Machine Learning for Data-Driven Smart Grid Applications

May 22, 2018

Patrick Glauner, Jorge Augusto Meira and Radu State
Interdisciplinary Centre for Security, Reliability and Trust,
University of Luxembourg
Motivation: non-technical losses (NTL)

Example of NTL: Two assumed occurrences of NTL due to significant consumption drops followed by inspections (visualized by vertical bars).
About us

• We work on real-world machine learning problems together with industry partners
• Recent research includes detection of electricity theft/non-technical losses, correction of biases in data and augmented reality
Goals of this tutorial

• Providing an introduction to machine learning
• Understanding the three pillars of machine learning
• Knowing when to use which model
Contents

• Introduction
• Supervised learning
• Unsupervised learning
• Reinforcement learning
• Deep Learning
• Conclusions
"Artificial intelligence is the science of knowing what to do when you don't know what to do." (Peter Norvig)

https://www.youtube.com/watch?v=rtmQ3xlt-4A
Introduction

Arthur Samuel (1959): “Field of study that gives computers the ability to learn without being explicitly programmed”.

Introduction
Introduction

What do customers buy after viewing this item?

Best Selling

Lenovo
N22 11.6-Inch HD Chromebook Laptop (Black) - (Intel Celeron N3060, 2 GB RAM, 32 GB EMMC, Chrome OS)

🌟🌟🌟🌟🌟 271

£109.99 ✔️Prime

Top Rated

Inateck
13-13.3 Inch Macbook Air/ Macbook Pro / Pro Retina Sleeve Case Cover Protective Bag Ultrabook Netbook Carrying Protector

🌟🌟🌟🌟🌟 35

£16.99 ✔️Prime
Introduction
Introduction

“Machine Learning is a subset of Artificial Intelligence techniques which use statistical models to enable machines to improve with experiences”

Use cases: data mining, autonomous cars, recommendation...

Tom Mitchell (1998): "A computer program is said to learn from experience $E$ with respect to some class of tasks $T$ and performance measure $P$ if its performance at tasks in $T$, as measured by $P$, improves with experience $E"."
Introduction

Alan Turing creates the “Turing Test” to determine if a computer has real intelligence. 1950

Frank Rosenblatt designed the first neural network for computers (the perceptron). 1957

Arthur Samuel wrote the first computer learning program. 1952

IBM’s Deep Blue beats the world champion at chess. 1997

Geoffrey Hinton coins the term “deep learning” to explain new algorithms that let computers “see” and distinguish objects and text in images. 2006

AlphaGo became the first computer Go program to beat a human professional Go player. 2015

source: https://www.forbes.com
Introduction: three pillars of machine learning

- Supervised learning: induce a function that maps from input to output.
- Unsupervised learning: find hidden structure in data.
- Reinforced learning: reward-based learning.
Supervised learning
Supervised learning

Regression

Classification
Supervised learning

Regression

Classification
Supervised learning

Regression

Classification
Supervised learning

Regression

Classification

Continuous values

Discrete values (categories)
Supervised learning: use cases

• Detection of anomalies
• Forecasting
• Medical diagnosis
• ...

Supervised learning: models

- Linear/logistic regression
- Decision tree, random forest
- Support vector machine
- ...
- Neural networks, Deep Learning
Supervised learning: workflow

- Training set
- Learning algorithm
- $x$: features
- $h$: hypothesis
- $y$: prediction
Supervised learning: regression

How to represent ‘h’ (hypothesis)
Supervised learning: regression

How to represent ‘h’ (hypothesis)

\[ h_\theta(x) = \theta_0 + \theta_1 x \]

\[ y = ax + b \]
Supervised learning: regression

How to represent ‘h’ (hypothesis)

For example, housing price.

<table>
<thead>
<tr>
<th></th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>~180k</td>
<td>~182k</td>
<td>~184k</td>
<td>~186k</td>
<td>???</td>
</tr>
</tbody>
</table>

\[ h_\theta(x) = \theta_0 + \theta_1 x \]
\[ y = ax + b \]
Supervised learning: regression

How to represent ‘h’ (hypothesis)

\[ h_0(x) = 1.5 + 0x \]

\[ h_0(x) = 0.5x \]

\[ h_0(x) = 1 + 0.5x \]
Supervised learning: regression

How to represent ‘h’ (hypothesis)

\[ h_0(x) = \theta_0 + \theta_1 x \]

Choose \( \theta_0, \theta_1 \) so that \( h_0(x) \) is close to \( y \) for our training set.
Supervised learning: regression

How to represent ‘h’ (hypothesis)

The idea is to minimize \( \theta_0 \theta_1 \), so that \( h_\theta(x) - y \) tends to decrease.

Thus, we can define the cost function \( J(\theta_0 \theta_1) \) aiming to minimize \( \theta_0 \theta_1 \):

\[
J(\theta) = \frac{1}{2} \sum (h_\theta(x) - y)^2
\]
Supervised learning: classification

Example of non-technical losses (NTL): Two assumed occurrences of NTL due to significant consumption drops followed by inspections (visualized by vertical bars).
Supervised learning: decision tree

Weather
Supervised learning: decision tree
Supervised learning: decision tree

- **Weather**
  - Sunny
    - High Humidity
  - Cloudy
    - Normal Humidity
    - Weak Wind
    - Strong Wind
  - Rain
    - no
Supervised learning: decision tree

- **Weather**
  - Sunny
    - High Humidity (yes)
    - Normal Humidity (no)
  - Cloudy
    - Weak Wind (yes)
    - Strong Wind (no)
  - Rain (no)
Unsupervised learning

Supervised

Known labels
Unsupervised learning

Supervised

x_2

x_1

Known labels

Unsupervised

x_2

x_1

Unknown labels
Unsupervised learning

Supervised

Known labels

Unsupervised

Unknown labels
Unsupervised learning

Supervised

Known labels

Unsupervised

Unknown labels
Unsupervised learning: clustering

K-means algorithm

1: Define K centroids randomly.
Unsupervised learning: clustering

K-means algorithm

1: Define K centroids randomly.
2: Associate every observation according to the nearest centroid.
K-means algorithm

1: Define K centroids randomly.
2: Associate every observation according to the nearest centroid.
Unsupervised learning: clustering

K-means algorithm

1: Define $K$ centroids randomly.
2: Associate every observation according to the nearest centroid.
3: Define new centroids according to the mean of the clusters.
Unsupervised learning: clustering

K-means algorithm

1: Define K centroids randomly.
2: Associate every observation according to the nearest centroid.
3: Define new centroids according to the mean of the clusters.
4: Repeat step 2 and 3 to converge.
Unsupervised learning: use cases

• Market segmentation
• Clustering of customers, news, etc.
• Dimensionality reduction of data
Unsupervised learning: models

- k-means clustering
- Expectation-maximization clustering
- Principal component analysis
- ...
Reinforcement learning
Reinforcement learning: use cases

• Planning
• Playing games, e.g. the game of Go
• ...
Reinforcement learning: models

- Value/policy iteration
- Q-learning
- Deep reinforcement learning
- ...
Deep Learning: neural network
Deep Learning
Conclusions

• Machine Learning allows to learn complex statistical patterns from data
• Not much domain knowledge required
• Many applications in daily life
• Tell us more about your workflows so that we can figure out how Machine Learning can help you!
References