Journal of Al-Qadisiyah for Computer Science and Mathematics Vol. 14(4) 2022 , pp Comp. 90–98

University Of AL-Qadisiyah

Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Explicit aspect extraction techniques: Review

Arwa Akram[a], Aliea Salman Sabir[b]

University of Basra, Computer science and information technology college, Information Systems depadepartment, Basra, Iraq..
itpg.arwa.akram@uobasrah.edu.iq

University of Basra, Computer science and information technology college, Information Systems depadepartment, Basra, Iraq.
aliea.sabir@uobasrah.edu.iq

ARTICLE INFO

Article history:
Received: 06 /11/2022
Rrevised form: 12 /12/2022
Accepted : 15 /12/2022
Available online: 31 /12/2022

Keywords:

sentiment analysis,
Natural language processing,
Aspect extraction,
Sentiment classification,
Explicit aspects

ABSTRACT

Sentiment analysis is gathering opinion keywords, such as aspects, opinions, or features- and figuring out their semantic perspective relations. More specifically, aspect-based sentiment analysis – (ABSA) is a subfield of natural language that focuses on phrases associated with aspects and detects the sentiment that belongs to each one. Aspect extraction and sentiment analysis are the two fundamental functions in ABSA, and the different classes of aspects are explicit and implicit. To shed some light on this problem, we discuss aspect extraction tasks and classify explicit aspect extraction strategies into two categories: supervised, unsupervised and semi-supervised. This article explores prior research and their approaches by reviewing works from 2016 to 2021 and comparing numerous elements comprehensively, including classifying systems, classifier methods, datasets, and performance.

MSC..

https://doi.org/ 10.29304/jqcm.2022.14.4.1116

1. Introduction

Aspect-Based Sentiment Analysis (ABSA) is one of the analysis types that separates data into aspects and extracts the sentiment associated with each aspect. It indicates the author's or speaker's attitude on a particular subject. It is necessary to comprehend the many sentiment analysis methods to determine which is most applicable. Text-based views and evaluations are assessed computationally to determine the polarity of the text and whether it holds a positive, negative, or neutral sentiment. A positive opinion indicates a favorable viewpoint compared to a negative one, which indicates an unfavorable perspective while a neutral opinion is a non-biased viewpoint that does not sympathize with or denigrate its topic [1].

In recent years, there has been a great deal of interest in ABSA which can be identified as the task that classifies opinion polarity towards a given target aspect [2]. The articles in this review span the years 2016 to 2021, when ABSA was particularly popular and people used it more frequently than ever. ABSA delivers significantly more context information than sentiment analysis in general. Due to technological improvements, aspect-based sentiment analysis is now more trustworthy, accurate, and accessible than ever. To improve sentiment analysis automation, modern-day applications include Deep Learning (DL), Natural

Email addresses:

Communicated by ‘sub etitor’

Language Processing (NLP), Machine Learning (ML) Text Processing, and much other cutting-edge technology. It enables rating within a quantifiable range the positive, negative, or neutral attitude of a text with minimal human effort [3].

Even though aspect-based sentiment analysis has been around for quite some time, it has recently become a prevalent task because of the massive technological advancement. Machine learning algorithms are more reliable and accurate at classifying and identifying statements with an opinion. ABSA consists of two subtasks, sentiment classification and aspect detection. Companies use sentiment analysis to increase their product sales and services. Table 1 clearly shows how the two tasks of ABSA work.

Sentence	Aspect Term	Sentiment	Polarity
The food is great but the price is a little bit expensive.	Food	great	positive
The food is great but the price is a little bit expensive.	Price	expensive	negative

Table 1- Implementation of ABSA

First, the aspect phrase is extracted, and then its sentiment is identified. The following step is to determine whether this sentiment is positive, negative, or neutral. The work of extracting aspects focuses on finding explicit and implicit aspect terms within a sentence and then categorizing them. Since it is common knowledge that every company's primary goal is to serve its customers, it is crucial to do these tasks precisely and deliver results that will enable businesses to adjust their services and goods accordingly.[4] The following figure illustrates the tasks of ABSA.

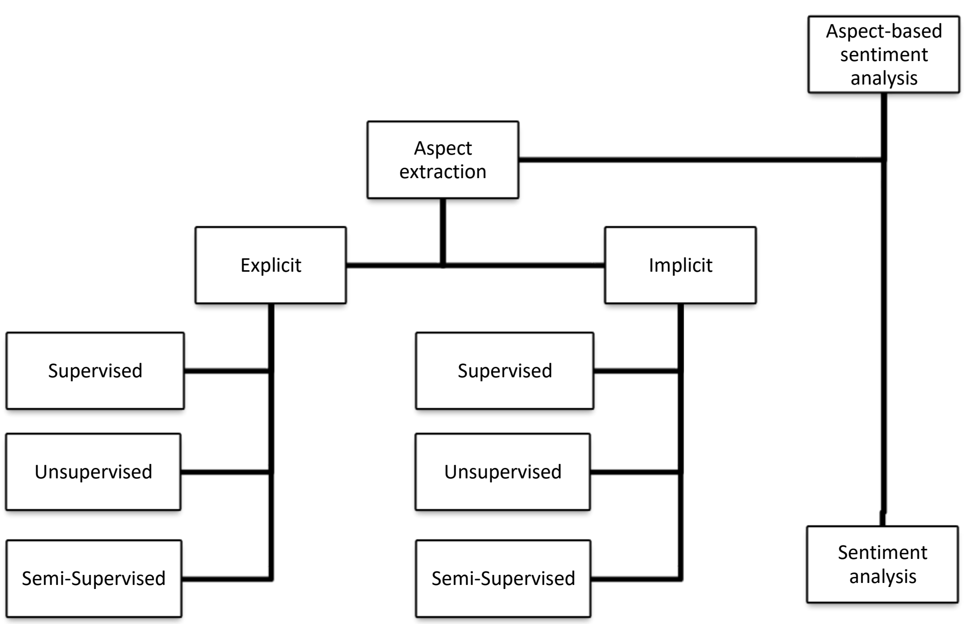

Fig. 1: Aspect-based sentiment (ABSA) analysis taxonomy

2. Aspect extraction techniques

Aspect extraction is a sentiment analysis sub-task that seeks to locate targets of opinions in text documents by identifying the particular features of a service or item that the opinion owner is either appreciating or criticizing[5]. Primarily, there are implicit and explicit features. For instance, "The smart TV is quite large, but its resolution is poor." In this scenario, there are explicit and implicit components of the smart TV. The explicit aspect, "resolution" has a negative polarity, while the implicit aspect, "big," refers to a feature of smart TV that has a positive polarity.

One of the major aims of : Aspect-based sentiment (ABSA) is to extract implicit and explicit elements. In this paper, the attempt is to divide explicit aspect extraction methods into three primary categories: supervised, unsupervised and semi-supervised as shown in Fig. 2 and further explain the different approaches inside each one of them. Aspect extraction of review content is handled on a sentence-by-sentence basis, with each sentence including one to many product aspects.

In this review, we will discuss twenty previous researchers that applied their techniques in extracting explicit aspects from the data and further determining their polarity. As shown in **table (2),** the paper's approach, the domain of work determined by the dataset used, and the performance achieved by applying their techniques.

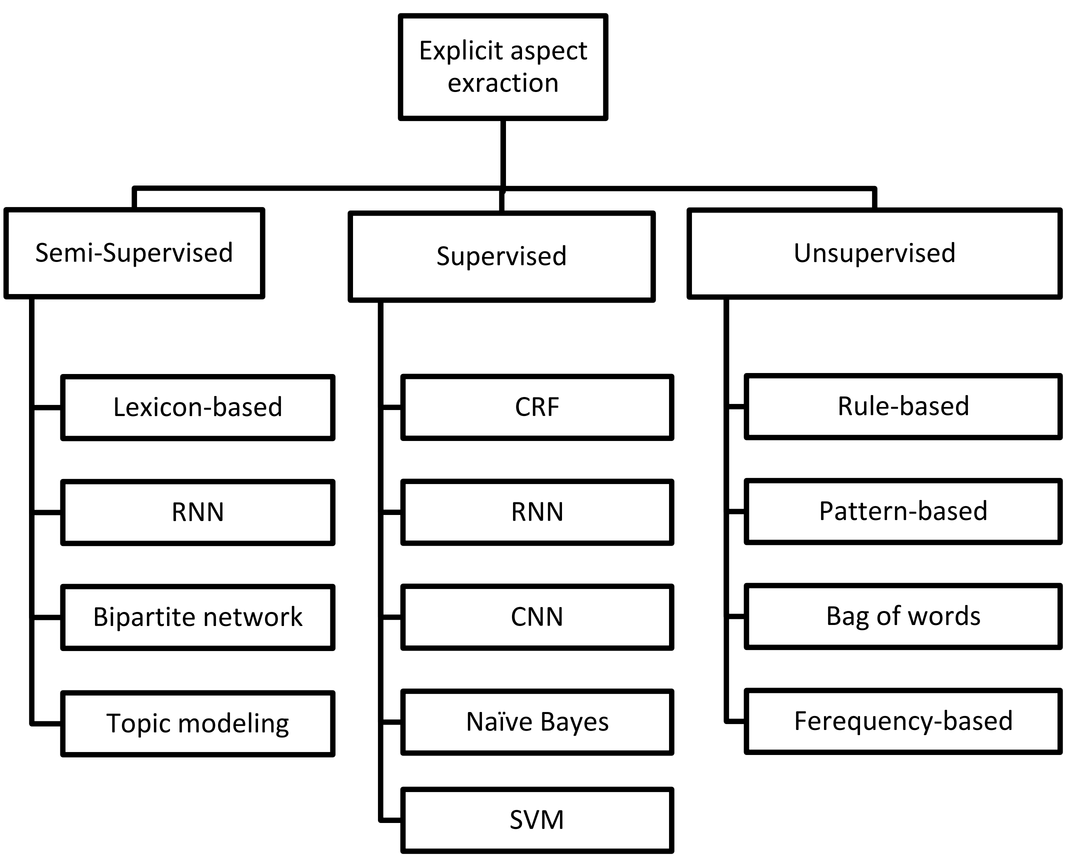

Fig. 2: Explicit aspect extraction taxonomy

Table 2- Previous work of explicit aspect extraction methods

ID	Reference	Year of publication	approach	Method / Technique	Datasets	Accuracy
1	[6]	2016	Supervised	CRF	Trip Advisor's restaurant reviews	0.88
2	[7]	2018	Supervised	CRF	Yelp Dataset : Restaurants Laptops	 0.68 0.70
3	[8]	2017	Supervised	CRF	Yelp Restaurant Yelp Hotel Amazon product Laptop	0.83 0.78 0.82
4	[9]	2020	Supervised	RNN	SemEval2014 restaurant Laptop twitter	0.80 0.72 0.72
5	[10]	2018	Supervised	CNN	SEMEVAL-2016 TASK 5 restaurant Laptop	 0.75 0.51
6	[11]	2021	Supervised	CNN	AI Challenger 2018	0.88
7	[12]	2017	Supervised	Naïve Bayes	social media website (Twitter)	0.81
8	[13]	2016	Supervised	SVM	SemEval 2016 Hotel reviews	 0.76
9	[14]	2016	Supervised	SVM	SemEval 2015	0.73
10	[15]	2017	Unsupervised	Rule-based	IMDb Movie reviews	0.74
11	[16]	2018	Unsupervised	Rule-based	SemEval- 2014 Restaurant Laptop	 0.72 0.64
12	[17]	2019	Unsupervised	Rule-based	Trip Advisor's restaurant reviews	N/A
13	[18]	2017	Unsupervised	Pattern-based	customer review dataset	0.89
14	[19]	2018	Unsupervised	Pattern-based	customer review dataset	0.86
15	[20]	2015	Unsupervised	Bag of words	SemEval- 2015 Restaurant Laptop	 0.71 0.43
16	[21]	2015	Unsupervised	Frequency-based	Electronic products Online reviews	0. 74
17	[22]	2017	Semi-Supervised	Lexicon-based	Naver Movie datasets Yelp Restaurants and Shopping datasets	0.76 0.88

18	[23]	2017	Semi-Supervised	RNN	YELP LAPTOP	0.65 0.73
19	[24]	2016	Semi-Supervised	Bipartite network	SEMEVAL 2014	N/A
20	[25]	2014	Semi-Supervised	Topic modeling	Crawl product review	N/A

3. Supervised:

It is distinguished by the way it trains computers to accurately classify data or predict the outcome using labeled datasets. The fundamental aim is to build an estimator that can forecast an object's label based on a set of characteristics. Supervised learning assists enterprises in finding scalable and flexible solutions to a number of real-world issues[26]. The methods of supervised learning are listed below, along with examples of how the researchers employed them:

3.1 Conditional random field (CRF)

3.1.1 Gojali and Khodra (2016) used the CRF algorithm to obtain the best tag sequence for a given sentence in their paper. To train the CRF, they used two features: the first is the linguistic part, and the second is its part-of-speech (POS) tags. The work of the Mallet tool achieved remarkable results in training and generating the best tag sequence.

3.1.2 Xu, Liu, Wang, and Yin In (2018) reported an enhanced system for opinion target recognition based on Conditional Random Fields and a developed sequence labeling technique with Conditional Random Fields. The CRF++ tool and training document was given for the modeling process, with the training document serving as the CRF's input.

3.1.3 The idea behind the work of Nasim and Haider (2017) is to use CRF to train the aspect term extractor. They studied modeling the probability distribution over the entire sequence labels and feature functions likely to appear in the data.

3.2 Recurrent neural network (RNN)

3.2.1 It has been noted that using RNNs in ABSA has several drawbacks, such as the lack of position invariance and the requirement for sensitivity to crucial local patterns. Liu N and Shen B (2020) suggest what is known as a Gated Alternate Neural Network (GANN), a novel framework to solve the shortcomings of RNNs utilized by ABSA.

3.3 Convolutional neural network (CNN)

3.3.1 Jihan, Senarath, and Ranathunga (2018) enhanced the aspect extraction task by presenting an improved CNN architecture. Their work demonstrates that employing non-static CNN rather than static CNN yields superior outcomes in terms of fine-tuning the word embedding characteristics in response to consumer input.

3.3.2 Zhao N, Gao H, Li H, and Wen X (2021) combined Gated Recurrent Unit (GRU) the Convolutional Neural Network (CNN). Their proposed approach uses a mix of local characteristics developed by CNN and long-term reliance learned by GRU. They use a blend of regional traits developed by CNN and long-term dependency defined by GRU.

3.4 Naïve Bayes (NB)

3.4.1 The framework proposed by Ahmed, Hina, Atwell, and Ahmed (2017) uses TF-IDF and Naive Bayes machine learning ML algorithms. They tested their work by implementing it on social media websites (Twitter) obtained by API. The aim was to handle the huge amounts of tweets submitted rapidly.

3.5 Support vector machine (SVM)

3.5.1 AL-Smadi, Qwasmeh, Talafha, Al-Ayyoub, Jararweh, and Benkhelifaa (2016) created a model that uses SVM, a Support Vector Machine, and a linear kernel. Unigrams used sentences connected to the training data classes to train their methodology. When used with customer reviews, it produced good accuracy.

3.5.2 The idea behind Pannala N, Nawarathna C, Jayakody J, Rupasinghe L, and Krishnadeva K (2016) research was to generally examine the binary SVM (Support Vector Machine) related to each category with numerous parameters based on the performance of the dataset.

4 Unsupervised:

A machine learning technique that enables models to be trained on unlabeled datasets and then permitted to act on that data with no monitoring. The aim of unsupervised learning is finding the underlying components of a dataset, classifying the data into groups based on similarities, and compressing the dataset and displaying it[26]. Here is a collection of unsupervised learning methods along with examples of how they were used by the researchers:

4.1 Rule-based

4.1.1 In Piryani, R., Gupta, V., Singh, V. K., & Ghose, U. (2017) work, they proposed a model that creates an aspect-level opinion summary. Their approach relies on a rule-based linguistic method that detects the aspects in the movie reviews data set from IMDb. The suggested design achieved good accuracy results, showing high possibilities applied in an integrated opinion profiling system.

4.1.2 The proposed approach by Wu, C., Wu, F., Wu, S., Yuan, Z., & Huang, Y. (2018) aimed to extract nominal phrase chunks by deploying chunk-level linguistic rules. Following the treatment of the chunks as potential opinion targets and aspects, unrelated candidates are then eliminated using domain correlation.

4.1.3 Firmanto and Sarno (2019) use a combination of SentiCircle, grammatical rules, and word similarity to propose an (ABSA) aspect-based sentiment analysis method. They extract aspects of rules based on phrase identification in constituency parse.

4.2 Pattern-based

4.2.1 In another paper, Rana, T. A., & Cheah, Y.-N. (2017) recommended a two-level aspect trimming strategy to remove irrelevant aspects. The suggested method employed an aspect extraction paradigm based on sequential patterns to extract noun terms and phrases where noun words are related.

4.2.2 The main contribution behind Rana, T. A., & Cheah, Y.-N. (2018) article is to identify links between aspects and opinions by learning from user behavior. For aspect detection, they suggested employing consecutive patterns. When the model was applied to the dataset of customer reviews, it produced great accuracy.

4.3 Bag of words (BOW)

4.3.1 In their research, Jiménez-Zafra, S. M., Martín-Valdivia, M. T., Ureña-López, L. A, E., and Martínez-Cámara. (2015) approach the use of a bag of words. They significantly improved the classification process by combining multiple linguistic resources that use a lexicon-based method.

4.4 Frequency-based

4.4.1 The proposed approach by Li, S., Zhou, L., & Li, Y. (2015) developed an extraction strategy based on frequency using PMI-IR. Their approach suggests utilizing web search to evaluate the semantic similarity in multiple aspects of possible target entities.

5 Semi-supervised:

A form of machine learning algorithm known as semi-supervised learning falls within supervised and unsupervised learning methods. During the training phase, it employs a mix of both labeled and unlabeled datasets[26]. The following list of semi-supervised learning techniques includes instances of how they were applied by the researchers.

5.1 Recurrent neural network (RNN)

5.1.1 Ding Y., Yu C., and Jiang J. produced words using an embedded vector in their paper. Additionally, they implemented a sequencing labeler built upon Recurrent Neural Networks (RNNs) into the developed neural network to take advantage of review phrases and their feature terms.

5.2 Topic modeling

5.2.1 In Wang T, Cai Y, Leung HF, Lau RY, Li Q, Min H work, to get over the constraints of current approaches, two novel semi-supervised strategies for aspect extraction are proposed with the help of the construction of efficient textual settings for topic modeling.

5.3 Lexicon-based

5.3.1 In their paper, Amplayo, R. K. & Hwang, S W suggested a framework known as the (MicroASM) which is short for Micro Aspect Sentiment Model. MicroASM is based on the concept that small reviews can be grouped into larger evaluations by using emotion aspect word collections as informational key components.

5.4 Bipartite network

5.4.1 Matsuno, I.P., Rossi, R.G., Marcacini, R.M provided a method that uses semi-supervised approaches, in which much less labeled data is needed for learning to take place. In order to accomplish semi-supervised learning, we model the data into networks by using bipartite networks to represent the data.

Conclusion

Aspect-based sentiment analysis is vital for any business to create a customer-centric experience. Each one of the techniques we talk about gives a different accuracy rate in various domains. In supervised approaches, Conditional random field (CRF) and naïve Bayes (NB) achieve the highest accuracies, while pattern-based strategies have scored the highest results in unsupervised techniques. In semi-supervised learning, the best outcomes have been achieved with lexicon-based approach. Fact remains that the accuracy results can vary depending on the implementation of the algorithms. Finally, numerous possible applications have yet to be researched, as well as additional domains, which are significant and need further investigations.

References

1. Yu, H. and V. Hatzivassiloglou. *Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences*. in *Proceedings of the 2003 conference on Empirical methods in natural language processing*. 2003.
2. Das, A. and W.E. Zhang. *ABSA-Bench: Towards the Unified Evaluation of Aspect-based Sentiment Analysis Research*. in *Proceedings of the The 18th Annual Workshop of the Australasian Language Technology Association*. 2020.
3. Khudeyer, R.S., M. Alabbas, and M. Radif, *Multi-font arabic isolated character recognition using combining machine learning classifiers.* journal of southwest jiaotong university, 2020. **55**(1).
4. Hamoud, A.K., et al. *Improving service quality using consumers' complaints data mart which effect on financial customer satisfaction*. in *Journal of Physics: Conference Series*. 2020. IOP Publishing.
5. Poria, S., E. Cambria, and A. Gelbukh, *Aspect extraction for opinion mining with a deep convolutional neural network.* Knowledge-Based Systems, 2016. **108**: p. 42-49.
6. Gojali, S. and M.L. Khodra. *Aspect based sentiment analysis for review rating prediction*. in *2016 International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA)*. 2016. IEEE.
7. Xu, L., et al., *Aspect based sentiment analysis for online reviews*, in *Advances in computer science and ubiquitous computing*. 2017, Springer. p. 475-480.
8. Nasim, Z. and S. Haider, *ABSA toolkit: An open source tool for aspect based sentiment analysis.* International Journal on Artificial Intelligence Tools, 2017. **26**(06): p. 1750023.
9. Liu, N. and B. Shen, *Aspect-based sentiment analysis with gated alternate neural network.* Knowledge-Based Systems, 2020. **188**: p. 105010.
10. Jihan, N., Y. Senarath, and S. Ranathunga. *Aspect extraction from customer reviews using convolutional neural networks*. in *2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer)*. 2018. IEEE.
11. Zhao, N., et al., *Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis.* IEEE Access, 2021. **9**: p. 15561-15569.
12. Ahmed, S., et al., *Aspect based sentiment analysis framework using data from social media network.* IJCSNS Int. J. Comput. Sci. Netw. Secur, 2017. **17**: p. 100-105.
13. Mohammad, A.-S., et al. *An enhanced framework for aspect-based sentiment analysis of Hotels' reviews: Arabic reviews case study*. in *2016 11th International conference for internet technology and secured transactions (ICITST)*. 2016. IEEE.
14. Pannala, N.U., et al. *Supervised learning based approach to aspect based sentiment analysis*. in *2016 IEEE international conference on computer and information technology (CIT)*. 2016. IEEE.
15. Piryani, R., et al., *A linguistic rule-based approach for aspect-level sentiment analysis of movie reviews*, in *Advances in computer and computational sciences*. 2017, Springer. p. 201-209.
16. Wu, C., et al., *A hybrid unsupervised method for aspect term and opinion target extraction.* Knowledge-Based Systems, 2018. **148**: p. 66-73.
17. Firmanto, A. and R. Sarno, *Aspect-based sentiment analysis using grammatical rules, word similarity and sentiCircle.* International Journal of Intelligent Engineering and Systems, 2019. **12**(5): p. 190-201.
18. Rana, T.A. and Y.-N. Cheah. *Improving aspect extraction using aspect frequency and semantic similarity-based approach for aspect-based sentiment analysis*. in *International conference on computing and information technology*. 2017. Springer.
19. Rana, T.A. and Y.-N. Cheah, *Sequential patterns rule-based approach for opinion target extraction from customer reviews.* Journal of Information Science, 2019. **45**(5): p. 643-655.
20. Jiménez-Zafra, S.M., et al., *Combining resources to improve unsupervised sentiment analysis at aspect-level.* Journal of Information Science, 2016. **42**(2): p. 213-229.
21. Li, S., L. Zhou, and Y. Li, *Improving aspect extraction by augmenting a frequency-based method with web-based similarity measures.* Information Processing & Management, 2015. **51**(1): p. 58-67.
22. Amplayo, R.K. and S.-w. Hwang. *Aspect sentiment model for micro reviews*. in *2017 IEEE International Conference on Data Mining (ICDM)*. 2017. IEEE.
23. Ding, Y., C. Yu, and J. Jiang. *A neural network model for semi-supervised review aspect identification*. in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. 2017. Springer.
24. Matsuno, I.P., et al., *Aspect-based sentiment analysis using semi-supervised learning in bipartite heterogeneous networks.* Journal of Information and Data Management, 2016. **7**(2): p. 141-141.
25. Wang, T., et al., *Product aspect extraction supervised with online domain knowledge.* Knowledge-Based Systems, 2014. **71**: p. 86-100.
26. Berry, M.W., A. Mohamed, and B.W. Yap, *Supervised and unsupervised learning for data science*. 2019: Springer.