EuroSDR

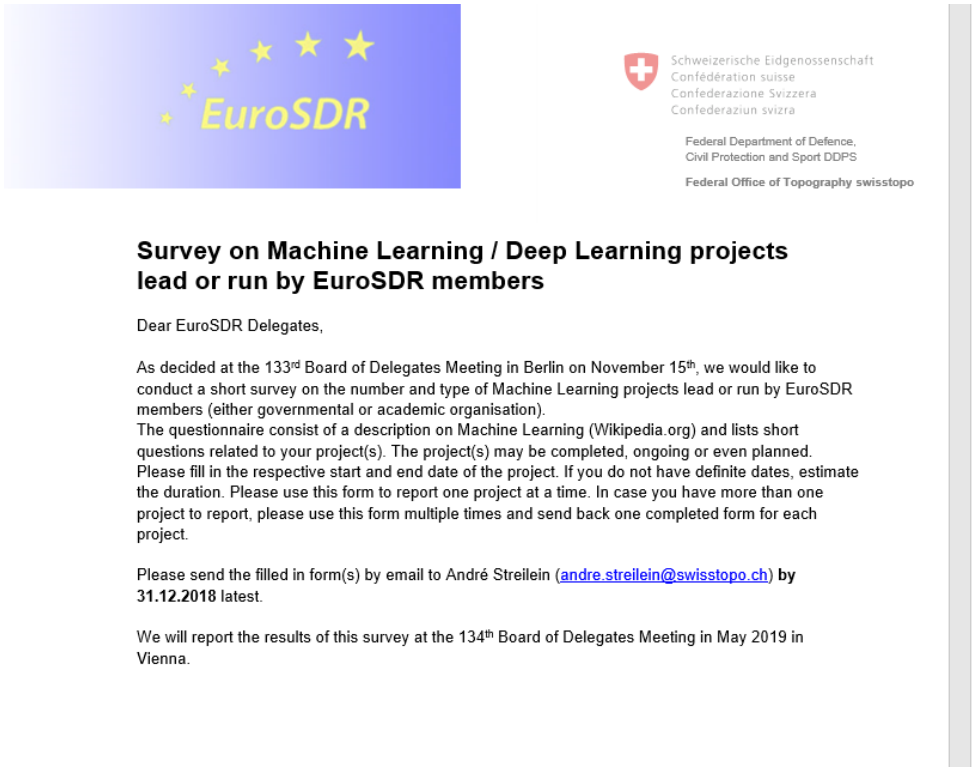
European Spatial Data Research

EuroSDR survey on ML/AI 2018/2019

André Streilein swisstopo

EuroSDR/EuroGeographics, February 2021

Survey



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Defence,
Civil Protection and Sport DDPS

Federal Office of Topography swisstopo

Survey on Machine Learning / Deep Learning projects lead or run by EuroSDR members

Dear EuroSDR Delegates,

As decided at the 133rd Board of Delegates Meeting in Berlin on November 15th, we would like to conduct a short survey on the number and type of Machine Learning projects lead or run by EuroSDR members (either governmental or academic organisation).
The questionnaire consist of a description on Machine Learning (Wikipedia.org) and lists short questions related to your project(s). The project(s) may be completed, ongoing or even planned.
Please fill in the respective start and end date of the project. If you do not have definite dates, estimate the duration. Please use this form to report one project at a time. In case you have more than one project to report, please use this form multiple times and send back one completed form for each project.

Please send the filled in form(s) by email to André Streilein (andre.streilein@swisstopo.ch) **by 31.12.2018** latest.

We will report the results of this survey at the 134th Board of Delegates Meeting in May 2019 in Vienna.

Definitions

Machine Learning (ML)
Machine learning is a field of Artificial Intelligence (AI) that uses statistical techniques to give computer systems the ability to "learn" (e.g., progressively improve performance on a specific task) from data, without being explicitly programmed.
Machine learning tasks are typically classified into several broad categories:

- Supervised learning: The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule that maps inputs to outputs. As special cases, the input signal can be only partially available, or restricted to special feedback.
- Semi-supervised learning: The computer is given only an incomplete training signal: a training set with some (often many) of the target outputs missing.
- Active learning: The computer can only obtain training labels for a limited set of instances (based on a budget), and also has to optimize its choice of objects to acquire labels for. When used interactively, these can be presented to the user for labeling.
- Unsupervised learning: No labels are given to the learning algorithm, leaving it on its own to find structure in its input. Unsupervised learning can be a goal in itself (discovering hidden patterns in data) or a means towards an end (feature learning).
- Reinforcement learning: Data (in form of rewards and punishments) are given only as feedback to the program's actions in a dynamic environment, such as driving a vehicle or playing a game against an opponent.

Deep Learning (DL)

- Deep learning is a class of machine learning algorithms that:use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input.
- learn in supervised (e.g., classification) and/or unsupervised (e.g., pattern analysis) manners.
- learn multiple levels of representations that correspond to different levels of abstraction; the levels form a hierarchy of concepts.

(Wikipedia.org)

November/December 2018

EuroSDR

Survey

Questionnaire

Please use this form to report one project at a time. In case, you have more than one project to report, use this form multiple times and send back one completed form for each project.

Organisation

Organisation	Click or tap here to enter text.
Country	Click or tap here to enter text.
Responsible Person(s)	Click or tap here to enter text.
Postal Address	Click or tap here to enter text.
Email Address	Click or tap here to enter text.
Contact Phone	Click or tap here to enter text.

Project

Name of the Project	Click or tap here to enter text.
Short description	Click or tap here to enter text.
Three main questions to be answered in the project	1. Click or tap here to enter text. 2. Click or tap here to enter text. 3. Click or tap here to enter text.
Project time frame	☐planned ☐running ☐completed Start date: Click or tap to enter a date. End date: Click or tap to enter a date.
Project contractor / research entity	☐internal ☐external ☐both
Type of Machine Learning	☐Supervised learning ☐Semi-supervised learning. ☐Active learning ☐Unsupervised learning ☐Reinforcement learning
Input data type	Geodata: Click or tap here to enter text. Auxiliary data: Click or tap here to enter text. Others: Click or tap here to enter text.
Pre-requisites	Click or tap here to enter text.
Main findings	Click or tap here to enter text.
Implemented in production-line	☐ yes ☐ no

Please send the filled form(s) by email to André Streilein (andre.streilein@swisstopo.ch) by **31.12.2018** latest.

Outcome

- First response: call 09:24, response 09:56 (->35min after call !!!)

- 11 member countries reported projects
- 3 member countries reported that they have no projects

- In total: responses from 14 member countries

Project 1	
Responsible person	
Country	
Organisation	
Postal address	
Email address	
Contact phone	
Name of project	
Short description	supervised ML (random forests, dictionary learning, probabilistic label smoothing, contextual classification) - Classification (pure) or integration of classification and change detection - Urban or river areas or interior areas - Image matching and laser scanning point clouds - Method development vs. large area application DL, supervised - Classification of region wide laser scanning point clouds
Three main questions to be answered in the project	1. Three main questions to be answered in the project 1. Quality of classification and amount of training data 2. Workflow design (point vs. raster, formulation of problem as ML task, etc.) 3. Transfer to super computer (effort, gain, costs)
Project time frame	planned running completed 01.01.2015
Project contractor research entity	internal both
Type of Machine Learning	supervised learning active learning
Input data type	Geodata: Point clouds Auxiliary Data: Reference data in the form of maps
Pre-requisites	-

Main findings	Main findings - additional effort in training data generation pays of in comparison to expert designed classification rules - Black box of machine learning still requires a lot of expertise to solve the classification task - lack of theoretical understanding (in our group) how to describe a problem to obtain results of predictable quality - usage of super computer power becomes additional field of expertise to build
Implemented in production line	Yes

Outcome

- 28 projects have been reported
- Only 3 of them were considered as operational
- The responses showed the heterogeneity of the current research
- There seemed to be no coordination/synergy – every organisation is running its own project with its own focus
- Of course this was not a complete overview of projects running – it was an indication!

Break out session – Leading questions

- How do you judge the overall relevance of ML/AI for the task of Mapping Agencies?
- Which results do you expect from AI processes?
  - With respect to: Transparent processes, reliable results, incremental updating, change detection (heat maps etc.), etc.
- Where would you see ML/AI better applied?
  - ML/AI can be used / applied in various steps of our processing pipeline. E.g. point cloud classification, image segmentation, cartographic generalization, 3D modeling, ontologies/schema matching, …

Q1: How do you judge the overall relevance of ML/AI for the task of Mapping Agencies?

- There is no other option (e.g. how to deal with the amount of data, cost effectiveness)
- Every technical development is relevant
- Automation is the task of today (increased resolution, increased update circle)
- The technology will be applied anyhow – either by NMAs or others
- Will the products of NMAs change?
- Impact on Delivery time of products

Q2: Which results do you expect from AI processes?

- Same results as with conventional methods
- More results (e.g. more classes, more lod)
- More automatic information for decision support
- Transparent processes or a redundant quality control
- Predicitions (likelihood)

Q3: Where would you see ML/AI better applied?

- Everything in the workflow that hinder automation (e.g. water / snow detection, feature matching, feature extraction)
- Image interpretation
- Data currency / update cycles

Expectations, Desires, Reality

Deep Learning Approaches for Geospatial Data Evaluation in Mapping

Expectations and wishes

- No human interaction
- Fully automatic
- Faster
- Cheaper
- More content

- DL is based on statistics!
  - Designed for Euclidean 2D GeoData (continua, not vectors) (because simple operations like convolutions)
  - Geometric Deep Learning in its beginning
  - no 3D approaches in DL (yet)
- appealingly suitable for object detection
- moderately suitable for spatial point statistics
- conditionally suitable for spatial area mapping

What's next?

- EuroSRD will repeat this survey in 2021 !
- Questionnaire is also open for EuroGeographics members (if interested)