Challenges and perspectives in machine learning

Alexis Zubiolo
alexis.zubiolo@gmail.com

Data Science Team Lead @ Adcash

February 16, 2017

Purpose of this talk

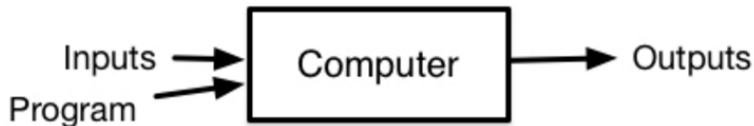Illustrate recent **trends** and **challenges** in machine learning on **practical examples**.

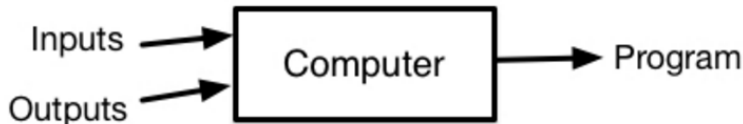# Supervised learning paradigm

Supervised learning paradigm

Goal: Build algorithms that can
- ▶ **learn** from data
- ▶ **make predictions** on data

Traditional Programming

Inputs ⟶
Program ⟶
| Computer | ⟶ Outputs

Machine Learning

Inputs ⟶
Outputs ⟶
| Computer | ⟶ Program

Example 1: Machine translation

Machine translation

Machine learning is a way to perform automated translation:

- Input = reference text
- Output = translated text

Machine translation

Machine learning is a way to perform automated translation:
- Input = reference text
- Output = translated text

There are differences between languages at different levels:
- Genders, e.g. neutral, masculine, feminine
- Conjugation
- Cases
- . . .

Look at these two sentences:

Machine translation: A tricky example

Look at these two sentences:

The trophy could not fit in the bag because **it** is too small.

and

The trophy could not fit in the bag because **it** is too big.

Machine translation: A tricky example

Look at these two sentences:

The trophy could not fit in the bag because **it** is too small.

and

The trophy could not fit in the bag because **it** is too big.

What does the word "**it**"refer to in both cases?

Machine translation: A tricky example

Take a close look at these two sentences:

The trophy could not fit in **the bag** because **it** is too small.

and

The trophy could not fit in the bag because **it** is too big.

What does the word "**it**"refer to in both cases?

Machine translation: A tricky example

Based on the context, we can guess what "**it**"refers to.

Machine translation: A tricky example

Based on the context, we can guess what "**it**"refers to.

This is key because, if we wanted to translate this sentence into Bulgarian, we would be faced to the following challenge:

- Trophy = трофей is masculine
- Bag = чанта is feminine

Machine translation: A tricky example

Based on the context, we can guess what "**it**"refers to.

This is key because, if we wanted to translate this sentence into Bulgarian, we would be faced to the following challenge:

- Trophy = трофей is masculine
- Bag = чанта is feminine

Consequence: "it"would refer to a masculine or a feminine noun, depending on the context.

Machine translation: A tricky example

Let's translate this sentence **from English to Bulgarian**... Here is my attempt:

Machine translation: A tricky example

Let's translate this sentence **from English to Bulgarian**… Here is my attempt:

The trophy could not fit in the **bag** because **it** is too small.
= Трофеят не може да се побере в **чантата**, защото (**тя**) е твърде **малка**.

Machine translation: A tricky example

Let's translate this sentence **from English to Bulgarian**... Here is my attempt:

The trophy could not fit in the **bag** because **it** is too small.
= Трофеят не може да се побере в **чантата**, защото (**тя**) е твърде **малка**.

The **trophy** could not fit in the bag because **it** is too big.
= **Трофеят** не може да се побере в чантата, защото (**той**) е твърде **голям**.

Machine translation: A tricky example

Let's translate this sentence **from English to Bulgarian**... Here is my attempt:

The trophy could not fit in the **bag** because **it** is too small.
= Трофеят не може да се побере в **чантата**, защото (**тя**) е твърде **малка**.

The **trophy** could not fit in the bag because **it** is too big.
= **Трофеят** не може да се побере в чантата, защото (**той**) е твърде **голям**.

This is often referred to as **long short-term memory** (LSTM).

Example 2: Acquiring data

Acquiring data

Let us switch to another example:

Acquiring data

Let us switch to another example: Reading house numbers in Google street view

Acquiring data

Let us switch to another example: Reading house numbers in Google street view

Acquiring data

Conclusion

Thank you! Questions?

alexis.zubiolo@gmail.com