

1. Record Nr.	UNINA9910741140603321
Autore	Edwards Kieran Jay
Titolo	Astronomy and Big Data : A Data Clustering Approach to Identifying Uncertain Galaxy Morphology // by Kieran Jay Edwards, Mohamed Medhat Gaber
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2014
ISBN	3-319-06599-8
Edizione	[1st ed. 2014.]
Descrizione fisica	1 online resource (112 p.)
Collana	Studies in Big Data, , 2197-6503 ; ; 6
Disciplina	520.222
Soggetti	Computational intelligence Artificial intelligence Observations, Astronomical Astronomy—Observations Data mining Computational Intelligence Artificial Intelligence Astronomy, Observations and Techniques Data Mining and Knowledge Discovery
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Introduction -- Astronomy, Galaxies and Stars: An Overview -- Astronomical Data Mining -- Adopted Data Mining Methods -- Research Methodology -- Development of Data Mining Models -- Experimentation Results -- Conclusion and Future Work.
Sommario/riassunto	With the onset of massive cosmological data collection through media such as the Sloan Digital Sky Survey (SDSS), galaxy classification has been accomplished for the most part with the help of citizen science communities like Galaxy Zoo. Seeking the wisdom of the crowd for such Big Data processing has proved extremely beneficial. However, an analysis of one of the Galaxy Zoo morphological classification data sets has shown that a significant majority of all classified galaxies are labelled as "Uncertain". This book reports on how to use data mining, more specifically clustering, to identify galaxies that the public has

shown some degree of uncertainty for as to whether they belong to one morphology type or another. The book shows the importance of transitions between different data mining techniques in an insightful workflow. It demonstrates that Clustering enables to identify discriminating features in the analysed data sets, adopting a novel feature selection algorithms called Incremental Feature Selection (IFS). The book shows the use of state-of-the-art classification techniques, Random Forests and Support Vector Machines to validate the acquired results. It is concluded that a vast majority of these galaxies are, in fact, of spiral morphology with a small subset potentially consisting of stars, elliptical galaxies or galaxies of other morphological variants.
