Using Semantic Similarity for Input Topic Identification in Crawling-based Web Application Testing Project Software Contract

1. Training stage: 2. Inference stage:

Output:

Topic

Inference with Corpus, Models and Topics

Vector Transformation

Bag-of-words, TF-IDF, LSI

Feature Extraction

(Proposed Algorithm)

Input:

DOMs

Input:

DOMs

Feature Extraction

(Proposed Algorithm)

Output:

Corpus, Models and Topics

Topic Identification and Labeling

Vector Transformation

Bag-of-words, TF-IDF, LSI