Minimum Redundancy and Maximum Relevance Feature Selection and Its Applications

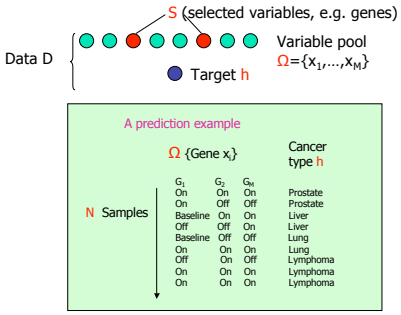
Hanchuan Peng

Janelia Farm Research Campus,
Howard Hughes Medical Institute

Outline

- What is mRMR feature selection
- Applications in cancer classification
- Applications in image pattern recognition
- Theoretical basis of mRMR
- Combinations with other methods
- How to use mRMR programs

Feature Selection Problem



Redundancy in Features

- **Most used methods**: select top-ranking features based on mutual info., *F*-test, etc., without considering relationships among features.
- **Problems**:
  - Selected features are correlated;
  - Selected features cover narrow regions in space.

mRMR: Discrete Variables

- **Minimize Redundancy**:

$$\min W_I, \qquad W_I = \frac{1}{|S|^2} \sum_{i,j \in S} I(i,j)$$

S is the set of features.

$I(i,j)$ is **mutual information** between features i and j

- **Maximize Relevance**:

$$\max V_I, \qquad V_I = \frac{1}{|S|} \sum_{i \in S} I(h,i)$$

h = target classes (e.g. types of different cancers, or annotations)

mRMR: Continuous Variables

- **Maximize relevance** : F-statistic: $F(i,h)$

$$\max V_F, \qquad V_F = \frac{1}{|S|} \sum_{i \in S} F(i,h).$$

- **Minimize redundancy** : Correlation $c(i,j)$

$$\min W_c, \qquad W_c = \frac{1}{|S|^2} \sum_{i,j} |\, c(i,j) \,|$$

- **Relevance and redundancy can also be defined using mutual info of hybrid variables (continuous or categorical) (Peng, Long, & Ding, TPAMI, 2005)**

Combine Redundancy and Relevance

- **Additive combination**

$$\max(V - W)$$

- **Multiplicative combination**

$$\max(V / W)$$

MRMR Selection Schemes

Type	Acronym	Full Name	Formula
Discrete	MID	Mutual information difference	$\max_{i\in\Omega_S}[I(i,h)-\frac{1}{\lvert S\rvert}\sum_{j\in S} I(i,j)]$
	MIQ	Mutual information quotient	$\max_{i\in\Omega_S}\{I(i,h)/[\frac{1}{\lvert S\rvert}\sum_{j\in S} I(i,j)]\}$
Continuous	FCD	F-test correlation difference	$\max_{i\in\Omega_S}[F(i,h)-\frac{1}{\lvert S\rvert}\sum_{j\in S} \lvert c(i,j)\rvert]$
	FCQ	F-test correlation quotient	$\max_{i\in\Omega_S}\{F(i,h)/[\frac{1}{\lvert S\rvert}\sum_{j\in S} \lvert c(i,j)\rvert]\}$
	FDM	F-test distance multiplicative	$\max_{i\in\Omega_S}[F(i,h)\cdot\frac{1}{\lvert S\rvert}\sum_{j\in S} d(i,j)]$
	FSQ	F-test similarity quotient	$\max_{i\in\Omega_S}\{F(i,h)/[\frac{1}{\lvert S\rvert}\sum_{j\in S}\frac{1}{d(i,j)}]\}$

Most Related Methods

- Most used feature selection methods: top-ranking features without considering relationships among features.
- Yu & Liu, 2003/2004. information gain, essentially similar approach.
- Wrapper: not filter approach, classifier-involved and thus features do not generalize well.
- PCA and ICA: Features are orthogonal or independent, but not in the original feature space.

Outline

- What is mRMR feature selection
- Applications in cancer classification
- Applications in image pattern recognition
- Theoretical basis of mRMR
- Combinations with other methods
- How to use mRMR programs

Microarray Gene Expression Data Sets for Cancer Classification

Table 1. Two-class datasets used in our experiments

Dataset	Leukemia		Colon cancer	
Source	Golub et al (1999)		Alon et al (1999)	
# Gene	7070		2000	
# Sample	72		62	
Class	Class name	# Sample	Class name	# Sample
C1	ALL	47	Tumor	40
C2	AML	25	Normal	22

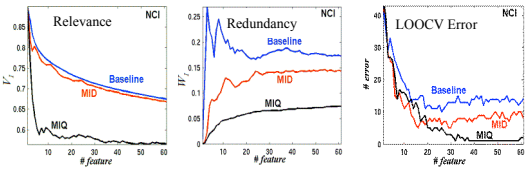
Table 2. Multi-class datasets used in our experiments (#S is the number of samples)

Dataset	NCI		Lung cancer		Lymphoma		Child Leukemia	
Source	Ross et al (2000) Scherf et al (2000)		Garber et al (2001)		Alizadeh et al (2000)		Yoeh et al (2002) Li et al (2003)	
# Gene	9703		918		4026		4026	
# S	60		73		96		96	
# Class	9		7		9		9	
Class	Class name	# S	Class name	# S	Class name	# S	Class name	# S
C1	NSCLC	9	AC-group-1	21	Diffuse large B cell lymphoma	46	BCR-ABL	9/6
C2	Renal	9	Squamous	16	Chronic Lympho. leukemia	11	E2A-PBX1	18/9
C3	Breast	8	AC-group-3	13	Activated blood B	10	Hyperdiploid>50	42/22
C4	Melanoma	8	AC-group-2	7	Follicular lymphoma	9	MLL	14/6
C5	Colon	7	Normal	6	Resting/ activated T	6	T-ALL	28/15
C6	Leukemia	6	Small-cell	5	Transformed cell lines	6	TEL-AML1	52/27
C7	Ovarian	6	Large-cell	5	Resting blood B	4	Others	52/27
C8	CNS	5			Germinal center B	2		
C9	Prostate	2			Lymph node/tonsil	2		

Lymphoma Results

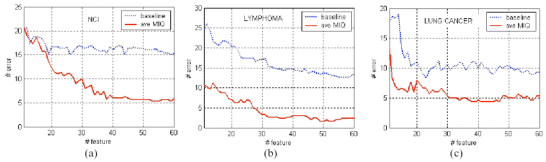
Classifier	Data Type	M / Method	3	6	9	12	15	18	21	24	27	30	36	42	48	54	60
NB	Discrete	Baseline	38	39	25	29	23	22	22	19	20	17	19	18	18	17	17
		MID	31	15	10	9	9	8	6	7	7	7	4	7	5	5	8
		MIQ	38	26	17	14	14	12	8	8	6	7	5	6	4	3	3
LDA	Discrete	Baseline	40	42	28	26	20	21	21	20	18	19	14	15	13	14	15
		MID	32	15	14	10	7	5	4	5	4	6	5	3	3	4	3
		MIQ	40	29	12	8	8	7	5	6	4	1	1	2	1	2	2
	Continuous	Baseline	66	26	26	17	17	18	18	18	15	11	14	12	11	11	13
		FCD	33	17	16	10	13	11	11	9	8	8	8	8	7	10	9
		FCQ	32	18	11	7	7	8	8	7	8	9	9	9	8	6	6
		FDM	40	21	17	17	16	13	14	12	12	11	12	9	10	10	11
		FSQ	40	14	14	13	11	12	10	11	10	9	10	8	8	9	10
SVM	Discrete	Baseline	32	29	25	23	20	22	18	13	14	15	11	10	10	8	9
		MID	24	10	7	4	2	3	3	3	3	3	3	3	3	3	3
		MIQ	26	21	13	9	8	7	6	5	5	2	1	1	2	1	2
	Continuous	Baseline	30	24	14	13	12	13	10	11	13	6	8	9	5	6	7
		FCD	24	19	11	13	11	9	10	8	7	8	7	6	5	4	5
		FCQ	31	17	9	7	6	6	8	8	6	7	7	8	7	4	4
		FDM	25	11	9	11	11	8	10	9	7	8	8	8	8	5	6
		FSQ	25	11	10	11	8	9	6	6	7	9	8	6	6	6	6

What is the Role of Redundancy Reduction?



(a) Relevance and (b) redundancy for MRMR features on discretized NCI dataset.
(c) The respective LOOCV errors obtained using the Naïve Bayes classifier

Do mRMR Features Better Cover the Data Distribution Space and Thus Perform Well on Different Classifiers?



Average LOOCV errors of three different classifiers, NB, SVM, and LDA on three multi-class datasets

Do mRMR Features Generalize Well on Unseen Data?

Classifier	Method \ M	3	6	9	12	15	18	24	30	40	50	60	70	80	90	100
LDA	Baseline	55	47	46	38	34	27	19	28	22	19	15	14	11	8	8
	MID	50	43	32	29	30	29	22	15	13	10	10	9	7	8	9
	MIQ	43	43	34	27	23	21	18	16	11	11	6	4	6	6	4
SVM	Baseline	56	55	49	37	33	33	27	35	29	30	23	20	18	14	13
	MID	45	42	33	33	25	25	29	25	26	22	20	13	10	12	9
	MIQ	38	30	34	33	27	26	24	21	14	15	17	10	7	11	9

Child Leukemia data (7 classes, 215 training samples, 112 testing samples) testing errors. M is the number of features used in classification

What is the Relationship of mRMR Features and Various Data Discretization Schemes?

Data Sets	Method \ M	3	6	9	12	15	18	21	24	27	30	36	42	48	54	60
NCI	Baseline	34	25	23	25	19	17	18	15	14	12	12	12	13	12	10
	MRMR (MIQ)	35	22	22	16	12	11	10	8	5	3	4	4	2	2	3
Lymphoma	Baseline	58	52	44	39	44	17	17	14	16	13	11	10	13	10	12
	MRMR (MIQ)	24	17	7	8	4	2	1	2	4	3	2	2	2	2	2

LOOCV testing results (#error) for binarized NCI and Lymphoma data using SVM classifier

Comparison with Other Work

Data	Method	NB	LDA	SVM	LR	Literature
NCI	Baseline	18.33	26.67	25.00	--	14.63 [a]
	MRMR	1.67	13.33	11.67	--	5-class: 0 [b], 0 [b]
Lymphoma	Baseline	17.71	11.46	5.21	--	3-class: 2.4 [c], 0 [c]
	MRMR	3.13	1.04	1.04	--	
Lung	Baseline	10.96	10.96	10.96	--	--
	MRMR	2.74	5.48	5.48	--	
Child Leukemia	Baseline	29.46	7.14	11.61	--	5.36 [d]
	MRMR	13.39	2.68	6.25	--	
Leukemia	Baseline	0	1.39	1.39	1.39	0 [e]
	MRMR	0	0	0	0	1.39 [f]
Colon	Baseline	11.29	11.29	11.29	11.29	9.68 [e]
	MRMR	6.45	8.06	9.68	9.68	6.45 [g]

Comparison of the best results (lowest error rates in percentage) of the baseline and MRMR features. Also listed are results in literature (the best results in each paper). [a] Ooi & Tan, using the genetic algorithm [26]. [b] Nguyen and Rocke [25] used a 5-class subset of NCI dataset and obtained 0% error rate; using the same 5-class subset, our NB achieves also 0% error rate. [c] Nguyen & Rocke used 3-class subset in lymphoma dataset and obtain 2.4% error rate. Using the same 3 classes, our NB led to zero errors. [d] Li et al, using prediction by collective likelihood [21]. [e] Furey et al, using SVM [11]. [f] Lee & Lee, using SVM [20]. [g] Nguyen & Rocke, using PLS [24].

Outline

- What is mRMR feature selection
- Applications in cancer classification
- Applications in image pattern recognition
- Theoretical basis of mRMR
- Combinations with other methods
- How to use mRMR programs

Automatic Annotation of in situ Gene Expression Patterns in Fly Embryos

- Minimal redundant wavelet-embryo features were used to train classifiers.
- 456 genes, 80 ontological annotation terms, 6 developmental ranges.
- Consistent automatic annotations with those manually generated in BDGP.
- 99.5% accuracy in predicting developmental stages of embryos by gene expression patterns.

Gene Name	Expression Pattern Image	Annotation Terms					Our Automatic Annotations	Manual Annotations in BDGP
		ECNS	VNC	EH	EM	EDE		
CG4133		1	0.99	-	-	-	ECNS, VNC	ECNS, VNC
CG6930		0.96	0.98	-	-	0.99	ECNS, VNC, EDE	ECNS, VNC, EDE
Caki(CG6703)		0.99	1	-	-	-	ECNS, VNC	ECNS, VNC
CG33071		-	-	0.96	0.98	-	EH, EM	EH, EM
CG9598		-	-	0.99	0.92	0.98	EH, EM, EDE	EH, EM, EDE
CG13875		-	-	0.96	-	1	EH, EDE	EH, EDE
CG17786		-	-	0.99	-	0.97	EH, EDE	EH, EDE
CG4532		0.99	0.94	-	-	0.91	ECNS, VNC, EDE	ECNS, VNC
NetB (CG10521)		1	1	-	-	0.87	ECNS, VNC, EDE	ECNS, VNC

Prediction confidence score

Short terms: ECNS: embryonic central nervous system; VNC: ventral nerve cord; EH: embryonic hindgut; EM: embryonic midgut; EDE: embryonic dorsal epidermis.

5-hours 5-years

Zhou & Peng, Bioinformatics, 2007

19

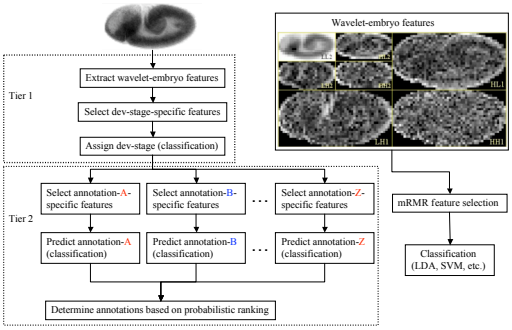
Wavelet-embryo Features

Gene	Image pattern	Image features and applications: * GMM-Blob: pattern extraction, comparison and retrieval * Eigen-profile: pattern clustering * Wavelet-profile: pattern annotation/classification		
Bub3	Segmented embryo	GMM-blob-profile	Wavelet-profile	
	Extracted pattern	Eigen-profile		
CG10489	Segmented embryo	GMM-blob-profile	Wavelet-profile	
	Extracted pattern	Eigen-profile		

Peng et al, BMC Cell Bio, 2007

20

Annotation System



21

Tier 1 Accuracy: Developmental Stage Prediction

Actual stage range/Predicted stage range	1–3	4–6	7–8	9–10	11–12	13–16
1–3	447	3	0	0	0	0
4–6	0	446	0	0	0	0
7–8	1	0	379	0	0	0
9–10	0	0	0	372	1	0
11–12	1	1	0	0	435	0
13–16	2	0	0	0	2	362

Overall accuracy >99%

Peng et al, BMC Cell Bio, 2007

22

Tier 2 Accuracy: Specific Annotation Prediction

Table 1. Recognition rates (%) of feature extraction/selection methods on synthetic data sets with mutually exclusive (M.E.) annotations

Stage range	Set size	Set details	Score	Eigen embryo	LeNet	LeNet + mRMR	Wavelet embryo	Wavelet embryo + mRMR
4–6	98	78CB + 20Subset	R_0	70	75	70	70	100
			R_A	82	85	83	82	98
7–8	69	49TMA + 20PEA	R_0	60	50	55	60	100
			R_A	67	59	80	67	100
9–10	69	49TMP + 20PEP	R_0	65	50	60	65	100
			R_A	71	52	83	71	100
11–12	61	41HPP + 20PMP	R_0	45	50	90	55	95
			R_A	59	69	90	64	98
13–16	32	20EBNS + 12VNC	R_0	83	58	67	83	100
			R_A	84	59	81	84	100

Top 20 eigen features are used. Top 20 features are selected when using mRMR feature selection. R_0: recognition rate on the smaller class only. R_A: overall recognition rate on the whole data set for all classes. (Annotation terms: CB, cellular blastoderm; TMA, trunk mesoderm anlage; PEA, posterior endoderm anlage; TMP, trunk mesoderm primordium; PEP, proxephalic ectoderm primordium; HPP, hindgut proper primordium; PMP, posterior midgut primordium; EBNS, embryonic central nervous system; VNC, ventral nerve cord.)

Table 2. Recognition rates (%) of feature extrction/selection methods on synthetic data sets with multi-objective (M.O.) annotations

Stage range	Set size	Set details	Score	Eigen embryo	LeNet	LeNet + mRMR	Wavelet embryo	Wavelet embryo + mRMR
4–6	152	78CB + 20Subset + 54Both	R_0	49	42	55	49	76
			R_A	64	63	71	64	82
7–8	89	49TMA + 20PEA + 20Both	R_0	45	50	58	45	83
			R_A	40	46	62	40	85
9–10	87	49TMP + 20PEP + 18Both	R_0	42	34	42	39	79
			R_A	49	33	55	47	80
11–12	75	41HPP + 20PMP + 14Both	R_0	38	32	59	38	88
			R_A	33	36	55	37	89
13–16	97	20EBNS + 12VNC + 65Both	R_0	57	53	70	57	87
			R_A	55	48	70	55	86

Short terms and testing parameters are the same as Table 1.

Zhou & Peng, Bioinformatics, 2007

23

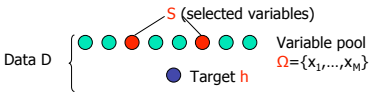
Outline

- What is mRMR feature selection
- Applications in cancer classification
- Applications in image pattern recognition
- Theoretical basis of mRMR
- Combinations with other methods
- How to use mRMR programs

24

Maximum Dependency Criterion

- **Statistical association**
- **Definition**
  - Mutual information I(S,h)



Mutual Information

- For two univariate variables x and y

$$I(x;y) = \iint p(x,y)\log\frac{p(x,y)}{p(x)p(y)}dxdy$$

- For multivariate variable S_m and the target h

$$I(S_m;h) = \iint p(S_m,h)\log\frac{p(S_m,h)}{p(S_m)p(h)}dS_m dh$$

not

$$J(x_1,x_2,...,x_m,h) = \int\cdots\int p(x_1,...,x_m,h)\log\frac{p(x_1,x_2,...,x_m,h)}{p(x_1)\cdots p(x_m)p(h)}dx_1\cdots dx_m dh$$
$$= \int\cdots\int p(S_m,h)\log\frac{p(S_m,h)}{p(x_1)\cdots p(x_m)p(h)}dS_m dh$$

High-Dimensional Mutual Information

- For multivariate variable S_m and the target h

$$I(S_m;h) = \iint p(S_m,h)\log\frac{p(S_m,h)}{p(S_m)p(h)}dS_m dh$$
$$= \iint p(S_{m-1},x_m,h)\log\frac{p(S_{m-1},x_m,h)}{p(S_{m-1},x_m)p(h)}dS_m dh$$
$$= \int\cdots\int p(x_1,\cdots,x_m,h)\log\frac{p(x_1,\cdots,x_m,h)}{p(x_1,\cdots,x_m)p(h)}dx_1\cdots dx_m dh.$$

- Estimating high-dimensional $I(S_m,h)$ is difficult
  - An ill-posed problem to find inverse of large co-variance matrix
  - Insufficient number of samples

Maximum Dependency Feature Selection is Combinatorial !

- Number of subspaces
  - Total $2^{|\Omega|}$ or 2^M
  - For given number of features |S| $\binom{|\Omega|}{|S|}$

Example:

|Ω|	|S|	#configurations of selected variables
1000	1	10^3
1000	2	0.5×10^6
1000	3	$\sim 1.66 \times 10^8$
5000	3	$\sim 2.08 \times 10^{10}$

Heuristic search algorithms are necessary.
Simplest case: the incremental search.

Factorize the Mutual Information

Mutual information for multivariate variable S_m and the target h

$$I(S_m;h) = \iint p(S_m,h)\log\frac{p(S_m,h)}{p(S_m)p(h)}dS_m dh$$

Define:

$$J(x_1,x_2,...,x_m) = \int\cdots\int p(x_1,...,x_m)\log\frac{p(x_1,x_2,...,x_m)}{p(x_1)\cdots p(x_m)}dx_1\cdots dx_m$$

It can be proved:

$$I(S_m,h) = J(h,S_m) - J(S_m)$$

Upper & Lower Bounds of J(.)

Lower bound

$$J(x_1,x_2,...,x_m) \geq 0.$$

when variables are maximally independent

Upper bound

$$J(x_1,x_2,...,x_n) \leq \min\{\sum_{i=2}^{n}H(x_i), \sum_{i=1,i\neq 2}^{n}H(x_i),\cdots, \sum_{i=1,i\neq n-1}^{n}H(x_i), \sum_{i=1}^{n-1}H(x_i)\}$$

when variables are maximally dependent

Factorize $I(S_m,h)$

- Relevance of S={x_1, x_2, …} and h, or $R_L(S,h)$
- Redundancy among variables {x_1, x_2, …}, or $R_D(S)$

$$R_L = \frac{1}{|S|}\sum_{x_i \in S} I(x_i,h) \qquad R_D = \frac{1}{|S|^2}\sum_{x_i,x_j \in S} I(x_i,x_j)$$

$$I(S_m,h) = J(S_{m-1},x_m,h) - J(S_{m-1},x_m).$$

- For incremental search, max I(S,h) is "equivalent" to max $[R_L(S,h) - R_D(S)]$, i.e. combination of min-Redundancy-Max-Relevance (mRMR).

Advantages of mRMR

- Both relevance and redundancy estimation are low-dimensional problems (i.e. involving only 2 variables). This is much easier than directly estimating multivariate density or mutual information in the high-dimensional space!
- Faster speed
- More reliable estimation

- mRMR is an optimal first-order approximation of I(.) maximization
- Relevance-only ranking only maximizes J(.)!

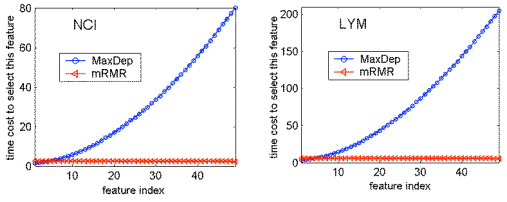
Search Algorithm of mRMR

- **Greedy search algorithm**
  - In the pool Ω find the variable x^1 that has the largest I(.,h). Exclude x^1 from Ω.
  - Search x^2 so that it maximizes $I(.,h) - \Sigma I(.,x^1)/|\Omega|$.
  - Iterate this process until an expected number of variables have been obtained, or other constraints are satisfied.
- **Complexity** O(|S|*|Ω|)

Outline

- What is mRMR feature selection
- Applications in cancer classification
- Applications in image pattern recognition
- Theoretical basis of mRMR
- Comparison/Combinations with other methods
- How to use mRMR programs

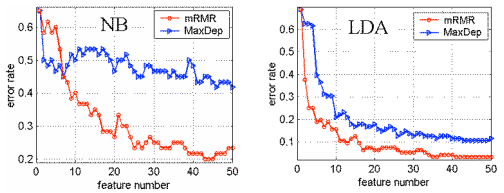
Comparing Max-Dep and mRMR: Complexity of Feature Selection



Time cost (seconds) for selecting individual features based on mutual information estimation for continuous feature variables.

(Parallel experiments on a cluster of eight 3.06G Xeon CPUs running Redhat Linux 9, with the Matlab implementation)

Comparing Max-Dep and mRMR: Accuracy of Feature Selected in Classification

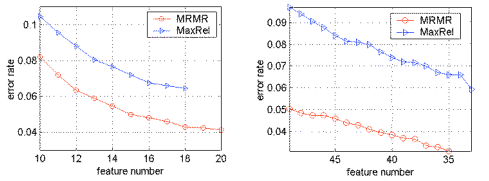


Leave-One-Out cross validation of feature classification accuracies of mRMR and MaxDep

Use Wrappers to Refine Features

- mRMR is a filter approach
  - Fast
  - Features might be redundant
  - Independent of the classifier
- Wrappers seek to minimize the number of errors directly
  - Slow
  - Features are less robust
  - Dependent on classifier
  - Better prediction accuracy
- Use mRMR first to generate a short feature pool and use wrappers to get a least redundant feature set with better accuracy

Use Wrappers to Refine Features



Forward wrappers (incremental selection)

Backward wrappers (Decremental selection)

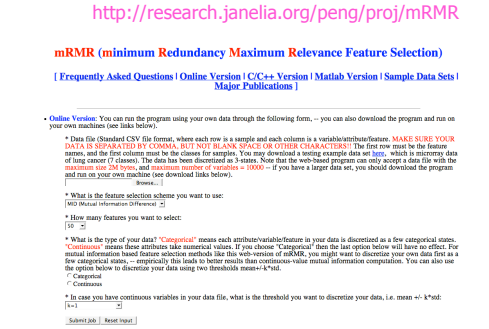
NCI data

Outline

- What is mRMR feature selection
- Applications in cancer classification
- Applications in image pattern recognition
- Theoretical basis of mRMR
- Combinations with other methods
- How to use mRMR programs

mRMR Website

http://research.janelia.org/peng/proj/mRMR


mRMR (minimum Redundancy Maximum Relevance Feature Selection)

[Frequently Asked Questions | Online Version | C/C++ Version | Matlab Version | Sample Data Sets | Major Publications]

Available Versions

- Matlab version
  - All source codes. Can be embedded in your programs easily.
- C/C++ version
  - For Linux/Unix/Mac, simple command line executable.
- Online version
  - Upload the data sets (csv format: comma separated values) and get the results right away.

Available Datasets

- NCI data (9 cancers, discretized as 3-states)
- Lung Cancer data (7 subtypes, discretized as 3-states)
- Lymphoma data (9 cancer-subtypes, discretized as 3-states)
- Leukemia data (2 subtypes, discretized as 3-states)
- Colon Cancer data (2 subtypes, discretized as 3-states)
- The continuous-value raw data should be obtained from the original sources.

Citations

- [TPAMI05] Hanchuan Peng, Fuhui Long, and Chris Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 8, pp.1226-1238, 2005.
  - This paper presents a theory of mutual information based feature selection. Demonstrates the relationship of four selection schemes: maximum dependency, mRMR, maximum relevance, and minimal redundancy. Also gives the combination scheme of "mRMR + wrapper" selection and mutual information estimation of continuous/hybrid variables.
- [JBCB05] Chris Ding, and Hanchuan Peng, "Minimum redundancy feature selection from microarray gene expression data," Journal of Bioinformatics and Computational Biology, Vol. 3, No. 2, pp.185-205, 2005.
  - This paper presents a comprehensive suite of experimental results of mRMR for microarray gene selection on many different conditions. It is an extended version of the CSB03 paper.
- [CSB03] Chris Ding, and Hanchuan Peng, "Minimum redundancy feature selection from microarray gene expression data," Proc. 2nd IEEE Computational Systems Bioinformatics Conference (CSB 2003), pp.523-528, Stanford, CA, Aug, 2003.
  - This paper presents the first set of mRMR results and different definitions of relevance/redundancy terms.
- [Bioinfo07] Jie Zhou, and Hanchuan Peng, "Automatic recognition and annotation of gene expression patterns of fly embryos," Bioinformatics, Vol. 23, No. 5, pp. 589-596, 2007.
  - One application of mRMR in selecting good wavelet image features.

Conclusions

- The Max-Dependency feature selection can be efficiently implemented as the mRMR algorithm.
- Significantly outperforms the widely used max-relevance selection method: mRMR features cover a broader feature space with less features.
- mRMR is very efficient and useful for gene selection and many other applications. The programs are ready!

mRMR website:
http://research.janelia.org/peng/proj/mRMR

Acknowledgment

Collaborators:


Chris Ding


Fuhui Long


Jie Zhou