Popular Python Libraries

Numpy/Scipy
(numerical and scientific computing)

Pandas
(different data format)

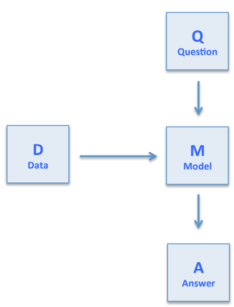
Matplotlib/Pandas/Seaborn/Plotly
(visualization: base, dataframe, stat, interactive)

SciKit-Learn
(basic machine learning algorithms)

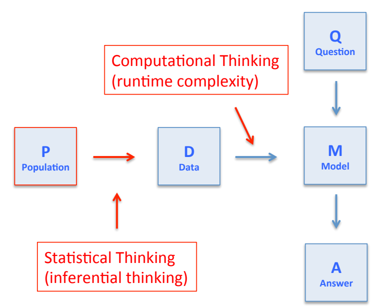
PySpark
(big data)

ML in Cartoon

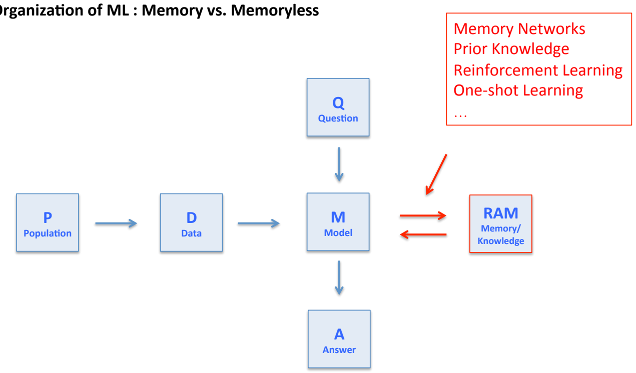
Organization of ML : Basic Backbone



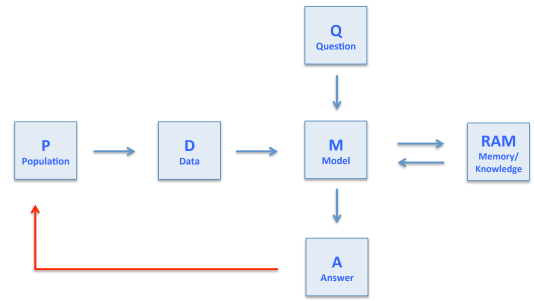
Organization of ML : Statistical vs. Computational Thinking



Organization of ML : Memory vs. Memoryless



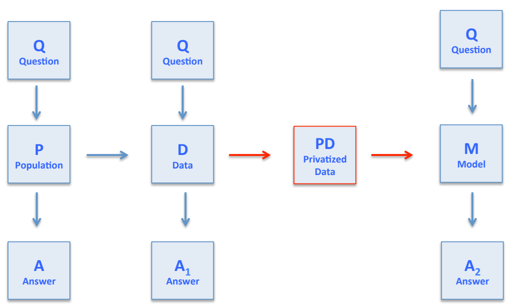
Organization of ML : Physical vs. Social



Physical: e.g., gravity invariant, no feedback

Social: e.g., good trading strategy becomes less useful due to active feedback

Organization of ML : Physical vs. Social



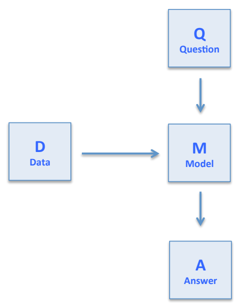
Privatized data or deliberately altered data

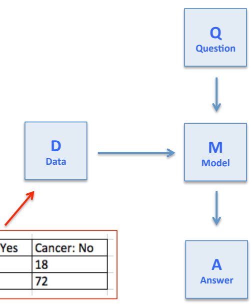
A_1 ~ A_2 ~ A

Example

Cancer Tests

Organization of ML : Basic Backbone



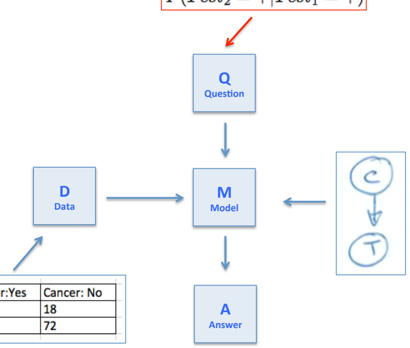


	Cancer:Yes	Cancer: No
Test:+	9	18
Test:-	1	72

$P(Cancer = Yes|Test = +)$

$P(Cancer = Yes|Test_1 = +, Test_2 = +)$

$P(Test_2 = +|Test_1 = +)$

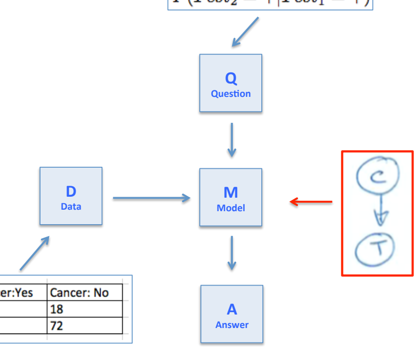


	Cancer:Yes	Cancer: No
Test:+	9	18
Test:-	1	72

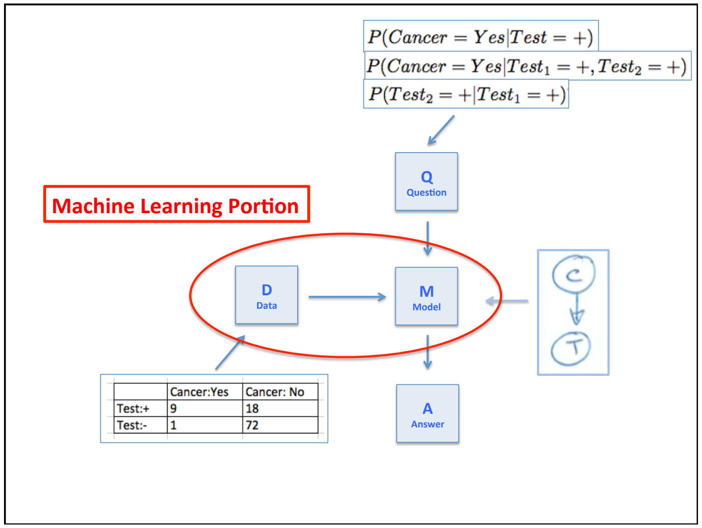
$P(Cancer = Yes|Test = +)$

$P(Cancer = Yes|Test_1 = +, Test_2 = +)$

$P(Test_2 = +|Test_1 = +)$



	Cancer:Yes	Cancer: No
Test:+	9	18
Test:-	1	72

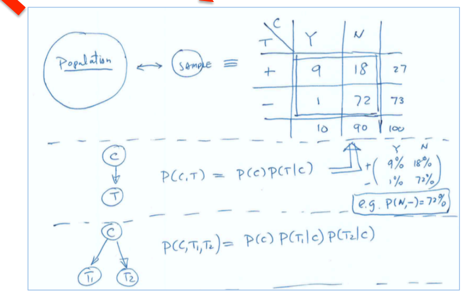


Machine Learning Portion

P(T=+, C=Yes) = 9/100 = 9%

P(T=+|C=Yes)= 9/(9+1)=90%

	Cancer:Yes	Cancer: No
Test:+	9	18
Test:-	1	72



P(C)

Y	10%
N	90%

P(C) ⇒ $P(Y) = \sum_T P(Y,T) = P(Y,+)+P(Y,-)$

$= 9\% + 1\% = 10\%$

$P(N) = \sum_T P(C=N,T) = P(N,+)+P(N,-)$

$= 18\% + 72\% = 90\%$

P(T|C)

T	C	
+	Y	90%
-	Y	10%
+	N	20%
-	N	80%

P(T|C) ⇒ $P(+|Y) = \frac{P(+,Y)}{P(+,Y)+P(-,Y)\ (\equiv P(Y))} = \frac{9\%}{10\%} = 90\%$

$P(-|Y) = 10\%$

$P(+|N) = \frac{P(+,N)}{P(N)} = \frac{18\%}{90\%} = 20\%$

$P(-|N) = 80\%$

$P(C,T_1,T_2)$

C	T_1	T_2		
Y	+	+	8.1%	10%
Y	+	-	0.9%	
Y	-	+	0.9%	
Y	-	-	0.1%	
N	+	+	3.6%	90%
N	+	-	14.4%	
N	-	+	14.4%	
N	-	-	57.6%	

$P(Y,+,+) = P(Y)\,P(+|Y)\,P(+|Y)$

$= 10\%\ 90\%\ 90\% = 8.1\%$

$P(Y,+,-) = P(Y)\,P(+|Y)\,P(-|Y)$

$= 10\%\ 90\%\ 10\% = 0.9\%$

$P(Y,-,+) = 0.9\%$

$P(Y,-,-) = 0.1\%$

$P(N,+,+) = P(N)\,P(+|N)\,P(+|N)$

$= 90\% \cdot 20\% \cdot 20\% = 3.6\%$

$P(N,+,-) = 90\%\ 20\%\ 80\% = 14.4\%$

$P(N,-,+) = 14.4\%$

$P(N,-,-) = P(N)\,P(-|N)\,P(-|N)$

$= 90\% \cdot 80\% \cdot 80\% = 57.6\%$

$$P(+) = \sum_{c} P(c,+) = \underbrace{P(Y,+)}_{9\%} + \underbrace{P(N,+)}_{18\%} = 27\%$$

$$= \sum_{C T_1} P(C,T,+) = P(Y,+,+) + P(Y,-,+) + P(N,+,+) + P(N,-,+)$$

$$= 8.1\% + 0.9\% + 3.6\% + 14.4\% = 27\%$$

$$P(+|+) = \frac{P(T_1=+, T_2=+)}{P(T_1=+, T_2=+) + P(T_1=-, T_2=+)}$$

But Then $P(+,+) = P(Y,+,+) + P(N,+,+) = \sum_c P(c,+,+)$

$$= 8.1\% + 3.6\% = 11.7\%$$

$$P(-,+) = P(Y,-,+) + P(N,-,+)$$

$$= 0.9\% + 14.4\% = 15.3\%$$

$$P(+|+) = \frac{11.7\%}{11.7\% + 15.3\%} = \frac{11.7}{27} = 43.3\%$$

One ML School

Bayesian Network

Independence/Factorization

$$p(x, y) = p(x)p(y)$$

Correlation
Causal Relation

Understanding!

Independence

Variables x and y are independent if knowing one event gives no extra information about the other event. Mathematically, this is expressed by

$$p(x, y) = p(x)p(y)$$

Independence of x and y is equivalent to

$$p(x|y) = p(x) \Leftrightarrow p(y|x) = p(y)$$

If $p(x|y) = p(x)$ for all states of x and y, then the variables x and y are said to be independent. We write then $x \perp\!\!\!\perp y$.

interpretation

Note that $x \perp\!\!\!\perp y$ doesn't mean that, given y, we have no information about x. It means the only information we have about x is contained in $p(x)$.

factorisation

If

$$p(x, y) = kf(x)g(y)$$

for some constant k, and positive functions $f(\cdot)$ and $g(\cdot)$ then x and y are independent.

Conditional Independence

$$\mathcal{X} \perp\!\!\!\perp \mathcal{Y} | \mathcal{Z}$$

denotes that the two sets of variables $\mathcal{X}$ and $\mathcal{Y}$ are independent of each other given the state of the set of variables $\mathcal{Z}$. This means that

$$p(\mathcal{X}, \mathcal{Y}|\mathcal{Z}) = p(\mathcal{X}|\mathcal{Z})p(\mathcal{Y}|\mathcal{Z}) \text{ and } p(\mathcal{X}|\mathcal{Y}, \mathcal{Z}) = p(\mathcal{X}|\mathcal{Z})$$

for all states of $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$. In case the conditioning set is empty we may also write $\mathcal{X} \perp\!\!\!\perp \mathcal{Y}$ for $\mathcal{X} \perp\!\!\!\perp \mathcal{Y} | \emptyset$, in which case $\mathcal{X}$ is (unconditionally) independent of $\mathcal{Y}$.

Conditional independence does not imply marginal independence

$$p(x, y) = \sum_z \underbrace{p(x|z)p(y|z)}_{\text{cond. indep.}} p(z) \neq \underbrace{\sum_z p(x|z)p(z)}_{p(x)} \underbrace{\sum_z p(y|z)p(z)}_{p(y)}$$

Conditional dependence

If $\mathcal{X}$ and $\mathcal{Y}$ are not conditionally independent, they are conditionally dependent. This is written

$$\mathcal{X} \top\!\!\!\top \mathcal{Y} | \mathcal{Z}$$

$$p(x, y) = p(x)p(y)$$

Correlation
Causal Relation

Understanding!

Factorization (locality)

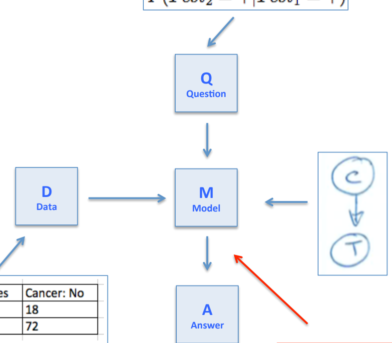
Efficient Computation!

$$\sum_{x,y,z=\{1,2\}} xyz = (\sum_{x=\{1,2\}} x)(\sum_{y=\{1,2\}} y)(\sum_{z=\{1,2\}} z)$$

$P(Cancer = Yes|Test = +)$

$P(Cancer = Yes|Test_1 = +, Test_2 = +)$

$P(Test_2 = +|Test_1 = +)$



	Cancer:Yes	Cancer: No
Test:+	9	18
Test:-	1	72

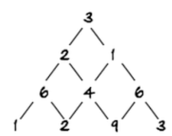
Factorization
Belief Propagation Algorithm

Schools of ML

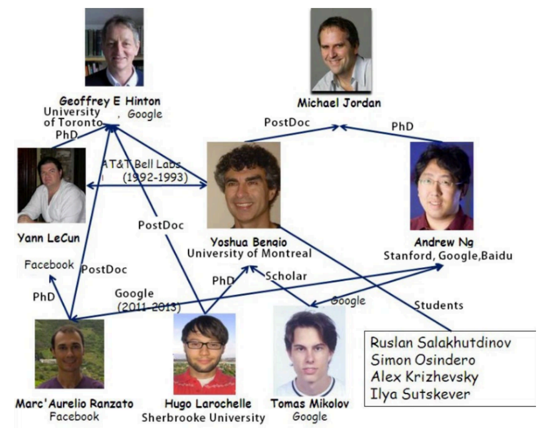
People and Kinds

Machine Learning

Bayesian Network	Neural Network
Prior Knowledge	No Prior Knowledge
Reasoning	Black Box
BP = Belief Propagation	BP = Back Propagation
(dynamic Programming)	(dynamic Programming)



CS: Dynamical Programming
PHYS: Transfer Matrix
Math: Recursive Relations
STAT: Message Passing

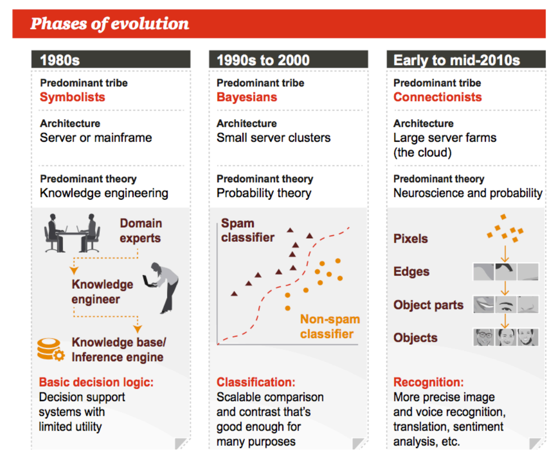


Source: internet

Top 50 authors in computer science

Author	Influencial	Citation	S2 Citations	Citations per	Influential	Citation
Michael I. Jordan	1185	5745	24452	203.89	3.82	18.53
Andrew Y. Ng	1125	7247	20634	282.63	6.25	40.26
Andrew Zisserman	1089	6457	25621	198.95	3.11	18.45
Christopher D. Manning	1087	5318	17499	177.68	4.99	24.39
Jitendra Malik	942	5043	22799	261.14	3.94	21.1
Geoffrey E. Hinton	834	6094	15338	198.65	4.37	31.91
Scott Shenker	780	4092	25975	226.07	2.28	11.96
Bernhard Schölkopf	768	3213	15114	140.62	2.38	9.95
Jon M. Kleinberg	757	4135	16219	258.43	4.53	24.76
Judea Pearl	747	1594	13681	141.27	2.8	5.97
David N. C. Tse	732	2939	12542	439.53	7.55	30.3
Cordelia Schmid	716	5114	16031	235.05	3.83	27.35
Jiawei Han	697	4310	22178	93.4	0.9	5.55
Rakesh Agrawal	696	2456	19327	289.27	3.15	11.11
Hari Balakrishnan	695	3307	20809	286.02	2.98	14.19
Richard S. Sutton	684	1899	9657	334.23	8.14	22.61
Emmanuel J. Candès	644	4713	10992	462.06	9.91	72.51
Daphne Koller	627	2188	13801	142.16	2.55	8.89
David M. Blei	622	3952	11075	234.18	5.6	35.6
Thorsten Joachims	620	2771	12829	232.99	4.28	19.11
Jeffrey Dean	617	4249	10190	478.02	11.02	75.88
Robert E. Schapire	608	2111	13885	317.73	4.83	16.75
Yoshua Bengio	607	5999	13147	98.89	1.94	19.17
Yoram Singer	597	2199	10888	235.02	5.1	18.79
Ilya Sutskever	588	3297	5716	231.35	12	67.29
Noga Alon	584	978	10909	50.93	1.01	1.69
David E. Culler	577	1969	18586	272.34	2.77	9.47
Andrew G. Barto	570	1698	8223	261.13	6.13	18.26
Robert Morris	564	1659	15257	266.2	3.2	9.43
Anil K. Jain	555	3012	19954	136.19	1.01	5.51
Dan Klein	552	1348	7192	72.43	2.89	7.06
Michael Collins	547	933	6692	108.03	4.38	7.46
Philipp Koehn	545	1725	8243	105.68	3.66	11.58
Alexander J. Smola	540	2840	11233	160.18	2.73	14.34
Thomas M. Cover	517	2988	12531	705.53	9.75	56.38
Andrew McCallum	517	2367	13190	199.85	3.29	15.08
Vladimir Vapnik	514	3228	13220	673.27	7.67	48.18
John D. Lafferty	514	2044	10821	280.75	5.65	22.46
Trevor Hastie	506	2132	8760	216.82	4.56	19.21
Christos Faloutsos	501	2930	17609	90.91	0.91	5.3
Jason Weston	499	2206	7176	169.5	4.46	19.7
David G. Lowe	493	4386	15331	686.33	7.14	63.57
Martin J. Wainwright	492	1395	5972	98.06	3.46	9.82
William T. Freeman	491	2280	11375	215.47	3.48	16.17
Christos H. Papadimitriou	473	1643	16184	143.35	1.47	5.1
Hans-Peter Seidel	473	579	9361	39.15	0.69	0.84
Deborah Estrin	472	2020	19669	231.18	1.66	7.11
Yi Ma	464	3148	6872	73.86	1.78	12.06
Francis R. Bach	458	2485	7011	112.6	3.05	16.57
Ion Stoica	452	3574	17281	193.75	1.74	13.75

Who's the Michael Jordan of computer science? New tool ranks researchers' influence



http://usblogs.pwc.com/emerging-technology/a-look-at-machine-learning-infographic/