Lab 6

Bayes classification

Useful information

Alex Becheru

irlab.becheru.net

irlab@becheru.net

Laboratory scope

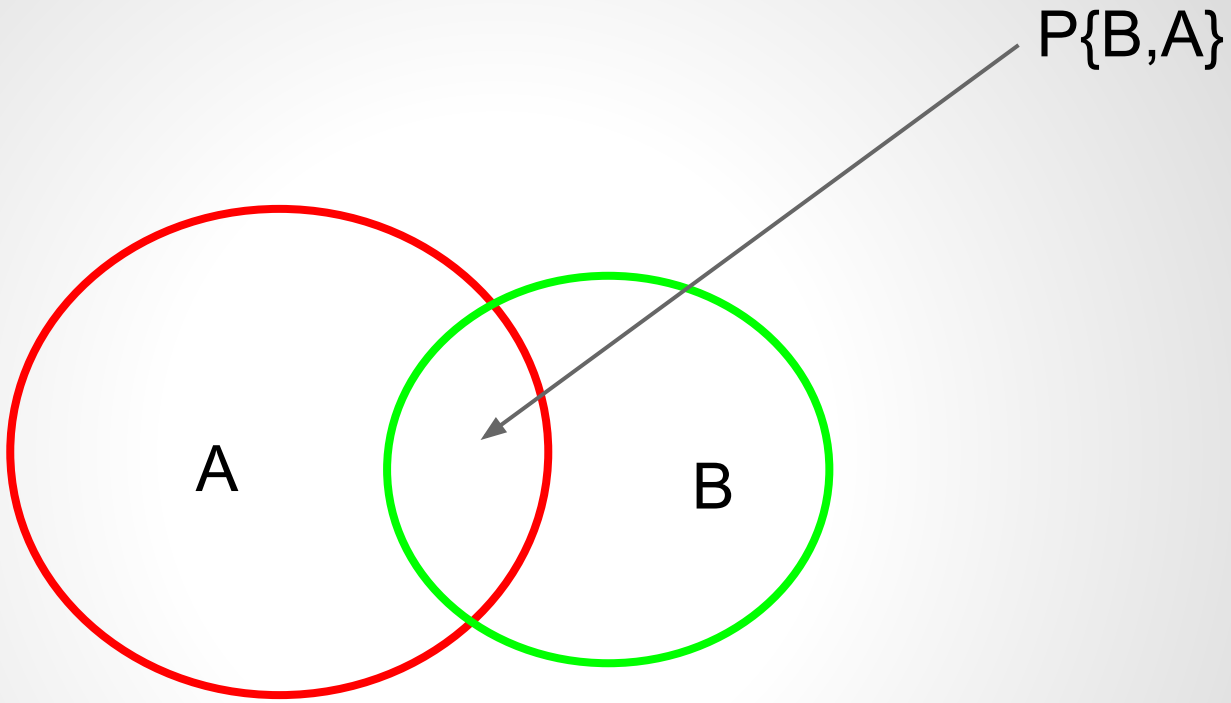
Naive Bayes Classification

Bayesian probabilistics

What is the probability of an event A to take place if it is conditioned by event B?

$$P\{B|A\} = \frac{P\{A|B\} \cdot P\{B)}{P\{A\}}.$$

$$posterior = \frac{likelihood \times prior\ probability}{evidence}.$$

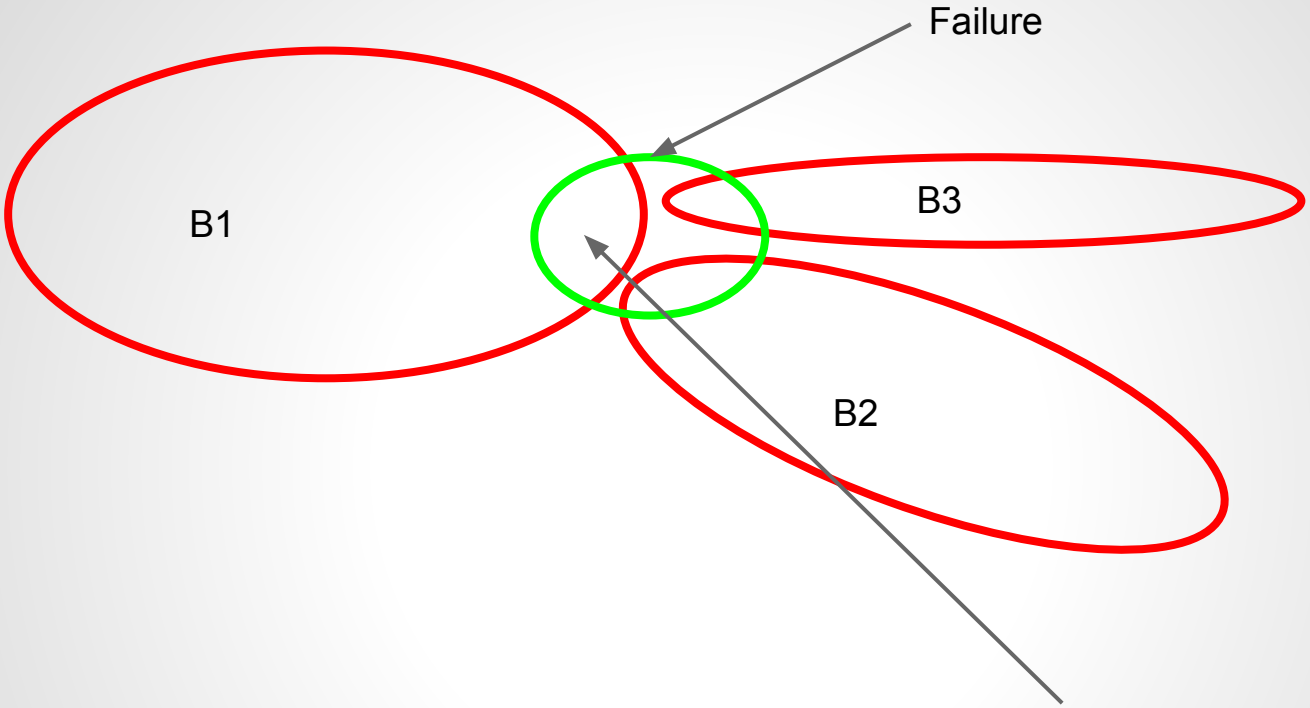

P{B,A}
A
B

Simplification of formula

- There are 2 neighbors A and B that share a common land.
- We know that the common land is ⅓ of B's territory: P {A|B}
- We know both A's and B's surface: P{A} , P{B}
- How much those the common land represent of A's territory (percent)? : P{B|A}?

Example 1

- A factory produces the same items using machines B1, B2 and B3.
- B1 produces 60% of items with 6% failure
- B2 produces 30% of items with 3% failure
- B3 produces 10% of items with 5% failure
- If i pick a failed product randomly, what is the probability to be produced from B1,B2 or B3?



Probability that an element made by B1 to be failed? = P{B1|F}
Probability that an element made by B2 to be failed? = P{B2|F}
Probability that an element made by B3 to be failed? = P{B3|F}

- P{B1|F}= P{F|B1}*P{B1}/P{F}= number of items broken made with B1 / total number of broken items
- P{F|B1}= 6% of items made by B1
- P{B1}= 60% of total items made by all machines
- P{F|B1}*P{B1} = 0.06 * 0.6 = 0.036% of all elements made are broken and made with B1
- P{F}= P{F|B1} + P{F|B2} + P{F|B3}= 0.06*0.6 + 0.03*0.3 + 0.05*0.1= 0.05 % of all elements are broken
- P{B1|F}= 0.036/0.05 = 72% chance to pick an element made by B1 if we chose randomly from the broken elements

Ex 1.

What about P{B2|F} and P{B3|F}?

Naive bayes classification

It is called naive or idiot because it supposes that there is no dependency between attributes.

Is that ok?

Usually attributes depend in some way on the other attributes!

The horsepower and a car's speed depend on each other, but naive bayes algorithm does not consider that.

The greatest advantage of Naive Bayes algorithm is that it needs a small training set!

Classification

Classification of instance I having attributes Ia1 ... Ian.

Classes: ClassA, ClassB

P{ClassA|I}= P{Ia1|ClassA} * P{Ia2|ClassA}* … * P{Ian|ClassA} * P{ClassA}

P{ClassB|I}= P{Ia1|ClassB} * P{Ia2|ClassB}* … * P{Ian|ClassB} * P{ClassB}

If P{ClassA|I} > P{ClassB|I} then : I is classified as ClassA

else: I is classified as ClassB

Example 2

Salary	Has car?	Marital status	Happy
100k	Yes	No	Yes
100k	Yes	Yes	No
120k	Yes	Yes	Yes
60k	No	No	Yes
60k	No	Yes	No

A guy with a salary of 100k with a wife and no car is he happy?

P{100k,happy=yes} = ½

P{100k,happy=no} = ½

P{car=no,happy=yes}=½

P{car=no,happy=no}=½

P{wife=yes,happy=yes}=⅓

P{wife=yes,happy=no}=⅔

P{happy=yes}=⅗

P{happy=no}=⅖

P{100k,car=no,wife=yes|happy=yes}=½*½*⅓*⅗=3/60

P{100k,car=no,wife=yes|happy=no}=½*½*⅔*⅖=4/60

No he is not happy

Ex. 2

Salary	Has car?	Marital status	Happy
100k	Yes	No	Yes
100k	Yes	Yes	No
120k	Yes	Yes	No
60k	No	No	Yes

A guy with a salary of 100k with a wife and no car is he happy?

Problem

With the current method of calculation when

P{married=yes,happy=yes}=0 which makes

P{100k,married=yes,car=no|happy=no}=0

The occurrence of 0 in calculations might mess up the classification!

The solution was to calculate probabilities from frequencies!

Ex. 3

- open iris.arff
- classify using NaiveBayes classifier
- see the result

Ex. 4

- open iris.arff
- classify using the BayesNet classifier
- where does the difference in precision of classification come?

Ex. 5

- compare NaiveBayes, J48 and ZeroR on:
  - diabetes.arff
  - supermarket.arff

- which is the best algorithm?