INF264 - Ex1

Pierre Gillot Natacha Galmiche*

Week 34 - 2021

In this exercise, we learn how to identify and formulate machine learning algorithms that are relevant to solve different types of real-world problems. Recall Mitchell's definition of a well-defined machine learning problem:

> "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**".

Imagine that you are a consultant specialized in machine learning solutions. For each problem given below, identify and formulate a machine learning algorithm that could be used to solve this task. To do this, shortly explain what **T**, **P** and **E** could be in those contexts. Note that a problem may be formulated in several reasonable ways.

- Task **T** is what the machine learning solution does. It may be helpful to think of it as a computational problem that has an input and an output which you should identify. You need not think how that program would work.
- Performance measure **P** tells how well the program works. The performance measure should reflect your customer's goal and it should be measurable in practice.
- Experience **E** specifies the data that is used for learning. Think about what kind of data you would collect.

1 Getting rich with Rema 1000

The grocery store Rema 1000 is considering opening a new store in Marineholmen. They know that popularity of different products varies depending on the location of a store. Now, they want to know what kind of assortment they should choose for the new store. The managers at Rema 1000 have heard that machine learning can help solve this kind of problems.

*Not a TA for this course this year.

2 Automating passion

John is conflicted: this long-time video game fan has become too bored playing video games himself, but at the same time he feels really ashamed admitting watching playthroughs on youtube. Fortunately, John also happens to know an experienced machine learning engineer (you), thus he decided to implement machine learning algorithms that would secretly play his favorite games for him. As an early experimentation, he wants to build an AI that would learn to beat as fast as possible old school 2D scrolling platformers like Super Mario World.
Help John become a proud gamer again.

3 Refill John's wallet

John is paid to change scooters' batteries when they are about to die. To do so he first finds groups of scooters whose batteries are running low, parks his car nearby and then walks to each of these scooters. He is paid per battery replaced so he wants to optimize the number of batteries changed per hour. Note that it sometimes takes longer to drive and park than walk. In addition, some parking slots are not free and John doesn't get reimbursed for parking tickets. Thus he wants to find the right number of stops, where to park his car as well as which scooters to walk to for each stop.
Help John get more money.

4 How to ~~lay off~~ protect workers in modern industry

A company is specialized in the destruction of industrial waste. Workers are employed to control mechanic arms that sort the different types of waste in a chain for proper destruction. A recent security investigation showed that the workers were exposed to dangerous substances and the company is expected to fully automate the waste destruction chain. They heard machine learning could help them automate the chain efficiently.

5 Empty John's wallet

John bought an electric scooter 3 weeks ago and he still gets ads for scooters every day on his favorite search engine. He would like to get ads for things he actually needs such as a helmet or larger size clothes for his incoming weight gain (master's student life).
Help John's favorite search engine empty his wallet.

6 The sheep of Wall Street

Due to the high risk of a new recession, the ever-growing Norwegian bank DNB decided internally to "play it safe": they fired without further notice all of their stockbrokers and set up a machine learning algorithm instead, with the promise from machine learning experts that all financial actions validated by the algorithm would be in the best interest of the bank itself rather than in the interest of the traders' personal wallets.

7 The phantom menace

In a galaxy far, far away... there is a ongoing pandemic. Aliens living there are willing to follow some rules but they get easily demotivated as soon as there are more than 3 rules to respect. So far the pandemic is not well understood, many factors could influence epidemic spread but they don't know which ones are most important and they don't have time to try every 3-combination possible.
Help Aliens survive.

8 Avoiding a bloodshed with Tesla's new self-driving car

Tesla wants to build reliable self-driving cars that don't crash. Cars use different sensors (cameras, radars, GPS) to gather knowledge of their surroundings. Now Tesla is building a steering system to avoid collisions, that involves machine learning.

9 Passing INF-264 exam

While collecting batteries, John realizes machine learning exam is tomorrow and there are lots of definitions to learn. John doesn't know where to start because he can't really understand a word if he doesn't understand all the words in its definition first. Usually he builds by hand a tree containing all chapter's definitions, the top being the most general concepts and the bottom specific terms such as methods, applications, etc...

This time he is running out of time so he wants to automatize the tree's construction to make sure he starts in the "right" order.
Help John hierarchize his definitions.

Note: In a definition, general concepts it belongs to are usually mentioned. But sometimes examples and/or similar concepts are mentioned too and need not be studied first!

Examples from Wikipedia:

> **Support-vector machines**: In machine learning, SVM are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis

SVM < machine learning, supervised learning, classification, regression analysis

> **Classification**: In the terminology of machine learning, classification is considered an instance of supervised learning, i.e., learning where a training set of correctly identified observations is available. The corresponding unsupervised procedure is known as clustering, and involves grouping data into categories based on some measure of inherent similarity or distance

Classification < machine learning, supervised learning BUT classification NOT < clustering and unsupervised!

> **Supervised learning**: Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs

Supervised learning < machine learning