# Kunskapskontroll: Teorifrågor

1. Hur är AI, Maskininlärning och Deep Learning relaterat?

AI är en metod eller ett koncept där datorprogramvara efterliknar mänsklig eller annan levande intelligens och beteendemönster.

Maskininlärning är en teknik inom AI där datorer lär sig från data istället för att man manuellt skulle behöva skriva en oerhörd mängd regler. T ex skapa en modell som kan upptäcka och kategorisera om epost är skräppost eller ej.

Deep Learning en del av maskininlärning där man använder neurala nätverk för att bygga modeller, inspererat av hur hjärnan fungerar. Många neurala nätverk skapar djupa neurala nätverk (deep learning). Deep learning är särskilt användbart för tex bild- och röstigenkänning.

1. Hur är Tensorflow och Keras relaterat?

Tensorflow är ett library och ett ramverk som kan användas för att bygga och träna modeller inom deep learning. Keras är ett enkelt och användarvänligt API som kan användas på/med Tensorflow, som alltså utför beräkningarna. Om man ska likna det med en bil är Tensorflow motorn, medan Keras är ratten som styr bilen.

1. Vad är en parameter? Vad är en hyperparameter?

Parametrarna i deep learning eller neurala nätverk är vikterna och biaserna som lärs under träningen utefter input data. Hyperparametrar styr hur modellen tränas, t ex antal neuroner i hidden layers, eller antalet många hidden layers. Hyperparametrarna bestäms manuellt. De kan dock automatiseras via exempelvis KerasTuner eller Scikit-learn wrapper.

1. När man skall göra modellval och modellutvärdering så kan man använda ett tränings, validerings och test data. Förklara hur de olika delarna kan användas.

Träningsdatan utgör majoriteten av datan och är vad vi använder för att träna vår modell på. Valideringsdatan används för att utvärdera hur väl modellen presterar för att sedan kunna finustera modellen så att den presterar bättre, genom att t ex justera hyperparametrarna, lägga till eller ta bort hidden layers, använda regulariseringstekniken etc. Testdatan används i slutskedet när vi är färdiga med våra finjusteringar av modellen och återigen utvärderar hur väl modellen presterar på ny data.

1. En bild som visar text, Teckensnitt, skärmbild

   Automatiskt genererad beskrivningFörklara vad nedanstående kod gör:

* Antal kolumner i x-datan sparas i variabeln n\_cols som sedan kommer användas till storleken på input-lagret på modellen
* En Sequential-modell initieras under namnet nn\_model
* Ett Dense-lager (ett lager där alla neuroner är kopplade till varje neuron i föregående lager) läggs till som har 100 neuroner och aktiveringsfunktionen ReLu. Här specificerar vi input shape = antalet kolumner i x-datan som vi sparade i n\_cols
* Ett Dropout lager läggs till med dropoutrate på 0.2, vilket innebär att 20 % av neuronerna slumpmässigt kommer droppas.
* Ett Dense lager läggs till med 50 neuroner och aktiveringsfunktonen ReLu
* Ett sista Dense lager läggs till med en neuron som output layer och Sigmoid som aktiveringsfunktion. Eftersom det bara är en neuron i output layer innebär det att modellen hanterar ett binärt klassificeringsproblem (två utfall), därför Sigmoid är lämplig som aktiveringsfuntion då output blir ett värde mellan 0 och 1.
* Vi kompilerar modellen med Adam som optimerare, binary crossentropy som loss function och accuracy (hur stor andel rätta prediktioner modellen gör) som metrics.
* Vi ”fit” modellen med träningsdatan med 20 % som används som valideringsdata. Vi sätter att vi tränar 100 epoker, med early stopping med patience 5. Med andra ord kommer vi gå igenom hela data setet med 20 % som valideringsdata. Om vår loss function fortsätter minska kommer den göra detta 100 gånger, men om valideringsfelet slutar minska 5 epoker i rad stoppar träningen.

1. Vad är syftet med att regularisera en modell?

Syftet med regularisering är att förhindra overfitting genom att minska på antalet parametrar i modellen. Deep learning-modeller kan generera väldigt många parametrar som innebär att den blir väldigt flexibel på träningsdatan men sämre på att generalisera på ny data.

1. ”Dropout” är en regulariseringsteknik, vad är det för något?

Dropout är en regulariseringsteknik där man ”droppar” neuroner(noder) i input layer och/eller hidden layers. Man sätter en dropout rate *p* (vanligtvis mellan 10-50 % ) som i sin tur ”droppar” (”stänger av”) slumpmässigt *p* procent neuroner i lagret som man applicerat dropout. Dropout-tekniken förhindrar co-adaptations, alltså att neuroner ”samarbetar” mellan lagerna, och gör att neuronerna måste förlita mer på sig själva under träningen.

1. ”Early stopping” är en regulariseringsteknik, vad är det för något?

Early stopping innebär att vi stoppar träningen av vår modell när valideringsfelet slutar minska efter *x* antal epoker (som vi själva väljer genom att sätta *patience*). Om vi till exempel väljer patience = 3 kommer modellen gå igenom hela datasetet och mäta valideringsfelet. När valideringsfelet inte minskar 3 epoker i rad stoppar träningen.

1. Din kollega frågar dig vilken typ av neuralt nätverk som är populärt för bildanalys, vad svarar du?

CNN (Convolutional Neural Network)

1. Förklara översiktligt hur ett ”Convolutional Neural Network” fungerar.

I en CNN-modell som används för t ex bildigenkänning identifierar modellen först ”low level features” som kurvor och kanter som sedan kombineras för att identifiera high level features som ögon, näsa, arm osv. Detta gör den genom convolutional layers som är som ett filter som fokuserar på olika features (t ex horisontella feautures eller vertikala feautures), samt pooling som fokuserar på de viktigaste delarna i dessa filter. Baserat på dessa features predikterar modellen vilken klass bilden tillhör, med andra ord - vad är det en bild på?

1. Din vän har ett album med 100 olika bilder som innehåller t.ex. tennisbollar och zebror. Hur hade han/hon kunnat klassificera de bilderna trots att han/hon inte har någon mer data att träna en modell på?

Om man har ett litet dataset att använda kan man istället använda sig av en förtränad modell som redan tränats på ett stort dataset men istället kan finjusteras för hans/hennes syfte.

1. En bild som visar text, Teckensnitt, skärmbild, linje

   Automatiskt genererad beskrivningVad gör nedanstående kod?

Koden sparar den tidigare skapade och tränade modellen ’model\_file’ i h5-filformat, för att kunna användas senare. Den andra raden kod laddar modellen för att kunna användas direkt istället för att behöva bygga och träna modellen på nytt, vilket är väldigt tidskrävande.

1. Deep Learning modeller kan ta lång tid att träna, då kan GPU via t.ex. Google Colab skynda på träningen avsevärt. Läs följande artikel: <https://blog.purestorage.com/purely-informational/cpu-vs-gpu-for-machine-learning/> och skriv mycket kortfattat vad CPU och GPU är.

CPU, central processing unit, är en datorprocessor som hanterar datorns grundläggande uppgifter och processer, kan liknas vid datorns hjärna.

GPU, graphics processing unit, är en datorprocessor som ursprungligen skapats för grafikbearbetning av grafik såsom 2D och 3D-bilder, videos och animationer. GPUs använder till skillnad från CPUs ”parallell processing”, vilket gör att den arbetar snabbare på mer specialiserade datoruppgifter.

CPU hanterar sekventiella uppgifter bäst och snabbast, medan GPUs kan hantera parallella processer samtidigt, vilket gör att GPUs är väldigt användbart inom deep learning då det kan skynda på träningen av modeller avsevärt