

Künstliche Intelligenz

Ich werde schreiben meiner **Ingenieurarbeit** im **Bereich** von Künstliche Intelligenz und ehrlich gesagt sagen wir gar nicht **künstliche Intelligenz**. Das ist ein bisschen *populärwissenschaftlicher Begriff*. Wir sind da genauer wenn das **Maschinenlernen** und **Deep Learning** heißen. Und genau darum geht es in diesem **Vortrag**. Ich erkläre euch was das aktuelle state-of-the-art Tool ist um Maschinen zu trainieren.

Hier mal ein paar **Beispiele** für das worum es geht. Bestimmt habt ihr schon mal die **Google Bildersuche** probiert. Ihr geht auf Google, ergibt ein Wort ein und ihr bekommt Bilder geliefert. Aber woher weiß Google dann wie eine Katze aussieht? Wenn ich Katze eingabe? Oder das gleiche funktioniert auch *umgekehrt*. Man kann bei Google auch Bilder hochladen und bekommt dann ein **Stichwort** was auf diesem Bild zu sehen ist, beziehungsweise bekommt noch andere Bilder von dieser **Kategorie**. Und ein aktuelles Beispiel dass ihr vielleicht gesehen habe die App **Prisma**. Da kann ich ein Foto von mir hoch laden und ein **Gemälde** auswählen und dann rendert diese App mein Foto im **Stil** des Gemäldes. Alle diese Tools funktionieren mit sogenannten **künstlichen neuronalen Netzen**.

Dazu klären wir erstmal was ist eigentlich ein **Neuron**. Was haben vielleicht *die ein oder anderen von euch* schon mal aus dem **Biologieunterrichts** mitbekommen das ist ein ganze großes Thema von dem Neuron. Das ist eine **Zelle** und hier zählt die jetzt ganz viele **Eingänge**. Das könnt ihr euch wieso kleine **Kanäle** vorstellen. Da können **Signale** ankommen. Und wenn diese Eingangssignale irgendeinen bestimmten **Schwellenwert** überschreiten dann sagt man dann feuert dieses Neuron. Das heißt, es gibt auf seinem großen **Axon** ein Signal aus. Also könnt ihr euch vorstellen viele Eingänge. Da kommt Signal rein und wenn die **Stimmen** laut genug in das Neuron rein brüllen sozusagen, dann gibt es was aus.

Jetzt kann ich diese Neuronen zusammen bauen. Da kann ich also ganz viele davon nebeneinander setzen. Und bekomme *dementsprechend* ganz viele Eingänge. So ähnlich sieht es *möglicherweise* in unserem **Gehirn** aus. Das ist natürlich sehr schwer im Computer nachzubauen. Deshalb macht man sich das einfacher. Man bildet einfach **Schichten** aus Neuronen. Also das sind 4 Neuronen. Sie haben *jeweils* 4 Eingänge und *jeweils* 1 Ausgang.

Und jetzt könnt ihr euch vorstellen dass diese Eingänge hier verbunden sind mit den **Pixeln** von einem Bild. Jeder dieser Pixel entspricht also einen **Bildpunkt**. Wenn ihr bei eurem Display ganz genau hinschaut auf dem Handy oder auf dem Computer, dann seht ihr das sind viele kleine Punkte. Und *letztendlich* sind es einfach nur **Zahlen**. Diese Zahlen setzt der **Bildschirm** dann in **Farbwert** um, und wir erkennen sofort: “aha, das ist wie eine Katze”, oder so.

Für den Computer ist es aber, wie gesagt, nur Zahlen. Damit kann der Computer nicht viel anfangen. Außer, man macht es *eben* das hier. Diese Pixeln hier, wie gesagt, es sind kleine Zahlen und diese Zahlen sind jetzt das, was in die Neuronen rein gefüttert wird. Und wenn diese **Zahlenwerte** dann den **Schwellwert** dieses Neurons überschreiten dann gibt das Neuron ein **Ausgabewert** aus.

Jetzt gibt es hier verschiedene **Parameter**. Vielleicht ist es so, dass dieses eine Pixel hier viel wichtiger als dieses Pixel da ist. Das heißt, für jeden **Eingangskanal** dieser Neuronen gibt es ein sogenanntes **Gewicht**, also wie wichtig ist dieser Kanal. Mit diesem Gewicht wird es einfach multipliziert und das ist einfach nur eine Zahl. Und dann kann man noch was einstellen, *und zwar* diesen Schwellwert hier vielleicht braucht dieses speziellen Neuron ganz-ganz viel **Input** dass es was aus gibt. Ein anderes Neuron, das ist vielleicht viel *sensibler* und da reicht es schon wenn ein ganz-ganz *geringer* Schwellenwert überschritten wird. Das sind also für die Neuronen die zwei *wesentlichen* **Parameter**: die Gewichte (1) und der Schwellwert der *eingestellt* werden muss(2).

Jetzt habe ich also hier eine tolle **Schicht** aus Neuronen gebaut. Und ihr könnt schon sehen dass aus meinen 16 **Eingabepixeln** (wahrscheinlich mehr in der Realität - ihr versteht was ich meine). Dass aus diesen 16, also vielen, Eingabepixeln nur 4 **Ausgabewerte** gemacht werden. Also das war jetzt der *allerwichtigste fundamentale Bausteine* für so ein künstliches neuronales Netz. Eine Schicht aus Neuronen die viele Eingänge haben und deren Eingänge mit Gewichten multipliziert werden. Wenn diese **Eingangssignale** irgendeinen bestimmten Schwellenwert überschreiten dann gibt es den Ausgabewert. Und das kann ich jetzt zusammen packen und noch weitere von diesen Schichten aufeinander stapeln.

Merkt euch einfach künstliche neuronale Netze bestehen aus Neuronen. Diese Neuronen sind *letztendlich* **nicht-lineare Funktionen** und die baue ich zusammen zu einem großen Stapel von Schichten. Eine andere **Art** das zu interpretieren ist, kann sich vorstellen das sind **Filtern** durch die das **Eingabebild** durchläuft. Je weiter man hier hochkommt, *desto komplexer* ist das, worauf das Neuron im **Eingabebildbereich** reagiert. Also *beispielsweise*, hier sind es **Ecken** und **Kanten**, hier sind es schon ein bisschen komplexere **Muster**, irgendwo kommen dann **Gesichter**. Und ganz-ganz *oben* bei der höchsten **Stufe der Abstraktion**, da ist dann alles Unwichtige *raus gefiltert*. Und nur noch das, worauf das Netz trainiert wurde, das bleibt dann *übrig*. Und das sind dann hier vielleicht diese 1000 Klassen mit Katzen, Hunden, Autos und Tischen und Stühlen.

Diese künstlichen neuronalen Netze haben die tolle **Eigenschaft**, dass man sie trainieren kann. Was bedeutet das? Also wenn ihr euch dieses Ding hier anguckt, dann kann man das mathematisch auch als **Funktionen** interpretiert. Also $f(x) = \text{irgendwas}$. Und hier ist es dann die Funktion die nimmt als Eingabewert so ein Bild, und als Ausgabe kommt so ein **Vektor** aus. Also klingt es mal so ein bisschen sehr kompliziert. Vor allem, ist jetzt noch nicht ganz genau klar was will ich jetzt eigentlich trainieren. Aber vielleicht erinnert ihr euch die Dinge die man hier einstellen kann, das sind die Gewichte der *einzelnen* Neuronen und ihr Schwellenwert. Bzw die Eigenschaften der nicht-linearen Funktionen

jedes Neurons. Das sind die Parameter die man einstellen kann. Was mache ich mit den Parameter?

Ganz einfaches Beispiel, das haben vielleicht einige von euch schon mal gesehen - **lineare Regression**. *Um Gottes willen*, was ist das denn? Also stellt euch vor ihr habt eine **Ladung Punkte** so. Und jetzt wollt ihr eine **Gerade** darein fitten, sagt man. Also ihr wollt eine Gerade finden, die diese Punkte so gut wie möglich approximiert, also die eine **Näherung** für diese Punkte darstellt. Und jetzt könnt ihr verschiedene Geraden malen. Also vielleicht diese Gerade hier, noch diese Gerade. Aber ich glaube das was ich als erstes gemalt hab, werdet ihr wahrscheinlich zustimmen, die repräsentiert diese **Punktwolke** am besten. Und eine Gerade hat zwei Parameter: die hat eine **Y-Achsen Abschnitt** und die hat eine **Steigung**. Und diese zwei Parameter, die kann ich jetzt so einstellen, dass die Punkte *optimal* durch die Gerade repräsentiert werden.

Beim unseren künstlichen neuronalen Netz hat man *möglicherweise* Millionen von Parametern. Das ist leider nicht ganz so einfach zu optimieren. Wenn man das trainiert, dann braucht man *wahnsinnig* viele **Trainingsbeispiele**. Also ganz viele Bilder, von denen man schon weiß was drauf ist. Dann muss irgendein Mensch diese Bilder mal “*gelabelt*” haben, so nennt man das. Also sich die **Klasse** überlegen, für die das Bild steht. Und dann habe ich am Anfang irgendeine Netz, das ist einfach untrainiert. Habe ich diese ganzen Parameter einfach mal zufällig initialisiert. Und jetzt nehme ich ein **Bild** und schiebe das *hierdurch*. Und dann wird es wahrscheinlich oben irgendeinen **Unsinn** ausgeben. Dann wird das sagen: “Ja, das ist *womöglich* 50%-Katze, 50%-Hund...”, - und *dabei* ist eigentlich Stuhl.

Was ich dann machen kann ist ich kann sagen: “*Moment mal*, ich weiß aber welche Klasse die richtige sein müsste”. Und jetzt kann ich gucken welche Neuronen haben am meisten dazu beigetragen, dass die **falsche Klasse** aktiviert wurde und welche Neuronen müssten mehr aktiv sein dass die **richtige Klasse** aktiviert wird. Und dann stelle ich

deren Parameter so ein, bis die **Vorhersage** in richtiger wird für dieses Bild. Und dann mache ich es mit dem nächsten Bild, am besten gleich mit 1000 Bildern *gleichzeitig*. Und wenn ihr dann eine Million von Parametern habt, dann könnt ihr euch vorstellen, dass man noch viel mehr **Beispielbilder** braucht um diese Gewichte und die ganze Schwellwerte einzustellen. Aber wenn ich das *eben richtig* mache und sehr viele **Trainingsdaten** habe, ich dann alle *immer wieder* durch das Netzwerk durchschiebe. Irgendwann habe ich dieses Netzwerk trainiert. Das heißt, ich hab die Parameter so eingestellt dass die **Vorhersagen** des Netzwerkes möglichst oft richtig sind.

Und so ein **trainiertes Netzwerk** das ist cool. Das kann ich dann anwenden auf irgendwelche Bilder, von denen ich keine Ahnung habe welches Label die haben (welche Klasse die repräsentieren). Und da kann ich dann, zum Beispiel, mit einer 1,000,000 Bildern (ok, klingt erstmal viel, aber im **Netz** gibt es *deutlich* mehr), so ein Netzwerk trainieren. Und vielleicht habe ich aber 2,000,000,000 Bildern und ich will nicht *jedes einzelne Bild* durchgehen und das *manuell* klassifizieren. Wenn ich aber dieses Netz trainiert hätte, dann kann ich die Milliarde Bilder nehmen, schön durch das Netzwerk durch schieben, dann habe ich die alle klassifiziert. **Vorausgesetzt**, sie beinhalten die gleichen Klassen wie mein **Trainingsdatenset** - wenn die andere Klassen dabei sind, kommt *möglicherweise* **Unsinn** raus.

Um zusammenzufassen, in diesem Vortrag habt ihr gelernt, dass neuronale Netze sind **Funktionsapproximatoren**. Wir versuchen Funktionen zu nähern, auf die sie trainiert wurden. Vielleicht können sie noch ein-zwei andere Dinge. Aber neuronale Netze können meistens nur eine Sache *richtig gut*. Zum Beispiel, Bilder richtig gut erkennen - besser als Menschen *möglicherweise*.

Mich würde aber natürlich auch eure Meinung interessieren. Vielleicht gibt es *einige* die sagen: “Das ist mir alles neu und ich finde das super *spannend*(1), total *abschreckend*(2) oder irgendwas anderes(3)”. Was denkt ihr zu *diesem Thema*?