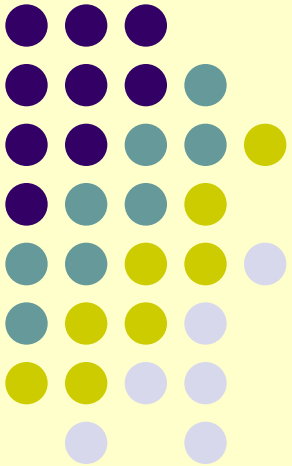
یادگیری عمیق
Deep Learning



استاد درس : محمد امین‌زاده

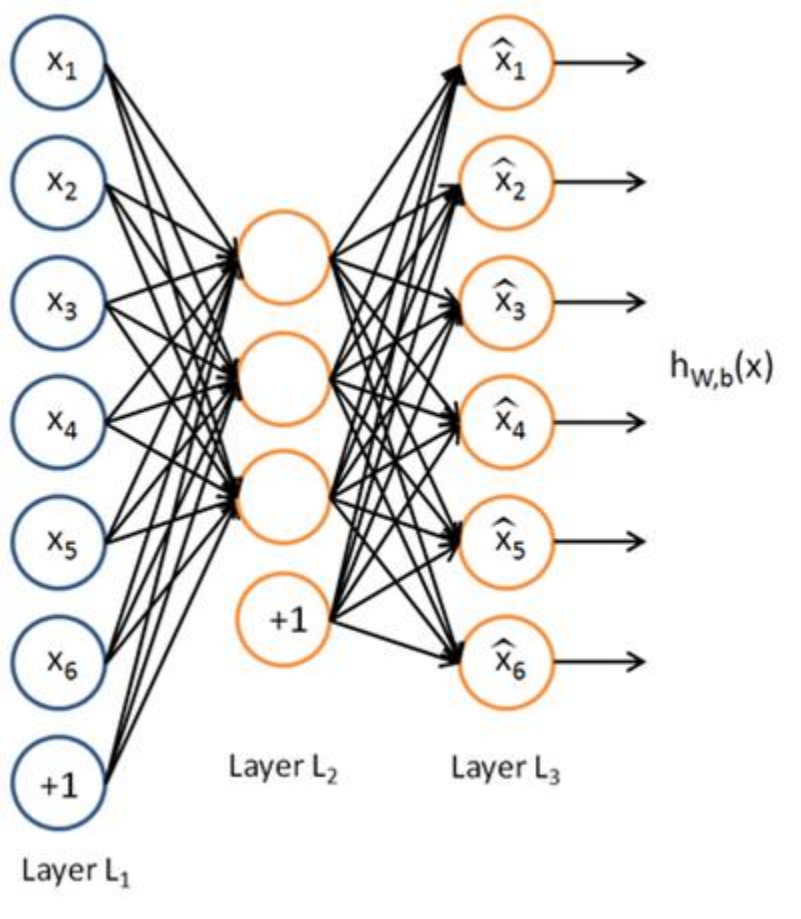
فصل چهارم
معرفی اتوانکدر

اتوانکدر

- اتوانکدرها شبکه های ساده‌ی یادگیری هستند که با هدف تبدیل ورودی به خروجی بدون کمترین تغییر پیاده‌سازی می‌شوند.
- اتوانکدرها از روش یادگیری بدون نظارت بهره می‌برند.
- اتوانکدر یک شبکه‌ی عصبی است که مجموعه‌ای از داده‌های بدون برچسب را دریافت می‌کند و با رمزکردن آن‌ها سعی در بازنمایی مجدد ورودی‌ها در خروجی می‌کند به صورتی که کمترین اختلاف ممکن را با مقدار ورودی داشته باشند.
- ساختار اتوانکدر به دو بخش رمزگذاری (encoding) و رمزگشایی (decoding)تقسیم می‌شود. در بخش رمزگذاری داده‌های ورودی به فضای ویژگی‌ها نگاشت می‌شوند و در بخش رمزگشایی از فضای ویژگی مجددا به حالت ابتدایی خود تبدیل می‌شوند. در واقع بخش اصلی یک اتوانکدر لایه‌ی پنهان میانی است که به عنوان ویژگی استخراج شده برای دسته‌بندی استفاده می‌شود.

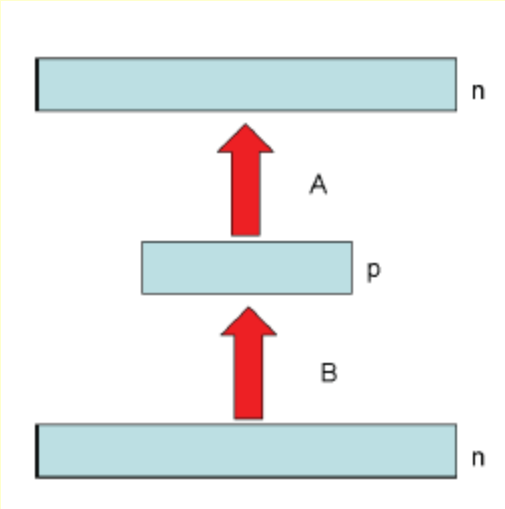
ساختار اتوانکدر ساده

- خروجی بازسازی ورودی است.
- لایه میانی لایه ویژگی‌ها است.
- رفتن از فضای ورودی به فضای ویژگی انکد نام دارد.
- رفتن از فضای ویژگی به فضای خروجی دیکد نام دارد.

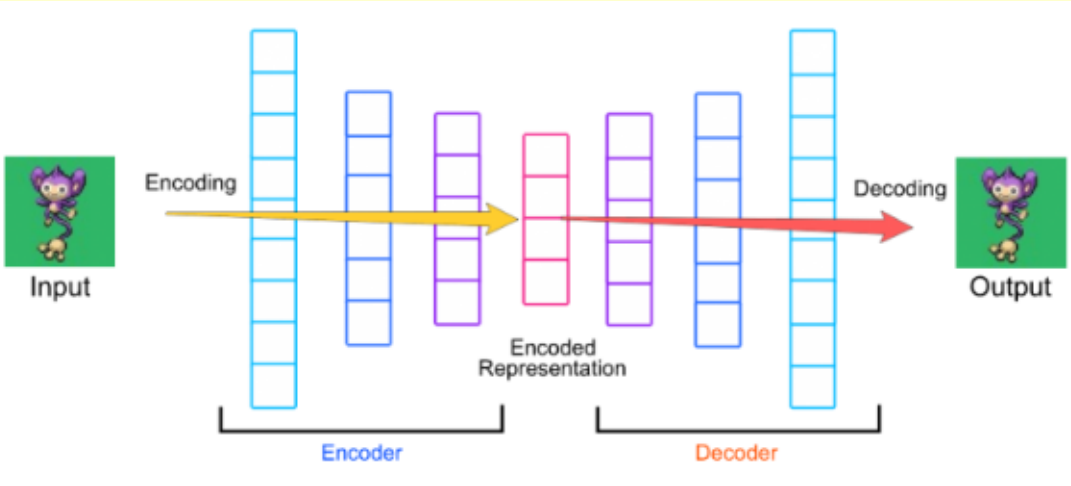


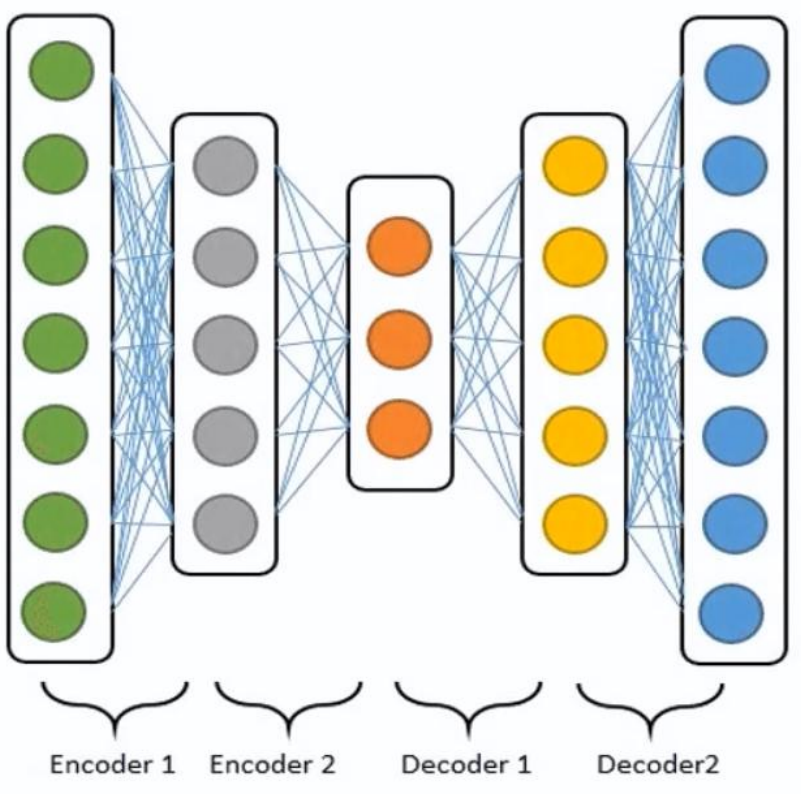
ساختار پشته‌ای

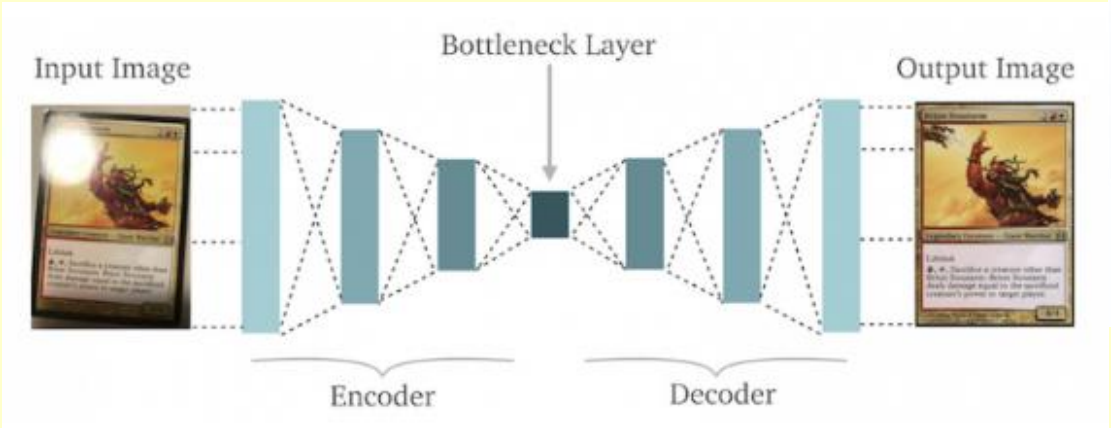
- لایه‌ی ورودی پایین‌ترین لایه‌ی پشته و آخرین لایه‌ی پشته نقش خروجی را ایفا می‌کند.
- n بیانگر مجموعه داده های ورودی و در نتیجه خروجی است. p نیز به عنوان مجموعه ویژگی ها شناخته می‌شود



Deep Auto-Encoder ساختار اتوانکدر عمیق

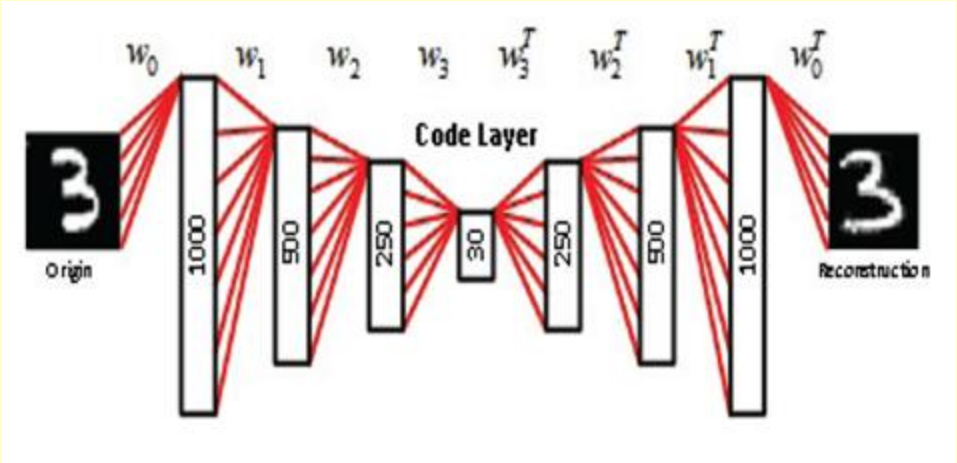






اتوانکدر عمیق

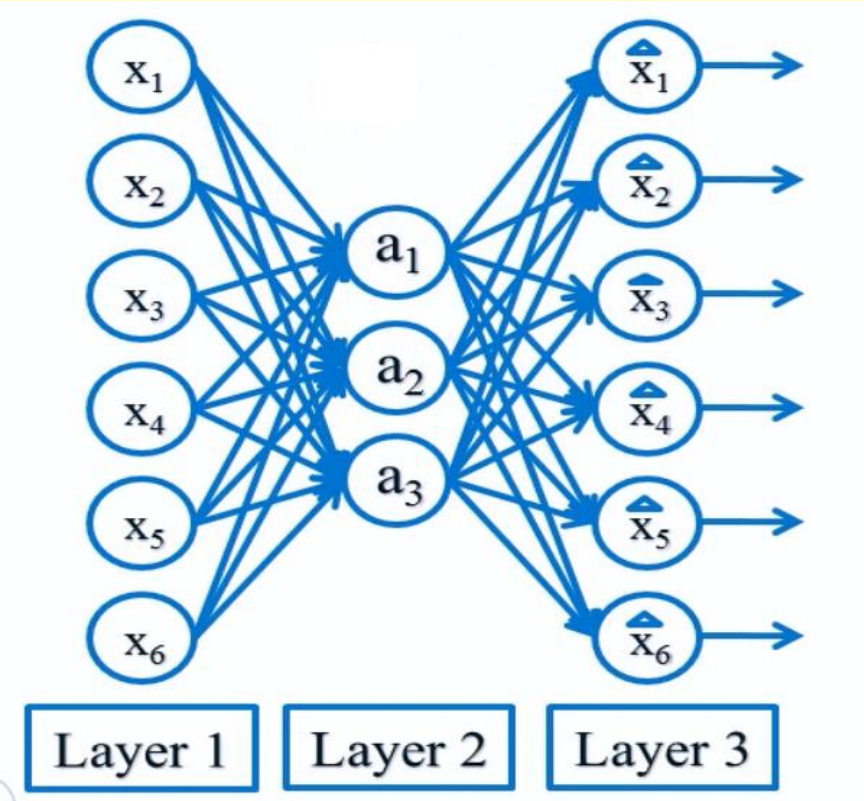
- اگر شبکه‌ی عصبی سازنده‌ی یک اتوانکدر، شبکه‌ای عمیق باشد آن را اتوانکدر عمیق گوییم. در این معماری تعداد لایه‌های پنهان شبکه بیش از یک می‌باشد.
- در این ساختار به عنوان نمونه می‌توان از هفت لایه‌ی میانی برای شبکه استفاده کرد. لایه‌ی پنهان اول از ۱۰۰۰ نورون، لایه‌ی پنهان دوم از ۵۰۰ نورون و به همین ترتیب در لایه‌ی چهارم پنهان از ۳۰ نورون استفاده شده‌است. سپس تا لایه‌ی آخر بخش رمزگشایی نیز به نمایش درآمده‌است.



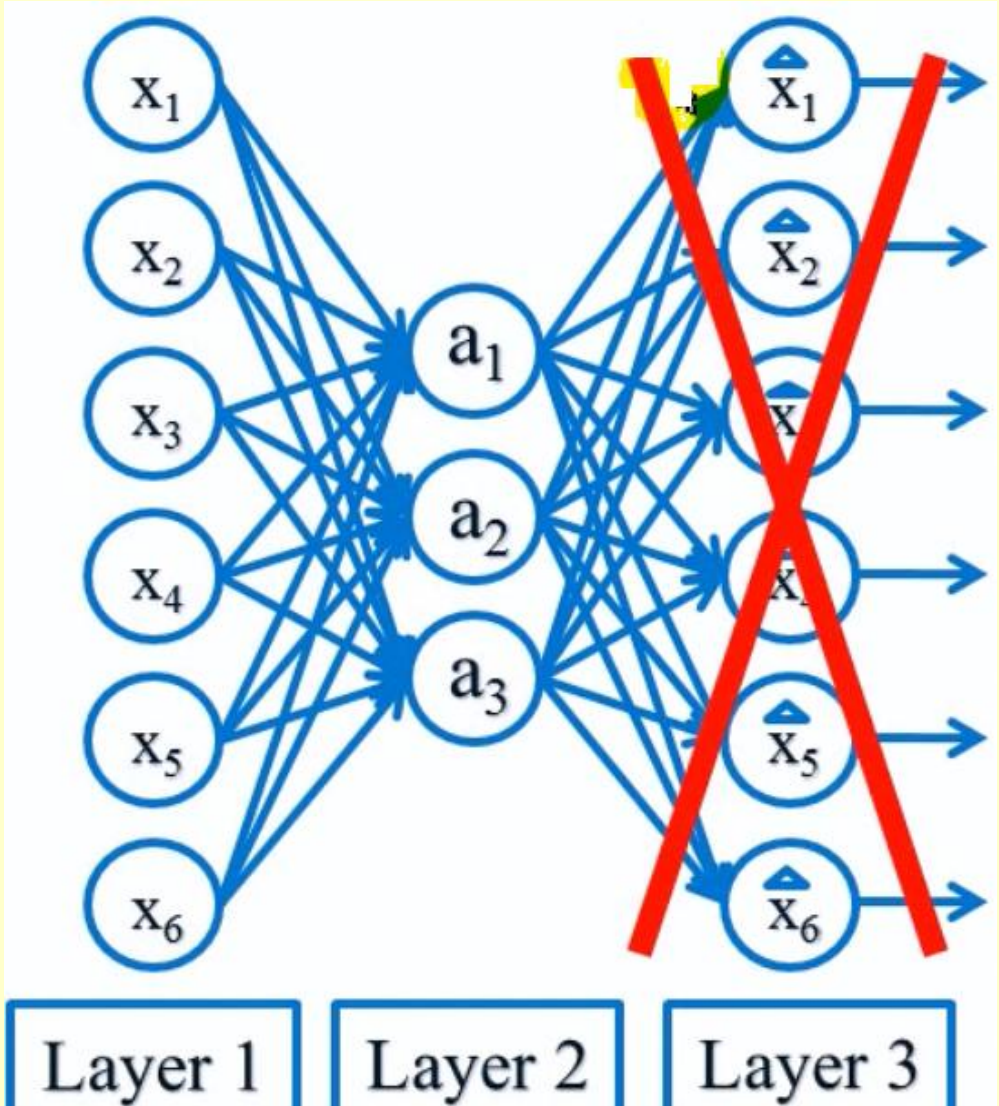
- برای مجموعه‌ی داده‌ی تشخیص صورت اگر از یک اتوانکدر عمیق استفاده کنیم نتیجه‌ی نهایی به صورت تصاویر زیر خواهد بود.

آموزش خودرمزنگار

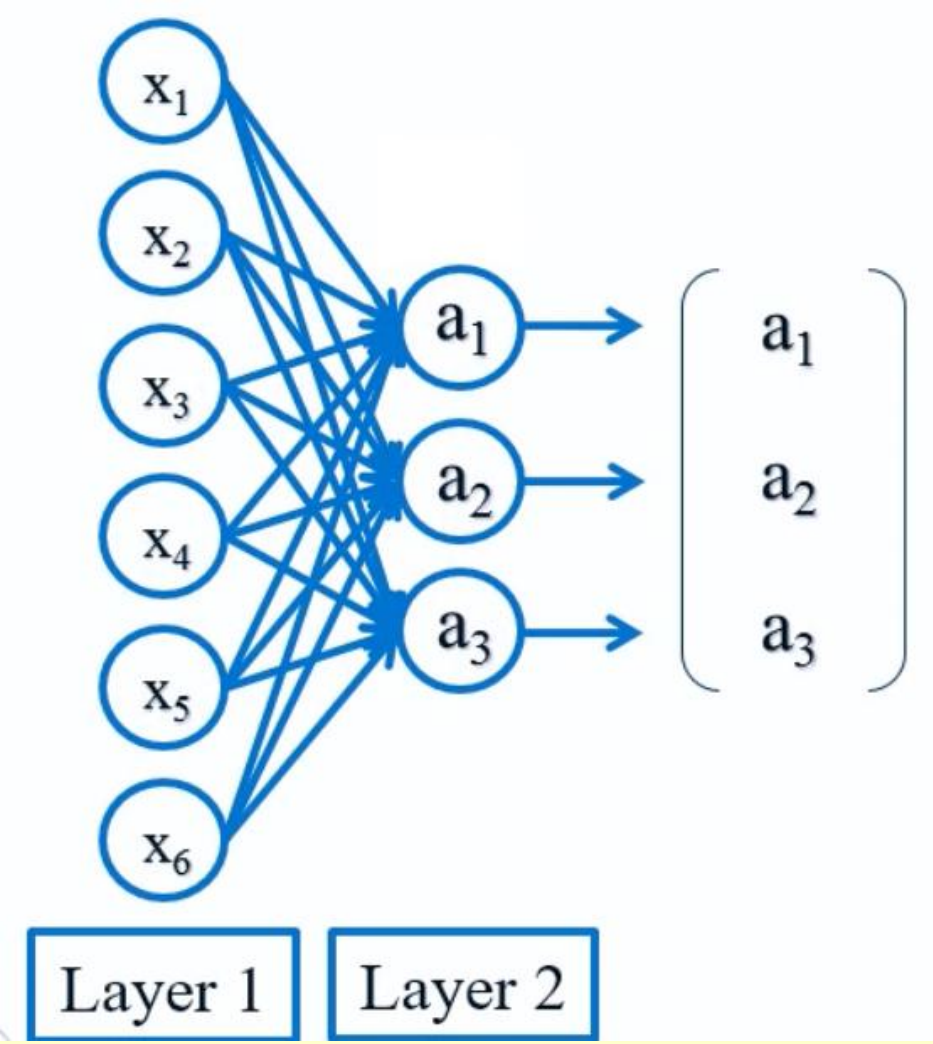
- آموزش لایه‌به لایه، نوع جدید از آموزش که در سال ۲۰۰۶ در این شبکه بکار گرفته شد.
- هر لایه از اتوانکدر در ساختار یک اتوانکدر ساده آموزش می‌بینند.
- لایه اول



پس از آموزش در این مرحله، بخش دیکدر کنار گذاشته می‌شود.

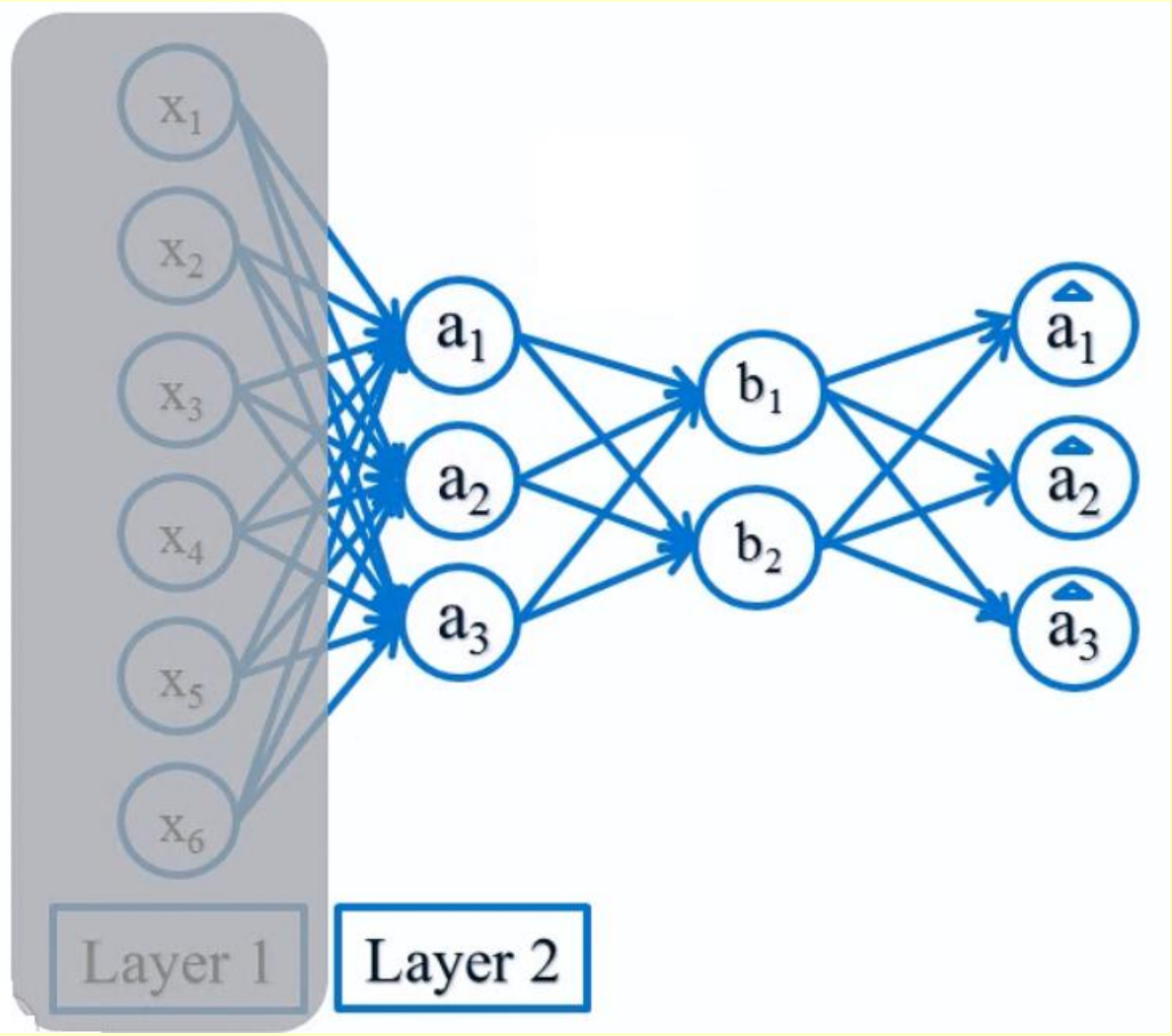


الان ویژگی‌های سطح یک در دسترس هستند

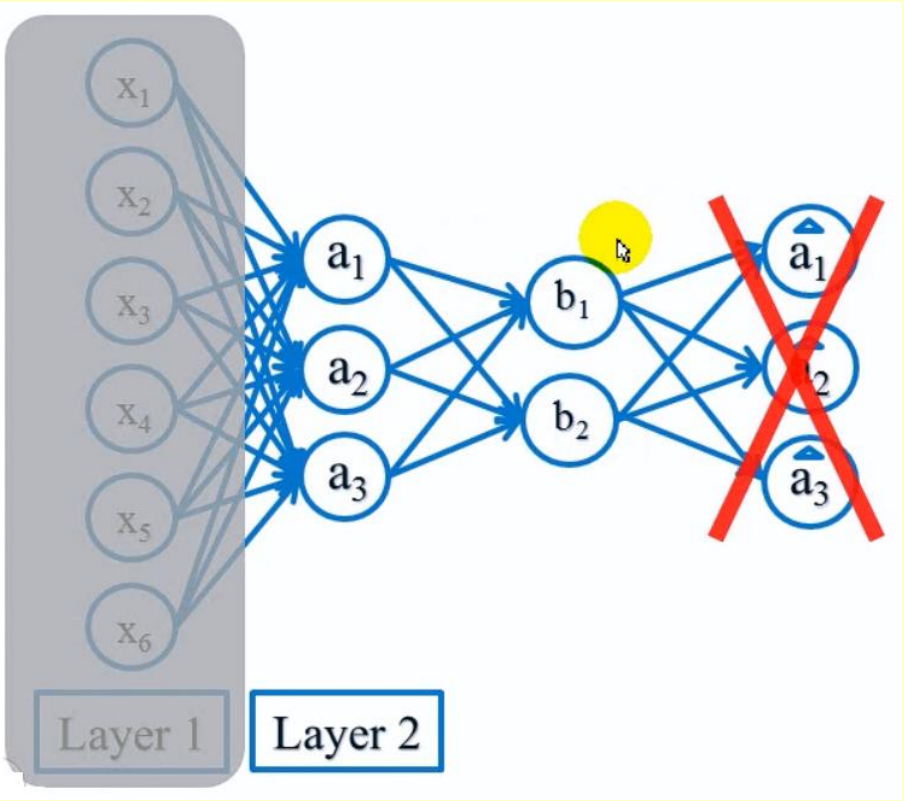


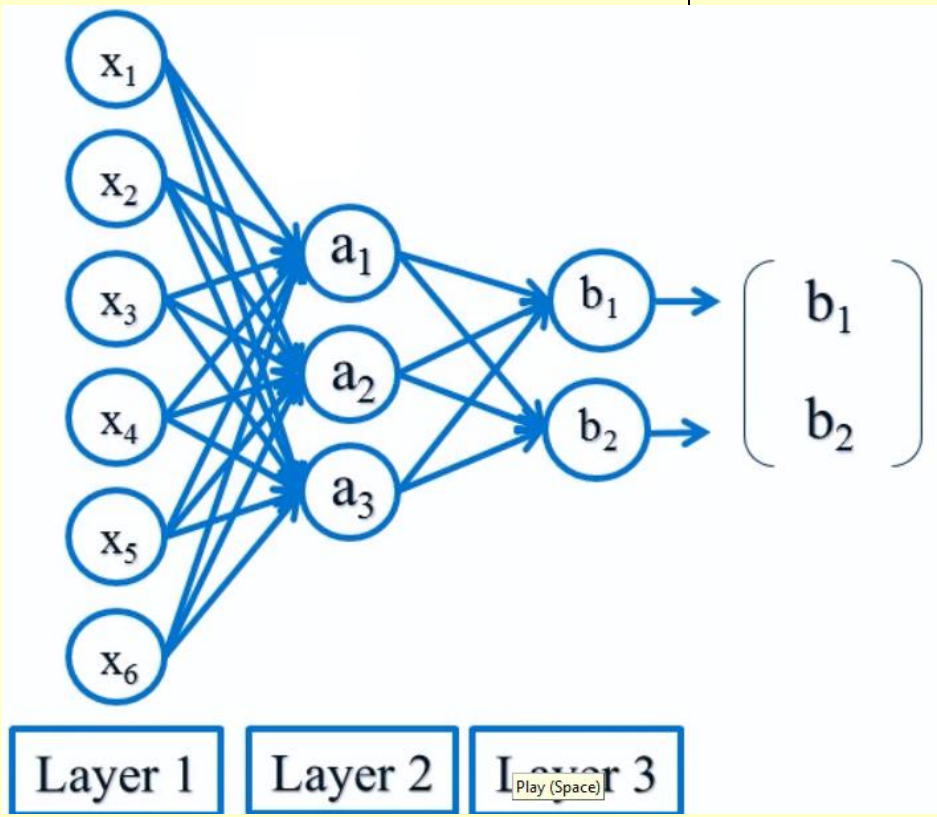
آموزش لایه دوم

- ورودی ویژگی‌های a
- لایه دوم، ویژگی‌های جدید
- خروجی بازسازی a



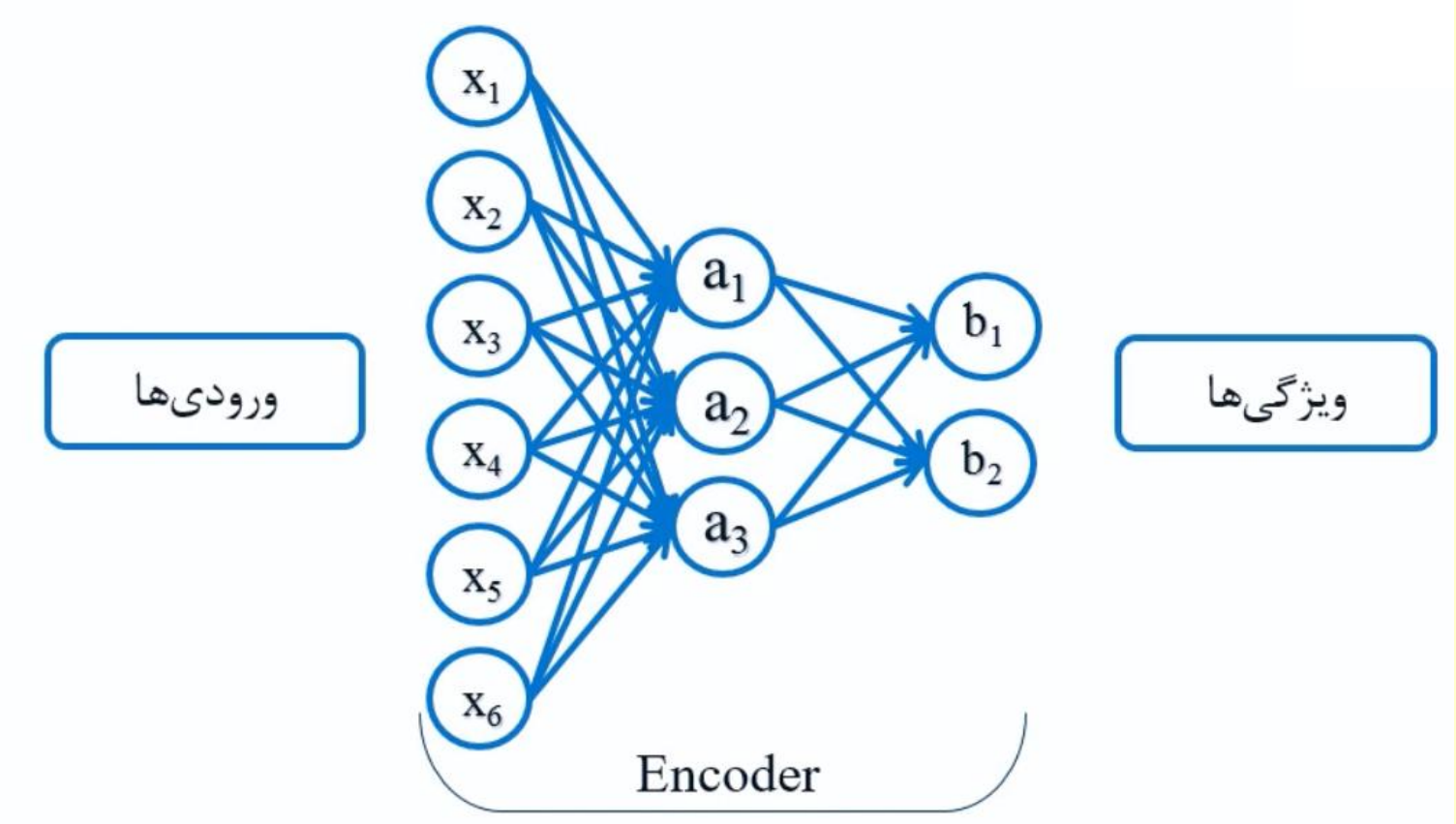
ادامه روند آموزش مثل مرحله قبل





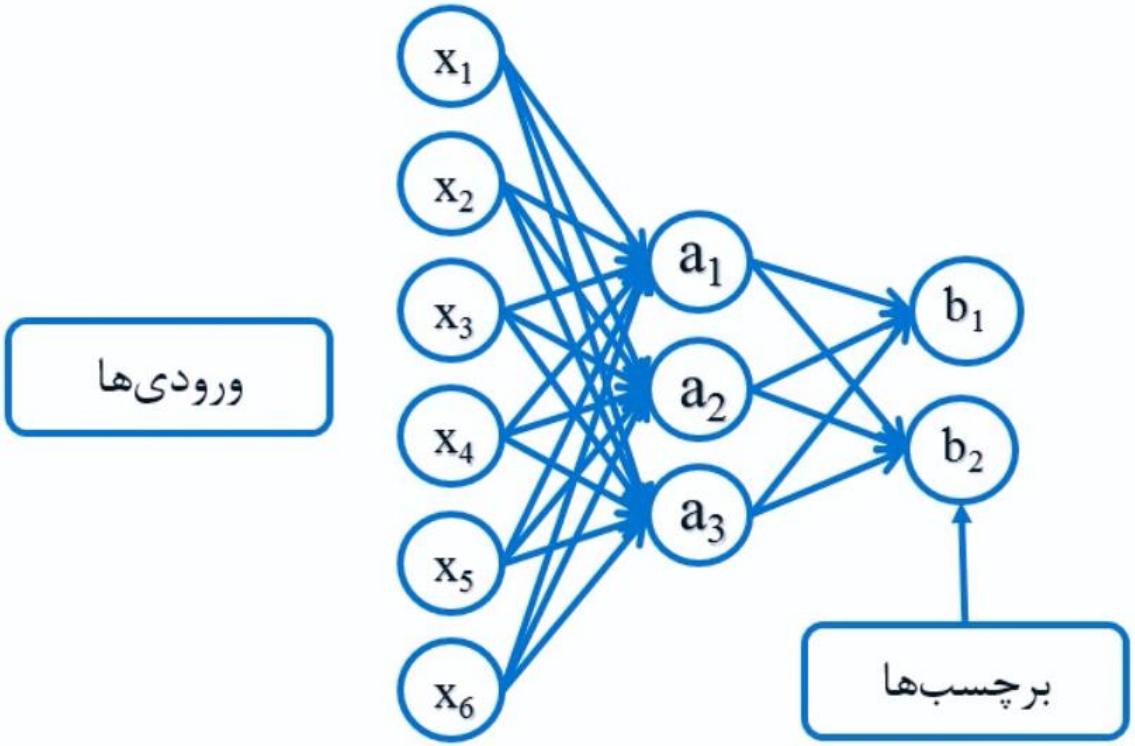
کاربردهای این شبکه

- استخراج ویژگی بدون برچسب‌گذاری
- کاهش بعد



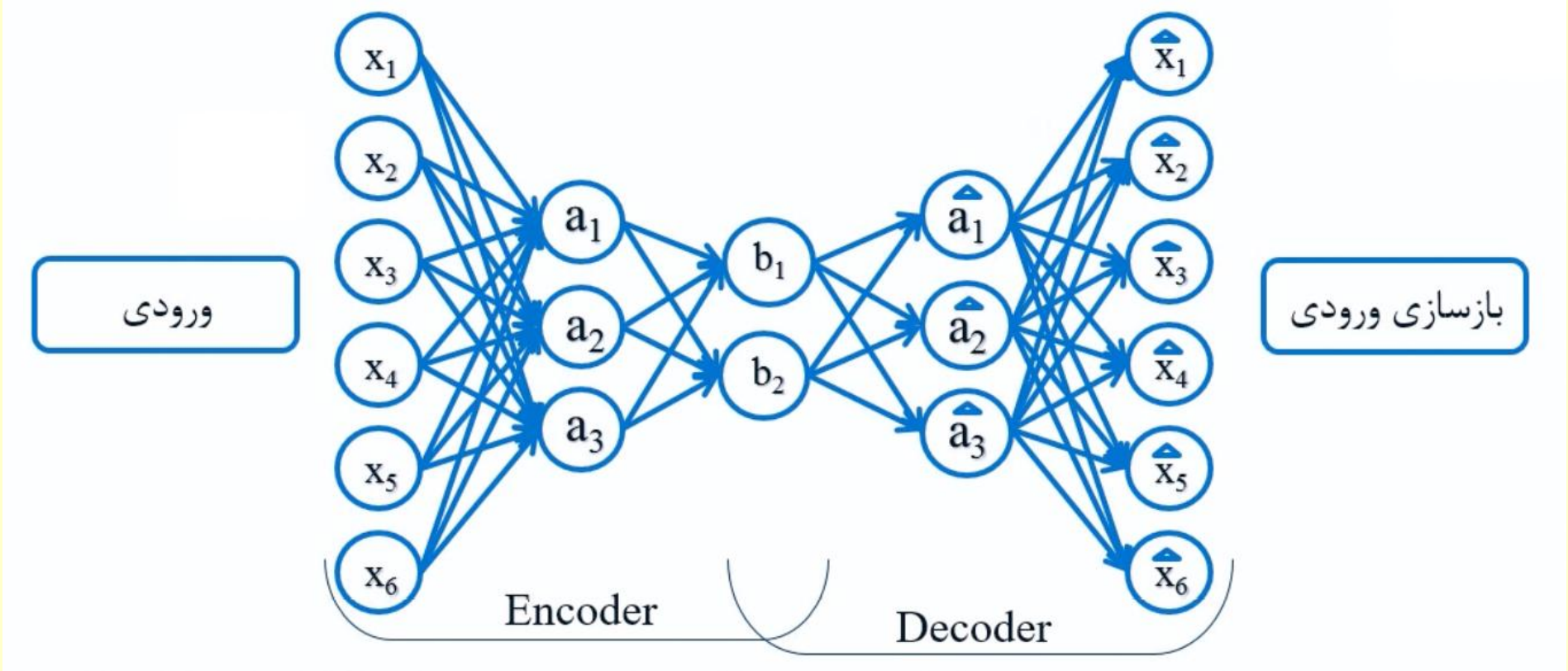
کاربردهای این شبکه

- کلاسه‌بندی
- اصلاح آموزش با داده‌های برچسب‌دار
- برای دسته‌بندی یک لایه تابع فعال‌سازی (Softmax) به قسمت encoder اضافه می‌کنید و کل شبکه رو با داده های برچسب‌دار finetune می‌کنید.



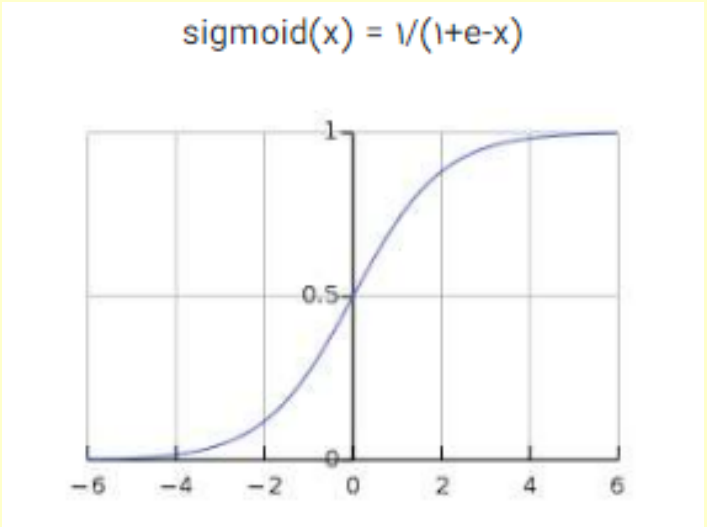
کاربرد دیگر: بازسازی

- اصلاح داده‌ها

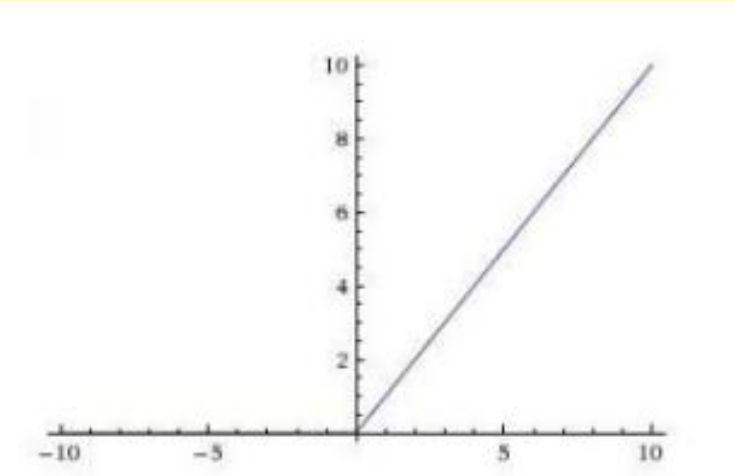


توابع فعال‌ساز مهم در ساختار شبکه‌های عصبی عمیق

- تابع سیگموئید



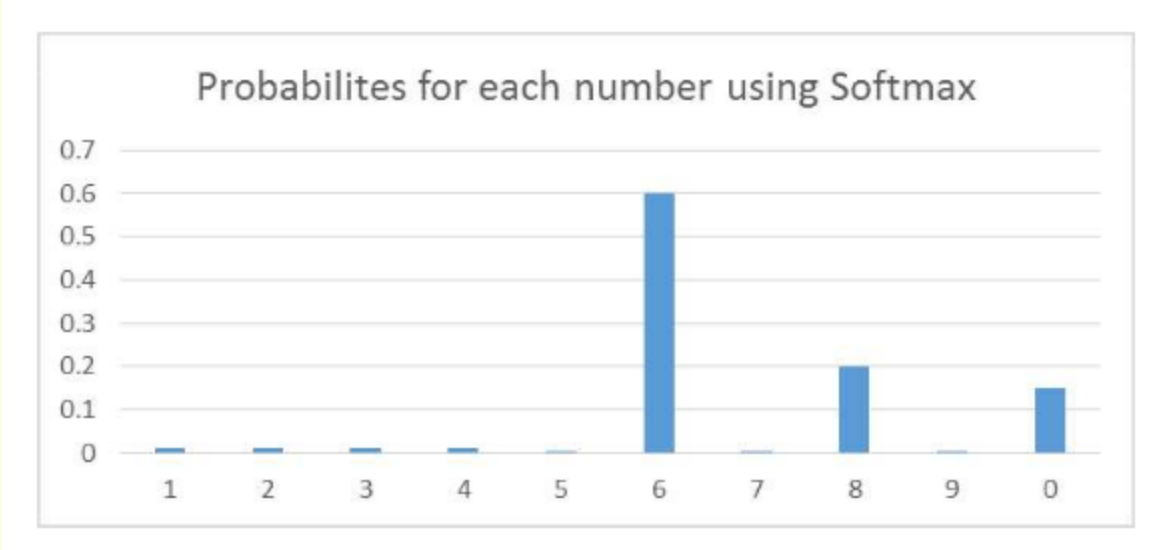
- تابع RelU : شبکه‌های جدید ترجیح می‌دهند به جای sigmoid، از توابع فعال‌سازی ReLU برای لایه‌های مخفی استفاده کنند.



توابع فعال‌ساز مهم در ساختار شبکه‌های عصبی عمیق

- تابع **Softmax** :این توابع معمولا در لایه‌ی خروجی و برای حل مشکلات مربوط به دسته‌بندی استفاده می‌شوند.

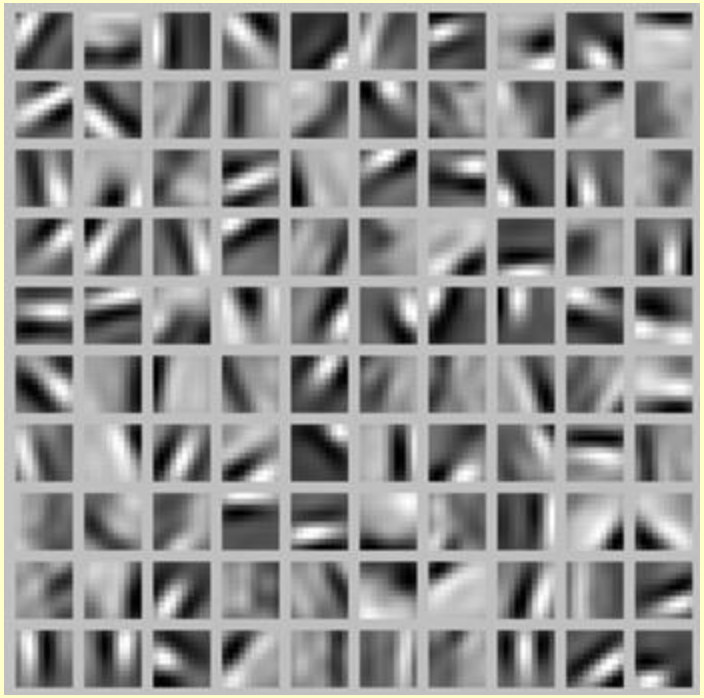
- خروجی‌های این تابع به گونه‌ای نرمال‌سازی شده‌اند که مجموع آن‌ها عدد یک باشد.
- فرض کنید می‌خواهید یک عدد «۶» را که کمی نیز شبیه «۸» است، شناسایی کنید. این تابع همانند تصویر زیر به هر عدد، مقداری می‌دهد. در نهایت، به سادگی می‌توانید ببینیم که بیشترین احتمال به «۶» داده شده‌است و بالاترین احتمال بعدی «۸» است و الی آخر.



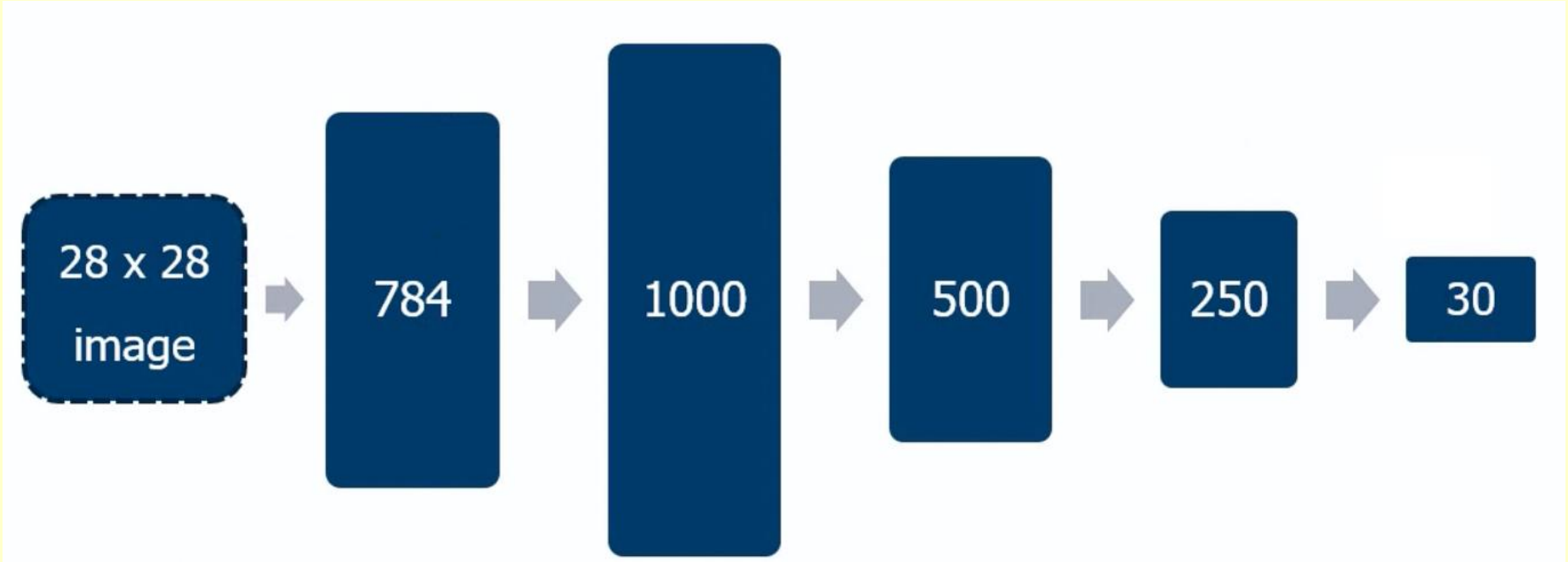
مثال: یادگیری مجموعه داده MNIST

- **MNIST : این مجموعه از اعداد انگلیسی تشکیل شده‌است که شامل ۶۰۰۰ داده برای تمرین و ۱۰۰۰۰ برای مجموعه آزمایشی می‌باشد.**
- **تصاویر این مجموعه از ماتریس‌های ۲۸ در ۲۸ تشکیل شده‌اند.**
- **برای رمزگذاری این مجموعه از یک شبکه‌عصبی با یک لایه‌ی مخفی استفاده شده است.**
- **لایه‌ی مخفی از ۱۰۰ نورون ساخته شده‌است.**
- **هر سلول تصویر، نشان می‌دهد هریک از نورون‌های لایه‌ی پنهان به چه ویژگی‌ای حساسند و با دیدن کدام ویژگی فعال می‌شود.**

- تصویر زیر ویژگی‌های استخراج شده از مجموعه داده‌ی MNIST را نشان می‌دهد.
- همان‌طور که از تصویر مشخص است اتوانکدر در اولین لایه‌ی خود همانند یک تشخیص‌دهنده‌ی لبه رفتار می‌کند و نسبت به لبه‌های موجود در تصویر حساسیت نشان می‌دهد.
-



- تغییر انکدر : برای رمزگذاری این مجموعه از یک شبکه‌عصبی ۴ لایه استفاده می‌کنیم.



مقایسه با PCA

- Input Data →
- 30D AE →
- 30D PCA →

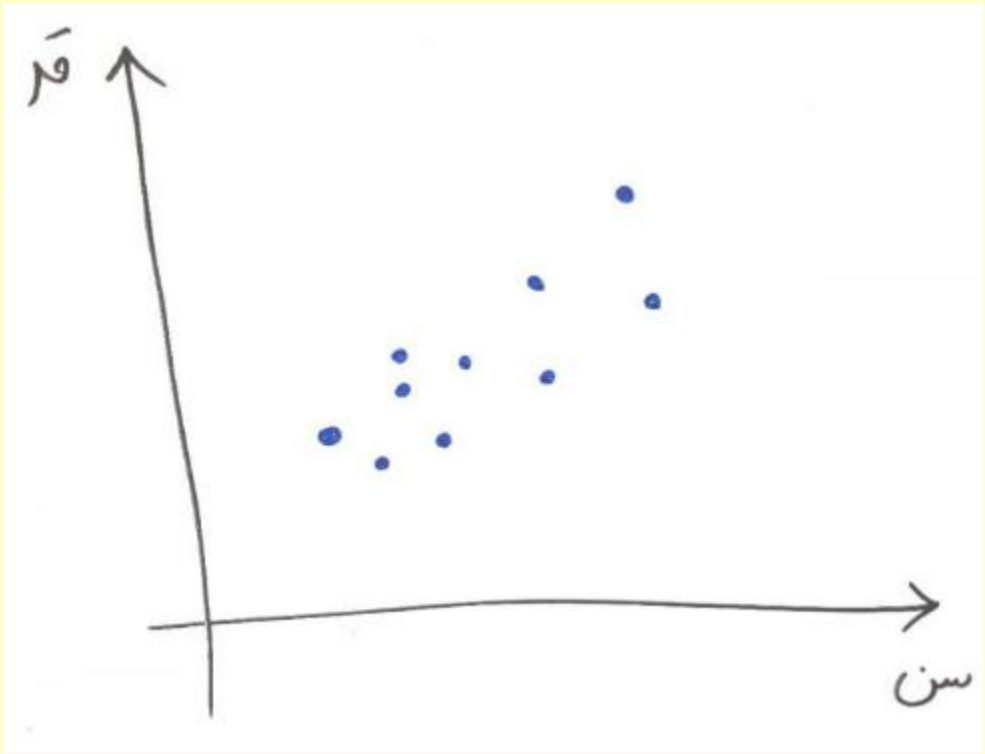
آنالیز مولفه اصلی (Principal Component Analysis)

- کاهش ویژگی و کاهش بعد.
- این روش، می‌تواند مولفه‌های اصلی را شناسایی کند و به ما کمک می‌کند تا به جای اینکه تمامیِ ویژگی‌ها را مورد بررسی قرار دهیم، یک سری ویژگی‌های با ارزشِ بیشتر را تحلیل کنیم.
- تولید ویژگی‌های جدید به کمک ترکیب خطی از ویژگی‌های ورودی
- با مثالی این مبحث را باز می‌کنیم.
- فرض کنید یک فروشگاه می‌خواهد ببیند که رفتار مشتریانش در خریدِ یک محصول خاص (مثلا یک کفش خاص) چطور بوده است. این فروشگاه، اطلاعات زیادی اَز هر فرد دارد. برای مثال این فروشگاه، از هر مشتری ویژگی‌های زیر را جمع‌آوری کرده است:
- سن، قد، جنسیت، محل تولد شخص (غرب، شمال، شرق یا جنوب ایران)، میانگین تعداد افراد خانواده، میانگین درآمد، اتومبیل شخصی دارد یا خیر و در نهایت اینکه این شخص بعد از بازدید کفش خریده است یا خیر. ۷ ویژگیِ اول ابعاد ورودی و ویژگیِ آخر مقدار خروجی است.

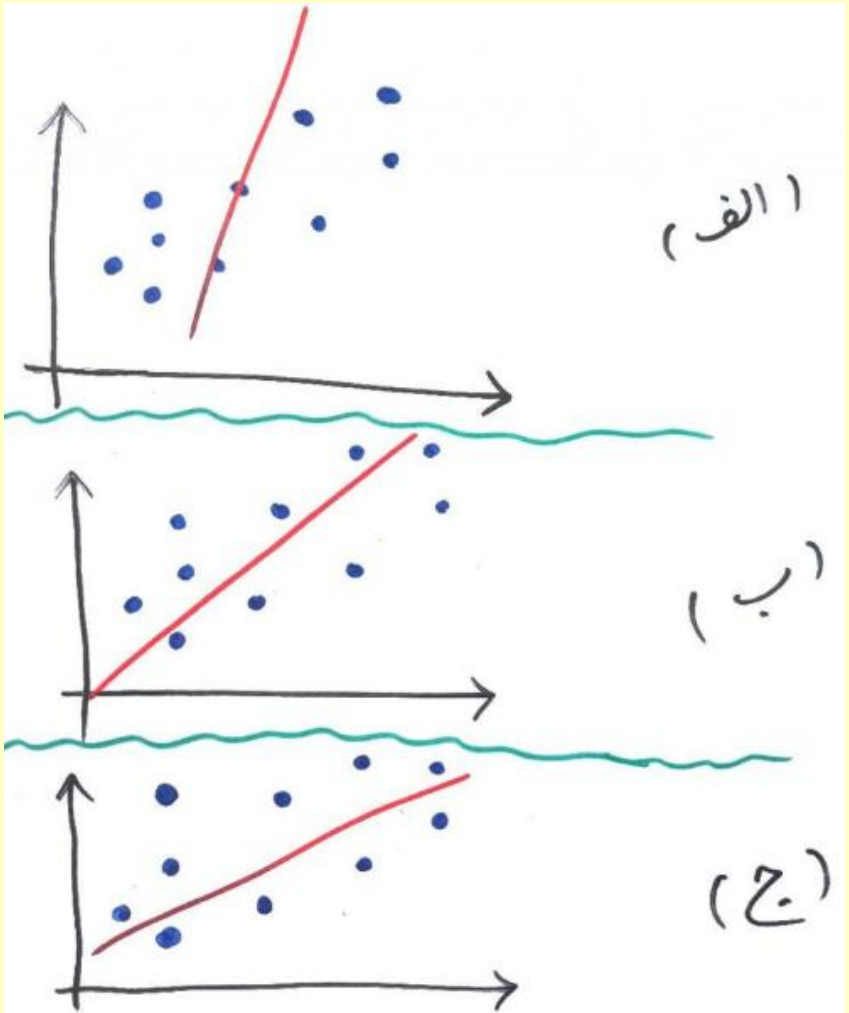
ادامه

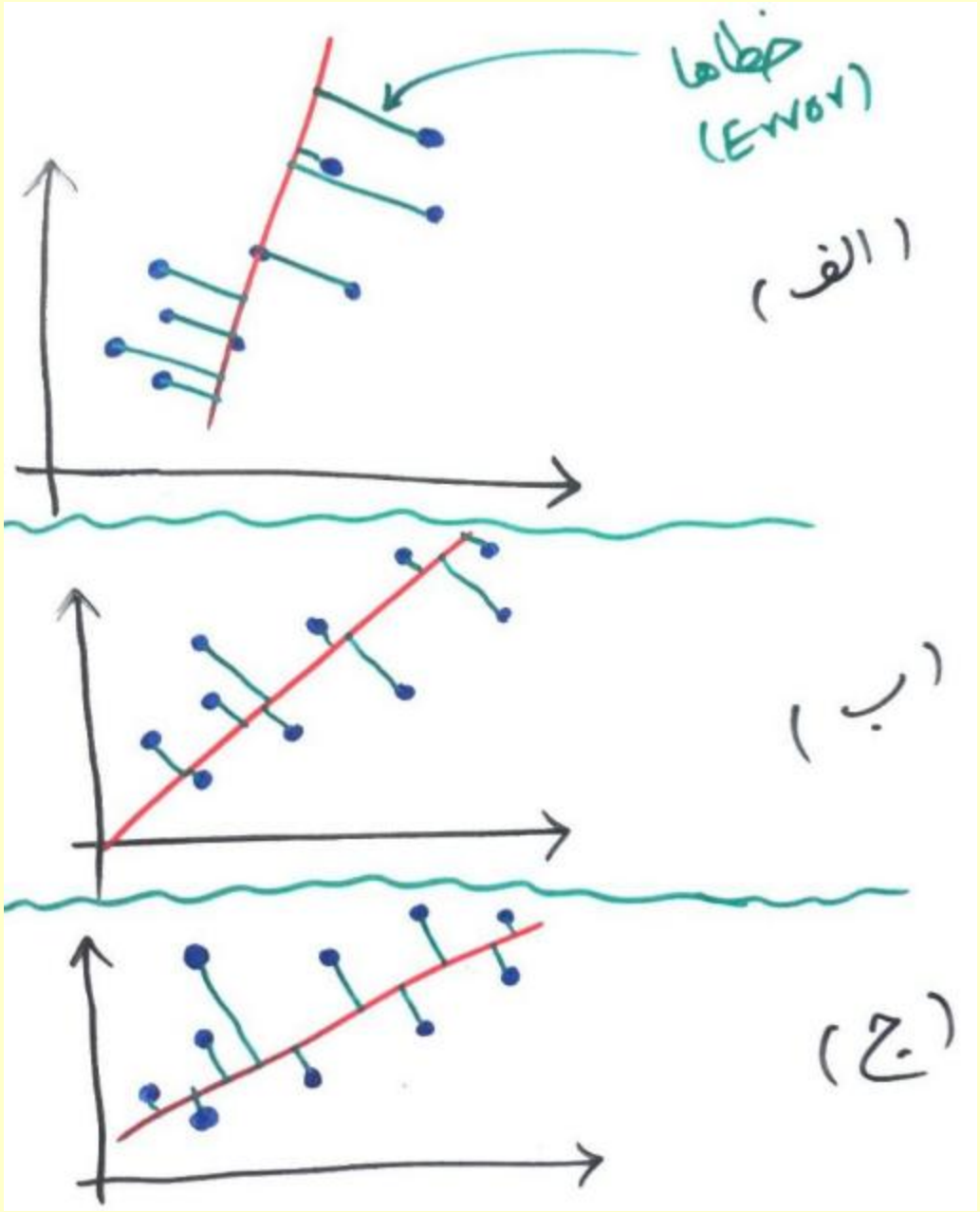
- مدیرِ فروش به ما گفته است که به جای اینکه هر ۷ بُعد را در تصمیم‌گیری دخالت دهیم، نیاز به ۳ بُعد (۳ویژگی) داریم.
- PCA به کمک ترکیب خطی ویژگی‌ها این کار را انجام می‌دهد.
- فرض کنید ویژگی قد، تاثیر زیادی در اینکه یک فرد از فروشگاه خرید کند نداشته باشد. PCA این قضیه را متوجه می‌شود و در الگوریتمِ خود ویژگیِ قد را تا جای ممکن حذف می‌کند. این‌گونه است که ویژگی‌های مهمتر از نظر PCA وزن بیشتری در تولیدِ ویژگی‌های کاهش یافته پیدا می‌کنند.
- PCA توان این را دارد که به یک سری ویژگی جدید برسد. مثلا این الگوریتم ممکن است به این نتیجه برسد که افرادی که در شمال و غرب ایران زندگی می‌کنند و سنِ آن‌ها بالای ۴۰سال است، احتمال خرید بالایی دارند. در واقع این‌جا PCA به یک ویژگی ترکیبی از محل تولد شخص و سن رسیده است.

- **برای سادگی فرض کنید داده‌های مشتریان ما کلا ۲بُعد دارند. سن و قد.**



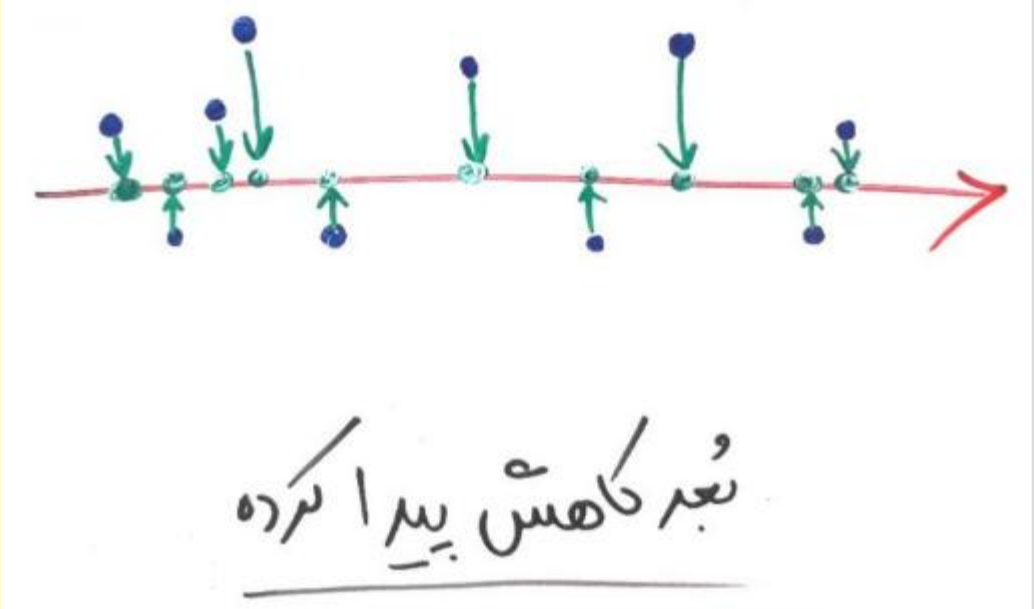
- خطی که کمترین فاصله با نقاط را دارد رسم می‌کنیم.





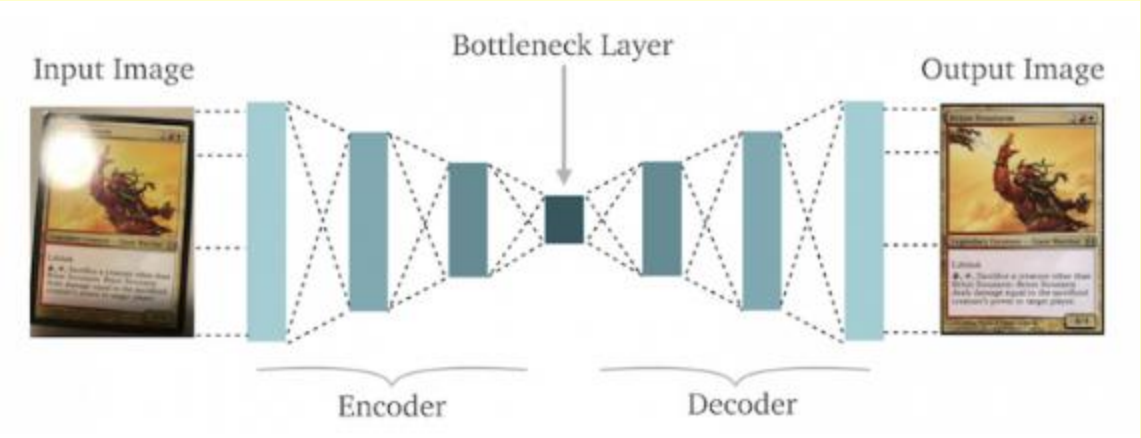
- محاسبه خطا

- هر کدام از نقاط بر روی خط قرمز نگاشت می‌شوند و در تصویرِ قبل که ۲ بُعدی است می‌توان این ۲ بُعد را به ۱ بُعد (که همان خط قرمز رنگ است) نگاشت کرد.

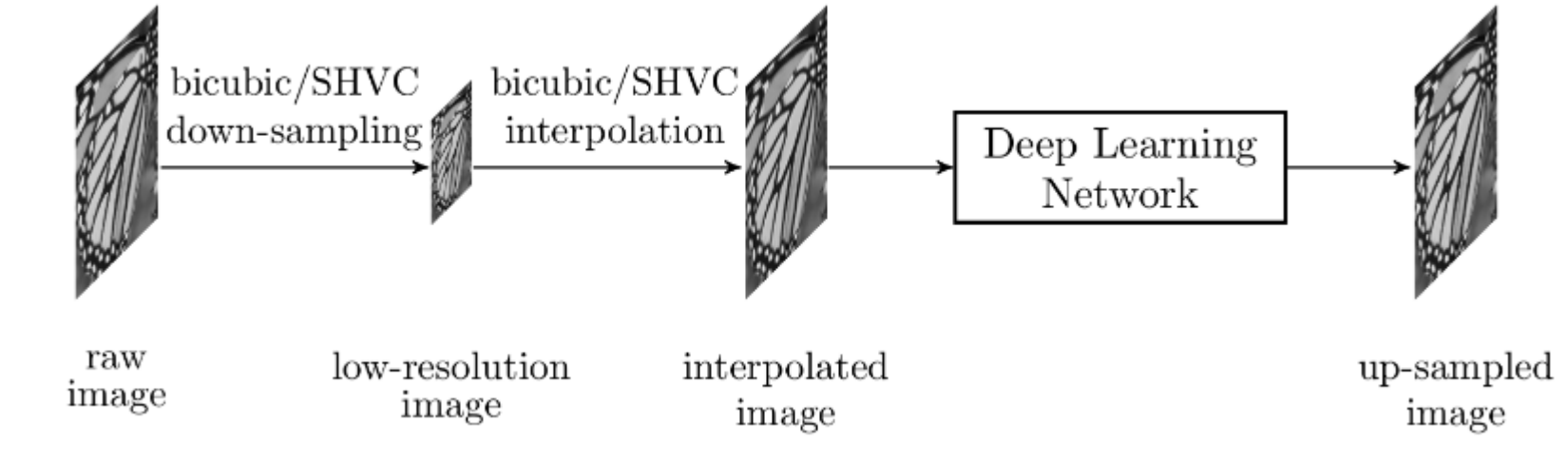


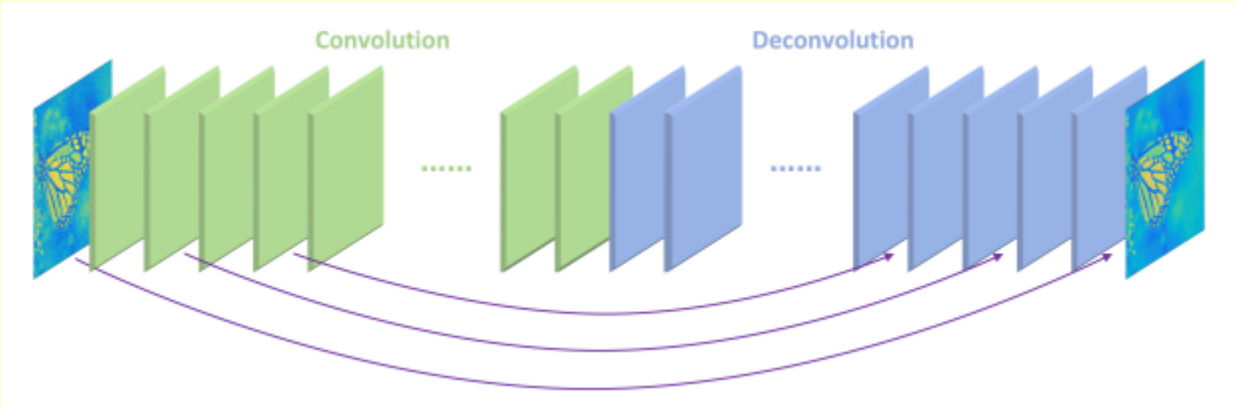
کاربردها

- رفع نویز

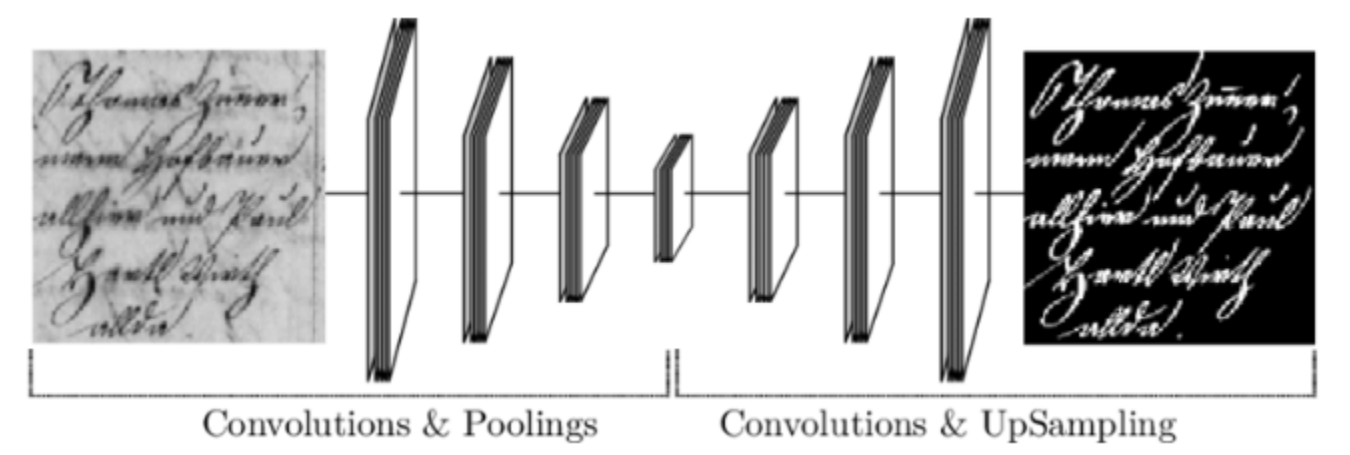


سوپر رزولوشن

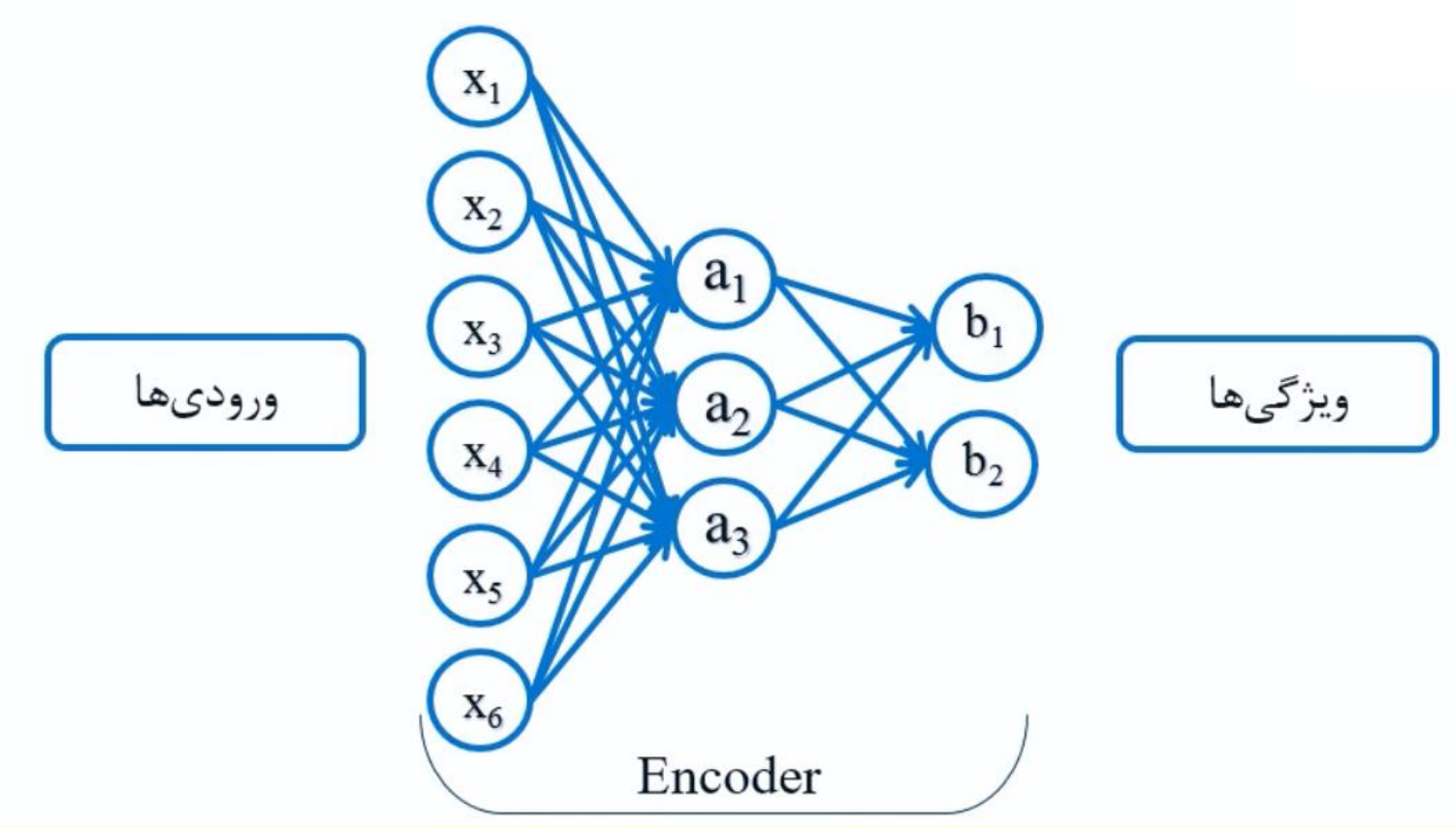




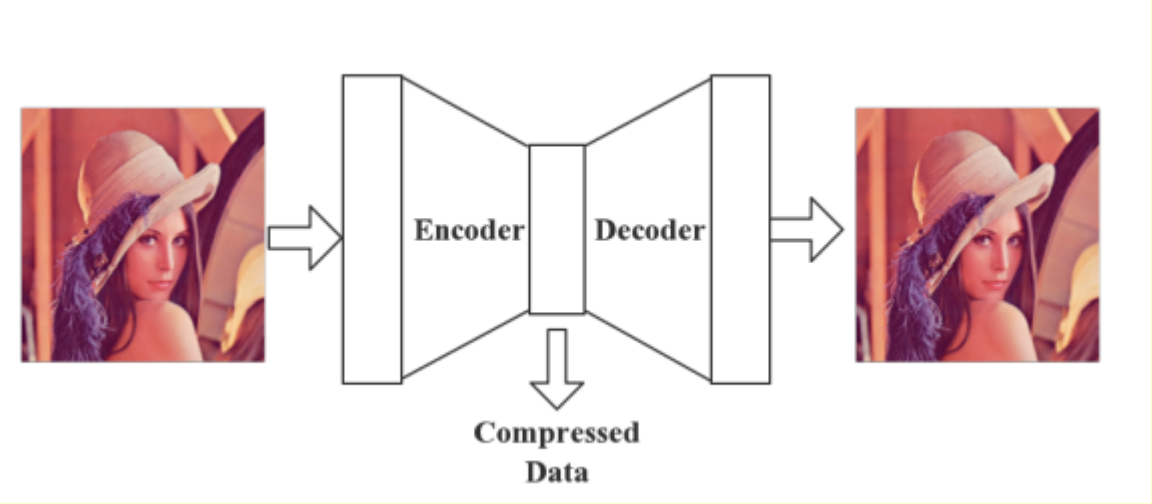
باینری‌سازی



استخراج بدون نظارت ویژگی‌ها

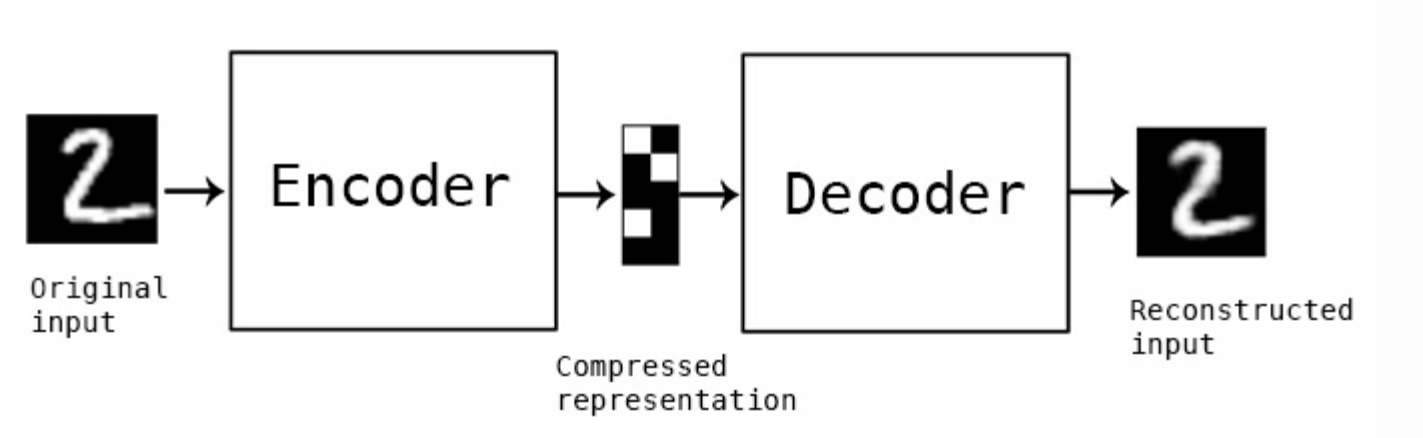


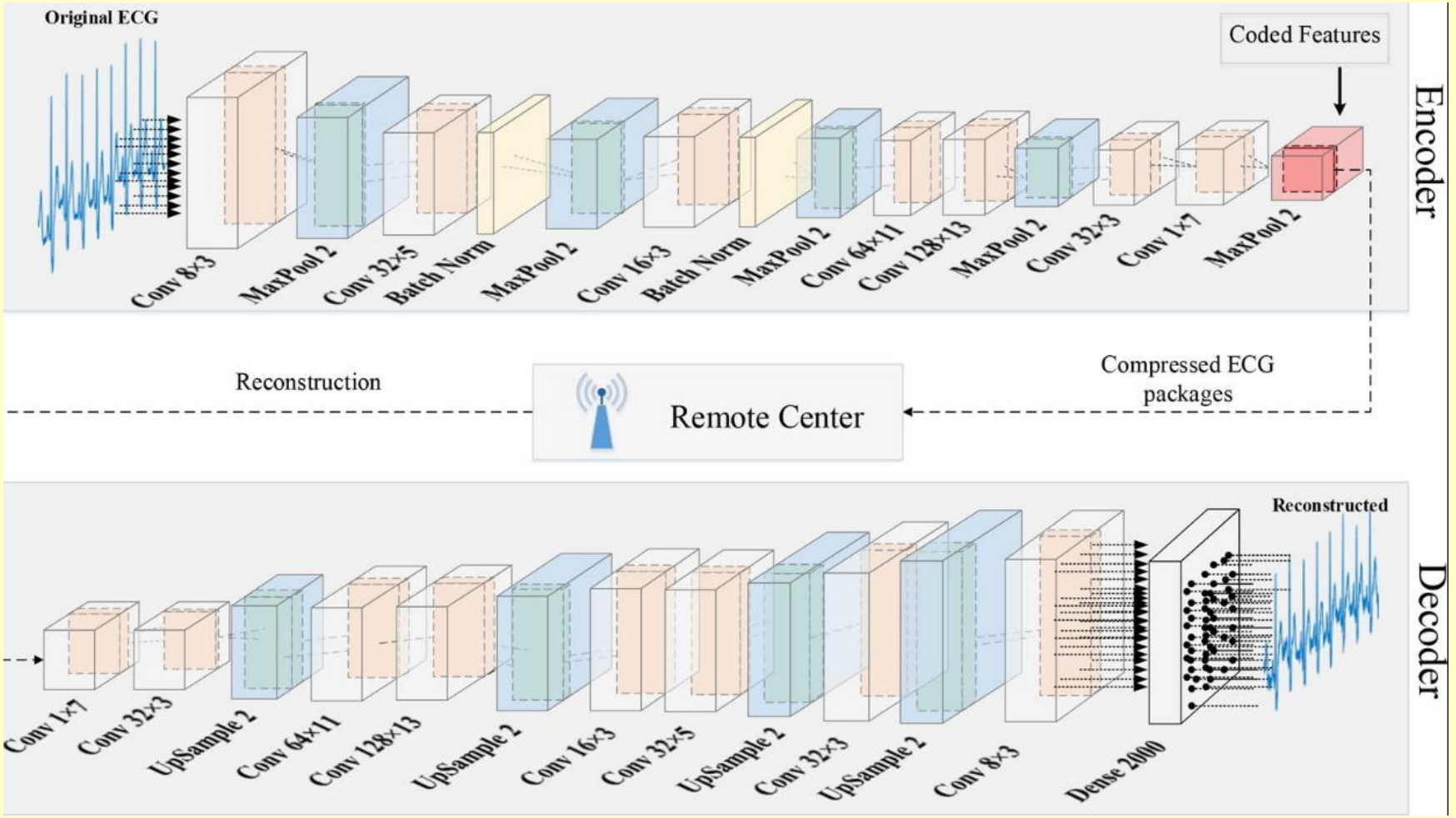
فشرده‌سازی با اتوانکدر: مرحله آموزش



فشرده‌سازی با اتوانکدر: مرحله آزمون

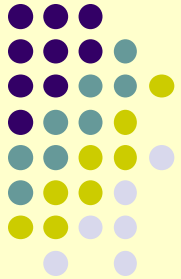
- بخش انکدر از دیکدر جدا می‌شود.
- خروجی بخش انکدر اطلاعات فشرده است.
- خروجی بخش دیکدر اطلاعات بازسازی شده است.




Original ECG
Coded Features
Conv 8×3
MaxPool 2
Conv 32×5
Batch Norm
MaxPool 2
Conv 16×3
Batch Norm
MaxPool 2
Conv 64×11
Conv 128×13
MaxPool 2
Conv 32×3
Conv 1×7
MaxPool 2
Encoder
Reconstruction
Remote Center
Compressed ECG packages
Reconstructed
Conv 1×7
Conv 32×3
UpSample 2
Conv 64×11
Conv 128×13
UpSample 2
Conv 16×3
Conv 32×5
UpSample 2
Conv 32×3
UpSample 2
Conv 8×3
Dense 2000
Decoder

انواع اتوانکدر

- Stack Auto Encoder (SAE)
  - نوع ساده‌ای که معرفی شد.
- Denoising Auto Encoder (DAE)
  - ورودی داده نویزدار است. در بازسازی داده بدون نویز را می‌خواهیم.
- Sparse Auto Encoder (SAE)
- Contractive Auto Encoder (CAE)
  - مقداری داده در حین آموزش برای آموزش بهتر اضافه می‌شوند.
- Convolutional Auto Encoder (CAE)
  - لایه‌ها به صورت تمام متصل نیستند. لایه‌ها به صورت کانولوشنی هستند.
- Variational Auto Encoder (VAE)
  - از الگوریتم‌های آماری برای تولید داده استفاده می‌کنند.



شبکه‌های تولید کننده داده جدید

- از آموزش بدون نظارت برای تولید داده جدید استفاده می‌کنند.

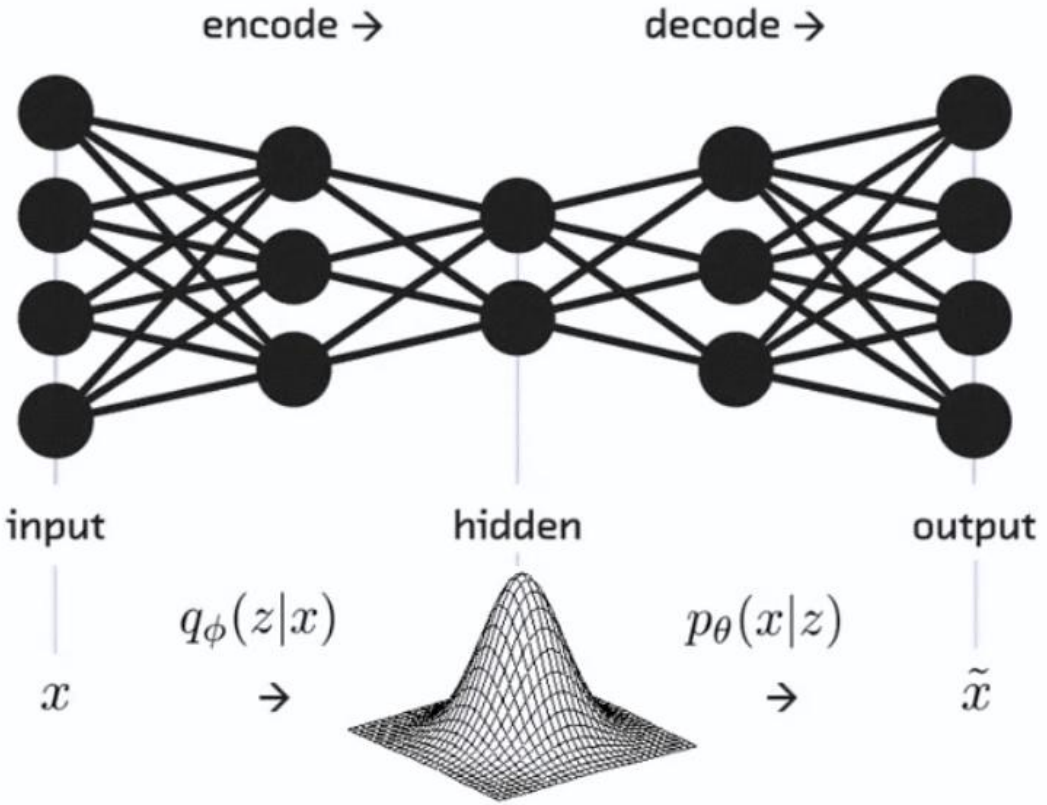
- Variational Auto Encoder (VAE)
- Generative Adversarial Network (GAN)

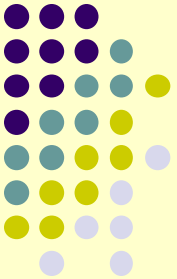
Variational Auto Encoder (VAE)

- برای آموزش شبکه از توابع آماری و احتمالاتی استفاده می‌کنند.
- از تئوری ریاضی خوب و قوی در آن استفاده شده است.
- داده‌های خروجی شبیه داده‌های ورودی است ولی داده‌ی جدیدی است که تولید می‌شود.

Encoder: q(z|x)

Decoder: p(x|z)

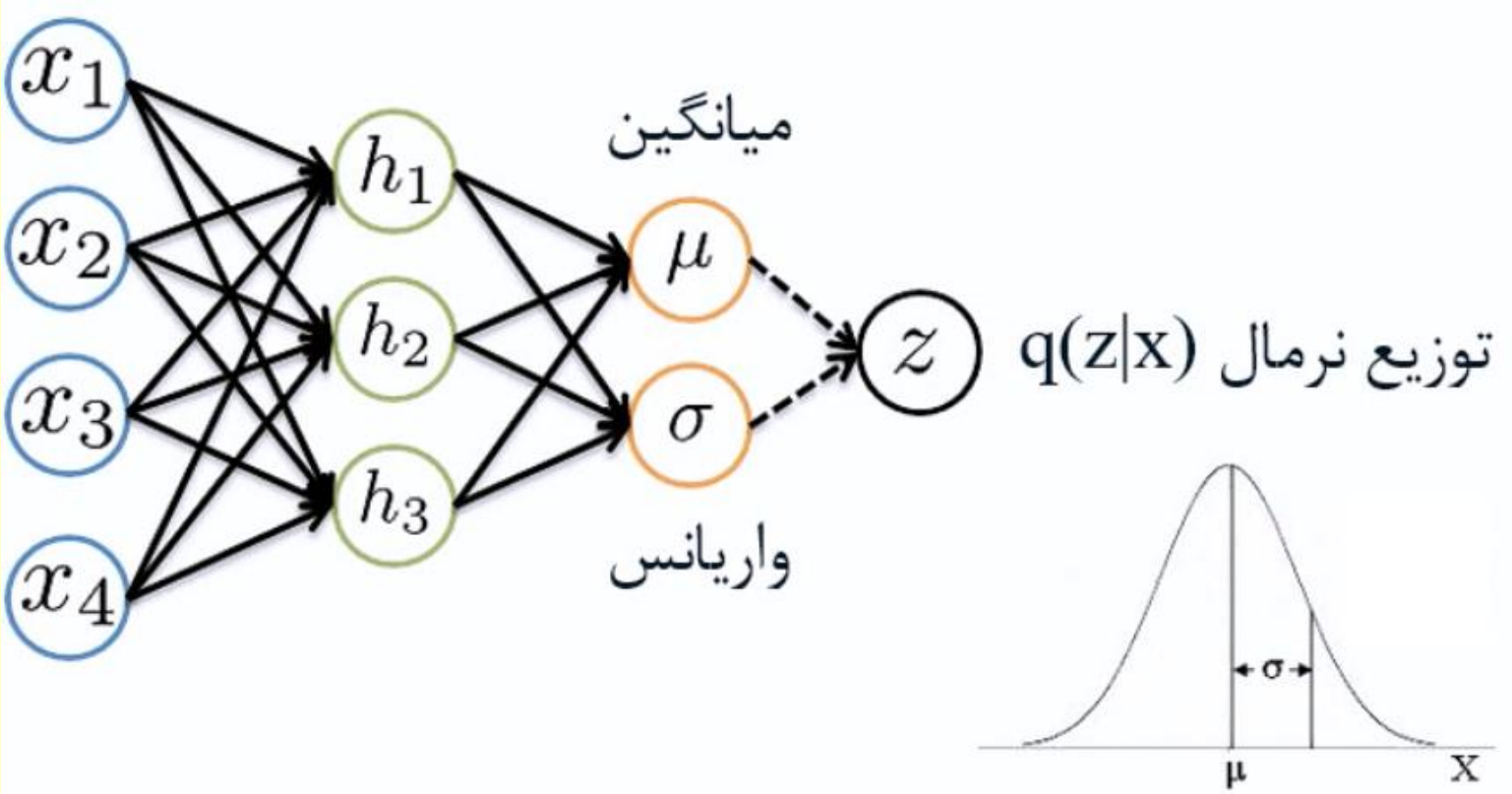




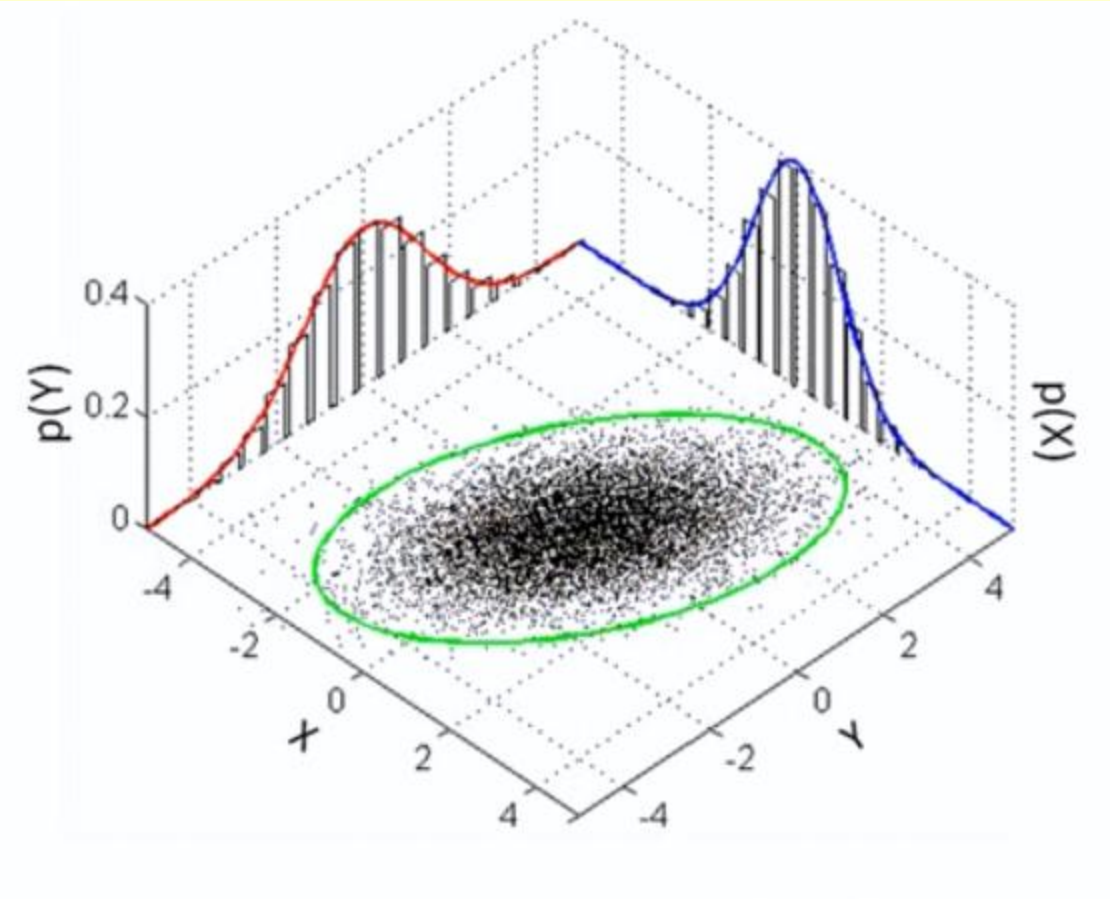
- می‌خواهیم داده خروجی شبیه به داده ورودی باشد، ولی داده‌ای جدید باشد.
- به بخش انکدر به صورت یک احتمال شرطی نگاه می‌شود.
- به بخش دیکدر نیز به صورت یک احتمال شرطی نگاه می‌شود. احتمال وقوع **Z** به شرط داشتن **X** .
- فرض می‌کنیم داده‌های ورودی از توزیعی خاص تبعیت می‌کند و فضای ویژگی نیز دارای توزیعی خاص است.
- در شکل قبل توزیع فضای ویژگی، توزیع نرمال یا گوسی است. اما توزیع داده‌های ورودی را نمی‌دانیم.
- با این فرض‌ها و فرمول‌های ریاضی شبکه را آموزش می‌دهیم.

مرحله آموزش

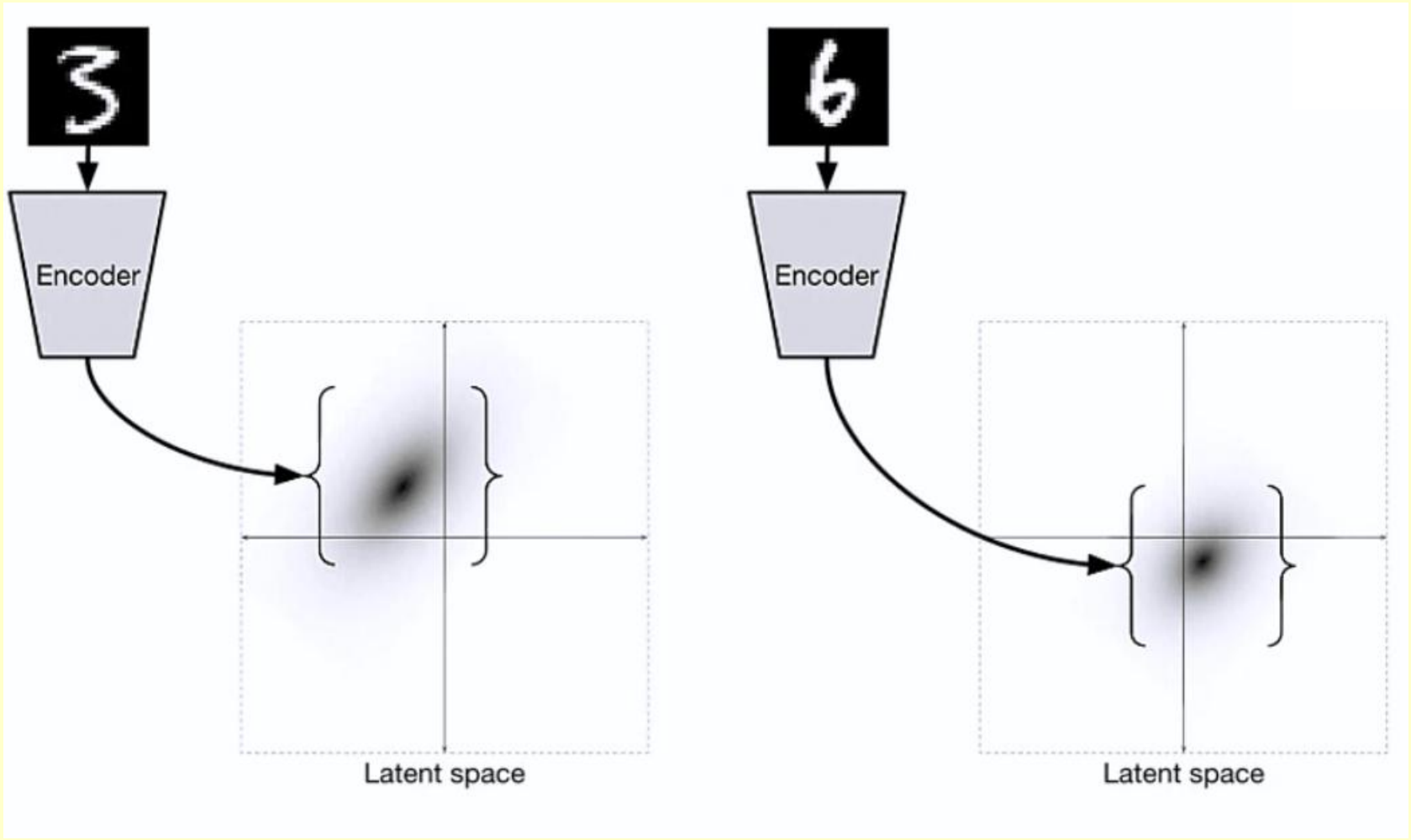
- آموزش فضای ویژگی‌ها (Latent) : یعنی بخش انکدر شبکه.
- فرض می‌کنیم ویژگی‌ها از توزیع گوسی تبعیت می‌کنند. نتیجه احتمال Z به شرط وقوع X یک توزیع گوسی است. یعنی با توجه به ورودی‌های مختلف، مجموع ویژگی‌ها دارای توزیع نرمال می‌شوند.
- این شبکه توزیع یک بعدی ویژگی‌ها را در اختیار ما قرار می‌دهد.

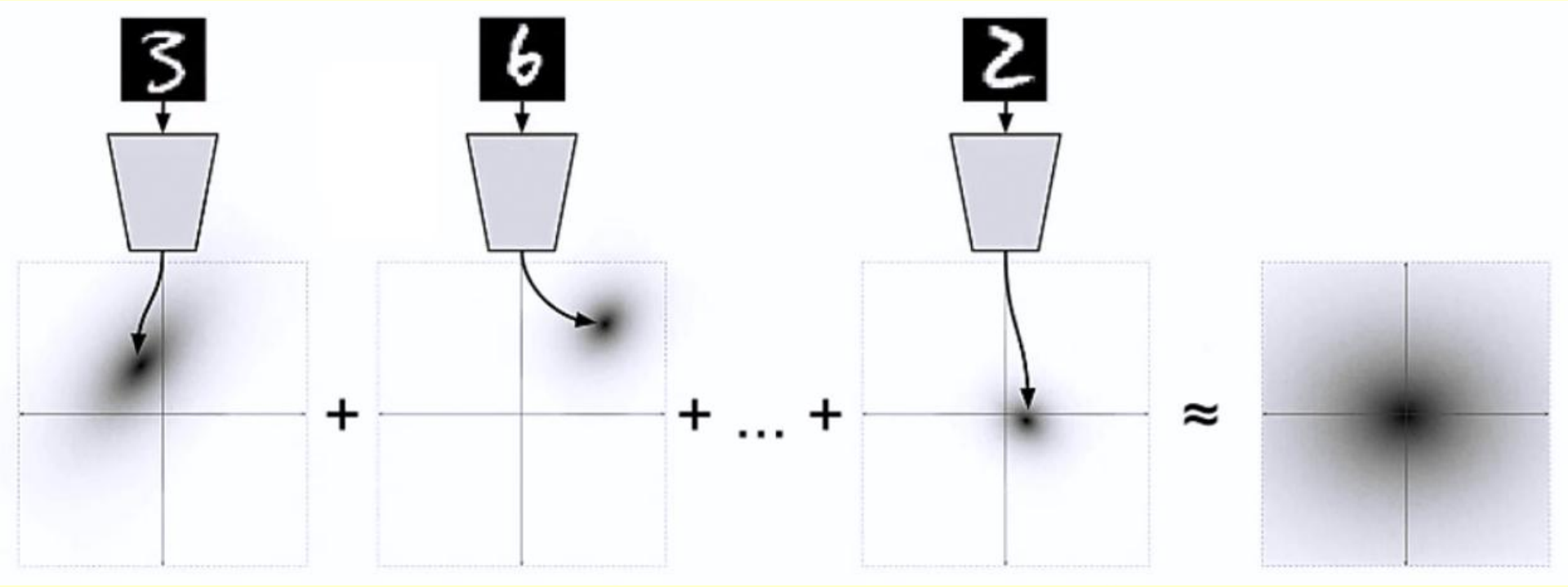


- برای ارائه توزیع دوبعدی نیاز به دو نورون برای میانگین و دو نورون برای واریانس است.

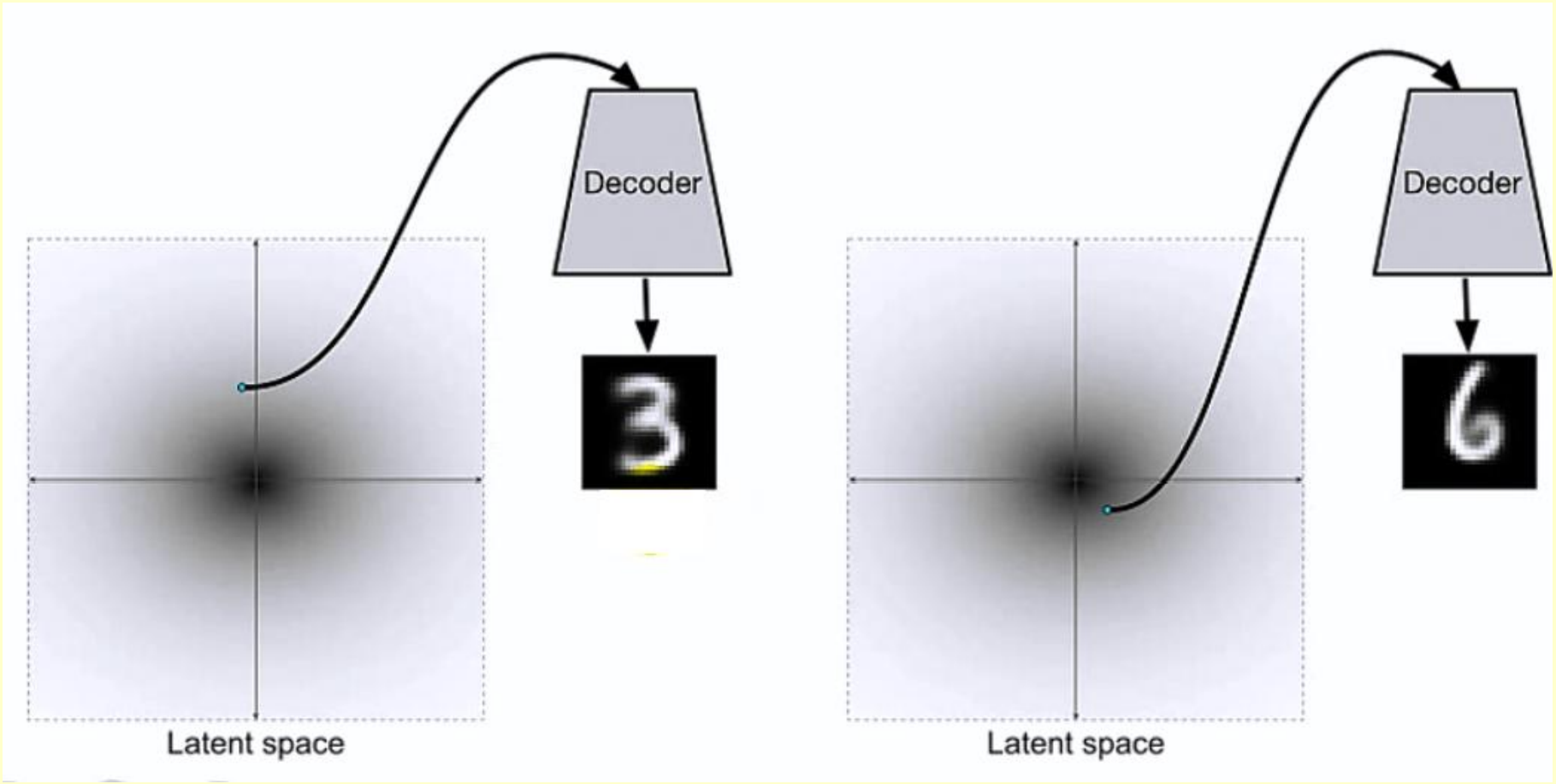


آموزش فضای ویژگی یعنی چه؟



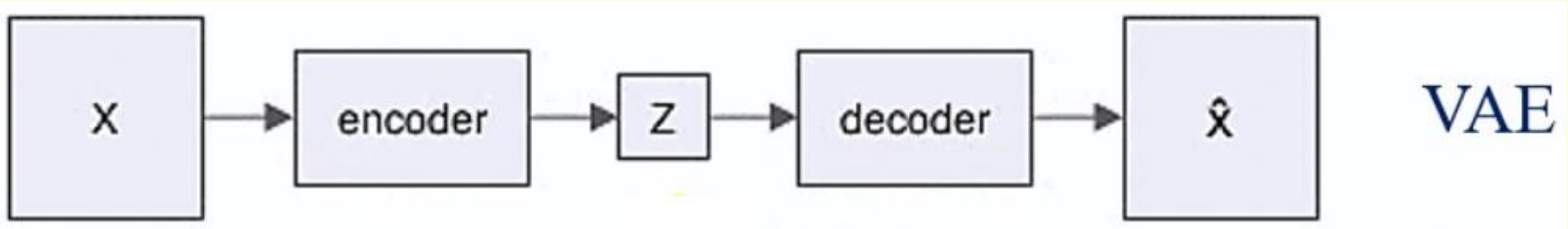

3
6
2
+
+
...
+
≈

تولید داده جدید

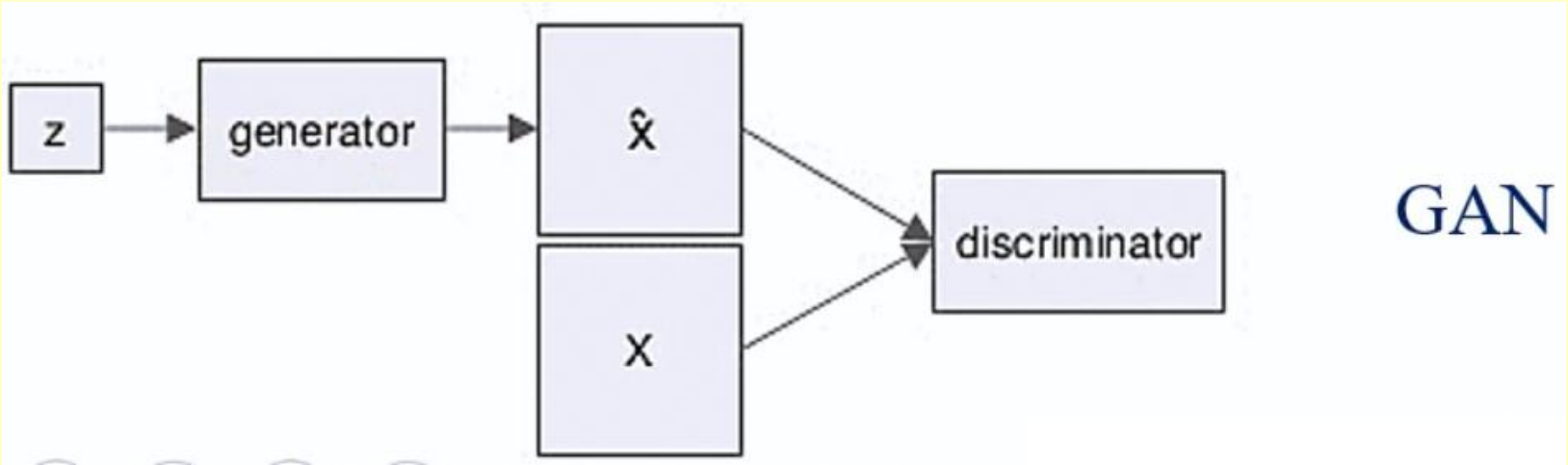


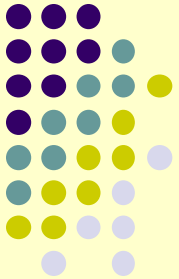
Generative Adversarial Network

- در سال ۲۰۱۴ ارائه شده است.
- از VAE برای تولید داده استفاده می‌شود. البته مکانیزمی دیگر به فرآیند تولید داده اضافه شده که موجب شده داده واقعی‌تر به نظر برسد.



- شبکه GAN از دو بخش تشکیل شده است.

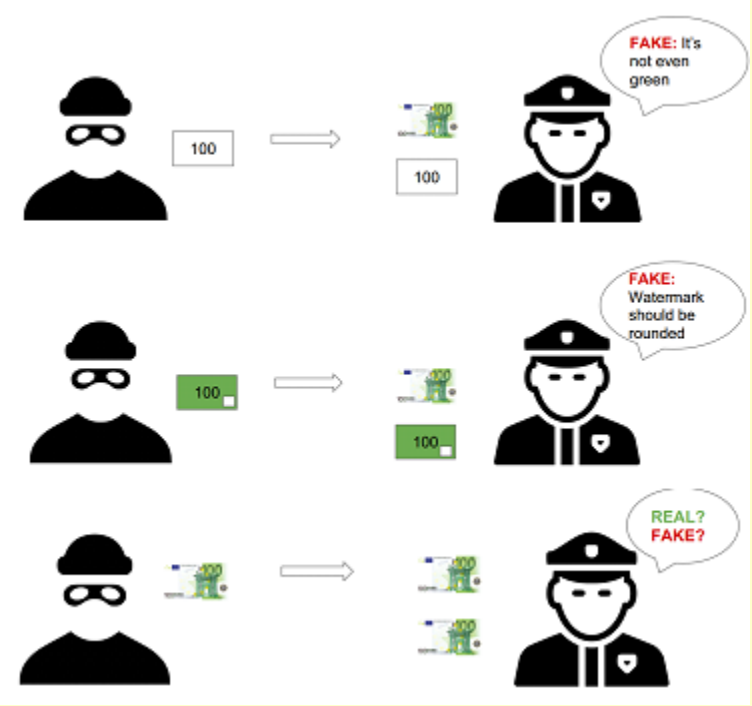




- این شبکه ها بر اساس رویکرد تئوری بازی ها بناگذاری شده اند که در آن یک شبکه یادگیری عمیق که مولد نامیده می‌شود با یک روند تخاصمی رقابت می‌کند.
- شبکه عمیق دیگری که متمایزکننده نامیده می‌شود، سعی می‌کند نمونه‌های تولید شده از شبکه مولد را از داده های اصلی متمایز کند. رقابت بین این دو شبکه در نهایت باعث یادگیری بهتر و بهبود عملکرد هر دو می‌شود.

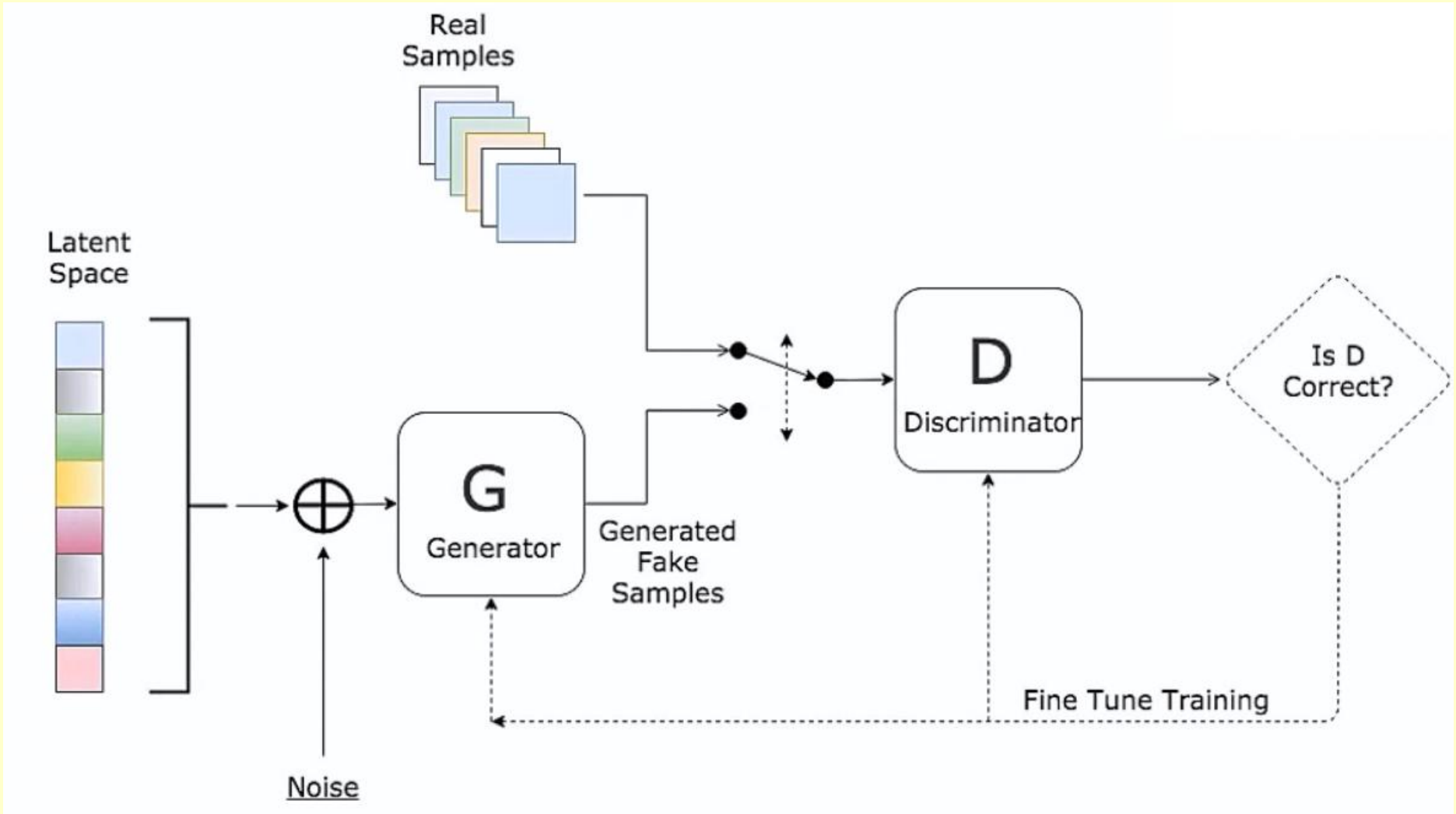
مثال: پول تقلبی

- فرض کنید یک جعل کننده اسکناس (G) سعی دارد پول‌هایی را چاپ کند که پلیس (D) قادر به تشخیص آن از پول اصلی نباشد.
- در اینجا D شبکه‌ای است که وظیفه آن تشخیص داده ساختگی است. به عبارت دیگر، پارامتر های این شبکه ویژگی‌های متمایزکننده پول‌های ساختگی و اصلی را فرامی‌گیرند. پارامترهای شبکه G نیز هر بار سعی می‌کنند توزیع داده‌های اصلی را فراگرفته و هر بار نمونه‌های بهتری تولید کنند.



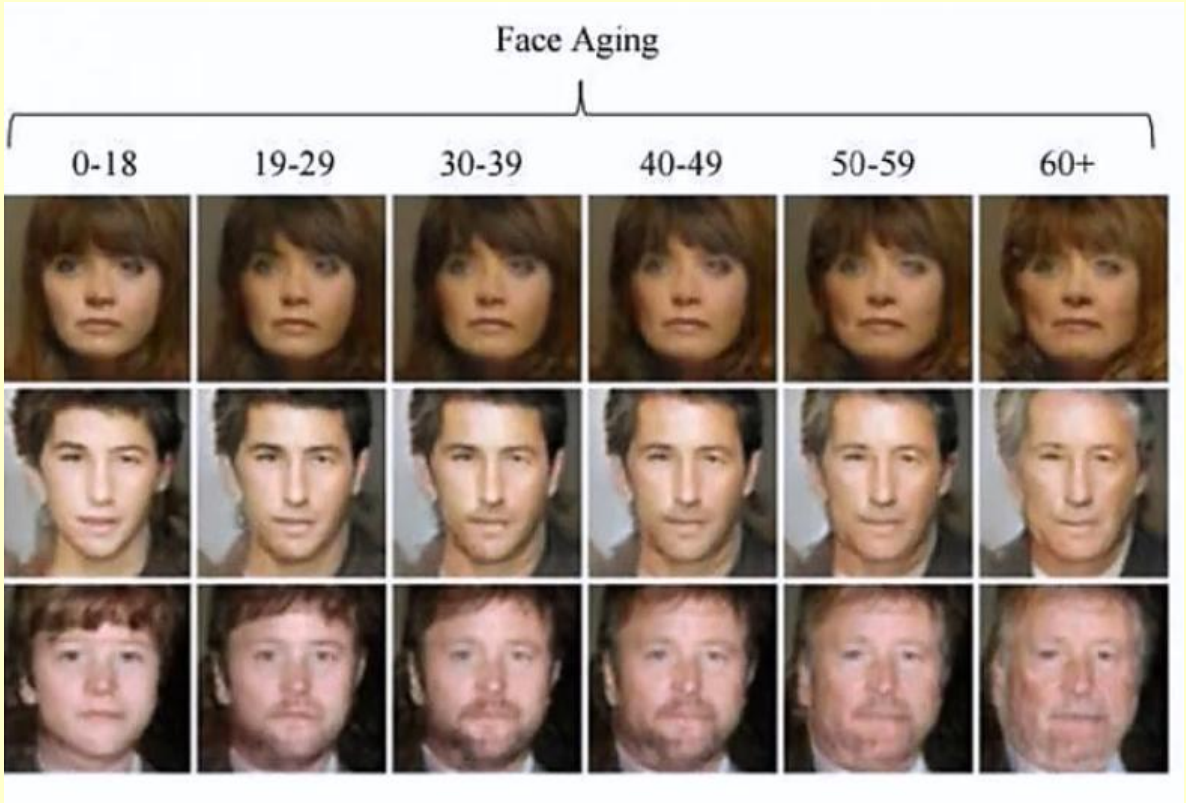
- در مساله پول تقلبی که اشاره شد، در ابتدا جاعل G به اندازه کافی در تولید پول تقلبی قوی عمل نمی‌کند. ولی با تکرار سعی و خطا و گرفتن فیدبک از پلیس D، رفته رفته قوی‌تر شده و ویژگی‌های پول اصلی را فرامی‌گیرد تا جایی که دیگر D قادر به تشخیص پول جعلی از اصلی نیست.
- برای یادگیری توزیع شبکه مولد بر روی داده‌های ورودی، ابتدا از یک توزیع نویز به عنوان ورودی استفاده می‌شود. هدف یافتن تابع مولد (G) می‌باشد. شبکه عصبی دوم متمایزکننده (D) نام دارد و خروجی آن اسکالر است.
- به ازای هر k بار بهینه‌سازی تابع D یک بار تابع G بهینه می‌شود.

آموزش شبکه GAN

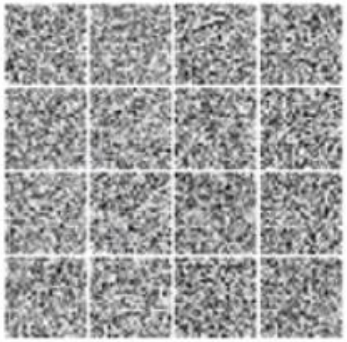


کاربردهای GAN

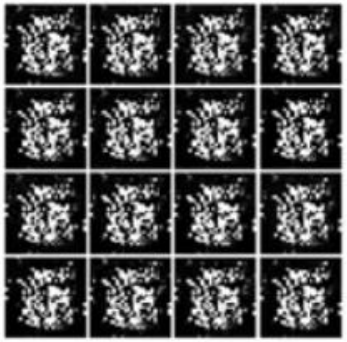
- تولید داده جدید: مثلا تولید داده پزشکی
- اصلاح داده‌ها: داده مخدوش را اصلاح می‌کند. مثلا کاهش نویز یا سوپر رزولوشن
- تغییر شکل داده‌ها: در بخش تولید داده یک شرط نیز می‌گذاریم که داده‌ها با این شرط تولید شوند.
- مثلا پیر کردن چهره



تولید ارقام



تکرار اول



تکرار سوم

تکرار دهم

تکرار ۲۰

تکرار ۳۰

تکرار ۳۷

ایجاد محصولات هنری

- گرافیک : از ترکیب عکس‌ها یک عکس هنری تولید می‌کند. مثلا از بافت یک عکس در عکس دیگر استفاده می‌کنند.
- متن : تولید متن ادبی با استفاده از چندین متن. مثلا شعرهای حافظ به این شبکه داده شده و این شبکه شعر تولید می‌کند.
- موسیقی : تولید موسیقی جدید