Natural Language Processing

Motivation

What is Machine Learning?

How many components does Machine Learning consist of?

Motivation

Machine learning addresses the problem of automatically learning computer programs from data. A typical machine learning system consists of three components:

$$\text{Machine Learning} = \text{Representation} + \text{Objective} + \text{Optimization}$$

Motivation

What specifically makes Deep Learning different from Machine Learning?

Motivation

Deep learning is a typical approach for representation learning, which has recently achieved great success in speech recognition, computer vision, and natural language processing. Deep learning has two distinguishing features:

- Distributed Representation
- Deep Architecture

Motivation

Distributed Representation: Deep learning algorithms typically represent each object with a low-dimensional real-valued dense vector, which is named as distributed representation. As compared to one-hot representation in conventional representation schemes (such as bag-of-words models).

Deep Architecture: Deep learning algorithms usually learn a hierarchical deep architecture to represent objects, known as multilayer neural networks. The deep architecture is able to extract abstractive features of objects from raw data, which is regarded as an important reason for the great success of deep learning for speech recognition and computer vision.

The role of Representation Learning in NLP.

NLP aims to build linguistic-specific programs for machines to understand languages. Natural language texts are typical unstructured data, with multiple granularities, multiple tasks, and multiple domains.

The role of Representation Learning in NLP.

Multiple Granularities: NLP concerns about multiple levels of language entries, including but not limited to characters, words, phrases, sentences, paragraphs, and documents. Representation learning can help to represent the semantics of these language entries in a unified semantic space, and build complex semantic relations among these language entries.

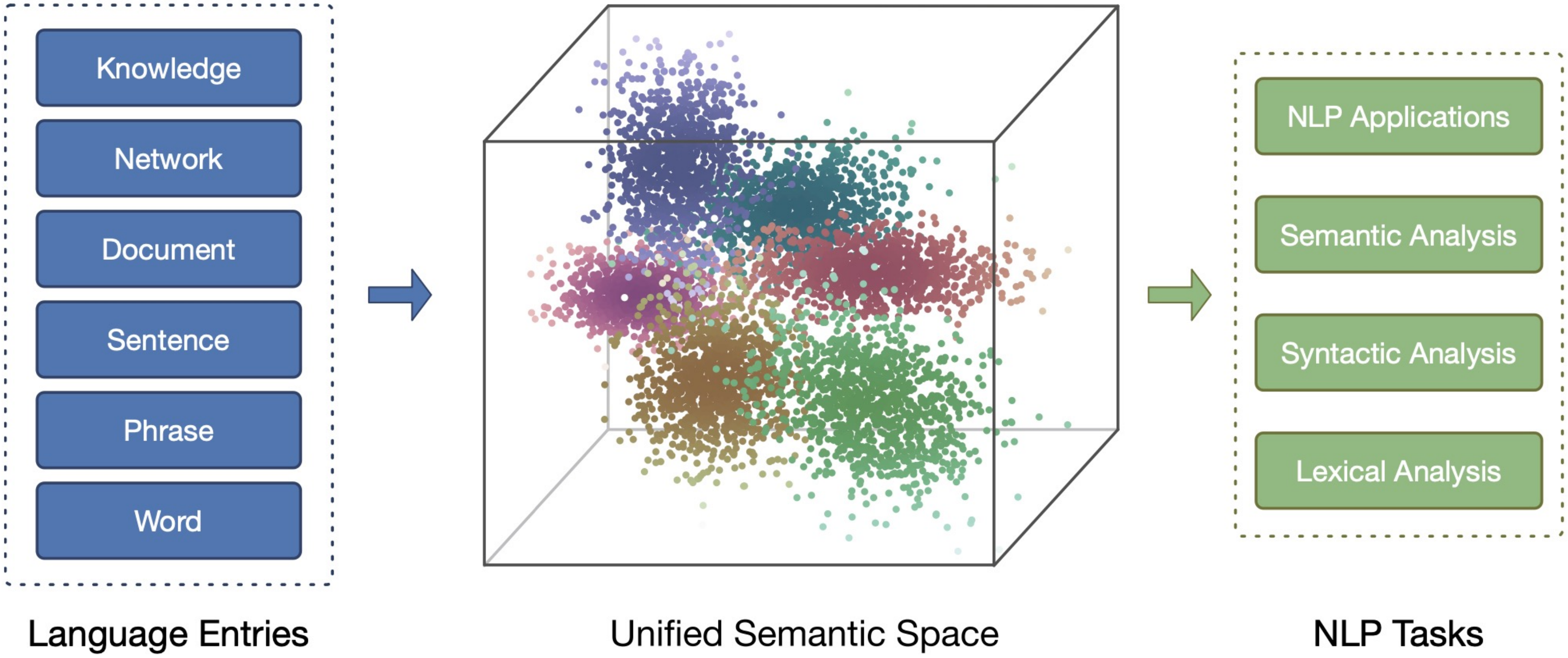
The role of Representation Learning in NLP.

Multiple Tasks: There are various NLP tasks based on the same input. For exam- ple, given a sentence, we can perform multiple tasks such as word segmentation, part-of-speech tagging, named entity recognition, relation extraction, and machine translation

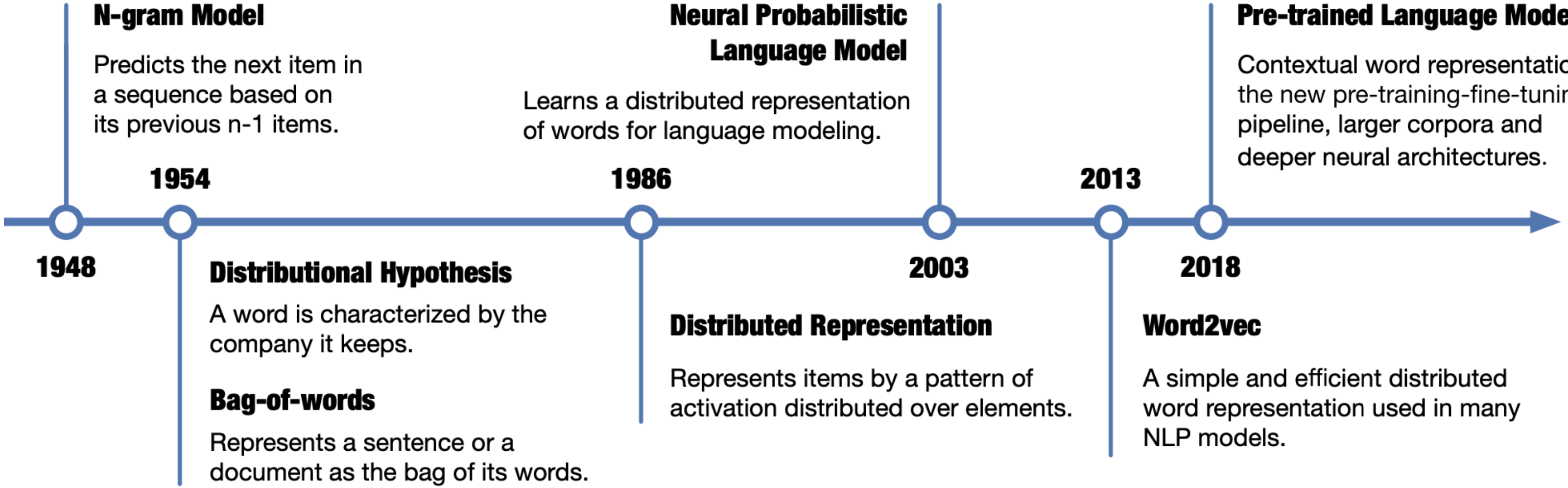
The role of Representation Learning in NLP.

Multiple Domains: Natural language texts may be generated from multiple domains, including but not limited to news articles, scientific articles, literary works, and online user-generated content such as product reviews

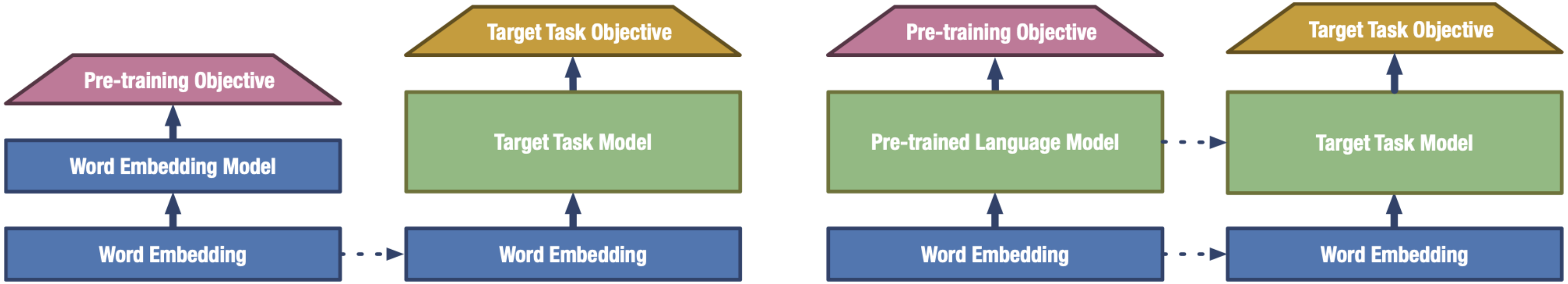
The role of Representation Learning in NLP.



Development of Representation Learning for NLP



Development of Representation Learning for NLP



Word Embedding

Pre-trained Language Model

Learning Approaches to Representation Learning for NLP

Statistical Features: As introduced before, semantic representations for NLP in the early stage often come from statistics, instead of emerging from the optimization process

Hand-craft Features: In certain NLP tasks, syntactic and semantic features are useful for solving the problem.

Supervised Learning: Distributed representations emerge from the optimization process of neural networks under supervised learning.

Self-supervised Learning: In some cases, we simply want to get good represen- tations for certain elements, so that these representations can be transferred to other tasks