How Grammarly Teaches Machines

by Oleksii Sliusarenko

Problem statement

It is og course a good idea to og for a walk.

~~og~~ → of

~~og~~ → go

If it will rain tomorrow, I won't go for a run.

~~will rain~~ → rains

Problem statement

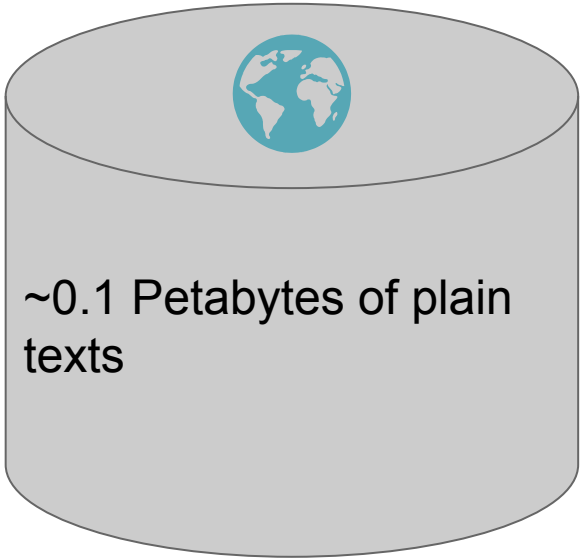
Tom is resembled by Andy.

Intransitive verb in passive voice
Andy resembles Tom

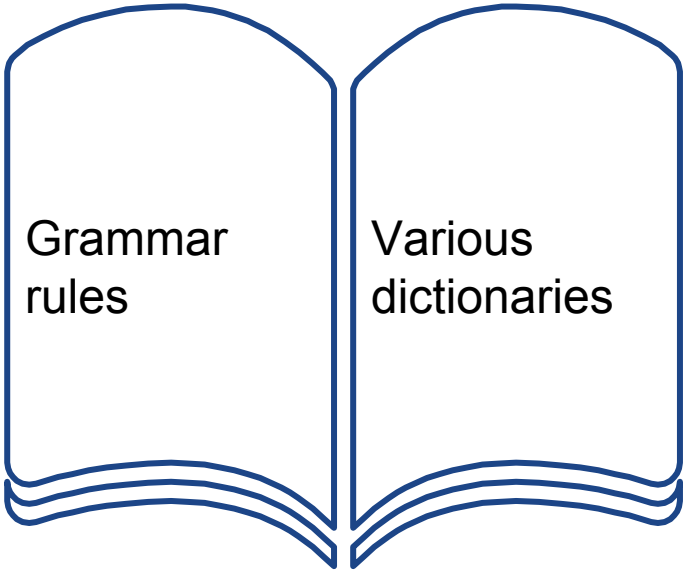
If I cannot understand the material, the teacher will me.

will help

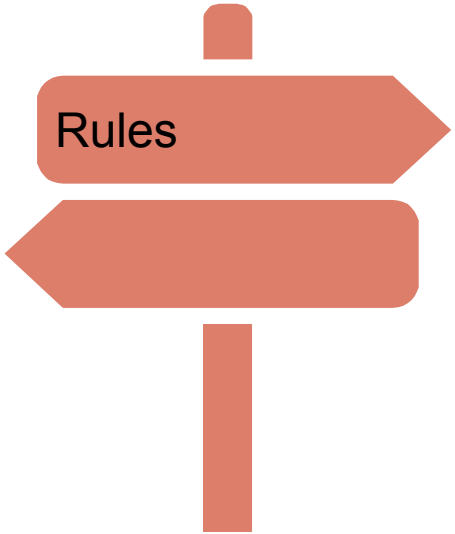
Available data

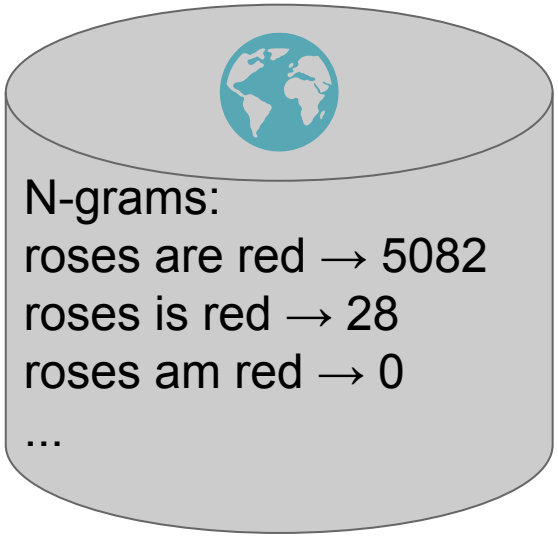


Available data

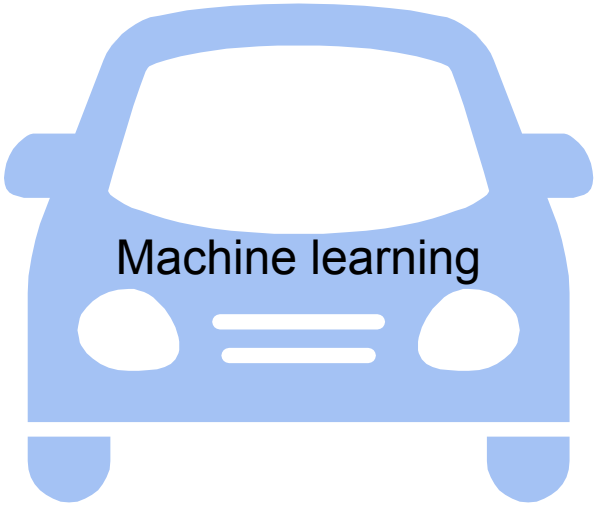


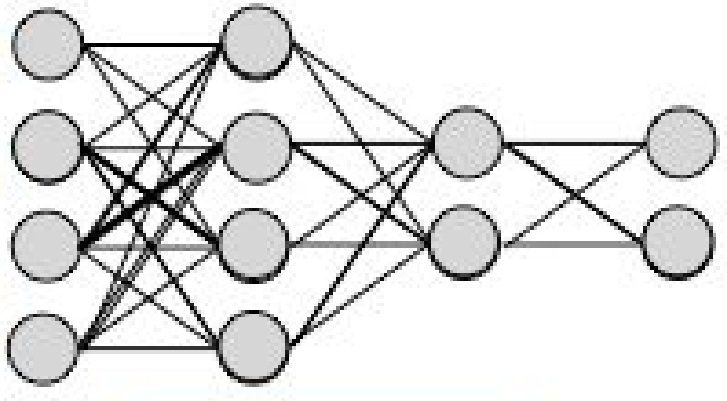
Grammarly uses all possible approaches





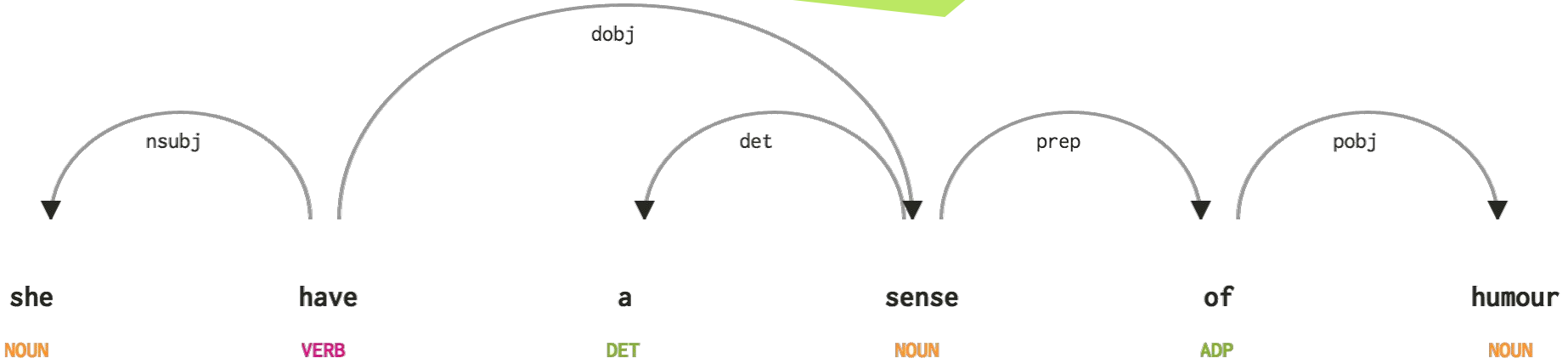
Grammarly uses all possible approaches

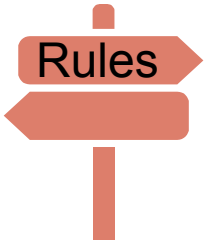




Deep learning

Rule based approach



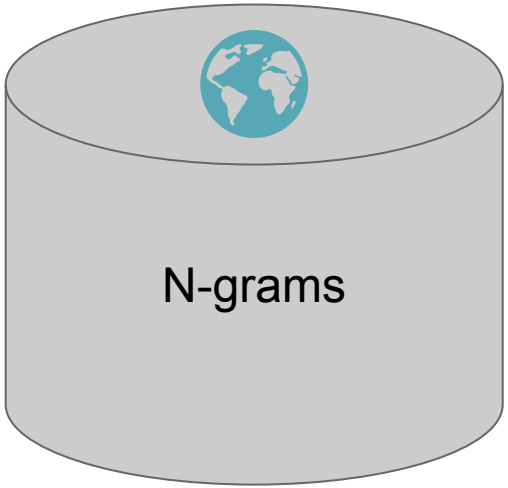


she → third-person pronoun

have → first-person verb

N-gram approach

Shane Bergsma, Dekang Lin, and Randy Goebel. 2009. Web-scale N-gram models for lexical disambiguation.



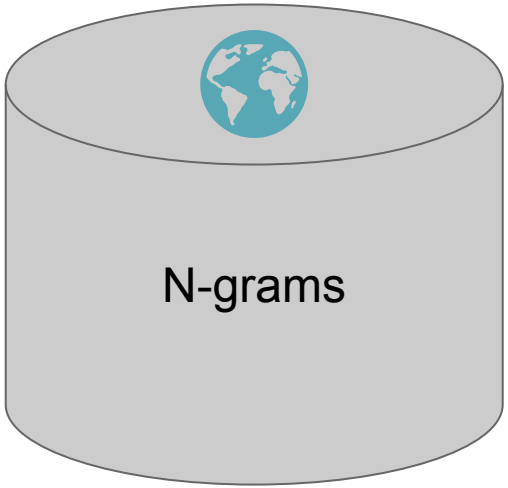
Of course it helps that the **whether** is almost always sunny and dry.

Of course it helps that the **weather** is almost always sunny and dry.

N-gram approach

a(5,0) = freq(*it helps that the* ***whether***)
a(5,1) = freq(*helps that the* ***whether*** *is*)
a(5,2) = freq(*that the* ***whether*** *is almost*)
a(5,3) = freq(*the* ***whether*** *is almost always*)
a(5,4) = freq(***whether*** *is almost always sunny*)

score = sum log(a(i, j))

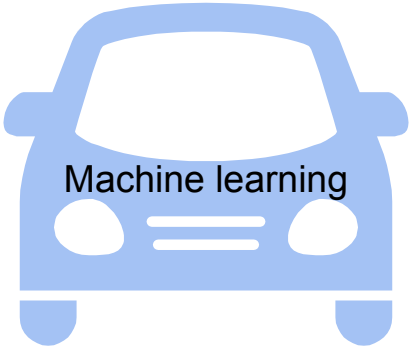


ML approach

Matthew Shardlow. 2013. A Comparison of Techniques to Automatically Identify Complex Words

Complex:
Mike got spam today that contained **gibberish**.

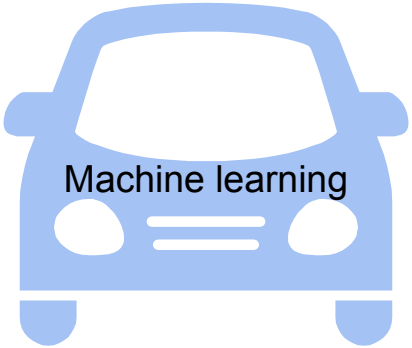
Non-complex:
The book tells about people's **irrationality**.



ML approach

Features:

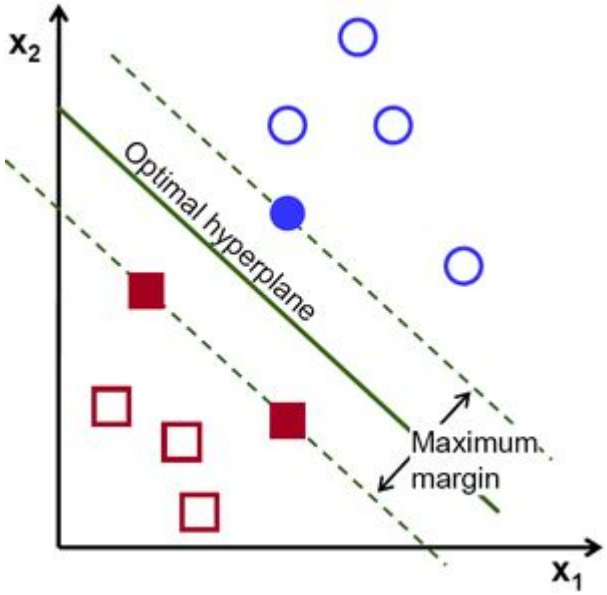
- x_1 = word frequency
- x_2 = word length
- x_3 = syllable count
- x_4 = sense count
- x_5 = synonym count

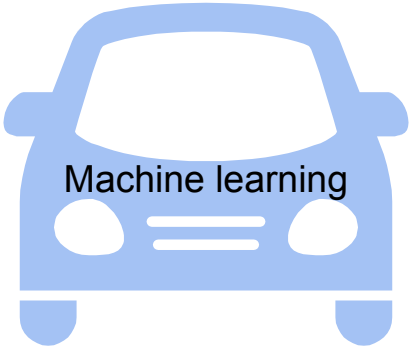


ML approach

Support vector machine:

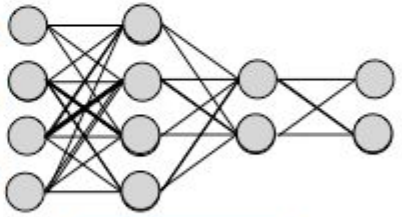
$w_1 * x_1 + w_2 * x_2 + \ldots + w_5 * x_5 > 0$





Deep learning approach

Chengjie Sun, Xiaoqiang Jin, Lei Lin, Yuming Zhao, and Xiaolong Wang. 2015. Convolutional Neural Networks for Correcting English Article Errors



Deep learning

wrong:

I watched **a** film you were talking about yesterday.

correct:

I watched **the** film you were talking about yesterday.

Deep learning approach

- No features! One-hot encoding:

rose = (0,0,0,0,1,0,0,0)

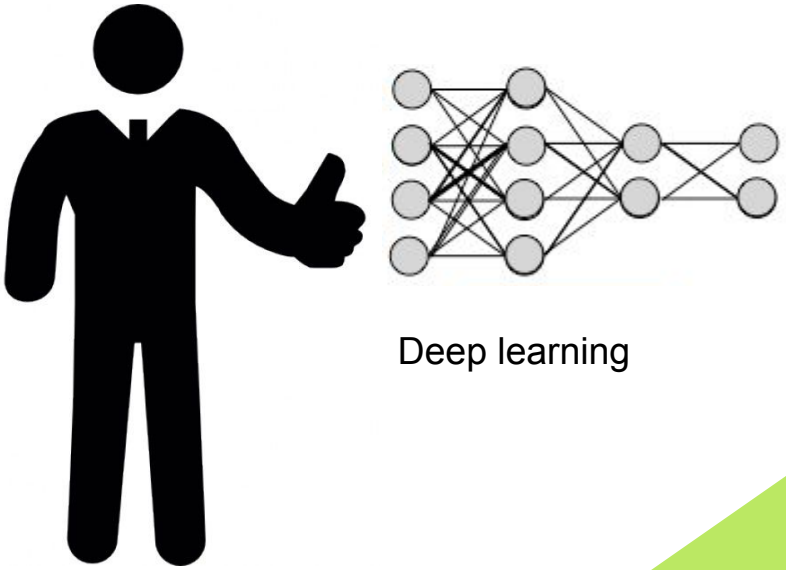
red = (0,1,0,0,0,0,0,0)

violet = (0,0,0,0,0,0,0,1)

…

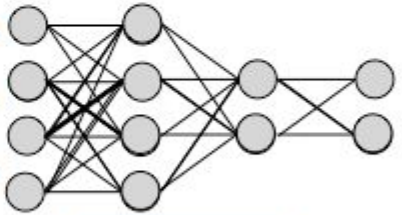
(in fact, word embeddings are used)

- Better quality
- Works for all: NLP, computer vision…



Deep learning

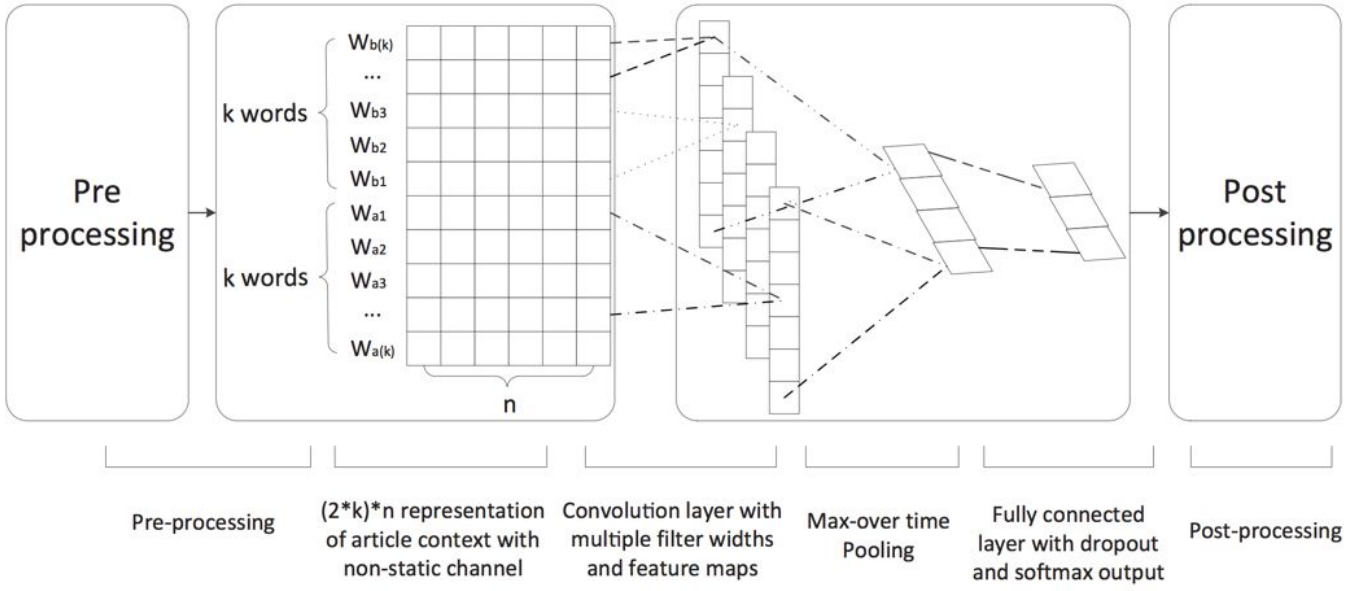
What's the catch?

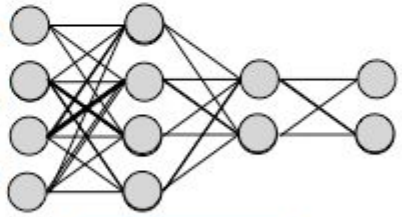


Deep learning

Deep learning approach

Much harder algorithm:

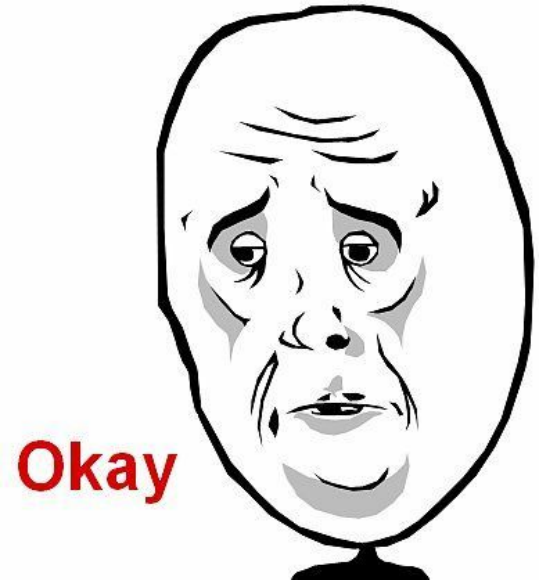


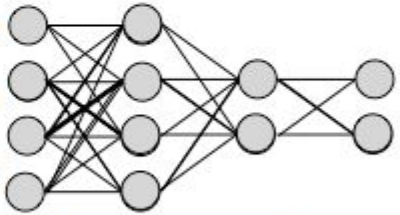


Deep learning

Deep learning approach

- Much harder algorithm
- Needs much more data
- Needs more computation





Deep learning

Difference from academia

Big data

High precision requirement

High speed requirement

Thanks!

Any questions?

Credits

Special thanks to all the people who made and released these **awesome resources** for free:

- Presentation template by SlidesCarnival
- Photographs by Unsplash