Anlemphase Bayes Spam Filter Ham Spam Start Zipfiles entpacken Mail parsen HTML entfernen Wörter zählen

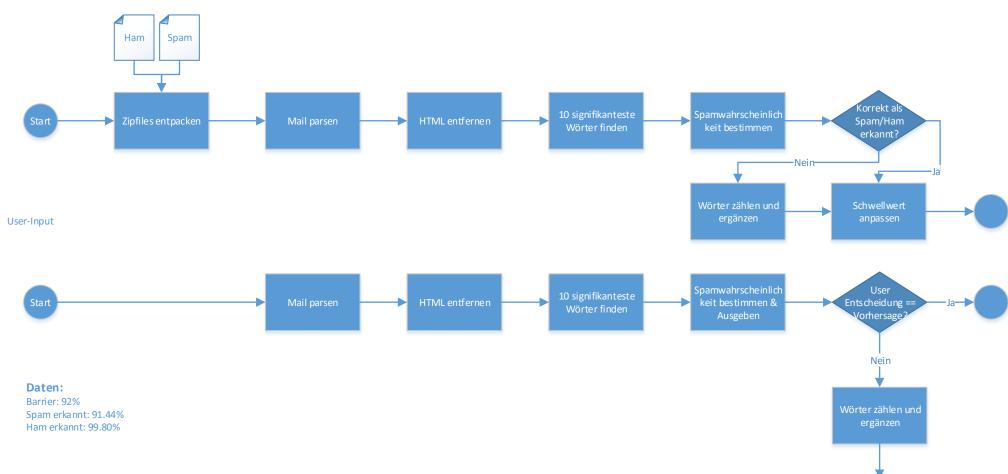
Vorteile

- Hohe Spam erkennungsrate
- Direkte Einspeisung von Zips
- Nachträgliches Prüfen einfach möglich

Nachteile

- Falls Spam häufig auf Englisch und Ham selten in Englisch -> Englischer Ham wird dann als Spam erkannt.

Kalibierungsphase



Autoren:

Janis Peyer, Samuel Merki, Benjamin Leber https://github.com/TheJP/dist