High-Ku

Beatriz A. Miranda | Project Fletcher

Hello!

I’m High-Ku

Forgive me for being
nonsensical
sometimes, I like
nature too.

You can find me in twitter at:
@the_high_ku

1.
Initial Process

Gather & Prep Data

Data Collection & Storage

Phase 1:

- Scraped from www.haikuguy.com (10,000)
- Haikuzao ‘haiku.txt’ (4,978)

Phase 2:

- Initial cleaning pre-processing
- Stored in MongoDB

Topic Modeling

Seasons & Date | Nature

Topic 1 : “on new year's day a cute little pilgrim at the gate”

Topic 2 : “plum blossoms fall in the hairdo lowering the basket”

2.
LSTM Model

Creating & Training

Understanding LSTM (Long Short-Term Memory)

- Recurrent Neural Network (RNN) architecture
- Contains feedback connections (loops)
- Weights of connections determine how “gates” operate

Key Components

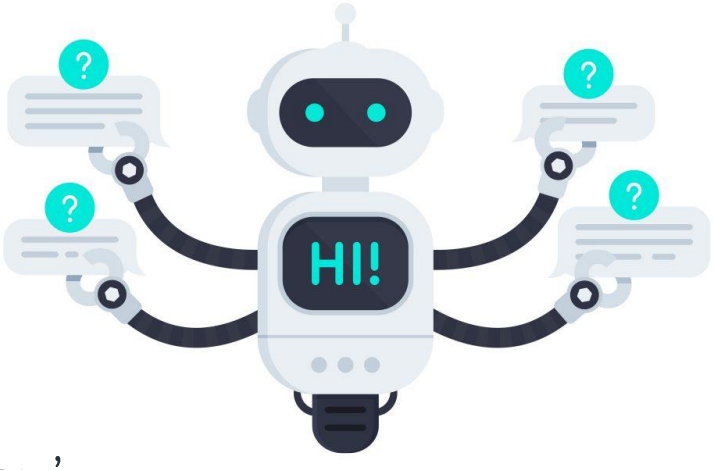
- Cells
- Input Gate
- Forget Gate (‘Keep’)
- Output Gate

“Le passé qui ne passe pas”

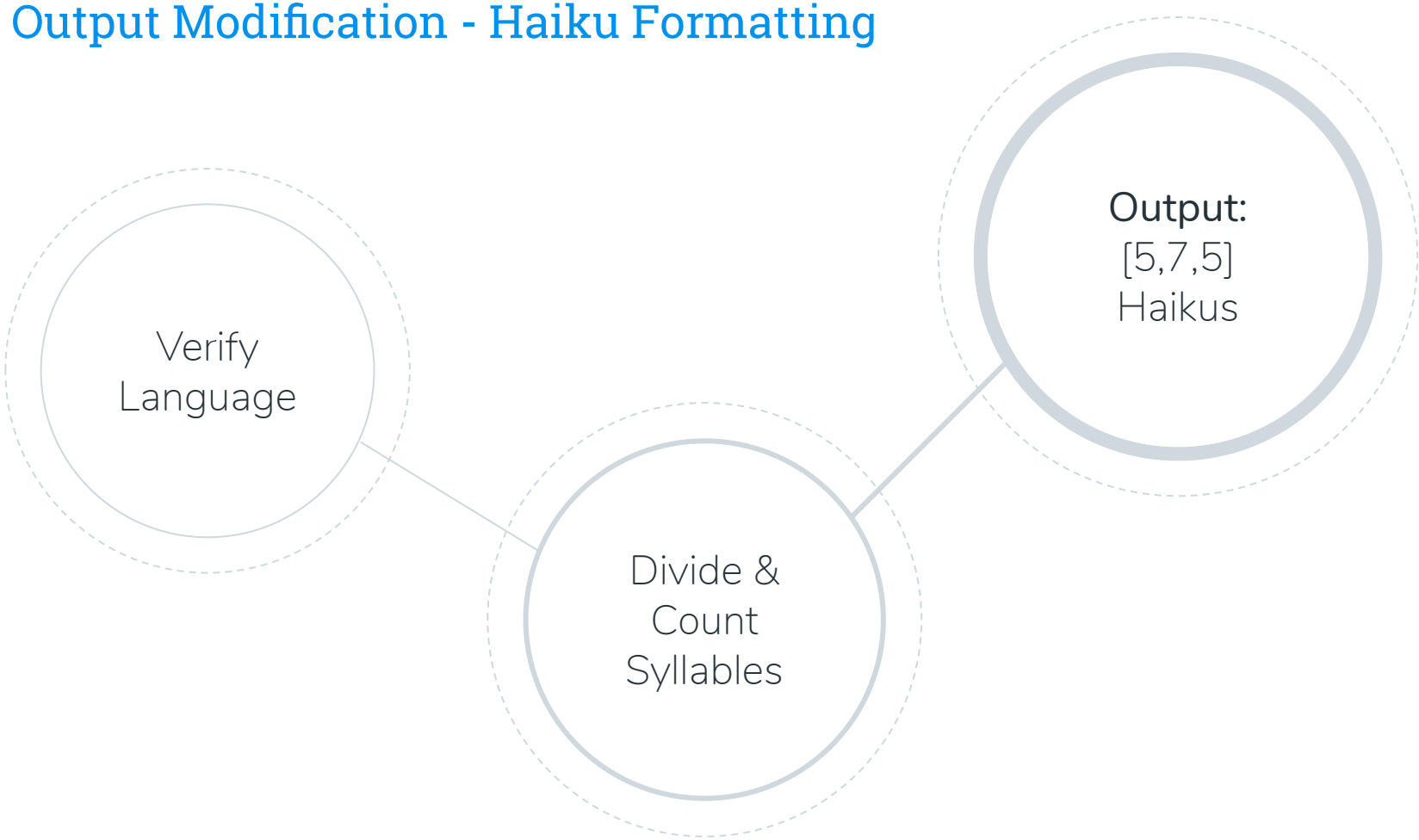
“The past that does not pass away”

Model Deets

- Sequential
- LSTM = 256 units
- Dropout = 0.2
- Dense = activation('softmax')
- Loss = 'categorical_crossentropy'
- Optimizer = 'adam'



Output Modification - Haiku Formatting

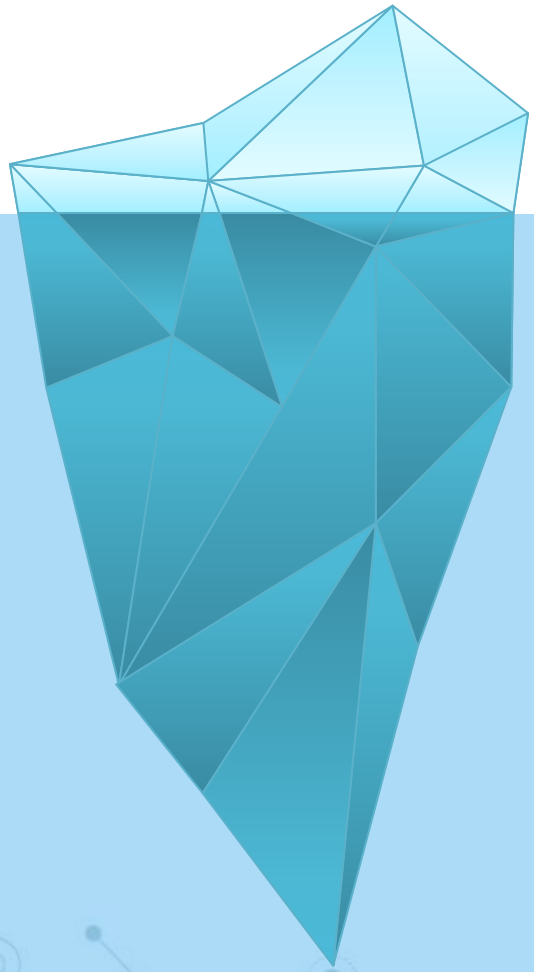


the_high_ku

Twitter bot that automatically generates haikus every 10 minutes!

Lessons Learned / Ideal Situation

CLEAN YOUR DATA | CLEAN YOUR DATA | CLEAN YOUR DATA | CLEAN YOUR DATA



WANTED TO

Spend more time preprocessing data

Train model with proper syllable count

Experiment with + layers

WILL DO

More ‘logical’ constraints

Reply haiku when tweeted at

Thanks!

Any questions?