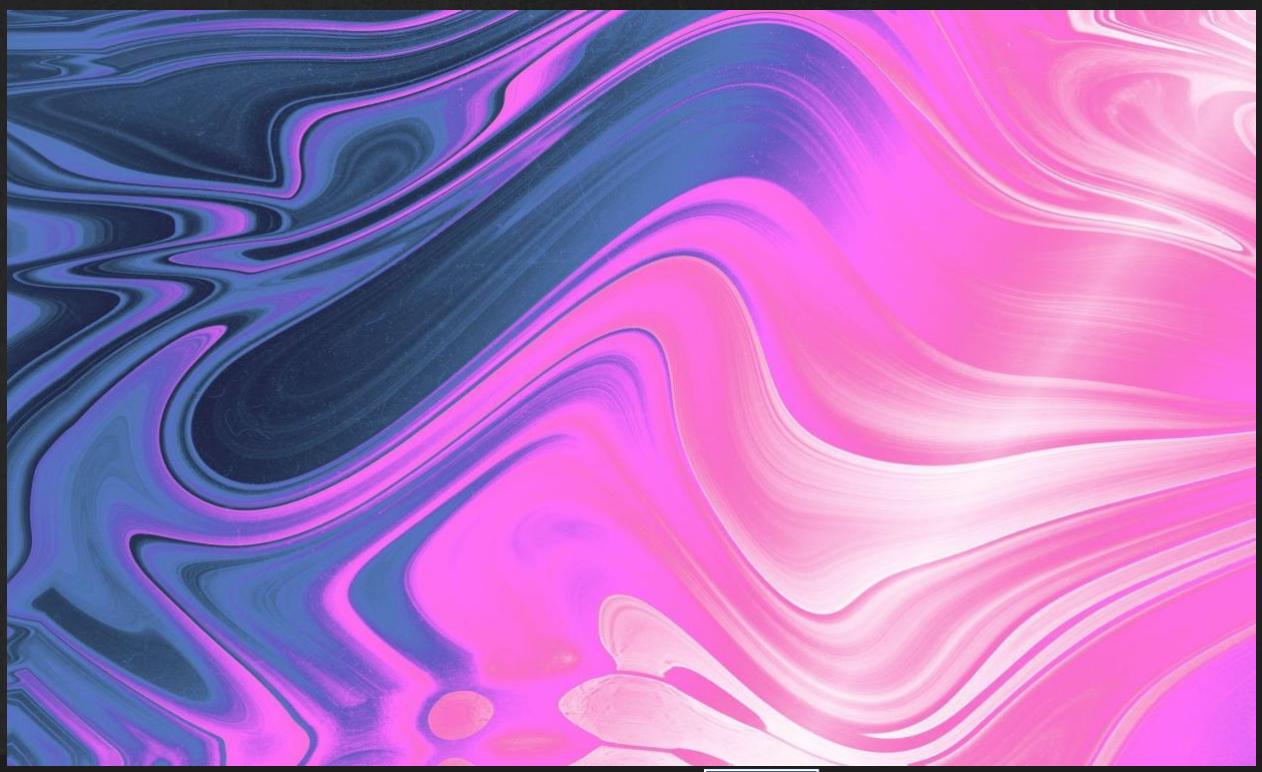


Funded by
the European Union

STYRKER UNGE GRUNDERE MED KI

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

Funded by
the European Union

INNHOLDSSIDE

Funded by the European Union

INTRODUKSJON

I denne modulen diskuteres det grunnleggende innen kunstig intelligens, og det gis informasjon om grunnpilarene som trengs for å lage en kunstig intelligens-modell. Flaskehalsene ved å lage kraftige kunstig intelligens-modeller diskuteres, inkludert designbeslutninger på et arkitektonisk nivå samt nyanser som må vurderes når man trener modeller.

Etter å ha lest denne modulen bør du ha en dypere forståelse av temaene knyttet til KI. Denne modulen fokuserer hovedsakelig på den teknologiske siden av KI, men foreslår også hvordan man kan bruke KI mest effektivt.

- Du vil lære om hvor viktig treningsdata er og hva du bør se etter.
- Du vil lære om de ulike arkitekturene som er tilgjengelige i dag, samt deres styrker og svakheter.
- Du vil lære om de mest produktive måtene å bruke tilgjengelige AI-modeller på.

Funded by the European Union

1: AI OG MASKINLÆRING

Etter hvert som bransjen har utviklet seg, har oppfatningen av betydningen av AI endret seg. I moderne tid, når folk sier AI, refererer de først og fremst til store maskinlæringsmodeller, som er det som vil bli diskutert i denne modulen. Konseptet maskinlæring refererer til en algoritme som kan finne mønstre i inndata og forbedrer seg selv etter hvert som den trenes mer omfattende. Nøkkelen er at den ikke blir fortalt hvordan den skal tenke. Noen modeller fungerer ved å etterligne den biologiske hjernen.

I noen tilfeller er det bare et moteord som brukes for å fange oppmerksomhet og selge produkter, så det er avgjørende å forstå hvordan denne mystiske og ofte misforståtte teknologien fungerer. Ved å forstå nøkkelkonseptene vil du få en bedre ide om hva som er mulig og ikke bli villedet av hypen. Du vil også være bedre rustet til å dra nytte av kraften som AI har å tilby.

I denne modulen vil vi dekke forståelsen av sentrale temaer innen maskinlæring. Vi vil fremheve viktigheten av treningsdata av høy kvalitet. Deretter vil vi dekke håndplukkede eksempler på AI-arkitekturer siden det finnes så mange, inkludert informasjon om hvordan de fungerer og hva du kan forvente av dem som styrker og svakheter. Følgelig vil vi se på hvordan ulike AI-modeller kan trenes og hvilke typer trening som finnes, og dermed blande arkitekturen og treningsdataene. Til slutt vil vi se på den nåværende tilstanden til AI og kommentere hvordan man får mest mulig ut av produktene som er tilgjengelige på markedet.

Funded by the European Union

2: VIKTIGHETEN AV TRENINGSDATA

Som i de fleste systemer er AI bare så sterk som det svakeste leddet. Hvis du har gode data, men ingen måte å behandle dem på, har du ingenting, og selv om du har den mest avanserte arkitekturen, men du forsyner den med data av lav kvalitet, vil resultatene være skuffende.

I dette kapittelet vil vi diskutere viktigheten av å tilby data av høy kvalitet når man trener en AI-modell, i tillegg til å evaluere hvilke kriterier data må oppfylle for å bli ansett som høy kvalitet.

Det finnes systemer som bruker simulert læring der en AI samhandler med et simulert miljø og lærer ved prøving og feiling gjennom forsterkning (dette vil bli forklart i kapittel 4). For andre AI-modeller er treningsdata den eneste kilden til forståelse for AI-en. Alle mønstrene som observeres av modellen er basert på dataene den blir matet med. Dette er fordi AI-en, i motsetning til mennesker, ikke har noen kroppslig eller personlig erfaring å trekke kunnskap fra og er avhengig av input for å definere dens sannhet. Den har ingen ytterligere måte å teste konklusjonene sine på med mindre den senere får beskjed om at informasjonen er feil. Dette er grunnen til at data av høy kvalitet er nødvendige for at AI-modeller skal fungere. Når det er sagt, hva utgjør data av høy kvalitet?

Den første vurderingen er nøyaktighet. Selv om det finnes både merkede og umerkede datasett, er nøyaktighet avgjørende for begge. Når merkede data brukes i kombinasjon med veiledet læring (mer om dette i kapittel 4), er nøyaktigheten av dataene relatert til hvor detaljerte og korrekte etikettene er i forhold til dataene. For eksempel kan du trene

Funded by the European Union

en AI til å identifisere objekter i et bilde.

Eksempel på inndatabilde https://www.rspb.org.uk/birds-and-wildlife/robin

Ta dette bildet for eksempel. Det kan brukes som input for å trene en AI. Bildet forblir det samme, men taggene kan variere avhengig av ønsket kompleksitet for AI-en. Én etikett kan være «fugl»; dette vil begrense forståelsen av AI-en. Et annet mer avansert datasett kan merke dette bildet med taggene «fugl, rødstrupe, sittende, stående på gren, oransje bryst, uskarp bakgrunn» og så videre. Du kan tenke deg forskjellen i de resulterende AI-modellene. Hvis du bruker et upålitelig datasett, kan dette bildet til og med ha en etikett som «katt». Dette er selvfølgelig overdrevet, men hvis du lærer AI-en at dette er en katt, kan den referere til den som en katt, selv om den forstår hva en fugl er. Av denne grunn er det avgjørende å bruke detaljerte og nøyaktige merkede data for riktig funksjonalitet når man trener en AI.

Funded by the European Union

Hvis vi ønsker å gi et eksempel på nøyaktighet i konteksten av umerkede data brukt i uovervåket trening, kan vi forestille oss en stor språkmodell (LLM). LLM-er som ChatGPT kan trenes på umerkede data som artikler, tweets, Reddit-innlegg, vitenskapelige tidsskrifter osv. Dette er umerkede data, og AI-en bruker dem til å identifisere mønstre i setningsstrukturen og danner relasjoner mellom ord. Ved å trene den på nok av disse dataene, begynner AI-en å produsere setninger som et menneske kan skrive. Hvis du derimot trener den på skolelekser skrevet av barn på 6 år, kan du få uønskede resultater fra AI-en. Dette er hva kvalitet betyr i konteksten av umerkede data.

Kort sagt, hvis du lærer den ting som ikke er sanne, har ikke AI-en noen måte å korrigere dette på. Sørg for at dataene som brukes til å trene en AI er av høy kvalitet, ellers kan den identifisere katter som hunder. Hvis du trener den på tekst, sørg for at den er skrevet i sammenhengende og grammatisk korrekt tekst, ellers kan resultatene være meningsløse. Selv om et lite delsett av dataene er av lav kvalitet, kan det forurense modellen siden AI jobber med mønstre og vil gjøre feil koblinger. Kvaliteten på inndataene er direkte korrelert med kvaliteten på resultatene.

En annen faktor når det gjelder data av høy kvalitet er innebygde skjevheter. Selv om det teknisk sett ikke er noe galt med dataene, kan du fortsatt få store problemer hvis du overser skjevheter. I en PR-skandale ser det ut til at en kunstig intelligens for bildegjenkjenning hovedsakelig ble trent på kaukasiske menneskeansikter, noe som førte til at den identifiserte mennesker fra andre raser som primater. Dette kan løses ved å ta hensyn til de to neste kriteriene for data av høy kvalitet – volum og mangfold.

Jo større hjernen til AI-en er – målt i antall parametere – desto mer data trengs for å kode verdier nøyaktig og effektivt inn i nodene, siden hver inngang bare vil endre nodeverdiene litt. Tenk på noder som

Funded by the European Union

matematiske funksjoner – de tar litt inngang og gir noe utdata etter å ha gjort litt matematikk. Av denne grunn er det viktig å ha tilstrekkelig volum når du trener en AI. Det finnes forskjellige forslag på nettet for hvor mye data som er nok, men den eneste virkelige måten å vite det på er å teste AI-en. Dette tester selvfølgelig hele systemet og ikke bare mengden inngangsdata, så det er vanskelig å identifisere det svake punktet, men ved å teste modellen vil du få en følelse av resultatene du kan forvente.

Den siste viktige faktoren for treningsdata er mangfoldet. Hvis du gir AI-en for mye av én type input, vil den overtilpasse og ikke gjenkjenne generelle mønstre, men bare optimaliseres for det ene spørsmålet. Dette avhenger selvfølgelig av formålet med AI-en, men for store språkmodeller er det avgjørende at de har en bred forståelse og prøver å resonnere på egenhånd. La oss se på noen eksempler: Hvis du bare viser den ved å legge til spørsmål, vil den ikke vite hvordan den skal multiplisere. Hvis du viser den for mange artikler, vil den høres ut som en reporter i alle sammenhenger, så du vil kanskje inkludere Reddit-tråder eller tweets. I en AI for bildegjenkjenning kan du mate den med for mange lignende bilder – hvis alle bilder med en fugl har en blå himmelbakgrunn, kan den anta at den blå fargen bidrar til at bildet er en fugl.

Funded by
the European Union

3: AI-ARKITEKTURER

Selv om treningsdata er viktige for den endelige funksjonaliteten til en AI-modell, er det ikke den eneste faktoren. For at disse dataene skal være nyttige, er det avgjørende å ha en passende arkitektur å basere modellen på. Tenk på arkitekturen som et skjelett, bare en haug med bein som må kobles sammen av muskler. Det finnes mange typer arkitekturer som kan brukes til forskjellige formål. Noen AI-modeller er smale, kjent som kunstig smal intelligens (ANI), og er optimalisert for ett veldig spesifikt formål, som sjakk-AI-er, mens andre kalles kunstig generell intelligens (AGI) og forsøker å etterligne menneskelig forståelse på alle felt. AGI er ikke oppnådd ennå, og det er mange debatter om definisjoner og bevissthet. Store språkmodeller (LLM-er) som ChatGPT er et forsøk på å lage AGI og er betydelig mer avanserte enn teknologien var for noen år siden. Når det er sagt, finnes det for begge kategorier mange forskjellige arkitekturer, hver med sine superkrefter, begrensninger og spesifikke brukstilfeller der disse avveiningene gir mening. For å gi en grunnleggende forståelse av hva som er mulig, vil vi bare diskutere to av disse - genetiske algoritmer og nevrale nettverk.

3.1: Genetiske algoritmer

Genetiske algoritmer (GA-er) er én type AI-arkitektur, litt annerledes enn det du sannsynligvis er vant til å se i mainstream-verdenen. For å utvide forståelsen av ulike typer AI skal vi se på hva de er nyttige til og hvordan de fungerer. Dette er en svært spesialisert type arkitektur som ikke er bredt nyttig for mange forskjellige typer oppgaver, men som er svært effektiv for bestemte problemer. Vanligvis brukes GA-er i optimaliseringsproblemer der det er flere variabler enn det som kan brutalt tvinges frem. Det er en elegant lettvektsløsning for å teste

Funded by the European Union

mulige løsninger på strengt definerte problemer. Alt dette kan høres litt abstrakt ut, så la oss undersøke det mest populære eksemplet – problemet med den reisende selgeren.

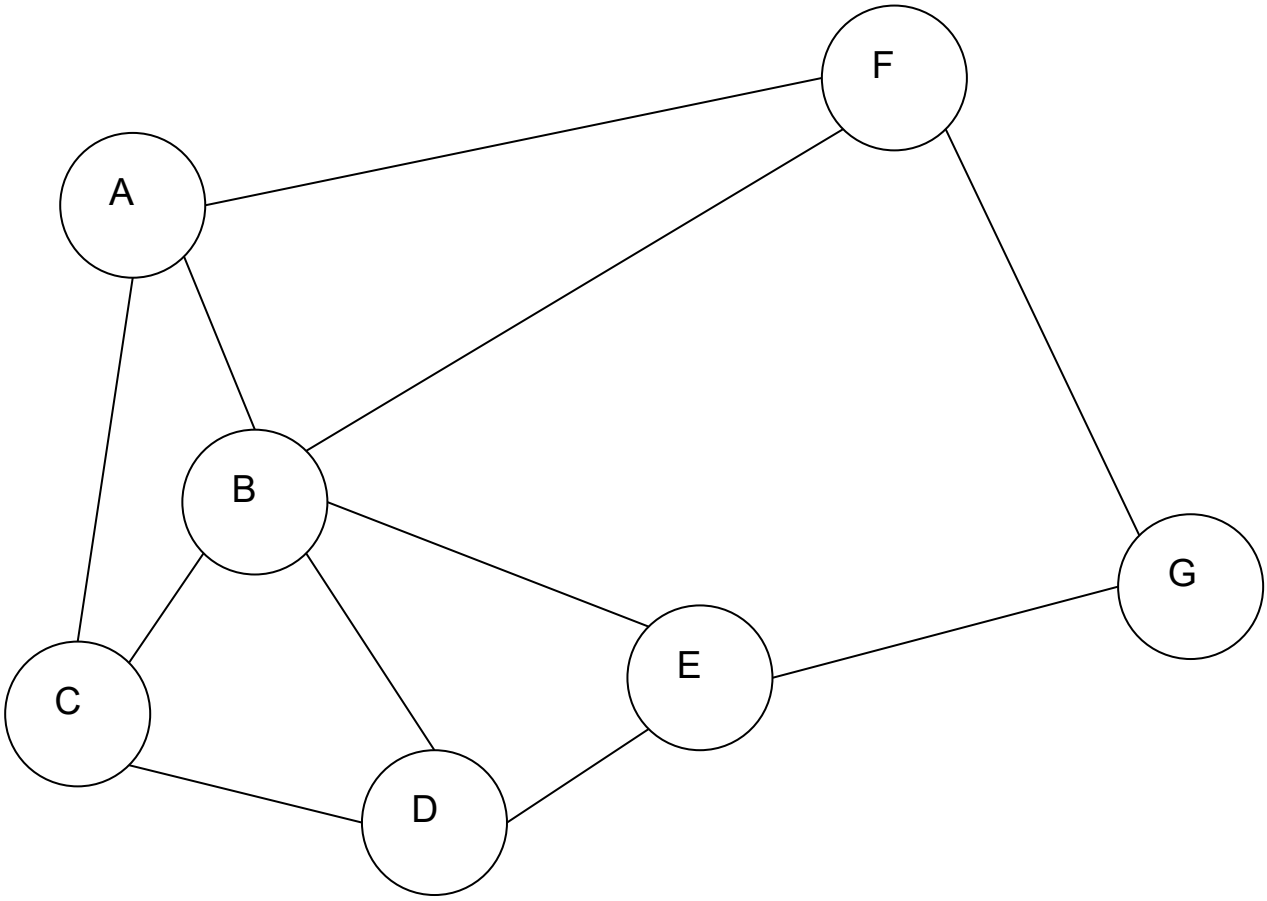


Diagram over potensielt scenario

I tegningen ovenfor kan du se 7 noder. Problemet med den reisende selgeren forventer en løsning der hver node eller by besøkes nøyaktig én gang, og målet er å finne den korteste veien. For enkelhets skyld er det ingen avstander mellom nodene i dette eksemplet. En potensiell løsning er ABCDEFG. En annen løsning er ABCDEGF. Selv om det finnes andre måter å løse dette på, presterer reisehandlere veldig bra i slike problemer.

Funded by the European Union

Så hvordan fungerer de? I dette eksemplet er hver løsning 7 bokstaver lang og inneholder hver bokstav nøyaktig én gang. Tenk på dette som DNA-et som genetikken er basert på, plasseringen av hver bokstav er en egenskap ved løsningen. Hver løsning er et medlem av populasjonen.

I utgangspunktet genereres en tilfeldig populasjon med alle typer løsninger, hvorav noen er ineffektive og andre er raskere. En fitnessfunksjon er et rangeringssystem som brukes til å sammenligne kvaliteten på løsninger. I dette eksemplet er fitnessfunksjonen bare avstanden til banen, der kortere avstand er bedre. Dette betyr at løsninger med kortere totalavstand vil bli rangert bedre og brukt til å produsere neste generasjon. Med 3 noder, hvis løsningen er ABC, er avstanden AB + AC, der AB er avstanden fra A til B.

Når populasjonen er generert, bør den sorteres etter fitnessfunksjonen. Deretter kan neste generasjon produseres fra foreldrene til den nåværende generasjonen. Selv om høyere rangerte løsninger har større sannsynlighet for å reprodusere seg, er det en viss tilfeldighet for å unngå for tidlig konvergens og lokale maksima. Dette emnet er for nyansert til å gå dypere inn i innenfor rammen av dette kurset, men ekstra ressurser kan finnes for nysgjerrige elever.

Reproduksjon gjøres ved å ta de to foreldrene og kombinere dem for å variere avkommet. Selv om det finnes mange forskjellige måter å kombinere genetikken på, er den generelle ideen at den nye populasjonen i gjennomsnitt er bedre siden de minst effektive løsningene fra hver generasjon fjernes. Ved å kjøre algoritmen i flere generasjoner kan du holde oversikt over de beste løsningene til én er god nok til formålet, da det aldri er noen garanti for at det er den best mulige løsningen.

Funded by the European Union

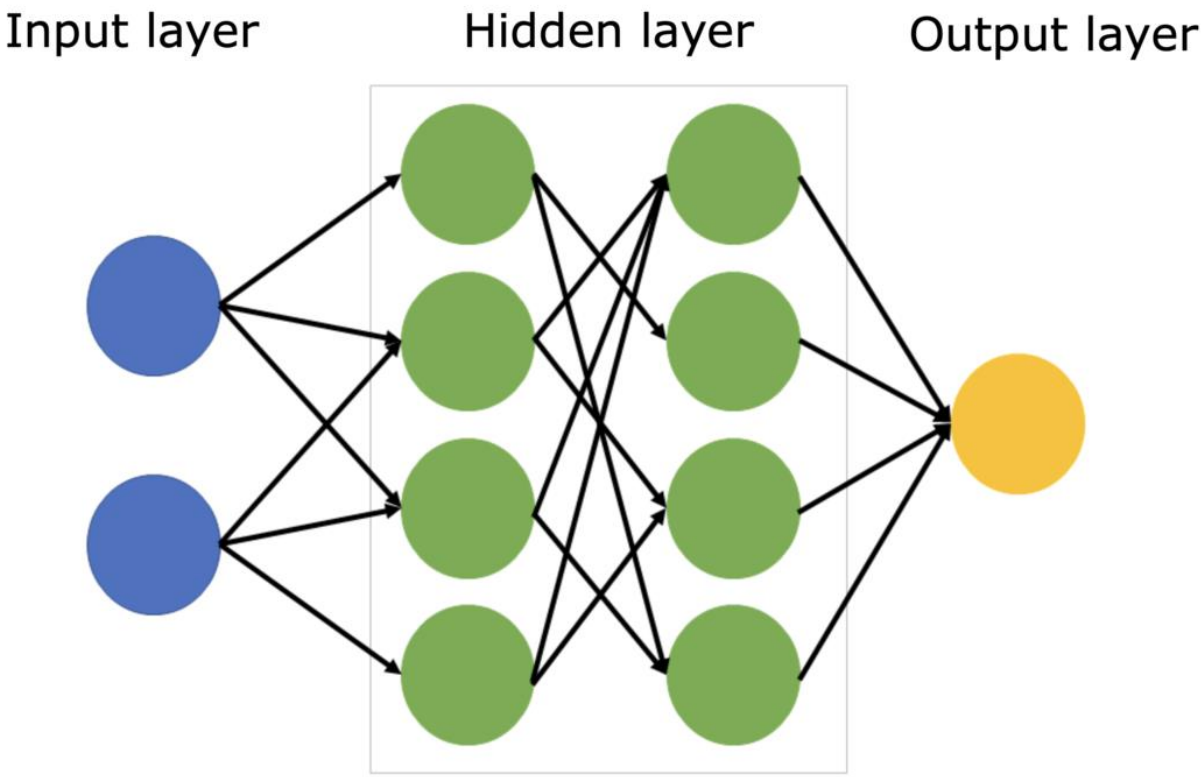
3.2: Nevrale nettverk

Nevrale nettverk (NN-er) er hjørnesteinen i moderne kunstig intelligens. Konseptet er at de etterligner hvordan menneskelige hjerner fungerer med nevroner (hjerneceller) og forbindelsene mellom dem. I kunstig intelligens kalles hvert nevron en node, og datamaskinrepresentasjonen er en funksjon for behandling av data. Et enkelt tilfelle er der noden har en verdi mellom 1 og 0, og inputen multipliseres med verdien, dette er en funksjon. Forbindelsene har også ofte verdier kalt vekter, som representerer innflytelsen fra én node på den neste.

Det finnes et stort antall forskjellige strukturer for nevrale nettverk, inkludert transformatorarkitekturen – dette er hva de fleste LLM-er som ChatGPT er basert på. I dette kapittelet vil vi diskutere den brede ideen om et nevralt nettverk snarere enn detaljene som utgjør mer avanserte arkitekturer.

Funded by
the European Union

For å utvikle en intuisjon for hvordan NN-er fungerer, vil vi bruke et visuelt eksempel på en AI for bildegjenkjenning.

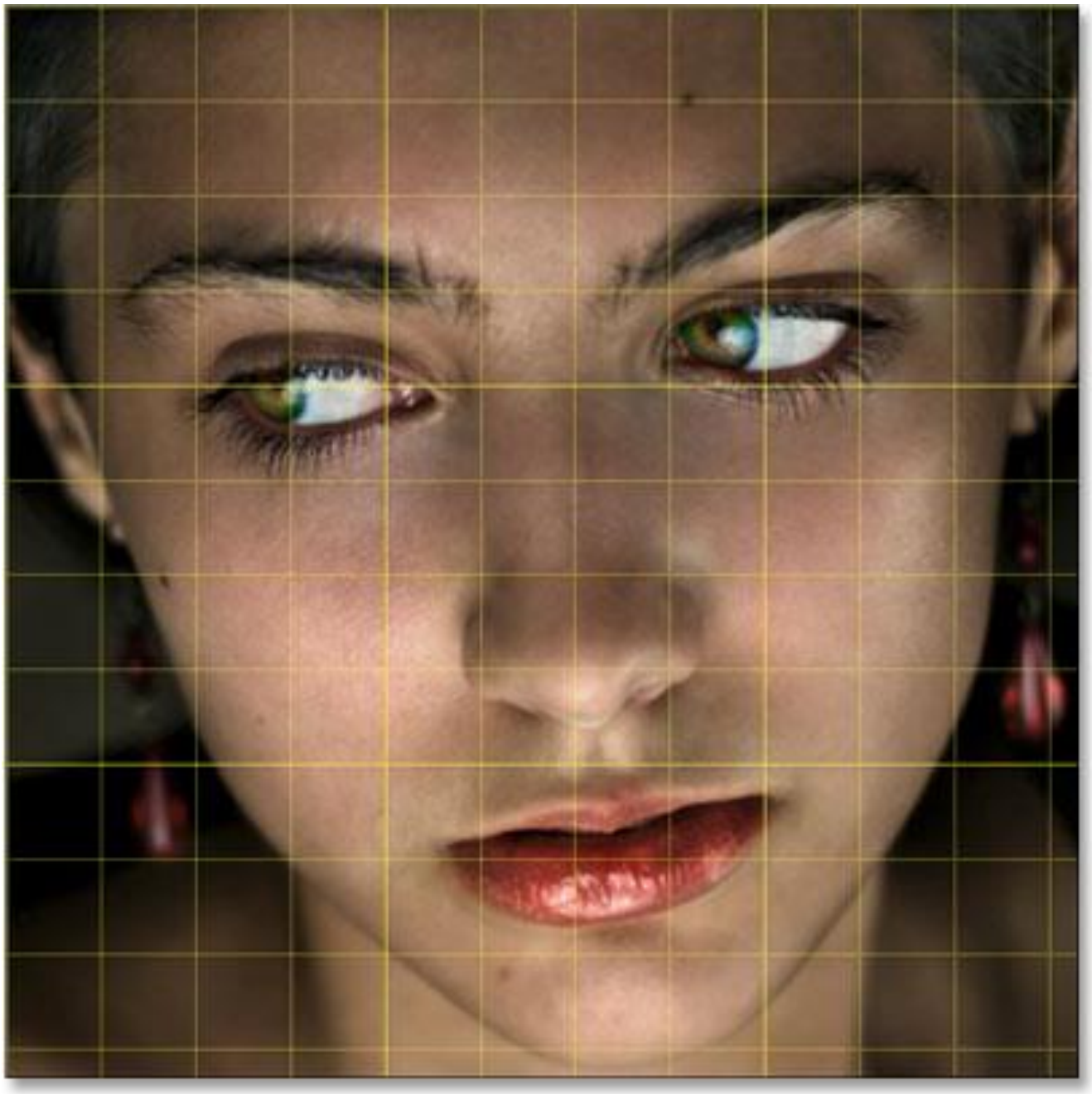


Artificial neural networks

Potensiell representasjon av nevralt nettverk
https://www.sciencelearn.org.nz/images/5156-neural-network-diagram

I dette eksemplet er nodene sirklene og pilene er forbindelsene. Vekten av forbindelsene og nodene påvirker begge resultatene. Vi vil forstå hvor disse verdiene kommer fra i kapittel 4 – trening av AI.

Funded by the European Union



Bilde med rutenett https://www.photoshopessentials.com/photo-effects/photo-strips/

Tenk deg nå et rutenett som dette over et bilde. I stedet for disse firkantene er imidlertid hver piksel en blå inputsirkel i NN. Disse inputene behandles deretter av nettverket, der hver forbindelse

Funded by the European Union

analyserer mønstre mellom de nærliggende pikslene. Når inputen passerer gjennom nok skjulte lag, er den i stand til å identifisere strukturer basert på inputpikslene. Den vil da kunne beskrive hva disse strukturene er.

Kapasitetene avhenger av størrelsen på det nevrale nettverket, ettersom det bare kan ta så mange input som det har noder. Tenk på det som et puslespill: hvis hver node kunne fortelle deg hvor en brikke skal være, ville du trenge like mange noder som du har brikker. Dette er imidlertid forenklet, og i realiteten er det fordelaktig å ha flere noder for nøyaktighet og nyanser. Selv om vi lærte at treningsdata er viktig i forrige kapittel, er størrelsen på det nevrale nettverket også en flaskehals for ytelse. Hvis du hadde én node, uansett hvor mye data du matet den med og hvor gode dataene dine er, ville alt du ville gjort være å oppdatere én verdi, og til slutt ville ikke AI-en din være nyttig.

Med forståelsen fra dette kapittelet burde det være klart at arkitekturen til en AI-modell vil direkte påvirke hva den er i stand til. Nevrale nettverk er mer allsidige ettersom inndataene kan være fleksible hvis hjernens størrelse (antall parametere) er tilstrekkelig. Genetiske algoritmer er bruksspesifikke på grunn av den tvungne input og output, men har absolutt sine styrker når de brukes i riktig kontekst. Det er langt flere arkitekturer å utforske, noen veldig avanserte og noen mer grunnleggende. Generelt er grunnleggende modeller raskere å kjøre og foretrekkes for veldefinerte forventninger. Avanserte modeller er dyre å trene og krever toppmoderne maskinvare for å kjøre, men produserer fantastiske resultater hvis de er konstruert effektivt. Sørg alltid for å analysere hvilken arkitektur som er egnet, da det vil avgjøre resultatene. Stockfish er en god sjakk-AI, bedre enn noe menneske i historien, mens ChatGPT fortsatt gjør ulovlige trekk gjennom spillene sine. Begge er AI, men de er trent med forskjellige intensjoner.

Funded by
the European Union

4: TRENING AV KUNSTIG INTELLIGENS

Nå som vi har dekket viktigheten av å velge riktig treningsdata og arkitektur, er det på tide å se hvordan alt henger sammen. Det siste trinnet er prosessen med å trene AI-en. Du bygde bygningen, du kjøpte møblene, og nå er det på tide å sette det i stand. I dette kapittelet vil vi diskutere flere relevante emner angående trening av AI. I sammenheng med GA-er vil vi forklare viktigheten av treningsfunksjoner og også relatere dette bredere til forsterkningslæring. Deretter vil vi se på forskjellen mellom veiledet og uveiledet læring.

4.1: Treningsfunksjoner og forsterkningslæring

Selv om det er noen forskjeller mellom disse to konseptene, har de blitt gruppert sammen siden AI-en lærer av å samhandle med et miljø. Treningsfunksjoner er fremtredende i GA, ettersom løsningene må rangeres på en eller annen måte slik at foreldrene kan velges. Det er derfor viktig å vurdere hva som er den mest effektive treningsfunksjonen.

La oss et øyeblikk late som om Google Maps brukte en AI for å finne den korteste veien. Hva ville fitnessfunksjonen være? Ville det være den korteste avstanden? Minst mulig drivstofforbruk? Minst mulig tid å komme til destinasjonen? Det er alltid avveininger – det er ganske enkelt å beregne avstanden, men hvis du vil beregne tiden du trenger, må du vurdere maksimalhastigheten på hver vei, trafikklys, veitrafikk, veisperringer osv. Dette kan ta lengre tid for systemet å score

Funded by the European Union

løsningene og bremse utviklingsprosessen. Det er imidlertid viktig å ha en passende fitnessfunksjon, da den bestemmer resultatene du får. Svarene for raskeste vei og drivstoffeffektivitet kan til og med være forskjellige.

Hvis du for eksempel trener en sjakk-AI og forteller den at antallet brikker på brettet er kondisjonsfunksjonen, kan du bli satt i sjakkmatt til tross for at du ikke har mistet noen brikker. Derfor må det vurderes nøye hvilken kondisjonsfunksjon som skal brukes.

Forsterkningslæring er på noen måter relatert til dette, ettersom det inkluderer treningsfunksjoner. Det er fremtredende i simulert læring – der AI-en er i et simulert miljø og prøver å oppnå et mål. Konseptet med forsterkningslæring er at AI-en blir belønnet for å gjøre det rette og straffet for å gjøre det gale. Hvis for eksempel AI-en er en bil i et racingspill, får den flere poeng for å nå målstreken raskere, og den kan også miste poeng for å kjøre i feil retning. På denne måten beholdes de bedre løsningene i hver neste generasjon, og de verste reproduseres ikke. Igjen bør funksjonen som brukes til å score AI-en være nøye utformet.

AI-en bør få nok rom til å finne ut av ting, siden den kan finne en bedre løsning enn du forventer, men hvis du er for streng med treningsfunksjonen, vil den finne løsningen du ønsker, som du kan finne selv. I samme eksempel fra racingspillet, hvis du gir den poeng for å kjøre tett rundt svingene, er du streng og lar den ikke eksperimentere. Det kan være slik at det å kjøre langt fra ett sving og opprettholde farten er mer effektivt i sammenheng med hele løpet.

Funded by
the European Union

4.2: Veiledet og uveiledet læring

GA-er, kondisjonsfunksjoner og forsterkningslæring er alle relaterte siden de bruker interaksjoner med et miljø for å trene AI-en ved å fortelle den hva som scores høyere som et resultat. Det andre paradigmet er ikke å fortelle AI-en om den gjør det bra, men heller bare trene den på en mengde data og se hva som skjer.

Til tross for navnet er det ingen mennesker som direkte overvåker opplæringsprosessen i veiledet læring. Forskjellen kommer fra typen data som mates til AI-en. Veiledet læring er når dataene er merket (vanligvis av mennesker). Dette betyr at noen har gitt de «riktige svarene», omtrent som å lese til en eksamen ved å bruke løsningene på tidligere eksamensspørsmål. Dette er fordelaktig når det er et klart resultat som forventes, og det produserer AI-modeller som er mer forutsigbare og mindre volatile (har mindre variasjon i nøyaktighet og konsistens).

Uovervåket læring er når dataene ikke er merket, og AI-en skanner etter mønstre i selve dataene i stedet for å koble sammen dataene og merkelappene. Dette er nyttig når forventningene er mer subjektive, som i skriftlig engelsk – det finnes mange måter å si det samme på, og ingen av dem er iboende feil. I dette tilfellet er det upraktisk å merke dataene, og det er mer effektivt å la AI-en oppdage mønstre på egenhånd.

Nå som vi vet at forskjellen mellom veiledet og uveiledet læring er typen data som gis til AI-en, la oss se på bruksområdene til begge. I avanserte modeller som LLM-er (ChatGPT) brukes begge for å gi AI-en mer kapasitet, men hva er hver av dem nyttig for?

Funded by the European Union

Skillet er viktig, ettersom begge læringstypene har spesielle trekk. Veiledet læring kan være dyrt, ettersom det krever at dataene merkes – dette er en intensiv prosess på grunn av mengden data som trengs for å trene en kompleks modell. Det har den fordelen at AI-en vil kjenne til forventet utdata, ettersom det er godt definert og tydelig strukturert. Uveiledet læring, derimot, kan bruke alle data du gir den og prøve å konstruere mønstre ut fra disse. Den er derfor mindre eksplisitt og strukturert i resultatene, men er lettere å trene, ettersom den ikke krever spesielle data. Den er god til å finne små nyanser og mønstre som ellers ikke er åpenbare.

Funded by
the European Union

5: Hvordan bruke AI effektivt

Nå som du har utviklet en forståelse for hva som kreves for å lage en AI-modell, er det på tide å se på det nåværende landskapet av eksisterende produkter og hvordan de kan brukes mest effektivt. Det er ingen hemmelighet at bruk av AI kan øke produktiviteten din, så i dette kapittelet vil vi diskutere de beste bruksområdene for LLM-er og dekke i hvilke situasjoner AI kan være kontraproduktivt. LLM-er som ChatGPT kan være nyttige i mange situasjoner, men siden dette kurset er for unge gründere, vil vi fokusere på de mest forretningsorienterte bruksområdene.

Idémyldring:

Fordeler

- Rask generering av relaterte emner
- Kjent med mange konsepter og god til å koble sammen relaterte punkter
- Kan sprette ideer med deg, på samme måte som du ville snakket med en annen person

Begrensninger

- Noen ganger forstår ikke hva du forventer, så du må være mer spesifikk
- Kun trent på tilgjengelige data, så hvis det er noe proprietært eller utenfor datasettet, har du flaks

Funded by the European Union

Skrivetekst:

Fordeler

- Skriver vanligvis mer sammenhengende enn folk flest, spesielt på sitt andrespråk
- Kjent med mange konsepter og god til å koble sammen relaterte punkter
- Kan sprette ideer med deg, på samme måte som du ville snakket med en annen person

Begrensninger

- Noen ganger er tonen ikke relevant for teksttypen
- Hvis du jobber med noen som ikke godtar AI-innhold, kan de bruke en AI-skanner.
- Det kan ta lengre tid å forklare konteksten og forventningene enn å skrive det selv.

Gi tilbakemelding:

Fordeler

- Oppdager feil du kanskje overser
- Raskere enn du leser ditt eget arbeid, spesielt for større tekster
- Kan foreslå forbedringer når kritikkpunkter er identifisert

Begrensninger

- Programmert til å gi et svar uansett hvor påtvunget det er, noe som noen ganger resulterer i triviell tilbakemelding
- Noen ganger mye mer kritisk enn nødvendig for oppgaven
- Bruker ofte flere ord enn nødvendig for budskapet den prøver å formidle, med

Funded by the European Union

mindre du instruerer den til å ikke gjøre det

- Kan forklare koden linje for linje

Koding:

Fordeler

- Rask kodegenerering
- Kjent med mange språk og rammeverk
- Kan fikse sin egen kode hvis det ikke fungerer på første forsøk

Begrensninger

- Kan ikke generere filsystemer og komplekse integrasjoner
- Kan kaste bort tid ved å generere kode du ikke forstår

Noen generelle hensyn ved bruk av AI:

- Hallusinerer noen ganger falske fakta
- Begrenset til opplæringsdata som er oppgitt
- Kjenner bare den kontekstuelle informasjonen du oppgir

LLM-er fungerer basert på å forutsi det neste ordet som skal genereres. Selv om nyere modeller har mer avanserte systemer for sitering og informasjonsformidling, bør du fortsatt dobbeltsjekke viktige fakta selv, da feil kan oppstå.

Hvis emnene du prøver å få informasjon om eller diskutere med AI-en ikke er fremtredende i treningsdataene, vil den ikke gjøre en god jobb i noen av spørsmålene relatert til disse emnene.

Hvis du for eksempel prøver å skrive en tekst til bedriftens nettsted, må du gi AI-en mye informasjon om visjonen, driften og så videre. Hvis denne informasjonen ikke allerede er i tilgjengelig tekst som du kan kopiere og lime inn, kan det ta lengre tid for deg å skrive den ned enn å skrive nettsideteksten selv. Dette gjelder spesielt når du trenger en kort tekst med veldig presis betydning. I slike tilfeller kan det være effektivt

Funded by the European Union

å skrive den første teksten og deretter spørre AI-en om den ser noe rom for forbedringer.

Funded by
the European Union

6: OPPSUMMERING AV VIKTIGE POENGER

Du har nådd slutten av modulen, gratulerer! La oss nå gå gjennom noen av de viktigste punktene, slik at du ikke glemmer de viktigste delene.

Fra kapittel 2 lærte vi hvorfor treningsdata har en betydelig innvirkning på resultatene fra en AI-modell. Vi lærte også at data av høy kvalitet bør:

- Vær nøyaktig
- Ha tilstrekkelig mengde
- Vær mangfoldig
- Vurder utilsiktede skjevheter

I kapittel 3 dekket vi deretter hvordan arkitekturtypen du valgte er relatert til typen problem som skal løses. Vi diskuterte at genetiske algoritmer er effektive for problemer med et definert resultatformat, spesielt optimaliseringsproblemer. Vi utviklet også en grunnleggende intuisjon for hvordan nevrale nettverk (NN) fungerer og hvordan det finnes mange typer NN. NN-ene er fleksible i sitt formål og funksjonalitet.

I kapittel 4 lærte vi hvordan forsterkningslæring og kondisjonsfunksjoner henger sammen, hvordan de fungerer og hva formålet deres er. Vi skilte også mellom veiledet og uveiledet læring.

Til slutt, i kapittel 5, snakket vi om hvordan man best kan bruke AI og hva man kan vurdere for å forbedre arbeidsflyten. Vi skisserte hvordan AI fungerer i oppgaver som koding, idémyldring, tekstskriving og tilbakemeldinger, samt dens styrker og svakheter i hver kategori.

Forhåpentligvis er denne forståelsen nyttig i din reise med AI, lykke til.

Funded by the European Union

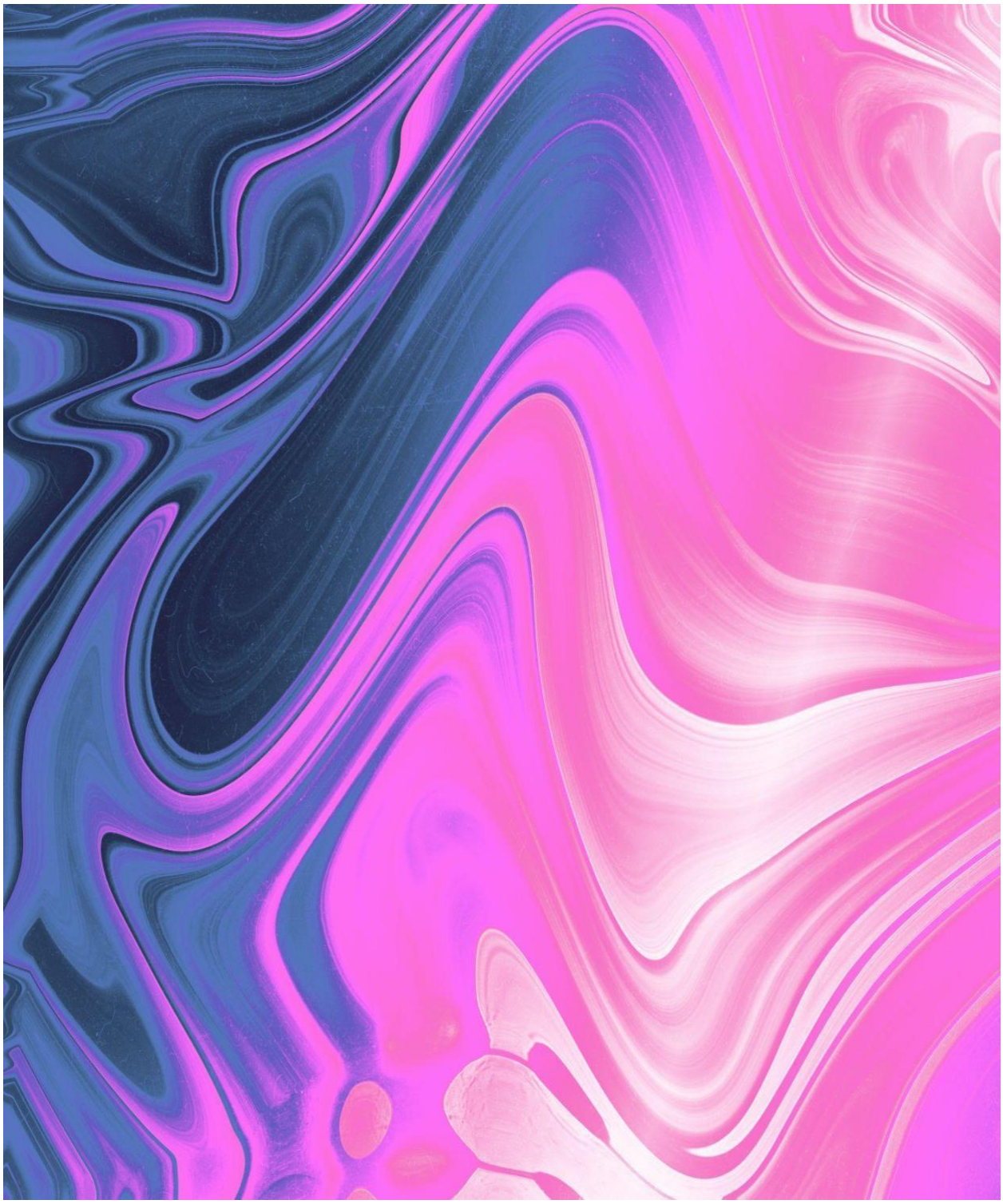
1. BIBLIOGRAFI

https://www.oracle.com/nl/artificial-intelligence/ai-model-training/#rolle-of-data

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

https://www.geeksforgeeks.org/genetic-algorithms/

Funded by the European Union



Funded by
the European Union