Application and Implementation of different deep learning

Jie Zou 12 hours ago · 9 min read

D**eep learning** is a part of machine learning that learns data representations. Many of the fundamental tasks (with appropriate reformulation) in deep learning are relevant to a much broader array of domains, and in particular, have tremendous potential in aiding the investigation of central scientific questions. There are many different Applications and Implementation of deep learning in different fields.

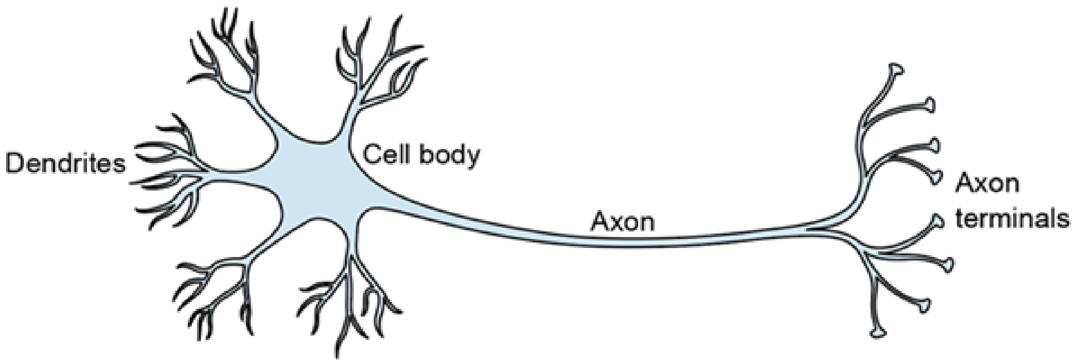
The Application and Implementation of deep learning in more and more wise fields is a hot topic in machine learning, neural networks, and deep learning. There have been many relevant studies on how to correctly select a deep learning approach. The story aims to propose a short story about the application and implementation of different deep learning. According to the author of the article "A Survey of Deep Learning for Scientific Discovery ", numerous studies have been published resulting in various models in the field of deep learning (DL) has recently begun to receive Widely concerned, this is mainly due to its better performance than the classic model. In recent years, most of the fundamental breakthroughs in the core problems of machine learning have been driven by the progress of deep neural networks.

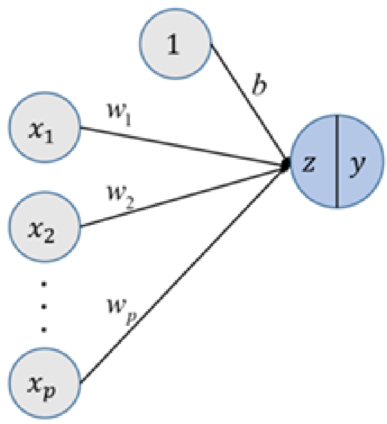
To supplement the information in this short story, other resources explaining the application and implementation of deep learning in different fields will be used. In addition, other sources will be used to compare the application of the deep learning (DL) field to other studies on similar topics. For example, the short story will include other papers in this field, such as the newest application and implementation of Meta-Learning, to compare with other existing studies.

The short story will begin with the introduction of Deep Learning, then briefly introduce existing research on the application, implementation of DL that has been published for several years, and the emerging technologies for DL. While discussing the overall quality of the application and implementation, this short story will also explain the whole situation of DL application and the source papers have used for other research, as well as the critique of the paper, what can be done to improve and future directions and suggestions.

What?

> *What is Deep Learning?*





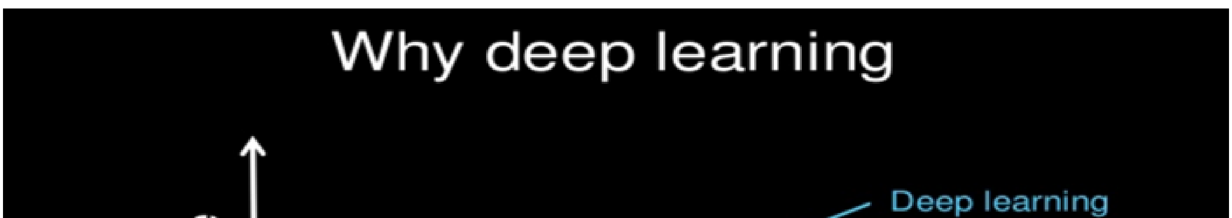
Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called **artificial neural networks**.

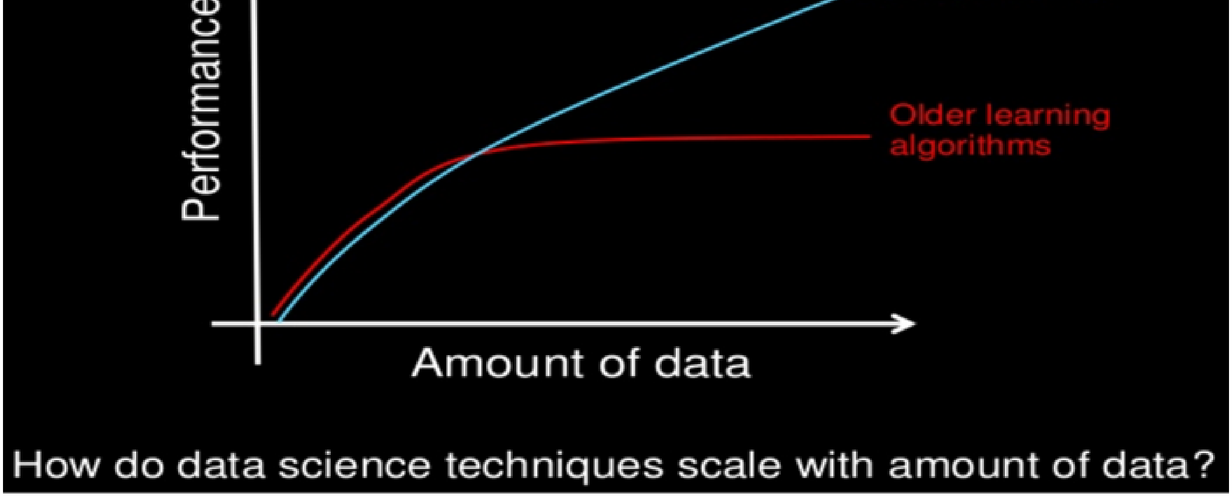
"Deep learning algorithms seek to exploit the unknown structure in the input distribution in order to discover good representations, often at multiple levels, with higher-level learned features defined in terms of lower-level features"

> ***Deep Learning is Hierarchical Feature Learning** or* **layered representations learning** *and* **hierarchical representations learning**
>
> Source： https://machinelearningmastery.com/what-is-deep-learning/

Why?

> *In the field of machine learning(ML), Numerous studies have been published resulting in various models. And deep learning (DL) has recently begun to receive Widely concerned, this is mainly due to its better performance than the classic model.*





This Photo by Unknown Author is licensed under CC BY-SA

Most of the fundamental breakthroughs in the core problems of machine learning have been driven by the progress of deep neural networks.

- **Prediction Problems**: the most straightforward way to apply deep learning
- **From Predictions to Understanding**: One fundamental difference between scientific questions and core machine learning problems.
- **Complex Transformations of Input Data**: In many scientific fields, the amount of generated data has increased dramatically, and deep learning can provide efficient analysis and automated processing.

Deep Learning can be used to used in all of these scientific fields efficiently!

Generally, the more data you provide, the more accurate the results in Deep Learning. Coincidentally, there are huge data sets in the scientific field, such as PB-level data of animal experiments, physical and chemical experiments, patient data, papers, and reports. But sometimes we need '**Doing More with Less Data'.** We will introduce them below:

How?

> • *Which DL models are preferred (and more successful) in different applications?*
> • *How do DL models pare against traditional soft computing / ML techniques?*
> • *What is the future direction for DL research in scientific?*
> *Our focus was on DL implementations for scientific applications*

Existing research on Deep Learning

In the Survey "A Survey of Deep Learning for Scientific Discovery", they tried to provide a state-of-the-art snapshot of the developed DL models for scientific applications, as of today. They not only categorized the works according to their intended subfield in scientific but also analyzed them based on their DL models. In addition, they also aimed at identifying possible future implementations and highlighted the pathway for the ongoing research within the field.

And I will introduce other studies of the applications of DL in the scientific discovery field on similar topics, such as:

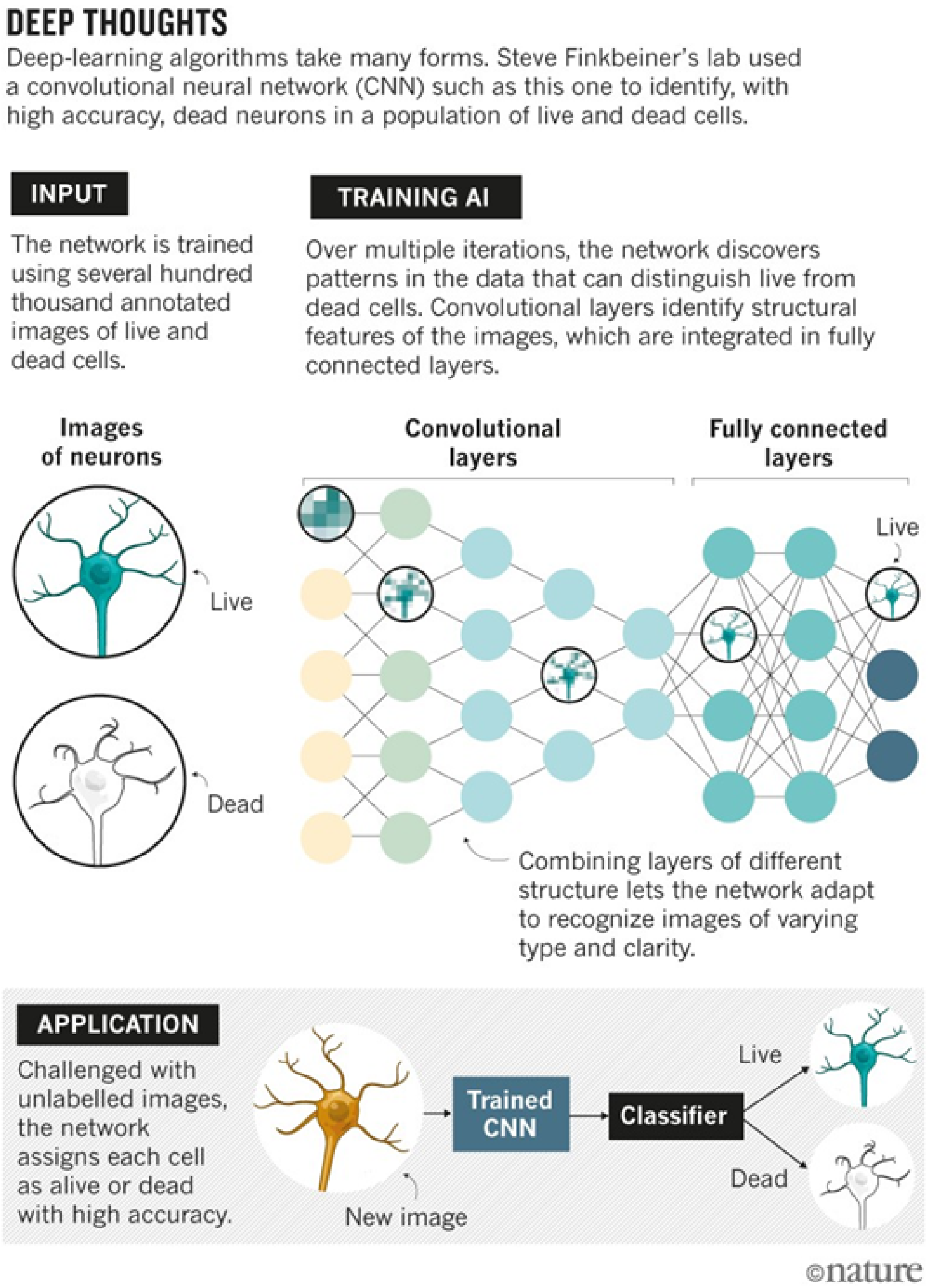
— Deep learning applications in biological research

— The application of deep learning in medicine

— Set up structured data, such as organic compounds

— Model different invariances-apply the invariance of specific Lie groups to molecular property prediction.

These topics involve a very wide range of research areas, it is hoped that the materials and references provided can help inspire novel contributions to these very exciting and rapidly developing research directions.

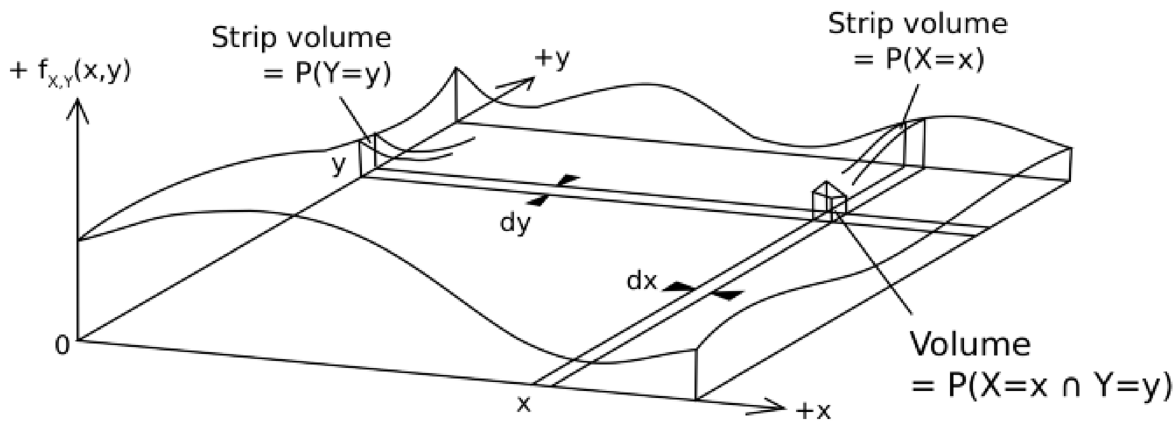


Source: Jeremy Linsley/Drew Linsley/Steve Finkbeiner/Thomas Serre

Bayes’ theorem

link: here

The basis of price prediction is probability theory and statistics, one of the most important assumptions is the Bayers’ theorem.



$$P(Y=y|X=x) = \frac{P(X=x \cap Y=y)}{P(X=x)} \qquad P(X=x|Y=y) = \frac{P(X=x \cap Y=y)}{P(Y=y)}$$

Deep Learning Libraries and Resources that the source papers have used for their research.

- Software Libraries for Deep Learning

— **PyTorch** : with a high-level API called Lightning

— **TensorFlow**: offers Keras as a high-level API

— **fast.ai:** https://course:fast:ai/,https://github:com/fastai/

http://www:cs:toronto:edu/~rgrosse/courses/csc421_2019/

https://paperswithcode:com/sota

- Research Overviews, Code, Discussion

https://paperswithcode:com/

http://www:arxiv-sanity:com/top

https://www:reddit:com/r/MachineLearning/

https://www:paperdigest:org/

https://www:ipam:ucla:edu/programs/workshops/deep-learning-and-medical-applications/?tab=schedul

- Models, Training Code, and Pretrained Models:

(i) Pytorch and TensorFlow have a collection of pretrained models, found at https://github:com/tensorflow/models and https://pytorch:org/docs/stable/torchvision/models:html.

(ii) https://github:com/huggingface Hugging Face (yes, that really is the name), offers a huge collection of both pretrained neural networks and the code used to train them. Particularly impressive is their library of Transformer models, a one-stop-shop for sequential or language applications.

(iii) https://github:com/rasbt/deeplearning-models offers many standard neural network architectures,including multilayer perceptrons, convolutional neural networks, GANs and Recurrent Neural Networks.

(iv) https://github:com/hysts/pytorch_image_classification does a deep dive into image classification architectures, with training code, highly popular data augmentation techniques such as cutout, and careful speed and accuracy benchmarking. See their page for some object detection architectures also.

(v) https://github:com/openai/baselines provides implementations of many popular RL algorithms.

(vi) https://modelzoo:co/ is a little like paperswithcode, but for models, linking to implementations of neural network architectures for many different standard problems.

(vii) https://github:com/rusty1s/pytorch_geometric. Implementations and paper links for many graph neural network architectures.

- Data Collection, Curation, and Labelling Resources

https://github:com/tzutalin/labelImg

https://github:com/wkentaro/labelme

https://rectlabel:com/

https://github:com/doccano/doccano

- Visualization, Analysis, and Compute Resources

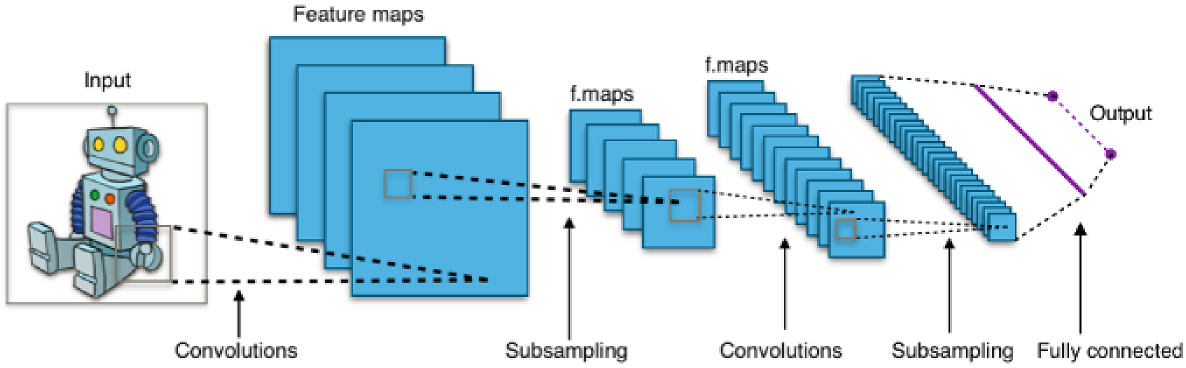
When training deep neural network models, we can visualize important metrics such as loss and accuracy while training.

— **Tensorboard**: https://www:tensorflow:org/tensorboard

— https://colab:research:google:com/notebooks/welcome:ipynb

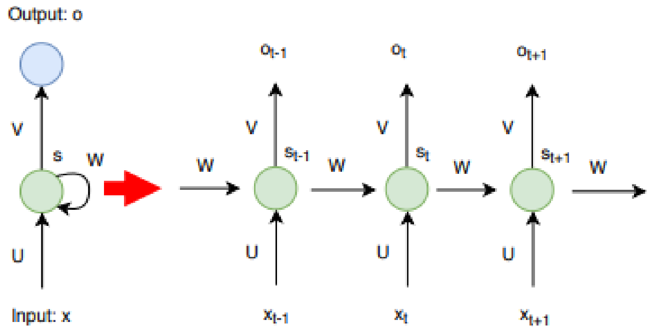
The proposed deep learning (DL) models that the source papers have used for their research.

- **Convolutional neural network(CNN)**



Typical CNN architecture

- **Recurrent Neural Networks (RNN)**

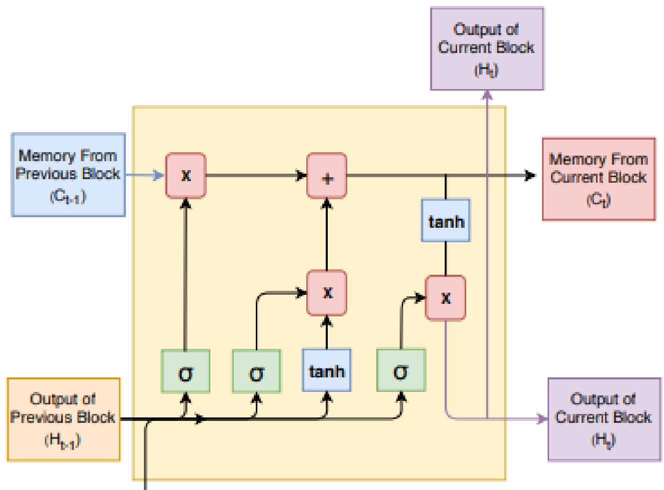


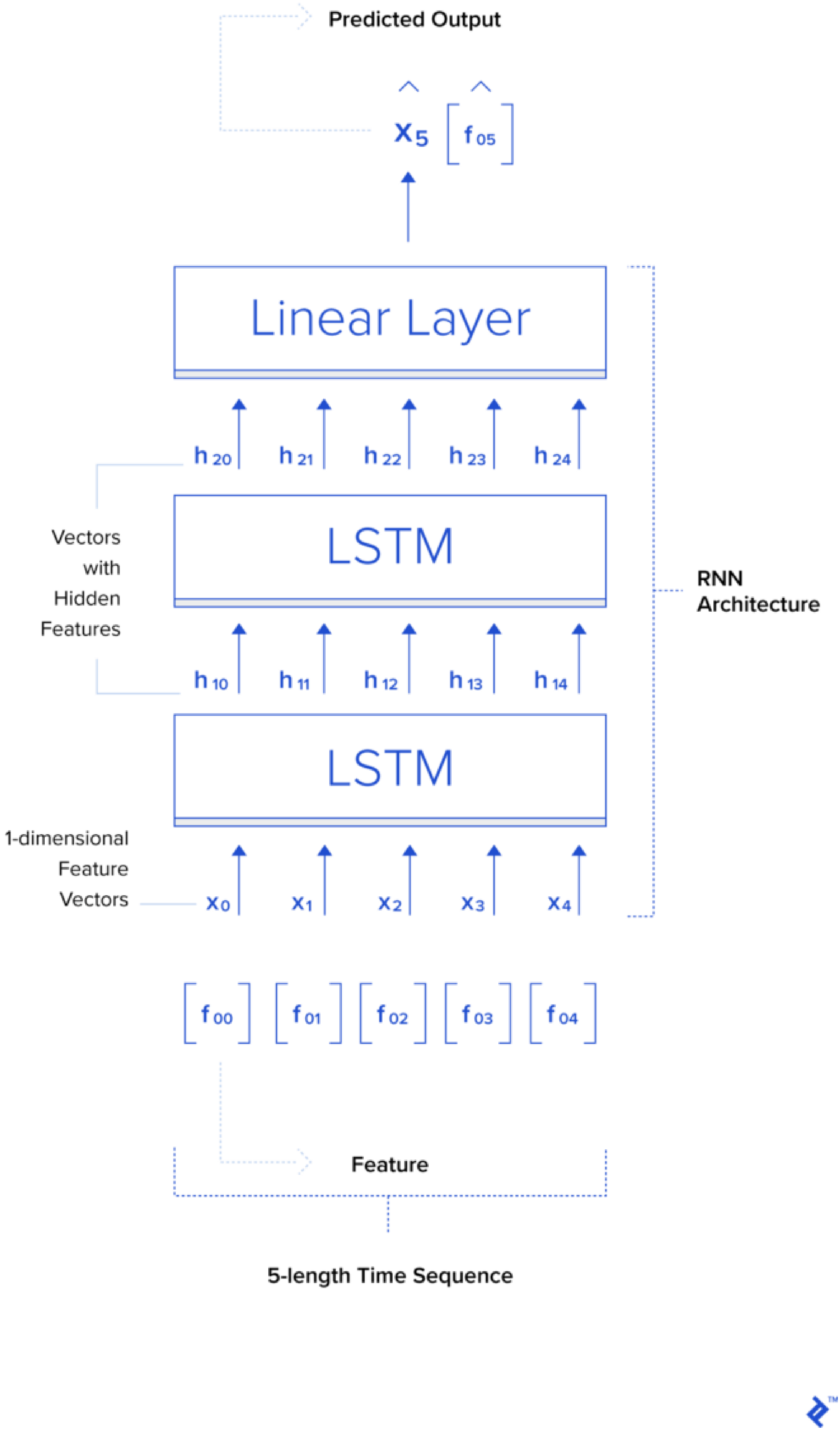
RNN

In contrast to feed-forward neural networks, in recurrent neural networks, data can flow in any direction. They can learn the time series dependency well.

- **Long Short Term Memory (LSTM)**

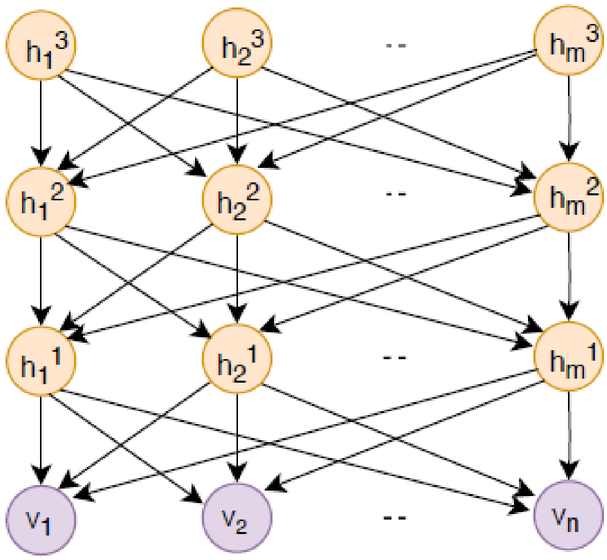
LSTM are also good at learning from sequential data, i.e. time series.





- **Deep Belief Networks (DBNs)**

DBN is a type of ANN that consists of a stack of RBM layers. DBN is a probabilistic generative model that consists of latent variables.



Deep Belief Networks (DBNs)

CNN should be the main method.

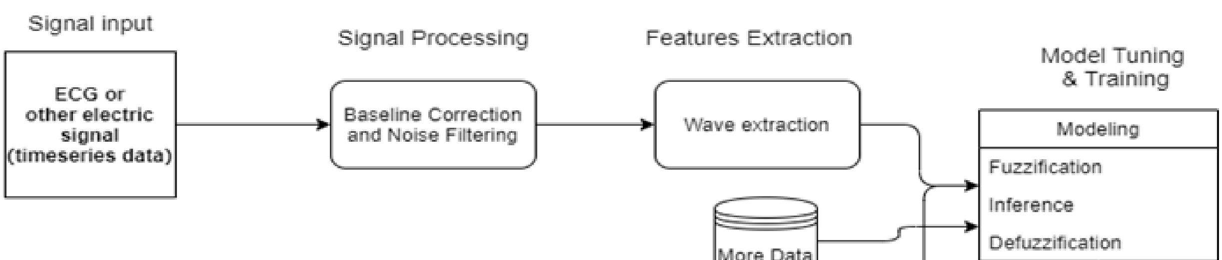
Key Methods (Supervised Learning)

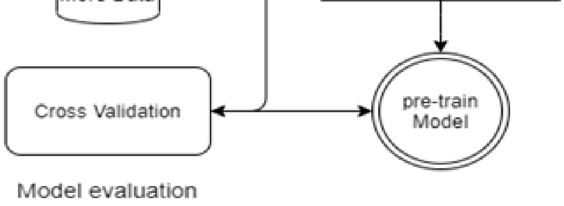
- **Transfer Learning**
- **Domain Adaptation**
- **Multitask Learning**
- **Weak Supervision (Distant Supervision)**

Transfer learning is a classic scientific research model using deep learning, and a similar model will be used in our master project (AI for healthcare).

Instead of randomly initializing parameters and directly training the target task, we first perform pre-training steps on some diversified general tasks. This leads to the neural network parameters converges to a set of values, called pre-training weights. If the pre-training tasks are sufficiently diverse, these pre-training weights will contain useful functions, which can be used to learn the target task more effectively. Starting with the pre-trained weights, we then train the network on the target task (called fine-tuning), which provides us with the final model.

Pre-train model Architecture





Instead of randomly initializing parameters and directly training the target task, we first perform pre-training steps on some diversified general tasks. This leads to the neural network parameters converges to a set of values, called pre-training weights. If the pre-training tasks are sufficiently diverse, these pre-training weights will contain useful functions, which can be used to learn the target task more effectively. Starting with the pre-trained weights, we then train the network on the target task (called fine-tuning), which provides us with the final model.

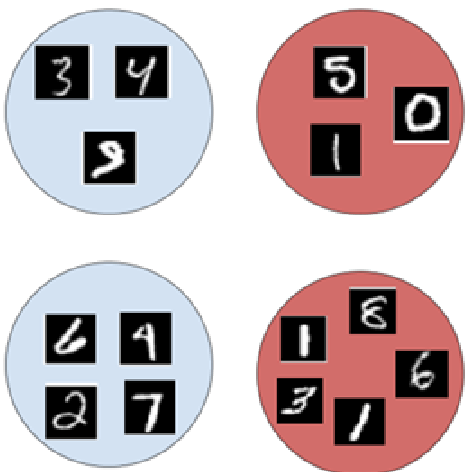
But sometimes we need — -

Doing More with Less Data

We can implement the techniques from below:

- Self-Supervised Learning
- Semi-Supervised Learning
- Co-training
- Data Augmentation
- **MIL**
- **GAN**
- **Meta-learning**

MIL(**Multi-instance learning)** is a supervised learning problem where individual examples are unlabeled; instead, bags or groups of samples are labeled.[5]



It is used in the medical field, drug activity prediction, text categorization, image retrieval, and classification.[6]

Meta-learning, or Learning to Learn, has become another important research branch after Reinforcement Learning (hereinafter referred to as Meta-Learning).

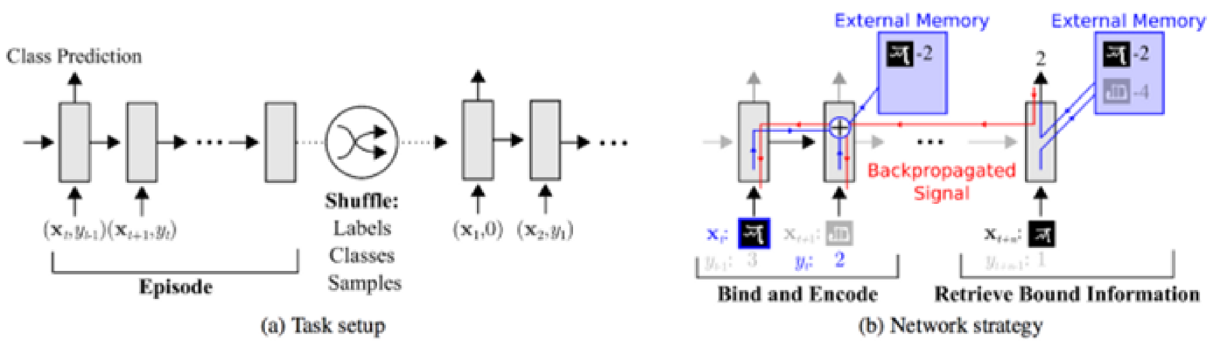
The theoretical research on artificial intelligence has shown this trend:

> *Artificial Intelligence → Machine Learning → Deep Learning → Deep Reinforcement Learning → Deep Meta-Learning*

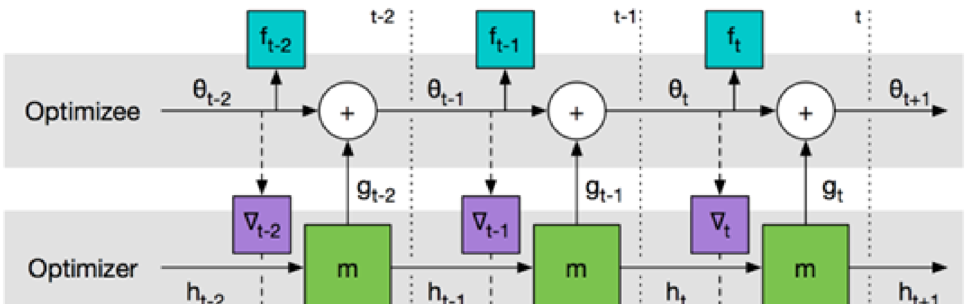
There are many ways to implement the Meta-learning now:

- Memory-based approach
- Method based on predicted gradient
- Method of using an Attention attention mechanism
- Learn from the LSTM method
- Meta-Learning Method for RL
- Through the method of training a good base model, and simultaneously applying to supervised learning and reinforcement learning
- Ways to use WaveNet
- Ways to predict Loss

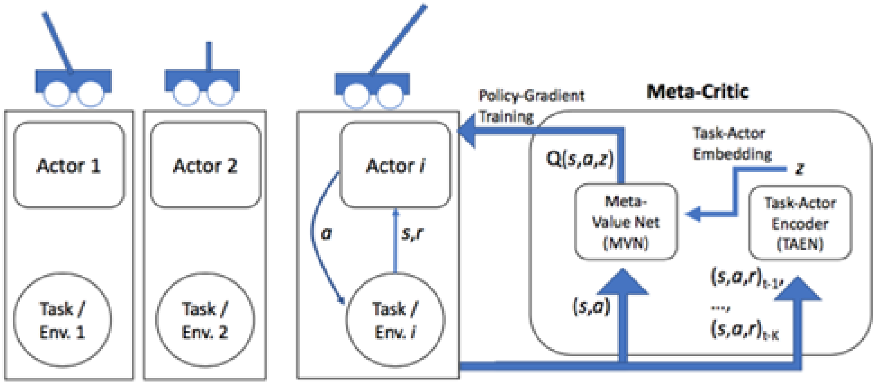
Meta-Learning is in the ascendant, and various magical ideas emerge endlessly.



Meta-learning with memory-augmented neural networks



Learning to learn by gradient descent by gradient descent



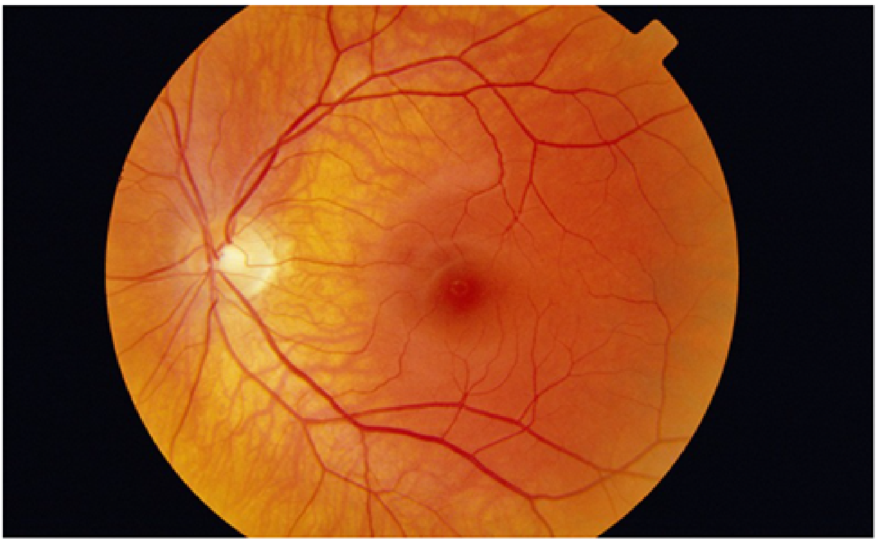
Learning to Learn: Meta-Critic Networks for Sample Efficient Learning

•Interpretability, Model Inspection, and Representation Analysis

- Feature Attribution and Per Example Interpretability
- Model Inspection and Representational Analysis

We can research on the open-sourced book： Interpretable Machine Learning.

- **Other studies**



Retinal images could allow computers to predict a person's risk of an imminent heart attack. Credit: Paul Parker/SPL

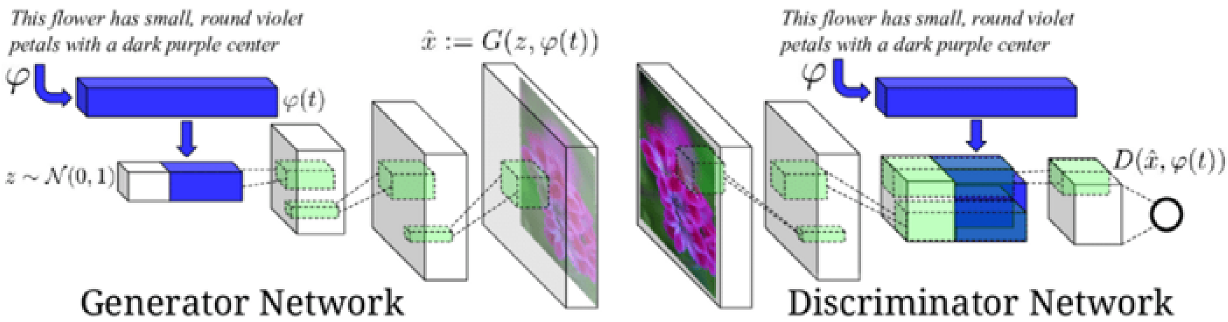
Predicted vs actual systolic blood pressure on the UK Biobank validation dataset, with the y=x line in black.

Model	AUC (95% CI)
Age	0.66 (0.61-0.71)
Systolic blood pressure (SBP)	0.66 (0.61-0.71)
Body mass index (BMI)	0.62 (0.56-0.67)
Gender	0.57 (0.53-0.62)
Current smoker	0.55 (0.52-0.59)
Algorithm	0.70 (0.65-0.74)

Algorithm	0.70 (0.65-0.74)
Age + SBP + BMI + gender + current smoker	0.72 (0.68-0.76)
Algorithm + age + SBP + BMI + gender + current smoker	0.73 (0.69-0.77)
Systematic COronary Risk Evaluation (SCORE)[6,7]	0.72 (0.67-0.76)
Algorithm + SCORE	0.72 (0.67-0.76)

Conclusion

Improvement, Future Directions, and Suggestions



We can pay more attention to the latest developments, including GAN, Meta-learning, and other related content, which will play a huge role in scientific research.

References:

[1] A Survey of Deep Learning for Scientific Discovery

[2] Deep learning sharpens views of cells and genes

[3] Recent Advances of Deep Learning in Bioinformatics and Computational Biology

[4] Deep learning for biology

[5] Multiple Instance Learning with MNIST dataset using Pytorch

[6] Review of Multi-Instance Learning and Its applications

[7] Santoro, Adam, Bartunov, Sergey, Botvinick, Matthew, Wierstra, Daan, and Lillicrap, Timothy. Meta-learning with memory-augmented neural networks. In Proceedings of The 33rd International Conference on Machine Learning, pp. 1842–1850, 2016.

[8] Munkhdalai T, Yu H. Meta Networks. arXiv preprint arXiv:1703.00837, 2017.

[9] Ravi, Sachin and Larochelle, Hugo. Optimization as a model for few-shot learning. In International Conference on Learning Representations (ICLR), 2017.

[10] Flood Sung, Zhang L, Xiang T, Hospedales T, et al. Learning to Learn: Meta-Critic Networks for Sample Efficient Learning. arXiv preprint arXiv:1706.09529, 2017.https://arxiv.org/pdf/1706.09529.pdf

[11] Ravi, Sachin and Larochelle, Hugo. Optimization as a model for few-shot learning. In International Conference on Learning Representations (ICLR), 2017.

Deep Learning