

I denna uppgift ska huspriser förutsägas med hjälp av linjär regression.
Notera att detta inte är en inlämning och det kommer ej ges någon rättning. Däremot kan du fråga om hjälp, om du tänkt rätt samt dela ditt resultat.

Uppgiften är uppdelad i en godkänt del och en väl godkänt del.
De obligatoriska inlämningarna som kommer framöver kommer att vara uppdelade på liknande vis. Första delen kommer mestadels vara att implementera algoritmen och andra delen kommer handla om djupare förståelse.

Godkänt:

Använd datasetet *area_price.csv* för att förutsäga huspriser.
Följande punkter skall behandlas och besvaras

- Analysera datan
- Dela upp datan i träning och testset
- Träna en linjär regressionsmodell
- Vad får du för träningsscore?
- Vad får du för testscore?
- Varför skiljer sig träning och testscore?
- Rita ut linjen som modellen skapar

Väl Godkänt:

Använd datasetet *house_price.csv* för att förutsäga huspriser.
Datasetet innehåller attribut som är strängar. Dessa kan antingen kastas bort eller förvandlas till kategoriska värden.

Följande punkter skall behandlas och besvaras

- Analysera datan
- Dela upp datan i träning och testset
- Vad får du för träningsscore?
- Vad får du för testscore?
- Vilka attribut i datan påverkar modellen mest?
- Kan du plocka bort några attribut för att göra modellen bättre?
- Kan du hantera några extremvärden för att göra modellen bättre?