Center for Artificial Intelligence Foundations & Scientific Applications Artificial Intelligence Research Laboratory 1 PennState PennState Institute for Computational and Data Sciences Clinical and Translational Science Institute Data Science for Researchers and Scholars Vasant G. Honavar Dorothy Foehr Huck and J. Lloyd Huck Chair in Biomedical Data Sciences and Artificial Intelligence Professor of Data Sciences, Informatics, Computer Science and Engineering, Bioinformatics & Genomics, Public Health Sciences and Neuroscience Director, Center for Artificial Intelligence Foundations and Scientific Applications Associate Director, Institute for Computational and Data Sciences Pennsylvania State University vhonavar@psu.edu http://faculty.ist.psu.edu/vhonavar http://ailab.ist.psu.edu



e for Researchers and Sch

Vasant Honavar, Fall 2023























































PennState Institute for Computational and Data Sciences	Applications	PennState Clinical and Translational Science Institute						
Bayesian Networks								
$S \in \{no, light, heat$	avy} Smok	king		→ Can	cer			
P(S=no)	0.80	$C \in$	malignant}					
P(S=light)	0.15							
P(S=heavy)) 0.05							
	Smoki	ng=	no	light	heavy			
	P(C=none)		0.96	0.88	0.60			
	P(C=benig	ŋn)	0.03	0.08	0.25			
	P(C=malig	<i>a)</i>	0.01	0.04	0.15			
PennState Galage of Lifermation Sciences And Technology	ata Science for Research	ers and Scholar	s	١	/asant Honavar, Fall 2023			

PennState Institute for Co and Data Scien	PennState Institute for Computational and Data Sciences							
Pro	Product Rule							
• $P(C,S) = P(C S) P(S)$ Smoking Cancer								
$S \Downarrow$	$C \Rightarrow$	none	benign	malignant				
no		0.768	0.024	0.008				
ligh	t	0.132	0.012	0.006				
hear	vy	0.035	0.010	0.005				
PennState Cellage of Information Sciences And Technology	Vasant Honavar, Fall 2023							

PennState Institute for Computational and Data Sciences Center for Artificial Intelligence Foundations & Scientific Applications Artificial Intelligence Research Laboratory PennState Clinical and Translational Science Institute								
Marginalization								
	$S \Downarrow C \Rightarrow$	none	benign	malig	total	`		
	no	0.768	0.024	0.008	.80			
	light	0.132	0.012	0.006	.15	P(Smoke)		
	heavy	0.035	0.010	0.005	.05			
	total	0.935	0.046	0.019				
P(Cancer)								
PennState College of Information Sciences And Technology	Data Science for Researchers and Scholars					nt Honavar, Fall 2023		

PennState Institute for Computational Artificial Intelligence Research Laboratory Clinical and Translational Science Institute								
Bayes Rule Revisited								
$P(S \mid C) = \frac{P(C \mid S)P(S)}{P(C)} = \frac{P(C,S)}{P(C)}$								
	$S \Downarrow C \Rightarrow$	none		benign	malig			
	no	0.768/.935		0.024/.046	0.008/.019			
	light	0.132/.935		0.012/.046	0.006/.019			
	heavy	0.030/.935		0.015/.046	0.005/.019			
r								
_	Cancer=		none	benign	malignant			
	P(S=no)		0.821	0.522	0.421			
	P(S=light)		0.141	0.261	0.316			
	P(S=heavy)		0.037	0.217	0.263			
PennState Data Science for Researchers and Scholars Vasant Honavar, Fall 2023								























































































Pen Institu and I	nState due for Computational Artificial Intellig The Learning Probl	icial Intelligence Foundations & gence Research Laboratory EM	Scientific Applications Pen Clinic Scien	nState al and Translational ce Institute			
		Known Structure	Unknown Structure				
	Complete Data	Statistical parameter estimation (closed-form eq.)	Discrete optimization over structures (discrete search)				
	Incomplete Data	Parametric optimization (EM, gradient descent)	Combined (Structural EM, mixture models)				
PennState Data Science for Researchers and Scholars Vasant Honavar, Fall 2023							

PennState Institute for Computational and Data Sciences Center for Artificial Intelligence Foundations & Scientific Applications Artificial Intelligence Research Laboratory								
Learning Problem								
	Known Structure	Unknown Structure						
Complete Data	Statistical parametric estimation (closed-form eq.)	Discrete optimization structures (discrete search)	over	-				
Incomplete Data	Parametric optimization (EM, gradient descent) E, B, A <y,n,n> <y,y,y></y,y,y></y,n,n>	Combined (Structural EM, mixtu models…)	ire	-				
E B P(A E,B) e b ? ? e 5 2 2	<n,n,y> <n,y,y></n,y,y></n,n,y>		E e	B P(A b 0.9	E,B) 0.1			
	<n,y,y></n,y,y>		e	Б 0.2	0.8			
ĒБ??			le le	b 0.9 5 0.01	0.1 0.99			
PennState Gebra et Information Sciences And Technology	Data Science for Researchers and Sch	olars	Vasant	: Honavar, Fa	III 2023			

PennState Institute for Computational and Data Sciences	PennSta Clinical and Science Inst	te Translational itute						
Learning Problem								
	Known Structure	Unknown Structure						
Complete Data	Statistical parametric estimation (closed-form eq.)	Discrete optimization structures (discrete search)	over					
Incomplete Data	Parametric optimization (EM, gradient descent)	Combined (Structural EM, mixtu models)	ire	-				
<u>F BP(A E,B)</u> e b ? ? e Б ? ?	<pre><pre><pre><pre><pre><pre><pre><pre></pre></pre></pre></pre></pre></pre></pre></pre>		E e	<u>B</u> P(A b 0.9	<u> E,B)</u> 0.1			
eb??	<n,y,y></n,y,y>		e e	Б 0.2	0.8 0.1			
			le	<u>Б</u> 0.01	0.99			
PennState Celever of Information Sciences and Inconvergery	Data Science for Researchers and Sch	olars	Vasant	: Honavar, F	all 2023			

PennState Institute for Computational and Data Sciences	PennSta Clinical and Science Insti	te Translational tute						
Learning Problem								
	Known Structure	Unknown Structure						
Complete Data	Statistical parametric estimation (closed-form eq.)	Discrete optimization structures (discrete search)	over					
Incomplete Data	Parametric optimization (EM, gradient descent) E, B, A <y,n,?></y,n,?>	Combined (Structural EM, mixtu models…)	re	-				
F B P(A E,B) e b ? ? e 5 ? ? ē b ? ?	<pre><',?,Y> <n,n,y> <?,Y,Y> <n,?,y></n,?,y></n,n,y></pre>		E e e	<u>В</u> Р(А b 0.9 Б 0.2	E,B) 0.1 0.8			
еБ??			e e	b 0.9 5 0.01	0.1 0.99			
PennState Critere of Information Scientific And Technicalogy	Data Science for Researchers and Sch	nolars	Vasant	: Honavar, Fa	all 2023			

PennState Institute for Computational and Data Sciences					ate d Translational titute			
Learning Problem								
	Known Structure	Unknown Structure						
Complete Data	Statistical parametric estimation (closed-form eq.)	Discrete optimization structures (discrete search)	over					
Incomplete Data	Parametric optimization (EM, gradient descent)	Combined (Structural EM, mixtu models)	re	-				
	E, B, A <y,n,?> <y,?,y></y,?,y></y,n,?>							
E BP(A E,B)	<n,n,y> <?,Y,Y></n,n,y>		E	B P(#	4 E,B)			
еБ??			e	b 0.9	0.1			
eb??	<n,?,y></n,?,y>		e	Б 0.2	2 0.8			
еБ??	B		ē	b 0.9	0.1			
			e	<u>Б</u> 0.0	1 0.99			
PennState Cleaner af Information Sciences And Technology	Data Science for Researchers and Sch	olars	Vasant	: Honavar, I	Fall 2023			





PennState Institute for Computational and Data Sciences					i e Translational tute			
Learning Problem								
	Known Structure	Unknown Structure						
Complete Data	Statistical parametric estimation (closed-form eq.)	Discrete optimization structures (discrete search)	over					
Incomplete Data	Parametric optimization (EM, gradient descent)	Combined (Structural EM, mixtu models)	re	-				
Е ВР(А Е.В) е b ? ? е Б ? ?	<pre></pre>		E	<u>В Р(А</u> b 0.9	<u> E,B)</u> 0.1			
ēb??	<n,y,y></n,y,y>		e	Б 0.2	0.8			
ēБ??			e	b 0.9	0.1			
			ē	<u>Б 0.01</u>	0.99			
PennState Gelange of Information Sciences And Rechnickopy	Data Science for Researchers and Sch	olars	Vasant	Honavar, Fa	II 2023			



















