Pitkä ja monitieteinen tie kohti modernia tekoälyä

The broad and long way towards AI

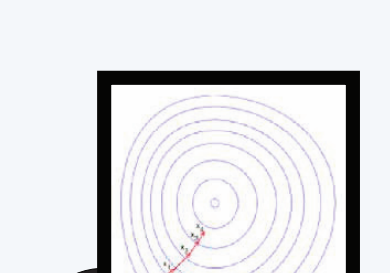
● **Tietojenkäsittelytiede, koneoppiminen**
Computer Science ML

● **Neurotiede**
Neuroscience

● **Julkiset tapahtumat**
Public Events

● **Vahvistusoppiminen**
Reinforcement Learning

● **Matematiikka**
Mathematics



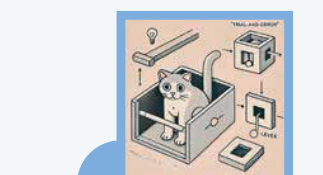
1847

Gradienttimenetelmä

Gradienttimenetelmä on ensimmäisen asteen iteratiivinen algoritmimenetelmä rajoittamattomaan matemaattiseen optimointiin ja paikallisten minimien löytämiseen erityisesti tilanteissa, joissa muuttujia on paljon ja yhtä eksplisiittistä ratkaisua ei ole tai se on erittäin työlästä löytää.

Gradient descent

Gradient descent is a method for mathematical optimization. It is a first-order iterative algorithm for finding a local minimum of a differentiable multivariate function.



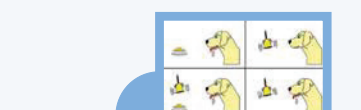
1894

Yritys ja erehdys

Eläintieteilijä ja psykologi C.L.Morgan 1894 käytti termiä *yritys ja erehdys* kuvaamaan eläinten kykyä ratkaista ongelmia. Tästä näkökulmasta oppiminen ei edellytä ylemmän tason älykkyyttä tai tietoisuutta.

Trial and Error

”Experimental Psychology” by Robert Sessions Woodworth and Harold Schlossberg, published in 1954, describe animal behavior using the term ”trial-and error”, which first have been introduced by Conway Lloyd Morgan 1894.



1927

Ehdollistetut refleksit

Termi *vahviste* ilmestyi ensimmäisen kerran Pavlovin ehdollistumisrefleksitutkimusten englanninkielisessä käännöksessä. Termi oli tärkeä siksi, että se rakensi siltaa Pavlovin Venäjälle tekemän pioneerityön ja kasvavan englanninkielisen käyttäytymistieteen tutkimuksen välillä.

Conditioned Reflexes

The term ”reinforcement” 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 with the growing field of behaviorism in the English-speaking world.

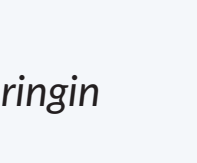
1936

Turingin kone

Alan Turingin vuonna 1936 esittelemä *Turingin kone* on teoreettinen malli tietokoneen toiminnasta. Koneella manipuloidaan symboleja erilaisten sääntöjen mukaan, ja sen avulla voidaan tutkia sitä, mitä ylipäätään voidaan koneellisesti laskea. Turingin koneen periaatteet ovat käytössä kaikissa moderneissa digitaalisissa tietokoneissa. Se loi myös teoreettista perustaa tietojenkäsittelytieteelle.

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 theoretical informatics.



1948

Älykkäät koneet

Alan Turing esitteli ajatuksen koneoppimisesta, jossa oppiminen perustuu palkkioon ja rangaistukseen.

Intelligent Machinery

Alan Turing, introduced the concept of machine learning by effect of the outcome and refered to it as the ”pleasure-pain system”.

1950

Voivatko koneet ajatella?

Teoksessaan *Computing machinery and intelligence* Alan Turing esitti kysymyksen ”Voivatko koneet ajatella?” ja analysoi mahdollisia vastauksia sekä niiden seurauksia. Tämä kuuluisa julkaisu johti täällä esiteltyyn *Turingin peliin* (*imitation game*).

Can machines think?

In his work, Alan Turing asked the question ”Can machines think” and analysed the possible answers and consequences thereof. This famous publication led to the here presented Turing Game.

1954

Dynaaminen ohjelmointi ja uusi formaali lähestymistapa variaatiolaskennassa

Richard E. Bellman esitteli vuonna 1954 julkaistussa artikkelissaan automatisoidun menetelmän suurten ongelmien dynaamiseen pilkkomiseen pienempiin osiin. Menetelmä tunnetaan nykyään nimellä *dynaaminen ohjelmointi*.

Dynamic Programming and a new formalism in the calculus of variations

Richard E. Bellman introduces in his 1954 paper an automatized method for the dynamic decomposition of a large problem in to smaller ones, today well known as *Dynamic Programming*.

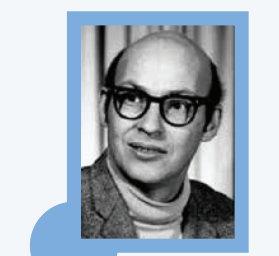
1957

Markovilainen päätöksentekoprosessi

Tässä vuonna 1957 julkaistussa artikkelissa Richard E. Bellman kehitti merkittävästi teoriaa epävarmuuden alaisesta päätöksenteosta. Tämä teoria on edelleen keskeinen monille tekoälyalgoritmeille.

A Markovian decision process

With this 1957 paper Richard E. Bellman significantly advanced the theory about decision-making under uncertainty, crucial for many AI algormithms still today.



1961

Askeleita kohti tekoälyä

Artikkelissaan *Steps toward Artificial Intelligence* Marvin Minsky käsittelee tässä artikkelissa useita tekoälyn keskeisiä kysymyksiä. Yhtenä niistä hän esittää ansioiden jakamisen ongelman: ”Minkä päätöksen ansioksi onnistuminen luetaan, kun monet eri päätökset ovat vaikuttaneet lopputulokseen?”

Steps toward Artificial Intelligence

Marvin Minsky discusses in this paper several core issues around AI. Amonst 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?”

1962

Adaline-algoritmi

Bernard Widrow kehitti yhdessä Ted Hoffin kanssa ADALINE:n (*Adaptive Linear Neuron*), varhaisen neuroverkkomenetelmän, jossa käytettiin lineaarista aktivointifunktiota ja oppimissääntöä, joka minimoi keskineliövirheen (MSE, Mean Square Error).

Adaline Algorithm

Bernard Widrow created - together with Ted Hoff - ADALINE (*Adaptive Linear Neuron*), an early neural network model with linear activation function and a learning rule that minimizes the mean square error (MSE) loss.

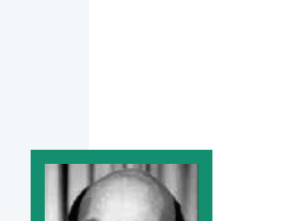
1969

ReLU-aktivointifunktio

Kunihiko Fukushima esitteli ensimmäisenä neuroverkoissa käytetyn ReLU-funktion, jonka laskentakaava on f(x) = max(x, 0). Vaikka funktio yleistyi vasta vuonna 2010 Geoffrey Hintonin ja muiden tutkijoiden työn myötä, se on nykyään käytössä lähes kaikissa koneoppimisen malleissa.

The ReLU function

Kunihiko Fukushima is the first to introduce the rectifying unit with f(x) = max(x, 0) in the neural model. This activation function was made popular by Geoffrey Hinton etal. only in 2010 and is now ubitquitous in modern ML models.



1969

Perseptronit ja XOR-haaste

Marvin Minskyn kirja *Perceptrons* vaikutti suuresti tekoälytutkimukseen siirtämällä huomion yksinkertaisista neuroverkoista monikerroksisiin malleihin. Yksi kirjan tunnetuimmista havainnoista oli, että matalarakenteinen perseptroni ei pysty ratkaisemaan yksinkertaista XOR-logiikkaongelmaa (eksklusiivinen disjunktio, ”joko/tai”).

Perceptrons and the XOR Problem The book

”Perceptrons” by Minsky was a key event as it shifted attentions from shallow perceptrons to deeper models. Notably, the authors showed that it is impossible to solve the XOR problem with a shallow perceptron.

1976

Vastavirta-algoritmi

Seppo Linnainmaa kehitti menetelmän, jolla voidaan analysoida ja vähentää pyöristysvirheitä numeraalisessa laskennassa. Menetelmä lähti laskemaan pyöristysvirheitä takaperoisesti lopputuloksesta alkaen, päätellen pyöristysvirheen suuruuden ainoastaan edellisestä laskutoimituksesta. Näin koko ketjua ei tarvinnut laskea aina uudelleen. Tämä loi perustan nykyaikaiselle vastavirta-algoritmimille.

Backpropagation

Seppo Linnainmaa introduced a method for analyzing and minimizing rounding errors in numerical computations using Taylor series expansion. The Base for modern backpropagation.

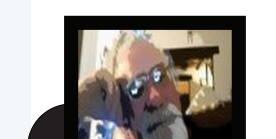
1989

Pidemmän aikavälin hyödyistä oppiminen

Christopher Watkins esitteli dynaamisen ohjelmoinnin ja oppimismenetelmien täyden integroinnin Cambridgen yliopistossa tekemässään väitöskirjassa.

Learning from delayed rewards

Christopher Watkins presents a full integration of dynamic programming with learning methods in his PhD-thesis at the Cambridge University.



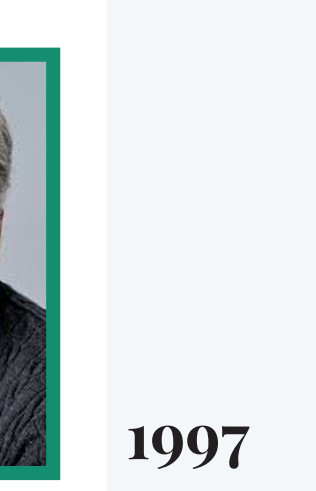
1990

Ensimmäinen DropOut-versio

Stephen J. Hanson julkaisi stokastisen optimoinnin menetelmän, jonka huomattiin myöhemmin olevan eräänlainen *DropOut*-menetelmän varhainen versio. DropOut julkaistiin virallisesti vasta vuonna 2014 Geoffrey Hintonin tutkimusryhmän toimesta.

First version of dropout

Stephen J. Hanson published a stochastic optimization technique that was later identified as a generalization of DropOut, published only in 2014 by the Geoffrey Hinton group.



1997

Pitkäkestoinen lyhytkestomuisti -arkkitehtuuriverkot

Sepp Hochreiter kehitti menetelmän, joka ratkaisi *häviävän gradientin ongelman*. Ongelma esti syviä takaisinkytkeytyviä neuroverkkoja oppimasta pitkiä riippuvuuksia.

Long Short Term Memory Networks

Sepp Hochreiters seminal work that solves the problem of vanishing gradient when training deep recurrent neural network Architectures.

2011

Ensimmäinen CNN GPU:lla

Jürgen Schmidhuber työryhmineen esitteli ensimmäisen toteutuksen, joka rinnakkain laskee konvoluutioneuroverkon (CNN) koulutukseen vaadittavat laskutoimitukset grafiikkaprosessorin (GPU) avulla.

First CNN on GPU

Jurgen Schmidhuber's group shows the first implementation that parallelized the computations required for the training of a CNN on a GPU (Graphical Processing Unit).

2012

ImageNet

Alex Krizhevsky kumppaneineen paransi merkittävästi kuvantunnistusta jakamalla koulutuksen useille GPU:ille.

ImageNet

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

2017

Transformer

Ashish Vaswani käytti huomiomekanismia (*attention mechanism*) koko sekvenssin käsittelemiseen. Transformer-mallit ja niiden variantit muodostavat nykyisten suurten kielimallien (LLM) perustan.

Transformer

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

2018

BERT

Tekijät esittivät arkkitehtuurin, joka on suunniteltu esikouluttamaan syviä kaksisuuntaisia representaatioita merkitsemättömästä tekstistä, ottamalla kaikissa kerroksissa samanaikaisesti huomioon sekä sanaa edeltävän että sitä seuraavan kontekstin.

BERT

The authors proposed an architecture designed to pretrain deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers.

2023

GPT-4

OpenAI:n kehittämä multimodaalinen kielimalli, joka kykenee käsittelemään sekä teksti- että kuvasyötteitä ja tuottamaan niistä tekstimuotoisia vastauksia.

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.

2023

Human or Not

Verkossa pelattava peli, joka on saanut inspiraationsa Turingin testistä ja mittaa tekoälychatbottien kykyä matkia ihmisiä sekä ihmisten kykyä tunnistaa botteja 1:1-tilanteessa. Peli houkutteli kuukaudessa yli 1,5 miljoonaa käyttäjää.

Human or Not

An online game inspired by the Turing test, measuring AI chatbot's ability tomimic humans and humans's ability to identify bots in a 1:1 setting, attracted over 1.5 million users in a month.

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

1676

Differentiaalilaskennan ketjusääntö

Gottfried Wilhelm Leibniz kehitti ketjusäännön, joka on vastavirtagradienttilaskennan ytimessä (esimerkiksi: Jos auto liikkuu kaksi kertaa nopeammin kuin pyörä, ja pyörä liikkuu neljä kertaa nopeammin kuin kävelijä, liikkuu auto kahdeksan kertaa kävelijää nopeammin. Tämän laskemiseen voidaan käyttää kaavaa 2*4 = 8. Vastavirtamenetelmä on tehokas ketjusäännön sovellus neuroverkkojen kouluttamiseen.

Chain Rule of Differential Calculus

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

1847

Lineaarinen regressioanalyysi

Lineaarinen regressioanalyysi on empiirisen tutkimuksen yleisesti käytetty tilastollinen analysointimenetelmä, jonka avulla arvioidaan tutkittavan ilmiön lineaarisia riippuvuuksia selittävistä muuttujista.

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

1850

Hapuilu ja kokeilu

Psykologi ja filosofi Alexander Bain käytti termiä *hapuilu ja kokeilu* pyrkiessään selittämään, miten ihmiset hankkivat tietämystä ja kehittävät uusia taitoja.

Groping and Experiment

Alexander Bains concept of learning by ”groping and experiment” is a pivotal aspect of his work in psychology and philosophy, particularly in his approach to understanding how individuals acquire knowledge and develop skills.

1911

Yksilöllisyys

Edward Thorndiken *vaikutuksen laki* kiteytti käyttäytymistieteen peruskäsitteisiin kuuluvan yritys ja erehdys -oppimisen. Toiminta, joka tuottaa tietyssä tilanteessa tyydyttävän vaikutuksen, toistuu todennäköisemmin kuin toiminta, joka tuottaa epämiellyttävän vaikutuksen.

Individuality

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

1920

Lenz-Ising-malli

Fyysikot Ernst Ising ja Wilhelm Lenz esittelivät ensimmäisen ei-oppivan *Ising-mallin*. Se on tilastollisen mekaniikan teoreettinen perusmalli, jota käytetään kuvaamaan ilmiöitä, joissa on vuorovaikutuksia kahden tilan välillä. Malli asettuu eri tasapainotiloihin vastauksena syöttöehtoihin. Se on ensimmäisten oppivien takaisinkytkeytyvien neuroverkkojen perusta.

Ising's model

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

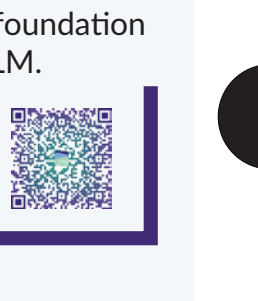
1943

McCulloch-Pitts-hermosolumalli

McCulloch ja Pitts käyttivät lauselogiikkaa tarkastellakseen hermoston toimintaa verkostona. He oivalsivat verkoston toiminnan perustuvan viime kädessä siihen, reagoiko yksittäinen hermosolu saamaansa ärsykkeeseen. Keksijöidensä mukaan nimetty *McCulloch-Pitts-neuroni* on siten formaali malli hermosolusta. Mallinnus on ollut perustana keinotekoisille hermoverkoille ja moderneille laajoille kielimalleille.

McCulloch-Pitts Neuron

The McCulloch-Pitts neuron, named after its inventors, 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 laid the foundation for artificial neural networks and modern LLM.



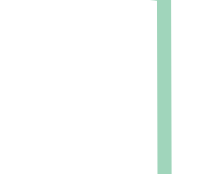
1951

Stokastinen gradienttimenetelmä

Stokastinen gradienttimenetelmä päivittää mallin parametreja käyttämällä yksittäisiä havaintoja tai pieniä datan eräjoukkoja. Tämä tekee siitä laskennallisesti tehokkaamman ja nopeamman kuin tavallinen gradienttimenetelmä, erityisesti suurilla aineistoilla. Tämä on olennainen osa nykyaikaisia tekoälymenetelmiä.

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.



1956

Dartmouthin tekoälyseminaari

Dartmouthin kesällä 1956 järjestettyä tekoälyseminaaria pidetään yleisesti tekoälytutkimuksen lähtölaukauksena. Seminaarissa otettiin ensimmäistä kertaa käyttöön termi *tekoäly* (*Artificial Intelligence*). Noin tusina eri alojen tutkijaa osallistui tapahtumaan ja edisti merkittävällä tavalla tekoälyn tutkimuksen varhaista kehitystä.

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.

1958

Perseptroni – varhaisin oppiva neuroverkkomalli

Frank Rosenblatt kehitti McCulloch-Pitts-neuronista inspiroituneena perseptronin vuonna 1958. Perseptroni on eräänlainen keinotekoinen neuroni, joka oppii luokittelemaan syötteitä eri luokkiin säätämällä painojaan virhepalautteen perusteella. Tämä osoitti, että kone voi oppia kokemuksen kautta ilman, että kaikki säännöt ohjelmoidaan suoraan.

Perceptron - the earliest learning neural network model

Inspired by the McCulloch-Pitts neuron, the perceptron was developed by Frank Rosenblatt in 1958. The perceptron is a type of 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.

1964

Mallintavat ohjausjärjestelmät

Bernard Widrow ja Ted Smith esittelivät *vaunu-tanko-ongelman* (*Cart-Pole-problem*). Siitä tuli klassinen testi säätöteoriassa ja vahvistusoppimisessa. Se koostuu radalla liikkuvasta kärrystä ja siihen kiinnitetyistä tangosta, jonka pystysuuntainen tasapaino pyritään ylläpitämään ohjaamalla kärryä.

Pattern-recognizing control systems

The Cart-Pole problem, formulated by Bernard Widrow and Ted Smith, is a classic benchmark in control theory and reinforce-ment learning. It involves a cart that can move along a track and a pole attached to the cart via a pivot. The objective is to apply forces to the cart to keep the pole balanced vertically.

1966

Ensimmäinen läpimurto syvien neuroverkkojen oppimisessa

Alexey G. Ivakhnenko onnistui opettamaan monikerroksisen neuroverkon käyttämällä toistuvia painojen päivityksiä – vaikkei kyse ollut vielä tunnetusta vastavirta-algoritmista (*backpropagation*), joka kehitettiin vasta vuonna 1970. Siihen asti ajateltiin, että yli kahden kerroksen neuroverkkoja ei voi tehokkaasti opettaa.

First successful learning of deeper architectures

Alexey G. Ivakhnenko employed iterative weigth updates for training a multilayered network, but that was not yet the backpropa-gation rule (1970). Up until this point it was widely assumed that it is close to impossible to train neural networks (or perceptrons) with more than 2 layers.

1970

Vastavirta-algoritmi

Seppo Linnainmaa loi pro gradu -tutkielmassaan perustan vastavirta-algoritmin nykyaikaiselle versiolle.

Backpropagation

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

1972

Mukautuva Ising-malli

Shun'ichi Amari kehitti menetelmän, jossa itsejärjestäytyvät kynnysarvoelementeistä koostuvat verkot kykenivät sarjaoppimiseen ja tunnistamaan kuvioita ilman valvottua opetusta. Tämä olitärkeä askel kohti itseohjautuvia neuroverkkoja ja ohjaamatonta koneoppimista.

Adaptive Ising Model

Shun'ichi Amari introduced a method for pattern recognition and sequence learning using self-organizing networks composed of threshold elements. This work laid the foun-dation for the development of self-organizing neural networks and unsupervised learning algorithms.

1980

Ensimmäiset konvoluutioneuroverkot

Kunihiko Fukushima esitteli ensimmäisen neuroarkkitehtuurin, joka pystyi tunnistamaan eri ärsykekuvioita geometrisen samankaltaisuuden perusteella, riippumatta niiden asemoinnista lähdekuvassa.

First Convolutional Neural Networks, CNN

Kunihiko Fukushima presented the first neural architecture with an ability to recognize stimulus patterns based on the geometrical similarity of their shapes without being affected by their positions in the input image.

1985

Boltzmannin koneet

Dave Ackley, Geoffrey Hinton ja Terry Sejnowski loivat oppivan Boltzmannin koneen. Se oli saanut inspiraationsa Ising-malleista. Oppivasta Boltzmannin koneesta on toistaiseksi ollut vain rajallisesti käytännön hyötyä, mutta sitä käytetään silti tietyissä erityistapauksissa.

Boltzmann Machines

Dave Ackley, together with Geoffrey Hinton and Terry Sejnowski, presented a learning Boltzmann machine. Inspired by the physical Ising models, these have so far been of limited practical value but are still useful for certain Applications.

1989

Yann LeCunin neuroverkot

Yann LeCun kehitti vastavirta-algoritmista ensimmäisen sovelluksen, jota pystyttiin hyödyntämään käsin kirjoitetun tekstin tunnistamisessa. Neuroverkkoa koulutettiin oikeilla kuvilla, mikä poikkesi tuolloin yleisestä tavasta hyödyntää kouluttamisessa valmiiksi laskettuja piirrevektoreita.

Conv Nets by Yann LeCun

Yann LeCun creates the first application of the backpropagation algorithm to handwriting recognition, by training directly on real images instead of the pre-calculated feature vectors commonly used at the time.

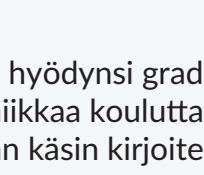
1997

Deep Blue voittaa Garry Kasparovin

Shakkia pidettiin älykkyyttä vaativana lajina, jossa vain ihminen saattoi tulla mestariksi. Deep Bluen voitto maailmanmestari Garry Kasparovista kuitenkin osoitti, että jopa ihmisistä parhaan pystyi päihittämään huippuunsa hiotulla koneella. Deep Blue teki tämän hyödyntäen niin kutsuttua ahnetta algoritmia valitakseen tarjolla olevista siirroista lyhyellä aikavälillä parhaan.

Deep Blue Defeats Garry Kasparov

Chess was considered an activity that requires intelligence and where only humans can excel. The defeat of Garry Kasparov by Deep Blue showed that it is possible to make a highly specialised machine that can defeat the best of humans, even though by performing greedy tree search of possible moves.



1998

LeNet

Yann LeCun hyödynsi gradientti-pohjaista oppimistekniikkaa kouluttaakseen syväverkon tunnistamaan käsin kirjoitettuja numeroita. Menetelmä otettiin käyttöön mm. pankeissa, joiden järjestelmät käsittelivät miljoonia käsin kirjoitettuja sekkejä päivässä.

Le Net

Yann LeCun successfully applied gradientbased learning technique for training a practically useful deep network for handwritten digit recognition. It was employed commercially for reading bank checks (several million checks per day).

2006

Nopeasti oppiva algoritmi syväuskomusverkoille

Geoffrey Hintonin johtama tutkijaryhmä esitteli kerroksittain etenevän koulutusmenetelmän Boltzmannin koneiden kaltaisille malleille. Menetelmässä käytetään valvomatta tapahtuvaa esikoulutusta ja valvottua hienosäätöä.

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.

2014

Generatiivinen kilpaileva verkosto

Ian Goodfellow kehitti menetelmän, jossa kaksi neuroverkkoa kilpailee toisiaan vastaan. Näistä kuvageneraattoreista tuli monien nykyaikaisten generatiivisten verkkojen perusta.

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.

2015

ResNet

ResNetin kehittäjät laajensivat Highway networks-julkaisun ideaa jopa 1000 kerroksen syvyisiin neuroverkkoihin. Tämä mahdollisti erittäin syvien mallien tehokkaan kouluttamisen. ResNet on edelleen yksi koneoppimisen historian eniten siteeratuista tutkimuksista.

ResNet (Residual Network)

The authors extended the results presented in Highway networks paper to networks with 1000 layers. Until 2023, this was the most cited paper in machine learning.

2022

ChatGPT

Marraskuussa 2022 lanseeratun ChatGPT-sovelluksen käyttö laajeni ennätyksellistä vauhtia, saavuttaen 100 miljoonaa uutta käyttäjää vain kahdessa kuukaudessa. Suurten kielimallien kilpajuoksu alkoi.

ChatGPT

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

2023

Llama

Meta AI:n julkaisu LLaMA (*Large Language Model Meta AI*) esitteli Metan suurten kielimallien kehitystyötä. Meta AI:n LLaMA-mallit suoriutuvat poikkeuksellisen hyvin luonnollisen kielen käsittelyssä.

Llama

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

2023

Chatbot Arena (nyk. LMArena)

LCM eli Large Model Systems Organization esitteli avoimen ja interaktiivisen alustan, joka mahdollisti suurten kielimallien vertailukelpoisen arvioinnin.

Chatbot Arena

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

Sisällön alkuperäinen tekijä: Software Competence Center Hagenberg GmbH
http://play.turinggame.ai
Suomennokset: Tekniikan Museo