One-Class Learning for Graph Neural Networks

Marcos Paulo Silva Gôlo[1] **and Ricardo Marcondes Marcacini**[1]

[1]University of São Paulo (ICMC-USP) - São Carlos, São Paulo, Brazil

marcosgolo@usp.br, ricardo.marcacini@icmc.usp.br

Abstract. *One-Class Learning (OCL) for graphs addresses problems in which only nodes from a class of interest are available for training, while the remaining graph nodes space is broad, unknown, or continuously evolving. Although Graph Neural Networks (GNNs) provide a natural mechanism for exploiting relational information, existing one-class approaches still present limitations regarding end-to-end optimization, interpretability, graph streams, edge-level prediction, and heterogeneous structures. This thesis introduces the OLGA family, a set of end-to-end, one-class, GNN-interpretable methods. OLGA establishes the methodological core by combining a novel sphere-based loss, which continuously compacts interest nodes in a low-dimensional latent space, with graph autoencoder reconstruction. OLGASTREAM extends the thesis to stream graphs by weighting the losses and employing FAConv to mitigate concept drift. ECHOLGA extends the thesis to causal edge classification by transforming relations into nodes of a heterogeneous hypergraph. FOLGA addresses heterogeneous node classification through topology rewiring and layer-wise fusion. Experiments on 19 datasets from textual, visual, tabular, streaming, causal, and heterogeneous domains show that the proposed methods achieved the best f_1-macro in 14. Overall, the thesis advances the field of OCL for GNNs.*

Institution where title was obtained: Institute of Mathematics and Computer Sciences (ICMC) - University of São Paulo (USP).

Advisor: Ricardo Marcondes Marcacini, ICMC/USP, São Paulo, Brazil.

Track: PhD.

Date of defense: 30/04/2026.

Examining board members (Prof. Dr.):

1. Luis Gustavo Nonato, ICMC/USP, São Paulo, Brazil.
2. Murillo Guimarães Carneiro, UFU, Minas Gerais, Brazil
3. Daniel Carlos Guimarães Pedronette, UNESP, São Paulo, Brazil.

Link to the full thesis: https://bit.ly/tese-golo-usp

Link to Google Scholar profile: https://bit.ly/scholar-golo

Link to Lattes CV: http://lattes.cnpq.br/7402825146038143

Links to the publications derived from the thesis: [Gôlo et al. 2022, da Silva et al. 2022, Coutinho et al. 2023, Gôlo et al. 2023a, Gôlo et al. 2023b, Gôlo et al. 2024a, Gôlo et al. 2024b, Gôlo et al. 2024c, Gôlo et al. 2024d, de Souza et al. 2024, Gôlo and Marcacini 2024, Filho et al. 2025, do Carmo et al. 2025, Marx et al. 2025, Gôlo and Marcacini 2025, Gôlo et al. 2025a, Gôlo et al. 2025b, Gôlo and Utino 2025, Miranda et al. 2025, Utino and Gôlo 2025, Gôlo et al. 2026]

1. Introduction

Graphs are a practical way to model real-world systems because they explicitly represent relationships rather than treating each data point in isolation [Khemani et al. 2024]. In a graph, nodes represent entities and edges represent the relationships or interactions between them [Diestel 2025]. For instance, in a social network, nodes can be users and edges can be friendships or follows; in a brain graph, a graph of neurons connected by synapses; in a biological network, nodes can be proteins and edges can represent interactions [Veličković 2023]. This entity–relation view captures dependencies and influence: who connects to whom, what interacts with what, which is often exactly what drives prediction. Heterogeneous graphs go a step further by modeling multiple entity and relation types within a single structure, which naturally leads to multiple node and edge types [Bing et al. 2023]. In practice, heterogeneity is how some real systems actually work: data is not a single table, but a collection of connected objects with different roles and semantics [Wang et al. 2022]. Thus, heterogeneous graphs enrich the signal available for learning, allowing models to exploit both within-type patterns and cross-type interactions [Wang et al. 2022]. Many applications rely on this structure, including event sensing (where events are linked to people, locations, and topics) and causal discovery, where relations are the object of discovery between heterogeneous events [Bing et al. 2023].

Most graph classification problems are modeled in the usual multi-class way: each node belongs to one of several categories [Wu et al. 2020]. However, many real scenarios are not driven by multi-class learning, but by identifying a class of interest defined by a user, company, or organization [Seliya et al. 2021]. One-Class Learning (OCL) is designed for exactly this supervision pattern where the goal is to determine whether an instance falls within the scope of interest [Tax 2001, Emmert-Streib and Dehmer 2022]. The idea is to learn only from instances of the interest class and then predict whether a new instance belongs to that class [Zadeh 2024]. This is particularly suitable for open-domain settings, where the non-interest space is large, diverse, constantly changing, or simply unknown in advance, and where the goal is inherently focused on one class [Zadeh 2024]. Typical applications include fake news detection [de Souza et al. 2024], event sensing [Nguyen and Grishman 2018], hit song prediction [da Silva et al. 2022], fraud detection [Dou et al. 2020], and anomaly detection [Fan et al. 2024].

Graph Neural Networks (GNNs) arise as one of the main tools because they learn representations by repeatedly aggregating information from a node's neighbors and combining it through neural transformations, allowing the model to capture patterns that depend on connectivity and context [Khemani et al. 2024, Gkarmpounis et al. 2024]. The literature on OCL for GNNs can be divided into three families of approaches. Some methods are unsupervised and focus on anomaly detection, flagging unusual patterns, and often capturing the class of interest indirectly [Feng et al. 2022, Deldar et al. 2022, Hu et al. 2023, Chen et al. 2025]. Another line follows a two-step pipeline. First, an unsupervised GNN learns node embeddings, then a separate one-class algorithm is applied to detect the interest class [Huang et al. 2022, Ganz et al. 2023, Shi et al. 2023]. More recent approaches are end-to-end. They learn representations and the one-class decision jointly, directly optimizing for compactness and separation of the interest class, frequently using hypersphere-based objectives to shape the learned space around the interest class [Zhang et al. 2022a, Zhang et al. 2022b, Fan et al. 2024, Li et al. 2024].

2. Related Work, Research Gaps, and Challenges

Unsupervised GNN-based methods. Unsupervised GNN-based methods typically learn node representations by reconstructing the graph structure, node attributes, or multiple structural views. In this family, the model typically produces a reconstruction score, and the final decision is made by applying a threshold to this score. [Feng et al. 2022] and [Deldar et al. 2022] explored graph autoencoders and reconstruction-based anomaly scores, while SAMCL [Hu et al. 2023] introduced a multiview contrastive framework with reconstruction objectives. More recently, FIAD [Chen et al. 2025] injected perturbations into latent representations and reconstructed both topology and attributes.

Two-step one-class GNN methods. Two-step methods separate representation learning from one-class learning. In the first step, a GNN is trained to learn node embeddings, usually through graph reconstruction. In the second step, an external one-class classifier, such as OCSVM or DeepSVDD, is applied. [Huang et al. 2022] proposed a one-class directed heterogeneous GNN combined with DeepSVDD. [Ganz et al. 2023] used collaborative software graphs and variational graph autoencoders followed by DeepSVDD to detect backdoor software. [Shi et al. 2023] proposed EAGLE, a heterogeneous graph approach for tax evasion detection based on contrastive learning followed by OCSVM.

End-to-end one-class GNN methods. End-to-end methods jointly optimize representation learning and the one-class objective in a single training process. These methods directly shape the latent space according to the one-class decision criterion. OCGNN [Wang et al. 2021] introduced a DeepSVDD-inspired hypersphere loss to compact interest nodes in the latent space, while OCTGNN [Huang et al. 2021] extended this idea to dynamic graphs by incorporating temporal information. Later, DUALSVDAE [Zhang et al. 2022a] and STEAM [Zhang et al. 2022b] combined SVDD-style losses with graph and feature autoencoders, adding reconstruction constraints. JAANE [Fan et al. 2024] further integrated structural and attribute embeddings with adversarial regularization, reconstruction, and SVDD objectives. MULDUALGNN [Li et al. 2024] combined global heterogeneous structure with learned local views and multiple hypersphere objectives to capture multimodal patterns in the interest class.

Limitations, Research Gaps, and Challenges: Despite recent advances, one-class learning with GNNs still presents limitations, gaps, and important open challenges. Many studies evaluate their methods using multi-class datasets adapted to one-class scenarios or by injecting artificial anomalies. Unsupervised anomaly-detection approaches based on reconstruction errors usually depend on threshold selection, which can be arbitrary and unstable. Two-step methods learn embeddings independently from the final one-class objective, producing representations that may not be optimal for the downstream classifier. End-to-end methods reduce this limitation but rely on the DeepSVDD hypersphere loss, which may stop improving once interest nodes are within the hypersphere. Moreover, combining multiple losses remains challenging due to different objectives and scales, even in the OCL scenario of one label to train. A major gap is interpretability, since most methods rely on high-dimensional latent spaces or post-hoc projections. Finally, graph streams, edge-level one-class classification, and heterogeneous graphs remain underexplored in end-to-end OCL settings, especially when the graph evolves over time, the prediction target is an observed relation rather than a node, or multiple node and edge types must be explicitly modeled under one-class supervision.

3. The OLGA Family

This thesis proposes a family of end-to-end methods for One-Class Learning with Graph Neural Networks. The central idea is to align graph representation learning with the One-Class Learning (OCL) objective. The proposal is developed through four main methods: OLGA, OLGASTREAM, ECHOLGA, and FOLGA. OLGA establishes the methodological core of the thesis; OLGASTREAM extends this core to data streams; ECHOLGA extends it to edge classification in heterogeneous hypergraphs; and FOLGA explores heterogeneous graphs by fusing multiple rewired graphs.

The thesis begins with **OLGA**, a method proposed for node classification in homogeneous graphs. Let $G = (V, E)$ be a homogeneous graph, where V is the set of nodes, E is the set of edges, and $\boldsymbol{X} \in \mathbb{R}^{|V| \times d}$ is the node feature matrix. In the OCL setting, only a subset $V_I \subset V$ is labeled as positive, i.e., as belonging to the class of interest. The goal is to learn a function capable of deciding whether a node $v_i \in V$ belongs to the class of interest or not. A GNN learns node representations through successive neighborhood aggregation steps. Considering $\boldsymbol{H}^{(0)} = \boldsymbol{X}$ and $\boldsymbol{A}$ as the adjacency matrix (representing E), the propagation rule at layer t can be written as:

$$\boldsymbol{H}^{(t+1)} = gnn(\boldsymbol{H}^{(t)}, \boldsymbol{A}; \boldsymbol{W}^{(t)}), \tag{1}$$

where $W^{(t)}$ denotes the trainable parameters of the layer. In OLGA, this process is instantiated using Graph Convolutional Network (GCN) layers, defined as: $\boldsymbol{h}_{v_i}^{t} = \sigma(\boldsymbol{W}^{(t)} \cdot MEAN\{\boldsymbol{h}_{v_j}^{t-1} : v_j \in \{N_{v_i} \cup v_i\}\})$, where N_{v_i} represents the neighborhood of node v_i and σ is an activation function. At the end of the encoder, the model produces a latent representation matrix $\boldsymbol{Z}$, in which each row $\boldsymbol{z}_{v_i}$ represents a node in the graph. OLGA adopts the structure of a Graph Autoencoder (GAE), in which a GNN serves as the encoder and the adjacency matrix is reconstructed by an inner-product decoder (Equation 2). OLGA explores the reconstruction of the adjacency matrix (Equation 3).

$$GAE = \begin{cases} Encoder : \boldsymbol{Z} = gnn(\boldsymbol{X}, \boldsymbol{A}; \boldsymbol{W}) \\ Decoder : \hat{\boldsymbol{A}} = \sigma\big(\boldsymbol{Z}\boldsymbol{Z}^{\top}\big) \end{cases}, \tag{2}$$

$$\mathcal{L}_{rec}(W) = -\sum_{i=1}^{|V|}\sum_{j=1}^{|V|}\Big(A_{ij}\log\hat{A}_{ij} + (1 - A_{ij})\log(1 - \hat{A}_{ij})\Big). \tag{3}$$

In addition to reconstruction, OLGA proposes a new sphere-based loss for OCL. The proposed loss continues to encourage positive examples to move toward the center. Let $\boldsymbol{c}$ be the sphere center and R its radius. For each interest node $v_i \in V_I$, we define $o_i = \|\boldsymbol{z}_{v_i} - \boldsymbol{c}\|_2^2 - R^2$. Negative values indicate that the node is inside the sphere, whereas positive values indicate that it is outside. The proposed OCL loss is defined as:

$$\mathcal{L}_{\text{ocl}}(\boldsymbol{W}) = \frac{1}{|\mathcal{V}_I|}\sum_{i=1}^{|\mathcal{V}_I|}\begin{cases} o_i + 1, & \text{if } o_i > 0, \\ \exp(o_i), & \text{otherwise.} \end{cases} \tag{4}$$

This formulation strongly penalizes interest nodes outside the sphere while preserving a nonzero gradient for those already inside it. Thus, the model continues to adjust

the representations to bring positive examples closer to the center while preserving intra-class similarity. The final objective function of OLGA is: $\mathcal{L}_{OLGA} = \mathcal{L}_{ocl} + \mathcal{L}_{rec}$. This combination allows the reconstruction loss to act as a structural constraint over the sphere loss. While $\mathcal{L}_{ocl}$ forces the interest nodes to concentrate in a compact region of the latent space, $\mathcal{L}_{rec}$ preserves the graph structure and helps prevent all nodes from being improperly collapsed into the sphere. Moreover, OLGA learns low-dimensional representations, especially in 2D or 3D, making it possible to visually interpret the learning process and the decision boundary induced by the sphere.

OLGASTREAM extends the thesis to data stream scenarios, in which instances arrive continuously, and the data distribution may change over time. In this setting, the graph is constructed over successive windows or iterations. At each iteration, the model uses previously observed examples from the class of interest and receives new nodes to be classified as either of interest or not of interest. The method preserves the central principles of OLGA: learning node representations with GNNs, reconstructing the graph structure via a graph autoencoder objective, and compacting interest examples into a low-dimensional sphere. However, data streams introduce an additional challenge: *concept drift*. Concept drift may alter the neighborhood composition and increase heterophilic connections. In traditional GNN layers, this can be harmful because messages from non-interest nodes may be aggregated into interest-node representations. OLGASTREAM therefore employs a Frequency Adaptive Convolutional layer (FAConv), whose signed adaptive coefficients can reduce the influence of neighbors with dissimilar representations. FAConv is used as a message-passing mechanism intended to improve robustness to the resulting heterophily. Formally, FAConv updates the representation of v_i as follows:

$$\boldsymbol{h}_{v_i}^{t} = \epsilon \cdot \boldsymbol{h}_{v_i}^{0} + \sum_{v_j \in N_{v_i}} \frac{\alpha_{i,j}}{\sqrt{d_i d_j}} \boldsymbol{h}_{v_j}^{t-1} \qquad (5) \qquad\qquad \alpha_{i,j} = \tanh\left(\boldsymbol{a}^{T}[\boldsymbol{h}_{v_i}, \boldsymbol{h}_{v_j}]\right) \qquad (6)$$

where $\boldsymbol{h}_{v_i}^{0}$ is the initial representation of node v_i, ϵ controls the preservation of the initial representation, d_i and d_j are node degrees, and $\alpha_{i,j}$ is an adaptive coefficient that controls the influence of node v_j on node v_i. Thus, FAConv makes OLGASTREAM more suitable for streaming scenarios where concept drift may induce connections between semantically distinct nodes. OLGASTREAM modifies the loss combination strategy used in OLGA. Since the relevance of structural reconstruction and one-class compactness may vary over time, OLGASTREAM introduces a weighting parameter λ_{ocl} to control the contribution of each objective: $\mathcal{L}_{OLGASTREAM} = \lambda_{ocl}\mathcal{L}_{ocl} + (1 - \lambda_{ocl})\mathcal{L}_{rec}$.

OLGASTREAM extends OLGA in two main directions. First, it replaces the GCNConv with a FAConv, thereby making the method more robust to heterophilic patterns arising from concept drift. Second, it introduces a weighting parameter λ_{ocl}, enabling the model to control the influence of the OCL and reconstruction objectives over time. OLGASTREAM also learns low-dimensional latent representations, preserving interpretability by allowing the one-class decision boundary and the evolution of node embeddings to be directly visualized during training.

ECHOLGA extends the thesis to a different task: edge classification under the OCL setting. The application explored in this thesis is causal detection between event pairs. In this problem, each event is represented by a textual description, and the goal is

to determine whether a directed relation between two events indicates a causal relation. Since causal relations form the class of interest, the task is naturally modeled as one-class learning: the model is trained using only positive causal relations and then classifies unlabeled relations as causal or non-causal. The first step of ECHOLGA is to obtain a directed causal graph, in which nodes represent events and edges represent observed causal or unlabeled relations between event pairs. However, because the prediction target is an edge, the method transforms this original causal graph into a hypergraph, in which the original edges are treated as nodes. More specifically, for each directed edge (v_i, v_j) in the original causal graph, ECHOLGA creates a new node v_o, called an edge-node, and adds two directed edges connecting v_i to v_o and v_o to v_j. Thus, classifying an edge in the original causal graph becomes equivalent to classifying an edge-node in the hypergraph.

ECHOLGA preserves the two central losses of OLGA. The sphere loss $\mathcal{L}_{ocl}$ is applied to the edge-nodes that represent causal relations, encouraging them to move toward the center of the sphere. The reconstruction loss $\mathcal{L}_{rec}$ preserves the structure of the heterogeneous hypergraph. Events are heterogeneous since they are connected with other node types (who, where, when, and what). Thus, ECHOLGA adds a third loss to encourage the separation of node types. This loss is defined as a cross-entropy loss:

$$\mathcal{L}_{nt}(W) = -\sum_{v_i \in V} \sum_{k=1}^{K} y_{ik} \log(\hat{y}_{ik}) \quad (7) \qquad \mathcal{L}_{ECHOLGA} = \lambda_{ocl}\mathcal{L}_{ocl} + \lambda_{rec}\mathcal{L}_{rec} + \lambda_{nt}\mathcal{L}_{nt} \quad (8)$$

where K is the number of node types, y_{ik} indicates the true type of node v_i, and $\hat{y}_{ik}$ is the predicted probability for that type. The third loss introduces an additional constraint on the latent space: in addition to bringing causal relations closer to the sphere center and reconstructing the hypergraph structure, the model must also preserve semantic differences among node types. Thus, ECHOLGA adapts the OLGA core for an edge classification scenario, in which the object of interest is not a traditional node but an edge.

FOLGA extends the thesis to node classification in heterogeneous graphs. In this setting, the class of interest is associated with a primary node type, while the remaining node types provide contextual and relational evidence. Instead of introducing an additional loss term, FOLGA focuses on enriching the graph topology before representation learning. Its main idea is to generate complementary rewired graphs and then fuse the information learned from all of them into a unified latent representation. Let $G = (V, E)$ be the original heterogeneous graph. FOLGA preserves the same node set V and generates multiple rewired graphs by adding new edges according to rewiring strategies. The first strategy creates additional primary-to-primary edges by considering distances in the heterogeneous graph. At a rewiring step ℓ, an edge between two primary nodes v_i and v_j is sampled by: $p_\ell(v_i, v_j) = \frac{1}{d_G(v_i,v_j)^\alpha |V_\tau|^\beta \ell^\gamma}$, where $d_G(v_i, v_j)$ is the distance between the nodes in the original heterogeneous graph, V_τ is the set of primary nodes, and α, β, and γ control the rewiring probability. Since $d_G(v_i, v_j)$ may be computed through paths that traverse secondary nodes, this strategy injects heterogeneous evidence into new primary-to-primary relations.

The second strategy is similarity-based and creates additional relations within node-type partitions. Given the initial node features $\boldsymbol{X}$, FOLGA builds kNN graphs for the primary and secondary nodes. This produces additional edge sets E_τ^{knn} and E_{sec}^{knn}.

In the thesis formulation, this yields a collection of three graphs: $G^{(0)}, G^{(1)}, G^{(2)} = G, (V, E \cup E_{\tau}^{knn}), (V, E \cup E_{sec}^{knn})$. The first graph is the original heterogeneous graph. The second enriches the topology with similarity relations among primary nodes. The third enriches the topology with similarity relations among secondary nodes. These complementary graphs allow message passing to exploit both the original schema and additional semantic similarities.

Given the rewired graphs, FOLGA introduces a fusion-rewired GNN encoder. Let ${G^{(\ell)}}_{\ell=0}^{L-1}$ be the list of graphs produced by the rewiring procedure, with adjacency matrices ${\boldsymbol{A}^{(\ell)}}_{\ell=0}^{L-1}$. Differently from a simple late fusion of final embeddings, FOLGA performs fusion inside each GNN layer. At each layer t, the model applies a GNN sublayer to each rewired graph and sums their outputs:

$$\boldsymbol{H}^{(t+1)} = folga\left(\sum_{\ell=0}^{L-1} gnn\Big(\boldsymbol{H}^{(t)}, \boldsymbol{A}^{(\ell)}; \boldsymbol{W}^{(t,\ell)}\Big)\right). \tag{9}$$

Each adjacency matrix $\boldsymbol{A}^{(\ell)}$ defines a different message-passing structure, and each sublayer has its own trainable parameters $\boldsymbol{W}^{(t,\ell)}$. $folga(\cdot)$ represents the activation and transformation applied after adding the outputs from the rewired graphs. Thus, information from all graphs is fused progressively across layers, rather than only after the final representation is produced. FOLGA is then instantiated as a graph autoencoder, where the encoder receives the node features and the list of rewired adjacency matrices, and the decoder reconstructs the original graph structure using an inner product over the learned representations: $\boldsymbol{Z} = folga\left(\boldsymbol{X}, \{\boldsymbol{A}^{(\ell)}\}_{\ell=0}^{L-1}\right)$. After obtaining $\boldsymbol{Z}$, FOLGA uses the same learning core introduced by OLGA with the 3D embedding space for interpretability: $\mathcal{L}_{FOLGA} = \mathcal{L}_{ocl} + \mathcal{L}_{rec}$. Its main contribution lies in the topology and encoder: by generating rewired graphs and fusing them at each encoder layer, FOLGA exploits heterogeneity without requiring specialized heterogeneous GNN layers or additional loss components.

4. Results

Table 1 summarizes the performance of OLGA against two-step baselines and end-to-end one-class GNN methods. The comparison uses the average f_1-macro over 10 folds, considering graph datasets from different domains, including text, image, and tabular data. The results show that OLGA achieved the best performance in four out of eight datasets and the second-best performance in most of the remaining cases. This behavior indicates that combining the sphere-based OCL loss with graph reconstruction yields representations better aligned with the class-of-interest objective than purely reconstruction-based or two-step strategies. Although DeepWalk and OCGAT are competitive on some datasets, OLGA exhibits more consistent behavior across different data types. These results support the thesis's main hypothesis: graph reconstruction serves as a structural constraint, while the OCL loss shapes the latent space toward a compact interest class.

Table 2 presents the data stream results for OLGASTREAM and the compared one-class baselines. The evaluation considers four stream datasets and reports the f_1-macro. OLGASTREAM achieved the best results across all datasets, demonstrating the effectiveness of extending OLGA to streaming graphs. This suggests that OLGASTREAM is particularly useful in the one-class stream scenario. The improvements are

Table 1. OLGA results. Best results are in bold, and second best underlined.

References	Baselines (+OCSVM)			[Wang et al. 2021]			OLGA
Datasets	GAE	Deepwalk	Node2Vec	OCGCN	OCGAT	OCSAGE	
FakeNews	**0.942**	0.879	0.829	0.746	0.884	0.824	<u>0.940</u>
Terrorism	0.727	**0.981**	0.800	0.845	0.900	0.785	<u>0.921</u>
Reviews	<u>0.728</u>	0.598	0.644	0.592	0.691	0.623	**0.750**
Food	0.982	0.752	0.960	0.901	<u>0.994</u>	0.985	**0.997**
Pneumonia	0.639	<u>0.878</u>	0.645	0.505	0.752	0.655	**0.914**
Strawberry	0.516	**0.961**	0.584	0.435	0.629	0.635	<u>0.647</u>
Musk	0.623	0.774	0.640	0.666	0.620	0.603	**0.785**
Tuandromd	0.904	0.689	0.905	0.898	**0.978**	<u>0.974</u>	<u>0.974</u>

less expressive in the Bank and Elect. datasets, where all methods present closer results, but the results are more expressive in the Agrawal and Twitter datasets. FAConv helps reduce the negative effect of heterophilic connections caused by concept drift, while λ_{ocl} controls the trade-off between preserving graph structure and compacting the class of interest. Therefore, OLGASTREAM demonstrates that the OLGA core can be adapted to evolving data without losing its one-class and interpretable nature.

Table 2. OLGASTREAM results. Bold for best results and underlined for 2nd best.

Methods	References	Bank	Elect.	Agrawal	Twitter
OCSVM	**Baseline**	<u>0.491</u>	0.590	<u>0.681</u>	<u>0.541</u>
LOF	**Baseline**	0.260	0.408	0.308	0.519
Isolation Forest	**Baseline**	0.421	<u>0.619</u>	0.652	0.386
GraphLOF	**Baseline**	0.229	0.377	0.253	0.519
CMA	[Jodelka et al. 2021]	0.491	0.593	0.641	0.448
FMA	[Jodelka et al. 2021]	0.491	0.568	0.655	0.442
CDD	[Mendelson and Lerner 2020]	0.328	0.390	0.281	0.356
OLGASTREAM	**Proposal**	**0.493**	**0.627**	**0.927**	**0.858**

Table 3 compares ECHOLGA with language-model-based methods and LLM-based approaches through f_1-macro. The task is causal discovery from event pairs, evaluated as an edge-classification problem in the one-class graph setting. ECHOLGA achieved the best result in three out of four datasets: Risk, Headlines, and FinCausal. In FinCausal, ECHOLGA obtains the highest f_1-macro, although its result is close to Gemma 2. In Risk and Headlines, the larger improvements suggest that heterogeneous relational context may provide complementary evidence to pairwise text-based inference. The Twitter dataset is the main exception, where text-based causal methods remain superior. Even so, the overall results show that transforming edges into nodes and enriching the graph with heterogeneous semantic information are promising strategies for edge classification.

Table 4 reports the results of FOLGA on three datasets and compares it with all families of one-class GNN methods using f_1-macro. FOLGA achieved the best f_1-macro

Table 3. ECHOLGA results. Best results in bold, and second best underlined.

Method	References	Risk	Twitter	Headlines	FinCausal
NLM-BERT	[Hassanzadeh et al. 2019]	0.669	-	-	-
NLM-BERT++	[Kayesh et al. 2020]	0.667	<u>0.668</u>	-	-
CF-Context	[Kayesh et al. 2023]	0.624	**0.673**	-	-
LLaMa 3 (8b)	[Liu et al. 2024]	<u>0.727</u>	0.570	0.610	0.761
Phi 3 (14b)	[Liu et al. 2024]	0.659	0.509	0.626	0.695
Qwen 2 (7b)	[Liu et al. 2024]	0.691	0.472	0.654	0.719
Gemma 2 (27b)	[Liu et al. 2024]	0.720	0.630	<u>0.734</u>	<u>0.822</u>
LLaMa 3 (70b)	[Liu et al. 2024]	0.726	0.623	0.728	0.767
ECHOLGA	**Proposal**	**0.776**	0.564	**0.806**	**0.824**

in all datasets, outperforming both homogeneous and heterogeneous one-class GNN baselines. The strongest performance occurs in Agriculture, where FOLGA outperforms the second-best method by a clear margin. In Immigration and Terrorism, the improvements are smaller but still consistent, which is relevant because several baselines already achieve high performance. These results indicate that rewiring and fusion are effective mechanisms for exploiting heterogeneous information without requiring additional loss terms. Thus, FOLGA validates the final step of the thesis progression: the OLGA learning core can be preserved while the graph topology is enriched to better represent heterogeneity.

Table 4. FOLGA results. Best results are in bold and the second-best underlined.

Methods	References	Agriculture	Immigration	Terrorism
OCGNN	[Wang et al. 2021]	0.791	0.663	<u>0.954</u>
DUAL-SVDAE	[Zhang et al. 2022a]	<u>0.865</u>	0.878	0.935
STEAM	[Zhang et al. 2022b]	0.330	<u>0.889</u>	0.939
SAMCL	[Hu et al. 2023]	0.516	0.416	0.402
EAGLE	[Shi et al. 2023]	0.536	0.699	0.934
JAANE	[Fan et al. 2024]	0.858	0.879	0.931
MULDUALGNN	[Li et al. 2024]	0.756	0.695	0.913
FIAD	[Chen et al. 2025]	0.671	0.474	0.409
FOLGA	**Proposal**	**0.933**	**0.902**	**0.962**

Table 5 summarizes the scientific production associated with the thesis. Table 5 shows the conference/journal acronym. For full names, consult the references section. The scientific production shows that the thesis generated publications across national and international venues, including journals and conferences with high impact or selective acceptance rates. Among the listed works, seven publications are classified as *Qualis* A1, with both methodological contributions and applied studies. The production also includes publications in A3, A4, and B2 venues, reflecting a continuous research trajectory from preliminary two-step graph-based OCL studies to mature end-to-end methods. Overall, these results indicate that the thesis produced not only a coherent methodological contribution but also sustained and diverse scientific output.

Table 5. Summary of the characteristics of the publications.

Papers	Event/Journal	AR[a]/IF[b]	*Qualis*[c]	Citations[d]
[Gôlo et al. 2022]	KDMILE	43%	B2	4
[da Silva et al. 2022]	KDMILE	43%	B2	6
[Coutinho et al. 2023]	ENIAC	61%	A4	1
[Gôlo et al. 2023a]	Webmedia	33%	A3	7
[Gôlo et al. 2023b]	EEAI	7.5	A1	37
[Gôlo et al. 2024a]	ENIAC	42%	A4	5
[Gôlo et al. 2024b]	JIDM	-	A4	0
[Gôlo et al. 2024c]	JIS	-	A4	1
[Gôlo et al. 2024d]	BRACIS	41%	A3	0
[de Souza et al. 2024]	Information Sciences	8.1	A1	22
[Filho et al. 2025]	NCAA	4.5	A1	3
[do Carmo et al. 2025]	ACM SAC	23%	A1	2
[Gôlo et al. 2025a]	Information Sciences	8.1	A1	3
[Gôlo and Marcacini 2025]	Scientific Reports	3.9	A1	1
[Gôlo et al. 2026]	IJCNN	39%	A1	0

[a] Acceptance rate, [b] Impact factor, [c] Qualis 2021-2024, [d] Google Scholar

5. Conclusion

This thesis advanced OCL for GNNs by proposing a coherent family of end-to-end methods that align graph representation learning with the one-class objective. Starting from OLGA, the thesis introduced a sphere-based loss combined with graph autoencoder reconstruction. This core was then extended to data streams with OLGASTREAM, to causal edge classification with ECHOLGA, and to heterogeneous graphs with FOLGA. The proposed methods achieved competitive, often state-of-the-art results, demonstrating that reconstruction, sphere-based compactness, topology enrichment, and interpretability can be combined into a robust methodological framework for OCL with GNNs.

From a practical perspective, the proposed methods can be applied to scenarios in which the objective is to detect a specific class of interest while the remaining space is broad, unknown, dynamic, or expensive to label. This includes fake news detection, event monitoring, anomaly and fraud detection, causal relation discovery, recommendation, and strategic profile matching. In social applications, these methods can support early detection of relevant or harmful events and improve decision-making. In business and industrial contexts, they can help organizations monitor risks, identify rare patterns, detect opportunities, and build graph-based intelligent systems in domains where only relevant examples are available.

The scientific impact of the thesis is also reflected in the awards received by publications derived from this research trajectory. This thesis received three second-best papers, one at KDMiLe through the paper [da Silva et al. 2022], one at ENIAC [Gôlo et al. 2024a], and one at BRACIS [Gôlo et al. 2024d]. Furthermore, in 2023, this thesis was awarded the Google Latin America Fellowship. These awards, together with publications at select conferences and in high-impact journals, indicate that the thesis made both methodological advances and recognized scientific contributions.

References

Bing, R., Yuan, G., Zhu, M., Meng, F., Ma, H., and Qiao, S. (2023). Heterogeneous graph neural networks analysis: a survey of techniques, evaluations and applications. *Artificial Intelligence Review*, 56(8):8003–8042.

Chen, A., Wu, J., and Zhang, H. (2025). Fiad: Graph anomaly detection framework based feature injection. *Expert Systems with Applications*, 259:125216.

Coutinho, G., Lima, A., Yugoshi, J., Moraes Junior, M., Gôlo, M., and Marcacini, R. (2023). Multimodal audio emotion recognition with graph-based consensus pseudolabeling. In *Proceedings of the National Meeting on Artificial and Computational Intelligence*, Porto Alegre, RS, Brasil. SBC.

da Silva, A., Gôlo, M., and Marcacini, R. (2022). Unsupervised heterogeneous graph neural network for hit song prediction through one class learning. In *Proceedings of the Symposium on Knowledge Discovery, Mining and Learning*, Porto Alegre, RS, Brasil. SBC.

de Souza, M. C., Gôlo, M. P. S., Jorge, A. M. G., de Amorim, E. C. F., Campos, R. N. T., Marcacini, R. M., and Rezende, S. O. (2024). Keywords attention for fake news detection using few positive labels. *Information Sciences*, 663:120300.

Deldar, F., Abadi, M., and Ebrahimifard, M. (2022). Android malware detection using one-class graph neural networks. *ISeCure*, 14(3).

Diestel, R. (2025). *Graph theory*. Springer, Berlin, Germany.

do Carmo, P. V., Silva Gôlo, M. P., Gwozdz, J., Marx, E., and Marcondes Marcacini, R. (2025). Improving natural product knowledge extraction from academic literature with enhanced pdf text extraction and large language models. In *Proceedings of the Symposium on Applied Computing*, pages 980–987, New York, NY, USA. ACM.

Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., and Yu, P. S. (2020). Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. In *Proceedings of the International Conference on Information & Knowledge Management*, pages 315–324, New York, NY, USA. ACM.

Emmert-Streib, F. and Dehmer, M. (2022). Taxonomy of machine learning paradigms: A data-centric perspective. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, page e1470.

Fan, H., Wang, R., Huang, X., Zhang, F., Li, Z., and Su, S. (2024). Deep joint adversarial learning for anomaly detection on attribute networks. *Information Sciences*, 654:119840.

Feng, Y., Chen, J., Liu, Z., Lv, H., and Wang, J. (2022). Full graph autoencoder for one-class group anomaly detection of iiot system. *Internet of Things Journal*.

Filho, I. J. R., Gôlo, M. P., Marcacini, R. M., and Rezende, S. O. (2025). How do financial time series enhance the detection of news significance in market movements? a study using graph neural networks with heterogeneous representations. *Neural Computing and Applications*, 37(3):1307–1319.

Ganz, T., Ashraf, I., Härterich, M., and Rieck, K. (2023). Detecting backdoors in collaboration graphs of software repositories. In *Proceedings of the Conference on Data and Application Security and Privacy*, pages 189–200, New York, NY, USA. ACM.

Gkarmpounis, G., Vranis, C., Vretos, N., and Daras, P. (2024). Survey on graph neural networks. *IEEE Access*.

Gôlo, M., Moraes, L., Goularte, R., and Marcacini, R. (2022). One-class recommendation through unsupervised graph neural networks for link prediction. In *Proceedings of the Symposium on Knowledge Discovery, Mining and Learning*, pages 146–153, Porto Alegre, RS, Brazil. SBC.

Gôlo, M., Moraes Junior, M., Goularte, R., and Marcacini, R. (2023a). On the use of early fusion operators on heterogeneous graph neural networks for one-class learning. In *Proceedings of the Brazilian Symposium on Multimedia and the Web*, New York, NY, USA. ACM.

Gôlo, M. P., Mori, A. L., Oliveira, W. G., Barbosa, J. R., Graciano-Neto, V. V., de Lima, E. A., and Marcacini, R. M. (2024a). On the use of large language models to detect brazilian politics fake news. In *Proceedings of the National Meeting on Artificial and Computational Intelligence*, pages 1–12, Porto Alegre, RS, Brasil. SBC.

Gôlo, M. P. S., de Medeiros Junior, J. G. B., Silva, D. F., and Marcacini, R. M. (2025a). One-class graph autoencoder: A new end-to-end, low-dimensional, and interpretable approach for node classification. *Information Sciences*, page 122060.

Gôlo, M. P. S., de Medeiros Junior, J. G. B., Vilas Boas, G. S., Lobato, F. M. F., Silva, D. F., and Marcacini, R. M. (2025b). Large language models ensemble for biochemical properties detection in scientific articles. In *Proceedings of the Text2KG Workshop at European Semantic Web Conference*, Aachen, Germany. CEUR-WS.

Gôlo, M. P. S., de Moraes, L. G., Goularte, R., and Marcacini, R. M. (2024b). The impact of representation learning on unsupervised graph neural networks for one-class recommendation. *Journal of Information and Data Management*, 15(1):112–122.

Gôlo, M. P. S., de Moraes Junior, M. I., Goularte, R., and Marcacini, R. M. (2024c). Unsupervised heterogeneous graph neural networks for one-class tasks: Exploring early fusion operators. *Journal on Interactive Systems*, 15(1):517–529.

Gôlo, M. P. S., de Souza, M. C., Rossi, R. G., Rezende, S. O., Nogueira, B. M., and Marcacini, R. M. (2023b). One-class learning for fake news detection through multimodal variational autoencoders. *Engineering Applications of Artificial Intelligence*, 122:106088.

Gôlo, M. P. S., Gama, J., and Marcacini, R. M. (2024d). One-class learning for data stream through graph neural networks. In *Proceedings of the Brazilian Conference on Intelligent Systems*, pages 61–75, Cham, Switzerland. Springer.

Gôlo, M. P. S. and Marcacini, R. M. (2024). One-class learning for text causal discovery through hypergraph neural networks. In *Proceedings of the Workshop on Causal Discovery*, pages 1–12.

Gôlo, M. P. S. and Marcacini, R. M. (2025). One-class edge classification through heterogeneous hypergraph for causal discovery. *Scientific Reports*.

Gôlo, M. P. S. and Utino, M. Y. R. (2025). One-class lightweight interpretable filtering for academic profiles and strategic themes affinity. In *Proceedings of the International Symposium on Computer Architecture and High Performance Computing Workshops*, pages 147–154, Piscataway, NJ, USA. IEEE.

Gôlo, M., Gama, J., Marx, E., and Marcacini, R. (2026). Fusion-rewired one class graph autoencoder. In *Proceedings of the International Joint Conference on Neural Networks*, Piscataway, NJ, USA. IEEE. Accepted for publication.

Hassanzadeh, O., Bhattacharjya, D., Feblowitz, M., Srinivas, K., Perrone, M., Sohrabi, S., and Katz, M. (2019). Answering binary causal questions through large-scale text mining an evaluation using cause-effect pairs from human experts. In *Proceedings of the International Joint Conferences on Artificial Intelligence*, pages 5003–5009, Washington, DC, USA. AAAI.

Hu, J., Xiao, B., Jin, H., Duan, J., Wang, S., Lv, Z., Wang, S., Liu, X., and Zhu, E. (2023). Samcl: Subgraph-aligned multiview contrastive learning for graph anomaly detection. *Transactions on Neural Networks and Learning Systems*.

Huang, B., Wang, X., Cui, P., Jiang, W., Yang, Y., and Fan, Q. (2021). One-class temporal graph attention neural network for dynamic graph anomaly detection. In *Proceedings of the International Conference on Electronics, Communications and Information Technology*, pages 783–790, Piscataway, NJ, USA. IEEE.

Huang, Z., Gu, Y., and Zhao, Q. (2022). One-class directed heterogeneous graph neural network for intrusion detection. In *Proceedings of the International Conference on Innovation in Artificial Intelligence*, pages 178–184, New York, NY, USA. ACM.

Jodelka, O., Anagnostopoulos, C., and Kolomvatsos, K. (2021). Adaptive novelty detection over contextual data streams at the edge using one-class classification. In *Proceedings of the International Conference on Information and Communication Systems*, Piscataway, NJ, USA. IEEE.

Kayesh, H., Islam, M. S., and Wang, J. (2023). Answering binary causal questions using role-oriented concept embedding. *Transactions on Artificial Intelligence*, 4(06):1426–1436.

Kayesh, H., Islam, M. S., Wang, J., Anirban, S., Kayes, A., and Watters, P. (2020). Answering binary causal questions: A transfer learning based approach. In *Proceedings of the International Joint Conference on Neural Networks*, pages 1–9, Piscataway, NJ, USA. IEEE.

Khemani, B., Patil, S., Kotecha, K., and Tanwar, S. (2024). A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *Journal of Big Data*, 11(1):18.

Li, Q., Wu, G., Ni, H., and You, T. (2024). Anomaly detection with dual-channel heterogeneous graph based on hypersphere learning. *Information Sciences*, 681:121242.

Liu, X., Xu, P., Wu, J., Yuan, J., Yang, Y., Zhou, Y., Liu, F., Guan, T., Wang, H., Yu, T., et al. (2024). Large language models and causal inference in collaboration: A comprehensive survey. *arXiv preprint arXiv:2403.09606*.

Marx, E., do Carmo, P. V., Gôlo, M., and Tramp, S., editors (2025). *Preface of the First International TEXT2SPARQL Challenge (TEXT2SPARQL'25)*, Aachen, Germany. CEUR-WS.

Mendelson, S. and Lerner, B. (2020). Online cluster drift detection for novelty detection in data streams. In *Proceedings of the International Conference on Machine Learning and Applications*, Piscataway, NJ, USA. IEEE.

Miranda, A. C. D., Utino, M. Y. R., Gôlo, M. P. S., Santos, M. A. A. D., and de Souza, M. C. (2025). Reducing costs in large-scale classification: A hybrid bert–llm strategy. In *Proceedings of the International Symposium on Computer Architecture and High Performance Computing Workshops*, pages 131–138, Piscataway, NJ, USA. IEEE.

Nguyen, T. and Grishman, R. (2018). Graph convolutional networks with argument-aware pooling for event detection. In *Proceedings of the Conference on Artificial Intelligence*, Washington, DC, USA. AAAI.

Seliya, N., Abdollah Zadeh, A., and Khoshgoftaar, T. M. (2021). A literature review on one-class classification and its potential applications in big data. *Journal of Big Data*, 8:1–31.

Shi, B., Dong, B., Xu, Y., Wang, J., Wang, Y., and Zheng, Q. (2023). An edge feature aware heterogeneous graph neural network model to support tax evasion detection. *Expert Systems with Applications*, 213:118903.

Tax, D. M. J. (2001). *One-class classification: Concept learning in the absence of counter-examples*. PhD thesis, Technische Universiteit Delft.

Utino, M. Y. R. and Gôlo, M. P. S. (2025). Efficient filtering with bert embeddings for researcher–topic affinity prediction in hpc pipelines. In *Proceedings of the International Symposium on Computer Architecture and High Performance Computing Workshops*, pages 155–162, Piscataway, NJ, USA. IEEE.

Veličković, P. (2023). Everything is connected: Graph neural networks. *Current Opinion in Structural Biology*, 79:102538.

Wang, X., Bo, D., Shi, C., Fan, S., Ye, Y., and Philip, S. Y. (2022). A survey on heterogeneous graph embedding: methods, techniques, applications and sources. *Transactions on Big Data*, 9:415 – 436.

Wang, X., Jin, B., Du, Y., Cui, P., Tan, Y., and Yang, Y. (2021). One-class graph neural networks for anomaly detection in attributed networks. *Neural Computing and Applications*, 33(18):12073–12085.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *Transactions on Neural Networks and Learning Systems*, 32(1):4–24.

Zadeh, A. A. (2024). *Advancing One-Class Classification: A Comprehensive Analysis From Theory To Novel Applications*. PhD thesis, Florida Atlantic University Boca Raton, FL.

Zhang, F., Fan, H., Wang, R., Li, Z., and Liang, T. (2022a). Deep dual support vector data description for anomaly detection on attributed networks. *International Journal of Intelligent Systems*, 37(2):1509–1528.

Zhang, J., Wang, X., Yu, Z., Huang, M., Duan, L., and Zuo, J. (2022b). Steam: Svdd-based anomaly detection on attributed networks via multi-autoencoders. In *Proceedings of the International Conference on High Performance Computing and Communications*, pages 595–602, Piscataway, NJ, USA. IEEE.

Apendix

Figures 1, 2, 3, and 4 jointly illustrate the interpretability of OLGA, OLGASTREAM, ECHOLGA, and FOLGA, respectively. For each method, representative stages from initialization to convergence are shown using the low-dimensional embeddings directly produced by the models, without dimensionality-reduction post-processing. The visualizations cover different application domains and graph structures: review classification, streaming social-media data, causal relation detection in heterogeneous hypergraphs, and heterogeneous agricultural event detection. Colors distinguish interest and non-interest instances or, in the ECHOLGA visualization, the different node types and causal relation classes. The learned hypersphere is also explicitly displayed, providing a geometric interpretation of the one-class decision function.

Figure 1. OLGA interpretable results.

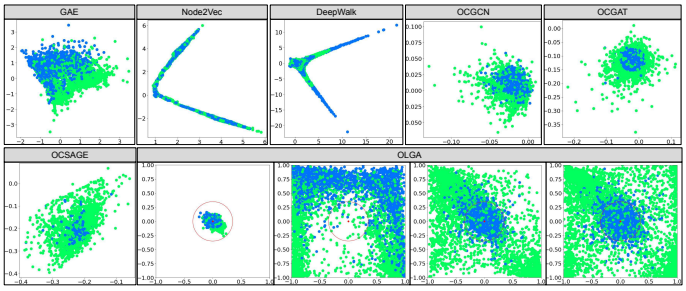


Figure 2. OLGASTREAM interpretable results.

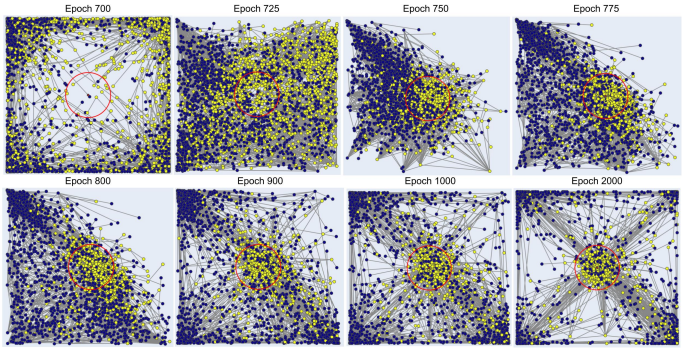


Figure 3. ECHOLGA interpretable results.

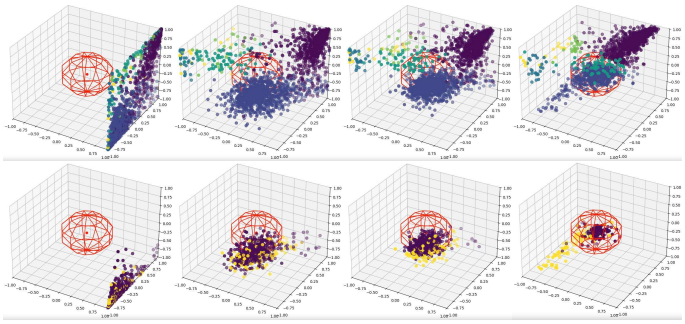
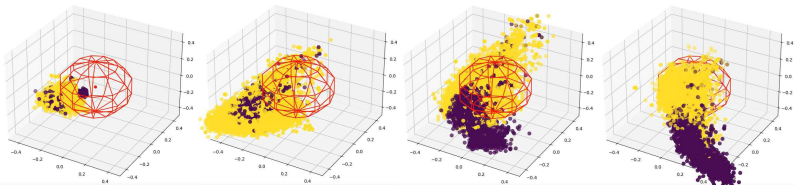


Figure 4. FOLGA interpretable results.



Across the four methods, the visualizations reveal a consistent multitask learning behavior. During the initial stages, reconstruction-related objectives dominate the optimization, preserving graph structure and preventing representation collapse. As training progresses, the one-class loss increasingly attracts interest instances toward the sphere center, while non-interest instances tend to remain outside or near its boundary. In ECHOLGA, the node-type loss additionally maintains the separation among heterogeneous node categories, whereas in OLGASTREAM the gradual variation of the loss weights adapts this balance to the streaming setting. FOLGA further benefits from the fused rewired representations, producing a compact interest cluster and a visibly separated group of outliers. These results show that the proposed models learn representations specifically tailored to one-class learning, unlike two-step methods whose embeddings are agnostic to the subsequent classifier. The visualizations provide a geometric interpretation of how the representations evolve during training and illustrate a reduced overlap between interest and non-interest instances in the selected examples. This behavior is consistent with the predictive results reported for the proposed methods.

This thesis also summarizes contributions and research resources developed throughout the Ph.D. that complement the end-to-end one-class graph neural networks presented in this thesis. These contributions comprise two-step one-class methods that motivated the integrated approaches, as well as reusable datasets, benchmarks, and experimental settings. These complementary contributions are not part of the four core OLGA methods; they document preliminary studies and reusable research resources developed during the Ph.D. that may support future one-class graph learning research. Furthermore, before developing the end-to-end methods, three studies investigated two-step methods, in which graph representation learning and one-class classification were performed separately. These studies demonstrated the importance of graph structure, multimodal information, and representation fusion while providing the conceptual and empirical basis for the methods proposed in the thesis.

The first study formulated hit song prediction as a one-class problem and represented songs and artists through a heterogeneous graph [da Silva et al. 2022]. Song nodes were initialized with lyric-based BERT representations, whereas artist nodes contained acoustic information and were connected according to their semantic relationships. An unsupervised GNN generated embeddings that were subsequently used by a One-Class Support Vector Machine (OCSVM). The results showed that propagating modality through graph relations was more effective than concatenating multimodal features.

The second study proposed one-class recommendation using heterogeneous graphs containing users, items, and associated metadata [Gôlo et al. 2024b, Gôlo et al. 2022]. Unsupervised GNNs generated user and item embeddings, which were concatenated and classified by an OCSVM. Experiments were performed on three rec-

ommendation datasets. The method improves the recommendation performance. These findings reinforced that representations explicitly optimized for graphs are particularly valuable when only positive user–item interactions are available.

The third study investigated early fusion operators for multimodal graph representations in one-class classification [Gôlo et al. 2024c, Gôlo et al. 2023a]. A Graph Autoencoder learned representations for fake news, music, and recommendation graphs, after which addition, subtraction, average, multiplication, minimum, maximum, and concatenation operators were evaluated. Graph Autoencoder representations generally outperformed regularized initial representations. Addition, average, and subtraction presented the most consistent behavior, while concatenation increased dimensionality without consistently improving classification. The results motivated the integration of representation learning, fusion, and one-class objectives into the end-to-end architectures developed in the thesis.

The Ph.D. also produced four families of resources suitable for future graph-based and one-class investigations. The first comprises Brazilian fake news datasets, including a labeled political news collection used to compare proprietary and open-source language models [Gôlo et al. 2024a]. Complementary studies represented news and terms as graphs and incorporated attention mechanisms to identify relevant linguistic elements under limited supervision [de Souza et al. 2024]. These datasets enable document–term, document-similarity, and heterogeneous graph constructions in which fake news can be treated as the class of interest. The second family concerns knowledge graphs. The Text2SPARQL challenge and its resulting benchmark provide curated natural-language question and SPARQL query pairs over DBpedia [Marx et al. 2025]. In the biochemical domain, the NatUKE benchmark organizes entities, textual evidence, and biochemical properties through a knowledge graph derived from NUBBE$_{DB}$ [do Carmo et al. 2025]. Subsequent participation in the BIKE challenge extended this setting through prompting, fine-tuning, and ensemble strategies for biochemical property extraction [Gôlo et al. 2025b]. These resources support semantic parsing, knowledge graph embeddings, information extraction, and graph-based anomaly or interest detection.

The third family addresses affinity between academic profiles and strategic themes. A real-world dataset contains 460 strategic themes and $42{,}046$ Portuguese theses and dissertations defended in Brazil in 2023 [Utino and Gôlo 2025, Gôlo and Utino 2025]. Profiles and themes can be modeled as a bipartite or heterogeneous graph enriched with institutions, research areas, keywords, and entities. This setting offers a large-scale environment for one-class link prediction, ranking, filtering, and researcher–project matching. Finally, multimodal resources include audio–text graphs for unsupervised emotion recognition [Coutinho et al. 2023] and datasets connecting news articles with financial time series [Filho et al. 2025]. Both settings naturally support heterogeneous graphs that connect representations from different modalities. Consequently, they provide practical scenarios for evaluating multimodal graph fusion and one-class classification, such as detecting a target emotion or identifying news relevant to financial and agricultural markets. Together, these contributions demonstrate the broader applicability of one-class graph learning and provide reusable foundations for extending the methods introduced in this thesis to recommendation, misinformation detection, knowledge extraction, scientific matching, and multimodal analysis.