Learning Physical Commonsense Knowledge

Carlos E. Jimenez
University of Utah

UUCS-20-005

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

27 April 2020

Abstract

As speakers and writers omit details about common human experiences, a great deal of relevant information eludes concrete attestation in the corpora that many language models and other NLP systems rely on. Physical information about everyday objects is a perfect representative of this type of elusive commonsense knowledge. Through extracting dependency-based contextual representations for training classifiers, we present a means of learning the physical attributes of scores of new words from corpora using only a small number of seed words.