Target class classification recursion preliminaries

Levon Aslanyan, Karen Gishyan and Hasmik Sahakyan

Institute for Informatics and Automation Problems of National Academy of Sciences of Armenia

An Important Problem

In Machine Learning the Supervised Classification scenario [3, 4] supposes a number of geometrically "compact" sets of elements (clusters) given by the sets of class representatives (elements of the so-called learning set), asking for universal procedures that may correctly classify new objects into the classes. After the learning / training stage, recognition / classification of the trial object is, in general, a static, one-step process. In our application, which is fundamentally different from the traditional learning model, one of the classes is marked, especially, and the aim is in learning, allocating all objects, by a step-by-step procedure to this special class [5, 2]. In an individual step of the recursion, when the temporal class of the object is determined, a predetermined class action is applied, that transfers this object to the same or to some other class.

Introduction

This article is the first comprehensive introduction to the research idea in the Project 21SC-BRFFR-1B029. Initial publications [5, 2] preceded the official start of the project. Since the general idea is presented in terms of classification algorithms and is referred to as a completely novel research postulate, we consider it necessary to point out the innovation of this idea, as well as to recall the existing associations with current research tasks in the field of classification algorithms. Sequential learning [6] is aimed at narrowing down classes, step-by-step, until finding the required class for the object. In the case of our task, sequential training/classification is aimed at a step-by-step approximation of the object being classified to the predetermined target class. Also here the classification structure is limited, basing on the totality of available data and rules of the subject area, and it is required to build an optimal strategy / algorithm that produces a correct classification, to the target class.

Analysis of Classification Framework

Classification problem in the field of machine learning belongs to the domain of pattern recognition [8, 12, 11, 9]. The classification problem objective is to identify the class (property, category) to which an observed object belongs, based on the set of associated feature values. A classification problem can be further categorized into (a) binary classification problem in case the class label can take only two values [13], (b) multiclass classification problem [7, 15, 14, 10, 1] in case the class label can take more than two different values, and (c) multi-label classification problem in case that each observation is associated with multiple classes. The current classification is a mixture of discrete mathematical procedures, heuristics, approximations, and estimations, with random sampling, average and worst case analyses, and statistical estimations that are as a rule an overestimation. The problem considered in this paper is a new, specific classification paradigm, but it can still be interpreted in association with the properties of unbalanced classes and sequential learning. Normal imbalance is meant to protect small classes from miss-classification while target class classification focuses on one and the main class which we call the target class. The task is to build an algorithm that will assign all objects to the target class. Classification means are limited, and in order to achieve the goal, it is necessary to apply them a consecutive number of times.

Results

Consider a graph $G = (V, E)$ with vertices that correspond to classes, edges with marked actions of classes, and an oriented transition to class vertices determined by the action. The normal class is allocated to the vertex at v_0. Structure of graph G contains a so-called normal tree rooted at v_0. We construct such a tree as follows. Consider the vertex at v_0. A number of oriented edges enter into it, for example, from the vertices $v_1, v_2, \ldots, v_{k_1}$. The vertices $v_1, v_2, \ldots, v_{k_1}$, having given an outgoing edge to the normal vertex, exhaust their outgoing edges and therefore they will not appear further up the tree. At the next step, we sort out the vertices $v_{i_1}, v_{i_2}, \ldots, v_{i_{k_1}}$ and all the edges entering them from the vertices $v_{i_1}, v_{i_2}, \ldots, v_{i_{k_1}}$. This procedure repeated recursively ends at some step l with the construction of a tree rooted at v_0, with one edge going down from all internal (non-root) vertices to the adjacent layer of vertices, but it is clear that all vertices except the root can contain multiple oriented ingoing edges. Now, if the graph is connected, then its structure is defined. Otherwise, consider its connected components $G_0, G_1, G_2, \ldots, G_m$. The component G_0 corresponds to the normal tree constructed at the first step of the graph interpretation. Let's consider an arbitrary one of the new connected components. First, there must be a cycle in them. This is explained by the finite number of vertices in the components, where each vertex can give an outgoing oriented edge to one other vertex. We may construct such a cycle as follows: start from an arbitrary vertex, go by the outgoing edge of this vertex to the target vertex, and repeat this step sequentially. The resulting chain is forced to close the cycle, as a result of the limited expansion space. Further, suppose that there is a second cycle in a component. If two cycles do not have common vertices, then we get a contradiction with the connectedness of the component. After all, a vertex of a cycle only descends to the next vertex within the cycle, and the only outgoing edge of each vertex is given to provide the formation of the cycle itself. Other structures may enter the vertices of the cycle but there is no way out of the cycle. Thus, any vertex of one cycle cannot be connected with the vertices of the second cycle. This asserts that in the presence of more than one cycle in the connected component, these cycles cannot have some connected vertices.

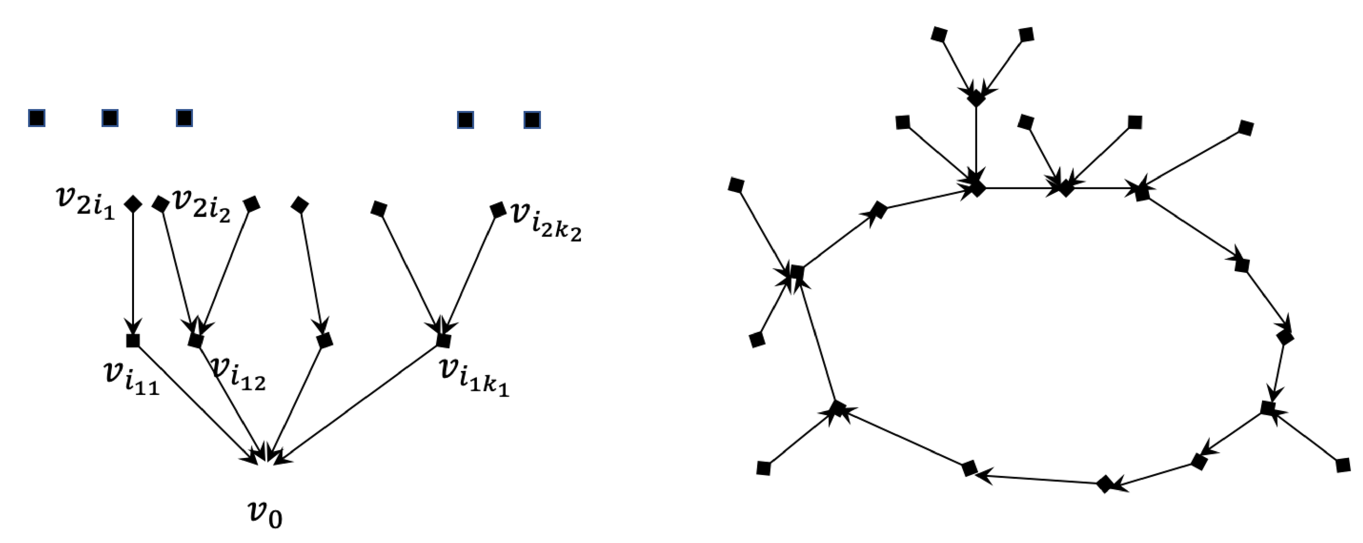

Fig. 1: Transition Graph Structure.

The vertex intersection of two cycles implies their identity. Indeed, starting from an arbitrary such vertex and following the outgoing oriented edges, we construct one single cycle. There is one outgoing edge and there is no branching space. It turns out that each of the connectivity components has a single cycle. Trees similar to a normal tree can converge at the vertices of this cycle, and thus the structure of the component itself is provided by the one cycle cactus graph and completes the proof of the proposition.

Conclusion

Among the imbalanced and recursive classification algorithms, there is a practical need to develop new algorithms that, through successive classifications and transformations, classify objects into one predetermined class. At a practical level, these studies are related to issues of precision medicine. Algorithmically, research starts with graph theory and discrete optimization, and continues towards stochastic modelling and reinforcement learning. An initial result is obtained in terms of graph theory, where the simplest models of problems with deterministic transition outcomes are considered.

Acknowledgements

This paper is partially supported by by grant №21SC-BRFFR-1B029 of the Science Committee of MESCS RA.

References

[1] Erin L Allwein, Robert E Schapire, and Yoram Singer. "Reducing multiclass to binary: A unifying approach for margin classifiers". In: *Journal of machine learning research* 1.Dec (2000), pp. 113–141.

[2] L Aslanyan, V Ryazanov, and H Sahakyan. "On Logical-Combinatorial Supervised Reinforcement Learning". In: *International Journal "Information Theories and Applications", Vol. 27, Number 1,* (2020).

[3] Levon Aslanyan and Juan Castellanos. "Logic based pattern recognition-ontology content (1)". In: *Statistics in medicine* 14.3 (2007), pp. 206–210.

[4] Levon Aslanyan and Vladimir Ryazanov. "Logic based pattern recognition-ontology content (2)". In: *Information Theories and Applications* 15.4 (2007), pp. 314–318.

[5] Levon Aslanyan et al. "Logical-Combinatorial Approaches in Dynamic Recognition Problems". In: *arXiv preprint arXiv:2101.11066* (2021).

[6] Yair Even-Zohar and Dan Roth. "A sequential model for multi-class classification". In: *arXiv preprint cs/0106044* (2001).

[7] Sokol Koço and Cécile Capponi. "On multi-class classification through the minimization of the confusion matrix norm". In: *Asian Conference on Machine Learning.* PMLR. 2013, pp. 277–292.

[8] Bongard M. M. *Problem of Cognition (in Russian).* Moscow: Fizmatgiz, 1967.

[9] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning.* MIT press, 2018.

[10] Anderson Rocha and Siome Klein Goldenstein. "Multiclass from binary: Expanding one-versus-all, one-versus-one and ecoc-based approaches". In: *IEEE Transactions on Neural Networks and Learning Systems* 25.2 (2013), pp. 289–302.

[11] A. Ya. Chervonenkis V. N. Vapnik. *Pattern recognition Theory (in Russian).* Moscow: Nauka, 1974.

[12] Zhuravlev Yu.I. *Selected research work.* Moscow: Magister, 1998.

[13] Yu I Zhuravlev, LA Aslanyan, and VV Ryazanov. "Analysis of a training sample and classification in one recognition model". In: *Pattern recognition and image analysis* 24.3 (2014), pp. 347–352.

[14] Yu I Zhuravlev et al. "Comparison of different dichotomous classification algorithms". In: *Pattern Recognition and Image Analysis* 30.3 (2020), pp. 303–314.

[15] Yu I Zhuravlev et al. "On a classification method for a large number of classes". In: *Pattern Recognition and Image Analysis* 29.3 (2019), pp. 366–376.