

AP1

Ionescu M. Emi-Marian

Ianuarie 2025

Descriere proiect

Scopul este prezicere datelor de pe luna Decembrie 2024 , antrenând un model de învățare automată Naive Bayes și ID3. Modelul se antrenează pe datele din Ianuarie 2022 până Noiembrie 2024

Justificare abordare

$Sold = Consum - Producție$ Preprocesez listele de attribute transformându-le în 'liste de diferențe' $dA[i] = A[i] - A[i-1]$, este o transformare reversibilă/bijectivă știind $A[0]$, se poate antrena și prezice direct pe aceste liste. Setul de date este dat la interval de 10 min , transform datele luând media valorilor să estimez intervale de o oră. Fac binning pe valorile de creștere $[-300, 100, 0, 100, 300]$. De exemplu eticheta/labelul -300 înseamnă că din ultima oră a crescut cu aproximativ -300 sau mai mult valoarea.

Rezultate

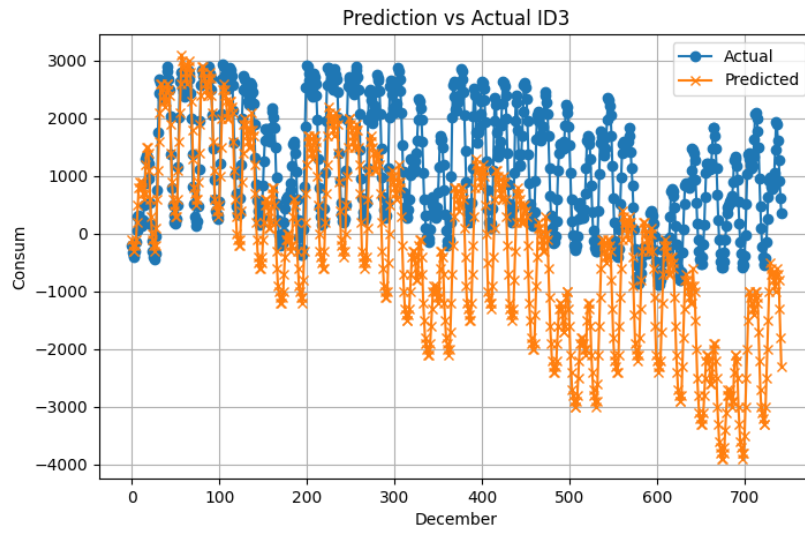


Figure 1: Consum ID3

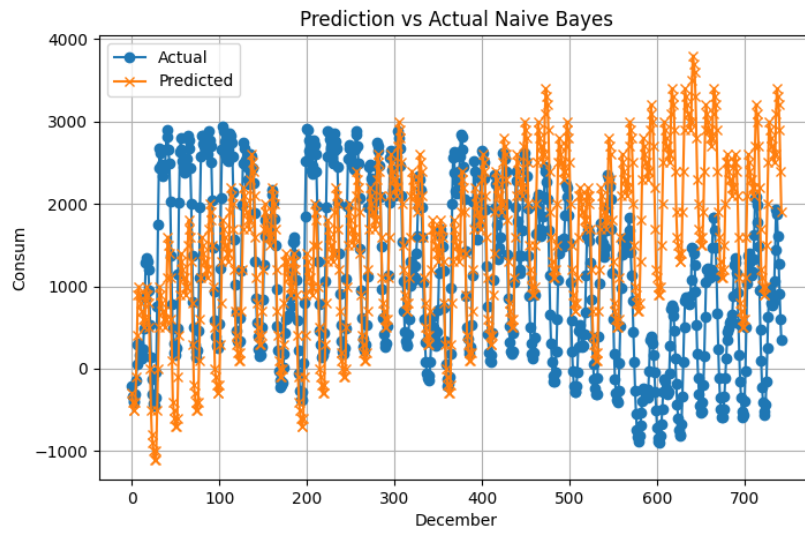


Figure 2: Consum Naive Bayes

Consum MAE ID3 : 1500.6191117092867
 Consum MAE Naive Bayes : 917.9650067294751

Figure 3: Consum MAE

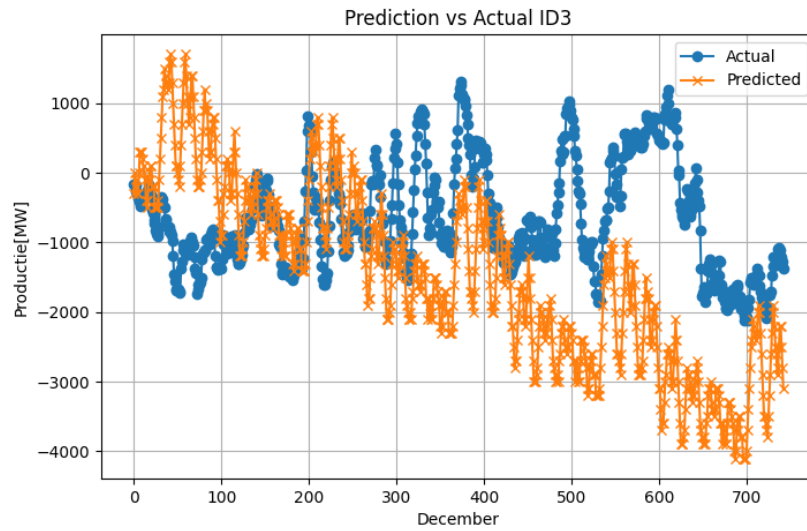


Figure 4: Productie ID3

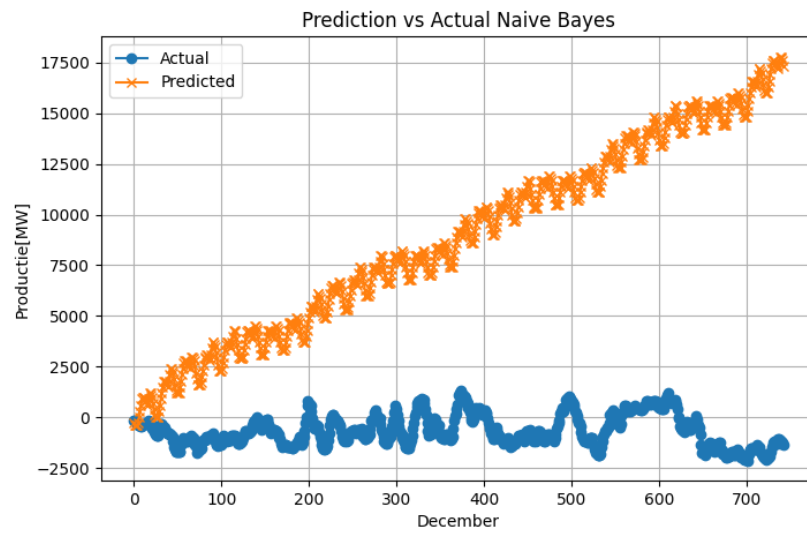


Figure 5: Productie Naive Bayes

```
Productie[MW] MAE ID3 : 1365.2001201691987
Productie[MW] MAE Naive Bayes : 9517.189202396972
```

Figure 6: Productie MAE

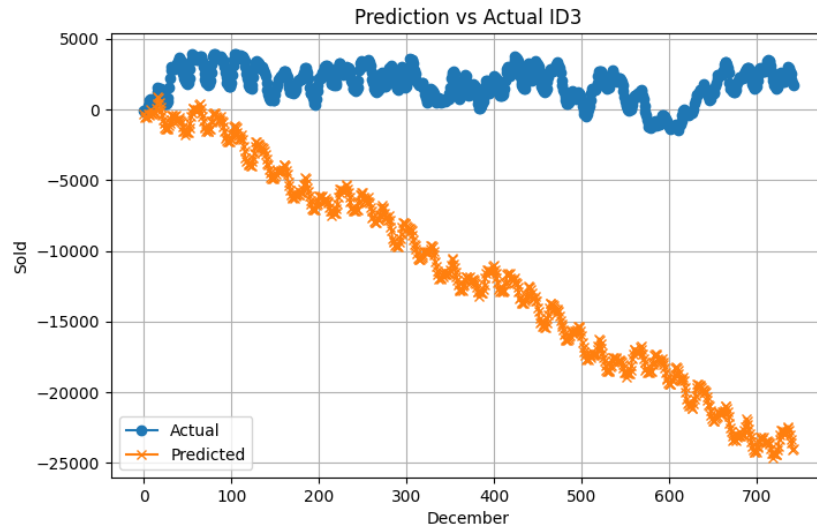


Figure 7: Consum ID3

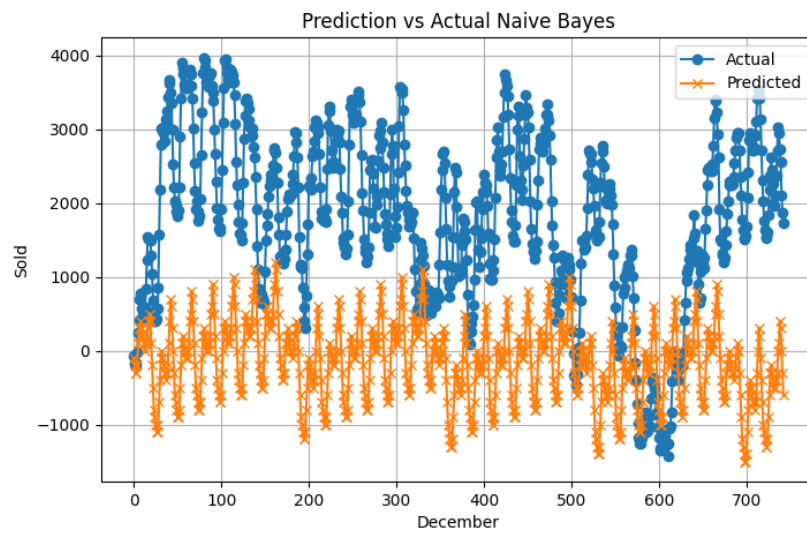
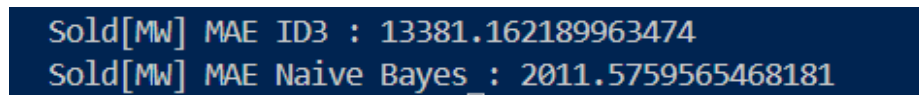


Figure 8: Consum Naive Bayes



```
Sold[MW] MAE ID3 : 13381.162189963474
Sold[MW] MAE Naive Bayes : 2011.5759565468181
```

Figure 9: Sold MAE

Concluzii

Aș fi putut adăuga o etapă de model selection/validare , antrenând până Noiembrie 2023 de exemplu și evaluând luna Decembrie 2023, modificând hyperparametrii (binsize și alți parametri specifici Naive Bayes sau ID3) până obțin o eroare mică și după să evaluez pe test setul din Decembrie 2024. Binsize-ul s-ar putea selecta adaptiv după setul de date.