Lehrstuhl für Datenverarbeitung
Fakultät für Elektrotechnik und Informationstechnik
Technische Universität München

CLSM - Convolutional Latent Semantic Model

A Latent Semantic Model with Convolutional-Pooling Structure for Information Retrieval

G10 - Moustafa AboulAtta, Martin Achtner, Ridon Arifi, Daniel Ehrhardt and Lukas Fichtner

Introduction

- Modern search engines rely on semantic models for the retrieval of Web documents with search queries.
- These perform better than simple lexical models by combining words that appear in a similar context into semantic clusters.
- However existing models do mostly not consider the context of the words which can lead to unwanted results. (e.g. microsoft *office* ↔ apartment *office*)
- CLSM is taking the context into account and is thus able to achieve a better performance.

CLSM Architecture

The CLSM consists of five layers (c. Figure 1) in the following order:

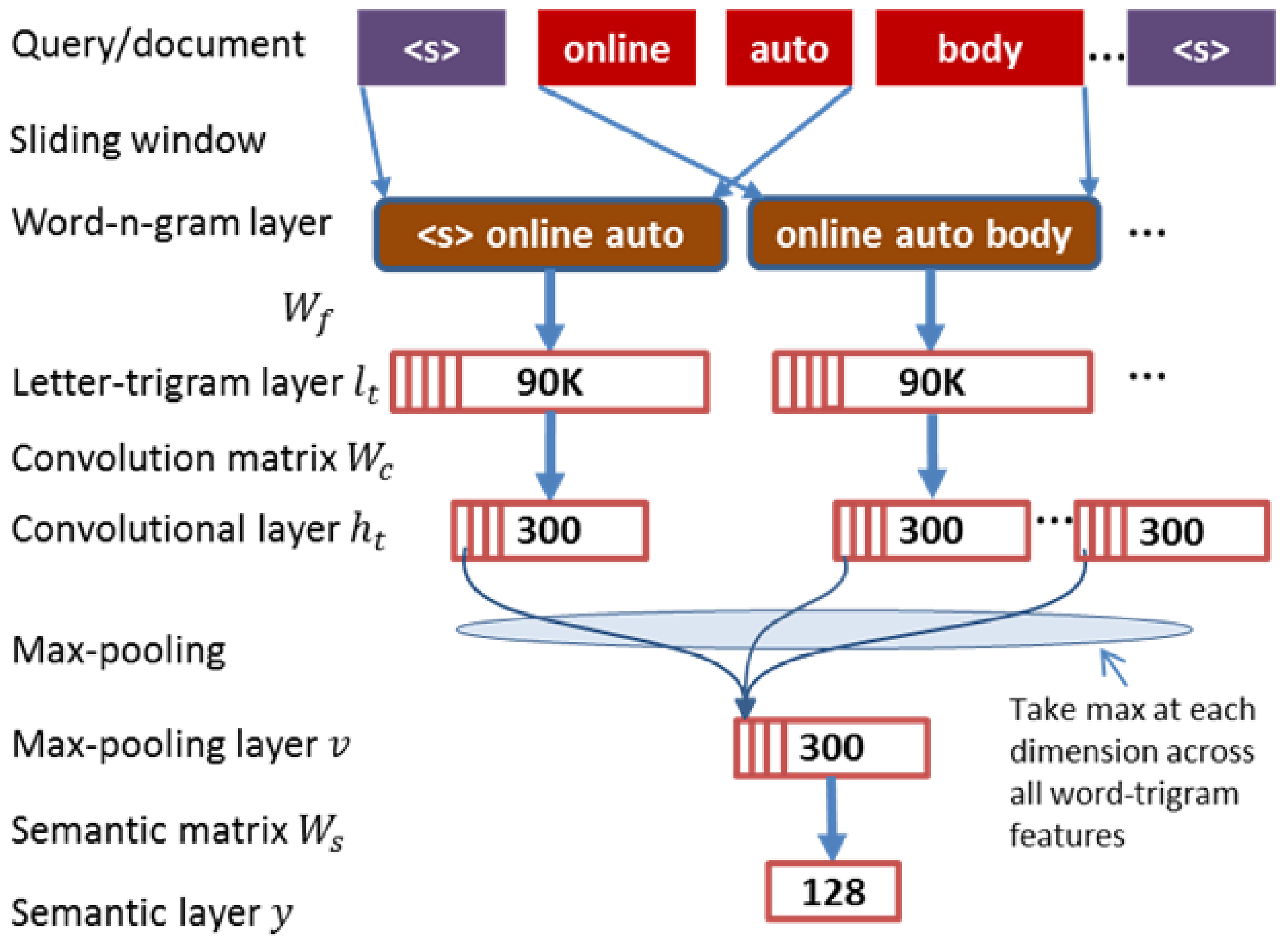

Figure 1: The Architecture of CLSM, Shen et al.

Usage

- Use the CLSM model on the query and the documents
- Determine the matching documents via the cosine distance of the resulting vectors

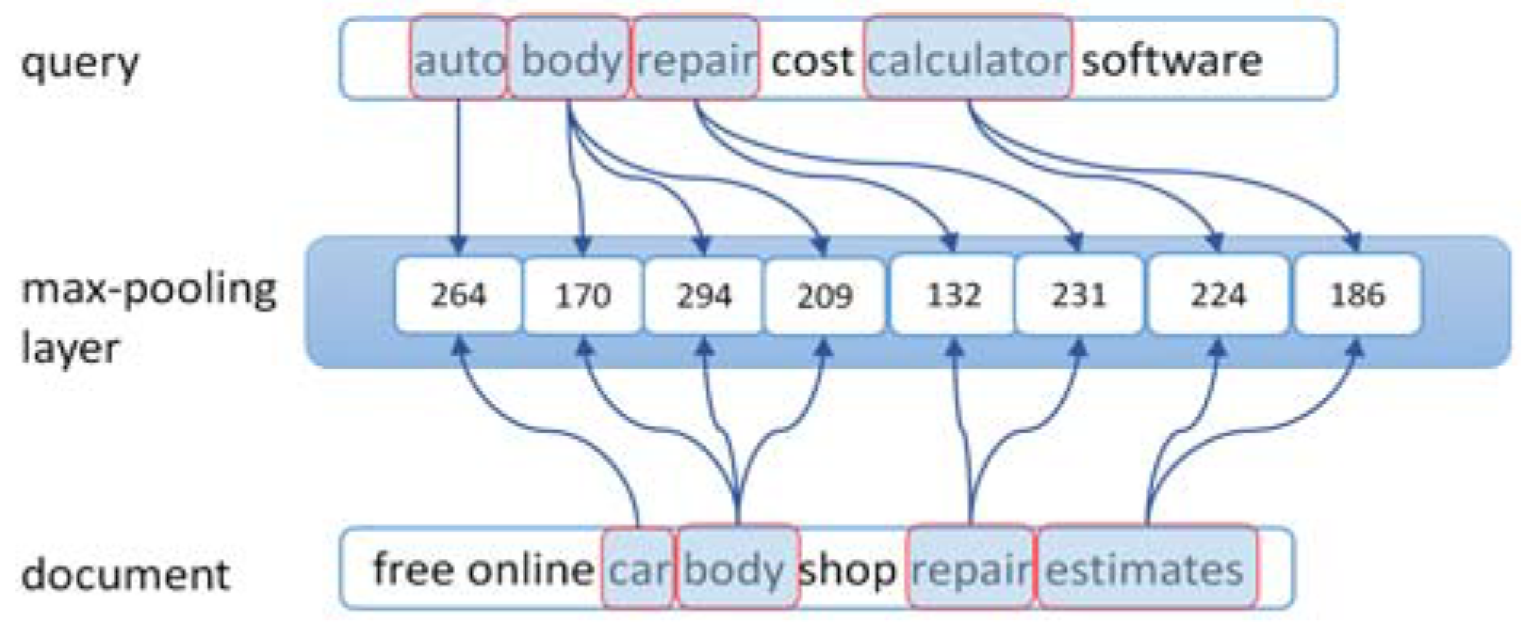

Figure 2: Semantic Matching on Maxpooling Layer, Shen et al.

Learning

- Model paramaters are trained to maximize the likelihood P(D+|Q) through the loss function.

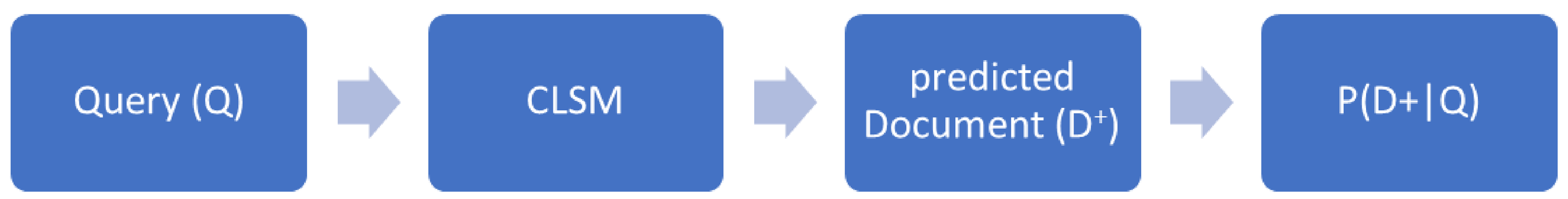

Figure 3: CLSM learning process

Experiments and Results

#	Models	NDCG@1	NDCG@3	NDCG@10
1	BM25	0.305	0.328	0.388
2	ULM	0.304	0.327	0.385
.				
.				
.				
11	PTM (maxlen = 3)	0.319^{α}	0.347^{α}	0.413^{α}
12	DSSM (J = 4)	0.320^{α}	0.355^{α}	0.431^{α}
13	DSSM (J = 50)	$0.327^{\alpha\beta}$	$0.355^{\alpha\beta}$	$0.431^{\alpha\beta}$
14	CLSM (J = 4)	$0.342^{\alpha\beta\gamma}$	$0.374^{\alpha\beta\gamma}$	$0.447^{\alpha\beta\gamma}$
15	**CLSM (J = 50)**	$\mathbf{0.348}^{\alpha\beta\gamma}$	$\mathbf{0.379}^{\alpha\beta\gamma}$	$\mathbf{0.449}^{\alpha\beta\gamma}$

Table 1: Comparison between state-of-the-art approaches. Superscripts α, β, and γ indicate statistically significant improvements over **BM25, PTM,** and **DSSM (J = 50)**, respectively.

- DSSM and CLSM are closely related in terms of the architecture and it is therefore feasible to compare them closer:

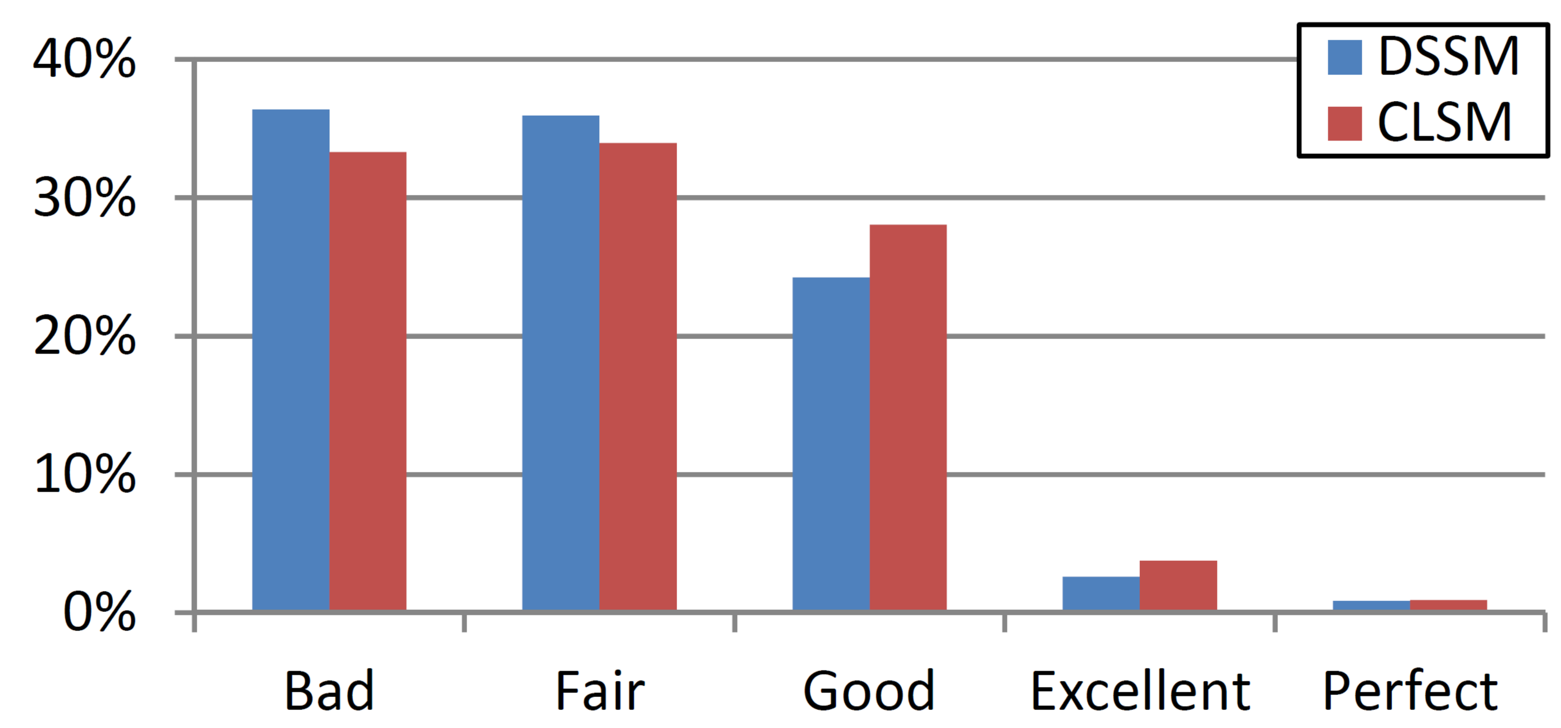

Figure 4: Comparison of DSSM [a] and CLSM regarding qualification quality

[a] Deep Structured Semantic Model