Multi-label Learning (MLL)

Classifier Chain

speaker Suhe Ye

College of Mathematics And Statistics, ShenZhen University

October 8, 2018
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
Multi-Label Object

- Sunset
- Clouds
- Trees
- Countryside
- ....
Multi-label Learning (MLL)
MLL- Applications

- Text Categorization
- Automatic annotation for multimedia contents
  - Image, Audio, Video
- Bioinformatics
- World Wide Web
- Information Retrieval
- Directed marketing
Formal Definition of MLL

Settings

\( \mathcal{X} : \) d -dimensional feature space \( \mathbb{R}^d \)

\( \mathcal{Y} : \) label space with L labels \([1, 2, 3, \ldots, L]\)

Inputs

\( \mathcal{D} : \) training set with N examples \((x_i, Y_i)|1 < i < N\)

\( x_i \in \mathcal{X} \) is a d -dimensional feature vector \((x_{i1}, x_{i2}, x_{i3}, \ldots, x_{id})^T\)

\( Y_i \in \mathcal{Y} \) is the label set associated with \( x_i \).

Outputs

\( h : \) multi-label predictor \( \mathcal{X} \rightarrow 2^\mathcal{Y} \)
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
The Major Challenge

The MLL Mapping

- **input space**
  - features

- **output space**
  - label sets

- **Exponential number of possible label sets!**

- $q=5 \rightarrow 32$ label sets
- $q=10 \rightarrow \sim 1k$ label sets
- $q=20 \rightarrow \sim 1M$ label sets
The Basic Philosophy

Exploiting Label Correlations

For example
An image labeled as lions and grassland would be likely annotated with label Africa
A document labeled as politics would be unlikely labeled as entertainment

Order Of Correlations

- **First-Order Strategy**: Tackle MLL problem in a label-by-label style, ignore the co-existence of other labels.
- **Second-Order Strategy**: Tackle MLL problem by considering pairwise relations between labels.
- **High-Order Strategy**: Tackle MLL problem by considering high-order relations between labels.
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
The Major Evaluation Metrics

**Example-based**
- **Classification**
  - \( \text{Subset Accuracy, Hamming Loss,} \)
  - \( \text{Accuracy}_{\text{exam}}, \text{Precision}_{\text{exam}}, \text{Recall}_{\text{exam}}, F^\beta_{\text{exam}} \)

- **Ranking**
  - \( \text{One-error, Coverage, Ranking Loss, Average Precision} \)

**Label-based**
- **Classification**
  - \( B_{\text{macro}}, B_{\text{micro}} \) \text{ (macro/micro-averaging)}
  - \( B \in \{ \text{Accuracy, Precision, Recall, } F^\beta \} \)

- **Ranking**
  - \( AUC_{\text{macro}}, AUC_{\text{micro}} \)
The Major Learning Algorithm

Multi-label learning algorithms

Problem transformation

- Transform to binary classification
  - Binary Relevance
  - Classifier Chains

- Transform to label ranking
  - Calibrated Label Ranking

- Transform to multi-class classification
  - Random k-labelsets

Algorithm adaptation

- Lazy learning
  - ML-kNN

- Decision tree
  - ML-DT

- Kernel learning
  - Rank-SVM

- Information-theoretic
  - CML
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
Basic Idea

The Classifiers Chain model (CC) involves $|L|$ binary classifiers as in BM. Classifiers are linked along a chain where each classifier deals with the binary relevance problem associated with label $l_j \in L$. The feature space of each classifier in the chain is extended with the 0/1 label associations of all previous classifiers. Table 1 shows a simple example of CC model with the input $x_1 = [1, 0, 1, 0, 0, 1]$. Finally, it forms a classifiers chain.

<table>
<thead>
<tr>
<th>$h_i$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1 : [1, 0, 1, 0, 0, 1]$</td>
<td>1</td>
</tr>
<tr>
<td>$h_2 : [1, 0, 1, 0, 0, 1, 1]$</td>
<td>0</td>
</tr>
<tr>
<td>$h_3 : [1, 0, 1, 0, 0, 1, 1, 0]$</td>
<td>1</td>
</tr>
<tr>
<td>$h_4 : [1, 0, 1, 0, 0, 1, 1, 0, 1]$</td>
<td>0</td>
</tr>
<tr>
<td>$h_5 : [1, 0, 1, 0, 0, 1, 1, 0, 1, 0]$</td>
<td>0</td>
</tr>
<tr>
<td>$h_6 : [1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0]$</td>
<td>1</td>
</tr>
</tbody>
</table>

- **Pros**: more appropriate for realistic correlations.
- **Cons**: high model complexity, less scalable.
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
My Proposed Method

Framework

1. The conditional entropy builds a complete label-correlated graph, according the maximum conditional entropy principle of the deleted loop, we obtain a directed acyclic graph, i.e. the initial network.

2. The initial network is sorted by topology, and the initial order is obtained. According to algorithm 1, the optimal parent label sets is obtained.

3. To get the optimal Bayesian networks(BN) structure according to Algorithm 2.

4. Topological sorting of BN labels, training and testing.
Conditional Entropy

If $H(Y|X = x)$ is the entropy of the discrete random variable $Y$ conditioned on the discrete random variable $X$ taking a certain value $x$, then $H(Y|X)$ is the result of averaging $H(Y|X = x)$ over all possible values $x$ that $X$ may take.

\[
H(L_j|L_i) = \sum_{l_i \in L_i} p(l_i) H(L_j|L_i = l_i)
\]

\[
= - \sum_{l_i \in L_i} p(l_i) \sum_{l_j \in L_j} p(l_j|l_i) \log(p(l_j|l_i))
\]

\[
= - \sum_{l_i \in L_i} \sum_{l_j \in L_j} p(l_i, l_j) \log p(l_j|l_i)
\]

\[
= - \sum_{l_i \in L_i} \sum_{l_j \in L_j} p(l_i, l_j) \log \frac{p(l_i, l_j)}{p(l_i)}
\]

(1)

correlational entropy is measured correlation between labels.
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
Bayesian information Criterion Scoring Function

\[ S(G) = L_D(G) - \frac{\text{Dim}_G}{2} \log N \]  

(2)

\[ L_D(G) = -N \sum_{i=1}^{L} H(l_i | \text{PA}(l_i)) \]  

(3)

\[ S(G) = \sum_{i=1}^{L} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} N_{ijk} \log \frac{N_{ijk}}{N_{ij}} - \sum_{i=1}^{L} q_i (r_i - 1) \log N \]  

(4)
**Algorithm 1** learning the optimal parent label sets

**input:** $D$: training data set containing $L$ labels $(l_1, l_2, l_3, \ldots, l_L)$

**initial order:** $l_1, l_2, l_3, \ldots, l_L$

**output:** $PA$: optimal parent label sets $PA$ of each label

1: $PA = \emptyset$
2: for $i = 1$ to $L$ do
3: \quad $Pred_i = [l_1, l_2, \ldots, l_{i-1}, l_{i+1}, \ldots, l_L]$;
4: \quad $PA(l_i) = \emptyset$;
5: \quad $maxBIC = BIC(l_i|PA(l_i))$;
6: \quad **while** $j \leq L \cup N_{PA(l_i)} \leq \log_2 N$ **do**
7: \quad \quad let $l_z$ be the label in $Pred_i$ $PA(l_i)$ that maximizes $BIC(l_i|PA(l_i) \cup l_z)$;
8: \quad \quad $BIC_{new} = BIC(l_i|PA(l_i) \cup l_z)$;
9: \quad \quad if $BIC > maxBIC$ then
10: \quad \quad \quad $PA(l_i) = PA(l_i) \cup l_z$;
11: \quad \quad end if
12: \quad end while
13: \quad $PA = PA \cup PA(l_i)$
14: end for
15: return $PA$;
Learning Bayesian networks structure

Algorithm 2 learning the optimal Bayesian networks structure

input: G: A directed graph built by best parent label sets
output: G’: the optimal Bayesian networks (a directed acyclic graph)

1: while G has circles do
2:    CG=circle
3:    maxweight=-inf;
4:    for i = 1 to size(CG, 1) do
5:        for j = 1 to size(CG, 1) do
6:            if CG(i,j) > maxweight then
7:                maxweight = CG(i,j);
8:                row=i;
9:                column=j;
10:           end if
11:        end for
12:    end for
13:    CG(row,column)=0;
14: end while
15: G’=G;
16: return G’;
Outline

1. Multi-label Learning (MLL)
   - What is MLL
   - Challenge and Philosophy
   - Evaluation Metrics And Learning Algorithms

2. Classifiers Chain
   - What is Classifiers Chain

3. My Proposed Method
   - Initial network
   - Bayesian networks structure
   - Experimental result
**Table: characteristics of data sets**

<table>
<thead>
<tr>
<th>data sets</th>
<th>Instances</th>
<th>Labels</th>
<th>Features</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>emotions</td>
<td>593</td>
<td>6</td>
<td>72</td>
<td>music</td>
</tr>
<tr>
<td>scene</td>
<td>2407</td>
<td>6</td>
<td>294</td>
<td>image</td>
</tr>
<tr>
<td>flags</td>
<td>194</td>
<td>7</td>
<td>19</td>
<td>image</td>
</tr>
<tr>
<td>yeast</td>
<td>2417</td>
<td>14</td>
<td>103</td>
<td>biology</td>
</tr>
<tr>
<td>art</td>
<td>5000</td>
<td>26</td>
<td>462</td>
<td>text</td>
</tr>
<tr>
<td>genbase</td>
<td>662</td>
<td>27</td>
<td>1185</td>
<td>biology</td>
</tr>
<tr>
<td>education</td>
<td>5000</td>
<td>33</td>
<td>550</td>
<td>text</td>
</tr>
<tr>
<td>science</td>
<td>5000</td>
<td>40</td>
<td>743</td>
<td>text</td>
</tr>
<tr>
<td>medical</td>
<td>978</td>
<td>45</td>
<td>1149</td>
<td>text</td>
</tr>
<tr>
<td>enron</td>
<td>1702</td>
<td>53</td>
<td>1001</td>
<td>text</td>
</tr>
<tr>
<td>CAL500</td>
<td>502</td>
<td>174</td>
<td>68</td>
<td>music</td>
</tr>
</tbody>
</table>
### Hamming Loss

<table>
<thead>
<tr>
<th>datasets</th>
<th>BR</th>
<th>LR</th>
<th>CC</th>
<th>GCC</th>
<th>LPLC</th>
<th>BNCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>emotions</td>
<td>0.1805 ± 0.0065</td>
<td><strong>0.1795 ± 0.0103</strong></td>
<td>0.1827 ± 0.0094</td>
<td>0.3163 ± 0.0351</td>
<td>0.3832 ± 0.0236</td>
<td>0.1815 ± 0.0038</td>
</tr>
<tr>
<td>flags</td>
<td>0.2780 ± 0.0127</td>
<td>0.2681 ± 0.0105</td>
<td>0.2688 ± 0.0153</td>
<td>0.3220 ± 0.0336</td>
<td>0.4411 ± 0.0303</td>
<td><strong>0.2645 ± 0.0020</strong></td>
</tr>
<tr>
<td>scene</td>
<td>0.0787 ± 0.0035</td>
<td>0.0775 ± 0.0035</td>
<td>0.0800 ± 0.0038</td>
<td>0.1030 ± 0.0091</td>
<td>0.1046 ± 0.0076</td>
<td>0.1508 ± 0.0021</td>
</tr>
<tr>
<td>yeast</td>
<td>0.1948 ± 0.0042</td>
<td>0.1981 ± 0.0037</td>
<td>0.1934 ± 0.0041</td>
<td><strong>0.1884 ± 0.0228</strong></td>
<td>0.2267 ± 0.0088</td>
<td>0.0168 ± 0.0021</td>
</tr>
<tr>
<td>medical</td>
<td>0.0094 ± 0.0006</td>
<td>0.0249 ± 0.0012</td>
<td>0.0004 ± 0.0006</td>
<td>0.7745 ± 0.0404</td>
<td>0.6898 ± 0.0076</td>
<td><strong>0.0092 ± 0.0004</strong></td>
</tr>
<tr>
<td>enron</td>
<td>0.0506 ± 0.0028</td>
<td>0.0518 ± 0.0028</td>
<td>0.0501 ± 0.0026</td>
<td><strong>0.0487 ± 0.0019</strong></td>
<td>0.0632 ± 0.0020</td>
<td>0.0008 ± 0.0001</td>
</tr>
<tr>
<td>genbase</td>
<td>0.0005 ± 0.0007</td>
<td>0.0265 ± 0.0050</td>
<td>0.0009 ± 0.0008</td>
<td>0.0073 ± 0.0006</td>
<td>0.0415 ± 0.0076</td>
<td>0.0334 ± 0.0007</td>
</tr>
<tr>
<td>science</td>
<td><strong>0.0330 ± 0.0011</strong></td>
<td>0.0347 ± 0.0007</td>
<td>0.0337 ± 0.0007</td>
<td>0.0330 ± 0.0016</td>
<td>0.1014 ± 0.0186</td>
<td>0.0539 ± 0.0003</td>
</tr>
<tr>
<td>art</td>
<td><strong>0.0523 ± 0.0009</strong></td>
<td>0.0570 ± 0.0008</td>
<td>0.0539 ± 0.0012</td>
<td>0.0650 ± 0.0146</td>
<td>0.3692 ± 0.0581</td>
<td>0.1794 ± 0.0020</td>
</tr>
<tr>
<td>CAL500</td>
<td>0.1376 ± 0.0043</td>
<td>0.1464 ± 0.0033</td>
<td>0.1888 ± 0.0150</td>
<td>0.1429 ± 0.0043</td>
<td>0.0482 ± 0.0055</td>
<td>0.0380 ± 0.0006</td>
</tr>
<tr>
<td>education</td>
<td>0.0376 ± 0.0009</td>
<td>0.0415 ± 0.0007</td>
<td>0.0385 ± 0.0004</td>
<td>0.0434 ± 0.0102</td>
<td>0.0482 ± 0.0055</td>
<td>0.0380 ± 0.0006</td>
</tr>
</tbody>
</table>

### Exam Accuracy

<table>
<thead>
<tr>
<th>datasets</th>
<th>BR</th>
<th>LR</th>
<th>CC</th>
<th>GCC</th>
<th>LPLC</th>
<th>BNCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>emotions</td>
<td>0.5627 ± 0.0180</td>
<td>0.5518 ± 0.0217</td>
<td>0.5830 ± 0.0233</td>
<td>0.4024 ± 0.0490</td>
<td>0.4953 ± 0.0279</td>
<td><strong>0.5837 ± 0.0057</strong></td>
</tr>
<tr>
<td>flags</td>
<td>0.5770 ± 0.0229</td>
<td>0.5703 ± 0.0204</td>
<td>0.5842 ± 0.0158</td>
<td>0.5070 ± 0.0440</td>
<td>0.5103 ± 0.0339</td>
<td>0.5864 ± 0.0058</td>
</tr>
<tr>
<td>scene</td>
<td>0.7153 ± 0.0144</td>
<td>0.6975 ± 0.0094</td>
<td>0.7361 ± 0.0128</td>
<td>0.6814 ± 0.0289</td>
<td>0.6451 ± 0.0255</td>
<td><strong>0.7414 ± 0.0046</strong></td>
</tr>
<tr>
<td>yeast</td>
<td>0.5068 ± 0.0983</td>
<td>0.4778 ± 0.0985</td>
<td>0.5251 ± 0.0694</td>
<td>0.2602 ± 0.0102</td>
<td>0.5246 ± 0.0193</td>
<td>0.5350 ± 0.0035</td>
</tr>
<tr>
<td>medical</td>
<td>0.7610 ± 0.0138</td>
<td>0.1857 ± 0.0419</td>
<td>0.7727 ± 0.0140</td>
<td><strong>0.8187 ± 0.0410</strong></td>
<td>0.6706 ± 0.0435</td>
<td>0.7844 ± 0.0014</td>
</tr>
<tr>
<td>enron</td>
<td>0.4416 ± 0.0267</td>
<td>0.3599 ± 0.0397</td>
<td>0.4493 ± 0.0222</td>
<td><strong>0.4626 ± 0.0211</strong></td>
<td>0.6088 ± 0.0763</td>
<td>0.4517 ± 0.0046</td>
</tr>
<tr>
<td>genbase</td>
<td>0.9919 ± 0.0104</td>
<td>0.3695 ± 0.2038</td>
<td>0.9899 ± 0.0889</td>
<td>0.9917 ± 0.0085</td>
<td>0.5715 ± 0.0149</td>
<td><strong>0.9923 ± 0.0023</strong></td>
</tr>
<tr>
<td>science</td>
<td>0.3135 ± 0.0177</td>
<td>0.1150 ± 0.0088</td>
<td>0.3175 ± 0.0123</td>
<td><strong>0.3671 ± 0.0206</strong></td>
<td>0.2570 ± 0.0193</td>
<td>0.3254 ± 0.0137</td>
</tr>
<tr>
<td>art</td>
<td>0.2990 ± 0.0081</td>
<td>0.1456 ± 0.0143</td>
<td>0.3114 ± 0.0065</td>
<td><strong>0.3713 ± 0.0730</strong></td>
<td>0.0521 ± 0.0462</td>
<td>0.3128 ± 0.0055</td>
</tr>
<tr>
<td>CAL500</td>
<td>0.2051 ± 0.0083</td>
<td>0.0909 ± 0.0102</td>
<td>0.2021 ± 0.0192</td>
<td><strong>0.2106 ± 0.0158</strong></td>
<td>0.2376 ± 0.0137</td>
<td>0.2054 ± 0.0034</td>
</tr>
<tr>
<td>education</td>
<td>0.3171 ± 0.0104</td>
<td>0.1092 ± 0.0104</td>
<td>0.3260 ± 0.0070</td>
<td><strong>0.3596 ± 0.0571</strong></td>
<td>0.3159 ± 0.0262</td>
<td>0.3292 ± 0.0066</td>
</tr>
</tbody>
</table>
### My Proposed Method

#### Experimental result

<table>
<thead>
<tr>
<th>datasets</th>
<th>BR</th>
<th>LR</th>
<th>CC</th>
<th>GCC</th>
<th>LPLC</th>
<th>BNCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>emotions</td>
<td>0.5558 ± 0.0210</td>
<td>0.6571 ± 0.0155</td>
<td>0.6724 ± 0.0201</td>
<td>0.4775 ± 0.0687</td>
<td>0.5120 ± 0.0408</td>
<td>0.6736 ± 0.0075</td>
</tr>
<tr>
<td>flags</td>
<td>0.5912 ± 0.0361</td>
<td>0.6449 ± 0.0303</td>
<td>0.6081 ± 0.0240</td>
<td>0.4361 ± 0.0489</td>
<td>0.6926 ± 0.0295</td>
<td>0.6219 ± 0.0214</td>
</tr>
<tr>
<td>scene</td>
<td>0.7809 ± 0.0991</td>
<td>0.7750 ± 0.0086</td>
<td>0.7769 ± 0.0119</td>
<td>0.7317 ± 0.0254</td>
<td>0.6531 ± 0.0220</td>
<td>0.7891 ± 0.0051</td>
</tr>
<tr>
<td>yeast</td>
<td>0.3355 ± 0.0061</td>
<td>0.3181 ± 0.0040</td>
<td>0.3028 ± 0.0079</td>
<td>0.3535 ± 0.0156</td>
<td>0.4501 ± 0.0207</td>
<td>0.3717 ± 0.0034</td>
</tr>
<tr>
<td>medical</td>
<td>0.3885 ± 0.0158</td>
<td>0.0837 ± 0.0272</td>
<td>0.3375 ± 0.0124</td>
<td>0.3185 ± 0.0289</td>
<td>0.1918 ± 0.0223</td>
<td>0.3401 ± 0.0110</td>
</tr>
<tr>
<td>enron</td>
<td>0.1940 ± 0.0230</td>
<td>0.1607 ± 0.0261</td>
<td>0.1942 ± 0.0205</td>
<td>0.1642 ± 0.0137</td>
<td>0.0971 ± 0.0126</td>
<td>0.1852 ± 0.0070</td>
</tr>
<tr>
<td>genbase</td>
<td>0.6550 ± 0.0435</td>
<td>0.3073 ± 0.1402</td>
<td>0.6357 ± 0.0697</td>
<td>0.6387 ± 0.0655</td>
<td>0.5801 ± 0.0238</td>
<td>0.6582 ± 0.0023</td>
</tr>
<tr>
<td>science</td>
<td>0.2073 ± 0.0246</td>
<td>0.1301 ± 0.0137</td>
<td>0.2147 ± 0.0155</td>
<td>0.1743 ± 0.0172</td>
<td>0.1712 ± 0.0172</td>
<td>0.2068 ± 0.0075</td>
</tr>
<tr>
<td>art</td>
<td>0.2113 ± 0.0348</td>
<td>0.1326 ± 0.0156</td>
<td>0.2212 ± 0.0150</td>
<td>0.1802 ± 0.0171</td>
<td>0.0967 ± 0.0099</td>
<td>0.2185 ± 0.0067</td>
</tr>
<tr>
<td>CAL500</td>
<td>0.0613 ± 0.0088</td>
<td>0.0376 ± 0.0043</td>
<td>0.1088 ± 0.0110</td>
<td>0.0552 ± 0.0070</td>
<td>0.2194 ± 0.0105</td>
<td>0.1162 ± 0.0037</td>
</tr>
<tr>
<td>education</td>
<td>0.1901 ± 0.0115</td>
<td>0.1240 ± 0.0093</td>
<td>0.1883 ± 0.0155</td>
<td>0.1660 ± 0.0282</td>
<td>0.1066 ± 0.0185</td>
<td>0.1901 ± 0.0109</td>
</tr>
</tbody>
</table>

#### macroF

<table>
<thead>
<tr>
<th>datasets</th>
<th>BR</th>
<th>LR</th>
<th>CC</th>
<th>GCC</th>
<th>LPLC</th>
<th>BNCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>emotions</td>
<td>0.5872 ± 0.0165</td>
<td>0.0831 ± 0.0191</td>
<td>0.5990 ± 0.0189</td>
<td>0.5093 ± 0.0663</td>
<td>0.5312 ± 0.0400</td>
<td>0.6979 ± 0.0068</td>
</tr>
<tr>
<td>flags</td>
<td>0.7176 ± 0.0206</td>
<td>0.7233 ± 0.0164</td>
<td>0.7236 ± 0.0140</td>
<td>0.6007 ± 0.0339</td>
<td>0.6741 ± 0.0282</td>
<td>0.7310 ± 0.0059</td>
</tr>
<tr>
<td>scene</td>
<td>0.7723 ± 0.0604</td>
<td>0.7692 ± 0.0099</td>
<td>0.7697 ± 0.0167</td>
<td>0.7683 ± 0.0265</td>
<td>0.6892 ± 0.0221</td>
<td>0.7586 ± 0.0045</td>
</tr>
<tr>
<td>yeast</td>
<td>0.5403 ± 0.0070</td>
<td>0.4174 ± 0.0074</td>
<td>0.5166 ± 0.0080</td>
<td>0.5434 ± 0.0186</td>
<td>0.6468 ± 0.0161</td>
<td>0.6601 ± 0.0034</td>
</tr>
<tr>
<td>medical</td>
<td>0.8218 ± 0.0112</td>
<td>0.1919 ± 0.0025</td>
<td>0.8209 ± 0.0104</td>
<td>0.8202 ± 0.0034</td>
<td>0.6785 ± 0.0080</td>
<td>0.8202 ± 0.0065</td>
</tr>
<tr>
<td>enron</td>
<td>0.5580 ± 0.0217</td>
<td>0.4810 ± 0.0344</td>
<td>0.5614 ± 0.0214</td>
<td>0.5785 ± 0.0167</td>
<td>0.4967 ± 0.0188</td>
<td>0.5941 ± 0.0016</td>
</tr>
<tr>
<td>genbase</td>
<td>0.9934 ± 0.0078</td>
<td>0.5530 ± 0.2397</td>
<td>0.9906 ± 0.0081</td>
<td>0.9926 ± 0.0069</td>
<td>0.9657 ± 0.0188</td>
<td>0.9941 ± 0.0016</td>
</tr>
<tr>
<td>science</td>
<td>0.4016 ± 0.0178</td>
<td>0.1955 ± 0.0142</td>
<td>0.4156 ± 0.0192</td>
<td>0.4318 ± 0.0250</td>
<td>0.4044 ± 0.0130</td>
<td>0.4044 ± 0.0130</td>
</tr>
<tr>
<td>art</td>
<td>0.3822 ± 0.0081</td>
<td>0.2222 ± 0.0212</td>
<td>0.3877 ± 0.0118</td>
<td>0.3952 ± 0.0491</td>
<td>0.0577 ± 0.0054</td>
<td>0.3858 ± 0.0064</td>
</tr>
<tr>
<td>CAL500</td>
<td>0.3338 ± 0.0114</td>
<td>0.1652 ± 0.0158</td>
<td>0.3322 ± 0.0257</td>
<td>0.3380 ± 0.0210</td>
<td>0.3813 ± 0.0170</td>
<td>0.3365 ± 0.0045</td>
</tr>
<tr>
<td>education</td>
<td>0.4206 ± 0.0107</td>
<td>0.1920 ± 0.0152</td>
<td>0.4217 ± 0.0085</td>
<td>0.4234 ± 0.0456</td>
<td>0.3855 ± 0.0281</td>
<td>0.4243 ± 0.0104</td>
</tr>
</tbody>
</table>
Thank you