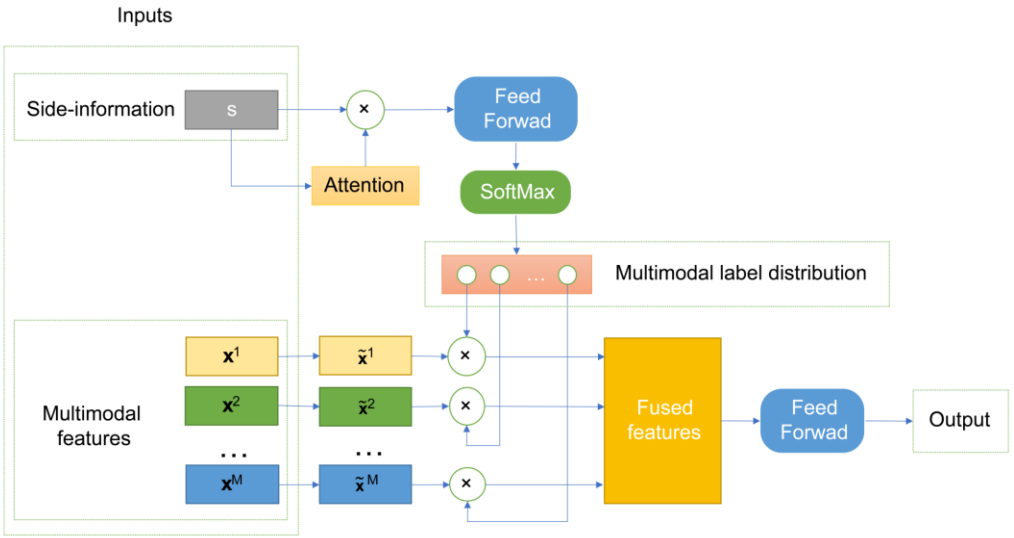
Label Distribution for Multimodal Machine Learning

Yi REN, Ning XU, Miaogen LING, Xin GENG

Frontiers of Computer Science, DOI: 10.1007/s11704-021-0611-6

Problems & Ideas

- Problem: The fusion of multiple modalities is always the key challenge of Multimodal machine learning (MML).
- Ideas: Considering the side-information, which reflects the situation and influences the fusion of multi-modalities.
- A novel framework named Multimodal Label Distribution Learning (MLDL) is proposed:
  - Recover Multimodal Label Distribution (MLD) by leveraging the side-information
  - fuse the multi-modalities with its guidance to learn an in-depth understanding of the jointly feature representation.



Main Contributions

- **Sentiment Analyses**

Table 3: Sentiment prediction results on MOSEI dataset with different combinations of modalities. (A-audio, V-video, T-text)

Modality	Methods	Average precision	Micro-F1	ROC-AUC	Ranking loss	Coverage
A	EF-LSTM	0.6359	0.2529	0.8341	0.1650	0.1467
	LF-LSTM	0.6302	0.2772	0.8518	0.1579	0.1404
	MARN	0.6376	0.2994	0.8612	0.1465	0.1302
	Graph-MFN	0.6406	0.2815	0.8605	0.1410	0.1253
	RAVEN	0.6283	0.2755	0.8417	0.1582	0.1406
	EF-MLDL	**0.6455**	**0.3035**	**0.8676**	**0.1389**	**0.1235**
V	EF-LSTM	0.6290	0.2367	0.8310	0.1721	0.1530
	LF-LSTM	0.6265	0.2295	0.8395	0.1658	0.1474
	MARN	0.6283	0.2695	0.8487	0.1582	0.1406
	Graph-MFN	0.6406	0.2879	0.8638	0.1421	0.1263
	RAVEN	0.6283	0.2709	0.8417	0.1582	0.1406
	EF-MLDL	**0.6492**	**0.3079**	**0.8694**	**0.1372**	**0.1219**
T	EF-LSTM	0.6345	0.3310	0.8655	0.1467	0.1304
	LF-LSTM	0.6482	0.3150	0.8712	0.1403	0.1247
	MARN	0.6469	0.3198	0.8722	0.1387	0.1233
	Graph-MFN	0.6499	0.2987	0.8689	0.1401	0.1245
	RAVEN	0.6283	0.2697	0.8452	0.1582	0.1406
	EF-MLDL	**0.6597**	**0.3221**	**0.8808**	**0.1313**	**0.1167**
A+V	EF-LSTM	0.6477	0.2988	0.8682	0.1387	0.1233
	LF-LSTM	0.6444	0.2986	0.8684	0.1382	0.1228
	MARN	0.6466	0.2848	0.8671	0.1381	0.1227
	Graph-MFN	0.6460	0.3064	0.8705	0.1374	0.1221
	RAVEN	0.6283	0.2680	0.8417	0.1583	0.1406
	EF-MLDL	**0.6528**	**0.3228**	**0.8713**	**0.1354**	**0.1203**
A+T	EF-LSTM	0.6380	0.2457	0.8570	0.1509	0.1341
	LF-LSTM	0.6357	0.3134	0.8633	0.1461	0.1298
	MARN	0.6283	0.2721	0.8463	0.1583	0.1407
	Graph-MFN	0.6550	0.3014	0.8738	0.1332	0.1184
	RAVEN	0.6347	0.2826	0.8521	0.1529	0.1359
	EF-MLDL	**0.6604**	**0.3385**	**0.8840**	**0.1291**	**0.1148**
V+T	EF-LSTM	0.6479	0.3264	0.8735	0.1365	0.1213
	LF-LSTM	0.6284	0.3087	0.8503	0.1578	0.1403
	MARN	0.6370	0.2868	0.8548	0.1510	0.1342
	Graph-MFN	0.6548	0.3036	0.8737	0.1351	0.1201
	RAVEN	0.6487	0.2961	0.8690	0.1388	0.1234
	EF-MLDL	**0.6603**	**0.3353**	**0.8821**	**0.1297**	**0.1153**
A+V+T	EF-LSTM	0.6533	0.3341	0.8749	0.1367	0.1216
	LF-LSTM	0.6454	0.3470	0.8723	0.1377	0.1224
	MARN	0.6378	0.2512	0.8543	0.1488	0.1322
	Graph-MFN	0.6487	0.3500	0.8747	0.1375	0.1222
	RAVEN	0.6339	0.2765	0.8509	0.1537	0.1366
	EF-MLDL	**0.6753**	**0.3587**	**0.8932**	**0.1192**	**0.1049**

- **Disease prediction**

Table 4: Disease prediction results on MIMIC-WHOLE dataset.

Methods	Average precision	Micro-F1	ROC-AUC	Ranking loss	Coverage
Linear regression	0.5194	0.4543	0.6930	0.6324	0.5411
Bayes classifier	0.5181	0.4544	0.6860	0.6323	0.5413
Decision tree	0.5688	0.4665	0.6105	0.5275	0.5682
LSTM	0.7473	0.5308	0.8662	0.1356	0.3540
Text-TF-IDF-CNN	0.7418	0.4909	0.8643	0.1378	0.3535
LF-MLDL	**0.7627**	**0.5412**	**0.8778**	**0.1261**	**0.3414**

Table 5: Disease prediction results on MIMIC-SEQ dataset.

Methods	Average precision	Micro-F1	ROC-AUC	Ranking loss	Coverage
Linear regression	0.4699	0.5132	0.6715	0.6959	0.6824
Bayes classifier	0.6242	0.5659	0.7063	0.5428	0.6807
Decision tree	0.5411	0.4600	0.6444	0.6772	0.7350
LSTM	0.7195	0.5478	0.8319	0.1639	0.4720
Doctor AI	0.7227	0.4966	0.8257	0.1585	0.4633
GRAM	0.7385	0.5358	0.8266	0.1551	0.4644
Text-TF-IDF-CNN	0.6948	0.4946	0.8159	0.1758	0.4859
LF-MLDL	**0.7401**	**0.5730**	**0.8466**	**0.1485**	**0.4551**

- **Some examples**

(a) (b) (c)