Appendix for
Implicit Relative Labeling-Importance Aware Multi-Label Metric Learning

Jun-Xiang Mao[1,2,3], **Yong Rui**[4], **Min-Ling Zhang**[1,2*]

[1]School of Computer Science and Engineering, Southeast University, Nanjing 210096, China
[2]Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, China
[3]Information Technology and Data Management Department of China Mobile Communications Group Zhejiang Co., Ltd
[4]Lenovo Research, Lenovo Group Ltd., Beijing, China
maojx@seu.edu.cn, yongrui@lenovo.com, zhangml@seu.edu.cn

A The Complete Procedure of ILIA

The complete procedure of the proposed ILIA approach is summarized in Algorithm 1. Firstly, the manifold structure in the feature space is exploited by local linear reconstruction (Step 1-3), and then the implicit RLIs are recovered by transferring such structure to the label space (Step 4). After that, the instance matrix $\mathbf{X}$ and the recovered RLI matrix $\mathbf{F}$ are centered for ease of solution (Step 5), and the metric $\mathbf{M}$ to be learned is initialized with an identity matrix (Step 6). Finally, an alternating optimization procedure is performed to obtain $\mathbf{\Theta}$ and $\mathbf{M}$ (Step 7-10).

Algorithm 1: The proposed ILIA approach

Input: $\mathcal{D}$: multi-label training set $\{(\boldsymbol{x}_i, \boldsymbol{y}_i) \mid 1 \le i \le n\}$;
μ, η, γ: trade-off parameters;
k: the number of nearest neighbors;
κ: the kernel function.
Output: $\mathbf{\Theta}$: the learned predictive modeling coefficients;
$\mathbf{M}$: the learned discriminative metric.

1: **for** $i = 1$ to n **do**
2: Calculate the closed-form solution $\boldsymbol{s}_i$ according to Eq.(8);
3: **end for**
4: Obtain the implicit RLI matrix $\mathbf{F}$ according to Eq.(12);
5: Centering the instance matrix $\mathbf{X}$ and RLI matrix $\mathbf{F}$ to obtain their centered version $\hat{\mathbf{X}}$ and $\hat{\mathbf{F}}$;
6: Initialize $\mathbf{M} = \mathbf{I}_q$;
7: **repeat**
8: Calculate $\mathbf{\Theta}$ by solving the Sylvester equation in Eq.(19);
9: Calculate $\mathbf{M}$ by solving the Riccati equation according to Eq.(24);
10: **until** convergence
11: **return** $\mathbf{\Theta}$ and $\mathbf{M}$.

After the above optimization procedure converges, the semantic similarity between a pair multi-label instances $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ can be explicitly formalized in the form of the Mahalanobis distance shown in Eq.(25) of the main body.

*Corresponding author.

B Main Experiments

B.1 Compared Methods

We compare ILIA with five state-of-the-art multi-label metric learning methods shown below. For their parameters, we use the suggested configurations in respective literature:

- LM (Liu and Tsang 2015): LM employs a large margin formulation to establish a unified metric space, maintaining the correlation between feature and label spaces [suggested configuration: $\eta = 0.4, C = 10$].
- LJE (Gouk, Pfahringer, and Cree 2016): LJE learns a metric that projects instances into a space where the Euclidean distance closely mirrors the Jaccard similarity of multiple labels [suggested configuration: $t = 32, e = 5$].
- COMMU (Sun and Zhang 2021): COMMU constructs a compositional metric by modeling structural interactions between feature and label spaces, exploring the integrated semantics of all labels [suggested configuration: $\alpha, \theta \in \{0.2, 0.4, \ldots, 0.8\}, C = 10$].
- LIMIC (Mao, Wang, and Zhang 2023): LIMIC learns one global and multiple label-specific metrics for each label by exploiting label-specific side information [suggested configuration: $\gamma = 2, \lambda_1, \lambda_2 \in \{10^{-3}, 10^{-1}, \ldots, 10^{3}\}$].
- LSMM (Mao, Hang, and Zhang 2024): LSMM learns multiple label-specific local metrics for different multi-label instances on the shoulder of a global one [suggested configuration: mode = '*semantic-based partition*', $k_t = k_i = 20, \alpha = 0.4, \gamma = 2, \lambda_1 \in \{10^{-1}, 1, \ldots, 10^{3}\}, \lambda_2 \in \{10^{-3}, 10^{-2}, \ldots, 10\}$].

B.2 More Results

Table B.1 reports detailed empirical results with five state-of-art multi-label metric learning algorithms in terms of *Ranking Loss*, *Coverage*, *Macro-F1*, and *Macro-averaging AUC*, which are not covered in the main body due to page limit. The best and second best results are highlighted in **boldface** and underline respectively. Furthermore, to statistically evaluate the improvement significance, pairwise t-test at 5% significance level is conducted on the empirical results. •/○ indicates whether $\mathcal{A}$-ILIA achieves significantly superior/inferior to other compared approaches. The resulting win/tie/loss counts are reported in Table B.2. These results clearly demonstrate the superiority of our proposed ILIA approach in learning effective similarity metrics between multi-label examples.

Compared Algorithms	Datasets									
	CAL500	emotions	medical	image	scene	arts	corel5k	education	health	entertainment
					Ranking Loss ↓					
BRKNN	.255$_{\pm.011}$●	.272$_{\pm.048}$●	.092$_{\pm.029}$	.186$_{\pm.020}$●	.096$_{\pm.011}$○	.377$_{\pm.040}$●	.614$_{\pm.022}$●	<u>.208</u>$_{\pm.011}$●	.201$_{\pm.014}$●	.288$_{\pm.014}$●
BRKNN-LM	.258$_{\pm.007}$●	.256$_{\pm.030}$●	.116$_{\pm.027}$●	.187$_{\pm.020}$●	.107$_{\pm.012}$	.267$_{\pm.013}$●	<u>.481</u>$_{\pm.015}$	.230$_{\pm.010}$●	.196$_{\pm.011}$●	.284$_{\pm.014}$●
BRKNN-LJE	.258$_{\pm.011}$●	.202$_{\pm.035}$●	.114$_{\pm.037}$●	.205$_{\pm.020}$●	.124$_{\pm.019}$●	.265$_{\pm.015}$●	.553$_{\pm.009}$●	.234$_{\pm.012}$●	.214$_{\pm.014}$●	.284$_{\pm.012}$●
BRKNN-COMMU	.253$_{\pm.011}$●	.272$_{\pm.048}$●	**.087**$_{\pm.028}$	.185$_{\pm.020}$●	.096$_{\pm.011}$○	.246$_{\pm.015}$●	**.480**$_{\pm.017}$	.198$_{\pm.011}$●	.201$_{\pm.014}$●	.288$_{\pm.014}$●
BRKNN-LIMIC	.252$_{\pm.014}$●	.197$_{\pm.028}$●	.089$_{\pm.021}$	.178$_{\pm.026}$●	.113$_{\pm.011}$●	.359$_{\pm.024}$●	.483$_{\pm.020}$	.223$_{\pm.015}$●	.167$_{\pm.010}$●	.251$_{\pm.013}$●
BRKNN-LSMM	<u>.247</u>$_{\pm.016}$●	<u>.188</u>$_{\pm.033}$●	.091$_{\pm.018}$	<u>.170</u>$_{\pm.021}$	.110$_{\pm.015}$	<u>.263</u>$_{\pm.028}$●	.488$_{\pm.014}$●	.210$_{\pm.010}$●	<u>.145</u>$_{\pm.009}$●	<u>.244</u>$_{\pm.009}$●
BRKNN-ILIA (Ours)	**.221**$_{\pm.009}$	**.183**$_{\pm.018}$	<u>.088</u>$_{\pm.026}$	**.160**$_{\pm.021}$	**.106**$_{\pm.019}$	**.175**$_{\pm.010}$	**.480**$_{\pm.013}$	**.142**$_{\pm.010}$	**.132**$_{\pm.009}$	**.206**$_{\pm.005}$
MLKNN	**.183**$_{\pm.005}$○	.258$_{\pm.038}$●	**.033**$_{\pm.012}$○	.176$_{\pm.019}$●	<u>.077</u>$_{\pm.010}$	.149$_{\pm.010}$●	.134$_{\pm.003}$○	<u>.078</u>$_{\pm.004}$	.070$_{\pm.004}$●	**.130**$_{\pm.004}$
MLKNN-LM	<u>.184</u>$_{\pm.004}$○	.240$_{\pm.026}$●	.038$_{\pm.012}$	.176$_{\pm.018}$●	.084$_{\pm.010}$	.130$_{\pm.007}$●	<u>.126</u>$_{\pm.004}$○	.079$_{\pm.003}$	**.066**$_{\pm.003}$	**.130**$_{\pm.006}$
MLKNN-LJE	.184$_{\pm.005}$	.201$_{\pm.030}$●	.053$_{\pm.017}$	.201$_{\pm.022}$●	.110$_{\pm.017}$●	.145$_{\pm.009}$●	.148$_{\pm.004}$	.089$_{\pm.003}$●	.084$_{\pm.005}$●	.136$_{\pm.004}$●
MLKNN-COMMU	**.183**$_{\pm.005}$○	.258$_{\pm.038}$●	.041$_{\pm.013}$	.175$_{\pm.020}$●	<u>.077</u>$_{\pm.009}$	.151$_{\pm.010}$●	.137$_{\pm.003}$○	**.077**$_{\pm.004}$	.070$_{\pm.004}$●	**.130**$_{\pm.004}$
MLKNN-LIMIC	**.183**$_{\pm.004}$○	<u>.197</u>$_{\pm.020}$	.039$_{\pm.012}$	.158$_{\pm.025}$	.080$_{\pm.011}$	.138$_{\pm.006}$●	**.123**$_{\pm.008}$○	<u>.078</u>$_{\pm.004}$	.068$_{\pm.004}$	.132$_{\pm.005}$
MLKNN-LSMM	.185$_{\pm.004}$	**.188**$_{\pm.024}$	<u>.034</u>$_{\pm.015}$○	<u>.157</u>$_{\pm.021}$	**.072**$_{\pm.004}$○	<u>.133</u>$_{\pm.008}$●	.136$_{\pm.005}$○	<u>.078</u>$_{\pm.004}$	<u>.067</u>$_{\pm.003}$	**.130**$_{\pm.004}$
MLKNN-ILIA (Ours)	.189$_{\pm.005}$	**.188**$_{\pm.017}$	.048$_{\pm.014}$	**.154**$_{\pm.018}$	.083$_{\pm.010}$	**.120**$_{\pm.007}$	.151$_{\pm.005}$	**.077**$_{\pm.004}$	**.066**$_{\pm.004}$	<u>.131</u>$_{\pm.003}$
					Coverage ↓					
BRKNN	.840$_{\pm.025}$●	.378$_{\pm.032}$●	.085$_{\pm.027}$●	.197$_{\pm.021}$●	<u>.087</u>$_{\pm.011}$	.335$_{\pm.032}$●	.425$_{\pm.010}$●	.157$_{\pm.007}$●	.167$_{\pm.008}$●	.208$_{\pm.005}$●
BRKNN-LM	.833$_{\pm.015}$●	.364$_{\pm.032}$●	.094$_{\pm.023}$●	.194$_{\pm.021}$●	.095$_{\pm.009}$	.268$_{\pm.009}$●	<u>.376</u>$_{\pm.012}$○	.169$_{\pm.010}$●	.160$_{\pm.009}$●	.212$_{\pm.007}$●
BRKNN-LJE	.835$_{\pm.022}$●	.330$_{\pm.026}$	.094$_{\pm.031}$●	.211$_{\pm.020}$●	.108$_{\pm.016}$●	.277$_{\pm.012}$●	.440$_{\pm.008}$●	.170$_{\pm.001}$●	.187$_{\pm.009}$●	.213$_{\pm.006}$●
BRKNN-COMMU	.837$_{\pm.024}$●	.378$_{\pm.032}$●	.081$_{\pm.026}$	.197$_{\pm.021}$●	<u>.087</u>$_{\pm.011}$	.261$_{\pm.013}$●	**.374**$_{\pm.010}$○	.156$_{\pm.008}$●	.166$_{\pm.009}$●	.208$_{\pm.005}$●
BRKNN-LIMIC	.835$_{\pm.022}$●	<u>.318</u>$_{\pm.022}$	**.072**$_{\pm.017}$	.184$_{\pm.027}$	.089$_{\pm.009}$	.312$_{\pm.018}$●	.385$_{\pm.014}$○	.161$_{\pm.011}$●	.152$_{\pm.006}$●	<u>.195</u>$_{\pm.006}$●
BRKNN-LSMM	<u>.827</u>$_{\pm.018}$●	.329$_{\pm.021}$	.078$_{\pm.025}$	<u>.182</u>$_{\pm.018}$	**.085**$_{\pm.012}$	<u>.279</u>$_{\pm.021}$●	.379$_{\pm.010}$○	<u>.150</u>$_{\pm.015}$●	<u>.147</u>$_{\pm.009}$●	.200$_{\pm.004}$●
BRKNN-ILIA (Ours)	**.801**$_{\pm.014}$	**.310**$_{\pm.026}$	<u>.075</u>$_{\pm.019}$	**.174**$_{\pm.021}$	.089$_{\pm.016}$	**.216**$_{\pm.007}$	.400$_{\pm.012}$	**.135**$_{\pm.006}$	**.137**$_{\pm.007}$	**.190**$_{\pm.004}$
MLKNN	**.745**$_{\pm.018}$	.377$_{\pm.026}$●	**.047**$_{\pm.014}$○	.195$_{\pm.021}$●	**.078**$_{\pm.010}$	.207$_{\pm.010}$●	.304$_{\pm.011}$○	**.104**$_{\pm.004}$	.111$_{\pm.005}$●	**.164**$_{\pm.004}$
MLKNN-LM	.755$_{\pm.018}$	.360$_{\pm.033}$●	.054$_{\pm.014}$○	.194$_{\pm.020}$●	.084$_{\pm.009}$	<u>.189</u>$_{\pm.007}$●	**.288**$_{\pm.013}$○	.107$_{\pm.005}$	**.107**$_{\pm.005}$	**.164**$_{\pm.005}$
MLKNN-LJE	.752$_{\pm.014}$	.329$_{\pm.030}$	.070$_{\pm.021}$	.213$_{\pm.021}$●	.106$_{\pm.015}$●	.205$_{\pm.008}$●	.331$_{\pm.011}$	.118$_{\pm.006}$●	.130$_{\pm.007}$●	.170$_{\pm.004}$●
MLKNN-COMMU	<u>.747</u>$_{\pm.018}$	.377$_{\pm.026}$●	.059$_{\pm.018}$	.194$_{\pm.022}$●	**.078**$_{\pm.011}$	.208$_{\pm.010}$●	.310$_{\pm.012}$	**.104**$_{\pm.005}$	.111$_{\pm.005}$●	**.164**$_{\pm.004}$
MLKNN-LIMIC	.756$_{\pm.016}$	**.304**$_{\pm.015}$	.057$_{\pm.014}$	.185$_{\pm.024}$	.086$_{\pm.010}$	.198$_{\pm.007}$●	.298$_{\pm.017}$○	.107$_{\pm.006}$	<u>.109</u>$_{\pm.008}$	**.164**$_{\pm.006}$
MLKNN-LSMM	.758$_{\pm.017}$	.318$_{\pm.026}$	<u>.053</u>$_{\pm.015}$○	<u>.179</u>$_{\pm.018}$	.085$_{\pm.008}$	.194$_{\pm.007}$●	<u>.289</u>$_{\pm.015}$○	<u>.105</u>$_{\pm.002}$	.110$_{\pm.007}$	<u>.165</u>$_{\pm.005}$
MLKNN-ILIA (Ours)	.750$_{\pm.016}$	<u>.314</u>$_{\pm.021}$	.069$_{\pm.015}$	**.176**$_{\pm.016}$	<u>.083</u>$_{\pm.009}$	**.178**$_{\pm.007}$	.324$_{\pm.013}$	<u>.105</u>$_{\pm.004}$	**.107**$_{\pm.005}$	<u>.165</u>$_{\pm.003}$
					Macro-F1 ↑					
BRKNN	.095$_{\pm.005}$●	.463$_{\pm.045}$●	.154$_{\pm.017}$●	.591$_{\pm.034}$●	.721$_{\pm.014}$●	.117$_{\pm.018}$●	.005$_{\pm.001}$●	.136$_{\pm.015}$●	.087$_{\pm.012}$●	.129$_{\pm.019}$
BRKNN-LM	**.103**$_{\pm.006}$	.490$_{\pm.033}$●	.331$_{\pm.045}$●	.609$_{\pm.033}$●	.738$_{\pm.026}$●	.186$_{\pm.013}$●	.003$_{\pm.004}$●	.154$_{\pm.015}$●	<u>.157</u>$_{\pm.013}$●	.111$_{\pm.013}$●
BRKNN-LJE	.086$_{\pm.006}$●	.583$_{\pm.047}$●	.273$_{\pm.041}$●	.554$_{\pm.040}$●	.664$_{\pm.028}$●	.148$_{\pm.014}$●	.005$_{\pm.002}$●	.122$_{\pm.015}$●	.051$_{\pm.005}$●	.103$_{\pm.016}$●
BRKNN-COMMU	.101$_{\pm.008}$	.463$_{\pm.045}$●	.296$_{\pm.042}$●	.591$_{\pm.034}$●	.721$_{\pm.014}$●	.182$_{\pm.016}$●	<u>.025</u>$_{\pm.006}$	**.182**$_{\pm.027}$	.112$_{\pm.018}$●	**.142**$_{\pm.023}$
BRKNN-LIMIC	.098$_{\pm.007}$	.609$_{\pm.016}$●	.414$_{\pm.046}$●	.619$_{\pm.029}$●	<u>.760</u>$_{\pm.021}$●	.181$_{\pm.018}$●	.023$_{\pm.005}$●	.157$_{\pm.016}$	.149$_{\pm.015}$●	.132$_{\pm.024}$
BRKNN-LSMM	.098$_{\pm.008}$	<u>.611</u>$_{\pm.037}$	<u>.428</u>$_{\pm.052}$●	<u>.631</u>$_{\pm.027}$●	.755$_{\pm.023}$●	<u>.189</u>$_{\pm.018}$	.020$_{\pm.004}$●	<u>.172</u>$_{\pm.015}$	.155$_{\pm.012}$	.129$_{\pm.021}$
BRKNN-ILIA (Ours)	<u>.102</u>$_{\pm.009}$	**.637**$_{\pm.035}$	**.476**$_{\pm.055}$	**.663**$_{\pm.025}$	**.787**$_{\pm.017}$	**.203**$_{\pm.017}$	**.030**$_{\pm.004}$	.171$_{\pm.021}$	**.163**$_{\pm.016}$	<u>.137</u>$_{\pm.026}$
MLKNN	.053$_{\pm.003}$●	.361$_{\pm.050}$●	.215$_{\pm.027}$●	.575$_{\pm.030}$●	.737$_{\pm.021}$●	.095$_{\pm.012}$●	.009$_{\pm.004}$●	.121$_{\pm.016}$●	.083$_{\pm.006}$●	.096$_{\pm.020}$●
MLKNN-LM	.053$_{\pm.003}$●	.438$_{\pm.036}$●	.284$_{\pm.031}$●	.584$_{\pm.025}$●	.735$_{\pm.026}$●	.190$_{\pm.014}$●	.029$_{\pm.004}$●	.160$_{\pm.022}$●	.146$_{\pm.011}$●	.095$_{\pm.014}$●
MLKNN-LJE	.049$_{\pm.002}$●	.530$_{\pm.052}$●	.245$_{\pm.044}$●	.530$_{\pm.037}$●	.651$_{\pm.023}$●	.125$_{\pm.012}$●	.004$_{\pm.002}$●	.101$_{\pm.018}$●	.035$_{\pm.007}$●	.059$_{\pm.011}$●
MLKNN-COMMU	<u>.062</u>$_{\pm.002}$●	.361$_{\pm.050}$●	.340$_{\pm.065}$●	.575$_{\pm.030}$●	.735$_{\pm.019}$●	.083$_{\pm.012}$●	.010$_{\pm.005}$●	.155$_{\pm.025}$●	.106$_{\pm.011}$●	.107$_{\pm.025}$●
MLKNN-LIMIC	.052$_{\pm.004}$●	.587$_{\pm.028}$	.392$_{\pm.034}$●	<u>.642</u>$_{\pm.033}$	<u>.765</u>$_{\pm.017}$	.198$_{\pm.015}$●	.030$_{\pm.005}$●	.177$_{\pm.021}$	.150$_{\pm.016}$●	<u>.112</u>$_{\pm.017}$
MLKNN-LSMM	.061$_{\pm.008}$●	<u>.591</u>$_{\pm.031}$	<u>.415</u>$_{\pm.052}$●	.632$_{\pm.034}$●	.761$_{\pm.025}$	<u>.203</u>$_{\pm.027}$	<u>.034</u>$_{\pm.005}$	<u>.179</u>$_{\pm.026}$	<u>.153</u>$_{\pm.019}$●	.110$_{\pm.021}$●
MLKNN-ILIA (Ours)	**.093**$_{\pm.010}$	**.608**$_{\pm.028}$	**.506**$_{\pm.057}$	**.665**$_{\pm.023}$	**.769**$_{\pm.020}$	**.212**$_{\pm.018}$	**.036**$_{\pm.006}$	**.186**$_{\pm.023}$	**.161**$_{\pm.020}$	**.122**$_{\pm.024}$
					Macro-averaging AUC ↑					
BRKNN	.540$_{\pm.016}$	.737$_{\pm.041}$●	.867$_{\pm.034}$	.840$_{\pm.019}$●	**.936**$_{\pm.006}$	.610$_{\pm.020}$●	.547$_{\pm.008}$●	.692$_{\pm.019}$●	.618$_{\pm.024}$●	.603$_{\pm.010}$●
BRKNN-LM	.541$_{\pm.014}$	.742$_{\pm.024}$●	.866$_{\pm.035}$	.826$_{\pm.020}$●	.922$_{\pm.009}$●	.659$_{\pm.017}$●	<u>.591</u>$_{\pm.008}$○	.651$_{\pm.019}$●	.627$_{\pm.022}$●	.593$_{\pm.016}$●
BRKNN-LJE	.526$_{\pm.014}$●	.800$_{\pm.034}$●	.855$_{\pm.041}$●	.817$_{\pm.021}$●	.904$_{\pm.016}$●	.643$_{\pm.010}$●	.538$_{\pm.005}$●	.656$_{\pm.020}$●	.579$_{\pm.014}$●	.587$_{\pm.014}$●
BRKNN-COMMU	<u>.542</u>$_{\pm.018}$	.737$_{\pm.041}$●	.864$_{\pm.040}$	.840$_{\pm.019}$●	**.936**$_{\pm.007}$○	<u>.676</u>$_{\pm.022}$●	**.597**$_{\pm.007}$○	<u>.697</u>$_{\pm.017}$	.618$_{\pm.026}$●	.603$_{\pm.010}$●
BRKNN-LIMIC	<u>.542</u>$_{\pm.017}$	.817$_{\pm.022}$	.877$_{\pm.031}$	.852$_{\pm.020}$●	.924$_{\pm.007}$●	.630$_{\pm.015}$●	.581$_{\pm.006}$○	.685$_{\pm.015}$●	.630$_{\pm.016}$●	.608$_{\pm.008}$
BRKNN-LSMM	.540$_{\pm.016}$	<u>.820</u>$_{\pm.030}$	**.883**$_{\pm.041}$	<u>.853</u>$_{\pm.018}$	.928$_{\pm.010}$	.654$_{\pm.021}$●	.586$_{\pm.008}$○	.694$_{\pm.020}$	<u>.635</u>$_{\pm.019}$	<u>.610</u>$_{\pm.013}$
BRKNN-ILIA (Ours)	**.544**$_{\pm.014}$	**.823**$_{\pm.022}$	<u>.878</u>$_{\pm.031}$	**.861**$_{\pm.020}$	<u>.930</u>$_{\pm.013}$	**.708**$_{\pm.021}$	.571$_{\pm.010}$	**.702**$_{\pm.021}$	**.642**$_{\pm.018}$	**.612**$_{\pm.011}$
MLKNN	<u>.526</u>$_{\pm.011}$	.720$_{\pm.041}$●	.864$_{\pm.035}$	.834$_{\pm.019}$●	**.933**$_{\pm.006}$○	.604$_{\pm.017}$●	.545$_{\pm.007}$●	.690$_{\pm.018}$○	.615$_{\pm.023}$	.600$_{\pm.012}$
MLKNN-LM	.524$_{\pm.016}$	.729$_{\pm.025}$●	.865$_{\pm.035}$	.827$_{\pm.021}$●	<u>.921</u>$_{\pm.009}$	.659$_{\pm.017}$●	**.590**$_{\pm.007}$○	.650$_{\pm.019}$●	**.631**$_{\pm.021}$	<u>.600</u>$_{\pm.015}$
MLKNN-LJE	.514$_{\pm.011}$●	.788$_{\pm.035}$●	.852$_{\pm.041}$	.810$_{\pm.021}$●	.901$_{\pm.016}$●	.638$_{\pm.009}$●	.534$_{\pm.005}$●	.655$_{\pm.020}$●	.577$_{\pm.013}$●	.583$_{\pm.014}$●
MLKNN-COMMU	<u>.526</u>$_{\pm.012}$	.720$_{\pm.041}$●	.846$_{\pm.039}$	.834$_{\pm.019}$●	**.933**$_{\pm.006}$○	.597$_{\pm.018}$●	.534$_{\pm.006}$●	**.695**$_{\pm.017}$○	.615$_{\pm.024}$●	<u>.600</u>$_{\pm.012}$
MLKNN-LIMIC	.523$_{\pm.014}$	.798$_{\pm.022}$●	**.879**$_{\pm.030}$○	.847$_{\pm.021}$	.914$_{\pm.008}$	.628$_{\pm.015}$●	.572$_{\pm.008}$○	.661$_{\pm.018}$	.618$_{\pm.020}$●	.597$_{\pm.010}$
MLKNN-LSMM	.525$_{\pm.014}$	<u>.800</u>$_{\pm.028}$	<u>.869</u>$_{\pm.043}$	<u>.848</u>$_{\pm.021}$	.913$_{\pm.006}$	<u>.673</u>$_{\pm.023}$●	<u>.579</u>$_{\pm.005}$○	<u>.672</u>$_{\pm.022}$	.620$_{\pm.024}$	**.601**$_{\pm.013}$
MLKNN-ILIA (Ours)	**.529**$_{\pm.016}$	**.807**$_{\pm.014}$	.858$_{\pm.037}$	**.850**$_{\pm.018}$	.918$_{\pm.014}$	**.686**$_{\pm.015}$	.554$_{\pm.008}$	.668$_{\pm.017}$	<u>.625</u>$_{\pm.018}$	.596$_{\pm.010}$

Table B.1: Predictive performance (mean$_{\pm\text{std}}$) of $\mathcal{A} \in \{$BRKNN, MLKNN$\}$ coupled with ILIA and state-of-the-art multi-label metric learning approaches in terms of *Ranking Loss*, *Coverage*, *Macro-F1*, and *Macro-averaging AUC*. ↑ (↓) indicates the larger (smaller) the value, the better the performance. The best and second best results are highlighted in **boldface** and <u>underline</u>, respectively. In addition, ●/○ indicates whether $\mathcal{A}$-ILIA achieves significantly superior/inferior to other compared approaches on each dataset in terms of different evaluation metrics (pairwise t-test at 5% significance level).

C Additional Comparison

C.1 Compared Methods

To underscore the significance of learning similarity metrics for multi-label examples, we compare ILIA-enhanced BRKNN and MLKNN against four well-established metric-free multi-label learning approaches that consider different orders of label correlations:

- LIFT (Zhang and Wu 2014): A first-order multi-label learning approach that exploits label-specific features for each class label [suggested configuration: $r = 0.1$].
- RELIAB (Zhang et al. 2021): A high-order multi-label learning approach that induces maximum entropy model to align the outputs with the enriched label information [suggested configuration: mode = *global*, $\tau \in \{0.1, 0.2, \dots, 0.5\}$, and $\lambda \in \{10^{-3}, 10^{-2}, \dots, 10\}$].
- WRAP (Yu and Zhang 2021): A second-order multi-label learning approach that conducts label-specific feature generation and model induction in a wrapped learning framework [suggested configuration: $\lambda_1, \lambda_2 \in \{0, 1, \dots, 10\}$, $\lambda_3 = 10$, and $\alpha = 0.9$].
- HOMI (Si et al. 2023):A high-order multi-label learning approach that explicitly formalizes label correlations and maintains the high-rank property of the label matrix during model training [suggested configuration: $\beta, \gamma, \in \{10^{-2}, 10^{-1}, \dots, 10^{2}\}$, $\lambda = 1$, $s = 10$].

C.2 Empirical Results

Table C.1 reports detailed empirical results on ten benchmark datasets in terms of six evaluation metrics. The results clearly demonstrate that although the performance of BRKNN and MLKNN are inferior to that of second-order and high-order multi-label learning methods, the ILIA-enhanced versions have the potential to approach or even surpass state-of-the-art multi-label learning methods.This outcome not only reaffirms the superiority of ILIA in characterizing the similarity of multi-label examples, but also emphasizes the significance of learning similarity metrics for multi-label examples.

D Sensitivity Analysis

We perform sensitivity analyses on kernel function κ and trade-off parameters μ, η, and γ, which are not included in the main body due to page limit. When analyzing the sensitivity of one parameter, the other parameters are fixed as the same in the 'Configuration' paragraph of the main body.

D.1 Impact of the Trade-off Parameter μ

Figure D.1 illustrates how the performance of $\mathcal{A}$-ILIA fluctuates with different values of μ, i.e., the trade-off parameters mentioned in Eq.(9) of the main body. (Datasets: emotions, image; Evaluation metrics: *Hamming loss*, *Average precision*). It is shown that the performance of $\mathcal{A}$-ILIA is relatively flat, while it stands out when $\mu = 10^{-3}$. Therefore, μ is fixed to 10^{-3} as a default parameter in this paper.

D.2 Impact of the Trade-off Parameter η

Figure D.2 illustrates how the performance of $\mathcal{A}$-ILIA fluctuates with different values of η, i.e., the trade-off parameters mentioned in Eq.(15) of the main body. (Datasets: emotions, image; Evaluation metrics: *Hamming loss*, *Average precision*). It is shown that the performance of $\mathcal{A}$-ILIA is poor when eta is within the range of $\{10^0, 10^1, \dots, 10^4\}$, but it is better when η is between 10^{-4} and 10^{-2}. Therefore, we choose $\eta = 10^{-2}$ as a default parameter in this paper.

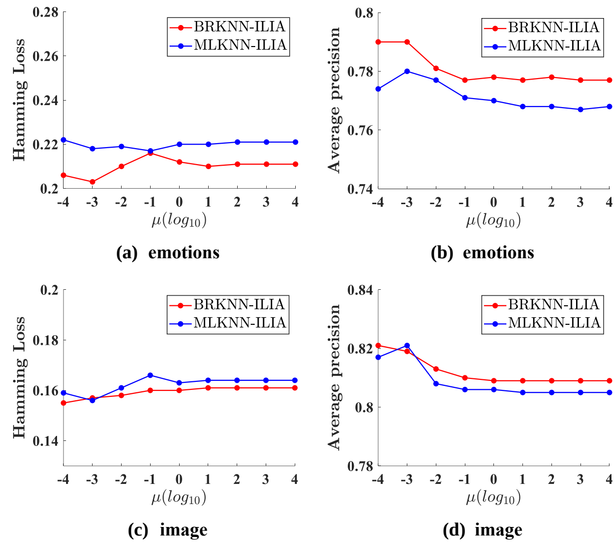

Figure D.1: Performance of $\mathcal{A}$-ILIA changes as the trade-off parameters μ varies in the range of $\{10^{-4}, 10^{-3}, \dots, 10^{4}\}$.

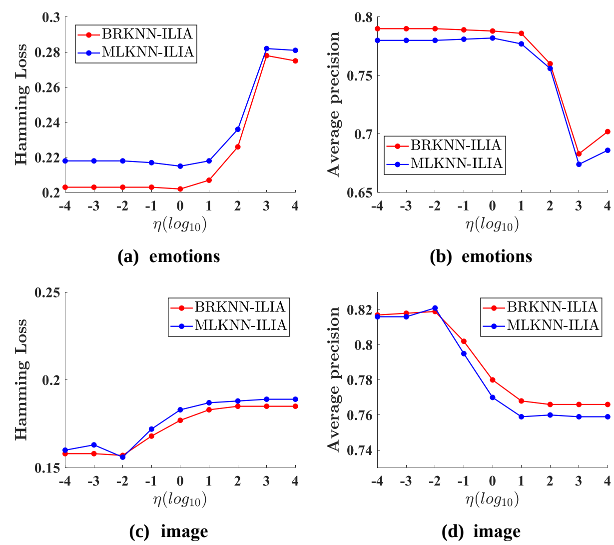

Figure D.2: Performance of $\mathcal{A}$-ILIA changes as the trade-off parameters η varies in the range of $\{10^{-4}, 10^{-3}, \dots, 10^{4}\}$.

D.3 Impact of the Trade-off Parameter γ

Figure D.3 illustrates how the performance of $\mathcal{A}$-ILIA fluctuates with different values of γ, i.e., the trade-off parameters mentioned in Eq.(16) of the main body. (Datasets: emotions, image; Evaluation metrics: *Hamming loss*, *Average precision*). It is shown that the performance of $\mathcal{A}$-ILIA is relatively

Evaluation Metrics	BRKNN-ILIA **against** BRKNN	BRKNN-LM	BRKNN-LJE	BRKNN-COMMU	BRKNN-LIMIC	BRKNN-LSMM
Hamming Loss	6/4/0	6/4/0	9/1/0	6/4/0	4/6/0	2/8/0
Ranking Loss	8/1/1	8/2/0	10/0/0	7/2/1	8/2/0	7/3/0
Coverage	9/1/0	8/1/1	9/1/0	7/2/1	5/4/1	5/4/1
Average precision	10/0/0	8/1/1	10/0/0	10/0/0	5/4/1	6/3/1
Macro-F1	9/1/0	9/1/0	10/0/0	6/4/0	7/3/0	4/6/0
Macro-averaging AUC	7/3/0	7/2/1	10/0/0	5/3/2	5/4/1	1/8/1
In Total	**49/10/1**	**46/11/3**	**58/2/0**	**41/15/4**	**34/23/3**	**25/32/3**

Evaluation Metrics	MLKNN-ILIA **against** MLKNN	MLKNN-LM	MLKNN-LJE	MLKNN-COMMU	MLKNN-LIMIC	MLKNN-LSMM
Hamming Loss	5/5/0	3/6/1	8/2/0	4/5/1	3/6/1	2/8/0
Ranking Loss	4/3/3	3/5/2	7/3/0	4/4/2	1/7/2	1/6/3
Coverage	4/4/2	3/5/2	7/3/0	4/4/2	1/7/2	1/6/3
Average precision	7/2/1	4/5/1	8/2/0	8/1/1	3/5/2	3/6/1
Macro-F1	10/0/0	10/0/2	10/0/0	10/0/0	5/5/0	5/450
Macro-averaging AUC	4/4/2	4/5/1	9/1/0	5/3/2	3/5/2	1/8/1
In Total	**34/18/8**	**27/26/7**	**48/12/0**	**35/19/6**	**16/36/8**	**13/40/7**

Table B.2: Win/tie/loss counts (pairwise t-test at 5% significant level) for ILIA against other state-of-the-art multi-label metric learning algorithms coupled with $\mathcal{A}$ ($\mathcal{A} \in \{\text{BRKNN}, \text{MLKNN}\}$).

stable, but it tends to be better in most cases when $\gamma = 10^{-2}$. Therefore, γ is set to 10^{-2} in this paper.

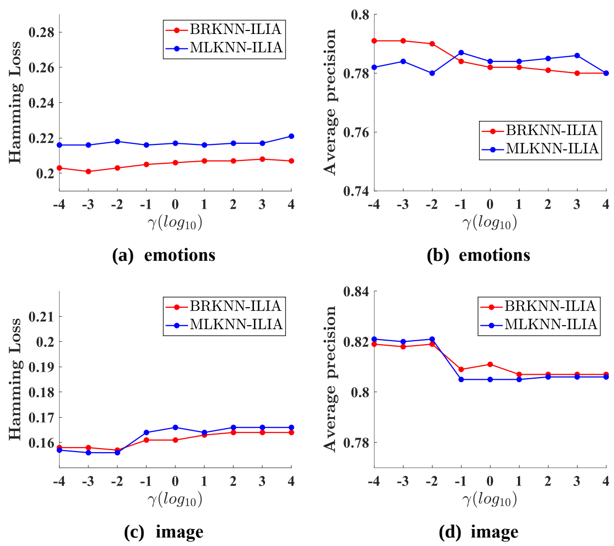

Figure D.3: Performance of $\mathcal{A}$-ILIA changes as the trade-off parameters γ varies in the range of $\{10^{-4}, 10^{-3}, \ldots, 10^{4}\}$.

D.4 Impact of the Kernel Function κ

We investigate the impact of the choice of kernel function κ on the predictive performance of $\mathcal{A}$-ILIA. In this paper, we study four commonly used kernel functions with different parameter settings, which is shown as follows:

- Linear Kernel:

$$\kappa(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{x} \cdot \boldsymbol{x}'. \tag{1}$$

- RBF Kernel:

$$\kappa(\boldsymbol{x}, \boldsymbol{x}') = \exp(-a \left\| \boldsymbol{x} - \boldsymbol{x}' \right\|_2^2). \tag{2}$$

- Laplacian Kernel:

$$\kappa(\boldsymbol{x}, \boldsymbol{x}') = \exp(-a \left\| \boldsymbol{x} - \boldsymbol{x}' \right\|_1). \tag{3}$$

- Polynomial Kernel:

$$\kappa(\boldsymbol{x}, \boldsymbol{x}') = (a\boldsymbol{x} \cdot \boldsymbol{x}' + b)^c. \tag{4}$$

Here, a, b, and c is the hyper-parameters to be determined.

Table D.1 illustrates how the performance of $\mathcal{A}$-ILIA fluctuates with different choices of κ, i.e., the kernel function mentioned in Eq.(15) of the main body. (Datasets: emotions, image, scene, health, and education; Evaluation metrics: *Average precision*). ↑ (↓) indicates the larger (smaller) the value, the better the performance. The best and second best results are highlighted in **boldface** and underline respectively. We observe that there is no universally optimal setting for all cases. However, the Polynomial kernel (a=0.1, b=10, c=2) consistently performs well in most cases and produces satisfactory results. To avoid severe training burden, we adopt this setting throughout the paper.

E Convergence Analysis

We study the convergence property of the proposed ILIA approach. Specifically, we record $\mathbf{M}$ after each iteration of alternating optimization and then compute the Frobenius norm of the difference between two adjacent $\mathbf{M}$. Figure E.1 illustrates the convergence curve of ILIA on ten benchmark multi-label datasets, where the vertical axis represents the resulting Frobenius norm and the horizontal axis represents the number of iterations. We can observe that the value of the Frobenius norm decreases quickly to near zero in 30 iterations (some datasets may initially increase due to the reason that $\mathbf{M}$ is initialized with an identity matrix), which demonstrates the fast convergence property of ILIA.

Compared Algorithms	Datasets									
	CAL500	emotions	medical	image	scene	arts	corel5k	education	health	entertainment
Hamming Loss ↓										
LIFT	.144$_{\pm .004}$	.249$_{\pm .015}$	.012$_{\pm .002}$	.159$_{\pm .014}$	.079$_{\pm .007}$	**.053**$_{\pm .001}$	**.009**$_{\pm .001}$	.042$_{\pm .001}$	.047$_{\pm .002}$	.074$_{\pm .002}$
RELIAB	.148$_{\pm .005}$	.243$_{\pm .009}$	.019$_{\pm .003}$	.178$_{\pm .015}$	.110$_{\pm .015}$	.062$_{\pm .004}$	.014$_{\pm .001}$	**.039**$_{\pm .002}$	.046$_{\pm .001}$	.066$_{\pm .002}$
WRAP	.143$_{\pm .004}$	.220$_{\pm .027}$	**.010**$_{\pm .001}$	.186$_{\pm .009}$	.114$_{\pm .007}$	.054$_{\pm .002}$	**.009**$_{\pm .000}$	.044$_{\pm .001}$	.054$_{\pm .001}$	.068$_{\pm .005}$
HOMI	**.137**$_{\pm .005}$	.226$_{\pm .022}$	.013$_{\pm .002}$	.162$_{\pm .017}$	.085$_{\pm .006}$	.063$_{\pm .002}$	**.009**$_{\pm .000}$	**.037**$_{\pm .001}$	.045$_{\pm .002}$	.065$_{\pm .002}$
BRKNN-ILIA (Ours)	.142$_{\pm .004}$	**.203**$_{\pm .018}$	.011$_{\pm .002}$	.157$_{\pm .015}$	**.078**$_{\pm .007}$	**.053**$_{\pm .001}$	.010$_{\pm .000}$	**.037**$_{\pm .001}$	**.044**$_{\pm .002}$	**.064**$_{\pm .001}$
MLKNN-ILIA (Ours)	.142$_{\pm .004}$	.218$_{\pm .014}$	.011$_{\pm .002}$	**.156**$_{\pm .014}$	.085$_{\pm .008}$	**.053**$_{\pm .001}$	.010$_{\pm .000}$	**.037**$_{\pm .002}$	.045$_{\pm .002}$	**.064**$_{\pm .001}$
Ranking Loss ↓										
LIFT	.190$_{\pm .006}$	.254$_{\pm .046}$	.028$_{\pm .011}$	.177$_{\pm .018}$	**.074**$_{\pm .009}$	.124$_{\pm .006}$	**.122**$_{\pm .004}$	.084$_{\pm .003}$	.081$_{\pm .006}$	.155$_{\pm .029}$
RELIAB	.434$_{\pm .028}$	.256$_{\pm .040}$	**.014**$_{\pm .009}$	.192$_{\pm .025}$	.089$_{\pm .008}$	.145$_{\pm .006}$	.279$_{\pm .008}$	.097$_{\pm .008}$	.122$_{\pm .032}$	.194$_{\pm .022}$
WRAP	.279$_{\pm .013}$	.229$_{\pm .029}$	.060$_{\pm .013}$	.171$_{\pm .023}$	.089$_{\pm .012}$	.128$_{\pm .009}$	.250$_{\pm .008}$	.086$_{\pm .006}$	.075$_{\pm .003}$	.140$_{\pm .009}$
HOMI	**.177**$_{\pm .004}$	.237$_{\pm .032}$	.041$_{\pm .012}$	.196$_{\pm .021}$	.085$_{\pm .011}$	.186$_{\pm .010}$	.210$_{\pm .011}$	.132$_{\pm .004}$	.077$_{\pm .035}$	.139$_{\pm .004}$
BRKNN-ILIA (Ours)	.221$_{\pm .009}$	**.183**$_{\pm .018}$	.088$_{\pm .026}$	.160$_{\pm .021}$	.106$_{\pm .019}$	.175$_{\pm .010}$	.480$_{\pm .013}$	.142$_{\pm .010}$	.132$_{\pm .009}$	.206$_{\pm .005}$
MLKNN-ILIA (Ours)	.189$_{\pm .005}$	.188$_{\pm .017}$	.048$_{\pm .014}$	**.154**$_{\pm .018}$	.083$_{\pm .010}$	**.120**$_{\pm .007}$	.151$_{\pm .005}$	**.077**$_{\pm .004}$	**.066**$_{\pm .004}$	**.131**$_{\pm .003}$
Coverage ↓										
LIFT	.758$_{\pm .015}$	.360$_{\pm .027}$	**.052**$_{\pm .015}$	.181$_{\pm .022}$	**.075**$_{\pm .010}$	**.173**$_{\pm .006}$	.287$_{\pm .011}$	.112$_{\pm .005}$	.124$_{\pm .005}$	.178$_{\pm .005}$
RELIAB	.763$_{\pm .014}$	.317$_{\pm .025}$	.063$_{\pm .013}$	.199$_{\pm .027}$	.091$_{\pm .009}$	.184$_{\pm .011}$	**.264**$_{\pm .009}$	.130$_{\pm .009}$	.184$_{\pm .045}$	.173$_{\pm .008}$
WRAP	.751$_{\pm .023}$	**.299**$_{\pm .038}$	.059$_{\pm .009}$	.192$_{\pm .013}$	.086$_{\pm .009}$	.184$_{\pm .015}$	.463$_{\pm .019}$	.128$_{\pm .008}$	.120$_{\pm .003}$	.237$_{\pm .026}$
HOMI	.792$_{\pm .029}$	.314$_{\pm .025}$	.055$_{\pm .016}$	.188$_{\pm .020}$	.085$_{\pm .008}$	.234$_{\pm .013}$	.321$_{\pm .009}$	.140$_{\pm .004}$	.123$_{\pm .006}$	.175$_{\pm .004}$
BRKNN-ILIA (Ours)	.801$_{\pm .014}$	.310$_{\pm .026}$	.075$_{\pm .019}$	**.174**$_{\pm .021}$	.089$_{\pm .016}$	.216$_{\pm .007}$	.400$_{\pm .012}$	.135$_{\pm .006}$	.137$_{\pm .007}$	.190$_{\pm .004}$
MLKNN-ILIA (Ours)	.750$_{\pm .016}$	.314$_{\pm .021}$	.069$_{\pm .015}$	.176$_{\pm .016}$	.083$_{\pm .009}$	.178$_{\pm .007}$	.324$_{\pm .013}$	**.105**$_{\pm .004}$	**.107**$_{\pm .005}$	**.165**$_{\pm .003}$
Average precision ↑										
LIFT	.501$_{\pm .007}$	.735$_{\pm .035}$	.856$_{\pm .028}$	.817$_{\pm .020}$	**.876**$_{\pm .014}$	.613$_{\pm .014}$	.270$_{\pm .006}$	.629$_{\pm .011}$	.632$_{\pm .023}$	.533$_{\pm .019}$
RELIAB	.498$_{\pm .008}$	.788$_{\pm .046}$	.837$_{\pm .018}$	.781$_{\pm .014}$	.860$_{\pm .013}$	.605$_{\pm .012}$	.251$_{\pm .013}$	.633$_{\pm .014}$	.649$_{\pm .010}$	.548$_{\pm .011}$
WRAP	**.509**$_{\pm .012}$	.782$_{\pm .040}$	.609$_{\pm .041}$	.794$_{\pm .025}$	.870$_{\pm .014}$	.613$_{\pm .017}$	.288$_{\pm .005}$	.628$_{\pm .016}$	.648$_{\pm .010}$	.542$_{\pm .014}$
HOMI	.475$_{\pm .006}$	.770$_{\pm .035}$	.859$_{\pm .023}$	.795$_{\pm .024}$	.866$_{\pm .013}$	.617$_{\pm .015}$	**.293**$_{\pm .011}$	**.634**$_{\pm .011}$	.664$_{\pm .013}$	.551$_{\pm .016}$
BRKNN-ILIA (Ours)	.482$_{\pm .012}$	**.790**$_{\pm .022}$	.856$_{\pm .022}$	.819$_{\pm .021}$	.868$_{\pm .016}$	.616$_{\pm .012}$	.239$_{\pm .010}$	.631$_{\pm .012}$	.661$_{\pm .013}$	.554$_{\pm .008}$
MLKNN-ILIA (Ours)	.485$_{\pm .009}$	.780$_{\pm .026}$	**.862**$_{\pm .020}$	**.821**$_{\pm .017}$	.871$_{\pm .011}$	**.620**$_{\pm .014}$	.248$_{\pm .008}$	**.634**$_{\pm .012}$	**.673**$_{\pm .015}$	**.573**$_{\pm .008}$
Macro-F1 ↑										
LIFT	.049$_{\pm .003}$	.418$_{\pm .043}$	.265$_{\pm .034}$	.606$_{\pm .036}$	.759$_{\pm .016}$	.123$_{\pm .009}$	.030$_{\pm .005}$	.156$_{\pm .011}$	.084$_{\pm .018}$	.025$_{\pm .007}$
RELIAB	.063$_{\pm .004}$	.547$_{\pm .053}$	.374$_{\pm .042}$	.555$_{\pm .016}$	.595$_{\pm .034}$	.160$_{\pm .018}$	.060$_{\pm .006}$	.176$_{\pm .004}$	.093$_{\pm .010}$	.040$_{\pm .006}$
WRAP	.044$_{\pm .004}$	.626$_{\pm .055}$	.205$_{\pm .032}$	.497$_{\pm .041}$	.635$_{\pm .034}$	.120$_{\pm .010}$	.015$_{\pm .002}$	.194$_{\pm .005}$	.046$_{\pm .007}$	.099$_{\pm .039}$
HOMI	.050$_{\pm .003}$	.602$_{\pm .040}$	.488$_{\pm .070}$	.663$_{\pm .022}$	.766$_{\pm .018}$	**.271**$_{\pm .023}$	**.066**$_{\pm .005}$	.183$_{\pm .016}$	.146$_{\pm .026}$	.108$_{\pm .026}$
BRKNN-ILIA (Ours)	**.102**$_{\pm .009}$	**.637**$_{\pm .035}$	.476$_{\pm .055}$	.663$_{\pm .025}$	**.787**$_{\pm .017}$	.203$_{\pm .017}$	.030$_{\pm .004}$	.171$_{\pm .021}$	**.163**$_{\pm .016}$	**.137**$_{\pm .026}$
MLKNN-ILIA (Ours)	.093$_{\pm .010}$	.608$_{\pm .028}$	**.506**$_{\pm .057}$	**.665**$_{\pm .023}$	.769$_{\pm .020}$	.212$_{\pm .018}$	.036$_{\pm .006}$	.186$_{\pm .023}$	.161$_{\pm .020}$	.122$_{\pm .024}$
Macro-averaging AUC ↑										
LIFT	.535$_{\pm .017}$	.769$_{\pm .030}$	**.926**$_{\pm .050}$	.856$_{\pm .022}$	.930$_{\pm .007}$	.673$_{\pm .032}$	.647$_{\pm .010}$	.687$_{\pm .022}$	.633$_{\pm .043}$	.580$_{\pm .035}$
RELIAB	.492$_{\pm .025}$	**.823**$_{\pm .044}$	.864$_{\pm .045}$	.832$_{\pm .024}$	.927$_{\pm .006}$	.704$_{\pm .023}$	.672$_{\pm .011}$	.716$_{\pm .020}$	.638$_{\pm .041}$	.588$_{\pm .030}$
WRAP	.477$_{\pm .011}$	.810$_{\pm .037}$	.913$_{\pm .061}$	.830$_{\pm .022}$	.918$_{\pm .009}$	**.717**$_{\pm .017}$	.464$_{\pm .015}$	.610$_{\pm .038}$	.533$_{\pm .019}$	.572$_{\pm .043}$
HOMI	.526$_{\pm .016}$	.813$_{\pm .032}$	.915$_{\pm .044}$	.842$_{\pm .027}$	.925$_{\pm .005}$	.707$_{\pm .036}$	**.695**$_{\pm .007}$	**.721**$_{\pm .017}$	.622$_{\pm .036}$	.611$_{\pm .023}$
BRKNN-ILIA (Ours)	**.544**$_{\pm .014}$	**.823**$_{\pm .022}$	.878$_{\pm .031}$	**.861**$_{\pm .020}$	**.930**$_{\pm .013}$	.708$_{\pm .021}$	.571$_{\pm .010}$	.702$_{\pm .021}$	**.642**$_{\pm .018}$	**.612**$_{\pm .011}$
MLKNN-ILIA (Ours)	.529$_{\pm .016}$	.807$_{\pm .014}$	.858$_{\pm .037}$	.850$_{\pm .018}$	.918$_{\pm .014}$	.686$_{\pm .015}$	.554$_{\pm .008}$	.668$_{\pm .017}$	.625$_{\pm .018}$	.596$_{\pm .010}$

Table C.1: Predictive performance (mean$_{\pm \text{std}}$) of $\mathcal{A}$-ILIA and four well-established metric-free multi-label learning approaches in terms of six evaluation metrics. ↑ (↓) indicates the larger (smaller) the value, the better the performance. The best and second best results are highlighted in **boldface** and underline respectively.

Kernel Function	Parameter Setting	Datasets				
		emotions	image	scene	health	education
		Average precision ↑				
Linear	N/A	$\mathbf{.791}_{\pm.030}/\mathbf{.792}_{\pm.036}$	$.804_{\pm.016}/.795_{\pm.014}$	$\underline{.863}_{\pm.014}/.861_{\pm.016}$	$.629_{\pm.011}/.651_{\pm.012}$	$.625_{\pm.015}/.630_{\pm.014}$
RBF	a=0.1	$.695_{\pm.027}/.676_{\pm.029}$	$.806_{\pm.021}/.798_{\pm.023}$	$.851_{\pm.015}/.855_{\pm.020}$	$.577_{\pm.011}/.600_{\pm.010}$	$.614_{\pm.013}/.614_{\pm.016}$
	a=1	$.554_{\pm.037}/.545_{\pm.032}$	$.780_{\pm.031}/.784_{\pm.029}$	$.567_{\pm.031}/.564_{\pm.029}$	$.525_{\pm.024}/.612_{\pm.012}$	$.625_{\pm.015}/.632_{\pm.015}$
	a=10	$.553_{\pm.037}/.542_{\pm.032}$	$.525_{\pm.018}/.521_{\pm.013}$	$.441_{\pm.033}/.437_{\pm.027}$	$.572_{\pm.016}/.600_{\pm.012}$	$.403_{\pm.021}/.428_{\pm.015}$
Lapla	a=0.1	$.767_{\pm.035}/.759_{\pm.026}$	$\underline{.816}_{\pm.022}/\underline{.817}_{\pm.022}$	$.862_{\pm..013}/.865_{\pm.014}$	$.612_{\pm.009}/.628_{\pm.011}$	$.623_{\pm.015}/.627_{\pm.016}$
	a=1	$.553_{\pm.037}/.542_{\pm.032}$	$.520_{\pm.018}/.515_{\pm.013}$	$.441_{\pm.033}/.437_{\pm.027}$	$.524_{\pm.017}/.614_{\pm.010}$	$.383_{\pm.014}/.464_{\pm.012}$
	a=10	$.553_{\pm.037}/.542_{\pm.032}$	$.515_{\pm.016}/.509_{\pm.016}$	$.441_{\pm.033}/.430_{\pm.026}$	$.572_{\pm.016}/.600_{\pm.012}$	$.403_{\pm.014}/.413_{\pm.016}$
Poly	a=0.1 b=1 c=1	$.789_{\pm.037}/.782_{\pm.032}$	$.794_{\pm.020}/.792_{\pm.029}$	$\mathbf{.868}_{\pm.015}/\underline{.866}_{\pm.017}$	$\underline{.645}_{\pm.017}/\underline{.660}_{\pm.017}$	$.611_{\pm.014}/.611_{\pm.015}$
	a=0.1 b=10 c=2	$\underline{.790}_{\pm.022}/.780_{\pm.026}$	$\mathbf{.819}_{\pm.021}/\mathbf{.821}_{\pm.017}$	$\mathbf{.868}_{\pm.016}/\mathbf{.871}_{\pm.011}$	$\mathbf{.661}_{\pm.013}/\mathbf{.673}_{\pm.015}$	$\mathbf{.631}_{\pm.012}/\underline{.634}_{\pm.012}$
	a=1 b=10 c=2	$.779_{\pm.035}/.776_{\pm.028}$	$.735_{\pm.023}/.761_{\pm.023}$	$812_{\pm.019}/.826_{\pm.021}$	$.636_{\pm.011}/.651_{\pm.012}$	$.585_{\pm.017}/.620_{\pm.014}$
	a=10 b=1 c=1	$\mathbf{.791}_{\pm.034}/\underline{.785}_{\pm.028}$	$.783_{\pm.019}/.779_{\pm.017}$	$.852_{\pm.020}/.849_{\pm.019}$	$.630_{\pm.011}/.655_{\pm.010}$	$\underline{.630}_{\pm.014}/\mathbf{.638}_{\pm.013}$
	a=10 b=10 c=2	$.762_{\pm.026}/.762_{\pm.027}$	$.732_{\pm.021}/.757_{\pm.024}$	$.818_{\pm.023}/.831_{\pm.020}$	$.625_{\pm.009}/.643_{\pm.012}$	$.551_{\pm.020}/.562_{\pm.021}$

Table D.1: Predictive performance (mean$_{\pm\text{std}}$) of $\mathcal{A}$-ILIA under different kernel function κ choices in terms of *Average precision*. In A/B, A represents the result of BRKNN-ILIA, and B represents the result of MLKNN-ILIA. ↑ (↓) indicates the larger (smaller) the value, the better the performance. The best and second best results are highlighted in **boldface** and underline, respectively.

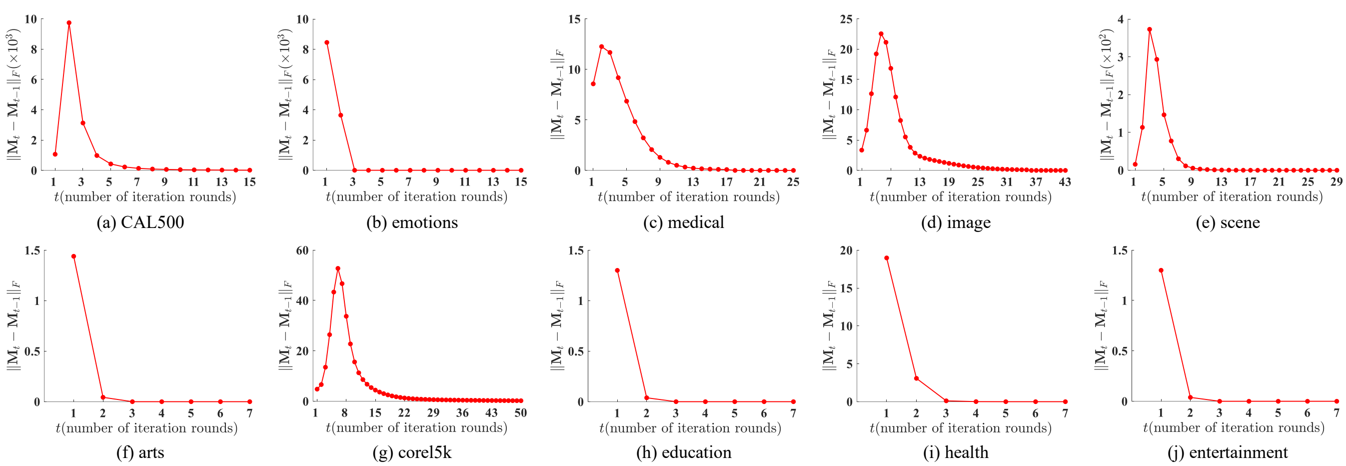

Figure E.1: Convergence curves of ILIA on ten benchmark multi-label datasets.

References

Gouk, H.; Pfahringer, B.; and Cree, M. 2016. Learning distance metrics for multi-label classification. In *Proceedings of the 8th Asian Conference on Machine Learning*, 318–333. Hamilton, New Zealand.

Liu, W.; and Tsang, I. W. 2015. Large margin metric learning for multi-label prediction. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, volume 29, 2800–2806. Austin, Tex.

Mao, J.-X.; Hang, J.-Y.; and Zhang, M.-L. 2024. Learning label-specific multiple local metrics for multi-label classification. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence*, 4742–4750. Jeju, South Korea.

Mao, J.-X.; Wang, W.; and Zhang, M.-L. 2023. Label specific multi-semantics metric learning for multi-label classification: Global consideration helps. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence*, 4055–4063. Macao, SAR, China.

Si, C.; Jia, Y.; Wang, R.; Zhang, M.-L.; Feng, Y.; and Qu, C. 2023. Multi-label classification with high-rank and high-order label correlations. *IEEE Transactions on Knowledge and Data Engineering*, 36(8): 4076–4088.

Sun, Y.-P.; and Zhang, M.-L. 2021. Compositional metric learning for multi-label classification. *Frontiers of Computer Science*, 15(5): 1–12.

Yu, Z.-B.; and Zhang, M.-L. 2021. Multi-label classification with label-specific feature generation: A wrapped approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9): 5199–5210.

Zhang, M.-L.; and Wu, L. 2014. LIFT: Multi-label learning with label-specific features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(1): 107–120.

Zhang, M.-L.; Zhang, Q.-W.; Fang, J.-P.; Li, Y.-K.; and Geng, X. 2021. Leveraging implicit relative labeling-importance information for effective multi-Label Learning. *IEEE Transactions on Knowledge and Data Engineering*, 33(5): 2057–2070.