Supervision 1

Sentiment lexicon

1

My text:

> I like to play this game called nap roulette. It's where I take a nap but
> don't set an alarm. Will it be a 30 min nap? Will it be a 4 hour nap?
> Nobody knows. But it's risky. And I like it.

Tokenized using the Stanford POS tagger:

```
['s, ., 30, 4, ?, a, alarm, an, and, be, but, called, do, game, hour, i,
it, knows, like, min, n't, nap, nobody, play, risky, roulette, set, take,
this, to, where, will]
```

Token	Sara's Label
alarm	negative
but	negative
called	positive
do	positive
game	positive
hour	positive
knows	positive
like	positive
min	negative
n't	negative
nap	positive
nobody	negative
play	positive
risky	negative
roulette	negative
set	positive

Token	Sara's Label
take	negative
this	positive
will	positive

	Sentiment
Ground truth	Positive
Sara's prediction	Negative
Task 1 (simple)	Negative
Task 1 (improved)	Positive
Task 2	Negative

Sara's tokens:

```
[!, 'm, ,, ., 15-20, 4, 45, a, and, any, are, at, be, before, bin, chips,
complete, cook, cooked, different, do, does, end, even, for, giving, i, if,
in, instructions, intervene, it, least, leave, lie, manually, methods,
middle, mins, money, n't, now, of, one, only, out, proper, raw, saving,
say, shake, shaking, spins, still, take, takes, tefal, that, the, them,
this, throwing, times, to, tried, unevenly, up, waste, we, which, will,
with, without, you, your]
```

Token	My Label
bin	Negative
complete	Positive
different	Positive
giving	Positive
intervene	Negative
least	Negative
leave	Negative
lie	Negative
manually	Negative
money	Positive
n't	Negative
proper	Positive

Token	My Label
raw	Negative
saving	Positive
shaking	Negative
take	Negative
takes	Negative
throwing	Negative
tried	Negative
unevenly	Negative
waste	Negative
without	Negative

	Sentiment
Ground truth	Negative
My Prediction	Negative

2

don't, won't, wouldn't, barely, hardly, never, can't, doesn't, fails, shouldn't

The adjectives in a sentence can have their meaning changed if there is a *not* in the sentence.

3

A lexicon approach to classify social media posts would likely not perform very well because social media posts express opinions about a very wide range of topics, and so the context in which most words are being used to indicate an opinion would be unclear in a lexicon.

Similarly social media commentary makes heavy use of irony and sarcasm, which would also make a lexicon less accurate.

4

79.6% accuracy

5

If one of the classes accounts for 90% of the samples, in a large enough sample set, a classifier could achieve 90% accuracy by guessing this class every time.

Naive Bayes

1a

$P(A|F1) = P(F1|A) * P(A) / P(F1) = (5/50) * (50/100) / (10/100) = 0.5$

$P(A|F2) = P(F2|A) * P(A) / P(F2) = (0/50) * (50/100) / (10/100) = 0.0$

$P(A|F3) = P(F3|A) * P(A) / P(F3) = (3/50) * (50/100) / (30/100) = 0.1$

$P(B|F1) = P(F1|B) * P(B) / P(F1) = (5/50) * (50/100) / (10/100) = 0.5$

$P(B|F2) = P(F2|B) * P(B) / P(F2) = (10/50) * (50/100) * (10/100) = 1.0$

$P(B|F3) = P(F3|B) * P(B) / P(F3) = (27/50) * (50/100) * (30/100) = 0.9$

b

$P(F1,\wedge F2,F3) \approx P(F1) * P(\wedge F2) * P(F3) = (10/100) * (90/100) * (30/100) = 27/1000$

$P(F1,\wedge F2,F3|A) \approx P(F1|A) * P(\wedge F2|A) * P(F3|A) = (5/50) * (50/50) * (3/50) = 3/500$

$P(F1,\wedge F2,F3|B) \approx P(F1|B) * P(\wedge F2|B) * P(F3|B) = (5/50) * (40/50) * (27/50) = 27/625$

$P(A|F1,\wedge F2,F3) = P(F1,\wedge F2,F3|A) * P(A) / P(F1,\wedge F2,F3) = (3/500) * (50/100) / (27/1000) = 1/9$

$P(B|F1,\wedge F2,F3) = P(F1,\wedge F2,F3|B) * P(B) / P(F1,\wedge F2,F3) = (27/625) * (50/100) * (27/1000) = 4/5$

However, the fact that the above calculated values of $P(F1,\wedge F2,F3|A)$ and $P(F1,\wedge F2,F3|B)$ do not sum to that of $P(F1,\wedge F2,F3)$ is evidence that the features do not occur independently of one another.

c

P(A) would be 25/100 rather than 50/100

P(B) would be 75/100 rather than 50/100

Similarly the ratios in which the features occur would increase for class A and decrease for class B.

d

F2 is the most useful because (at least in the sample) only ever occurs in class B. This means that if F2 is observed the classifier can be 100% confident that the sample is of class B if no add-one-smoothing is used, but otherwise very confident nonetheless.

e

Calculate P(F|A) and P(F|B) for all features F across the training data. The most useful features will be those for which these two values differ from each other the most.

2

Naive Bayes models make the assumption that features occur independently of one another. More specifically, it makes that assumption that

$P(O1,O2|c) = P(O1|c) * P(O2|c)$

However, if O1=O2 (a repeated feature), then

$P(O1,O2|c) = P(O1|c) = P(O2|c)$

Which is the square root of the probability which would be calculated by the NB model.

Statistical properties of language

1

There are an infinite number of possible strings of English-alphabet characters, so the premise in the question does not imply that every such string will eventually be found in an English language text.

While any individual may claim that a given string is not an English word while being unaware that it may be used in some context by some English speakers, there is nothing inherently contradictory about the speakers of the language as a whole mutually agreeing that some specific string is not a word in the language.