Marrying Graphical Models with Deep Learning

Welling, M.

Publication date
2016

Document Version
Final published version

Published in
ERCIM News

License
CC BY

Citation for published version (APA):
Welling, M. (2016). Marrying Graphical Models with Deep Learning. *ERCIM News*, *107*, 20-21. https://ercim-news.ercim.eu/en107

Marrying Graphical Models with Deep Learning

by Max Welling (University of Amsterdam)

In our research at the University of Amsterdam we have married two types of models into a single comprehensive framework which we have called "Variational Auto Encoders". The two types of models are: 1) generative models where the data generation process is modelled, and 2) discriminative models, such as deep learning, where measurements are directly mapped to class labels.

Deep learning is particularly successful in learning powerful (e.g., predictive/discriminative) features from raw, unstructured sensor data. Deep neural networks can effectively turn raw data streams into new representations that represent abstract, disentangled and semantically meaningful concepts. Based on these, a simple linear classifier can achieve the state of the art. But to learn them one needs very large quantities of annotated data. They are flexible input-output mappings but do not incorporate a very sophisticated inductive bias about the world. An important question is how far will this take us?

If we are asked to analyse a scene depicted in an image we seek a story that can explain the things we see in the image. Yes, there is a fast feedforward pipeline that quickly segments out the objects and classifies them into object classes. But when you need to truly understand a scene you will try to infer a story about which events caused other events, which in turn led to the image you are looking at. This causal story is also a powerful tool to predict how the events may unfold into the future.

So, to understand and reason about the world we need to find its causal atoms and their relationships. Now this is precisely what Bayesian networks [1] were intended to do. Each random variable connects to other random variables and their directed relations model their causal relationships. (Bayesian networks do not necessarily represent the causal relationships, but an extension called "structural equation models" does.) Another key advantage of interpretable models like Bayesian networks is that they can express our expert knowledge. If we know X causes Y then we can simply hard-code that relation into the model. Relations that we do not know will need to be learned from the data. Incorporating expert knowledge (e.g., the laws of physics) into models is the everyday business of scientists. They build sophisticated simulators with relatively few unidentified parameters, for instance implemented as a collection of partial differential equations (PDEs).

Generative models can also be used for classification by inverting the relationship they model from class label to input features. When you have a lot of (labelled) data at your disposal this type of classifier will generally speaking not work as well as a direct mapping from input features to labels (such as a deep neural network). But when the amount of data is small relative to the complexity of the task, the opportunity to
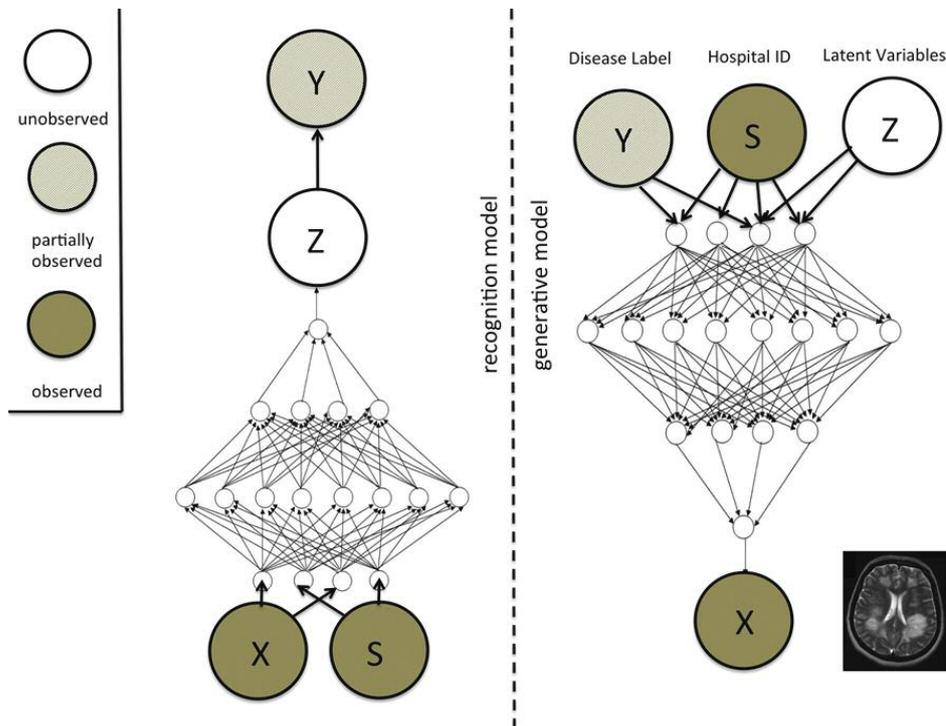
Figure 1: Example of a Variational Auto-encoder model. On the right we have the generative model with three groups of variables at the top: class labels Y, nuisance variables S and latent variables Z. In this example, we may think of Y as disease labels, S as hospital identifiers and Z as size, shape and other relevant properties of a brain. These variables are input to the generative process that generates a pseudo data case; in the example an MRI image of a brain. The discriminative or recognition model on the left takes all the observed variables as input and generates posterior distributions over the latent variables and possibly (if unobserved) the class labels. The models are trained jointly using the variational EM framework.

inject expert knowledge may pay back. Concluding, for very complex tasks (causal) generative models should in my opinion be part of the equation.

Can graphical models and deep neural networks be meaningfully combined into a more powerful framework? The variational auto-encoder (VAE) naturally combines generative models with discriminative models where the generative model can be a Bayesian network or a simulator and the discriminative model a deep neural network [2]. The discriminative model performs inference of the unobserved (latent) variables necessary to perform the (variational EM) learning updates. In this view the discriminative model approximately inverts the generative model. However, one can also interpret the VAE differently if we are more interested in the latent representation itself (based on which we can for instance perform classification). Now the generative model guides the discriminative model to learn interesting, semantically meaningful representations. They represent the fundamental sources of variation that are the input for the generative model. Thus, the

generative model may be viewed as an informed way to regularize the discriminative model.

VAEs as described above are a framework for unsupervised learning. However, they are easily extended to semi-supervised learning by incorporating labels in the generative model [3]. In this case the input data are generated by instantiating a label and some latent variables and sampling from the Bayesian network. In contrast, the discriminative model inverts this relationship and learns a mapping from input directly to class labels and latent variables (see Figure 1). Semi-supervised learning is now easy, because for the unlabelled examples, the label variable is treated as latent, while for a labelled data-case it is treated as observed.

In summary, marrying (discriminative) deep learning with causality and probabilistic reasoning in graphical models may be an important component in reaching the ambitious goals of Artificial General Intelligence. However, most likely completely new ideas are needed as well.

References:
[1] Pearl, Judea: "Probabilistic reasoning in intelligent systems: networks of plausible inference" Morgan Kaufmann, 2014.
[2] Kingma, P. Diederik, and M. Welling: "Auto-encoding variational bayes", the International Conference on Learning Representations (ICLR), Banff, 2014.
[3] Kingma, P. Diederik et al.: "Semi-supervised learning with deep generative models", Advances in Neural Information Processing Systems 2014: 3581-3589.
https://www.youtube.com/watch?v=XNZIN7Jh3Sg
http://dpkingma.com/?page_id=277

Please contact:
Max Welling
University of Amsterdam (UvA)
+31 (0)20 525 8256
welling.max@gmail.com
http://www.ics.uci.edu/~welling/staff.fnwi.uva.nl/m.welling/