**Intro**

I’m Joshua Banks Mailman. For my Metis Data Science Bootcamp final project, I chose to make a streamlit web app exploiting machine learning, specifically Neural Nets, and Word Embeddings, to classify and then caption images in an amusing way. Here’s an example generated by my Streamlit app.

Guitar

This caption comes from the idiom ‘Close, but no cigar’ The neural net detected the guitar, but it’s caption brings us to the singer who isn’t playing guitar but IS near someone who is. Perhaps she wishes she played guitar. Or: If *only* she did play guitar, she’d be more successful, and so on.

**Motivation**

I had three inspirations:

The New Yorker cartoon caption contest, where an uncaptioned image prompts readers to invent captions to be evaluated and voted on. People now do it with all kinds of photos on social media. The most amusing captions refer to something not actually pictured.

Composer Milton Babbitt titles his pieces with puns, which use idioms to reference musical features: A piece in a tango rhythm based on 12 tones; anotherbased on a pitch series that loops around, and an intimate violin and piano coupling based on six tones.

Rene Magritte’s titled his surrealist paintings so the meaning comes not just from the image, but rather from how we relate it to its title. He isn’t painting the object he sees with his eyes, but rather what anticipates: seeing the egg’s future is to be a bird.

And the clever code that is Cockney Rhyming Slang: ‘Bees and honey’ means ‘money’; ‘Borrow and beg’ for ‘egg’ means so forth.

All these entail *indirect* reference, so that we enjoy the process of bring it to mind the missing piece. We enjoy that *thinking around the corrner,* and so I aimed my app at doing this.

Although it could upload any image, I have it set so you choose one of 5500 images with a slider. Let’s see…

Hmmm, Perhaps you don’t have time for a *full* breakfast this morning?

**Examples**

Beetle

PAUSE.

From ‘let the genie out of the bottle,’ it’s cute to think the ladybug has some special secret, or a magical power we don’t yet know about.

Suit

With humans pictured, the captions a touch of ironic. PAUSE.

From ‘Don’t shoot yourself in the foot,’ we think: one of these two—perhaps he because he’s wearing the suit, or she, also because he’s wearing the suit—is about to do something they’ll later regret. Through the idiom and rhyme, the algorithm prompts us to draw our own conclusions.

**Data**

The source data was this: 5500 images from Imagenet. 1500 English Idioms webscraped, which I colloqualized with text substituions. To analyze and transform image labels, my algos rely on the SentiWordNet corpus of 117,000 sentences, the Gensim Google News and Text 8 Word2Vec corpuses, Carnegie Mellon corpus of English words, and the International Phonetic Alphabet.

**Nuance/Subteley**

The bees knees or Shangra la of my project’s original intent is to generate captions that point outside the picture, and I succeeded in tailoring the algorithm to do this.

Screeching

Since we know owls screech, the caption amusuingly *inverts* the original meaning of the idiom ‘preaching to the choir’. Notice that screeching is not contained in the image, but rather is brought in by the algorithm, and we make the connection between this caption and what’s pictured.

**Algorithms**

My algorithm and its components are worth explaining. From the 5500 images in IMAGENET, choose any one. The pre-trained Inception Neural Network categorizes it with one of 1000 labels, in this case *owl*.

Then three algos supply related words: NLTK Synset supplies synonyms if any.

Term frequency-Inverse Document Frequency correlation matrix I computed, with over 3 Billion datapoints, answers the question: what words are most likely to occur in a sentence with owl?

Gensium, word embedding, works by analogy. I used human action as a model, asking the Word2Vec, based on your modeling of text corpuses: if a ‘man walks’ what does an owl do?

Together these generate a Semantic Family of words: stump, hoot, screeching, chasing.

Then Phonetics algos come in. These translate each Semantic Family word into a standardized phonetic spelling and then compute the Levenshsetin edit-distance to phonetic spellings of all the words in the English dictionary, to find close matches. and I tailored it it emphasize rhymes and assonances. This gives phonetic families of words: bump, lump, stop, stem, and so on.

When running in its *nuanced* or *subtelty* mode, the original label and synonyms are supressed, so direct reference to image content is avoided.

1500 English Idioms are searched for a matches of each phonetic family word. For speed, I precomputed this as a lookup dictionary. From semantic and phonetic similarity, suitability scores are computed.

These are sorted, to inspect all the high scores. The idiom ‘Preaching to the choir’ ranks best, so it’s selected, and the rhyming word ‘screeching’ is subsituted, to produce the caption, which is then placed under the image.

Here are some of my favorites so far.

**Final examples**

peel

Here the algo settled on ‘The real McCoy’, meaning the only real thing. The photo with caption conjur the peeling an apple, and thereby indirectly suggest that perhaps apples, or some specific varieties of apple, are the only real fruit, or the best fruit, or maybe it’s apple computers or New York City.

Deisel

This one might be a headline for a news article about the geopolitics of renewable energy.

Trow

sers

This one one’s a touch ironic. Based on the idiom: ‘All talk and no trousers, which means something like ‘talk is cheap,’’all talk an no action’, or ‘you’ve talked the talk but can you walk the walk?’ which is funny because the photo presents a bunch of prominant signage in a laundreyroom, indicating where to install money for soap and commanding “DO NOT…” do this or that, yet without much evidencc of any launderying going on—or trousers for that matter.

Bones

Sometimes I was momentarily fooled into thinking the algorithm really does have a mischevous sense of humor, with perhaps, more than ourselves, an insightful appreciation of the seemingly carefree quality of invertibrate animals.

Thanks for watching.