



AI/ML är vägen till energieffektiv Edge IoT

AI och IoT är två av de starkaste tekniktrenderna idag. Även om de är olika, överlappar de varandra allt oftare vid kanten av näten. AI förbättrar prestanda och ger nya funktioner samtidigt som den introducerar smartare mekanismer för att hantera och till och med minska energiförbrukningen.

Tidigare utvecklades teknik i allmänhet i en förväntad takt. Om ingenjörer till exempel behövde en viss mängd processorkraft till en viss prisnivå för att åstadkomma en tänkt produkt, kunde de konsultera Moores lag och avgöra – ibland ner till ett kvartal – när deras idéer sannolikt skulle bli genomförbara.

Med hastigheten på AI och Edge IoT-utvecklingen har allt förändrats. Utvecklare av IoT-enheter för kanten kan använda AI och ekonomiskt åstadkomma mycket mer än vad många av dem inser. Kraftfulla verktyg och tekniker för att konstruera avancerade AI-drivna Edge IoT-enheter ger ett överflöd av marknadsmöjligheter för IoT-utvecklare.

Nätverkets kant är abstrakt och dess begreppsmässiga plats har flyttats flera gånger under det senaste decenniet. För bara några år sedan ansågs ett datacenter, som ägdes av ett telekomföretag, vara en del av kanten eftersom det tekniskt sett inte var en del av molnet. Inte långt därefter ansågs lokala serverrack och routrar installerade av små och medelstora företag vara nätverkets kant.

Nu definieras den av betydligt mindre enheter som samlar in och bearbetar data på

platser som är fysiskt avlägsna från de största datacentrena. Dessa enheter kan vara allt från övervakningskameror och uppkopplade apparater till bärbara enheter och individuella sensorer – kort sagt saker som definierar IoT. I takt med att IoT har expanderat har kanten förflyttats, vare sig det är smarta städer, smarta hem, jordbruksfält eller till och med oljeplattformar till sjöss.

Nästan varje enhet i denna breda miljö har dubbla utmaningar. Det första är extremt begränsade resurser när det gäller vikt, storlek, pris och driftskostnad (inklusive energiförbrukning). Dessutom omfattar IoT-kanten en uppsättning hårt konkurrensutsatta marknader där mindre, billigare och mer ekonomiska alternativ alltid finns runt hörnet.

AI och Edge IoT kombineras

AI har blivit allt vanligare vid nätverkskanten av flera anledningar. Till att börja med är användning av AI ett av de bästa sätten att få mindre, billigare och mer ekonomiska enheter. Samtidigt rör sig affärslogiken mot datakällan av ekonomiska och säkerhetsmässiga skäl och användning av AI främjar båda målen.

En annan viktig drivkraft är energieffektivitet. Majoriteten av alla IoT-enheter är trådlöst anslutna till nätverket och ur en IoT-enhets synvinkel förbrukar trådlös kommunikation en betydande mängd energi. Med de senaste processorerna med extremt låg strömförbrukning och AI kan det vara mer energieffektivt för en medvetet utformad

Av Tamas Daranyi, Silicon Labs

Tamas Daranyi arbetar med produktmarknadsföring och är ansvarig för AI/ML-initiativet. Han har mer än tio års erfarenhet av IoT-noder och molntjänster.



Edge-enhet (till exempel utrustad med hårdvaruacceleratorer för maskininlärning) att bearbeta data lokalt än att ladda upp rådata via radio för vidare bearbetning.

Detta är också direkt kopplat till energiförbrukningen. Många Edge-enheter behöver drivas med minimala mängder energi. En viktig energisparteknik är att sätta enheten i viloläge där den bara drar minsta mängd energi. Enheter i viloläge måste vara redo att vakna och lika viktigt – att bara vakna när det verkligen är motiverat.

Att avgöra om en ingång motiverar att hela enheten tas ur viloläge blir mycket effektivare genom att använda AI. I vissa fall kan AI till och med göra det möjligt att implementera ett effektivt viloläge redan från början.

I sådana fall kan det genom att lägga till AI i IoT-enheten göra skillnaden mellan att använda ett stort batteri som gör produkten för otymplig för att vara verkligt bärbar och att använda ett litet, diskret knappbatteri som gör enheten bekvämt bärbar. Omvänt kan det vara den variabel som gör att ett batteri bara räcker i några dagar istället för att samma batteri fungerar i veckor eller längre.

AI kan ge prestandaförbättringar och nya funktioner till en produkt som ibland var användbar till att göra något som verkligen är viktigt. Till exempel gör AI i kombination med framsteg inom sensorteknik att det går att utveckla bärbara hälsoenheter som mäter ett fåtal vitala parametrar för att indikera allmän hälsa till enheter som ger tillförlitlig data som sjukvårdspersonal kan använda för diagnostik och behandlingar.

AI-aktiverad Edge IoT:s maskininlärningsutveckling

Modeller för Edge AI kan vara relativt enkla, särskilt jämfört med de typer av stora komplexa språkmodeller (LLM) som chattverktyg baseras på. Det är ett sätt att minska bear-



AI/ML är särskilt användbart vid kanten eftersom majoriteten av data som samlas in där inte är användbar någon annanstans och att det därför inte finns någon anledning att föra den vidare.



betningsbördan för AI i Edge IoT-applikationer. Det finns andra involverade i skapandet av smarta algoritmer men poängen är att AI enkelt kan skalas ner för att köras på mindre kraftfulla processorer som är lämpligt dimensionerade och prissatta för Edge-bearbetning. Dessa tenderar att vara styrkretsar som körs på lägre frekvenser och med bara megabyte minne, ibland bara kilobyte.

På ungefär samma sätt som modeller kan ha rätt storlek för Edge AI i allmänhet kan ML också ha rätt storlek. Det finns en branschsträngning för att kunna köra ML på samma små processorer. Ursprungligen kallat TinyML, nu kallas det vanligtvis EdgeAI – ytterligare en indikation på att AI och Edge IoT inte bara överlappar varandra utan även smälter samman.

AI/ML är särskilt användbart vid kanten eftersom den stora majoriteten av data som samlas in där inte är användbar någon annanstans och att det därför inte finns någon anledning att föra den vidare. Det finns flera skäl att aktivt undvika att göra det:

- Överföring av rådata förbrukar energi
- Användning av nätverk har vanligtvis en kostnad
- Bandbredden är alltid begränsad och överföring av rådata förbrukar bandbredd som kunde ha använts till viktigare trafik
- Databehandling i datacentraler utförs av bearbetningssystem som förbrukar mycket mer energi än IoT-processorer vid kanten
- Returöverföring av de bearbetade resultaten förbrukar energi och bandbredd
- Tur- och returöverföringsprocessen medför nätverkslatens; den introducerar en tidsfördröjning mellan inmatning och resultat

DETTA INNEBÄR att bearbetning av data som samlas in vid kanten är en fråga om affärslogik, och affärslogiken rör sig mot kanten

eftersom AI/ML-bearbetning vid kanten producerar de nödvändiga resultaten samtidigt som den förbrukar minsta möjliga resurser (bandbredd, energi, pengar, tid) och är mer effektiv.

Om någon ber Siri att tända lampor bör svaret vara omedelbart; nätverkslatens är oacceptabelt.

Svarstider kan snabbas upp genom att undvika nätverkslatens, men nätverksavbrott förekommer; lokal bearbetning innebär att ett svar fortfarande kommer att komma.

En sista faktor är att lokal bearbetning är säkrare eftersom det minskar risken för att information avlyssnas jämfört med att skicka den genom flera överföringskanaler till och från en datacentral. Myndigheter på flera stora marknader har infört rättsliga skyldigheter för att skydda personuppgifter och därför har det blivit en princip inom data-skydd och integritet att säkerställa att lokala data förblir lokala.

AI Edge IoT till, från och vänteläge

AI/ML spelar en roll för att minska energiförbrukningen vilket har viktiga konsekvenser för batterier och batteritid.

Det finns flera sätt att minska energiförbrukningen. Edge IoT-enheter tenderar att vara anslutna trådlöst och kommunicera med hjälp av en mängd olika protokoll däribland Wi-Fi och Bluetooth, men vanligare är alternativ som är utformade för att använda mindre energi, som Bluetooth Low Energy (BLE), Zigbee och Thread. Men även när man använder dessa kostar varje överföring fortfarande energi – ganska mycket för en batteridrivna enhet.

Ta exemplet med säkerhetsenheter som lyssnar efter glaskrossning. Fram till för några år sedan var dessa enheter tvungna att alltid vara påslagna och redo att utvärdera varje ljud för att avgöra om det verkligen var

glaskrossning eller något annat. Falska positiva resultat tenderade att inkludera andra skarpa ljud som handklappningar. Dessa enheter förbrukade sedan ännu mer energi och varnade i onödan för irrelevanta ljud.

AI/ML eliminerar detta problem. Med TinyML (EdgeAI) utför enheten lokal bearbetning och lär sig att upptäcka skillnaden mellan krossat glas och andra ljud men också ljud som delar några av samma ljudegenskaper som krossat glas men som inte är av intresse. Först när den fastställer ett krossat glasljud initierar den processen att slå på resten av enheten, inklusive den relativt strömslukande trådlösa radion, för att utlösa ett larm.

Detta är fullt möjligt även utan AI/ML – om enheten har en tillräckligt kraftfull processor. Den kommer dock sannolikt att kosta alldeles för mycket för en Edge IoT-enhet. Med AI/ML är denna funktion möjlig på en processor som har en lämplig nivå för en Edge IoT-produkt.

Eftersom dessa enheter förväntas användas i tre till fem år på ett enda batteri är dessa energibesparande processer omvälvande.

Bara början

AI/ML gör det möjligt att göra extraordinära saker på Edge IoT-enheter idag på ett ekonomiskt sätt med hjälp av verktyg som gör konstruktionsarbetet mycket enklare än man kan förvänta sig.

Dessutom utvecklas AI/ML inom Edge IoT fortfarande. Kontinuerlig kretsintegration innebär att även de blygsamma processorer som används på denna nivå i dag kommer att bli kraftfullare.

Optimerade algoritmer och ökad Edgeberäkningskraft kommer säkerligen att leda till betydande framsteg i användarupplevelsen på lokala enheter. TinyML (EdgeAI) är på väg att föra ner konversations-AI-funktioner till Edge-enheter. ■