Abstract:
Certain types of classification problems may be performed at multiple levels of granularity; for example, we might want to know the sentiment polarity of a document or a sentence, or a phrase. Making more localized prediction (e.g., words or phrases), however, is relatively harder because the role of smaller text units depends on the context in which they are used (e.g., sentences or paragraphs), and training a supervised model to take the context into account, requires labeled training corpora, which is not available in many problem domains.

Often, the global prediction at a greater context may be informative for a more localized prediction at a smaller semantic unit. However, directly inferring the most salient local features from the corresponding, easier to learn, global prediction may overlook the semantics of this relationship. This thesis argues that inference along the contraposition relationship of the local prediction and the corresponding global prediction makes a more robust and accurate inference scheme and show how it can be implemented as a transfer function that rewrites a greater context from one class to another.

We study the generalizability of the proposed framework to problem domains with varying data availability profiles and different levels of inference granularity. We demonstrate the robustness of the contrapositive inference to the noisy data and how data augmentation can facilitate the generation of weakly-labeled training data for resource-constrained problem domains. We discuss the transferability and adaptability of the contrapositive relationship to the problem domains with limited amount of training data. In addition, we show the robustness of the contrapositive inference scheme to variability in the size of the local and global contexts: from paragraphs to sentences, and from sentences to words and phrases.