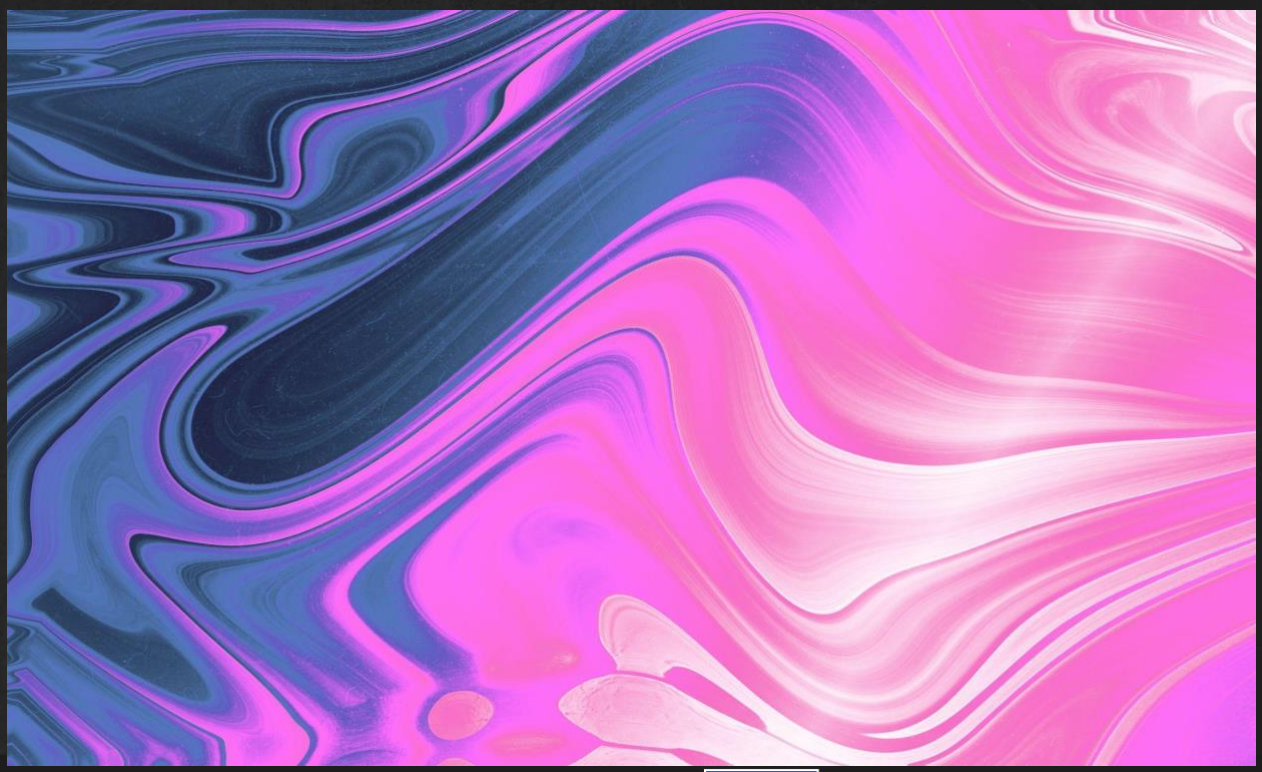


Funded by
the European Union

JONGE ONDERNEMERS KRACHTIG MAKEN MET AI

nr. 2024-1-BG01-KA220-YOU-000249027

Funded by
the European Union

INHOUDSOPGAVE

Funded by the European Union

INVOERING

In deze module worden de basisprincipes van AI besproken en wordt informatie gegeven over de pijlers die nodig zijn om een AI-model te creëren. Ook worden de knelpunten bij het creëren van krachtige AI-modellen besproken, waaronder ontwerpbeslissingen op architectuurniveau en nuances waarmee rekening moet worden gehouden bij het trainen van modellen.

Na het lezen van deze module heb je een dieper begrip van de thema's die verband houden met AI. Deze module richt zich voornamelijk op de technologische kant van AI, maar geeft ook suggesties voor hoe je AI het meest effectief kunt inzetten.

- U leert hoe belangrijk trainingsgegevens zijn en waar u op moet letten.
- U leert over de verschillende architecturen die momenteel beschikbaar zijn en hun sterke en zwakke punten.
- U leert over de meest productieve manieren om de momenteel beschikbare AI-modellen te gebruiken.

Funded by the European Union

1: AI EN MACHINAAL LEREN

Naarmate de industrie zich ontwikkelde, veranderde ook de perceptie van de betekenis van AI. Tegenwoordig verwijzen mensen met AI vooral naar grote machine learning-modellen, en dat is precies wat in deze module aan bod komt. Het concept machine learning verwijst naar een algoritme dat patronen in invoergegevens kan vinden en zichzelf verbetert naarmate het uitgebreider wordt getraind. De sleutel is dat het niet wordt geleerd hoe het moet denken. Sommige modellen werken door het biologische brein na te bootsen.

In sommige gevallen is het slechts een modewoord om de aandacht te trekken en producten te verkopen. Het is daarom cruciaal om de werking van deze mysterieuze en vaak verkeerd begrepen technologie te begrijpen. Door de belangrijkste concepten te begrijpen, krijgt u een beter beeld van wat er mogelijk is en laat u zich niet misleiden door hype. U bent ook beter bereid om de kracht van AI te benutten.

In deze module behandelen we de belangrijkste onderwerpen binnen machine learning. We benadrukken het belang van hoogwaardige trainingsdata. Vervolgens bespreken we zorgvuldig geselecteerde voorbeelden van AI-architecturen, aangezien er zoveel zijn, inclusief informatie over hoe ze werken en wat je ervan kunt verwachten, inclusief sterke en zwakke punten. Vervolgens bekijken we hoe verschillende AI-modellen getraind kunnen worden en welke soorten training er zijn, waarbij we architectuur en trainingsdata combineren. Tot slot bekijken we de huidige stand van zaken op het gebied van AI en bespreken we hoe je de beschikbare producten optimaal kunt benutten.

Funded by
the European Union

2: HET BELANG VAN TRAININGSGEGEVENS

Zoals in de meeste systemen is AI zo sterk als de zwakste schakel. Als je goede data hebt, maar geen manier om die te verwerken, heb je niets. En zelfs als je de meest geavanceerde architectuur hebt, maar je voert data van lage kwaliteit in, zullen de resultaten teleurstellend zijn.

In dit hoofdstuk bespreken we het belang van het leveren van gegevens van hoge kwaliteit bij het trainen van een AI-model. Daarnaast evalueren we aan welke criteria gegevens moeten voldoen om als gegevens van hoge kwaliteit te worden beschouwd.

Er zijn systemen die gebruikmaken van gesimuleerd leren, waarbij een AI met een gesimuleerde omgeving om gaat en leert door middel van trial-and-error door middel van bekrachtiging (dit wordt uitgelegd in hoofdstuk 4). Voor andere AI-modellen is trainingsdata de enige bron van begrip voor de AI. Alle patronen die het model observeert, zijn gebaseerd op de data die het krijgt. Dit komt doordat de AI, in tegenstelling tot mensen, geen lichaam of persoonlijke ervaring heeft om kennis uit te halen en afhankelijk is van de input om zijn waarheid te bepalen. Het heeft geen verdere manier om zijn conclusies te testen, tenzij het later te horen krijgt dat de informatie onjuist is. Daarom zijn data van hoge kwaliteit noodzakelijk voor het functioneren van AI-modellen. Maar wat zijn nu eigenlijk data van hoge kwaliteit?

De eerste overweging is nauwkeurigheid. Hoewel er zowel gelabelde als ongelabelde datasets zijn, is nauwkeurigheid voor beide cruciaal. Bij gebruik van gelabelde data in combinatie met supervised learning (meer hierover in hoofdstuk 4), hangt de nauwkeurigheid van de data af van

Funded by the European Union

hoe gedetailleerd en correct de labels zijn ten opzichte van de data. Je kunt bijvoorbeeld een AI trainen om objecten in een afbeelding te identificeren.

Voorbeeld invoer afbeelding https://www.rspb.org.uk/birds-and-wildlife/robin

Neem bijvoorbeeld deze afbeelding; deze kan worden gebruikt als input om een AI te trainen. De afbeelding blijft hetzelfde, maar de tags kunnen variëren afhankelijk van de gewenste complexiteit van de AI. Eén label kan 'vogel' zijn; dit beperkt het begrip van de AI. Een andere, meer geavanceerde dataset kan deze afbeelding labelen met tags als 'vogel, roodborstje, zittend, staand op tak, oranje borst, wazige achtergrond', enzovoort. U kunt zich het verschil in de resulterende AI-modellen voorstellen. Als u een onbetrouwbare dataset gebruikt, kan deze afbeelding zelfs het label 'kat' hebben. Dit is natuurlijk overdreven, maar als u de AI leert dat dit een kat is, kan hij, zelfs als hij begrijpt wat een vogel is, ernaar verwijzen als een kat. Daarom is het gebruik van

Funded by the European Union

gedetailleerde en nauwkeurige gelabelde data cruciaal voor een goede functionaliteit bij het trainen van een AI.

Als we een voorbeeld willen geven van nauwkeurigheid in de context van ongelabelde data die gebruikt worden in ongeleide training, kunnen we ons een groot taalmodel (LLM) voorstellen. LLM's zoals ChatGPT kunnen getraind worden met ongelabelde data, zoals artikelen, tweets, Reddit-berichten, wetenschappelijke tijdschriften, enz. Dit zijn ongelabelde data en de AI gebruikt deze om patronen in de zinsstructuur te identificeren en relaties tussen woorden te leggen. Door de AI te trainen met voldoende van deze data, begint de AI zinnen te produceren die een mens zou kunnen schrijven. Als je de AI echter traint met huiswerk van zesjarige kinderen, kun je ongewenste resultaten van de AI krijgen. Dit is wat kwaliteit betekent in de context van ongelabelde data.

Kortom, als je hem dingen leert die niet waar zijn, kan de AI dit niet corrigeren. Zorg ervoor dat de data die gebruikt wordt om een AI te trainen van hoge kwaliteit is, anders zou hij katten als honden kunnen identificeren. Als je hem traint met tekst, zorg er dan voor dat deze coherent en grammaticaal correct is geschreven, anders kan de uitkomst onzinnig zijn. Zelfs een kleine subset van de data van lage kwaliteit kan het model verontreinigen, omdat AI werkt met patronen en daardoor onjuiste verbanden legt. De kwaliteit van de invoerdata is direct gecorreleerd met de kwaliteit van de uitkomsten.

Een andere overweging voor hoogwaardige data is de aanwezigheid van vooroordelen. Zelfs als er technisch niets mis is met de data, kun je nog steeds enorme problemen krijgen als je vooroordelen over het hoofd ziet. In een PR-schandaal lijkt een AI voor beeldherkenning voornamelijk te zijn getraind op blanke gezichten, waardoor mensen van andere rassen als primaten werden herkend. Dit kan worden opgelost door rekening te houden met de twee volgende criteria voor hoogwaardige data: volume en diversiteit.

Funded by the European Union

Hoe groter het brein van de AI – gemeten in het aantal parameters – hoe meer data er nodig is om waarden nauwkeurig en effectief in de nodes te coderen, aangezien elke input de nodewaarden slechts lichtjes zal verschillen. Beschouw nodes als wiskundige functies – ze ontvangen input en geven output na wat rekenwerk. Daarom is het belangrijk om voldoende volume te hebben bij het trainen van een AI. Er zijn online verschillende suggesties te vinden over hoeveel data voldoende is, maar de enige echte manier om dit te bepalen is door de AI te testen. Dit test natuurlijk het hele systeem en niet alleen de hoeveelheid inputdata, waardoor het moeilijk is om het zwakke punt te identificeren. Door het model te testen, krijg je echter een idee van de resultaten die je kunt verwachten.

De laatste belangrijke overweging bij het trainen van data is de diversiteit. Als je de AI te veel van één type input geeft, zal deze overfitten en algemene patronen niet herkennen, maar alleen geoptimaliseerd worden voor die ene vraag. Dit hangt natuurlijk af van het doel van de AI, maar voor grote taalmodellen is het cruciaal dat ze een breed begrip hebben en zelfstandig proberen te redeneren. Laten we eens een paar voorbeelden bekijken: als je de AI alleen vragen laat over toevoegen laat maken, zal het niet weten hoe het moet vermenigvuldigen. Als je de AI te veel artikelen laat zien, klinkt hij in alle contexten als een verslaggever, dus je kunt beter Reddit-threads of tweets opnemen. In een AI voor beeldherkenning kun je hem te veel vergelijkbare foto's geven - als alle foto's met een vogel een blauwe lucht als achtergrond hebben, kan de AI ervan uitgaan dat het blauw bijdraagt aan het feit dat de afbeelding een vogel is.

Funded by
the European Union

3: AI-ARCHITECTUREN

Hoewel het trainen van data belangrijk is voor de uiteindelijke functionaliteit van een AI-model, is het niet de enige overweging. Om die data uiteindelijk bruikbaar te maken, is een geschikte architectuur cruciaal. Beschouw de architectuur als een skelet, een verzameling botten die met spieren met elkaar verbonden moeten worden. Er zijn vele soorten architecturen die voor verschillende doeleinden kunnen worden gebruikt. Sommige AI-modellen zijn beperkt, bekend als kunstmatige smalle intelligentie (ANI) en zijn geoptimaliseerd voor één heel specifiek doel, zoals schaak-AI's, terwijl andere kunstmatige algemene intelligentie (AGI) worden genoemd en proberen het menselijk begrip op alle gebieden na te bootsen. AGI is nog niet bereikt en er zijn veel discussies over definities en bewustzijn. Grote taalmodellen (LLM's) zoals ChatGPT zijn een poging om AGI te creëren en zijn aanzienlijk geavanceerder dan de technologie een paar jaar geleden was. Dat gezegd hebbende, zijn er voor beide categorieën veel verschillende architecturen, elk met hun eigen superkrachten, beperkingen en specifieke use cases waarvoor deze afwegingen zinvol zijn. Om een basisbegrip te geven van wat mogelijk is, bespreken we er slechts twee: genetische algoritmen en neurale netwerken.

3.1: Genetische algoritmen

Genetische algoritmen (GA's) zijn een type AI-architectuur, net iets anders dan wat u waarschijnlijk gewend bent in de mainstream. Om ons begrip van verschillende soorten AI te verbreden, zullen we kijken waar ze voor nuttig zijn en hoe ze werken. Dit is een zeer gespecialiseerd type architectuur dat niet breed bruikbaar is voor veel verschillende soorten taken, maar zeer efficiënt is voor specifieke problemen. GA's worden

Funded by the European Union

meestal gebruikt bij optimalisatieproblemen waarbij er meer variabelen zijn dan er met brute force-force kunnen worden uitgevoerd. Het is een elegante, lichtgewicht oplossing voor het testen van mogelijke oplossingen voor strikt gedefinieerde problemen. Dit klinkt misschien allemaal wat abstract, dus laten we het meest populaire voorbeeld eens bekijken: het handelsreizigersprobleem.

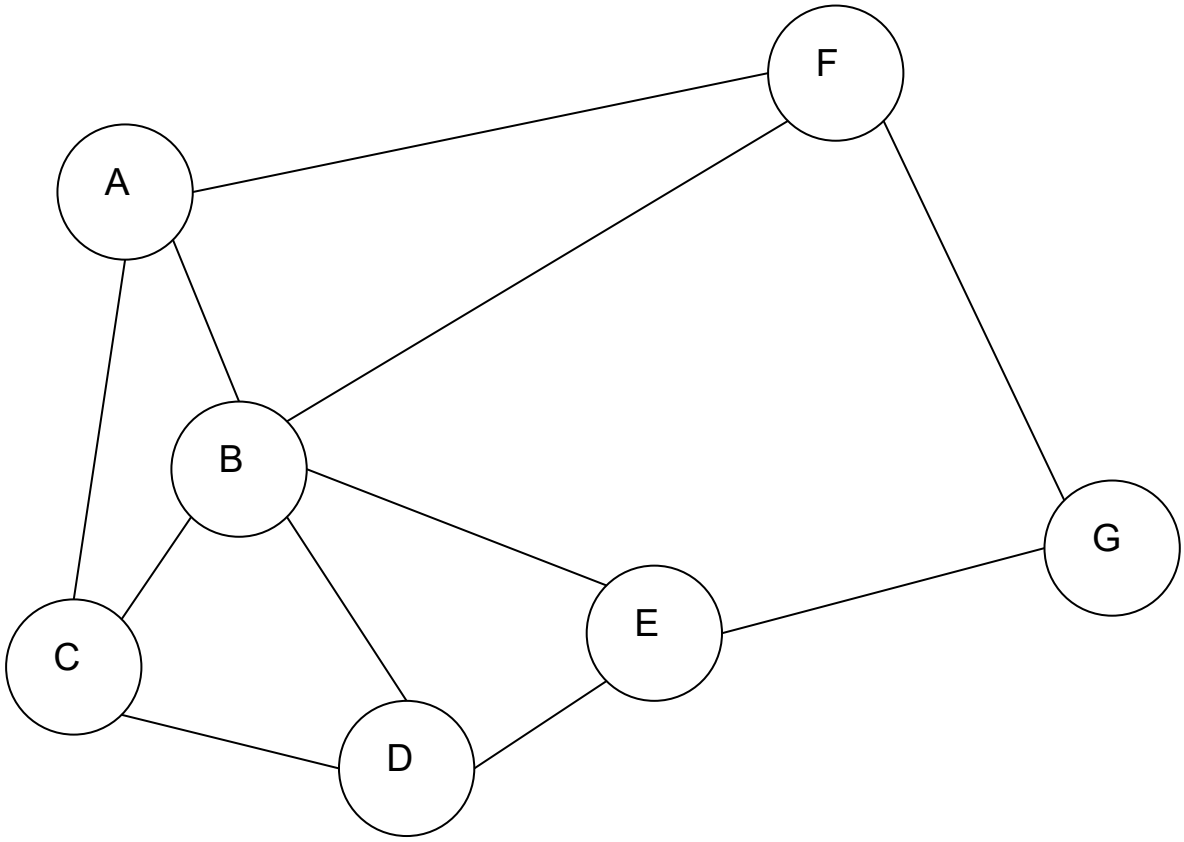


Diagram van een mogelijk scenario

In de bovenstaande tekening ziet u 7 knooppunten. Het handelsreizigersprobleem verwacht een oplossing waarbij elk knooppunt of elke stad precies één keer wordt bezocht, en het doel is om het kortste pad te vinden. Voor de eenvoud zijn er in dit voorbeeld geen afstanden tussen de knooppunten. Een mogelijke oplossing is ABCDEFG. Een andere oplossing is ABCDEGF. Hoewel er andere manieren zijn om dit op te lossen, presteren GA's erg goed in dergelijke problemen.

Funded by the European Union

Hoe werken ze? In dit voorbeeld is elke oplossing 7 letters lang en bevat elke letter precies één keer. Zie dit als het DNA waarop de genetica is gebaseerd; de locatie van elke letter is een kenmerk van de oplossing. Elke oplossing is een lid van de populatie.

Aanvankelijk wordt een willekeurige populatie gegenereerd met allerlei soorten oplossingen, waarvan sommige inefficiënt en andere sneller zijn. Een fitnessfunctie is een rangschikkingssysteem dat wordt gebruikt om de kwaliteit van oplossingen te vergelijken. In dit voorbeeld is de fitnessfunctie gewoon de afstand van het pad, waarbij een lagere afstand beter is. Dit betekent dat oplossingen met een lagere totale afstand beter gerangschikt zullen worden en gebruikt zullen worden voor de volgende generatie. Met 3 knooppunten, als de oplossing ABC is, is de afstand AB + AC, waarbij AB de afstand van A naar B is.

Zodra de populatie is gegenereerd, moet deze worden gesorteerd op basis van de fitnessfunctie. Vervolgens kan de volgende generatie worden geproduceerd uit de ouders van de huidige generatie. Hoewel oplossingen met een hogere rangorde zich waarschijnlijker reproduceren, is er enige willekeur om voortijdige convergentie en lokale maxima te voorkomen. Dit onderwerp is te genuanceerd om er binnen het bereik van deze module dieper op in te gaan, maar er zijn extra bronnen beschikbaar voor nieuwsgierige cursisten.

Voortplanting gebeurt door de twee ouders te combineren om zo de nakomelingen te variëren. Hoewel er veel verschillende manieren zijn om genetica te combineren, is het algemene idee dat de nieuwe populatie gemiddeld beter is, omdat de minst effectieve oplossingen uit elke generatie worden verwijderd. Door het algoritme meerdere generaties lang te laten draaien, kun je de beste oplossingen bijhouden

Funded by the European Union

totdat er één goed genoeg is voor het doel, omdat er nooit een garantie is dat het de best mogelijke oplossing is.

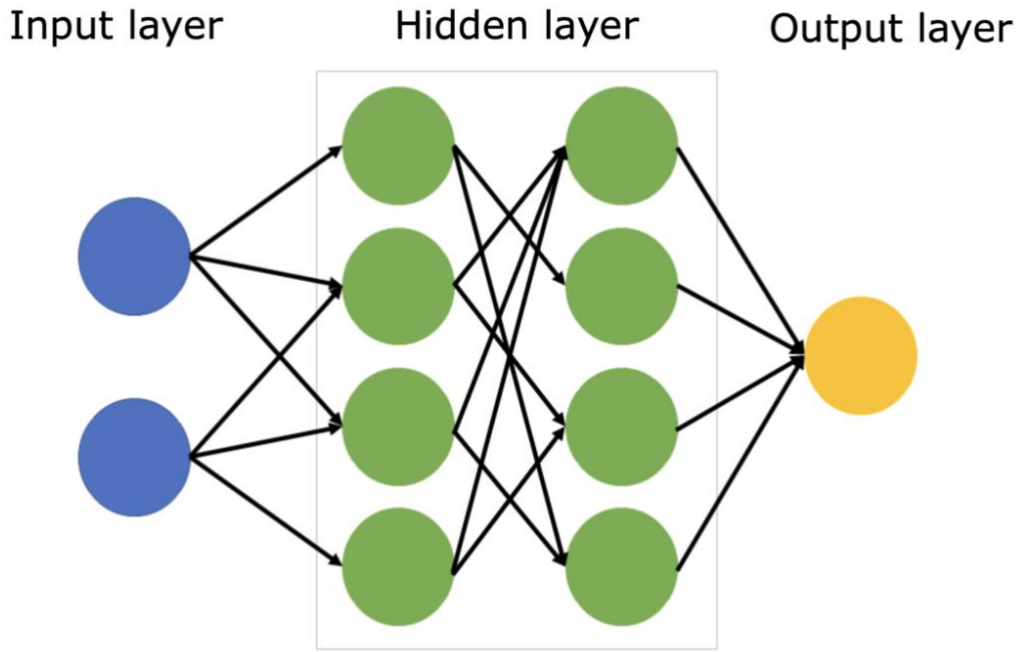
3.2: Neurale netwerken

Neurale netwerken (NN's) vormen de hoeksteen van moderne AI. Het concept is dat ze nabootsen hoe menselijke hersenen werken met neuronen (hersencellen) en de verbindingen daartussen. In de context van AI wordt elk neuron een knooppunt genoemd en is de computerrepresentatie ervan een functie voor dataverwerking. Een eenvoudig geval is wanneer het knooppunt een waarde tussen 1 en 0 heeft en de invoer met die waarde wordt vermenigvuldigd; dit is een functie. De verbindingen hebben vaak ook waarden die gewichten worden genoemd en die de invloed van het ene knooppunt op het volgende weergeven.

Er bestaan talloze verschillende structuren voor neurale netwerken, waaronder de transformerarchitectuur – hierop zijn de meeste LLM's zoals ChatGPT gebaseerd. In dit hoofdstuk bespreken we het algemene idee van een neuraal netwerk in plaats van de details die geavanceerdere architecturen vormen.

Funded by the European Union

Om inzicht te krijgen in de werking van NN's, gebruiken we een visueel voorbeeld van een AI voor beeldherkenning.

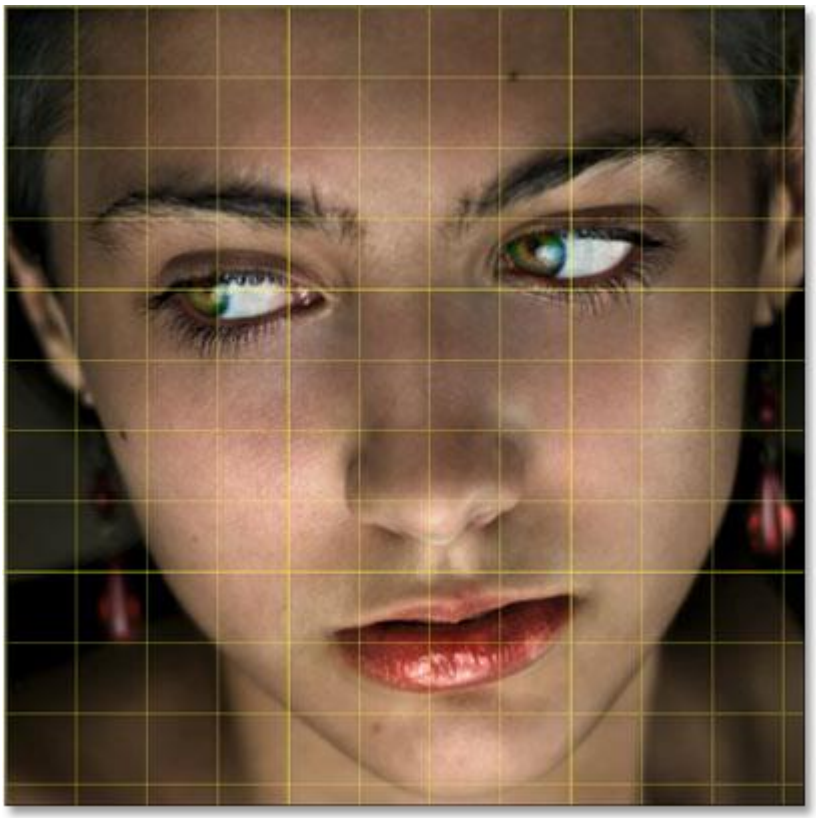


Artificial neural networks

Mogelijke weergave van een neuraal netwerk
https://www.sciencelearn.org.nz/images/5156-neural-network-diagram

In dit voorbeeld zijn de knooppunten de cirkels en de pijlen de verbindingen. De gewichten van de verbindingen en de knooppunten hebben beide invloed op de resultaten. We zullen in hoofdstuk 4 - AI trainen - begrijpen waar deze waarden vandaan komen.

Funded by the European Union



Afbeelding met raster https://www.photoshopessentials.com/photo-effects/photo-strips/

Stel je nu een raster als dit voor over een afbeelding. In plaats van deze vierkanten is elke pixel echter een blauwe invoercirkel in het NN. Deze invoer wordt vervolgens verwerkt door het netwerk, waarbij elke verbinding patronen tussen de aangrenzende pixels analyseert. Wanneer de invoer door voldoende verborgen lagen gaat, is het in staat om structuren te identificeren op basis van de invoerpixels. Het zou dan kunnen beschrijven wat deze structuren zijn.

De mogelijkheden zijn afhankelijk van de grootte van het neurale netwerk, aangezien het slechts zoveel inputs kan verwerken als er knooppunten zijn. Zie het als een puzzel: als elk knooppunt je kon vertellen waar een onderdeel hoort, zou je evenveel knooppunten nodig hebben als er onderdelen zijn. Dit is echter een te simplistische

Funded by the European Union

voorstelling van zaken en in werkelijkheid is het gunstig om meer knooppunten te hebben voor nauwkeurigheid en nuance. Hoewel we in het vorige hoofdstuk hebben geleerd dat trainingsdata belangrijk is, is de grootte van het neurale netwerk ook een knelpunt voor de prestaties. Als je één knooppunt had, ongeacht hoeveel data je eraan hebt toegevoegd en hoe goed je data is, zou je slechts één waarde bijwerken en uiteindelijk zou je AI niet nuttig zijn.

Met de inzichten uit dit hoofdstuk zou het duidelijk moeten zijn dat de architectuur van een AI-model direct van invloed is op wat het kan. Neurale netwerken zijn veelzijdiger omdat de invoergegevens flexibel kunnen zijn als de grootte van de hersenen (het aantal parameters) voldoende is. Genetische algoritmen zijn per geval specifiek vanwege de gedwongen invoer en uitvoer, maar hebben zeker hun sterke punten wanneer ze in de juiste context worden gebruikt. Er zijn veel meer architecturen te verkennen, sommige zeer geavanceerd en andere meer basaal. Over het algemeen zijn basismodellen sneller te gebruiken en hebben ze de voorkeur voor goed gedefinieerde verwachtingen. Geavanceerde modellen zijn duur om te trainen en vereisen state-of-the-art hardware, maar leveren verbluffende resultaten op als ze effectief worden ontworpen. Zorg er altijd voor dat u analyseert welke architectuur geschikt is, aangezien dit de resultaten zal bepalen. Stockfish is een goede schaak-AI, beter dan welke mens dan ook in de geschiedenis, terwijl ChatGPT nog steeds illegale zetten maakt tijdens partijen. Beide zijn AI, maar ze worden met verschillende intenties getraind.

Funded by
the European Union

4: AI TRAINEN

Nu we het belang van het selecteren van de juiste trainingsdata en -architectuur hebben besproken, is het tijd om te zien hoe alles samenkomt. De laatste stap is het trainen van de AI. Je hebt het gebouw gebouwd, je hebt de meubels gekocht en nu is het tijd om alles in te richten. In dit hoofdstuk bespreken we verschillende relevante onderwerpen met betrekking tot het trainen van AI. In de context van GA's leggen we het belang van fitnessfuncties uit en relateren we dit ook breder aan reinforcement learning. Daarna bekijken we het verschil tussen supervised en unsupervised learning.

4.1: Fitnessfuncties en versterkend leren

Hoewel er enkele verschillen zijn tussen deze twee concepten, zijn ze samengevoegd omdat de AI leert van interactie met een omgeving. Fitnessfuncties spelen een prominente rol in GA, omdat de oplossingen op een bepaalde manier moeten worden gerangschikt, zodat de bovenliggende oplossingen kunnen worden geselecteerd. Het is daarom belangrijk om te overwegen wat precies de meest effectieve fitnessfunctie is.

Laten we even veronderstellen dat Google Maps een AI gebruikt om de kortste route te vinden. Wat zou de fitnessfunctie dan zijn? Zou het de kortste afstand zijn? Het minste brandstofverbruik? De kortste tijd om de bestemming te bereiken? Er zijn altijd afwegingen: het is vrij eenvoudig om de afstand te berekenen, maar als je de tijd wilt berekenen, moet je rekening houden met de maximumsnelheid van elke weg, verkeerslichten, verkeer, wegafsluitingen, enz. Dit kan langer

Funded by the European Union

duren voordat het systeem de oplossingen scoort en het evolutieproces vertragen. Het is echter belangrijk om een geschikte fitnessfunctie te hebben, omdat deze de resultaten bepaalt die je krijgt. De antwoorden op de vragen over de snelste route en brandstofefficiëntie kunnen zelfs verschillen.

Als je bijvoorbeeld een schaak-AI traint en hem vertelt dat het aantal stukken op het bord de fitnessfunctie is, kun je schaakmat worden gezet, ook al verlies je geen stukken. Daarom moet je goed nadenken over welke fitnessfunctie je moet gebruiken.

Reinforcement learning is hier in sommige opzichten mee verwant, omdat het fitnessfuncties omvat. Het is prominent aanwezig in gesimuleerd leren – waarbij de AI zich in een gesimuleerde omgeving bevindt en probeert een doel te bereiken. Het concept van reinforcement learning is dat de AI beloond wordt voor het juiste en gestraft voor het verkeerde. Als de AI bijvoorbeeld een auto in een racegame is, krijgt hij meer punten voor het sneller bereiken van de finishlijn, maar kan hij ook punten verliezen voor het in de verkeerde richting rijden. Op deze manier blijven in elke volgende generatie de betere oplossingen behouden en worden de slechtste niet gereproduceerd. Nogmaals, de functie die gebruikt wordt om de AI te beoordelen, moet zorgvuldig worden ontworpen.

De AI moet voldoende ruimte krijgen om dingen uit te zoeken, want hij kan een betere oplossing vinden dan je verwacht. Maar als je te streng bent met de fitnessfunctie, vindt hij de oplossing die je zoekt, en die kun je zelf ook vinden. In hetzelfde voorbeeld van de racegame: als je hem punten geeft voor het scherp nemen van de bochten, ben je streng en laat je hem niet experimenteren. Het kan zijn dat ver van één bocht afgaan en je snelheid behouden effectiever is in de context van de hele race.

Funded by
the European Union

4.2: Begeleid en onbegeleid leren

GA's, fitnessfuncties en reinforcement learning zijn allemaal aan elkaar gerelateerd, omdat ze interacties met een omgeving gebruiken om de AI te trainen door hem te vertellen wat als resultaat een hogere score oplevert. Het andere paradigma is niet om de AI te vertellen of hij het goed doet, maar hem gewoon te trainen met een heleboel data en te kijken wat er gebeurt.

Bij supervised learning is er, ondanks de naam, geen mens die het trainingsproces direct monitort. Het verschil zit in het type data dat aan de AI wordt doorgegeven. Bij supervised learning worden de data gelabeld (meestal door mensen). Dit betekent dat iemand de "juiste antwoorden" heeft gegeven, vergelijkbaar met studeren voor een examen door de antwoorden op eerdere examenvragen te gebruiken. Dit is nuttig wanneer er een duidelijk resultaat wordt verwacht, en het levert AI-modellen op die voorspelbaarder en minder volatiel zijn (met minder variantie in nauwkeurigheid en consistentie).

Ongesuperviseerd leren vindt plaats wanneer de data niet gelabeld is en de AI zoekt naar patronen in de data zelf in plaats van verbanden te leggen tussen de data en de labels. Dit is handig wanneer de verwachtingen subjectiever zijn, zoals in geschreven Engels - er zijn veel manieren om hetzelfde te zeggen en geen daarvan is inherent fout. In dit geval is het onpraktisch om de data te labelen en is het effectiever om de AI zelf patronen te laten ontdekken.

Nu we weten dat het onderscheid tussen supervised en unsupervised learning het type data is dat aan de AI wordt verstrekt, laten we de

Funded by the European Union

toepassingen van beide bekijken. In geavanceerde modellen zoals LLM's (ChatGPT) worden beide gebruikt om de AI meer mogelijkheden te geven, maar waar is elk nuttig voor?

Dit onderscheid is belangrijk, omdat beide vormen van leren specifieke kenmerken hebben. Supervised learning kan duur zijn omdat de data gelabeld moeten worden – dit is een intensief proces vanwege de hoeveelheid data die nodig is om een complex model te trainen. Het heeft het voordeel dat de AI de verwachte output kent, omdat deze goed gedefinieerd en duidelijk gestructureerd is. Unsupervised learning daarentegen kan alle data die je geeft gebruiken en proberen daar patronen uit te construeren. De resultaten zijn daardoor minder expliciet en gestructureerd, maar gemakkelijker te trainen omdat er geen speciale data voor nodig is. Het is goed in het vinden van kleine nuances en patronen die anders niet voor de hand liggen.

Funded by
the European Union

5: Hoe gebruik je AI effectief

Nu je inzicht hebt in wat er allemaal komt kijken bij het creëren van een AI-model, is het tijd om de huidige bestaande producten te bekijken en te kijken hoe deze het meest effectief kunnen worden ingezet. Het is geen geheim dat AI je productiviteit een boost kan geven. Daarom bespreken we in dit hoofdstuk de beste toepassingen voor LLM's en in welke situaties AI contraproductief kan zijn. LLM's zoals ChatGPT kunnen in veel situaties nuttig zijn, maar aangezien deze cursus bedoeld is voor jonge ondernemers, richten we ons op de meest bedrijfsgerichte toepassingen.

Brainstormen:

Voordelen

- Snelle generatie van gerelateerde onderwerpen
- Vertrouwd met veel concepten en goed in het verbinden van gerelateerde punten
- Kan ideeën met je bespreken, vergelijkbaar met hoe je met iemand zou praten

Beperkingen

- Soms begrijpt hij niet wat je verwacht, dus moet je specifieker zijn
- Alleen getraind met beschikbare data. Als het om iets bedrijfseigens gaat of buiten de dataset valt, heb je pech.

Funded by the European Union

met hoe je met iemand zou praten

Tekst schrijven:

Voordelen

- Schrijft doorgaans coherenter dan de meeste mensen, vooral in hun tweede taal
- Vertrouwd met veel concepten en goed in het verbinden van gerelateerde punten
- Kan ideeën met je bespreken, vergelijkbaar

Beperkingen

- Soms is de toon niet relevant voor het type tekst
- Als u samenwerkt met iemand die geen AI-content accepteert, kan hij of zij een AI-scanner gebruiken.
- Het kan langer duren om de context en de verwachting uit te leggen dan wanneer u het zelf schrijft

Feedback geven:

Voordelen

- Ontdek fouten die u misschien over het hoofd ziet
- Sneller dan wanneer u uw eigen werk leest, vooral bij grotere teksten
- Kan verbeteringen voorstellen zodra kritiek is geïdentificeerd

Beperkingen

- Geprogrammeerd om een antwoord te geven, hoe geforceerd ook, wat soms resulteert in triviale feedback
- Soms veel kritischer dan nodig voor de taak die voorhanden is

Funded by the European Union

- Gebruikt vaak meer woorden dan nodig zijn voor de boodschap die het probeert over te brengen, tenzij u het de opdracht geeft dit niet te doen.

- Kan zijn eigen code repareren als het bij de eerste poging niet werkt
- Kan de code regel voor regel uitleggen

Beperkingen

- Bestandssystemen en complexe integraties kunnen niet worden gegenereerd
- Kan tijdverspilling zijn door het genereren van code die u niet begrijpt

Codering:

Voordelen

- Snelle codegeneratie
- Vertrouwd met veel talen en frameworks

Enkele algemene overwegingen bij het gebruik van AI:

- Soms hallucineert hij over valse feiten
- Beperkt tot trainingsgegevens die zijn verstrekt
- Kent alleen de contextuele informatie die u verstrekt

LLM's werken op basis van het voorspellen van het volgende woord dat gegenereerd moet worden. Hoewel nieuwere modellen geavanceerdere systemen hebben voor citaten en informatievoorziening, is het toch verstandig om belangrijke feiten zelf te controleren, aangezien er fouten kunnen ontstaan.

Als de onderwerpen waarover u informatie wilt krijgen of waarover u met de AI wilt discussiëren, niet prominent aanwezig zijn in de trainingsgegevens, zal de AI het niet goed doen in prompts die betrekking hebben op deze onderwerpen.

Als u bijvoorbeeld een tekst voor de website van uw bedrijf probeert te schrijven, moet u de AI een heleboel informatie geven over uw visie,

Funded by the European Union

activiteiten, enzovoort. Als deze informatie niet al in toegankelijke tekst staat die u kunt kopiëren en plakken, kan het langer duren om deze op te schrijven dan wanneer u de tekst voor de website zelf schrijft. Dit geldt vooral wanneer u een korte tekst met een zeer precieze betekenis nodig hebt. In die gevallen kan het effectief zijn om de initiële tekst te schrijven en de AI vervolgens te vragen of er ruimte voor verbetering is.

Funded by
the European Union

6: SAMENVATTING VAN DE BELANGRIJKSTE LEERPUNTEN

Gefeliciteerd, je hebt het einde van de module bereikt! Laten we nu een aantal belangrijke punten doornemen, zodat je de belangrijkste onderdelen niet vergeet.

Uit hoofdstuk 2 hebben we geleerd waarom trainingsdata een significant effect heeft op de output van een AI-model. We hebben ook geleerd dat data van hoge kwaliteit:

- Wees nauwkeurig
- Zorg voor voldoende hoeveelheid
- Wees divers
- Houd rekening met onbedoelde vooroordelen

In hoofdstuk 3 hebben we besproken hoe het gekozen architectuurtype verband houdt met het type probleem dat moet worden opgelost. We hebben besproken dat genetische algoritmen efficiënt zijn voor problemen met een gedefinieerd resultaatformaat, met name optimalisatieproblemen. We hebben ook een basisinzicht ontwikkeld in hoe neurale netwerken (NN's) functioneren en dat er veel verschillende soorten NN's zijn. De NN's zijn flexibel in hun doel en functionaliteit.

In hoofdstuk 4 leerden we hoe reinforcement learning en fitnessfuncties met elkaar samenhangen, hoe ze werken en wat hun doel is. We maakten ook onderscheid tussen supervised en unsupervised learning.

Tot slot hebben we in hoofdstuk 5 besproken hoe je AI het beste kunt inzetten en welke overwegingen je kunt gebruiken om de workflow te verbeteren. We hebben beschreven hoe AI presteert bij taken zoals coderen, brainstormen, copywriting en feedback geven, en wat de sterke en zwakke punten van elke categorie zijn.

Ik hoop dat dit inzicht nuttig is op uw reis met AI. Veel succes.

Funded by the European Union

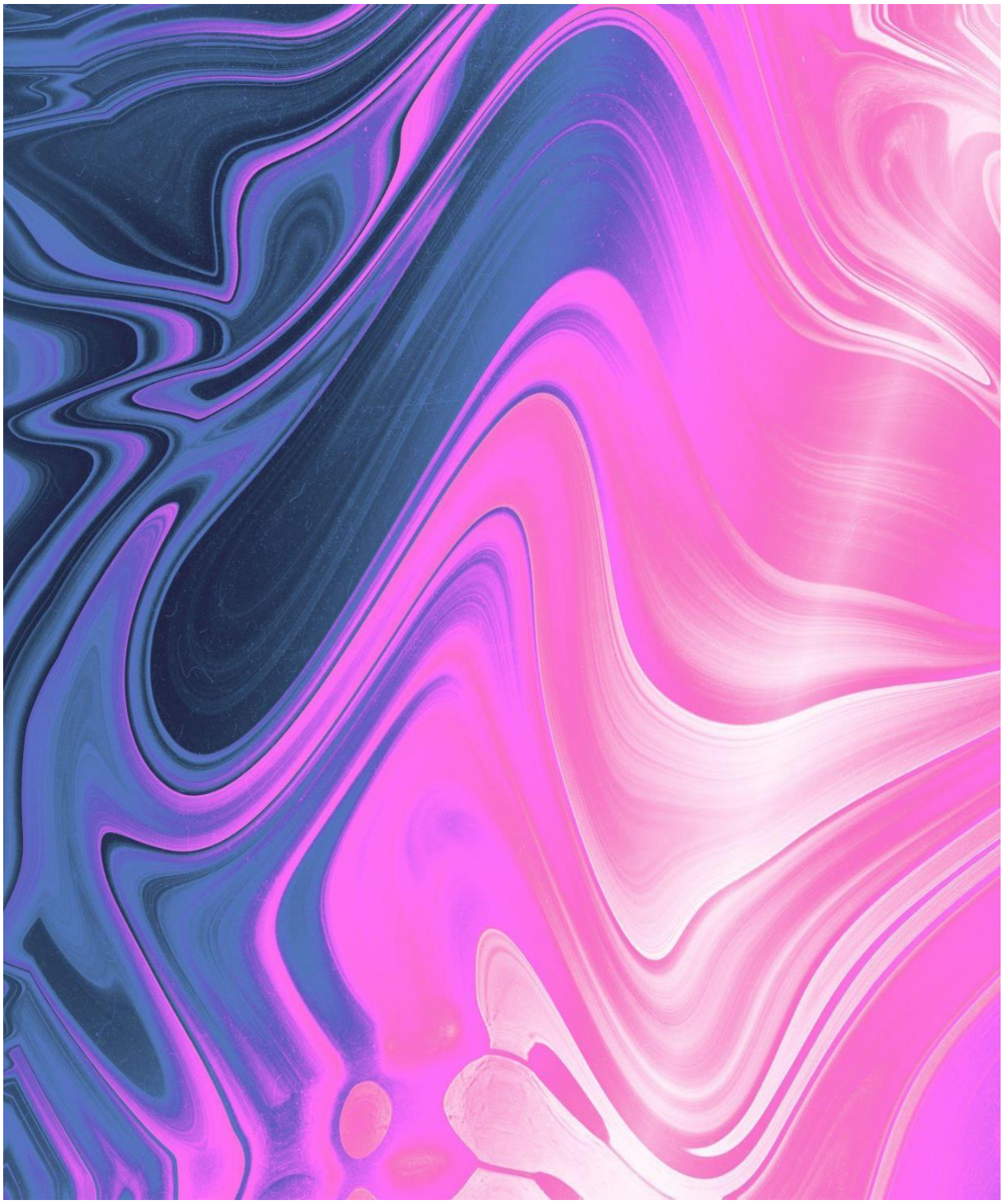
1. BIBLIOGRAFIE

https://www.oracle.com/nl/kunstmatige-intelligentie/ai-model-training/#role-of-data

https://www.nytimes.com/2021/09/03/technology/facebook-ai-race-primates.html

https://www.geeksforgeeks.org/genetische-algoritmen/

Funded by the European Union



Funded by
the European Union