The broad and long way towards modern AI

Der breite und lange Weg zur modernen KI

- Computer Science *Informatik*
- Neuroscience *Neurowissenschaften*
- Public Events *Öffentliche Ereignisse*
- Reinforcement Learning *Bestärkendes Lernen*
- Mathematics *Mathematik*

1650 1700 1750 1800 1850 1900 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030

1676

Chain Rule of Differential Calculus
Gottfried Wilhelm Leibniz developed the chain rule of differentiation. It is the key ingredient of backpropagation, which is the heart of all training of neural networks.

Kettenregel des Differenzierens
Gottfried Wilhelm Leibniz entwickelte die Kettenregel des Differenzierens. Sie ist die wesentliche Komponente des Backpropagation-Verfahrens, das den Kern des Trainierens von neuronalen Netzwerken bildet.

1847

Linear regression
Linear regression is the standard method for regressing linear dependencies between phenomena. It is the first out-of-shelf method, by, for example, econometrics. It estimates the linear relationship between a scalar response and one or more explanatory variables.

Lineare Regression
Die lineare Regression ist die Standardmethode zur Erklärung linearer Abhängigkeiten zwischen Phänomenen. Zum Beispiel in der Ökonometrie, sie schätzt die lineare Beziehung zwischen einer skalaren Antwortvariablen und einer oder mehreren erklärenden Variablen.

1894

Experimental Psychology
"Experimental Psychology" by Robert Sessions Woodworth and Harold Schlosberg, published in 1954, rewrites annual review articles of the field "experimental psychology", which first have been introduced by Conway Lloyd 1894.

Experimentelle Psychologie
"Experimentelle Psychologie" von Robert Sessions Woodworth und Harold Schlosberg, veröffentlicht im Jahr 1954, überarbeitet die Artikel von Jahres zur Jahresübersicht "Experimental Psychology" von Conway Lloyd 1894 erstmalig veröffentlicht wurden.

1927

Conditioned Reflexes
"Conditioned Reflexes" first appeared in an english translation of the work of Pavlov on conditioned reflexes. This introduction of the term was significant because it helped bridge Pavlov's pioneering research in Russia and the emerging field of behaviorism in the English-speaking world.

Konditionierung
Der Begriff "Konditionierte Lernen" taucht erstmalig in einer englischen Übersetzung von Pawlows Werk über bedingte Reflexe auf. Die Einführung dieses Begriffs war bedeutsam, da sie Pawlows bahnbrechende Forschung in Russland mit dem aufkeimenden Feld des Behaviorismus in der englischen Welt verbindet.

1936

The Turing Machine
The Turing machine, introduced by Alan Turing in 1936, is a theoretical model that formalizes the concept of computability. It is the foundation of theoretial informatics.

Die Turing Maschine
Turing Maschine: Die 1936 von Alan Turing eingeführte Turingmaschine ist ein theoretisches Modell, welches das Konzept der Berechenbarkeit formalisiert. Es ist die Basis der Theoretischen Informatik.

1943

McCulloch-Pitts Neuron
This model of a simplified neuron appeared after its creation, was introduced in 1943. It is one of the earliest formal models of a neuron that used binary outputs and weighted inputs to mimic neural processing and decision-making. It is a foundational concept in neural networks and modern LLMs.

McCulloch-Pitts Neuron
Das McCulloch-Pitts-Neuron, so benannt nach seinen Erfindern, erschien nach seiner Vorgestellt. Es ist einer der frühesten formalen Modelle eines Neurons, das binäre Ausgaben und gewichtete Eingaben verwendet, und legte damit die Grundlage für künstliche Neuronale Netzwerke und LLMs.

1950

Grouping and Experiment
"Grouping and experiment" of learning by "grouping and experiment" is a pivotal aspect of his work in psychology and philosophy, particularly in his approach to epistemology, where he emphasizes acquire knowledge and develop skills.

Turkey and Experimentieren
Das Konzept des Lernens durch "Turkey and Experimentieren" von Alexander Bain ist ein zentraler Aspekt seiner Arbeit in der Psychologie und Philosophie, insbesondere in seinem Ansatz, wie Individuen Wissen erwerben und Fähigkeiten entwickeln.

1911

Individuality
The concept of trial-and-error learning, encapsulated in Edward Thorndike's "Law of Effect", is a foundational concept in behavioral psychology.

Individualität
Das Konzept des Lernens durch Versuch und Irrtum, zusammengefasst im "Gesetz der Wirkung" von Edward Thorndike, ist ein grundlegendes Konzept der Verhaltenspsychologie.

1920

Ising's model
Ising's model, predating recurrent NN architectures (the Ising model or Lenz-Ising model) was introduced and analyzed by physicists Ernst Ising and Wilhelm Lenz in the 1920s. It settles into different equilibrium states in response to input conditions, and is the foundation of the first learning RNNs.

Ising Modell
Die erste nicht-lernende rekurrente NN-Architektur (das Ising-Modell oder Lenz-Ising-Modell) wurde in den 1920er Jahren von den Physikern Ernst Ising und Wilhelm Lenz entwickelt und analysiert. Es nimmt bei Eingabe verschiedene stand einen Gleichgewichtszustand und ist die Basis der ersten lernenden RNNs.

1948

Intelligent Machinery
Alan Turing introduced the concept of a reinforcement-like machine learning and referred to it as the "pleasure-pain system".

Intelligente Maschinerie
Alan Turing führte das Konzept eines Bestärkungs-artigen maschinellen Lernens ein und bezeichnete dieses als "Vergnügen-Schmerz-System".

1950

Can machines think?
In this work, Alan Turing asked the question "Can machines think" and analyzed the possible answers and consequences thereof. This famous publication lead to the here presented Turing Game.

Können Maschinen denken?
In dieser Arbeit stellte Alan Turing die Frage, "Können Maschinen denken" und analysierte die möglichen Antworten und deren Konsequenzen. Dieses berühmte Veröffentlichung führte zu dem hier präsentierten Turing Game.

1954

Dynamic Programming and a new formulation in the calculus of variations
Richard E. Bellman introduces in his 1954 paper an automated method for the dynamic decomposition of a large problem in several smaller ones, today well known as Dynamic Programming.

Dynamische Programmierung und ein neuer Formulierung in der Variationsberechnung
In seinem 1954 veröffentlichten Artikel führt Richard E. Bellman eine automatische Methode zur dynamischen Zerlegung eines großen Problems in kleinere ein, heute weitbekannt als Dynamic Programming.

1956

Dartmouth Workshop on AI
The Dartmouth Workshop on AI is considered the founding event of Artificial Intelligence as a field (that's also where the term AI was proposed). A group of around a dozen scientists made important contributions to the then-nascent field throughout this summer workshop.

Dartmouth Workshop zu KI
Der Dartmouth Workshop über KI wird als das Gründungsereignis der Künstlichen Intelligenz als Fachgebiet angesehen (hier wurde auch der Begriff KI vorgeschlagen). Eine Gruppe von etwa einem Dutzend Wissenschaftlern leistete während dieses Sommerworkshops wichtige Beiträge zu dem damals völlig neuen Thema.

1957

A Markovian decision process
With this 1957 paper Richard E. Bellman significantly advanced the theory of dynamic programming, which, in turn, lay the groundwork for many AI algorithms still today.

A Markovscher Entscheidungsprozess
Mit diesem Papier aus dem Jahr 1957 hat Richard E. Bellman die Theorie der Entscheidungsfindung weiter vorangetrieben, was wiederum einen großen Teil der heutigen KI-Algorithmen erst ermöglicht hat.

1951

Stochastic Gradient Descent
Stochastic Gradient Descent (SGD) updates model parameters using single data points or footnotesize batches, making it computationally more efficient and faster than simple GD especially for large datasets A crucial ingredient for modern AI methods.

Stochastischer Gradientenabstieg
Stochastischer Gradientenabstieg (SGD) aktualisiert Modellparameter mittels einzelner Datenpunkte oder kleiner Batches, was es rechnerisch effizienter und schneller macht, speziell für große Datensätze; ein unverzichtbarer Bestandteil moderner KI-Methoden.

1958

Perceptron - the earliest learning neural network
Inspired by the McCulloch-Pitts neuron, the perceptron was the first artificial neuron that could learn its weights. It is a type of an artificial neuron that can learn to classify inputs into different categories by adjusting its weights based on error feedback. This adaptation or learning capability was the significant advancement.

Perceptron - das erste lernende neuronie Network
Inspiriert vom McCulloch-Pitts-Neuron wurde das Perceptron 1958 von Frank Rosenblatt entwickelt. Das Perceptron ist eine künstliches Neuron, das lernen kann, Eingaben in verschiedene Kategorien einzuordnen, indem es seine Gewichte Anpassung aufgrund von Fehler-Feedback erlernt. Diese Anpassungs- oder Lernfähigkeit war der entscheidende Fortschritt.

1960

Steps towards Artificial Intelligence
Marvin Minsky discusses in this issue around AI-Amongst them, he formulates the credit assignment problem: "How do you distribute credit for success among the many decisions that may have been involved in producing it?"

Schritte in Richtung KI
Marvin Minsky diskutiert mehrere Kernprobleme der KI. Unter anderem formuliert er das sogenannte Credit Assignment Problem: "Wie verteilt man das Verdienst für den Erfolg auf die vielen Entscheidungen, die möglicherweise an diesem Ergebnis beteiligt waren?"

1962

Adaline Algorithm
Bernard Widrow created - together with Ted Hoff - ADALINE (Adaptive Linear Neuron). It is a single layer NN with a linear activation function and a learning rule that minimizes the mean square error (MSE) loss.

Adaline Algorithmus
Bernard Widrow erschuf - zusammen mit Ted Hoff - ADALINE (Adaptive Linearer Neuron), ein frühes Modell eines neuronalen Netzwerks mit linearer Aktivierungsfunktion und einer Lernregel zur Minimierung des mittleren quadratischen Fehlers (MSE).

1961

Steps towards Artificial Intelligence
Marvin Minsky discusses in this issue around AI-Amongst them, he formulates the credit assignment problem: "How do you distribute credit for success among the many decisions that may have been involved in producing it?"

Schritte in Richtung KI
Marvin Minsky diskutiert mehrere Kernprobleme der KI. Unter anderem formuliert er das sogenannte Credit Assignment Problem: "Wie verteilt man das Verdienst für den Erfolg auf die vielen Entscheidungen, die möglicherweise an diesem Ergebnis beteiligt waren?"

1964

Pattern-recognizing control system
"Pattern-recognizing control system" by Bernard Widrow and Ted Hoff, is a classic benchmark problemin control theory and reinforcement learning. It involves a cart that can move along a track and a pole attached to the cart by a hinge. The goal is to apply forces to the cart to keep the pole balanced vertically.

Mustererkennung in Steuerungssystemen
Das "Cart-Pole-Problem, formuliert von Bernard Widrow und Ted Hoff, ist ein klassisches Benchmarkproblem in Bestärkendem und von Bestärkendem Lernen. Es handelt sich um einen Wagen, der sich entlang einer Schiene bewegen kann, und einen daran über ein Gelenk verbundenen Stab. Das Ziel besteht darin, Kräfte auf den Wagen anzuwenden, um den Stab vertikal im Gleichgewicht zu balancieren.

1966

First successful learning of deeper architecture
Alexey G. Ivakhnenko employed iterative weight updates for training a multilayered network, but this was not yet the backpropagation algorithm (1970). Up until this point it was the deepest learning algorithm (in contrast to brain animal networks (or perceptrons) with more than 2 layers.

Erstes erfolgreiches tiefes Lernverfahren
Alexey G. Ivakhnenko nutzte iterative Gewichtsverbesserungen für das Trainieren eines mehrschichtigen Netzwerks, aber das war noch nicht der Backpropagation-Algorithmus (1970). Bis zu diesem Zeitpunkt wurde allgemeinen, dass es mehrere Schichten (im Gegensatz zu tierischen Netzwerke (oder Perceptronen) mit mehr als 2 Schichten zu trainieren.

1969

The ReLU function
Kunihiko Fukushima is the first to introduce the rectifying unit with $f(x) = \max(x, 0)$ in a NN-context. This simple activation function was made popular by Geoffrey Hinton et al. - only in 2010 and is now ubiquitous in modern ML models.

Die ReLU-Funktion
Kunihiko Fukushima führte als erster die rektifizierende Einheit $f(x) = \max(x, 0)$ in einem neuronalen Modell ein. Diese Aktivierungsfunktion wurde erst 2010 von Geoffrey Hinton et al. populär gemacht und ist nun in modernen ML Modellen allgegenwärtig.

1969

Perceptrons and the XOR Problem
The book "Perceptrons" by Minsky was a key work that highlighted the limitations of perceptrons to deeper models. Notably, the authors showed that it is impossible to solve the XOR problem with a shallow perceptron.

Perceptrons und das XOR- Problem
Das Buch "Perceptrons" von Minsky war ein entscheidendes Werkzeug, da es die Aufmerksamkeit von flachen Perceptrons auf tiefere Modelle lenkte. Besonders bemerkenswert ist, dass die Autoren zeigten, dass es unmöglich ist, das XOR-Problem mit einem flachen Perceptron zu lösen.

1970

Backpropagation
In his master thesis, Seppo Linnainmaa laid ground for the modern version of the backpropagation algorithm.

Backpropagation
In seiner Masterarbeit legte Seppo Linnainmaa die Grundlagen für die moderne Version des Backpropagation-Algorithmus.

1972

Adaptive Ising Model
Shun'ichi Amari introduced a method for patterns to adaptively and sequence learning using self-organizing networks composed of threshold elements. This work laid the foundation for the development of self-organizing neural networks and unsupervised learning algorithms.

Adaptives Ising Modell
Shun'ichi Amari führte eine Methode zur adaptiven und sequenziellen Musterlernen mittels selbstorganisierender Netzwerke aus Schwellwert-Elementen basiert. Diese Arbeit legte die Grundlage für die Entwicklung selbstorganisierender neuronaler Netzwerke und unüberwachter Lernalgorithmen.

1980

First Convolutional Neural Network (CNN)
Kunihiko Fukushima presented the first neural architecture with an ability to recognize stimulus patterns based on the geometrical similarity of the stimulus without being affected by their positions in the input image.

Erstes Convolutional Neural Network (CNN)
Kunihiko Fukushima präsentierte die erste neuronale Architektur mit der Fähigkeit, Stimulusmuster anhand der geometrischen Ähnlichkeit ihrer Formen zu erkennen, unpanicked ihrer Position im Eingabefeld.

1985

Boltzmann Machines
Dave Ackley, together with Geoffrey Hinton and Terry Sejnowski, presented a learning algorithm for stochastic NNs. Unlike feed-forward learning models, these have so far been of limited practical value, but are still useful for certain applications.

Boltzmann Maschinen
Dave Ackley, gemeinsam mit Geoffrey Hinton und Terry Sejnowski eine lernende Boltzmann-Maschine. Inspiriert von den physikalischen Ising-Modellen wird dieses stochastische neuronale Netzwerk verwendet für bestimmte Anwendungen immer noch nützlich.

1989

Learning from delayed rewards
Christopher Watkins presents a full integration of dynamic programming with learning agents in his PhD-thesis at the Cambridge University.

Lernen von verzögerten Belohnungen
Christopher Watkins präsentiert eine vollständige Integration von dynamischer Programmierung mit Lernmethoden in seiner PhD Arbeit an der Universität Cambridge.

1989

ConvNets by Yann LeCun
Yann LeCun creates the first application of the backpropagation algorithm to handwritten digit recognition. It is a convolutional NN that operates on images instead of the pre-calculated feature vectors commonly used at the time.

ConvNets von Yann LeCun
Yann LeCun erstellt die erste Anwendung des Backpropagation-Algorithmus auf Bildverarbeitung und trainiert dabei direkt an den Bildern, anstelle der damals üblichen vorberechneten Merkmalsvektoren.

1990

First version of dropout
Geoff Hinton et al. introduced a stochastic optimization technique that was later identified as a generalization of DropOut, published only in 2014 by the Geoffrey Hinton's group.

Erste Version von Dropout
Hinton, J. Hassan veröffentlichte eine stochastische Optimierungstechnik, die später als eine Verallgemeinerung von DropOut identifiziert wurde und erst 2014 von der Gruppe von Geoffrey Hinton veröffentlicht wurde.

1997

Long Short Term Memory (LSTM) Networks
Sepp Hochreiters seminal work that solves the problem of vanishing gradient when training deep recurrent neural network architectures.

Long Short Term Memory (LSTM) Netzwerke
Sepp Hochreiters wegweisende Arbeit, die das Problem des verschwindenden Gradienten beim Trainieren tiefer rekurrenter neuronaler Architekturen löst.

1998

LeNet
Yann LeCun unveiled refinements the Toolset for gradientbasierten Lernen on, um ein praktisch nützliches tiefes Netzwerk für die Erkennung handgeschriebener Ziffern zu trainieren. Es wurde kommerziell von Lucent von mehreren Millionen Bank-Schecks pro Tag eingesetzt.

LeNet
Yann LeCun wandte verfeinerte die Technik des gradientenbasierten Lernens an, um ein praktisch nützliches tiefes Netzwerk für die Erkennung handgeschriebener Ziffern zu trainieren. Es wurde kommerziell von Lucent von mehreren Millionen Bank-Schecks pro Tag eingesetzt.

1997

Deep Blue Defeats Garry Kasparov
Chess was considered an activity that requires intelligence and where only humans can excel. Deep Blue, developed by IBM and Carnegie Mellon, showed that it is possible to make a highly specialized machine that can win over the best of humans.

Deep Blue schlägt den Schach-Großmeister Garry Kasparov
Schach wurde als eine Aktivität betrachtet, die Intelligenz erfordert und in der nur Menschen hervorragend zum Schauen. Die Niederlage von Garry Kasparov gegen Deep Blue zeigte, dass es möglich ist, eine hochspezialisierte Maschine zu entwickeln, die die besten Menschen besiegen kann.

2006

Fast Learning Algorithm for Deep Belief Nets
The authors around Geoffrey Hinton present a layer-by-layer training method for Boltzmann-like machines with unsupervised pre-training and supervised fine-tuning.

Schneller Lernalgorithmus für Deep Belief Netze
Die Autoren um Geoffrey Hinton präsentieren eine schichtweise Trainingsmethode für Boltzmann-artige Maschinen mit unüberwachtem Vortraining und überwachter Feinabstimmung.

2011

First CNN on GPU
Jürgen Schmidhuber's Group *wrote* the *first implementation*, *this die für das Training einen CNN (Convolutional Neural Network) erforderlichen Berechnungen multiple einer GPU (Grafikprozessor) parallelisierte.*

Erstes CNN auf GPU
Jürgen Schmidhuber's Gruppe zeigt die erste Implementation, die die für das Training eines CNN (Convolutional Neural Network) erforderlichen Berechnungen multiple einer GPU (Grafikprozessor) parallelisierte.

2012

ImageNet
Alex Krizhevsky et al. significantly improves image recognition through training parallelization on multiple GPUs.

ImageNet
Alex Krizhevsky et al. verbessern die Bilderkennung entscheidend durch die Parallelisierung des Trainings auf mehreren GPUs.

2014

Generative Adversarial Networks
Ian Goodfellow lets two neural networks compete with each other. These resulting image generators are the basis of many modern generative Networks.

Generative Adversarial Networks
Ian Goodfellow lässt zwei neuronale Netzwerke gegeneinander konkurrieren. Die daraus resultierenden Bildgeneratoren sind die Basis der modernen Generativen Netzwerke.

2015

ResNet
Kaiming Hu et al. extended the results presented in Highway networks paper to networks with 1000 layers. Until today, this is the most used paper in machine learning.

ResNet
Die Autoren erweiterten die in der Arbeit über Highway-Netzwerke präsentierten Ergebnisse auf Netzwerke mit 1000 Schichten. Die heute mit über das am häufigsten zitierte Papier im Bereich des maschinellen Lernens.

2017

Transformer
Ashish Vaswani uses an attention mechanism to process entire sequences. Variants of the Transformer are the current basis of all LLMs.

Transformer
Ashish Vaswani verwendet einen Aufmerksamkeitsmechanismus, um gesamte Sequenzen zu verarbeiten. Varianten des Transformers sind die Basis aller aktuellen LLMs.

2022

ChatGPT
Launched in November, 2022, it quickly became the fastest growing application in the human history, with 100 million unique users in a record time. It also marks the beginning of the LLM race.

ChatGPT
Im November 2022 gestartet, wurde es schnell zur am schnellsten wachsenden Anwendung in der Menschheitsgeschichte, mit 100 Millionen einzigartigen Nutzern in rekordverdächtiger Zeit. Es markiert auch den Beginn des LLM-Wettlaufs.

2023

GPT-4
The GPT-4 Technical Report by OpenAI describes GPT-4 as a multimodal model that processes text and image inputs to generate text outputs.

GPT-4
Der technische Bericht zu GPT-4 von OpenAI beschreibt GPT-4 als ein multimodales Modell, das Text- und Bildeingaben verarbeitet, um Textausgaben zu generieren.

2023

Human or Not
An online game inspired by the Turing test, measuring AI chatbot's ability to mimic humans and humans's ability to identify bots in a 3-5 setting, attracted over 15 million users in a month.

Human or Not
Ein Online-Spiel, das vom Turing-Test inspiriert ist und die Fähigkeit von KI-Chatbots misst, Menschen zu imitieren, sowie die Fähigkeit von Menschen, Bots zu identifizieren, zog innerhalb eines Monats über 15 Millionen Nutzer an.

2023

Llama
The LLaMA (Large Language Model Meta AI) paper outlines Meta AI's development of large language models that excel in various NLP tasks.

Llama
Das LLaMA (Large Language Model Meta AI) Papier beschreibt die Entwicklung von großen Sprachmodellen durch Meta AI, die in verschiedenen NLP-Aufgaben hervorragende Leistungen erbringen.

2023

Chatbot Arena
Introduces an open platform for evaluating LLMs based on human preferences.

Chatbot Arena
Stellt eine offene Plattform zur Bewertung von LLMs basierend auf menschlichen Präferenzen vor.