

Läsåret 24/25

FÖRSÄTTSLAD TILL SLUTUPPGIFT I KURSEN

LL142U: Artificiell intelligens i gymnasieskolan: Samhällsperspektiv

Lektionsplanering promptning och agenter

Håkan Elderstig

2024-01-04

Lärare/kursansvarig: Per Norström, KTH

Lektionsplanering promptning och agenter	2
Sammanfattning	2
Bakgrund	2
Syfte	3
Tidsåtgång	3
Användning	3
Förkunskaper	3
Innehållet kommer att vara under förändring	3
Innehåll	4
Lektion 1 - Promptning och vad man kan lära av det	5
Effektiv promptning	5
Olika modeller	5
Inbyggda problem	5
Lektion 2 - Promptning för lärande	6
Undervisningsmaterial som används	6
Uppgifter	6
Koppling till Centralt innehåll	6
Bedömning	7
Litteratur	7

Lektionsplanering promptning och agenter

Sammanfattning

Denna lektionsplanering syftar till att lära elever hur man använder AI på ett konstruktivt sätt i sitt lärande. Målet är att eleverna ska utveckla sin förmåga att prompta effektivt och öka sin förståelse för hur stora språkmodeller fungerar. Lektionsplaneringen består av två lektioner, en om effektiv promptning och en om att skapa agenter. Eleverna kommer att lära sig hur man kan använda AI för att förbättra sin studieteknik och hur man kan skapa agenter som kan hjälpa dem i framtiden.

Bakgrund

Min skola, Stockholm Science & Innovation School, driftar språkmodeller på vår egen server och det är en tjänst som alla inom Stockholm stad kan logga in på och använda. Vi har sedan våren 2023 undervisat våra elever i AI-litteracitet men en erfarenhet av detta är att det skulle vara bra om vi hade en del om promptning, vilket jag passar på att utveckla här.

Tidigt våren 2023 insåg vi problematiken kring fusk och hemuppgifter vilket ledde till fler examinationer i skolan på datorer i säkert provläge vilket missgynnar många elever. Det finns goda skäl att lära eleverna hur de använder AI på ett konstruktivt sätt i sitt lärande. Kompetent och medveten promptning är en del i det.

Syfte

Eleverna ska lära sig använda AI på ett sätt som gynnar deras studier. Det ska ske genom att de utvecklar sin förmåga att prompta effektivt och ökar sin förståelse för hur stora språkmodeller fungerar.

Tidsåtgång

Denna planering beskriver två lektioner om studieteknik med promptning och agenter. Lektion 1 fokuserar på effektiv promptning, medan lektion 2 handlar om att skapa agenter. Materialet är tänkt att kombineras med annat undervisningsstoff för att ge en verklighetsnära nytta och något att öva med. Du kan anpassa materialet efter dina behov och hinna med det viktigaste på två lektioner, för att sedan återkomma till valda delar senare.

Användning

Lektionsplaneringen passar även att använda i andra kurser än AI-kursen, och även på introduktionsdagar, AI-dagar eller för lärarfortbildning. Även om jag planerar detta för teknikprogrammet på gymnasiet så är det användbart och angeläget inom alla kurser på gymnasiet och på grundskolan. Det är lämpligt att samverka med andra ämnen för att visa på allmängiltigheten och nytta av att förstå promptning i olika sammanhang.

Förkunskaper

Vi utgår från att eleverna har varierande förkunskaper och börjar från noll för att alla ska hänga med.

Koppling till Centralt innehåll i kursen Artificiell Intelligens 1

Promptning är en viktig kompetens och bör således ingå i en AI-kurs men det har inte ett tydligt stöd i kursens centrala innehåll men genom promptningsövningar kommer vi att arbeta med många centrala begrepp. Lektion två handlar om att skapa agenter (ett slags digitala lär-assistenter) för lärande och det får ju anses vara problemlösning med AI inom lärande.

Artificiell intelligens

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

Praktisk användning av AI

- Översikt över processen för **problemlösning** med AI.

Denna promptskola tar ett brett grepp om hur språkmodeller fungerar i praktiken och berör centrala begrepp som **prompt**, **temperatur**, **tokens**, **bias**, **agent**, **systemprompt**. Vidare jämförs olika modeller och begrepp som **inferens**, **instruct** och **parameter** studeras i praktiken.

Innehållet kommer att vara under förändring

Ansvarsfriskrivning

Den här lektionsplaneringen innehåller inga minutplaneringar, arbetsblad, instruktioner eller detaljplaner. Däremot ges ramar, riktlinjer, hållpunkter och uppslag. Skälet till denna avgränsning är att AI och stora språkmodeller utvecklas så fort att det som planeras nu kommer att vara föråldrat när detta ska användas (exempelvis hösten 2025). Det kommer hela tiden nya språkmodeller vilket kräver anpassningar. Skolor och lärare har olika lösningar på hur och vad man använder för AI tillsammans med eleverna. I texten nedan har jag haft de språkmodeller som finns tillgängliga på SSIS och i hela Stockholm stad i åtanke men förhoppningsvis går planeringen att använda i andra sammanhang.

Det som ges i den här lektionsplaneringen är punktlistor med saker som man kan ta upp. Man måste inte arbeta igenom alla punkter utan man kan välja att fokusera på det som man finner intressant, som passar undervisningsgruppen. Gör gärna eleverna delaktiga i den urvalsprocessen och de kan även vara behjälpliga i att ta fram underlag för övningar och eleverna kan exempelvis dela med sig av bra promptar i ett delat dokument på Google Drive eller ta AI till hjälp för att planera sin egen undersökning. Låt de här två lektionerna bli ett inspirerande tillfälle för gemensamt samlärande.

Den snabba utvecklingen kräver Just-in-Time-planering

Den här lektionsplaneringen innehåller dels teorin, alltså exempel på vad man kan berätta för eleverna, och dels övningar som man kan använda i lektionerna.

Eftersom AI har förändrats i så snabb takt har jag inte valt att göra detaljerade lektionsplaneringar utan jag exemplifierar hur jag tycker att man kan agera utifrån läget idag, december 2024. Det finns många möjligheter att ta AI:s hjälp med att planera olika delmoment i en lektion. Internet och sociala medier ger också många uppslag till delar som kan användas i undervisningen. Jag har fokuserat på att

ge riktlinjer för hur man kan lägga upp undervisningen i stort. I praktiken kommer detaljerna så klart att variera beroende på undervisningsgruppens egenskaper och lärarens förkunskaper, intressen och mål.

Hur man undervisar om AI kommer att förändras över tid, speciellt nu när kursen Artificiell intelligens är ny. Men framför allt förändras förutsättningarna beroende på att generativ AI förändras i en rasande takt så en del av det som jag beskriver här kommer förhoppningsvis att fungera när du läser detta, men mycket annat kommer inte att fungera så det som jag har försökt visa i denna lektionsplanering är exempel på hur man kan arbeta och principer för hur man kan använda AI för att skapa undervisning om AI.

Jag hoppas det framgår av mina instruktioner för hur undervisning ska bedrivas att det är mycket viktigt att man introducerar, engagerar och inkluderar eleverna i planeringen och utformningen av undervisningen. Det är ju trots allt de som ska använda den här kompetensen, och vår framtid kommer att påverkas av AI-utvecklingen, men tänk på situationen för våra elever. Det kommer inte att ha hunnit utbilda sig klart om de väljer en traditionell utbildning till exempelvis civilingenjör innan AI har gjort så stora framsteg. Arbetsmarknaden kommer att se helt annorlunda ut på många områden än den gjorde när de började sin utbildning. Detta perspektiv finns som en bakomliggande närvaro i dessa två lektioner även om det inte ingår explicit. Ju mer man inser potentialen i den AI som nu sköljer över oss desto mer inser man vilka konsekvenser utvecklingen kan ha på några års sikt varför jag kommer att göra en liten utvikning på temat.

Det finns personer med god insyn i AI-utvecklingen, exempelvis Leopold Aschenbrenner (Aschenbrenner, 2024), som menar att det finns en risk för att vi når den teknologiska singulariteten (Wikipedia) där AI "tar över" redan 2028 och att artificiell intelligens kommer att kunna ta över vårt jobb och utföra det på forskarnivå redan 2027.

Perspektivet är svindlande. I denna lektionsplanering använder vi promptningar mot modeller som är relativt begränsade i sin kapacitet, även om de är otroligt duktiga. Men samtidigt finns det prognoser inför framtiden som pekar på att vi har en helt annorlunda situation med superintelligenser om 3, 4, 5 år. Så vad är meningen med denna lektion? Jo, det är viktigt därför att vi måste börja förstå AI från de mer vardagsnära och enkla utgångspunkterna för att senare kunna komma till de viktiga frågorna om bias, etik, transparens, alignment och var AI kommer att ta oss i framtiden.

Jag hoppas att denna lektionsplanering kommer att vara en bra utgångspunkt för dig att börja utforska AI och dess möjligheter och utmaningar. Kom ihåg att AI är ett område som är i ständig utveckling, och det är viktigt att vara öppen för nya idéer och perspektiv.

Denna lektionsbeskrivning innehåller få detaljerade beskrivningar av hur och vad som kan göras minut för minut. Det är upp till dig som lärare att välja hur du utformar undervisningen i detalj. Det finns som sagt många länkar och resurser men det enklaste och viktigaste är att du använder AI som stöd i din förfining av denna planering. Det är när du använder AI ofta i din vardag som du blir duktig på att använda den.

Innehållet i de två lektionerna

Vi går i två steg så att man först lär sig om språkmodellerna genom att bara prompta. I andra steget skapar man en agent/assistent som man kan använda som stöd för sitt lärande.

Lektion 1 - Promptning och vad man kan lära av det

Lär känna modellerna

Denna introduktion lämpar sig för en **gemensam övning** där du som lärare visar och eleverna gör efter och provar själva.

Det finns flera kända kommersiella språkmodeller, exempelvis ChatGPT, Gemini, Claude. Men det finns även fritt tillgängliga modeller (open weights) med prestanda i närheten av de kommersiella modellerna, exempelvis Qwen, DeepSeek, Mixtral och Meta Llama. SSIS använder den sistnämnda i storleken 70 B (70 miljarder parametrar/vikter).

Vad har språkmodellen för egenskaper? Vi undersöker någon stor språkmodell, exempelvis Meta-Llama-3.3-70B, Vi försöker utröna om den har en **personlighet**. Hur kan vi beskriva den i så fall? Vi undersöker om den har moral, undersöker om den har politiska åsikter eller tro. Kan vi få den att ändra personlighet?

Övning: Ge instruktioner till modellen så att den beter sig på ett visst sätt med en personlighet och andra egenskaper som du bestämmer.

Sen ska vi undersöka skillnaden mellan modeller. Vi jämför den breda generella standardmodellen med specialiserade modeller för kodning eller matematik. Kan vi se skillnad i kvalitet och hur kan man bedöma skillnaden? Än tydligare skillnader finns om man jämför med mindre modeller (0,3 - 3 GB är lagom). Identifiera några frågor där det är tydligt att svaren beror på **sannolikheten** att ett ord följer på föregående, det vill säga svaren blir slumpmässiga. Syftet är att påvisa att modellerna inte är så fantastiskt bra utan att svaren vi får beror på sannolikheter och slump.

Övning: Andelen korrekta svar. Helst ska man ställa samma fråga typ tio gånger och ladda om sidan emellan för att få en siffra på sannolikheten för det felaktiga svaret. När en hel klass ställer samma frågor kan det bli en fin sammanställning.

Effektiv promptning

Det här är en öppen laboration där eleverna själva skapar frågor och arbetar utifrån punktlistan nedan. Eleverna uppmuntras att utveckla punkterna och utforska enligt eget intresse. De antecknar sina erfarenheter och reflektioner i en **loggbok**.

- Hur promptar man? Använd [Promptskolan](#) (EDU AI, 2024)
- Undersök några verktyg som integrerar sökningar, exempelvis:
 - Perplexity
 - Googles NotebookLM.
- Testa promptning för bild, film, musik. Tillgängligheten varierar men sök på [EDU AI](#) eller Google.

Kända inbyggda problem

- Kan vi upptäcka bias. Går det att prompta bort bias?

Det finns många ingångar till att undersöka bias, exempelvis the American smile (Gurfinkel, 2023), vilket är värt att diskutera även om modellerna idag är förbättrade så man kan inte återskapa just detta.

Det finns ett utmärkt material i: Undersök mansrollen; En lektion skapad av Joel Rangsjö (Rangsjö, 2024.). Denna [lektionsbeskrivning](#) innehåller så pass många frågor att eleverna kommer att vara sysselsatta en god timme.

- Hallucinationer. Testa frågorna i [artikeln](#) *De gör fortfarande fel* i Computer Sweden (Elgan, 2024), Är det så? Testa exempelfrågorna på några olika språkmodeller, gärna mindre varianter. Kan du hitta andra typer av frågor som är enkla för människor men svåra för språkmodeller? Läs mer om hur man skapar tester för språkmodeller, exempelvis *SimpleQA* (OpenAI, 2024)

Exempel: Fråga vad ett "Gustavfår" är till några olika modeller. Här får du ofta fantasifulla svar trots att det inte finns något som heter Gustavfår, vilket en snabb googling visar.

Exempel 2: Den gör sitt bästa men når inte ända fram: Lös $x^2 = 2^x$

Lektion 2 - Promptning för lärande

Nu ska vi skapa digitala assistenter eller agenter med ett annat ord. OpenAI kallar dem GPTs, (OpenAI, 2023). På SSIS kallar vi dem agenter och beskrivningen nedan utgår från vår tjänst [TeachGPT](#) men det fungerar snarlikt på OpenAI eller andra tjänster.

Hur använder jag AI utan att ta genvägar? Det är ju ingen poäng om AI gör hela jobbet. Om du däremot får lite hjälp på vägen men gör det mesta själv så har du varit delaktig i processen och sannolikt lärt dig en del på vägen. Du ska vara delaktig i skapandet och samarbeta med AI. Forskaren Ethan Mollick har beskrivit detta i sin bok *Co-Intelligence..* (Mollick, 2024). Andra talar om kentaureer och refererar till schacktävlingar där människor använder datorer för att bli bättre och till och med slå rena schackdatorer (Sundin 2023). Som förberedelse för framtida arbetsliv är detta en värdefull förmåga (Myndigheten för arbetsmiljökunskap, 2022).

Uppgift: Skapa din egen agent. Detta kan vara en uppgift som bedöms.

- Välj typ av agent?

Vad finns det för olika typer av agenter, exempelvis en allmän lärare, en lärkompis, en hjälp för ett specifikt ämnesområde, att träna språk, skrivande, programmering, Du kan definiera agenten smalt som exempelvis för att träna svensk språkhistoria eller bredare som att den kan hjälpa till med allt inom matematik.

- Skriv en systemprompt. Den ställer in hur agenten ska bete sig och vad den kan bäst. Du kan exempelvis skriva

"Du är en tekniklärare som är bra på elektronik och programmering av Arduino. Du ger mig inte färdig kod utan frågar vad jag har för idéer först och så ger du återkoppling och små förslag- Du förklarar hur alla delar fungerar och kontrollerar med små frågor att jag förstår."

Testa en systemprompt, modifiera den, förbättra testa igen och försök skapa din extralärare, lärlärare, coach, bollplank eller vad du trivs bäst med. Målet är att du ska lära dig på ett effektivare sätt och samtidigt ha en mer personlig relation än att bara chatta med en slumpmässig språkmodell.

- Några inställningar du kan göra i agenten är: **Temperatur** och max antal **tokens**.

Temperaturen är en parameter som styr kreativiteten hos en språkmodell och därmed i förlängningen hur benägen den är att hallucinera. Om man ändrar till ett extremt höga värden på temperaturen så ser man hur modellen kraschar totalt. Men testa standardmodellen med värden som är lite lägre och lite högre än det normala, säg 0,3 respektive 1,2 och ge den två olika uppgifter som där en kräver stramhet och en kräver mer kreativitet. Som exempel kan vi fråga om matematiska begrepp respektive att skriva en dikt.

Parametern Max antal tokens anger ett högsta antal tokens som en agent kan använda i sitt svar. En token är grovt uppskattat fyra tecken långt.

Eftersnack

Det kan vara på sin plats att diskutera risker som finns med att skapa digitala assistenter om de blir för personliga. Det finns risker med virtuella kompisar (Falk, (2024). Tag en stund att diskutera detta och var lyhörd och uppmärksam om du ser tecken på elever som har ett osunt förhållande till sin agent eller AI i något sammanhang.

Undervisningsmaterial som används

Först kommer du att använda olika resurser på nätet för dessa lektioner. Det finns mer material än man hinner använda och om det saknas något kan du låta AI generera det du behöver. Det viktiga är snarare att söva och samla det som är viktigt att förmedla. Samtidigt är det viktigt att lämna lite öppet för eleverna att utforska och upptäcka. Eleverna behöver lära sig själva att lära med AI och om AI. De länkar som finns här fungerar också som startpunkter för att forska vidare kring vad som hänt efter att detta skrevs.

Tips: Jag skulle göra en Powerpoint utifrån denna instruktion och lyfta de viktigaste momenten som jag vill ha med. Därtill skulle jag använda ett delat textdokument för att visa exempel, förbereda agenter och låta eleverna ge respons och reflektioner samt bidra till utvecklingen.

Bedömning

Det här är ett litet område i kursen och det är inte nödvändigt med examinerande uppgifter på detta moment. Däremot är det bra att öva eleverna på att redovisa sina kunskaper och förmågor inom AI-området samtidigt som vi vill säkerställa att alla elever tillgodogjort sig grundläggande prompt-kunskaper (i bred bemärkelse).

Uppgift

1. Utveckla din egen agent (AI-assistent). Inlämning av agentens systemprompt och relevanta inställningar samt en exemplifierande chatt från när du använt agenten.

Uppgiften bedöms summativt med möjlighet för eleverna att lämna in på nytt efter att de fått respons på uppgiften. Det finns en inbyggd formativ process där elevernas interaktion med språkmodellerna kommer att utveckla deras förmåga.

I bedömningen kan följande kriterier användas (för betyg E):

Eleven visar godtagbara kunskaper om AI, dess tillämpningar, tekniker och beståndsdelar.

Eleven använder med viss säkerhet AI för att lösa enkla problem.

Det viktigaste resultatet av detta moment är att eleverna har skapat agenter som de har nytta av och kan skapa fler användbara agenter i framtiden - inte deras betyg på momentet..

Diskussion

Denna lektionsplanering har begränsningar som beror på den snabba utveckling inom AI som förändrar förutsättningarna för undervisningen. Det går helt enkelt inte att göra detaljerade planeringar för undervisning som ska ske i framtiden. Vad som däremot skulle gå att göra är att skapa en levande samling dokument med löpande uppdateringar, klassificeringar, omdömen och utvärderingar. Så en framtida utvecklingsväg för förbättring vore att skapa en databas med all information. Eller så tar man fasta på att dokumenten utvecklas på ett sätt som liknar mjukvaruprojekt. Det vore värt att prova att lägga upp detta och andra dokument på GitHub eller **GitLab**.

Litteratur

1. Aschenbrenner, L. (2024, juni.). *Situational Awareness*. Hämtad från <https://situational-awareness.ai/>
2. EDU AI. (2024). *Promptskolan*. Hämtad från <https://www.eduai.se/promptskolan/>
3. Elgan, M. Computer Sweden. (2024, november 18). *AI är dummare än du tror*. Hämtad från <https://computersweden.se/article/3607869/ai-ar-dummare-an-du-tror.html>
4. Falk, J. (2024). [Inte säkert att AI-vännen är din bästa kompis](#) . *Dagens nyheter*, 30 september 2024.
5. Gurfinkel, J. (2023, mars 27). *AI and the American Smile. How AI misrepresents culture through a facial expression*. Hämtad från: <https://medium.com/@socialcreature/ai-and-the-american-smile-76d23a0fbfaf>.
6. Mollick, E. (2024). *Co-intelligence*. Penguin Random House. Hämtad från <https://www.penguinrandomhouse.com/books/741805/co-intelligence-by-ethan-mollick/>
7. Myndigheten för arbetsmiljökunskap. (2022). *Artificiell intelligens, robotisering och arbetsmiljön*. Kunskapssammanställning 2022:1 978-91-986882-6-9
8. OpenAI. (2023, September 6). Introducing GPTs. Hämtat från <https://openai.com/index/introducing-gpts/>
9. OpenAI. (2024, oktober 30). *Introducing SimpleQA*. Hämtad från <https://openai.com/index/introducing-simpleqa>

10. Rangsjö, J. (2024.). *Instruktion: Utforska mansrollen och könsstereotyper i AI-system*. Hämtad från <https://docs.google.com/document/d/10-sjPZVAPzCN0qjw4UhT8srjU-i3msDX/mobilebasic>
11. Sundin, M. (2023, november 23). *Så slog två amatörer stormästarna i schack*. WARP NEWS. Hämtad från <https://www.warpnews.se/kentaurens-fordel/sa-slog-tva-amatorer-stormastarna-i-schack-utdrag-ur-nya-ai-boken-kentaurens-fordel/>.
12. Wikipedia. (u.d). *Teknologisk singularitet*. Hämtad från https://sv.wikipedia.org/wiki/Teknologisk_singularitet.