Neural Networks
Applications

André C P L F de Carvalho
Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo, Brazil

1

Topics

- Applications
- Challenges
  - Design
  - Interpretability
- Final comments

2

ANN applications

- ANN is already present in our daily life
  - Recommend which messages to show in our social network app
  - Filter spam from your e-mail
  - Decide which results (and ads) to show in our web search
    - and in which order
  - Suggest films or books that someone will like
  - Decide if we can have a bank loan
  - Diagnose if you have a particular disease

3

Classical ANN applications

- Learn to read aloud
  - NETtalk (Sejnowski and Rosenberg, 1986)
- Learn to recognize spoken words
  - SPHINX (Lee, 1989)
- Learn to drive a car
  - ALVINN (Pomerleau, 1989)
- Learn to play backgammon
  - TD-GAMMON (Thesaurus, 1992)

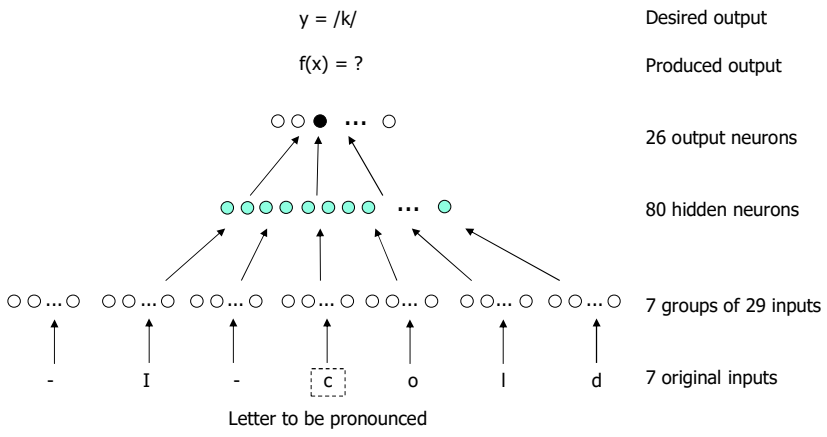
4

NETtalk

- *NETtalk: a parallel network that learns to read aloud* (1986)
  - Automatic system that learns to speak what is written in an English text
  - Maps texts in phonemes
    - Without using natural language processing or rules of phonetics
    - Uses a neural network with 1 hidden layer
      - Input layer: 7 groups with 29 inputs
      - Hidden layer: 80 neurons
      - Output layer: 26 neurons

5

NETtalk



6

NETtalk

- Input layer
  - Sliding window
  - 7 groups for the 7 text positions
    - 29 inputs = 26 letters + 3 punctuation/pause marks
- Hidden layer
  - Most neurons recognize more than one combination of letters

7

NETtalk

- Output layer
  - Code of the phoneme to pronounce the letter in the center of the input window
    - Besides silence, sound omission (syllabi), pause and end of reading
  - Output code depends only on the letters
    - Doe not use information about neighbour phonemes

8

NETtalk

9

NETtalk

- Dataset available in the UCI repository
  - 20008 English words
    - Together with their phonetic representations with stresses
  - 250 person-hours were dedicated to create and test the dataset

10

NETtalk

- Dataset has four predictive attribute for each word:
  - Letter representation
  - Phonemic representation
  - Stress and syllabic structure
  - An integer indicating foreign and irregular words

 11

11

NETtalk

- Predictive performance
  - 95% for training data and 78% for test data
  - Compared with Dectalk
    - Rule-based expert systems developed by linguists
    - Dectalk Presented a better predictive performance
      - Developed after a decade of studies by linguists

 12

12

Reading a poem

The crow crooked on more beautiful and free,
He journeyed off into the quarter sea.
his radiant ribs girdled empty and very –
least beautiful as dignified to see.

13

ANN poet

- Jack Hopkins, Researcher from University of Cambridge, UK
- ANN was trained with thousands of lines of poetry
  - More than 7 million words of 20th-century English poetry
    - Most of the poems came from online books

14

ANN poet

- After trained, created several poems
- Researcher asked more than 70 people whether the poem was written by
  - Human
  - ANN
- Poem with more human votes was written by an ANN

15

Sophia

16

ALVINN

Dean Pomerleau
CMU

17

ALVINN

- *Autonomous Land Vehicle In a Neural Network* (1989)
  - Autonomous car based on neural networks
    - CMU PhD Thesis
  - Input received by a camera on top of the vehicle
  - Travelled 4500 km at 110 Km/h, coast to coast, in American public roads
    - Except for 80 Km

18

ALVINN

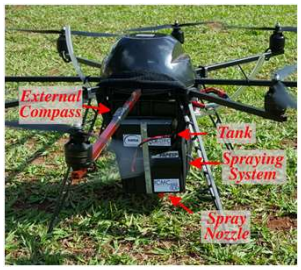
https://www.youtube.com/watch?v=r4JrcVEkink

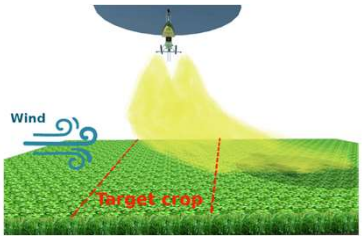
19

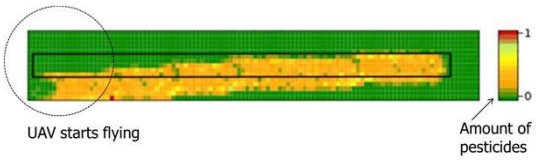
Autonomous vehicles

20

Pesticide spread in a crop field

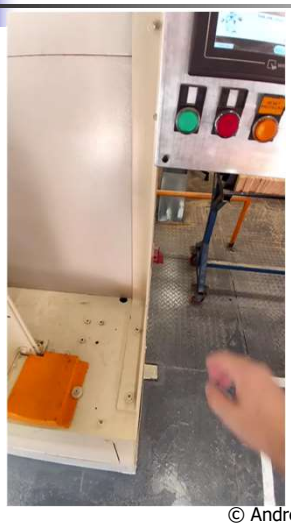






21

Wood classification



- Prototype of system to classify wooden board
- 100 boards
  - 50% good and 50% bad
    - 14 types of fault
  - 150 boards/minute or 27,7 meters/minute
  - Predictive accuracy: 97%
    - Human operator predictive accuracy: 93%

22

ANN architecture size

- Large networks:
  - Learning can be faster
  - Minimum local is easier to avoid
  - Solution space is less restricted
  - Fault tolerance is higher
- Small networks:
  - Generalization may be better
  - Require less computational resources
  - Knowledge extraction is easier

 23

23

ANN Architecture design

- ANN architecture usually has a strong impact on predictive performance
  - Neither too large, nor too small
- Alternatives
  - Train and error
  - Constructive
  - Pruning
  - Evolutionary

 24

24

Explainable AI (XAI)

- ML algorithms induce models with different levels of interpretation
  - Black box
    - Difficult to interpret
  - White box
    - Easy to interpret
    - Explain how decisions are made
  - Gray box

25

Explainable AI

- Although ANNs have been successfully employed in several problems
  - They still suffer from significant limitations
- One of them is the lack of comprehensibility of what they learned
  - How they make decisions
  - Essential in applications that affect human life
    - GDPR and LGPD
  - There are methods for knowledge extraction from ANNs

26

Explainable AI

- The methods for extraction knowledge from ANNs have some restrictions:
  - Distributed learning representation
  - Some methods work only with features with discrete values
  - Difficulty of their application to large networks

27

Data Age

Tech Giants Are Paying Huge Salaries for Scarce AI Talent

Newsweek

h & Science | Culture | Newsgeek | Sports | Health | Opinion

NEWSWEEK MAGAZINE

Need a Job? Why Artificial Intelligence Will Help Human

Artificial Intelligence: Salaries Heading Skyward

While the average salary for a Software Engineer is around $100,000 to $150,000, to make the big bucks you want to be an AI or Machine Learning (Specialist/Scientist/Engineer)

Stacy Stanford Follow
Aug 29, 2018 · 6 min read ★

BW BUSINESSWORLD

Businesses Expect AI To Have A Positive Impact On Growth, Productivity, Innovation, Job Creation: The Economist - Microsoft Study

Surprise! Machine learning jobs are high-paying and in-demand

By Ryan Daws
Editor of AI News. A gadget lover, music purveyor, and ex-host of a consumer technology show.

Posted on March 15, 2019

28

End-to-End Machine Learning

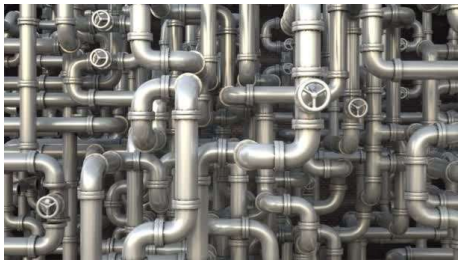
They are not independent

Modify ML algorithm

Oversample/ Undersampling

Re-code Features

Feature selection



Tune hyper-parameters

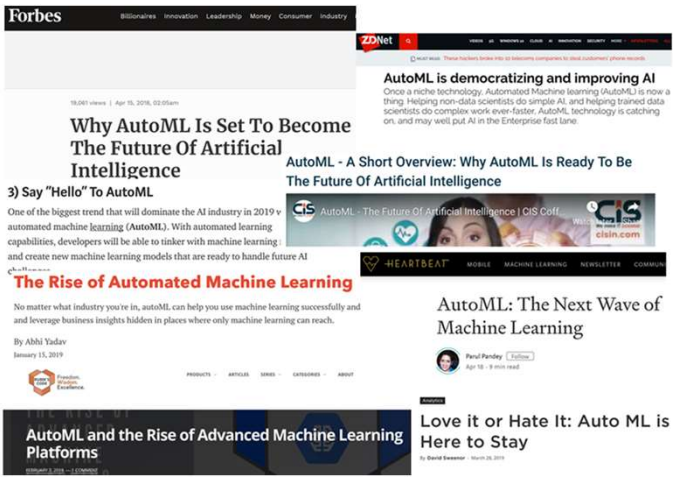
Deal with missing values

Check overfitting

Discover bugs

Adapted from Rick Caruana, Research opportunities in AutoML Microsoft Research

AutoML



AutoML tools

ML JAR

DataRobot

Auto-Sklearn

Amazon Rekognition

TPOT

H2O.ai

Cloud AutoML Vision

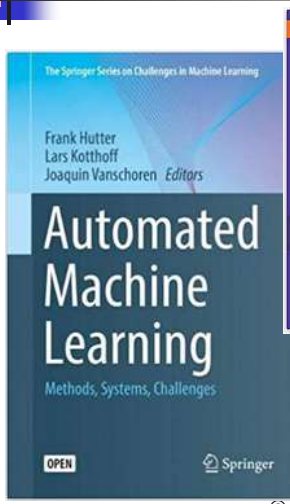
clarifai

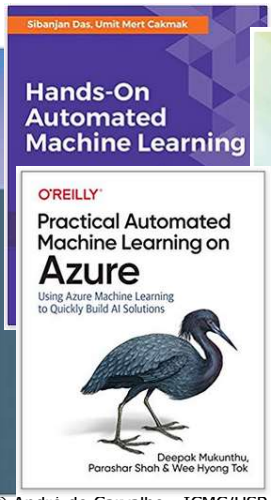
CreateML Apple

AutoML.org Freiburg

31

AutoML Books

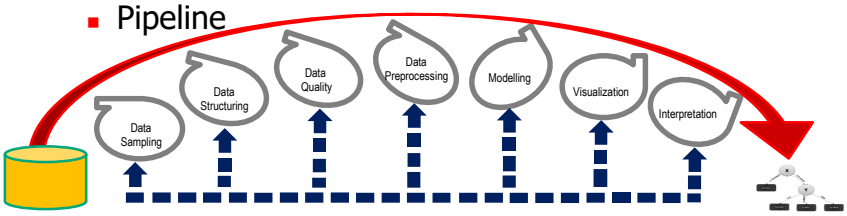




32

Pajé AutoML

- End-to-end AutoML (AutoDS)
- Main focus
  - Data pre-processing
  - Expandable
  - Explainable ML
  - Pipeline



Pajé Timeline (PhD Thesis)

Ricardo Sovat. Uma Abordagem Híbrida Baseada em Casos e Redes Neurais. Uma aplicação: escolha e configuração de modelos de redes neurais.

Mariá C. V. Nascimento. Meta-heurísticas para o problema de agrupamento de dados em grafo.

Rodrigo C.Barros. Automatic design of decision tree induction algorithms.

Davi P. Santos. Seleção e controle do viés de aprendizado ativo.

2002 2003 2010 2013 2016 2018

Estéfane G. Lacerda. Model Selection of RBF Networks via Genetic Algorithms.

Claudia Regina Milaré. Extração de Conhecimento de Redes Neurais Artificiais.

Bruno F. de Souza. Meta-aprendizagem aplicada à classificação de dados de expressão gênica.

André L. D. Rossi. Meta-aprendizado aplicado a fluxos contínuos de dados.

Luis Paulo Garcia. Noise detection in classification problems.

Rafael G. Mantovani. Use of meta-learning for hyperparameter tuning of classification problems.

Conclusion

- Challenges
- ANN design
- Hyperparameter tuning
- Continuous update
- Interpretability
- Mathematical foundation

35

36

Support

37

Questions?

38