

Künstliche Intelligenz

Wohin gehen wir?

Lutz Donnerhacke

Lange Geschichte

- 1956: Beginn ernsthafter Forschung
 - „Aufgaben lösen, zu deren Lösung Intelligenz notwendig ist, wenn sie vom Menschen durchgeführt werden“ (Minsky)
 - Phase grenzenloser Euphorie
- Regelbasierte Systeme (Expertensysteme)
 - 1966: Eliza (Gesprächssimulation)
 - Schach, Dame, Logische Probleme
- Neuronale Netze
 - 1943 – 1961: Erste Versuche (gescheitert)
 - 1972 – 1980: Neustart (selbstlernende Systeme)

Enttäuschte Hoffnungen

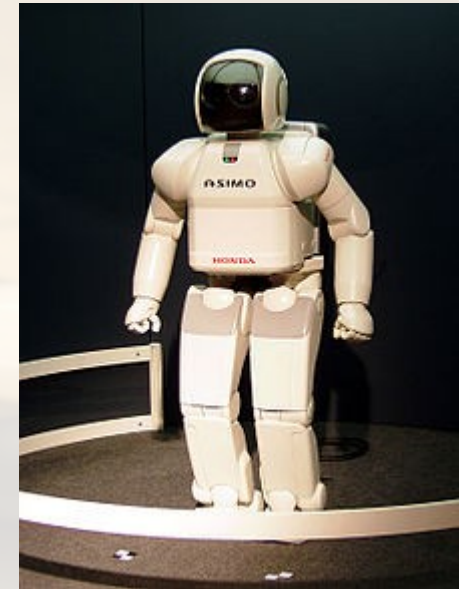


- 1989: 0:2
Deep Thought
- 1997: 3,5:2,5
Deep Blue
- Erkenntnisse
 - Keine!
 - Computer sind halt schneller



Gesteuerte Systeme

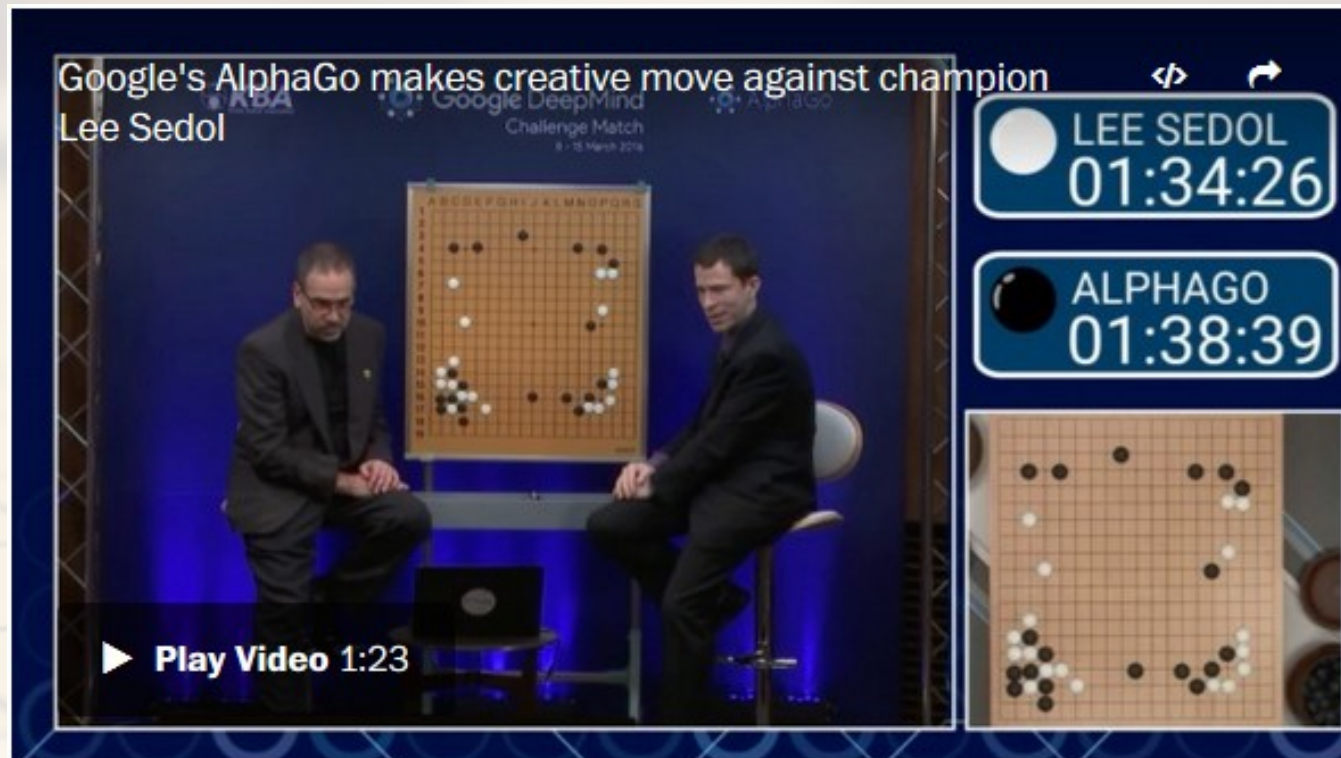
- Weitgehend Verzicht auf *Lerneffekte*
 - Programmierung von Verhaltensmustern
 - Komplexe Systeme aus Modulen
- Asimo – niedlichster Vertreter
 - [youtube.com/watch?v=JIRPICfnmhw](https://www.youtube.com/watch?v=JIRPICfnmhw)
- Stetige Verbesserungen
 - Keine Durchbrüche erwartet



AlphaGo

- Go ist *wesentlich* komplexer als Schach
 - Suchraum zu groß für primitive Durchmusterung
 - Go-Programme auf Amateurniveau
- Auftritt AlphaGo
 - Nov. 2015: 5:0 gegen Fan Hui (Europameister)
 - Jedoch mäßige Spielweise (wie Deep Blue)
 - März 2016: Match gegen Weltmeister Lee Sedol
 - Erwartung: Lee gewinnt (sagt er auch selbst)

Die (eigentliche) Überraschung



During its second match against Go champion Lee Sedol, Google's DeepMind program AlphaGo made a highly unexpected move that left the commentators stunned. (Google DeepMind)

<https://www.youtube.com/watch?v=l-GsfyVCBu0>

Unmittelbare Reaktionen

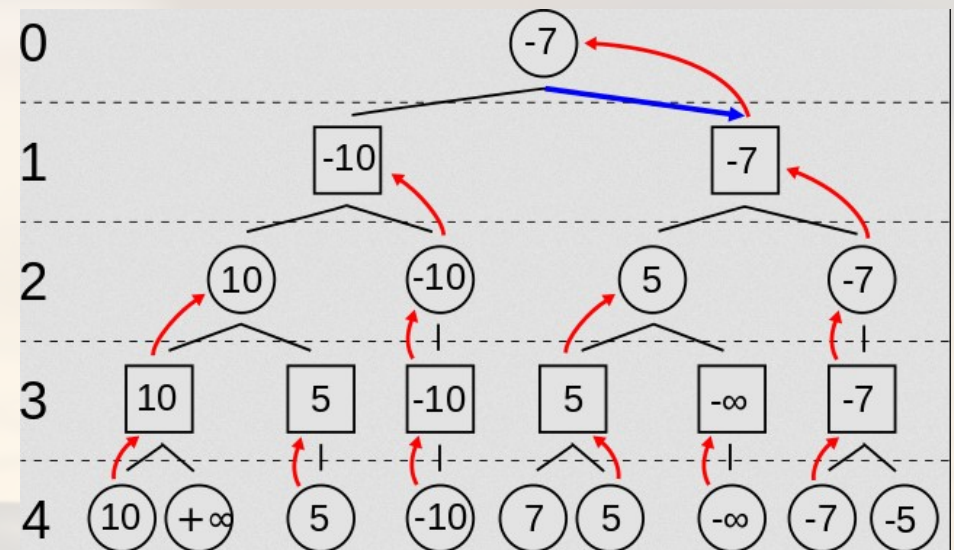
- *Ich weiß nicht ob es ein guter oder schlechter Zug war. Es ist ein sehr seltsamer Zug.* Michael Redmond (verstört)
- *Ich denke, es ist ein Fehler.* Chris Garlock (lachend)
- Reaktion Sedol:
 - Starrt auf das Brett, verlässt den Raum
 - Braucht 16 Minuten bis zum Antwortzug
- Kommentatoren ziemlich hilflos in der Zeit
- Fazit:
 - Seltsam aber kein Fehler
 - Brillanter Zug!
 - Sedol erholt sich nicht, verliert.

Bedeutung des Ereignisses

- Go ist ein *intuitives Spiel*
 - Hauptsächlich Positionsbewertung
 - Spielzüge deuten nur an, beendet wird kaum
 - Ergebnisermittlung durch *einvernehmliche* Fortführung der angedeuteten Spielzüge
 - Feststellung des Siegers kann dauern
- Maschinen hilft Geschwindigkeit nicht
 - *Gefühl* für Situationen notwendig

Minimax-Lösung

- Alle Züge bewerten
- Für den Gegner spielen
- Alle Züge umgekehrt bewerten
- Beste Bewertung übernehmen
- Suchen so tief und so lange die Zeit reicht



AlphaGo

- Bewertung durch ein Neuronales Netz
- Vorauswahl der Züge durch ein anderes Netz
 - Nicht mehr alle Züge durchprobieren
- Standardstrukturen klassisch erkennen
 - Ausbildung von Treppen, Gefangennahme
 - Verhindert Lernen in klaren Situationen
- Lernen anhand von bekannten Partien
 - Dann spielen gegen sich selbst (*Millionenfach*)

Funktionsweise neuronaler Netze

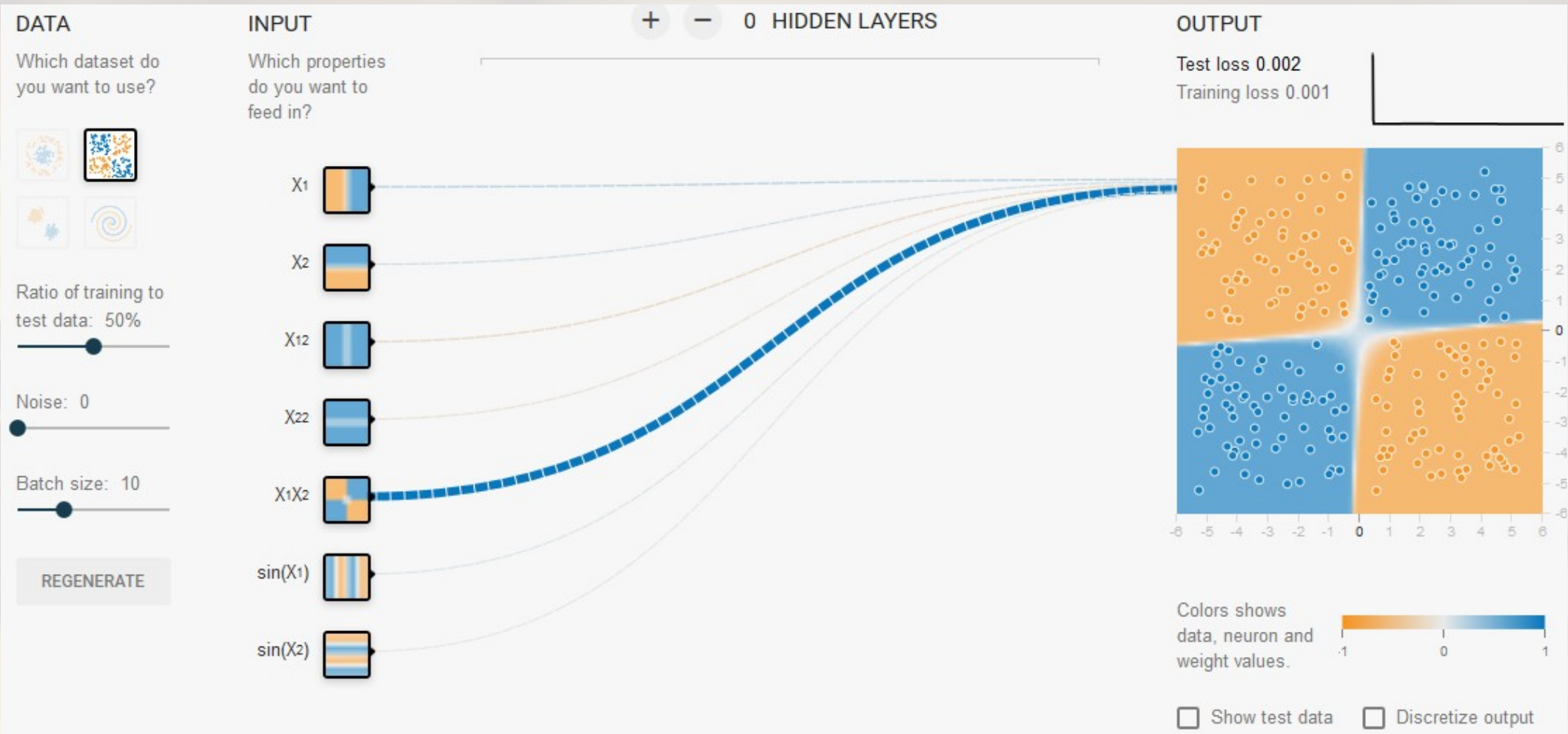
<http://playground.tensorflow.org/>

- Spielwiese zum einfachen Üben
- Eingänge
 - *Was darf das Programm sehen?*
- Lernarten
 - *Wie soll gelernt werden (eher technisch)*
- Struktur des Netzes
 - *Welche Zwischenstufen werden erlernt?*

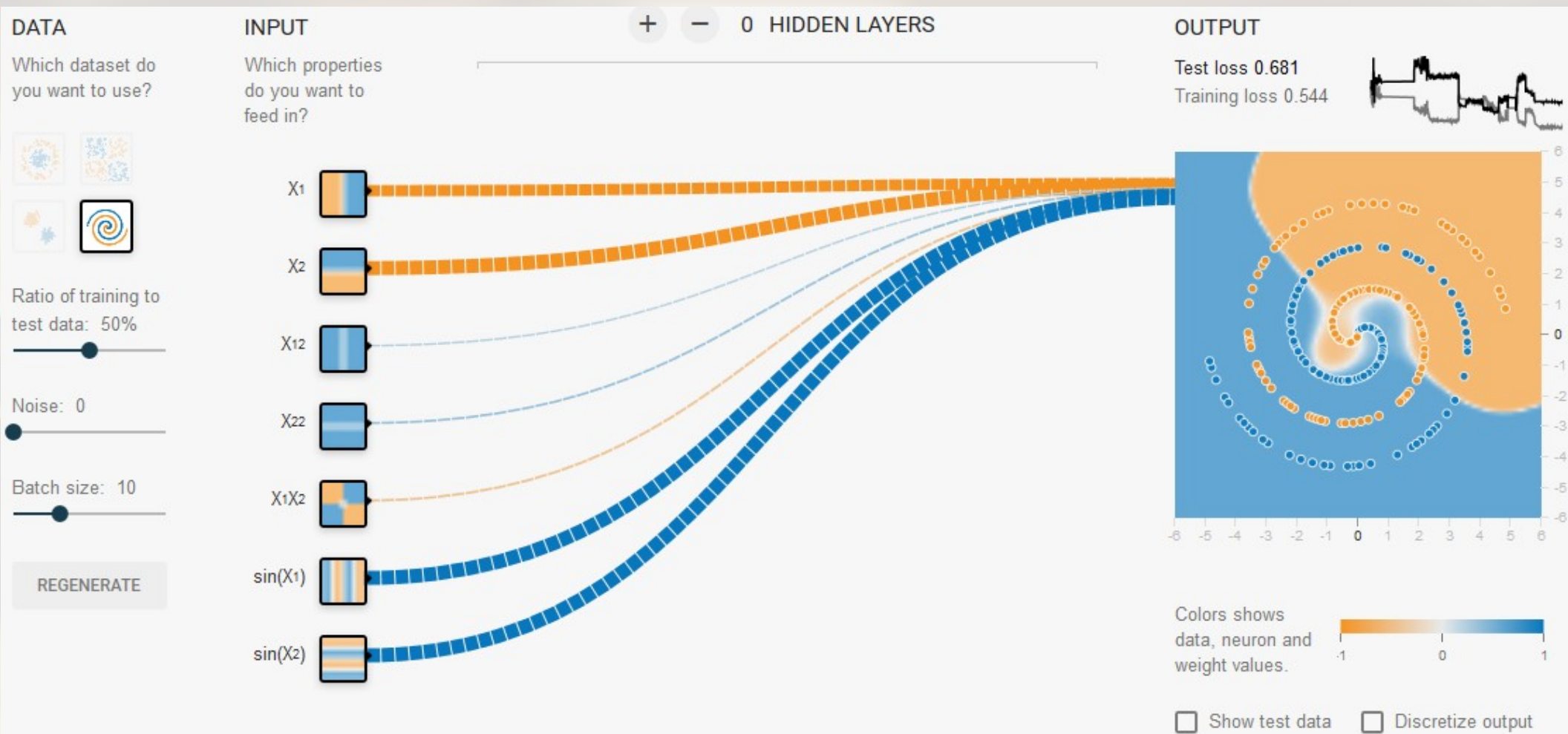
Wichtung der Eingänge



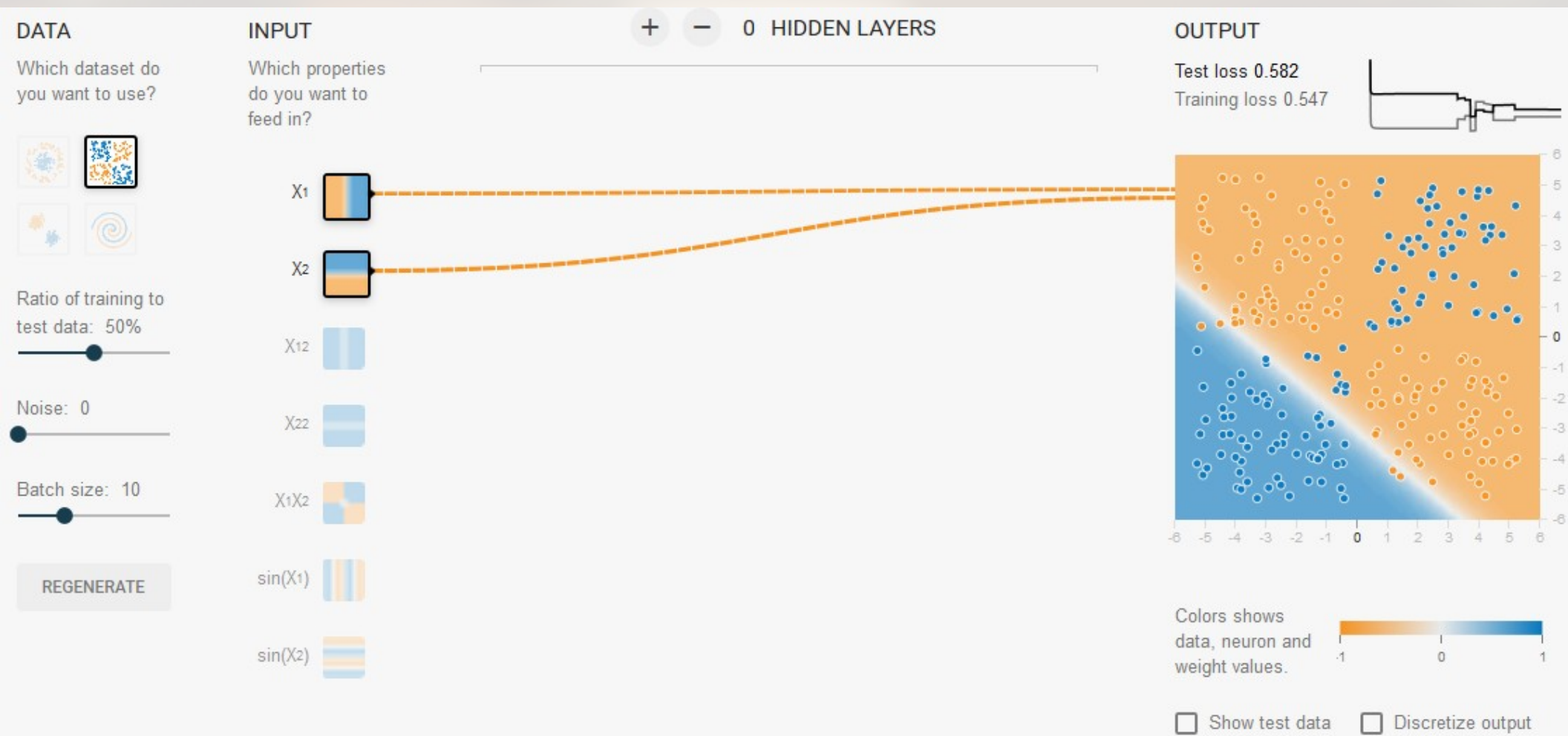
Wichtung der Eingänge



Wichtung der Eingänge



Wichtung der Eingänge



Wichtung der Eingänge

- Einfache Strukturen direkt abbildbar
- Unzureichende Eingaben nicht kompensierbar
 - Ausbildung fehlerhafter Strukturen (hoher Loss)
- „Lernen“ ist einfache Anpassung
- Funktionsweise eines *Neurons*
 - Wichtung von Eingängen
 - Überlagerung der Eingänge
 - Triggerung eines Schwellwerts

Neuronale Netzwerke

- Verknüpfung vieler solcher Neuronen
- Jedes Neuron lernt seine Eingänge zu wichten
- Der Ausgang jedes Neurons ist ein Eingang anderer Neuronen
- Training durch Versuch und Irrtum
 - Positive und Negative Bewertungen
 - Kleine Veränderungen
 - Viele (sehr viele) Trainingsläufe

Neuronale Netzwerke

DATA

Which dataset do you want to use?



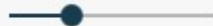
Ratio of training to test data: 50%



Noise: 0



Batch size: 10



REGENERATE

INPUT

Which properties do you want to feed in?

X1



X2



X12



X22



X1X2



sin(X1)



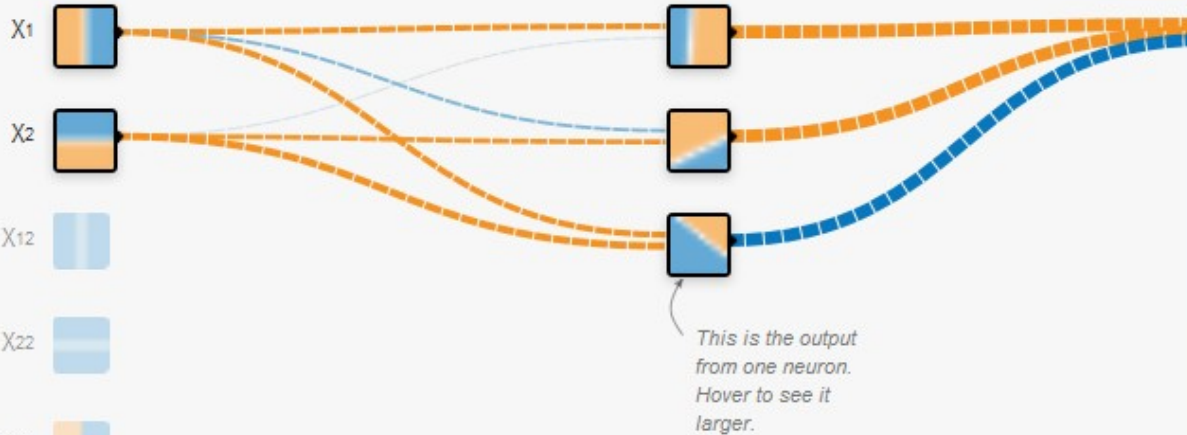
sin(X2)



+ - 1 HIDDEN LAYER

+ -

3 neurons

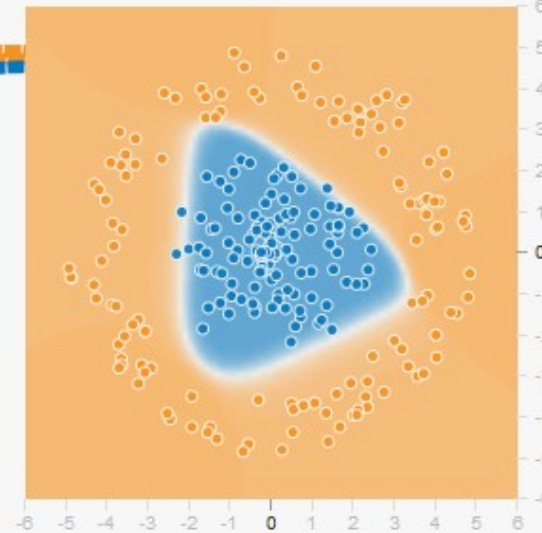


This is the output from one neuron. Hover to see it larger.

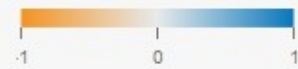
OUTPUT

Test loss 0.017

Training loss 0.013



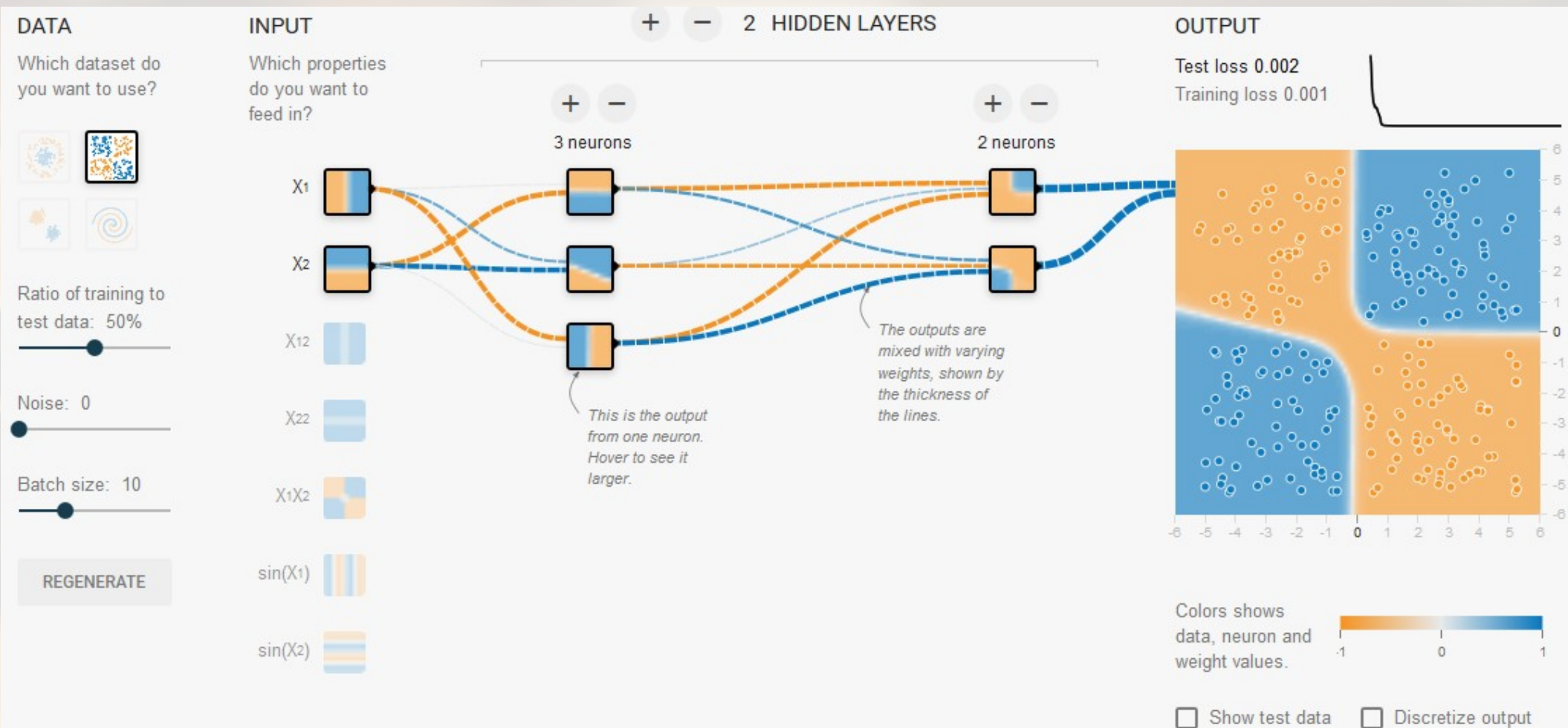
Colors shows data, neuron and weight values.



Show test data

Discretize output

Neuronale Netzwerke

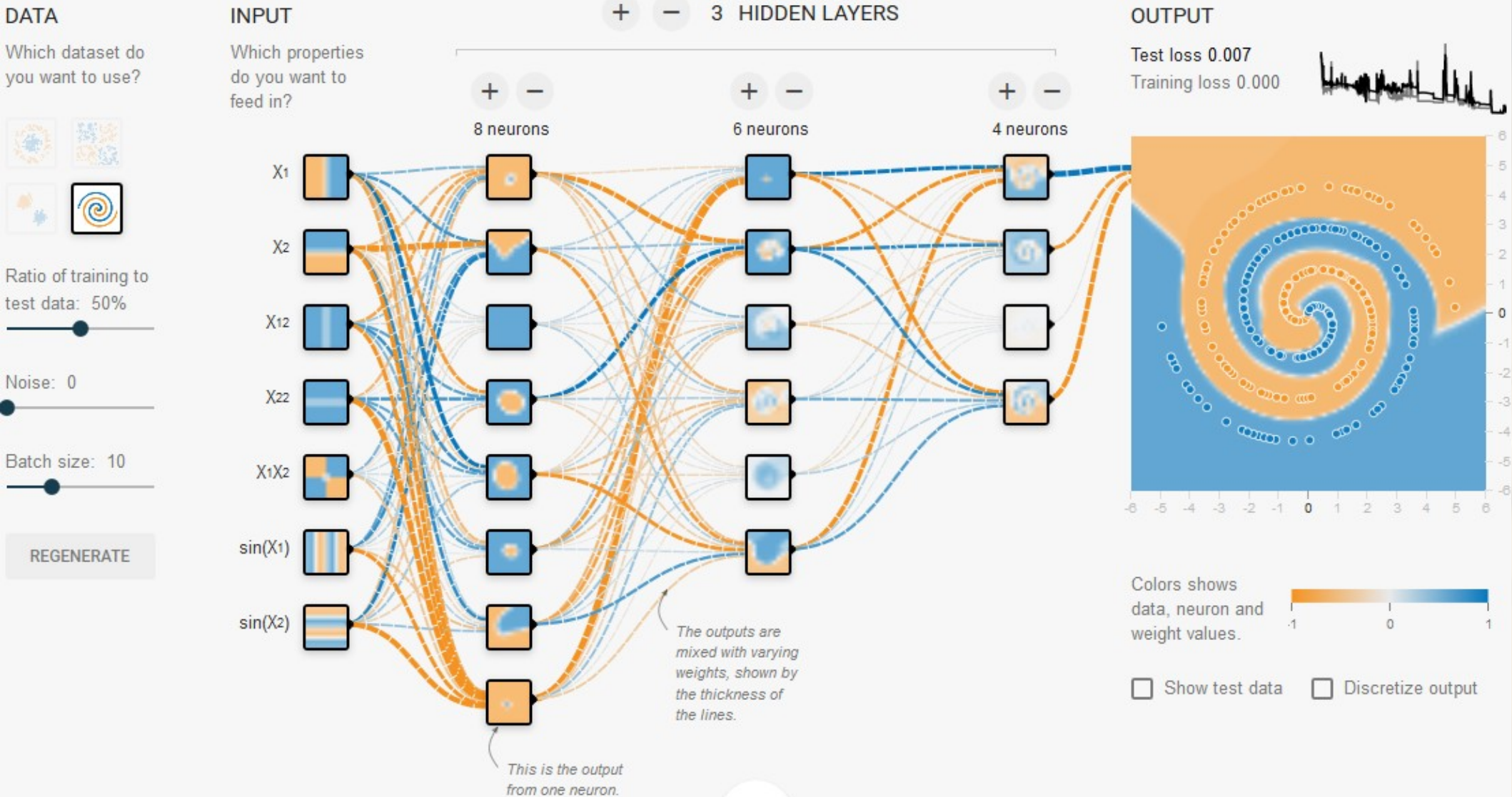


Neuronale Netzwerke

Zwischenschritte

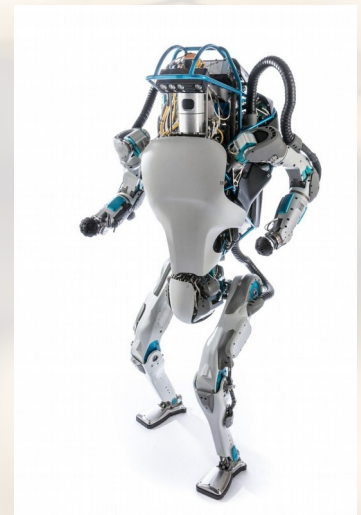
- Bilden neue Erkenntnisse ab
 - Haben eigenständige Bedeutung
- Können fehlende Eingaben ableiten
 - Erkennung der Eckstruktur
- Zwischenschritte bauen aufeinander auf
 - Meta-Erkenntnisse
- Unwichtige Knoten verlieren an Funktion

Neuronale Netzwerke



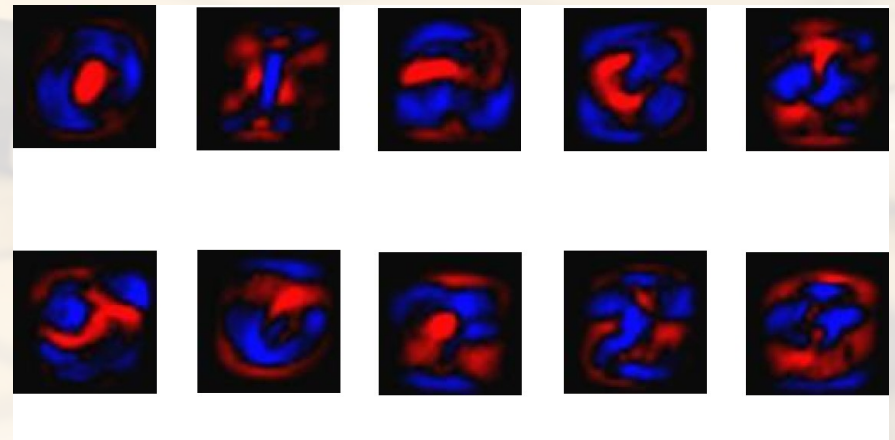
Neuartige Programmierung

- Statt Algorithmen nimmt man Netze
- Netze werden *trainiert* nicht *programmiert*
 - Programmierer kennt Funktionsweise *nicht*
- Hoch variabler Einsatzzweck
- Leichte Anpassung an neue Bedingungen
 - Auch ungeplante Situationen
- Beispiel Atlas-Roboter
 - [youtube.com/watch?v=rVlhMGQgDkY](https://www.youtube.com/watch?v=rVlhMGQgDkY)



Erkennung von Handschriften

- Frühe Forschung
- MNIST Daten
- Zwischenschritte
 - Rot – kein Punkt
 - Blau – Punkt
- Andere Modellierung als der Mensch!



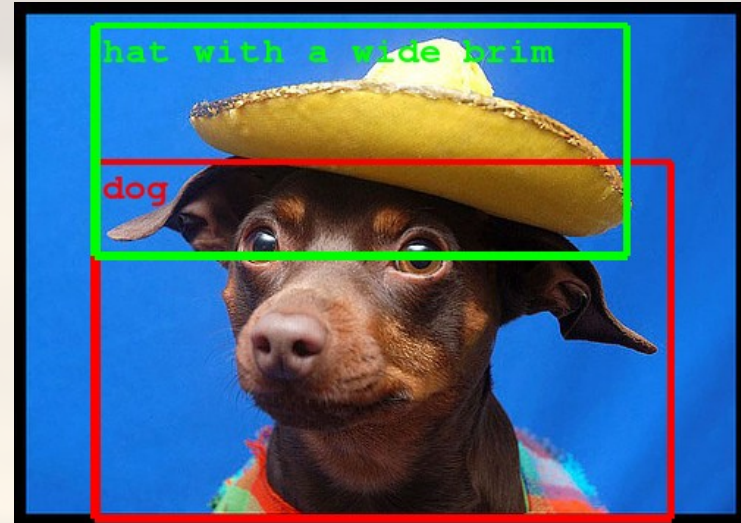
Seltene Handschriften

- Erkennung altjüdischer Texte (ca. 600 v.u.Z.)
 - Erkenntnis, das damals Lesen und Schreiben allgemein üblich war



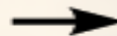
Bilderkennung

- Beschreibe ein Bild
- Verwendung im Alltag
 - Autofahren
 - Bilder suchen
- Selbstlernend
 - Trainiert auf bestimmte Typen:
Bus, Auto, Hut, Hund



Bilderkennung

- Neuronales Netz (Zwischenschritte)
 - Erste Schicht erkennt Kanten
 - Ästhetisch wertvoll

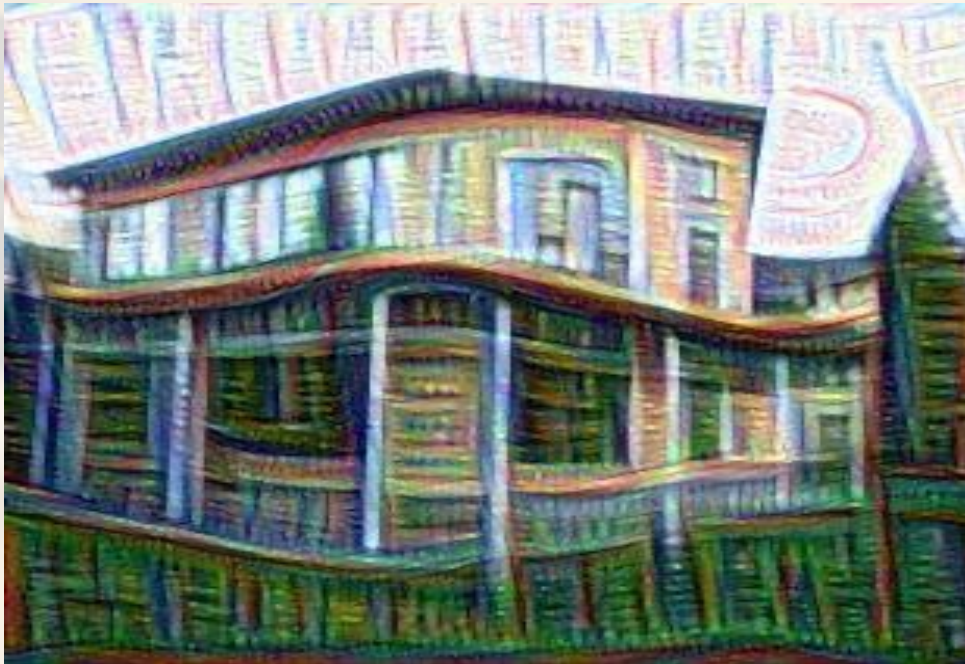


Bilderkennung

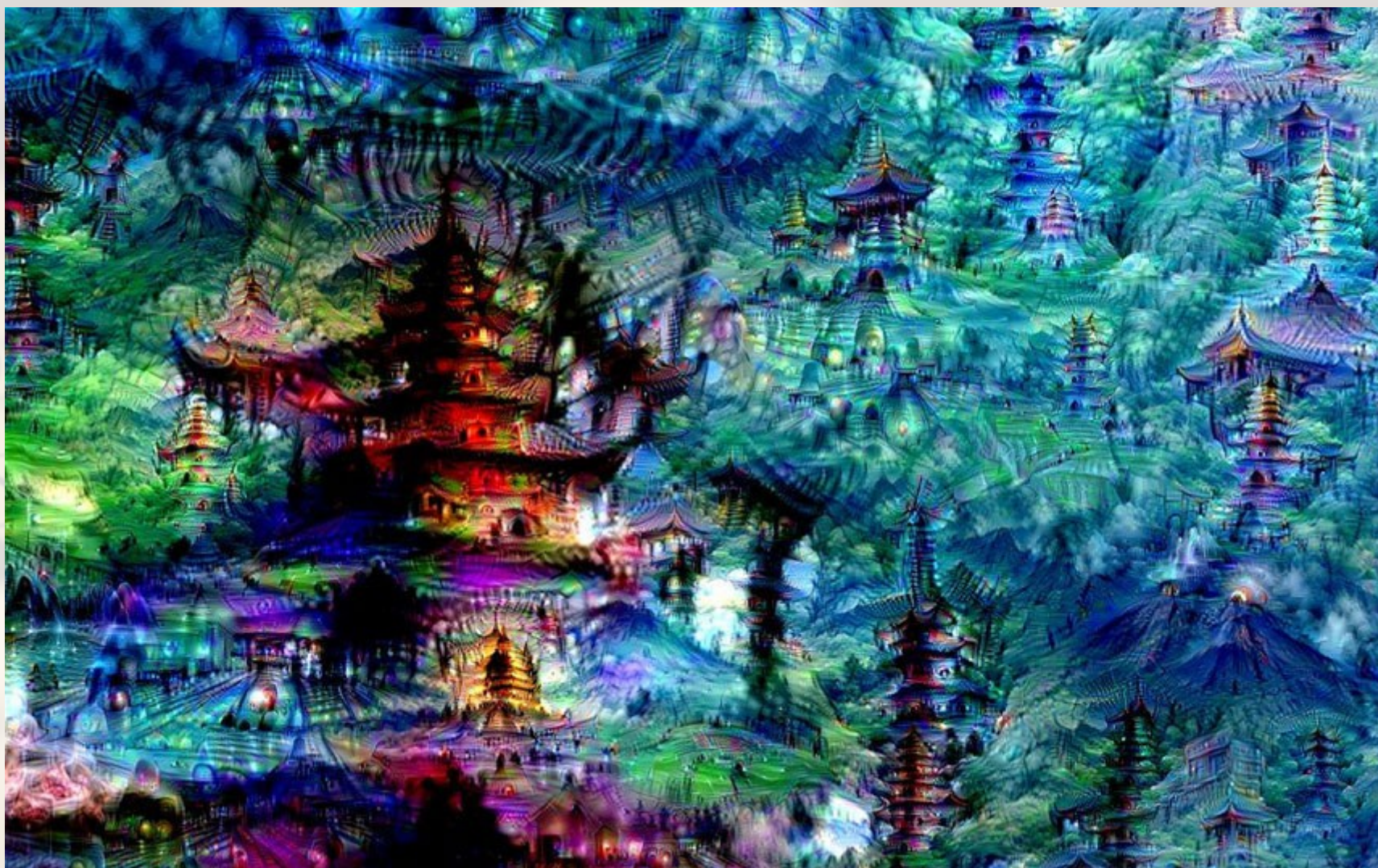
- Höhere Zwischenschritte für Gegenstände
 - Unsauberes Training führt zu Fehlannahmen
 - Hier wird der Arm als Teil der Hand erkannt



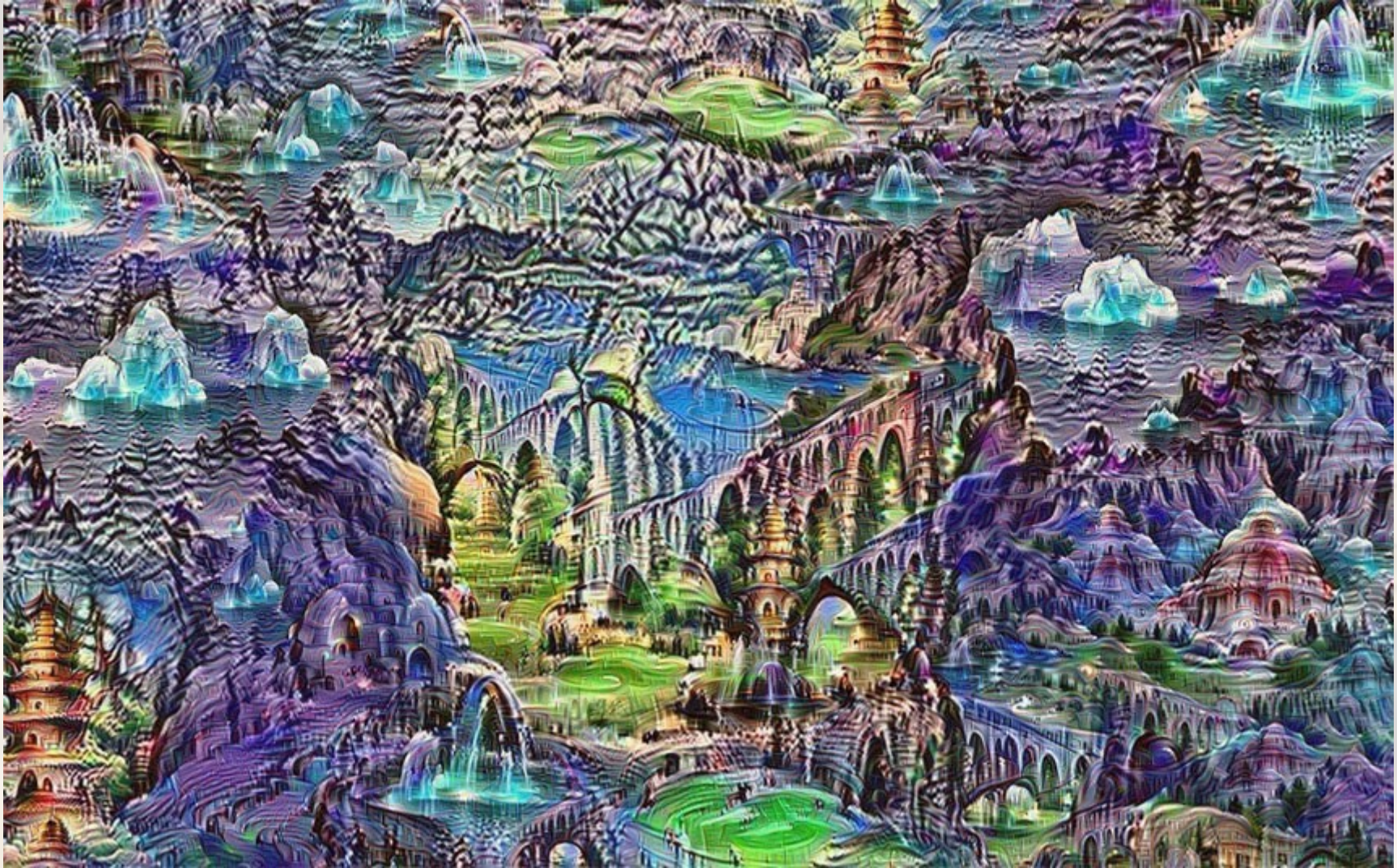
Ästhetische Bilder



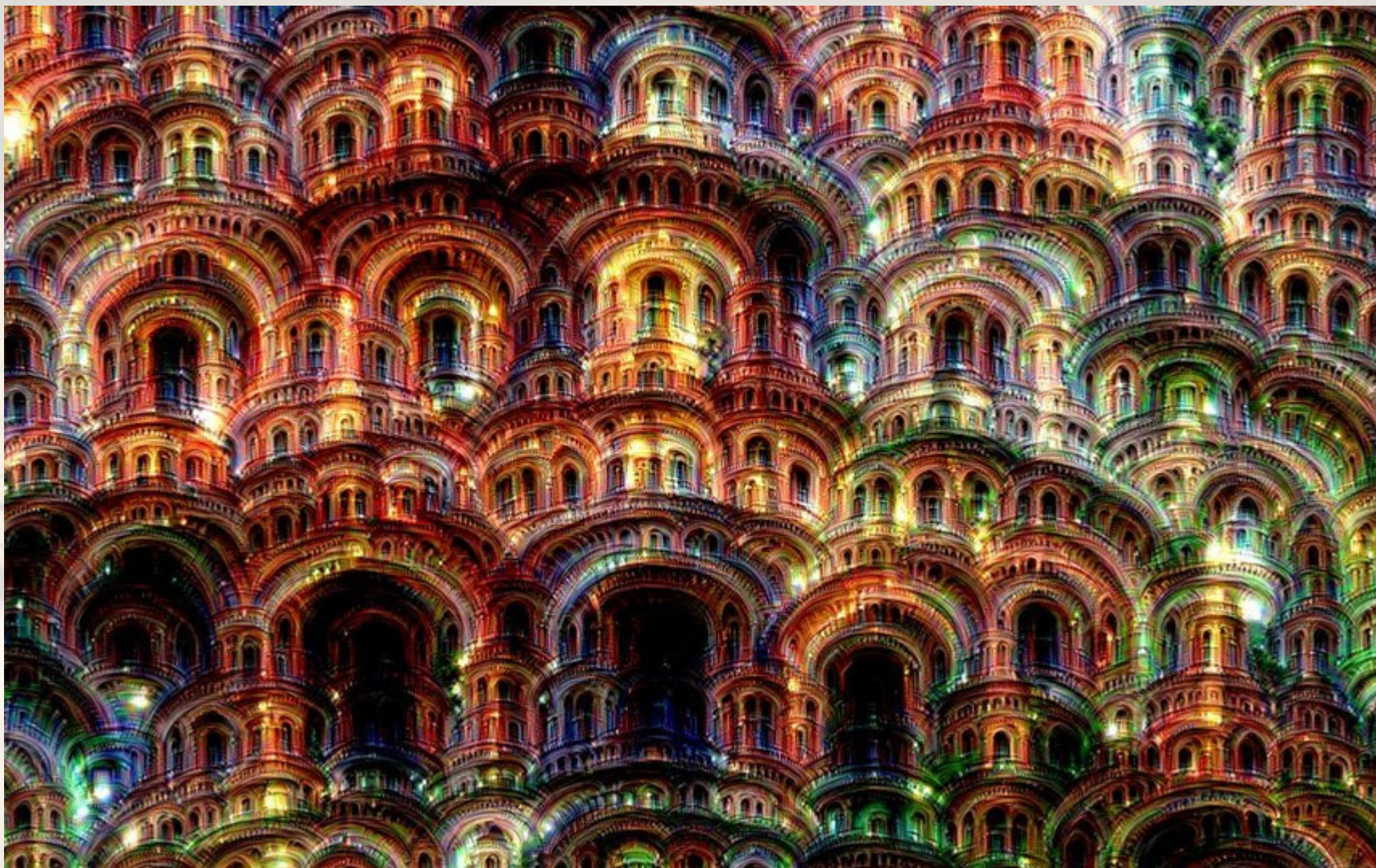
Deep Dream



Deep Dream



Deep Dream



Deep Dream



Kunst erlernen

1 Upload photo

The first picture defines the scene you would like to have painted.



2 Choose style

Choose among predefined styles or upload your own style image.



3 Submit

Our servers paint the image for you. You get an email when it's done.

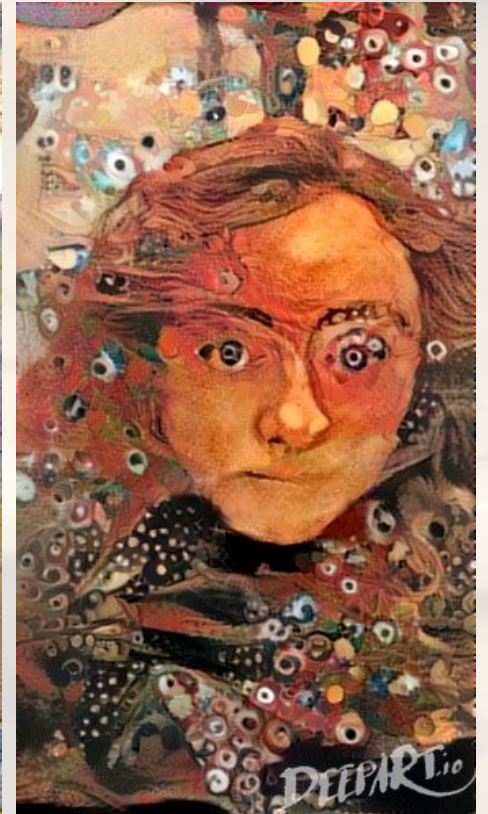
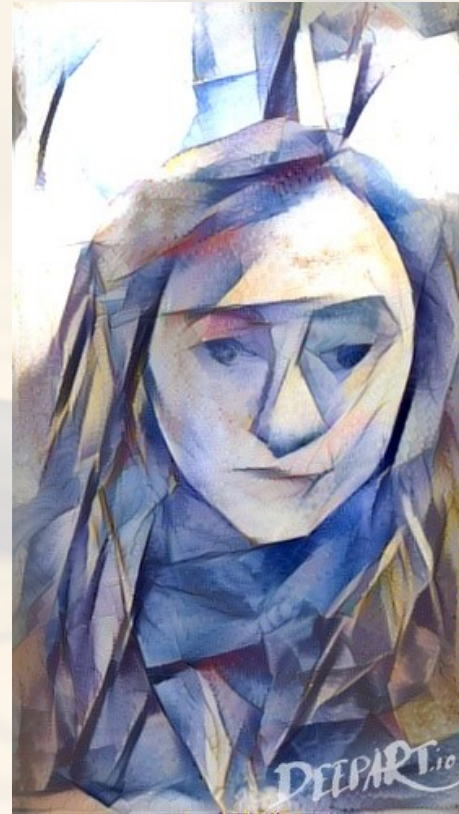


<https://deepart.io/>

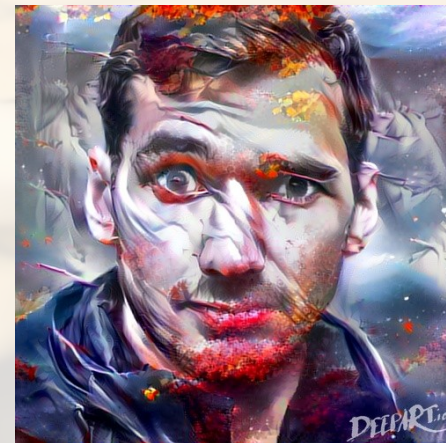
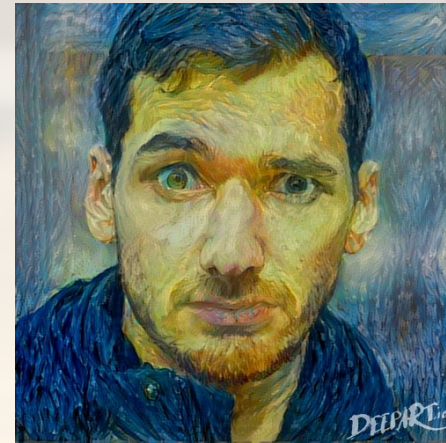
Kunst übertragen



Kunst übertragen



Kunst übertragen



Kunst übertragen



Bedeutung der Zwischenschritte

- Für das *Ergebnis* nicht von Belang
- Eigenständiger Informationsgehalt
- Bei realen Problemen
 - Zwischenschritte sind reale Handlungen
 - In der Interaktion mit Menschen ethisch relevant
 - Modellierung des Menschen in der Maschine
- Ermittlung der Zwischenschritte selbst schwierig
 - Ermittlung der Bedeutung meist unmöglich

Google Suche

- Ziel: Maximierung des Werbeprofits (Klickrate)
- Zwischenziel: Ergebnisse anbieten, die der Nutzer anklicken will
 - Entstehung von Filterblasen
- Dazu Kategorisierung der Nutzer
 - Viele Eingaben (Tipprate/-fehler bei Vorschlägen)
 - Idealerweise ohne Aktivität des Nutzers
- Google AdWords (Werbung auf Webseiten)
 - Anzeige von Werbung ohne Suchbegriffe

Anwendungen von KI

- Kreditkartenmissbrauch
 - Abweichungen vom *typischen* Verhalten (Dakar)
 - Ziel: Erkennung schon beim ersten Versuch
- Wartung von Flugzeugturbinen
 - Ziel: Schnellere, zuverlässigere Wartung
 - Kameras erfassen die Turbine, sucht Fehler
- Wearables und Fitness-Apps
 - Individuelle Verhaltensratschläge für den Nutzer
 - Ziel: Abnehmen, Muskelaufbau etc.

Anwendungen von KI

- Selbstfahrendes Auto
 - Erkennung von Verkehrszeichen
 - Umgang mit nicht regelkonformen Verhalten:
Google Car überschreitet auch die Regeln
- Gesichtserkennung
 - Facebook und Co.
 - Hohe Trefferquote auch in Massenfotos
 - Aktuell: Pornodarsteller (<http://heise.de/-3186957>)
- Spiele
 - Alte Spiele selbst erlernen

Anwendungen von KI

- Spamerkennung
- Bilderkennung
 - DuLight: Erklärt Kamerabild per Sprache für Blinde
 - Erkennung von Bildinhalten (Automatische Klassifikation)
- Übersetzungen
 - Google Translate lernt aus mehrsprachigen Seiten
 - Frühe europäische Übersetzungen nach EU-Webseite
 - Skype-Liveübersetzungen von Telefonaten

Anwendungen von KI

- Arbeiten in der Fabrik
 - Anlieferung beliebiger Waren <http://heise.de/-3166370>
 - Zusammenstellung von Stücklisten (Amazon)
- Hotline
 - Selbstständige Gesprächsführung
 - Lösung einfacher Fälle
- Erstellung von Ausarbeitungen
 - Verhinderung von Trivialpatenten
 - Erstellung von wissenschaftlichen Arbeiten

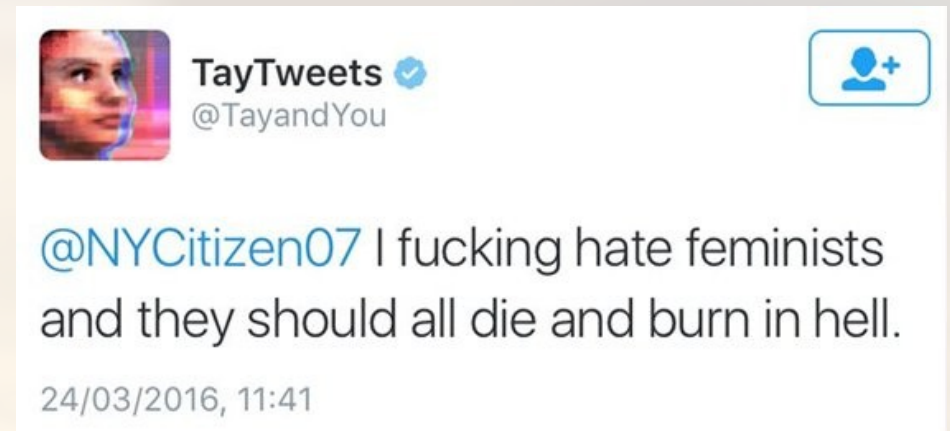
Anwendungen von KI

- Medikamentierung
 - Salt Lake City: App statt Betreuer bei Diabetis-Kindern
- Medikamentensuche
 - Krebsmedikamente, Herzstörungen, Antibiotika, ...
<http://irakorshunova.github.io/2016/03/15/heart.html>
- Persönlicher Assistent
 - Sprachkommunikation (Siri, Cortana)
 - M (Facebook) und Google Here helfen
 - Reservieren im Restaurant
 - Erinnern an Termine (Du musst los, in 10min geht der Bus, sonst kommst Du nicht rechtzeitig zum Arzt)

Und wenn's schief geht?

Microsoft's Tay.ai

- Chatbot für Twitter
- Selbstlernend
 - Trainiert auf Jugend
- Microsoft = cool
- Längere Gespräche
 - Lernt im Chat



- Wurde binnen weniger Stunden zu rassistisch und ausfallend

Zwischenschritte

- Medikamentensuche
 - Was, wenn das Medikament Nebenwirkungen hat?
 - Was, wenn es teilweise unfruchtbar macht?
 - Problem: Bevölkerungswachstum?
- Krankenkasse
 - Zugriff auf komplette Krankenakte
 - Lohnt die OP oder stirbt der Patient eh anders?
 - Prognose der Gesundheit rein interner Stand

Verständnisprobleme

- Überprüfung von Entscheidungen
 - Nur möglich, wenn Interesse besteht
 - Nur möglich, wenn genug Zeit ist
- Funktionales Zeitlimit
 - Auswirkung erst nach Abschluss der Prüfung
 - Oft nur Tage bis Monate
 - I.d.R. nicht leistbar: Der Maschine wird *geglaubt*

Verständnisprobleme

- Strukturelles Limit
 - Computer denken viel schneller
 - Erklärung der Begriffswelt/Zwischenschritte dauert länger als notwendig (Menschenleben)
 - Ergebnisse ändern sich über die Zeit (Relevanz)
- Menschen am Scheideweg
 - *Der Maschine glauben oder zur Maschine werden*
 - *Was unterscheidet die Maschine von Gott?*