



LUNDS UNIVERSITET
Lunds Tekniska Högskola

Kursplan för

Minnesteknologi för maskininlärning Memory Technology for Machine Learning

EITP25, 7,5 högskolepoäng, A (Avancerad nivå)

Gäller för: Läsåret 2023/24

Fakultet: Lunds tekniska högskola

Beslutad av: Programledning C/D

Beslutsdatum: 2023-04-18

Allmänna uppgifter

Huvudområde: Nanovetenskap.

Valfri för: E4-is, F4, F4-fel, MNAV1, MSOC2, N4-hn

Undervisningsspråk: Kursen ges på engelska

Syfte

Målet med denna kurs är att ge djup förståelse för fysiken bakom de vanligaste minneskomponentteknologierna med stort fokus på icke-volatila minnen. Vidare behandlas hur dessa kan integreras för att skapa neuromorf hårdvara för tillämpning inom maskininlärning och artificiell intelligens. Slutligen ger kursen även en introduktion till de arkitekturer och algoritmer som används inom maskininlärning för att ge en grundläggande förståelse för vilka behov minneskomponenter och dess kopplingar behöver uppfylla för denna tillämpning.

Mål

Kunskap och förståelse

För godkänd kurs skall studenten

- förklara översiktligt hur minneshierarkin i en modern dator ser ut och i detalj hur dess minneskomponenter fungerar.
- förstå de fysiska processer som ligger bakom funktionen i vanliga icke-volatila minnestyper som RRAM, PCM, FeRAM och STT-MRAM.
- förstå generellt vilka begränsningar och fördelar som de olika minnestyperna som behandlas i kursen har.
- förstå generellt hur integrationen av minneskomponenter till neurala nätverkskretsar kan göras, och fördelar och begränsningar med dessa tillvägagångssätt.

- förklara i detalj hur inlärning sker i ett spikande neuralt nätverk, samt hur memristorer kan användas i sådana.
- förklara generellt hur träning genomförs i neurala nätverk med hjälp av bakåtpropagering och gradiell nedstigning.

Färdighet och förmåga

För godkänd kurs skall studenten

- utföra mätningar och analys av ström-spänningskaraktistiken från en RRAM komponent.
- mäta upp samt utifrån ett uppmätt polarisation-fält-diagram kunna extrahera viktiga parametrar för en ferroelektrisk kondensator.
- kunna ge förslag på hur hastighet, pålitlighet samt energiförbrukning kan förbättras i de olika minnestyperna som behandlas i kursen.
- designa och träna ett spikande neuralt nätverk för att utföra bildigenkänning.

Värderingsförmåga och förhållningssätt

För godkänd kurs skall studenten

- inse behovet av energieffektiv och skalbar neuromorf hårdvara för maskininlärning och AI.
- utvärdera lämpligheten för en given minnesteknologi för diverse användningsområden utifrån dess fördelar och nackdelar.

Kursinnehåll

Minneskomponenter i datorn: SRAM, DRAM, NAND

Icke-volatila minneskomponenter: Memristorn, resistiva minnen (RRAM), fasändringsminnen (PCM), ferroelektriska minnen (FeRAM), magnetiska minnen (MRAM).

Integration av minneskomponenter: 3D stackning för skalbarhet, crossbar-arkitekturen.

Neurala nätverksarkitekturer: Fullt kopplade neurala nätverk, konvolutionella nätverk, rekursiva nätverk, spikande nätverk.

Maskininlärningsmetoder: Bakåtpropagering, gradiell nedstigning, STDP.

Kursens examination

Betygsskala: TH - (U,3,4,5) - (Underkänd, Tre, Fyra, Fem)

Prestationsbedömning: Skriftlig examen, projektarbete, laborationer med rapport.

Om så krävs för att en student med varaktig funktionsnedsättning ska ges ett likvärdigt examinationsalternativ jämfört med en student utan funktionsnedsättning, så kan examinator efter samråd med universitetets avdelning för pedagogiskt stöd fatta beslut om alternativ examinationsform för berörd student.

Delmoment

Kod: 0119. **Benämning:** Skriftlig tentamen.

Antal högskolepoäng: 4. **Betygsskala:** TH. **Prestationsbedömning:** Prov Delmomentet omfattar: Skriftlig examen vars omfattning ska täcka alla kursens ämnen.

Kod: 0219. **Benämning:** Laboration.

Antal högskolepoäng: 1. **Betygsskala:** UG. **Prestationsbedömning:** Godkänt moment vid närvaro vid laborationen samt godkänd labbrapport. **Delmomentet omfattar:** Praktisk laboration med efterföljande labbrapport. **Övrig information:** Labbarbete och rapportskrivande kan genomföras självständigt eller i grupper om två personer.

Kod: 0319. **Benämning:** Projektuppgift.

Antal högskolepoäng: 2,5. **Betygsskala:** UG. **Prestationsbedömning:** Förståelsenivån i den skriftliga rapporten, samt kontroll av programkod avgör huruvida godkänt resultat har erhållits. **Delmomentet omfattar:** Projektet består i att designa samt träna upp ett neuralt nätverk för att kunna känna igen över 90% av bilderna från en fördefinierad databas. Projektet ska utföras enskilt eller i par. Projektet ska redovisas med en skriftlig rapport och fungerande MATLAB-kod.

Antagningsuppgifter

Förutsatta förkunskaper: Grundläggande kunskaper i fasta tillståndets fysik.

Begränsat antal platser: Nej

Kurslitteratur

- An Chen, James Hutchby, Victor Zhirnov, George Bourianoff: Emerging Nanoelectronic Devices. Wiley, 2014, ISBN: 9781118958254.

Kontaktinfo och övrigt

Kursansvarig: Mattias Borg, mattias.borg@eit.lth.se

Hemsida: <http://www.eit.lth.se/kurs/eitp25>

Övrig information: Kursen använder sig av Canvas för kommunikation, inlämning samt studiematerial.