

But what do we even
mean by “Tree”?

...

Gray Davidson | May 30th 2018

The Young Men Walked through the Trees.

| | |
|-----------------|--|
| Me: | The Young Men Walked through the Trees . |
| Austen/Dickens: | The youthful sailors rode through the hedges |
| Rand: | The brilliant lives started through the granite . |
| Thoreau: | The recent minds steered through the pin es. |
| Tolkien/Lewis: | The unhappy dwarves ran through the roots . |

Data

Ayn Rand
Thoreau & Emerson
Austen & Dickens
Tolkien & Lewis

“At length I entered within
the skirts of the cloud which
seemed forever drifting over
the summit....”

[length, entered, skirts,
cloud, forever, drifting,
summit, ..., ..., ...],

LDA

| Topic for Project Gutenberg and the Internet Archive: | Topic associated with Ayn Rand: | Topic associated with Lord of the Rings: |
|---|--|---|
| class, div, archive, org, archive org, https, https archive, href, href https, span | men, roark, like, know, think, man, keating, wynand, thought | frodo, bilbo, gandalf, long, ring, know, come, sam, old |

Gensim Word2Vec()

```
model = Word2Vec(bigrams[documents], sg=1, size=200, window=5, min_count=2, workers=4)
```

Bigrams are 2 word long phrases (frequently occurring terms) with both (containing) in the ('mr, darcy') to (mr_darcy).

Skip-gram attempts to predict the context around a given word as opposed to CBOW which predicts the word from context.

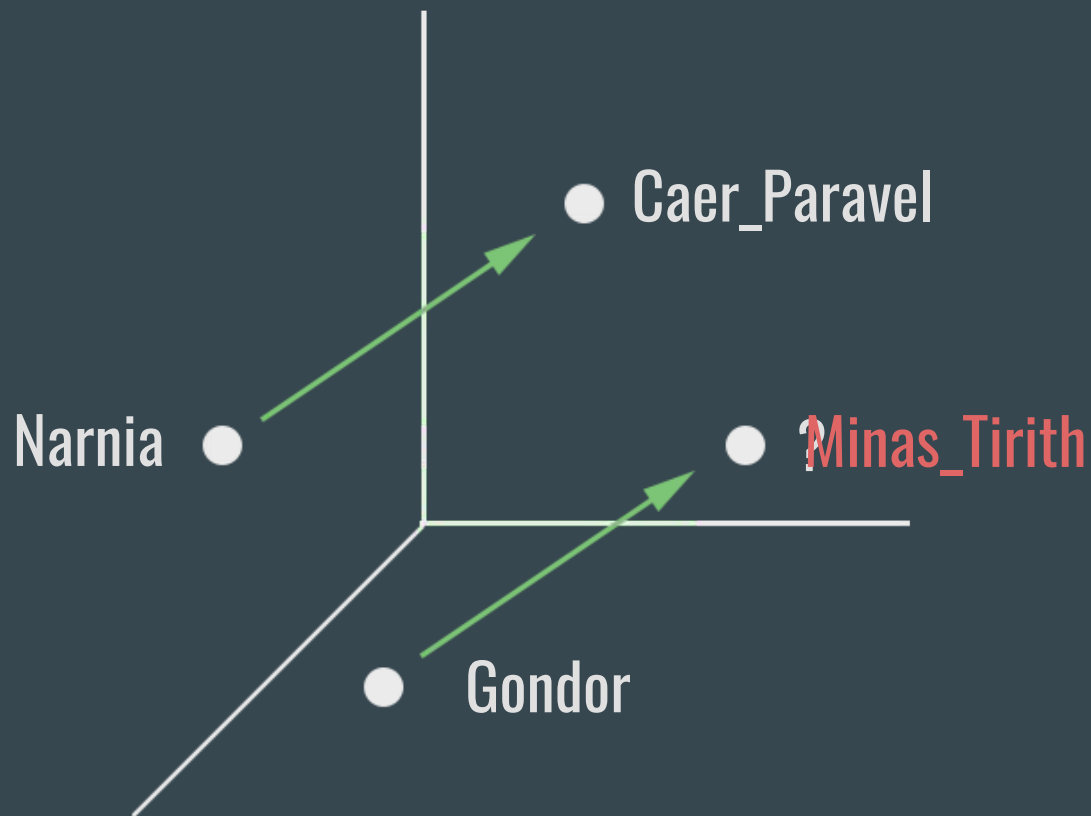
Skip-gram is better for smaller training data sets.

Embeddings

“Tree”

| Austen & Dickens | Ayn Rand | Thoreau & Emerson | Tolkien & Lewis |
|---|--|--|---|
| Hill Hedge Mountain Arch Church Wilderness Field Rook Ivy Ruin | Rock Granite Branch Streak Scattered Shape Narrow Haze Weed Spark | Pine Trees Pines Oak Forest White Roots Branches Grass Plants | Trees Roots Wood Higher Forest Wave Glade Larger Shelter Holes |

Word Math



Future Directions

1. Map embeddings to each other >> style transfer.
2. Compare the 4 types of embeddings (see appendix).

Appendix 1: FastText

“Tree”

| Austen & Dickens | Ayn Rand | Thoreau & Emerson | Tolkien & Lewis |
|--|--|--|--|
| Hedge Tree Lane Palace Wild_Flower Pane Trench Torch Wall Frost | Ranch Strip Pile Crane Streak Trees Piles Streams Stretch Curve | Tree Pine Maple Birch Bunch Pinch Pineapple Pines Spruce Bush | Trees Lane Fir_Tree Pine_Tree Beech_Tree Bee Lake Treetop Grasse Wave |

Topic for Project
Gutenberg and
the Internet Archive:

LDA Slide, Alternate Layout (Appendix)

[class, div, archive, org, archive org, https, https archive, href, href https, span]

Topics associated with
specific books:

[men, roark, like, know, think, man, keating, wynand, thought]

[frodo, bilbo, gandalf, long, ring, know, come, sam, old]