Unsupervised learning: Generative Models

The defining characteristic of unsupervised learning models is that they are learned from a set of observed data $\{x_i\}$ without the presence of labels.

All unsupervised models share this property, but they have diverse goals.

- Generative: to generate plausible new samples from the dataset
- Outlier Detection: distinguish whether new examples belong to the same data set or are outliers.
- manipulate, denoise, interpolate between, or compress examples.
- Reveal the internal structure of a dataset (e.g., by dividing it into coherent clusters)

Deep unsupervised = "generative"

Focus: Generative

Type of generative

- generative adversarial net- works,
- variational autoencoders,
- normalizing flows, and
- diffusion models.

Generative vs discriminative

Discriminative:

- map data to labels

Generative:

- Map a *latent* variable z to data.
- Latent means?

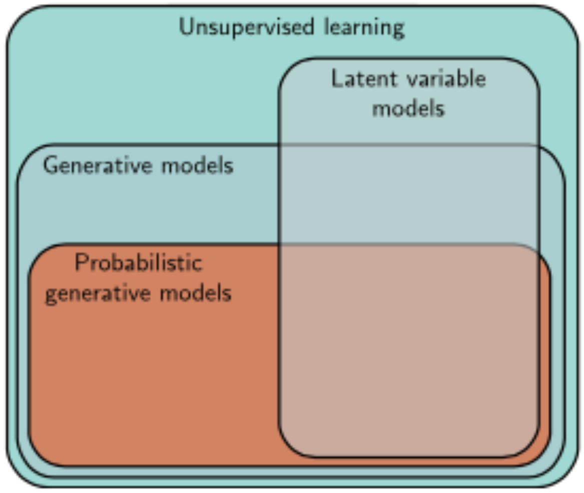


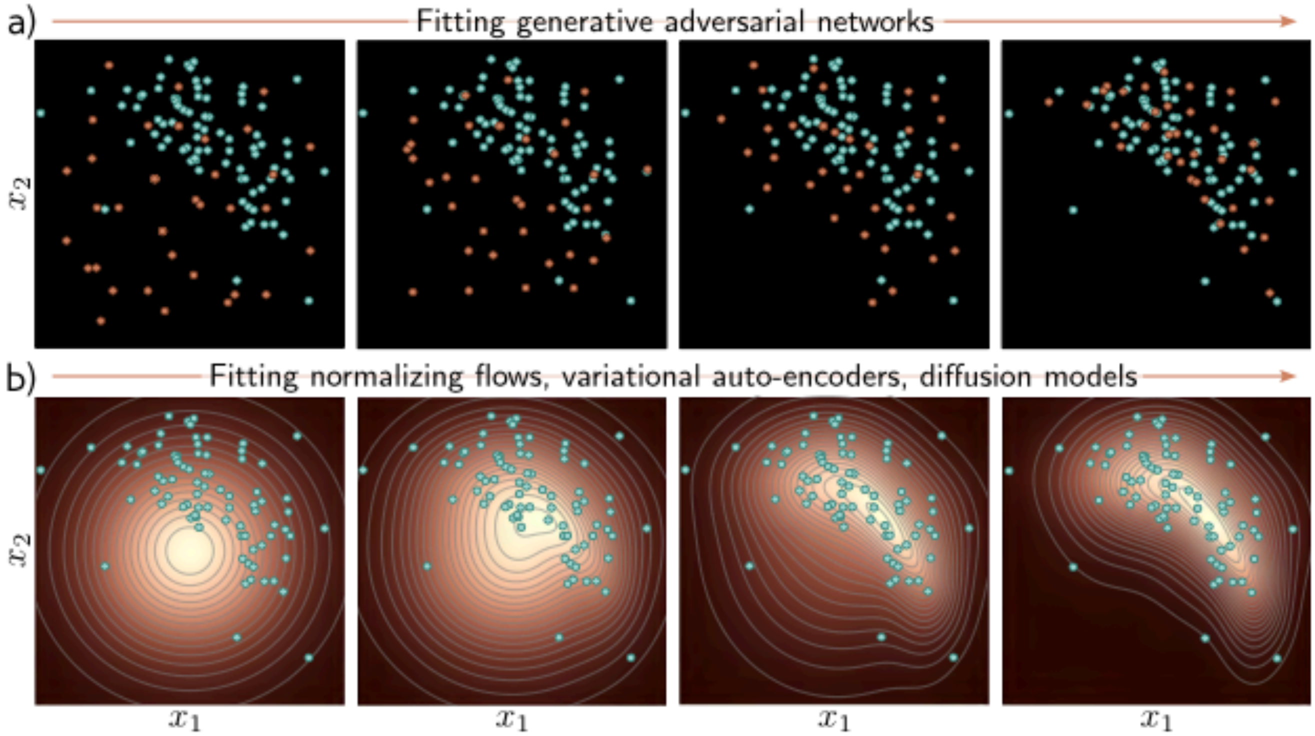
Figure: Unsupervised learning can refers to any model applied to datasets without labels. Generative models can synthesize or generate new examples that have the same statistics as the train- ing data. A subset of these are proba- bilistic and also define a distribution over the data; to generate new examples, we draw samples from this distribution. Latent variable models define a mapping between an underlying explanatory (la- tent) variable z and the data x. They may fall into any of the above categories

What makes a good generative model?

Generative models based on latent variables should have the following properties:

• Efficient sampling: Generating samples from the model should be computation- ally inexpensive and take advantage of the parallelism of modern hardwar

- High-quality sampling: The samples should be indistinguishable from the real data that the model was trained with.
- Coverage: Samples should represent the entire training distribution. It is insuf- ficient to only generate samples that all look like a subset of the training data.
- Well-behaved latent space: Every latent variable z should correspond to a plausible data example x and smooth changes in z should correspond to smooth changes in x.
- Interpretable latent space: Manipulating each dimension of z should corre- spond to changing an interpretable property of the data. For example, in a model of language, it might change the topic, tense, or verbosity.
- Efficient likelihood computation: If the model is probabilistic, we would like to be able to calculate the probability of new examples efficiently and accurately.



Quantifying performance

Test likelihood: One way to compare probabilistic models is to measure the likelihood of a set of test data. It is not effective to measure the likelihood of the training data because a model could simply assign a very high probability to each training point and very low probabilities in between. This model would have a very high training likelihood but could only reproduce the training data. The test likelihood captures how well the model generalizes from the training data and also the coverage; if the model assigns a high probability to just a subset of the training data, it must assign lower probabilities elsewhere, and so a portion of the test examples will have low probability.

Test likelihood is a good way to quantify probabilistic models, but unfortunately, it is not relevant for generative adversarial models (which do not assign a probability), and is expensive to estimate for variational autoencoders and diffusion models (although it is possible to compute a lower bound on the log-likelihood). Normalizing flows are the only type of model for which the likelihood can be computed exactly and efficiently.

Inception score: The inception score (IS) is specialized for images, and ideally for generative models trained on the ImageNet database. The score is calculated using a pre-trained classification model – usually the “Inception” model, from which the name is derived. It is based on two criteria. First, each generated image should look like one and only one of the 1000 possible categories y. Hence, the the probability distri- bution Pr(y|xn) should be highly peaked at the correct class. Second, the entire set of generated images should be assigned to the classes with equal probability, so Pr(y) should be flat when averaged over all generated examples.

Notes:

Popular generative models include

- generative adversarial networks (Goodfellow et al., 2014),
- variational autoencoders (Kingma & Welling, 2014),
- normalizing flows (Rezende & Mohamed, 2015),
- diffusion models (Sohl-Dickstein et al., 2015; Ho et al., 2020),
- auto-regressive models (Bengio et al., 2000; Van den Oord et al., 2016b), and
- energy-based models (LeCun et al., 2006). All except energy models are discussed in this book. Bond-Taylor et al. (2022) provide a recent survey of generative models.

Evaluation: Salimans et al. (2016) introduced the inception score and Heusel et al. (2017) introduced the Fréchet inception distance, both of which are based on the Pool-3 layer of the Inception V3 model (Szegedy et al., 2016). Nash et al. (2021) used earlier layers of the same network that retain more spatial information to ensure that the spatial statistics of images are also replicated. Kynkäänniemi et al. (2019) introduced the manifold precision/recall method. Barratt & Sharma (2018) discuss the inception score in detail and point out its weaknesses. Borji (2022) discusses the pros and cons of different methods for assessing generative models.