Artificial Intelligence

Dr. Mubashir Ahmad (Ph.D.)

Probabilistic reasoning

- Probabilistic reasoning is a key aspect of artificial intelligence (AI) that allows for handling uncertainty and ambiguity in decision-making. It is a powerful technique that enables AI systems to make informed decisions even when faced with incomplete or noisy data. Probabilistic reasoning is widely used in various AI applications such as machine learning, natural language processing, robotics, computer vision, and many more.

Probabilistic reasoning

- AI has made great strides in recent years, and probabilistic reasoning has played a significant role in many of these advancements. Probabilistic reasoning is making decisions based on probabilities and likelihoods rather than absolute facts.
- At its core, probabilistic reasoning involves reasoning about probabilities or likelihoods of events or outcomes. It allows AI systems to model and reason about uncertain situations in a principled and quantitative manner, considering the inherent uncertainty in real-world data and evidence.

Uncertainty

- From self-driving cars to virtual personal assistants, AI technologies have become integral to our daily routines. However, one of the key challenges that AI systems face is dealing with uncertainty. Uncertainty arises due to various factors such as unreliable sources of Information, experimental errors, equipment faults, temperature variations, and climate change, among others. To address this challenge, probabilistic reasoning techniques have gained significant importance in AI, allowing machines to make decisions and predictions in uncertainty.

Causes of Uncertainty

- **Information Occurred from Unreliable Sources:**
- AI systems rely on data to make decisions and predictions. However, data obtained from various sources may not always be reliable. Data can be incomplete, inconsistent, or biased, leading to uncertainty in the outcomes generated by AI systems.
- **Experimental Errors:**
- In scientific research and experimentation, errors can occur at various stages, such as data collection, measurement, and analysis. These errors can introduce uncertainty in the results and conclusions drawn from the experiments.

Causes of Uncertainty

- **Equipment Fault:**
- In many AI systems, machines and sensors are used to collect data and make decisions. However, these machines can be subject to faults, malfunctions, or inaccuracies, leading to uncertainty in the outcomes generated by AI systems.
- **Temperature Variation:**
- Many real-world applications of AI, such as weather prediction, environmental monitoring, and energy management, are sensitive to temperature variations. However, temperature measurements can be subject to uncertainty due to factors such as sensor accuracy, calibration errors, and environmental fluctuations.

Causes of Uncertainty

- **Climate Change:**
- Climate change is a global phenomenon that introduces uncertainty in various aspects of our lives. For example, predicting the impacts of climate change on agriculture, water resources, and infrastructure requires dealing with uncertain data and models.

Probabilistic reasoning

- Probabilistic reasoning is a technique used in AI to address uncertainty by modeling and reasoning with probabilistic Information. It allows AI systems to make decisions and predictions based on the probabilities of different outcomes, taking into account uncertain or incomplete Information. Probabilistic reasoning provides a principled approach to handling uncertainty, allowing machines to reason about uncertain situations in a rigorous and quantitative manner.

Need for Probabilistic Reasoning in AI

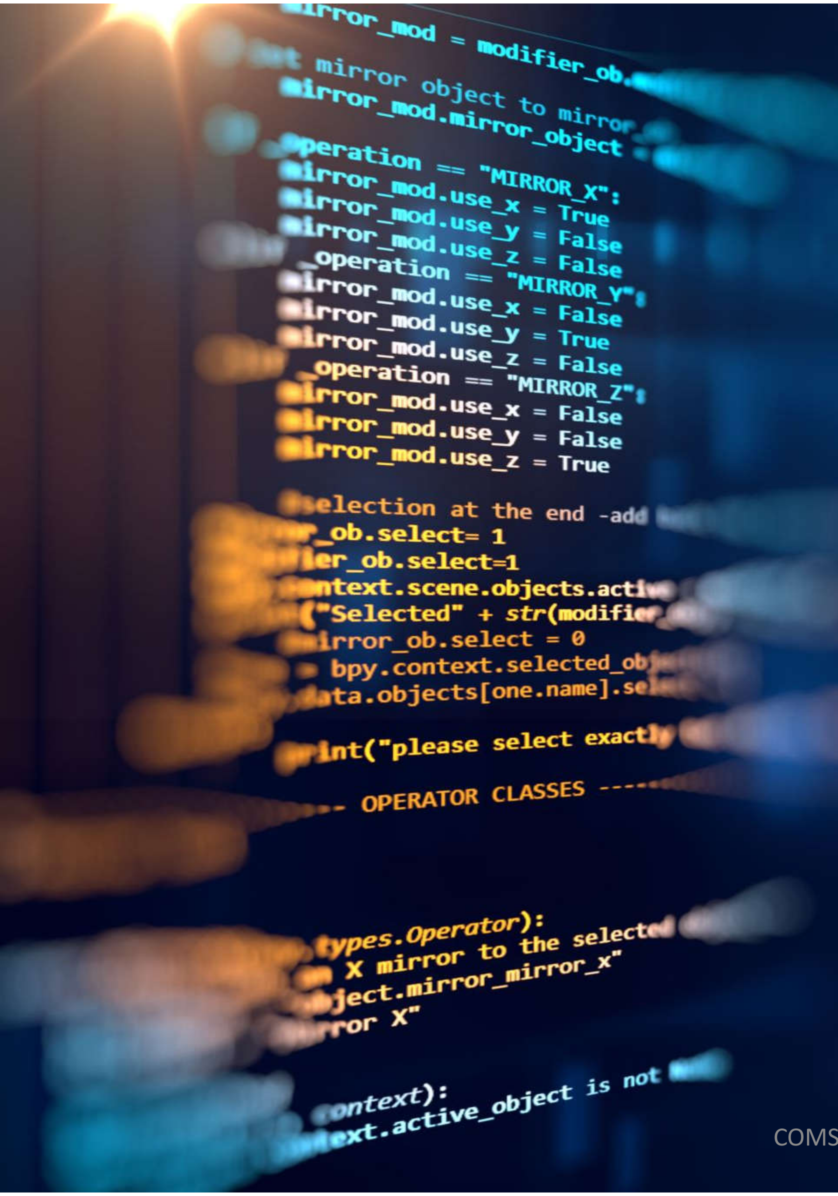
- The need for probabilistic reasoning in AI arises because uncertainty is inherent in many real-world applications. **For example,** there is often uncertainty in the symptoms, test results, and patient history in medical diagnosis. In autonomous vehicles, there is uncertainty in the sensor measurements, road conditions, and traffic patterns. In financial markets, there is uncertainty in stock prices, economic indicators, and investor behavior. Probabilistic reasoning techniques allow AI systems to deal with these uncertainties and make informed decisions.

Solve Problems with Uncertain Knowledge

- **Bayes' Rule:**
- Bayes' rule is a fundamental theorem in probability theory that allows updating probabilities based on new evidence. It provides a principled way to combine prior knowledge with new data to update the probabilities of different outcomes. Bayes' rule has been widely used in AI for classification, prediction, and decision-making tasks where uncertainty needs to be addressed.
- Mathematically, Bayes' Theorem is expressed as:
- $P(A|B) = (P(B|A) * P(A)) / P(B)$

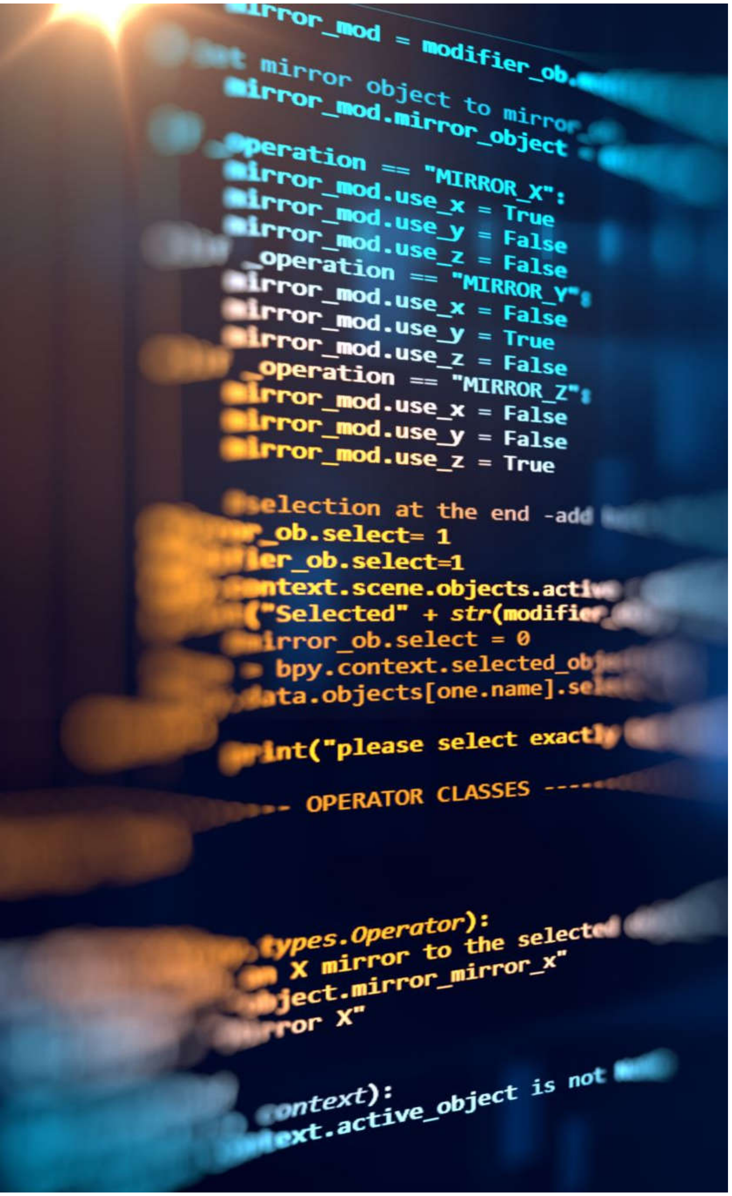
Bayes' Rule

- **P(A|B)** represents the posterior probability, which is the probability of event A occurring given that event B has occurred.
- **P(B|A)** represents the likelihood, which is the probability of observing event B given that event A has occurred.
- **P(A)** represents the prior probability, which is the initial probability of event A occurring before considering any new evidence.
- **P(B)** represents the marginal likelihood, which is the probability of observing event B, regardless of whether event A has occurred.
- In the context of AI, Bayes' Theorem is used to update the probabilities of different hypotheses or predictions based on new data or evidence. It is particularly useful in handling uncertainty and making decisions when there is incomplete or ambiguous Information.



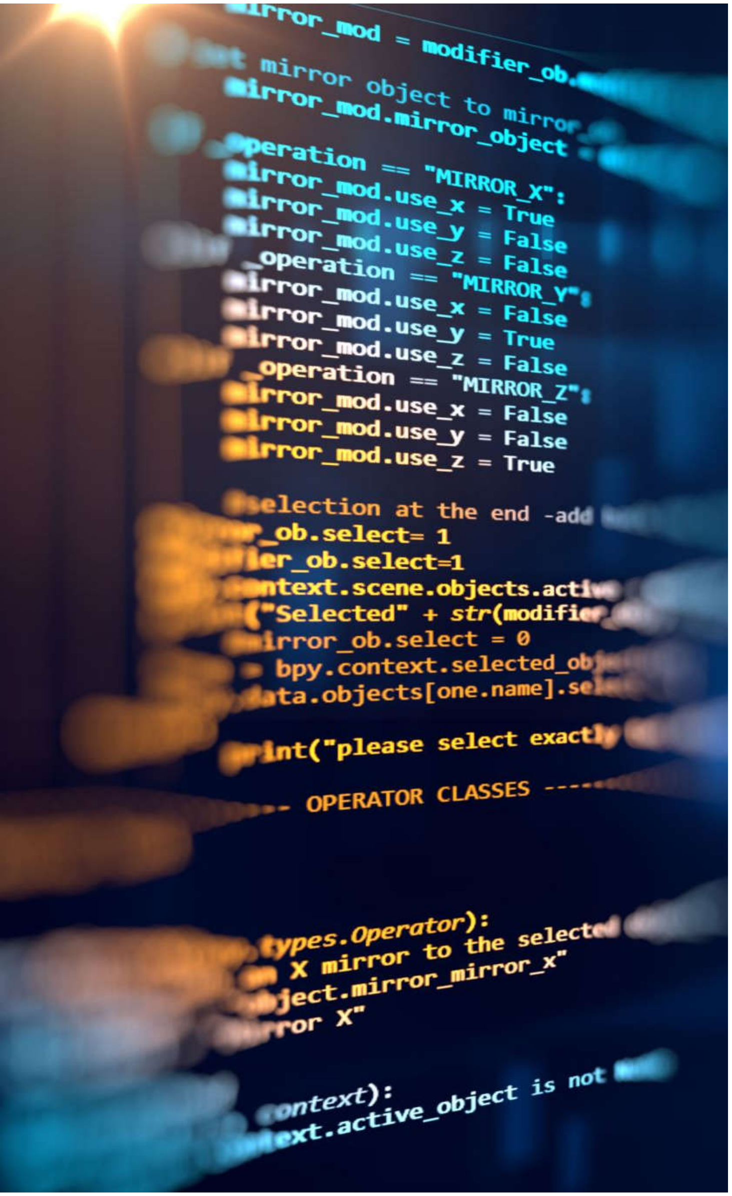
Naïve Bayes Classifier Algorithm

- Naïve Bayes algorithm is a supervised learning algorithm, which is based on **Bayes theorem** and used for solving classification problems.
- It is mainly used in *text classification* that includes a high-dimensional training dataset.
- Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.



Naïve Bayes Classifier Algorithm

- It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.
- Some popular examples of Naïve Bayes Algorithm are spam filtration,
- Sentimental analysis and classifying articles.
- Naïve: It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the basis of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. Hence each feature individually contributes to identify that it is an apple without depending on each other.
- Bayes: It is called Bayes because it depends on the principle of Bayes' Theorem.



Working of Naïve Bayes' Classifier:

- Suppose we have a dataset of **weather conditions** and the corresponding target variable "**Play**". So using this dataset we need to decide whether we should play or not on a particular day according to the weather conditions. So to solve this problem, we need to follow the following steps:

1. Convert the given dataset into frequency tables.
2. Generate a Likelihood table by finding the probabilities of given features.
3. Now, use Bayes theorem to calculate the posterior probability.

Naïve Bayes' Base Theorem

- Bayes theorem is the cornerstone of Naïve Bayes Classifier because it provides a way to calculate the posterior probability **P(h|D), from the prior** probability **P(h), together with P(D) and P(D|h).**

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)}$$

Posterior probability

(P(D) = P(y) + P(N)

```
D is the datset and h is the hypothesis, for example yes/no
The hypothesis which give the maximum value, you consider it
as classification results.
```

Naïve Bayes' Base Theorem

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)}$$

Posterior probability

$$\begin{aligned} h_{MAP} &\equiv \underset{h \in H}{\operatorname{argmax}}\ P(h|D) \\ &= \underset{h \in H}{\operatorname{argmax}}\ \frac{P(D|h)P(h)}{\cancel{P(D)}} \\ &= \underset{h \in H}{\operatorname{argmax}}\ P(D|h)P(h) \end{aligned}$$

maximum Hypothesis

P(D) is divided with every hypothesis

Naïve Bayes' Base Theorem

- In real world we can not use D as individual, because it's a dataset and datasets have multiple number of attributes.

- The Bayesian approach to classifying the new instance is to assign the most probable target value, $\mathbf{v}_{\mathbf{MAP}}$ given the attribute values **<a1, a2 . . .an> that describe** the instance.

$$v_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} P(v_j | a_1, a_2 \dots a_n)$$

- We can use Bayes theorem to rewrite this expression as

$$v_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} \frac{P(a_1, a_2 \dots a_n | v_j) P(v_j)}{\cancel{P(a_1, a_2 \dots a_n)}}$$

$$= \underset{v_j \in V}{\operatorname{argmax}} P(a_1, a_2 \dots a_n | v_j) P(v_j)$$

$$\boxed{\begin{aligned} h_{MAP} &\equiv \underset{h \in H}{\operatorname{argmax}} P(h|D) \\ &= \underset{h \in H}{\operatorname{argmax}} \frac{P(D|h)P(h)}{P(D)} \\ &= \underset{h \in H}{\operatorname{argmax}} P(D|h)P(h) \end{aligned}}$$

- Naïve Bayes Classifier: $v_{NB} = \underset{v_j \epsilon V}{\operatorname{argmax}} \, p(v_j) \prod_i p(a_i | v_j)$

Hypothesis

Solved Example 1 Naïve Bayesian classifier

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

$\langle Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong \rangle$?

Solved Example 1: Get the prior probability

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

$$\langle Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong \rangle \text{ ?}$$

$$P(PlayTennis = yes) = 9/14 = .64$$

$$P(PlayTennis = no) = 5/14 = .36$$

Prior Probabilities

Solved Example 1: conditional probability of individual attribute.

Total "yes" are 9 and total "no" are 5

Conditional Probabilities

Outlook	Y	N
sunny	2/9	3/5
overcast	4/9	0
rain	3/9	2/5
Tempreature		
hot	2/9	2/5
mild	4/9	2/5
cool	3/9	1/5

Humidity	Y	N
high	3/9	4/5
normal	6/9	1/5
Windy		
Strong	3/9	3/5
Weak	6/9	2/5

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

$\langle Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong \rangle$?

The new instance is given in the first row, on the basis of these assumptions we can say yes or no.

$$\langle Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong \rangle$$?

$$v_{NB} = \underset{v_j \in \{yes, no\}}{\operatorname{argmax}} P(v_j) \prod_i P(a_i | v_j)$$

set of attribute

Possible outcome

$$= \underset{v_j \in \{yes, no\}}{\operatorname{argmax}} P(v_j) \quad P(Outlook = sunny|v_j) P(Temperature = cool|v_j) \cdot P(Humidity = high|v_j) P(Wind = strong|v_j)$$

$$v_{NB}(yes) = P(yes)\ P(sunny|yes)\ P(cool|yes)\ P(high|yes)\ P(strong|yes) = .0053$$

$$v_{NB}(no) = P(no)\ P(sunny|no)\ P(cool|no)\ P(high|no)\ P(strong|no) = .0206$$

$$\boxed{v_{NB}(yes) = \frac{v_{NB}(yes)}{v_{NB}(yes) + v_{NB}(no)} = 0.205 \qquad v_{NB}(no) = \frac{v_{NB}(no)}{v_{NB}(yes) + v_{NB}(no)} = 0.795}$$

Normalize

Outlook	Y	N
sunny	2/9	3/5
overcast	4/9	0
rain	3/9	2/5

Tempreature		
hot	2/9	2/5
mild	4/9	2/5
cool	3/9	1/5

Humidity	Y	N
high	3/9	4/5
normal	6/9	1/5

Windy		
Strong	3/9	3/5
Weak	6/9	2/5

If add both probabilities then it will be equal to 1.

NAIVE BAYES CLASSIFIER Example - 2

- Estimate conditional probabilities of each attribute {color, legs, height, smelly} for the species classes: {M, H} using the data given in the table.
- Using these probabilities estimate the probability values for the new instance - (Color=Green, legs=2, Height=Tall, and Smelly=No).

NAIVE BAYES CLASSIFIER Example - 2

No	Color	Legs	Height	Smelly	Species
1	White	3	Short	Yes	M
2	Green	2	Tall	No	M
3	Green	3	Short	Yes	M
4	White	3	Short	Yes	M
5	Green	2	Short	No	H
6	White	2	Tall	No	H
7	White	2	Tall	No	H
8	White	2	Short	Yes	H

NAIVE BAYES CLASSIFIER Example - 2

To calculate the prior and conditional probabilities

No	Color	Legs	Height	Smelly	Species
1	White	3	Short	Yes	M
2	Green	2	Tall	No	M
3	Green	3	Short	Yes	M
4	White	3	Short	Yes	M
5	Green	2	Short	No	H
6	White	2	Tall	No	H
7	White	2	Tall	No	H
8	White	2	Short	Yes	H

$$P(M) = \frac{4}{8} = 0.5 \quad P(H) = \frac{4}{8} = 0.5$$

Color	M	H
White	2/4	3/4
Green	2/4	1/4

Legs	M	H
2	1/4	4/4
3	3/4	0/4

Height	M	H
Tall	1/4	2/4
Short	3/4	2/4

Smelly	M	H
Yes	3/4	1/4
No	1/4	3/4

NAIVE BAYES CLASSIFIER Example - 2

$$P(M) = \frac{4}{8} = 0.5 \quad P(H) = \frac{4}{8} = 0.5$$

Color	M	H
White	2/4	3/4
Green	2/4	1/4

Legs	M	H
2	1/4	4/4
3	3/4	0/4

Height	M	H
Tall	1/4	2/4
Short	3/4	2/4

Smelly	M	H
Yes	3/4	1/4
No	1/4	3/4

$$p(M|New\ Instance) = p(M) * p(Color = Green|M) * p(Legs = 2|M) * p(Height = tall|M) * p(Smelly = no\ |M)$$

$$p(M|New\ Instance) = 0.5 * \frac{2}{4} * \frac{1}{4} * \frac{1}{4} * \frac{1}{4} = 0.0039$$

$$p(H|New\ Instance) = p(H) * p(Color = Green|H) * p(Legs = 2|H) * p(Height = tall|H) * p(Smelly = no\ |H)$$

$$p(H|New\ Instance) = 0.5 * \frac{1}{4} * \frac{4}{4} * \frac{2}{4} * \frac{3}{4} = 0.047$$

$$p(H|New\ Instance) > p(M|New\ Instance)$$

Hence the new instance belongs to Speices H

NAIVE BAYES CLASSIFIER Example - 3

Record	A	B	C	Class
1	0	0	0	+
2	0	0	1	−
3	0	1	1	−
4	0	1	1	−
5	0	0	1	+
6	1	0	1	+
7	1	0	1	−
8	1	0	1	−
9	1	1	1	+
10	1	0	1	+

Consider the dataset given in the table and

1. Estimate the conditional probabilities of P(A|+), P(B|+), P(C|+), P(A|-), P(B|-) and P(C|-)
2. Use the conditional probability estimates and predict the class

$$P(+) = \frac{5}{10} = 0.5$$

$$P(-) = \frac{5}{10} = 0.5$$

NAIVE BAYES CLASSIFIER Example - 3

Record	A	B	C	Class
1	0	0	0	+
2	0	0	1	−
3	0	1	1	−
4	0	1	1	−
5	0	0	1	+
6	1	0	1	+
7	1	0	1	−
8	1	0	1	−
9	1	1	1	+
10	1	0	1	+

A	0	1
+	2/5	3/5
-	3/5	2/5

B	0	1
+	4/5	1/5
-	3/5	2/5

C	0	1
+	1/5	4/5
-	0/5	5/5

NAIVE BAYES CLASSIFIER Example - 3

Record	A	B	C	Class
1	0	0	0	+
2	0	0	1	−
3	0	1	1	−
4	0	1	1	−
5	0	0	1	+
6	1	0	1	+
7	1	0	1	−
8	1	0	1	−
9	1	1	1	+
10	1	0	1	+

$$New \Rightarrow P(A=0, B=1, C=0)$$

$$P(+|New) = P(+) * P(A=0|+) * P(B=1|+) * P(C=0|+)$$

$$= 0.5 * \frac{2}{5} * \frac{1}{5} * \frac{1}{5} = 0.008$$

$$P(-|New) = P(-) * P(A=0|-) * P(B=1|-) * P(C=0|-)$$

$$= 0.5 * \frac{3}{5} * \frac{2}{5} * \frac{0}{5} = 0.0$$

$$\boxed{P(+|New) > P(-|New)}$$

+