

# Justificarea Alegerii Algoritmului Naive Bayes pentru Clasificarea Email-urilor Spam

Apricopoi Andrei-Constantin, Roman Lavinia

12.01.2024

Acest raport justifică alegerea algoritmului Naive Bayes pentru problema de clasificare a email-urilor spam. Analizăm atât aspectele teoretice, cât și rezultatele experimentale, comparând Naive Bayes cu alți algoritmi de clasificare.

## 1 Introducere

Clasificarea email-urilor spam este o provocare majoră în domeniul procesării textului și al învățării automate. Acest raport explorează utilizarea algoritmului Naive Bayes pentru această sarcină și justifică alegerea sa prin analiza teoretică și compararea performanței cu alte metode.

## 2 Fundamentele Teoretice ale Naive Bayes

Naive Bayes este un clasificator probabilistic bazat pe aplicarea Teoremei Bayes cu presupunerea de independență naivă între predictorii. Această simplificare face Naive Bayes surprinzător de eficient pentru clasificarea textelor, cum ar fi email-urile, datorită abilității sale de a gestiona un număr mare de caracteristici și de a se adapta rapid la noi date.

## 3 Descrierea Setului de Date Ling-Spam

Setul de date Ling-Spam este folosit extensiv în clasificarea email-urilor pentru a distinge între spam și mesaje legitime. Acesta conține email-uri colectate de pe o listă de discuții despre lingvistică, incluzând atât mesaje relevante (non-spam), cât și mesaje nesolicitate (spam).

### 3.1 Compoziția Setului de Date

Setul este compus din email-uri clasificate în două categorii principale:

- **Spam:** Email-uri nesolicitate, adesea conținând publicitate sau alte mesaje irelevante.

- **Non-Spam:** Mesaje legitime, relevante pentru subiectul listei de discuții.

Fiecare email din set este marcat corespunzător, cu email-urile de spam având adesea un prefix distinctiv în titlu, cum ar fi "spm".

### 3.2 Prelucrarea și Pregătirea Datelor

Prelucrarea setului de date Ling-Spam implică mai mulți pași esențiali pentru a-l pregăti pentru analiza de clasificare:

1. **Curățarea Textului:** Eliminarea elementelor nerelevante din email-uri, cum ar fi anteturile, semnăturile și orice alt text care nu contribuie la identificarea naturii spam a mesajului.
2. **Extragerea Conținutului:** Separarea textului efectiv al email-ului de metadate și structurarea acestuia într-un format adecvat pentru procesare.
3. **Etichetarea Datelor:** Asigurarea că fiecare email este corect etichetat ca fiind spam sau non-spam, bazat pe marcajele existente sau pe evaluări manuale.

Acești pași sunt cruciali pentru a asigura acuratețea și eficiența algoritmilor de clasificare care vor fi aplicați ulterior pe acest set de date.

### 3.3 Metodologia Experimentală

Metodologia experimentală adoptată pentru evaluarea algoritmului Naive Bayes și compararea acestuia cu alți algoritmi implică mai mulți pași, de la prelucrarea inițială a datelor până la evaluarea finală a performanței modelului.

#### 3.3.1 Setările de Antrenament

Pentru antrenarea modelului, am folosit următoarele setări:

- **Împărțirea Datelor:** Setul de date a fost împărțit într-un set de antrenament și un set de testare, utilizând o proporție de 90% pentru antrenament și 10% pentru testare.
- **Parametrii Modelului:** În cazul Naive Bayes, am utilizat setări implicite.
- **Cross-Validare:** Pentru a evalua robustețea modelului, am aplicat tehnici de cross-validare, cum ar fi k-fold sau Leave-One-Out, pe setul de antrenament.

### 3.3.2 Metodele de Evaluare

Evaluarea performanței fiecărui model s-a bazat pe mai multe metrice:

- **Acuratețea:** Procentul de clasificări corecte realizate de model.
- **Timpul de Antrenament:** Evaluarea eficienței temporale a fiecărui algoritm.

Această metodologie ne-a permis să realizăm o evaluare cuprinzătoare a performanței Naive Bayes în comparație cu alte tehnici de clasificare, oferind astfel o bază solidă pentru selecția algoritmului cel mai potrivit pentru problema noastră.

## 4 Concluzii

Concluzionăm de ce Naive Bayes este o alegere adecvată pentru această problemă, subliniind echilibrul său unic între eficiență, acuratețe și simplitate.