

1. Record Nr.	UNINA9910678248403321
Autore	Sarang P. G (Poornachandra G.)
Titolo	Thinking Data Science : A Data Science Practitioner's Guide // by Poornachandra Sarang
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2023
ISBN	9783031023637 9783031023620
Edizione	[1st ed. 2023.]
Descrizione fisica	1 online resource (366 pages) : illustrations
Collana	The Springer Series in Applied Machine Learning, , 2520-1301
Disciplina	006.31 005.7
Soggetti	Machine learning Artificial intelligence - Data processing Artificial intelligence Machine Learning Data Science Artificial Intelligence
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Chapter. 1. Data Science Process -- Chapter. 2. Dimensionality Reduction - Creating Manageable Training Datasets -- Chapter. 3. Classical Algorithms - Over-view -- Chapter. 4. Regression Analysis -- Chapter. 5. Decision Tree -- Chapter. 6. Ensemble - Bagging and Boosting -- Chapter. 7. K-Nearest Neighbors -- Chapter. 8. Naive Bayes -- Chapter. 9. Support Vector Machines: A supervised learning algorithm for Classification and Regression -- Chapter. 10. Clustering Overview -- Chapter. 11. Centroid-based Clustering -- Chapter. 12. Connectivity-based Clustering -- Chapter. 13. Gaussian Mixture Model -- Chapter. 14. Density-based -- Chapter. 15 -- BIRCH -- Chapter. 16. CLARANS -- Chapter. 17. Affinity Propagation Clustering -- Chapter. 18. STING -- Chapter. 19. CLIQUE -- Chapter. 20. Artificial Neural Networks -- Chapter. 21. ANN-based Applications -- Chapter. 22. Automated Tools -- Chapter. 23. DataScientist's Ultimate Workflow.
Sommario/riassunto	This definitive guide to Machine Learning projects answers the

problems an aspiring or experienced data scientist frequently has: Confused on what technology to use for your ML development? Should I use GOFAI, ANN/DNN or Transfer Learning? Can I rely on AutoML for model development? What if the client provides me Gig and Terabytes of data for developing analytic models? How do I handle high-frequency dynamic datasets? This book provides the practitioner with a consolidation of the entire data science process in a single "Cheat Sheet". The challenge for a data scientist is to extract meaningful information from huge datasets that will help to create better strategies for businesses. Many Machine Learning algorithms and Neural Networks are designed to do analytics on such datasets. For a data scientist, it is a daunting decision as to which algorithm to use for a given dataset. Although there is no single answer to this question, a systematic approach to problem solving is necessary. This book describes the various ML algorithms conceptually and defines/discusses a process in the selection of ML/DL models. The consolidation of available algorithms and techniques for designing efficient ML models is the key aspect of this book. Thinking Data Science will help practising data scientists, academicians, researchers, and students who want to build ML models using the appropriate algorithms and architectures, whether the data be small or big.

---