Broad Learning Introduction

1.1 What Is Broad Learning

It was six men of Indostan
To learning much inclined,
"Who went to see the Elephant"
(Though all of them were blind),
That each by observation
Might satisfy his mind.

The First approached the Elephant,
And happening to fall
Against his broad and sturdy side,
At once began to bawl:
"God bless me! but the Elephant
Is very like a WALL!"

The Second, feeling of the tusk,
Cried, "Ho, what have we here,
So very round and smooth and sharp?
To me 'tis mighty clear
This wonder of an Elephant
Is very like a SPEAR!"

The Third approached the animal,
And happening to take
The squirming trunk within his hands,
Thus boldly up and spake:
"I see," quoth he, "the Elephant
Is very like a SNAKE!"

The Fourth reached out an eager hand,
And felt about the knee
"What most this wondrous beast is like
Is mighty plain," quoth he:
"Tis clear enough the Elephant
Is very like a TREE!"

The Fifth chanced to touch the ear,
Said: "E'en the blindest man
Can tell what this resembles most;
Deny the fact who can,
This marvel of an Elephant
Is very like a FAN!"

The Sixth no sooner had begun
About the beast to grope,
Than seizing on the swinging tail
That fell within his scope,
"I see," quoth he, "the Elephant
Is very like a ROPE!"

And so these men of Indostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong,
Though each was partly in the right,
And all were in the wrong!

— John Godfrey Saxe (1816–1887)

We would like to start this book with an ancient story about "*The Blind Men and the Elephant*" from *John Godfrey Saxe*. This story is a famous Indian fable about six blind sojourners who come across different parts of an elephant in their life journeys. In turn, each blind man creates his own version of reality from that limited experiences and perspectives. Instead of explaining its philosophical meanings, we indent to use this story to illustrate the current situations that both the academia and industry are facing about *artificial intelligence*, *machine learning*, and *data mining*.

In the real world, on the same information entities, e.g., products, movies, POIs (points-of-interest), and even human beings, a large amount of information can actually be collected from various sources. These information sources are usually of different varieties, like Walmart vs. Amazon for commercial products, IMDB vs. Rotten Tomatoes for movies, Yelp vs. Foursquare for POIs, and various online social media websites for human beings. Each information source provides a specific signature of the same entity from a unique underlying aspect. Meanwhile, in most of the cases, these information sources are normally separated from each other without any correlations. An effective fusion of these different information sources will provide an opportunity for researchers and practitioners to understand the entities more comprehensively, which renders *broad learning* to be introduced in this book an extremely important problem.

Broad learning initially proposed in [52, 54, 56] is a new type of learning task, which focuses on fusing multiple large-scale information sources of diverse varieties together and carrying out synergistic data mining tasks across these fused sources in one unified analytic. Fusing and mining multiple information sources are both the fundamental problems in broad learning studies. Broad learning investigates the principles, methodologies, and algorithms for synergistic knowledge discovery across multiple fused information sources, and evaluates the corresponding benefits. Great challenges exist in broad learning for the effective fusion of relevant knowledge across different aligned information sources, which depends upon not only the relatedness of these information sources but also the target application problems. Broad learning aims at developing general methodologies, which will be shown to work for a diverse set of applications, while the specific parameter settings can be effectively learned from the training data. A recent survey article about broad learning is available at [39].

1.2   Problems and Challenges of Broad Learning

Broad learning is a novel yet challenging learning task. The main problems covered in broad learning include *information fusion* and *knowledge discovery* of multiple data sources. Meanwhile, there exist great challenges to address these two tasks due to both the *diverse data inputs* and *various application scenarios* in the real-world problem settings.

1.2.1   Cross-Source Information Fusion

One of the key tasks in broad learning is the fusion of information from different sources, which can be done at different levels, e.g., *raw data level*, *feature space level*, *model level*, and *output level*. The specific fusion techniques used at different levels may have significant differences.

- *Raw data level*: In the case where the data from different sources are of the same modality, e.g., textual data, and have no significant information distribution differences, such kinds of data can be effectively fused at the raw data level.
- *Feature space level*: Based on the data from different sources, a set of features can be effectively extracted, which can be fused together via simple feature vector concatenation (if there exist no significant distribution differences) or feature vector embedding and transformation (to accommodate the information distribution differences).
- *Model level*: In some cases, with the information from different sources, a set of learning models can be effectively trained for each of the sources, so as to fuse the multi-source information at the model level.
- *Output level*: Based on the learned results from each of the sources, they can be effectively fused to output a joint result, which will define the information fusion task at the output level.

In broad learning, the "multiple sources" term is actually a general concept, which may refer to the different information *views*, *categories*, *modalities*, *concrete sources*, and *domains* depending on the specific application settings. We will illustrate these concepts with examples as follows, respectively.

- *Multi-view*: In the traditional webpage ranking problem, information about webpages can be represented in two different views: *textual contents* and *hyperlinks*, which provide complimentary information for learning the ranking scores of the webpages.
- *Multi-categories*: In online social networks, users can perform various categories of social activities, e.g., *making friends, writing posts, checking-in at locations, etc*. Each of these social activities will generate a category of information providing crucial information to help discover the social patterns of users.
- *Multi-modality*: In recent multimedia studies, the information entities (e.g., news articles) usually have information represented in different modalities, including the *textual title, textual content, images*, and *live interview videos*. Integration of such diverse information modalities together will improve the learning performance of models designed for the news articles greatly.
- *Multi-source*: Nowadays, to enjoy more social network services, users are usually involved in multiple online social networks simultaneously, e.g., Facebook, LinkedIn, Twitter, and Foursquare. Information in each of these social network sources can reveal the users' social behaviors from different perspectives.
- *Multi-domain*: Compared with the concepts aforementioned, *domain* is usually used in mathematics, which refers to a set of possible values of independent variables that a function/relation is defined on. Multi-domain learning is a special case of broad learning tasks which fuses the information from different groups of relevant or irrelevant information entities.

Meanwhile, as illustrated in Fig. 1.1, depending on the information flow directions, information fusion in broad learning can be done in different manners, e.g., *information immigration*, *information exchange*, and *information integration*.

- *Information immigration*: In many applications of broad learning, there will exist a specific target source, which can be very sparse and short of useful information. Immigration of information from external mature sources to the target source can hopefully resolve such a problem.
- *Information mutual exchange*: Mutual information exchange is another common application scenario in broad learning, where information will be immigrated among all these sources mutually. With the information from all these sources, application problems studied in all these data sources can benefit from the abundant data simultaneously.
- *Information integration*: Another common application setting of broad learning is the integration of information from all the data sources together, where there exists no target source at all. Such an application setting is normally used in the profiling problem of the information entities shared across different data sources, the fused information from which will lead to a more comprehensive knowledge about these entities.

In this book, we will take online social networks as an example to illustrate the broad learning problems and algorithms. Formally, the information fusion across multiple online social networks is also called the *network alignment* problem [18, 42, 43, 47, 50, 51], which aims at inferring the set of *anchor links* [18, 46] mapping users across networks.

Information Immigration   **Information Exchange**   **Information Integration**

Fig. 1.1 Information immigration vs. information exchange vs. information integration

1.2.2 Cross-Source Knowledge Discovery

Based on the fused data sources, various concrete application problems can be studied, which will also benefit from the fused data greatly. For instance, with the multiple fused online social network data sets, research problems, like *link prediction* [21, 44–46, 51, 53], *community detection* [7, 15, 40, 41], *information diffusion* [10, 37, 38, 48, 49], *viral marketing* [16, 37, 38, 48, 49], and *network embedding* [2, 27, 52] can all be improved significantly.

- *Link prediction*: With the fused information from multiple aligned network data sources, we will have a better understanding about the connection patterns among the social network users. The link prediction results will be greatly improved with the fused network data sets.
- *Community detection*: According to the connections accumulated from the multiple online social networks, we can also obtain a clearer picture about the network community structures formed by the users. For the social networks with extremely sparse social connections, the data transferred from external sources can hopefully recover the true communities of users in the network.
- *Information diffusion*: Via the social interactions among users, information of various topics can propagate among users through various diffusion channels. Besides the intra-network information diffusion channels, due to the cross-network information sharing, information can also propagate across social networks as well, which will lead to broader impacts actually.
- *Viral marketing*: To carry out commercial advertising and product promotion activities, a necessary marketing strategy is required (e.g., the initial seed users selected to propagate the product information), which can guide the commercial actions during the promotion process. Due to the cross-source information diffusion, the marketing strategy design should consider the information from multiple networks simultaneously.
- *Network embedding*: In recent years, due to the surges of deep learning, representation learning has become a more and more important research problem, which aims at learning the feature vectors characterizing the properties of information entities. Based on the fused social networks, more information about the users can be collected, which can be used for more effective representation feature vector learning for the users.

1.2.3 Challenges of Broad Learning

In the two tasks covered in the broad learning problem mentioned above, great challenges may exist in both fusing and mining the multiple information sources. We categorize the challenges into two main groups as follows:

- *How to fuse*: The data fusion strategy is highly dependent on the data types, and different data categories may require different fusion methods. For instance, to fuse image sources about the same entities, a necessary entity recognition step is required; to combine multiple online social networks, inference of the potential anchor links mapping the shared users across networks will be a key task; meanwhile, to integrate diverse textual data, concept entity extraction or topic modeling can both be the potential options. In many cases, the fusion strategy is also correlated with the specific applications to be studied, which may pose extra constraints or requirements on the fusion results. More information about related data fusion strategies will be introduced later in Part II of this book.
- *How to learn*: To learn useful knowledge from the fused data sources, there also exist many great challenges. In many of the cases, not all the data sources will be helpful for certain application tasks. For instance, in the social community detection task, the fused information about the users' credit card transaction will have less correlation with the social communities formed by the users. On the other hand, the information diffusion among users is regarded as irrelevant with the information sources depicting the daily commute routes of people in the real world. Among all these fused data sources, picking the useful ones is not an easy task. Several strategies, like meta path weighting/selection, information source importance scaling, and information domain embedding, will be described in the specific application tasks to be introduced in Part III of this book.

1.3 Comparison with Other Learning Tasks

There exist great differences of broad learning with other existing learning tasks, e.g., *deep learning* [9], *ensemble learning* [55], *transfer learning* [26], *multi-task learning* [4], and *multi-view* [36], *multi-source* [5], *multi-modal* [25], *multi-domain* [24, 35] learning tasks. In this part, we will provide a brief comparison of these learning tasks to illustrate their correlations and differences.

1.3.1 Broad Learning vs. Deep Learning

As illustrated in Fig. 1.2, deep learning [9] is a rebranding of multi-layer neural network research works. Deep learning is "deep" in terms of the model hidden layers connecting the input to the output space. Generally, the learning process in most machine learning tasks is to achieve a good projection between the input and the output space. In addition, for lots of the application tasks, such a mapping can be very complicated and is usually non-linear. With more hidden layers, the deep learning models will be capable to capture such a complicated projection, which has been demonstrated with the successful application of deep learning in various areas, e.g., speech and audio processing [6, 13], language modeling and processing [1, 23], information retrieval [11, 29], objective recognition and computer vision [20], as well as multi-modal and multi-task learning [32, 33].

However, broad learning actually focuses on a very different perspective instead. Broad learning is "broad" in terms of both its input data variety and the learning model components. As introduced before, the input data sources of broad learning are of very "broad" varieties, including text, image, video, speech, and graph. Meanwhile, to handle such diverse input sources, the broad learning model should have broad input components to process them simultaneously. In these input raw data representation learning components, deep learning and broad learning can also work hand in hand. Deep learning can convert these different modality representations of text, image, video, speech, and graph all into feature vectors, which can make them easier to fuse in broad learning via the
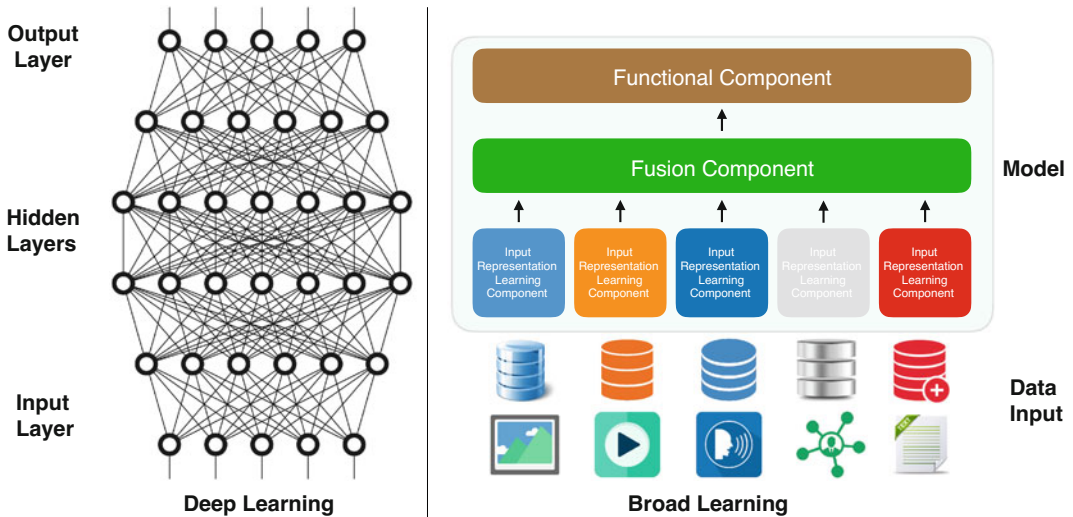
Fig. 1.2 An comparison of deep learning vs. broad learning

fusion component as shown in Fig. 1.2. Viewed in such a perspective, broad learning is broad in both its model input components and information fusion component, where deep learning can serve an important role for the broad input representation learning.

1.3.2   Broad Learning vs. Ensemble Learning

As shown in Fig. 1.3, ensemble learning [55] is a machine learning paradigm where multiple unit models will be trained and combined to solve the same problem. Traditional machine learning models usually employ one hypothesis about the training data, while ensemble learning methods propose to construct a set of hypotheses on the training data and combine them together for the learning purposes. In most of the cases, ensemble learning is carried out based on one single training set, with which a set of unit learning models will be trained and combined to generate the consensus output.

Generally speaking, broad learning works in a very different paradigm compared with ensemble learning. Broad learning aims at integrating diverse data sources for learning and mining problems, but ensemble learning is usually based on one single data input. Meanwhile, broad learning also has a very close correlation with ensemble learning, as ensemble learning also involves the "information integration" step in the model building. Depending on the specific layers for information fusion, many techniques proposed in ensemble learning can also be used for broad learning tasks. For instance, when the information fusion is actually done at the output level in broad learning, the existing techniques, e.g., *boosting* and *bagging*, used in ensemble learning can be adopted as the fusion method for generating the output in broad learning.

1.3.3   Broad Learning vs. Transfer Learning vs. Multi-Task Learning

Traditional transfer learning [26] focuses on the immigration of knowledge from one source to the other source, which may share the same information entities or can be totally irrelevant. Transfer learning problems and models are mostly proposed based on the assumption that the data in different
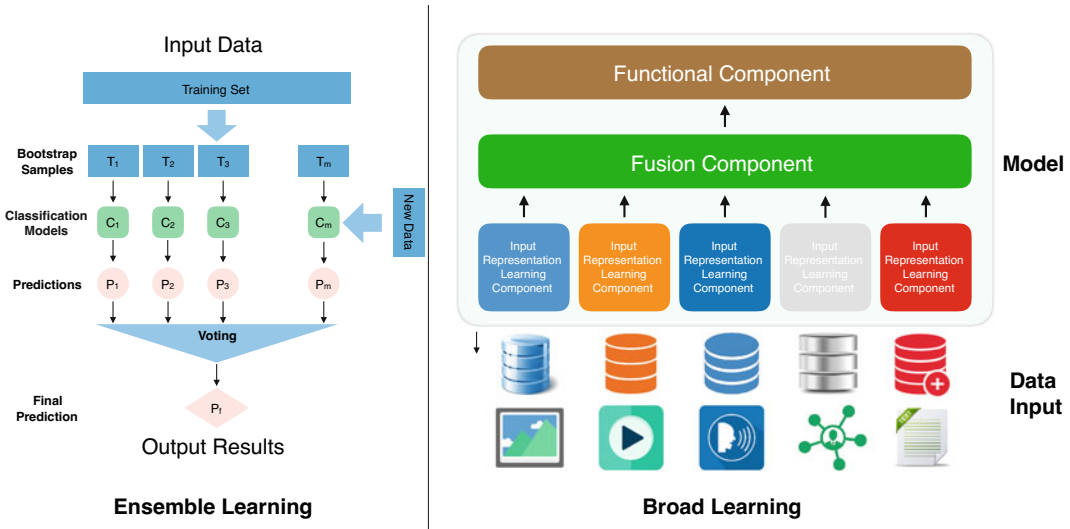
Fig. 1.3 A comparison of ensemble learning vs. broad learning

domains may follow a similar distribution (either in the original feature space or in a certain latent feature space), where the models trained based on external data sources can also be used in the target source. Transfer learning has demonstrated its advantages in overcoming the information shortage problems in many applications.

Meanwhile, for the multi-task learning task [4], it focuses on studying the simultaneous learning of multiple tasks at the same time, where each task can optimize its learning performance through some shared knowledge with other tasks. Furthermore, in the case where there exists a sequential relationship among these learning tasks, the problem will be mapped to the lifelong learning [30] instead.

In a certain sense, both transfer learning and multi-task learning can be treated as a special case of broad learning. Transfer learning aims at immigrating knowledge across sources, which is actually one of the learning paradigms of broad learning as introduced before. Meanwhile, according to the descriptions in Sect. 1.2.1, broad learning allows more flexible information flow directions in the learning process. For the multi-task learning, it studies the learning problem of multiple tasks simultaneously, which can be reduced to the broad learning problem with each task studied based on a separate data set, respectively.

1.3.4   Broad Learning vs. Multi-View, Multi-Source, Multi-Modal, Multi-Domain Learning

Considering that the multiple data "sources" studied in broad learning can have different physical meanings depending on the specific learning settings as introduced in Sect. 1.2.1, which provides a great flexibility for applying broad learning in various learning scenarios. Generally, as illustrated in Fig. 1.4, the multi-view [36], multi-source [5], multi-modal [25], and multi-domain [24, 35] learning tasks can all be effectively covered in the broad learning concept, and the research works in these learning tasks will also greatly enrich the development of broad learning.

Fig. 1.4 An comparison
of ensemble learning vs.
broad learning
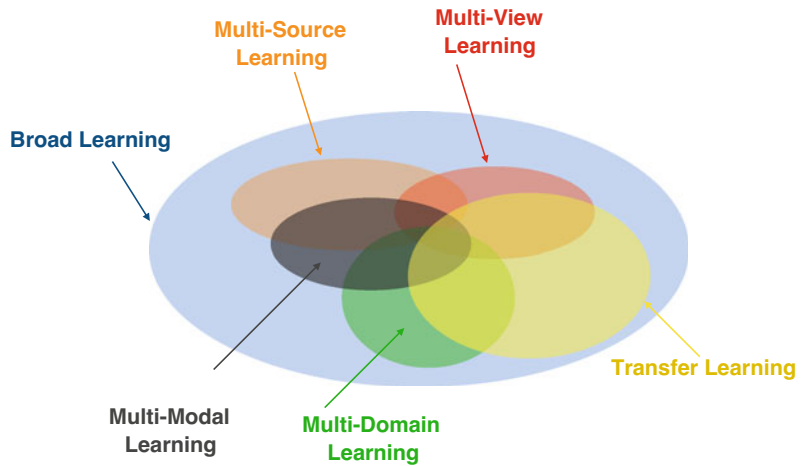
1.4   Book Organization

This book has four main parts. The first part, covering Chaps. 1–3, will provide the basic essential background knowledge about *broad learning*, *machine learning*, and *social networks*. In the following parts of this book, we will use online social networks as an application setting to introduce the problems and models of broad learning. In the second part, i.e., Chaps. 4–6, this book mainly focuses on introducing the existing online social network alignment concepts, problems, and algorithms, which is also the prerequisite step for broad learning across multiple online social networks. After that, Chaps. 7–11 will make up the third part of this book, which will talk about the application problems that can be studied across multiple online aligned social networks, including *link prediction*, *community detection*, *information diffusion*, *viral marketing*, and *network embedding*. Finally, in the fourth part of this book, i.e., Chap. 12, some potential future research directions and opportunities about broad learning will be provided as a conclusion for this book.

1.4.1   Part I

The first part of the book covers three chapters, and will provide some basic background knowledge of *broad learning*, *machine learning*, and *social network* to make this book self-contained.

- **Chapter 1: Broad Learning Introduction.** Broad learning is a new type of learning and knowledge discovery task that emerge in recent years. The first chapter of this book illustrates the definitions of broad learning, the motivations to study broad learning problems, and also provides the main problems and challenges covered in broad learning in various concrete applications. To help the readers understand this book better, Chap. 1 also includes two sections of reading instructions, including the potential readers of this book as well as how to read this book.
- **Chapter 2: Machine Learning Overview.** Broad learning introduced in this book is based on the existing machine learning works. Chapter 2 introduces some basic background knowledge about machine learning, including data representation, supervised learning, unsupervised learning, deep learning, and some frequently used evaluation metrics.

- **Chapter 3: Social Network Overview.** Online social networks can be formally represented as graphs involving both various kinds of nodes and complex links. Before talking about the problems and models, some essential knowledge about social networks will be provided in Chap. 3, including graph essentials, network categories and measures, network models, and the meta path concept.

1.4.2 Part II

The second part of the book includes Chaps. 4–6. Depending on the availability of training data, different categories of network alignment models have been proposed to solve the social network alignment problem based on the supervised, unsupervised, and semi-supervised learning settings, respectively.

- **Chapter 4: Supervised Network Alignment.** In the case where a set of existing and non-existing anchor links are labeled and available, supervised network alignment models can be built based on the labeled data, where the existing and non-existing anchor links can be used as the positive and negative instances, respectively. Chapter 4 will introduce several online social network alignment models based on the supervised learning setting.
- **Chapter 5: Unsupervised Network Alignment.** In the real scenarios, the anchor links are very expensive to label, which will introduce very large costs in terms of time, money, and labor resources. In such a case, unsupervised network alignment models requiring no training data at all can be a great choice, which will be introduced in Chap. 5 in great detail.
- **Chapter 6: Semi-supervised Network Alignment.** In many other cases, a small proportion of the anchor link labels can be retrieved subject to certain cost constraints, while majority of the remaining anchor links are still unlabeled. Inferring the other potential anchor links aligning the networks with both labeled and unlabeled anchor links is called the semi-supervised network alignment problem, which will be introduced in Chap. 6.

1.4.3 Part III

By aligning different online social networks together, various application tasks can benefit from the information fused from these different information sources. This part will introduce five different knowledge discovery application problems across aligned online social networks.

- **Chapter 7: Link Prediction.** Given a screenshot of an online social network, inferring the connections to be formed among users in the future is named as the link prediction problem. Various real-world services can be cast as the link prediction problem. For instance, friend recommendation can be modeled as a friendship link prediction problem, while location check-in inference can be treated as a location link prediction problem. Chapter 7 covers a comprehensive introduction of the link prediction problem within and across aligned online social networks.
- **Chapter 8: Community Detection.** Social communities denote groups of users who are strongly connected with each other inside each groups but have limited connections to users in other groups. Discovering the social groups formed by users in online social networks is named as the community detection problem, and correctly detected social communities can be important for many social network services. Chapter 8 will introduce the community detection problem within and across aligned online social networks in detail.
- **Chapter 9: Information Diffusion.** Via the social interactions among users, information will propagate from one user to another, which is modeled as the information diffusion process in online

social networks. Chapter 9 focuses on introducing the information diffusion problem specifically, where different categories of diffusion models will be talked about. Across aligned social networks, several state-of-the-art information diffusion models will also be covered in Chap. 9.

- **Chapter 10: Viral Marketing.** Based on the information diffusion model, influence can be effectively propagated among users. By selecting a good group of initial seed users who will spread out the information, real-world product promotions and election campaigns usually aim at achieving the maximum influence in online social networks. Chapter 10 covers the viral marketing problems across online social networks.
- **Chapter 11: Network Embedding.** Most existing machine learning algorithms usually take the feature representation data as the input, which can hardly be applied to handle the network structured data directly. One way to resolve such a problem is to apply network embedding to extract the latent feature representations of the nodes, where those extracted features should also preserve the network structure information. Chapter 11 will focus on the network embedding problem in detail.

1.4.4   Part IV

Broad learning is a novel yet important area, and there exist adequate opportunities and new research problems to be studied. The current broad learning research works also suffer from several big problems, which will be the future research directions as well. In the last part of the book, a big picture of broad learning will be provided and some potential future development directions in this area will also be illustrated, which altogether will form the conclusion of this book.

- **Chapter 12: Frontier and Future Work Directions.** Chapter 12 will provide some other frontier broad learning applications in various areas. Besides the works introduced in the previous chapters of this book, several promising future development directions of broad learning will be illustrated in Chap. 12. The data sets available in the real world are usually of a very large scale, which can involve millions even billions of data instances and attributes. Broad learning across multiple aligned data sources renders the scalability problem more severe. What's more, existing broad learning algorithms are mostly based on pairwise aligned data sources. The more general broad learning problems across multiple (more than two) aligned data sources simultaneously can be another potential development direction. In addition, this book mainly focuses on broad learning in online social networks, but broad learning is a general multiple aligned source learning problem, can actually be applied in many other different domains as well. Finally, at the end of this book, Chap. 12 will draw a conclusion about broad learning.

1.5   Who Should Read This Book

The readers of this book are oriented at the senior undergraduate, graduate students, academic researchers, industrial practitioners, and project managers interested in broad learning and social network mining, data mining, and machine learning.

Readers with a computer science background and have some basic knowledge about data structure, programming, graph theory, and algorithms will find this book to be easily accessible. Some basic knowledge about linear algebra, matrix computation, statistics, and probability theory will be helpful for readers to understand the technique details of the models and algorithms covered in the book. Having prior knowledge about data mining, machine learning, computational social science is a plus, but not necessary.

This book can be used as the textbook in the computational social science, social media mining, data science, data mining, and machine learning application courses for both undergraduate and graduate students. This book is organized in a way that can be taught in one semester to students with background about computer science, statistics, and probability. This book can also be used as the seminar course reading materials for graduate students interested in doing research about data mining, machine learning, and social media mining. The learning materials (including course slides, tutorials, data sets, and toolkit package) will be provided at the IFM Lab broad learning webpage.[1]

Moreover, this book can also be used as the reference book for researchers, practitioners, and project managers of related fields who are interested in learning the basics and tangible examples of this emerging field. This book can help these readers to understand the potentials and opportunities of broad learning in academic research, system/model development, and applications in the industry.

1.6 How to Read This Book

As introduced in the previous sections, this book has four parts. Part I (Chaps. 1–3) covers the introductory materials of broad learning, machine learning, and social networks. Part II (Chaps. 4–6) covers the network alignment problems and models based on different learning settings. Part III (Chaps. 7–11) covers the application problems across multiple aligned social networks. Part IV (Chap. 12) covers the discussions about the future development directions and opportunities of broad learning.

1.6.1 To Readers

We recommend the readers to read this book in the order as how this book is organized from Part I–IV, but the chapters covered in each part can be read in any order (Exception: Chap. 9 had better be read ahead of or together with Chap. 10). Depending on the expertise of the students, some chapters can be skipped or skimmed through quickly. For the readers with background about machine learning, Chap. 2 can be skipped. For readers having background about social media and graph theory, Chap. 3 can be skipped. For readers doing research in social media mining projects on link prediction, community detection, information diffusion, marketing strategy, and network embedding, the introductory materials of traditional single-network based models provided at the beginning of Chaps. 7–11 can be skipped.

1.6.2 To Instructors

This book can be taught in one semester (28 courses, 1.25 h per course) in the order as the how the book is organized from Parts I to IV. The chapters covered in each part can be taught in any sequence (Exception: Chap. 9 had better be delivered ahead of or together with Chap. 10). Depending on the prior knowledge that the students have, certain chapters can be skipped or leave for the students to read after class. The first three introductory chapters of Part I can be delivered in six courses, and Chaps. 2 and 3 together can take five courses. The three chapters about network alignment in Part II can be taught in six courses, where each chapter takes two courses. These five chapters covered in Part III can be delivered in 15 courses, and each chapter takes 3 courses. The last chapter can leave free for the students to read if they are interested in and also plan to further explore this area.

---

[1]http://www.ifmlab.org/BroadLearning.html.

1.6.3 Supporting Materials

Updates to chapters, teaching materials (including source slides, tutorials, slides, courses, data sets, learning toolkit packages, and other resources) are available at the IFM Lab broad learning webpage (see footnote 1) as mentioned in Sect. 1.5.

1.7   Summary

In this chapter, we have provided the introduction to *broad learning*, which is a new type of learning tasks focusing on fusing multiple information sources together for synergistic knowledge discovery. The motivation to study *broad learning* is due to the scattered distributions of data in the real world. An effective fusion of these different information sources will provide an opportunity for researchers and practitioners to understand the target information entities more comprehensively.

We presented the main problems covered in *broad learning*, which include *information fusion* and *cross-source knowledge discovery*. Great challenges exist in addressing these two tasks due to both the diverse data inputs and various application scenarios in the real-world problem settings. We briefly introduced various information fusion methods and manners. We also took online social networks as an example to illustrate five different application problems that we can study across the fused information sources.

We showed the comparison of *broad learning* with several other learning tasks, including *deep learning*, *ensemble learning*, *transfer learning*, *multi-task learning*, as well as *multi-view*, *multi-source*, *multi-modal*, and *multi-domain* learning tasks. We introduced the differences of these learning tasks, and also illustrated the correlations among these different learning problems.

We provided a brief introduction about the organization of this book, which involves four main parts: (1) Part I (three chapters): background knowledge introduction; (2) Part II (three chapters): social network alignment problem and algorithms; (3) Part III (five chapters): application problems studied across the aligned social networks; and (4) Part IV (one chapter): potential future development directions.

We also indicated the potential readers of this book, including senior undergraduate, graduate students, academic researchers, industrial practitioners, and project managers interested in broad learning. For both the readers and the instructors, we provided the reading and teaching guidance as well. Useful supporting materials would also be provided at the book webpage, which could help the readers, instructors, and students to follow the contents covered in this book more easily.

1.8   Bibliography Notes

The *broad learning* concept initially proposed by the authors of this textbook in their research papers [52, 54, 56] is an important problem in both academia and industry. Lots of prior works have been done by the book authors to study both the *fusion* of multiple information sources, e.g., the fusion of online social networks via alignment [18, 42, 43, 47, 50, 51], and the application problems that can be analyzed based on the fused information sources, e.g., *link prediction* [44–46, 51, 53], *community detection* [15, 40, 41], *information diffusion* [37, 38, 48, 49], *viral marketing* [37, 38, 48, 49], and *network embedding* [52]. In the following chapters of this book, we will introduce these research works in great detail.

The essence of *deep learning* is to compute hierarchical feature representations of the observational data [9, 20]. With the surge of deep learning research in recent years, lots of research works have

appeared to apply the deep learning methods, like deep belief network [12], deep Boltzmann machine [29], deep neural network [14, 17, 19], and deep autoencoder model [31], in various applications. *Ensemble learning* [22, 28] aims at using multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone. There exist a number of common ensemble types, which include *bootstrap aggregating (bagging)* [3], *boosting* [8], and *stacking* [34]. A more detailed introduction to the *ensemble learning* is available in the book [55].

Transfer learning focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. A comprehensive survey about *transfer learning* problems and algorithms is available in [26]. Meanwhile, by studying multiple tasks together, the *multi-task learning* problem was introduced. For the formal introduction to *multi-task learning*, please refer to [4]. Several other learning problems mentioned in this chapter include *multi-view* [36], *multi-source* [5], *multi-modal* [25], and *multi-domain* [24, 35] learning tasks, which are all strongly correlated with the *broad learning* task covered in this book.

1.9 Exercises

1. (Easy) Please describe what is *broad learning*, and why *broad learning* is an important learning problem.
2. (Easy) What are the two main problems covered in *broad learning* tasks?
3. (Easy) Please enumerate the information fusion methods at different levels, and try to list some application examples in which these information fusion methods can be used.
4. (Easy) Based on the information flow directions, how many different information fusion manners exist? Please enumerate these approaches, and provide some application settings that these approaches can be applied.
5. (Easy) Besides the *link prediction*, *community detection*, *information diffusion*, *viral marketing*, and *network embedding* problems, can you list some other learning problems that can be studied based on the *broad learning* setting? Please also mention what are the input information sources.
6. (Medium) What are the differences and correlations between *deep learning* and *broad learning*?
7. (Medium) What are the differences and correlations between *ensemble learning* and *broad learning*?
8. (Medium) Can *ensemble learning* models be used to handle *broad learning* problems? Please briefly explain how to do that.
9. (Medium) As introduced in this chapter, *transfer learning* and *multi-task learning* can both be viewed as a special case of *broad learning*. Can you briefly introduce why?
10. (Medium) Please briefly describe the relationship between the *multi-view, multi-modal, multi-source, multi-domain learning* with *broad learning*.

References

1. E. Arisoy, T. Sainath, B. Kingsbury, B. Ramabhadran, Deep neural network language models, in *Proceedings of the NAACL-HLT 2012 Workshop: Will We Ever Really Replace the N-gram Model? On the Future of Language Modeling for HLT (WLM '12)* (Association for Computational Linguistics, Stroudsburg, 2012), pp. 20–28
2. A. Bordes, N. Usunier, A. Garcia-Durán, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, in *Advances in Neural Information Processing Systems* (2013)
3. L. Breiman, Bagging predictors. Mach. Learn. **24**(2), 123–40 (1996)
4. R. Caruana, Multitask learning. Mach. Learn. **28**(1), 41–75 (1997)

5. W. Dai, Q. Yang, G. Xue, Y. Yu, Boosting for transfer learning, in *Proceedings of the 24th International Conference on Machine Learning* (ACM, New York, 2007), pp. 193–200
6. L. Deng, G. Hinton, B. Kingsbury, New types of deep neural network learning for speech recognition and related applications: an overview, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, Piscataway, 2013)
7. S. Fortunato, Community detection in graphs. Phys. Rep. **486**(3–5), 75–174 (2010)
8. Y. Freund, R. Schapire, A short introduction to boosting. J. Jpn. Soc. Artif. Intell. **14**(5), 771–780 (1999)
9. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, Cambridge, 2016). http://www.deeplearningbook.org
10. D. Gruhl, R. Guha, D. Liben-Nowell, A. Tomkins, Information diffusion through blogspace, in *Proceedings of the 13th International Conference on World Wide Web* (ACM, New York, 2004), pp. 491–501
11. S. Hill, Elite and upper-class families, in *Families: A Social Class Perspective* (2012)
12. G. Hinton, S. Osindero, Y. Teh, A fast learning algorithm for deep belief nets. Neural Comput. **18**(7), 1527–1554 (2006)
13. G. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, B. Kingsbury, Deep neural networks for acoustic modeling in speech recognition. IEEE Signal Process. Mag. **29**(6), 82–97 (2012)
14. H. Jaeger, Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach, Technical report (2002)
15. S. Jin, J. Zhang, P. Yu, S. Yang, A. Li, Synergistic partitioning in multiple large scale social networks, in *2014 IEEE International Conference on Big Data (Big Data)* (IEEE, Piscataway, 2014)
16. D. Kempe, J. Kleinberg, É Tardos, Maximizing the spread of influence through a social network, in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, New York, 2003), pp. 137–146
17. Y. Kim, Convolutional neural networks for sentence classification, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Association for Computational Linguistics, 2014), pp. 1746–1751
18. X. Kong, J. Zhang, P. Yu, Inferring anchor links across multiple heterogeneous social networks, in *Proceedings of the 22nd ACM international conference on Information and Knowledge Management* (ACM, New York, 2013), pp. 179–188
19. A. Krizhevsky, I. Sutskever, G. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems* (2012)
20. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. Nature **521**, 436–444 (2015). http://dx.doi.org/10.1038/nature14539
21. D. Liben-Nowell, J. Kleinberg, The link prediction problem for social networks, in *Proceedings of the Twelfth International Conference on Information and Knowledge Management* (ACM, New York, 2003), pp. 556–559
22. R. Maclin, D. Opitz, Popular ensemble methods: an empirical study. J. Artif. Intell. Res. (2011). arXiv:1106.0257
23. A. Mnih, G. Hinton, A scalable hierarchical distributed language model, in *Advances in Neural Information Processing Systems 21 (NIPS 2008)* (2009)
24. H. Nam, B. Han, Learning multi-domain convolutional neural networks for visual tracking. Comput. Vis. Pattern Recognit. (2015). arXiv:1510.07945
25. J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, A. Ng, Multimodal deep learning, in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)* (2011), pp. 689–696
26. S. Pan, Q. Yang, A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2010)
27. B. Perozzi, R. Al-Rfou, S. Skiena, Deepwalk: online learning of social representations, in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, New York, 2014), pp. 701–710
28. R. Polikar, Ensemble based systems in decision making. IEEE Circuits Syst. Mag. **6**(3), 21–45 (2006)
29. R. Salakhutdinov, G. Hinton, Semantic hashing. Int. J. Approx. Reason. **50**(7), 969–978 (2009)
30. S. Thrun, *Lifelong Learning Algorithms* (1998)
31. P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P. Manzagol, Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J. Mach. Learn. Res. **11**, 3371–3408 (2010)
32. J. Weston, S. Bengio, N. Usunier, Large scale image annotation: learning to rank with joint word-image embeddings. J. Mach. Learn. Res. **81**(1), 21–35 (2010)
33. J. Weston, S. Bengio, N. Usunier, Wsabie: scaling up to large vocabulary image annotation, in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)* (2011)
34. D. Wolpert, Stacked generalization. Neural Netw. **5**, 241–259 (1992)
35. T. Xiao, H. Li, W. Ouyang, X. Wang, Learning deep feature representations with domain guided dropout for person re-identification. Comput. Vis. Pattern Recognit. (2016). arXiv:1604.07528
36. C. Xu, D. Tao, C. Xu, A survey on multi-view learning. Mach. Learn. (2013). arXiv:1304.5634

37. Q. Zhan, J. Zhang, S. Wang, P. Yu, J. Xie, Influence maximization across partially aligned heterogeneous social networks, in *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (Springer, Berlin, 2015), pp. 58–69

38. Q. Zhan, J. Zhang, X. Pan, M. Li, P. Yu, Discover tipping users for cross network influencing, in *2016 IEEE 17th International Conference on Information Reuse and Integration (IRI)* (IEEE, Piscataway, 2016)

39. J. Zhang, Social network fusion and mining: a survey. Soc. Inf. Netw. (2018). arXiv:1804.09874

40. J. Zhang, P. Yu, Community detection for emerging networks, in *Proceedings of the 2015 SIAM International Conference on Data Mining* (Society for Industrial and Applied Mathematics, Philadelphia, 2015)

41. J. Zhang, P. Yu, MCD: mutual clustering across multiple social networks, in *2015 IEEE International Congress on Big Data* (IEEE, Piscataway, 2015)

42. J. Zhang, P. Yu, Multiple anonymized social networks alignment, in *2015 IEEE International Conference on Data Mining* (IEEE, Piscataway, 2015)

43. J. Zhang, P. Yu, PCT: partial co-alignment of social networks, in *Proceedings of the 25th International Conference on World Wide Web* (ACM, New York, 2016), pp. 749–759

44. J. Zhang, X. Kong, P. Yu, Predicting social links for new users across aligned heterogeneous social networks, in *2013 IEEE 13th International Conference on Data Mining* (IEEE, Piscataway, 2013)

45. J. Zhang, X. Kong, P. Yu, Transferring heterogeneous links across location-based social networks, in *Proceedings of the 7th ACM International Conference on Web Search and Data Mining* (ACM, New York, 2014), pp. 303–312

46. J. Zhang, P. Yu, Z. Zhou, Meta-path based multi-network collective link prediction, in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, New York, 2014), pp. 1286–1295

47. J. Zhang, W. Shao, S. Wang, X. Kong, P. Yu, PNA: partial network alignment with generic stable matching, in *2015 IEEE International Conference on Information Reuse and Integration* (IEEE, Piscataway, 2015)

48. J. Zhang, S. Wang, Q. Zhan, P. Yu, Intertwined viral marketing in social networks, in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (IEEE, Piscataway, 2016)

49. J. Zhang, P. Yu, Y. Lv, Q. Zhan, Information diffusion at workplace, in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management* (ACM, New York, 2016), pp. 1673–1682

50. J. Zhang, Q. Zhan, P. Yu, Concurrent alignment of multiple anonymized social networks with generic stable matching, in *Theoretical Information Reuse and Integration* (Springer, Cham, 2016), pp. 173–196

51. J. Zhang, J. Chen, J. Zhu, Y. Chang, P. Yu, Link prediction with cardinality constraints, in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (ACM, New York, 2017), pp. 121–130

52. J. Zhang, C. Xia, C. Zhang, L. Cui, Y. Fu, P. Yu, BL-MNE: emerging heterogeneous social network embedding through broad learning with aligned autoencoder, in *Proceedings of the 2017 IEEE International Conference on Data Mining* (IEEE, Piscataway, 2017)

53. J. Zhang, J. Chen, S. Zhi, Y. Chang, P. Yu, J. Han, Link prediction across aligned networks with sparse and low rank matrix estimation, in *2017 IEEE 33rd International Conference on Data Engineering (ICDE)* (IEEE, Piscataway, 2017)

54. J. Zhang, L. Cui, P. Yu, Y. Lv, BL-ECD: broad learning based enterprise community detection via hierarchical structure fusion, in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (ACM, New York, 2017), pp. 859–868

55. Z. Zhou, *Ensemble Methods: Foundations and Algorithms*, 1st edn. (Chapman & Hall/CRC, London, 2012)

56. J. Zhu, J. Zhang, L. He, Q. Wu, B. Zhou, C. Zhang, P. Yu, Broad learning based multi-source collaborative recommendation, in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (ACM, New York, 2017), pp. 1409–1418