



IK2221 Nätverkssystem för maskininlärning 7,5 hp

Networked Systems for Machine Learning

Fastställande

Betygsskala

A, B, C, D, E, FX, F

Utbildningsnivå

Avancerad nivå

Huvudområden

Datalogi och datateknik

Särskild behörighet

Kunskaper och färdigheter i programmering, 6 hp, motsvarande slutförd kurs DD1337/DD1310-DD1319/DD1321/DD1331/DD1333/DD100N/ID1018/ID1022.

Kunskaper inom minst ett av följande områden i syfte att kunna genomföra kursens fördjupningsuppgift:

- datornätverk, 6 hp, motsvarande slutförd kurs EP1100/IK1203/IK2218/EP2120
- distribuerade system, 7,5 hp, motsvarande slutförd kurs ID2201
- operativsystem, 6 hp, motsvarande slutförd kurs ID1200/ID1206
- realtidssystem, 6 hp, motsvarande slutförd kurs IL2235.

Lärandemål

Efter godkänd kurs ska studenten kunna

- beskriva komponenterna i virtualiserade nätverksfunktioner och hur de samarbetar
- konfigurera ett exempel på en tjänst med användning av virtualisering av nätverksfunktioner (NFV)
- förklara i detalj de viktiga fördelarna med direkt fjärrminnesåtkomst (RDMA) och hur det fungerar
- analysera metoder för in- och utmatning i nätverksdirekt till/från grafikprocessorer (GPU)
- i detalj förklara metoder för att förbättra slutledningsfördröjning
- beskriva lastbalanseringens roll för servrar
- ge exempel på och beskriva aktuella forskningsproblem inom betjäning av maskininlärningsberäkningar i datacenter.

Kursinnehåll

Nätverksfunktioner. Virtualisering. Kernel bypass-tekniker (t.ex. DPDK) för nätverk med över 100 gigabit per sekund. Avlastning till Smart Network Interface Cards (Smart-NIC). Snabbt nätverk med liten eller ingen CPU-inblandning med hjälp av direkt fjärrminnesåtkomst (RDMA). Nätverksaspekter av maskininlärningslutledning med hjälp av grafiska processorer (GPU). Lastuppskattning och belastningsutjämnning. Begäran om utskick och schemaläggning. Effektiv, storskalig maskininlärningslutledning. Slutledning med stora språkmodeller (LLM).

Examination

- PRO1 - Projektuppgifter, 2,5 hp, betygsskala: A, B, C, D, E, FX, F
- SEM1 - Rapportsammanfattningar, 2,5 hp, betygsskala: A, B, C, D, E, FX, F
- TEN1 - Skriftlig tentamen, 2,5 hp, betygsskala: A, B, C, D, E, FX, F

Examinator beslutar, baserat på rekommendation från KTH:s handläggare av stöd till studenter med funktionsnedsättning, om eventuell anpassad examination för studenter med dokumenterad, varaktig funktionsnedsättning.

Examinator får medge annan examinationsform vid omexamination av enstaka studenter.

När kurs inte längre ges har student möjlighet att examineras under ytterligare två läsår.

Etiskt förhållningssätt

- Vid grupparbete har alla i gruppen ansvar för gruppens arbete.
- Vid examination ska varje student ärligt redovisa hjälp som erhållits och källor som använts.

- Vid muntlig examination ska varje student kunna redogöra för hela uppgiften och hela lösningen.