

Projektnummer/Project no Uppdragsgivare/Client
E38027 Försvarets materielverk
FoT-område
Inget FoT-områdeFörfattare/Author
Johan SchubertDatum/Date Memo nummer/Number
2023-12-13 FOI Memo 8386

Förslag på AI teknikutveckling

1 Introduktion

I detta memo studerar vi viktiga AI-tekniker¹ primärt för ledningssystemområdet. Dessa metoder ligger till grund för beslutsstöd² som stödjer beslutsfattande på alla nivåer. AI finns i hela kedjan från deskriptivt beslutsstöd (lägesbild), över prediktivt beslutsstöd (prognos) till preskriptivt beslutsstöd (hantering av handlingsalternativ). Dessa tre beslutsstödsprocesser kan beskrivas som om de följer efter varandra rent logiskt, men processerna pågår egentligen parallellt hela tiden. Vid en jämförelse med OODA-loopen³ så motsvarar det faserna observera, besluta och agera.

I Schubert (2017) [1] drogs slutsatsen att ”AI idag kommer starkt på alla fronter. Två exempel på olika militära nivåer är AI i beslutsstödsystem för central informationsbehandling i ett högkvarter, respektive AI i autonoma system som är distribuerade över slagfältet och kan fatta egna beslut.” Det gäller fortfarande. Sedan dess har maskininlärningsområdet (ML) fortsatt vara på start frammarsch. Enbart maskininläring är dock inte tillräckligt som metod för militärt beslutsstöd. Det behövs metoder som kan binda ihop en sub-symbolisk nivå, där flera maskininlärningsmetoder opererar, med en symbolisk nivå där system kan resonera kring vilka handlingsalternativ som är mest lämpade och presentera dessa för beslutsfattare.

För att få en AI att lösa rätt problem är det viktigt att konstruera en korrekt kunskapsrepresentation för den aktuella frågeställningen [2]. Kunskapsrepresentationen är den ram inom vilket en lösning på ett specifikt problem söks. Utan en god kunskapsrepresentation riskerar både AI-tekniker på symbolisk och sub-symbolisk nivå att hantera en felaktig problemställning. En sådan kunskapsrepresentation beskrivs alltid på symbolisk nivå. För att överföra resultat från en sub-symbolisk nivå till den symboliska nivån behövs metoder som neuro-symboliskt resonerande⁴ och förklaringsbar AI⁵, och för att kunna resonera om händelser och deras konsekvenser behöver man kunna resonera kausalt⁶.

Dessa metodiker och flera andra nödvändiga metodiker för beslutsstöd i en stab och för beslutsfattande i en autonom plattform diskuteras i förslagen om AI-områden för teknikutveckling i Kap. 2. Avslutningsvis dras i Kap. 3 några slutsatser om behovet av teknikutveckling.

¹ Med AI avses de tekniker där intelligenta agenter kan resonera om och lära sig från sin omgivning.

² Med beslutsstöd avses datorsystem som sammanställer, analyserar och resonerar utifrån information från flera källor för att presentera ett underlag över vad situationen är, kommer att bli och hur man bör agera.

³ OODA-loopen refererar till beslutcykeln av att observera, orientera, besluta och agera (eng. observe, orient, decide, act).

⁴ Neuro-symbolic AI. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Neuro-symbolic_AI (accessed 13 Dec. 2023).

⁵ Explainable artificial intelligence. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Explainable_artificial_intelligence (accessed 13 Dec. 2023).

⁶ Causal AI. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Causal_AI (accessed 13 Dec. 2023).

Titel/Title
Förslag på AI teknikutveckling

Memo nummer/Number
FOI Memo 8386

Memot har tagits fram och arbetet har genomförts på uppdrag av FMV.

Titel/Title
Förslag på AI teknikutvecklingMemo nummer/Number
FOI Memo 8386

2 Teknikområden

I detta kapitel redovisas tio forskningsområden inom AI med olika utvecklingsläge där alla är av betydelse för utvecklingen inom ledningssystemteknik, och mestadels i behov av omfattande teknikutveckling. Varje område har en beskrivning av områdets huvuddrag, en kort diskussion om möjliga militära tillämpningar och ett diagram som visar på trend och volym i forskningen.

2.1 Kausal AI

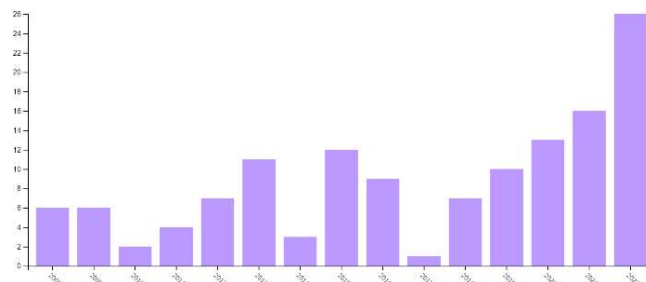
Kausal AI (eng. Causal AI) är ett framväxande område av artificiell intelligens med metoder som fokuserar på att lära sig orsak-och-verkan-samband inom data som kan förklara orsak och verkan [3]. Detta skiljer sig från de maskininlärningsmodeller som bygger på korrelationsbaserade modeller. Metoderna används för att förklara underlag för beslutsfattande och orsakerna till rekommendation till beslut. Kausal AI erbjuder bättre transparens och förklararbarhet för AI (XAI) och hjälper till att förstå resultaten som produceras av algoritmer. Områdets metoder överbryggas klyftan mellan prognos och beslutsfattande och gör det möjligt att variera indata för att dra slutsatser om orsakssamband genom att använda redan tillgängliga data.

Kausala grafmodeller (eng. Causal Graph Models) hjälper till att kartlägga de olika stegen i ett kausalt resonemang fram till ett intressant resultat och visar hur olika variabler förhåller sig till varandra. En vanlig metod är strukturell ekvationsmodellering (eng. Structural Equation Model) [4], där man specificerar de variabler som kan interagera, hur de kan göra det, och låter modellen analysera data för att avslöja om de faktiskt gör det.

Strukturell ekvationsmodellering involverar en modell som representerar hur olika aspekter av en händelse tros ha kausalt samband med varandra. Modellens begränsningar är att den endast testar kopplingar mellan de specificerade variablerna. En annan kausal grafmetod är kausala Bayesianska nätverk (eng. Causal Bayesian Networks) [5], som uppskattar sambanden mellan alla variabler i en datamängd. De består av en riktad acyklisk graf som representerar orsakssamband och en uppsättning sannolikhetsstabeller som anger den sammansatta sannolikheten för de variabler som representeras som noder i grafen.

Om man vill tillåta användning av maskininläring i militära system som normalt bara identifierar samband mellan händelser och konsekvenser måste man känna till alla möjliga orsakssamband i förväg. Detta kräver en hög förståelse för alla möjliga händelser och konsekvenser. Om denna kunskap inte finns tillgänglig är det istället nödvändigt att använda metoder för att lära sig orsakssamband [6]. Exempel på militära tillämpningar med kausal AI finns vid prediktion av motståndarens beteende där kausal AI ligger till grund för att både förutspå och att förklara en händelseutveckling vid beslutsstöd [7].

Kausal AI har funnit som ett aktivt forskningsområde på låg nivå under många år. De senaste två åren har det sett ett uppsving pga. behovet av att gå bortom modeller som endast bygger på korrelation, figur 1.



Figur 1. Antalet forskningspublikationer per år inom forskningsområdet kausal AI.

Titel/Title
Förslag på AI teknikutvecklingMemo nummer/Number
FOI Memo 8386

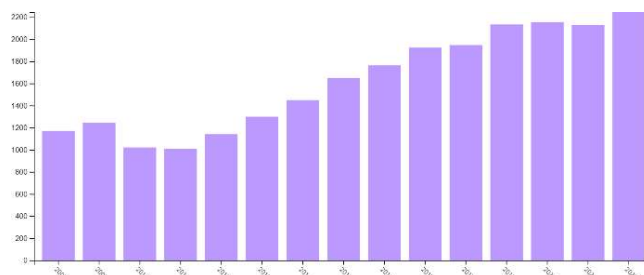
2.2 Multiagentsystem

Ett multiagentsystem (eng. Multi-agent systems) är ett datoriserat system som består av flera interagerande intelligenta agenter [8] och deras miljö. Multiagentsystem kan lösa problem som är svåra eller omöjliga för en enskild agent att lösa. Vanligtvis syftar forskning om flera agenter på programvaruagenter. Det är en aktuell utvecklingstrend mot små modulära bitar programkod där varje modul utför en väldefinierad, fokuserad uppgift eller uppsättning uppgifter. Men agenterna i ett multiagentsystem kan lika gärna vara robotar, människor eller team av människor. Ett multiagentsystem kan även innehålla kombinerade människa-maskin team.

Agenter kan antingen vara passiva, aktiva eller kognitiva [9]. Passiva agenter är agenter utan mål (som hinder), aktiva agenter har däremot enkla mål (som fåglar i en flock) och kognitiva agenter är agenter med beteenden som bygger på komplexa beräkningar.

Militära tillämpningar finns t.ex. inom koordinering av autonoma farkoster i alla domäner [10]. Vid U.S. Army DEVCOM Army Research Laboratory⁷ har man distribuerade algoritmer för koordinering av flera agenter och tillhandahåller ett ramverk för informationsinsamling i scenarier med begränsad bandbredd för att förbättra situationsmedvetenheten på slagfältet. Syftet är att utveckla tekniker som skulle tillåta flera UAV:er att samordna sig så att de tillsammans ger bästa följning av ett manövrerande mål.

Forskningsområdet är etablerat sedan lång tid med en stadig tillväxt över de senast tolv åren, figur 2.



Figur 2. Antalet forskningspublikationer per år inom forskningsområdet multiagentsystem.

2.3 Neuro-symboliskt resonerande

Neuro-symboliskt resonerande (NSR) (eng. Neuro-symbolic reasoning) är ett framväxande forskningsfält som syftar till att kombinera styrkan hos symboliskt resonemang och djupinlärning [11]. Det har fått stor uppmärksamhet de senaste åren på grund av dess potential att lösa komplexa problem som ligger utanför räckvidden för traditionell maskininlärning.

Neuro-symbolisk AI är ett allt mer aktivt forskningsfält som beskriver kombinationen av symbolisk AI, som ofta inkluderar logik och regelbaserade tillvägagångssätt och probabilistisk resonerande, med neurala nätverk och djupinlärning. De främsta fördelarna med neuro-symbolisk AI beskrivs som att uppnå jämförbar prestanda med nuvarande neuronnätsbaserade metoder samtidigt som de möjliggör tolkning av resultaten [12]. En avskanning av forskningsfronten inom området vid FOI [13] visar att:

- NSR tros kunna bidra till att lösa kritiska utmaningar såsom förklarbarhet.
- NSR är ett relevant teknikområde för försvar. Det finns många olika tillämpningsområden och en relativt stor del av forskningen är finansierad av internationella försvarsorganisationer.

Tillämpningsområden av NSR för försvar, finns inom klassificering, resonerande AI, beslutsstöd, robotik, autonomi, inlärning med hjälp av expert, simulering samt textanalys. Andra militära

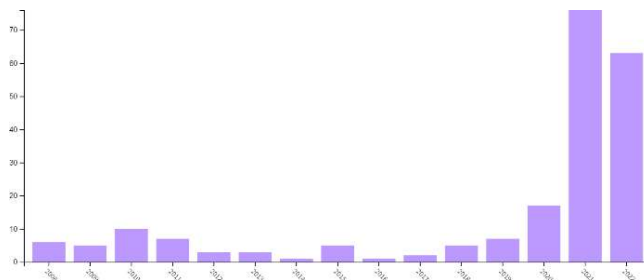
⁷ DEVCOM Army Research Laboratory – The Army's foundational research laboratory. Army.mil. [Online]. Available: <https://arl.devcom.army.mil/> (accessed 13 Dec. 2023).

Titel/Title
Förslag på AI teknikutveckling

Memo nummer/Number
FOI Memo 8386

tillämpningar finns inom avancerad målföljning i mark- och sjöstrid. NSR kan även användas tillsammans med XAI för att förklara text och bilder inom underrättelseanalys. På högre konceptuell nivå kan NSR användas för att förstå situationer och händelser som visar sig i videodata.

Neuro-symbolisk AI har en lång historia på låg nivå, men har fått ett stort uppsving de senaste tre åren pga. behovet av att kombinera maskininlärningsmetoder med symboliskt och probabilistiskt resonerade, figur 3.



Figur 3. Antalet forskningspublikationer per år inom forskningsområdet neuro-symboliskt resonande.

2.4 Prediktiv analys

Prediktiv analys (eng. Predictive analytics) använder statistiska metoder som datamodellering, maskininläring, AI och djupinläring för att analysera aktuella fakta och göra förutsägelser om framtida eller okända händelser. Det utgör en stor delmängd av maskininläring och används ofta vid bedömningar av händelser. Prediktiva modeller utnyttjar de mönster som finns i data för att identifiera risker. De hittar relationer mellan olika faktorer för att möjliggöra en bedömning av de risker eller den potential som är förknippade med vissa situationer inför ett beslutsfattande.

Prediktiv analys omfattar teknologier som avslöjar relationer och mönster inom stora mängder data och kan användas för att förutsäga beteende och händelser. Kärnan i prediktiv analys bygger på att fånga relationer mellan förklarande variabler och de förutspådda variablerna utifrån tidigare händelser, och att utnyttja dem för att förutsäga ett nytt okänt resultat [14].

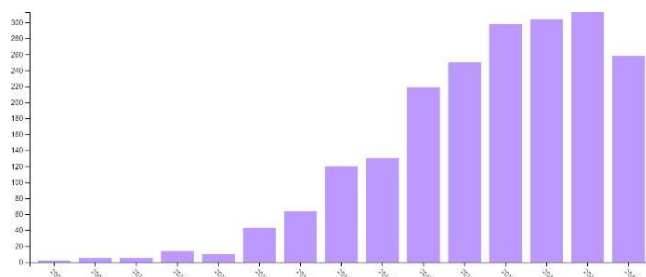
I militära scenarier kan prediktiv analys användas för att förutsäga stridsberedskap, optimera logistik och förbättra beslutsfattande [15]. Användningen av prediktiv analys i försvaret har dock vissa utmaningar:

- Prediktiv analys förlitar sig på stora mängder data av god kvalitet för att kunna göra korrekta förutsägelser. Militär data är dock ofta ofullständig, inkonsekvent eller föråldrad, vilket riskerar att leda till felaktiga förutsägelser.
- Modeller för prediktiv analys är bara så bra som det data de tränas på. Om data är partisk eller ofullständig kommer modellen att göra felaktiga förutsägelser.
- Prediktiv analys kräver specialiserad kunskap. Om sådan saknas kan det leda till en bristande förståelse för hur man använder prediktiv analys på ett effektivt sätt.

De tillvägagångssätt som används för att utföra prediktiv analys kan i stort sett grupperas i regressionsanalys⁸ och maskininläring. Maskininläring ger fördelen med att anpassa sig snabbt och med hög effektivitet.

Forskningsområdet har sett en betydande tillväxt under de senaste elva åren, figur 4.

⁸ Regression analysis. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Regression_analysis (accessed 13 Dec. 2023).

Titel/Title
Förslag på AI teknikutvecklingMemo nummer/Number
FOI Memo 8386

Figur 4. Antalet forskningspublikationer per år inom forskningsområdet prediktiv analys.

2.5 Generativ AI

Generativ AI (eng. Generative AI) är en teknik som använder maskininlärning för att skapa nytt innehåll, som text, bilder, video, ljud och programvara [16]. I militära scenarier kan generativ AI användas för att öka situationsmedvetenheten och ge bättre beslutsfattande genom att analysera data från flera källor och identifiera mönster [17].

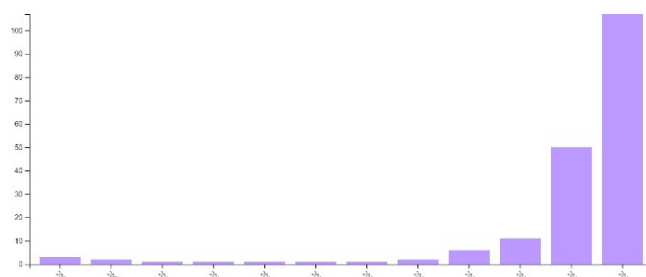
Generativ AI kan användas för att utöka mängden träningsdata för att träna maskininlärningsmodeller som används i olika militära applikationer som objektigenkänning, målidentifiering och för autonoma system. I försvaret kan generativ AI förbättra situationsmedvetenheten genom att analysera data från flera källor och identifiera mönster. Metodiken kan även proaktivt identifiera cybersårbarheter i programvara och infrastruktur för att hjälpa till att stärka cybersäkerhetsskyddet.

Generativ AI kan påskynda utvecklingen av militär utrustning och minska kostnaderna genom att effektivisera utvecklingscyklerna. Funktioner som snabb prototypframställning, skapandet av simuleringar för komplexa system och design av virtuella prototyper hjälper till att påskynda utrustningsutvecklingen.

Simuleringar som drivs av generativ AI kan ge realistiska träningsmiljöer för militär personal. T.ex. kan generativ AI användas för att skapa realistiska träningsmiljöer för soldater, så att de kan öva i många olika scenarier [18].

Generativ AI kan också generera olika scenarier för att förutsäga potentiella resultat under planering, krigsspel och beslutsfattande. Modellerna kan här användas för att simulera olika militära scenarier. Dessa simuleringar hjälper planerare att undersöka möjliga resultat och identifiera optimala strategier. Metoderna kan också generera alternativa handlingsätt för militära operationer baserat på tillgängliga data och historiska mönster. Det kan ske genom att kombinera analysen av historiska data med aktuell underrättelseinformation för att göra prediktioner om motståndarens handlingar, vilket hjälper planerare att förutse alternativa händelseutvecklingar [19].

Generativ AI är ett nytt forskningsområde med kraftig tillväxt under de senaste fyra åren, som i forskningslitteraturen till 90% handlar om stora språkmodeller, figur 5.



Figur 5. Antalet forskningspublikationer per år inom forskningsområdet generativ AI.

Titel/Title
Förslag på AI teknikutveckling

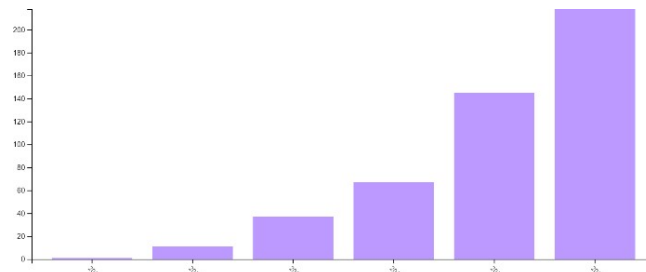
Memo nummer/Number
FOI Memo 8386

2.6 Edge AI

Edge AI syftar på kombinationen av edge computing och AI för att utföra maskininlärningsuppgifter direkt på sammankopplade edge-enheter som sensorer eller Internet of Things-enheter. Edge computing gör att data kan lagras nära enhetens plats, och AI-algoritmer gör att data kan bearbetas direkt på nätverksskanten. Detta underlättar bearbetningen av data inom millisekunder och ger feedback i realtid [20].

Edge AI har ett brett utbud av möjliga tillämpningar i militära scenarier, inklusive autonoma plattformar, drönare och robotar som kan arbeta i avlägsna eller fientliga miljöer, och intelligenta sensorer som kan upptäcka och svara på hot i realtid. Speciellt kan edge AI förbättra situationsmedvetenhet, beslutsfattande och uppdragseffektivitet för flygvapnet [7].

Forskningsområdet är nytt med en första forskningspublikation 2016 och därefter en snabbt stigande trend de senaste fem åren, figur 6.



Figur 6. Antalet forskningspublikationer per år inom forskningsområdet edge AI .

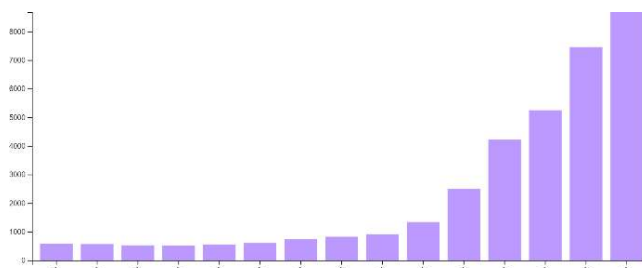
2.7 Förstärkningsinläring

Förstärkningsinläring (eng. Reinforcement learning) är ett område inom maskininläring som handlar om hur en intelligent agent bör vidta lämpliga åtgärder i en dynamisk miljö för att maximera belöningen [21] i en viss situation. I förstärkningsinläring bestämmer agenten själv vad som ska göras för att utföra den givna uppgiften. Fokus ligger på att hitta en balans mellan utforskning av okänt område och utnyttjande av aktuell kunskap.

Djup förstärkningsinläring har visat bred tillämpning inom militär underrättelsetjänst. Beslutsfattande i luftkrigföring, uppdragstilldelning och militärt schack⁹ (eng. Military chess deduction) är typiska militära tillämpningar för förstärkningsinläring [22,23]. Inom flygkrigföring kan man använda metoder för djup förstärkningsinläring med en ”trial-and-error”-metod för att interagera med omgivningen och förbättra dess beslutsfattande genom kontinuerligt lärande av taktik för flygstrid. Den genererar en serie beslutssekvenser för luftstrider genom självträning. Metoden kan till exempel användas för att träna drönare att flyga autonomt och fatta beslut utifrån den aktuella miljön. I uppdragstilldelning bygger djup förstärkningsinläring vanligtvis på intelligenta system och interagerar med komplexa militära uppgiftstilldelningsalgoritmer. Militärt schack är simuleringen av militära operationer av två eller flera motsatta sidor. Djup förstärkningsinläring kan nå hög beslutsförmåga på grund av dess flexibilitet när det gäller att inhämta kunskap från stora mängder data.

Forskningsområdet har en relativt lång historik med en snabbt växande trend under de senaste sju åren, figur 7.

⁹ Militärt schack är ett samlingsbegrepp för simulering av militära operationer med två eller flera motsatta sidor.

Titel/Title
Förslag på AI teknikutvecklingMemo nummer/Number
FOI Memo 8386

Figur 7. Antalet forskningspublikationer per år inom forskningsområdet förstärkningsinlärning.

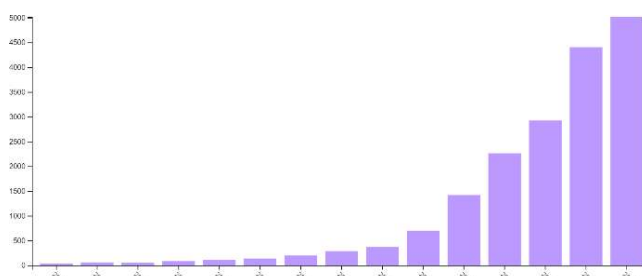
2.8 Överföringsinlärning

Överföringsinlärning (eng. Transfer learning) är en maskininlärningsmetodik där en modell utvecklad för en viss uppgift återanvänds som utgångspunkt för en modell för en annan uppgift. Det tillåter oss att överföra kunskap från en uppgift för att förbättra prestanda på en närliggande relaterad uppgift. Metodiken används tillsammans med djupinlärning där förtränade modeller används som utgångspunkt.

Överföringsinlärning kan hanteras genom att man tar ett förinlärt neuralt nätverk och anpassar det neurala nätverket till en ny, annorlunda datamängd. Efter det att inlärning gjorts med den initiala mängden data sker en avslutande inlärning med en mindre mängd skarpa data från den nya datamängden. Detta tillvägagångssätt är särskilt användbart när den andra uppgiften liknar den första uppgiften eller när tillgången på data är begränsad för den andra uppgiften [24,25,26]. Praktiskt så kan vikterna för vissa lager i ett neuronät fixeras efter den första inlärningen och sedan lärs de återstående lagren in på nytt vid den avslutande inlärningen [27]. Ett alternativt sätt att hantera bristen på data för maskininlärning är att generera den initiala mängden data genom simulering innan en slutlig inlärning görs med den mindre mängden skarpa data.

Metodiken för överföringsinlärning är generisk och militära tillämpningar finns inom flera områden. Ett exempel är inom bildklassificering där bristen på träningsdata kan vara en begränsande faktor.

Överföringsinlärning är ett relativt nytt forskningsområde med en kraftig tillväxt under de senaste sju åren, figur 8.



Figur 8. Antalet forskningspublikationer per år inom forskningsområdet överföringsinlärning.

2.9 Övervakat lärande

Oövervakad inlärning (eng. Unsupervised learning) är en typ av maskininlärning där algoritmer lär sig mönster uteslutande från omärkta data. Dessa algoritmer hittar mönster och samband inom data utan förkunskaper om datastrukturen [28]. Två av de vanligaste algoritmerna som används vid oövervakad inlärning är klustring och upptäckt av anomalier.

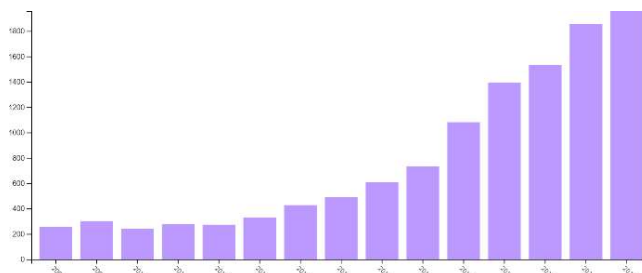
Oövervakat lärande bortom klustring och upptäckt av anomalier är ännu underutvecklade forskningsområden. Det finns ett behov av att utveckla nya metoder som fungerar på ett oövervakat

Titel/Title
Förslag på AI teknikutveckling

Memo nummer/Number
FOI Memo 8386

sätt med data som samlas in av agenter som verkar i en okänd miljö [29]. På lång sikt blir oövervakad inlärning ett av de viktigaste områdena inom maskininlärning.

Oövervakad inlärning har en lång historik inom delområdena klustring och anomalidetektion men är i övrigt ett relativt nytt område. En betydande tillväxt har skett under de senaste tio åren, figur 9.



Figur 9. Antalet forskningspublikationer per år inom forskningsområdet oövervakad inlärning.

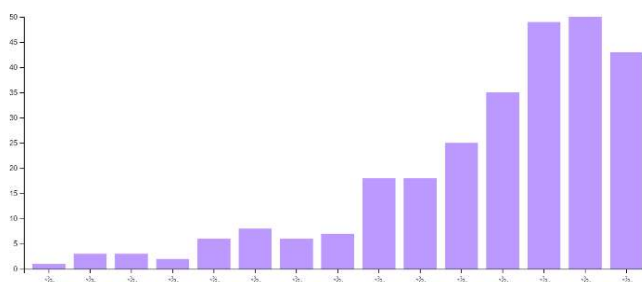
2.10 Principen om fri energi

Principen om fri energi (eng. Free energy principle) är en informationsteoretisk metodik om beteenden hos system som upprätthåller en separation från sin omgivning. Metodik kan användas för att beskriva en autonom agents val av aktioner. Det sker genom att agenten minskar sin osäkerhet genom att göra förutsägelser baserade på interna modeller och uppdatera dem med hjälp av sensorer [30].

Principen om den fria energin ger en beskrivning av en agents aktioner. Principen omfattar två parallella och samverkande processer. *Perceptionsinferens* är en informationsfusionsprocess hos agenten som uppdaterar en intern probabilistisk modell vilken tar hänsyn till sensorers indata, medan *aktiv inferens* låter agenten genomföra aktioner som förändrar omvärlden på så sätt att sensorernas indata därefter bättre överensstämmer med förutsägelser från samma interna modell, dvs. agenten genomför de aktioner som minimerar den förväntade fria energin [31].

Militära tillämpningar finns bl.a. inom team sammansatta av människor och autonoma system som tillsammans löser uppgifter, styrning av system för informationsinsamling eller insatser och som en intelligent komponent inom ramen för ett ledningssystem som kan generera externa aktioner.

Forskningsområdet har sitt ursprung inom psykologi inom vilket det finns en lång historik [32]. Inom ramen för AI är det dock ett ungt område som uppvisar en kraftig tillväxt sedan 2016, figur 10.



Figur 10. Antalet forskningspublikationer per år inom forskningsområdet principen om fri energi.

Titel/Title
Förslag på AI teknikutvecklingMemo nummer/Number
FOI Memo 8386

3 Slutsatser

Detta memo beskriver tio delområden inom AI som alla är viktiga teknikutvecklingsområden. De tio områdena för teknikutveckling har inbördes en del likheter varför vi grupperar dem i tre grupper AI, ML respektive Övriga. Grupp AI beskriver tre områden för att få maskininläring att fungera tillsammans med AI på symbolisk nivå för beslutsstöd. Grupp ML beskriver tre områden vilka kan ge användaren asymmetriska fördelar gentemot en motståndare som inte har samma teknik. Grupp Övriga grupperar fyra olika tekniker i skilda utvecklingsfaser.

3.1 Grupp AI

Tre områden som är viktiga för att få maskininläring att fungera tillsammans med resonerande-processer för beslutsstöd är (i) neuro-symboliskt resonerande (Kap. 2.3), (ii) kausal AI (Kap. 2.1) och (iii) överföringsinläring (Kap. 2.8). Den första kopplar ihop den sub-symboliska nivån med den symboliska där beslutsfattande sker. Detta är även relevant för motsvarande delar av förklaringsbar AI (XAI). Metodiken möjliggör att maskininlärningsresultat integreras på en symbolisk nivå som underlag för beslutsfattande. Den andra tar steget bort från de vanliga korrelationsmodellerna i maskininläring till att inkludera ett kausalt resonerande med orsak och verkan. Den tredje hanterar problemet med bristande tillgång på data i militära scenarier genom att införa möjligheten att förträna med större närliggande datamängder när sådana finns tillgängliga eller när sådana kan framställas genom simulering. Dessa metoder syftar till att få maskininläring och annan AI att fungera tillsammans.

3.2 Grupp ML

Tre metodiker, (iv) principen om fri energi (Kap. 2.10), (v) förstärkningsinläring (Kap. 2.7) och (vi) multiagentsystem (Kap. 2.2) handlar om metoder för att få självständiga system och interagerande AI-komponenter att samarbeta. Förstärkningsinläring lär en agent ett beteende genom att optimera en signal. Metodiken kan användas för att lära in taktiska regler i situationer där det är möjligt att optimera med avseende på en enskild belöning. Principen om fri energi går ett steg längre genom att låta en agent sampla tillstånd i omvärlden i alla riktningar och kombinera traditionell perceptionsinferens med en ny idé om aktiv inferens, där aktiv inferens utgör grunden för ett eget agerande gentemot omvärlden. På detta sätt kan agenten både uppdatera sin egen modell så att den överensstämmer med omvärlden och förändra omvärlden så att den överensstämmer med den egna modellen. Det gör agenter baserade på principen om fri energi robusta. Tekniker för multiagentsystem tillåter flera agenter att samarbeta. Dessa metoder kan ge asymmetriska fördelar gentemot en motståndare som släpar efter med teknikutvecklingen inom AI.

3.3 Grupp Övriga

Bland övriga tekniker finns (vii) prediktiv analys (Kap. 2.4) och (viii) edge AI (Kap. 2.6). Den första behandlar stora datamängder och mönster i dessa. Metodiken som är relativt mogen kan exempelvis användas för prediktivt beslutsstöd. Edge AI är kombinationen av edge computing med AI. Tillämpningar finns för plattformar nära fronten där beräkningar måste vara mycket snabba och metoderna är något enklare än centralt baserade metoder.

På längre sikt kan det vara intressant med teknikutveckling om (ix) militärt inriktad generativ AI (Kap. 2.5) bortom genererande av text, bild, video och ljud och (x) oövervakat lärande (Kap. 2.9) annat än för klustring och anomalidetektion. Dessa specifika inriktningar av generativ AI och oövervakad inläring är dock mindre mogna än tekniker i grupperna AL respektive ML.

Titel/Title
Förslag på AI teknikutveckling

Memo nummer/Number
FOI Memo 8386

3.4 Rekommendation

Prioriterad teknikutveckling bör främst ske inom grupperna AI respektive ML: Teknikutveckling inom grupp AI för att integrera maskininlärning med AI så att maskininlärning blir till nytta i beslutsstöd, teknikutveckling inom grupp ML för att nå asymmetriska fördelar gentemot en motståndare.

4 Referenser

- [1] Schubert, J. (2017). Artificiell intelligens för militärt beslutsstöd. Swedish Defence Research Agency, Stockholm, Sweden, Tech. Rep. FOI-R--4552--SE.
- [2] Brynielsson, J., Nilsson, M., Schubert, J. and Svenmarck, P. (2018). Artificiell intelligens för beslutsstöd i ledningssystem. Swedish Defence Research Agency, Stockholm, Sweden, Tech. Rep. FOI-R--4678--SE.
- [3] Casual AI. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Causal_AI (accessed 13 Dec. 2023).
- [4] Structural equation modeling. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Structural_equation_modeling (accessed 13 Dec. 2023).
- [5] Bayesian network. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Bayesian_network (accessed 13 Dec. 2023).
- [6] Schubert, J. and Johansson, R. (2020). Learning causal structures from data, Swedish Defence Research Agency, Stockholm, Sweden, Tech. Rep. FOI-R--4882--SE.
- [7] Morgan, F. E., Boudreaux, B., Lohn, A. J., Ashby, M., Curriden, C., Klima, K. and Grossman, D. (2020). Military applications of artificial intelligence: Ethical concerns in an uncertain world. Santa Monica, CA: RAND Corporation.
- [8] Multi-agent system. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Multi-agent_system (accessed 13 Dec. 2023).
- [9] Kubera, Y., Mathieu, P. and Picault, S. (2010). Everything can be Agent!, in *Proc. Ninth Int. J. Conf. Auton. Agents and Multi-Agent Syst.*, 10–14 May 2010, Toronto, Canada, pp. 1547–1548.
- [10] Army researchers enhance communications for multi-agent teaming. Army.mil. [Online]. Available: https://www.army.mil/article/236261/army_researchers_enhance_communications_for_multi_agent_teaming (accessed 13 Dec. 2023).
- [11] Neuro-symbolic AI. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Neuro-symbolic_AI (accessed 13 Dec. 2023).
- [12] d’Avila Garcez, A. and Lamb, L. C. (2020). Neurosymbolic AI: the 3rd wave, in arXiv preprint arXiv:2012.05876.
- [13] Lindberg, A., Doring, C., Johansson, R., Schubert, J., Kindvall, G. and Dalberg, E. (2022). Avskanning av AI området neuro-symboliskt resonerande – Resultat och kortfattat om metod, Swedish Defence Research Agency, Stockholm, Sweden, Tech. Rep. FOI-R--5224--SE.
- [14] Predictive analytics. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Predictive_analytics (accessed 13 Dec. 2023).
- [15] Abadicio, M. (2019). Predictive analytics in the military – Current applications, Emerj.com. [Online]. Available: <https://emerj.com/ai-sector-overviews/predictive-analytics-in-the-military-current-applications/> (accessed 13 Dec. 2023).
- [16] Generative artificial intelligence. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Generative_artificial_intelligence (accessed 13 Dec. 2023).
- [17] Limbasiya, J. (2023). Generative AI’s potential as a force multiplier in defense. Cio.com. [Online]. Available: <https://www.cio.com/article/651428/generative-ais-potential-as-a-force-multiplier-in-defense.html> (accessed 13 Dec. 2023).

Titel/Title
Förslag på AI teknikutveckling

Memo nummer/Number
FOI Memo 8386

- [18] Frąckiewicz, M. (2023). Generative AI: The new frontier for the US military. Ts2.space. [Online]. Available: <https://ts2.space/en/generative-ai-the-new-frontier-for-the-us-military/> (accessed 13 Dec. 2023).
- [19] Nair, J. N. (2023). Generative AI for military and aerospace applications. Dras.in. [Online]. Available: <https://dras.in/generative-ai-for-military-and-aerospace-applications/> (accessed 13 Dec. 2023).
- [20] Edge computing. https://en.wikipedia.org/wiki/Edge_computing (accessed 13 Dec. 2023).
- [21] Reinforcement learning. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Reinforcement_learning (accessed 13 Dec. 2023).
- [22] Wang, N., Li, Z., Liang, X., Hou, Y. and Yang, A. (2023). A Review of Deep Reinforcement Learning Methods and Military Application Research, *Math. Probl. Eng.* **2023**, Article ID 7678382. doi:10.1155/2023/7678382.
- [23] Wang, H., Tang, H., Hao, J., Hao, X., Fu, Y. and Ma, Y. (2020). Large Scale Deep Reinforcement Learning in War-games, in *Proc. 2020 IEEE Int. Conf. Bioinform. Biomed.*, 16–19 2020, Seoul, Korea, pp. 1693–1699. doi:10.1109/BIBM49941.2020.9313387.
- [24] Transfer learning. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Transfer_learning (accessed 13 Dec. 2023).
- [25] Transfer learning. Deepai.org. [Online]. Available: <https://deepai.org/machine-learning-glossary-and-terms/transfer-learning> (accessed 13 Dec. 2023).
- [26] Donges, N. (2022). What Is Transfer Learning? Exploring the Popular Deep Learning Approach. BuiltIn.com. [Online]. Available: <https://builtin.com/data-science/transfer-learning> (accessed 13 Dec. 2023).
- [27] Yang, Z., Yu, W., Liang, P., Guo, H., Xia, L., Zhang, F., Ma, Y. and Ma, J. (2019). Deep transfer learning for military object recognition under small training set condition, *Neural Comput. Appl.* **31**(10):6469–6478. doi:10.1007/s00521-018-3468-3.
- [28] Unsupervised learning. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Unsupervised_learning (accessed 13 Dec. 2023).
- [29] Fossaceca, J. M. and Young, S. H. (2018). Artificial intelligence and machine learning for future army applications, in *Proc. SPIE 10635 Ground/Air Multisens. Interoper., Integr., Netw. Persistent ISR IX*, paper 1063507. doi:10.1117/12.2307753.
- [30] Free energy principle. Wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Free_energy_principle (accessed 13 Dec. 2023).
- [31] Schubert, J. (2021). Aktiv inferens, Swedish Defence Research Agency, Stockholm, Sweden, Tech. Rep. FOI-R--5221--SE.
- [32] Friston, K. J. (2009). The free-energy principle: A rough guide to the brain?, *Trends in Cogn. Sci.* **13**:293–301. doi:10.1016/j.tics.2009.04.005.