

Rapport om Energiförbrukningsanalys

1. Introduktion

1.1 Bakgrund

Energiförbrukning är ett avgörande ämne med betydande miljö- och ekonomiska implikationer. För att minska kostnader och miljöpåverkan är det viktigt att förstå och optimera energiförbrukningen. Denna rapport presenterar en analys av energiförbrukningsdata med syftet att identifiera mönster och förutsäga framtida förbrukning.

1.2 Syfte och Frågeställning

Syftet med detta projekt är att utföra en utförlig analys av historisk energiförbrukningsdata och utveckla en modell för att förutsäga framtida förbrukning. Vi söker svar på frågor som:

Hur varierar energiförbrukningen över tid?

Finns det några tydliga mönster i data?

Kan vi bygga en modell som kan förutsäga energiförbrukningen med hög precision?

2. Databeskrivning / EDA (Exploratory Data Analysis)

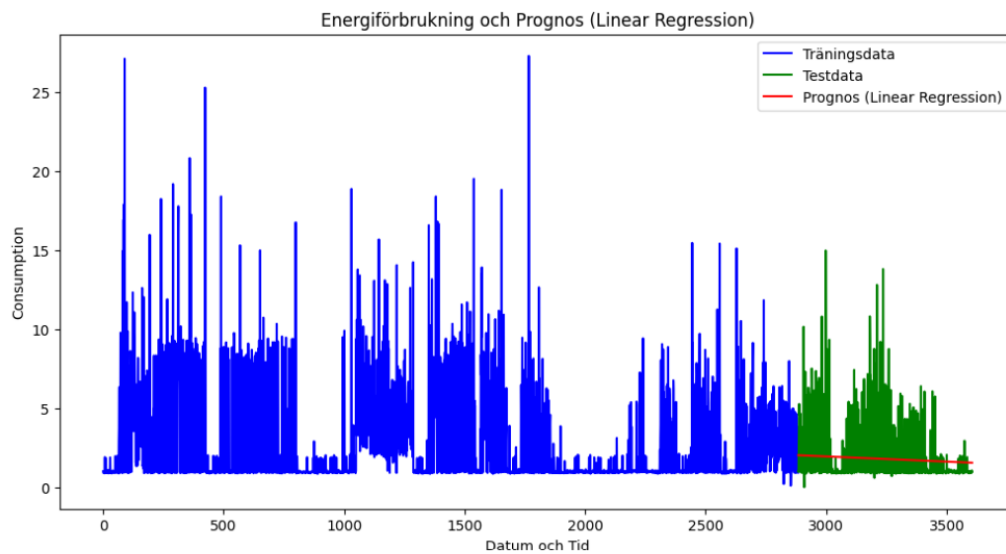
Databasen består av tre separata dataset som har kombinerats till ett större dataset. Initial Exploratory Data Analysis (EDA) visade att det inte fanns några saknade värden i datat, vilket är en fördel. Jag observerade en tydlig variation i energiförbrukningen över tid och använde EDA för att skapa visuella representationer av dessa variationer, inklusive grafer som visar den dagliga genomsnittliga energiförbrukningen.

3. Metod och Modeller (Teori)

I detta projekt har jag använt två olika modeller för att förutsäga energiförbrukningen: linjär regression och beslutsträdregressor.

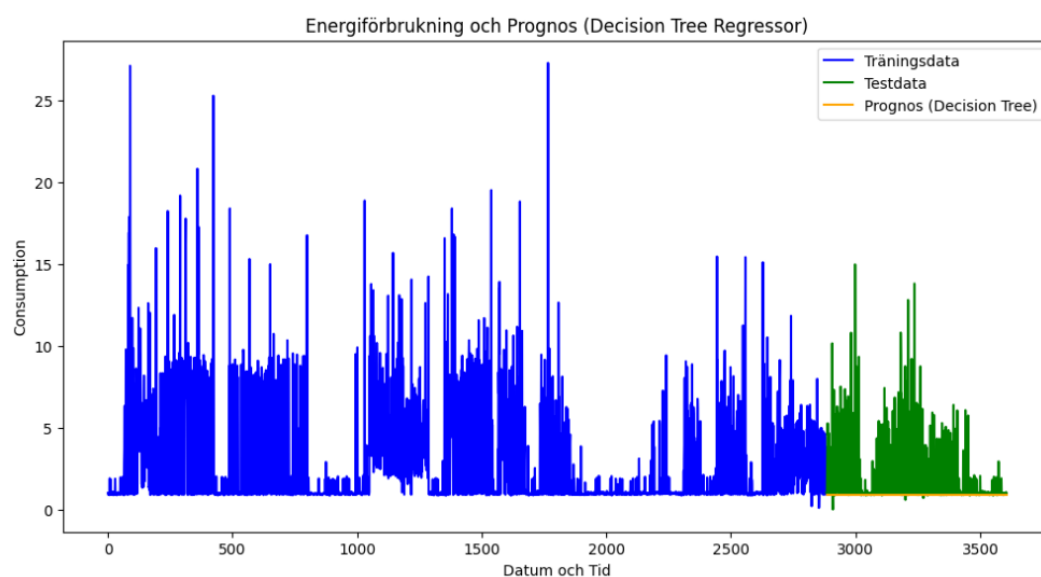
3.1 Linjär Regression

Jag använde linjär regression, en enkel och effektiv modell, med hjälp av Scikit-Learn i Python. Resultatet visade en medel kvadratisk fel (MSE) på 3.95, vilket indikerar en rimlig noggrannhet för förutsägelserna.



3.2 Decision Tree Regressor

Beslutsträdregressorn användes också, men resultatet visade ett högre MSE på 5.26 jämfört med linjär regression. Notera att visualiseringen av trädet inte kunde genereras, vilket indikerar ett potentiellt problem.



4. Projekt Resultat och Analys

Jag delade datasetet i tränings- och testset och tränade båda modellerna med linjär regression och presterade bättre med lägre MSE. De visuella resultaten för båda modellerna var rätt lika, vilket kan vara ett resultat av den specifika datan eller en möjlig felkälla i Decision Tree Regressor.

5. Slutsats och förslag på potentiell vidareutveckling

Slutsatsen är att linjär regression verkar vara mer lämpad för detta ändamål baserat på de lägre MSE-resultaten. För framtida arbete föreslår jag:

En djupare analys av faktorer som påverkar energiförbrukningen.

Utforskning av andra modelltyper och maskininlärningsmetoder.

Insamling av mer data för att ytterligare förbättra modellens prestanda.

6. Reflektion över arbetet

Arbetet var utmanande, särskilt när det gällde att implementera och felsöka modeller. Patricia's feedback, särskilt angående att inkludera mer information om datan och förbättra visualiseringen, har beaktats och resulterat i en mer strukturerad och informativ rapport.

När det gäller varför trädet i beslutsträdregressorn inte visas korrekt, kan det bero på flera faktorer. En möjlig orsak kan vara överanpassning, där trädet blir för komplext och svårt att visualisera. Det kan också finnas problem med biblioteket eller versionen som används. För att lösa detta, kan du överväga att justera modellparametrar eller använda olika versioner av de använda biblioteken.

Det har även varit en utmaning att planera arbetet på egen hand, hitta en bra avgränsning och slutligen slutföra. Jag har stött på flertal error under tiden jag har jobbat på koden, däribland att få streamlit att fungera. Jag får med delar av funktioner, se bild men jag lyckas inte få med alla delar och funktioner jag önskar.

