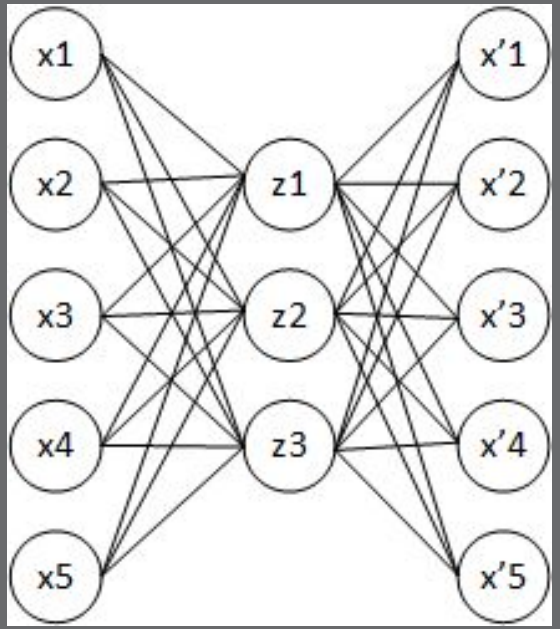
Autoencoders

Brent Lemieux

galvanize



galvanize

Resources for Further Study

More resources:

https://blog.keras.io/building-autoencoders-in-keras.html (Examples w/ code)

http://www.deeplearningbook.org/contents/autoencoders.html (Conceptual)

http://kvfrans.com/variational-autoencoders-explained/ (Generative models)

galvanize

Objectives

1. Introduce autoencoders
2. Draw an example of a standard autoencoder
3. State some other uses

galvanize

What is an Autoencoder?

Autoencoders are a type of data compression algorithm.

More well known compression algorithms: JPEG, MP3, etc.

Are Autoencoders good at compression?

What is an Autoencoder?

Autoencoders are a type of data compression algorithm.

More common compression algorithms: JPEG, MP3, etc.

Are Autoencoders good at compression?

No, not really…

But that is okay for our purposes -- in fact it's a feature

Autoencoders - Why use them?

Reduce the dimensions of our feature space to:

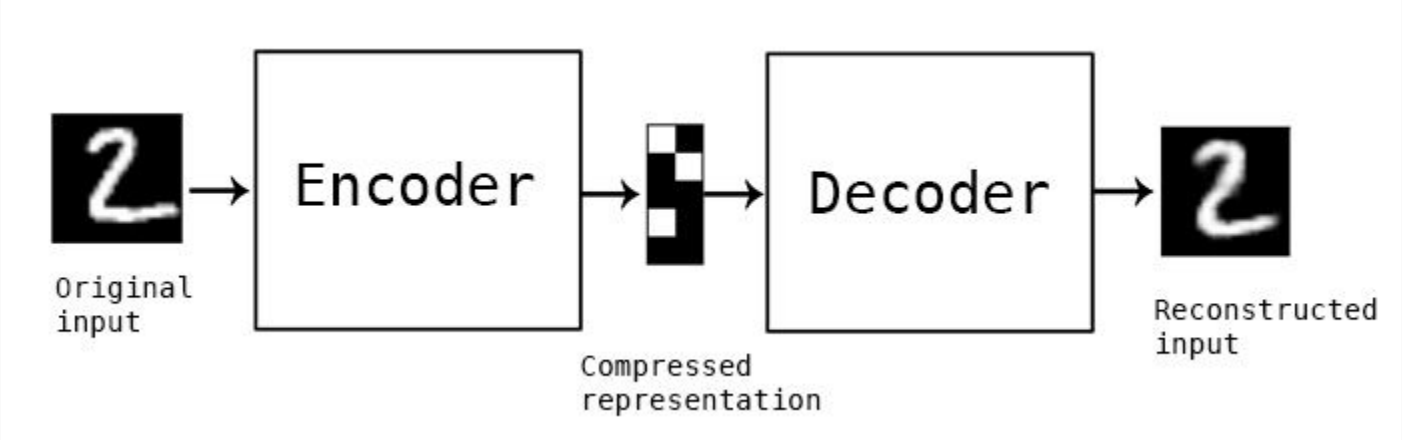
- Learn latent features
- Increase efficiency

Terminology

- Embed (in lower dimensional feature space)
- Encode
- Compress

All of these terms are referring to the same thing… Don't confuse with word embeddings which Adam will talk about tomorrow.

Autoencoder - Basic Functionality



galvanize

Autoencoder Trivia

Autoencoders are lossy -- what does that mean?

Autoencoder trivia

Autoencoders are lossy -- what does that mean?

The data is degraded when encoded into lower dimensions.

Autoencoder trivia

Autoencoders are lossy -- what does that mean?

The data is degraded when encoded into lower dimensions.

Autoencoders are data specific -- any intuition what this might mean?

Autoencoder trivia

Autoencoders are lossy -- what does that mean?

The data is degraded when encoded into lower dimensions.

Autoencoders are data specific -- any intuition what this might mean?

They work much better when trained on similar data.

Autoencoder trivia

Autoencoders are lossy -- what does that mean?

Autoencoders are data specific -- any intuition what this might mean?

Autoencoders are self-supervised machine learning -- what does that mean?

The data is degraded when encoded into lower dimensions.

They work much better when trained on similar data.

Autoencoder trivia

Autoencoders are lossy -- what does that mean?

The data is degraded when encoded into lower dimensions.

Autoencoders are data specific -- any intuition what this might mean?

They work much better when trained on similar data.

Autoencoders are self-supervised machine learning -- what does that mean?

They train themselves by learning from the reconstruction error.

Objectives

Introduce autoencoders

2. Draw an example of a simple autoencoder

3. State some other uses

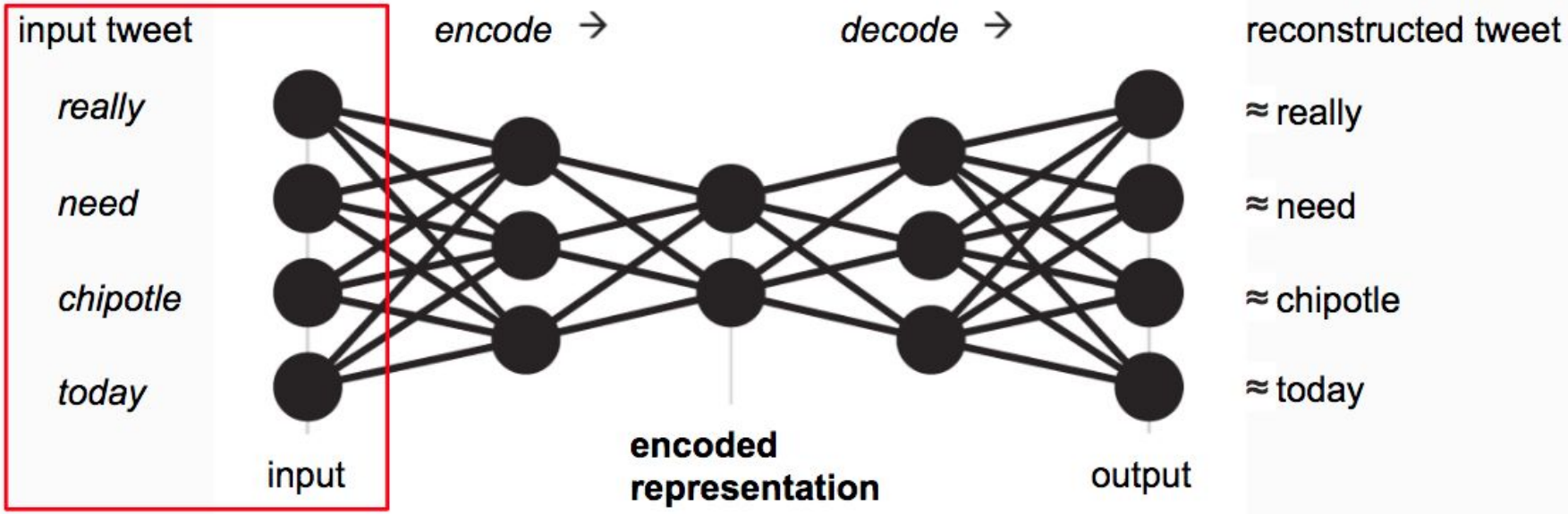
Dimensionality Reduction

We are primarily focused on how they can reduce dimensionality to learn latent features.

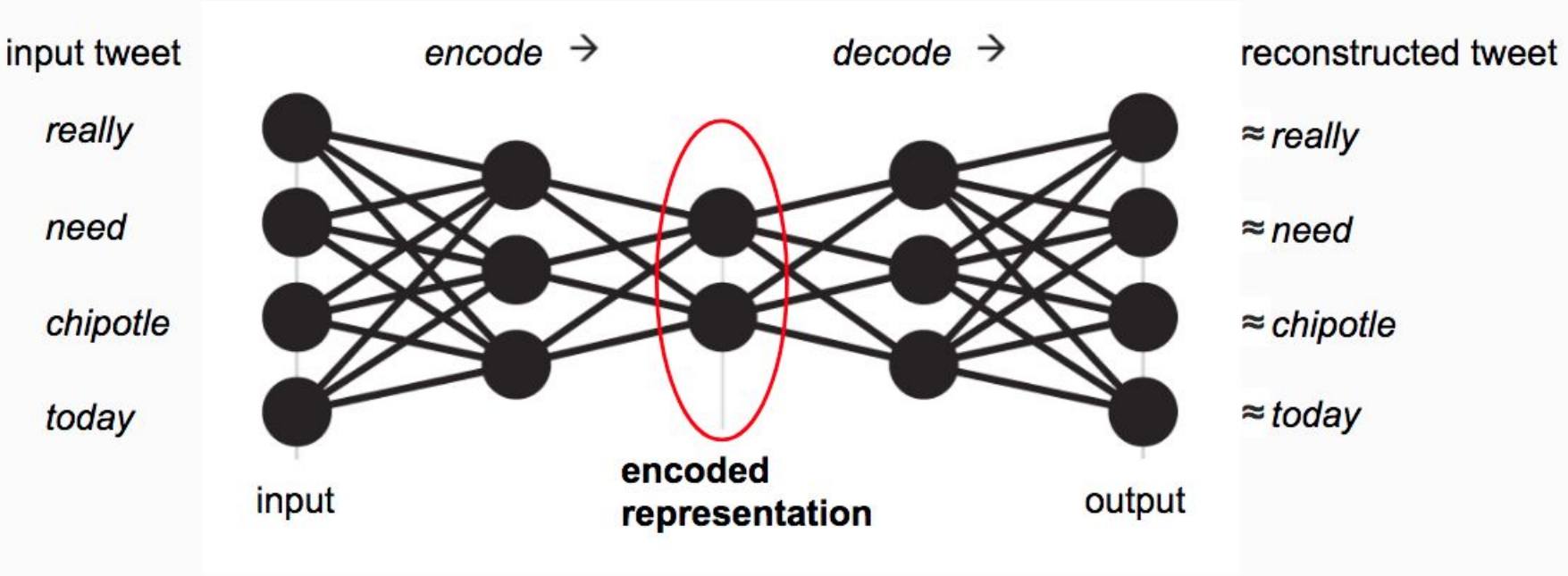
This is accomplished by embedding inputs into lower dimensional feature space.

Note: Don't give them too much capacity.

Example: Embedding (Encoding) Tweets

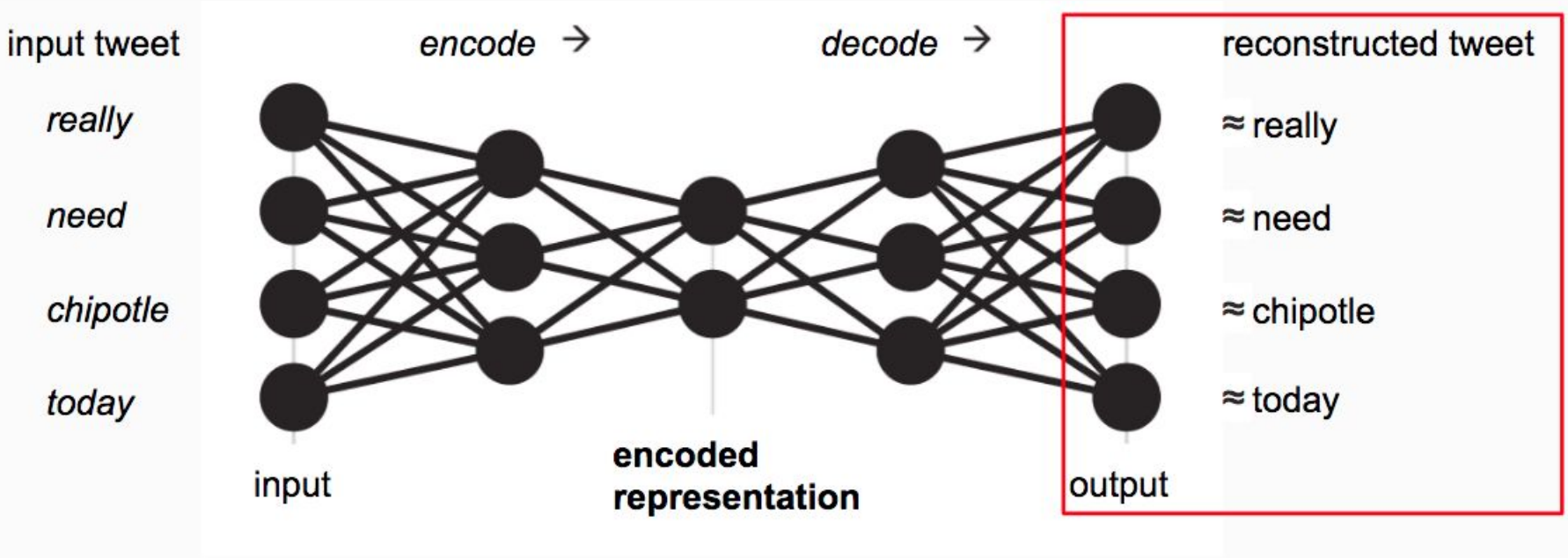


Example: Embedding (Encoding) Tweets

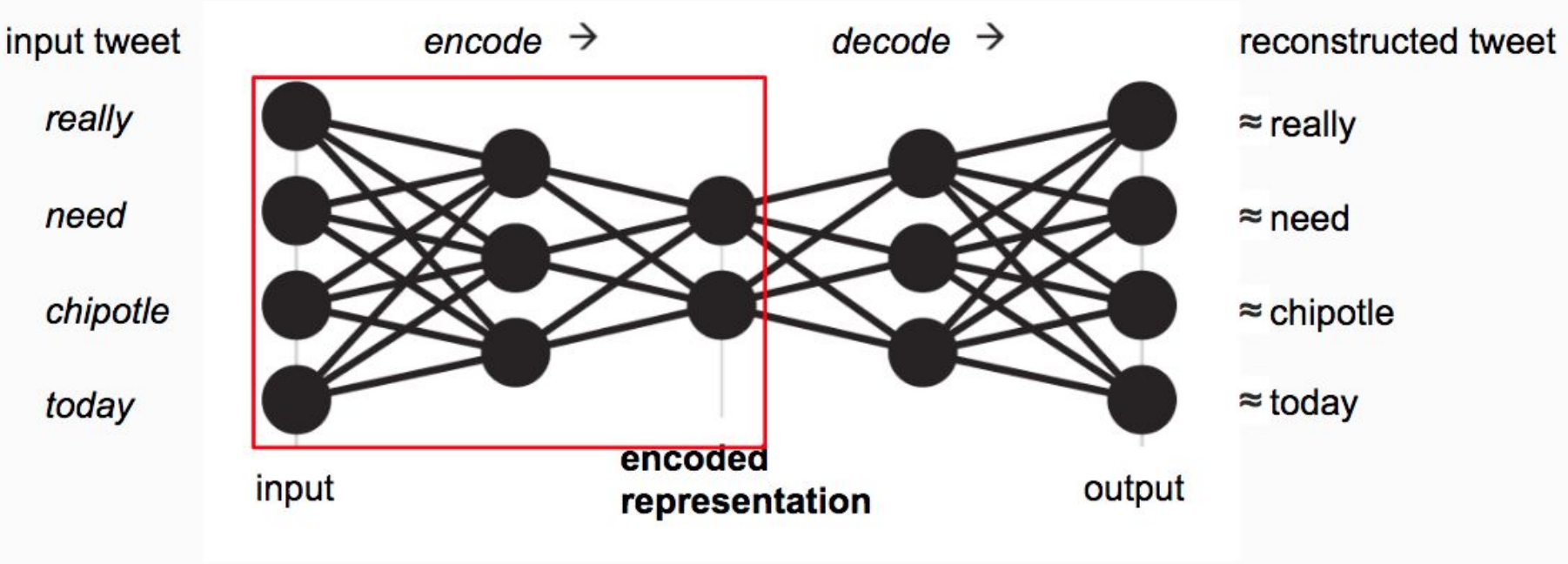


galvanize

Example: Embedding (Encoding) Tweets



Example: Embedding (Encoding) Tweets



galvanize

Objectives

- ✓ Introduce autoencoders
- ✓ Show an example of a standard autoencoder
3. State some other uses

Are they only used for text?

No!

Autoencoders can be used for images, audio and other data types, as well.

They can also be used for tasks other than dimensionality reduction.

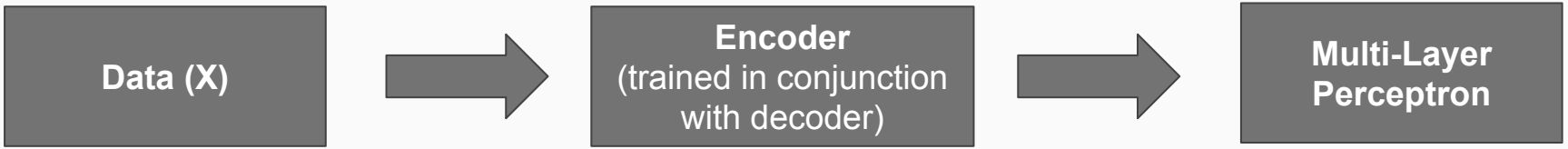
galvanize

Other types explained...

- Sparse Autoencoders
- Denoising Autoencoders
- Variational Autoencoders

Other Applications

The encoded representation can be fed into another supervised neural net -- or other machine learning algorithm



Note: can feed into other types of models as well...

galvanize

Autoencoder Trivia Round 2

Why might we like that autoencoders are lossy/can't re-represent the input perfectly?

Autoencoder Trivia Round 2

Why might we like that autoencoders are lossy/can't re-represent the input perfectly?

Their lossy-ness helps them learn attributes/latent features about the data.

galvanize

Autoencoder Trivia Round 2

Why might we like that autoencoders are lossy/can't re-represent the input perfectly?

Their lossy-ness helps them learn attributes/latent features about the data.

What are the two functions that make up an autoencoder? What are their purposes?

Autoencoder Trivia Round 2

Why might we like that autoencoders are lossy/can't re-represent the input perfectly?

Their lossy-ness helps them learn attributes/latent features about the data.

What are the two functions that make up an autoencoder? What are their purposes?

The ***encoder*** compresses the data into lower dimensions. The ***decoder*** reconstructs the code into an approximation of the input.