


[Get started](#)
[Open in app](#)

# Venkata Krishna Jonnalagadda

32 Followers About

[Follow](#)


## Sparse, Stacked and Variational Autoencoder



Venkata Krishna Jonnalagadda Dec 6, 2018 · 11 min read

### Autoencoder:

An Autoencoder is a neural network which is an unsupervised learning algorithm which uses back propagation to generate output value which is almost close to the input value. Lets see now how an autoencoder actually works in detail. It takes input such as image or vector anything with a very high dimensionality and run through the neural network and tries to compress the data into a smaller representation with two principal components. The first one is the encoder which is simply a bunch of layers that are full connected layers or convolutional layers which are going to take the input and compress it to a smaller representation which has less dimensions then the input which is known as bottleneck. Now from this bottleneck it tries to reconstruct the input using full connected layers or convolutional layers.

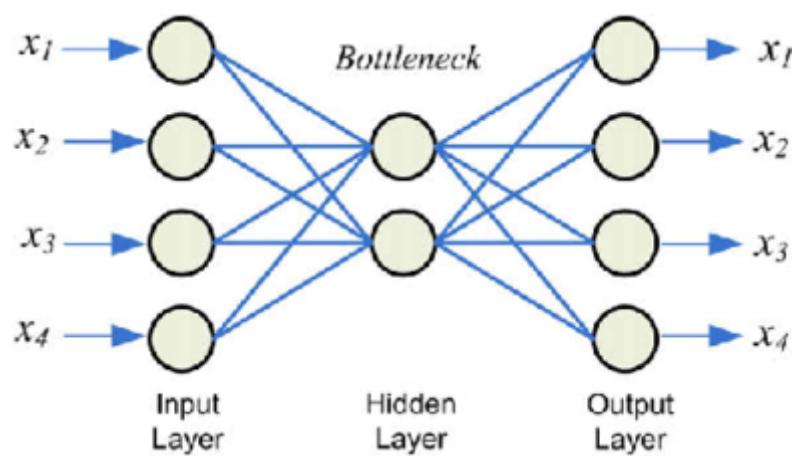


Figure 1: Layers in an autoencoder [1]


[Get started](#)
[Open in app](#)

An autoencoder takes the input image or vector and learns code dictionary that changes the raw input from one representation to another. Where in sparse autoencoders with a sparsity enforcer that directs a single layer network to learn code dictionary which in turn minimizes the error in reproducing the input while restricting number of code words for reconstruction[8].

The sparse autoencoder consists a single hidden layer, which is connected to the input vector by a weight matrix forming the encoding step. The hidden layer then outputs to a reconstruction vector, using a tied weight matrix to form the decoder[8].

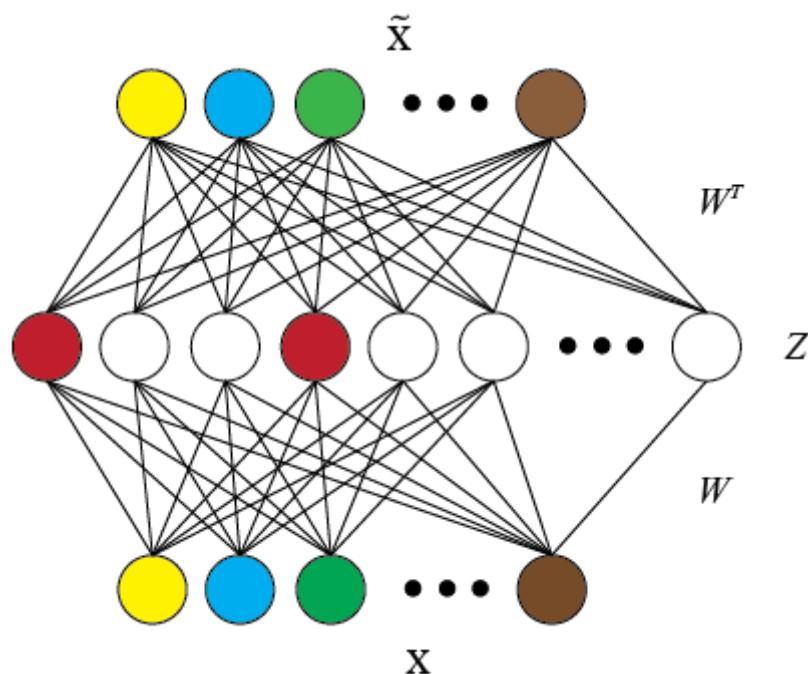


Figure 2: Sparse autoencoder[8]

An advancement to sparse autoencoders is the k-sparse autoencoder. Here we choose k neurons with highest activation functions ignoring other activation functions using ReLU activation functions and adjusting the threshold to find the largest neurons. This tune the value of k to obtain sparsity level best suited for the dataset[8].

### Stacked Autoencoder:

A stacked autoencoder is a neural network consist several layers of sparse autoencoders where output of each hidden layer is connected to the input of the successive hidden layer.

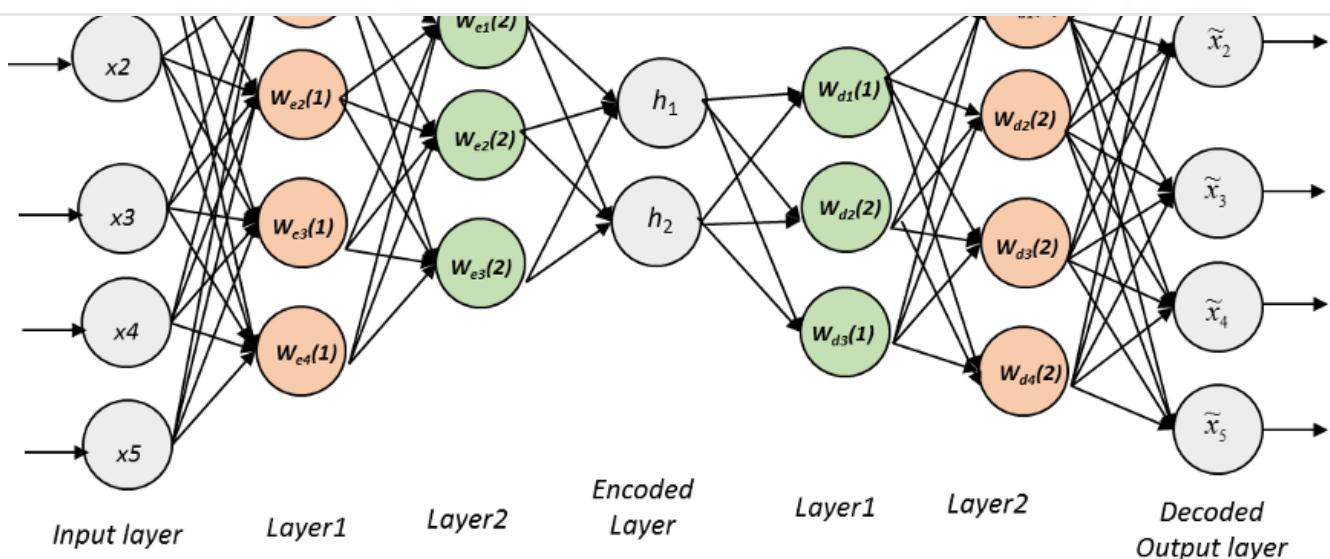

[Get started](#)
[Open in app](#)


Figure 3: Stacked Autoencoder[3]

As shown in Figure above the hidden layers are trained by an unsupervised algorithm and then fine-tuned by a supervised method. Stacked autoencoder mainly consists of three steps[4].

- Train autoencoder using input data and acquire the learned data.
- The learned data from the previous layer is used as an input for the next layer and this continues until the training is completed.
- Once all the hidden layers are trained use the backpropagation algorithm to minimize the cost function and weights are updated with the training set to achieve fine tuning.

The recent advancements in Stacked Autoencoder is it provides a version of raw data with much detailed and promising feature information, which is used to train a classifier with a specific context and find better accuracy than training with raw data.

Stacked autoencoder improving accuracy in deep learning with noisy autoencoders embedded in the layers [5].

Stacked autoencoder are used for P300 Component Detection and Classification of 3D Spine Models in Adolescent Idiopathic Scoliosis in medical science. Classification of the rich and complex variability of spinal deformities is critical for comparisons between treatments and for long-term patient follow-ups.


[Get started](#)
[Open in app](#)

The basic idea behind a variational autoencoder is that instead of mapping an input to a fixed vector, input is mapped to a distribution. The only difference between the autoencoder and variational autoencoder is that bottleneck vector is replaced with two different vectors one representing the mean of the distribution and the other representing the standard deviation of the distribution.

Loss function for variational autoencoder

$$li(\theta, \phi) = -E_{z \sim q\theta(z|x)} [\log p\phi(x_i|z)] + KL(q\theta(z|x)||p(z))$$

The loss function in variational autoencoder consists of two terms. First, one represents the reconstruction loss and the second term is a regularizer and KL means Kullback-Leibler divergence between the encoder's distribution  $q\theta(z|x)$  and  $p(z)$ . This divergence measures how much information is lost when using  $q$  to represent  $p$ .

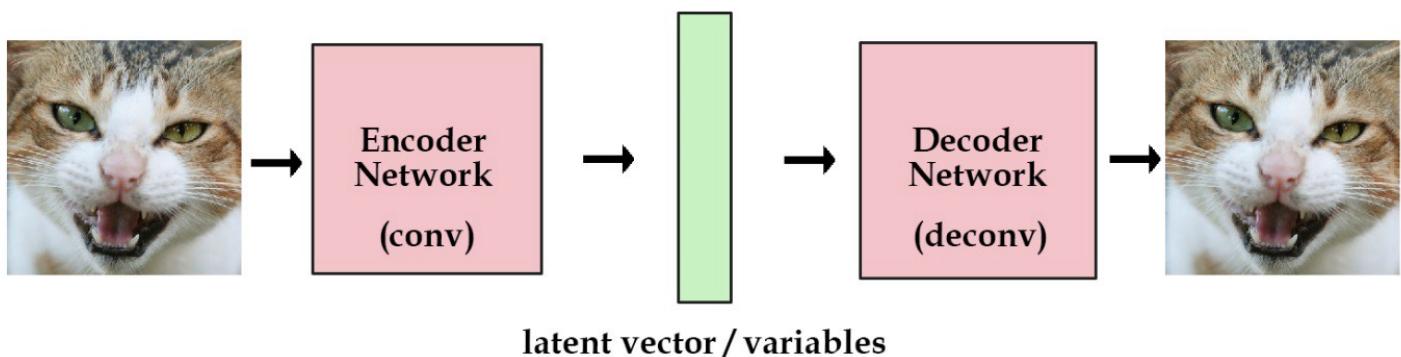


Figure 4: Layers in an Autoencoder [2]

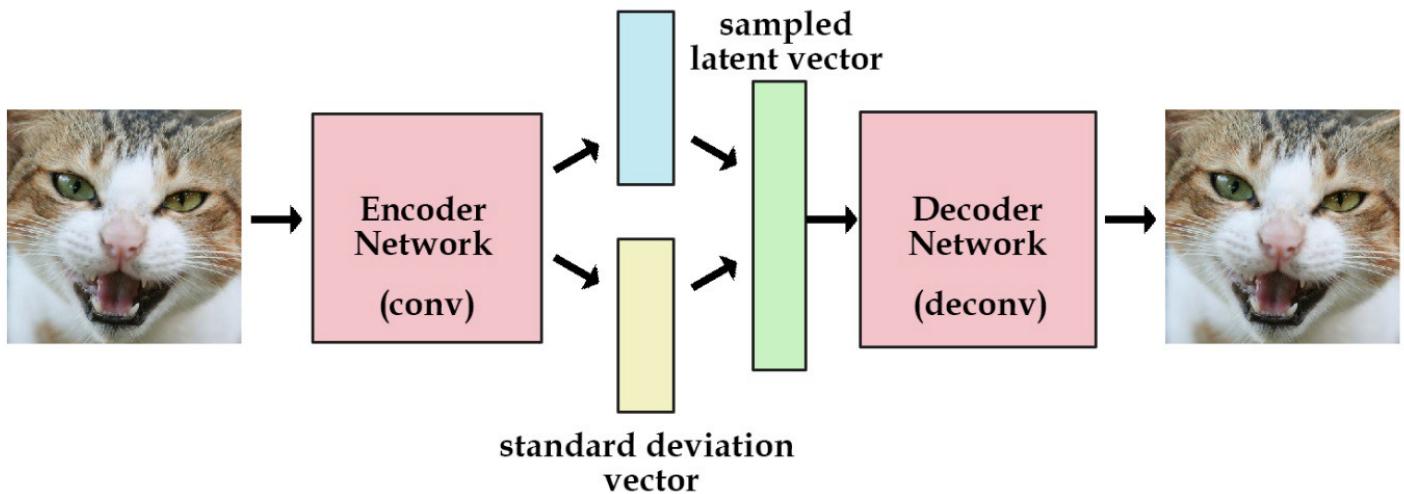

[Get started](#)
[Open in app](#)


Figure 5: Layers in Variational Autoencoder [2]

Recent advancements in VAE as mentioned in [6] which improves the quality of VAE samples by adding two more components. Firstly, a pre-trained classifier as extractor to input data which aligns the reproduced images. Secondly, a discriminator network for additional adversarial loss signals. Other significant improvement in VAE is Optimization of the Latent Dependency Structure by [7]. In this VAE parameters, network parameters are optimized with a single objective. Interference is formed through sampling which produces expectations over latent variable structures and incorporates top-down and bottom-up reasoning over latent variable values.

### Autoencoders Application:

Today data denoising and dimensionality reduction for data visualization are the two major applications of autoencoders. With dimensionality and sparsity constraints, autoencoders can learn data projections which is better than PCA.[11]

Previously Autoencoders are used for dimensionality reduction or feature learning. In recent developments with connection with the latent variable models have brought autoencoders to forefront of the generative modelling. [11]

Autoencoders or its variants such as stacked, sparse or VAE are used for compact representation of data. For example a 256x256 pixel image can be represented by 28x28 pixel. Google is using this type of network to reduce the amount band width you use it on your phone. If you download an image the full resolution of the image is


[Get started](#)
[Open in app](#)

Autoencoders are used in Natural Language Processing, where NLP enclose some of the most difficult problems in computer science. With advancement in deep learning and indeed, autoencoders are been used to overcome some of these problems[9].

**Word Embedding:** Words or phrases from a sentence or context of a word in a document are sorted in relation with other words.

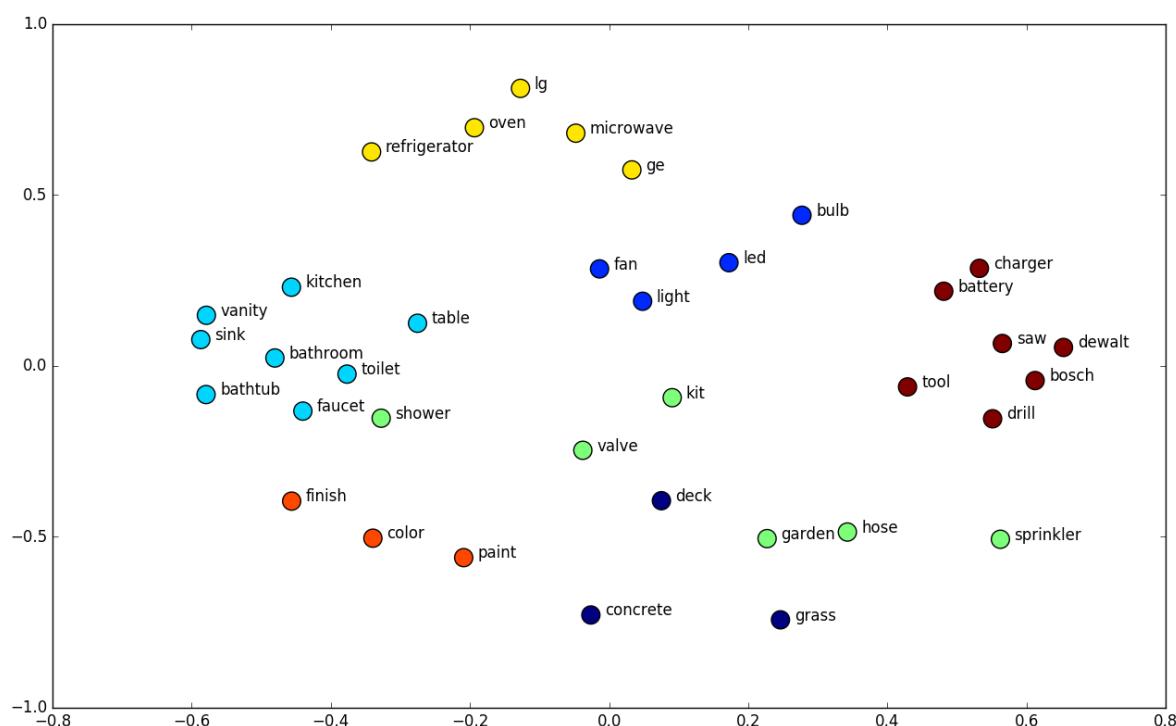


Figure 6: Some example word vectors mapped to two dimensions. Source:

<http://suriyadeepan.github.io/img/seq2seq/we1.png>

**Document Clustering:** classification of documents such as blogs or news or any data into recommended categories. The challenge is to accurately cluster the documents into categories where there actually fit. Hinton used autoencoder to reduce the dimensionality vectors to represent the word probabilities in newswire stories[10]. The Figure below shows the comparisons of Latent Semantic analysis and an autoencoder based on PCA and non linear dimensionality reduction algorithm proposed by Roweis where autoencoder outperformed LSA.[10]

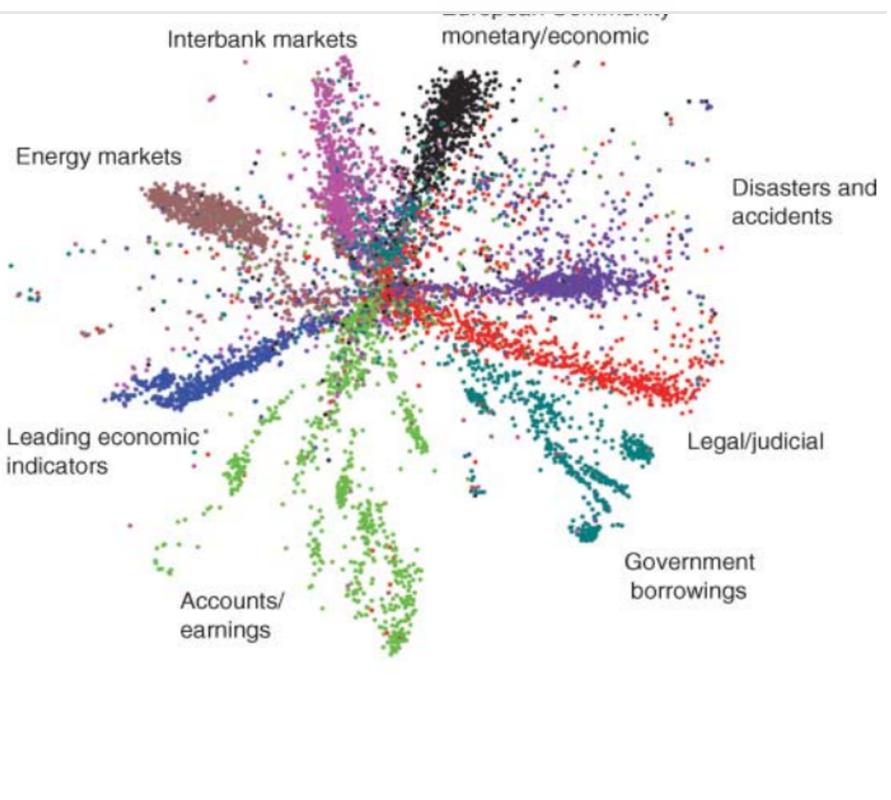
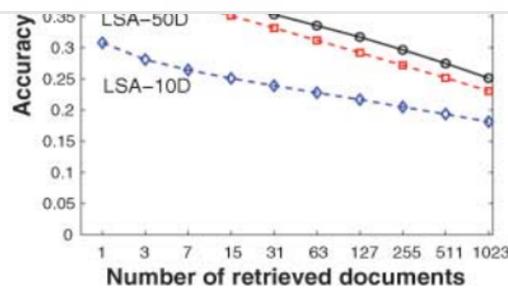

[Get started](#)
[Open in app](#)


Figure 7: Clustering documents using (B) LSA and © an autoencoder. Source: [10].

Machine translation: it has been studied since late 1950s and an incredibly difficult problem to translate text from one human language to another human language. With the use of autoencoders machine translation has taken a huge leap forward to accurately translate text from one language to another.

Autoencoders to extract speech: A deep generative model of spectrograms containing 256 frequency bins and 1,3,9 or 13 frames has been created by [12]. This model has one visible layer and one hidden layer of 500 to 3000 binary latent variables. [12]

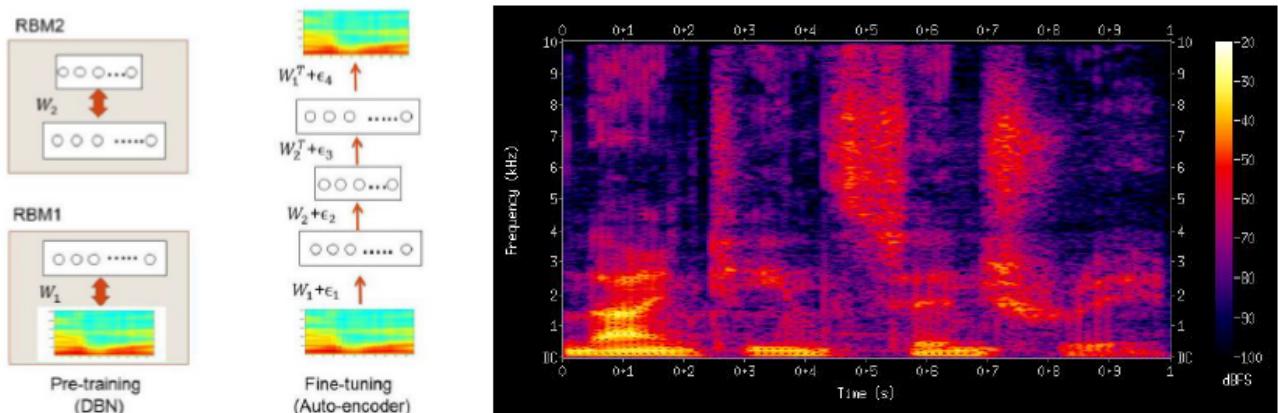


Figure 8: Speech extraction [12].


[Get started](#)
[Open in app](#)

problem for a long time in Automatic Speech Recognition(ASR). This has been implemented in various smart devices such as Amazon Alexa.

Reverberant speech recognition using deep learning in front end and back of a system. This model is built by Mimura, Sakai and Kawahara, 2015 where they adopted a deep autoencoder(DAE) for enhancing the speech at the front end and recognition of speech is performed by DNN-HMM acoustic models at the back end [13].

Denoising of speech using deep autoencoders:

In actual conditions we experience speech signals are contaminated by noise and reverberation. If this speech is used by SR it may experience degradation in speech quality and in turn effect the performance.

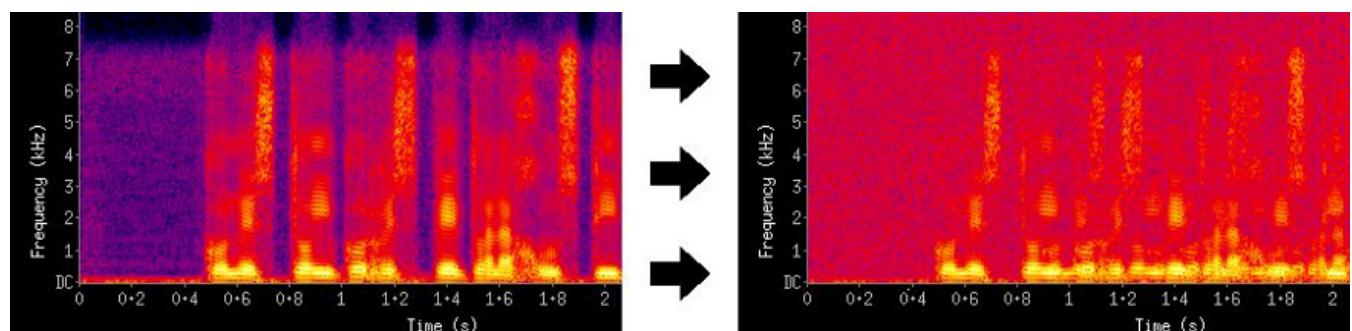
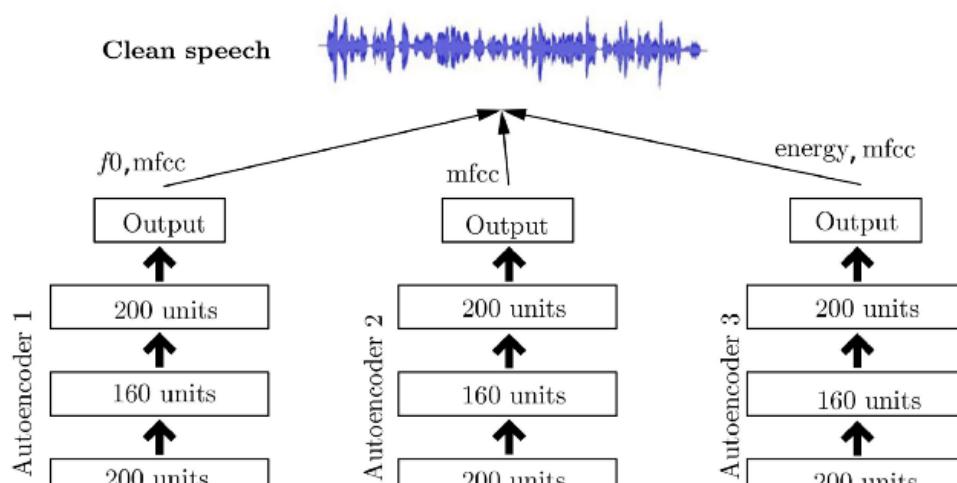


Figure 9: (left) clean speech signal (right) noisy speech signal.

In order to improve the accuracy of the ASR system on noisy utterances, will be trained a collection of LSTM networks, which map features of a noisy utterance to a clean utterance. The figure below shows the model used by (Marvin Coto, John Goddard, Fabiola Martínez) 2016.



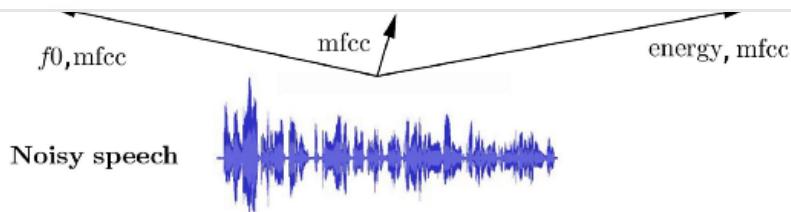

[Get started](#)
[Open in app](#)


Figure 10: source [(Marvin Coto, John Goddard, Fabiola Martínez) 2016]

### Spatio-Temporal AutoEncoder for Video Anomaly Detection:

Anomalous events detection in real-world video scenes is a challenging problem due to the complexity of “anomaly” as well as the cluttered backgrounds, objects and motions in the scenes. In (Zhao, Deng and Shen, 2018) they proposed model called Spatio-Temporal AutoEncoder which utilizes deep neural networks to learn video representation automatically and extracts features from both spatial and temporal dimensions by performing 3-dimensional convolutions. They introduced a weight-decreasing prediction loss for generating future frames, which enhances the motion feature learning in videos. Since most anomaly detection datasets are restricted to appearance anomalies or unnatural motion anomalies. Figure below shows the architecture of the network. An encoder followed by two branches of decoder for reconstructing past frames and predicting the future frames.

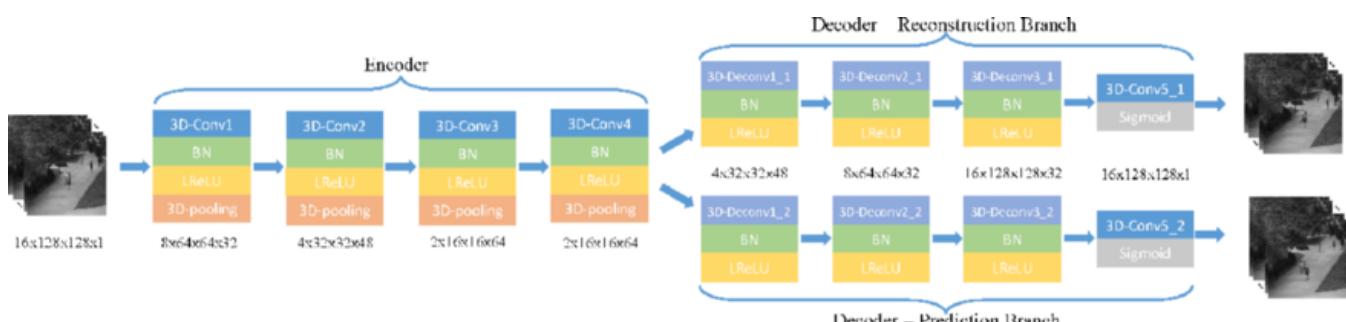


Figure 11: source[18]

**Reconstruction image using Convolutional Autoencoders:** CAE are useful in reconstruction of image from missing parts. For this the model has to be trained with two different images as input and output. The input image can rather be a noisy version or an image with missing parts and with a clean output image. During training process the model learns and fills the gaps in the input and output images. Here is an example below how CAE replace the missing part of the image.

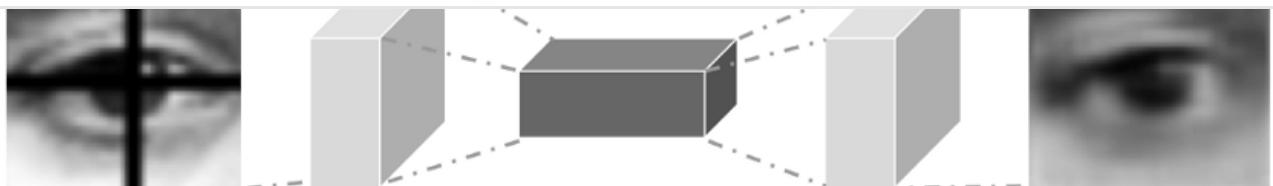

[Get started](#)
[Open in app](#)


Figure 12: Reconstructed image from missing image [14]

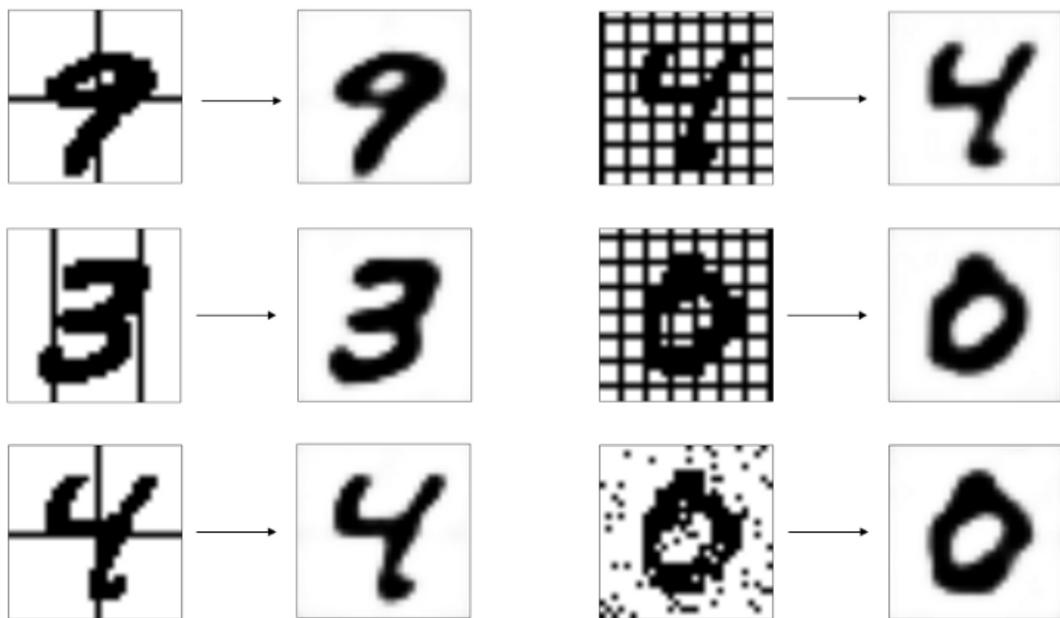


Figure 13: Source [15]

**Paraphrase Detection:** in many languages two phrases may look differently but when it comes to the meaning they both mean exactly same. Deep learning autoencoders allow us to find such phrases accurately.

Many other advanced applications includes full image colorization, generating higher resolution images by using lower resolution as input.

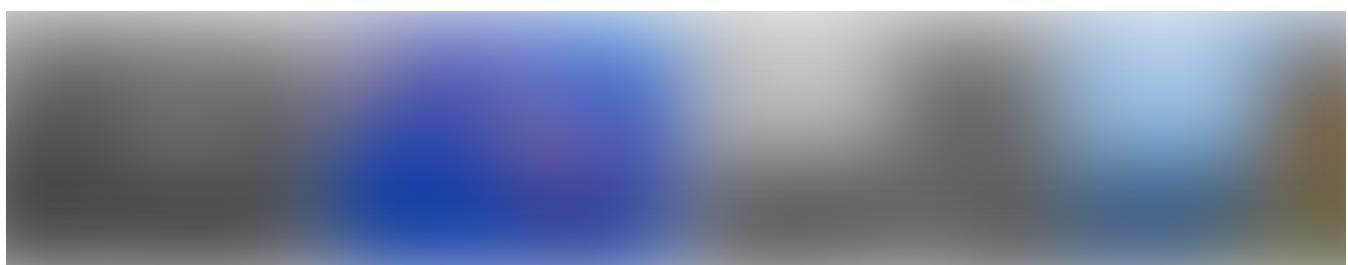


Figure 14: Colorful Image Colorization by Richard Zhang, Phillip Isola, Alexei A. Efros

## Autoencoders advantage over PCA/SVD for dimensionality reduction:

[Get started](#)[Open in app](#)

An autoencoder can learn non-linear transformations, unlike PCA, with a non-linear activation function and multiple layers.

An autoencoder doesn't have to learn dense (affine) layers; it can use convolutional layers to learn too, which could be better for video, image and series data.

It may be more efficient, in terms of model parameters, to learn several layers with an autoencoder rather than learn one huge transformation with PCA.

An autoencoder gives a representation as the output of each layer, and maybe having multiple representations of different dimensions is useful.

An autoencoder could let you make use of pre trained layers from another model, to apply transfer learning to prime the encoder/decoder.

Figure below from the 2006 Science paper by Hinton and Salakhutdinov show a clear difference between Autoencoder vs PCA. It shows dimensionality reduction of the MNIST dataset ( $28 \times 28 \times 28$  black and white images of single digits) from the original 784 dimensions to two.

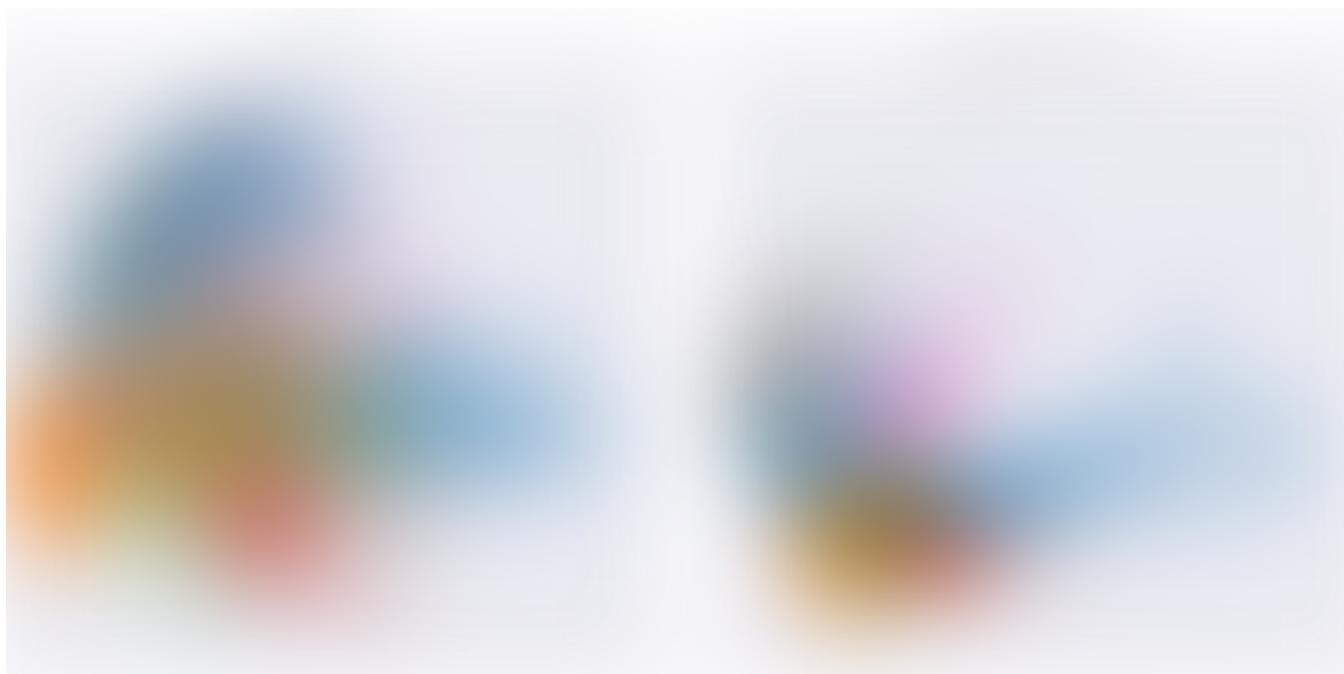


Figure 15: from the 2006 Science paper by Hinton and Salakhutdinov

## Autoencoders vs GAN

[Get started](#)[Open in app](#)

a smaller representation, and then transforms it back into a tensor with the same shape as its input over several neural net layers. Autoencoders are trained to reproduce the input, so it's kind of like learning a compression algorithm for that specific dataset.

A GAN looks kind of like an inside out autoencoder — instead of compressing high dimensional data, it has low dimensional vectors as the inputs, high dimensional data in the middle.

Another difference: while they both fall under the umbrella of unsupervised learning, they are different approaches to the problem. A GAN is a generative model — it's supposed to learn to generate realistic new samples of a dataset. Variational autoencoders are generative models, but normal “vanilla” autoencoders just reconstruct their inputs and can't generate realistic new samples. [16]

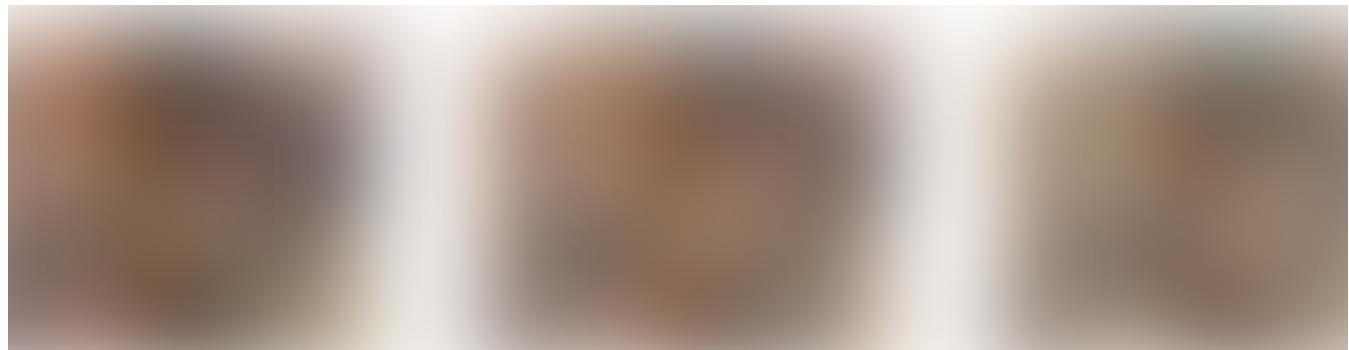


Figure 16: Image reconstructed by VAE and VAE-GAN compared to their original input images [17]

Autoencoders are an extremely exciting new approach to unsupervised learning, and for virtually every major kind of machine learning task, they have already surpassed the decades of progress made by researchers handpicking features.

## References

- [1] et al N. A dynamic programming approach to missing data estimation using neural networks; Available from: [https://www.researchgate.net/figure/222834127\\_fig1](https://www.researchgate.net/figure/222834127_fig1).
- [2] Kevin frans blog. (2018). *Variational Autoencoders Explained*. [online] Available at: <http://kvfrans.com/variational-autoencoders-explained/> [Accessed 28 Nov. 2018].
- [3] Packtpub.com. (2018). {{metadataController.pageTitle}}. [online] Available at: [https://www.packtpub.com/mapt/book/big\\_data\\_and\\_business\\_intelligence/978178](https://www.packtpub.com/mapt/book/big_data_and_business_intelligence/978178)

[Get started](#)[Open in app](#)

- [4] Liu, G., Dao, T. and Tran, D. (2018). A Stacked Autoencoder-Based Deep Neural Network for Achieving Gearbox Fault Diagnosis. [online] Hindawi. Available at: <https://www.hindawi.com/journals/mpe/2018/5105709/> [Accessed 23 Nov. 2018].
- [5] V., K. (2018). Improving the Classification accuracy of Noisy Dataset by Effective Data Preprocessing. *International Journal of Computer Applications*, 180(36), pp.37–46.
- [6] Hou, X. and Qiu, G. (2018). IMPROVING VARIATIONAL AUTOENCODER WITH DEEP FEATURE CONSISTENT AND GENERATIVE ADVERSARIAL TRAINING. *Workshop track — ICLR*.
- [7] Variational Autoencoders with Jointly Optimized Latent Dependency Structure. (2018). *ICLR 2019 Conference Blind Submission*.
- [8] Wilkinson, E. (2018). *Deep Learning: Sparse Autoencoders*. [online] Eric Wilkinson. Available at: <http://www.erichwilkinson.com/blog/2014/11/19/deep-learning-sparse-autoencoders> [Accessed 29 Nov. 2018].
- [9] Doc.ic.ac.uk. (2018). *Autoencoders: Applications in Natural Language Processing*. [online] Available at: <https://www.doc.ic.ac.uk/~js4416/163/website/nlp/> [Accessed 29 Nov. 2018].
- [10] Hinton G, Salakhutdinov R. Reducing the Dimensionality of Data with Neural Networks. *Science*. 2006;313(5786):504–507. Available from: <https://www.cs.toronto.edu/~hinton/science.pdf>.
- [11] Autoencoders: Bits and bytes, <https://medium.com/towards-data-science/autoencoders-bits-and-bytes-of-deep-learning-eaba376f23ad>
- [12] Binary Coding of Speech Spectrograms Using a Deep Auto-encoder, L. Deng, et al.
- [13] Mimura, M., Sakai, S. and Kawahara, T. (2015). Reverberant speech recognition combining deep neural networks and deep autoencoders augmented with a phone-class feature. *EURASIP Journal on Advances in Signal Processing*, 2015(1).
- [14] Towards Data Science. (2018). *Autoencoders — Introduction and Implementation in TF..* [online] Available at: <https://towardsdatascience.com/autoencoders->

[Get started](#)[Open in app](#)

[15] Towards Data Science. (2018). Autoencoder Zoo — Image correction with TensorFlow — Towards Data Science. [online] Available at:

<https://towardsdatascience.com/autoencoder-zoo-669d6490895f> [Accessed 27 Nov. 2018].

[16] Anon, (2018). [online] Available at: <https://www.quora.com/What-is-the-difference-between-Generative-Adversarial-Networks-and-Autoencoders> [Accessed 30 Nov. 2018].

[17] Towards Data Science. (2018). What The Heck Are VAE-GANs? — Towards Data Science. [online] Available at: <https://towardsdatascience.com/what-the-heck-are-vae-gans-17b86023588a> [Accessed 30 Nov. 2018].

[18] Zhao, Y., Deng, B. and Shen, C. (2018). Spatio-Temporal AutoEncoder for Video Anomaly Detection. *MM '17 Proceedings of the 25th ACM international conference on Multimedia*, pp.1933–1941.

[Machine Learning](#)[Autoencoder](#)[About](#) [Write](#) [Help](#) [Legal](#)[Get the Medium app](#)