Repetitions- och kontrollfrågor. Föreläsning Generalisering, val av modellstruktur.

1. Förklara med hjälp av ett exempel på vad som menas med att en modell har goda *generaliseringsegenskaper*.

2. Förklara utförligt vad som menas med "the Bias-Variance dilemma"?

3. Vad menas med *regularisering*?

4. Hur påverkas termerna $Bias^2$ respektive $Variance$ (vanligtvis) om man utför regulariserad parameterskattning? (Jämför med oregularised skattning)

5. Illstrera med ett matematiskt exempel hur man kan tolka regulariserad/icke regulariserad parameterskattning som MAP-skattning/ML-skattning.

6. Illustrera med figurer hur fenomenet överanpassning skulle kunna yttra sig i följande situationer (utnyttja figurer där så passar. Anta konsekvent att antalet parametrar är av samma storleksordning som antalet träningsdatapunkter.):

   (a) anpassning av ett polynom av högt gradtal.

   (b) träning av MLP med många neuroner för klassificering.

   (c) träning av MLP med många neuroner för funktionsapproximation.

   (d) klassificering av data från två normalfördelningar där vi först estimerar väntevärdesvektorer och kovariansmatriser med hjälp av träningsdata.

   (e) linjär regression med OLS-metoden.

7. Beskriv arbetsgången för hur man kan använda *validering* för att bestämma en lämplig modellstruktur (tex antal lager och antal neuroner i en MLP) givet en träningsdatamängd av begränsad storlek.

8. Två varianter av validering är hold out-metoden och korsvalidering. Beskriv hur dessa fungerar samt vilka fördelar respektive nackdelar metoderna har gentemot varandra.