Chapter 1
Introduction

Abstract When we say "learning," one of the words that comes into our mind may be "mystery"; when we talk about "large scale networks," we may associate it to the word "complexity." What happens when we put together these two concepts? In this chapter, we present an overview on complex network-based machine learning. Throughout the entire book, we show the diversity of approaches for treating such a subject.

1.1 General Background

Human beings are born with the fascinating gift of learning. With the aid of such ability, they absorb and assimilate knowledge throughout their entire life. In an attempt to simulate such notorious characteristic in a computational environment, the research area entitled *machine learning* arose, which aims at developing computational methods that are capable of "learning" with past accumulated experiences [11, 19, 33].

Based on the computational data representation obtained from a wide range of domains, machine learning techniques can generate models apt to organize the existing knowledge or, yet, mimic the behavior of a human expert in an automatically manner. Generally speaking, these techniques are traditionally divided into two paradigms: the supervised and the unsupervised learning [11, 19]. In the *supervised learning* case, the goal is to infer concepts regarding the data using both the data distribution and external knowledge in the form of labels. In essence, the supervised learning process aims to construct a mapping function conditioned to the provided training data set. When these labels comprise discrete values, the inference problem is denominated *classification*, whereas when label values are continuous, *regression*. In contrast, in the *unsupervised learning* case, the main task consists in finding intrinsic data structures. The learning process, in this case, is solely guided by the data relationships as no external knowledge about the existing labels is required [33].

Supervised learning requires a labeled data set for training. However, the task of manual labeling is, in the majority of the cases, a cumbersome and expensive process, which usually involves the work of human experts. In order to soften this shortcoming, a third learning paradigm denominated *semi-supervised learning*

T.C. Silva, L. Zhao, *Machine Learning in Complex Networks*,
DOI 10.1007/978-3-319-17290-3_1

has emerged. The main distinct feature of semi-supervised algorithms is that they employ both labeled and unlabeled data in the prediction process by means of, for example, a diffusive labeling process. Frequently in real-world situations, data sets consist of a great amount of unlabeled data and a few labeled data. In this way, semi-supervised learning can considerably reduce human experts' efforts. Moreover, empirical results have shown that the usage of unlabeled data can improve the performance of fully supervised classification [15, 50].

Over the last decade, there has been an increasing interest in network research, with the focus drifting away from analyzing small graphs to considering large-scale graphs, called *complex networks*. Such networks have emerged as a unified representation of complex systems in various branches of science. In general, they are used to model systems that have nontrivial topologies and are composed of a large amount of vertices [1, 8, 36].

The study of networks began with the development of the graph theory, inaugurated by Leonhard Euler in 1736 with the solution of the seven bridges problem of Königsberg, today Kaliningrad, Russia. The problem, much discussed at the time, recorded that there were seven bridges crossing the river Pregel, with two intermediate islands. The residents desired to know whether it was possible to cross all of these seven bridges, without repetition, and then return to the starting point. Euler demonstrated, in an analytical manner, for the Russian Academy of Sciences in St. Petersburg, that it was not possible to complete such a walk. For this end, he made use of a graphical representation consisting of points and curves connecting these points. It was the beginning of the formal representation of a graph or network,[1] known until today, with vertices and edges. Thereafter, many researchers began studying this branch of research in search of new theories and applications [36].

In fact, the first major step in the study of complex networks was driven by Paul Erdös and Alfréd Réyni, who analyzed a certain type of network called *random networks* in their work published in 1959 [20]. That investigation opened doors to a novel area of study termed the theory of random networks, which represents a mixture of the graph and probabilistic theories to generate and analyze large-scale graphs.

Following the chronology, in 1967, Stanley Milgram decided to accept the challenge posed by Frigyes Karinthy, which, inspired by the conjectures of Guglielmo Marconi in 1909, dared one to find another person to whom he/she could not be transitively connected by using at most five intermediate people [32]. And it was exactly because of this problem that the concept denominated *separation in six degrees* was born, which was the first seed for the study of the *small-world networks*. To address this challenge, Milgram conducted experiments in order to try to discover the probability of any two arbitrary people to know each other. For this, letters were sent to random people living in predetermined regions of the United States, whose inner content dealt with information about any other arbitrary person. If the person

[1] Since graph and network share the same definition, these two terms are interchangeable in this book.

referred to in the letter was known by the reader, then he/she would mail the letter for the recipient. On the other hand, if he/she did not know, then the letter should be sent to someone else known. At the end of experiment, Milgram found that the average number of referrals of one person to another reached 5.5 persons. He was, therefore, empirically discovering the property of small world, which states that even though there are millions of persons interconnected in a social network, the average distance between them is only a small amount, in the example, 5.5 persons [32].

Despite of the findings of Milgram, it was only in the late 1990s that surveys on this field were retaken. In 1998, Watts and Strogatz found that the average shortest paths in a network can be drastically reduced by a random alteration of few links, starting from a regular network [56]. The resulting network is called *small-world network*, which, as we have seen, had already been empirically discovered by Milgram. In 1999, Barabási and Albert discovered that many real networks have a degree distribution of vertices that obeys a power law: $P(k) \sim k^{-\gamma}$, where k is the number of connections of a randomly chosen vertex and γ is a scaling exponent [7]. This heterogeneous distribution describes the existence of a small number of vertices that has a large number of connections, while the majority only shares few connections. Such networks are called *scale-free networks*.

Driven by the technological advances and also by the increasing number of data to be jointly analyzed, the complex networks area has emerged as a unifying topic in complex systems and is present in various branches of science [13]. Structurally, complex networks are represented by a graph $\mathscr{G} = \langle \mathscr{V}, \mathscr{E} \rangle$, where $\mathscr{V}$ represents the set of vertices and $\mathscr{E}$, the set of edges. Complex networks can be conceived as a general modeling scheme for heterogenous systems with arbitrary sizes [3], as they naturally incorporate the usual nontrivial aspects of the system agents.

Evidences of complex networks in real-world settings abound. Among some examples of real-world systems in which the network representation is perfectly plausible, we can highlight: the *Internet* [22], the *World Wide Web* [2], biological neural networks [52, 57], financial networks [14, 43, 51], information networks [59], social networks among individuals [27, 46] and between companies and organizations [34], food webs [35], metabolic networks [16, 26] and distribution as the bloodstream [58], protein-protein networks [55], postal delivery and electricity distribution networks [3] etc.

The data representation in complex networks inherently presents some positive characteristics of which we can cite [53]:

- *The structural complexity*—which translates into the heterogeneous and nontrivial data connections that are shared between vertices in the network. This feature can be easily understood by taking into account the difficulty in visualizing the network properties.
- *The evolution*—which marks the constant changes in the network structure due to the inclusion and removal of vertices and connections [18].
- *The diversity of connections*—because the connections between vertices can have various physical meanings, such as capacity, length, width and direction. These features are often operationalized via multilayer networks [12]. These networks

have several network layers, each of which representing different aspects of the possible connections.
- *The dynamical nature*—which affects, at a large scale, the states of a network, as can be construed as traffic information [63], occurrences of failures in communications [60, 61, 63], similarity relations between vertices, the distribution of functions [36], among others.

1.2 Focus of This Book

Machine learning and data mining techniques using complex networks have triggered increased attention. This is because networks are ubiquitous in nature and everyday life, as several real-world problems can be directly represented in terms of networks. Moreover, many other kinds of data sets can be transformed into network representations using a suitable network formation technique. For instance, a set of items described by feature (or attribute) vectors can be transformed into a network by simply connecting each sample to its k nearest neighbors. In addition, the complex network representation unifies the structure, dynamics, and functions of the system that it represents. It does not only describe the interaction among vertices (structure) and the evolution of such interactions (dynamics), but also reveals how the structure and dynamics affect the overall functions of the network [39]. For example, it is well known that there is a strong connection between the structure of protein-protein interaction networks and protein functions [42]. The main motivation of network research is the ability of describing topological structures of the original system. In the machine learning domain, it has been shown that the topological structure is quite useful to detect clusters of arbitrary forms in data clustering [23, 29].

This book brings as main goals the review and the development of machine learning techniques that are based on complex networks. Our intention is twofold:

1. To provide a thorough coverage on this topic. In this way, the book not only gives a general vision on this topic but also facilitates the effort of readers in finding out relevant materials for their own development. For this purpose, comprehensive reviews on complex networks and network-based machine learning are explored in the first part of this book.
2. To describe and bridge the connection between machine learning and complex networks. To this end, we focus on several up to date developments of these two topics in the second part of the book (the last three chapters). Here, the three branches of the machine learning area are explored, i.e., the unsupervised, the semi-supervised, and the supervised learning areas.

In the first part of the book, we give an overview of relevant concepts and some technical details on network-based machine learning. Since both complex networks and machine learning are well developed and multidisciplinary research areas, we would like to follow this feature and present diversity of ideas. However,

due to the vast research results developed so far, we must make choices on the content of the book. For example, synchronization in complex networks has not been reviewed in this book. We have not gone to technical details on some topics, for examples, epidemic spreading and graph kernels. The interested readers may get further knowledge with specific materials cited in this book.

In the review of complex networks in Chap. 2, we first introduce basic concepts and notations on graphs, then we present the classic complex network models. The focus of this chapter is to provide a comprehensive review on network measures in a categorized way. Many such measures have been applied to develop machine learning techniques, while others may be used for future developments probably by the readers of this book and can be easily found in classical materials.

Then we go to the review of machine learning in Chap. 3, where we introduce the three traditional paradigms: supervised, unsupervised, and semi-supervised learning. When exploring them, we pay attention to the basic concepts and characterization of machine learning instead of technical details, which are out of the scope of this book.

Many data sets are not already in a network format. In order to apply network-based techniques for data analysis, it is necessary to transform the original data sets into networks. Intuitively, a single data set can be represented by different networks, which may lead to different qualities of final results. Therefore, network construction from original data sets is a crucial issue for data mining and machine learning. For this reason, Chap. 4 is dedicated to reviewing and compiling various methods that can be employed for network construction. It is worth mentioning that network construction is still at its infant stage and there is a large space for exploration. The readers are invited to contribute to this interesting topic.

Given the background knowledge of complex networks and machine learning, we start to present to the readers an overview of network-based machine learning. As we will see, the majority of the supervised and semi-supervised learning techniques is within-network methods, i.e., probabilistic or deterministic inference of class labels of some vertices takes place through finding out the "best" route from the labeled to unlabeled vertices within the network. (In the second part of this book, we will present an across-network supervised learning technique, in which global information of the underlying network is taken into account). We will also show that network-based unsupervised learning is really a community detection task. The presence of *communities* is a striking phenomena of complex networks. The notion of community in networks is straightforward: each community is defined as a subnetwork whose vertices are densely connected within itself, but sparsely connected with the rest of the network. Community detection in complex networks has turned out to be an important topic in graph mining and in data mining [17, 23, 30, 40]. In graph theory, community detection corresponds to graph partition, which is an NP-complete problem [23]. For this reason, a lot of efforts has been spent to developing efficient but suboptimal solutions, such as the spectral method [38], the technique based on the "betweenness" measure [40], modularity optimization [37], community detection based on the Potts model [45], synchronization [6], information theory [24], and random walks [64].

As community detection is directly related to unsupervised learning, a review of this topic will be presented in Chaps. 2 and 6.

The second part of this book is devoted to presenting concrete ideas and technical details on specific realizations of the three machine learning paradigms. When investigating these developments, we are concerned in addressing:

- The development of novel computational methods in machine learning are expected to be equipped, whenever possible, with an underlying mathematical framework. The construction of models that rely on mathematical grounds opens way to characterizing the short- and long-run behaviors of these techniques in a concise and formal manner. We believe that this is an important step to better understand the dynamics of the models and, as a consequence, to better enable one to perceive the potentialities and the shortcomings offered by them. Nonetheless, whenever it is possible, empirical studies are likewise conducted to consolidate and to confirm the validity of the analytical predictions.
- The design of these machine learning techniques, in contrast to traditional methods, is expected to provide alternative and novel ways to solve the challenging problems posed by the machine learning area. In this way, novel and efficient algorithms modeled for tasks of clustering and classification are explored. In other words, we intend to investigate new features and possible advantages offered by complex networks in the machine learning domain. In fact, we do show that the network-based approach really brings interesting features for machine learning.
- Having in mind the possibility of utilizing the methods described herein in real-world applications, we also take into account the design of techniques that are complementary in terms of performance and computational complexity. Consequently, we must find equilibrium between quality and efficiency when designing network-based machine learning techniques.

In order to give a concrete vision and present some know-hows on designing complex network-based machine learning techniques, this book details the following up to date developments:

- With regard to the unsupervised learning area:

    1. The fundamental mechanism and technical details of the particle competition model in complex networks are presented. The particle competition model was originally proposed and applied in community detection in [44]. Then, it was reformulated in a formal dynamical system and extended to solve data clustering problems by [49]. The model consists of several particles walking within the network and competing with each other to occupy as many vertices as possible, while attempting to reject intruder particles. The particle's walking rule is composed of a stochastic combination of random and preferential movements. As we will see in Chap. 9, computer simulations reveal that this model presents high community detection rates, as well as low computational complexity. One strong argument for delving into the process of particle competition is that of its similarity to many social and

natural processes, such as: competition among animals, territorial exploration by humans (animals), election campaigns, among others. Furthermore, the random-preferential movement incorporated into the particles' movement policy can substantially improve the model's performance, as we will see in Chap. 9 and 10. This model corroborates the importance of the randomness role in evolutionary systems whose primary function is to prevent particles from falling into local traps in an automatic manner. Besides that, it endows particles with the ability to explore unknown territories. Therefore, a certain amount of randomness is essential for the learning process. This randomness is charged with representing the state "I do not know" and lends itself as an effective "explorer of new features."

2. The particle competition model is constructed under a stochastic nonlinear dynamical system. In this regard, we provide an analytical analysis of the model, deriving probabilistic expressions that are able to predict the model's behavior as time progresses. A numerical validation confirms the theoretical predictions. In addition, we show that the model generalizes the process of single random walks to multiple interacting random-preferential walks in a competitive way. Such generalization is realized by calibrating the parameters of the model. A convergence analysis of the particle competition model is supplied. Therein, we show that the model does not converge to a fixed point, but instead it is confined within a certain region with a finite diameter. Furthermore, an upper bound of this region is estimated. Such a feature is similar to real-world systems due to the presence of noises and other uncontrolled variables.
3. A fuzzy index for detecting overlapping cluster or community structures in the network is explored. Most of the traditional community detection methods aim at assigning each vertex to a single community [23]. However, in real networks, vertices are often shared among different communities [23]. For example, in the language network composed of words as vertices, the word "bright" might be a member of several communities, such as those representing words related to the following subjects: "light," "astronomy," "color," "intelligence," and so on [42]. In a social network, each person naturally belongs to the communities of the company where he/she works and also to the community of his/her family at the same time. Therefore, uncovering overlapping community structure is important not only for network mining, but also for data analysis in general, once a data set is transformed into a network [21, 31, 41, 42, 47, 54, 62]. A drawback of traditional techniques that identify overlapping communities is that the overlap detection procedure is performed as a separated or dedicated process apart from the standard community detection technique. In this way, additional computational time is often required. As a result, the entire process may have high computational complexity. As we will see, the particle competition technique detects overlap community structures during the community detection procedure by using the dynamic variables generated by the particle competition process. As a result,

the particle competition method does not increase the overall model's time complexity order when employed to identify overlapping structures.

- With regard to the semi-supervised learning area:

  1. The model based on cooperation and competition of multiple particles is reviewed in detail [48]. This technique performs semi-supervised learning for classification tasks, as we will see in Chap. 10. A rigorous definition is provided, in which the particle competition is formally modeled from a stochastic nonlinear dynamical system. In essence, particles of the same class navigate in the network in a cooperative manner to propagate their labels, while particles of different classes compete with each other to determine class borders. Given that the model of several interacting particles corresponds to many natural and artificial systems, the study of this topic stands as an important task.
  2. Another interesting feature is that the particle competitive-cooperative model has a local label-spreading behavior. This property arises due to the competitive mechanism in which particles only visit portions of vertices potentially belonging to their teams. This can be roughly understood as a "divide-and-conquer" effect embedded in the scheme of competition and cooperation. In this way, many long-range redundant operations are avoided. As a result, the method has low computational complexity order. In contrast, traditional techniques of network-based semi-supervised learning normally rely on minimizing cost functions that ultimately lead to several matrix multiplication operations. Thus, the computational complexity of these techniques is usually of the order $\mathcal{O}(V^3)$ or higher [9, 10, 65], in which V is the number of vertices. Even though methods for enhancing matrix multiplication have been extensively studied,[2] the minimization of cost functions, which are usually based on a regularization framework, slows down the entire process [65]. It is expected that the models generated using competition of particles will be more efficient, which is important to treat large-scale databases.
  3. Since we construct the underlying network, in which the learning process is conducted, directly from the input data set, the correspondence between the input data and the processing result (the final network) is maintained. Consequently, the "black box" effect of artificial neural networks can be avoided to a large extent.
  4. The reliability of the labels is a crucial factor in a semi-supervised learning environment, because mislabeled samples may propagate wrong labels to a portion of or even to the entire data set. Here, we address the error propagation problem originated by these mislabeled samples by presenting detection and prevention processes embedded within the particle competition-cooperation model. Though this is an important topic, it has not received

[2]For instance, c.f. the generalized iterated matrix-vector multiplication technique, GMIV-M, proposed by [28].

much attention from researchers and there are still few works devoted to the study of semi-supervised learning from imperfect training data [4, 5, 25]. Usually, in supervised or semi-supervised learning, the input label information of the training data set is supposed to be completely reliable. However, in real situations, this is not always true and mislabeled samples are commonly found in the data sets due to instrumental errors, corruption from noise, or even human mistakes in the labeling process. For example, in a medical diagnostic system, the diagnostic results in the training set provided by doctors may be wrong. If these kinds of wrong labels are used to further classify new data (in the supervised learning case) or are propagated to unlabeled data (in the semi-supervised learning case), severe consequences may occur. This situation becomes more critical in autonomous learning, in which no external or minimal external intervention is involved. Thus, if the prior knowledge presented to the autonomous learning system contains errors, the performance of the learning system will get worse and worse because of the error propagation. Therefore, considering and designing mechanisms to prevent error propagation is important in the machine learning study and especially in the autonomous learning. Specifically, the prevention of error propagation can benefit the learning systems from two different aspects:

- Improvements of the performance of the learning system, i.e., the system can learn from errors;
- Avoidance of a system's catastrophe by limiting the spreading of wrong labels (input and generated errors).

To our knowledge, many semi-supervised learning techniques have been proposed [15], but the great majority considers that the label information of the labeled set is totally correct, i.e., there is no error prevention mechanism. In this way, such a mechanism make a clear contribution to general machine learning and especially to autonomous learning research.

- With regard to the supervised learning area:

1. A hybrid supervised learning framework composed of a convex combination of low- and high-level classifiers is studied in Chap. 8. Traditional supervised data classification considers only physical features (e.g., distance or similarity) of the input data. Here, this type of learning is called *low-level classification*. The human (animal) brain, in contrast, performs both low and high orders of learning and it has facility in identifying patterns according to the semantic meaning of the input data. Data classification that considers not only physical attributes but also the pattern formation is here referred to as *high-level classification*. The idea behind introducing the high-level term in a learning process is that data items often have patterns or organizational features that are left hidden within the numerous interrelationships among them. These, in turn, are not very well explored by traditional low-level classification techniques, as they are concerned with smoothness or cluster assumptions, which are in essence physical constraints. The high-level term precisely comes into play to fill in this gap by trying to find organizational

or structural features among different classes. Thus, it abstracts its decisions from physical constraints. As we will see, both the low- and high-level classifiers are often necessary to provide good accuracy rates, suggesting that they are complementary in the learning process.

2. Motivated by the intrinsic ability to describe topological structures among the data items, two types of high-level classification techniques have been proposed, all of which running in a networked environment. As a common goal, both realize the prediction scheme by extracting the features of the underlying network constructed from the input data. The high-level classification techniques are comprised of:
   - Three classical network measures borrowed from the complex network theory: assortativity, clustering coefficient, and network connectivity. The combination of these three measures can capture local to global structural patterns from the network topology, when utilized with reasonable low-level classifiers.
   - A weighted linear combination of tourist walks processes with different memory lengths. For this end, variations of dynamic variables generated by this deterministic process, the transient and cycle lengths, for different values of the tourist's memory length are employed. We show that, by adjusting the memory length of the tourist walker, we can systematically capture structural network features that range from local to global.
3. An interesting phenomenon uncovered using this hybrid classification framework is that, as the class complexity increases, a larger portion of the high-level term is required to get correct classification. The class complexity may be understood in terms of the mixture or overlap that exists among different classes. This feature confirms that the high-level classification has a special importance in complex situations of classification.

1.3 Organization of the Remainder of the Book

The remainder of this book is organized as follows. In Chap. 2, we review the complex network theory and dynamic processes that run in networks. In Chap. 3, we present the basic definitions of the machine learning area. These two first chapters elucidate the fundamental concepts that make way for the understanding of the developments reviewed and developed in this book.

Once the elementary theory is presented, in Chap. 4, we discuss the problem of constructing a network from unstructured data, which is a required step whenever we are dealing with data that is not yet in a network format. This task is crucial for network-based learning methods. Following that, we give a comprehensive review on network-based supervised learning, unsupervised learning, and semi-supervised learning in Chaps. 5–7, respectively.

In the second part of this book, we present the new developments of network-based machine learning in Chaps. 8–10. In Chap. 8, we delve into the supervised learning domain by showing a hybrid classification framework that derives its decision based on a convex combination of low- and high-level classifiers. This method is a pioneer across-network supervised learning technique. The model is analyzed in an empirical way and significant results are obtained. A real-world application (handwritten digits recognition) is supplied. In Chap. 9, we explore the unsupervised learning domain using a technique that is based on a particle competition model. Several experiments and mathematical investigations are performed, as well as a real-world application (handwritten digits and letters clustering). In Chap. 10, the particle competition model is extended to the semi-supervised learning domain. Likewise the previous chapter, we also display many experiments and mathematical investigations, as well as another real-world application (detection and prevention of mislabeled vertices).

References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47–97 (2002)
2. Albert, R., Jeong, H., Barabási, A.L.: Diameter of the world wide web. Nature **401**, 130–131 (1999)
3. Albert, R., Albert, I., Nakarado, G.L.: Structural vulnerability of the north american power grid. Phys. Rev. E **69**, 025103 (2004)
4. Amini, M.R., Gallinari, P.: Semi-supervised learning with explicit misclassification modeling. In: IJCAI 03: Proceedings of the 18th International Joint Conference on Artificial Intelligence, pp. 555–560. Morgan Kaufmann, San Francisco, CA (2003)
5. Amini, M.R., Gallinari, P.: Semi-supervised learning with an imperfect supervisor. Knowl. Inf. Syst. **8**(4), 385–413 (2005)
6. Arenas, A., Guilera, A.D., Pérez Vicente, C.J.: Synchronization reveals topological scales in complex networks. Phys. Rev. Lett. **96**(11), 114102 (2006)
7. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science (NY) **286**(5439), 509–512 (1999)
8. Barrat, A., Barthélemy, M., Vespignani, A.: Dynamical Processes on Complex Networks. Cambridge University Press, Cambridge (2008)
9. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. **15**(6), 1373–1396 (2003)
10. Belkin, M., Matveeva, I., Niyogi, P.: Regularization and semi-supervised learning on large graphs. In: Shawe-Taylor, J., Singer, Y. (eds.) Learning Theory. Lecture Notes in Computer Science, vol. 3120, pp. 624–638. Springer, Berlin, Heidelberg (2004)
11. Bishop, C.M.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer, Berlin (2007)
12. Boccaletti, S., Bianconi, G., Criado, R., del Genio, C., Gómez-Gardeñes, J., Romance, M., Sendiña-Nadal, I., Wang, Z., Zanin, M.: The structure and dynamics of multilayer networks. Phys. Rep. **544**(1), 1–122 (2014)
13. Bornholdt, S., Schuster, H.G.: Handbook of Graphs and Networks: From the Genome to the Internet. Wiley-VCH, Weinheim (2003)
14. Castro Miranda, R.C., Stancato de Souza, S.R., Silva, T.C., Tabak, B.M.: Connectivity and systemic risk in the brazilian national payments system. J. Complex Networks **2**(4), 585–613 (2014)

15. Chapelle, O., Schölkopf, B., Zien, A. (eds.): Semi-Supervised Learning. Adaptive Computation and Machine Learning. MIT, Cambridge, MA (2006)
16. da Silva, M., Ma, H., Zeng, A.P.: Centrality, network capacity, and modularity as parameters to analyze the core-periphery structure in metabolic networks. Proc. IEEE **96**(8), 1411–1420 (2008)
17. Danon, L., Díaz-Guilera, A., Duch, J., Arenas, A.: Comparing community structure identification. J. Stat. Mech. Theory Exp. **2005**(09), P09008 (2005)
18. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of Networks: From Biological Nets to the Internet and WWW (Physics). Oxford University Press, USA (2003)
19. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley-Interscience (2000)
20. Erdös, P., Rényi, A.: On random graphs I. Publ. Math. Debr. **6**, 290–297 (1959)
21. Evans, T.S., Lambiotte, R.: Line graphs, link partitions, and overlapping communities. Phys. Rev. E **80**(1), 016105 (2009)
22. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the internet topology. In: SIGCOMM 99: Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, vol. 29, pp. 251–262. ACM, New York (1999)
23. Fortunato, S.: Community detection in graphs. Physics Reports **486**, 75–174 (2010)
24. Fortunato, S., Latora, V., Marchiori, M.: Method to find community structures based on information centrality. Phys. Rev. E **70**(5), 056104 (2004)
25. Hartono, P., Hashimoto, S.: Learning from imperfect data. Appl. Soft Comput. **7**(1), 353–363 (2007)
26. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.L.: The large-scale organization of metabolic networks. Nature **407**(6804), 651–654 (2000)
27. Jiang, Y., Jiang, J.: Understanding social networks from a multiagent perspective. IEEE Trans. Parallel Distrib. Syst. **25**(10), 2743–2759 (2014)
28. Kang, U., Tsourakakis, C.E., Faloutsos, C.: PEGASUS: mining peta-scale graphs. J. Knowl. Inf. Syst. **27**(2), 303–325 (2011)
29. Karypis, G., Han, E.H., Kumar, V.: Chameleon: hierarchical clustering using dynamic modeling. Computer **32**(8), 68–75 (1999)
30. Lambiotte, R., Delvenne, J.C., Barahona, M.: Random walks, Markov processes and the multiscale modular organization of complex networks. IEEE Trans. Netw. Sci. Eng. **1**(2), 76–90 (2014)
31. Lancichinetti, A., Fortunato, S., Kertész, J.: Detecting the overlapping and hierarchical community structure in complex networks. New J. Phys. **11**(3), 033015 (2009)
32. Milgram, S.: The small world problem. Psychol. Today **2**, 60–67 (1967)
33. Mitchell, T.M.: Machine Learning. McGraw-Hill Science/Engineering/Math, New York (1997)
34. Mizruchi, M.S.: The American corporate network. Sage **2**, 1904–1974 (1982)
35. Montoya, J.M., Solée, R.V.: Small world patterns in food webs. J. Theor. Biol. **214**, 405–412 (2002)
36. Newman, M.E.J.: The structure and function of complex networks. SIAM Rev. **45**(2), 167–256 (2003)
37. Newman, M.E.J.: Fast algorithm for detecting community structure in networks. Phys. Rev. E **69**(6), 066133 (2004)
38. Newman, M.E.J.: Modularity and community structure in networks. Proc. Natl. Acad. Sci. **103**(23), 8577–8582 (2006)
39. Newman, M.E.J.: Networks: An Introduction. Oxford University Press, Oxford (2010)
40. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. Lett. **69**, 026113 (2004)
41. Nicosia, V., Mangioni, G., Carchiolo, V., Malgeri, M.: Extending the definition of modularity to directed graphs with overlapping communities. J. Stat. Mech. Theory Exp. **2009**(03), 03024 (2009)
42. Palla, G., Derenyi, I., Farkas, I., Vicsek, T.: Uncovering the overlapping community structure of complex networks in nature and society. Nature **435**(7043), 814–818 (2005)

43. Poledna, S., Molina-Borboa, J.L., Martínez-Jaramillo, S., van der Leij, M., Thurner, S.: The multi-layer network nature of systemic risk and its implications for the costs of financial crises. J. Financ. Stab. **20**, 70–81 (2015)
44. Quiles, M.G., Zhao, L., Alonso, R.L., Romero, R.A.F.: Particle competition for complex network community detection. Chaos **18**(3), 033107 (2008)
45. Reichardt, J., Bornholdt, S.: Detecting fuzzy community structures in complex networks with a potts model. Phys. Rev. Lett. **93**(21), 218701(1–4) (2004)
46. Scott, J.P.: Social Network Analysis: A Handbook. SAGE, Beverly Hills, CA (2000)
47. Shen, H., Cheng, X., Cai, K., Hu, M.B.: Detect overlapping and hierarchical community structure in networks. Physica A **388**(8), 1706–1712 (2009)
48. Silva, T.C., Zhao, L.: Network-based stochastic semisupervised learning. IEEE Trans. Neural Netw. Learn. Syst. **23**(3), 451–466 (2012)
49. Silva, T.C., Zhao, L.: Stochastic competitive learning in complex networks. IEEE Trans. Neural Netw. Learn. Syst. **23**(3), 385–398 (2012)
50. Singh, A., Nowak, R.D., Zhu, X.: Unlabeled data: now it helps, now it doesn't. In: The Conference on Neural Information Processing Systems NIPS, pp. 1513–1520 (2008)
51. Souza, S.R., Tabak, B.M., Silva, T.C., Guerra, S.M.: Insolvency and contagion in the brazilian interbank market. Physica A **431**, 140–151 (2015)
52. Sporns, O.: Networks analysis, complexity, and brain function. Complexity **8**(1), 56–60 (2002)
53. Strogatz, S.H.: Exploring complex networks. Nature **410**(6825), 268–276 (2001)
54. Sun, P.G., Gao, L., Shan Han, S.: Identification of overlapping and non-overlapping community structure by fuzzy clustering in complex networks. Inf. Sci. **181**, 1060–1071 (2011)
55. Wang, P., Yu, X., Lu, J.: Identification and evolution of structurally dominant nodes in protein-protein interaction networks. IEEE Trans. Biomed. Circuits Syst. **8**(1), 87–97 (2014)
56. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**(6684), 440–442 (1998)
57. Weng, J., Luciw, M.: Brain-inspired concept networks: Learning concepts from cluttered scenes. IEEE Intell. Syst. **29**(6), 14–22 (2014)
58. West, G.B., Brown, J.H., Enquist, B.J.: A general model for the structure, and algometry of plant vascular systems. Nature **400**, 122–126 (1999)
59. Yang, J., Leskovec, J.: Overlapping communities explain core-periphery organization of networks. Proc. IEEE **102**(12), 1892–1902 (2014)
60. Zhao, L., Park, K., Lai, Y.C.: Attack vulnerability of scale-free networks due to cascading breakdown. Phys. Rev. E **70**, 035101(1–4) (2004)
61. Zhao, L., Park, K., Lai, Y.C.: Tolerance of scale-free networks against attack-induced cascades. Phys. Rev. E (Rapid Commun.) **72**(2), 025104(R)1–4 (2005)
62. Zhang, S., Wang, R.S., Zhang, X.S.: Identification of overlapping community structure in complex networks using fuzzy C-Means clustering. Physica A **374**(1), 483–490 (2007)
63. Zhao, L., Cupertino, T.H., Park, K., Lai, Y.C., Jin, X.: Optimal structure of complex networks for minimizing traffic congestion. Chaos **17**(4), 043103(1–5) (2007)
64. Zhou, H.: Distance, dissimilarity index, and network community structure. Phys. Rev. E **67**(6), 061901 (2003)
65. Zhou, D., Bousquet, O., Lal, T.N., Weston, J., Schölkopf, B.: Learning with local and global consistency. In: Advances in Neural Information Processing Systems, vol. 16, pp. 321–328. MIT, Cambridge, MA (2004)