

Deep Learning for Aspect-Based Sentiment Analysis

A Comparative Review

Úvod

- Aspect-based Sentiment Analysis(ABSA), Hu a kol. (2004) - 3 úrovně
 - dokument
 - věta
 - entita
- Vyšší přesnost pomocí hlubokých konvolučních neuronových sítí
- Rané přístupy k DL zkoumaly lingvistické rysy, gramatické vztahy
- Cíl představit a porovnat novější vývoj v přístupech DL

Aspect-Based Sentiment Analysis

- Extrakce cíle (OTE),
- Detekce kategorie aspektů (ACD)
- Sentiment Polarity (SP)
- ABSA se používá hlavně na zákaznické recenze z webových stránek
- Dřívější přístupy k identifikaci OTE a ACD (např. Hu et al., 2004) byly založeny na četnosti podstatných jmen
- Mezinárodní workshopy o sémantickém hodnocení podporují rozvoj analýzy sentimentu na úrovni aspektů

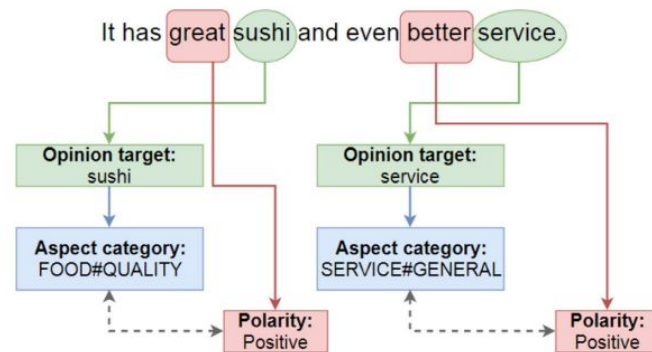


Fig. 1. Three tasks of ABSA in a sample sentence from SemEval ABSA dataset 2016. The sentence has two opinion targets: sushi & service. The category of "sushi" is "Food", with the attribute being "Quality" and polarity "Positive". The category is "Service", with an attribute of "General" and polarity of "Positive".

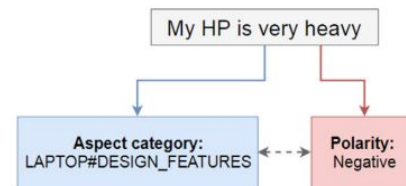


Fig. 2. The opinion target is not explicitly stated, but the category and polarity of the sentence can still be inferred. Sentence from SemEval data 2016.

Hluboké učení

- Metoda strojového učení založená na reprezentaci dat učení pomocí algoritmů – umělých neuronových sítí
- Hluboké neuronové sítě jsou typy umělých neuronových sítí (algoritmů), které zahrnují značný počet vrstev „neuronů“ nebo připojených procesorů, aktivovaných buď senzory z okolí, nebo váženými výpočty
- Přístupy hluboké neuronové sítě (DNN) k NLP se vyznačují
 - hustým vkládáním slov
 - více skrytých vrstev mezi vstupem a výstupem
 - výstupní jednotky

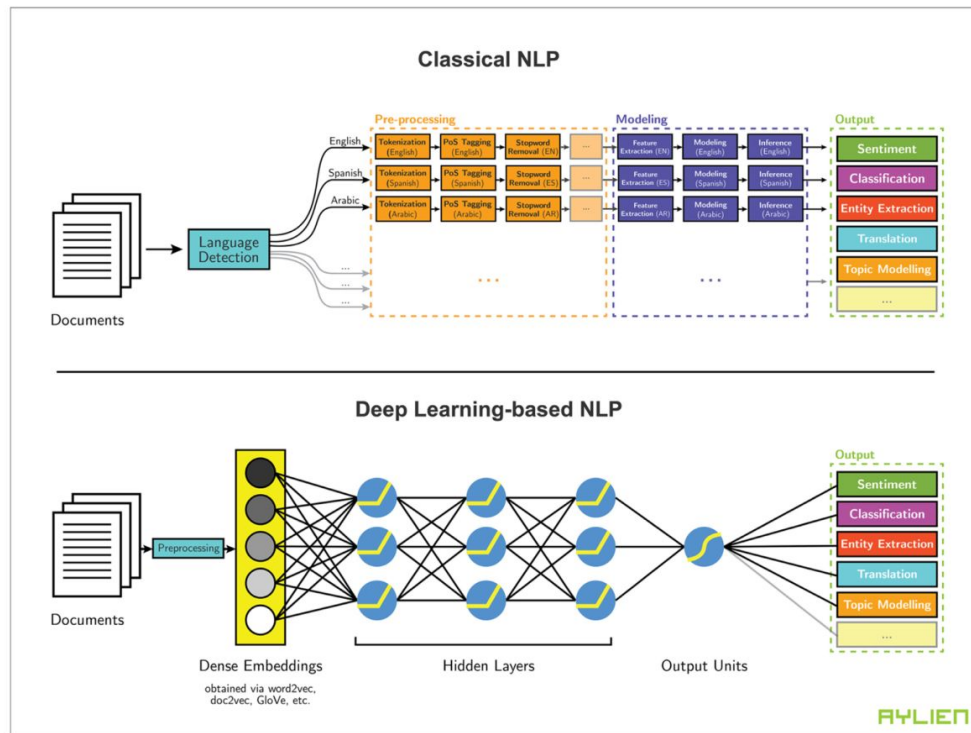


Fig. 3. Comparison of classic machine learning and deep learning processes for NLP. Deep learning architecture is characterized by dense embeddings and hidden layers - adapted from [Thanaki \(2017\)](#).

Convolutional Neural Network Model

- schopnosti extrahovat lokální vzory (tj. nejdůležitější n-gram)
- velké soubory tréninkových dat a vyžadují značné množství přesně vyladěných parametrů
- předpoklad, že klíčová slova mohou obsahují výraz aspektu a označují kategorii nebo určují polaritu bez ohledu na jejich polohu

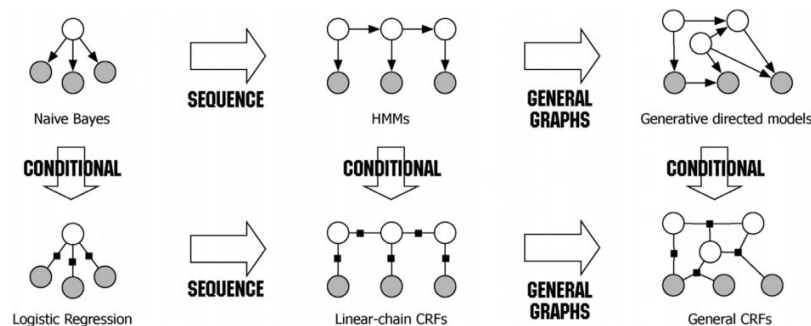


Fig. 5. Conditional random fields with other models. Taken from (Sutton, 2012).

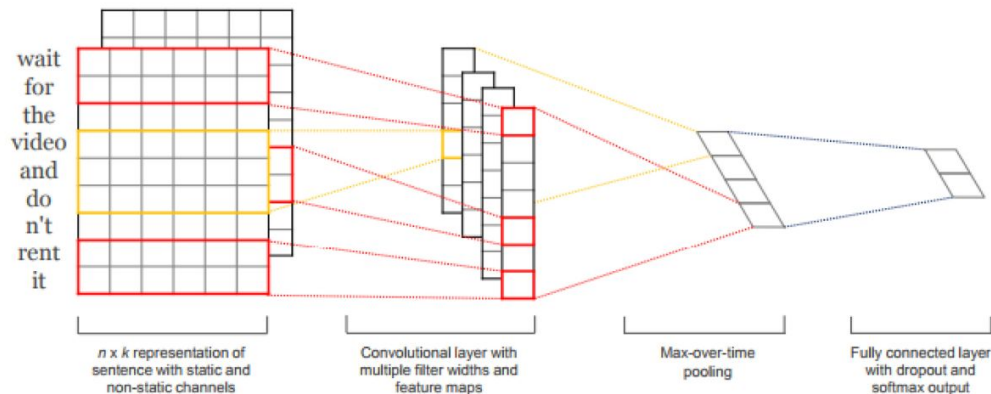


Fig. 6. Basic CNN model with 4 layers. Adapted from Kim (2014).

Recurrent neural network

- Populární v úlohách analýzy sentimentu
- Základem modelů RNN je, že vektor s pevnou velikostí představuje jednu sekvenci
- Závislý na předchozích výpočtech, takže je schopen zachytit kontextové závislosti v jazyce

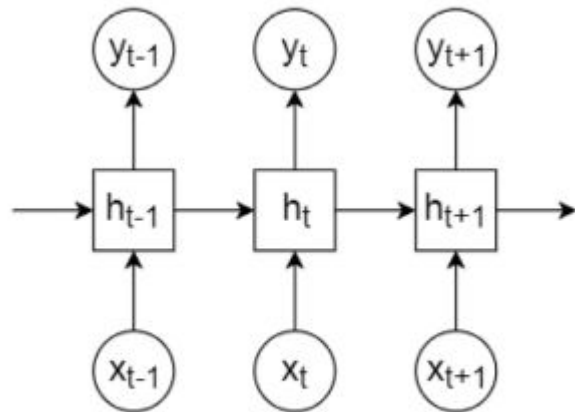


Fig. 8. Basic RNN model.

Modely rekurzivní neuronové sítě (RecNN)

- Lingvisticky motivovány tím, že zkoumají stromové struktury
- Cílem je naučit se elegantně kompoziční sémantiku

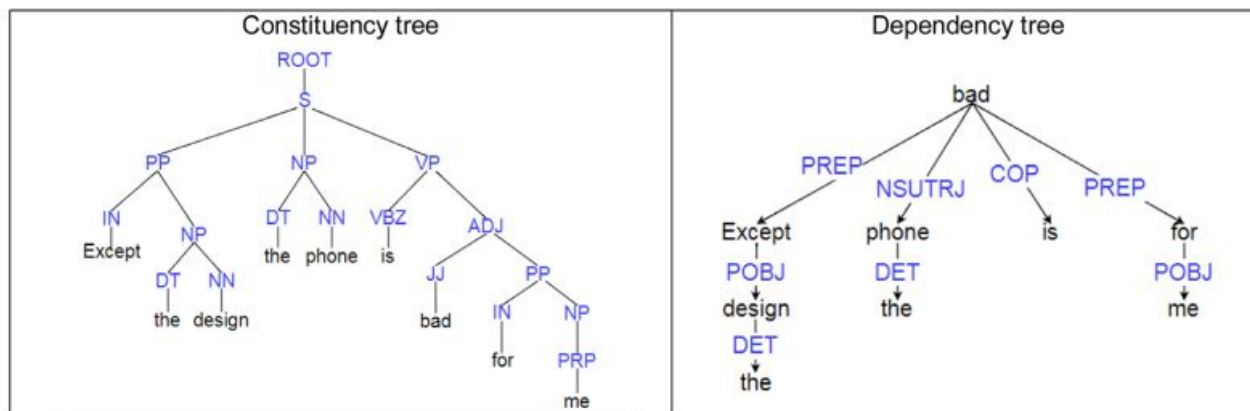


Fig. 16. Constituency tree and dependency tree. Adapted from [Nguyen & Shirai \(2015\)](#).

Srovnání

- CNN mohou být vysoce efektivní díky své schopnosti extrahovat lokální vzory
- RNN jsou výkonné, protože kombinují dvě vlastnosti:
 - Distribuované skryté stavy, které jim umožňují efektivně ukládat informace z minulých výpočtů
 - Nelineární dynamika, která lépe odpovídá nelineární povaze dat
- RecNN jsou jednoznačnými výhodami jednoduchá architektura a schopnost učit se stromové struktury vět a nová slova