



# MeltingPotQA: Multi-hop question answering

Nathan Dass, Saelig Khattar, Ankit Mathur

## PROBLEM

- Reading comprehension requires model innovations
- Baseline model is designed to work on single context document, with no methodology to focus on more relevant documents
- Hypothesis: understanding which documents are more likely to be relevant to a question will guide the model in the right direction

## HOTPOTQA – MULTI-HOP QUESTION ANSWERING

### Paragraph 1:

[1] *Return to Olympus* is the only album by the alternative rock band Malfunkshun. [2] It was released after the band had broken up and after lead singer Andrew Wood (later of Mother Love Bone) had died of a drug overdose in 1990. [3] Stone Gossard, of Pearl Jam, had compiled the songs and released the album on his label, Loosegroove Records.

### Paragraph 2:

[4] *Mother Love Bone* was an American rock band that formed in Seattle, Washington in 1987. [5] The band was active from 1987 to 1990. [6] Frontman Andrew Wood's personality and compositions helped to catapult the group to the top of the burgeoning late 1980s/early 1990s Seattle music scene. [7] Wood died only days before the scheduled release of the band's debut album, "Apple", thus ending the group's hopes of success. [8] The album was finally released a few months later.

**Question:** What was the former band of the member of Mother Love Bone who died just before the release of "Apple"?

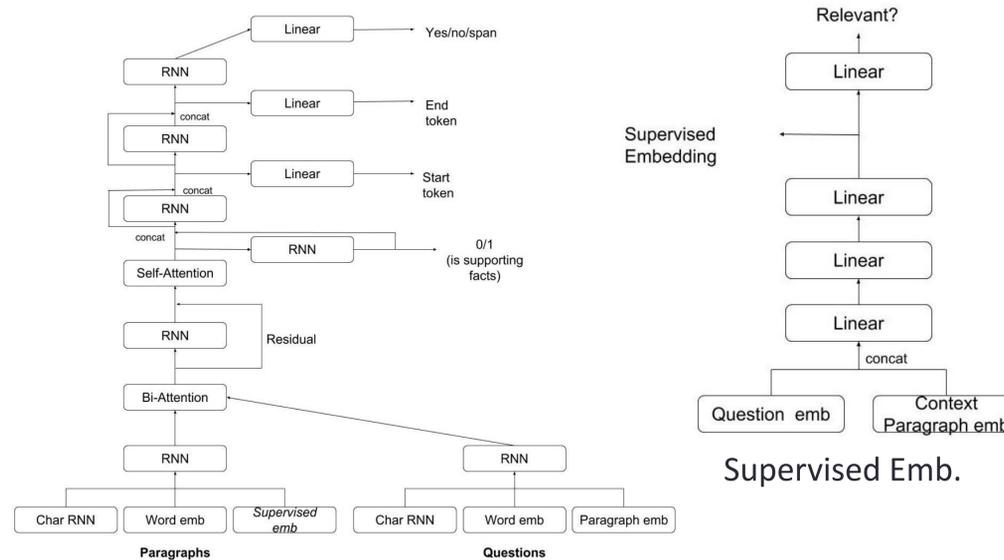
**Answer:** Malfunkshun

**Supporting facts:** 1, 2, 4, 6, 7

- SQuAD contains single paragraph context, with an answer derived from a single span
- HotPotQA focuses on reading comprehension task – synthesizing disparate sources of information

## APPROACH: RELEVANT PARAGRAPHS

- Document Embeddings (doc2vec)
  - Averaging
  - Per-word paragraph embeddings
- Supervised Document Embeddings
- Other improvements to the baseline model



MeltingPotQA Architecture, built on [3]

## RESULTS

	F1	EM
Baseline	54.93	40.90
Baseline, patience 7	56.38	42.26
Baseline (clean data, patience 7)	<b>62.14</b>	<b>47.37</b>
Supervised Emb. (Step 6700*)	51.51	38.31
Doc2Vec50	55.03	41.08
Doc2Vec150	55.44	41.15
Doc2Vec300	55.16	41.36

## ANALYSIS

### Preprocessing

- **Before:** `<t> Greisen - -OOV- - Kuzmin limit - OOV- The Greisen - -OOV- - Kuzmin limit ( GZK limit ) is a theoretical upper limit on the energy of cosmic rays (high energy charged particles from space ) coming from " distant " sources .-OOV- The limit is , or about 8 - OOV- joules .`
- **After:** `<title> Greisen - -OOV- - Kuzmin limit <p> The Greisen - -OOV- - Kuzminlimit ( GZK limit ) is a theoretical upper limit on the energy of cosmicrays ( high energy charged particles from space ) coming from " distant" sources . <p> The limit is , or about 8 -OOV- joules .`

### Supervised document embeddings

- Cross-entropy vs. weighted cross entropy
- Are the embeddings actually useful? Autoencoder/train the learned parameters in the full model
- Margin ranking loss

## CONCLUSION

- Document embeddings are promising additions to the model and can have an even bigger impact in the full-wiki setting
- Hyperparameter tuning makes a big difference
- Properly cleaned data makes an even bigger difference

## REFERENCES

1. Christopher Clark and Matt Gardner. Simple and effective multi-paragraph reading comprehension. CoRR, abs/1710.10723, 2017.
2. Quoc V. Le and Tomas Mikolov. Distributed representations of sentences and documents, 2014.
3. Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W. Cohen, Ruslan Salakhutdinov, and Christopher D. Manning. HotpotQA: A dataset for diverse, explainable multi-hop question answering. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018.