Människan, AI och energianvändning

JERKER BJÖRKQVIST

Artificiell intelligens har idag blivit en etablerad del av många verksamheter. AI används bland annat som verktyg för att korrekturläsa, översätta och sammanfatta texter, i arbete med bilder, musik med mera. AI används också alltmer som ett verktyg i undervisning. Samtidigt medför AI också problem, det kan locka till fusk och kan stärka fördomar. En stor fråga idag är det enorma energibehov som datacenter behöver. Jerker Björkqvist skriver.

Artificiell intelligens (AI) är precis vad begreppet säger – alltså artificiell. Artificiell har synonymerna konstgjort, syntetisk, oäkta, konstlad. AI är när vi får en maskin att verka intelligent, i praktiken är det fråga om en datamaskin, som idag för det mesta utför sina beräkningar som molntjänst. Dataingenjörer har programmerat den datamaskin som erbjuder AI-tjänster åt oss. Knappast någon har väl inte hört om ChatGPT eller Copilot, skapade av företagen OpenAI respektive Microsoft.

Det stora genombrottet för varje-mans-AI var ett nytt sätt att ”lära sig” från existerande data, beskrivet i en forskningsartikel med titeln ”Attention Is All You Need” från 2017. I den beskrivs det som de kallar en ”Transformer”-arkitektur, vilken möjliggör en mycket snabbare träningsprocess. Med träning menas här att programvara bygger upp en intern representation utgående från existerande texter, för att sedan kunna generera texter som efterapar det som mänskor tidigare skrivit. Denna representation kallas en språkmodell, och man talar idag om stora språkmodeller (Large Language Models, LLM) för att de är tränade på stora mängder data.

När språkmodeller är tränade på stora mängder data, så känns tjänster som ChatGPT facinerande, för de verkar kunna skriva bra text. Och inte bara skriva, utan också skriva om våra egna texter, korrekturläsa, och göra sammanfattningar på existerande texter. Hur mycket text ChatGPT är tränat på är inte offentligt, men gissningar säger ungefär 45 TB. Om man räknar om denna mängd data i bokform, så blir det ungefär 135 miljoner böcker. Men träningsdata är inte bara böcker, utan källor är också GitHub (källkod till programvara), Wikipedia, diskussionsforum och liknande.

Men det är viktigt att komma ihåg att en språkmodell såsom ChatGPT är en maskin, och precis som alla andra maskiner gör de det som tillverkaren har byggt in, inget annat. En stor språkmodell är programmerad att reagera på vår input genom att söka en statistiskt sannolik fortsättning på det vi har skrivit. Så det är inte fråga om någon som helst kognitiv verksamhet, bara simpel statistik och korrelation, just sådan verksamhet som datorer slår oss mänskor med siffrorna 100-0. Men det är fascinerande, precis som antagligen de första bilarna var, eller första gången man kunde sätta sig i ett flygplan och flyga till andra sidan världen.

Det har diskuterats en del om hur man skall förhålla sig till stora språkmodeller i undervisningen. Skall man tillåta elever och studerande att använda ChatGPT för att producera texter, skall man använda språkmodeller för att tolka existerande texter? Det som är viktigt att komma ihåg är att för elever och studerande är inte den producerade texten det viktiga, utan den kognitiva process som ledde till texten är det viktiga. Själva texten har för inlärandet inte något annat värde än att det är ett bevis på en lärprocess. Men om texten är genererad av en språkmodell så finns det inget värde i den, för den är då inget bevis på en lärprocess.

Språkmodeller är verktyg för språkhantering. En bil är ett verktyg för att ta sig fram. En cirkelsåg är ett verktyg för att såga olika

material. För att använda verktyg tryggt och effektiv, måste vi lära oss använda dem. Så är det också med språkmodeller. Förstår vi att de är verktyg och förstår vi hur man kan använda dem kan de vara mycket effektiva. Använder vi verktyg fel kan de vara destruktiva. Så också med AI i form av språkmodeller.

AI system likt språkmodeller kan också vara tränade på annan data, såsom bilder, musik och video. Med en bildmodell, kan vi skapa AI-bilder, som tar intryck av den stora mängd klassificerade bilder som de stora sociala mediajättarna har tillgång till. Liksom med texter är det fascinerande att kunna skapa egna bilder, som är någon slags statistisk kopia av tidigare bildmaterial. Men någon ny **Picasso** eller **Rembrandt** skapas nog inte via dataingenjörens programkod.

En väsentlig skillnad på AI och den intelligens vi mänskor har är förmågan att avgöra vad som är bra. Det som är bra är olika i olika kulturer, olika demografier och olika världsdelar. Våghalsiga stunts kan av ungdomar anses häftiga och bra, medan den äldre generationen tycker att det inte alls är bra. En text som skrivs på sociala media kan upplevas som bra i skrivandets stund, men samma text kan visa sig vara vinklad och bidra till att sprida fördomar efter att ha varit ute på nätet en stund.

Ett ytterligare problem är att fördomar kan förstärkas genom hur sociala media fungerar. Sociala media finansieras i första hand av reklam. Faktureringen av reklam sker via mått på visningar och engagemang av reklaminslag, och för att de skall visas i ett flöde krävs att användaren surfar vidare. För att surfa vidare skall plattformen visa sådant innehåll som på något sätt engagerar. Speciellt material som stöder egna fördomar lockar till att surfa vidare, vilket då också ger reklamintäkter. För mediajättarna är reklamintäkter bra. Därför blir det lätt så att sociala media stärker fördomar.

Det att maskiner ger oss möjlighet att skapa texter och bilder kallas allmänt generativ artificiell intelligens (GenAI). GenAI har vuxit enormt de senaste åren, och uppskattningar gällande hur mycket investeringar det kommer att göras kring detta rör sig i hundratals miljarder euro. Eftersom GenAI utförs i datacenter, är behovet av nya datacenter stort. I Finland finns det planer på investeringar i datacenter på 20 miljarder euro, med effektbehov flerfalt vårt största kraftverk Olkiluoto 3 på 1,6 GW.

Ett datacenter använder elenergi för att göra beräkningar. Elenergin driver miljardtals transistorer att utföra de återigen miljardtals logiska operationer som i slutändan för användaren ser ut som om maskiner har någon form av intelligens. Den elenergi som används per logisk operation är mycket liten, den är i storleksordningen pJ (picojoule, där joule är SI-enheten för energi), ett tal med 12 nollor efter decimalkommat. Som jämförelse en kilokalori som de som följer med sitt ätande känner till, en kilokalori motsvarar 4 kilojoule, med en kilokalori kan man därför utföra miljontals miljarder logiska operationer i en dator. Som jämförelse krävs det 20 kilokalorier för att värma upp en kopp té. För att skapa texter och bilder krävs enorma mängder av logiska operationer, och därför använder datacenter mycket energi. Idag använder datacenter nästan 2 procent av elenergin globalt.

All den energi som beräkningarna utför omvandlas i de elektriska kretsarna till värme. Samma mängd elenergi som matas in i ett datacenter måste föras ut som en värmeström från datacentret. Det traditionella sättet är att installera stora fläktar som bara blåser ut värmen till omgivningen. Därför har det varit fördelaktigt att bygga datacenter i kalla klimat, då behövs det mindre energi till att fläkta ut värmen ur datacenter. De nordiska länderna har aktivt marknadsfört sig som bra ställen för datacenter. Vi har stabilt elnät, generellt billig elenergi, stabilt politiskt klimat, inga jordbävningar och möjlighet till billig kylning. I Finland har gamla pappersbruk varit bra ställen för att bygga datacenter. Pappersbruk har också använt sig av mycket elenergi, och har haft behov av vatten för själva pappersprocessen. Datacenter behöver det

samma, elenergi för att driva verksamheten, vatten fungerar utmärkt för kylningen. Städer som Fredrikshamn, Kotka och Kajana har dragit stor nytta av etablering av datacenter.

Den stora mängd elenergi som datacenter använder har förstås väckt uppmärksamhet. På EU-nivå har man satt krav på energieffektivitet i datacenter. Energieffektivitet i datacenter betyder att så mycket som möjligt av elenergin används till beräkningar, istället för att driva processer som kylning, belysning, luftväxling och annan infrastruktur. Termen Power Usage Effectiveness (PUE) definieras som total eleffekt i ett datacenter dividerat med eleffekt för beräkningar. En PUE nära 1 betyder alltså ett mycket energieffektivt datacenter. Googles datacenter i Kotka rapporteras till PUE=1.09, vilket är ett av de bästa i världen.

Samtidigt är det ineffektivt att mata ut stora mängder värme direkt i luften eller i vattendrag i ett land som Finland där uppvärmning behövs stora delar av året. De stora datacenteroperatörerna har redan flera år arbetat på detta och överloppsvärmen från flera datacenter i Finland är nu integrerade i det lokala fjärrvärmenätet. Men detta är inte ännu standard och kunskap gällande hur man skall integrera datacenter, värmeåtervinning, samt hur man skall styra datacenter för att öka energieffektiviteten, saknas ännu många gånger.

Åbo Akademi bygger tillsammans med Vasa universitet, Yrkeshögskolan Novia och Vasa yrkeshögskola ett forskningsdatacenter WSTAR, som forskar i hur datacenter skall integreras och styras för att öka energieffektiviteten i de datacenter som nu håller på att byggas upp. Forskningsdatacentret kommer att vara ett jämförelsevis mycket litet datacenter med en eleffekt på en bråkdel av de stora datacentren. Datacentret möjliggör forskningsexperiment där man ser hur framtidens energiintegration, borde göras för att bygga datacenter med förbättrad energieffektivitet. Detta datacenter kommer att invigas hösten 2025.

Artificiell intelligens upplever för tillfället en egen teknikbubbla, eventuellt överdrivna förväntningar på vad allt som skall hända och hur samhället förändras, men samtidigt kommer det som vid alla teknikbubblor att bli kvar en bestående förändring i hur vi lever på detta jordklot. AI kommer att bli en integrerad del av hur vi arbetar och hur vi lever, och datacenter kommer att behövas för att driva denna sak. Finland har med sin fossilfria elenergi, behov av användning av överloppsvärme och goda ingenjörskunskaper en utmärkt möjlighet att bli en mycket relevant plats för att bedriva den dator i Finland beräkningsverksamhet som hela jordens befolkning kommer att använda i sitt dagliga liv. Här i Finland har vi både möjligheterna och viljan att göra det miljömässigt mycket rätt.

Jerker Björkqvist är professor i datateknik vid Åbo Akademi, med inriktning på digitalisering av industrin. Speciellt intresse är hur man kan använda datadrivna metoder för att styra både industriprocesser och samhälle för att uppnå högre energieffektivitet och hållbar användning av naturresurser. Han är medlem av Svenska Tekniska Vetenskapsakademin i Finland och jobbar aktivt med att främja teknikintresse bland barn och ungdomar.