

Kommentar till ämnesplanen i artificiell intelligens på gymnasial nivå

Om ämnet

Artificiell intelligens (AI) är en högaktuell teknik som i allt större utsträckning påverkar oss och vårt samhälle. Vikten av källkritisk granskning ökar när informationsflödet både kan styras av AI och ges ett AI-genererat innehåll. Sociala medier, handel, infrastruktur och kontakter med samhällsinstitutioner är exempel på områden där AI på olika sätt används och därmed påverkar våra liv. Kunskaper om AI är viktiga för att förstå den tekniska utvecklingen, dra nytta av fördelarna med AI och bli medveten om riskerna. Att lära sig om AI har betydelse för fortsatta studier, för kommande arbetsliv och för ett liv som samhällsmedborgare.

Ämnet artificiell intelligens är till sin karaktär tvärvetenskapligt. Innehållet i ämnet kan delas in i tre delar; en teknisk, en samhällsvetenskaplig och en tillämpad del. Undervisningen ska behandla AI och dess funktionssätt, primära beståndsdelar och uppbyggnad. Undervisningen ska också behandla AI-utvecklingen ur ett samhällsperspektiv och utforska vilka konsekvenser AI kan ha för det dagliga livet. Dessutom ska eleverna få använda AI i problemlösning tillämpad på deras program, intresse eller andra önskemål.

Artificiell intelligens ingår i programfördjupningen på de flesta nationella program. Eleverna kan också välja ämnet som individuellt val. Artificiell intelligens kan alltså läsas av alla elever oavsett programval. Ämnet kan med fördel undervisas av flera lärare med olika kunskaper, så att de kompletterar varandra. Beroende på när ämnet erbjuds under studietiden, och elevernas förkunskaper i matematik, kan undervisningen behöva ta upp vissa matematiska begrepp. Det centrala innehållet kan till stora delar anpassas utifrån elevernas program, intressen och förkunskaper. Ämnesplanen är skriven för att olika typer av produkter och tjänster ska kunna användas och är inte bunden till specifika verktyg.

Syfte

Artificiell intelligens är i hög grad en socioteknisk innovation. Den tekniska utvecklingen av AI flätas samman med samhället på olika nivåer. Detta samspel behöver därför ständigt uppmärksammas och vara en återkommande reflektion i

undervisningen. Utvecklingen av AI innebär både möjligheter och risker. Genom att stärka elevernas kunskaper om AI, vad som styr utvecklingen och hur tekniken tillämpas medverkar undervisning i AI till att balansera farhågor och rusta eleverna till kloka avvägningar och val.

Ämnets utformning möjliggör att mer eller mindre programmering används i undervisningen, för att konkretisera algoritmer, utforska olika tekniker för AI eller för att lösa problem. Urval, presentation och tillämpning av algoritmer kan anpassas utifrån elevernas matematikkunskaper. De elever som har kunskaper i kodning kan använda och fördjupa dessa i ämnet. För elever som saknar kunskaper och vana i textbaserad programmering kan blockprogrammering vara ett alternativ. Undervisningen kan också utgå ifrån färdiga kodexempel där valda delar kompletteras eller modifieras. Programmering kan ske i onlinemiljöer utan krav på att programvaror installeras lokalt.

Kopplingen mellan tekniken och påverkan på individ och samhälle är en viktig del i undervisningen. Den tillämpade delen i ämnet kan utgå ifrån elevernas program eller intressen och anpassas till deras kunskaper i programmering och andra ämneskunskaper.

De delar ur syftestexten som ska vara underlag för lärarens bedömning återfinns i de fem mål som avslutar denna del av ämnesplanen.

Centralt innehåll

Det centrala innehållet i ämnet återfinns under tre rubriker; *artificiell intelligens*, *praktisk användning av AI* och *förhållningssätt till AI*. Rubrikerna har framför allt kommit till för att underlätta läsningen av ämnesplanen; de är inte avsedda för att dela in undervisningen i separata delar. Innehållet går många gånger i vartannat och löper parallellt genom hela nivån. Kommentarer till det centrala innehållet följer under dessa rubriker och syftar till att utveckla innehållet och ge exempel på vad som kan tas upp. De kursiverade raderna är centrala innehållspunkter i ämnesplanen.

I det centrala innehållet förekommer uttrycket *till exempel*. De exempel som ges är valda för att ge en anvisning om innehållets svårighetsgrad, omfattning eller avgränsning. Exempelen är inte obligatoriska att behandla i undervisningen, utan kan ersättas av annat likvärdigt innehåll. I det centrala innehållet förekommer också uttrycket *däribland*. Det mer detaljerade innehåll som nämns är då obligatoriskt att behandla i undervisningen.

Nivå 1

Artificiell intelligens

Under denna rubrik finns sex innehållspunkter som ska ge möjlighet att ge undervisning om vad AI är, hur den används och vad den kan användas till.

- *Definition av AI samt centrala begrepp inom området.*

Det finns ingen gemensam och enhällig definition av AI. Undervisningen kan därför ta upp några olika definitioner av artificiell intelligens, från såväl nationella som internationella instanser och organisationer.

Artificiell intelligens omfattar ett stort område. Det behöver också avspeglas i undervisningen så att den inte begränsas till att hantera en enskild form av AI, såsom generativ AI eller mänskliga robotar.

Även om AI har en kort historik jämfört med mycket annan teknik, kan ett historiskt perspektiv om hur AI vuxit fram ändå vara angeläget att lyfta. Här kan också finnas plats för historiska definitioner av AI, till exempel Turingtestet och expertsystem, och att lyfta historiska och aktuella händelser som fått genomslag i massmedia.

Undervisningen kan även undersöka hur synen på AI förändras över tid, liksom gränsen för vad som kallas AI. Troligen kan eleverna identifiera situationer som innehåller AI utan att tänka på det, till exempel rekommendationssystem vid e-handel och streamingtjänster.

Genom att diskutera tekniker som är viktiga inom AI går det också att närma sig vad AI är, till exempel genom begreppen robotik och behandling av naturligt språk (NLP).

- *Vikten av data, datakvalitet för AI och val av data.*

Data är grundläggande för AI. Några avgörande faktorer för framgångsrika AI-system är val av data, dess kvalitet och kostnaden för att ta fram eller få tillgång till den data som behövs. AI-modellens förmåga att ge korrekta svar är beroende av både kvantitet och kvalitet för att lära sig mönster och fatta beslut.

Urvalet av data är avgörande för kvaliteten och behöver därför vara balanserad, representativ och opartisk, det vill säga inte innehålla bias. Här finns intressanta etiska frågor att samtala om i klassrummet. Vilka effekter får fördomsfulla data för oss som mottagare av budskap och information?

Även situationer där andra urval är önskvärda kan förekomma. Sällsynta händelser, som olyckor eller ovanliga sjukdomar, är exempel där representativa data inte innehåller tillräckligt med exempel för att kunna göra en meningsfull analys. Då kan data behöva rensas och förberedas och bristfälliga, felaktiga eller inkonsekventa värden hanteras.

- *Drivkrafter bakom utvecklingen av AI.*

Utvecklingen av AI möjliggörs av kraftfulla datorer och tillgång till stora datamängder. Detta gör det möjligt att träna komplexa modeller och algoritmer. Att förstå vad som driver utvecklingen av AI kan hjälpa eleverna att förstå varför AI utvecklas och i vilken riktning.

Samhällets demografi och förändrad konsumtion är några faktorer som driver på behovet av ökad automatisering. Viljan hos företag och organisationer att investera i AI-forskning och utveckling driver också fram innovation inom

området. Efterfrågan på ny teknik utifrån säkerhet och arbetsmiljö har blivit större och framsteg inom bildigenkänning och förståelse av naturligt språk har ökat antalet tillämpningar för AI.

För eleverna kan det även vara intressant att diskutera vilka effekter användningen av AI-teknik får. Hur påverkar övervakningsteknologi eller AI-tillverkad konst vår integritet respektive immateriella rättigheter? Dessa kopplingar kan göras löpande när olika tekniker hanteras, och ur flera aspekter, exempelvis sociala, ekonomiska eller etiska. Genom att samtala om bland annat risker, människosyn och kommersiella intressen kan kritiska perspektiv lyftas.

- *Översikt över användningen av AI, däribland prediktion, robotik, vision och generativ AI samt olika tekniker och metoder som ligger bakom dessa.*

Användningen av AI är omfattande och ingår bland annat i sökmotorer, bildigenkänning och översättningstjänster. För dessa finns olika tekniker och metoder och de som anges i innehållspunkten ovan ska ingå i undervisningen. I samband med detta kan även stora språkmodeller tas upp i undervisningen, dels för att förklara vad de är och hur de fungerar på en övergripande nivå, dels för att belysa deras svagheter.

- *Översikt av tekniker för AI, däribland sökning, klassificering och objektigenkänning.*

Här kan olika tekniker för AI presenteras på en konceptuell och övergripande nivå, men det finns också utrymme för fördjupning inom någon eller några tekniker som kan studeras mer i detalj. De AI-tekniker som skrivs fram i innehållspunkten ska ingå i undervisningen. Utöver dessa kan andra tas upp i undervisningen, som till exempel textklassificering, rekommendationssystem, ruttplanering och spelagenter.

- *Översikt av metoder och algoritmer inom AI, däribland beslutsträd, regression samt övervakat och oövervakat lärande.*

Metoder och algoritmer är grundläggande för att lösa problem med hjälp av AI. De begrepp som anges i innehållspunkten ovan ska tas upp i undervisningen. Beroende på elevernas förkunskaper i matematik kan läraren välja några av följande exempel på metoder: naiva bayesianska klassificerare, k-närmaste-grannar-algoritmen eller linjär regression.

Praktisk användning av AI

Under denna rubrik finns fyra innehållspunkter som ska ge möjlighet att ge undervisning i praktisk användningen av AI.

- *Översikt över processen för problemlösning med AI.*

Ett systematiskt arbetssätt och en väl beprövad process är centralt vid problemlösning med AI. I undervisningen kan jämförelser göras till andra situationer då eleverna mött och arbetat med problemlösning, till exempel inom

design eller matematik. Några steg som ingår i processen, och kan beröras, är problemformulering, datainsamling och modellval. Efter utvärdering av resultatet kan det också finnas behov av förbättringar. Som en avslutande del i processen ingår bland annat implementering och driftsättning.

- *Enklare typ av problemlösning med hjälp av AI, till exempel klassificering, objektigenkänning, prediktion, tolkning och bearbetning av naturligt språk (NLP), enklare maskininlärning och användning av spelagent.*

Eleverna ska ges möjlighet att lösa enklare problem med hjälp av AI. De tekniker och metoder som anges i innehållspunkten är exempel, och läraren kan välja andra. Denna del kan genomföras med hjälp av färdiga programmeringsexempel, egen kodning med hjälp av befintliga programmeringsbibliotek eller tillgängliga tjänster och plattformar för AI.

- *Principer för, och tillämpningar av, vanliga tekniker inom AI, däribland maskininlärning och robotik.*

Maskininlärning och robotik är två principer och tillämpningar inom AI som ska ingå i undervisningen, men fler kan presenteras. Några exempel som kan lyftas är hur AI kan eliminera riskfyllda arbetsmoment, hur generativ AI förändrar åtkomst av information, när det är en fördel att implementera AI eller när själva AI:n kan bli en risk.

- *Metoder för enklare träning av AI.*

Undervisningen ska ge eleverna möjlighet att träna enklare AI med några olika metoder. Grundläggande begrepp att utgå ifrån kan vara tränings- och testdata samt märkt och omärkt data. Genom att låta eleverna dela upp data i tränings- och testmängder kan de därefter diskutera hur man kan använda träningsdata för att anpassa sin modell och sina testdata för att utvärdera datans prestanda. Här finns också möjlighet att diskutera hur AI-modeller kan tränas på olika sätt, till exempel övervakat och oövervakat lärande eller förstärkningsinlärning samt hur dessa kan kombineras.

Förhållningssätt till AI

Under denna rubrik finns fem centrala innehållspunkter som ska ge möjlighet att ge undervisning om samhällsvetenskapliga aspekter av AI.

- *Jämförelse mellan hur enklare lösningar med AI fungerar och hur en människa löser samma problem.*

Förutom att diskutera olika sätt där AI och människor löser problem på liknande eller helt olika sätt, kan undervisningen ta upp vilka konsekvenser detta kan få. Några förslag på likheter mellan hur AI och människor löser samma problem skulle kunna vara dataanvändning, mönsterigenkänning, användning av algoritmer och verktyg. Förslag på skillnader som kan lyftas i undervisningen är hastighet, kreativitet och etiska överväganden.

- *Några situationer där AI är överlägsen människan och tvärtom.*

Det finns en rad situationer då AI är överlägsen människan. Några exempel som kan tas upp i undervisningen är snabb dataanalys, repetitiva uppgifter och självkörande fordon. AI kan bearbeta stora mängder data på kort tid, vilket är användbart inom medicinsk diagnostik, finansanalyser och klimatmodellering. AI kan också utföra monotona och repetitiva arbetsuppgifter utan att tröttna, till exempel sortera och kategorisera stora datamängder. Dessutom kan AI köra fordon mer exakt och reagera snabbare än människor.

Den motsatta situationen, då människan är överlägsen AI, kan vara vid etiska beslut och social interaktion eller vid tillfällen som kräver kreativitet och intuition. Människor kan göra avvägningar baserat på moraliska och etiska överväganden, medan AI agerar enligt programmerade regler. Dessutom har människor förmågan att förstå känslor, bygga relationer och kommunicera på ett sätt som AI inte kan. Ett exempel på detta är kroppsspråk. Människor är också bättre på att lösa komplexa problem.

Genom fallstudier, där eleverna får analysera verkliga exempel, kan undervisningen belysa tillfällen där AI har överträffat människor eller vice versa. Genom att eleverna får genomföra projekt där de skapar egna AI-lösningar och jämför dem med mänskliga metoder, kan ytterligare situationer belysas. Det kan också vara lämpligt att låta eleverna reflektera över och samtala om när AI är ett lämpligt val och när mänskligt omdöme är nödvändigt. I sådana samtal är det angeläget att även diskutera och kritiskt reflektera över konsekvenser av att låta AI respektive människan fatta beslut.

Ytterligare exempel på vad som kan lyftas i undervisningen är situationer där AI uppenbart gör fel eller inte förstår, samt AI i förhållande till vad som är enkelt respektive svårt att automatisera. I detta sammanhang kan Moravecs paradox diskuteras.

- *Normer samt lagar och andra bestämmelser som gäller AI och dess användning.*

Lagstiftningen som gäller AI sker samtidigt med AI-utvecklingen. De lagar och andra bestämmelser, normer och principer som finns behöver eleverna känna till. Vissa är nationella, andra internationella. De bestämmelser som gäller inom EU kan vara av särskilt intresse att belysa. På samma sätt som AI påverkar lagstiftningen har även juridiken och rättsliga frågor en påverkan på AI, bland annat när det gäller upphovsrätt för genererat material och ansvarsfrågor.

- *Etiska dilemman med användandet av AI, däribland transparens.*

Användning av AI, och kanske även själva tekniken, väcker ett antal etiska frågor. Några av dessa är breda och svårare att fånga, därför kan ett sätt att närma sig dem vara att ta upp mer specifika frågeställningar som sedan generaliseras. Som innehållspunkten anger är transparens en aspekt som ska tas upp i undervisningen, medan andra aspekter av etiska dilemman väljs utifrån elevgrupp, program och andra omständigheter.

Några exempel på etiska frågeställningar som kan tas upp är bristen på transparens, AI som svart låda och vem som bär ansvar när AI betar sig på ett sätt som inte kunnat förutsägas. I vilka situationer är förklarbarhet viktigare än prestanda i ett AI-system och omvänt? Ett annat område som är värt att diskutera är upphovsrätt. Vilken data är det etiskt försvarbart att träna AI-modeller på? Hur ställer vi oss till AI-skapad media som härmar särskilda konstnärer? Vem, om någon, borde ses som upphovsperson för AI-skapad media? Ytterligare ett intressant område att samtala om är bias och fördomar. Det kan finnas risk att vissa grupper gynnas eller missgynnas utifrån till exempel kön, ålder eller hudfärg. Samtidigt finns även bias och fördomar hos människor, och med AI-modeller finns möjligheter att systematiskt undersöka och motverka bias. Hur bör man se på att använda AI-stöd där fördomar hos både människor och AI sannolikt påverkar beslut?

- *Demokratiska, sociala, ekonomiska, miljömässiga och säkerhetsmässiga möjligheter och risker med AI-användning samt dess konsekvenser för samhället.*

Möjligheter och risker med AI samt kopplingen till demokrati och hållbarhet är minst sagt ett stort område. Flertalet tekniska AI-lösningar kan diskuteras i ljuset av dessa aspekter. De perspektiv som skrivs fram i innehållspunkten ska belysas i undervisningen och ge eleverna tillfällen till reflektion och diskussion. Några ytterligare exempel på möjligheter och risker följer nedan.

Tillgängligheten till myndigheter har på flera sätt ökat genom att enskilda medborgare kan få svar på sina frågor snabbare via hemsidor och chatfunktioner. Samtidigt utestängs grupper som av olika anledningar inte kan använda eller ta del av digital kommunikation.

Informationsflödet påverkas genom ökade möjligheter till stärkt nyhetsproduktion, dataanalys och tillgång till information. Samtidigt främjar algoritmerna reaktioner som ger större utrymme för sensationsnyheter, men också desinformation. Hur påverkas demokratin av trollfabriker och alternativa fakta? Vilka konsekvenser har det för det demokratiska samtalet? Och hur kan vi navigera i ett förändrat informationslandskap? Flera beslut fattas idag helt eller delvis av AI. Det kan handla om låneansökningar och villkor för räntor och återbetalning, beslut om bidrag eller sortering vid ansökan om jobb. Beslut som får både sociala och ekonomiska konsekvenser för den enskilda medborgaren.

Inom hälso- och sjukvård har användningen av AI inneburit snabbare och många gånger säkrare diagnostik. Robotar kan på sikt överta delar av hemsjukvården. Men vilka etiska frågor väcker det? Hur stor är risken för feldiagnostik och kan en robot ersätta mänsklig kontakt?

På miljöområdet används AI på flera sätt för att få säkra och snabba sammanställningar, modeller och analyser. Automatisk skadebedömning, energibesparingar och realtidsövervakning av klimatsystem är några exempel där AI är en stor tillgång för välgrundade beslut. Samtidigt är användningen av AI i sig energikrävande och lagring av data förutsätter stora serverhallar.

Utöver hållbarhetsaspekterna finns en rad säkerhetsfrågor kopplat till AI. Ökade möjligheter till brottsförebyggande arbete, till exempel för att förebygga bedrägerier, väcker samtidigt frågor om personlig integritet. Olika lösningar för att öka säkerheten diskuteras bland forskare, till exempel genom inbyggda säkerhetsmekanismer i öppet tillgängliga AI-verktyg och vattenmärkning av AI-genererat material.

En angelägen fråga är hur vår syn på kunskap påverkas av AI. Vad är och vad blir viktigt att kunna? Vad behöver vi själva kunna och vad kan vi lämna över till maskiner? Även om maskiner kan utföra saker åt oss behöver vi diskutera vad vi själva behöver kunna för att AI inte ska göra oss svagare. Här finns en rad psykologiska och sociala aspekter att samtala om.

Nivå 2

Nivå 2 i ämnet har en tydlig progression från nivå 1, där det centrala innehållet både breddar och fördjupar det innehåll som eleverna tidigare mött.

Artificiell intelligens

Inom detta område fördjupas den generella kunskapen ur motsvarande rubrik i nivå 1. Detta kan göras genom att i undervisningen fördjupa kunskaper relaterat till ett specifikt tillämpningsområde för AI. Fördjupningen handlar exempelvis om olika begrepp och tekniker, om dataurval och bias eller om AI-användningens effekter inom ett område kopplat till elevens program.

Praktisk användning av AI

Progressionen består i att kunskaperna breddas genom att fler steg i processen för problemlösning behandlas samt genom att olika metoder och algoritmer tydligare värderas. Nivå 2 ger också möjlighet att hantera inte bara enklare typer och principer utan även fördjupning inom ett urval av metoder och deras användning i specifika situationer.

Förhållningssätt till AI

Det centrala innehållet i nivå 2 fördjupas genom att i undervisningen inte enbart jämföra utan även värdera och analysera konsekvenser och risker samt att fokusera på svagheter i AI-lösningar. Genom att utgå ifrån den praktiska användningen av AI som eleven genomför under nivå 2, kan lösningarnas konsekvenser analyseras på ett mer komplext sätt.