

Uppgiftsbeskrivning: Maskininlärning med Iris-datasetet

Introduktion

Denna uppgift syftar till att ge dig grundläggande färdigheter i datahantering, analys och maskininlärning genom arbete med det klassiska Iris-datasetet. Du kommer att utföra följande steg:

1. Utforska datasetet (EDA - Exploratory Data Analysis).
2. Rengöra data och förbereda det för modellering.
3. Bygga och utvärdera en enkel maskininlärningsmodell.

Steg 1: Utforska datasetet

- Analysera datasetets struktur och innehåll genom att använda ``Step_1_Explore_Data.ipynb``.
- Identifiera datasetets kolumner, storlek och datatyper.
- Utför grundläggande statistisk analys, till exempel medelvärde och standardavvikelse för numeriska kolumner.
- Kontrollera om det finns saknade värden eller utliggare.

Steg 2: Datastädning

- Använd ``Step_2_Data_Cleaning.ipynb`` för att förbereda datasetet.
- Hantera saknade värden och normalisera/standardisera data vid behov.
- Förbered datasetet så att det är redo för modellering.

Steg 3: Bygga och utvärdera en modell

- Använd ``Exercise_3_Simple_Model.ipynb`` för att träna och utvärdera en maskininlärningsmodell.
- Skapa en modell med en enkel algoritm, t.ex. beslutsträd, logistisk regression eller K-Nearest Neighbors.
- Utvärdera modellens prestanda med måtvärden som noggrannhet och en förvirringsmatris.
- Visualisera resultaten för att få insikter i modellens prestanda.

Instruktioner och material

- Du hittar datasetet (iris.csv) och Notebooks i materialet som tillhandahålls.
- Utför uppgiften i tre steg, enligt beskrivningen ovan.
- Reflektera över dina resultat och dokumentera vilka insikter du har fått.