

(۱)

- **Unsupervised Learning** یا یادگیری بدون ناظر به نوعی از یادگیری میگویند که داده های آموزشی در آن لیبل ندارند. در واقع برای مسائلی که به صورت **Unsupervised** حل میشوند، جواب درستی وجود ندارد و بیشتر برای کلاسترینگ و دسته بندی داده ها استفاده میشود.
- **Self-Supervised Learning** در واقع حالتی از **Unsupervised Learning** است، یعنی در این نوع یادگیری نیز داده ها لیبل ندارند. در واقع در این روش مدل با استفاده از بخشی از داده سعی در پیشبینی بخش دیگر و برچسب گذاری آن دارد. برای مثال برای پیش بینی بخش بریده شده ای از عکس میتوان از مدل های **Self-Supervised Learning** استفاده کرد.
- **Representation Learning** یا یادگیری بازنمایی که به آن **Feature Learning** (یادگیری ویژگی) نیز گفته میشود به مجموعه ای از تکنیک هاست که به سیستم اجازه میدهد به طور خودکار ویژگی ها و بازنمایی های مورد نیازش را از داده خام بدست آورد. حال یادگیری های باز نمایی همانطور که در تصویر صورت سوال مشاهده میکنید، میتوانند با ناظر یا بدون ناظر باشند. به عنوان نمونه **pre-train** کردن **image-net** یک نوع یادگیری بازنماییست که باعث میشود که ما وزن هایمان را به صورت رندوم مقدار دهی نکنیم و بتوانیم راحت تر مسئله را حل کنیم. همچنین مثال کامل کردن عکس های بریده شده که در بالا گفته شد نیز نوعی یادگیری بازنمایی است.

منابع:

<https://towardsdatascience.com/supervised-semi-supervised-unsupervised-and-self-supervised-learning-7fa79aa9247c>

[https://en.wikipedia.org/wiki/Feature\\_learning](https://en.wikipedia.org/wiki/Feature_learning)

(۲)

- **Utility** در واقع به عنوان معیاری از سودمندی یادگیری خود نظارتی تعریف میکنیم. از **Utility** برای صرفه جویی در استفاده از داده های لیبل زده شده استفاده میشود که به این صورت تعریف میشود. اگر دقت مدلی را که از ابتدا با  $n$  داده برچسب زده شده آموزش دیده باشد را  $a(n)$  و دقت مدل **fine tune** شده را  $aft(n)$  در نظر بگیریم، **Utility** مدل به صورت  $U(n) = \frac{n_1}{n} - 1$  محاسبه میشود که در آن  $n_1$  برابر با عددی است که  $a(n_1) = a(n)$  شود.

- در مقاله دو روش Downstream Task زیر آمده است:
  - Object Classification : در این روش مدل را آموزش می‌دهیم تا بتوان داده را از بین ۱۰ کلاس ShapeNet که برای ارائه داده های مصنوعی استفاده میشود، تفکیک کرده و لیبل بزند. تصاویر به گونه ای انتخاب شده اند که فقط یکی از ۱۰ کلاس را شامل میشوند و توضیح تصاویر مربوط به هر کلاس به صورت یکنواخت است. عملکرد نهایی مدل را نیز با استفاده از Accuracy آن اندازه گیری میکنیم.
  - Object Pose Estimation : در این روش هم داده های مورد استفاده دارای یک شی هستند و آن در وسط تصویر میباشد.
- در این مقاله در باره ی ۴ الگوریتم Self-Supervised برای Pretrain بحث شده است:
  - Variational autoencoder یا به اختصار VAE یک استاندارد برای نگاشت تصاویر به فضاهای با بعد کمتر است.
  - Rotation یک روش ساده و موثر برای pretraining است. در این روش شبکه پیشبینی میکند که آیا تصویر به صورت ۰، ۹۰، ۱۸۰ یا ۲۷۰ درجه چرخیده است یا خیر.
  - Contrastive Multiview Coding یا به اختصار CMC یک روش نسبتاً جدید برای یادگیری خودنظارتی است که با تقسیم یک تصویر به کانال های مختلف مانند L و ab کار میکند. دو کانال از دو شبکه نصف شده عبور داده میشوند خروجی های آن ها با خروجی های تصاویر دیگر مقایسه میشوند.
  - Augmented Multiscale Deep InfoMax یا به اختصار AMDIM مانند روش قبل عمل میکند با این تفاوت که به جای مقایسه بین کانال های تصویر، نمایش دو نسخه تقویت شده تصویر را با نمایش های لایه های میانی شبکه مقایسه میکند.