Deep API Learning

ABSTRACT

Developers often wonder how to implement a certain functionality (e.g., how to parse XML \_les) using APIs. Obtaining an API usage sequence based on an API-related natural language query is very helpful in this regard. Given a query, existing approaches utilize information retrieval models to search for matching API sequences. These approaches treat queries and APIs as bags-of-words and lack a deep understanding of the semantics of the query.

We propose DeepAPI, a deep learning based approach to generate API usage sequences for a given natural language query. Instead of a bag-of-words assumption, it learns thesequence of words in a query and the sequence of associated APIs. DeepAPI adapts a neural language model named RNN Encoder-Decoder. It encodes a word sequence (user query) into a \_xed-length context vector, and generates an API sequence based on the context vector. We also augment the RNN Encoder-Decoder by considering the importance of individual APIs. We empirically evaluate our approach with more than 7 million annotated code snippets collected from GitHub. The results show that our approach generates largely accurate API sequences and outperforms the related approaches.

我们提出了一种基于深度学习的方法DeepAPI，它可以为给定的自然语言查询生成API使用序列。它学习查询中单词的均数和相关api的序列，而不是一堆单词的假设。