



Naïve Bayes Discussion

Computational Linguistics: Jordan Boyd-Graber
University of Maryland

LECTURE 1A

Roadmap

- Content Questions
- Administrivia Questions
- NB Exercise

Content Questions

Administrivia Announcements

- Use Piazza
- submit server
- TA office hours

Administrivia Questions

Documents

D1: Spam

abuja man

D3: Spam

cialis deal

D5: Spam

abuja deal

D7: Spam

cialis dog

D2: Ham

man dog

D4: Ham

logistic mother logistic abuja

D6: Ham

bagel deal

Documents

D1: Spam

abuja man

D3: Spam

cialis deal

D5: Spam

abuja deal

D7: Spam

cialis dog

D2: Ham

man dog

D4: Ham

logistic mother logistic abuja

D6: Ham

bagel deal

What's $|C|$ and $|V|$?

Documents

D1: Spam

abuja man

D3: Spam

cialis deal

D5: Spam

abuja deal

D7: Spam

cialis dog

$|C| = 2$ (spam vs. ham)

D2: Ham

man dog

D4: Ham

logistic mother logistic abuja

D6: Ham

bagel deal

Documents

D1: Spam

abuja man

D3: Spam

cialis deal

D5: Spam

abuja deal

D7: Spam

cialis dog

D2: Ham

man dog

D4: Ham

logistic mother logistic abuja

D6: Ham

bagel deal

$|V|=8$: 'deal', 'dog', 'bagel', 'logistic',
'mother', 'cialis', 'abuja', 'man'

Background Probabilities

D1: Spam

abuja man

D3: Spam

cialis deal

D5: Spam

abuja deal

D7: Spam

cialis dog

D2: Ham

man dog

D4: Ham

logistic mother logistic abuja

D6: Ham

bagel deal

Background Probabilities

D1: Spam

abuja man

D3: Spam

cialis deal

D5: Spam

abuja deal

D7: Spam

cialis dog

D2: Ham

man dog

D4: Ham

logistic mother logistic abuja

D6: Ham

bagel deal

What's $\hat{P}(c_j)$?

Background Probabilities

- For spam:

(1)

Background Probabilities

- For spam:

$$\hat{P}(c_j = \text{spam}) = \frac{N_c + 1}{N + |C|} \quad (1)$$

(2)

Background Probabilities

- For spam:

$$\hat{P}(c_j = \text{spam}) = \frac{N_c + 1}{N + |C|} \quad (1)$$

$$= \frac{4 + 1}{7 + 2} \quad (2)$$

$$= \frac{5}{9} \quad (3)$$

Background Probabilities

- For spam:

$$\hat{P}(c_j = \text{spam}) = \frac{N_c + 1}{N + |C|} \quad (1)$$

$$= \frac{4 + 1}{7 + 2} \quad (2)$$

$$= \frac{5}{9} \quad (3)$$

- For ham:

$$(4)$$

Background Probabilities

- For spam:

$$\hat{P}(c_j = \text{spam}) = \frac{N_c + 1}{N + |C|} \quad (1)$$

$$= \frac{4 + 1}{7 + 2} \quad (2)$$

$$= \frac{5}{9} \quad (3)$$

- For ham:

$$\hat{P}(c_j = \text{ham}) = \frac{N_c + 1}{N + |C|} \quad (4)$$

$$(5)$$

Background Probabilities

- For spam:

$$\hat{P}(c_j = \text{spam}) = \frac{N_c + 1}{N + |C|} \quad (1)$$

$$= \frac{4 + 1}{7 + 2} \quad (2)$$

$$= \frac{5}{9} \quad (3)$$

- For ham:

$$\hat{P}(c_j = \text{ham}) = \frac{N_c + 1}{N + |C|} \quad (4)$$

$$= \frac{3 + 1}{7 + 2} \quad (5)$$

$$= \frac{4}{9} \quad (6)$$

Conditional Probabilities

D1: Spam

abuja man

D3: Spam

cialis deal

D5: Spam

abuja deal

D7: Spam

cialis dog

D2: Ham

man dog

D4: Ham

logistic mother logistic abuja

D6: Ham

bagel deal

What's the conditional probability $\hat{P}(w = \text{dog} | c)$?

Conditional Probabilities

- For spam:

(7)

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (7)$$

(8)

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (7)$$

$$= \frac{1 + 1}{8 + 8} \quad (8)$$

$$(9)$$

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (7)$$

$$= \frac{1 + 1}{8 + 8} \quad (8)$$

$$= \frac{1}{8} \quad (9)$$

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (7)$$

$$= \frac{1 + 1}{8 + 8} \quad (8)$$

$$= \frac{1}{8} \quad (9)$$

- For ham:

(10)

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (7)$$

$$= \frac{1 + 1}{8 + 8} \quad (8)$$

$$= \frac{1}{8} \quad (9)$$

- For ham:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (10)$$

$$(11)$$

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (7)$$

$$= \frac{1 + 1}{8 + 8} \quad (8)$$

$$= \frac{1}{8} \quad (9)$$

- For ham:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (10)$$

$$= \frac{1 + 1}{8 + 8} \quad (11)$$

$$(12)$$

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (7)$$

$$= \frac{1 + 1}{8 + 8} \quad (8)$$

$$= \frac{1}{8} \quad (9)$$

- For ham:

$$\hat{P}(w = \text{dog} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (10)$$

$$= \frac{1 + 1}{8 + 8} \quad (11)$$

$$= \frac{1}{8} \quad (12)$$

Prediction

What if you saw a document with the word “dog”?

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

(14)

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

$$= \frac{5}{9} \cdot \frac{1}{8} \quad (14)$$

$$(15)$$

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

$$= \frac{5}{9} \cdot \frac{1}{8} \quad (14)$$

$$= 0.07 \quad (15)$$

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

$$= 0.07 \quad (14)$$

- For ham:

$$(15)$$

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

$$= 0.07 \quad (14)$$

- For ham:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (15)$$

$$(16)$$

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

$$= 0.07 \quad (14)$$

- For ham:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (15)$$

$$= \frac{4}{9} \cdot \frac{1}{8} \quad (16)$$

$$(17)$$

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

$$= 0.07 \quad (14)$$

- For ham:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (15)$$

$$= \frac{4}{9} \cdot \frac{1}{8} \quad (16)$$

$$= 0.06 \quad (17)$$

Prediction

What if you saw a document with the word “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (13)$$

$$= 0.07 \quad (14)$$

- For ham:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (15)$$

$$= 0.06 \quad (16)$$

These aren't probabilities? What if we wanted the real probabilities?

Conditional Probabilities

- For spam:

(17)

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{logistic} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (17)$$

(18)

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{logistic} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (17)$$

$$= \frac{0 + 1}{8 + 8} \quad (18)$$

$$= \frac{1}{16} \quad (19)$$

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{logistic} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (17)$$

$$= \frac{0 + 1}{8 + 8} \quad (18)$$

$$= \frac{1}{16} \quad (19)$$

- For ham:

$$(20)$$

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{logistic} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (17)$$

$$= \frac{0 + 1}{8 + 8} \quad (18)$$

$$= \frac{1}{16} \quad (19)$$

- For ham:

$$\hat{P}(w = \text{logistic} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (20)$$

$$(21)$$

Conditional Probabilities

- For spam:

$$\hat{P}(w = \text{logistic} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (17)$$

$$= \frac{0 + 1}{8 + 8} \quad (18)$$

$$= \frac{1}{16} \quad (19)$$

- For ham:

$$\hat{P}(w = \text{logistic} | c) = \frac{T_{cw} + 1}{(\sum_{w' \in V} T_{cw'}) + |V|} \quad (20)$$

$$= \frac{2 + 1}{8 + 8} \quad (21)$$

$$= \frac{3}{16} \quad (22)$$

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (23)$$

(24)

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (23)$$

$$= \frac{5}{9} \cdot \frac{1}{8} \cdot \frac{1}{16} \cdot \frac{1}{16} \quad (24)$$

$$(25)$$

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (23)$$

$$= \frac{5}{9} \cdot \frac{1}{8} \cdot \frac{1}{16} \cdot \frac{1}{16} \quad (24)$$

$$= 0.0002 \quad (25)$$

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (23)$$

$$= 0.0002 \quad (24)$$

- For ham:

$$(25)$$

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (23)$$

$$= 0.0002 \quad (24)$$

- For ham:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (25)$$

$$(26)$$

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (23)$$

$$= 0.0002 \quad (24)$$

- For ham:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (25)$$

$$= \frac{4}{9} \cdot \frac{1}{8} \cdot \frac{3}{16} \cdot \frac{3}{16} \quad (26)$$

$$(27)$$

Prediction

What if you saw a document with the words “logistic” “logistic” “dog”?

- For spam:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (23)$$

$$= 0.0002 \quad (24)$$

- For ham:

$$P(c|d) \propto P(c) \prod_{1 \leq i \leq n_d} P(w_i|c) \quad (25)$$

$$= \frac{4}{9} \cdot \frac{1}{8} \cdot \frac{3}{16} \cdot \frac{3}{16} \quad (26)$$

$$= 0.002 \quad (27)$$