



Machine Learning

2024

Inlämning 1

Data Preprocessing

1- Data Exploration:

- a- Förstå varje kolumn och dess datatyp.
- b- Identifiera unika värden, saknade data och potentiella fel.

2- Data Cleaning:

- a- Hantera saknade värden och datainkonsekvenser.
- b- Standardisera textdata för enhetlighet (t.ex. Category, Size, Style).
- c- Konvertera 'Date' till ett lämpligt datetime-format.

3- Feature Engineering:

- a- Derivera nya funktioner, såsom 'Month' eller 'Year' från 'Date'.
- b- Skapa binära funktioner från 'B2B'.
- c- Gruppera 'Amount' i olika försäljningsintervall för kategorisering.

4- Data Transformation:

- a- Normalisera/standardisera numeriska värden som 'Qty' och 'Amount'.
- b- Kodifiera(Encode) kategoriska variabler med tekniker som one-hot encoding.

5- Presentation and Documentation:

- a-** Dokumentera varje steg och resonemanget bakom det.
- b-** Förbered en kort presentation för att visa deras arbete.