

Kunskapskontroll 2

DS24 Deep Learning

Geisol Urbina

Del 1: Teoriska frågor

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

AI handlar om att skapa system som kan utföra uppgifter som normalt kräver mänsklig intelligens. Maskininlärning är en metod inom AI där system lär sig från data snarare än att programmeras explicit, det vill säga att systemen använder algoritmer för att upptäcka mönster i data och fatta beslut baserat på det. Deep Learning är en typ av maskininlärning som använder neuronala nätverk med många lager. Det är en avancerad teknik inom maskininlärning och en metod för att realisera AI

2. Hur är Tensorflow och Keras relaterat?

Tensorflow tillhandahåller den grundläggande motorn för beteckningar, medan Keras lägger ett användarvänligt till ovanpå, som standardiserar och snabbar upp utveckling. Tillsammans gör både kraftfull och tillgänglig djupinlärning. Keras är inbyggt i TensorFlow som dess officiella högnivå-API, vilket gör dem till ett sammanhängande verktyg.

3. Vad är en parameter? Vad är en hyperparameter?

En parameter är en del av modellen som lärs under träning. Till exempel: vikterna och bias-värdena i ett neuronnät. Dessa justeras automatiskt under träning för att minimera fel. Medan en hyperparameter är en inställning som bestäms före träning och styr hur modellen lär sig. Till exempel: antal lager, inlärningstakt, osv. Dessa väljs manuellt eller via optimering. På ett enkelt sätt kan man säga att hyperparametrar väljs av mig och parametrar lärs av datorn.

4. När man skall göra modellval och modellutvärdering kan man använda tränings-, validerings- och testdataset. Förklara hur de olika delarna kan användas.

Träningsdata används för att träna modellen och justera parametrar som vikter och bias. Det bör noteras att den inte får användas för att utvärdera modellens generalisering, eftersom det finns risk att modellen memorerar. Validering används för att justera hyperparametrar och kontrollera överanpassning. Testdata används endast en gång för att slutgiltigt utvärdera modellen efter träning och validering. Den simulerar hur modellen presterar på helt nya data.

5. Förklara vad nedanstående kod gör:

Koden bygger ett neuralt nätverk med Keras.

1. Bestämmer antalet kolumner i träningsdata där `x_train.shape[1]` hämtar antalet kolumner i träningsdata (`x_train`) och `n_cols` ska användas som input-dimension i det neurala nätverket.
2. `nn_model = Sequential ()` skapar en sequential modell.

3. Vi lägger till ett Dense lager med ReLU activation, där *Dense(100, ...)* visar att det finns ett anslutet lager med 100 neuroner, *activation='relu'* som introducerar icke linjäritet och *input_shape = (n_cols)* definiera input-dimensionen.
 4. Vi lägger till Dropout för regularisering, där *Dropout(rate=0.2)* inaktiverar slumpmässigt 20% av neuronerna under träningen för att förhindra överpassning.
 5. Ett Dense lager med ReLu, där finns ett anslutet lager med 50 neuroner och ReLu activation.
 6. Utdata-lager med Sigmoid-activation, där *Dense(1)* ska använda ett utdataneuron för binär klassificering, och *activation = 'sigmoid'* omvandlar utdata till ett värde mellan 0 och 1
 7. Vi kompilerar modellen med *compile()* funktion där vi använder Adam-optimizer som justerar inlärningstakten för varje parameter individuellt. Vi använder också binär korsentropi (*loss= binary_crossentropy*) som förlustfunktion, samt mäter träffsäkerhet under träning och validering (*accuracy*).
 8. Vi använder en *early stopping* för att förhindra överanpassning. Träningen stoppar om förlusten inte förbättras på 5 epoker i rad.
 9. Till sist tränar vi modellen.
- 6. Vad är syftet med att regularisera en modell?**
- Syftet med regularisering är att förhindra överanpassning, detta gör genom att begränsa modellens komplexitet, främja enklare modeller och genom att förbättra generaliseringsförmågan.
- 7. "Dropout" är en regulariseringsteknik, vad är det för något?**
- Dropout är en metod där man slumpmässigt stänger av neuroner under träningen. Detta görs i varje träningssteg för att förhindra att neuronerna blir beroende av varandra. Detta gör att varje neuron inaktiveras temporärt under träningsomgång och när träningen är klar används alla neuroner igen, men deras vikter har anpassats för att vara mer robusta.
- 8. "Early stopping" är en regulariseringsteknik, vad är det för något?**
- Early stopping innebär att man avbryter träningen i förtid när modellens prestanda på valideringsdata slutar förbättras. I stället för att köra alla epoker fortsätter träningen endast tills valideringsfelet börjar öka medan träningsfelet fortsätter minska, för att det är ett tecken på att modellen börjar överanpassa, varpå träningen avbryts.
- 9. Din kolla frågor din vilken typ av neuralt nätverk som är populärt för bildanalys, vad svarar du?**
- För bildanalys är konvolutionella neurala nätverk (CNN) det mest effektiva valet. De använder speciella lager som automatiskt lär sig känna igen enkla saker som kanter och färger, och sedan mer avancerade mönster som till exempel öron på en hund. Genom att använda små filter som fokuserar på delar av bilden i taget, i stället för att koppla allt till allt, sparar modellen mycket beräkningskraft och fungerar bra även med stora och detaljerade bilder. Dessutom är CNN robusta – de kan känna igen objekt även om de flyttas, roteras eller ändrar storlek, vilket gör dem mycket användbara i praktiken.

10. Förklara översiktligt hur ett "Convolutional Neural Network" fungerar.

Ett konvolutionella neurala nätverk (CNN) fungerar som en digital hjärna som lär sig känna igen mönster i bilder, steg för steg. Först använder det små filter som glider över bilden och letar efter enkla saker som kanter och färger. Varje filter skapar en karta som visar var just det mönstret finns. Sedan använder nätverket en funktion som kallas ReLU, som hjälper det att fokusera på det viktiga genom att ta bort onödiga (negativa) värden. Därefter förenklas bilden med något som kallas pooling, där bara det viktigaste i varje del bevaras, vilket gör modellen snabbare och mer robust. Till sist plattas all information ut och skickas till ett vanligt neuralt nätverk som bestämmer vad bilden föreställer, till exempel om det är en blomma eller en bil.

11. Vad gör nedanstående kod?

Koden handlar om att spara och ladda en maskininlärningsmodell med hjälp av Keras-biblioteket.

12. Deep Learning modeller kan ta lång tid att träna, då kan GPU via t. ex Google Colab skynda på träningen avsevärt. Skriv mycket kortfattat vad CPU och GPU är.

En CPU (Central Processing Unit) är datorns "hjärna" som hanterar många olika uppgifter, men inte alltid så snabbt med stora datamängder. En GPU (Graphics Processing Unit) är specialiserad på att göra många enkla beräkningar samtidigt, vilket gör den mycket snabbare för att träna deep learning-modeller.