



NanoEdge AI Studio:

Maskininlärning för edgetillämpningar



Större företag som vill dra nytta av maskininlärning behöver idag anställa en expert på data science (analys av big data) om de vill samla in stora datamängder. Arbetet kan ta månader och inkluderar städning av data och att skapa AI-modeller.

Därefter anlitas inbyggdprogrammerare för att portera implementationerna till styrkretsar eller med hjälp av särskilda verktyg konvertera de neurala nätverken till optimerad kod för styrkretsen.

För ett företag med snäv budget kan det vara omöjligt att vara sig anställa en dataanalytiker eller att lägga ut arbetet på en underkonsult.

ÄVEN MED RÄTT MÄNNISKOR och all tid i världen är det fortfarande ett problem att få tag på data av god kvalitet att träna systemet med. Tag som exempel ett system som ska upptäcka ett avvikande beteende – men så saknas detta beteende i träningsdata. Medan många datauppsättningar är användbara för klassificeringsproblemet att



Av Joël Rubino, ST Microelectronics

Joël Rubino började sin karriär på IBM i Nice och arbetade en tid på företagets kontor i New York. År 2012 lämnade han IBM för att startade Apicube som specialiserade sig på big data och analys av sociala media med inriktning på marknadsföring. 2016 var han med och startade Cartesiam med fokus på AI i inbyggdprogrammering. Företaget köptes av ST Microelectronics i juli 2021. Idag är han ansvarig för Edge AI på ST Microelectronics.

upptäcka en avvikelse, är de värdelösa på att upptäcka nya situationer.

Det är också viktigt att ha kvalitet i data, en aspekt som är svår att verifiera. Även om samplen inte har stavfel eller informationsluckor kan det krävas betydande investeringar för att samla in användbara data och sätta konsekventa etiketter på dem.

NanoEdge AI Studio är ett verktyg för utvecklare av inbyggda system som saknar djup kompetens i data science. Magin ligger i att det på en och samma enhet dels går att köra träningsprocessen för att få fram viktorna och dels beräkna inferenser i den

färdiga modellen.

Hela processen kan således utföras på samma ST-processor. Dessutom är interaktionen för slutanvändaren enkel, det kan räcka med att trycka på en knapp.

Resultatet är att utvecklarna kan skraddarsy sina system för sin lokala tillämpning vilket gör dem robustare och enklare att installera.

NanoEdge AI Studio kan köras på Windows 10 eller Ubuntu och är det bästa sättet att dels processa data och dels att hitta relevanta AI-bibliotek. Verktyget är utformat för att passa inbyggda system och för att enkelt kunna integreras med C-kod.

NanoEdge AI Studio väger in vilka komponenter som används – processor, minne och sensorer – och letar reda på ett lämpligt NanoEdge AI-bibliotek som körs på processorfamiljen STM32 och som utvecklaren kan integrera i sin egen inbyggnadstillämpning.

Den senaste versionen har fått fler algoritmer för maskininlärning. Förutom att upptäcka anomalier och att göra klassificering innehåller det dessutom två nya algoritmfamiljer: extrapolation och ytterlighetsfall (outliers).

Det förra underlättar uppgiften att förutspå beteenden under otestade förhållanden. Det kallas regression och handlar om att hitta kopplingen mellan olika variabler i data. Exempelvis kan en uppsättning data avspegla en fläkts egenskaper vid 100°C, 110°C och 150°C. Tack vare regressionsalgoritmen kan maskininlärningsalgoritmen extrapolera vad som händer vid 160°C.

Extrapoleringsalgoritmen i NanoEdge AI Studio täcker inte enbart linjär regression. Den erbjuder även mer avancerad analysteknik för att tackla komplexa situationer. Som ett resultat av detta kan utvecklarna nu skapa nya tillämpningar som övervakar saker som en dataanalytiker inte klarar att testa på egen hand.

Den andra algoritmen upptäcker utanföriggare (outliers) och baseras på en enda klass av värden. Systemet lär sig de normala beteendena. Allt som därefter avviker från det räknas som anomalier.

Tidigare versioner använde istället metoden att registrera både normalt beteende och beteendet vid ett eller flera simulerade problem eller störningar. Som sagt var det i princip möjligt att lära sig alla beteenden på samma processor vilket förenklade arbetet. I vissa fall var det dock omöjligt att åter skapa anomalierna. I dessa fall går det nu att använda normala data för att härleda att en anomali föreligger.

Även om det inte existerar bättre data än de som samlas in under verkliga förhållanden går det ibland tyvärr inte att få tag på sådana. Eller så finns det inte tid att få tag på tillräckligt många datapunkter.

En ny datainspelningsfunktion förvandlar



alla industriella och trådlösa STWIN-noder till enkla verktyg för att samla in data. Användarna ansluter kortet till sin dator och använder NanoEdge AI Studio för att växla över till datainsamling, som därefter kan göras per automatik.

Utvecklarna kan ansluta STWIN-kortet till den utrustning de vill övervaka. Sensorerna spelar in data som utvecklarna kan namnge och studera för att skapa mer noggranna tillämpningar.

En annan förbättring i den nya versionen är användargränssnittet. När det tillkom nya algoritmer och metoder för att samla in data, blev det nödvändigt att förbättra användarupplevelsen. Det var dessutom nödvändigt att optimera utvecklarnas arbetsgång.

Verktaget är avsett för utvecklare som vill göra maskininlärning i Edge-tillämpningar. Biblioteken är små, ned till 1 kbyte, och högt optimerade. Det var därför nödvändigt att även förbättra tillgången till algoritmerna för att garantera att det är enkelt för utvecklarna att välja projektkategori för att snabbt kunna generera bibliotek.

Innan NanoEdge AI Studio fanns var utvecklare tvungna att kontakta mjukvaruleverantörer, undersöka sin hårdvarukonfiguration och gå igenom de beteenden de ville övervaka. Nu har utvecklare fått ett verktyg som låter dem skräddarsy, generera och validera ett eget maskininlärningsbibliotek.

Man börjar med att välja en krets i Cortex-M-familjen och bestämma vilka sensorer som behövs. Sen importerar man en fil med värden som beskriver systemets typiska beteende. Det kan vara data från en accelerometer som sitter på en fläkt eller elektriska data från industriutrustning. Därefter testas, optimeras och sorterar NanoEdge AI Studio automatiskt de bästa kombinationerna av algoritmer bland de hundratal miljoner som finns och levererar ett skräddarsytt bibliotek som utvecklarna kan validera med hjälp av inbyggnadsemulatorn.

NanoEdge AI Studio v3 stödjer alla ST:s utvecklingskort direkt från användargränssnittet. Tillgången till optimerade och gratis bibliotek betyder att det är enkelt att testa ett koncept.

Ta exemplet med den smarta vibrations-sensorn. Där kan användarna välja kortet NUCLEO-L432KC för att fånga fläkts normala beteende. Sedan kan de mata in data i NanoEdge AI Studio för att skapa ett bibliotek som de kan anropa i huvudloopen för att köra ett minimalt antal träningscykler, tidigare definierat i prestandamått i den nya mjukvaran, innan de går över till att göra inferenser.

På så sätt kan NanoEdge AI-bibliotek hjälpa till att snabbt skapa applikationer för bland annat förebyggande underhåll och smart säkerhet. ■

