# Förutsättningar

Jag har provat mig fram med Linear Regression, Decision Trees samt Random Forest för att se vad som verkar träffa bäst. Datasetet som använts sträcker sig från januari 2006 till och med februari 2023 och består av 206 rader då det endast håller en mätserie per månad.

Det slutliga datasetet är uppbyggt av följande kolumner:

Priser på:

* Bensin
* Diesel
* Etanol
* Snitt på rörliga elpriset över el-områdena

Antal sålda bilar fördelat på:

* Bensin
* Diesel
* El
* Elhybrid
* Laddhybrid
* Etanol

Då jag hade svårt att hitta bra informationskällor gällande gaspriser och med tanke på hur liten del av försäljningssiffrorna gasdrivna bilar representerar har jag valt att bortse från dessa.

I stället för att försöka förutsäga en exakt försäljningssiffra har jag lagt till en extra kolumn för respektive drivmedel (sålda bilar) som håller andelen (antal sålda av den specifika kategorin/totala antalet). Det finns med andra ord en ”bensin\_andel”, ”diesel\_andel” osv.

# Utfall

Generellt har jag kunnat konstatera att ren Linear Regression är ”spretigast” medan Random Forest samlar träffarna aningen bättre än de Decision Trees jag satt upp. Det som i mina ögon påverkar mest är hur långt fram i tiden jag försöker göra mina förutsägelser. Som synes i graferna nedan händer något runt två-tre månader där punkterna samlas närmare idealet vilket faktiskt skulle kunna vara rimligt med tanke på ledtiden från beställning till leverans av en bil. Vid fyra månader glider förutsägelserna isär igen men samlas åter ytterligare en månad framåt vilket till viss del talar emot att ”sweet spoten” ligger just vid tre månader.

En faktor att ha i åtanke är att datamängden modellen har att jobba med minskar i takt med att horisonten framåt i tiden ökar. Om jag tittar två månader framåt i tiden kan jag inte nyttja mina två sista X i serien som i så fall skulle peka ut ett y som inte existerar.

Sannolikt hade det gått att se en tydligare trend med ett större underlag. Skulle man backa längre tillbaka i tiden hade bensin och diesel dock snart varit ensamma kategorier men det kunde vara en intressant övning att fokusera endast på dessa.

## X och y samma månad

En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivningEn bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivningEn bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning

## En månad framåt

En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivningEn bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivningEn bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning

## Två månader framåt

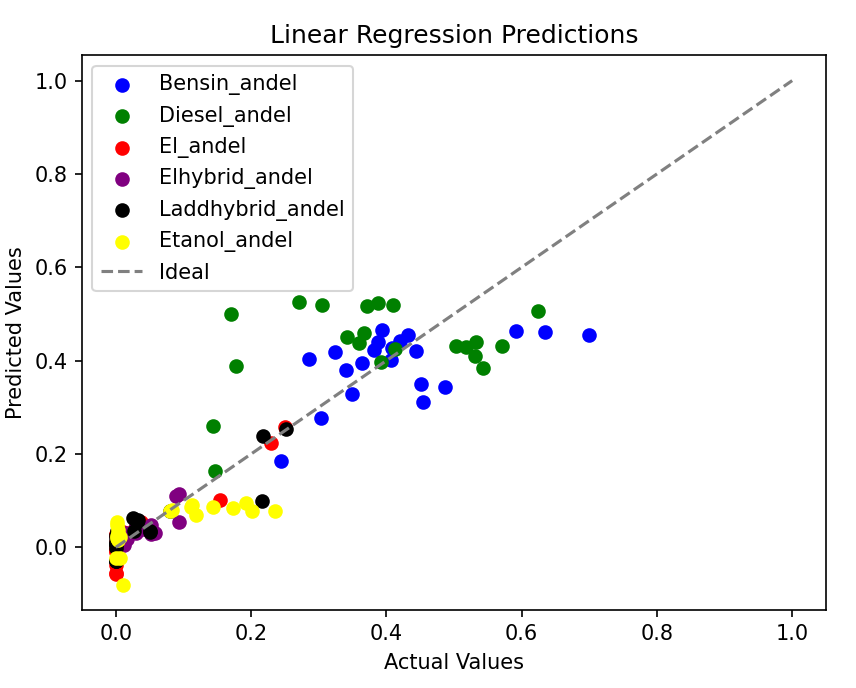
En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning

## Tre månader framåt

 En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning

## Fyra månader framåt

En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning

## Fem månader framåt

En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning

## Sex månader framåt

En bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivningEn bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivningEn bild som visar text, skärmbild, diagram, linje

Automatiskt genererad beskrivning