Machine Learning,ML
Prof. Carl Gustaf Jansson
Prof. Henrik Boström
Prof. Fredrik Kilander
Department of Computer Science and Engineering
KTH Royal Institute of Technology, Sweden

Lecture 2
Foundation of Artificial Intelligence and Machine Learning

Welcome to this second lecture of the first week the machine learning course. The theme for this lecture is foundation of artificial intelligence and machine learning. The purpose of this lecture is not just by a single phrase or a sentence but through a more substantial narrative convince you that artificial intelligence is not a new hype, it's a old established area sub area of computer science which has a long history and from the beginning when this area was founded also that machine learning already at that time formed an integral part of the area. So artificial intelligence as a research area has 62 year old roots. It was established already in 1956 where a small group of researchers gathered at Dartmouth College in New Hampshire US. It was just a summer workshop during six to eight weeks that year and as you make an infer this happened only 10 year or less later than the advent of the new of the first computer. So why did this workshop take place at this point in time, so short time after the birth of computer science in general? Part of the answer could be that the bulk of the work in the early days of computer science became focused on very low-level matters where when try to apply the new computer tools on very standardized tasks. In contrast to that a lot of the people who were instrumental in the development of computer science already from the beginning had other ambitions and other visions and that they probably very strongly felt that what they really intended with the new computational devices didn't really happen, not the exciting things they wanted to happen but rather applying the new computers to very much day-to-day use for standardized things. So probably after this decade of preliminary work they thought that maybe it's time to raise the flag again, to raise the flag for the original ambitions for the use of computers and essentially that the potential of computers is to do very advanced things and that this should be put up very high on the agenda again. So at this workshop the group of researchers who gathered there they managed to make a defining statement for the area for artificial intelligence which they agreed on and that statement, I will go through together with you because I think it's important statement and a lot of what is said there still holds so the statement starts 'the study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that the machine can be made to simulate it' one part of this first part of this of the statement that is good to observe on a course on machine learning like this is that the first example given of these kinds of features of intelligence that one want to mimic and simulate is

the aspects of learning, so that statement continues 'an attempt will be made to find how to make machines, use language, form abstractions and concepts, solve kinds of problems now reserved for humans and improve themselves' so similarly in this course it's interesting to observe that our two items on this in this list that relates to learning first number two which is to form abstractions and concepts and number four improve themselves so even in the second paragraph of the statement learning takes a major role the statement ends with a kind of funny paragraph which said it goes as follows 'we think that the significant advance can be made in one or more of these problems if carefully selected group of scientists work on it together for a summer'. It's kind of funny because after 60 years we haven't really made significant advance we have made some advances but still significant advance would say we still wait for but the ambitions at this point in time was very high in this group. So now let's take a look at the people who met at this summer school. they have called be called the founding fathers of artificial intentions one of them Claude Shannon is the founder of information and communications he and he published his seminal work on that in the late 1940s he's also you may know known for having published and written the most famous master's thesis in the history of computer science which was a precursor for his later publications in information and communication theory. DM Mckay, was a well-known British researcher who worked already at this point worked on the borderline between information theory and cognitive science. Julian Bigelow, was the chief engineer for the phenomenal computer at Princeton in 1946 the IAS computer or as is more popular it was called 'the maniac', Nathaniel Rochester, author of the first assembler and a key person in the development of the first commercial computer IBM 701. Oliver Selfridge was named the father of machine perception for his very early work on trying to automate processes similar to the way the human vision system works. Ray Solomonoff, the inventor of algorithmic probability and one of the key persons that very early understood the importance not only to try to create practical learning system but also understand the theoretical limits and restrictions of this processes. John Holland, the inventor of genetic algorithms. Marvin Minsky, one of the key mighty researchers, the founder of the MIT early lab who was very influential in the early development of AI. Allen Newell, the champion for symbolic AI, an inventor of many central AI techniques, his colleague Herbert Simon who was a pioneer in decision-making theory and a Nobel prize winner in economy and finally John McCarty, the founder of the Stanford AI lab and the inventor of the LISP programming language so not only the Dartmouth College summer school happened in the 1950s which is or what is were of relevance for artificial intelligence. A lot of early work already occurred in this decade and i will first mention a few general things that was important for the development of artificial intelligence as a field. So unfortunately one of the key persons from the computer science era was not present in at the Dartmouth College, Alan Turing. But already 1950 Alan Turing had published a paper called 'computing machinery and intelligence' and the program proposed in that paper have been of very big importance for the development of artificial intelligence and also in that paper automation of learning was given a prominent role. Another result that was very early even before the Dartmouth conference was the work at Carnegie Mellon by Allen Newell and Herbert Simon on what is called the 'Logic Theorist'. So the logic theorist is a computer program who tries to mimic the problem-solving

skills of a human being actually it's being called the first artificial intelligence program ever. The purpose of that program was to be able to prove theorems in Whitehead and Russell's principia Mathematica and this program even though some of the techniques wasn't even named at that point introduced techniques that later had a big important for the area such as LISP processing, means-ends analysis and heuristic search and so this is what was a key research coming very early. Later in the 50s Simon and Newell continue that work and when they published results of what they call 'the general problem solver' which is a computer program intended to work as universal problem solver machine which of course then was in some way weaker than a logic theorist because it wasn't so dedicated to one kind a specific category of problem but had at most a wider range of applicability. Also in the 1950s work started on defining programming languages which could be specifically useful to develop artificial intelligence systems and in 1958 John McCarthy created the first version of LISP which had that purpose today is the second-oldest high-level programming language still used only fortran is older by one year. The final result i want to mention is that Oliver Selfridge also one of the Dartmouth participants created a system what he called Pandemonium which was one the first attempt to create a computational model that can mimic pattern recognition of images of course inspired by the models at that time of how the human images recognition system works. So what is interesting to observe is on top the early work as described in artificial intelligence there is a substantial list of things happening in this first decade that is essential for machine learning. So maybe first it's appropriate to mention the person who coined the term machine learning that was Arthur Samuel and he did that in 1959 and Arthur Samuel worked at that point at one of the IBM research labs and one of his tasks there was to look at computer programs that could play check, so this is one of the first example from game playing programs actually and in the context of these work of producing a program that could play checker and compete with human players obviously he observed that he needed learning algorithms so he included in his system some learning mechanisms and when he published these results he was the first to use the tape term 'machine learning'. Going back a little actually before the 50s this the real starting point for work on neural networks was as early as 1943 when McCulloch and Pitts in published their work called 'neural networks as a model of computation', so this was actually the first attempt to look at how we view neural activity in the brain and try to mimic that in a computational model. So this was very early but actually this work also was followed up in the 1950s. So people like Marvin Minsky did some work in the mid 1950s on a system called SNARK which is assumed to be actually that the first really computer implementation of a neural network machine and not much later Frank Rosenblatt published his work on the Perceptron which is today considered the starting point for the artificial neural network development. He was working then at Cornell laboratory and it was a very successful work at first but it turned out that the way at least Frank Rosenblatt presented the Perceptron and the initial applications had severe limitations so after that it took a time since that area really got the pace. Not much later there was another kind of work focused on sub symbolic representations so John Holland introduced in 1960 his first work on genetic algorithms inspired not by neural activities but by Darwin's theory of evolution so this was an even bolder step to look wider for the inspiration for computational models and finally

also in the mid-1950s Ray Slolomonoff published his first work on machine learning which where we had a system which is termed in the inductive inference machine and as I already said Ray Solomonoff became one of the keepers and not only to look at practical application of machine learning but also on a theory of the field. So to sum up this lecture the first point is that this lecture wanted to convey to you what kind of agenda the area artificial intelligence had when it started 60 years ago, the second point is that already from the start machine learning had a key role in the development of this area, the third point is that many of the kinds of machine learning we see today had its roots already at this time so there were several contributions initial contributions to development of neural network at this point. There were several contributions to symbolic computation and learning in symbolic representations already at this time and even this kind of machine learning algorithm based on evolutionary theory had its roots at this time and finally machine learning theory also originated at this time. So thank you thank you for your attention and the next lecture in this series will continue to give you a picture of artificial intelligence and the role of machine learning.

 Thank you