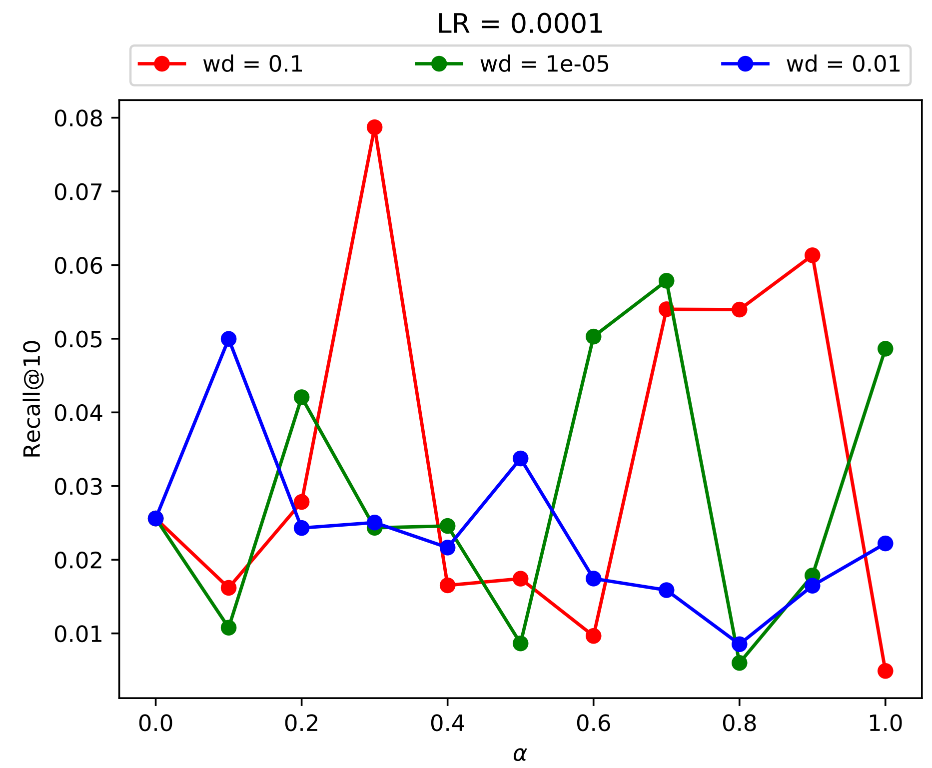
LOSS FUNCTION

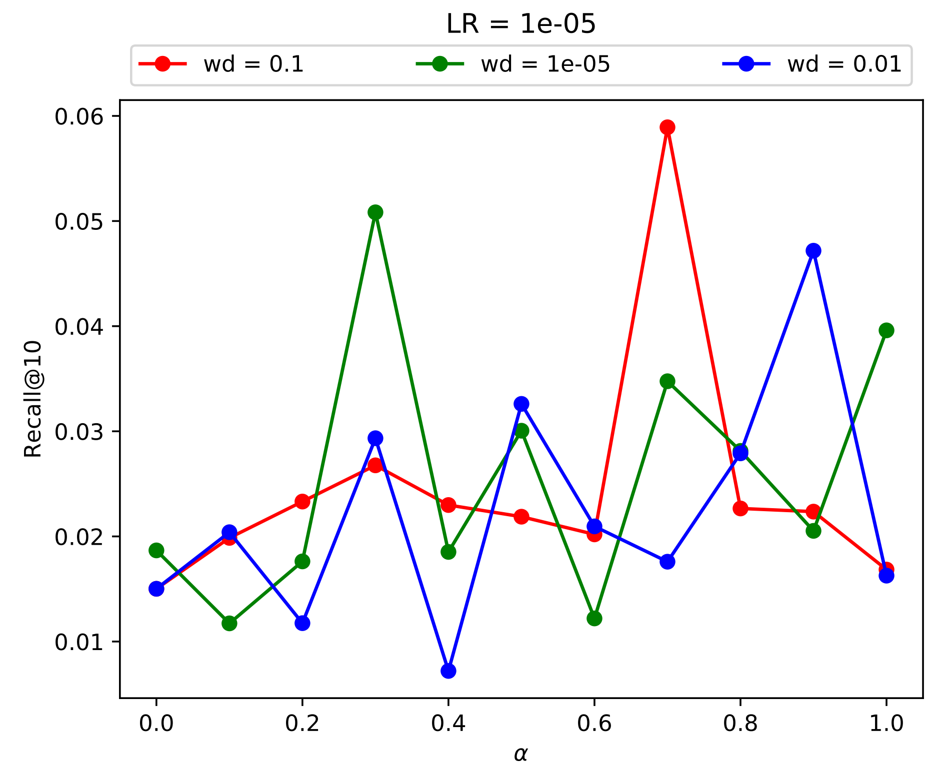
$$\sum_{n=1}^{N} \alpha L_{ce}(y_n^e, \hat{y}_n^e) + (1-\alpha) \mid\mid y_n^c - \hat{y}_n^c \mid\mid_2^2$$

y_n^e: Ground truth entity similarity
y_n^c: Ground truth context similarity
L_{ce}: Cross entropy loss
$\hat{y}_n^e = f^e(x_n, \theta)$: Predicted entity similarity
$\hat{y}_n^c = f^c(x_n, \theta)$: Predicted context similarity
x_n: Data vector
θ: Learnable parameters
α: Trade-off parameter

RESULTS







DATASET & BASELINE

Dataset: TREC [a] 2005 & 2006 List QA (Train: 268200, Val: 16980, Test: 78920)
Baseline: Average Sentence Embedding

[a] trec.nist.gov

TAKEAWAYS & FUTURE WORK

- Auxiliary task improves primary objective for contextual entity retrieval.
- Weak supervision for auxiliary task is effective.
- We plan to incorporate RNN and LSTM for capturing sequence in text.
- We will explore selective parameter sharing instead of full parameter sharing.

REFERENCES

[1] J. Mueller and A. Thyagarajan. Siamese recurrent architectures for learning sentence similarity. In *AAAI*, pages 2786–2792, 2016.

[2] Y. Sun et al. Modeling mention, context and entity with neural networks for entity disambiguation. In *IJCAI*, pages 1333–1339, 2015.