

Question 1)

۱) چگونی ساخته شدن درخت ها: در Random forest، هر درخت یک دسته میگیرد که درخت های تکاره بازدید و تا ساخت درخت تمام شده، نه توانند درخت بعدی را بازدید نمایند. درخت های تکاره بازدید و تا ساخت درخت تمام شده، نه توانند و همین باعث بهبود معايي مدل می شوند. معمولاً درخت هایی می شوند. ✓

ترکیب نتایج: در Random forest، تایخ (رای تکی) روسه باهم ترکیب می‌شوند (با وسیلهٔ میانیان رفتن و با "majority rules" در "gradient boosting" (مدل روسه)، تایخ باهم ترکیب می‌شوند.

کاپیدار داد چون دین دو کاپیدار باشد noise یعنی مهارم می شوند.
دی روشنها در جا های متفاوت استاده می شوند. برای مثال
multi-class object detection Random forest و bioinformatics (B)
کاپیدار داد چون دین دو کاپیدار باشد noise یعنی مهارم می شوند.
دی روشنها در جا های متفاوت استاده می شوند. برای مثال
real-time risk in unbalanced-data assessment کاپیدار پیشتری دارد.
Gradient boosting

ص

پرسی حل این سوال، باید root level و root Entropy را درست تفهیم کنیم. و زیرا جایگاه در درست تفهیم از بالا پایین، Entropy بیکاره محسوس میدارد (نحوه این label های نهایی برسی) پس آنکه باید این اسیه که level خاص از طبقه درست، Entropy افزایش میدارد، یعنی در طبقه نهاده است. همچنانکه، میکه هستکل وجود دارد است.

① level 1 :

$$\begin{array}{c} \text{orange} \quad \text{green} \\ 10 \quad + V \end{array} \Rightarrow E(r_1) = - \sum_{i=1}^N p_i \log p_i = -\frac{10}{14} \log \frac{10}{14} - \frac{V}{14} \log \frac{V}{14} = 0.2942$$

② level 2 :

$$\begin{array}{c} \text{orange} \quad \text{green} \\ 1 \quad + 9 \end{array} \Rightarrow E(r_2) = -\frac{1}{10} \log \frac{1}{10} - \frac{9}{10} \log \frac{9}{10} = 0$$

$$\begin{array}{c} \text{orange} \quad \text{green} \\ 9 \quad + V \end{array} \Rightarrow E(r_2) = -\frac{9}{19} \log \frac{9}{19} - \frac{V}{19} \log \frac{V}{19} = 0.24$$

③ level 3 :

$$\begin{array}{c} \text{orange} \quad \text{green} \\ 1 \quad + 8 \end{array} \Rightarrow E(r_3) = -\frac{1}{8} \log \frac{1}{8} - \frac{7}{8} \log \frac{7}{8} = 0.2174$$

$$\begin{array}{c} \text{orange} \quad \text{green} \\ 1 \quad + 8 \end{array} \Rightarrow E(r_3) = -\frac{1}{8} \log \frac{1}{8} - \frac{7}{8} \log \frac{7}{8} = 0.2442 *$$

④ level 4 :

$$\begin{array}{c} \text{orange} \quad \text{green} \\ 1 \quad + 7 \end{array} \Rightarrow E(r_4) = -\frac{1}{7} \log \frac{1}{7} - \frac{6}{7} \log \frac{6}{7} = 0.2748 *$$

آسترودیجی (نیمه) رسیه ها و level اینکه به

level 1 رسیده اند، به این صفات است.

همانطوره که در بالا می بینیم، آسترودیجی (سی) سمت راست در level 2 باید 0.2442 است. و می آسترودیجی اسیه سمت راست در level 3 باید 0.2642 است. یعنی بحسب پایین درست راهنمای دی آسترودیجی افزایش میدارد. و این همان مسئله است که درست سهیل بیان شده است. از طرفی این استناد در رسیه سمت دیگر در level 4 درست هم می شود. میان آسترودیجی اسیه 0.2748 و 0.2442 است. $0.2748 > 0.2442$ است. $0.2442 < 0.2748$ است. $0.2442 < 0.2748$ است.

$$\Rightarrow \text{accuracy} = \frac{10 + 100}{10 + 100 + 10 + 10} = \sqrt{p_1 p_2}$$

مدل‌های است فله کننده /
بیان فنی مدل‌های classifier
ایمیل های ایمیل های spam
کروماتوگرام دسته بندی کردن
دقت دهنده دهنده metric
accuracy

فیلم و ویدیو های اینستاگرام "No Spam" میگویند (پیشنهاد و روش) که از فقط در یک کلاس نباشد تا همچنان

(original)	Classified Positive	Classified Negative	
positive class	$o \Rightarrow TP$	$\text{No } \Rightarrow FN$	$\Rightarrow \text{total cases} = 140 \rightarrow 16, \text{No } 16$
Negative class	$o \Rightarrow FP$	$140 \Rightarrow TN$	

$$\text{recall} = \frac{TP}{TP + FN} \quad (1)$$

$$F_\beta\text{-Score} = \frac{(1+\beta^2) \times \text{Precision} \times \text{Recall}}{(\beta^2 \times \text{Precision} + \text{Recall})} \quad (2)$$

$$F1\text{-score} = \frac{P \times \text{precision} \times \text{recall}}{P + \text{precision} + \text{recall}} \quad (1)$$

Confusion Matrix (2)

✓

Question 4)

Y₀ classifiers

$N = \text{var } \mathcal{U}_1 \cup \mathcal{U}_2$ classifier \mathcal{L}

چون صاحب کدن این مقدار \rightarrow بجهالت دستی طبق صورت این جواب را می‌کند در پاسخ این جواب را می‌کند که در پاسخ داده که د. ن. فایصل را در حقیقت عنوان "question4-theory.py" در فایل "supporting Materials" مارداده ام

$$\textcircled{1} \quad E(\text{Edible}) = H(\alpha, \beta) = -\frac{\alpha}{16} \log \frac{\alpha}{16} - \frac{\beta}{16} \log \frac{\beta}{16} = 0.9404 = \text{class}$$

Question 5)

Edible:

	Poisonous	Edible
Woods	1/4	3/4
Grasses	0	1
Leaves	1/4	3/4

$$\textcircled{1} \quad \text{Habitat:} \quad \Rightarrow E(\text{Habitat} = \text{Woods}) = -\frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4} = 0.941 \rightarrow$$

$$\Rightarrow E(\text{Habitat} = \text{Grasses}) = -\frac{0}{4} \log \frac{0}{4} - \frac{1}{4} \log \frac{1}{4} = 0$$

$$\Rightarrow E(\text{Habitat} = \text{Leaves}) = E(1/4, 3/4) = 0.941$$

$$\Rightarrow E(\text{Edible} | \text{Habitat}) = P(\text{Habitat} = \text{Woods}) E(\text{Edible} | \text{Habitat} = \text{Woods}) + P(\text{Habitat} = \text{Grasses}) E(\text{Edible} | \text{Habitat} = \text{Grasses}) + P(\text{Habitat} = \text{Leaves}) E(\text{Edible} | \text{Habitat} = \text{Leaves}) = \frac{1}{3} (0.941) + \frac{1}{3} (0) + \frac{1}{3} (0.941) = 0.944$$

$$\Rightarrow \text{Information Gain} = E(\text{Edible}) - E(\text{Edible} | \text{Habitat}) = 0.9404 - 0.944 = 0.004$$

Edible:

	Poisonous	Edible	$= E(1/4, 3/4)$
Red	1/4	3/4	$\Rightarrow E(\text{Cap color} = \text{Red}) = -\frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4} = 0.9999$
Green	1/4	3/4	$\Rightarrow E(\text{Cap color} = \text{Green}) = -\frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4} = 0.9111$
White	1/2	1/2	$\Rightarrow E(\text{Cap color} = \text{White}) = -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} = 0.1111$

$$\textcircled{2} \quad \text{Cap color:} \quad \underbrace{P(\text{Cap color} = \text{Red})}_{\text{Probability}} \underbrace{E(\text{Edible} | \text{Cap color} = \text{Red})}_{\text{Information}}$$

$$\Rightarrow E(\text{Edible}, \text{Cap color}) = P(\text{Red}) E(\text{Cap color} = \text{Red}) + P(\text{Green}) E(\text{Cap color} = \text{Green}) + P(\text{White}) E(\text{Cap color} = \text{White})$$

$$\underbrace{E(\text{Edible} | \text{Cap color})}_{\text{Probability}} = I(\text{Cap color}) = \text{Information about Cap color} \text{ for class } \text{Edible}$$

$$= \frac{1}{3} (0.9999) + \frac{1}{3} (0.9111) + \frac{1}{3} (0.1111) = 0.9109$$

$$\Rightarrow \text{Information Gain} = E(\text{Edible}) - E(\text{Edible}, \text{Cap color}) = 0.9404 - 0.9109 = 0.0294$$

Edible:

	Poisonous	Edible
flat	1/2	1/2
convex	1/2	1/2

$$\textcircled{3} \quad \text{Cap shape:} \quad \Rightarrow E(\text{Cap shape} = \text{flat}) = -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} = 0.9109$$

$$\Rightarrow E(\text{Cap shape} = \text{convex}) = -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} = 0.9109$$

$$\Rightarrow E(\text{Edible}, \text{Cap shape}) = P(\text{flat}) E(\text{Edible}, \text{Cap shape} = \text{flat}) + P(\text{convex}) E(\text{Edible}, \text{Cap shape} = \text{convex}) =$$

$$\frac{1}{2} (0.9109) + \frac{1}{2} (0.9109) = 0.9109$$

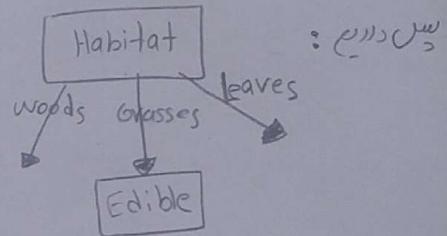
$$\Rightarrow \text{Information Gain} = 0.9404 - 0.9109 = 0.0294$$

	Poisonous	Edible
None	Y	Y
Foul	W	W

$$\frac{1}{10} (0,11110) + \frac{4}{10} (0,9999) = 0,18910$$

$$\Rightarrow \text{Information Gain} = 0,94544 - 0,17921 = \underline{0,76623}$$

Habitat	...	Habitat	...	جدول ١١
Leaves		woods		
N :	...	N :	...	
Leaves		woods		



① class = Woods = class

9 cap color
"woods" (5) N.

$$\Rightarrow \text{Information Gain} = \underline{\underline{0.19V1}} - 0.1K = \boxed{0.19V1}$$

Edible:

		Poisonous	Edible	
		flat	o	$\Rightarrow E(\text{cap shape} = \text{flat}) = 0$
		Convex	o	$\Rightarrow E(\text{cap shape} = \text{convex}) = 0$

$$\Rightarrow E(\text{Edible, cap shape}) = P(\text{flat})E(\text{Edible, cap shape} = \text{Edible}) + P(\text{convex})E(\text{Edible, cap shape} = \text{convex})$$

$$= \frac{1}{2}(0) + \frac{1}{2}(0) = 0$$

$$\Rightarrow \text{Information Gain} = 0.9V1 - 0 = 0.9V1$$

Edible:

		Poisonous	Edible	
		None	1	$\Rightarrow E(\text{odor} = \text{None}) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 0.911V1$
		Foul	1	$\Rightarrow E(\text{odor} = \text{foul}) = 1$

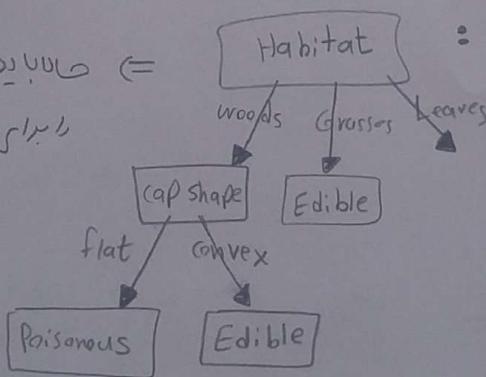
$$\Rightarrow E(\text{Edible, odor}) = P(\text{none})E(\text{Edible, odor} = \text{none}) + P(\text{foul})E(\text{Edible, odor} = \text{foul}) =$$

$$\frac{1}{2}(0.911V1) + \frac{1}{2}(1) = 0.9V1$$

$$\Rightarrow \text{Information Gain} = 0.9V1 - 0.9V1 = 0.9V0$$

رسی مفاسدی Information Gain سی انتخاب کنیم. رسی مفاسدی Sub-tree :

Root "cap shape" "woods" "Leaves" =



② Class = Leaves = class

$$\Rightarrow E(\text{Edible}) = E(1, 1) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 0.9V1$$

Leaves (1,1)

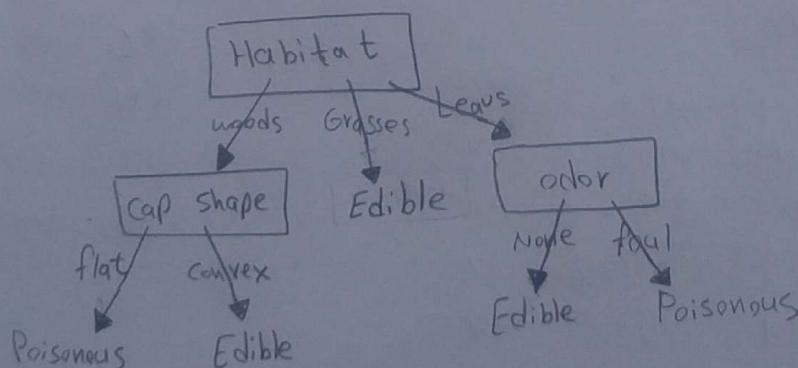
Edible:

		Poisonous	Edible	
		Red	—	$\Rightarrow E(\text{cap color} = \text{Red}) = E(1, 1) = 0.911V1$
		Green	1	$\Rightarrow E(\text{cap color} = \text{Green}) = E(1, 1) = 0.911V1$
(i)	Cap color:	(i) Red	—	$\Rightarrow E(\text{cap color} = \text{white}) = 1$

$$= \frac{Y_1(1)}{\omega} + \frac{Y_1(0, 91\pi\omega)}{\omega} = 0,9\omega$$

$$\frac{2}{3}(0) + \frac{1}{3}(0) = 0$$

$$\Rightarrow \text{Information Gain} = 0,9V1 - 0 = 0,9V1$$



9

Question 6)

Bayes
rule

Chain rule

independencibl
a, b, c ous

$$1) P(y|a, b, c) = \frac{P(a, b, c|y) P(y)}{P(a, b, c)} = \frac{P(a|y) P(b|y) P(c|y) P(y)}{P(a) P(b) P(c)}$$

2)

P(each class)

: $\sum_{i=1}^3 P(c_i) = 1$ و $P(c_i) = P(Y=ok) \cdot P(c_i|Y=ok)$ $\begin{cases} A = \text{false} \rightarrow 1/2 \\ B = \text{True} \rightarrow 1/2 \\ C = \text{false} \rightarrow 1/2 \end{cases}$

$$P(c_j) : \begin{cases} P(Y=ok) = \frac{\omega}{\lambda} = 0.44 \\ P(Y=bad) = \frac{\mu}{\lambda} = 0.56 \end{cases}$$

P(each feature | each class)

$$P(x_j|c_j) : \begin{cases} P(A=\text{false} | Y=ok) = \frac{\xi}{\omega} = 0.1 \\ P(A=\text{true} | Y=ok) = \frac{1}{\omega} = 0.1 \\ P(A=\text{false} | Y=bad) = \frac{1}{\mu} = 0.18 \\ P(A=\text{true} | Y=bad) = \frac{\gamma}{\mu} = 0.44 \\ P(B=\text{false} | Y=ok) = \frac{\omega}{\lambda} = 0.4 \\ P(B=\text{true} | Y=ok) = \frac{\lambda}{\omega} = 0.1 \\ P(B=\text{false} | Y=bad) = \frac{1}{\mu} = 0.18 \\ P(B=\text{true} | Y=bad) = \frac{\gamma}{\mu} = 0.44 \end{cases}$$

$$\begin{cases} P(c=\text{false} | Y=ok) = \frac{\omega}{\lambda} = 0.1 \\ P(c=\text{true} | Y=ok) = \frac{1}{\omega} = 0.1 \\ P(c=\text{false} | Y=bad) = \frac{\mu}{\lambda} = 0.44 \\ P(c=\text{true} | Y=bad) = \frac{1}{\mu} = 0.56 \end{cases}$$

\Rightarrow Posterior Probability :

Naïve-Bayes

$\Rightarrow A, B, C \rightarrow \text{independent}$

پس چون $P(Y=ok | X)$ می‌شود و می‌تواند سیم بگیر X خاصیتی داشته باشد، احتمال اینکه $Y=ok$ باشد برابر آن سیم است. \Rightarrow