Learning to smell for wellness

Kehinde Owoeye, ucabowo@ucl.ac.uk

Problem

1. Learning to perceive smell automatically is becoming increasingly important for a variety of olfaction applications.
2. Applications in monitoring quality of food and drinks for healthy living.
3. Machine learning models are trained on sensor data collected to smell a variety of things.
4. However, models trained on one domain rarely perform well in another similar but related domain due to:

- Shift in the distribution of the input features.
- label distribution mismatch.

Contributions

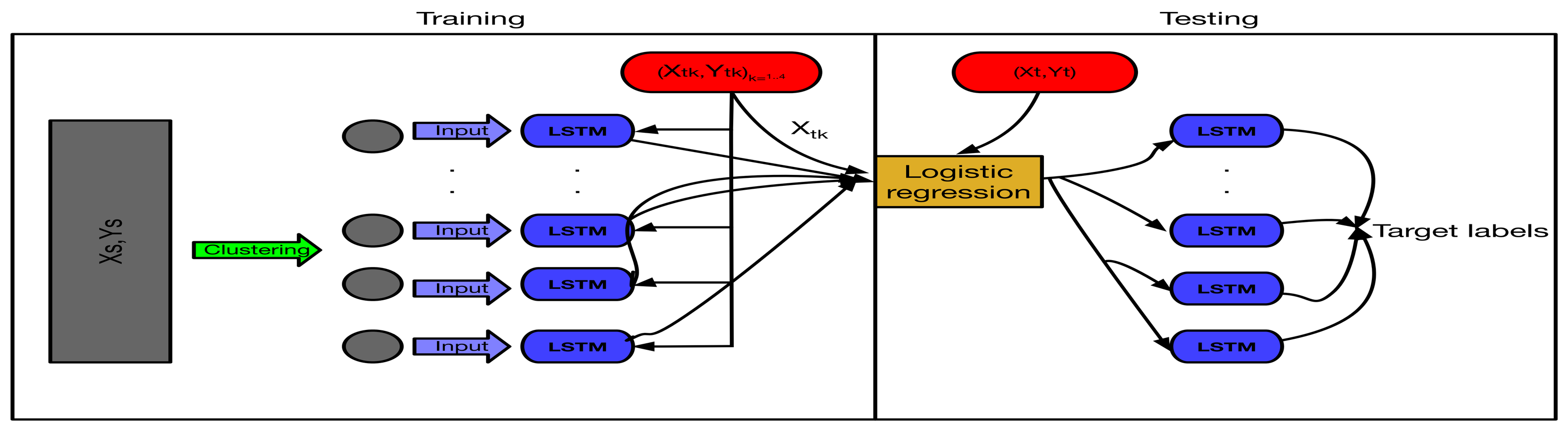
Our main contributions are:

1. We propose a weakly supervised domain adaptation framework where we demonstrate that by building multiple models in a mixture of supervised and unsupervised framework, we can generalise effectively from one domain to another.
2. We evaluate our approach on several datasets of beef cuts and quality collected across different conditions and environments[1].
3. We empirically show via several experiments that our approach perform better variety of baselines.

Training Objectives

$$\underset{\theta_1,\theta_2,\theta_3}{\text{argmin}}\ E(\theta_1,\theta_2,\theta_3) = \frac{1}{N+4n}\sum_{i=1....N+4n} L(y_i, f(X_i;\theta_3)|$$
$$\frac{1}{4n}\sum_{i=1....4n} L(y_i, f(X_i;\theta_2)|$$
$$\frac{1}{N}\sum_{i=1....N} L(y_i, f(X_i;\theta_1))))$$

Proposed Model



Architecture

- 2 RNN's to train data before adding few target data.
- 2 RNN's to train data after adding few target data.
- Logistic regression to handle few target training data.
- Gaussian mixture to cluster data.

Results

Source-Target	LR	AB	SVM	SS[2]	DNN	LSTM	R-DANN	Ours
$1_{1-5}-2$	19.59	69.43	11.26	36.24	5.95	46.02	47.18	**79.85**
1_1-3_{1-12}	46.58	22.00	33.60	14.65	5.30	37.55	44.86	**65.07**
1_2-3_{1-12}	33.97	54.73	25.89	13.59	7.97	41.21	37.61	**64.39**
1_3-3_{1-12}	36.65	54.73	26.37	13.34	3.54	60.62	33.79	**57.73**
1_4-3_{1-12}	39.41	31.37	28.42	13.63	10.53	13.31	32.62	**66.77**
1_5-3_{1-12}	58.99	64.53	30.02	69.32	12.24	23.74	35.43	**69.90**
$2-1_{1-5}$	52.06	59.44	49.03	38.81	11.69	14.08	42.77	**67.14**
$2-3_{1-12}$	75.41	**83.82**	73.51	27.11	7.21	61.60	63.51	78.59
3_1-1_{1-5}	54.14	46.66	45.31	**64.46**	8.34	49.06	38.91	52.19
3_2-1_{1-5}	59.34	57.83	45.09	**59.45**	19.44	44.27	38.91	54.85
3_3-1_{1-5}	58.08	62.22	44.99	**71.04**	19.44	51.30	42.26	52.84
3_4-1_{1-5}	**63.36**	62.22	47.75	47.66	14.27	43.67	39.37	62.98
3_5-1_{1-5}	61.99	**62.22**	43.88	41.04	15.58	50.44	34.33	59.89
3_6-1_{1-5}	54.97	**62.22**	44.56	34.50	14.38	45.37	37.59	60.23
3_7-1_{1-5}	55.84	62.22	48.28	20.72	14.26	39.43	41.79	**65.33**
3_8-1_{1-5}	48.93	62.22	47.35	59.46	22.07	37.93	45.19	**62.39**
3_9-1_{1-5}	53.40	62.22	43.48	18.96	20.17	33.31	47.12	**64.46**
$3_{10}-1_{1-5}$	58.65	62.22	47.80	**68.15**	15.52	47.32	30.66	61.37
$3_{11}-1_{1-5}$	55.69	62.22	43.67	**74.14**	13.31	48.44	35.92	56.77
$3_{12}-1_{1-5}$	74.21	51.11	44.01	**100**	14.69	34.20	36.69	63.46
$3_{1-12}-2$	51.56	88.08	60.28	91.33	20.52	33.60	78.74	**92.18**
Avg	52.99	59.23	42.12	46.55	13.16	40.78	43.59	**66.59**

Classification accuracy all in %. Our approach can be seen to outperform all baselines most of the time across all experiments suggesting that it is useful when there are both label shifts as well as covariate shifts in the input features.

References

[1] Rahman Wijaya, Riyanarto Sarno, and Enny Zulaika., GraphTrack: Electronic nose dataset for beef quality monitoring under an uncontrolled environment? *Mendeley Data '18b*

[2] Li Wei and Eamonn Keogh. Semi-supervised time series classification. In *SIGKDD '06*

Future Direction & Acknowledgment

While we have used the same architecture for all the experiments, we aim to explore different architectures to see if the performance will improve. Our approach can also benefit from meta learning allowing us to train the whole network end to end.

Author is grateful to Intel for their financial support.

Source Code & Data

See references for link to some of the datasets and the rest in the paper. The source code is available at `https://github.com/kehindeowoeye/ltsfw`