Naive_Bayes_Alogrithm

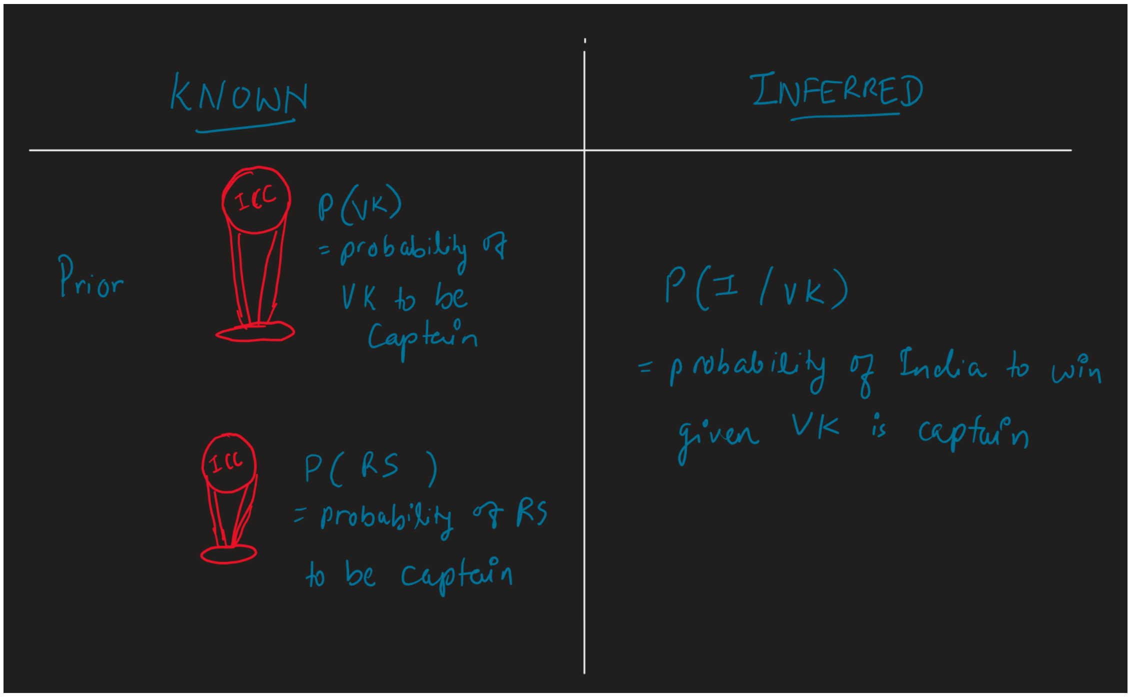
January 29, 2022

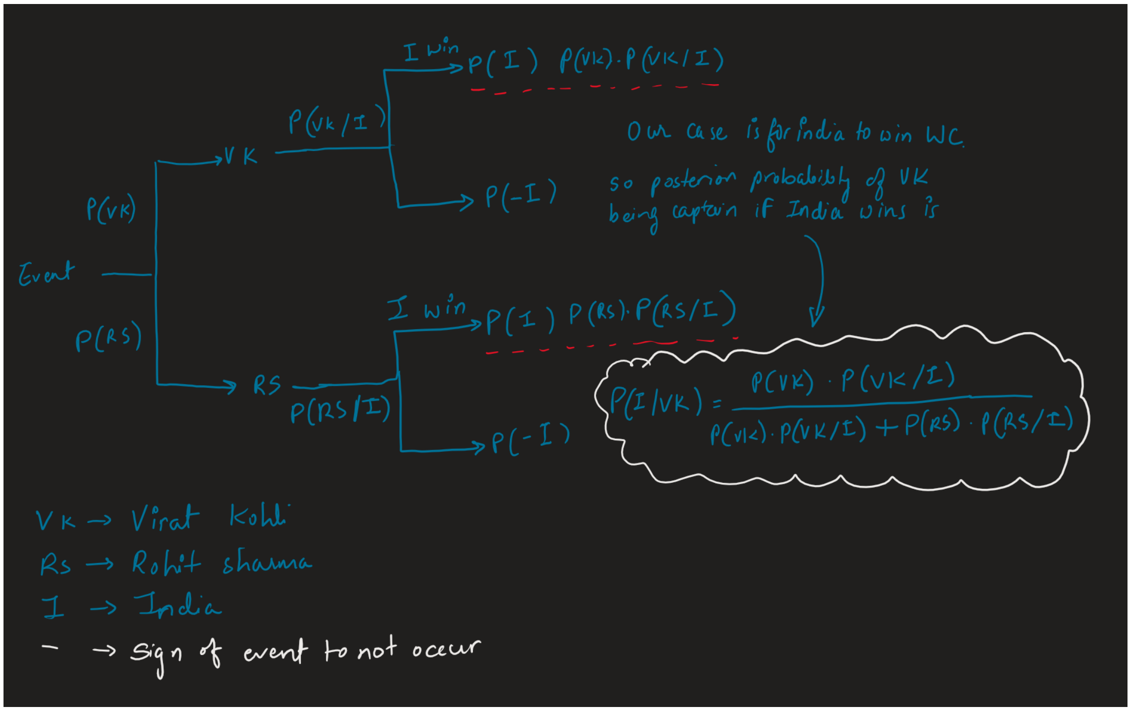
1 Naive Bayes (Bayesian Algorithm)

Naive Bayes algorithm is on one of the supervised learning techniques used for classification types of problems. The core of this algorithm is the probabilistic model which calculates the probabilities of input features and outputs the probalities distributions to each possible class.This algorithm is used for Natural Language processing(NLP) types of application.

1.0.1 Bayes Theorem

The main concept of the Bayes theorem is based on what probabilites we know and what we can infer from it. In this model, there are two probablities, prior and posterior. Prior is the probility of a given event to occur, say what is the probability of **Virat Kohli** or **Rohit Sharma** to be the Captain. Posterior is the probility of a given event to occur given a conditon, say what is the probability of Inida to win **given** Virat Kohli is the Captain. That's it, that's the Bayes Theorem.


KNOWN
INFERRED
Prior
ICC
P(VK)
= probability of VK to be Captain
ICC
P(RS)
= probability of RS to be captain
P(I / VK)
= probability of India to win given VK is captain


I win
P(I) P(VK)·P(VK/I)
P(VK/I)
VK
P(VK)
Event
P(-I)
Our case is for India to win WC.
so posterior probability of VK being captain if India wins is
I win
P(I) P(RS)·P(RS/I)
P(RS)
RS
P(RS/I)
P(-I)
P(I/VK) = P(VK) · P(VK/I) / P(VK)·P(VK/I) + P(RS) · P(RS/I)
VK → Virat Kohli
Rs → Rohit sharma
I → India
− → Sign of event to not occur

1.1 Naive Bayes Algorithm

The algorithm is based on the condition that two events for which we calculate the probablity should be different. So in P(A ∩ B) , event A and event B should be independent of each other. This is the naive assumption that we make for this algorithm.

$P(A \cap B) = P(A) \,.\, P(B)$

1.1.1 Problem Statement 1:

Suppose you have a bag with three standard 6-sided dice with face values [1,2,3,4,5,6] and two non-standard 6-sided dice with face values [2,3,3,4,4,5]. Someone draws a die from the bag, rolls it, and announces it was a 3. What is the probability that the die that was rolled was a standard die?

Input your answer as a fraction or as a decimal with at least three digits of precision.

[12]:
```
std_dice = 3
non_std_dice =2
P_A = std_dice / (std_dice+non_std_dice)
P_B = non_std_dice / (std_dice+non_std_dice)
print("Probability of picking a standard dice is {} and Probability of picking␣
 ↪a non-standard dice is {}".format(P_A, P_B) )

print("After rolling the dice the number was 3")

P_3_A = 1/6 * P_A
P_3_B = 2/6 * P_B
```

```python
print("Probability of having number 3 given its a std die: {:0.3f}␣
 ↪\nProbability of having number 3 given its a non std die: {:0.3f}".
 ↪format(P_3_A,P_3_B))
#Normalizing to return the probability of die that was rolled was a standard die
P_A_3 = P_3_A/(P_3_A+P_3_B)
print("Probablity having a std die given that the rolled dice is having number␣
 ↪3: {:0.3f}".format(P_A_3))
```

```
Probability of picking a standard dice is 0.6 and Probability of picking a non-
standard dice is 0.4
After rolling the dice the number was 3
Probability of having number 3 given its a std die: 0.100
Probability of having number 3 given its a non std die: 0.133
Probablity having a std die given that the rolled dice is having number 3: 0.429
```

1.1.2 Spam Classifier Modeling:

The modeling of spam classifier is done using the Naive Bayes algorithm and is carried out in a diffrent workspace notebook. Refer to the github link given below to see the notebook here: https://github.com/hseju/Udacity-Intro-To-ML-TensorFlow/tree/main/Classification/Naive%20Bayes

```
[ ]:
```