



Naive bayes




$p(\mathbf{N}) = 0.67$



Lunch Money Money Money Money

...and that means we will always classify the messages with **Lunch** in them as **normal**, no matter how many times we see the word **Money**.




$p(\mathbf{S}) = 0.33$

$p(\text{Dear} | \mathbf{N}) = 0.47$
 $p(\text{Friend} | \mathbf{N}) = 0.29$
 $p(\text{Lunch} | \mathbf{N}) = 0.18$
 $p(\text{Money} | \mathbf{N}) = 0.06$

$p(\mathbf{N}) \times p(\text{Lunch} | \mathbf{N}) \times p(\text{Money} | \mathbf{N})^4 = 0.000002$

 $p(\mathbf{S}) \times p(\text{Lunch} | \mathbf{S}) \times p(\text{Money} | \mathbf{S})^4 = 0$



Naivní je protože neuvažuje pořadí slov