

Multi-Granularity Open Intent Classification via Adaptive Granular-Ball Decision Boundary

Yanhua Li^{1,2}, Xiaocao Ouyang^{1,2}, Chaofan Pan^{1,2}, Jie Zhang^{1,2}, Sen Zhao³,
Shuyin Xia³, Xin Yang^{1,2*}, Guoyin Wang⁴, Tianrui Li⁵

¹School of Computing and Artificial Intelligence, Southwestern University of Finance and Economics, Chengdu 611130, China

²Engineering Research Center of Intelligent Finance, Ministry of Education, Chengdu 611130, China

³Key Laboratory of Big Data Intelligent Computing, Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

⁴National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China

⁵School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

yyhhliyanhua@163.com, oyxiaocao@swufe.edu.cn, pan.chaofan@foxmail.com, jiezhanglx@gmail.com,

zhaosen@cqupt.edu.cn, xiasy@cqupt.edu.cn, yangxin@swufe.edu.cn, wangy@cqu.edu.cn, trli@swjtu.edu.cn

Abstract

Open intent classification is critical for the development of dialogue systems, aiming to accurately classify known intents into their corresponding classes while identifying unknown intents. Prior boundary-based methods assumed known intents fit within compact spherical regions, focusing on coarse-grained representation and precise spherical decision boundaries. However, these assumptions are often violated in practical scenarios, making it difficult to distinguish known intents from unknowns using a single spherical boundary. To tackle these issues, we propose a **Multi-granularity Open intent classification method via adaptive Granular-Ball decision boundary (MOGB)**. Our MOGB method consists of two modules: representation learning and decision boundary acquiring. To effectively represent the intent distribution, we design a hierarchical representation learning method. This involves iteratively alternating between adaptive granular-ball clustering and nearest sub-centroid classification to capture fine-grained semantic structures within known intent classes. Furthermore, multi-granularity decision boundaries are constructed for open intent classification by employing granular-balls with varying centroids and radii. Extensive experiments conducted on three public datasets demonstrate the effectiveness of our proposed method.

Code — <https://github.com/Liyanhuaa/MOGB>

Introduction

The identification of user intent is critical for dialogue systems, especially as individuals increasingly rely on these systems to support both daily activities and professional tasks. In real-world applications, dialogue systems often encounter intents from unknown (open) classes that they have not seen during training, resulting in their inability to handle these intents effectively (Zheng, Chen, and Huang 2020). To address this, open intent classification aims to classify known intents while identifying unknown intents (Shu, Xu, and Liu 2017).

*Corresponding author

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

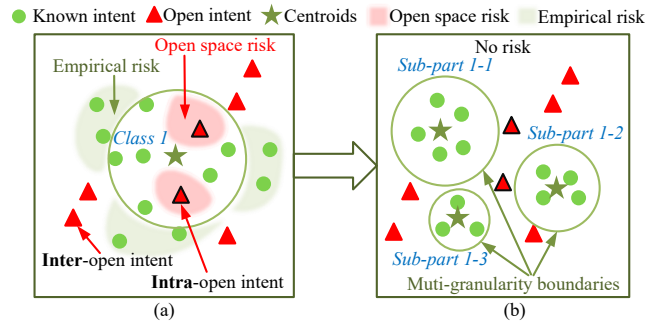


Figure 1: (a) Previous boundary-based methods struggle to learn a boundary for open intent classification by balancing open space risk and empirical risk. (b) Our proposed Multi-granularity decision boundary can effectively eliminate both two risks.

Recent researches endeavor to develop effective methods to tackle the issue of open intent classification. Early methods (Jain, Scheirer, and Boulton 2014; Hendrycks and Gimpel 2022) focused on establishing a K -class classifier and determining whether a sample belongs to unknowns by checking if its maximum probability exceeds a certain threshold. Due to the overfitting of deep learning approaches (Guo et al. 2017), K -class classifiers tend to overconfidence about known intent classes and are prone to misclassifying unknown intents as known ones. To overcome this limitation, some studies train a $K+1$ classifier directly, leveraging pseudo-unknown data to represent the open intents (Zheng, Chen, and Huang 2020; Cheng et al. 2022).

In recent years, boundary-based methods (Zhang, Xu, and Lin 2021; Zhang et al. 2023a; Yang et al. 2022) have been proposed to learn specific decision boundaries between known intents and unknowns. These methods often involve two stages, i.e. representation learning and boundary learning. They mainly focused on learning more accurate decision boundaries to balance open space risk (classify unknowns as known) and empirical risk (classify known

as unknowns). Initially, Zhang et al. (Zhang, Xu, and Lin 2021) acquired feature inputs through softmax-based representation learning and obtained adaptive decision boundary (ADB) for each known intent class. Following this work, Liu et al. (Liu et al. 2023) designed a K-center contrastive learning algorithm to learn discriminative intent representations and proposed adjustable boundary learning by expanding or shrinking the initial decision boundary. Zhang et al. (Zhang et al. 2023a) learned distance-aware intent representations and tight adaptive decision boundaries to improve open intent detection.

Existing boundary-based methods (Zhang, Xu, and Lin 2021; Zhang et al. 2023a) implicitly assumed the features of known intents distributed in a compact spherical region, while intents lying outside this distribution are regarded as open intents. It means that open intents do not exist within the distribution of a single known intent class and only appear between different known intent classes. However, as illustrated in Figure 1 (a), in real-world situations, known intents do not consistently adhere to a compact spherical distribution and open intents can exist between or within the distribution of known intents (named inter-open intents and intra-open intents). The investigation of Gaussian Hypothesis Testing concerning known intents distribution, as presented by Zhou et al. (Zhou, Liu, and Qiu 2022), revealed that merely 57% of the known intents within the CLINC-FULL dataset conform to the Gaussian assumption. This finding further implies that the assumption of a spherical feature distribution is unreasonable. When dealing with non-spherical intent distributions, it is difficult to distinguish known intents from unknowns using just one decision boundary. Furthermore, the open space risk resulting from intra-open intents can never be eliminated, regardless of how to adjust the boundary, as shown in Figure 1 (a).

In this paper, we propose a new straightforward but powerful solution: exploring the patterns within each known intent class, dividing intents into multiple sub-parts, and establishing decision boundaries for each sub-part, as illustrated in Figure 1 (b). This solution offers two benefits over earlier boundary-based methods. On the one hand, having several sub-parts can more accurately reflect the actual distribution of known intent classes. On the other hand, it creates fine-grained decision boundaries that can more efficiently exclude intra-open intents, reducing the open space risk.

In light of this, the following three challenges come to this paper: (1) How to divide each known intent class into multiple sub-parts that accurately reflect the intent distribution? (2) How to learn discriminative representations that capture different semantic aspects within class, facilitating the division of known classes into sub-parts? (3) How to obtain decision boundaries for these sub-parts within each known intent class?

To solve these challenges, we propose a **Multi-granularity Open intent classification method via adaptive Granular Ball decision boundary (MOGB)**. Our MOGB method consists of a hierarchical representation learning module and a decision boundary acquiring module. For **challenge (1)**, inspired by granular-ball computing (Xia et al. 2023), we propose a hierarchical representation approach through adaptive

granular-ball clustering, which reflects intent distribution using multiple granular-balls with varying sizes. For **challenge (2)**, the nearest sub-centroid classifier is proposed to learn discriminative representations, with the loss function encouraging each intent to move closer to its semantically closest sub-centroid. In the training of hierarchical representation learning, the adaptive granular-ball clustering and nearest sub-centroid classification are iteratively conducted to mutually complement each other, facilitating the acquisition of discriminative representations. For **challenge (3)**, we propose multi-granularity decision boundaries through granular-balls with varying radii and sub-centroids to distinguish the known and open intents. Our contributions are summarized as follows:

- We propose the hierarchical representation learning method, which adaptively reflects intra-class structure and intent distribution via granular-ball clustering.
- We propose the nearest sub-centroid classifier to acquire discriminative representations by gathering intents to its semantically nearest sub-centroid.
- We design multi-granularity decision boundaries for each known intent class to facilitate open intent classification, which provides a new way to solve this question.

Related Work

Open Intent Classification

Recognizing the intent behind user interactions is crucial in dialogue systems, as they aim to understand users' potential requirements. Most classification models currently work under the closed-world assumption, which may not align with real-world systems functioning in open settings. Such systems often face queries that fall outside the supported intents, referred to as out-of-domain queries. With the prevalent of dialogue systems, there has been an increasing emphasis on identifying open intents in recent times (Zhou, Liu, and Chen 2021; Zhan et al. 2021; Parmar et al. 2023; Yan et al. 2020; Zhang et al. 2021).

Existing open intent classification methods can be categorized into two primary groups. The first group involves directly training a classifier with $K+1$ classes, with an extra class representing the open intent (Zheng, Chen, and Huang 2020). The second group of methodologies focused on outlier detection algorithms, which can be categorized into two main subtypes: probability-based and decision boundary-based approaches. Probability-based methods use a pre-defined threshold to determine if a sample represents an open intent (Shu, Xu, and Liu 2017). Boundary-based approaches (Zhang, Xu, and Lin 2021; Liu et al. 2023; Zhang et al. 2023a) have been developed to overcome the limitations of probability-based methods, which struggle to establish a distinct boundary between known classes and unknowns.

However, current boundary-based approaches assume that the feature distribution of known intents is within a compact spherical region, but this is not always the case in reality. In addition, the single decision boundary they established includes intra-open intents, increasing the open space risks.

Granular-ball Computing

Building on the human cognitive mechanisms of “macro first” (Chen 1982) and multi-granularity cognitive computation (Wang 2017), Xia et al. (Xia et al. 2019) proposed granular-ball computing, which is efficient, robust, and interpretable. It adheres to a hierarchical rule of progressing from coarse-grained to fine-grained. In granular-ball computing, varying sizes of granular-balls are adaptively generated to depict data distribution. Each granular-ball is defined by a centroid and a radius, with the centroid representing the mean of all object vectors within the ball and the radius determined by the average distance of objects from the centroid. The effectiveness of this self-supervised data-covering method has been empirically validated (Xie et al. 2024c), and it has demonstrated successful applications in addressing diverse challenges within the field of AI (Liu et al. 2024; Wang et al. 2024a,b).

Current granular-ball computing methods encompass granular-ball classification (Xia et al. 2024b; Quadir and Tanveer 2024), granular-ball clustering (Xie et al. 2024a,b), granular-ball three-way decision (Yang et al. 2024; Xia et al. 2024a), and granular-ball rough sets (Xia et al. 2020, 2023; Zhang et al. 2023b). Furthermore, several applications have demonstrated the effectiveness of granular-ball representation, such as granular-ball representation in text adversarial defense (Wang et al. 2024a), label noise combating (Wang et al. 2024b), feature selection (Cao et al. 2024), and deep reinforcement learning (Liu et al. 2024).

In summary, granular-ball computing is an effective way to discover the structure of arbitrary data distribution, with the radius of the granular-ball serving as a natural decision boundary for open classification. Despite this, there remains a notable absence of research applying granular-ball computing to the specific domain of open intent classification. Our MOGB method will employ granular-ball clustering to achieve hierarchical representation and establish multi-granularity adaptive decision boundaries.

Our Methodology

Problem Statement and Model Overview

Problem Statement Given a dataset $S = \{(\mu_i, y_i)\}_{i=1}^N$ consisting of N intents, where each pair (μ_i, y_i) represents an intent μ_i and its corresponding label y_i . The labels $y_i \in \{1, \dots, K, K+1\}$ can be divided into two categories: $I_{\text{Known}} = \{1, \dots, K\}$ denotes a set of known intent labels with K representing the number of known intent classes, while $I_{\text{Unknown}} = K+1$ is used to label unknown (open) intents. During training and validation, only samples associated with known intents are utilized. In contrast, the testing dataset includes both known and open intents, and we need to classify the known intents into their respective classes and identify open intents as unknowns.

Model Overview Our proposed MOGB method consists of two modules: hierarchical representation learning and decision boundary acquiring. In module 1, the hierarchical representation of each known intent class is obtained via adaptive granular-ball clustering, as depicted in Figure 2

(a). Subsequently, we learn discriminative representation by the nearest sub-centroid classifier, as shown in Figure 2 (b). The iterative process of granular-ball clustering and nearest sub-centroid classification during training mutually reinforce each other to obtain discriminative representations. In module 2, multi-granularity decision boundaries are constructed for open intent classification, as illustrated in Figure 2 (c).

Hierarchical Representation Learning

In this section, the BERT pre-trained model is employed for feature extraction, followed by adaptive granular-ball clustering and the nearest sub-centroid classifier.

Feature Extraction Utilizing the pre-trained BERT model (Kenton and Toutanova 2019), we extract token embeddings $[CLS, T_1, T_2, \dots, T_M] \in \mathbb{R}^{(M+1) \times H}$ from the final hidden layer to represent each intent μ_i . Here, M represents the sequence length and H denotes the embedding dimension. Following previous studies (Zhang, Xu, and Lin 2021), a mean pooling operation is applied to these embeddings in order to construct a semantic representation of each sentence denoted as $x_i \in \mathbb{R}^H$:

$$x_i = \text{MeanPooling}([CLS, T_1, T_2, \dots, T_M]), \quad (1)$$

where CLS acts as a classification token. To further enhance the feature extraction process, x_i is subsequently passed through a dense layer h , transforming the representation to $z_i \in \mathbb{R}^D$:

$$z_i = h(x_i) = \text{ReLU}(W_h x_i + b_h), \quad (2)$$

where D denotes the dimension of the intent representation, $W_h \in \mathbb{R}^{H \times D}$ and $b_h \in \mathbb{R}^D$ represent the weight and bias parameters of the dense layer h , respectively.

Adaptive Granular-ball Clustering In order to accurately represent the true distribution of the known intent class and mitigate open space risk associated with intra-open intents, we need to group the known intent class into sub-classes that reflect inherent patterns within the class. Although conventional class-wise clustering methods can partially reveal data structure, their predefined cluster number setting limits the exploration of diverse patterns within each class. Granular-ball clustering is an efficient and adaptive clustering method to represent the true data distribution (Xia et al. 2021). Motivated by it, we represent the intent distribution using granular-balls generated through adaptive granular-ball clustering, as shown in Figure 2 (a), which eliminates the necessity for specifying cluster numbers.

The basic definitions of the granular-ball are introduced as following (Xia et al. 2021):

Definition 1. Given a dataset of intents $Z = \{(z_i, y_i) \mid i = 1, 2, \dots, N\}$ consisting of N feature representations z_i and their corresponding labels y_i , we have a set of granular-balls denoted as $G = \{gb_1, gb_2, \dots, gb_m\}$ generated on the dataset, where m denotes the number of granular-balls. Each granular-ball $gb_j = \{(z_i, y_i) \mid i = 1, 2, \dots, n_j\}$ represents a distinct subset containing n_j intents. The fundamental components of each granular-ball include sample

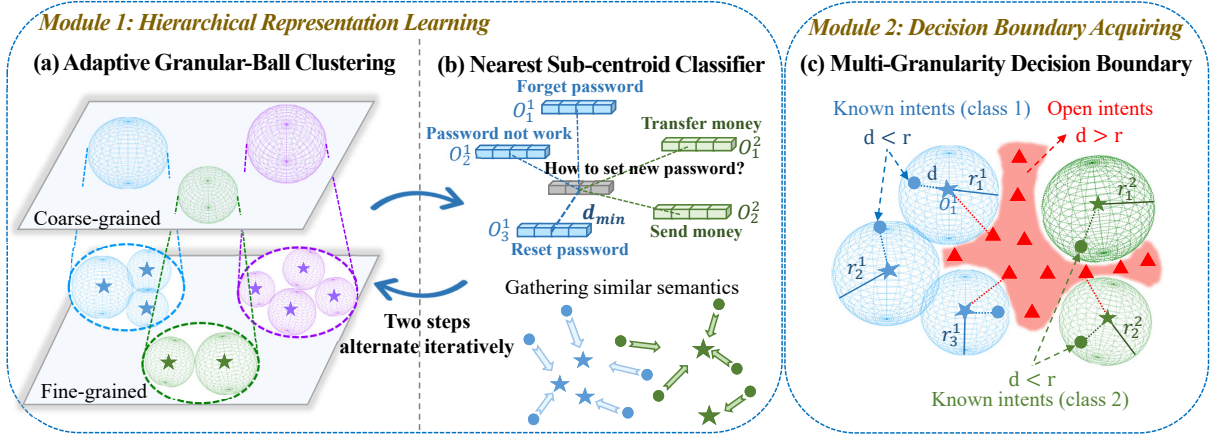


Figure 2: The architecture of MOGB. In the hierarchical representation learning module, we generate granular-balls on all known intents via adaptive granular-ball clustering and then use the nearest sub-centroid classifier to learn representation. In the boundary acquiring module, multi-granularity decision boundaries are established.

count n_j , centroid O_j , radius r_j , label l_j , and purity p_j . The centroid of gb_j is computed as the average representation $O_j = \frac{1}{n_j} \sum_{i=1}^{n_j} z_i$. The radius of gb_j is determined by calculating the average distance between all samples and the centroid: $r_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \|z_i - O_j\|$, where $\|z_i - O_j\|$ denotes the Euclidean distance between z_i and O_j . The label l_j assigned to granular-ball gb_j corresponds to the class that contains the highest number of samples within it. Purity p_j indicates the proportion of samples labeled with l_j in granular-ball gb_j .

The input of the clustering includes all representations $Z = \{(z_i, y_i) \mid i = 1, 2, \dots, N\}$ from different known intent classes. Adaptive granular-ball clustering is an iterative process to split granular-balls. First, the representations Z are initialized as a single granular-ball. Subsequently, the granular-balls that satisfy the split conditions undergo splitting (conditions described below). The process terminates when all granular-balls can no longer be further split.

To control the splitting process effectively, we set the condition of purity limit p_l and sample count limit n_l . Specifically, for a granular-ball gb_j with purity p_j and sample count n_j , if $p_j < p_l$ and $n_j > n_l$, the granular-ball will be split further until either $p_j \geq p_l$ or $n_j \leq n_l$. During the splitting phase, the granular-ball with a unique class label set L_j will be split into $|L_j|$ new granular-balls, where $L_j = \text{unique}(\{y_i \mid (z_i, y_i) \in gb_j\})$. To achieve this division accurately, one intent from each unique class within the original granular-ball is randomly selected as a pseudo centroid for creating a new granular-ball. Subsequently, other intents are assigned to their closest respective new granular-balls based on their Euclidean distances from pseudo centroids.

The adaptive granular-ball clustering algorithm produces a set of granular-balls denoted as $G = \{gb_1, gb_2, \dots, gb_m\}$, where each granular-ball gb_j ($j = 1, \dots, m$) is assigned a label $l_j \in I_{\text{Known}} = \{1, \dots, K\}$. To eliminate outliers and prevent overfitting, we select high-quality granular-balls with purity p_j above p_t and sample count n_j more than n_t to

represent known intents. Each known intent class $c \in I_{\text{Known}}$ corresponds to a number of n_c granular-balls labeled with c , where $n_c = \text{count}(l_j = c)$, and $\sum_{c=1}^K n_c = m$. The granular-balls representing class c are denoted as $\{gb_s^c\}_{s=1}^{n_c}$, where gb_s^c represents the s -th granular-ball with label c . The collection of all these granular-balls $\{gb_s^c\}_{c,s=1}^{K,n_c}$ reveals the detailed structure of the distributions for all known intents.

Nearest Sub-Centroid Classifier The effectiveness of adaptive granular-ball clustering is highly dependent on the quality of representation. Therefore, it is necessary to obtain discriminative representations that reflect the semantic structure of known intent classes. However, previous cross-entropy loss-based (Zhang, Xu, and Lin 2021) or contrastive loss-based (Liu et al. 2023) representation learning methods treat each class as an entity, failing to explore patterns within the class. To this end, we propose the nearest sub-centroid representation learning method to gather similar semantics and separate dissimilar semantics, as shown in Figure 2 (b).

According to the granular-balls $\{gb_s^c\}_{c,s=1}^{K,n_c}$, we compute their centroids as $\{O_s^c\}_{c,s=1}^{K,n_c}$, where O_s^c denotes the s -th sub-centroids of class c . These sub-centroids serve as the informative properties of known intents. In the classification of representation learning, the distances from test intent to all sub-centroids are calculated as follows:

$$\text{dis} \langle z_i, O_s^c \rangle = \frac{z_i^\top O_s^c}{\|z_i\| \|O_s^c\|}, \quad i = \{1, \dots, N\}, \quad (3)$$

where $\|z_i\|$ and $\|O_s^c\|$ denote the Euclidean norm vectors of z_i and O_s^c . We assign the label of test intent z_i corresponding to the label of the sub-centroid closest to z_i . Concretely, if $\text{dis} \langle z_i, O_{s^*}^{c^*} \rangle$ denotes the minimum distance among $\text{dis} \langle z_i, O_s^c \rangle$, the predicted label of z_i is c^* . The classifier rule is elaborated as follows:

$$\hat{y}_i = c^*, (c^*, s^*) = \arg \min_{c \in \{1, \dots, K\}, s \in \{1, \dots, n_c\}} \text{dis} \langle z_i, O_s^c \rangle. \quad (4)$$

Based on the classifier provided, the training loss associated with the nearest sub-centroid representation learning is

shown as follows:

$$\mathcal{L}_{gb} = \frac{1}{N} \sum_{i=1}^N -\log p(y_i|z_i), \quad (5)$$

$$p(y|z) = \frac{\exp(-\min(\{dis\langle z, O_s^y \rangle\}_{s=1}^{n_y}))}{\sum_{c=1}^K \exp(-\min(\{dis\langle z, O_s^c \rangle\}_{s=1}^{n_c}))}. \quad (6)$$

During hierarchical representation learning, adaptive granular-ball clustering produces informative sub-centroids and the nearest sub-centroid classifier optimizes the representation by adjusting the arrangement between intents and sub-centroids. The improved representation subsequently facilitates the identification of more informative sub-centroids. In summary, the two steps iteratively complement each other, eventually benefiting the open intent classification.

Multi-Granularity Decision Boundary in Inference

In our MOGB method, pairs of centroids and radii $\{O_s^c, r_s^c\}_{c,s=1}^{K,n_c}$ generated from training data of known intents are directly employed to construct multi-granularity decision boundaries for each class.

In inference, the distances between test sample z_i and all sub-centroids $\{O_s^c\}_{c,s=1}^{K,n_c}$ are first calculated as $dis\langle z_i, O_s^c \rangle$. Then, we compare each $dis\langle z_i, O_s^c \rangle$ with its corresponding radii $\{r_s^c\}_{c,s=1}^{K,n_c}$ to make open classification. If the test sample falls outside all multi-granularity boundaries, we classify it as unknown. Otherwise, the sample is predicted as the same class as the nearest sub-centroid that satisfies the distance condition. The inference rule is formulated as:

$$\hat{y}_i = \begin{cases} \text{unknown, if } dis\langle z_i, O_s^c \rangle > r_s^c, \forall c \in I_{\text{Known}}; \\ \arg \min_{c \in I_{\text{Known}}, dis\langle z_i, O_s^c \rangle \leq r_s^c} dis\langle z_i, O_s^c \rangle, \text{ Otherwise.} \end{cases} \quad (7)$$

Intuitively, compared to existing boundary-based methods that include two training processes, our MOGB method can simultaneously learn representation and acquire decision boundaries with just one training.

Experiments

Datasets

In accordance with prior studies (Zhang, Xu, and Lin 2021; Zhang et al. 2023a), we conduct experiments on three commonly used datasets: **StackOverflow** (Xu et al. 2015) is a dataset containing 3,370,528 programming question titles across 20 categories. Consistent with prior studies, a subset of 1,000 examples from each category was randomly selected for analysis. **SNIPS** (Coucke et al. 2018) is composed of 14,484 spoken utterances with 7 distinct categories of intent classes. **BANKING** (Casanueva et al. 2020) is an online banking inquiry with 13,080 instances and 77 intents classes. Table 1 presents detailed statistics of these datasets.

Baselines

We compare our MOGB method with the following open intent classification methods: **MSP** (Hendrycks and Gimpel

Dataset	#Class	#Train/Valid/Test	Length
StackOverflow	20	12,000/2,000/6,000	9.18
SNIPS	7	13,084/700/700	9.05
BANKING	77	9,003/1,000/3,080	11.91

Table 1: Statistics of datasets.

2022) classifies samples based on their maximum softmax probabilities; **DOC** (Shu, Xu, and Liu 2017) conducts open intent classification by establishing probability thresholds through Gaussian fitting. **OpenMax** (Bendale and Boulton 2016) implements open set recognition by transforming the softmax layer in a neural network into an OpenMax layer. **DeepUnk** (Lin and Xu 2019) utilizes margin loss that minimizes intra-class variance and maximizes inter-class variance to learn enhanced representation. **ADB** (Zhang, Xu, and Lin 2021) acquires feature representation by softmax loss and learns a single decision boundary for each class. **DA-ADB** (Zhang et al. 2023a) learns distance-aware intent representations to obtain appropriate decision boundary.

Evaluation Metrics

Following previous work (Zhang, Xu, and Lin 2021), we utilize F1-score and accuracy metrics to assess different approaches. Additionally, we also compute the fine-grained macro F1-score for both known classes and unknowns.

Experimental Settings

Following the experimental setting of prior study (Zhang, Xu, and Lin 2021), a predetermined percentage of classes are randomly designated as known classes, while the remaining classes are considered unknown (open). During testing, all unknown classes are collectively treated as a single additional class. Training exclusively utilizes samples from the known classes, with samples from the unknown classes excluded from the training dataset. The test dataset comprises both known and unknown samples. The experimental results are presented based on three datasets, with known class ratios set at 25%, 50%, and 75%, respectively.

During representation learning, we use the pre-trained BERT model (Kenton and Toutanova 2019), keeping all parameters fixed except for those in the final layer. The training batch size is set to 128, with a learning rate of $2e-5$ to fine-tune the final layer. We control the adaptive granular-ball clustering by two attributes of the granular-ball: purity limit p_l and sample count limit n_l . We set $p_l = 0.9$ for all datasets and varying n_l for different datasets depending on the number of samples within each class. In addition, granular-balls with high quality are selected to represent the distribution by new purity limit p_t and sample count limit n_t . We set $p_t = 1$ for all datasets and set varying n_l based on the dataset.

Main Results

Table 2 displays the results of open intent classification on three datasets (StackOverflow, SNIPS, and BANKING) at different proportions of known classes (25%, 50%, and 75%). We report the accuracy of all intent classes (denoted as ‘‘Acc’’) and F1-score for all intent classes, unknown, and

		StackOverflow				SNIPS				BANKING			
Methods		Acc	F1-All	F1-U	F1-K	Acc	F1-All	F1-U	F1-K	Acc	F1-All	F1-U	F1-K
25%	MSP†	28.67	37.85	13.03	42.82	28.57	37.65	0.00	56.48	43.67	50.09	41.43	50.55
	DOC†	42.74	47.73	41.25	49.02	45.63	55.10	34.67	65.32	56.99	58.03	61.42	57.85
	OpenMax†	40.28	45.98	49.29	52.60	59.57	49.68	61.92	43.56	49.94	54.14	51.32	54.23
	DeepUnk*	70.68	65.57	36.87	47.39	53.42	59.67	49.32	64.85	70.68	65.57	76.98	64.97
	ADB†	86.72	80.83	90.88	78.82	58.57	65.72	58.92	69.12	78.85	71.62	84.56	70.94
	DA-ADB*	87.01	80.05	91.16	77.82	47.09	52.75	44.38	56.93	78.39	70.67	84.27	69.95
	MOGB	91.48	84.43	94.42	87.80	68.00	72.20	71.36	72.63	83.08	75.19	88.29	74.50
50%	MSP†	52.42	63.01	23.99	66.91	57.97	63.34	7.12	77.40	59.73	63.60	41.19	71.97
	DOC†	52.53	62.84	25.44	66.58	72.50	78.15	54.89	83.97	64.81	73.12	55.14	73.59
	OpenMax†	60.35	68.18	45.00	70.49	59.82	65.50	14.23	78.31	65.31	74.24	54.33	74.76
	DeepUnk*	71.01	75.41	35.80	67.67	69.53	75.10	49.67	81.45	71.01	75.41	67.80	75.61
	ADB†	86.40	85.83	87.34	85.68	73.57	78.82	62.17	82.98	78.86	80.90	78.44	80.96
	DA-ADB*	86.02	85.21	87.21	85.01	73.69	77.35	66.58	80.05	79.09	81.03	78.78	81.09
	MOGB	88.67	87.49	89.71	87.27	73.86	78.91	62.42	83.03	80.58	81.52	81.04	81.53
75%	MSP†	72.17	77.95	33.96	80.88	71.49	72.56	6.52	85.77	75.89	83.60	39.23	84.36
	DOC†	68.91	75.06	16.76	78.95	79.96	83.11	51.67	89.40	76.77	83.34	50.60	83.91
	OpenMax†	74.42	79.78	44.87	82.11	72.20	73.63	11.03	86.15	77.45	84.07	50.85	84.64
	DeepUnk*	71.56	77.63	34.38	77.63	78.50	81.74	52.05	87.68	74.73	81.12	50.57	81.65
	ADB†	82.78	85.99	73.86	86.80	86.71	88.61	73.35	91.66	81.08	85.96	66.47	86.09
	DA-ADB*	82.89	86.11	74.06	86.92	81.57	84.29	68.62	87.42	81.31	86.01	67.22	86.36
	MOGB	84.37	87.37	75.52	88.16	87.71	89.39	73.62	92.55	82.69	86.82	71.27	87.09

Table 2: Results of MOGB across StackOverflow, SNIPS, and BANKING with different known class ratios (25%, 50%, and 75%). * means the results are not provided in the original paper and we get the results by running its released codes. The results with † are from (Zhang, Xu, and Lin 2021).

known classes (denoted as “F1-All”, “F1-U”, and “F1-K”, respectively). The most optimal performance is emphasized through the **bold** values, while the **underline** values indicate a suboptimal performance of our MOGB.

According to Acc and F1-All, our MOGB methodology exhibits superior performance across all three datasets, consistently surpassing all other baseline methods. Moreover, we observed that our MOGB approach surpasses ADB and DA-ADB by a greater margin when the known class ratio is low. Conversely, when the known class ratio is high, the enhancement in performance of our method is somewhat minimal. Specifically, in cases where the proportion of the known class is low (25%), compared with the strong baseline ADB, our MOGB improves four metrics (“Acc” / “F1-All” / “F1-U” / “F1-K”) by 4.76% / 3.60% / 3.54% / 8.98% on StackOverflow, by 9.43% / 6.48% / 12.44% / 3.51% on SNIPS, and by 4.23% / 3.57% / 3.73% / 3.56% on BANKING. Conversely, when the known class ratio is higher (75%), compared with ADB, our MOGB method enhances four metrics by 1.59% / 1.38% / 1.66% / 1.36% on StackOverflow, by 1.00% / 0.78% / 0.27% / 0.89% on SNIPS and by 1.61% / 0.86% / 4.80% / 1.00% on BANKING.

These above results indicate that our MOGB is more effective in situations with a high proportion of unknown intents. This can be ascribed to the fact that the test dataset includes more open intents under a low known class ratio. Compared with the single decision boundary methods (ADB and DA-ADB), the multi-granularity decision boundaries in MOGB allocate more space for open intents. Therefore, the

open space risk associated with intra-open intents is reduced effectively, resulting in enhanced performance.

Ablation Study

In this subsection, we examine the impact of two components of MOGB on three datasets with a known class ratio of 25%. From Table 3, we observed that removing any component will lead to performance degradation, emphasizing the essence of each independent component. Specifically, (1) w/o HRL refers to removing the Hierarchical Representation Learning (HRL) module and learning representation by cross-entropy loss; (2) w/o MB refers to removing the Multi-granularity decision Boundary (MB) and classification by a single boundary of each class; (3) w/o HRL+MB indicates the exclusion of both Hierarchical Representation Learning and Multi-granularity decision Boundary. The combination of HRL and MB can achieve the best performance.

Discussion

In the MOGB method, the multi-granularity decision boundaries are established based on granular-balls derived from the training dataset. Consequently, the quality of these granular-balls has a significant impact on the classification performance. The quality of a granular-ball largely depends on its purity and the sample count. We use granular-balls with purity exceeding p_t and sample count more than n_t to build the decision boundaries. Next, we report the effect of p_t and n_t .

Dataset	Method	Acc	F1-All
StackOverflow	MOGB	91.48	84.43
	w/o HRL	85.70	79.72
	w/o MB	82.38	75.82
	w/o HRL+MB	86.30	80.06
SNIPS	MOGB	68.00	72.20
	w/o HRL	62.57	68.94
	w/o MB	62.57	68.88
	w/o HRL+MB	60.29	67.54
BANKING	MOGB	83.08	75.19
	w/o HRL	77.40	71.55
	w/o MB	81.14	74.25
	w/o HRL+MB	78.80	72.20

Table 3: Ablation study on StackOverflow, SNIPS, and BANKING with a known class ratio of 25%.

p_t	Acc	F1-All	F1-U	F1-K
0.80	83.10	86.41	72.70	87.32
0.85	83.27	86.50	73.15	87.39
0.90	84.37	87.37	75.52	88.16
0.93	84.45	87.41	75.78	88.19
0.95	84.23	87.31	75.40	88.10
0.97	84.23	87.31	75.40	88.10

Table 4: Effect of p_t on StackOverflow with known class ratio of 75%.

Effect of the Purity Limit p_t

A higher p_t results in a small number of trustworthy granular-balls being used to represent the dataset and fewer decision boundaries being constructed, reducing the known intent space and increasing the chance of misclassifying known intents as unknowns. Conversely, a lower p_t expands the known space, reducing open space but increasing the risk of misclassifying unknown intents as known classes.

Table 4 presents the Acc, F1-All, F1-U, and F1-K as the purity threshold p_t increases from 0.80 to 0.97 on the StackOverflow dataset, with a known class ratio of 75%. Overall, the performance of our method in open classification remains relatively stable, highlighting its robustness. However, a slight trend in performance variation can be observed: as p_t increases, classification performance initially improves, peaking at $p_t = 0.93$. This improvement occurs because high-purity granular-balls better capture the distribution of known classes by reducing uncertainty. Beyond this point, further increasing p_t to 0.95 causes a performance decline. The stricter purity requirement excludes certain decision boundaries that represent known classes, leading to portions of known class spaces being misclassified as unknowns and increasing open space risk. This outcome is consistent with our expectations.

Effect of the Sample Count Limit n_t

Granular-balls containing samples exceeding a specified threshold n_t are chosen to represent data distribution and es-

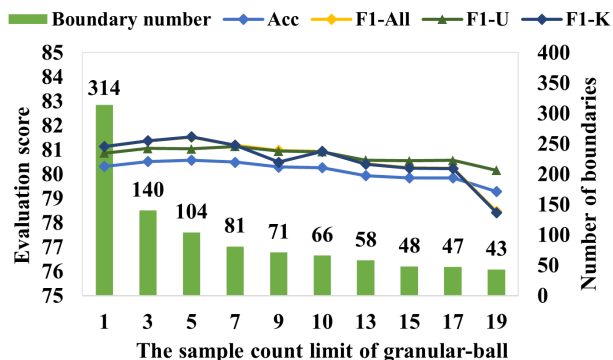


Figure 3: Effect of n_t on BANKING with 50% known class. X-axis represents the value of n_t , the left Y-axis denotes the values of four metrics, and the right Y-axis indicates the number of established decision boundaries for all known classes.

tablish decision boundaries for the open classification. Figure 3 shows the performance of MOGB on dataset BANKING with a known class ratio of 50% as n_t changes, along with the number of established decision boundaries at different n_t values. There are 39 known classes under a known class ratio of 50% on the BANKING dataset. From Figure 3, we observed that the number of decision boundaries for open classification decreases as n_t increases. In particular, when $n_t = 1$, a total of 314 decision boundaries are established, signifying that each known intent class is characterized by several distinct boundaries. Optimal performance is achieved at $n_t = 5$, where 104 decision boundaries are constructed. However, at $n_t = 19$, only 43 decision boundaries are constructed, with most classes represented by a single boundary, degrading the performance of four metrics.

The observed phenomenon aligns with our expectations. A strict n_t limit reduces decision boundaries, leading to incomplete feature space coverage and higher empirical risk. In contrast, a looser n_t limit introduces more boundaries that better fit the true distribution but risk overcovering, ultimately reducing performance. Although the performance of MOGB varies with changes in n_t , the overall fluctuation remains small, demonstrating the robustness of our method.

Conclusions

This paper proposes a multi-granularity open intent classification method via hierarchical representation learning and multi-granularity decision boundary (MOGB). Specifically, the granular-balls of diverse sizes are generated by adaptive granular-ball clustering to represent the known intent space during hierarchical representation learning. Additionally, the nearest sub-centroid classifier equips for learning fine-grained representation that reflects semantic structures within known intent classes. Furthermore, multi-granularity decision boundaries for each known intent class are constructed for open intent classification, which reduces both empirical and open-space risk. Finally, extensive experiments demonstrate the superiority of the MOGB method.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 62476228, 62406259, 62222601, 62450043, 62176033), the Sichuan Science and Technology Program (Nos. 2024ZYD0180, 2024YFHZ0024, MZGC20240153), and the Graduate Representative Achievement Cultivation Project of Southwest University of Finance and Economics (Nos. JGS2024066, JGS2024068).

References

- Bendale, A.; and Boulton, T. E. 2016. Towards open set deep networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1563–1572.
- Cao, X.; Yang, X.; Xia, S.; Wang, G.; and Li, T. 2024. Open Continual Feature Selection via Granular-Ball Knowledge Transfer. *IEEE Transactions on Knowledge and Data Engineering*, 36(12): 8967–8980.
- Casanueva, I.; Temčinas, T.; Gerz, D.; Henderson, M.; and Vulić, I. 2020. Efficient intent detection with dual sentence encoders. In *Proceedings of the Workshop on Natural Language Processing for Conversational AI*, 38–45.
- Chen, L. 1982. Topological structure in visual perception. *Science*, 218(4573): 699–700.
- Cheng, Z.; Jiang, Z.; Yin, Y.; Wang, C.; and Gu, Q. 2022. Learning to classify open intent via soft labeling and manifold mixup. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30: 635–645.
- Coucke, A.; Saade, A.; Ball, A.; Bluche, T.; Caulier, A.; Leroy, D.; Doumouro, C.; Gisselbrecht, T.; Caltagirone, F.; Lavril, T.; et al. 2018. Snips voice platform: an embedded spoken language understanding system for private-by-design voice interfaces. arXiv:1805.10190.
- Guo, C.; Pleiss, G.; Sun, Y.; and Weinberger, K. Q. 2017. On calibration of modern neural networks. In *International Conference on Machine Learning*, 1321–1330.
- Hendrycks, D.; and Gimpel, K. 2022. A Baseline for Detecting Misclassified and Out-of-Distribution Examples in Neural Networks. In *International Conference on Learning Representations*.
- Jain, L. P.; Scheirer, W. J.; and Boulton, T. E. 2014. Multi-class open set recognition using probability of inclusion. In *Computer Vision—ECCV*, 393–409. Springer.
- Kenton, J. D. M. W. C.; and Toutanova, L. K. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, volume 1, 2.
- Lin, T. E.; and Xu, H. 2019. Deep unknown Intent Detection with Margin Loss. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 5491–5496.
- Liu, J.; Jianye, H.; Ma, Y.; and Xia, S. 2024. Unlock the Cognitive Generalization of Deep Reinforcement Learning via Granular Ball Representation. In *International Conference on Machine Learning*.
- Liu, X. K.; Li, J. Q.; Mu, J. J.; Yang, M.; Xu, R. F.; and Wang, B. Y. 2023. Effective open intent classification with K-center contrastive learning and adjustable decision boundary. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 13291–13299.
- Parmar, J.; Chouhan, S.; Raychoudhury, V.; and Rathore, S. 2023. Open-world machine learning: applications, challenges, and opportunities. *ACM Computing Surveys*, 55(10): 1–37.
- Quadir, A.; and Tanveer, M. 2024. Granular ball twin support vector machine with pinball loss function. *IEEE Transactions on Computational Social Systems*, 1–10.
- Shu, L.; Xu, H.; and Liu, B. 2017. DOC: Deep open classification of text documents. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2911–2916.
- Wang, G. Y. 2017. DGCC: data-driven granular cognitive computing. *Granular Computing*, 2(4): 343–355.
- Wang, Z.; Li, J.; Xia, S.; Lin, L.; and Wang, G. 2024a. Text adversarial defense via granular-ball sample enhancement. In *Proceedings of the International Conference on Multimedia Retrieval*, 348–356.
- Wang, Z. L.; Zhang, T.; Xia, S. Y.; Lin, L. L.; and Wang, G. Y. 2024b. GBRAIN: Combating textual label noise by granular-ball based robust training. In *Proceedings of the International Conference on Multimedia Retrieval*, 357–365.
- Xia, D.; Wang, G.; Zhang, Q.; Yang, J.; and Xia, S. 2024a. Three-way approximations fusion with granular-ball computing to guide multi-granularity fuzzy entropy for feature selection. *IEEE Transactions on Fuzzy Systems*, 32(10): 5963–5977.
- Xia, S.; Lian, X.; Wang, G.; Gao, X.; Chen, J.; and Peng, X. 2024b. Gbsvm: an efficient and robust support vector machine framework via granular-ball computing. *IEEE Transactions on Neural Networks and Learning Systems*, 1–15.
- Xia, S.; Zhang, H.; Li, W.; Wang, G.; Giem, E.; and Chen, Z. 2020. GBNRS: A novel rough set algorithm for fast adaptive attribute reduction in classification. *IEEE Transactions on Knowledge and Data Engineering*, 34(3): 1231–1242.
- Xia, S.; Zheng, S.; Wang, G.; Gao, X.; and Wang, B. 2021. Granular ball sampling for noisy label classification or imbalanced classification. *IEEE Transactions on Neural Networks and Learning Systems*, 34(4): 2144–2155.
- Xia, S. Y.; Liu, Y. S.; Ding, X.; Wang, G. Y.; Yu, H.; and Luo, Y. G. 2019. Granular ball computing classifiers for efficient, scalable and robust learning. *Information Sciences*, 483: 136–152.
- Xia, S. Y.; Wang, C.; Wang, G. Y.; Gao, X. B.; Ding, W. P.; Yu, J. H.; Zhai, Y. J.; and Chen, Z. Z. 2023. GBRs: A unified granular-ball learning model of pawlak rough set and neighborhood rough set. *IEEE Transactions on Neural Networks and Learning Systems*, 1–15.
- Xie, J.; Hua, C.; Xia, S.; Cheng, Y.; Wang, G.; and Gao, X. 2024a. W-GBC: An Adaptive Weighted Clustering Method Based on Granular-Ball Structure. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*, 914–925.

- Xie, J.; Xiang, X.; Xia, S.; Jiang, L.; Wang, G.; and Gao, X. 2024b. MGNR: A multi-granularity neighbor relationship and its application in KNN classification and clustering methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12): 7956–7972.
- Xie, Q.; Zhang, Q.; Xia, S.; Zhao, F.; Wu, C.; Wang, G.; and Ding, W. 2024c. Gbg++: A fast and stable granular ball generation method for classification. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 8(2): 2022–2036.
- Xu, J. M.; Wang, P.; Tian, G. H.; Xu, B.; Zhao, J.; Wang, F. Y.; and Hao, H. W. 2015. Short text clustering via convolutional neural networks. In *Proceedings of the Workshop on Vector Space Modeling for Natural Language Processing*, 62–69.
- Yan, G.; Fan, L.; Li, Q.; Liu, H.; Zhang, X.; Wu, X.-M.; and Lam, A. Y. 2020. Unknown intent detection using Gaussian mixture model with an application to zero-shot intent classification. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 1050–1060.
- Yang, J.; Liu, Z.; Xia, S.; Wang, G.; Zhang, Q.; Li, S.; and Xu, T. 2024. 3WC-GBNRS++: A novel three-way classifier with granular-ball neighborhood rough sets based on uncertainty. *IEEE Transactions on Fuzzy Systems*, 32(8): 4376–4387.
- Yang, X.; Li, Y. J.; Meng, D.; Yang, Y. X.; Liu, D.; and Li, T. R. 2022. Three-way multi-granularity learning towards open topic classification. *Information Sciences*, 585: 41–57.
- Zhan, L. M.; Liang, H. W.; Liu, B.; Fan, L.; Wu, X. M.; and Lam, A. Y. 2021. Out-of-scope intent detection with self-supervision and discriminative training. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing*, 3521–3532.
- Zhang, H.; Xu, H.; and Lin, T.-E. 2021. Deep open intent classification with adaptive decision boundary. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 14374–14382.
- Zhang, H.; Xu, H.; Lin, T.-E.; and Lyu, R. 2021. Discovering new intents with deep aligned clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 14365–14373.
- Zhang, H.; Xu, H.; Zhao, S.; and Zhou, Q. 2023a. Learning discriminative representations and decision boundaries for open intent detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 31: 1611–1623.
- Zhang, Q.; Wu, C.; Xia, S.; Zhao, F.; Gao, M.; Cheng, Y.; and Wang, G. 2023b. Incremental learning based on granular ball rough sets for classification in dynamic mixed-type decision system. *IEEE Transactions on Knowledge and Data Engineering*, 35(9): 9319–9332.
- Zheng, Y. H.; Chen, G. Y.; and Huang, M. L. 2020. Out-of-domain detection for natural language understanding in dialog systems. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28: 1198–1209.
- Zhou, W. X.; Liu, F. Y.; and Chen, M. H. 2021. Contrastive out-of-distribution detection for pretrained transformers. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 1100–1111.
- Zhou, Y.; Liu, P.; and Qiu, X. 2022. KNN-contrastive learning for out-of-domain intent classification. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 5129–5141.