# Modeling Text and Language

## 6.1  Natural Language Processing

Relation extraction means that a system processes a text and can _____ _____.

The ELIZA chat bot works by _____

Preprocessing is crucial in language processing because _____ _____.

The difference between stemming and lemmatization is _____ _____.

## 6.2  Bags of Words

The main aspect of bag-of-words models is that _____ .

The issue with topic modeling and its resulting word lists is that _____ _____.

## 6.3  Language Models

The goal of language models is to _____ ____.

A 3-gram sequence of the sentence "I am watching a movie" would look like this: _____ _____.

## 6.4  Words as Vectors

The closer words are in a semantic space, the _____.

Documents can be represented as word vectors in order to _____
.

Cosine similarity means that _____.

The inverse document frequency factors in _____
_____. This is important because _____
_____

Word2vec makes it possible to do algebraic operations on the vector representations of words to _____
_____.