Reading Assignment 1: COMP 550, Fall 2024

Hamza Rashid

Introduction

The paper "Neural Architectures for Named Entity Recognition" by Lample et al. (2016)., addressed the reliance of state-of-the-art Named Entity Recognition (NER) systems on hand-crafted features and domain-specific knowledge, a common approach at the time due to the small, supervised training corpora that was available. The authors proposed two LSTM-based neural architectures designed to generalize without relying on external resources. Key components of their methodology included: IOBES tagging scheme (Inside, Outside, Beginning, End, Singleton), and forming character-sensitive word embeddings to capture orthographical and morphological details. While the proposed models acheive state-of-the-art performance, some aspects of the methodology limit the scope for generalization. We delve into further detail on these matters.

Paper Content Overview

The proposed neural architectures are (i) a bidirectional LSTM supplemented with a CRF layer, and (ii) a greedy chunking algorithm utilizing a Stack-LSTM (supports stack operations and embeddings for stack objects). The main components of the bidirectional LSTM-CRF architecture, depicted in Figure 1: word embeddings, Bi-LSTM encoder, and CRF layer. The word embeddings