

# Label Enhancement for Label Distribution Learning via Prior Knowledge

Yongbiao Gao, Yu Zhang and Xin Geng\*

School of Computer Science and Engineering, Southeast University, Nanjing, China

{gaoyb, zhang\_yu, xgeng}@seu.edu.cn

## Abstract

*Label distribution learning (LDL)* is a novel machine learning paradigm that gives a description degree of each label to an instance. However, most of training datasets only contain simple logical labels rather than label distributions due to the difficulty of obtaining the label distributions directly. We propose to use the prior knowledge to recover the label distributions. The process of recovering the label distributions from the logical labels is called *label enhancement*. In this paper, we formulate the label enhancement as a dynamic decision process. Thus, the label distribution is adjusted by a series of actions conducted by a reinforcement learning agent according to sequential state representations. The target state is defined by the prior knowledge. Experimental results show that the proposed approach outperforms the state-of-the-art methods in both age estimation and image emotion recognition.

## 1 Introduction

Multi-label learning (MLL) [Grigorios and Ioannis, 2006] can deal with the ambiguous case where one instance belongs to more than one classes [Zhou and Zhang, 2006; Zhang and Zhou, 2014; Gibaja and Eva, 2015], but it does not fit many real-world learning problems well where the importance of relevant labels is often different. *Label distribution learning (LDL)* is proposed to cover a certain number of labels, representing the degree to which each label describes the instance [Geng, 2016]. During the past decade, LDL techniques have been widely employed to learn age estimation [Xin *et al.*, 2013; Gao *et al.*, 2018], video parsing [Ling and Geng, 2017], emotion classification [Yang *et al.*, 2017b; Zhou *et al.*, 2016; Yang *et al.*, 2017a], facial landmark detection [Su and Geng, 2019], etc. However, most of the existing datasets only contain single logical labels or multi-labels rather than label distributions. It is difficult to obtain the label distribution directly. A popular method to obtain the label distribution is the vote strategy [Yang *et al.*, 2017b;

Liu *et al.*, 2017]. The vote strategy is not only costly, but also easily influenced by the annotators' individual subjective opinions, age, sex, occupation, social background, etc. This fact encourages researchers to explore new methods to recover label distributions.

*Label enhancement (LE)* [Xu *et al.*, 2018] is the way to recover the label distributions from the logical labels in the training set leveraging the topological information in the feature space and the correlation among the labels. Some label enhancement approaches, like fuzzy clustering [Gayar *et al.*, 2006], label propagation [Li *et al.*, 2016] or manifold learning [Hou *et al.*, 2016] are proposed to recover label distributions. But the pioneer works ignore the importance of the prior knowledge, which is widely existing in many applications. The prior knowledge indicates some natural properties of the label distribution. For example, the people usually predict another person's age in a way such as "around 25 years old", which indicates using not only 25, but also the neighboring age to describe the appearance of the face. Among the ages, 26 and 24 are closer to 25, 27 and 23 are farther away 25. The farther away from the chronological age, the less the importance it is. In addition, psychology theories, like Plutchik's wheel [Plutchik, 2001] and Mikels' wheel [Mikels *et al.*, 2005] can also be used as the prior knowledge. We propose a reinforcement learning based method for label enhancement (*RLLE*) via the prior knowledge. Reinforcement learning is much more focused on goal-directed learning [Sutton and Barto, 2018]. Given a goal, an agent can automatically learn a policy to reach it. Coincidentally, properties implied in the prior knowledge can be defined as the goal. And the target label distribution may not be obtained in one step. It is a more natural way to obtain the label distribution by sequential adjustment. Therefore, in this paper, label enhancement is formulated as a dynamic decision process that a set of actions conducted on the label distribution to adjust the description degrees automatically until it reaches the terminal state, which is defined by the prior knowledge. *RLLE* can automatically learn a policy to recover the label distribution from the single logical label directly based only on the prior knowledge. No additional information is needed anymore.

The proposed model defines instance as the environment. The features of an instance and the current label distribution are defined as the state. The action space includes *{bigger, smaller, constant}*. The multi-dueling Q-learning algorithm

---

\*Corresponding Author

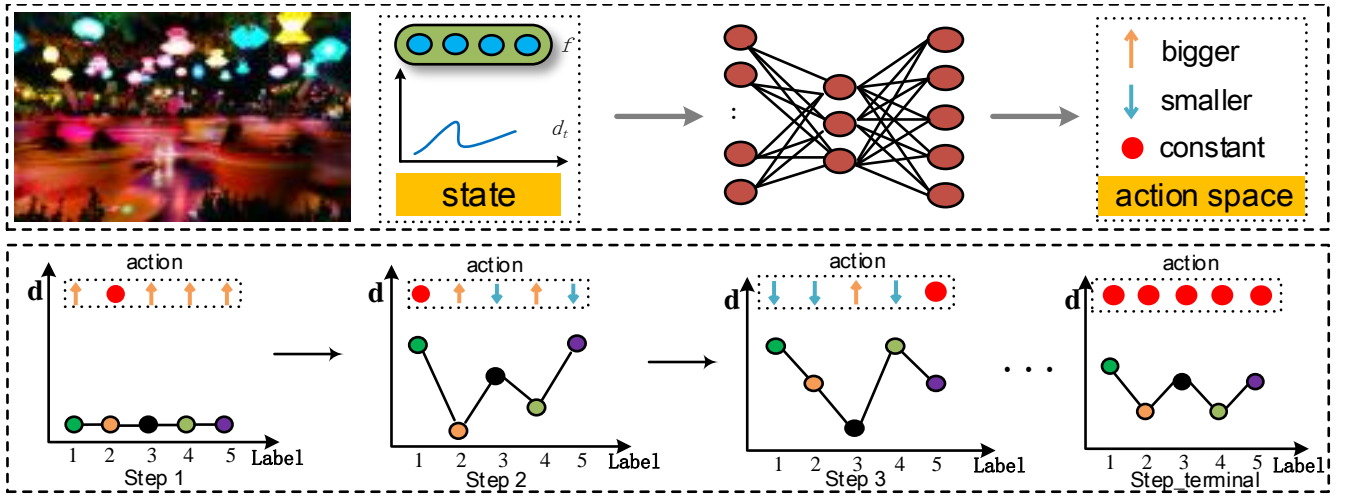


Figure 1: Overview of the *RLLE* framework. The state includes the features of the instance and the current label distribution. A set of actions is conducted on the description degrees of the label distribution respectively to adjust the label distribution at every step. The modified label distribution will replace the one contained in the state after the adjustment process.

is designed to output several actions at every step. The model starts by analyzing the initial label distribution whose description degrees are all zero and then sequentially to adjust the label distribution by the actions. Figure 1 illustrates several steps of the dynamic decision process to generate a label distribution. To evaluate the proposed method, we conduct a comprehensive experimental evaluation on both facial age estimation and image emotion classification datasets, obtaining competitive results in terms of MAE and accuracy. Our contributions are summarized as follows.

- We model the label enhancement as a dynamic decision process and propose a reinforcement learning based algorithm for label enhancement (*RLLE*). To the best of our knowledge, this paper is the first to apply reinforcement learning to label enhancement.
- We address the label enhancement for label distribution learning problem by using the prior knowledge and design the multi-dueling Q-learning algorithm to train the *RLLE* network.
- We conduct comprehensive experiments on two applications. The results show that our proposed approach not only outperforms the other state-of-the-art methods, but also achieves superior performance over the vote strategy.

## 2 Related Work

**Label distribution learning.** Label distribution learning is successfully applied to ambiguous tasks. According to the methodology in [Geng, 2016], the nature way to label an instance  $x$  is to assign a real number  $d_x^y$ , to each possible label  $y$ , representing the degree to which  $y$  describes  $x$ . Without of generality, assume that  $d_x^y \in [0, 1]$ . Further suppose that the label set is complete, *i.e.* using all the labels in the set can always fully describe the instance. Then,  $\sum_y d_x^y = 1$ . Such  $d_x^y$  is called the description degree of  $y$  to  $x$ . The learn-

ing process on the instances labeled by label distributions is therefore called *label distribution learning (LDL)*. Problem transfer (PT), algorithm adaption (AA), and specialized algorithms (SA) strategies are contained in [Geng, 2016]. LDL is successfully applied to age estimation [Xin *et al.*, 2013; Gao *et al.*, 2018], crowd counting [Zhang *et al.*, 2015], etc. LDSVR [Geng and Hou, 2015] is proposed to predict the crowd opinions about a movie. [Yi and Wu, 2019] proposes an end-to-end framework called PENCIL, which can update both network parameters and label estimations as label distribution to solve the problem of the noisy labels. DLDL [Bin-Bin *et al.*, 2017] is proposed to learn the label distribution using deep neural network for tasks with discrete and continuous labels. [Yang *et al.*, 2017b] develops binary conditional probability neural network (BCPNN) and augmented conditional probability neural network (ACPNN) for predicting sentiment distributions. [Yang *et al.*, 2017a] demonstrates a multi-task deep framework by jointly optimizing classification and label distribution prediction. As shown above, LDL applications are very broad. However, due to the difficulty of obtaining the label distributions directly, many training sets only contain simple logical labels rather than label distributions.

**Label enhancement.** Label enhancement (LE) [Xu *et al.*, 2018] is defined as recovering the label distributions from the logical labels in the training set via leveraging the topological information of the feature space and the correlation among the labels. The logical label vector of instance  $x_i$  is denoted by  $l_i = (l_{x_i}^{y_1}, l_{x_i}^{y_2}, \dots, l_{x_i}^{y_c})^T$ , where  $c$  is the number of possible labels. The description degree of  $y$  to  $x$  is denoted by  $d_x^y$ , and the label distribution of  $x_i$  is denoted by  $d_i = (d_{x_i}^{y_1}, d_{x_i}^{y_2}, \dots, d_{x_i}^{y_c})^T$ . Given a training set,  $S = \{(x_i, l_i) | l \leq i \leq n\}$ , where  $l_i$  is the logical labels, LE recovers the label distribution  $d_i$  of  $x_i$  from the logical labels, and thus transforms  $S$  into a LDL train set  $\varepsilon = \{(x_i, d_i) | l \leq i \leq n\}$ . Fuzzy clustering (FCM) [Gayar *et al.*, 2006] is one of the label en-

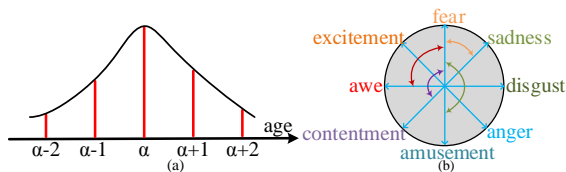


Figure 2: Two examples of the prior knowledge for age estimation and image emotion classification. (a) The farther away from the chronological age  $\alpha$ , the lower its description degree is. (b) Emotion pairs with 45 degrees have a positive relationship. More than 45 degrees between one emotion and the dominant votes emotion has no/negative relationship.

enhancement algorithms, which attempts to cluster feature vectors by iteratively minimizing an objective function. [Li *et al.*, 2016] recovers the label distributions from logical labels by using iterative label propagation. The topological structure of the feature space is considered for label enhancement in [Hou *et al.*, 2016], which is based on manifold learning. [Jiang *et al.*, 2006] demonstrates that a fuzzy support vector machine can effectively reduce the effects of outliers when solving the classification problem and uses the kernel function to calculate the center and radius. Generating the distribution with a univariate Gaussian function is widely used in multiple applications [Xin *et al.*, 2013; Gao *et al.*, 2018; Yang *et al.*, 2017a]. More recently, the Graph Laplacian Label Enhancement (GLLE) [Xu *et al.*, 2018] is proposed to generate label distributions. However, the related works ignore recovering the label distribution by using the prior knowledge, which always implies the important properties of the label distribution. In this paper, we consider using prior knowledge combining reinforcement learning algorithm to solve the label enhancement problem.

### 3 Methodology

We define the properties implied in prior knowledge as the target state. The agent can learn the policy to sequentially adjust the description degrees of the label distribution automatically to the target state.

#### 3.1 The Definition of Prior Knowledge

The prior knowledge always implies important properties of the label distribution. Figure 2 shows two examples of prior knowledge for facial age and image emotion respectively. In the age estimation application, due to the appearance similarity of the neighboring ages, both the chronological age 25 and the neighboring ages 24 and 26 can be used to describe the appearance of a 25-year-old face [Xin *et al.*, 2013]. This is consistent with the real-life experience that people usually predict another person’s facial age in a way such as “around 25 years old”. The description importance of neighboring ages is lower than the chronological age. Therefore, as shown in Figure 2(a), the prior knowledge for facial age estimation satisfies the following two properties: 1) The description degree of  $\alpha$  is the highest in the label distribution, which ensures the leading position of the chronological age; 2) the description degree of other age decreases with the increase of the distance away from  $\alpha$ , which makes the age closer to the

chronological age contribute more to the class description. In the adjustment process, if the generated label distribution meets these two properties, it means that the terminal state has been reached.

Psychological theories are widely adopted [Zhou *et al.*, 2016; Yang *et al.*, 2017a]. The distance among emotions can be defined by counting the number of steps from one emotion to another according to Mikels’ wheel [Mikels *et al.*, 2005]. In this paper, we use the label from the dominant votes as the single emotion label  $y_{dominant\_label}$ . The description degree of the single emotion label is the highest.  $d_x^{y_i} > d_x^{y_j}$  if the distance between the  $i$ -th emotion and the dominant votes label is smaller than the distance between the  $j$ -th emotion and the dominant votes label. Inspired by [Zhou *et al.*, 2016], we reproduce a wheel of eight emotions’ relationship, which is shown in Figure 2(b). In the reproduced emotion wheel, emotions sat at the opposite end have an opposite relationship, while emotions next to each other are more closely related. The correlation between the two emotions comes from psychological research [Mikels *et al.*, 2005]. We quantify the relations between each pair of emotions based on the angle between them in the wheel. For example, emotion pairs with 45 degrees have a positive relationship, while emotion pairs with 90 degrees have no relationship. Emotions pairs with 180 degrees and 135 degrees have negative relationship. To reduce the influence of the negative relationship, it would like to enforce the description degree  $d_x^{y_i} = 0$ , if the angle between the label  $y_i$  and the dominant votes label  $y_{dominant\_label}$  more than 90 degrees.

From the above analysis, the prior knowledge can be defined as the properties that the descriptive degree of the logical label is the highest, the farther the other labels are from the logical label, the lower the descriptive degree. The prior knowledge can be applied to any applications that meet these two properties. It is not limited to facial age estimation and image emotion classification.

#### 3.2 Label Enhancement as a Dynamic Decision Process

We cast the problem of label enhancement as a dynamic decision process since this setting provides a formal framework to model an agent that makes a sequence of decisions to adjust the single logical labels into label distributions. The properties implied in the prior knowledge can be formulated as the terminal state. A reinforcement learning agent can learn a policy to reach the terminal state. Formally, the dynamic decision process has a set of actions  $A$ , a set of states  $S$ , and a reward function  $R$ . This section presents details of these three components.

**Actions.** The set of actions  $A$  is composed of three actions that can be applied to the description degrees of the label distribution respectively. These actions are illustrated in Figure 1, action to make *bigger*, *smaller*, *constant*. In this way, the agent has three degrees of freedom to adjust the distribution during any interaction with the environment. A label distribution is defined as  $d_x = \{d_x^{y_1}, d_x^{y_2}, \dots, d_x^{y_c}\}$ , where  $c$  is the number of the possible labels. Any of the adjustment actions making a discrete change to the description degree by a factor

relative to its current value in the following way:

$$d_x^{y_i} = d_x^{y_i} + \delta, d_x^{y_i} = d_x^{y_i} - \delta, d_x^{y_i} = d_x^{y_i}, \quad (1)$$

In early exploration experiments, we noticed that smaller values make the agent slower to reach the target distribution. We set  $\delta = 1$  in our experiments. Since this value gives a good trade-off between speed and adjustment accuracy. According to the definition of label distribution learning,  $\sum_y d_x^{y_i} = 1$ . Every adjustment step is followed by a normalization operation on the label distribution to satisfy the sum of description degrees is 1, *i.e*

$$d_x^{y_i} = \frac{e^{d_x^{y_i}}}{\sum_c e^{d_x^{y_i}}}, \quad (2)$$

We do not set a *trigger* action like [Zhao *et al.*, 2018], because we found that the trigger action introduced extra errors. Unlike previous works, only one action is mapped from the state. In our work, a set of actions is selected at every step. The number of selected actions equals to the number of labels.

**State.** The state  $s_t$  is defined as a tuple  $(f, d_t)$ , where  $f \in \mathbb{R}$  is the features of an instance, and  $d_t \in \mathbb{R}^c$  is the label distribution at time step  $t$ . The feature vector  $f$  is extracted from a pre-processing stage. We will illustrate the details in Section 4.1 and 4.2. The label distribution  $d_t$  is defined as the result after the last adjustment step.  $c$  is the number of labels. So  $d_t$  is a  $c$ -dimension vector. In our earlier experiments, inspired by [Zhao *et al.*, 2018], the history actions were included in the state space. But the results were poor, so we removed the history actions from the state space. After the decision of the action  $a_t$  in the state  $s_t$ , the next state  $s_{t+1}$  is obtained by the state transition:  $s_{t+1} = (f, d_{t+1})$ . When the agent reaches the terminal state, which is defined by the prior knowledge, we finish the distribution adjustment process and obtain the target label distribution.

**Reward.** The reward function  $r$  is proportional to the improvement that the agent makes to adjust the label distribution after selecting a particular set of actions. The reward function is defined as  $r(s_t)$  since the agent obtains the reward by the state  $s$  regardless of the action  $a$ . Similar to [Silver *et al.*, 2016], at the terminal step  $T$ , that is, the agent reaches the terminal state,  $r(s_T)$  is assigned by 3. We give a larger award to encourage *RLLE* to produce the label distribution with quick convergence, which was shown effective in experiments.

### 3.3 Training with Multi-dueling Q-learning

The input of *RLLE* is the features of the instance and the logical label. The output is the required label distribution  $d_x = \{d_x^{y_1}, d_x^{y_2}, \dots, d_x^{y_c}\}$ . A set of actions is needed to recover the label distribution. We propose multi-dueling Q-learning named as *MQ-learning* that takes as input the state representation and gives as output the value of a set of actions based on dueling Q-learning [Wang *et al.*, 2015]. The multi-dueling Q-learning outputs  $c$  actions. And every action is selected by  $\text{argmax}_{a \in A} Q^*(s, a)$ . So  $c$  dueling Q functions are included in *MQ-learning*. Every dueling Q function represents two separate estimators: one for the state value function and one for the state dependent action advantage function, while sharing a common feature learning module. The multi-dueling

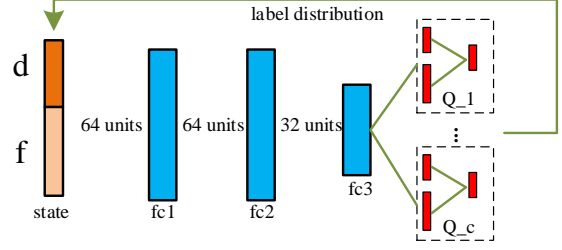


Figure 3: The multi-dueling Q-learning network includes three fully connected layers and outputs  $c$  dueling Q functions. Each dueling Q function represents two separate estimators: one for the value function and one for the advantage function.

Q-learning structure is combined by the all dueling Q functions jointly.

$$MQ(s, a) = \sum_i^c Q_i(s, a; \theta, \phi, \beta), \quad (3)$$

$$Q(s, a; \theta, \phi, \beta) = V(s; \theta, \beta) + (A(s, a; \theta, \phi) - \frac{1}{|A|} \sum_a A(s, a; \theta, \phi)), \quad (4)$$

where  $c$  indicates the number of labels.  $s$  represents the state,  $a$  denotes the action, and  $\theta, \phi, \beta$  mean the parameters in the feature learning module, advantage function and value function respectively.

Our network has three fully connected layers (fc1, fc2, fc3) combined with the ReLU layers, fc1 and fc2 have 64 nodes, fc3 has 32 nodes, as shown in Figure 3. In our task, the network outputs  $c$  actions at every step, where  $c$  equals to the number of labels. And every branch has two separate estimators. One for the value function  $V(s; \theta, \beta)$  and the other for the advantage function  $A(s, a; \theta, \phi)$ . The  $c$  actions will be applied to the description degrees respectively. The update for the network weights at the  $i$ -th iteration  $\theta_i$  given transition sample  $(s, a, r, s')$  is as follows:

$$\theta_{i+1} = \theta_i + \alpha(r + \gamma \max_a MQ(s', a; \theta_i) - MQ(s, a; \theta_i)) \nabla_{\theta_i} MQ(s, a; \theta_i), \quad (5)$$

where  $a$  represents the actions that can be taken at state  $s'$ ,  $\alpha$  is the learning rate and  $\gamma$  is the discount factor.

## 4 Experiments

We evaluate our method on two applications, *e.g.* *age estimation* and *image emotion classification*, with four datasets. For both models, the learning rate is 0.001, the batch size is 64, the discount factor  $\gamma$  is 0.9. And the size of the prioritized replay is 5000. We use the  $\epsilon$ -greedy method to select the action for exploration. For every instance, the number of episodes is 5. Though label enhancement is different from the standard supervised learning. The setting of episodes can make the agent reach the terminal state with a quick speed.

Algorithm	Method	Dataset	
		FG-NET	MORPH
IIS	Single	6.27	6.35± 0.17
	Gaussian	5.77	5.67± 0.15
	Triangle	5.90	6.09± 0.14
	RLLE	<b>5.32</b>	<b>5.34± 0.04</b>
BFGS	Gaussian	6.70	4.78± 0.08
	RLLE	<b>5.92</b>	<b>3.94± 0.07</b>
CPNN	Single	5.31	6.59± 0.31
	Gaussian	4.76	4.87± 0.31
	Triangle	5.07	4.91± 0.29
	RLLE	<b>4.58</b>	<b>4.33± 0.10</b>

Table 1: MAEs of *RLLE*, Single label, Gaussian and Triangle functions on the *FG-NET* and *MORPH* datasets.

## 4.1 Age Estimation

**Implementation Details.** Two datasets are used in this application. The first one is the FG-NET Aging dataset [Lanitis *et al.*, 2002]. There are 1,002 face images from 82 subjects in this dataset. Each image is labeled by its chronological age. The ages are distributed in a wide range from 0 to 69. The second dataset is the much larger MORPH dataset [Ricanek and Tesafaye, 2006]. There are 55,132 face images from more than 13,000 subjects in this dataset. The ages of the face images range from 16 to 77.

An agent selects an action according to the state representation. As is illustrated in Section 3.2, the state  $s_t$  is defined as a tuple  $(f, d_t)$ , where  $f \in \mathbb{R}$  is the features of the instance. For fair comparison, following [Xin *et al.*, 2013], the feature extractor used for the FG-NET dataset is the appearance model [Edwards *et al.*, 1998]. The first 200 model parameters are used as the extracted features. The features used for the MORPH dataset are the BIF [Guo *et al.*, 2009]. The dimensionality of the BIF vector is further reduced to 200 using marginal Fisher analysis [Yan *et al.*, 2005]. In our earlier experiment, we try to output actions for all ages at every step, but we find that the actions farther from the chronological age are noise. Therefore, 5 actions are output left and right respectively, centered on the chronological age. The features vector concatenates with the label distribution, resulting in a state vector  $s \in \mathbb{R}^{211}$  in the FG-NET and MORPH datasets. We compare the proposed method *RLLE* against the Gaussian and Triangle functions used in [Xin *et al.*, 2013; Gao *et al.*, 2018]. The standard deviation of Gaussian function and the bottom length of Triangle function are recommended by [Xin *et al.*, 2013]. We use the label distributions generated by *RLLE* as the ground-truth to train the age estimation model with three label distribution learning algorithms, *e.g.* *IIS*, *BFGS*, and *CPNN* [Xin *et al.*, 2013]. According to [Xin *et al.*, 2013; Gao *et al.*, 2018], mean absolute error (MAE) is used to evaluate the performance.

**Results.** Table 1 shows the MAE of the all compared algorithms. The standard deviations on the MORPH dataset are also given in the table. We randomly select 80% for training and the remaining 20% for testing. The results are reported by running this process 10 times. As can be seen, the overall performance of the *RLLE* method is better than that

of the Single, Gaussian and Triangle functions [Geng, 2016; Xin *et al.*, 2013; Gao *et al.*, 2018]. Specially, *RLLE* performs remarkably better than the baseline with the BFGS algorithm on the MORPH dataset, which reaches  $3.94 \pm 0.07$ . There are mainly two reasons for the good performance of *RLLE*. First, since Gaussian and Triangle functions are just simple distribution hypothesis. And there are many restrictions by using Gaussian and Triangle functions. The prior knowledge always implies natural properties. The using of the prior knowledge makes it possible that reduces the restrictions and learns more natural label distributions. Second, given the target properties, an agent of reinforcement learning can learn to achieve the implied properties automatically. Sequential sets of actions applied to adjust the description degrees of the label distribution by multi-dueling Q-learning is a more natural way for label enhancement. The experimental results demonstrate that the label distributions generated by *RLLE* are closer to the natural truth label distributions.

## 4.2 Image Emotion Classification

**Implementation Details.** We execute our experiments on two image emotion distribution datasets, Flickr\_LDL and Twitter\_LDL [Yang *et al.*, 2017b], which have multiple annotations. Flickr\_LDL and Twitter\_LDL contain 11,150 and 10,045 images respectively, whose labels fall in the typical eight emotions space [Mikels *et al.*, 2005] (*i.e.* *amusement*, *anger*, *contentment*, *awe*, *disgust*, *excitement*, *fear* and *sadness*). 11 viewers are hired to label the Flickr\_LDL dataset with the eight commonly used emotions and 8 viewers are hired to label Twitter\_LDL dataset. The votes from the workers are integrated to generate the label distribution for each image. We use the label from dominant votes as the single emotion label to generate label distributions. We compare the label distributions generated from our method with those from vote strategy and the other label enhancement methods to evaluate which result is more suitable for image emotion classification task.

As suggested by [Yang *et al.*, 2017b], deep features are extracted with VGGNET [Simonyan and Zisserman, 2014]. For each image, we extract the last fully connected layer output as the emotion representation and reduce it to 280 dimensions using principle component analysis (PCA). Because the label space contains eight labels, so the state vectors in Flickr\_LDL and Twitter\_LDL result in  $s \in \mathbb{R}^{288}$ . Figure 2 shows the relationship among emotions, emotion pairs with 45 degrees have a positive relationship, while emotion pairs with 90 degrees have no relationship. Emotions pairs with 135 degrees and 180 degrees have a negative relationship. To reduce the influence of the negative relationship, we enforce the description degrees of the label distribution generated by *RLLE* be zero if the angle between one emotion and the dominant votes emotion is more than 45 degrees. We use the processed results as the ground-truth to train the image emotion classification model via nine label distribution learning approaches, including PT-bayes, PT-SVM, AA-KNN, AA-BP, IIS, BFGS, CPNN [Xin *et al.*, 2013], BCPNN, and ACPNN [Yang *et al.*, 2017b]. What needs to be pointed out in particular is that BCPNN and ACPNN are specially designed for image emotion classification with Flickr\_LDL and Twitter\_LDL dataset-

Dataset	Method	BP	SVM	KNN	Bayes	IIS	BFGS	CPNN	BCPNN	ACPNN	Avg R
F_LDL	Vote	52.0(4)	37.3(2)	<b>61.4(1)</b>	46.9(2)	<b>57.9(1)</b>	50.1(4)	<b>57.7(1)</b>	59.7(2)	60.0(2)	2.11
	Gau_c1	56.9(2)	35.4(4)	56.6(4)	25.8(4)	36.3(4)	63.7(2)	42.5(2)	55.8(4)	58.0(3)	3.22
	Gau_c2	57.7(3)	36.5(3)	58.1(2)	29.9(3)	37.0(3)	62.9(3)	42.3(3)	56.8(3)	57.3(4)	3.0
	RLLE	<b>58.6(1)</b>	<b>38.6(1)</b>	58.0(3)	<b>47.4(1)</b>	52.2(2)	<b>64.3(1)</b>	55.3(2)	<b>60.7(1)</b>	<b>62.5(1)</b>	<b>1.44</b>
T_LDL	Vote	72.4(2)	40.4(2)	<b>72.6(1)</b>	45.1(2)	70.3(2)	57.0(4)	70.0(3)	<b>73.0(1)</b>	<b>74.2(1)</b>	2.0
	Gau_c1	71.8(3)	36.0(3)	68.2(4)	27.6(4)	46.2(3)	74.4(2)	71.0(2)	70.7(3)	70.2(4)	3.11
	Gau_c2	70.7(4)	35.6(4)	70.6(3)	38.4(3)	45.8(4)	74.3(3)	64.7(4)	67.0(4)	70.4(3)	3.55
	RLLE	<b>72.4(1)</b>	<b>42.5(1)</b>	70.9(2)	<b>49.8(1)</b>	<b>72.2(1)</b>	<b>74.9(1)</b>	<b>71.7(1)</b>	72.5(2)	72.6(2)	<b>1.33</b>

Table 2: Experimental Results of classification accuracy on two image emotion distribution datasets are shown. ‘‘Avg R’’ in the last column means the average rank, which is used to indicate the overall performance of the classification accuracy. The accuracy is calculated by the dominant votes emotional category from the label with the maximum description degree in the predicted label distribution. ‘‘Gau\_c1’’ and ‘‘Gau\_c2’’ represent the two restrictive label enhancement methods as illustrated in Eq. (6) and Eq. (7). ‘‘Vote’’ denotes the vote strategy.

s. Besides, we compare the proposed method against the state-of-the-art label enhancement approaches used in [Yang *et al.*, 2017a]. [Yang *et al.*, 2017a] considers two constraints to generate the label distributions. The first one is using the Gaussian function to assign the description degrees of all the emotion labels based on the distance to the dominant votes label according to Mikels’ wheel. The description degrees of the label distribution can be written as follows:

$$f(x, \mu, \sigma_{conf}) = \frac{1}{\sqrt{2\pi}\sigma_{conf}} \exp\left(-\frac{|i-\mu|^2}{2\sigma_{conf}^2}\right) + \frac{\varepsilon}{C}, \quad (6)$$

where  $\mu$  represents the label from the dominant votes, and the  $\sigma_{conf}$  denotes the level of influence of each emotion label determined by the confidence in the label annotations. And the fixed parameters  $\varepsilon$  and  $C$  ensure that the sum of description degrees is normalized to 1. The second constraint assumes that each affective image only evokes sentiment with the same valence in the label distribution, either positive or negative. So the description degrees of the label distribution are changed to:

$$f(x, \mu, \sigma_{conf}) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_{conf}} \exp\left(-\frac{|i-\mu|^2}{2\sigma_{conf}^2}\right), & i \in Y_\mu, \\ 0, & i \notin Y_\mu, \end{cases} \quad (7)$$

where  $Y_\mu$  denotes all the emotion labels of the same valence, either positive or negative. There exists label ambiguous cases among the emotions. Most works adopt vote strategy to get label distributions [Yang *et al.*, 2017b; Geng and Hou, 2015]. Vote strategy is always considered as the most objective method to solve the uncertainty of the ‘‘ground-truth label’’ problem. We also compare our method against the results from the vote strategy.

**Results.** The performance of our method and the other three methods is shown in Table 2. For Flickr\_LDL and Twitter\_LDL datasets, the accuracy is examined for evaluating the performance of the results. For testing, the accuracy is measured by the label with the maximum description degree in the label distribution. As can be seen, for Flickr\_LDL dataset, *RLLE* ranks 1st in 66.6% cases and 2nd in 22.2% cases across all the evaluation accuracy measures. And vote strategy ranks 1st in 33.3% cases and 2nd in 44.4% cases. The other two Gaussian constraints have lower ranks. The average rank of *RLLE* is 1.44, which achieves superior performance over the vote strategy and the other label enhancement

methods. For Twitter\_LDL dataset, *RLLE* ranks 1st in 66.6% cases and 2nd in 33.3% cases. And the vote strategy ranks 1st in 33.3% and 2nd in 44.4% cases. For the two image emotion distribution datasets, the average rank of our method is both 1st. We can conclude that our proposed method shows superiority in accuracy across all the two Gaussian constraints and the vote strategy. The experimental results present that using psychology theories as the prior knowledge is more suitable for generating the label distributions than the vote strategy and the other methods. There are mainly two reasons for the good performance of *RLLE*. First, using prior knowledge to recover the label distributions can highlight the essential properties, which is implied in the prior knowledge. Second, similar to the age estimation application, given the target properties, sequential decision making process to adjust the description degrees of the label distribution is a more suitable way for label enhancement.

## 5 Conclusion

This paper proposes the *RLLE* approach for label enhancement via the prior knowledge. We cast the problem of label enhancement as a dynamic decision process and design a multi-dueling Q-learning algorithm to train the proposed model. Given the target state defined by the prior knowledge, the agent can conduct a set of actions on the description degrees to adjust the initial label distribution into the target one. We conduct experiments on two scenarios. Comparing with several state-of-the-art methods, our proposed approach achieves the best performance. In the future, we will explore more prior knowledge to achieve better performance for different applications.

## Acknowledgements

This research was supported by the National Key Research and Development Plan of China (No. 2017YFB1002801), the National Science Foundation of China (61622203), the Jiangsu Natural Science Funds for Distinguished Young Scholar (BK20140022), the National Science Foundation of China (61702095), the Collaborative Innovation Center of Novel Software Technology and Industrialization, and the Collaborative Innovation Center of Wireless Communications Technology.

## References

- [Bin-Bin *et al.*, 2017] Gao Bin-Bin, Xing Chao, and Geng Xin. Deep label distribution learning with label ambiguity. *TIP*, 26(6):2825–2838, 2017.
- [Edwards *et al.*, 1998] Gareth J. Edwards, Andreas Lanitis, Christopher J Taylor, and Timothy F. Cootes. Statistical face models: Improving specificity. *Image Vision Computing*, 1998.
- [Gao *et al.*, 2018] Bin-Bin Gao, Hong-Yu Zhou, Jianxin Wu, and Xin Geng. Age estimation using expectation of label distribution learning. In *IJCAI*, 2018.
- [Gayar *et al.*, 2006] Neamat EI Gayar, Friedhelm Schwenker, and Palm. A study of the robustness of knn classifiers trained using soft labels. *ANNPR*, 2006.
- [Geng and Hou, 2015] Xin Geng and Peng Hou. Pre-release prediction of crowd opinion on movies by label distribution learning. In *IJCAI*, 2015.
- [Geng, 2016] Xin Geng. Label distribution learning. *TKDE*, 28(7):1734–1748, 2016.
- [Gibaja and Eva, 2015] Gibaja and Eva. A tutorial on multi-label learning. *ACM Computing Surveys*, 2015.
- [Grigorios and Ioannis, 2006] Tsoumakas Grigorios and Katakis Ioannis. Multi-label classification: An overview. *IJDWM*, 3(3):1–13, 2006.
- [Guo *et al.*, 2009] Guodong Guo, Guowang Mu, Yun Fu, and Thomas S. Huang. Human age estimation using bio-inspired features. In *CVPR*, 2009.
- [Hou *et al.*, 2016] Peng Hou, Xin Geng, and Min Ling Zhang. Multi-label manifold learning. In *AAAI*, 2016.
- [Jiang *et al.*, 2006] Xiufeng Jiang, Zhang Yi, and Jian Cheng Lv. Fuzzy svm with a new fuzzy membership function. *Neural Computing and Applications*, 15(3-4):268–276, 2006.
- [Lanitis *et al.*, 2002] Andreas Lanitis, Christopher J Taylor, and Timothy F. Cootes. Toward automatic simulation of aging effects on face images. *TPAMI*, 24(4):442–455, 2002.
- [Li *et al.*, 2016] Yu Kun Li, Min Ling Zhang, and Xin Geng. Leveraging implicit relative labeling-importance information for effective multi-label learning. In *ICDM*, 2016.
- [Ling and Geng, 2017] Miaogen Ling and Xin Geng. Soft video parsing by label distribution learning. In *AAAI*, 2017.
- [Liu *et al.*, 2017] Yujie Liu, Yongbiao Gao, Shihe Feng, and Zongmin Li. Weather-to-garment: Weather-oriented clothing recommendation. In *ICME*, 2017.
- [Mikels *et al.*, 2005] Joseph A. Mikels, Barbara L. Fredrickson, and Gregory R. Larkin. A reuter-lorenz. emotional category data on images from the international affective picture system. *Behavior research methods*, 37(4):626–630, 2005.
- [Plutchik, 2001] Plutchik. The nature of emotions. *Amer. Scientist*, 2001.
- [Ricanek and Tesafaye, 2006] Karl Ricanek and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *FG*, 2006.
- [Silver *et al.*, 2016] David Silver, Aja Huang, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.
- [Simonyan and Zisserman, 2014] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *Computer Science*, 2014.
- [Su and Geng, 2019] Kai. Su and Xin. Geng. Soft facial landmark detection by label distribution learning. In *AAAI*, 2019.
- [Sutton and Barto, 2018] Richard S. Sutton and Andrew G. Barto. Reinforcement learning: An introduction. *MIT Press*, 2018.
- [Wang *et al.*, 2015] Ziyu Wang, Nando De Freitas, and Marc Lanctot. Dueling network architectures for deep reinforcement learning. In *ICML*, 2015.
- [Xin *et al.*, 2013] Geng Xin, Yin Chao, and Zhou Zhi-Hua. Facial age estimation by learning from label distributions. *TPAMI*, 35(10):2401–2412, 2013.
- [Xu *et al.*, 2018] Ning Xu, An Tao, and Xin Geng. Label enhancement for label distribution learning. In *IJCAI*, 2018.
- [Yan *et al.*, 2005] Shuicheng Yan, Dong Xu, Benyu Zhang, and Hong Jiang Zhang. Graph embedding: A general framework for dimensionality reduction. In *CVPR*, 2005.
- [Yang *et al.*, 2017a] Jufeng Yang, Dongyu She, and Ming Sun. Joint image emotion classification and distribution learning via deep convolutional neural network. In *IJCAI*, 2017.
- [Yang *et al.*, 2017b] Jufeng Yang, Ming Sun, and Xiaoxiao Sun. Learning visual sentiment distributions via augmented conditional probability neural network. In *AAAI*, 2017.
- [Yi and Wu, 2019] Kun Yi and Jianxin Wu. Probabilistic end-to-end noise correction for learning with noisy labels. In *CVPR*, 2019.
- [Zhang and Zhou, 2014] Min Ling Zhang and Zhi Hua Zhou. A review on multi-label learning algorithms. *TKDE*, 26(8):1819–1837, 2014.
- [Zhang *et al.*, 2015] Zhaoxiang Zhang, Wang Mo, and Geng Xin. Crowd counting in public video surveillance by label distribution learning. *Neurocomputing*, 2015.
- [Zhao *et al.*, 2018] Sicheng Zhao, Hongxun Yao, Gao Yue, and Tat Seng Chua. Stackdrl: Stacked deep reinforcement learning for fine-grained visual categorization. In *IJCAI*, 2018.
- [Zhou and Zhang, 2006] Zhi-Hua Zhou and Min Ling Zhang. Multi-instance multi-label learning with application to scene classification. In *NIPS*, 2006.
- [Zhou *et al.*, 2016] Deyu Zhou, Xuan Zhang, Yin Zhou, Quan Zhao, and Xin Geng. Emotion distribution learning from texts. In *EMNLP*, 2016.