NEURAL CAUSAL INDUCTION FROM INTERVENTIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Neural networks have achieved astonishing success in many machine learning areas by relying on the assumption of independent and identically distributed (iid) data in training and testing settings. In particular, these methods rely on learning correlations from the data and cannot per se identify causal links. In many settings of interest, e.g. domain adaptation, transfer learning, this is a problem as the iid assumption does not hold. On the other hand, causal models efficiently and compactly encode the data-generating process under all possible interventions. Often these models are represented as Bayesian networks and learning them scales poorly with the number of variables. Furthermore, these approaches cannot leverage previously learned knowledge to help with learning new causal models. To tackle these challenges, we propose a novel framework called *causal relational networks* (CRN) for learning causal models using neural networks. Furthermore, we propose to use a decoding-based metric for evaluation. We test our method on synthetic data achieving high accuracy and quick adaptation to previously unseen causal models.

1 INTRODUCTION

While neural networks have very powerful learning capabilities and have achieved impressive empirical success on many tasks across a wide range of domains, e.g. computer vision (Krizhevsky et al., 2012; Simonyan & Zisserman, 2014; Szegedy et al., 2015; He et al., 2016), natural language processing (Sercu et al., 2016; Child et al., 2019) and reinforcement learning (Silver et al., 2016; Moravčík et al., 2017), their success crucially hinges on the assumption of independently and identically distributed (iid) data. Furthermore, most machine learning methods, especially those relying on neural networks learn based on associations present in the data but do not distinguish between correlation and causation. This is a problem when the iid assumption does not hold, for example in domain adaptation, transfer learning and meta-learning (Dasgupta et al., 2019; Bengio et al., 2019; Ke et al., 2019). In order to tackle these domains, and to correctly handle interventional data, we need to capture the causal mechanisms in the underlying generative process. Thus, learning causal models is an important step towards expanding the abilities of our algorithms.

In this paper, we focus on learning causal models from interventional data. Previous work in this area often focuses on learning representations for causal models in terms of Bayesian networks, e.g. (Heckerman et al., 1995). While such approaches provide an easily interpretable output, they rely on iterating over all possible graph structures and thus scale poorly as the number of possible graphs scales super exponentially with the number of variables. We propose a paradigm shift away from learning Bayesian networks to learning continuous representations of causal graphs using deep neural networks. In particular, we examine a meta-learning setup where learning is divided into episodes. At the beginning of each episode, a causal graph that governs the data generation during that episode is drawn. At every time step t, $t \leq k$, our algorithm receives the node values of a sample under intervention. Specifically, the intervention is picked at random, and the algorithm inputs both what variable the intervention affected, the value of the intervention, as well as the values consequently assumed by the remaining variables. Thus, during an episode of length k our algorithm observes k samples from the underlying causal graph under different interventions. To facilitate learning, we ask our model to predict outcomes of interventions at every time step. Note that interventions in our framework correspond to hard interventions, i.e. an intervention constitutes a hard assignment of a variable to a specific value. In order to enable learning of causal graphs in this setup, we propose a novel neural network architecture that efficiently learns to encode causal graphs in a compact, continuous representation.

When learning causal graphs explicitly as discrete structures, it is straightforward to evaluate accuracy by counting how many of the inferred edges match the ground truth. Our implicit, continuous representation of a causal graph cannot be directly be compared against the ground truth. One possibility is to compare the predictions of node values under all possible interventions, rather than comparing the edge structure. Unfortunately, comparing node predictions scales poorly and would be prohibitively expensive even for medium-sized graphs. Therefore, we instead learn a decoder from the continuous representation into a graph, and then measure its edge accuracy.

The main contribution of this paper is to introduce a new framework for learning to learn causal discovery models using continuous representations via unsupervised losses. We also introduce a relational encoding-decoding network that performs well in this task relative to a natural baseline.

2 Related work

Most approaches to learning causal models, i.e. causal induction, can be grouped into three categories – constraint-based, score-based and asymmetry-based methods. Constraint-based methods try to infer the a directed acyclic graph (representing the causal structure) by analyzing conditional independencies in the data (Spirtes et al., 2000; Sun et al., 2007; Zhang et al., 2011). These methods are mathematically well grounded, but their performance is critically dependent on the conditional independence methodology used and sensitive to the amount of data available. Score-based methods search the space of all DAGs by scoring the data fit using a predefined function. Examples of this approach include (Chickering, 2002; Tsamardinos et al., 2006; Goudet et al., 2017). Unfortunately, this class of methods does not scale well as the search space grows super-exponentially with the number of variables. Furthermore, there are not principled ways of selecting problem-specific scoring functions and search heuristics. Asymmetry-based methods assume that there is an inherent asymmetry between cause and effect, and try to infer the underlying causal structure by measuring this asymmetry. Examples of this class of methods includes (Shimizu et al., 2006; Hoyer et al., 2009; Daniusis et al., 2012; Budhathoki & Vreeken, 2017; Mitrovic et al., 2018). The work of (Peters et al., 2016) and use invariance

While most causal induction deal with purely observational data, a number of methods has been proposed for the setting when interventional data is available. For example, (Heckerman et al., 1995; Cooper & Yoo, 1999) have proposed score-based methods that explicitly take into account interventional data. The works from (Peters et al., 2016) and (Heinze-Deml et al., 2018) exploit invariances across enviroments to learn the causal structure. However, it is difficult to scale to large size graphs due to the super exponetial search through all possible graphs.Furthermore, there has been extensive research around the selecting informative interventions for causal discovery, e.g. (He & Geng, 2008; Kocaoglu et al., 2017; Zhu & Chen, 2019). The work in (Bengio et al., 2019) proposes a meta-learning framework for learning causal models from interventional data. However, it is also non-trivial to scale the proposed method in (Bengio et al., 2019) to any reasonable sized graphs and the experiments were performed on models with 2 variables only. The work of (Ke et al., 2019) extends that of (Bengio et al., 2019) to the case of larger variable graphs. However, each meta-sample is generated from an interventional distribution based on a single ground-truth structured causal model (SCM). Hence, the model learns only a single SCM at a time. This is in comparison with our model, which could adapt and learn a completely new SCM within each meta-sample. Most similar to our work, (Dasgupta et al., 2019) uses meta-learning to learn to make predictions under interventions that suggest (but do not prove) that some causal reasoning is occurring. The difference is that their approach does not induce a causal graph, neither explicitly nor via a decoding, thus it cannot be used for general causal discovery, but merely to make predictions of variable values.

3 Causal relational networks

In this paper, we propose a general framework for learning causal models from intervention data leveraging the power of neural networks. Unlike most previous methods that learn Bayesian network structure, we employ continuous representations to represent causal graphs. Furthermore, we train our model on many distinct causal graphs which allows it to learn new causal graphs very quickly. In order to evaluate the effectiveness of our method, we measure the ability of our model to predict the outcome of interventions as well as decode the structure of the underlying causal Bayes network.

3.1 PROBLEM FORMULATION

We propose a meta-learning setup to learn causal graphs. Here, training occurs in episodes, with a single ground-truth causal graph for the duration of each episode. Episodes have a fixed length k and at each time step the learner receives a sample of node values from the causal graph. Thus, in each episode, we first draw a new causal graph G, followed by generating k samples from this graph under random interventions. The resulting intervention samples x are fed into the model sequentially. At each time step t, the model is asked to predict the intervention outcome given the intervention (which variable has been intervened on and what was the value of the intervention).

In order to evaluate the performance of our model, we cannot directly compare the inferred graph structure to the ground truth as is usually done when learning Bayesian networks given that we have a continuous representation that implicitly encodes the underlying causal graph. Given that evaluating the predictions under all possible interventions does not scale to large graphs and state spaces, we propose to use a decoding-based evaluation metric for continuous representations of causal graphs, similar to (Gregor et al., 2019). Specifically, we propose to evaluate if the discrete graph structure is encoded in the continuous representation by training a neural network to decode it. Thus, at every time step t, we feed the hidden state of our model into a decoder trained to output the right causal graph. Note that we do not pass gradients from the ground-truth supervision signal back into the decoder.

3.2 MODEL DESCRIPTION

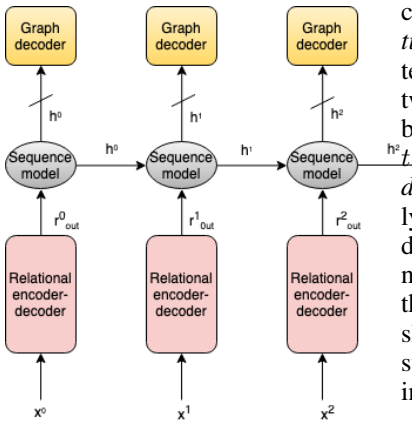

Figure 1: Causal relational model with all components.

Our learning architecture named *causal relational networks* comprises three modules shown in Figure 1. First, the *relational encoder-decoder* models the observed data through a latent representation and serves to preserve the relationship between variables. Second, the *sequence model* computes the belief state at every time step t, such that the belief state at t models the available information up to t. Third, the *graph decoder* outputs the estimated adjacency matrix of the underlying causal graph. As in a meta-learning setup, training is divided into episodes. Each episode begins with sampling a new causal graph. Data from each episode is sampled from the same causal graph. The samples are read as a sequence as shown in Figure 1. For a causal graph of size n, at each time step, one sample vector of size n along with the intervention information is passed into the model.

Relational encoder-decoder. At time step t, the encoder input x^t is a vector length $n+1$, where n is the number of variables in the ground-truth causal graph. The first element specifies what variable was intervened upon, by providing the index of the variable. The remaining n elements in the vector contain the sample from the intervened-on graph.

In our initial experiments, we have observed that a fully connected encoder (i.e. using just an LSTM, in section 4.3) does not manage to learn the causal relationships between variables. To tackle this, we propose a novel encoder architecture inspired by relational networks (Santoro et al., 2018). In particular, this architecture more appropriately models causal relationships between variables because it is trained to predict, given a particular variable in the sample, the value of all other variables.

Within an episode, let $x_0^t, x_1^t, ..., x_n^t$ denote the samples from the causal graph under different interventions, with x_i^t the value of the i-th variable in the t-th sample. For every variable, we instantiate a feedforward neural network that takes the values of this variable as input. Note that there are n MLPs in the model that are identical in shape but do not share parameters (see Figure 2). The task of each MLP is to predict the value of the remaining variables (i.e. output size $n-1$). Specifically, let f_k be the MLP that takes as input x_i^t, i.e. $x_0, x_1, ...x_{k,k\neq i} = f_k(x_i)$.

Using the intuition that a variable can usually better predict direct parents or children than unconnected variables, we augment our architecture with an attention model. Thus, for a given variable x_i^t, the MLP returns $n-1$ predictions (for all $x_{k,k\neq i}$), which are fed into the *attention MLP* (see Figure 2). This then outputs attention weight $w_{i,k}^t$, with k denoting the prediction of variable X_i^t from f_k. Next, we use the attention weights to compute the intermediate prediction O_i^t, where this prediction is being passed through 2 separate MLPs. The first MLP outputs the final predictions of the variable X_i^t. The input to the second MLP is a concatenation of O_i^t where $0 \leq i < n$ to form O^t. The second MLP then outputs the encoding r^t which is passed to the sequence model. Taking all of this together, the MLPs together with the attention weight encode the relationship between a single variable and the rest of the causal graph. There are 2 outputs for the encoder, one that predicts the intervention $\bar{X}_0^t, ..., \bar{X}_{n-1}^t$ and the other one r^t that is passed to the *sequence model* that is used for decoding the causal graph.

Sequence model. We use a sequence model to accumulate information produced by the encoder and update the belief state of the model. In our setup, we use a simple cumulative summation operator that adds current information to already accumulated information. Specifically, we take r^t that encode the observed data using the *Relational encoder-decoder*. Then the hidden state h^t of the sequence model at time t is updated as $h^{t+1} = h^t + r^t$. This simple additive update guarantees that the final representation h^t is invariant with respect to permutations in the order of presentation of the data points. A more complex recurrent model, such as LSTM ((Hochreiter & Schmidhuber, 1997)), could also be used but it is not trivial to preserve permutation invariance in this case.

Graph decoder. Having computed the belief state h^t, we use a LSTM (Hochreiter & Schmidhuber, 1997) conditioned on h^t to predict the adjacency matrix of the the causal graph. We also ensure that no gradient with supervision signal are passed back into the other components of the model.

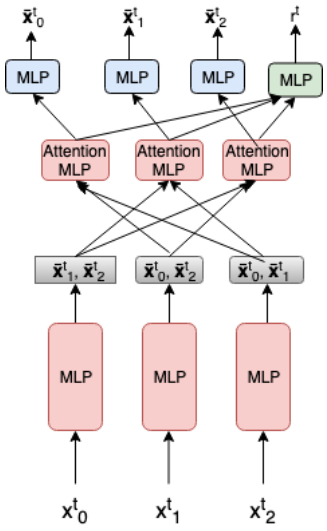

Figure 2: Relational encoder model for inputs with 3 variables.

4 EXPERIMENTS

4.1 BINARY CAUSAL BAYESIAN NETWORKS

We explore the setting of binary Bayesian networks with all variables observed such that there is no confounder. We focus on this setting as a proof of concept for our methodology; extensions to categorical Bayesian networks are straightforward. The ground-truth model defined a joint $p(X = x) = \prod_i p\left(X_i = x_i \mid x_{<i}, w_i\right)$ is formulated as follows,

$$p\left(X_i = x_i \mid x_{<i}, w_i\right) = \text{sigmoid}\left[w_i \beta \left(\sum_{j<i} U_{j,i} M_{j,i} x_j + b_i\right)\right], \tag{1}$$

where $w_i \in \{0, 1\}$ is an intervention mask sampled randomly ($w_i = 0$ corresponds to an intervention replacing the node distribution by an uniform distribution), $U_{i,j} \in \{-1, 1\}$ is the coupling potential, $b_i \in \{1 - \sum_{j<i} M_{ji}, \sum_{j<i} M_{ji} - 1\}$ is an integer potential bias, $M_{i,j} \in \{0, 1\}$ is the connectivity mask representing the Bayesian net causal graph, and $\beta = 5$ is a concentration parameter. This model simulates a Bayesian network over binary variables with randomized interventions. It can be thought of as a soft random Boolean network. For the ease of sampling, we restrict the ground-truth model to a directed acylic graph (DAG). We have fixed the number of variables in the graph to 5 for the experiments.

4.2 OUR MODEL

We perform the same hyper-parameter search for our baselines as well as our model. We use a batch size of 8, learning rate of 0.0002, all models are trained with the Adam optimizer (Kingma & Ba,

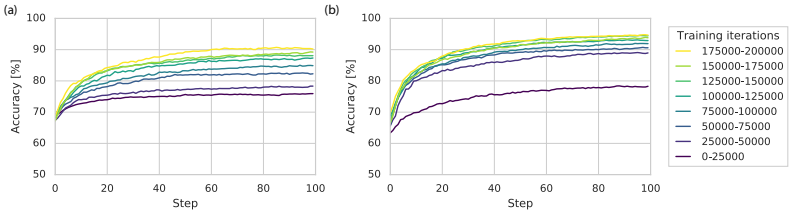

Figure 3: Accuracy for edge probability over the course of the episode, binned by different stages of training, for **a)** our causal relational network (CRN) model, and **b)** CRN with supervised graph learning.

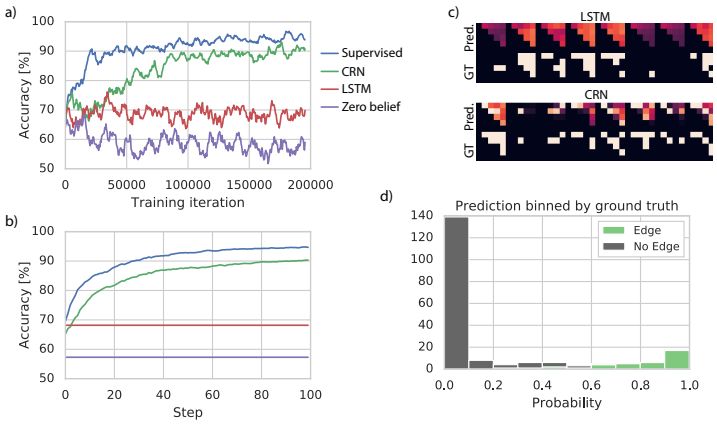

Figure 4: Comparing performance for all models. **a)** Accuracy of edge prediction after seeing the entire episode of samples, over the course of training. **b)** Edge accuracy over the course of the episode, averaged over 50 episodes at the end of training. **c)** Predicted causal graph (adjacency matrix format) vs. ground-truth graph for the (upper) LSTM baseline and (lower) CRN. **d)** Histogram of predicted edge weights at end of episode for ground-truth edges of weight = 1 vs. weight = 0, CRN model. CRN = Causal relational network; zero belief = belief state set to 0.

2014) for 20000 iterations. For the relational encoder, we set the MLP to have 2 hidden layers of size 128 and 64 with a ReLU (Nair & Hinton, 2010) in between. The relational encoder final output size is 64. The graph decoder LSTM has a hidden state size of 128.

We evaluate our model on the accuracy of the decoded graph with respect to the ground-truth graph. The accuracy is defined to be the percentage of decoded edges that align with the ground-truth edge. The edge is said to be aligned with the ground-truth edge if both of the *argmax* of the Bernoulli distribution over the learned edge and the ground-truth edge matches.

The edge accuracy for our model across the course of an episode, for different stages of training, is shown in Figure 3a. The model has learned to extract information about the causal graph structure from intervention samples, reflected in the improvements in edge accuracy as the model observes more intervention data from the causal graph. In the beginning of the episode, the model only knows the underlying causal graph is a DAG. As the model observes more intervention samples from the same graph, it can quickly infer underlying causal graph. Note that the edge accuracy also improves across training iterations.

4.3 BASELINES

There are 3 main components to our model as shown in Figure 1. First, we evaluate our encoder-decoder model against an LSTM with a naive fully-connected MLP. We then evaluate the amount of performance gain the model could achieve by receiving supervision signal from our graph decoder. Lastly, in order to verify that the information for the conditional graph decoder for decoding the graph is indeed contained in our model, we compare the performance of our model to one with zero information in its input to the graph decoder. The sequence model that accumulates information

across time is fairly straightforward and hence we have not performed comparisons against this component.

LSTM baseline. Our model uses a relatively sophisticated relational network to extract information in the input. In order to verify the necessity of this component, we compare our encoder to a LSTM with naive fully connected feed-forward neural network (MLP). For simplicity, we call this the LSTM baseline. For this baseline, we no longer need a decoder as the encoder-decoder could just be an identity model. As there are no decoders, the baseline is missing the unsupervised learning signal. To compensate for this, we pass the gradients from the supervised graph decoder to the encoder. The baseline model has an easier job, since it could learn directly from supervision signal. Interestingly, as shown in Figure 4, though the baseline model has an easier task, the edge prediction accuracy is only around 70% throughout training as compared to the 90% reached by our model as shown in Figure 4. The baseline model neither learns within an episode (as it observes more samples), nor does it learn with more training.

Supervised graph learning. We next evaluate the amount of performance gain the model gets by receiving supervision signal from the graph decoder. This is achieved by removing the stop gradient from the belief state that is passed in as inputs to the graph decoder. The results are shown in Figure 3b. The model with supervision signal achieved 95% accuracy as compared to the 90% achieved by the unsupervised model.

Zero belief state baseline. Lastly, we verify that the belief state in our model contains useful information for decoding the structured causal graph. This baseline uses the same relational encoder-decoder and graph decoder from our original model; the only difference is that the belief state is set to zero. If this baseline could also learn to decode the structured causal graph, it would indicate that it was not the information in the belief state that helped to decode the structured causal graph, it was rather the graph decoder itself. Figure 4 shows that the decoder was not in fact able to decode the causal graph structure. The accuracy remains around and below 60% throughout training compared to the 90% that our model achieved at the end of training (Figure 3a), indicating that the belief state was crucial to decoding the causal graph.

5 Conclusion

In this paper, we presented a new framework for learning causal models using neural networks. In particular, we have introduced a novel neural architecture, causal relational networks, and proposed to use a decoding-based metric for evaluation. We have tested our methodology on synthetic data and have shown that it can accurately learn causal structures. Furthermore, we have shown that causal relational networks can very quickly learn new causal models.

References

Yoshua Bengio, Tristan Deleu, Nasim Rahaman, Rosemary Ke, Sébastien Lachapelle, Olexa Bilaniuk, Anirudh Goyal, and Christopher Pal. A meta-transfer objective for learning to disentangle causal mechanisms. *arXiv preprint arXiv:1901.10912*, 2019.

Kailash Budhathoki and Jilles Vreeken. Causal inference by stochastic complexity. *arXiv:1702.06776*, 2017.

David Maxwell Chickering. Optimal structure identification with greedy search. *Journal of machine learning research*, 3(Nov):507–554, 2002.

Rewon Child, Scott Gray, Alec Radford, and Ilya Sutskever. Generating long sequences with sparse transformers. *arXiv preprint arXiv:1904.10509*, 2019.

Gregory F. Cooper and Changwon Yoo. Causal Discovery from a Mixture of Experimental and Observational Data. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, UAI'99, pp. 116–125, San Francisco, CA, USA, 1999.

Povilas Daniusis, Dominik Janzing, Joris Mooij, Jakob Zscheischler, Bastian Steudel, Kun Zhang, and Bernhard Schölkopf. Inferring deterministic causal relations. *arXiv preprint arXiv:1203.3475*, 2012.

Ishita Dasgupta, Jane Wang, Silvia Chiappa, Jovana Mitrovic, Pedro Ortega, David Raposo, Edward Hughes, Peter Battaglia, Matthew Botvinick, and Zeb Kurth-Nelson. Causal reasoning from meta-reinforcement learning. *arXiv preprint arXiv:1901.08162*, 2019.

Olivier Goudet, Diviyan Kalainathan, Philippe Caillou, Isabelle Guyon, David Lopez-Paz, and Michèle Sebag. Causal generative neural networks. *arXiv preprint arXiv:1711.08936*, 2017.

Karol Gregor, Danilo Jimenez Rezende, Frederic Besse, Yan Wu, Hamza Merzic, and Aaron van den Oord. Shaping belief states with generative environment models for rl. *arXiv preprint arXiv:1906.09237*, 2019.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

Yang-Bo He and Zhi Geng. Active learning of causal networks with intervention experiments and optimal designs. *Journal of Machine Learning Research*, 9(Nov):2523–2547, 2008.

David Heckerman, Dan Geiger, and David M Chickering. Learning bayesian networks: The combination of knowledge and statistical data. *Machine learning*, 20(3):197–243, 1995.

Christina Heinze-Deml, Jonas Peters, and Nicolai Meinshausen. Invariant causal prediction for nonlinear models. *Journal of Causal Inference*, 6(2), 2018.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8): 1735–1780, 1997.

Patrik O Hoyer, Dominik Janzing, Joris M Mooij, Jonas Peters, and Bernhard Schölkopf. Nonlinear causal discovery with additive noise models. In *Advances in neural information processing systems*, pp. 689–696, 2009.

Nan Rosemary Ke, Olexa Bilaniuk, Anirudh Goyal, Stefan Bauer, Hugo Larochelle, Chris Pal, and Yoshua Bengio. Learning neural causal models from unknown interventions. *arXiv preprint arXiv:1910.01075*, 2019.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Murat Kocaoglu, Karthikeyan Shanmugam, and Elias Bareinboim. Experimental design for learning causal graphs with latent variables. In *Advances in Neural Information Processing Systems*, pp. 7018–7028, 2017.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.

Jovana Mitrovic, Dino Sejdinovic, and Yee Whye Teh. Causal inference via kernel deviance measures. In *Advances in Neural Information Processing Systems*, pp. 6986–6994, 2018.

Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisỳ, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 356(6337):508–513, 2017.

Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814, 2010.

Jonas Peters, Peter Bühlmann, and Nicolai Meinshausen. Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(5):947–1012, 2016.

Adam Santoro, Ryan Faulkner, David Raposo, Jack Rae, Mike Chrzanowski, Theophane Weber, Daan Wierstra, Oriol Vinyals, Razvan Pascanu, and Timothy Lillicrap. Relational recurrent neural networks. In *Advances in Neural Information Processing Systems*, pp. 7299–7310, 2018.

Tom Sercu, Christian Puhrsch, Brian Kingsbury, and Yann LeCun. Very deep multilingual convolutional neural networks for lvcsr. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4955–4959. IEEE, 2016.

Shohei Shimizu, Patrik O Hoyer, Aapo Hyvärinen, and Antti Kerminen. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7(Oct):2003–2030, 2006.

David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484, 2016.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

Peter Spirtes, Clark Glymour, Richard Scheines, et al. Causation, prediction, and search. *MIT Press Books*, 2000.

Xiaohai Sun, Dominik Janzing, Bernhard Schölkopf, and Kenji Fukumizu. A kernel-based causal learning algorithm. In *Proceedings of the 24th international conference on Machine learning*, pp. 855–862. ACM, 2007.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

Ioannis Tsamardinos, Laura E Brown, and Constantin F Aliferis. The max-min hill-climbing bayesian network structure learning algorithm. *Machine learning*, 65(1):31–78, 2006.

K Zhang, J Peters, D Janzing, and B Schölkopf. Kernel-based conditional independence test and application in causal discovery. In *Proceedings of the 27th Annual Conference on Uncertainty in Artificial Intelligence (uai)*, 2011.

Shengyu Zhu and Zhitang Chen. Causal discovery with reinforcement learning. *arXiv preprint arXiv:1906.04477*, 2019.