# Resolving Limits Faced by Classical Machine Learning Approaches: Areas of Application for Collaborative Interactive Learning Techniques

Christoph Sonntag[1]
[1]*Intelligent Systems Group, University of Passau, Innstrasse 70, Passau, Germany*
*sonntagc@fim.uni-passau.de*

***Zusammenfassung*—The area of *Machine Learning (ML)* has experienced a high level of research interest in the last few years with it's underlying theory reaching back far into the 18Th century. Due to minimal costs for computation and the massive availability of data in the Internet recent research interest has been mainly focused on *automatic ML (aML)*. In this ML paradigm a model (e.g. classifier, . . . ) is being trained with a pre-labeled training set in order to make predictions on unknown unlabeled data. Problems arise in domains where data is rarely available or where the labeling process is too expensive (computationally or economically). Furthermore, real-world data can be uncertain, incomplete or contain noisy data. In these cases full automation is not reliable enough or even infeasible for specific domains. However, robust and trustworthy results are mandatory in a lot of applications like health informatics or autonomous vehicles. Therefore, systems are needed that interactively integrate knowledge from various experts, which can be collaborating humans or machines, in order to continuously improve results over a systems whole lifetime: So called *Collaborative Interactive Learning (CIL)* systems. This paper contributes in presenting an overview of classical ML approaches in comparison with CIL and outlines possible areas of application.**

## 1. Introduction

As with most concepts, there is no canonical definition for the term *Machine Learning* but at its most basic form it can be described as algorithms that learn from a given set of training examples $\{(x_i, y_i)\}$ of inputs $x_i$ and outputs $y_i$ in order to make predictions on unknown input data. According to Tom Mitchell, a Computer Scientist from Carnegie Mellon University,ML tries to answer how we can build algorithms that automatically improve through experience and what fundamental laws govern all these learning processes [1]. The area of ML in general is a fast-growing discipline at the intersection of statistics and computer science and has experienced massive research interest in the last decades. With its various application possibilities it has also been an interesting branch for economists and entrepreneurs. Since ML scenarios like *supervised*, *unsupervised* and *semi-supervised* learning, which we will cover later, are heavily data-driven, the Internet with its massive amount of data has contributed to the further development of research in the area of fully automated learning algorithms. Today, we can see the results in a variety of applications [2] [3], not limited to the list below.

- *Text classification, Natural Language Processing (NLP)*. Many current mail programs have built-in spam detection which originally used simple regular expressions in order to detect phrases commonly used in spam mails. With state-of-the-art ML techniques it is not only possible for a mail program to query for a number of words but to the pattern how spam mails are constructed and to adapt itself to new types of spam.
- *Speech recognition*. Most commercial applications for speech recognition use ML to train itself to recognize a users speech input. Umformen. Beschreiben, wofr ML genutzt wird.
- *Image recognition, OCR*. Machine Learning methods have also been successfully applied in domains where it's important to extract information from images such as detecting and classifiying objects or recognize handwritten only handwritten? characters. Image recognition and OCR are therefore for example used in medical diagnosis to detect cancer on radiographs, in autonomous vehicles to detect obstacles and to stay on track or in post-offices to sort envelopes with hand-written addresses. Nochmal eine Quellenangabe?
- *Games*. Computer Games ususally offer a Multiplayer mode where a user can interact and play the game even without a human opponent. For games with a small number of possible moves after each turn, and therefore a realtively small game tree conatining all possible moves, the minimax algorithm [4] can be sufficient. It figures out which next move would minimize the worst-case scenario for all subsequent moves. Therefore, it needs to know all possible moves, which can easily be computationally infeasible for games like Chess or AlphaGo where a general move faces between 35 (Chess) and 250 (AlphaGo) possible subsequent game states. In order to still have challenging opponents modern Games are using ML to gain experience by themselves.
- *Search engines, Recommendation systems*. Nearly all current search machines use ML in one fashion or

another. They mainly use these techniques for *(1) User classification*[q?] in order to offer personalized search results and for *(2) Query classification* in order to “understand’ a users query and to provide further meaningful information.

Despite the successfull application of ML in a lot of fields, most of these examples need a sufficient amount of labeled training data ($x_i$ with assigned $y_i$) in order to make correct predictions on or discover structured patterns in unknown data $x_j$. However, most data sets in the biomedical domain, in robotics or in other fields, where data is collected from sensor systems or other unreliable sources, are often either not available (e.g. rare diseases, borderline cases in road traffic, ...) or contain noisy data, dirty data or unwanted data due to dirty sensors or poor visibility conditions in camera applications. In addition, a machine learning algorithms prediction is solely based on the training data it has seen before. One might argue, that human decisions are also only based on experience they have made since their childhood but humans are often still superior to most algorithms in terms of the instinctive interpretation of complex patterns. Furthermore, they can learn to recognize structural patterns from very few training data.

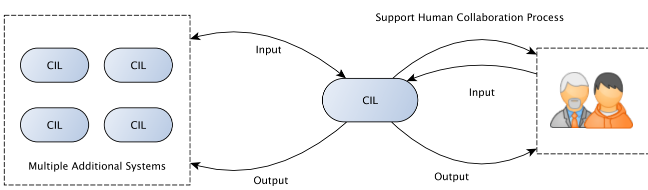


Abbildung 1. Collaborative Interactive Learning.

Hence, *Collaborative Interactive Learning (CIL)* includes domain experts (humans) and other CIL systems in the decision/prediction making process as shown in Figure 1. The relatively new term describes a new generation of systems with

- lifelong learning capabilities in order to continuously improve its knowledge base
- and the ability to exchange knowledge with other CIL systems as well as humans in order to improve the own knowledge and the one of other entities in a bi-directional way [5].

This article gives a brief introduction on possible areas of application for the CIL-approach. It starts with an overview of classical machine learning scenarios, discussing where they are facing limits and motivating the use of CIL. It then lists possble applications and concludes with a view on future research interests.Relation to Organic Computing?

## 2. Algorithmic Foundations

Bevor wir uns genauer den Anwendungen und Herausforderungen von CIL widmen, schauen wir uns zuerst vorhandene machine learning Strategien und Algorithmen an. Wie oben bereits angedeutet, gibt es eine Vielzahl an Definitionen fr den Begriff ML. Hauptschlich hat jedoch Arthur Samuel diesen Begriff im Jahr 1959 geprgt. Seitdem hat sich in diesem Bereich natrlich einiges gendert, jedoch gilt die Definition, dass ML Computern erlaubt Probleme zu lsen, ohne speziell dafr programmiert zu sein, heute immer noch [9].

ber die letzten Jahrzehnte hat sich ML stark weiterentwickelt und ist zu einem Bereich starken Forschungsinteresse vieler verschiedener Akteure geworden, was es schwierig macht, Lern-Strategien in verschiedene Kategorien einzuteilen. Unterschiedliche Forscher nutzen dazu verschiedene Anstze [2] [6]. Im folgenden wollen wir fnf Strategien vorstellen [2].

### 2.1. Supervised learning

Basically supervised learning is the way of learning from labeled examples and making predictions for all unseen points. This is the most basic concept of ML and used in a variety of applications where labeled data is easy to obtain. The spam mail example discussed in the introduction is an instance of supervised learning.

### 2.2. Unsupervised learning

Im Gegensatz zu supervised learning erhlt der Lernende mit dieser Technik ausschlielich unlabeled training data und macht Vorraussagen fr alle ungesehene Punkte. Diese Technik ist relativ hnlich zu der Art, wie Suglinge neue Sachen lernen (Lernen von Sprache) und findet zum Beispiel Einsatz im Clustering von Daten in Gruppen [8].

### 2.3. Semi-supervised learning

Wie der Name suggeriert erhlt der Lernende bei dieser Technik Trainingsdaten, die sowohl aus labeled als auch aus unlabeled data besteht und macht Vorraussagen fr alle ungesehenen Punkte. Diese Technik wird hauptshlich in Bereichen eingesetzt, in denen Daten leicht zu erreichen sind, aber das finden von passenden Labels teuer ist.

### 2.4. Reinforcement Learning

Reinforcement Learning ist der wohl lteste Ansatz und ein potentiell wichtiger Ansatz fr CIL. Der Lernende interagiert mit seiner Umgebung, um neue Informationen zu gewinnen. Fr jede Aktion erhlt er eine Belohnung. Sein Ziel ist es ber mehrere Wiederholungen seine Belohnung zu maximieren. Er befindet sich so jedoch im Zwiespalt zwischen der Entscheidung bereits gewonnene Information auszunutzen oder mehr Informationen durch weiter Aktionen zu erhalten.

### 2.5. Active Learning

Das Ziel von Active Learning ist es, durch geschicktes interaktives Agieren mit anderen Agenten bereits vorhandene Datenpunkte zu beschriften und so eine hnliche Performance bei weniger Informationen zu erreichen, wie in der Methode des supervised learning. Um das zu erreichen, bentigt der Lernende eine Auswahlstrategie [7], die entscheidet ob eine Aktion einen ausreichend groen fr das beschriften eines Datenpunkts bringt oder nicht.

## 3. When is using CIL helpful?

In etablierten ML Systemen muss fr jede spezielle Anwendung ein eigenes Modell im Voraus entworfen und trainiert werden. ML Systeme haben also einen relativ schmalen Anwendungsrahmen. Die Idee bei CIL-Systemen ist, dass sie ein Leben lang lernen und selbstorganisiert Wissen aus verschiedenen Quellen sammeln und auswerten, um so gemeinsam mit anderen technischen und menschlichen Agenten kollaborierend und austauschend Probleme lsen [5]. Der CIL-Ansatz integriert also ganz bewusst andere intelligente Systeme (Menschen und Maschinen) in den Entscheidungsprozess, um fr die einzelnen Entitten ansonsten schwierige Probleme einfacher zu lsen. Ein weiterer Unterschied zu klassischen Anstzen stellt der wechselseitig Nutzen dar. Maschinen profitieren nicht nur von Trainingsdaten anderer Agenten, sondern leisten einen eigenen Beitrag, um Kollaborationsprozesse von Menschen aktiv zu untersttzen, indem sie ihre Wnsche und Bedrfnisse erkennen und entsprechend reagieren [5].

## 4. Classification of CIL in Organic Computing

CIL-Systeme knnen sozusagen teilweise als organisch strukturierte Informationstechnologie verstanden werden, die sogenannte selbst-x-Eigenschaften erfllen [10] [11], sich also

- selbst konfigurieren, in dem Sinn, dass sie nicht von Entwicklern auf einen Anwendungsfall ausgerichtet werden,
- selbst optimieren, indem sie verschiedene Strategien des ML anwenden, um ihr Wissen zu erweitern,
- und sich selbst heilen und schtzen, da Kooperation mit anderen System mglich ist.

Je nachdem welche Methoden des ML eingesetzt werden sind CIL-Systeme mehr oder weniger selbst erklrend [10].

## 5. Application examples

- Clustering
  - (clustering communities on Facebook for group target advertising (politics))
  - Methods: k-Means, DJ-clustering? –¿ add paper ref
- Stock trading
- Health domain
  - Detecting cancer on radon, Prof. Sauer/Forwiss research?
  - Growing number of users are using smartphones sensors and related devices for collating a range of health information, track regularly and detect patterns (noise?, privacy?)
- Industry 4.0
  - quality assurance
- Agriculture
  - Detecting bad grains, → Hackzurich project

## 6. Limitations/Challenges

Challenges

## 7. Conclusion/Future Directions

The conclusion goes here.

## Literatur


[1] Mitchell, T. M. (1997). Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 45(37), 870-877.

[2] Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). Foundations of machine learning. MIT press.

[3] Mitchell, T. M. (2006). The discipline of machine learning (Vol. 9). Carnegie Mellon University, School of Computer Science, Machine Learning Department.

[4] Bachmaier, C. (2017). Lecture. Programming II. Faculty for Informatics and Mathematics, University of Passau, Germany.check this

[5] Sick, B., Oeste-Reiß, S., Schmidt, A., Tomforde, S., & Zweig, A. K. (2018). Collaborative Interactive Learning. Informatik-Spektrum, 41(1), 52-55.

[6] Corne, D., Dhaenens, C., & Jourdan, L. (2012). Synergies between operations research and data mining: The emerging use of multi-objective approaches. European Journal of Operational Research, 221(3), 469-479.

[7] Calma, A., Leimeister, J. M., Lukowicz, P., Oeste-Reiß, S., Reitmaier, T., Schmidt, A., . . . & Zweig, K. A. (2016, April). From active learning to dedicated collaborative interactive learning. In ARCS 2016; 29th International Conference on Architecture of Computing Systems; Proceedings of (pp. 1-8). VDE.

[8] Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop?. Brain Informatics, 3(2), 119-131.

[9] Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3), 210-229.

[10] Müller-Schloer, C., von der Malsburg, C., & Würt, R. P. (2004). Organic computing. Informatik-Spektrum, 27(4), 332-336.

[11] Schmeck, H. (2005, May). Organic computing-a new vision for distributed embedded systems. In Object-Oriented Real-Time Distributed Computing, 2005. ISORC 2005. Eighth IEEE International Symposium on (pp. 201-203). IEEE.