Supplementary Material
Multimodal Learning with Deep Boltzmann Machines

Anonymous Author(s)
Affiliation
Address
email

1 Examples of Text Generated by the DBM model

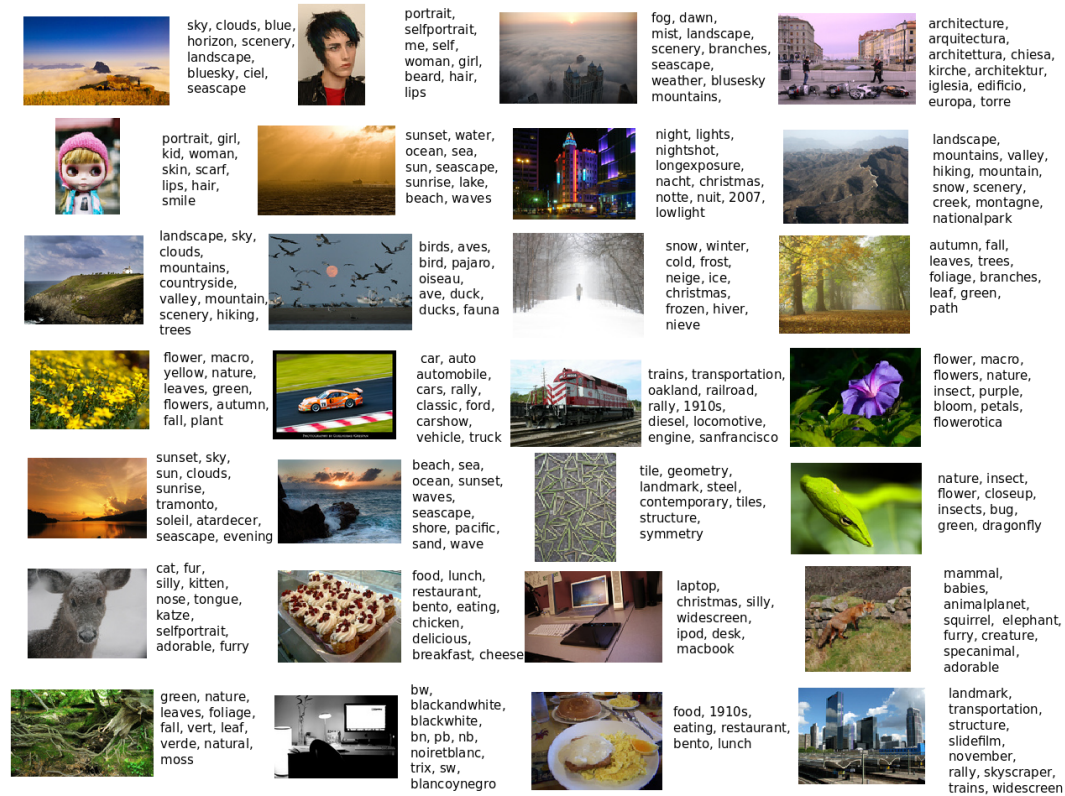Figure 1: Examples of text generated from images by the DBM model

2 Topic-wise classification results

The classification task considered in this paper involves classifying inputs into topics. There are 38 topics and each input may belong to several topics. In order to make our results comparable to previous results, we do one-vs-all classification for each topic separately. Table 1 reports class-wise average precision scores for the models described in the paper with multimodal inputs.

Table 2 reports class-wise average precision scores for the models with unimodal inputs.

Table 1: Comparison of Average Precision scores of different models on the MIR Flickr Dataset with multimodal inputs.

LABELS	ANIMALS	BABY	BABY*	BIRD	BIRD*	CAR	CAR*	CLOUDS	CLOUDS*	DOG
RANDOM	0.129	0.010	0.005	0.030	0.019	0.047	0.015	0.148	0.054	0.027
LDA	0.537	**0.285**	**0.308**	0.426	0.500	0.297	0.389	0.651	0.528	**0.621**
SVM	0.531	0.200	0.165	0.443	0.520	0.339	0.434	0.695	0.434	0.607
DBM-LABELLED	0.511	0.139	0.145	0.190	0.253	0.319	0.368	0.768	0.723	0.351
DBM-UNLABELLED	0.630	0.107	0.069	0.454	0.522	0.318	0.437	0.735	0.679	0.588
AUTOENCODER	0.602	0.156	0.121	**0.461**	0.547	0.366	**0.526**	0.735	0.684	0.605
DBN	0.625	0.115	0.128	0.382	0.459	0.341	0.486	0.772	0.739	0.457
DBM	**0.641**	0.130	0.111	0.426	**0.588**	**0.414**	0.500	**0.799**	**0.748**	0.603

LABELS	DOG*	FEMALE	FEMALE*	FLOWER	FLOWER*	FOOD	INDOOR	LAKE	MALE	MALE*
RANDOM	0.024	0.247	0.159	0.073	0.043	0.040	0.333	0.032	0.243	0.146
LDA	**0.663**	0.494	0.454	0.560	0.623	0.439	0.663	0.258	0.434	0.354
SVM	0.641	0.465	0.451	0.480	0.717	0.308	0.683	0.207	0.413	0.335
DBM-LABELLED	0.385	0.535	0.493	0.604	0.668	0.462	0.759	0.277	0.505	0.424
DBM-UNLABELLED	0.618	0.569	0.527	0.657	0.736	0.483	**0.764**	0.277	0.484	0.408
AUTOENCODER	0.642	0.557	0.542	0.613	0.723	0.558	0.730	0.271	0.491	0.388
DBN	0.515	0.588	0.564	0.643	0.765	0.491	0.754	0.281	**0.522**	**0.436**
DBM	0.662	**0.619**	**0.595**	**0.671**	**0.798**	0.575	0.764	0.314	0.522	0.422

LABELS	NIGHT	NIGHT*	PEOPLE	PEOPLE*	PLANT_LIFE	PORTRAIT	PORTRAIT*	RIVER	RIVER*	SEA
RANDOM	0.108	0.027	0.415	0.314	0.351	0.157	0.153	0.036	0.006	0.053
LDA	0.615	0.420	0.731	0.664	0.703	0.543	0.541	0.317	**0.134**	0.477
SVM	0.588	0.450	0.748	0.565	0.691	0.480	0.558	0.158	0.109	0.529
DBM-LABELLED	0.666	0.505	0.802	0.742	0.794	0.651	0.665	0.274	0.110	0.582
DBM-UNLABELLED	0.684	0.473	0.843	0.783	0.796	0.644	0.640	0.260	0.071	0.611
AUTOENCODER	0.657	0.464	0.791	0.742	0.769	0.655	0.656	0.240	0.016	0.608
DBN	**0.698**	0.567	0.837	0.788	0.823	0.691	0.690	**0.351**	0.103	0.647
DBM	0.687	0.585	0.805	**0.848**	**0.832**	**0.713**	0.695	0.312	0.107	**0.684**

LABELS	SEA*	SKY	STRUCTURES	SUNSET	TRANSPORT	TREE	TREE*	WATER	MEAN
RANDOM	0.009	0.316	0.400	0.085	0.116	0.187	0.027	0.133	0.124
LDA	0.197	0.800	0.709	0.528	0.411	0.515	0.342	0.575	0.492
SVM	0.201	0.823	0.695	0.613	0.369	0.559	0.321	0.527	0.475
DBM-LABELLED	0.260	0.883	0.796	0.659	0.423	0.668	0.492	0.628	0.513
DBM-UNLABELLED	0.259	0.875	0.808	0.662	0.426	0.607	0.518	0.696	0.545
AUTOENCODER	0.357	0.836	0.761	0.625	0.460	0.641	0.513	0.683	0.547
DBN	**0.359**	**0.888**	0.811	**0.679**	0.464	**0.679**	**0.539**	0.703	0.563
DBM	0.358	0.886	**0.819**	0.678	**0.495**	0.667	0.533	**0.719**	**0.587**

Table 2: Comparison of Average Precision scores of different models with unimodal inputs from the MIR Flickr Dataset.

LABELS	ANIMALS	BABY	BABY*	BIRD	BIRD*	CAR	CAR*	CLOUDS	CLOUDS*	DOG
IMAGE-SVM	0.278	0.084	0.088	0.128	0.129	0.179	0.227	0.651	0.511	0.155
IMAGE-DBN	0.348	**0.343**	**0.245**	**0.424**	0.384	**0.486**	0.407	0.601	0.403	0.106
IMAGE-DBM	0.407	0.078	0.094	0.064	0.077	0.234	0.227	0.723	0.648	0.486
DBM-ZEROTEXT	**0.546**	0.096	0.099	0.369	**0.440**	0.284	**0.439**	0.675	0.618	**0.533**
DBM-GENTEXT	0.546	0.135	0.106	0.210	0.243	0.302	0.388	**0.761**	**0.681**	0.334

LABELS	DOG*	FEMALE	FEMALE*	FLOWER	FLOWER*	FOOD	INDOOR	LAKE	MALE	MALE*
IMAGE-SVM	0.156	0.461	0.389	0.469	0.519	0.293	0.605	0.188	0.407	0.294
IMAGE-DBN	0.301	0.351	**0.625**	0.595	0.590	0.364	0.617	0.225	0.470	0.334
IMAGE-DBM	0.525	0.514	0.485	0.432	0.502	0.300	0.688	0.249	0.461	0.394
DBM-ZEROTEXT	**0.574**	0.500	0.476	0.545	0.651	**0.478**	0.661	0.237	0.432	0.377
DBM-GENTEXT	0.354	**0.595**	0.552	**0.600**	0.697	0.478	**0.757**	0.257	0.530	**0.444**

LABELS	NIGHT	NIGHT*	PEOPLE	PEOPLE*	PLANT_LIFE	PORTRAIT	PORTRAIT*	RIVER	RIVER*	SEA
IMAGE-SVM	0.554	0.390	0.631	0.558	0.687	0.493	0.493	0.179	0.102	0.366
IMAGE-DBN	0.337	0.240	0.420	0.389	0.481	0.415	0.609	**0.372**	**0.116**	0.318
IMAGE-DBM	0.563	0.462	0.752	0.693	0.713	0.604	0.598	0.223	0.106	0.549
DBM-ZEROTEXT	0.585	**0.493**	0.714	0.669	0.708	0.584	0.573	0.250	0.091	0.578
DBM-GENTEXT	**0.659**	0.459	**0.830**	**0.777**	**0.800**	**0.662**	**0.658**	0.208	0.085	**0.630**

LABELS	SEA*	SKY	STRUCTURES	SUNSET	TRANSPORT	TREE	TREE*	WATER	MEAN
IMAGE-SVM	0.126	0.775	0.626	0.588	0.298	0.514	0.205	0.448	0.375
IMAGE-DBN	**0.363**	0.622	0.586	0.579	0.352	0.600	0.218	0.457	0.413
IMAGE-DBM	0.282	0.827	0.721	0.562	0.363	0.630	0.424	0.505	0.452
DBM-ZEROTEXT	0.277	0.738	0.685	0.590	0.413	0.571	**0.475**	0.600	0.490
DBM-GENTEXT	0.254	**0.875**	**0.795**	**0.671**	**0.437**	**0.639**	0.465	**0.623**	**0.513**