1 Introducere

Clasificarea e-mailurilor spam reprezintă o problemă în cadrul sistemelor de filtrare a e-mailurilor, având ca obiectiv distingerea între e-mailurile considerate spam și cele non-spam. Această sarcină este esențială pentru a asigura clienților primirea comunicărilor critice, în timp ce elimină conținutul nedorit sau potențial periculos. În abordarea acestei probleme, s-au utilizat diverse tehnici de învățare automată, fiecare cu avantajele și limitele sale specifice. Între acestea, metoda Credulous Bayes a devenit o opțiune populară datorită eficienței sale, productivității și eficacității în sarcinile de clasificare a conținutului.

2 Alegerea unui algoritm

2.1 Individual

2.1.1 Teoretic

Algoritmul Naive Bayes a fost ales în primul rând pentru simplitatea și fundamentul teoretic robust. În ciuda presupunerii sale naive" de independență a caracteristicilor, care ar putea să nu fie adevărată în scenariile din lumea reală, a demonstrat performanțe remarcabile în sarcinile de clasificare a textului. Algoritmul folosește teorema lui Bayes pentru a calcula probabilitatea condiționată ca un e-mail să aparțină unei anumite clase, având în vedere caracteristicile sale. Această abordare teoretică solidă o face o alegere atractivă pentru clasificarea spam-ului.

2.1.2 Experimental

Evaluările empirice pe diverse seturi de date au demonstrat în mod constant eficiența Naive Bayes în clasificarea spam-ului. Capacitatea sa de a gestiona spații de caracteristici cu dimensiuni mari, inerente datelor text și eficiența sa de calcul îl fac o alegere atrăgătoare. Rezultatele experimentale au indicat o acuratețe ridicată și performanță fiabilă, chiar și cu resurse de calcul limitate, justificând și mai mult selecția sa.

2.2 Comparativ

2.2.1 Teoretic

Comparativ, Naive Bayes a fost favorizat față de alți algoritmi datorită simplității și capacității sale de a gestiona eficient spații mari de caracteristici. În timp ce modelele mai complexe, cum ar fi SVM (Support Vector Machines) sau metodele de ansamblu, pot oferi o expresivitate mai mare, Naive Bayes excelează în eficiență computațională, făcându-l mai potrivit pentru sarcinile de clasificare a textului la scară largă.

2.2.2 Experimental

În experimente comparative care implică algoritmi precum SVM, Decision Trees și Random Forests, Naive Bayes a demonstrat constant performanță competitivă sau chiar superioară. Eficiența sa de calcul și eficacitatea în manipularea datelor text cu dimensiuni mari au oferit un avantaj atât în ceea ce privește precizia, cât și viteza de antrenament.

3 Algoritmul Bayes Naiv pentru clasificarea spam

Implementarea Naive Bayes pentru clasificarea spam-ului implică de obicei mai mulți pași:

1. Preprocesare: Curățarea și preprocesarea datelor de e-mail prin eliminarea semnelor de punctuație, tokenizare, eliminarea cuvintelor redundante, stemming si lematizare.

2. Feature Extraction: Conversia datelor text în vectori de caracteristici numerice, folosind adesea pachete de cuvinte sau reprezentări TF-IDF.

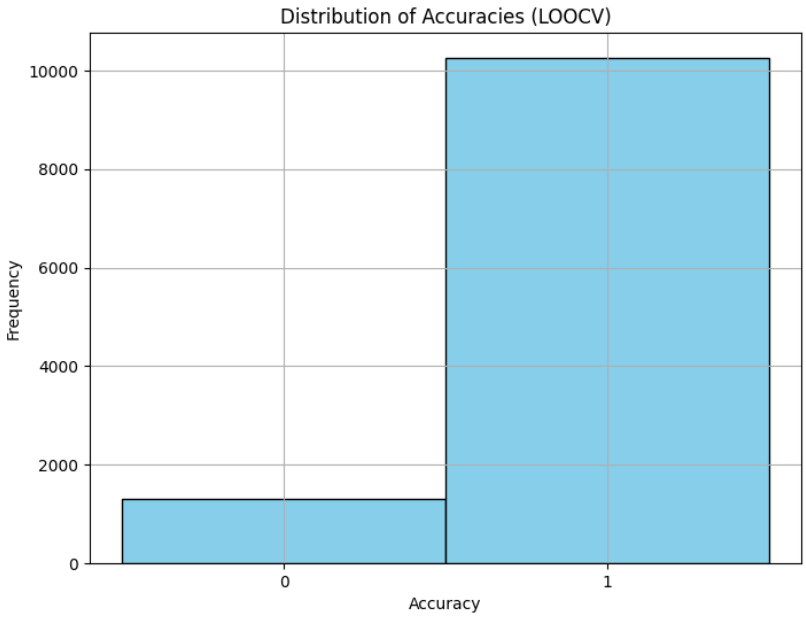
3. Instruire: Estimarea probabilităților caracteristicilor date etichete de clasă (spam sau non-spam) folosind setul de date de antrenament.

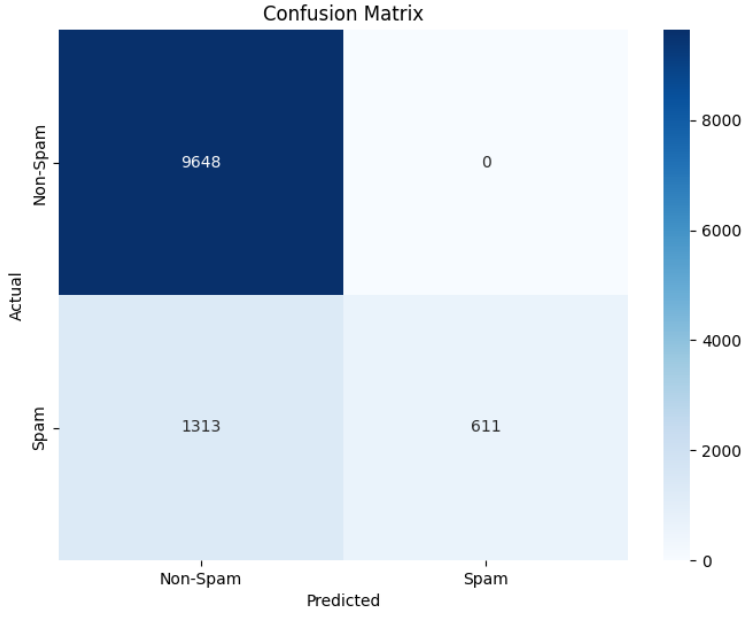
4. Predicție: Utilizarea acestor probabilități pentru a prezice eticheta de clasă a noilor e-mailuri pe baza caracteristicilor acestora.

5. Evaluare: Evaluarea performanței algoritmului folosind valori precum acuratețea, precizia, retragerea și scorul F1.

Naive Bayes, în ciuda simplității sale, s-a dovedit a fi un instrument puternic și eficient pentru clasificarea spam-ului, oferind un echilibru între eficiența de calcul și performanța respectabilă.

4 Grafice aferente experimentului


Distribution of Accuracies (LOOCV)
Frequency
Accuracy
0
2000
4000
6000
8000
10000
0
1


Confusion Matrix
Actual
Predicted
Non-Spam
Spam
9648
0
1313
611
8000
6000
4000
2000
0