A random walk in Machine Learning

Jean-Michel Begon

University of Liege

March 15th, 2017

Université de Liège

Overview

The bread and butter of Machine Learning

- A typology of Machine Learning
- Unsupervised learning
- Supervised learning

Selected topics

- ML-based games
- Enhancement/Restauration
- Automatic face detection and recognition
- Tracking
- Recommender systems
- Natural language processing and speech recognition

Generative models

- Procedural content generation
- Automatic image captioning
- Image generation
- Style transfer

Conclusion

Overview

The bread and butter of Machine Learning

- A typology of Machine Learning
- Unsupervised learning
- Supervised learning

Selected topics

- ML-based games
- Enhancement/Restauration
- Automatic face detection and recognition
- Tracking
- Recommender systems
- Natural language processing and speech recognition

Generative models

- Procedural content generation
- Automatic image captioning
- Image generation
- Style transfer

Conclusion

A typology of Machine Learning

Machine Learning (ML) encompasses several areas :

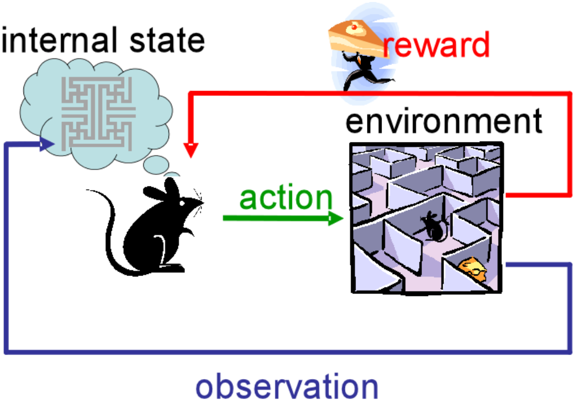
Unsupervised learning : finding structure/regularities in large set of data.

Supervised learning : building an input-output model from the observation of a large set of data in order to predict the target value of new examples.

Reinforcement learning : taking sequential actions in an uncertain environment to maximize some cumulative reward.

Common theme : the machine learns by itself ; the solution is not programmed *a priori*.

Reinforcement learning



It was the technique used by Google DeepMind to build `AlphaGo`, the program who beat a professional Go human player in March 2016.

Unsupervised learning

Goal :

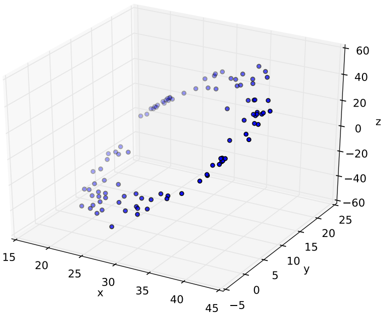
- Discover structure/regularities in large set of data.
- Preprocessing step for other techniques.

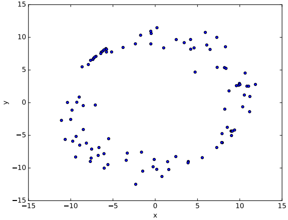
Common techniques :

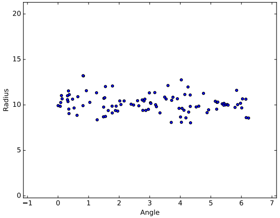
- Dimensionality reduction.
- Clustering.

Dimensionality reduction

Goal : find a simpler basis to express the data.

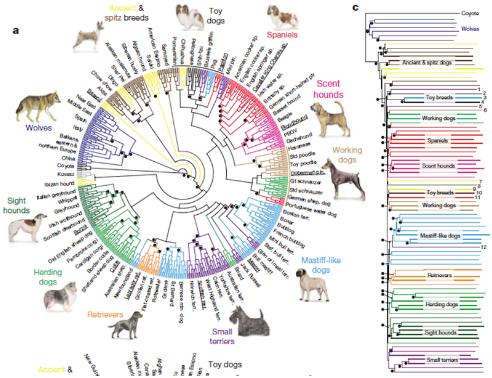






Clustering

Goal : group data that are similar together.



Dendrogram

Compression (7 colors)

Unsupervised learning — limitations

Limitations :

- Require lots of data.
- Not always well-posed :

  Dimensionality reduction: How many variables left ? What is a good approximation ?

  Clustering: How many clusters ? Similarity of complex data ?

Supervised learning

Goals :

- Build an input-output model from the observation of a large set of data in order to predict the target value of *new examples*.
  - Contrary to unsupervised learning, there is a clear target.

New input data　　　Model　　　output

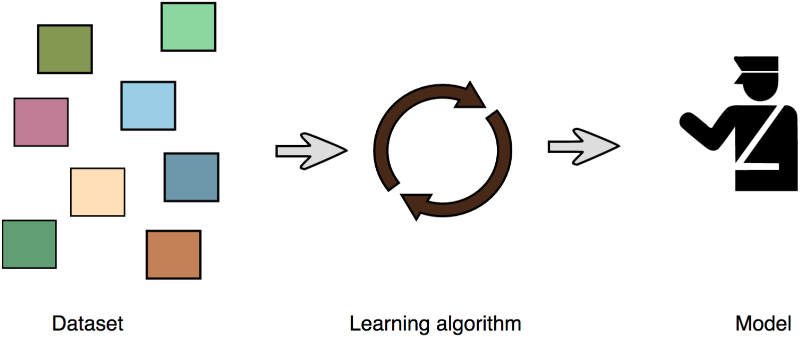
- Derive an understanding of the input-output relationship.

Depending on the nature of the output, we distinguish :

Classification : the output is discrete (label, class).

Regression : the output is a real value.

Supervised learning — learning part



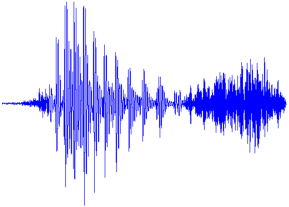
Classification

The output is discrete (label, class).

Medical diagnosis

Optical character recognition (ORC)

Spam detection



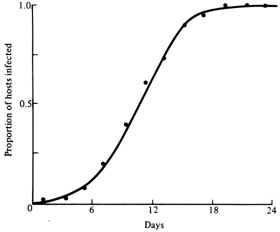
Voice identification

Regression

The output is a real value.

House price estimation

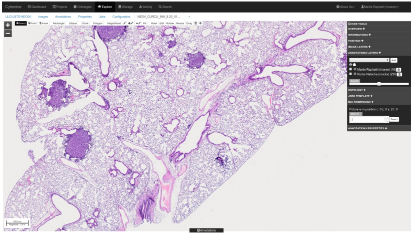
Sentiment analysis

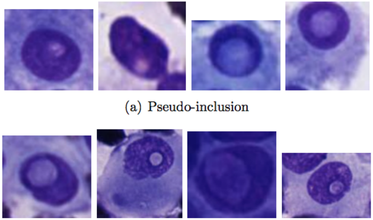


Epidemic diffusion

Stock market prediction

Ad time : Cytomine





(a) Pseudo-inclusion

(b) Inclusion

Segment, locate, dispatch and classify (SLDC)

http://www.cytomine.be/

https://github.com/waliens/sldc

Better understand the input-output relationship

From physiochemical properties of wine predict wine taste preference.

↪ understand why some wine are preferred over other.

Predict the delay of arriving flights.

↪ understand why planes arrive late.

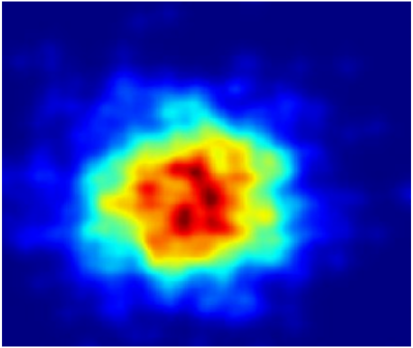
Multi-ouput

The model can predict several outputs at once :

Classification

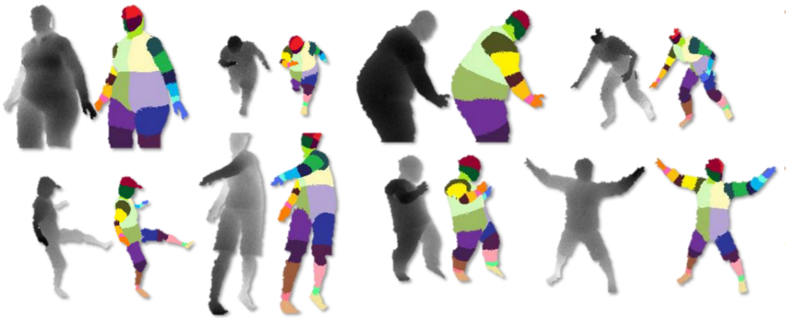
Automatic scene tagging

Regression



Chemical diffusion prediction

Case study : Microsoft's KINECT

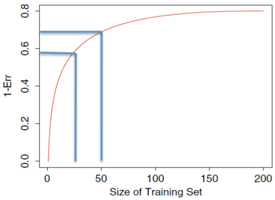


Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., & Moore, R. (2013). Real-time human pose recognition in parts from single depth images. Communications of the ACM, 56(1), 116-124.

Supervised learning — limitations

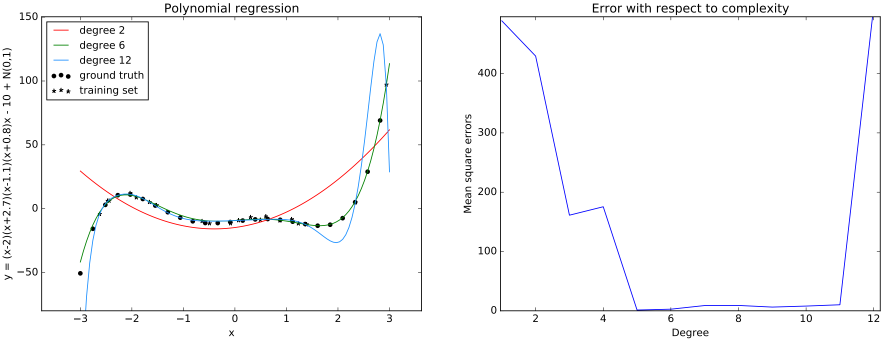
Limitations :

- Require lots of **labeled** data.



  - Tedious, time-consuming, expensive, difficult.
- Require relevant features.
- No free lunch : for every correct extrapolation a model makes, there exists at least one other consistent problem for which the model is wrong.
- Overfitting.

Supervised learning — limitations : overfitting



When the model is too powerful, it will not generalized well.

Overview

ML-based games

Guess who ?

http://en.akinator.com/

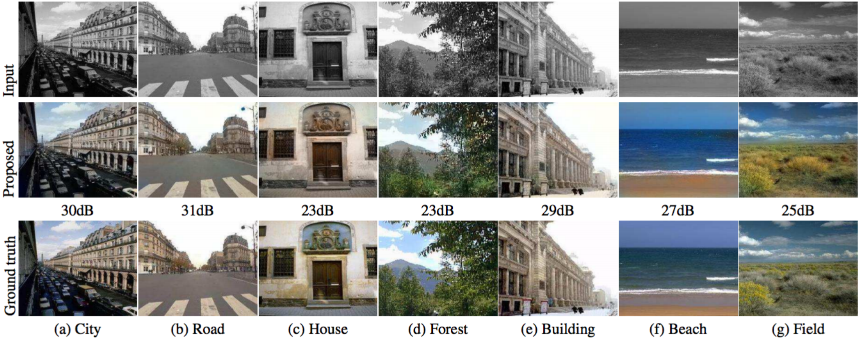
Any good at Pictionary ?

https://aiexperiments.withgoogle.com/quick-draw

Enhancement/Restauration — Image inpainting

Xie, J., Xu, L., & Chen, E. (2012). Image denoising and inpainting with deep neural networks. In Advances in Neural Information Processing Systems (pp. 341-349).

Enhancement/Restauration — Image (re-)coloration



Cheng, Z., Yang, Q., & Sheng, B. (2015). Deep colorization. In Proceedings of the IEEE International Conference on Computer Vision (pp. 415-423).

Automatic face detection and recognition : DeepFace

Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface : Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1701-1708).

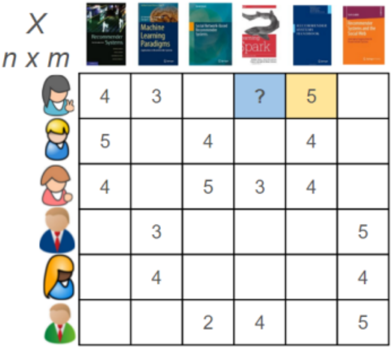
Tracking objects in video

https://www.youtube.com/watch?v=R1dAFjrzGC8

Kuen, J., Lim, K. M., & Lee, C. P. (2015). Self-taught learning of a deep invariant representation for visual tracking via temporal slowness principle. Pattern Recognition, 48(10), 2964-2982.

Recommender systems

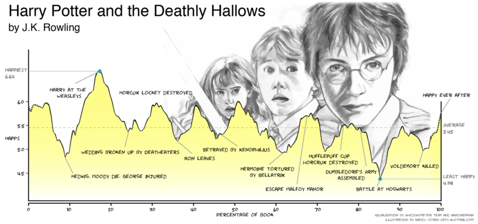
Given many user histories and possibly other relevant information, predict the rating a known user will give to a specific item.

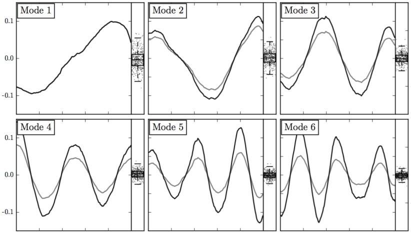


Use that prediction to recommend (new) items to a user.

Recommender systems

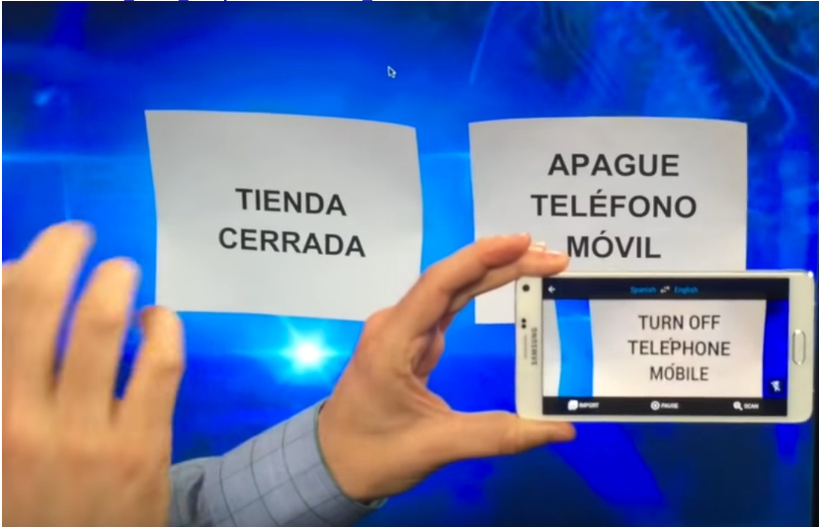
Natural language processing — The shapes of stories





Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. EPJ Data Science, 5(1), 31.

Natural language processing : Google translate



https://www.youtube.com/watch?v=Ro-HfETpzhc

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., & Klingner, J. (2016). Google's Neural Machine Translation System : Bridging

Natural language processing and speech recognition : amazon echo

https://www.youtube.com/watch?v=24Hz9qjTDfw

Overview

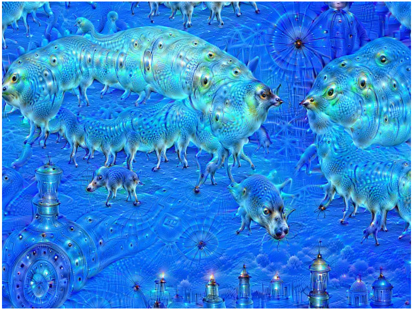
Procedural content generation

https://medium.com/@ageitgey/machine-learning-is-fun-part-2-a26a10b68df3#.yyrzqol0p

Automatic image captioning

Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). Show and tell : A neural image caption generator. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3156-3164).

Case study : DeepDream – Inception

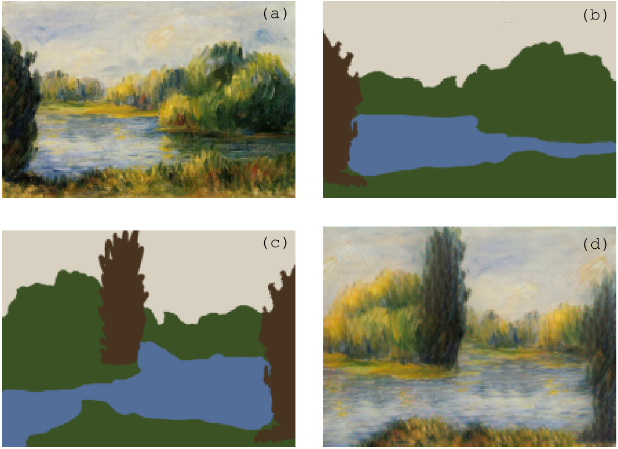


https://deepdreamgenerator.com
http://googleresearch.blogspot.ch/2015/06/inceptionism-going-deeper-into-neural.html

Style transfer

Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv :1508.06576.

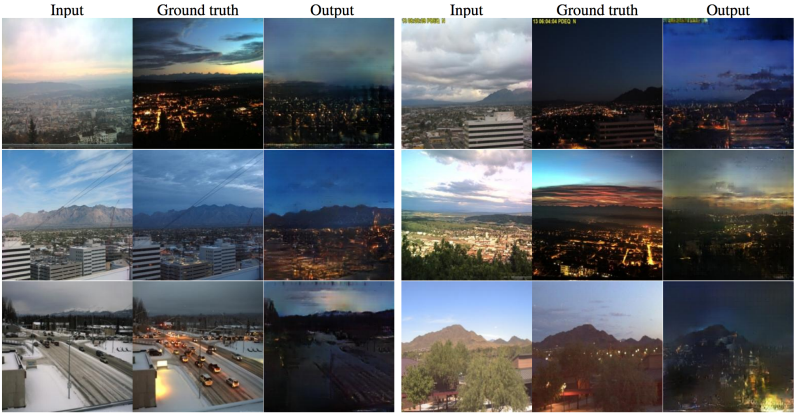
Style transfer with semantic map



(a) Original painting by Renoir, (b) semantic annotations, (c) desired layout, (d) generated output.

Champandard, A. J. (2016). Semantic style transfer and turning two-bit doodles into fine artworks. arXiv preprint arXiv :1603.01768.

Style transfer : day to night



Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2016). Image-to-image translation with conditional adversarial networks. arXiv preprint arXiv :1611.07004.

Style transfer for video

`https://www.youtube.com/watch?v=Khuj4ASldmU`

Overview

Conclusion

Machine learning (ML) is a versatile building block and is becoming widespread in our everyday life.

Despite its relatively long history, new application of ML are still coming to life nowadays.

It cannot solve everything but it is applicable in a surprisingly large number of scenarios.

It requires lots of (labeled) data.