EEP 596: LLMs: From Transformers to GPT || Lecture 13 (Part 1)

Dr. Karthik Mohan

Univ. of Washington, Seattle

February 20, 2024

Part 2
- Tomorrow
- 4.30PM - 5.50PM

Next Assignment
- APE
- Stable Diffusion

Final presentation
- Demo from MP2
- Also Stable diffusion

Final presentation
Tuesday of Finalsweek (March 12th)

Deep Learning and Transformers References

Deep Learning

Great reference for the theory and fundamentals of deep learning: Book by Goodfellow and Bengio et al Bengio et al

Deep Learning History

Embeddings

SBERT and its usefulness

SBert Details

Instacart Search Relevance

Instacart Auto-Complete

Attention

Illustration of attention mechanism

Generative AI References

Prompt Engineering

Prompt Design and Engineering: Introduction and Advanced Methods

Retrieval Augmented Generation (RAG)

Toolformer

RAG Toolformer explained

Misc GenAI references

Time-Aware Language Models as Temporal Knowledge Bases

Generative AI references

Stable Diffusion

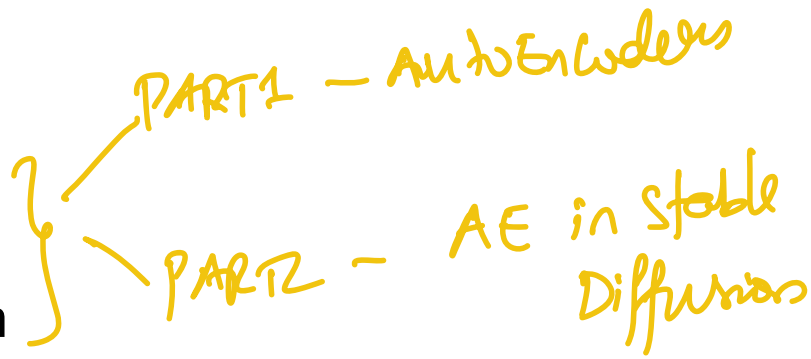
- Diffusion Explainer: Visual Explanation for Text-to-image Stable Diffusion
- The Illustrated Stable Diffusion

Previous Lecture

- Toolformer
- Introduction to Stable Diffusion

This Lecture

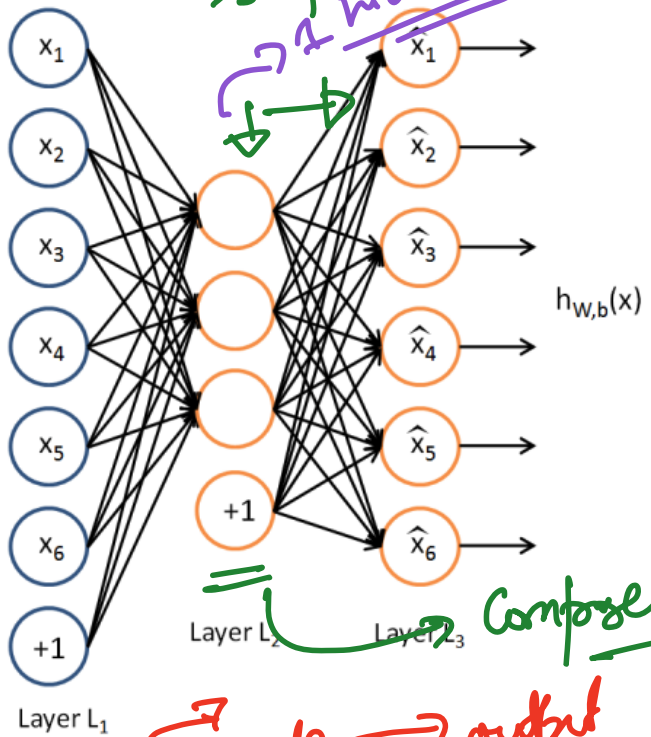
- Stable Diffusion model
- Understanding Stable Diffusion



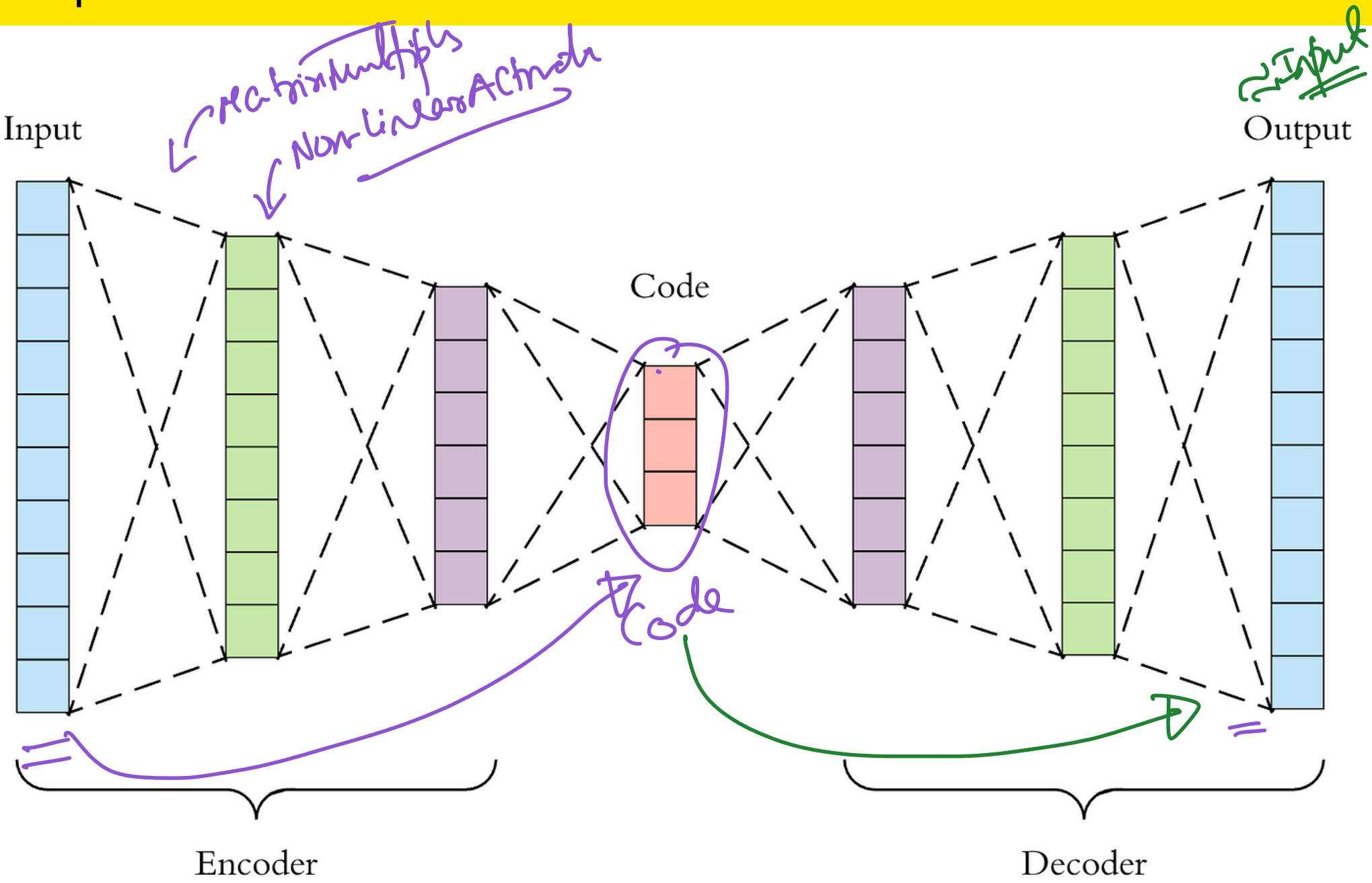
Stable Diffusion Explained

- Based on the concept of "de-noising auto encoders" and the use of text prompt to *guide the de-noising*
- Stable diffusion is also trained to successfully de-noise and increase the resolution of the image using text guidance

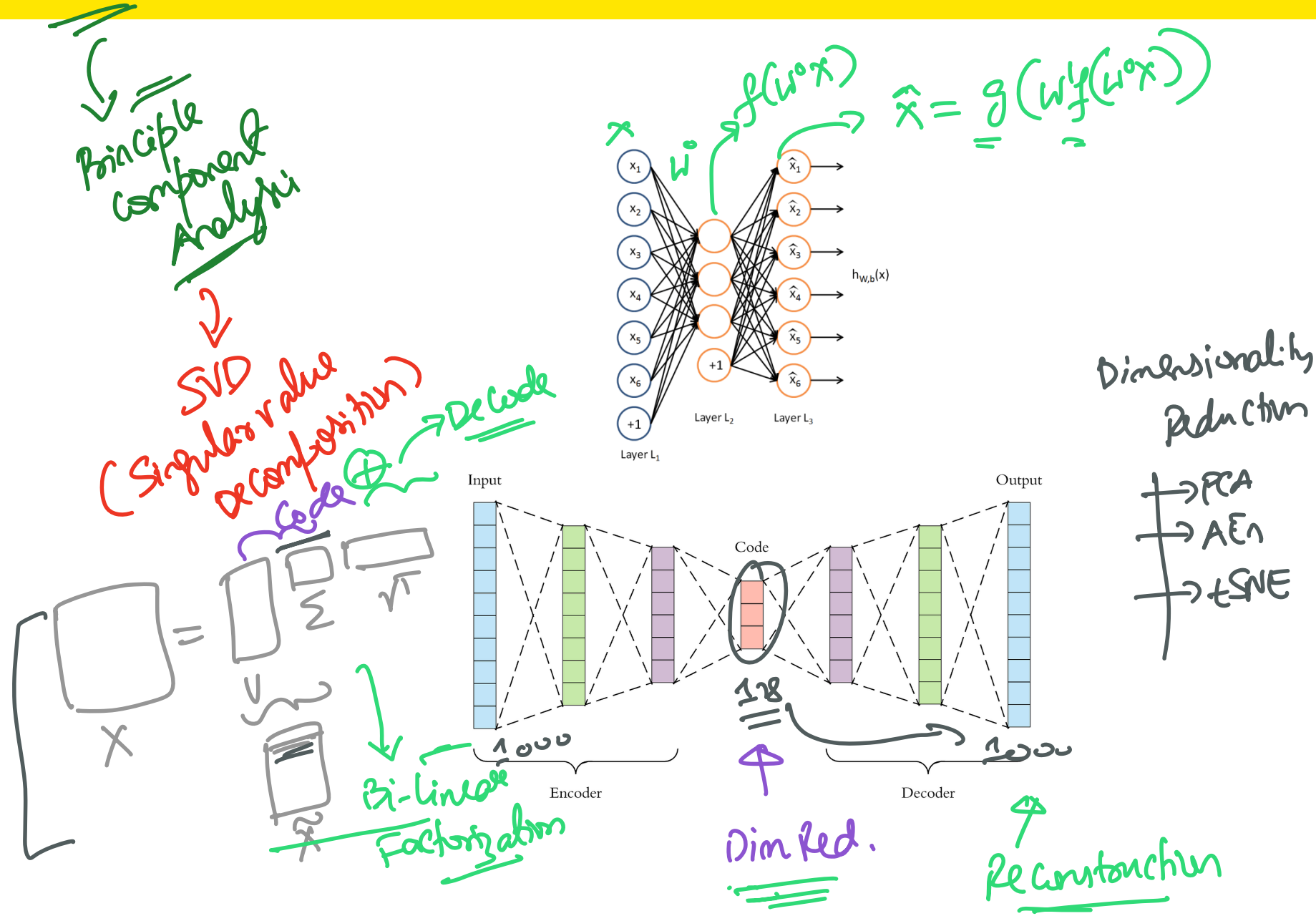
Auto Encoders



Deep Auto Encoders



PCA vs Auto-Encoders



ICE #1

PCA vs Auto Encoder

Which of the following statements are true ?

1. Both PCA and Auto Encoders serve the purpose of dimensionality reduction
2. They are both linear models but one uses a neural nets architecture and the other is based on projections
3. PCA is robust to outliers while Auto Encoders are not
4. Auto Encoders can compress images better than PCA

AutoEncoders and Dimensionality Reduction

Visualization Performance

Auto Encoder Reference Paper

(take a look on your own!)

AutoEncoders and Dimensionality Reduction

Reading Reference for AE Dimensionality Reduction

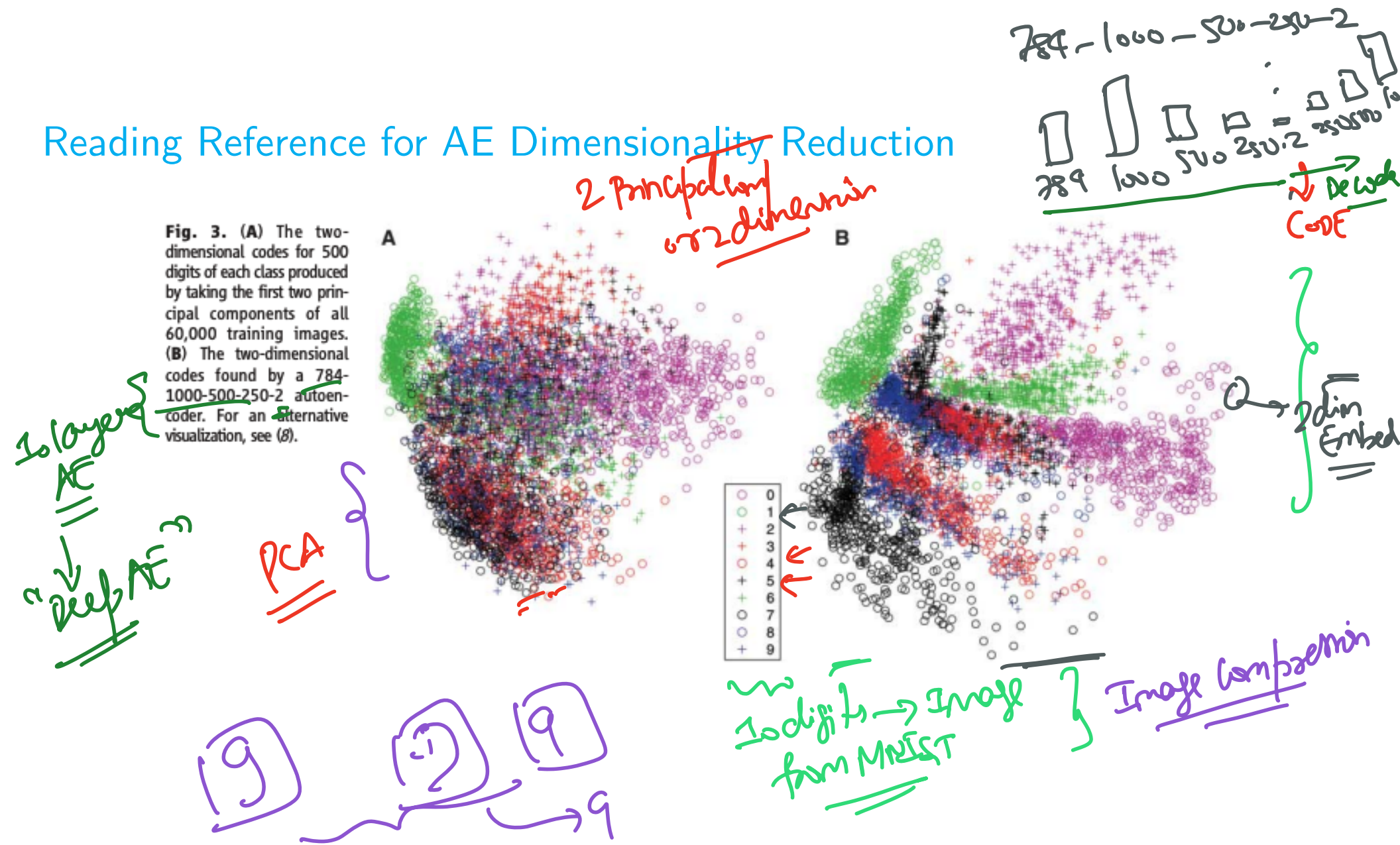
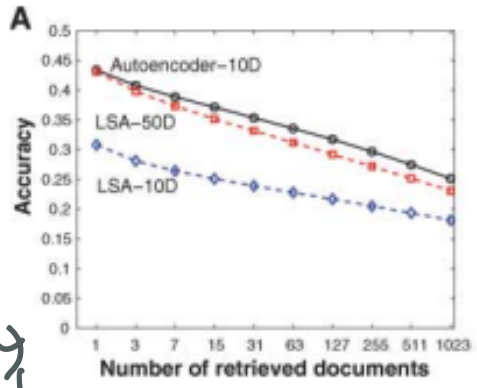


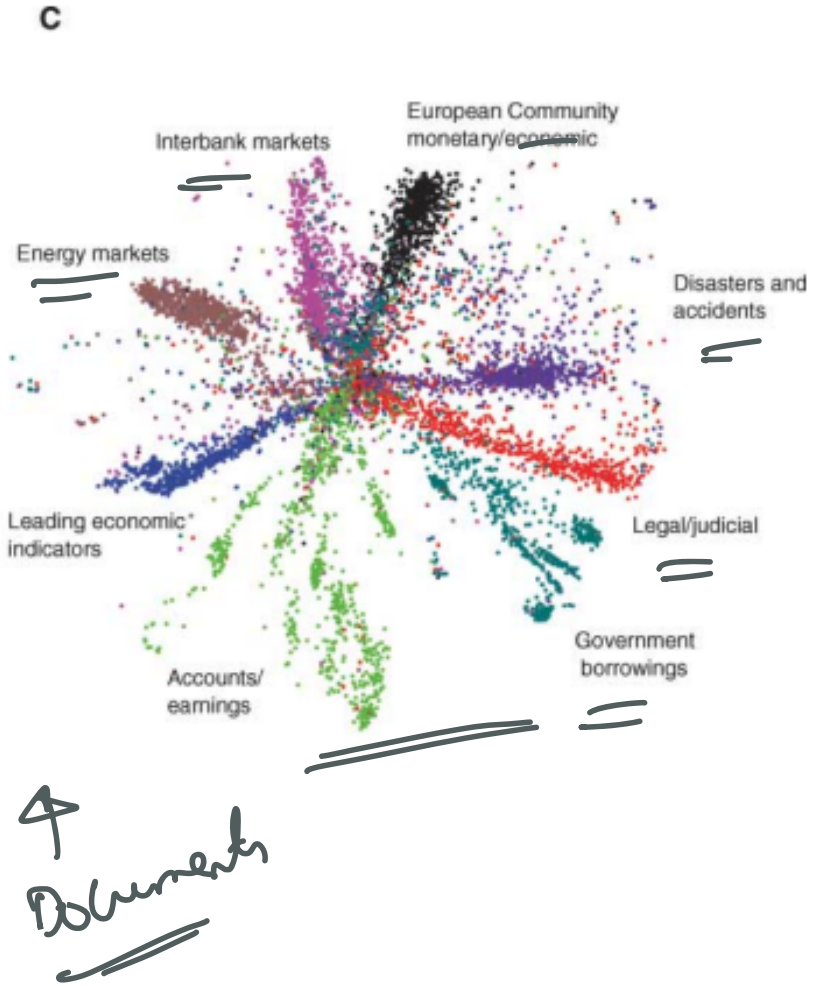
Fig. 3. **(A)** The two-dimensional codes for 500 digits of each class produced by taking the first two principal components of all 60,000 training images. **(B)** The two-dimensional codes found by a 784-1000-500-250-2 autoencoder. For an alternative visualization, see (*8*).

AutoEncoders and Dimensionality Reduction

Reading Reference for AE Dimensionality Reduction

Fig. 4. (**A**) The fraction of retrieved documents in the same class as the query when a query document from the test set is used to retrieve other test set documents, averaged over all 402,207 possible queries. (**B**) The codes produced by two-dimensional LSA. (**C**) The codes produced by a 2000-500-250-125-2 autoencoder.





AutoEncders Summary

1. Auto-Encoders are a method for dimensionality reduction and can do better than PCA for visualization

AutoEncders Summary

1. Auto-Encoders are a method for dimensionality reduction and can do better than PCA for visualization
2. Use Neural Networks architecture and hence can encode non-linearity in the embeddings

AutoEncders Summary

1. Auto-Encoders are a method for dimensionality reduction and can do better than PCA for visualization
2. Use Neural Networks architecture and hence can encode non-linearity in the embeddings
3. Anything else?

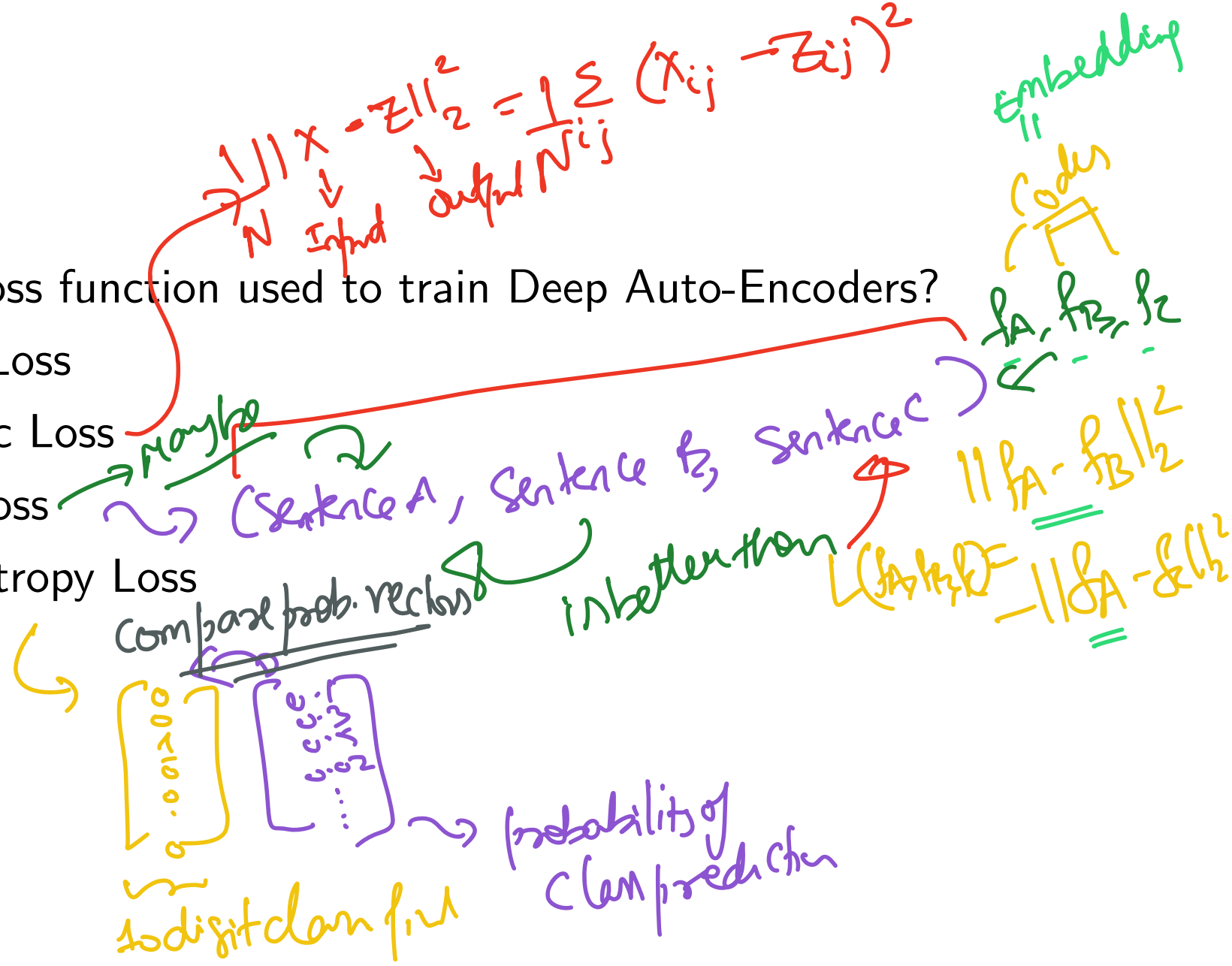
AutoEncders Summary

1. Auto-Encoders are a method for dimensionality reduction and can do better than PCA for visualization
2. Use Neural Networks architecture and hence can encode non-linearity in the embeddings
3. Anything else?
4. Auto Encoders can learn convolutional layers instead of dense layers - Better for images! More flexibility!!

ICE #2: Loss Function for Auto-Encoders

What is the loss function used to train Deep Auto-Encoders?

1. Logistic Loss
2. Quadratic Loss
3. Triplet Loss
4. Cross-Entropy Loss



Removing obstacles in images

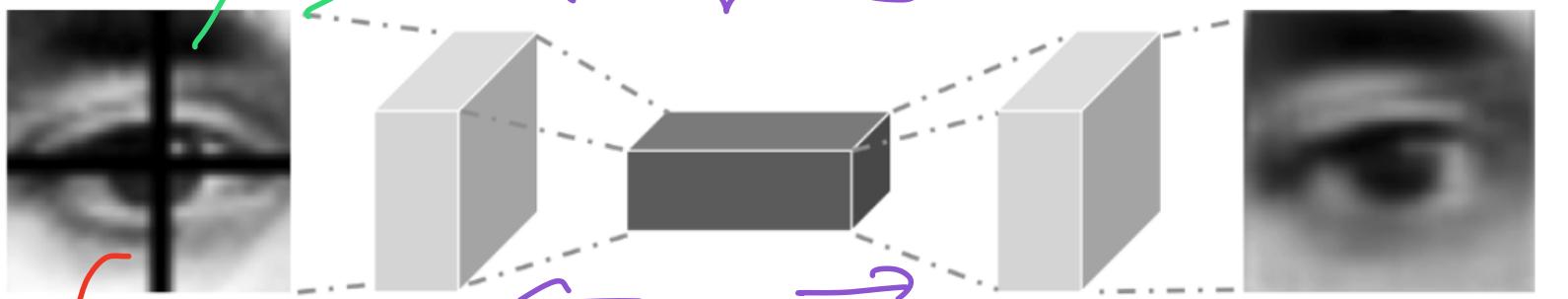


Figure 12: Reconstructed image from missing image [14]

Removing obstacles in images

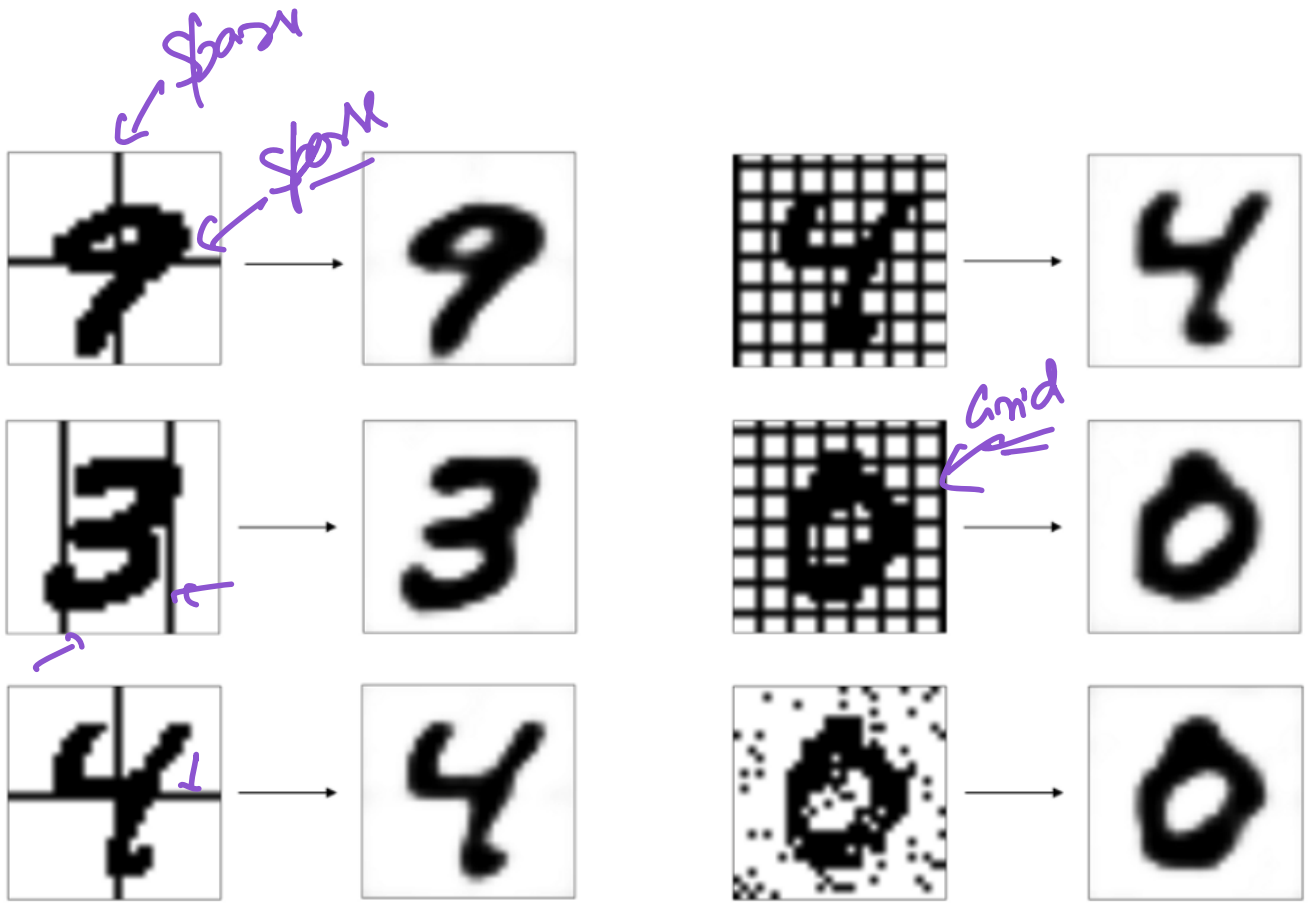
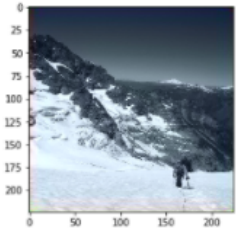
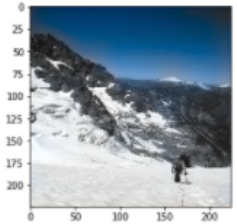
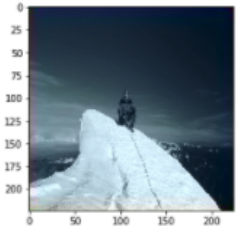
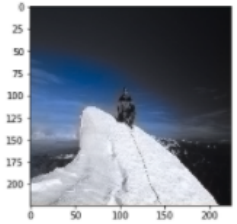
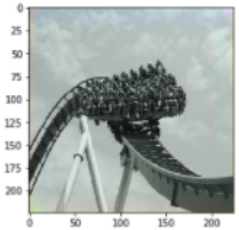
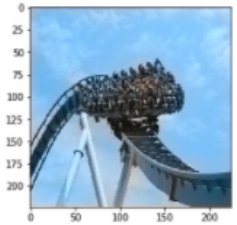


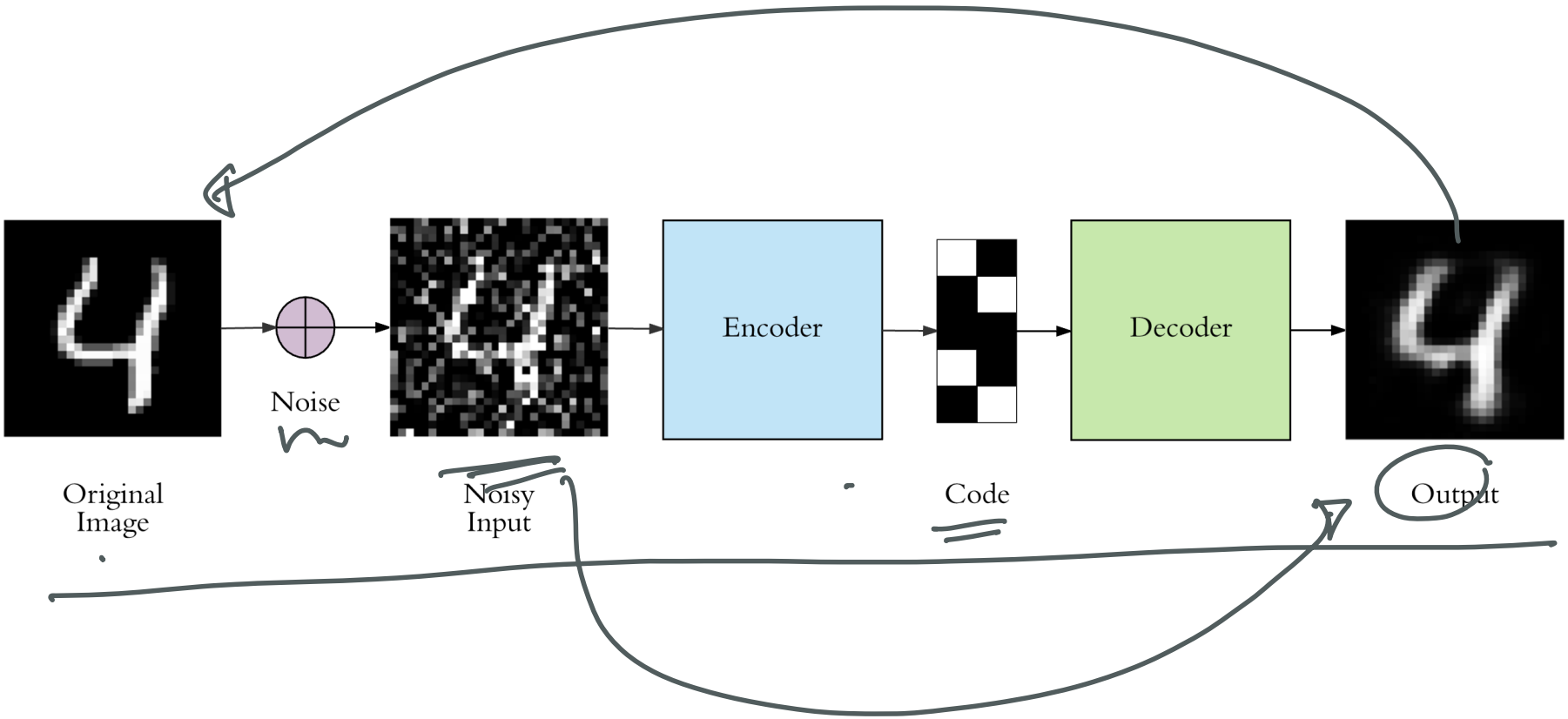
Figure 13: Source [15]

Coloring Images

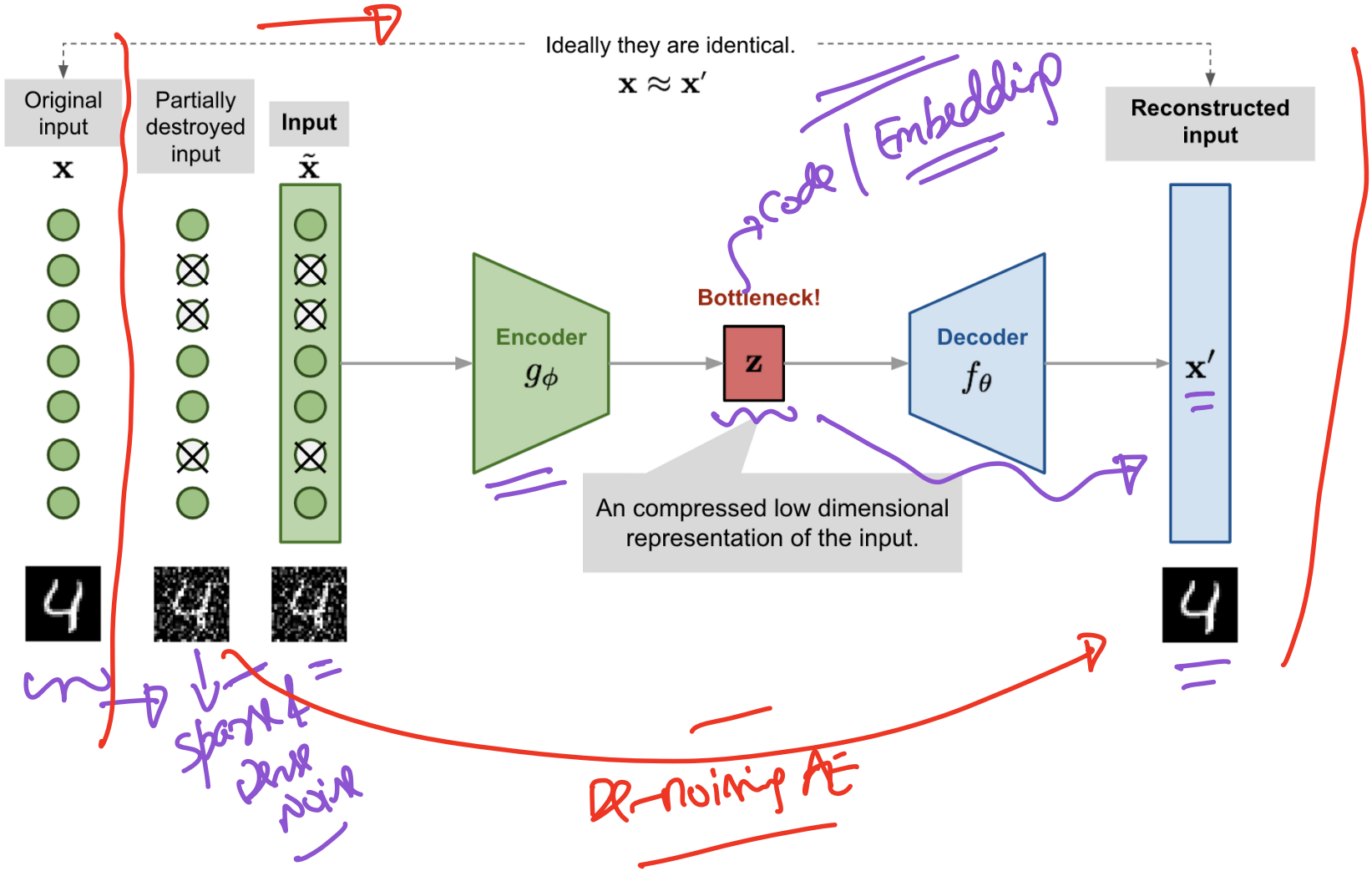
"Coloring" Auto Encoder

Gray Image	Vanilla Autoencoder	Merge Model (YCbCr)	Merge Model (LAB)	Original
				
				
				

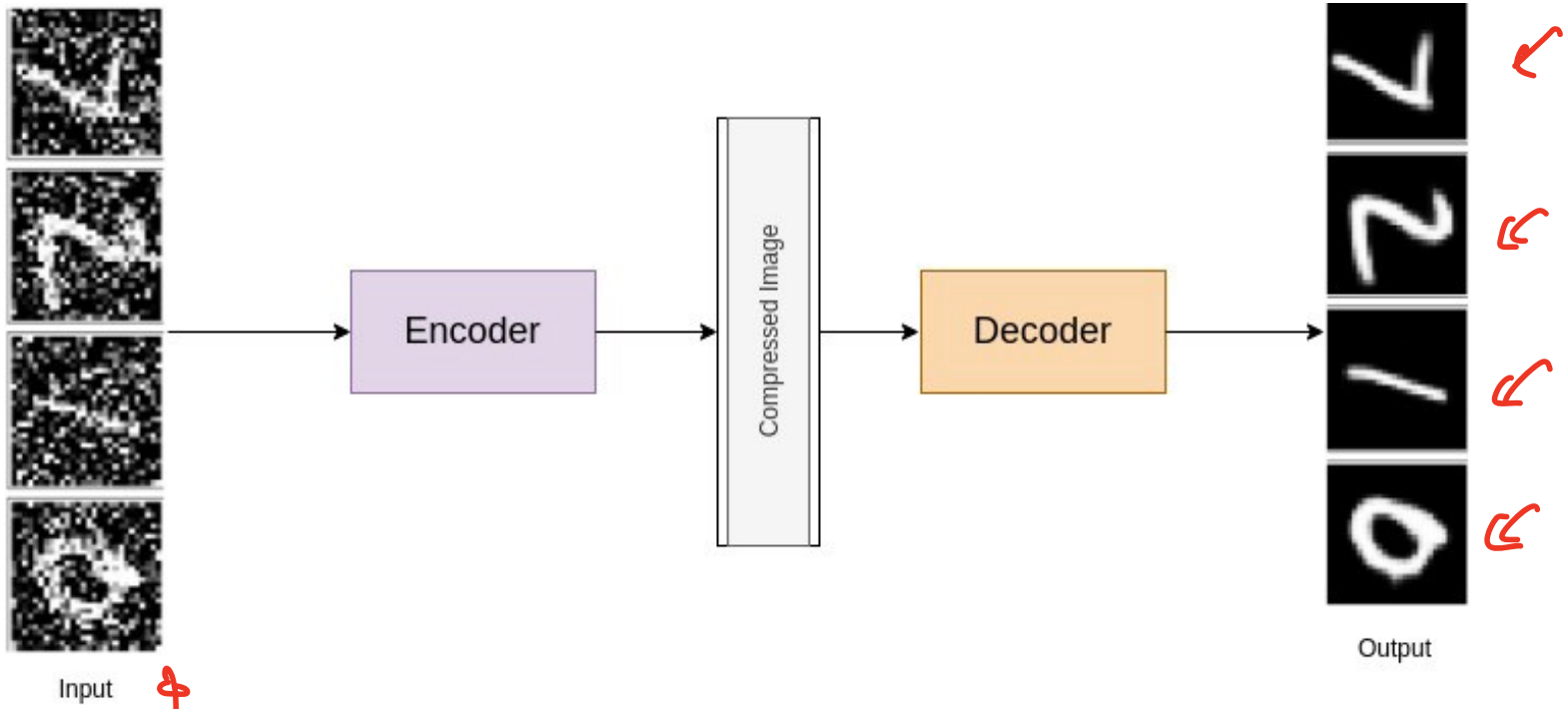
De-noising Auto Encoders



De-noising Auto Encoders



De-noising Auto Encoders



"Grainy" MNIST Digits

De-noising Auto Encoders

Details

- Just like an Auto Encoder

De-noising Auto Encoders

Details

- Just like an Auto Encoder
- Difference: Noise is injected in the inputs on purpose but output is a clean data point.

De-noising Auto Encoders

Details

- Just like an Auto Encoder
- Difference: Noise is injected in the inputs on purpose but output is a clean data point.
- This forces the Auto Encoder to "de-noise" data, esp. useful for images!

De-noising Auto Encoders

Details

- Just like an Auto Encoder
- Difference: Noise is injected in the inputs on purpose but output is a clean data point.
- This forces the Auto Encoder to "de-noise" data, esp. useful for images!
- Esp. useful for a category of objects or images (e.g. digit recognition or face recognition, etc)

↪ Also for "Stable Diffusion"

ICE #3

Unsupervised Learning

Which of these is NOT an example of unsupervised learning?

1. Perceptron
2. Auto Encoder
3. De-noising Auto Encoder
4. K-means++
5. None of the above
6. All of the above

Breakouts Time #1

5 mins

Discuss in your groups what are some real-world applications of any or many of the Auto Encoder Architectures we discussed so far you can think of in your area of work or in a standard context e.g. images.

Lecture 13 - Part 2

1. Use of De-noising Auto-Encoders for **Stable Diffusion**
2. Architecture behind Stable-Diffusion
3. Demo for understanding stable-diffusion process