

An Investigation into Character-Aware Neural Language Models

Anonymous

November 6, 2023

Abstract

This document outlines the conceptual approach and implementation details of a Character-Aware Neural Language Model, an exploration based on the work by Kim et al. (2016) [1]. It provides a discussion on the algorithm, the anticipated challenges in its implementation, particularly the design of the highway model, and the expected methodologies to overcome these obstacles.

1 Introduction

Character-Aware Neural Language Models represent a significant advancement in understanding the morphological subtleties in language processing tasks. By combining character-level information with traditional word embeddings, these models aim to effectively capture the syntactic and semantic nuances of language beyond the capabilities of word-level models alone. This report delves into the algorithmic foundation of such models and discusses the conceptual framework of their implementation.

2 Algorithm Discussion

The Character-Aware Neural Language Model is designed to utilize both character and word-level representations. At the character level, a convolutional neural network (CNN) extracts morphological features which are then processed through a highway network allowing for adaptive information flow. These features are combined with word-level embeddings before being passed to a recurrent neural network (RNN) that models language sequences.

The algorithm's strength lies in its ability to learn rich representations from character-level inputs, addressing out-of-vocabulary and rare words more effectively than word-level models. Additionally, the model can potentially capture language morphology, an asset for languages with rich inflectional systems.

3 Implementation Discussion

The expected implementation will involve the following components: a character embedding layer, a Char-CNN, a highway network, and an LSTM as the RNN. The embedding layer will transform character indices into vectors, using ascii characters as the vector size, which the Char-CNN will process. The most salient features from the CNN will be filtered through a highway network designed to control the flow of information. This combination of character and word embeddings aims to enhance the model’s prediction accuracy for the subsequent word.

4 Challenges and Design of the Highway Model

One of the primary challenges lies in designing the highway network, which must effectively combine character-level features with word embeddings while retaining crucial information. The highway network functions with gating mechanisms similar to those in LSTMs, allowing it to regulate the flow of information and facilitate training deep networks.

To address the complexity of training such a network, the design will incorporate a series of transformation and carry gates. These gates will be fine-tuned to ensure that they can manage the blend of non-linear character features with the linear word embeddings efficiently.

5 Conclusion

The Character-Aware Neural Language Model presents a nuanced approach to language modeling that leverages the granularity of characters and the contextuality of words. The implementation, while challenging, particularly in the aspect of the highway network design, promises a model that can grasp the intricacies of language with greater finesse than traditional models.

References

- [1] Yoon Kim, Yacine Jernite, David Sontag, and Alexander Rush. Character-aware neural language models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1), Mar 2016.