What kind of a graphical model is the brain?

Geoffrey Hinton (review by **Vrushali B20BB047**)

November 6, 2021

1 Pros and Cons of the paper

1.1 Pros

The paper in general:

1. The paper takes into account a lot of perspectives such as even if the model performs well, is it interpret-able or not.
2. On the MNIST dataset used by the paper, it explained small details of what happened during the various epochs that it was trained on. This is good as it provides an insight into the entire process.

The final model proposed by the paper has many pros:

1. A good model is found fairly quickly (as per its time, today it would be considered a slow model) even with millions of parameters and many hidden layers.
2. The algorithm builds a generative model which is able to approximate the data's distribution well.
3. The algorithm is unsupervised which is considered as true learning.
4. It combines both a directed model and an undirected model hence combining the best features of the two.
5. Being a hybrid model, it doesn't treat its parts as separate but uses the associative memory extracted from the RBM and uses it as feature space for the directed network which converts it to observables such as pixels.
6. The learning is local i.e. activation of a particular node depends only on the nodes attached to it. This is called a local Markov property, which is important because not only is this how human neurons fire, but also because it reduces computation greatly.

1.2 Cons

1. The paper only considers one kind of data. There is no discussion about whether the proposed model can use other kinds of data (eg. temporal data, natural language data) as input or output or both. The brain can handle all such kinds of data, hence there should have been some discussion about the same before labelling it as a graphical model of the brain.
2. The model was trained only on the MNIST dataset which gives a very small error rate (0.25) even for models(CNNs) that were proposed before the publication of this paper. To be an improvement, other more complex datasets should've been tested.
3. No code used in the paper has been provided.
4. There is no formal, quantitative comparative analysis between all the models such as: evaluation of stability of the model (cannot be tested as only one dataset was used), evaluation of accuracy, sensitivity of the methods to hyperparameter, and what effect different kinds of pre-processing has on them and evaluation of computing cost.
5. The entire paper was quite qualitative in nature.
6. There was no information about an entire pipeline. Which is more or less acceptable because we are using a toy dataset.

2 Technical Errors

Errors mean mistakes or inaccuracies in a process. As this paper is qualitative and our ML course is just an introductory course, it wasn't possible for me to recreate the results and be able to gauge any technical errors.

3 Directions of future research

The following are some directions where future research can be done:

1. Compatibility with different types of data:
   (a) Finding out whether the proposed hybrid model works for other kinds of data such as temporal data, natural language data.
   (b) If it doesn't work then why.
   (c) What inspiration can we take from existing models which handle these different kinds of data
   (d) Whether to integrate different algorithms based on different data types, or whether to integrate different data types such that just one algorithm can handle their learning?
      i. How to integrate different algorithms
      ii. How to integrate different data types into one which can be used for a single algorithm
2. Naturally our brains also do tasks based on a reward system. Decisions are made based on what will provide more pleasure and least pain. While this is not very subjective, many tasks of human learning are objective. Hence I think introducing a reinforcement learning model in the hybrid model will help improve its performance.
3. Commenting on the experimental aspect:
   (a) More varied datasets (even of the same image type) should be used
   (b) And calculation of different kinds of metrics for different kinds of tasks should be done and not just classification
   (c) Different kinds of feature selection and extraction methods can be compared alongside this hybrid model
4. Newer generative models such as GANs (generative adversarial networks), variational autoencoders, auto-regressive models, bayesian networks (although sigmoid belief network is a bayesian network), HMMs should be compared with the proposed model.