

4 Generic AI

4.1 Inleiding tot AI en Belangrijke Concepten

Kunstmatige Intelligentie (AI) vs. Machine Learning (ML)

- **AI:** Breed veld gericht op het creëren van systemen die menselijke intelligentie nabootsen.
- **ML:** Subset van AI waarbij algoritmen patronen uit data leren.
- **Deep Learning:** Gebruikt neurale netwerken met meerdere lagen.
- **Generatieve AI:** Creëert nieuwe inhoud (bijv. tekst, afbeeldingen).
- **Large Language Models (LLMs):** Geavanceerde modellen zoals ChatGPT die mensachtige tekst genereren.

4.2 Wiskundige Grondslagen van AI

Modelleren met Polynoomfuncties

- AI kan worden gezien als een wiskundige formule.
- Voorbeeld: Voorspellen van de prijs van brood over tijd:

$$\text{Prijs}(x) = 1.61 + x \cdot \frac{(2.94 - 1.61)}{20}$$

- Algemene polynoomfunctie:

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

Complexiteit in Multi-Variabele Modellen

- Voor m -variabele polynomen van orde n :

$$\text{Parameters} = \frac{(n + m)!}{n! \cdot m!}$$

- Het oplossen van hoogdimensionale systemen is rekenintensief.

4.3 Neurale Netwerken en Functiebenadering

Self-Organizing Maps (SOM)

- Neurale netwerken benaderen functies door gewichten (w) aan te passen:

$$\text{Uitvoer} = \text{SOM}(w \cdot O_x)$$

- **Training:** Minimaliseer fouten door iteratieve aanpassingen (bijv. gradient descent).

4.4 Large Language Models (LLMs)

Tekstgeneratieproces

1. **Tokenisatie:** Verdeel tekst in tokens (bijv. “Ongelooflijk” \rightarrow [“On”, “geloof”, “lijk”]).
2. **Embedding:** Map tokens naar hoogdimensionale vectoren (bijv. GPT-3 gebruikt 12.288 dimensies).
3. **Attention Mechanism:** Pas woordgewichten aan op basis van context (bijv. “koning” vs. “koningin”).
4. **Voorspelling:** Neuraal netwerk voorspelt waarschijnlijkheden van het volgende woord.

Technische Specificaties van GPT-3

- **Lagen:** 96
- **Attention Heads:** 96
- **Parameters:** 175 miljard
- **Trainingskosten:** ~75 miljoen euro (24.000 NVIDIA H100 GPU's).

4.5 Bouwen en Inzetten van Generatieve AI

Training vs. Inferentie

- **Training:**
 - Voorbeeld: Meta's LLaMA-3 duurde 3 maanden op 24.000 GPU's.
 - Energie: ~550 MW/dag (gelijk aan het voeden van 75.000 huishoudens).
- **Inferentie:**
 - OpenAI gebruikt 30.000 GPU's voor real-time taken.
 - Energie: ~615 MW/dag (gelijk aan het voeden van 110.000 huishoudens).

4.6 Uitdagingen in AI-ontwikkeling

Belangrijke Technische Uitdagingen

1. **Nauwkeurigheid van Algoritmen:** Fouten in beeld-/tekstgeneratie (bijv. verkeerd gelabelde objecten).
2. **Energieverbruik:** Hoge energiebehoefte voor training/inferentie.
3. **Gedistribueerde Intelligentie:** Coördinatie van AI over gedecentraliseerde systemen.
4. **Controle en Veiligheid:** Voorkomen van irrationeel gedrag (bijv. adversarial attacks op Tesla Autopilot).

4.7 Ethische Overwegingen en Valkuilen

Bias in AI-systemen

1. **Dataset Bias:**
 - Gezichtsherkenning getraind op niet-diverse datasets.
 - Historische data die oneerlijke praktijken in stand houden.
2. **Onethische Toepassingen:**
 - AI-modellen gebruikt voor schadelijke doeleinden (bijv. chatbots die zelfmoordgedachten aanmoedigen).

Casestudy's

- **Tesla Autopilot:** Kleine stickers misleiden het systeem om van baan te wisselen.
- **Dodelijke Ongevallen:** Zelfrijdende auto's betrokken bij fatale ongevallen (bijv. Uber's incident in 2018).

4.8 Toekomstige Richtingen en Verantwoord Gebruik

Actiepunten voor Ethische AI

- Train teams in ethische AI-implementatie.
- Valideer modellen met onafhankelijke teams om bias te verminderen.
- Monitor databronnen voor veranderingen in populatie.
- Gebruik diverse teams voor ontwikkeling en testen.

Publieke Perceptie en Veiligheid

- 52% van de Amerikanen vindt telefoongebruik acceptabel in autonome voertuigen.
- Traditionele auto's veroorzaken 1,3 miljoen doden per jaar vs. ~100 voor Tesla (geëxtrapoleerd).

Key Points to Remember

- **AI vs. ML:** AI is het bredere veld; ML richt zich op data-gedreven leren.
- **Wiskundige Kern:** AI-modellen zijn gebaseerd op polynoomfuncties en parameteroptimalisatie.
- **LLM Mechanica:** Tokenisatie, embedding en attention sturen tekstgeneratie aan.
- **Energiekosten:** Het trainen van GPT-3 verbruikt energie gelijk aan een kleine stad.
- **Bias Mitigatie:** Representatieve datasets en ethische validatie zijn cruciaal.
- **Veiligheidsuitdagingen:** Adversarial attacks en controleproblemen vereisen robuuste veiligheidsmaatregelen.
- **Ethische Inzet:** Diverse teams en continue monitoring zorgen voor verantwoord gebruik.