

Kernel-based Learning to Rank with Syntactic and Semantic Structures

Alessandro Moschitti
 Qatar Computing Research Institute
 Qatar Foundation, Doha, Qatar
 DISI, University of Trento, Italy
 amoschitti@qf.org.qa

ABSTRACT

In recent years, machine learning (ML) has been more and more used to solve complex tasks in different disciplines, ranging from Data Mining to Information Retrieval (IR) or Natural Language Processing (NLP). These tasks often require the processing of structured input. For example, NLP applications critically deal with syntactic and semantic structures. Modeling the latter in terms of feature vectors for ML algorithms requires large expertise, intuition and deep knowledge about the target linguistic phenomenon. Kernel Methods (KMs) are powerful ML techniques (see e.g., [5]), which can alleviate the data representation problem as they substitute scalar product between feature vectors with similarity functions (kernels) directly defined between training/test instances, e.g., syntactic trees, (thus features are not needed anymore). Additionally, kernel engineering, i.e., the composition or adaptation of several prototype kernels, facilitates the design of the similarity functions required for new tasks, e.g., [1, 2]. KMs can be very valuable for IR research, e.g., KMs allow us to easily exploit syntactic/semantic structures, e.g., dependency, constituency or shallow semantic structures, in learning to rank algorithms [3, 4]. In general, KMs can make easier the use of NLP techniques in IR tasks.

This tutorial aims at introducing essential and simplified theory of Support Vector Machines (SVMs) and KMs for the design of practical applications. It describes effective kernels for easily engineering automatic classifiers and learning to rank algorithms, also using structured data and semantic processing. Some examples are drawn from well-known tasks, i.e., Question Answering and Passage Re-ranking, Short and Long Text Categorization, Relation Extraction, Named Entity Recognition, Co-Reference Resolution. Moreover, some practical demonstrations are given with SVM-Light-TK (tree kernel) toolkit. More in detail, best practices for successfully using KMs for IR and NLP are presented according to the following outline:

(i) a very brief introduction to SVMs (explained from an

application viewpoint) and KM theory (the essential content for understanding practical procedures).

(ii) Presentation of kernel engineering building blocks, such as linear, polynomial, lexical, sequence and tree kernels, by focusing on their function, accuracy and efficiency rather than their mathematical characterization, so that they can be easily understood.

(iii) Illustration of important applications for which kernels achieve the state of the art, i.e., Question Classification, Question and Answer (passage) Reranking, Relation Extraction, coreference resolution and hierarchical text categorization. In this perspective kernels for reranking will be presented as an efficient and effective approach to learning dependencies between structured input and output.

(iv) Practical exercise on quick design of ML systems using SVM-Light-TK toolkit, which encodes several kernels in SVMs.

(v) Summary of the key points to engineer innovative and effective kernels starting from basic kernels and using systematic data transformations.

(vi) Presentation of the latest KM findings: kernel-based learning on large-scale with fast SVMs, generalized structural and semantic kernels and reverse kernel engineering.

Categories and Subject Descriptors

I.2.7 [Natural Language Processing]: [Language parsing and understanding, Text analysis]

General Terms

Algorithms, Experimentation

Keywords

Question Answering, Kernel Methods, Large-Scale Learning, Support Vector Machines, Structural Kernels

REFERENCES

- [1] A. Moschitti. Efficient convolution kernels for dependency and constituent syntactic trees. In *Proceedings of ECML*, 2006.
- [2] A. Moschitti. Kernel methods, syntax and semantics for relational text categorization. In *Proceeding of CIKM*, 2008.
- [3] A. Moschitti and S. Quarteroni. Linguistic kernels for answer re-ranking in question answering systems. *Information Processing and Management*, 2011.
- [4] A. Severyn and A. Moschitti. Structural relationships for large-scale learning of answer re-ranking. In *Proceedings of SIGIR*, 2012.
- [5] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGIR'13, July 28–August 1, 2013, Dublin, Ireland.
 ACM 978-1-4503-2034-4/13/07.