

# Dependency Parsing

Dataset: UD\_ENGLISH\_GUM

Universal Dependency dataset used UD\_ENGLISH\_GUM, this corpus contains 9124 sentences, 162227 tokens and 164396 syntactic words [1]. A good Recall & F1 score of 87 for UAS and 85 for LAS is provided by the parser.

Metrics	Precision	Recall	F1 Score	AligndAcc
Tokens	100.00	100.00	100.00	
Sentences	100.00	100.00	100.00	
Words	100.00	100.00	100.00	
UPOS	100.00	100.00	100.00	100.00
XPOS	100.00	100.00	100.00	100.00
Feats	100.00	100.00	100.00	100.00
AllTags	100.00	100.00	100.00	100.00
Lemmas	100.00	100.00	100.00	100.00
UAS	87.00	87.00	87.00	87.00
LAS	84.91	84.91	84.91	84.91

I manually cross verified on last 15 sentences. Most of them were correctly parsed except the following three sentences:

Sentence 1: *However, you might attract other types of owls or birds of prey with your nesting box.*

Review: in this sentence word '*nesting*' is VBG but parser recognizes as NN

Sentence 2: *Promote the nesting of birds of prey in your yard.*

Review: in this also '*nesting*' is VBG but parser recognizes as NN

Sentence 3: *If you have a large infestation, a cat may not be enough to totally get rid of it.*

Review: in this sentence word '*rid*' is VBN but parser recognizes it as JJ

## REFERENCES:

1. <https://universaldependencies.org/treebanks/en-comparison.html>