

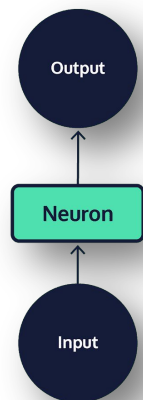
Jasveen S. Undergraduate Researcher,
Anusaaraka lab, LTRC, IIIT-Hyderabad (Prof
Vineet Chaitanya)

NLP(コンピュータソフトウェアとして)は、
近い将来、私たちの生活をどのくらい劇
的に変えますか

**オープンソースの祭典、Open Source
Conference 2020 Online/Hokkaido**

Long Short Term Memory Networks (長期短期記憶ネットワーク)

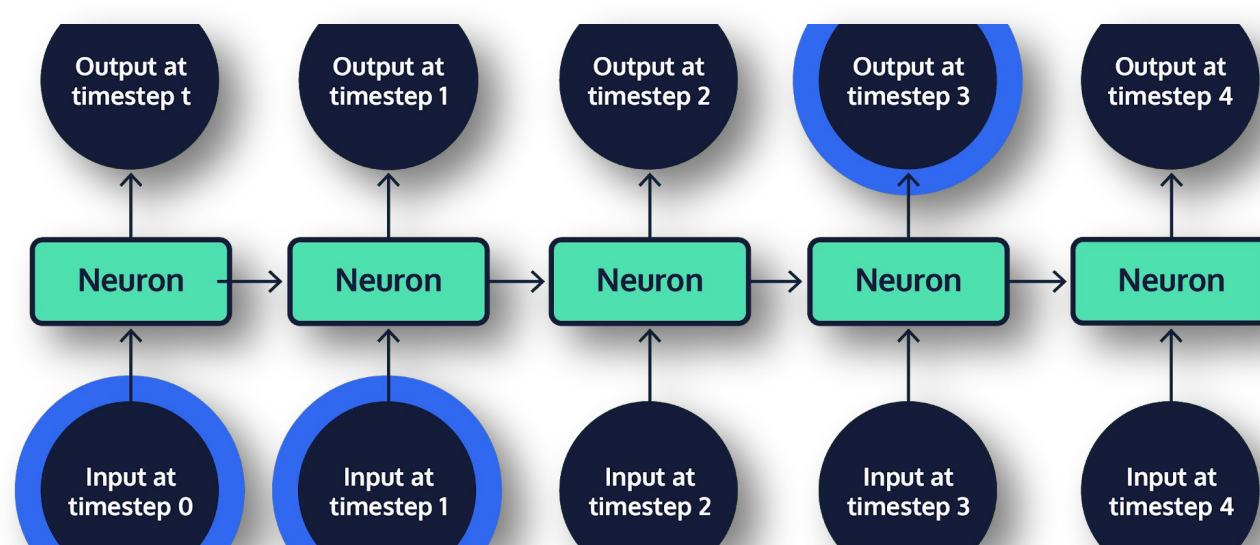
1



Neural Networks (ニューラルネットワーク)

The very first network models, called perceptrons, were relatively simple systems that used many combinations of simple functions, like computing the slope of a line.

2

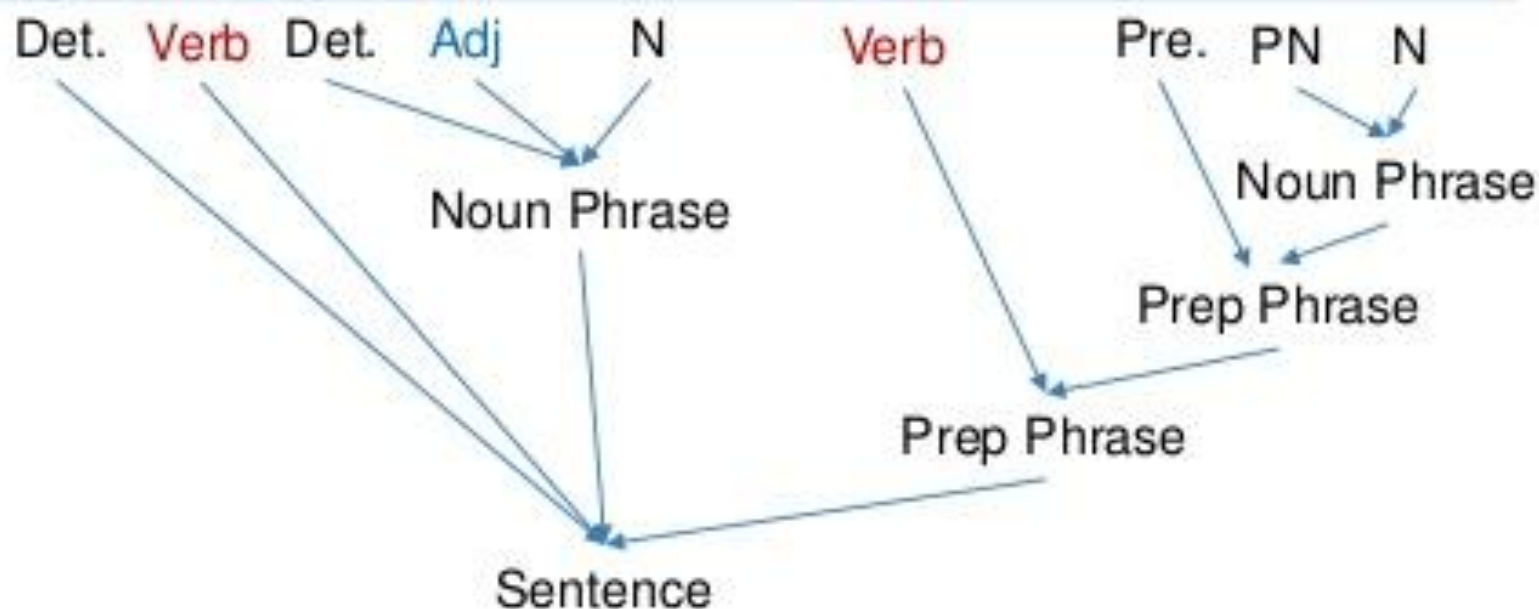


Long-term Dependencies (長期的な依存関係)

While the broad family of RNNs powers many of today's advanced artificial intelligence systems, one particular RNN variant, the long short-term memory network (LSTM), has become popular for building realistic chatbot systems. 会話の次の単語を予測する場合、最近言及された単語は、強い推測をするのに十分な情報を提供することがよくあります。

3

This is the best thing happened in my life.



辞彙分析

Lexical analysis

(Part-of Speech
Tagging 詞性標註)

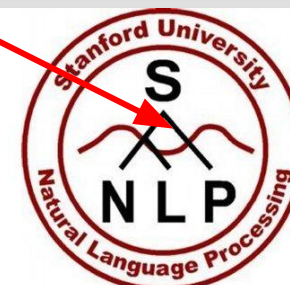
句法分析

Syntactic analysis

(Parsing)

1. The parser provides Universal Dependencies (v1) and Stanford Dependencies output as well as phrase structure trees. 型付きの依存関係は、別の方法では既知の文法関係です。この出力スタイルは、英語と中国語でのみ使用できます。
2. Dependency parsing is the task of extracting a dependency parse of a sentence that represents its grammatical structure and defines the relationships between “head” words and words, which modify those heads.

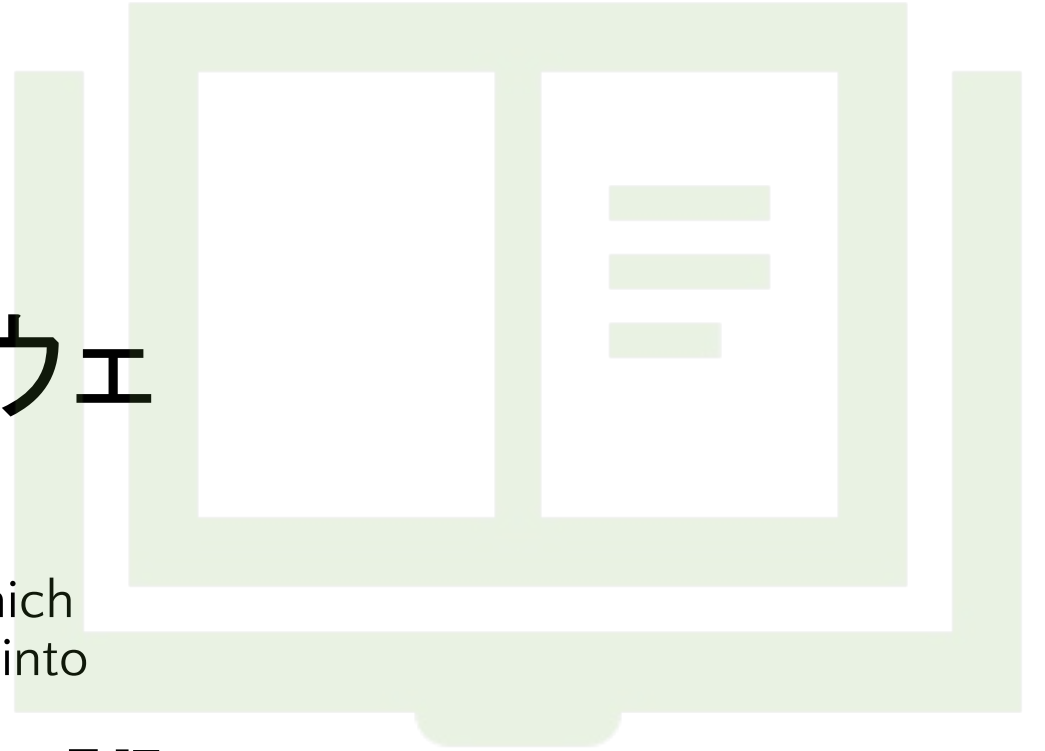
4





Anusaaraka (Software) (ソフトウェア ア)

- Anusaaraka is a computer software which renders text from one Indian language into another.
- Anusaarakaは、英語からヒンディー語への言語アクセスソフトウェアで、PāṇiniのAshtadhyayiから派生したアルゴリズムを採用しています。

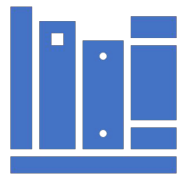




Machine Translation: Use Keras models with seq2seq neural networks to build a better translation tool. (機械翻訳)

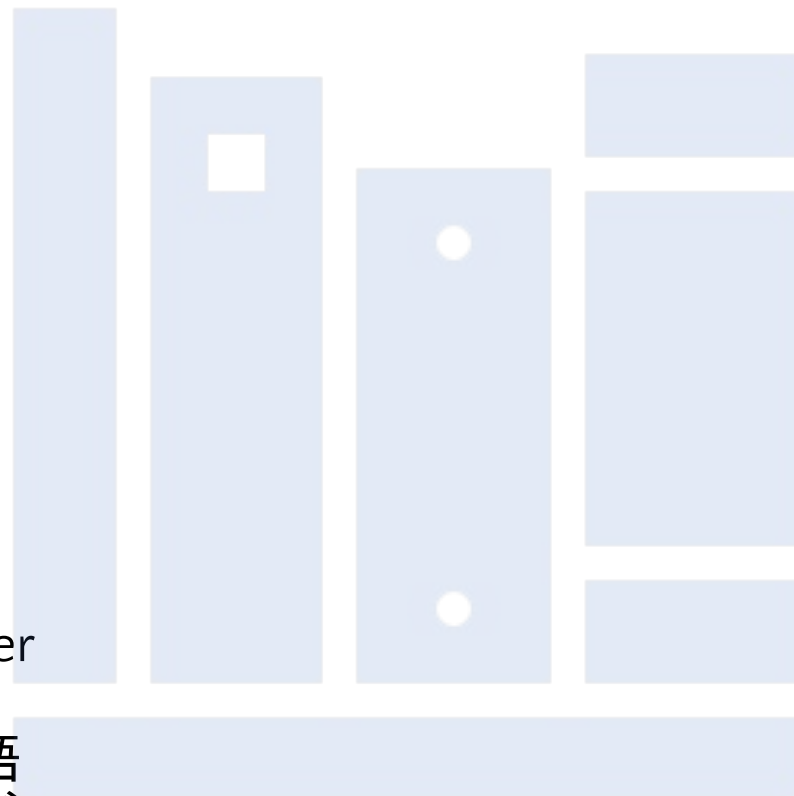
- Install Tensorflow, Keras, and NumPy
- Get yourself a dataset
- Set Up The Code (work on a machine translation code: preprocessing.py, training_model.py, test_function.py)
- Work on preprocessing, the training model, and the test model and function
- 最初のディープラーニングプログラムをローカルに設定できました。





Anusaaraka (Software): Features // 特徴

Throughout the various layers of Anusaaraka output, there is an effort to ensure that the user should be able to understand the information contained in the English sentence. हिन्दी語で完璧な文章を書くよりもこの取り組みのほうが重要です。読みやすい翻訳を用意しても意味がなく、原文の情報を正確に把握できないからです。



Some Exciting NLP Challenges // エキサイティングなNLPの課題

- **Human-like Language Understanding:** 比喩/イディオム、ユーモア、皮肉、礼儀正しさ/無礼
- **言語の生成と対話:** document summarization, database to language summary, coherent and intelligent conversation models
- **Grounded Language with Vision and Speech:** 画像とテキストの配置、画像による言語の明確化、画像/ビデオのキャプション、画像/ビデオの質問応答、テキストからイメージへの生成、ビジュアルストーリーの含意
- **ロボットタスクの具体化された言語:** instructions for navigation, articulation, manipulation, skill learning
- **Machine Learning Models:** 深層構造化モデル、解釈可能なモデル、敵対的モデル、報酬ベースのモデル (reinforcement learning)

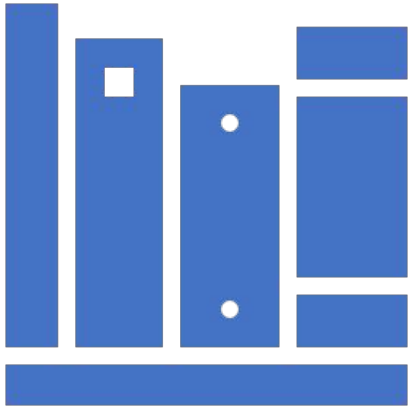




Anusaaraka (Software): Notes //

(ソフトウェア): メモ

- Rule-based, deep parser based, paninian framework based; all programs and language data are free and open-source
- ステップバイステップの翻訳レイヤーを表示すると、エンドユーザーがソースにさかのぼることができ、出力レイヤーの分析とコンテキストへの参照によって翻訳されたテキストを明確にすることができるため、エンドユーザーの信頼性が高まります。



Research Presentation End
私に連絡して: userjs76@gmail.com

ソフトウェア開発またはデータサイエンス(NLP)にエキサイティングな機会がある場合は、一言お願いします。